JMIR Public Health and Surveillance

Impact Factor (2023): 3.5
Volume 11 (2025) ISSN 2369-2960 Editor in Chief: Travis Sanchez, PhD, MPH

Contents

Original Papers

Engagement With Digital Adherence Technologies as Measures of Intervention Fidelity Among Adults With Drug-Susceptible Tuberculosis and Health Care Providers: Descriptive Analysis Using Data From Cluster-Randomized Trials in Five Countries (e62881)
Jason Alacapa, Amare Tadesse, Natasha Deyanova, Tanyaradzwa Dube, Andrew Mganga, Rachel Powers, Job van Rest, Norma Madden, Egwuma Efo, Salome Charalambous, Kristian van Kalmthout, Degu Jerene, Katherine Fielding. . . . 18

The Impact of a Mobile Money–Based Intervention on Maternal and Neonatal Health Outcomes in Madagascar: Cluster-Randomized Controlled Trial (e70182)
Lisa Bogler, Bítia Vieira, Harizaka Andriamasy, Zavaniarivo Rampanjato, Sebastian Vollmer, Till Bärnighausen, Andriamampianina Ralisimalala, Julius Emmrich, Samuel Knauss. . . . 33

Leveraging Administrative Health Databases to Address Health Challenges in Farming Populations: Scoping Review and Bibliometric Analysis (1975-2024) (e62939)
Pascal Petit, Nicolas Vuillerme. . . . 267

Health Equity of Hypertension Management Between Local Residents and Internal Migrants in Shenzhen, China: Cross-Sectional Study (e65548)
Jinyu He, Yanjun Li, Huatang Zeng, Haoran Sun, Liqun Wu, Zhenzhen Zhu, Ning Zhang, Wannian Liang. . . . 394

Accessibility of eHealth Before and During the COVID-19 Pandemic Among People With and People Without Impairment: Repeated Cross-Sectional Survey (e64707)
Linda Pettersson, Stefan Johansson, Ingrid Demmelmaier, Lena von Koch, Jan Gulliksen, Per-Olof Hedvall, Karl Gummesson, Catharina Gustavsson. . . . 410

Facilitated Telemedicine as a Patient-Centered, Sociotechnical Intervention to Integrate Hepatitis C Treatment Into Opioid Treatment Programs and Overcome the Digital Divide Among Underserved Populations: Qualitative Study (e68854)
Andrew Talal, Arpan Dharia, Marianthi Markatou, Lawrence Brown Jr, Kenneth Bossert, Zakiya Grubbs, Raktim Mukhopadhyay, Boatemaa Ntiri-Reid, Elisabeth Houtsmuller. . . . 425

Changes in Epidemiological Characteristics of Varicella and Breakthrough Cases in Ningbo, China, From 2010 to 2023: Surveillance Study (e71691)
Xingqiang Pan, Yan Zhang, Xuefei Zhao, Dandan Zhang. . . . 441

Epidemic Trends and Spatial Distribution Characteristics of Hepatitis B in China: Surveillance Study (e70888)
Xiaoxue Li, Jie Liang, Lanqing Ma, Shijie Shen, Xueying Han, Tao Sun, Heng Guo. . . . 456

XSL•FO
RenderX

Spatiotemporal Trends in Deaths From External Causes in Brazil: 23-Year Ecological and Population-Based Study (e68002)
Lucas Andrade, Wandklebson Silva da Paz, Luís Santos de Melo, Débora dos Santos Tavares, Alvaro Lopes de Sousa, Emerson Silva Camargo, Carla Arena Ventura, Regina Silva Souza, Karina Gomes Machado de Araújo, Carlos Freire de Souza, Allan dos Santos, Fagner Campos, Márcio Bezerra Santos. . . . 470

Epidemiology and Treatment Outcomes of Pulmonary Tuberculosis in Dazu District, Chongqing, China, 2005-2024: Surveillance Study (e78564)
Yu Yu, Xiu-juan Hu, Xiao-man Fang, Jing Wu. . . . 486

Impact of Primary Health Care Data Quality on Infectious Disease Surveillance in Brazil: Case Study (e67050)
Pilar Florentino, Juracy Bertoldo Junior, George Barbosa, Thiago Cerqueira-Silva, Vinicius Oliveira, Marcio Garcia, Gerson Penna, Viviane Boaventura, Pablo Ramos, Manoel Barral-Netto, Izabel Marcilio. . . . 508

Effectiveness of Frequent Point-of-Care Molecular COVID-19 Surveillance in a Rural Workplace: Nonrandomized Controlled Clinical Trial Among Miners (e59845)
Akshay Sood, William “Cotton” Jarrell, Xin Shore, Nestor Sosa, Alisha Parada, Nicholas Edwardson, Alexandra Yingling, Teah Amirkabirian, Qiuying Cheng, Ivy Hurwitz, Linda Cook, Shuguang Leng, Orrin Myers, Douglas Perkins. . . . 517

Updated Surveillance Metrics and History of the COVID-19 Pandemic (2020-2023) in East Asia and the Pacific Region: Longitudinal Trend Analysis (e53214)
Alexander Lundberg, Alan Soetikno, Scott Wu, Egon Ozer, Sarah Welch, Yingxuan Liu, Claudia Hawkins, Maryann Mason, Robert Murphy, Robert Havey, Charles Moss, Chad Achenbach, Lori Post. . . . 533

Wastewater Monitoring During the COVID-19 Pandemic in the Veneto Region, Italy: Longitudinal Observational Study (e58862)
Honoria Ocagli, Marco Zambito, Filippo Da Re, Vanessa Groppi, Marco Zampini, Alessia Terrini, Franco Rigoli, Irene Amoruso, Tatjana Baldovin, Vincenzo Baldo, Francesca Russo, Dario Gregori. . . . 575

Monitoring Public Health Through a Comprehensive Primary Care Database in the Netherlands: Overview of the Nivel Syndromic Surveillance System (e58767)
Christos Baliatsas, Jojanneke van Summeren, Sander van Beusekom, Amy Matser, Mariette Hooiveld. . . . 588

Experiences and Perceptions About Death Reporting and Notification Among Rural Communities on the Islands of Lake Victoria, Uganda: Qualitative Study (e77135)
Steven Kabwama, Douglas Bulafu, Rawlance Ndejjo, Rose Nampeera, Christine Kihembo, Chukwuma Umeokonkwo, Simon Antara, Sheba Gitta, Caroline Kyozira, Allan Muruta, Jean-Edgard Nguessan, Rhoda Wanyenze. . . . 596

Comparison of Virus Watch COVID-19 Positivity, Incidence, and Hospitalization Rates With Other Surveillance Systems: Surveillance Study (e69655)
Wing Fong, Vincent Nguyen, Sarah Beale, Thomas Byrne, Cyril Geismar, Ellen Fragaszy, Jana Kovar, Alexei Yavlinsky, Ibrahim Abubakar, Andrew Hayward, Robert Aldridge. . . . 607

Analyzing Satellite Imagery to Target Tuberculosis Control Interventions in Densely Urbanized Areas of Kigali, Rwanda: Cross-Sectional Pilot Study (e68355)
Mauro Faccin, Caspar Geenen, Michiel Happaerts, Sien Ombelet, Patrick Migambi, Emmanuel André. . . . 623

Balancing Human Mobility and Health Care Coverage in Sentinel Surveillance of Brazilian Indigenous Areas: Mathematical Optimization Approach (e69048)
Juliane Oliveira, Adriano Vasconcelos, Andrêza Alencar, Maria Cunha, Izabel Marcilio, Manoel Barral-Netto, Pablo P Ramos. . . . 635

Statistical Relationship Between Wastewater Data and Case Notifications for COVID-19 Surveillance in the United States From 2020 to 2023: Bayesian Hierarchical Modeling Approach (e68213)
Masahiko Haraguchi, Fayette Klaassen, Ted Cohen, Joshua Salomon, Nicolas Menzies. . . . 642

XSL•FO
RenderX

SARS-CoV-2 Detection in International Travelers Through Wastewater-Based Epidemiology at the Kigali International Airport: Genomic Surveillance Study (e71104)
Rwagasore Edson, Robert Rutayisire, Ziad El-Khatib, Olivier Nsekuye, Hugues Mucunguzi, Raissa Muvunyi, Esperance Umumararungu, Eric Remera. . . . 653

Identifying Gaps and Challenges in Acute Hepatitis B Surveillance in the Country of Georgia: Comprehensive Surveillance System Evaluation (e72888)
Lika Karichashvili, Ketevan Galdavadze, Khatuna Zakhashvili, Maia Tsereteli, Ekaterine Ruadze, Sophia Surguladze, Pawel Stefanoff. . . . 662

Equivalence of Type 2 Diabetes Prevalence Estimates: Comparative Study of Similar Phenotyping Algorithms Using Electronic Health Record Data (e79653)
Muchiri Wandai, Katie Allen, Ashley Wiensch, John Price, Brian Dixon. . . . 671

Assessing Usefulness of the Dashboard Instrument to Review Equity (DIRE) Checklist to Evaluate Equity in Public Health Dashboards: Reliability Study (e71094)
Paulina Sosa, Emir Syailendra, Harold Lehmann, Hadi Kharrazi. . . . 682

Nonrandom Missingness in Child Race and Ethnicity Records and the US Federal Data Standards: Pooled Analysis of Community-Based Child Health Studies (e65660)
Danielle Krobath, Norbert Wilson, Adolfo Cuevas, Elena Naumova, Jennifer Sacheck, Sydney Tyler, Christina Economos. . . . 698

Identifying Predictors of Cervical Cancer Screening Uptake in Sub-Saharan Africa Using Machine Learning: Cross-Sectional Study (e71677)
Nebebe Baykemagn, Mekuriaw Aweke, Amare Mesfin, Lemlem Baffa, Muluken Agimas, Habtamu Abuhay, Dagnew Adugna, Tewodros Alemu, Alemu Bicha, Gebrie Getu Alemu. . . . 714

Dietary Risk Factors and Associated Disease Burden Among Chinese Adults Aged 25 Years and Older: Systematic Analysis of the Global Burden of Disease Study 2021 (e72978)
Yuze Xin, Dong Shui, Guangcan Yan, Wei Tian, Ning Tang, Jinyu Liang, Junyi Peng, Hongru Sun, Anqi Ge, Xinyan Liu, Katrina Kissock, Kathy Trieu, Jing Zhang, Xinyi Zhang, Pengpeng Ye, Maoyi Tian. . . . 729

Characterizing US Spatial Connectivity and Implications for Geographical Disease Dynamics and Metapopulation Modeling: Longitudinal Observational Study (e64914)
Giulia Pullano, Lucila Alvarez-Zuzek, Vittoria Colizza, Shweta Bansal. . . . 742

Behavioral and Demographic Profiles of HIV Transmission and Exposure Networks in Florida: Network Analysis of HIV Contact Tracing Data (e65573)
Yiyang Liu, Christina Parisi, Chaoyue Sun, Rebecca Fisk-Hoffman, Marco Salemi, Diego Viteri, Brandi Danforth, Mattia Prosperi, Simone Marini. . . . 752

COVID-19–Relevant Insights Into the Elevated Risk of Accidental Injuries in Survivors of SARS and Their Relatives in Taiwan: Retrospective Cohort Study (e70608)
Chieh Sung, Chi-Hsiang Chung, Chien-An Sun, Chang-Huei Tsao, Daphne Ng, Tsu-Hsuan Weng, Li-Yun Fann, Fu-Huang Lin, Wu-Chien Chien. . . . 766

An Explainable Artificial Intelligence Text Classifier for Suicidality Prediction in Youth Crisis Text Line Users: Development and Validation Study (e63809)
Julia Thomas, Antonia Lucht, Jacob Segler, Richard Wundrack, Marcel Miché, Roselind Lieb, Lars Kuchinke, Gunther Meinlschmidt. . . . 781

A Cloud-Based Platform for Harmonized COVID-19 Data: Design and Implementation of the Rapid Acceleration of Diagnostics (RADx) Data Hub (e72677)
Marcos Martínez-Romero, Matthew Horridge, Nilesh Mistry, Aubrie Weyhmiller, Jimmy Yu, Alissa Fujimoto, Aria Henry, Martin O'Connor, Ashley Sier, Stephanie Suber, Mete Akdogan, Yan Cao, Somu Valliappan, Joanna Mieczkowska, Ashok Krishnamurthy, Michael Keller, Mark Musen, RADx Data Hub Team. . . . 799

XSL•FO
RenderX

Prevalence, Risk Factors, Disease-Related Knowledge, and Vaccination Attitudes and Behaviors for Long COVID Among French Civil Servants: Cross-Sectional Survey (e83323)
Florence Carrouel, Virginie-Eve Lvovschi, Benjamin du Sartz de Vigneulles, Maryem Rhanoui, Roger Salamon, Michel Lamure, Corélie Salque, Romain Lan, Claude Dussart. . . . 821

Using the Health Belief Model to Examine Parental Knowledge and Health Beliefs About Human Papilloma Virus (HPV) and iHPV Vaccine in Kuwait: Cross-Sectional Survey Study (e75818)
Ahmad Abuzoor, Md Jabin, Cyril Eshareturi, Rae Nesbitt, Cor Jonker. . . . 839

Using Social Media Marketing to Improve Retention of Children in the Special Supplemental Nutrition Program for Women, Infants, and Children: Implementation Study (e77172)
Eriko Robinson, Carla Sabugo, Carolina Crisafi, Susan Gross, Yunhee Kang, Elisabet Eppes, Marycatherine Augustyn, Karen Castellanos-Brown, David Paige, Laura Caulfield. . . . 860

Clustering of Lifestyle Behaviors and Their Association With Risk of Metabolic Syndrome Among Adults in Taiwan: Nationwide Cross-Sectional Study (e73114)
Ya-Hui Chang, Chung-Yi Li, Hon-Ping Ma, Chien-Yuan Wu, Yann-Yuh Jou, Chiachi Lee. . . . 874

Temporal Trends in Cervical Human Papillomavirus Prevalence Among Females in Xiamen, China (2016-2023): Cross-Sectional Study (e70507)
Zhuju Chen, Erjuan Lin, Mo Chen, Shihan Wang, Qiuyuan Lin, Qingquan Chen, Heng Xue, Guanbin Zhang. . . . 885

Development of Machine Learning–Based Risk Prediction Models to Predict Rapid Weight Gain in Infants: Analysis of Seven Cohorts (e69220)
Miaobing Zheng, Yuxin Zhang, Rachel Laws, Peter Vuillermin, Jodie Dodd, Li Wen, Louise Baur, Rachael Taylor, Rebecca Byrne, Anne-Louise Ponsonby, Kylie Hesketh. . . . 897

Prevalence of Chronic Back Pain and Associated Factors in Children and Adolescents: Secondary Analysis of the 2001–2019 Health Behavior in School-Aged Children Study (e67960)
Camila Miño, Héctor Gutiérrez-Espinoza, Jorge Olivares-Arancibia, Rodrigo Yáñez-Sepúlveda, Daniel Duclos-Bastías, Felipe Araya-Quintanilla, Lee Smith, José López-Gil. . . . 911

How Healthy and Unhealthy Lifestyle Behaviors Affect Cognitive Function—Evidence From Older Adults in Chinese Communities: Cross-Sectional Study (e73398)
Huixiu Hu, Yajie Zhao, Yuqing Hao, Huanhuan Luo, Lanying Xie, Chao Sun. . . . 925

Tenofovir to Prevent HIV Infection in Western China: Pragmatic Randomized Controlled Trial (e71494)
Yi Tao, Yan Zhang, Bin Peng, Aizhong Zeng, Jianghong Dai, Hao Liang, Juying Zhang, Huarui Shao, Shihan Feng, Xiaoni Zhong, Ailong Huang. . . . 938

Knowledge, Attitudes, and Practices to Periodontal Health of the Northeast Chinese Public: Cross-Sectional Study (e72069)
Chengde Jin, Ming Chi, Chunxiao Li, Yushan Li, Wanting Wang, Yushi Zhang, Qiao Liu. . . . 954

Gender-Specific Moderating Role of Physical Fitness and Cardiovascular Risk Factors in the Relationship Between BMI and C-Reactive Protein: Cross-Sectional Study (e76485)
Chunyuan Lai, Yunqing Liu, Yue Yan, Boli Cheng, Yakun Liu, Fanshu Yu, Chengyuan Hu, Liqiang Su, Fanghui Li. . . . 965

Epidemiological Features, Clinical Symptoms, and Environmental Risk Factors for Notifiable Japanese Encephalitis in Taiwan From 2008 to 2020: Retrospective Study (e63053)
Fu-Huang Lin, Yu-Ching Chou, Chi-Jeng Hsieh, Chia-Peng Yu. . . . 985

XSL•FO
RenderX

Effectiveness of a Health Education Program to Reduce Recurrence of Stroke by Controlling Modifiable Risk Factors in a Specialized Hospital in Bangladesh: Randomized Controlled Trial (e72233)
Mahabuba Afrin, K Huq, Sharif Khan, Subir Das, Mohammad Chowdhury, Yasuko Fukuoka, Yasuko Fukushima, Michiko Moriyama. 1001

Impact of Commercial Food Environments on Local Type 2 Diabetes Burden: Cross-Sectional and Ecological Multimodeling Study (e70045)
Kurubaran Ganasegeran, Mohd Abdul Manaf, Lance Waller, Nazarudin Safian, Muhammad Mohd Rizal, Wye Chiew, Feisul Mustapha, Khairul Abdul Maulud. 1015

Behavioral Profiles and Sociodemographic Predictors of Planetary Health Diet Engagement Among Health Care Professionals to Inform Public Health Promotion: Cross-Sectional Study (e67633)
Florence Carrouel, Rita Nugem, Laurie Fraticelli, Yohan Fayet, Lama Basbous, Corélie Salque, Claude Dussart, Marie-Thérèse Charreyre, Romain Lan, Audrey Murat-ringot. 1035

Assessing COVID-19 Mortality in Serbia's Capital: Model-Based Analysis of Excess Deaths (e56877)
Dane Cvijanovic, Nikola Grubor, Nina Rajovic, Mira Vucevic, Svetlana Miltenovic, Marija Laban, Tatjana Mostic, Radica Tasic, Bojana Matejic, Natasa Milic. 1049

Digital Survey–Based Tracing of COVID-19 Over the Early Pandemic: Comprehensive Geospatial and Symptomatic Analysis in Lebanon (e80331)
Youssef Bassim, Abir Abdelrahman, Amal Iaaly, Gavin Douglas, Patrick Daou, Mayda Finianos, Rim Hassan, Ibrahim Nassif, Layal Greige, Liza Dib, Jean Mardirossian, Mira El Chaar. 1061

Social Media Use, Influencer Status, and Outdoor Risk-Taking in Australian Adults: Cross-Sectional Survey (e73089)
Samuel Cornell, Amy Peden. 1101

Assessing the Dissemination of Federal Risk Communication by News Media Outlets During Enteric Illness Outbreaks: Canadian Content Analysis (e68724)
Hisba Shereefdeen, Lauren Grant, Vayshali Patel, Melissa MacKay, Andrew Papadopoulos, Leslie Cheng, Melissa Phypers, Jennifer McWhirter. 1109

Mindfulness Intervention for Health Information Avoidance in Older Adults: Mixed Methods Study (e69554)
Chenyu Gu, Liquan Qian, Xiaojie Zhuo. 1122

Alternative Presentations of Overall and Statistical Uncertainty for Adults' Understanding of the Results of a Randomized Trial of a Public Health Intervention: Parallel Web-Based Randomized Trials (e62828)
Christine Holst, Steven Woloshin, Andrew Oxman, Christopher Rose, Sarah Rosenbaum, Heather Munthe-Kaas. 1138

Evaluation of a Digital Media Campaign to Promote Knowledge and Awareness of the GPFirst Program for Nonurgent Conditions: Repeated Survey Study (e66062)
Rebecca Ong Hui Shan, Hong Oh, Priscilla Goh Sook Kheng, Lyndia Lee Sze Hui, Mas Riza Bte Mohd Razali, Edris Ahmad, Jagadesan Raghuram, Choon How, Steven Lim Hoon Chin. 1158

Validation of the Perceived Barriers to Antiretroviral Therapy Adherence (PEDIA) Scale Among Gay, Bisexual, and Other Men Who Have Sex With Men and Transgender and Nonbinary Persons: Cross-Sectional Study (e67005)
Victor Matos, Dyego Araújo, Thiago Torres, Paula Luz. 1190

Intention to Use Pre-Exposure Prophylaxis Among Age Groups of Brazilian Men Who Have Sex With Men: National Cross-Sectional Study (e58405)
Alvaro Sousa, Caique Ribeiro, Guilherme Santos, Emerson Camargo, Herica Carvalho, Guilherme Schneider, Leticia Vieira, Liliane Carneiro, Ana Fernandes, Talita Fernandes, Márcio Bezerra-Santos, Rita Nascimento, Lucas Andrade, Anderson Sousa, Inês Fronteira, Isabel Mendes.
1 2 1 0

XSL•FO
RenderX

Impact of Knowledge Access on Risky Sexual Behaviors Among Chinese Youths to Improve HIV Prevention: Cross-Sectional Study (e68339)
Yun Zhang, Jie Jin, Feng Cheng, Xingliang Zhang, Junfang Xu. 1227

HIV Incidence and Associated Risk Factors Among Young Men Who Have Sex With Men in Tianjin, China: Retrospective Cohort Study (e66487)
Yan Guo, Hui Gong, Xiangyu Yan, Xinying Zhang, Tielin Ning, Zhongquan Liu, Long Li, Jie Yang, Liu Li, Changxu Ma, Zhuang Cui, Maohe Yu. 1241

Impact of Disclosing to Patients the Use of Antiretroviral Resistance Testing Results for Molecular HIV Surveillance: A Randomized Experiment in 2 National Surveys (e64663)
Jeremy Sugarman, Juli Bollinger, Jose Agostini, Kevin Weinfurt, Gail Geller, Sheethal Jose, Marissa Hannah, O. Edwards, Leslie Henry, Travis Sanchez. 1255

The Effect of a Brief Video-Based Intervention to Improve AIDS Prevention in Older Men: Randomized Controlled Trial (e65674)
Tian Huaju, Xie Rendie, Xiao Lu, Li Mei, Luo Yue, Zhang Daiying, Chen Yanhua, Ren Jianlan. 1264

Awareness of Sexual Partner's HIV Status Among Men Who Have Sex With Men in China: Cross-Sec. tional Survey Study (e66139)
Qian Ma, Tingting Jiang, Wanjun Chen, Shaoqiang Jiang, Jinlei Zheng, Hui Wang, Lin He. 1282

The Effects of MyChoices and LYNX Mobile Apps on HIV Testing and Pre-Exposure Prophylaxis Use by Young US Sexual Minority Men: Results From a National Randomized Controlled Trial (e63428)
Katie Biello, Kenneth Mayer, Hyman Scott, Pablo Valente, Jonathan Hill-Rorie, Susan Buchbinder, Lucinda Ackah-Toffey, Patrick Sullivan, Lisa Hightow-Weidman, Albert Liu. 1295

Perceptions of Long-Acting Injectable Pre-Exposure Prophylaxis Among Men Who Have Sex With Men and Transgender Individuals in Europe Using Structural Text Modeling Technique: Cross-Sectional Study (e72491)
Haoyi Wang, Johann Kolstee, Julio Martinez, David van de Vijver, Jonathan Tosh, Melanie Schroeder, Ama Appiah, Hanne Zimmermann, Kai Jonas. 1344

Secondary Distribution of Dual HIV/Syphilis Self-Testing Among Men Who Have Sex With Men: Pragmatic Randomized Controlled Trial in China (e70775)
Min Wang, Yun Zhang, Weiyi Tian, Jinli Mo, Haimei Huang, Jiawen Zhu, Sumin Tan, Yingqiong Huang, Li Jiang, Ping Cen, Guanghua Lan, Hao Wang, Wei Pan, Joseph Tucker, Chuanyi Ning. 1358

Effects of Maternal Use of the Continuum of Care on Complementary Feeding Practices in Bangladesh: Cross-Sectional Study (e76666)
M Rifat, Rokibul Islam, Rinath Didar, Joya Bhowmick, Plabon Sarkar, Md Amin, Sanjib Saha. 1371

Use of Cigarettes, Cannabis, and Alcohol Among Asian American, Native Hawaiian, and Pacific Islander Adults: Community-Based National Survey Analysis (e76465)
Vuong Do, Van Ta Park, Nhung Nguyen, Pamela Ling, Marian Tzuang, Bora Nam, Marcelle Dougan, Oanh Meyer, Janice Tsoh. 1383

Association Between Social Distancing Compliance and Public Place Crowding During the COVID-19 Pandemic: Cross-Sectional Observational Study Using Computer Vision to Analyze Surveillance Footage (e50929)
Lasse Liebst, Wim Bernasco, Peter Ejbye-Ernst, Nigel van Herwijnen, Thomas van der Veen, Dennis Koelma, Cees Snoek, Marie Lindegaard.
1 3 9 8

Diagnostic Testing Preferences in Rural and Vulnerable Populations During a Pandemic: Discrete Choice Experiment (e68734)
Eline van den Broek-Altenburg, Jamie Benson, Yvonne Jonk, Abimbola Leslie, Jan Carney, Gary Stein. 1407

XSL•FO
RenderX

Mapping Key Populations to Develop Improved HIV and AIDS Interventions: Multiphase Cross-Sectional Observational Mapping Study Using a District and City Approach (e56820)
Pande Januraga, Endang Lukitosari, Lanny Luhukay, Rizky Hasby, Aang Sutrisna. 1420

Trends in Prescription of Stimulants and Narcoleptic Drugs in Switzerland: Longitudinal Health Insurance Claims Analysis for the Years 2014-2021 (e53957)
Tamara Scharf, Carola Huber, Markus Näpflin, Zhongxing Zhang, Ramin Khatami. 1488

Assessment of Bidirectional and Threshold-Response Associations Between Cognitive Function and Physical Performance: Nationwide Cross-Sectional Study (e80575)
Huixiu Hu, Lanying Xie, Yuqing Hao, Yajie Zhao, Huanhuan Luo, Kang Yu, Chao Sun. 1500

Knowledge, Perception, and Attitude Toward Voluntary Counseling and Testing for HIV Among Secondary School Students in Iringa Rural District: Descriptive Cross-Sectional Study (e66739)
Maiko Mkwambe, Shakilu Kayungo, Youping Deng. 1516

Severity Classification of Anxiety and Depression Using Generalized Anxiety Disorder Scale and Patient Health Questionnaire: National Cross-Sectional Study Applying Classification and Regression Tree Models (e72591)
Andre Faro, Julian Tejada, Wael Al-Delaimy. 1528

Trends in Avoidable Hospitalizations Before and During the COVID-19 Pandemic: Multiple Cross-Sectional Study Using Administrative Data From Beijing, China (e69768)
Xiangzhen Wang, Yin Chen, Yuqi Ta, Moning Guo, Hongqiao Fu. 1545

Factors Influencing the Maintenance of Public Health Behaviors After an Epidemic: Cross-Sectional Study (e66535)
Xingmin Wang, Hongli Yan, Lushaobo Shi, Ting Li, Yi Xia, Dong Wang. 1560

Impact of Perceived Unfairness, Uncertainty, and Life Stress on Urban Residents' Health in China: Cross-Sectional Study (e72094)
Jinsong Chen, Shuhan Jiang, Mingli Pang, Huan Zhou, Weifang Zhang. 1575

Knowledge and Beliefs Toward Mammography Screening Among Jordanian Women: Cross-Sectional Study (e75384)
Ahmad Abuabed, Hana Taha, Luz Garcia-Valdes, Mohammad Al-Share, Hadi Al-shar'e, Carmen Amezcua-Prieto. 1591

Health Indicators as Measures of Individual Health Status, Their Perceived Importance, and Associated Factors: Cross-Sectional Observational Study (e65616)
Temiloluwa Sokoya, Yuchun Zhou, Sebastian Diaz, Timothy Law, Lina Himawan, Francisca Lekey, Lu Shi, Sarah Griffin, Ronald Gimbel, Xia Jing. 1606

Migrant-Local Differences in the Relationship Between Oral Health, Social Support, and Loneliness Among Older Adults in Weifang, China: Cross-Sectional Study (e66061)
Hui Liu, Jieru Wang, Rui Chen, Xixing Xu, Mingli Pang, Kaiyuan Feng, Bingsong Li, Qinling Li, Ziwei Qin, Shuyi Yan, Nabila Ibn Ziyat, Fanlei Kong. 1622

Multidimensional Evaluation of the Process of Constructing Age-Friendly Communities Among Different Aged Community Residents in Beijing, China: Cross-Sectional Questionnaire Study (e66248)
Yingchun Peng, Zhiying Zhang, Ruyi Zhang, Yiyao Zhang, Runying Wang, Jiaying Zhang, Shaoqi Zhai, Qilin Jin, Jiaojiao Zhou, Jingjing Chen.
1 6 3 6

Long-Term Exposure to Fine Particulate Matter (PM2.5) Components and Precocious Puberty Among School-Aged Children: Cross-Sectional Study (e62861)
Xuelian Zhou, Xiaochi Zhang, Guannan Bai, Guanping Dong, Xinyi Li, Ruimin Chen, Shaoke Chen, Rongxiu Zheng, Chunlin Wang, Haiyan Wei, Bingyan Cao, Yan Liang, Hui Yao, Zhe Su, Mireguli Maimaiti, Feihong Luo, Pin Li, Min Zhu, Hongwei Du, Yu Yang, Lanwei Cui, Jinling Wang, Jinna Yuan, Zhuang Liu, Wei Wu, Qi Zhao, Junfen Fu. 1650

XSL•FO
RenderX

Interpreting the Influence of Using Blood Donor Residual Samples for SARS-CoV-2 Seroprevalence Studies in Japan: Cross-Sectional Survey Study (e60467)
Ryo Kinoshita, Sho Miyamoto, Tadaki Suzuki, Motoi Suzuki, Daisuke Yoneoka. 1663

Association Among BMI, Self-Esteem, and Nonsuicidal Self-Injury in Young Adults to Understand the Influence of Socioenvironmental Factors: Longitudinal Study (e52928)
Yi Zhang, Ruixue Ying, Wan Lu, Xuemeng Liu, Keyan Hu, Qing Feng, Zixiang Yu, Zhen Wang, Fangting Lu, Yahu Miao, Nanzhen Ma, Fangbiao Tao, Tian Jiang, Qiu Zhang. 1671

Mpox Prevention Self-Efficacy and Associated Factors Among Men Who Have Sex With Men in China: Large Cross-Sectional Study (e68400)
Quyige Gao, Shangbin Liu, Muzaibaier Tuerxunjiang, Huifang Xu, Jiechen Zhang, Gang Xu, Jianyu Chen, Yong Cai, Fan Hu, Ying Wang. . . .
1 6 8 5

Identifying Preferred Features of Influenza Vaccination Programs Among Chinese Clinicians Practicing Traditional Chinese Medicine and Western Medicine: Discrete Choice Experiment (e63314)
Liuren Zhang, Linchen Chu, Maria Sundaram, Yi Zhou, Xiu Sun, Zheng Wei, Chuanxi Fu. 1699

Attitudes Toward Psychotherapeutic Treatment and Health Literacy in a Large Sample of the General Population in Germany: Cross-Sectional Study (e67078)
Rebekka Schröder, Tim Hamer, Ralf Suhr, Lars König. 1712

Nationwide Trends in Screen Time and Associated Risk Factors by Family Structures Among Adolescents, 2008-2022: Nationwide Cross-Sectional Study (e57962)
Seokjun Kim, Hyesu Jo, Yejun Son, Min Shin, Kyeongmin Lee, Jaeyu Park, Hayeon Lee, Lee Smith, Elena Dragioti, Guillaume Fond, Laurent Boyer, Guillermo López Sánchez, Mark Tully, Masoud Rahmati, Damiano Pizzol, Selin Woo, Dong Yon. 1728

How Information Exposure Shapes Risk Perceptions and Vaccination Intentions Among Gay, Bisexual, and Other Men Who Have Sex With Men: Cross-Sectional Survey Study (e70635)
Doug Cheung, Luyao Xie, Lijuan Wang, Siyu Chen, Xinge Li, Zheng Zhang, Xinyue Chen, Shen Ge, Fuk-yuen Yu, Yuan Fang, Zihuang Chen, Zhennan Li, Fenghua Sun, Phoenix Mo, Yingjie Liu, Zixin Wang. 1743

Determinants of Health-Promoting Behaviors Among Indonesian Adolescents Living in Child Welfare Institutions: Cross-Sectional Survey Study (e75024)
Aloysia Ispriantari, Hyejung Lee, Sue Kim, Hyeonkyeong Lee, Anna Lee, Chang Park. 1764

Ultraprocessed Foods Associated With Depressive Symptoms Among Chinese Adolescents: Cross-Sectional Study (e75061)
Honglv Xu, Jieru Yang, Dongyue Hu, Xiaoxiao Li, Feihui Zhou. 1776

Associations Between Both HIV and Metabolic Comorbidity and Self-Reported Mpox Among Men Who Have Sex With Men: Multicenter Cross-Sectional Study (e83450)
Yujie Liu, Xin Ge, Yinghuan Zhang, Yuxuan Wang, Meihui Zhang, Jianyu Chen, Gang Xu, Jiechen Zhang, Ying Wang, Yong Cai. 1793

Association of Frailty Index at 66 Years of Age with Health Care Costs and Utilization Over 10 Years in Korea: Retrospective Cohort Study (e50026)
Jieun Jang, Anna Kim, Mingee Choi, Ellen McCarthy, Brianne Olivieri-Mui, Chan Park, Jae-Hyun Kim, Jaeyong Shin, Dae Kim. 1805

Impact of the National Essential Public Health Service Package on Blood Pressure Control in Chinese People With Hypertension: Retrospective Population-Based Longitudinal Study (e65783)
Feiran Wei, You Ge, Han Li, Yuan Liu. 1815

XSL•FO
RenderX

Incidence of Deliberate Self-Harm in Hong Kong Before and During the COVID-19 Pandemic: Population-Wide Retrospective Cohort Study (e57500)
Luke Luk, Christie Ching, Tsz Yip, Sunny Chan, Catherine Lam, Elizabeth Lam, Serena Yue, Hoi Pang, Janet Wong, Carlos Wong, Chak Tong, Tafu Yamamoto, Timothy Rainer, Abraham Wai, Joshua Ho. . . . 1832

Investigating the Trajectories of Poor Vision in Children and Adolescents in Wuhan, China From 2016 to 2019: Prospective Cohort Study (e53028)
Lijuan Xu, Hanjia Li, Fang Li, Tinghui Zhang, Jingyan Yan, Hong Yan, Lu He, Bin Yu. . . . 1846

Identifying Data-Driven Clinical Subgroups for Cervical Cancer Prevention With Machine Learning: Population-Based, External, and Diagnostic Validation Study (e67840)
Zhen Lu, Binhua Dong, Hongning Cai, Tian Tian, Junfeng Wang, Leiwen Fu, Bingyi Wang, Weijie Zhang, Shaomei Lin, Xunyuan Tuo, Juntao Wang, Tianjie Yang, Xinxin Huang, Zheng Zheng, Huifeng Xue, Shuxia Xu, Siyang Liu, Pengming Sun, Huachun Zou. . . . 1858

Survival Tree Analysis of Interactions Among Factors Associated With Colorectal Cancer Risk in Patients With Type 2 Diabetes: Retrospective Cohort Study (e62756)
Sarah Yau, Chi Hung, Eman Leung, Albert Lee, Eng Yeoh. . . . 1879

Area-Level Indices and Health Care Use in a Pediatric Brain and Central Nervous System Tumor Cohort: Observational Study (e66834)
Yvette Tran, Seho Park, Scott Coven, Eneida Mendonca. . . . 1894

Peer Relationships Are a Direct Cause of the Adolescent Mental Health Crisis: Interpretable Machine Learning Analysis of 2 Large Cohort Studies (e60125)
Heiner Stuke, Robert Schlack, Michael Erhart, Anne Kaman, Ulrike Ravens-Sieberer, Christopher Irrgang. . . . 1906

Validation and Refinement of Scores to Predict Stroke Risk: Prospective Cohort Study (e72497)
Hua Meng, Zhuo Liu, Dongfeng Pan, Xinya Su, Wenwen Lu, Xingtian Wang, Yuhui Geng, Xiaojuan Ma, Peifeng Liang. . . . 1921

Catch-Up of Routine School-Based Immunizations Since the COVID-19 Pandemic: 4-Year Observational Cohort Study (e79669)
Kelvin Nartey, Laura Reifferscheid, Marilou Kiely, Shannon MacDonald, Canadian Immunization Research Network. . . . 1931

The Association Between Posting WeChat Moments and the Risk of Depressive Symptoms Among Middle-Aged and Older Chinese Adults: Prospective National Cohort Study (e62730)
Wei Wang, Hui Wang, Xinru Hu, Qian Yu, Fangyi Chen, Xirui Qiu, Xiaoxiao Wang. . . . 1942

Identifying Community-Built Environment's Effect on Physical Activity and Depressive Symptoms Trajectories Among Middle-aged and Older Adults: Chinese National Longitudinal Study (e64564)
Kaili Zhang, Bowen Huang, Prasanna Divigalpitiya. . . . 1956

Factors Affecting Nonresponse Among Female Participants in the Korea Nurses' Health Study: Longitudinal Cohort Survey Study (e68038)
Young Kim, Chiyoung Cha, Gumhee Baek, Bohye Kim, Bo Song, Joong-Yeon Lim, Hyun-Young Park, Juh Shin. . . . 1972

Health Impact of Suspected Interpersonal Violence Against Adults: Retrospective Cohort Study (e65191)
Rui Barbosa, Rita Lopes, Tiago Taveira-Gomes, Carla Ponte, Teresa Magalhães. . . . 1983

Spatial Clusters of Condyloma Acuminata and the Regional Risk Factors in South Korea: Bayesian Spatial Regression Analysis (e76751)
Joonsu Jang, Seyul Park, Byung Chun. . . . 1996

Hypertension Prevalence Among People Lifted Out of Poverty in China in 2018-2023: Retrospective Spatiotemporal Analysis (e66501)
Ying Zhang, Dong Xia, Zhiyu Lv, Chennan Wu, Jiapeng Chen, Lulu Zhang. . . . 2010

XSL•FO
RenderX

Distribution and Risk Factors of Scrub Typhus in South Korea, From 2013 to 2019: Bayesian Spatiotemporal Analysis (e68437)
Jeehyun Kim, Penelope Vounatsou, Byung Chun. 2024

Identifying the Factors Associated With Spatial Clustering of Incident HIV Infection Cases in High-Prevalence Regions: Quantitative Geospatial Study (e75291)
Qiyu Zhu, Chunnong Jike, Chengdong Xu, Shu Liang, Gang Yu, Dan Yuan, Hong Mai, Yiping Li, Lin Xiao, Ju Wang, Hong Yang, Fengshun Yuan, Jing Hong, Muga Mao, Maogang Shen, Jing Liu, Lin He, Yuehua Wang, Huanyi Cheng, Peng Guan, Yan Jiang, Mengjie Han, Cong Jin, Zhongfu Liu. 2040

Long-Term Ambient Air Pollution Exposure and the Risk of Cardiovascular and Cerebrovascular Diseases in Rural Chinese Populations: 10-Year Follow-Up of a Multicenter Prospective Cohort Study (e81218)
Yaqi Zhao, Xuefang Cao, Jiang Du, Aiwei He, Jun Liang, Weitao Duan, Yuanzhi Di, Yijun He, Boxuan Feng, Linyu Shen, Juanjuan Huang, Zihan Li, Jianguo Liang, Hongzhi Li, Zisen Liu, Fang Liu, Shumin Yang, Zuhui Xu, Bin Zhang, Jiaoxia Yan, Yanchun Liang, Rong Liu, Fei Shen, Qi Jin, Henan Xin, Lei Gao. 2052

Data Parameters From Participatory Surveillance Systems in Human, Animal, and Environmental Health From Around the Globe: Descriptive Analysis (e55356)
Carrie McNeil, Nomita Divi, Charles Bargeron IV, Andrea Capobianco Dondona, Kacey Ernst, Angela Gupta, Olukayode Fasominu, Lucy Keatts, Terra Kelly, Onicio Leal Neto, May Lwin, Mvuyo Makhasi, Eric Mutagahywa, Diego Montecino-Latorre, Sarah Olson, Pranav Pandit, Daniela Paolotti, Matt Parker, Muhammad Samad, Kara Sewalk, Anita Sheldenkar, Lertrak Srikitjakarn, Channé Suy Lan, Michael Wilkes, Terdsak Yano, Mark Smolinski. 2067

Quality of Life Among Family Caregivers of Individuals With Rare Diseases: Web-Based Population Study on the Validity and Reliability of the Polish World Health Organization Quality of Life-BREF Questionnaire (e72590)
Piotr Jabkowski, Jan Domaradzki, Dariusz Walkowiak. 2081

Estimating the Population Size of People Who Inject Drugs in 3 Cities in Zambia: Capture-Recapture, Successive Sampling, and Bayesian Consensus Estimation Methods (e66551)
Lauren Parmley, Giles Reid, Joyce Neal, Brave Hanunka, Leigh Tally, Lophina Chilukutu, Tepa Nkumbula, Chipili Mulemfwe, Lazarous Chelu, Ray Handema, John Mwale, Kennedy Mutale, Lloyd Mulenga, Anne McIntyre, Neena Philip, Hannah Chung, Maria Lahuerta. 2099

Small-Area Lung Cancer Incidence and Mortality: Cross-Sectional Population-Based Study Using Hospital Discharge and Death Registration Data (e74062)
Yu He, Xinxin Xia, Qing Wang, Yaoyun Zhang, Ying Meng, Xiaokang Ji, Qingbo Zhao, Yongchao Wang, Yifu Zhao, Fuzhong Xue, Jin Xu. . . .
2 1 1 4

Correcting for the Inflated Adult Population Denominator in an English Nationwide Health Care Cohort: Database Analysis Study (e64788)
Sudhir Venkatesan, Mark Joy, Gavin Jamie, Debasish Kar, Robert Williams, Xuejuan Fan, Wilhelmine Meeraus, Ruby Tsang, Kathryn Taylor, Sylvia Taylor, F Hobbs, Sneha Anand, Chris Robertson, Simon de Lusignan. 2126

Population Size Estimation of Men Who Have Sex With Men in Low- and Middle-Income Countries: Google Trends Analysis (e58630)
Carly Malburg, Steve Gutreuter, Horacio Ruiseñor-Escudero, Abu Abdul-Quader, Wolfgang Hladik. 2138

Prevalence and Economic Impact of Acute Respiratory Failure in the Prehospital Emergency Medical Service of the Madrid Community: Retrospective Cohort Study (e66179)
Ana Cintora-Sanz, Cristina Horrillo-García, Víctor Quesada-Cubo, Ana Pérez-Alonso, Alicia Gutiérrez-Misis. 2155

Trends in Early-Onset Colorectal Cancer in Singapore: Epidemiological Study of a Multiethnic Population (e62835)
Hui Chen, Qingqing Chong, Brenda Tay, Siqin Zhou, Evelyn Wong, Isaac Seow-En, Ker Tan, Yi Wang, Adeline Seow, Kwong-Wei Tan, Bee Tan, Sze Tan. 2169

XSL•FO
RenderX

Building and Developing a Tool (PANDEM-2 Dashboard) to Strengthen Pandemic Management: Participatory Design Study (e52119)
Carlos Tighe, Lem Ngongalah, Alexis Sentís, Francisco Orchard, Gheorghe-Aurel Pacurar, Conor Hayes, Jessica Hayes, Adrian Toader, Máire Connolly. . . . 2224

Suspected Mpox Symptoms and Testing in Men Who Have Sex With Men in the United States: Cross-Sectional Study (e57399)
Kaitlyn Atkins, Thomas Carpino, Amrita Rao, Travis Sanchez, O Edwards, Marissa Hannah, Patrick Sullivan, Yasmin Ogale, Winston Abara, Kevin Delaney, Stefan Baral. . . . 2243

Effect of a Short, Animated Storytelling Video on Transphobia Among US Parents: Randomized Controlled Trial (e66496)
Doron Amsalem, Merlin Greuel, Shuyan Liu, Andrés Martin, Maya Adam. . . . 2255

Exploring Social-Ecological Pathways From Sexual Identity to Sleep Among Chinese Women: Structural Equation Modeling Analysis (e53549)
Chanchan Wu, Pui Chau, Edmond Choi. . . . 2269

Comparing Multiple Imputation Methods to Address Missing Patient Demographics in Immunization Information Systems: Retrospective Cohort Study (e73916)
Sara Brown, Ousswa Kudia, Kaye Kleine, Bryndan Kidd, Robert Wines, Nathanael Meckes. . . . 2285

A “Pandemic-Proof” Methodology for Outbreak Detection Adapted From COVID-19’s Impact on Notifications of Infectious Diseases in the Netherlands: Surveillance Study (e73953)
Babette van Deursen, Stijn Raven, Wolfer van den Bosch, Cornelia van Jaarsveld, Aura Timen. . . . 2298

Development and Validation of Prediction Models for Perceived and Unmet Mental Health Needs in the Canadian General Population: Model-Based Synthetic Estimation Study (e66056)
Jianli Wang, Heather Orpana, André Carrington, George Kephart, Helen-Maria Vasiliadis, Benjamin Leikin. . . . 2311

Strategies to Increase Response Rate and Reduce Nonresponse Bias in Population Health Research: Analysis of a Series of Randomized Controlled Experiments during a Large COVID-19 Study (e60022)
Christina Atchison, Nicholas Gilby, Galini Pantelidou, Sam Clemens, Kevin Pickering, Marc Chadeau-Hyam, Deborah Ashby, Wendy Barclay, Graham Cooke, Ara Darzi, Steven Riley, Christl Donnelly, Helen Ward, Paul Elliott. . . . 2324

Investigating Social Network Peer Effects on HIV Care Engagement Using a Fuzzy-Like Matching Approach: Cross-Sectional Secondary Analysis of the N2 Cohort Study (e64497)
Cho-Hee Shrader, Dustin Duncan, Redd Driver, Juan Arroyo-Flores, Makella Coudray, Raymond Moody, Yen-Tyng Chen, Britt Skaathun, Lindsay Young, Natascha del Vecchio, Kayo Fujimoto, Justin Knox, Mariano Kanamori, John Schneider. . . . 2338

Leveraging Electronic Health Records in International Humanitarian Clinics for Population Health Research: Cross-Sectional Study (e66223)
Sarah Draugelis, Jessica Hunnewell, Sam Bishop, Reena Goswami, Sean Smith, Philip Sutherland, Justin Hickman, Donald Donahue, George Yendewa, Amir Mohareb. . . . 2354

Opportunities and Challenges for an Organizational Digital Public Health Strategy in a Provincial Public Health Program in Canada: Qualitative Description of Practitioner Perspectives (e72588)
Ihoghosa Iyamu, Devon Haag, Anna Carson, Ivy Wang, Colin King, Ian Roe, Kristy Kerr, Sofia Bartlett, Geoffrey McKee, Mark Gilbert. . . . 2363

Building the Workforce’s Capacity to Support the Digital Transformation of Public Health: Environmental Scan of Training Programs for Digital Technologies in Public Health (e73088)
Ihoghosa Iyamu, Swathi Ramachandran, Hsiu-Ju Chang, Andre Kushniruk, Francisco Ibáñez-Carrasco, Catherine Worthington, Hugh Davies, Geoffrey McKee, Adalsteinn Brown, Mark Gilbert. . . . 2375

Estimating Chronic Hepatitis B Prevalence and Undiagnosed Proportion in Canada, 2007-2021: Mathematical Framework Development (e66309)
Julien Smith-Roberge, Farinaz Forouzannia, Abdullah Hamadeh, Zeny Feng, Nashira Popovic, William Wong. . . . 2389

XSL•FO
RenderX

Effectiveness of Antivirals Nirmatrelvir-Ritonavir and Molnupiravir in Viral Sepsis: Retrospective Cohort Study (e72124)
Teddy Lee, Alex Lyu, Ting Jiang, Sunny Chan, Crystal Chan, Edmond Yip, Luke Luk, Joshua Ho, Kevin So, Omar Tsui, Man Lam, Shi Lee, Tafu Yamamoto, Chak Tong, Man Wong, Eliza Wong, Abraham Wai, Timothy Rainer. 2403

Shifts in Influenza and Respiratory Syncytial Virus Infection Patterns in Korea After the COVID-19 Pandemic Resulting From Immunity Debt: Retrospective Observational Study (e68058)
Minah Park, Won Choi, Benjamin Cowling. 2413

Greater Improvements in Vaccination Outcomes Among Black Young Adults With Vaccine-Resistant Attitudes in the United States South Following a Digital Health Intervention: Latent Profile Analysis of a Randomized Control Trial (e67370)
Noah Mancuso, Jenna Michaels, Erica Browne, Allysha Maragh-Bass, Jacob Stocks, Zachary Soberano, C Bond, Ibrahim Yigit, Maria Comello, Margo Larsen, Kathryn Muessig, Audrey Pettifor, Lisa Hightow-Weidman, Henna Budhwani, Marie Stoner. 2422

Associations of Violence Against Women With Comorbid Symptoms of Depression and Anxiety Among Left-Behind Women in Rural China: Cross-Sectional Study (e72064)
Leying Hou, Zeyu Luo, Weidi Sun, Jiayao Ying, Jing Wu, Shiyi Shan, Wen Liu, Peige Song. 2436

Contribution of Travelers to Plasmodium Vivax Malaria in South West Delhi, India: Cross-Sectional Survey (e50058)
Deepali Savargaonkar, Bina Srivastava, Chander Yadav, Mrigendra Singh, Anup Anvikar, Amit Sharma, Himmat Singh, Abhinav Sinha. 2452

Potentially Inappropriate Use of Transdermal Fentanyl in Working-Age and Older Adult Populations with Non-Cancer Pain: Nationwide Cross-Sectional Study (e63960)
Seunghyun Cheon, Seo-Yong Choi, Kyu Kim, Han-Gon Choi, Jee-Eun Chung. 2463

The Impact of Point-of-Care Testing for Influenza on Antimicrobial Stewardship in UK Primary Care: Nested Cohort Study (e72322)
Uy Hoang, Jessica Smylie, Elizabeth Button, Jack Macartney, Cecilia Okusi, Rachel Byford, Filipa Ferreira, Charis Xie, Mark Joy, Tristan Clark, Simon de Lusignan. 2476

Assessment of Influenza Testing Distribution in the United States for the 2021-2022, 2022-2023, and 2023-2024 Influenza Seasons: Online Cross-Sectional Study (e76459)
Autumn Gertz, Juliana Sopko, Christopher Remmel, Benjamin Rader, John Brownstein. 2489

COVID-19 Testing Equity in New York City During the First 2 Years of the Pandemic: Demographic Analysis of Free Testing Data (e52972)
Daniel Rosenfeld, Sean Brennan, Andrew Wallach, Theodore Long, Chris Keeley, Sarah Kurien. 2501

Predictors of Loneliness and Psychological Distress in Older Adults During the COVID-19 Pandemic: National Cross-Sectional Study (e78728)
Rahela Orlandini, Antonela Matana, Deana Švaljug, Ivana Gusar, Vesna Anti evi 2509

Influenza Vaccination Coverage and Determinants of New Vaccinations During the COVID-19 Pandemic in Spain (ENE-COVID): Nationwide Population-Based Study (e60658)
Miguel de la Cámara, Nerea Fernández de Larrea-Baz, Roberto Pastor-Barriuso, Amparo Larrauri, Pablo Fernández-Navarro, Marina Pollán, Beatriz Pérez-Gómez, ENE-COVID Study Group. 2524

Meeting Global Health Needs via Infectious Disease Forecasting: Development of a Reliable Data-Driven Framework (e59971)
Ravikiran Keshavamurthy, Karl Pazdernik, Colby Ham, Samuel Dixon, Samantha Erwin, Lauren Charles. 2544

XSL•FO
RenderX

Preferences for Nonpharmaceutical Interventions During the Endemic Phase of COVID-19: Discrete Choice Experiment (e67725)
Yi Wang, Chee Har, Sharon Tan, Hooi Cheng, Ian Ang. . . . 2561

Identifying Success Factors for Optimizing COVID-19 Vaccine Uptake Among Indigenous Populations in Taiwan: Cross-Sectional Questionnaire Survey (e75278)
Jia-Chi Tu Abùs Isbabanal, Ray-E Chang, Liam O'Neill, Ting-Hung Chen, Po-Han Lee, Tsan-Teng Ou. . . . 2576

Effects of Perceived Scarcity on Mental Health, Time and Risk Preferences, and Decision-Making During and After COVID-19 Lockdown: Quasi-Natural Experimental Study (e69496)
Haiou Zhu, Fangzhou You, Thorsten Gruber, Hua Dong, Cees de Bont. . . . 2590

Consistency of Daily Number of Reported COVID-19 Cases in 191 Countries From 2020 to 2022: Comparative Analysis of 2 Major Data Sources (e65439)
Han Liu, Huiying Zong, Yang Yang, David Schwebel, Bin Xie, Peishan Ning, Zhenzhen Rao, Li Li, Guoqing Hu. . . . 2608

Association Between Triglyceride-Glucose Index, Blood Pressure Status, and Coronary Heart Disease Risk Among Chinese Adults With Disabilities: 10-Year Disability Health Survey Cohort Study (e78068)
Hui Liu, Yao Li, Yiyan Wang, Tao Wang, Chenghua Jiang, Hengjing Wu, Jing Wu. . . . 2617

Unmet Health Care Use Among Socially Withdrawn Youth (Hikikomori) in South Korea: Cross-Sectional Survey Study (e79043)
Eunjeong Choi, Seoyeong Choi, Sukyong Jang. . . . 2630

Vaccine Hesitancy and Associated Factors Among Caregivers of Children With Special Health Care Needs in the COVID-19 Era in China: Cross-Sectional Study (e67487)
Mingyan Li, Changxuan Sun, Chai Ji, Meiying Gao, Xia Wang, Dan Yao, Junxia Guo, Lidan Sun, Abdul Rafay, Antonita George, Sanduni Muhandiramge, Guannan Bai. . . . 2641

Prevalence and Associated Factors of Bullying Experienced by Community Patients With Serious Mental Disorders: Cross-Sectional Study (e75027)
Yazhuo Qi, Meiqi Wang, Zhen Wei, Wenyu Wang, Kaixian Wang, Long Sun. . . . 2655

The Association Between Cultural Tightness and COVID-19 Vaccine Confidence From 28 Countries: Cross-Sectional Study (e66872)
Qiang Wang, Ana Bolio, Leesa Lin. . . . 2668

Predictors of COVID-19 Vaccination Intention and Behavior Among Young People in a European Union Country With Low COVID-19 Vaccination Rates: Cross-Sectional Study (e64653)
Sara Atanasova, Tanja Kamin, Nina Perger. . . . 2681

SARS-CoV-2 Antibody Prevalence Across Unvaccinated Health Care Workers During the COVID-19 Pandemic in Yemen: Cross-Sectional Study (e67296)
Eihab Al-Herwi, Maryam Baras, Khadega Alhetar, Mohammed Farhan, Rashad Saleh, Amani As-suhbani, Manal Alshoaibi, Mansour Mohammed, Zainab Morshed, Asia Shujaa-Aldeen, Fatima Ghanem, Bra'ah AlQisi, Mohammed Alherwi, Tony Bruns. . . . 2698

Dissemination and Implementation Approach to Increasing Access to Local Pre-Exposure Prophylaxis (PrEP) Resources With Black Cisgender Women: Intervention Study With Vlogs Shared on Social Media (e67367)
Mandy Hill, Laurenia Mangum, Sandra Coker, Tristen Sutton, Diane Santa Maria. . . . 2733

Differences in Hemodialysis Claim Patterns Across Membership Types Among Patients With Renal Failure Based on National Health Insurance Data From 2017 to 2022: Cross-Sectional Analysis (e73731)
Aries Munandar, Syarif Hasibuan, Dian Kusuma. . . . 2763

XSL•FO
RenderX

Dental Caries Detection in Children Using Intraoral Scanners Featuring Fluorescence: Diagnostic Agreement Study (e78023)
Bree Jones, Tong Chen, Stavroula Michou, Nicky Kilpatrick, Nigel Curtis, David Burgner, Christoph Vannahme, Mihiri Silva. 2777

Characterizing Cross-Provincial High-Cost Patients in Rural China: Cross-Sectional Study (e54234)
Minjiang Guo, Xiaotong Jiang, Yang Liu, Yang Liu, Fangyuan Zhang, Zhongyuan Zhang, Yazi Li. 2791

Assessing the Impact of Distance Traveled and Birth Volumes of Hospital Maternity Units on Newborn Outcomes: Population-Based Cohort Study (e58944)
Anna Cantarutti, Riccardo Boracchini, Roberto Bellù, Raffaella Ronco, Federico Rea, Anna Locatelli, Rinaldo Zanini, Giovanni Corrao. 2804

Digital Surveillance of Mental Health Care Services in Saudi Arabia: Cross-Sectional Study of National e-Referral System Data (e64257)
Abdullah Alharbi, Nawfal Aljerian, Meshary Binhotan, Hani Alghamdi, Ali Alsultan, Mohammed Arafat, Abdulrahman Aldhabib, Yasser Alaska, Eid Alwahbi, Mohammed Muaddi, Ahmad Alqassim, Ronnie Horner. 2818

Viewpoints

Microorganisms Linked to Health Care–Associated Infections: Modernization of Terminology Resources for Reporting to the National Healthcare Safety Network (e66088)
Sheila Abner, Shuai Zheng, Caroline Macumber, Shaun Shakib, Jonathan Edwards, Andrea Benin. 53

Machine Learning Applications in Population and Public Health: Guidelines for Development, Testing, and Implementation (e68952)
Andrew Pinto, Sharon Birdi, Steve Durant, Roxana Rabet, Rahul Parekh, Shehzad Ali, David Buckeridge, Marzyeh Ghassemi, Jennifer Gibson, Ava John-Baptiste, Jillian Macklin, Melissa McCradden, Kwame McKenzie, Parisa Naraei, Akwasi Owusu-Bempah, Laura Rosella, James Shaw, Ross Upshur, Sharmistha Mishra. 60

Ethical Considerations for Wastewater Surveillance Conducted by the US Department of Defense (e67145)
Hunter Smith, Richard Agans, William Kowallis. 71

A Call for Action: Lessons Learned From a Pilot to Share a Complex, Linked COVID-19 Cohort Dataset for Open Science (e63996)
Clara Amid, Martine van Roode, Gabriele Rinck, Janko van Beek, Rory de Vries, Gijsbert van Nierop, Eric van Gorp, Frank Tobian, Bas Oude Munnink, Reina Sikkema, Thomas Jaenisch, Guy Cochrane, Marion Koopmans. 79

The Promise and Perils of Artificial Intelligence in Advancing Participatory Science and Health Equity in Public Health (e65699)
Abby King, Zakaria Doueiri, Ankita Kaulberg, Lisa Goldman Rosas. 91

Recent Use of Novel Data Streams During Foodborne Illness Cluster Investigations by the United States Food and Drug Administration: Qualitative Review (e58797)
Michael Bazaco, Christina Carstens, Tiffany Greenlee, Tyann Blessington, Evelyn Pereira, Sharon Seelman, Stranjae Ivory, Temesgen Jemaneh, Margaret Kirchner, Alvin Crosby, Stelios Viazis, Sheila van Twuyver, Michael Gwathmey, Tanya Malais, Oliver Ou, Stephanie Kenez, Nichole Nolan, Andrew Karasick, Cecile Punzalan, Colin Schwensohn, Laura Gieraltowski, Cary Chen Parker, Erin Jenkins, Stic Harris. 101

Scaling Up and Enhancing the Functionality of the Electronic Integrated Diseases Surveillance and Response System in Uganda, 2020-2022: Description of the Journey, Challenges, and Lessons Learned (e59783)
Rodney Mugasha, Andrew Kwiringira, Vivian Ntono, Lydia Nakiire, Immaculate Ayebazibwe, Caroline Kyozira, Allan Muruta, Juliet Kasule, Dathan Byonanebye, Judith Nanyondo, Richard Walwema, Francis Kakooza, Mohammed Lamorde. 112

XSL•FO
RenderX

Collaborative Surveillance: Using a Minimum Set of Key Data Parameters for One Health Participatory Surveillance (e77448)
Mark Smolinski, Nomita Divi, Onicio Leal Neto. . . . 121

Reviews

Effects of Injury Registry Data on Policymaking, Hospitalizations, and Mortality: Systematic Review (e67115)
Ana Medeiros-de-Souza, Luana Sinhori Lopes, Tayna Felicissimo Gomes de Souza Bandeira, Lucas Reis Correia, Naiza Bandeira de Sá, Bruno Zocca de Oliveira. . . . 130

Pathological Processes Among Content Creators on Social Media: Scoping Review (e76708)
Sergio Martinez-Aguirre, Javier Sanz-Valero, Elena Ronda-Pérez. . . . 145

Preparing for Potential Health and Safety Risks at the Olympic Games: Scoping Review (e66829)
Shaotong Ren, Tiantian Li, Yongzhong Zhang, Song Bai, Zichen Zhou, Shengxin Li. . . . 161

Association of Indoor Temperature Level and Mental Health Among Community-Dwelling Older Adults: Systematic Review (e78257)
Min Park, Bada Kang, Dahye Hong, Seolah Yoon, Minsub Kim, Yubeen Kim, Zhao Ni, Eleanor McConnell. . . . 180

Safety and Efficacy of Respiratory Syncytial Virus Vaccination in Older Adults: Systematic Review and Meta-Analysis of Randomized Controlled Trials (e74271)
Qingzhi Xiao, Ruizhe Yang, Lei Zhang, Ye Tian, Xu Wang, Wei Li. . . . 199

Minimum Data Set and Metadata for Active Vaccine Safety Surveillance: Systematic Review (e63161)
Mengdi Zhang, Junting Yang, Yan Li, Yuan Li, Tong Li, Ziqi Dong, Shuo Gong, Yahui Wu, Minrui Ren, Chunxiang Fan, Lina Zhang, Yi Wang, Yali Wang, Jingtian Ren, Feng Sun, Chuanyong Shen, Keli Li, Zhike Liu, Siyan Zhan. . . . 214

Identifying Health Care Services Offered in the HIV Care Continuum via a Machine Learning–Based Topic Modeling Approach: Exploratory Literature Review (e65081)
SangA Lee, Layoung Kim, Mi-So Shim, Gwang Kim. . . . 232

Symptom Trajectories and Clinical Subtypes in Post–COVID-19 Condition: Systematic Review and Clustering Analysis (e72221)
Mingzhi Hu, Tian Song, Zhaoyuan Gong, Qianzi Che, Jing Guo, Lin Chen, Haili Zhang, Huizhen Li, Ning Liang, Guozhen Zhao, Yanping Wang, Nannan Shi, Bin Liu. . . . 248

When Infodemic Meets Epidemic: Systematic Literature Review (e55642)
Chaimae Asaad, Imane Khaouja, Mounir Ghogho, Karim Baïna. . . . 310

Exploring the Use of Digital Educational Tools for Sexual and Reproductive Health in Sub-Saharan Africa: Systematic Review (e63309)
Ramatu Abdul Hamid Alhassan, Catherine Haggerty, Abimbola Fapohunda, Nabeeha Affan, Martina Anto-Ocrah. . . . 345

Uptake of the Second Dose of the Measles Vaccine and Its Determinants Among Children Aged Less Than 5 Years: Systematic Review and Meta-Analysis (e77195)
Molalign Adisu, Tesfaye Habtie, Tegene Kitaw, Alemu Zemariam. . . . 362

Risk Factors for Typhoid Fever: Systematic Review (e67544)
Portia Boakye Okyere, Sampson Twumasi-Ankrah, Sam Newton, Samuel Nkansah Darko, Michael Owusu Ansah, Eric Darko, Francis Agyapong, Hyon Jeon, Yaw Adu-Sarkodie, Florian Marks, Ellis Owusu-Dabo. . . . 376

XSL•FO
RenderX

Availability and Key Characteristics of National Early Warning Systems for Emerging Profiles of Antimicrobial Resistance in High-Income Countries: Systematic Review (e57457)
Jessica Iera, Claudia Isonne, Chiara Seghieri, Lara Tavoschi, Mariateresa Ceparano, Antonio Sciurti, Alessia D'Alisera, Monica Sane Schepisi, Giuseppe Migliara, Carolina Marzuillo, Paolo Villari, Fortunato D'Ancona, Valentina Baccolini. 553

Development of a Conceptual Framework of Health Misinformation During the COVID-19 Pandemic: Systematic Review of Reviews (e62693)
Javier Alvarez-Galvez, Jesus Carretero-Bravo, Carolina Lagares-Franco, Begoña Ramos-Fiol, Esther Ortega-Martin. 1081

Use of Behavior Change Techniques in Digital HIV Prevention Programs for Adolescents and Young People: Systematic Review (e59519)
Phoenix Mo, Luyao Xie, Tsz Lee, Angela Li. 1311

Multilevel Diabetes Prevention Interventions to Address Population Inequities in Diabetes Risk: Scoping Review (e70267)
Reshma Patel, Kathy Kornas, David Gerstle, Lori Diemert, Laura Rosella. 2715

Interventions to Foster Mental Health and Reintegration in Individuals Who Are Unemployed: Systematic Review (e65698)
Sophia Adam, Florian Junne, Svenja Schlachter, Miriam Mehler, Harald Gündel, Nicolas Rüsch, Jörn von Wietersheim, Katrin Giel, Stephan Zipfel, Rebecca Erschens. 2743

Protocols

The American Transformative HIV Study: Protocol for a US National Cohort of Sexual and Gender Minority Individuals With HIV (e66921)
Christian Grov, Alexa D'Angelo, Chloe Mirzayi, Michelle Dearolf, Elena Hoeppner, Yan Guo, Nicole Richards, Rifa Ehsan, Sarah Kulkarni, Denis Nash, Viraj Patel, Dustin Duncan, Meredith Ray, Tyler Bartholomew, Jennifer Manuzak, Jennifer Manuel, Kathryn McCollister, Drew Westmoreland, Adam Carrico. 1174

The PrEP Pharmacy Reach Study: Protocol for the Creation of Maps to Visualize the Impact of Expanding Access to HIV Prevention Services Through Pharmacies (e75077)
Kristin Harrington, Chante Hamilton, Daniel Alohan, Alexis Hudson, Henry Young, Natalie Crawford. 1434

An Enhanced Social Network Strategy to Increase the Uptake of HIV Services: Protocol for Type I Hybrid Implementation Study (Carolinas RESPOND) (e69495)
Ian Dale, Meagan Zarwell, Alicia Diggs, Jesse Strunk Elkins, Eunice Okumu, Sharon Weissman, Feng-Chang Lin, Victoria Mobley, Carol Golin, Abby Rudolph, Ann Dennis. 1443

Evaluating the Effectiveness of an Intelligent mHealth Intervention for Child Unintentional Injury Prevention: Protocol for a Cluster Randomized Controlled Trial (e76195)
Yang Yuan, Yuying Jing, Jiyu Li, Jieyi He, David Schwebel, Peishan Ning, Zhenzhen Rao, Li Li, Guoqing Hu. 1461

Substance-Related Acute Toxicity Deaths in Canada From 2016 to 2017: Protocol for a Retrospective Chart Review Study of Coroner and Medical Examiner Files (e49981)
Jenny Rotondo, Amanda VanSteelandt, Fiona Kouyoumdjian, Matthew Bowes, Tanya Kakkar, Graham Jones, Brandi Abele, Regan Murray, Emily Schleihauf, Jessica Halverson, Jennifer Leason, Dirk Huyer, Beth Jackson, Songul Bozat-Emre, Devanshi Shah, Erin Rees. 1474

XSL•FO
RenderX

Notes from the Fields

Social Media Recruitment in Indigenous and Native American Populations: Challenges in the AI Age (e76677)
Nadia Diamond-Smith, Alison Comfort, Anna Epperson, Alicia Riley, Natalie Beylin, Mary Garcia, Sarah Francis, Lucía Abascal Miguel. 2188

Amplifying the Voices of Youth for Equity in Wellness and Technology Research: Reflections on the Midwest Youth Wellness Initiative on Technology (MYWIT) Youth Advisory Board (e69013)
Linnea Laestadius, Fridarose Hamad, Leena Le, Rosemary Buchtel, Celeste Campos-Castillo. 2193

Implementing Social Media Strategies in Community-Partnered HIV Research: Practical Considerations From 3 Ongoing Studies (e73318)
Elena Rosenberg-Carlson, Tamra Loeb, Raquel Hernandez, Jane Lee, Aima Ahonkhai, Jessica Perkins, Chunqing Lin, Sung-Jae Lee, Sharon Hurt, Luke Johnsen, Alison Hamilton. 2204

Corrigenda and Addendas

Correction: Assessing the Impact of Frailty on Infection Risk in Older Adults: Prospective Observational Cohort Study (e78866)
Ya Yang, Kechun Che, Jiayan Deng, Xinming Tang, Wenyuan Jing, Xiuping He, Jiacheng Yang, Wenya Zhang, Mingjuan Yin, Congcong Pan, Xiaoling Huang, Zewu Zhang, Jindong Ni. 2217

Correction: Meal Timing and Depression Among Chinese Children and Adolescents: Cross-Sectional Study (e73716)
Huilun Li, Zhaohui Lu, Erliang Zhang, Jie Zhang, Shuheng Cui, Masaki Takahashi, Mi Xiang. 2219

Correction: Comparison of Virus Watch COVID-19 Positivity, Incidence, and Hospitalization Rates With Other Surveillance Systems: Surveillance Study (e85512)
Wing Fong, Vincent Nguyen, Sarah Beale, Thomas Byrne, Cyril Geismar, Ellen Fragaszy, Jana Kovar, Alexei Yavlinsky, Ibrahim Abubakar, Andrew Hayward, Robert Aldridge. 2220

Correction: Assessment of Bidirectional and Threshold-Response Associations Between Cognitive Function and Physical Performance: Nationwide Cross-Sectional Study (e88582)
Huixiu Hu, Lanying Xie, Yuqing Hao, Yajie Zhao, Huanhuan Luo, Kang Yu, Chao Sun. 2222

Research Letters

Net Reproduction Number as a Real-Time Metric of Population Reproducibility (e63603)
Chiara Achangwa, Changhee Han, Jun-Sik Lim, Seonghui Cho, Sangbum Choi, Sukhyun Ryu. 2706

Racial/Ethnic Inequities in Polysubstance Use Among Online Help–Seeking Sexual and Gender Minoritized People in San Francisco From 2022 to 2025: Cross-Sectional Study (e82313)
Jarett Maycott, Sean Arayasirikul. 2711

XSL•FO
RenderX

Engagement With Digital Adherence Technologies as Measures of Intervention Fidelity Among Adults With Drug-Susceptible Tuberculosis and Health Care Providers: Descriptive Analysis Using Data From Cluster-Randomized Trials in Five Countries

Jason Alacapa[1], MD, MBA, MHM, MPH; Amare Worku Tadesse[2], MD, PhD; Natasha Deyanova[3], MA; Tanyaradzwa Dube[4], MPH; Andrew Mganga[5], MSc; Rachel Powers[6], MSc; Job van Rest[6], MSc; Norma Madden[6], MSc; Egwuma Efo[6], MSc; Salome Charalambous[4], MBBS, PhD; Kristian van Kalmthout[6], MSc; Degu Jerene[6], MD, PhD; Katherine Fielding[2], PhD

[1]KNCV Tuberculosis Plus, Manila, Philippines

[2]TB Centre and Department of Infectious Disease Epidemiology and International Health, London School of Hygiene & Tropical Medicine, Keppel St, London, United Kingdom

[3]Program for Appropriate Technology in Health (PATH), Kyiv, Ukraine

[4]Implementation Research Division, The Aurum Institute, Johannesburg, South Africa

[5]KNCV Tuberculosis Plus, Dar es Salaam, United Republic of Tanzania

[6]Evidence and Impact, KNCV Tuberculosis Plus, The Hague, The Netherlands

Corresponding Author:
Amare Worku Tadesse, MD, PhD
TB Centre and Department of Infectious Disease Epidemiology and International Health, London School of Hygiene & Tropical Medicine, Keppel St, London, United Kingdom

Abstract

Background: Digital adherence technologies (DATs) are promising tools for supporting tuberculosis (TB) treatment. DATs can serve as reminders for people with TB to take their medication and act as proxies for adherence monitoring. Strong engagement with DATs, from both the person with TB and health care provider (HCP) perspectives, is essential for ensuring intervention fidelity. The Adherence Support Coalition to End TB (ASCENT) project evaluated 2 types of DATs, pillboxes and medication labels (99DOTS), in cluster-randomized trials across 5 countries.

Objective: This study aims to investigate participant and HCP engagement with DATs for TB treatment, stratified by DAT type and country.

Methods: This study is a subanalysis of data generated through the ASCENT trials, which enrolled adults with drug-susceptible TB. A digital dose was defined as either a pillbox opening (for pillbox users) or a dosing confirmation SMS text message sent by the participant (for label users), both of which were recorded on the adherence platform. Descriptive analysis was used to provide an overview of dose-day outcomes. DAT engagement was assessed from both participant and HCP perspectives. To enhance participant engagement, we summarized the frequency of digital engagement overall and by treatment phase, as well as the frequency of consecutive days without engagement. For HCP engagement, we summarized the frequency of doses added manually, the number of days between the actual dose day and when a manual dose was added, and instances of consecutive manual dosing lasting more than 3 and more than 7 days, where doses were added more than 1 week after the dose day.

Results: Of the 9511 participants included, 6719 (70.64%) were using the pillbox, 3544 (37.26%) were female, and the median age was 40 years. Across DAT types, there were 1,384,879 dose days, with 973,876 (70.32%) contributed by pillbox users. Of all dose days, 1,165,195 (84.14%) were recorded as digital, 156,664 (11.31%) as manual, 59,045 (4.26%) had no information, and 3975 (0.29%) were confirmed as missed. Digital dosing decreased slightly from the intensive to the continuation phase. The percentage of digital dose days was higher among pillbox users (851,496/973,876, 87.43%) compared with label users (313,699/411,003, 76.33%). Among label users, manual dosing was most common in the Philippines (37,919/171,786, 22.07%) and least common in Tanzania (11,108/76,231, 14.57%). Among pillbox users, manual dosing was most common in the Philippines (24,015/208,130, 11.54%) and Ukraine (13,209/111,901, 11.80%). Overall, 512 out of 2792 (18.34%) label users and 588 out of 6719 (8.75%) pillbox users experienced a run of more than 7 consecutive nondigital dose days that were resolved more than 1 week after the dose day. The highest occurrence was observed in the Philippines (368/1142, 32.22%, for label users and 224/1351, 16.58%, for pillbox users).

XSL•FO
RenderX

Conclusions: There was considerable variation in DAT engagement across countries and DAT types, reflecting differences in how the intervention was implemented. Further refinement of the intervention and improvements in its delivery may be necessary to enhance outcomes.

Trial Registration: Pan African Clinical Trials Registry PACTR2020087766949990; https://pactr.samrc.ac.za/TrialDisplay.aspx?TrialID=12241 and ISRCTN Registry ISRCTN17706019; http://www.isrctn.com/ISRCTN17706019

(JMIR Public Health Surveill 2025;11:e62881) doi:10.2196/62881

KEYWORDS

tuberculosis; digital adherence technology; DAT engagement; intervention fidelity; manual dosing; digital dosing

Introduction

Ending the tuberculosis (TB) epidemic is one of the globally agreed-upon development goals, with a target of reducing TB deaths by 90% by 2030 compared with the 2015 baseline [1]. However, poor treatment outcomes continue to pose significant challenges in high TB–burden settings [1]. Several interventions have been designed and implemented to improve treatment outcomes, including early diagnosis and treatment, counseling and psychosocial support, and managing comorbidities and coinfections [2]. These interventions are articulated within the broader directly observed therapy (DOTS) strategy, which emphasizes maximizing opportunities for direct observation of treatment. The standard treatment duration for drug-susceptible TB (DS-TB) is 6 months, comprising a 2-month intensive phase with 4 drugs followed by a 4-month continuation phase with 2 drugs. Treatment is taken daily during both phases, and direct in-person observation of dose intake remains the predominant recommendation worldwide [3].

With the increasing availability of technologies that facilitate remote treatment observation, alternatives to DOT, such as digital adherence technologies (DATs), have emerged as viable options. Several types of DATs have been used to support TB treatment, including SMS text message–based reminders, smart pillboxes, and video-supported therapy. A recent systematic review suggested that DAT use improves adherence in both DS-TB and drug-resistant TB treatment, although its effect on treatment outcomes remains mixed [4]. Scoping reviews of DAT implementations have documented intervention fidelity challenges, including technological and provider-facing issues, as well as complex cellular accessibility barriers [5,6]. While studies have summarized engagement with DATs from the perspective of individuals with TB, the perspective of health care providers (HCPs) has rarely been explored. A recent scoping review reported the weighted mean of digital engagement over dose days among people with TB as 87% (12 studies) for pillbox users and 61% (4 studies) for "99DOTS" users [6]. An analysis of 11 projects across 10 high TB-burden settings found that electronic doses added manually by HCPs (referred to as "manual doses") accounted for 21.6% of reported doses in the DS-TB population [7].

Under the Adherence Support Coalition to End TB (ASCENT) project, we conducted pragmatic cluster-randomized trials in 5 countries to determine whether the use of DATs can improve treatment outcomes [8,9]. In these studies, eligible participants in the intervention arms received either pillboxes or medication labels (similar to "99DOTS" [10]), with the latter requiring participants to send a toll-free SMS text message each time treatment was taken. These DATs were integrated into a digital adherence platform, where participants' digital engagement with the DATs (pillbox opened/SMS text messages sent) served as a proxy indicator of treatment adherence. Failure to engage with the DATs triggered a cascade of care actions, including automated SMS text message reminders, phone call reminders, and home visits, with escalating actions documented in response to higher levels of DAT nonengagement.

The objectives of this study were to describe the engagement of people with TB and HCPs with the DAT interventions, by country and DAT type. Our findings provide valuable insights into the implementation challenges of DAT interventions aimed at improving treatment adherence and outcomes for individuals with TB.

Methods

Study Design

We conducted pragmatic cluster-randomized trials in Ethiopia (PACTR2020087766949990), Ukraine, Tanzania, South Africa, and the Philippines (4 countries under ISRCTN17706019) to assess the effectiveness of DATs on TB treatment outcomes. The trial designs have been reported elsewhere [11,12]. Clusters, defined as health facilities (rayons in Ukraine), were purposively selected in consultation with the national TB programs based on previously notified TB cases, willingness and capacity to participate, and achieving a reasonable balance of urban and rural settings as well as large and small facilities. Clusters were selected from 5 oblasts in Ukraine, 4 regions in Tanzania, and 2 provinces in each of South Africa and the Philippines. Clusters were randomized (1:1:2 and 1:1:1 in Ethiopia) to receive either a smart pillbox, medication labels, or standard of care in all countries except Ukraine. As fixed-dose combinations were not available in Ukraine, the medication labels, which required fixed-dose combinations, were not used in the trial.

All adults with drug-sensitive TB starting the standard treatment regimen in facilities randomized to a DAT were offered enrollment into the study at treatment initiation by routine HCPs at the facility, following consent, using the DAT type allocated to their facility. Access to a mobile phone was a prerequisite for participation in the labels intervention. As such, participants in the medication labels arm without mobile phone access were offered the pillbox instead. Treatment supporter phone numbers could also be used in either DAT arm for participants without

XSL•FO
RenderX

personal mobile phone access. Designated HCPs at each health facility responsible for TB care were trained in enrollment and in delivering the intervention, including the use of the digital adherence platform [13].

Participants using pillboxes were asked to keep their medication in the pillbox, which provided a daily audio/visual reminder at a predefined time based on participants' preferred time to take their treatment. Box opening, a proxy for dose taken, was captured in real time on the cloud-based adherence platform, which could be viewed by HCPs at each facility. Participants using medication labels were asked to send a unique toll-free SMS text message code daily when taking their dose, with these data also captured on the adherence platform. A box opening or an SMS text message sent by a participant using labels was referred to as a digital dose. At a prearranged time, if no box-opening signal (for pillbox users) or SMS text message code (for label users) had been received, an automated SMS text message reminder to take and log the dose was sent from the adherence platform to participants with mobile phone access. No SMS text message reminders were used in Ukraine, based on a decision by the Ukrainian research team. HCPs reviewed dosing data on the adherence platform and employed strategies such as phone calls and home visits to promote medication adherence, based on a differentiated response algorithm. A dose not recorded digitally on the adherence platform but self-reported as taken by the participant following a phone call required manual adjustment of the recorded dose by the HCP on the adherence platform. Manual entries were expected to be completed promptly to meet adherence criteria and maintain the integrity of the treatment monitoring process, enabling timely intervention when needed.

Adults starting TB treatment in the standard-of-care arm received care according to their country's guidelines, which could have involved facility-based or treatment-supporter–observed therapy, or self-administration.

In this analysis, we constructed a cohort of trial participants from the intervention arms of the 5 trials and used longitudinal data on their engagement with the DAT during their 6-month treatment period.

Setting and Participants

We included adult participants enrolled in the trials who received either the pillbox or medication labels from June 1, 2021, to May 31, 2022. In South Africa, medication labels were discontinued early due to feasibility challenges; all participants enrolled after October 2021 were offered the pillbox. We excluded participants who switched from medication labels to pillboxes during their treatment. Participants were analyzed according to the DAT type received.

Dosing Data

Data from the adherence platform were used to construct a profile of DAT engagement for each participant over their treatment period, from the DAT start to the earliest of (1) the treatment outcome date or (2) 168 days after treatment start (assuming 6 months of treatment). We excluded dose-day records with errors where either the date the dose was received was missing (10 records) or the date the dose was received preceded the dose day (36 records).

For each dose day, the adherence platform documented 1 of 4 mutually exclusive outcomes: digital dose, manual dose, dose missed, or no information. A digital dose-day confirmation refers to when the participant opens the pillbox/sends an SMS text message on the dose day. A manual dose refers to when the pillbox was not opened or an SMS text message was not sent on the dose day, but the HCP confirmed with the participant that the dose was taken and marked it as such on the platform. This confirmation of the dose being taken could occur on or after the actual dose day, with the date of confirmation recorded on the platform. A missed dose refers to instances when the pillbox was not opened or an SMS text message was not sent on the dose day, and the HCP later confirmed with the participant that the dose was not taken. Confirmation of both manual doses and doses not taken was based on participant self-report. "No information" for a dose day refers to instances when the pillbox was not opened or an SMS text message was not sent on the dose day, and no further information was available from the HCP regarding whether the dose was taken.

In real time, participants' DAT engagement data were displayed in a calendar format on the platform, accessed by the HCPs (Figure S1 in Multimedia Appendix 1). Digital doses were shown in dark green. "No information" days, when the box was not opened or an SMS text message was not sent, were displayed in light red, starting from midnight the following day. If a "no information" dose day was later updated to a manual dose by the HCP, it was shown in light green. For confirmed missed doses, the dose day was changed to dark red. There were instances in the labels arm when participants correctly opened their pillbox or sent an SMS text message, but due to temporary connectivity issues, this information was not synced to the platform. As pillboxes maintain an on-device log of all openings, once connectivity was restored, the record of digital dosing would overwrite any interim activity, such as manual dosing entered by the HCP. This was not possible for medication labels, as the system did not support automatic resending of unreceived SMS text messages. There was no standardized approach to logging connectivity issues (except in Ethiopia), and when such issues did occur, they were generally facility-dependent.

The dataset used for the analysis was based on the last dose day recorded on November 15, 2022, and the data extract was dated December 7, 2022. A 3-week window was therefore allowed for any HCP activity, such as adding manual doses or overwriting data due to connectivity issues, to occur.

Analyses

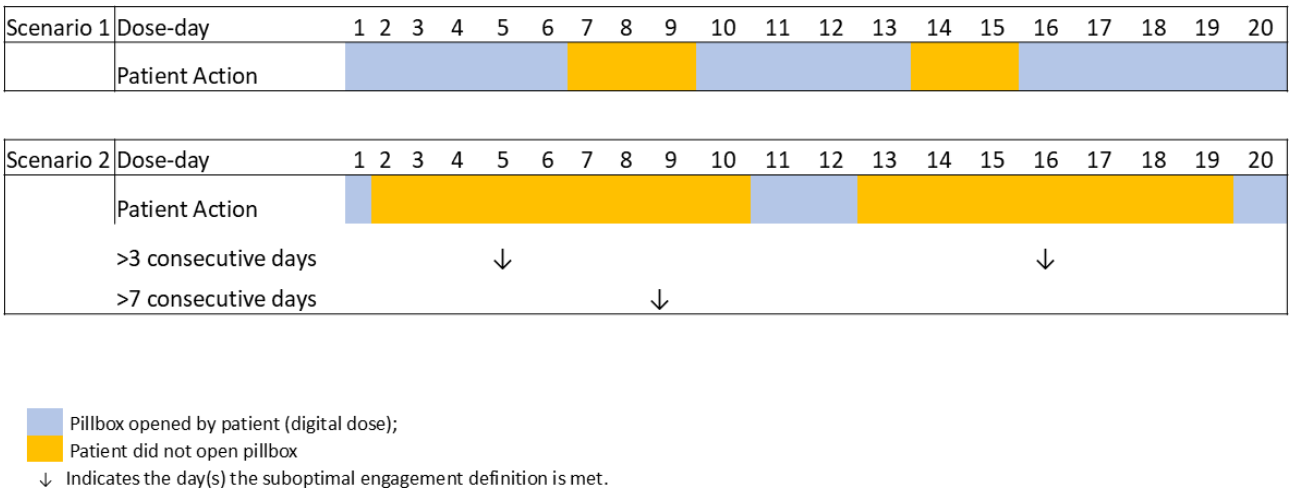
Descriptive analysis was used to provide an overview of dose-day outcomes. DAT engagement was categorized into 2 areas: participant engagement, represented by digital doses recorded on the platform; and HCP engagement, represented by dose days added manually. For participant engagement, we summarized the frequency of digital dosing overall, and separately for the intensive and continuation phases. For HCP engagement, we summarized the frequency of participants with manual dosing, as well as the timing of manual dose entries,

XSL•FO
RenderX

defined as the number of days from the actual dose day to when a manual dose was added.

We further defined scenarios reflecting "suboptimal engagement" with the DAT from both participant and HCP perspectives. From the participant's perspective, suboptimal engagement was characterized by inadequate, timely interaction with the DAT, indicated by nondigital doses, defined as manual, missed, or no information dose days. We considered suboptimal, timely interaction to be present when there were more than 3 or more than 7 consecutive nondigital dose days. Figure 1 illustrates a case where suboptimal engagement criteria were met (scenario 2), whereas scenario 1 does not meet these criteria.

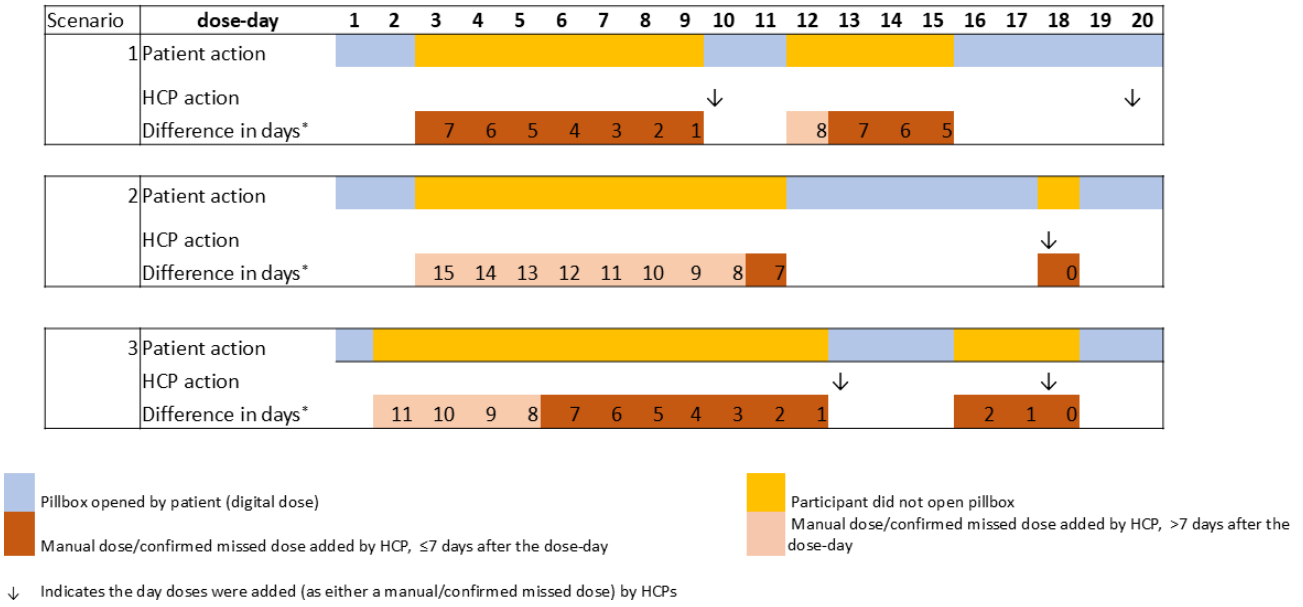
Figure 1. Illustration of 2 scenarios showing suboptimal engagement by person on tuberculosis treatment. Acceptable engagement is defined as ≤3 consecutive days with no digital confirmation. Suboptimal engagement is defined as (1) >3 consecutive dose days or (2) >7 consecutive dose days with no digital confirmation. Scenario 1: The patient had consecutive nondigital confirmation on days 7 to 9 (3 days; inclusive) and days 14 to 15 (2 days; inclusive); the suboptimal engagement condition of either >3 or >7 consecutive dose days is not met. Scenario 2: The patient had consecutive nondigital dose days on days 2 to 10 (9 days; inclusive) and days 13 to 19 (7 days; inclusive); the suboptimal engagement condition of (1) >3 dose days is met twice (on day 5 and day 16) and (2) >7 consecutive dose days is met once (on day 9).



From the HCP perspective, the scenario reflecting suboptimal engagement with the DAT was defined based on a combination of (1) delayed action, where manual or missed dose was added by the HCP, or there was no information about a dose; and (2) this occurring for periods of consecutive nondigital dose days. The first component was defined as a manual or missed dose added more than 7 days after the dose day, or the absence of any information more than 7 days after the dose day. A threshold of more than 1 week was chosen, as this was considered a sufficient time for the HCP to have successfully contacted the participant. For the second component, we considered periods of consecutive nondigital dose days, specifically more than 3 and more than 7 days, as instances when timely HCP interaction (ie, contact with the participant) would have been expected. See Figure 2 for illustrative scenarios. Scenarios 2 and 3 meet the criteria for suboptimal engagement, whereas scenario 1 does not. Analyses, stratified by country and DAT type, were conducted using Stata version 17 (StataCorp).

XSL•FO
RenderX

Figure 2. Illustration of 3 scenarios showing suboptimal engagement by health care professionals (HCPs). Suboptimal engagement is defined as (1) delayed action of manual/confirmed missed dose added by the HCP or no information about a dose and (2) this occurring for periods of consecutive days where non-digital dosing occurred. All scenarios show periods when the patient opened and did not open the pillbox. Scenario 1: HCP added manual/confirmed missed doses on days 10 and 20; of 11 doses added, 10 were within (≤) 7 days of the dose day; dose day 12 was added on day 20 (8 days late); the suboptimal engagement condition is not met. Scenario 2: HCP added manual/confirmed missed doses on day 18; of 10 doses added, 8 doses added were >7 days from the dose day; there is a run of 8 days where doses were added late (dose days 3-10; inclusive); the suboptimal engagement condition is met (consecutive dose days >3 and >7). Scenario 3: HCP added manual/confirmed missed doses on days 13 and 18; of 14 doses added, 4 were >7 days from the dose day; there is a run of 4 days where doses were added late (dose days 2-5; inclusive); the suboptimal engagement condition is met (consecutive dose days >3). *From actual dose day to when the dose was added (either as a manual or confirmed missed dose) by HCP.



Ethical Considerations

Written informed consent was obtained from all participants initiating a DAT in the ASCENT trials. The main trials were approved by the World Health Organization (WHO) Ethical Review Committee, Switzerland; Institutional Review Boards in Ethiopia (Addis Ababa City Administration Health Bureau Public Emergency and Health Research Directorate Institutional Review Board, and Oromia Regional Health Bureau Public Emergency and Health Research Directorate Institutional Review Board); the Ethics Committee for the Philippines (Single Joint Research Ethics Board), South Africa (University of the Witwatersrand Human Research Ethics Committee), and Tanzania (Tanzania Medical Research Coordinating Committee at the National Institute for Medical Research, Dar es Salaam); and the London School of Hygiene & Tropical Medicine Ethics Committee, United Kingdom. The analyses presented here are covered by the original ethics approvals and did not require additional consent. All data were deidentified and contained no personal identifiers. Trial participants did not receive any compensation for their participation.

Results

Participant Characteristics

Overall, 9511 participants from 163 health facilities/rayons contributed to the analysis, of whom 6719 (70.64%) were using the pillbox: 1243 out of 2268 (54.81%) in Ethiopia, 1698 out of 1759 (96.53%) in South Africa, 1611 out of 2185 (73.73%) in Tanzania, 1351 out of 2493 (54.19%) in the Philippines, and 816 out of 816 (100%) in Ukraine. In total, 3544 (37.26%) were female: 936 out of 2268 (41.27%) in Ethiopia, 676 out of 1749 (38.65%) in South Africa, 824 out of 2185 (37.71%) in Tanzania, 845 out of 2493 (33.89%) in the Philippines, and 263 out of 816 (32.23%) in Ukraine. The median age was 40 years (29 years in Ethiopia, 41 years in South Africa, 45 years in Tanzania, 45 years in the Philippines, and 44 years in Ukraine; see Table 1).

XSL•FO
RenderX

Table . Baseline characteristics of study participants by country of enrollment and DAT[a] received.

Country and DAT type	Number of participants (number of facilities), n	Male, n (%)	Female, n (%)	Transgender, n (%)	Age (years), median (Q1[b]-Q3)[c]
Ethiopia					
Labels	1025 (26)	607 (59.22)	418 (40.78)	0 (0)	28 (23 - 39)
Pillbox	1243 (49)	725 (58.33)	518 (41.67)	0 (0)	30 (24 - 41)
South Africa					
Labels	51 (9)	35 (68.63)	16 (31.37)	0 (0)	37 (31 - 44)
Pillbox	1698 (30)	1038 (61.13)	660 (38.87)	0 (0)	41 (33 - 51)
Tanzania					
Labels	574 (16)	369 (64.29)	205 (35.71)	0 (0)	41 (30 - 54)
Pillbox	1611 (36)	992 (61.58)	619 (38.42)	0 (0)	46 (34 - 60)
The Philippines					
Labels	1142 (16)	745 (65.24)	395 (34.59)	2 (0.18)	44 (29 - 55)
Pillbox	1351 (27)	901 (66.69)	450 (33.31)	0 (0)	46 (32 - 59)
Ukraine					
Pillbox	816 (12)	553 (67.77)	263 (32.23)	0 (0)	44 (35 - 54)
Overall	9511 (163)	5965 (62.72)	3544 (37.26)	2 (0.02)	40 (29 - 53)

[a]DAT: digital adherence technology.

[b]Q1: lower quartile.

[c]Q3: upper quartile.

Description of Dosing Type

During the follow-up period, a total of 1,384,879 dose days from 9511 participants—across both DAT types—contributed to the analysis. The median number of dose days per participant was 168 days. Of all dose days, 973,876 out of 1,384,879 (70.32%) were among participants using pillboxes. Overall, 1,165,195 out of 1,384,879 (84.14%) dose days were recorded digitally, 156,664 out of 1,384,879 (11.31%) manually, 59,045 out of 1,384,879 (4.26%) had no information, and fewer than 3975 out of 1,384,879 (0.29%) were confirmed as missed. A higher percentage of dose days were recorded digitally among pillbox users compared with label users (851,496/973,876, 87.43% vs 313,699/411,003, 76.33%), with patterns varying by country (Table 2). Manual dosing for labels was most common in the Philippines (37,919/171,786, 22.07%, dose days) and least common in Tanzania (11,108/76,231, 14.57%, dose days). Among pillbox users, manual dose days were most common in the Philippines and Ukraine (24,015/208,130, 11.54%, and 13,209/111,901, 11.80%, dose days, respectively) and least common in South Africa (9075/235,417, 3.85%, dose days). No information for the dose day was most common among participants using labels (1190/6154, 19.34%, dose days) and pillboxes (18,711/235,417, 7.95%, dose days) in South Africa. Manual doses were added a median of 4 (lower quartile-upper quartile 1 - 16) days after the dose day for participants using labels, and 6 (lower quartile-upper quartile 2 - 21) days for those using pillboxes. The median time from the dose day to manual dose being added was shortest for participants in Ethiopia (2 days for labels and 3 days for pillboxes) and longest for participants in the Philippines (9 days for labels and 12 days for pillboxes). The Philippines had the highest percentage of manual doses added more than 7 days after the dose day (20,831/37,919, 54.94%, for labels and 14,708/24,015, 61.25%, for pillboxes), followed by South Africa (541/1149, 47.08%, for labels and 3966/9075, 43.70%, for pillboxes) and Ukraine (6858/13,209, 51.92%, for pillboxes). Ethiopia had the lowest proportions (6016/28,783, 20.90%, for labels and 5075/16,845, 30.13%, for pillboxes; see Table 2).

Digital dosing decreased slightly from the intensive compared with the continuation phase (see Figure S2 in Multimedia Appendix 1). The median number of days to add manual doses increased slightly during the continuation phase compared with the intensive phase for label users in the Philippines (from 8 to 10 days) and for pillbox users in Ukraine (from 7 to 10 days). For other countries and DAT types, the median number of days remained similar (see Table S1 in Multimedia Appendix 1).

XSL•FO
RenderX

Table . Summary of dose-day types and time to manual dose entry over the 6-month treatment period, by country of enrollment and digital adherence technology received.

Country	Labels	Pillbox
Ethiopia		
Recording of dose, day level		
Digital, n (%)	126,718 (80.80)	177,599 (90.45)
Manual, n (%)	28,783 (18.35)	16,845 (8.58)
Missed, n (%)	309 (0.20)	144 (0.07)
No information, n (%)	1022 (0.65)	1764 (0.90)
Total, n	156,832	196,352
Days from the dose day to when the manual dose was added to the platform		
Median days (Q1[a]-Q3[b])	2 (1-5)	3 (1-10)
Same day, n (%)	5665 (19.68)	3559 (21.13)
1 - 7 days, n (%)	17,102 (59.42)	8211 (48.74)
>7 days, n (%)	6016 (20.90)	5075 (30.13)
South Africa		
Recording of dose, day level		
Digital, n (%)	3815 (61.99)	207,569 (88.17)
Manual, n (%)	1149 (18.67)	9075 (3.85)
Missed, n (%)	0 (0)	62 (0.03)
No information, n (%)	1190 (19.34)	18,711 (7.95)
Total, n	6154	235,417
Days from the dose day to when the manual dose was added to the platform		
Median days (Q1-Q3)	6 (2-20)	6 (2-19)
Same day, n (%)	149 (12.97)	727 (8.01)
1 - 7 days, n (%)	459 (39.95)	4382 (48.29)
>7 days, n (%)	541 (47.08)	3966 (43.70)
Tanzania		
Recording of dose, day level		
Digital, n (%)	63,862 (83.77)	202,382 (91.13)
Manual, n (%)	11,108 (14.57)	14,561 (6.56)
Missed, n (%)	213 (0.28)	657 (0.30)
No information, n (%)	1048 (1.37)	4476 (2.02)
Total, n	76,231	222,076
Days from the dose day to when the manual dose was added to the platform		
Median days (Q1-Q3)	3 (1-8)	4 (2-10)
Same day, n (%)	551 (4.96)	1174 (8.06)
1 - 7 days, n (%)	7651 (68.88)	8764 (60.19)
>7 days, n (%)	2906 (26.16)	4623 (31.75)
The Philippines		
Recording of dose, day level		
Digital, n (%)	119,304 (69.45)	172,234 (82.75)
Manual, n (%)	37,919 (22.07)	24,015 (11.54)
Missed, n (%)	1756 (1.02)	384 (0.18)

XSL•FO
RenderX

Country	Labels	Pillbox
No information, n (%)	12,807 (7.46)	11,497 (5.52)
Total, n	171,786	208,130
Days from the dose day to when the manual dose was added to the platform		
Median days (Q1-Q3)	9 (3-24)	12 (4 - 17)
Same day, n (%)	2381 (6.28)	932 (3.88)
1 - 7 days, n (%)	14,707 (38.79)	8375 (34.7)
>7 days, n (%)	20,831 (55)	14,708 (61.25)
Ukraine		
Recording of dose, day level		
Digital, n (%)	N/A[c]	91,712 (81.96)
Manual, n (%)	N/A	13,209 (11.80)
Missed, n (%)	N/A	450 (0.40)
No information, n (%)	N/A	6530 (5.84)
Total, n	N/A	111,901
Days from the dose day to when the manual dose was added to the platform		
Median days (Q1-Q3)	N/A	9 (3-24)
Same day, n (%)	N/A	599 (4.53)
1 - 7 days, n (%)	N/A	5752 (43.55)
>7 days, n (%)	N/A	6858 (51.92)
Overall		
Recording of dose, day level		
Digital, n (%)	313,699 (76.33)	851,496 (87.43)
Manual, n (%)	78,959 (19.21)	77,705 (7.98)
Missed, n (%)	2278 (0.55)	1697 (0.17)
No information, n (%)	16,067 (3.91)	42,978 (4.41)
Total, n	411,003	973,876
Days from the dose day to when the manual dose was added to the platform		
Median days (Q1-Q3)	4 (1-16)	6 (2-21)
Same day, n (%)	8746 (11.08)	6991 (9.00)
1 - 7 days, n (%)	39,919 (50.56)	35,484 (45.67)
>7 days, n (%)	30,294 (38.37)	35,230 (45.34)

[a]Q1: lower quartile.

[b]Q3: upper quartile.

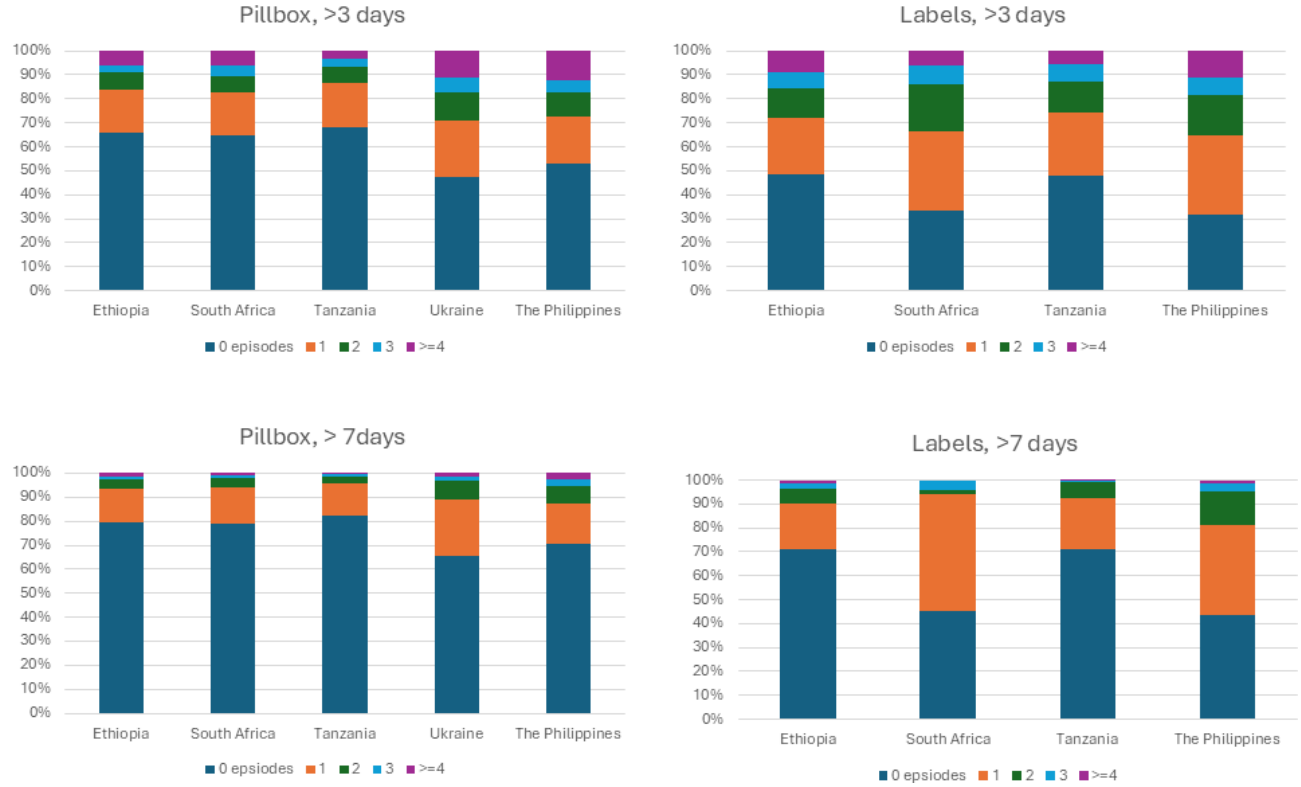
[c]N/A: not applicable.

Participant and HCP Suboptimal DAT Engagement

Participant suboptimal engagement analysis is summarized in Figure 3. From the participant's perspective, the Philippines had the highest number of episodes of more than 3 consecutive nondigital dose days for both pillboxes and labels, while Tanzania and Ethiopia had the lowest frequency. For episodes of more than 7 consecutive nondigital dose days, the Philippines again showed the highest frequency for both DAT types, with the lowest frequency observed in Ethiopia and South Africa.

XSL•FO
RenderX

Figure 3. Bar charts showing the distribution of the total number of participant-level episodes (grouped) of consecutive nondigital dose days lasting >3 and >7 days over the 6-month treatment period, stratified by country of enrollment and type of digital adherence technology received. Top panels: summary at the participant level of the number of episodes (grouped into 0, 1, 2, 3, and ≥4 episodes) where >3 consecutive non-digital dose-days were observed for pillbox (left) and labels (right), by country. Lower panels: summary at the participant level of the episodes (grouped into 0, 1, 2, 3, and ≥4 episodes) where >7 consecutive non-digital dose-days were observed for pillbox (left) and labels (right), by country.



Suboptimal engagement by the HCP is summarized in Table 3. Of the 2792 participants using labels, 820 (29.37%) and 512 (18.34%) met the criteria for suboptimal engagement of more than 3 days and more than 7 days, respectively, at least once. The Philippines had the highest percentage of HCP suboptimal engagement, with 501 out of 1142 (43.87%) and 368 out of 1142 (32.22%) participants experiencing suboptimal engagement, at least once, of more than 3 days and more than 7 days, respectively. South Africa ranked second, with 19 of 51 (37%) and 15 of 51 (29%) for suboptimal engagement of more than 3 days and more than 7 days, respectively. Combining across the 4 countries, 909 out of 1698 (53.53%) episodes lasting more than 3 days also met the criterion for more than 7 days.

XSL•FO
RenderX

Table . Suboptimal health care provider engagement with the digital adherence technology over the 6-month treatment period, by country of enrollment and type of digital adherence technology received.[a]

DAT type, country (number of participants), and consecutive nondigital runs[b]	Total number of episodes, n	Number of participants with at least one episode, n (%)	Summary of number of episodes, participant-level, median (Q1[c]-Q3[d])
Labels			
Ethiopia (n=1025)			
>3 days	368	208 (20.3)	1 (1-2)
>7 days	183	103 (10.0)	1 (1-2)
South Africa (n=51)			
>3 days	39	19 (37.3)	1 (1-3)
>7 days	30	15 (29.4)	2 (1-2)
Tanzania (n=574)			
>3 days	135	92 (16.0)	1 (1-2)
>7 days	31	26 (4.5)	1 (1-1)
The Philippines (n=1142)			
>3 days	1156	501 (43.9)	2 (1-3)
>7 days	665	368 (32.2)	1 (1-2)
Pillboxes			
Ethiopia (n=1243)			
>3 days	349	185 (14.9)	1 (1-2)
>7 days	157	98 (7.9)	1 (1-2)
South Africa (n=1698)			
>3 days	241	138 (8.2)	1 (1-2)
>7 days	114	70 (4.1)	1 (1-2)
Tanzania (n=1611)			
>3 days	284	175 (10.9)	1 (1-2)
>7 days	103	79 (4.9)	1 (1-1)
The Philippines (n=1351)			
>3 days	825	347 (25.7)	2 (1-3)
>7 days	407	224 (16.6)	1 (1-2)
Ukraine (n=816)			
>3 days	388	217 (26.6)	1 (1-2)
>7 days	158	117 (14.3)	1 (1-2)

[a]Suboptimal engagement was defined as (1) delayed action by the provider—either manually adding a missed dose or providing no dose information—more than 7 days after the dose day, and (2) this delay occurring during runs of more than 3 or more than 7 consecutive nondigital dose days.

[b]For doses added more than 7 days after the dose day.

[c]Q1: lower quartile.

[d]Q3: upper quartile.

Among participants using the pillbox (n=6719), 1062 (15.81%) and 588 (8.75%) met the criteria for suboptimal engagement of more than 3 days and more than 7 days, respectively, at least once. Ukraine had the highest proportion of participants with more than 3 days (217/816, 26.6%) and more than 7 days (117/816, 14.3%) of suboptimal engagement, followed by the Philippines (347/1351, 25.7% and 224/1351, 16.6%, respectively). Combining across the 5 countries, of the 2087 episodes lasting more than 3 days, 939 (45.0%) also met the definition for more than 7 days of suboptimal engagement. Typically, participants experienced a median of 1 episode of suboptimal engagement, except in the Philippines and South Africa. In the Philippines, the median number of episodes lasting more than 3 days was 2 for both DATs. In South Africa, the

XSL•FO
RenderX

median number of episodes lasting more than 7 days was 2 for the label DAT.

Discussion

Principal Findings

This study sought to understand the complex interactions between adults on DS-TB treatment, HCPs, and the DAT intervention, using daily DAT engagement data from 2 DAT types across 5 countries. We found a higher percentage of dose days recorded as digital among pillbox users compared with label users, with digital dosing lowest among pillbox users in Ukraine and the Philippines. The time taken by HCPs to add manual doses varied across countries: the Philippines had the longest intervals for both DAT types, followed by Ukraine and South Africa. Combining across DATs, 34,537 out of 45,628 (75.69%) manual doses in Ethiopia and 18,140 out of 25,669 (70.67%) in Tanzania were added within a 7-day time frame. There was also a lack of dose-taking information recorded on the adherence platform, most frequently observed in South Africa, reflecting another marker of poor HCP engagement with the intervention. Finally, our marker of suboptimal DAT engagement by HCPs showed variation by DAT type and country. Across countries, a range of label users (92/574, 16.0%, to 501/1142, 43.9%) and pillbox users (138/1698, 8.13%, to 217/816, 26.6%) experienced at least one run of nondigital dosing lasting more than 3 days, which was either resolved more than 7 days after the dose day or not resolved at all (no information). The Philippines had the highest percentage of participants experiencing this issue for labels (501/1142, 43.87%) and, along with Ukraine, the highest for pillboxes (347/1351, 25.7%, and 217/816, 26.6%, respectively).

We observed differences in our indicators by country, which may be explained by variations in participants, health care systems, and external factors and are consistent with our process evaluation [13]. The findings from the Philippines stand out, reflecting a pattern of lower engagement with the DATs by both participants and HCPs. Higher levels of manual dosing and more frequent episodes of consecutive nondigital dosing suggest lower participant engagement in this setting. This is further supported by findings from the ASCENT acceptability substudy, where participants in the Philippines reported lower mean scores for Capability, Opportunity, and Motivation, key domains of the Theoretical Domains Framework used to assess acceptability, compared with participants in other countries [14]. Participants also reported lower HCP engagement in the acceptability substudy, as indicated by a lower percentage reporting that HCPs showed them their adherence data during routine clinic visits. In the Philippines, intervention delivery was affected by the reassignment of TB focal staff to COVID-19 vaccination programs in 2021. This led to a decline in participant enrollment during that period and likely posed challenges in delivering the DAT intervention to those already enrolled. In a qualitative study involving people with TB and HCPs participating in the Philippines trial, HCPs cited insufficient training and lack of refresher training at the program design level, as well as power issues at the community level, as barriers to DAT implementation [15]. In South Africa, missing dose information on the platform may have been due to difficulties in contacting participants, such as changes in phone numbers, loss of phones, or provision of incorrect contact details, hindering HCPs' ability to confirm doses. Additionally, some participants were lost to follow-up due to mobility-related issues, such as relocating for work, which resulted in unsuccessful home visits for dose confirmation. During the study period, frequent power outages in South Africa may have contributed to suboptimal HCP engagement with the intervention. In Ukraine, the study coincided with disruptions caused by the military invasion by Russia, which likely affected the timeliness of HCP contact with participants following missed doses recorded on the platform. Regional differences in dosing patterns were observed within Ukraine; Mykolaivska and Donetska oblasts, occupied by Russia early in the invasion, had higher levels of manual dosing added late compared with other regions [16]. Furthermore, in Ukraine, the DAT intervention was implemented by HCPs within the context of staff reductions in TB care delivery, a challenge not reported in other countries. By the midpoint of trial implementation, fewer HCPs were providing DAT support than had been trained at the start of the trial, which may have contributed to suboptimal engagement with the DAT.

In the acceptability substudy, regular access to a mobile phone was identified as an important aspect of the intervention from both participant and HCP perspectives. Shared phone ownership may hinder timely dose taking in response to automated SMS text message reminders in either arm, as well as limit HCPs' ability to successfully contact participants following nonengagement. Ownership of a nonshared phone was lowest among participants in the Philippines compared with other countries, potentially contributing to suboptimal engagement from both participant and HCP perspectives [14]. Disparities in phone ownership are well-known challenges in implementing digitally mediated interventions [17,18].

Poor connectivity poses a significant barrier to optimal DAT engagement, particularly affecting the label intervention. Unlike the pillbox intervention, where box openings are transmitted once connectivity is restored, the label intervention cannot "correct" for periods of nonconnectivity. This limitation may have contributed to lower recorded digital doses for labels compared with pillboxes, as missed or delayed transmissions could not be retrospectively updated. Consequently, participants using labels may appear less engaged, even if their actual adherence was higher than reflected in the data. This discrepancy can also affect health care workers' engagement with DATs, as they may become less motivated to rely on or use the technology if they perceive it as unreliable [5,19,20]. Addressing connectivity issues and incorporating features to mitigate their impact are crucial for ensuring accurate data capture and improving overall engagement from both participants and health care workers.

Comparison With Prior Work

Various levels of participant engagement have been documented across different DAT types in various settings. A recent scoping review reported lower participant engagement with 99DOTS (similar to our labels interventions) compared with pillboxes

XSL•FO
RenderX

[6]. Furthermore, factors such as suboptimal DAT functionality, complex cellular acceptability, and less accurate adherence data may have contributed to these discrepancies [5]. This could be partly explained by differences in ease of use: pillbox users received an optional daily audio reminder, as well as an automated SMS text message reminder if the system did not receive a pillbox opening signal or SMS text message dose confirmation from the participant within a set time. Ready access to a phone among label users may have been a challenge for documenting dose-taking. In the 99DOTS stepped-wedge trial in Uganda, the median percentage of digital dosing was 58.4%, peaking in month 1 at 71.4% and declining to 46.4% by month 6. When manual dosing was included, the median percentage exceeded 99.4% [21]. Digital dosing among label users in the ASCENT trials was generally higher compared with this study. The percentage of digital dosing recorded in pillbox intervention arms in other trials ranges from 84.2% to 88%, aligning with the ASCENT trial data [22-24].

Few studies, however, have summarized digital/nondigital dosing patterns from both participant and HCP perspectives. A meta-analysis of data from 11 projects provided a valuable overview of DAT engagement patterns over the treatment period, but did not explore the nuances of manual dosing frequency and timing. This may be due to differences in how manual dosing was defined across studies, as well as limitations in data capture that prevented clear distinctions between digital and manual dosing [7]. The significance of consecutive nondigital doses, particularly when frequent and prolonged, may indicate participant nonengagement with the DAT. In such cases, manual doses added late by HCPs could also reflect poor HCP engagement. High fidelity in intervention delivery by HCPs is essential for the success of DATs. A multicomponent intervention conducted in Tibet among 276 individuals starting TB treatment, which included pillboxes and a switch to video-supported therapy for those with adherence challenges, reported high fidelity to the transition to video-supported therapy, with at least 89% maintained over the treatment period [25]. In our trials, the calendar format on the adherence platform provided HCPs with a clear visualization of patients' dose taking, intended to support timely action. However, concerns have been raised about the use of this format, particularly in busy health facilities. There may be a temptation for HCPs to manually record doses in the digital calendar without clear confirmation from the participant, similar to critiques of adherence documentation in TB register calendars used in standard care.

Limitations

Limitations of our study include the fact that DAT engagement reflects only dosing as recorded on the adherence platform; therefore, digital dosing should not be directly equated with actual pill consumption. What was recorded were proxies (either via SMS text messages or pillbox signals) of medication intake. Manual doses added late may have resulted from HCPs being unable to successfully contact participants, rather than from delayed action or lack of attempt. However, we were unable to distinguish between these possibilities. Hence, caution should be exercised when interpreting the findings. Furthermore, the external validity is limited, as the sampled participants and collected data were specific to the ASCENT trial. Nonetheless, the strengths of our study are its coverage of 5 countries, each with distinct epidemiological, socioeconomic, geographical, infrastructural, and health system contexts. Data were generated by routine staff delivering the intervention across 163 health facilities/rayons and included 9511 participants.

Conclusions

This study highlights the role of DATs in TB treatment, focusing on how adults with DS-TB and HCPs interact with the technology. We observed considerable variation in DAT engagement across countries and DAT types, reflecting differences in how the intervention was implemented. These findings underscore the need for further research to assess whether variations in DAT engagement affect the overall effectiveness of these interventions. Such research could help inform improvements in DAT design and implementation to foster higher engagement among people with TB and HCPs, thereby potentially enhancing intervention effectiveness.

Acknowledgments

This study was conducted as a part of the Adherence Support Coalition to End TB (ASCENT) trials funded under the ASCENT Consortium through a grant from Unitaid (Grant Agreement Number 2019-33-ASCENT).

Data Availability

The datasets generated or analyzed during this study are available in the DataCompass LSHTM repository [26].

Authors' Contributions

JA, AWT, SC, DJ, and KF developed the study concept. JA and KF contributed to the study design, with KF providing statistical expertise for the analysis. JA, AWT, ND, TD, and AM implemented the study. JA and KF wrote the first draft of the manuscript. All authors provided intellectual input into the research and its implications, contributed to revising earlier versions of the manuscript, and approved the final version. Generative artificial intelligence was not used in any part of the manuscript writing.

XSL•FO
RenderX

Conflicts of Interest

The authors received salary support and funding to implement the digital adherence technology (DAT) program in the 5 countries as part of the Adherence Support Coalition to End TB (ASCENT) project, led by KNCV. Additionally, KNCV receives funding to support the Global DAT Task Force Secretariat and, by extension, to provide technical assistance to countries that are planning or implementing DAT programs.

Multimedia Appendix 1
Additional analysis.
[DOCX File, 149 KB - publichealth_v11i1e62881_app1.docx]

Checklist 1
STROBE checklist.
[DOCX File, 41 KB - publichealth_v11i1e62881_app2.docx]

References

1. Global tuberculosis report 2024. World Health Organization (WHO). 2024. URL: https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024 [accessed 2025-07-08]
2. Guidelines for treatment of drug-susceptible tuberculosis and patient care, 2017 update. World Health Organization (WHO). 2017. URL: https://iris.who.int/bitstream/handle/10665/255052/9789241550000-eng.pdf [accessed 2025-07-08]
3. Dye C, Watt CJ, Bleed DM, Hosseini SM, Raviglione MC. Evolution of tuberculosis control and prospects for reducing tuberculosis incidence, prevalence, and deaths globally. JAMA 2005 Jun 8;293(22):2767-2775. [doi: 10.1001/jama.293.22.2767] [Medline: 15941807]
4. Mohamed MS, Zary M, Kafie C, et al. The impact of digital adherence technologies on health outcomes in tuberculosis: a systematic review and meta-analysis. medRxiv. Preprint posted online on Feb 3, 2024. [doi: 10.1101/2024.01.31.24302115]
5. Bahukudumbi S, Chilala CI, Foster N, et al. Contextual factors influencing implementation of tuberculosis digital adherence technologies: a scoping review guided by the RE-AIM framework. BMJ Glob Health 2025 Feb 13;10(2):e016608. [doi: 10.1136/bmjgh-2024-016608] [Medline: 39947716]
6. Chilala CI, Foster N, Bahukudumbi S, et al. Implementation outcomes of tuberculosis digital adherence technologies: a scoping review using the RE-AIM framework. BMJ Glob Health 2025 Feb 13;10(2):e016535. [doi: 10.1136/bmjgh-2024-016535] [Medline: 39947715]
7. de Groot LM, Straetemans M, Maraba N, et al. Time trend analysis of tuberculosis treatment while using digital adherence technologies-an individual patient data meta-analysis of eleven projects across ten high tuberculosis-burden countries. Trop Med Infect Dis 2022 Apr 22;7(5):65. [doi: 10.3390/tropicalmed7050065] [Medline: 35622692]
8. Tadesse AW, Sahile M, Foster N, et al. Cluster-randomized trial of digital adherence technologies and differentiated care to reduce poor end-of-treatment outcomes and recurrence among adults with drug-sensitive pulmonary TB in ethiopia. medRxiv. Preprint posted online on May 9, 2024. [doi: 10.1101/2024.05.09.24307117]
9. Jerene D, van Kalmthout K, Levy J, et al. Effect of digital adherence technologies on treatment outcomes in people with drug-susceptible tuberculosis: four pragmatic, cluster-randomised trials. Lancet 2025;405(10485):1155-1166. [doi: 10.1016/S0140-6736(24)02847-2]
10. 99DOTS deployments. 99DOTS. URL: https://www.99dots.org/ [accessed 2025-07-07]
11. Jerene D, Levy J, van Kalmthout K, et al. Effectiveness of digital adherence technologies in improving tuberculosis treatment outcomes in four countries: a pragmatic cluster randomised trial protocol. BMJ Open 2023 Mar 14;13(3):e068685. [doi: 10.1136/bmjopen-2022-068685] [Medline: 36918242]
12. Tadesse AW, Mohammed Z, Foster N, et al. Evaluation of implementation and effectiveness of digital adherence technology with differentiated care to support tuberculosis treatment adherence and improve treatment outcomes in Ethiopia: a study protocol for a cluster randomised trial. BMC Infect Dis 2021 Nov 10;21(1):1149. [doi: 10.1186/s12879-021-06833-x] [Medline: 34758737]
13. Madden N, Tadesse AW, Leung CL, et al. Process evaluation of pragmatic cluster-randomized trials of digital adherence technologies for tuberculosis treatment support: a mixed-method study in five countries. Trop Med Infect Dis 2025 Mar 6;10(3):68. [doi: 10.3390/tropicalmed10030068] [Medline: 40137822]
14. Tadesse AW, Mganga A, Dube TN, et al. Feasibility and acceptability of the smart pillbox and medication label with differentiated care to support person-centered tuberculosis care among ASCENT trial participants - a multicountry study. Front Public Health 2024;12:1327971. [doi: 10.3389/fpubh.2024.1327971] [Medline: 38444445]
15. Leung CL, Alacapa J, Tasca BG, et al. Digital adherence technologies and differentiated care for tuberculosis treatment and their acceptability among persons with tuberculosis, health care workers, and key informants in the Philippines: qualitative interview study. JMIR Hum Factors 2024 Jul 23;11:e54117. [doi: 10.2196/54117] [Medline: 39042889]

XSL•FO
RenderX

16. Deyanova N, Madden N, Fielding K, et al. Digital adherence monitoring of tuberculosis treatment in Ukraine: analysis of adherence levels and patient-provider communication before and during the war in Ukraine [Abstract]. The International Journal of Tuberculosis and Lung Disease (IJTLD) 2022 Nov;26:S52 [FREE Full text]
17. Bommakanti KK, Smith LL, Liu L, et al. Requiring smartphone ownership for mHealth interventions: who could be left out? BMC Public Health 2020 Jan 20;20(1):81. [doi: 10.1186/s12889-019-7892-9] [Medline: 31959145]
18. Doyle AM, Bandason T, Dauya E, et al. Mobile phone access and implications for digital health interventions among adolescents and young adults in Zimbabwe: cross-sectional survey. JMIR Mhealth Uhealth 2021 Jan 13;9(1):e21244. [doi: 10.2196/21244] [Medline: 33439136]
19. Mukora R, Maraba N, Orrell C, et al. Qualitative study exploring the feasibility of using medication monitors and a differentiated care approach to support adherence among people receiving TB treatment in South Africa. BMJ Open 2023 Mar 3;13(3):e065202. [doi: 10.1136/bmjopen-2022-065202] [Medline: 36868589]
20. Zary M, Mohamed MS, Kafie C, et al. The performance of digital technologies for measuring tuberculosis medication adherence: a systematic review. . 2024(7) p. e015633.
21. Kiwanuka N, Kityamuwesi A, Crowder R, et al. Implementation, feasibility, and acceptability of 99DOTS-based supervision of treatment for drug-susceptible TB in Uganda. PLOS Digit Health 2023 Jun;2(6):e0000138. [doi: 10.1371/journal.pdig.0000138] [Medline: 37390077]
22. Charalambous S, Maraba N, Jennings L, et al. Treatment adherence and clinical outcomes amongst in people with drug-susceptible tuberculosis using medication monitor and differentiated care approach compared with standard of care in South Africa: a cluster randomized trial. EClinicalMedicine 2024 Jul 27;75:102745. [doi: 10.1016/j.eclinm.2024.102745] [Medline: 39170937]
23. Liu X, Thompson J, Dong H, et al. Digital adherence technologies to improve tuberculosis treatment outcomes in China: a cluster-randomised superiority trial. Lancet Glob Health 2023 May;11(5):e693-e703. [doi: 10.1016/S2214-109X(23)00068-2] [Medline: 37061308]
24. Velen K, Nguyen TA, Pham CD, et al. The effect of medication event reminder monitoring on treatment adherence of TB patients. Int J Tuberc Lung Dis 2023 Apr 1;27(4):322-328. [doi: 10.5588/ijtld.22.0500] [Medline: 37035979]
25. Wei X, Hicks JP, Zhang Z, et al. Effectiveness of a comprehensive package based on electronic medication monitors at improving treatment outcomes among tuberculosis patients in Tibet: a multicentre randomised controlled trial. Lancet 2024 Mar 9;403(10430):913-923. [doi: 10.1016/S0140-6736(23)02270-5] [Medline: 38309280]
26. Data for: “engagement with digital adherence technologies among adults with drug-susceptible tuberculosis: data from cluster randomized trials in five countries” – data codebook. LSHTM Data Compass. URL: https://datacompass.lshtm.ac.uk/id/eprint/4695/3/Participant_doseday_data_codebook.html [accessed 2025-07-07]

Abbreviations

ASCENT: Adherence Support Coalition to End TB
DAT: digital adherence technology
DOTS: directly observed therapy
DS-TB: drug-susceptible tuberculosis
HCP: health care provider
TB: tuberculosis
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 03.06.24; peer-reviewed by A Mtenga, I Prasetyo, S Goyal, T Li; revised version received 05.06.25; accepted 06.06.25; published 28.07.25.

Please cite as:
Alacapa J, Tadesse AW, Deyanova N, Dube T, Mganga A, Powers R, van Rest J, Madden N, Efo E, Charalambous S, van Kalmthout K, Jerene D, Fielding K
Engagement With Digital Adherence Technologies as Measures of Intervention Fidelity Among Adults With Drug-Susceptible Tuberculosis and Health Care Providers: Descriptive Analysis Using Data From Cluster-Randomized Trials in Five Countries
JMIR Public Health Surveill 2025;11:e62881
URL: https://publichealth.jmir.org/2025/1/e62881
doi:10.2196/62881

© Jason Alacapa, Amare Worku Tadesse, Natasha Deyanova, Tanyaradzwa Dube, Andrew Mganga, Rachel Powers, Job van Rest, Norma Madden, Egwuma Efo, Salome Charalambous, Kristian van Kalmthout, Degu Jerene, Katherine Fielding. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.7.2025. This is an open-access article

XSL•FO
RenderX

distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Impact of a Mobile Money–Based Intervention on Maternal and Neonatal Health Outcomes in Madagascar: Cluster-Randomized Controlled Trial

Lisa Bogler[1,2], Dr rer pol; Bítia Vieira[2]; Harizaka Emmanuel Andriamasy[2,3], Dr med; Zavaniarivo Rampanjato[2,4,5], Dr med; Sebastian Vollmer[1], Prof Dr; Till Bärnighausen[4,6,7], Prof Dr; Andriamampianina Ralisimalala[8], Dr med; Julius Valentin Emmrich[2,4,9*], Dr med; Samuel Knauss[2,4,9*], Dr med

[1]Department of Economics & Centre for Modern Indian Studies, University of Goettingen, Göttingen, Germany

[2]Global Digital Last Mile Health Research Lab, Charité Center for Global Health, Charité - Universitätsmedizin Berlin, Berlin, Germany

[3]Doctors for Madagascar, Antananarivo, Madagascar

[4]Heidelberg Institute of Global Health, Medical Faculty and University Hospital, University of Heidelberg, Heidelberg, Germany

[5]Ministry of Public Health of the Republic of Madagascar, Antananarivo, Madagascar

[6]Department of Global Health and Population, Harvard T H Chan School of Public Health, Boston, MA, United States

[7]Africa Health Research Institute, Durban, South Africa

[8]Centre d'études et de recherche en économie de la santé, Faculté d'économie, de gestion et de sociologie, University of Antananarivo, Antananarivo, Madagascar

[9]Doctors for Madagascar, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Lisa Bogler, Dr rer pol
Department of Economics & Centre for Modern Indian Studies
University of Goettingen
Waldweg 26
Göttingen, 37073
Germany
Phone: 49 5513928966
Email: lisa.bogler@cemis.uni-goettingen.de

Abstract

Background: Financial barriers to accessing obstetric care persist in many low-resource settings. With increasing use of mobile phones, mobile money services appear as a promising tool to address this concern. Maternal health care is particularly suitable for a savings program using mobile money due to the predictable timing and costs of delivery. The mobile money–based Mobile Maternal Health Wallet (MMHW) intervention aimed to ease the burden of out-of-pocket expenses related to maternal health care by providing an accessible savings tool.

Objective: This study aimed to assess the impact of the MMHW on maternal and neonatal health outcomes.

Methods: We used a stratified cluster-randomized trial to assess the impact of the MMHW on maternal and neonatal health outcomes in the Analamanga region of Madagascar. All 63 eligible public sector primary care health facilities (*Centres de Santé de Base* [CSBs]) within 6 strata were randomized to either receive the intervention or not. We estimated intention-to-treat effects and contamination-adjusted effects following an instrumental variable approach. The primary outcomes included (1) delivery at a health facility, (2) antenatal care visits, and (3) total health care expenditure. Between March 2022 and December 2022, a total of 6483 women who had been pregnant between July 2020 and December 2021 were surveyed.

Results: Among women in catchment areas of treated CSBs, 38.79% (1297/3344) had heard of the MMHW, and 37.42% (485/1296) of them registered for the tool. There was considerable variation in uptake across treated CSBs. Descriptively, women in the catchment areas of treated CSBs were more likely to deliver in a facility and had more antenatal care visits and higher total health expenditures compared to women in control CSB catchment areas in the intention-to-treat and contamination-adjusted analyses. However, none of the effects were statistically significant.

XSL•FO
RenderX

Conclusions: While this study did not identify a statistically significant impact, the estimated contamination-adjusted effects suggest that the MMHW has potential to improve access to maternal care for women who are receptive to such a mobile money–based savings tool. Estimated population-level effects were much smaller, and this study was underpowered to detect such effects due to lower-than-anticipated uptake of the intervention.

Trial Registration: German Clinical Trials Register DRKS00014928; https://www.drks.de/search/de/trial/DRKS00014928

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-021-05694-8

(JMIR Public Health Surveill 2025;11:e70182) doi:10.2196/70182

KEYWORDS

mobile health; mHealth; health financing; maternal health; health wallet; mobile money; Madagascar; cluster-randomized trial

Introduction

Background

In several sub-Saharan African countries, health care continues to be primarily financed through out-of-pocket payments, exposing households to the risk of catastrophic health expenditures and impoverishment [1,2]. Household savings can mitigate the financial strain caused by unforeseen health care costs [1,3]. However, economically disadvantaged households frequently encounter unexpected and irregular expenses that hinder their efforts to achieve long-term saving goals [4]. Consequently, obstetric care access is often hindered by limited financial resources [5]. Expectant mothers from low- and middle-income households frequently refrain from seeking birth assistance or emergency obstetric care due to the substantial expenses associated with childbirth [6,7].

In Madagascar, ensuring access to and acceptance of health care services for pregnant women remains an ongoing challenge. The national health policy of Madagascar stipulates the free provision of antenatal care (ANC) and uncomplicated deliveries in all public health care facilities. Nevertheless, charges are often imposed for folic acid, iron supplements, laboratory tests, medications, and treatments, as well as for complications during delivery. These charges act as barriers limiting women's access to quality maternal care, particularly for those in economically disadvantaged circumstances [8,9]. The maternal mortality rate in 2020 was 392 per 100,000 live births [10] but is estimated to be up to 3 times higher in the poorest districts of the country [11]. Only approximately half of pregnant women in Madagascar complete 4 ANC visits as recommended by the World Health Organization, and over half of deliveries take place without qualified personnel [12].

Over the past 10 years, there has been astonishing growth in mobile phone subscription rates and mobile money (MM) accounts in Madagascar, as in several other sub-Saharan African countries [13-15]. MM, a commonly used term for mobile payment systems, allows individuals to securely store, send, and receive electronic money through a digital platform operated by a mobile phone carrier, serving as a viable alternative for physical cash. By leveraging simple technologies such as unstructured supplementary service data, these MM services can facilitate monetary transactions even in the absence of a traditional bank account, internet connectivity, or advanced smartphone features. It has been shown that providing a secure place for storing money can increase savings [16]. MM services could act as such a simple, accessible tool.

MM platforms do offer users the option to sign up for additional financial services, such as savings, credit, or insurance, either through mobile operators or third-party service providers. Given that nearly 80% of adults in sub-Saharan Africa lack access to formal banking services, economies in the region are increasingly reliant on mobile payment systems [17]. This technological advancement is also being harnessed to introduce mobile payment–based hospital insurance or savings mechanisms, allowing economically challenged households to allocate funds specifically for health care needs [14,18,19]. The swift accessibility of cash, remittances, electronic savings accounts, and insurance plans provided by MM is driving its growing use in the health care sector across sub-Saharan Africa [17]. MM users have a lower risk of catastrophic health care expenses during emergencies and are less prone to reducing expenditures on education or food [20-24]. This presents an opportunity for inclusive solutions to address health care coverage gaps.

Maternal and childbirth-related health care emerges as a particularly suitable candidate for a savings program due to its relatively predictable timing and costs.

In Madagascar, a significant proportion of pregnant women rely on personal cash savings to financially prepare for giving birth in health care institutions. However, this approach is particularly arduous for impoverished populations [8]. Addressing this issue, the nongovernmental organization Doctors for Madagascar developed a tool known as the Mobile Maternal Health Wallet (MMHW) to ease the burden of out-of-pocket expenses related to maternal health care [25]. Accessible via the unstructured supplementary service data menu on any mobile phone, this software platform offers expectant mothers the capability to save, pay, and receive MM and electronic vouchers for maternal health care services at participating health care providers. The personal information and funds stored in the MMHW are linked to an individual SIM card, which functions on any type of mobile device regardless of whether it is a smartphone or has internet connectivity. Additional incentives provided through the MMHW encompass access to free ambulance services in case of referrals and a free obstetric ultrasound checkup. To encourage pregnant women to save using the MMHW, Doctors for Madagascar offers 50% matching to all deposits made. Collaborating with the Malagasy Ministry of Public Health, Doctors for Madagascar randomly selected 31 out of 63 public

XSL•FO
RenderX

primary care health facilities and 4 reference hospitals in Antananarivo, located in the Analamanga region of central Madagascar, for the implementation of the MMHW. To facilitate the integration of this digital payment system within health care facilities, tablets or mobile phones were distributed to the health care providers operating in the participating facilities.

Objectives

In this study, we evaluated the impact of the MMHW intervention on maternal and neonatal health outcomes among women living in the catchment area of the health facilities. The evaluation strategy closely follows the analysis strategy outlined in the published study protocol [26].

Methods

Study Design

This cluster-randomized controlled trial was implemented in 3 districts of the Analamanga region in Madagascar (Antananarivo Renivohitra, Atsimondrano, and Avaradrano). Eligible for inclusion in the trial were all 63 public sector primary care health facilities (*Centres de Santé de Base* [CSBs]) that provided ANC at the time of the intervention. Due to budget constraints, not all of these could receive the intervention. Therefore, CSBs were randomized within 6 strata to either receive the intervention or not. The control group consisted of 32 CSBs, and the intervention group consisted of 31 CSBs as well as 4 public reference hospitals for maternal care. All health facilities in the intervention group received the intervention package by May 2020. Since then, registration for MMHW was open, and activities to encourage registration were ongoing until December 2022.

For the evaluation of the intervention, we used several data sources. A quantitative population-based survey was conducted during household visits between March 2022 and December 2022. For this survey, women who had completed their pregnancy between July 1, 2020, and December 31, 2021, were interviewed. They answered questions about socioeconomic characteristics, pregnancy, delivery, expenses, saving behavior, decision-making, postpartum depression, and their use of the MMHW. In addition, in households in which there was a woman matching the eligibility criteria but she was not present at the time of the visit, a short interview was conducted with 1 household member present to capture socioeconomic characteristics. This allowed us to compare the sample of women interviewed with those women who were absent but matched the eligibility criteria. The second data source was health facility records. Between August 2022 and September 2022 (with 3 exceptions for whom data were gathered in April 2023), a survey of all health facilities included in the randomization was conducted. Health facility staff was asked to share records on ANC visits, deliveries, and expenses from 2020 and 2021. In addition, respondents at facilities where the MMHW was implemented answered questions about their use of and satisfaction with the MMHW.

This trial was registered with the German Clinical Trials Register (identifier: DRKS00014928) on March 12, 2021. Registration took place after randomization and intervention implementation but before the start of data collection and outcome assessment. A study protocol outlining the evaluation design and estimation approach followed in this analysis was submitted on March 22, 2021, and published in October 2021 [26].

Participants

For inclusion in the randomization, all CSBs in the study region that provided ANC at the time of the intervention were eligible. Eligibility was determined based on 2017 records obtained from health authorities. In addition, the 4 public reference hospitals for maternal care in Antananarivo, as defined by the Malagasy Ministry of Public Health, were included in the intervention group. Each health facility represented 1 cluster.

All health facilities included in the randomization, as well as the reference hospitals, were included in the survey of health care providers.

For the quantitative population survey, we randomly selected 4 census enumeration areas (*fokontany*) in each catchment area of the included health facilities. In catchment areas with ≤4 *fokontanys*, all were included for data collection. In each *fokontany*, data collectors mapped out the area and knocked on every second door to identify women eligible for survey participation. Women were eligible for survey participation if they had completed their pregnancy between July 1, 2020, and December 31, 2021; were aged ≥18 years; and provided verbal consent to the interview.

Randomization and Masking

The unit of randomization for the intervention was the CSB. There was previous evidence suggesting that ANC and delivery quality varied significantly with a facility's patient volume and capacity to perform deliveries [27]. Therefore, all health facilities were stratified into 6 subgroups according to the facility's ANC visit volume (<1750, 1750-3500, and >3500 ANC visits per month in 2017) and the capacity to perform deliveries (none vs any deliveries). Within each stratum, the facilities were sorted in descending order of reported ANC visits. A senior biostatistician then performed pairwise randomization. Pairing 2 health facilities following each other in the ordered list, one was randomly assigned to the intervention and the other to the control group. In addition, 4 reference hospitals were assigned to receive the intervention to ensure that women could use the MMHW in case of referral for complications. Due to their nonrandom assignment to the intervention group, reference hospitals were excluded from the analysis of facility-level outcomes. As they do not have their own defined catchment areas as CSBs have from where women respondents were sampled, the reference hospitals do not affect the analysis of individual-level outcomes. During the intervention rollout, 3 CSBs assigned to the intervention group turned out not to have the technical capacity for the intervention. Of these 3 CSBs, 2 (67%) did not have the required internet connection, and the third (33%) did not manage any pregnancies. Therefore, 2 additional CSBs that were semiprivate and, therefore, ineligible for trial inclusion in 2017 and became fully public only in 2019 were added to the intervention group. These 2 additional CSBs were excluded from analysis as their

XSL•FO
RenderX

assignment to the intervention group was not random. At the individual level, women living in the catchment area of these 2 CSBs were excluded from analysis. Due to the nature of the intervention, neither health facility staff nor patients visiting the facilities were blinded to intervention allocation.

Procedures

The MMHW package consisted of 2 major components. The main component was directed at pregnant women attending ANC visits at health care facilities. Women who registered with the MMHW received a SIM card with which they were able to save in the MMHW and to pay using these savings at the health care facilities in the intervention group. Deposits could be made and used at any time during pregnancy and up to 30 days after childbirth. Additional benefits were provided for those who registered—costs for preventive drugs such as iron and folic acid tablets and mebendazole received during ANC visits, as well as costs for transfer via ambulance in the event of obstetric and neonatal emergencies to the reference hospitals, were covered. A savings incentive of 50% of the saved amount to be used for any pregnancy-related and newborn care up to 30 days after birth was provided through external donors. In addition, users could receive a free obstetric ultrasound checkup. Pregnant women were encouraged to register with the MMHW through a sensitization campaign. Sensitization took place in communities through awareness-raising mass sensitization events and home visits to pregnant women by community health workers and during group meetings organized by workers of Doctors for Madagascar at health care facilities during the facilities' official dates for ANC visits.

The second component was directed at CSBs and provided maternal and neonatal health training, capacity-building, and refresher courses for health workers in maternity wards and ANC departments. The trainings were conducted by trainers from the Ministry of Public Health. The goal of these trainings was to ensure quality of care matching the expected increased demand for health care services. Financial incentives to participate in the intervention were also provided. For each delivery of a beneficiary, the health care professional in charge of the delivery received a bonus of MGA 10,000 (US $2.44 according to the OANDA currency converter rate for January 1, 2021). The pharmacists responsible for taking photo evidence of claims received a bonus of MGA 2000 (US $0.51) per patient and month for each claim sent via the MMHW platform. In addition, medical equipment, including blood pressure monitors, stethoscopes, examination beds, bedpans, and baby scales, was distributed to the corresponding maternity units.

Outcomes

In this study, we analyzed the outcomes as described in the published study protocol, which are based on the quantitative household survey and the health facility records [26]. Primary outcomes were (1) facility-based delivery; (2) number of ANC visits per woman; and (3) total health expenditure during pregnancy, delivery, and the neonatal period per woman. Secondary outcomes at the individual level were related to ANC (number of diagnoses received and whether they had ultrasound checkups or received iron and folic acid tablets or syrup), the delivery process (whether they had complications, planned or emergency cesarean section, or neonatal mortality), finances (health savings, contributions to savings, relative health care expenditure, and indication of financial distress), and satisfaction after delivery (with the health facility, the health system, and life), as well as postpartum depression score. Secondary outcomes at the facility level included maternal and neonatal mortality, number of ANC drugs distributed, and public sector costs. Compared with the full list of proposed outcomes in the protocol, we excluded those outcomes that could not be analyzed sensibly following the planned estimation approach as they would only be relevant for women who actually registered with the MMHW (ie, MMHW use, recommendation of MMHW, and time of sign-up). Furthermore, maternal mortality was only analyzed based on data from health facility records as the outcome was not captured accurately in the quantitative household survey. A full list of outcomes and details regarding their coding can be found in Tables S1 and S2 in Multimedia Appendix 1.

Statistical Analysis

The analysis in this study followed the estimation strategy outlined in the published study protocol [26]. The primary analysis was an intention-to-treat (ITT) estimation of all outcomes. We regressed each individual-level outcome on an indicator of living in the catchment area of either a treated or control health facility and clustered SEs at the facility level. Outcomes at the facility level were regressed on an indicator of treatment assignment. For both individual- and facility-level outcomes, we added strata dummy variables to account for stratification during randomization. Because of the randomized treatment assignment, a comparison of postintervention outcomes between the intervention and control group, on average, provided the unbiased ITT effect. This ITT effect captured the real-life policy impact of the intervention. As a secondary analysis, we conducted a contamination-adjusted analysis for outcomes at the individual level using the random assignment to the intervention versus control group as an instrumental variable (IV). This estimation of the local average treatment effect provided a measure of the effect size that would have been attained without contamination assuming that those who did not take up the intervention would behave similarly once registered for the tool as those who did take up the intervention. Contamination in this case would occur if women living in an intervention area did not register for the MMHW and women living in a control area did. This is very possible in this context in which women do not necessarily choose the facility in the catchment area they live in. Actual registration with the MMHW was instrumented with treatment assignment in the IV analysis. In addition, we instrumented having heard of the MMHW and having used the MMHW for paying with treatment assignment. For ease of interpretation and comparison between ITT and IV estimation, we used linear regressions for all outcomes. We present absolute and relative effect sizes (relative to the mean of the outcome in the control group). As a robustness check, we reran estimations excluding the 3 CSBs that were assigned to the intervention but could not use it. As a further robustness check on individual-level outcomes, we reran estimations with controls for individual characteristics that were imbalanced (employment status, access to a mobile

XSL•FO
RenderX

phone, and television ownership). We assessed statistical significance at the 5% level. All statistical analyses were conducted using Stata (version 16; StataCorp).

Power Calculations

A power calculation was presented in the study protocol [26]. Minimum detectable differences between the intervention and control groups were calculated for the 3 primary outcomes using methods for cluster-randomized controlled trials. We assumed that 50% of survey participants in the catchment areas of intervention health facilities would adopt the MMHW. Power calculations showed that the study would have 80% power to detect a 4–percentage point increase in the rate of facility-based deliveries, an increase of 0.08 average ANC visits per woman, and an increase in total health expenditure of MGA 1700 (US $0.41 according to the OANDA currency converter rate for August 16, 2022) if 2300 women were surveyed per group. However, intervention uptake differed from this assumption, as discussed in the following sections.

Ethical Considerations

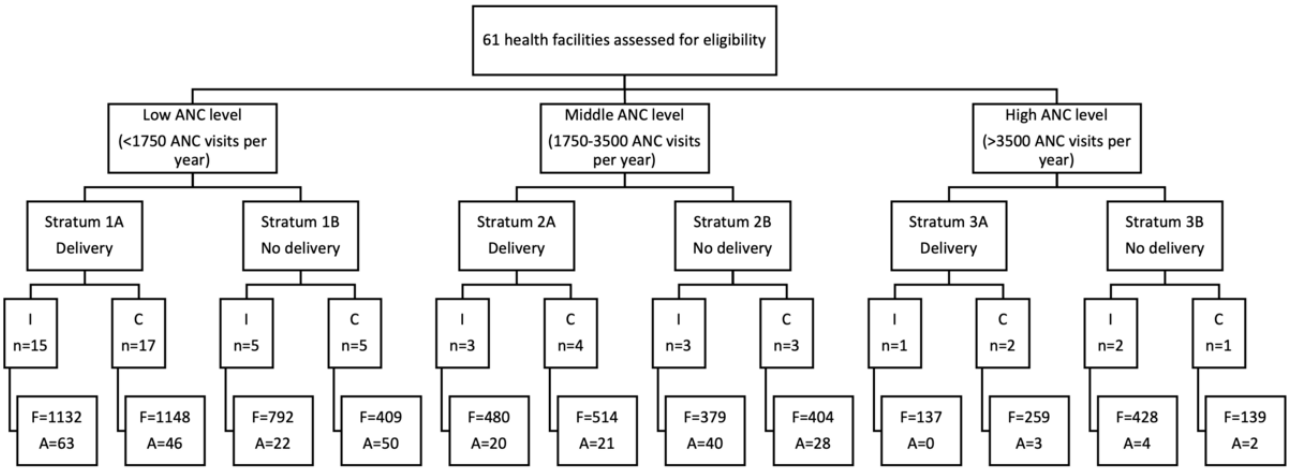
This trial was approved by the Institutional Review Board of Heidelberg University on February 3, 2020 (S-428/2019). The study was conducted in accordance with the WMA Declaration of Helsinki. Informed consent was obtained orally from all study participants before their interviews. Privacy and confidentiality of survey participants' data and identity was maintained. Survey participants received a bar of soap as a locally appropriate in-kind compensation for their time.

Results

Sample Description and Balance

A total of 6323 women in the catchment areas of the CSBs in the control and intervention groups were identified as eligible for participation and present at the time of the home visit, of whom 6243 (98.7%) completed the full interview. The interview about eligible women who were not present at the time of the visit was completed by 300 respondents (ie, household members who were present and could answer questions capturing the basic characteristics of the eligible women). Figure 1 provides a flowchart of the sample.

Figure 1. Flowchart of sample selection. A: number of interviews for absent eligible women; ANC: antenatal care; C: control group; F: number of full interviews; I: intervention group; n: number of facilities.



Most women surveyed were aged <35 years (2673/6243, 43.82% between 17 and 24 years; 2741/6243, 43.91% between 25 and 34 years) with 1 or 2 children (2282/6235, 37% with 1 child; 1990/6235, 31.92% with 2 children), and the large majority was married (see Table 1 for a summary of socioeconomic characteristics). Some women reported no children as their children had passed away and the eligibility criteria were based on the timing of past pregnancies and not childbirth. A large share of women (4437/6240, 71.11%) had completed secondary or higher education, but less than half (3061/6230, 49.13%) reported receiving regular income. Almost half (2820/6242, 45.18%) were currently not working, which was likely driven by the fact that these were women with very young children. Access to a mobile phone, a requirement for using the intervention, was common but not universal as 58.09% (3626/6242) reported having access to a mobile phone and 77.96% (2827/3626) of those with access reported that the phone was their own. The large majority of respondents (5634/6239, 90.3%) were not covered by any health insurance.

XSL•FO
RenderX

Table 1. Sample characteristics at the individual level (N=6243)[a].

Characteristic	Full interviews					Absent sample		
	Number of observations per characteristic, n(%)	Women in the total sample, n (%)	Women in the control group, n (%)	Women in the intervention group, n (%)	*P* value	Number of observations per characteristic, n (%)	Women, n (%)	*P* value
Age of the woman (y)	6243 (100)					299 (99.7)		
17-24		2673 (42.82)	1260 (43.52)	1413 (42.2)	.29		97 (32.4)	<.001
25-34		2741 (43.91)	1281 (44.25)	1460 (43.61)	.61		157 (52.5)	.003
35-55		829 (13.28)	354 (12.23)	475 (14.19)	.02		45 (15.1)	.38
Married	6242 (99.98)	5680 (91)	2649 (91.5)	3031 (90.56)	.19	—[b]	—	—
Number of children	6235 (99.87)					—	—	—
0		17 (0.27)	7 (0.24)	10 (0.3)	.67			
1		2282 (36.6)	1092 (37.77)	1190 (35.59)	.07			
2		1990 (31.92)	915 (31.65)	1075 (32.15)	.67			
3		1098 (17.61)	489 (16.91)	609 (18.21)	.18			
4		495 (7.94)	229 (7.92)	266 (7.95)	.96			
≥5		353 (5.66)	159 (5.5)	194 (5.8)	.61			
Level of education	6240 (99.95)					—	—	—
No education		104 (1.67)	47 (1.62)	57 (1.7)	.80			
Primary education		1699 (27.23)	800 (27.63)	899 (26.88)	.50			
Secondary education		3711 (59.47)	1749 (60.41)	1962 (58.65)	.16			
Higher education		726 (11.63)	299 (10.33)	427 (12.77)	.003			
Received regular income	6230 (99.79)	3061 (49.13)	1398 (48.32)	1663 (49.84)	.23			
Type of employment	6242 (99.98)					—	—	—
Wage earner		357 (5.72)	157 (5.42)	200 (5.98)	.35			
Self-employed		2152 (34.48)	915 (31.61)	1237 (36.96)	<.001			
Family farm, livestock, or fishing		899 (14.40)	572 (19.76)	327 (9.77)	<.001			
Student		29 (0.46)	14 (0.48)	15 (0.45)	.84			
Not working		2820 (45.18)	1257 (43.42)	1563 (46.7)	.009			
Other		131 (2.10)	65 (2.25)	66 (1.97)	.45			
Access to a mobile phone	6242 (99.98)	3626 (58.09)	1632 (56.39)	1994 (59.56)	.01	300 (100)	204 (68)	.001
Phone ownership	3626 (58.08)					—	—	—
Own phone		2827 (77.96)	1255 (76.90)	1572 (78.84)	.14			
Spouse's phone		642 (17.71)	298 (18.26)	344 (17.25)	.38			
Someone else's phone		157 (4.33)	79 (4.84)	78 (3.91)	.17			
Used mobile money (excluding mTOMADY)	5152 (82.52)	2371 (46.02)	1218 (46.68)	1153 (45.34)	.33			
Television ownership	6231 (99.81)	3184 (51.10)	1306 (45.24)	1878 (56.16)	<.001	300 (100)	204 (68)	<.001
Vehicle ownership	6234 (99.86)					298 (99.3)		

XSL•FO
RenderX

Characteristic	Full interviews					Absent sample		
	Number of observations per characteristic, n(%)	Women in the total sample, n (%)	Women in the control group, n (%)	Women in the intervention group, n (%)	*P* value	Number of observations per characteristic, n (%)	Women, n (%)	*P* value
No vehicle		4808 (77.13)	2201 (76.16)	2607 (77.96)	.09		214 (71.8)	.03
Bicycle		824 (13.22)	418 (14.46)	406 (12.14)	.007		32 (10.7)	.22
Motorcycle		568 (9.11)	268 (9.27)	300 (8.97)	.68		39 (13.1)	.02
Car		175 (2.81)	84 (2.91)	91 (2.72)	.66		29 (9.7)	<.001
Other vehicles		46 (0.74)	24 (0.83)	22 (0.66)	.43		3 (1)	.60
Insurance	6239 (99.94)					—	—	—
None		5634 (90.30)	2620 (90.53)	3014 (90.1)	.57			
Community-based medical health insurance		36 (0.58)	13 (0.45)	23 (0.69)	.22			
Microfinance-based scheme		4 (0.06)	1 (0.03)	3 (0.09)	.39			
Employment-based scheme		429 (6.88)	202 (6.98)	227 (6.79)	.76			
Association or NGO[c]		32 (0.51)	12 (0.41)	20 (0.6)	.31			
Other		111 (1.78)	48 (1.66)	63 (1.88)	.50			

[a]This table shows the shares of respondents for each characteristic in percentages, for the sample who took part in the full interviews (total sample, control, and intervention), and the sample who gave interviews for women absent at the time of the visit. The columns Number of observations per characteristic, n (%) provide the number of women in the respective sample for whom the information was available. The *P* values for the full interviews correspond to a 2-tailed *t* test of the difference between the control and intervention group. The *P* values for absent interviews correspond to a 2-tailed *t* test of the difference between the interviews about absent women and the full interviews.

[b]Information not available for absent sample.

[c]NGO: nongovernmental organization.

The sample of respondents who completed the full interview slightly differed from the eligible women who were absent at the time of the visit (Table 1). Absent women were less likely to be aged <25 years and more likely to be aged 25 to 34 years than women present at their household residence; more likely to have access to a mobile phone; and more likely to own a television, motorcycle, or car. This suggests that absent women were more likely to be working, possibly because it had not been their first birth as they were older and they were slightly wealthier.

Randomization of health facilities led to a balanced sample of health facilities (Table 2) and of respondents in terms of most characteristics (Table 1). However, some statistically significant differences between the control and intervention groups emerged at the individual level, especially regarding type of employment and asset ownership. Respondents in the intervention group were more likely to be self-employed or not working and less likely to work on a family farm or in livestock or fishing. Access to a mobile phone was slightly higher in the intervention group, as was the probability of owning a television.

XSL•FO
RenderX

Table 2. Facility characteristics (2020 and 2021)[a].

Characteristic	Number of observations		Values, mean (SD)		*P* value	2020			2021		
	2020	2021	2020	2021		Control	Intervention	*P* value	Control	Intervention	*P* value
Number of ANC[b] visits	57	56	1416 (1194)	1379 (1230)	.38	1453 (1270)	1429 (1128)	.94	1297 (1275)	1468 (1197)	.61
Number of deliveries	56	50	236 (333)	209 (273)	.10	234 (393)	222 (255)	.89	201 (301)	217 (248)	.83
Maternal mortality	39	36	0.003 (0)	0.005 (0)	.31	0.002 (0)	0.004 (0)	.43	0.007 (0)	0.003 (0)	.38
Newborn mortality	38	37	0.000 (0)	0.000 (0)	.32	0.001 (0)	0.000 (0)	.30	0.000 (0)	0.000 (0)	—[c]
Number of iron and folic acid supplements distributed	55	52	38,420 (31,086)	37,904 (31,984)	.76	40,530 (36,894)	35,317 (23,407)	.54	39,339 (38,026)	36,411 (25,228)	.74
Number of iron and folic acid supplements distributed per ANC visit	54	51	34 (40)	32 (26)	.79	37 (55)	28 (2)	.41	35 (37)	28 (2)	.34
Number of antiparasite medications distributed	56	51	312 (273)	310 (273)	.88	323 (303)	289 (239)	.64	300 (289)	315 (259)	.85
Expenses for ANC drugs (MGA)	57	51	1401 (285)	1399 (292)	.32	1403 (203)	1423 (355)	.79	1395 (212)	1406 (357)	.89
Expenses for delivery drugs (MGA)	41	36	7329 (4973)	7329 (4792)	—	6887 (3954)	8238 (5802)	.39	6915 (4210)	7660 (5296)	.65

[a]Data from the health facility survey excluding 2 *Centres de Santé de Base* (public sector primary care health facilities) and 4 reference hospitals assigned to the intervention group nonrandomly. The number of observations and means in the total sample are split by year and by control and intervention group. *P* values correspond to 2-tailed t tests of the difference in means across years and treatment.

[b]ANC: antenatal care.

[c]Values for the two groups, which are being compared, are equal.

Uptake of the MMHW Intervention

A total of 35 health facilities, including the 4 reference hospitals, received the intervention, whereas 32 facilities did not. Of the 35 health facilities that received the intervention, including the reference hospitals, 2 (6%) were dropped from the analysis because their assignment was not random. In the survey of health facilities, 3% (1/32) of the respondents in the control group reported that the MMHW was usable in their facility, whereas 12% (4/33) of the respondents in the intervention group reported that it was not usable. Of the latter, 75% (3/4) were those CSBs that could not implement the intervention due to technical issues. The fourth facility seemed to have stopped using the MMHW since intervention rollout.

Intervention uptake was lower among surveyed women than assumed in the power calculation. Among women in the catchment areas of treated CSBs, 38.79% (1297/3344) had heard of the MMHW (372/2891, 12.87% in the control group). Registration rates were at 14.51% (485/3344) in the intervention group and 3.46% (100/2890) in the control group, corresponding to 37.42% (485/1296) and 27% (100/371) of those who had heard of the MMHW, respectively. Among those who registered for the MMHW, 66.7% (320/480) in the intervention group and 65% (65/100) in the control group actually used it to pay for maternal care services. Uptake varied considerably across treated CSBs; the share of women who had heard of the MMHW ranged from 4% (1/24) to 78% (25/32), and the share of women who registered ranged from 0% (0/24) to 69% (22/32).

Impact on Primary and Secondary Outcomes

Results of the ITT and IV estimations for outcomes on the individual level are shown in Table 3 for primary outcomes, Table S3 of Multimedia Appendix 1 for secondary outcomes, and are illustrated in Figures 2-6. None of the coefficients for the 3 primary outcomes—delivery at a facility, number of ANC visits, and total expenditure for health—were statistically significant. Relative effect sizes were larger in the IV analysis than in the ITT analysis, and they were larger the stricter the instrumented indicator was (Figure 2). For delivery at a facility, the relative effect size in the ITT analysis was 0.009, meaning that the share of women delivering at health facilities increased

XSL•FO
RenderX

by 0.9% of the control mean (95% CI −0.049 to 0.067). When treatment was used as an instrument for having heard of the MMHW, the relative effect size increased to 0.033 (95% CI −0.174 to 0.239). When treatment was used as an instrument for having registered for the MMHW, the relative effect size was 0.077 (95% CI −0.408 to 0.562). When treatment was used as an instrument for having used the MMHW, the relative effect size was 0.110 (95% CI −0.616 to 0.836), suggesting an increase by 11% of the control mean. For the number of ANC visits, the respective relative effect sizes were 0.024 (95% CI −0.009 to 0.056) in the ITT analysis, meaning an increase of 2.4% over the control mean ANC visits, 0.088 (95% CI −0.026 to 0.203) for IV—heard of the intervention, 0.208 (95% CI −0.067 to 0.484) for the IV—registered with the intervention, and 0.326 (95% CI −0.088 to 0.739) for IV—used the intervention. For total expenditure for health, the respective relative effect sizes were 0.045 (95% CI −0.140 to 0.229) in the ITT analysis, meaning an increase of 4.5% over the control mean of total expenditure for health, 0.157 (95% CI −0.481 to 0.795) for IV—heard of the intervention, 0.374 (95% CI −1.160 to 1.908) for IV—registered with the intervention, and 0.555 (95% CI −1.789 to 2.898) for IV—used the intervention.

Figure 2. Relative effect sizes of intention-to-treat (ITT) and instrumental variable (IV) estimations for primary individual-level outcomes. The graph shows the estimated relative effect sizes with 95% CIs from the ITT analysis as well as the IV analysis where assignment to treatment was used as an instrument for having heard of the Mobile Maternal Health Wallet (MMHW), having registered with the MMHW, and having used the MMHW to pay. ANC: antenatal care.

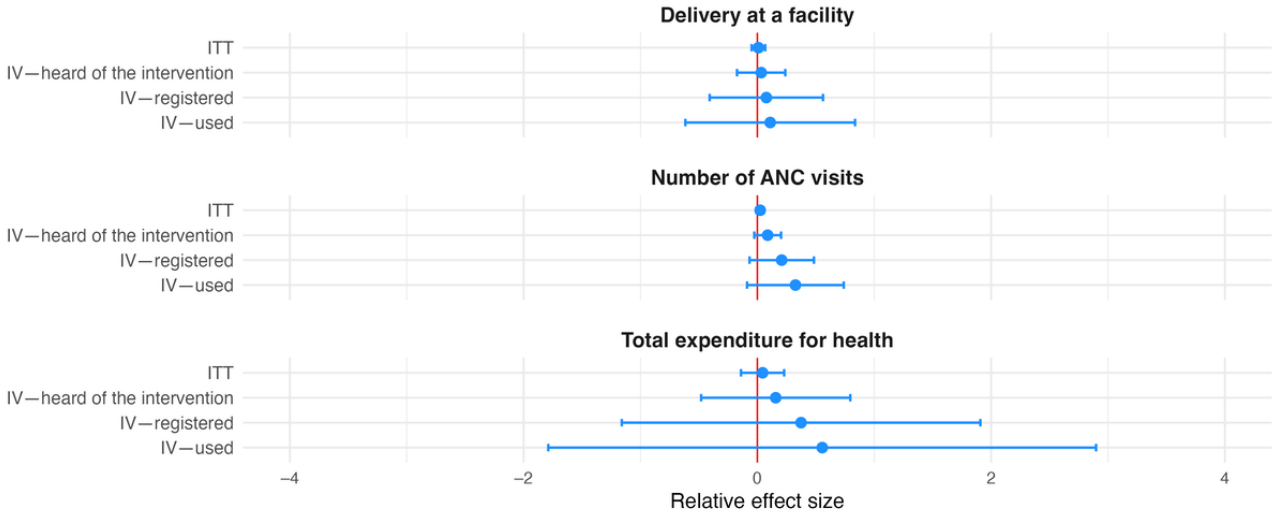
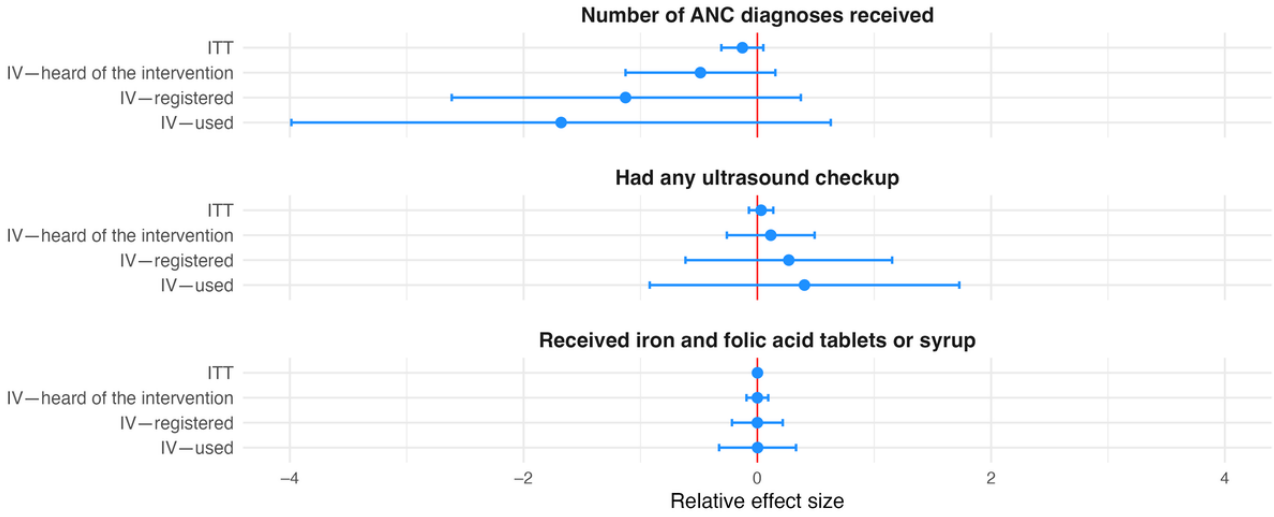


Figure 3. Relative effect sizes of intention-to-treat (ITT) and instrumental variable (IV) estimations for secondary individual-level antenatal care (ANC)–related outcomes. The graph shows the estimated relative effect sizes with 95% CIs from the ITT analysis as well as the IV analysis where assignment to treatment was used as an instrument for having heard of the Mobile Maternal Health Wallet (MMHW), having registered with the MMHW, and having used the MMHW to pay.



XSL•FO
RenderX

Table 3. Regression results of intention-to-treat (ITT) and instrumental variable (IV) estimations for individual-level outcomes[a].

Outcome	Number of clusters	ICC[b]	Number of observations	Mean (SD)	ITT			IV—heard of the intervention			IV—registered for the intervention			IV—used the intervention		
					Coefficient	*P* value	Relative effect size	Coefficient	*P* value	Relative effect size	Coefficient	*P* value	Relative effect size	Coefficient	*P* value	Relative effect size
Delivery at a facility	61	0.032	6185	0.780 (0.414)	0.007	.76	0.009	0.026	.76	0.033	0.060	.76	0.077	0.086	.77	0.110
Number of ANC[c] visits	61	0.021	6186	4.325 (1.515)	0.102	.15	0.024	0.381	.13	0.088	0.901	.14	0.208	1.408	.12	0.326
Total expenditure for health (MGA)	61	0.048	4692	208,021 (330,946)	9309	.63	0.045	32,891	.63	0.157	78,346	.63	0.374	116,198	.64	0.555

[a]This table shows the regression results of the ITT and IV estimations for primary outcomes. It reports the number of clusters and observations used for each outcome, the intraclass correlation coefficient, the mean value and SD of each outcome in the control group, the coefficient of being in the catchment area of a treated Centre de Santé de Base (CSB; public sector primary care health facility) with the corresponding *P* value, and the relative effect size (calculated as the coefficient divided by the mean) for both ITT and IV models. In all models, indicator variables for strata were included as controls, and SEs were clustered at the CSB level.

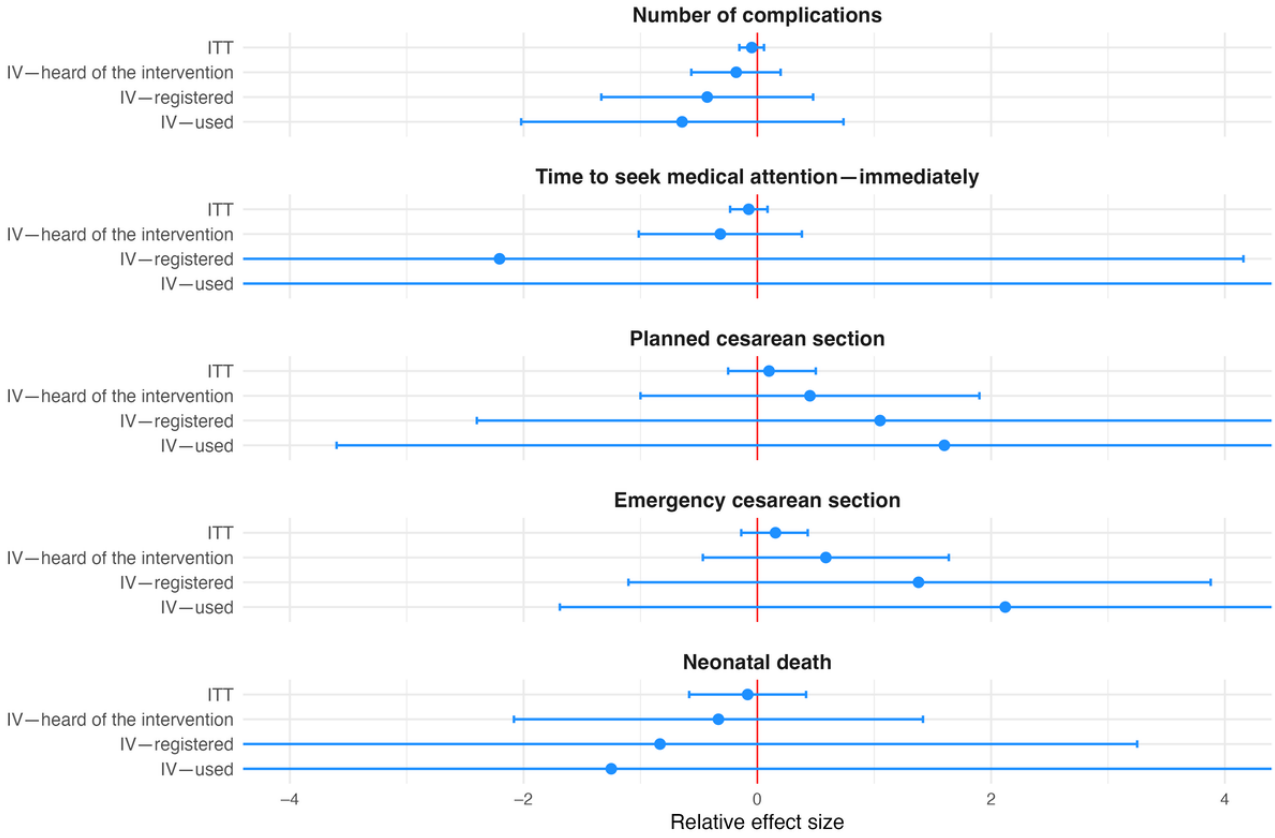
[b]ICC: intraclass correlation coefficient.

[c]ANC: antenatal care.

We grouped secondary outcomes into 4 categories. Coefficients for ANC- and delivery-related outcomes were all statistically insignificant. With the exception of the indicator for having received iron and folic acid tablets or syrup, they all followed the pattern described for primary outcomes (ie, larger relative effect sizes in the IV analysis; Figures 3 and 4). The same held true for the 4 outcomes grouped as being related to finances, including the amount of contributions to health savings, amount of health savings, relative health expenditure, and indication of financial distress (Figure 5). All coefficients for these outcomes went in the direction intended by the intervention, showing positive associations with the amount of savings and contributions and negative associations with relative health care expenditure and indication of financial distress, but no coefficient was statistically significantly different from 0 at the 5% level. The last group combined outcomes related to satisfaction and the postpartum depression score. None of the coefficients were statistically significant. Apart from the coefficient for life satisfaction, effect sizes increased across specifications, as described previously (Figure 6).

XSL•FO
RenderX

Figure 4. Relative effect sizes of intention-to-treat (ITT) and instrumental variable (IV) estimations for secondary individual-level delivery-related outcomes. The graph shows the estimated relative effect sizes with 95% CIs from the ITT analysis as well as the IV analysis where assignment to treatment was used as an instrument for having heard of the Mobile Maternal Health Wallet (MMHW), having registered with the MMHW, and having used the MMHW to pay. Elements not visible on the graph: time to seek medical attention in the IV—registered estimation (−2.206, 95% CI −8.573 to 4.159), time to seek medical attention in the IV—used estimation (−5.783, 95% CI −29.899 to 18.332), planned cesarean section in the IV—registered estimation (1.050, 95% CI −2.400 to 4.500), planned cesarean section in the IV—used estimation (1.600, 95% CI −3.600 to 6.850), emergency cesarean section in the IV—used estimation (2.121, 95% CI −1.690 to 5.931), neonatal death in the IV—registered estimation (−0.833, 95% CI −4.917 to 3.250), and neonatal death in the IV—used estimation (−1.250, 95% CI −7.417 to 5.000).



XSL•FO
RenderX

Figure 5. Relative effect sizes of intention-to-treat (ITT) and instrumental variable (IV) estimations for secondary individual-level finance-related outcomes. The graph shows the estimated relative effect sizes with 95% CIs from the ITT analysis as well as the IV analysis where assignment to treatment was used as an instrument for having heard of the Mobile Maternal Health Wallet (MMHW), having registered with the MMHW, and having used the MMHW to pay.

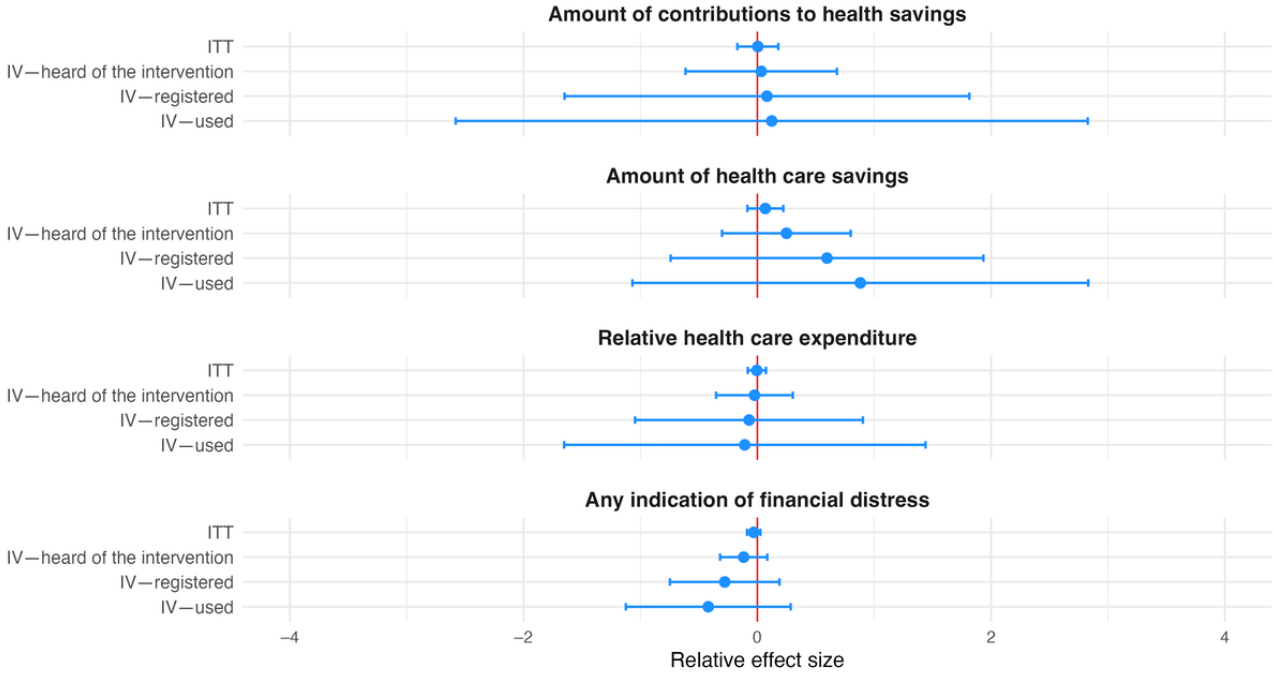
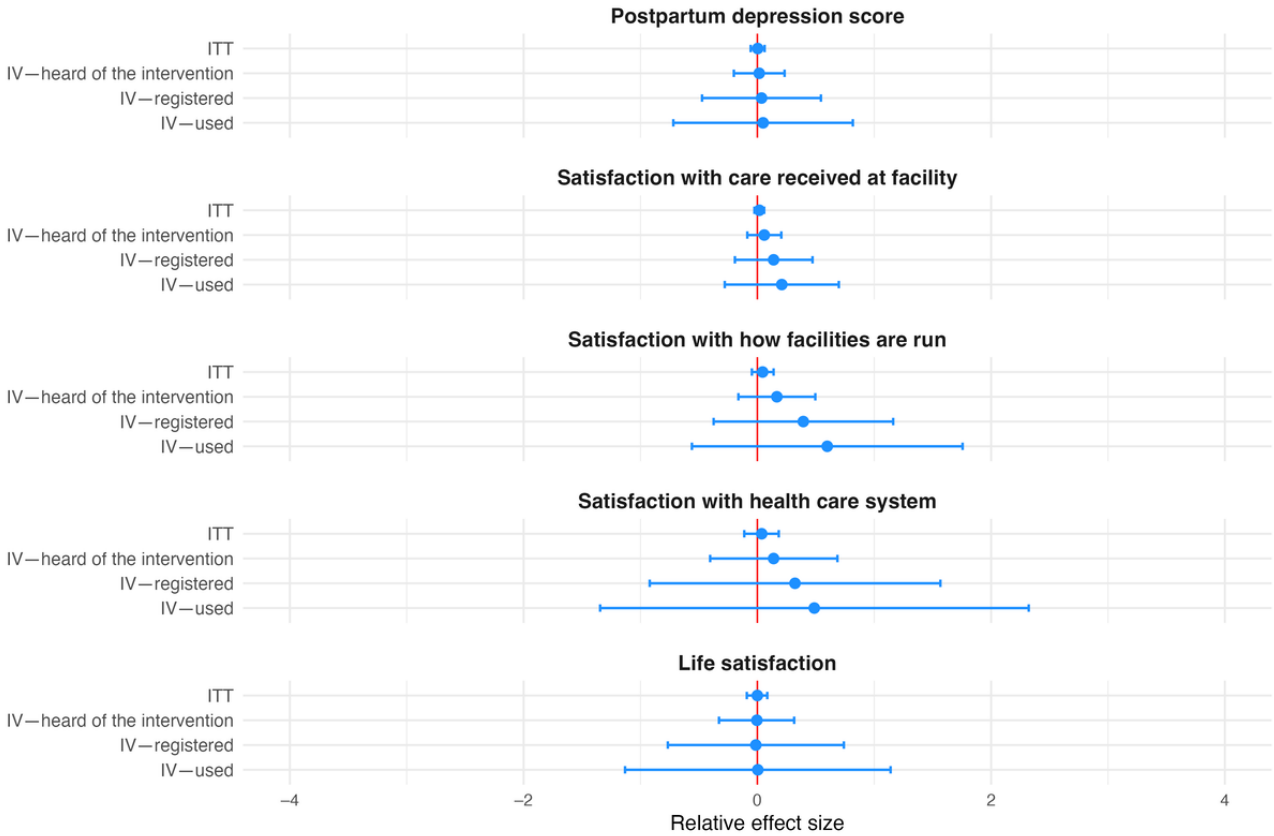


Figure 6. Relative effect sizes of intention-to-treat (ITT) and instrumental variable (IV) estimations for secondary individual-level satisfaction-related outcomes and postpartum depression score. The graph shows the estimated relative effect sizes with 95% CIs from the ITT analysis as well as the IV analysis where assignment to treatment was used as an instrument for having heard of the Mobile Maternal Health Wallet (MMHW), having registered with the MMHW, and having used the MMHW to pay.



XSL•FO
RenderX

In a robustness check, we excluded those CSBs that did not implement the MMHW for technical reasons. The results are shown in Table S4 in Multimedia Appendix 1. For all primary outcomes and several secondary outcomes (whether the women had any ultrasound checkups, whether they received iron and folic acid tablets or syrup, the number of complications, the time to seek medical attention, planned and emergency cesarean sections, the amount of health care savings, and postpartum depression score), the effect sizes went in the same direction as and were larger (in absolute terms) than in the full sample, but the coefficients were still not statistically significant. For other outcomes (number of ANC diagnoses, amount of contributions to health savings, indication of financial distress, and all indicators of facility and health care satisfaction), effect sizes were smaller but of the same sign. The effect sizes for relative health care expenditure and life satisfaction went in the opposite direction to that in the main specification. The effect size for children passing away before their first birthday was still negative but closer to 0.

In a further robustness check, we controlled for variables that appeared imbalanced. The results are shown in Table S5 in Multimedia Appendix 1. For 2 of the primary outcomes, delivery at a facility and total expenditure for health, the signs of the coefficients changed, indicating a reduction from the control mean. Among several secondary outcomes (whether the women had any ultrasound checkups, planned cesarean sections, 3 of 4 finance-related outcomes, and postpartum depression score), the sign of the coefficients also changed. Again, coefficients were not statistically significant.

The results of the ITT estimations at the facility level are shown in Table 4. Regarding outcomes at the individual level, no statistically significant impact on facility-level outcomes was found for outcomes either in 2020 or 2021. Effects on the number of iron and folic acid supplements and antiparasite medications went in the opposite direction of what was intended by the intervention. In the robustness checks excluding the CSBs in the intervention group that did not implement the MMHW, there were also no statistically significant effects (Table S6 in Multimedia Appendix 1).

XSL•FO
RenderX

Table 4. Regression results of intention-to-treat (ITT) estimations for facility-level outcomes[a].

Outcome	Number of observations	Values, mean (SD)	Coefficient	*P* value	Relative effect size
Outcomes based on 2020 records					
Number of ANC[b] visits	57	1429 (1128)	42	.81	0.029
Number of deliveries	56	222 (255)	41	.43	0.180
Maternal mortality	39	0.004 (0)	0.002	.42	0.667
Newborn mortality	38	0.000 (0)	−0.001	.34	—[c]
Number of iron and folic acid supplements distributed	55	35,317 (23,407)	−359	.95	−0.009
Number of iron and folic acid supplements distributed per ANC visit	54	28 (2)	−10	.40	−0.303
Number of antiparasite medications distributed	56	289 (239)	−23	.72	−0.075
Expenses for ANC drugs	57	1423 (355)	9	.90	0.006
Expenses for delivery drugs	41	8238 (5802)	2331	.08	0.308
Outcomes based on 2021 records					
Number of ANC visits	56	1468 (1197)	160	.35	0.116
Number of deliveries	50	217 (248)	44	.34	0.211
Maternal mortality	36	0.003 (0)	−0.004	.44	−0.800
Newborn mortality	37	0.000 (0)	0.000	—	—
Number of iron and folic acid supplements distributed	52	36,411 (25,228)	−3278	.50	−0.087
Number of iron and folic acid supplements distributed per ANC visit	51	28 (2)	−7	.37	−0.219
Number of antiparasite medications distributed	51	315 (259)	−14	.83	−0.046
Expenses for ANC drugs	51	1406 (357)	−7	.93	−0.005
Expenses for delivery drugs	36	7660 (5296)	952	.43	0.130

[a]This table shows the regression results of the ITT estimations for secondary outcomes at the facility level. Centres de Santé de Base (CSBs; public sector primary care health facilities) and reference hospitals assigned to the intervention group nonrandomly are excluded. The table reports the number of observations used for each outcome, the mean value and SD of each outcome in the control group, the coefficient of being treated with the corresponding *P* value, and the relative effect size (calculated as the coefficient divided by the mean). Indicator variables for strata were included as controls.

[b]ANC: antenatal care.

[c]Calculation of *P* value or relative effect size was not possible.

Discussion

Principal Findings

In this analysis, we quantitatively assessed the impact of a mobile savings tool for maternal and newborn care in the Analamanga region in Madagascar. With this tool, users could save money and pay for services at participating health facilities, including primary care health facilities as well as reference hospitals. The intervention was evaluated using a cluster-randomized controlled trial.

In the ITT analysis, we did not detect any statistically significant impact of the MMHW intervention on maternal and neonatal health outcomes. The lack of statistically significant effects could be due to the study being underpowered to detect these effect sizes given that the realized knowledge about the intervention and the uptake among the survey participants was below what was assumed in the initial power calculation. It could also mean that the intervention did not achieve population-level effects.

Relative effect sizes were larger in the contamination-adjusted analysis than in the ITT analysis, and they were meaningful though also statistically insignificant, especially when treatment was used as instrument for having used the MMHW for paying. The larger effects in the contamination-adjusted analysis indicate that the intervention potentially had an impact on those who were receptive to it and used the tool. Coefficients for the 3 primary outcomes were all positive, suggesting a higher likelihood of delivery at a facility, a higher number of ANC visits, and higher health care expenditure among MMHW users. The latter could indicate that, due to financial barriers, these women previously did not purchase all the services they required, leading to an initial increase in expenditure after

XSL•FO
RenderX

financial barriers were eased [9]. However, when controlling for employment status and asset ownership, the direction of the association changed for several outcomes, including delivery in a facility and health care expenditure. This could indicate a potential confounding bias in the main analysis.

A striking finding of this study was the level of expenditure for maternal care, which is in theory free in Madagascar. On average, women in the control group spent MGA 208,021 (US $48.45 according to the OANDA currency converter rate for August 16, 2022—middle of data collection) on health care during pregnancy and delivery. This amount could pose a barrier to many women who might be deterred from delivering at a health facility [8]. Given that MMHW use was associated with higher expenditure, women might actually limit their expenditure compared to their needs. As women continue to face charges for essential medication, treatment, and services beyond what is provided for free, this calls for demand-side financing solutions for maternal care.

Related Evidence

Relative effect sizes for an increase of 11%, 32%, and 55% compared to the control group for the 3 primary outcomes (in-facility delivery, number of ANC visits, and total expenditure for health, respectively) when treatment was used as an instrument for having used the MMHW were in line with effect sizes previously reported for general demand-side and mobile health (mHealth) interventions for maternal care [28,29]. However, mHealth is broadly defined and covers a range of interventions beyond MM interventions. Evidence on MM interventions specifically is much more limited.

Several studies have assessed the association between use of MM and health outcomes. They suggest that users of MM services are more likely to use formal health care services as a response to health shocks than nonusers [30] and that their health care use may be better [31,32]. However, these studies look at MM services in general, which are mostly not focused on saving tools and are not targeted at health spending, as is the case for the MMHW. Moreover, their identification is often not causal.

mHealth technologies have become frequently used tools to address challenges related to maternal and newborn health in low- and middle-income countries [33]. Many of these technologies aimed at improving ANC access seem to have informational functions through messaging and notification alerts [34]. Few seem to have functions related to financial transactions and incentives [35] that could be directly compared to the MMHW intervention assessed in this study. A systematic review of mHealth interventions and their impact on maternal and newborn health identified 1 study evaluating an intervention with the aforementioned financial-related functions [28]. In a randomized trial in Kenya, financial incentives paid through participants' mobile phones increased women's probability of attending ANC, whereas they did not increase the probability of delivering at a facility [36]. This intervention was primarily a conditional cash transfer that was paid for each health visit, whereas the mobile phone was the mode of payment of the transfer. Other studies have also shown that cash transfers and vouchers can have a positive impact on the use of maternal care services [37]. In Western Kenya, a program offering a full voucher and a conditional cash transfer to expectant mothers led to an increase in institutional delivery rates by 33% [38]. Similarly, a conditional cash transfer program in Nigeria paid pregnant women to deliver at a health facility and resulted in a 41% increase in facility deliveries [39].

The MMHW was primarily a savings tool, with a conditional cash transfer in the form of vouchers and matching of deposits made by the user as an incentive to use the tool. It did not constitute a full voucher, meaning that copayment and cosaving by the user was required. This may have reduced its effectiveness compared to a full voucher [38]. Instead, this tool added an aspect to mHealth technology interventions that has not been studied previously, namely, saving. A study protocol of a randomized controlled trial in Kampala, Uganda, described the use of an MM health savings account in their intervention, but the evaluation focused on the effect of SMS text messages to encourage saving behavior and did not analyze the uptake and use of the health savings account itself [40].

The Innovative Partnership for Universal Sustainable Healthcare intervention in Kenya provided fully subsidized, digital health insurance through women's own SIM cards to low-income women of reproductive age [41]. While saving was not a component of the intervention, the goal was to improve access to health insurance using mobile technology and increase the use of maternal health care services. All women in the treatment group received the digital health insurance on their mobile phones. This intervention did not improve maternal health care use [41].

To the best of our knowledge, this was the largest trial of a digital financing solution for maternal health. This trial was implemented in 63 CSBs and 4 reference hospitals in a central region of Madagascar. The evaluation of the intervention rigorously followed a randomized controlled trial protocol for a public health intervention. By conducting a large population-based survey instead of focusing on women visiting health care facilities, the evaluation used the highest possible bar to measure the relevant outcomes.

Limitations

Despite these strengths, this study faced several limitations. The main concern is related to statistical power. While the original power calculation assumed uptake of 50% of women in the intervention area, actual awareness of the intervention and uptake were lower. Uptake among women who had heard of the intervention was 37.42% (485/1296) and, thus, closer to the initially anticipated uptake. We addressed the lower-than-assumed uptake by increasing the sample size of the survey during data collection. However, given an uptake of 14.51% (485/3343), we would have needed to increase the sample size 44-fold to reach a power of 0.8 compared to complete uptake (1/[compliance rate]^2). We conducted a post hoc power analysis based on the collected data and realized uptake rate following the study by Duflo et al [42] for power calculations with grouped errors and imperfect compliance. Given the 14.51% (485/3343) registration rate in intervention areas, 31 clusters in each group, and an average of 103 respondents per cluster, we can only reject effect sizes at the

XSL•FO
RenderX

population level larger than a 69–percentage point increase in delivery at a health facility, larger than an increase of 2.5 ANC visits, and larger than an increase of MGA 558,731 (US $132.80 according to the OANDA currency converter rate for August 16, 2022) in health expenditure. This study was underpowered to detect smaller yet still meaningful effects. Therefore, we cannot reject that the study did not have the observed effect. One possible approach to address low uptake and the resulting low power in future studies would be to sample eligible individuals at health care facilities, randomly assign registration among this sample, and survey only these individuals before and after the intervention. This approach does not allow for an analysis of uptake in a real-world scenario of program rollout. We particularly chose the population-based design to be as close as possible to the real-world effect of the intervention. However, we are aware that this is a composite evaluation of the intervention itself and its implementation, which could have led to the observed lower-than-anticipated uptake. Further research is needed on how the implementation could have impacted uptake, but this is out of scope for this analysis.

No baseline data collection was conducted, which would have strengthened the power of the analysis. The lack of baseline data also meant that it was not possible to assess changes over time and control for potential confounding factors. Our estimation approach solely compared outcomes between the intervention and control groups at endline and relied on the randomized treatment assignment for causal identification.

The larger effect sizes in the IV analysis suggest that the intervention might have been beneficial for MMHW users. Unfortunately, the IV analysis had even greater demands on power, and no statistical significance was achieved. Assuming that all potential users behaved similarly when registered for the tool, the local average treatment effect showed the effect under full uptake. However, this is a strong assumption that may not be valid in the study context. Women who did not register for the tool voluntarily might have behaved differently once registered than those who did. The decision to register with the MMHW may have been influenced by characteristics that are also associated with the measured outcomes or how individuals may respond to the intervention. The factors that are associated with the decision to register have been analyzed quantitatively and qualitatively elsewhere (Schäfer, L, et al., unpublished data, June 2025). Therefore, the results of the IV analysis should be interpreted with caution. Moreover, the IV analysis assumed that treatment assignment affected outcomes only through MMHW uptake. This assumption may be violated if sensitization campaigns influenced health-seeking behavior independently of registration with the MMHW. However, sensitization activities strongly focused on information about the MMHW and did not include general messages such as a call to deliver at a health facility. Therefore, it seems implausible that health-seeking behavior was influenced independently of registration.

Our analysis was an evaluation of the entire intervention package, which combined financial, patient-focused components (eg, savings incentives and free ultrasound checkups) with nonfinancial, health care provider-level components (eg, health care worker training). We are not able to clearly disentangle the effects of these different components. However, when treatment was used as an instrument for having registered with the MMHW, effect sizes were often larger than when treatment was used as an instrument for having heard of the MMHW. This indicates that the components that were dependent on registration (eg, access to free ultrasound checkups and savings incentives) were more relevant than those that were not dependent on registration (ie, health care worker training). Similarly, the larger effect sizes when treatment was used as an instrument for having used the MMHW to pay compared to having registered with the MMHW suggest that the ability to save and collect contributions using the tool may be more beneficial than the vouchers for ultrasound checkups.

Further limitations relate to intervention assignment at the individual level and sample selection into the population-based survey. The analysis treated women in the catchment area of the intervention CSBs as treated. However, these women did not necessarily attend the CSB of their *fokontany* (administrative unit) for various reasons. This CSB might not be the one closest to their residence, they might visit another CSB for other personal reasons irrespective of the CSB's ability to offer payment through the MMHW, or they might not visit any CSB during pregnancy or for delivery. We only observed visits to the CSB after the intervention, and the decision to visit any CSB or a particular CSB might be affected by the intervention. Therefore, it is not possible to capture where women would have attended in a scenario without the intervention, and assigning treatment status through residence in the catchment area was the only option to link women with CSBs given our data.

Furthermore, the analysis was based on survey data from women who were home at the time of the visit, which took place during daytime on weekdays. These women were different from eligible women who were not present at their home, as shown previously. The latter were found to be slightly older and wealthier and have a higher probability of having access to a mobile phone. Studies of MM adoption have shown that wealthier and better educated individuals and those who own a mobile phone are more likely to adopt MM [43]. Therefore, it is possible that uptake of the MMHW might have been higher among those eligible women who were not found at their home at the time of the visit and, therefore, did not participate in the survey. While this means that we potentially underestimated intervention uptake, it is unclear how the sample selection bias impacted the estimated effects of the intervention. On the one hand, wealthier and better educated individuals who own a mobile phone could be in a better position to use the MMHW, leading to an underestimation of the effect using the available survey sample. On the other hand, slightly poorer individuals might benefit more from the MMHW as their financial barriers are more severe, resulting in an overestimation of the population-level effect. A population-based survey always faces difficulties in surveying individuals who are not present at their homes. An alternative approach would be sampling at health care facilities. However, this would lead to a different selection bias as women who do not visit health care facilities for ANC or delivery would not be included. Our study had the explicit goal of measuring the ITT effect at the population level.

XSL•FO
RenderX

A last concern relates to the quality of facility-level data based on facility records, which seemed rather imprecise, particular regarding expenses for medication and deaths during delivery. Expenses were not tracked consistently in facility records, and respondents often reported estimates, which was indicated by respondents reporting the same values for 2020 and 2021 and reporting intervals instead of a single value. Deaths of mothers or children were often not reported at all. This limits the usability of the health care provider record data. Underreported maternal and neonatal deaths could underestimate true mortality rates, thereby masking differences between facilities in the intervention and control groups. Imprecise reporting of expenses may negatively affect the precision of effect estimates on ANC and delivery drug expenses. While we followed the published protocol [26] and analyzed the facility-level data, our interpretation puts a larger weight on the data from the population-based survey.

Conclusions

While this study did not identify a statistically significant impact, the estimated contamination-adjusted effects suggest that the MMHW has potential to improve access to maternal care for women who are receptive to such an MM-based savings tool. Estimated population-level effects were much smaller, and this study was underpowered to detect such effects due to lower-than-anticipated uptake of the intervention. It is now of great interest to investigate the reasons for low uptake and strategies to increase adoption of the MMHW or similar tools. The MMHW intervention consisted of several components that could be adjusted to increase the tool's suitability and attractiveness to potential users. Further research could explore which information channels are most effective in informing potential users about the tool, which information sources are locally trusted by potential users to encourage uptake, and whether different incentive schemes could increase saving among users. Adding reminders to save as a function of the MMHW could potentially increase its use and, eventually, its effectiveness. On the health care providers' side, a better understanding of their perception of the tool and their motivation for its use is required. While health care providers received monetary incentives to file claims via the MMHW platform, these incentives might have been insufficient to compensate for the perceived increased workload or adjustment to the new process. Similar interventions should include health care providers in the design of tools and use processes to ensure their acceptance of and comfort with the interventions. In conclusion, while the potential of the MMHW lies in its promise to lower barriers to accessing obstetric care, this potential is not yet fully realized.

Acknowledgments

This trial was financially supported by the German Federal Ministry of Education and Research (01DG21036GA). The intervention was funded by the Else Kröner-Fresenius Foundation, the Berlin Institute of Health, and Doctors for Madagascar. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The individual-level datasets analyzed during this study are available in the Göttingen Research Online repository [44]. The facility-level data analyzed during this study are not publicly available due to identification risks.

Authors' Contributions

SK and JVE developed the intervention and study design and defined the analysis plan. AR, HEA, and BV were responsible for organizing the data collection. LB, in collaboration with HEA, ZR, and BV, designed the data collection tools. LB conducted the data analysis. LB, SK, JVE, and SV interpreted the findings with contributions from TB. LB and BV wrote the first draft of the manuscript. SK, JVE, and SV critically revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

JVE and SK are shareholders of the nonprofit company mTOMADY and co–principal investigators on the Federal Ministry of Education and Research–funded 4MOTHERS trial. All other authors declare that they have no conflicts of interest.

Multimedia Appendix 1
Additional descriptive tables and results tables.
[PDF File (Adobe PDF File), 190 KB - publichealth_v11i1e70182_app1.pdf]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.x).
[PDF File (Adobe PDF File), 90 KB - publichealth_v11i1e70182_app2.pdf]

References

1. Leive A, Xu K. Coping with out-of-pocket health payments: empirical evidence from 15 African countries. Bull World Health Organ 2008 Nov;86(11):849-856 [FREE Full text] [doi: 10.2471/blt.07.049403] [Medline: 19030690]

XSL•FO
RenderX

2. Eze P, Osaheni Lawani L, Agu UJ, Acharya Y. Catastrophic health expenditure in sub-Saharan Africa: systematic review and meta-analysis. Bull World Health Organ 2022 May 05;100(05):337-351J. [doi: 10.2471/blt.21.287673]
3. Honda A, Randaoharison PG, Matsui M. Affordability of emergency obstetric and neonatal care at public hospitals in Madagascar. Reprod Health Matters 2011 May 17;19(37):10-20 [FREE Full text] [doi: 10.1016/S0968-8080(11)37559-3] [Medline: 21555082]
4. Banerjee AV, Duflo E. The economic lives of the poor. J Econ Perspect 2007 Jan;21(1):141-167. [doi: 10.1257/089533007780095556]
5. Kyei-Nimakoh M, Carolan-Olah M, McCann TV. Access barriers to obstetric care at health facilities in sub-Saharan Africa-a systematic review. Syst Rev 2017 Jun 06;6(1):110 [FREE Full text] [doi: 10.1186/s13643-017-0503-x] [Medline: 28587676]
6. Sharma G, Mathai M, Dickson KE, Weeks A, Hofmeyr GJ, Lavender T, et al. Quality care during labour and birth: a multi-country analysis of health system bottlenecks and potential solutions. BMC Pregnancy Childbirth 2015;15 Suppl 2(Suppl 2):S2 [FREE Full text] [doi: 10.1186/1471-2393-15-S2-S2] [Medline: 26390886]
7. McIntyre D, Thiede M, Dahlgren G, Whitehead M. What are the economic consequences for households of illness and of paying for health care in low- and middle-income country contexts? Soc Sci Med 2006 Feb;62(4):858-865. [doi: 10.1016/j.socscimed.2005.07.001] [Medline: 16099574]
8. Muller N, Emmrich PM, Rajemison EN, De Neve J, Bärnighausen T, Knauss S, et al. A mobile health wallet for pregnancy-related health care in Madagascar: mixed-methods study on opportunities and challenges. JMIR Mhealth Uhealth 2019 Mar 05;7(3):e11420 [FREE Full text] [doi: 10.2196/11420] [Medline: 30457972]
9. Garchitorena A, Miller AC, Cordier LF, Ramananjato R, Rabeza VR, Murray M, et al. In Madagascar, use of health care services increased when fees were removed: lessons for universal health coverage. Health Aff (Millwood) 2017 Aug 01;36(8):1443-1451. [doi: 10.1377/hlthaff.2016.1419] [Medline: 28784737]
10. Trends in maternal mortality 2000 to 2020: estimates by WHO, UNICEF, UNFPA, World Bank Group and UNDESA/population division. World Health Organization. Geneva: World Health Organization; 2023. URL: https://www.who.int/publications/i/item/9789240068759 [accessed 2023-11-14]
11. Miller AC, Ramananjato RH, Garchitorena A, Rabeza VR, Gikic D, Cripps A, et al. Baseline population health conditions ahead of a health system strengthening program in rural Madagascar. Glob Health Action 2017 Jun 16;10(1):1329961 [FREE Full text] [doi: 10.1080/16549716.2017.1329961] [Medline: 28621206]
12. Enquête Par Grappes à Indicateurs Multiples. United Nations International Children's Emergency Fund (UNICEF). URL: https://www.unicef.org/madagascar/en/node/1116 [accessed 2024-07-04]
13. Mobile cellular subscriptions (per 100 people) - Sub-Saharan Africa, Madagascar. The World Bank. URL: https://data.worldbank.org/indicator/IT.CEL.SETS.P2?locations=ZG-MG&name_desc=true [accessed 2023-11-14]
14. Demirguc-Kunt A, Klapper L, Singer D, Ansar S, Hess J. Global Findex database 2017: measuring financial inclusion and the fintech revolution. The World Bank. 2018. URL: https://openknowledge.worldbank.org/bitstreams/d6dbf1fa-bef9-53e0-9b58-dafae1e4eaa1/download [accessed 2025-05-29]
15. State of the industry report on mobile money 2023. GSM Association. URL: https://www.gsma.com/mobilefordevelopment/resources/state-of-the-industry-report-on-mobile-money-2023-2/ [accessed 2023-11-14]
16. Dupas P, Robinson J. Why don't the poor save more? Evidence from health savings experiments. Am Econ Rev 2013 Jun 01;103(4):1138-1171 [FREE Full text] [doi: 10.1257/aer.103.4.1138] [Medline: 29533047]
17. Stein PB. Access to finance: Sub-Saharan Africa (English). The World Bank. URL: http://documents.worldbank.org/curated/en/599611468202144127/Access-to-finance-Sub-Saharan-Africa [accessed 2023-10-17]
18. Changamka. Innovations in Healthcare. 2013. URL: https://www.innovationsinhealthcare.org/profile/changamka/ [accessed 2023-10-17]
19. Haas S, Heymann M, Riley P, Taddese A. Mobile money for health. United States Agency for International Development. URL: https://www.hfgproject.org/wp-content/uploads/2013/12/Mobile-Money-for-Health-Case-Study.pdf [accessed 2023-10-17]
20. Jack W, Suri T. Risk sharing and transactions costs: evidence from Kenya's mobile money revolution. Am Econ Rev 2014 Jan 01;104(1):183-223. [doi: 10.1257/aer.104.1.183]
21. Suri T, Jack W. The long-run poverty and gender impacts of mobile money. Science 2016 Dec 09;354(6317):1288-1292. [doi: 10.1126/science.aah5309] [Medline: 27940873]
22. Suri T, Jack W, Stoker TM. Documenting the birth of a financial economy. Proc Natl Acad Sci U S A 2012 Jun 26;109(26):10257-10262 [FREE Full text] [doi: 10.1073/pnas.1115843109] [Medline: 22689945]
23. Geng X, Ide V, Janssens W, Kramer B, van der List M. Health insurance, a friend in need? Impacts of formal insurance and crowding out of informal insurance. World Dev 2018 Nov;111:196-210 [FREE Full text] [doi: 10.1016/j.worlddev.2018.07.004]
24. Ky S, Rugemintwari C, Sauviat A. Does mobile money affect saving behaviour? Evidence from a developing country. J Afr Econ 2017;27(3):285-320. [doi: 10.1093/jafeco/ejx028]
25. mTOMADY - powering access to healthcare and making UHC 2030 possible. mTOMADY. URL: https://www.mtomady.com [accessed 2023-10-17]

XSL•FO
RenderX

26. Lacroze E, Bärnighausen T, De Neve JW, Vollmer S, Ratsimbazafy RM, Emmrich PM, et al. The 4MOTHERS trial of the impact of a mobile money-based intervention on maternal and neonatal health outcomes in Madagascar: study protocol of a cluster-randomized hybrid effectiveness-implementation trial. Trials 2021 Oct 21;22(1):725 [FREE Full text] [doi: 10.1186/s13063-021-05694-8] [Medline: 34674741]
27. Kruk ME, Leslie HH, Verguet S, Mbaruku GM, Adanu RM, Langer A. Quality of basic maternal care functions in health facilities of five African countries: an analysis of national health system surveys. Lancet Glob Health 2016 Nov;4(11):e845-e855. [doi: 10.1016/S2214-109x(16)30180-2]
28. Knop MR, Nagashima-Hayashi M, Lin R, Saing CH, Ung M, Oy S, et al. Impact of mHealth interventions on maternal, newborn, and child health from conception to 24 months postpartum in low- and middle-income countries: a systematic review. BMC Med 2024 May 15;22(1):196 [FREE Full text] [doi: 10.1186/s12916-024-03417-9] [Medline: 38750486]
29. Hurst TE, Semrau K, Patna M, Gawande A, Hirschhorn LR. Demand-side interventions for maternal care: evidence of more use, not better outcomes. BMC Pregnancy Childbirth 2015 Nov 13;15(1):297 [FREE Full text] [doi: 10.1186/s12884-015-0727-5] [Medline: 26566812]
30. Ahmed H, Cowan B. Mobile money and healthcare use: evidence from East Africa. World Develop 2021 May;141:105392. [doi: 10.1016/j.worlddev.2021.105392]
31. Egami H, Matsumoto T. Mobile money use and healthcare utilization: evidence from rural Uganda. Sustainability 2020 May 05;12(9):3741. [doi: 10.3390/SU12093741]
32. Bukari C, Koomson I. Adoption of mobile money for healthcare utilization and spending in rural Ghana. In: Churchill SA, editor. Moving from the Millennium to the Sustainable Development: Goals Lessons and Recommendations. Cham, Switzerland: Springer; 2020:37-60.
33. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]
34. Mishra M, Parida D, Murmu J, Singh D, Rehman T, Kshatri JS, et al. Effectiveness of mHealth interventions for monitoring antenatal care among pregnant women in low- and middle-income countries: a systematic review and meta-analysis. Healthcare (Basel) 2023 Sep 27;11(19):2635 [FREE Full text] [doi: 10.3390/healthcare11192635] [Medline: 37830672]
35. Mohamed H, Ismail A, Sutan R, Rahman RA, Juval K. A scoping review of digital technologies in antenatal care: recent progress and applications of digital technologies. BMC Pregnancy Childbirth 2025 Feb 13;25(1):153 [FREE Full text] [doi: 10.1186/s12884-025-07209-8] [Medline: 39948493]
36. Vanhuyse F, Stirrup O, Odhiambo A, Palmer T, Dickin S, Skordis J, et al. Effectiveness of conditional cash transfers (Afya credits incentive) to retain women in the continuum of care during pregnancy, birth and the postnatal period in Kenya: a cluster-randomised trial. BMJ Open 2022 Jan 06;12(1):e055921 [FREE Full text] [doi: 10.1136/bmjopen-2021-055921] [Medline: 34992119]
37. Hunter BM, Harrison S, Portela A, Bick D. The effects of cash transfers and vouchers on the use and quality of maternity care services: a systematic review. PLoS One 2017 Mar 22;12(3):e0173068 [FREE Full text] [doi: 10.1371/journal.pone.0173068] [Medline: 28328940]
38. Grépin KA, Habyarimana J, Jack W. Cash on delivery: results of a randomized experiment to promote maternal health care in Kenya. J Health Econ 2019 May;65:15-30 [FREE Full text] [doi: 10.1016/j.jhealeco.2018.12.001] [Medline: 30878794]
39. Okeke EN, Wagner Z, Abubakar IS. Maternal cash transfers led to increases in facility deliveries and improved quality of delivery care in Nigeria. Health Aff (Millwood) 2020 Jun 01;39(6):1051-1059 [FREE Full text] [doi: 10.1377/hlthaff.2019.00893] [Medline: 32479220]
40. Hartwig L, Namukose V, Kiriya J, Onyancha C, Shibanuma A, Jimba M. The effectiveness of a behavioral science and design intervention for family savings on use of maternal health services and male involvement: study protocol for a randomized controlled trial. BMC Public Health 2022 Aug 19;22(1):1579. [doi: 10.1186/S12889-022-13985-1]
41. Abajobir A, de Groot R, Wainaina C, Pradhan M, Janssens W, Sidze EM. The impact of digital interventions on health insurance coverage for reproductive, maternal, newborn and child health services utilization in Kakamega, Kenya: a cluster randomized controlled trial. Health Policy Plan 2024 Nov 14;39(10):1007-1021. [doi: 10.1093/heapol/czae079] [Medline: 39185605]
42. Duflo E, Glennerster R, Kremer M. Using randomization in development economics research: a toolkit. In: Schultz TP, Strauss JA, editors. Handbook of Development Economics. New York, NY: Academic Press; 2007:3895-3962.
43. Aron J. Mobile money and the economy: a review of the evidence. World Bank Res Obs 2018;33(2):135-188. [doi: 10.1093/wbro/lky001]
44. Bogler L, Vieira B, Andriamasy HE, Rampanjato Z, Vollmer S, Bärnighausen T, et al. Replication data for: the impact of a mobile money-based intervention on maternal and neonatal health outcomes in Madagascar: a cluster-randomized controlled trial version 1.0. Göttingen Research Online. URL: https://data.goettingen-research-online.de/dataset.xhtml?persistentId=doi:10.25625/KWT5KG [accessed 2025-05-29]

Abbreviations

ANC: antenatal care

XSL•FO
RenderX

CSB: Centres de Santé de Base
ITT: intention-to-treat
IV: instrumental variable
mHealth: mobile health
MM: mobile money
MMHW: Mobile Maternal Health Wallet

Edited by H Taha; submitted 17.12.24; peer-reviewed by HC Moungui, S Zare; comments to author 01.04.25; revised version received 19.05.25; accepted 17.06.25; published 15.08.25.

Please cite as:
Bogler L, Vieira B, Andriamasy HE, Rampanjato Z, Vollmer S, Bärnighausen T, Ralisimalala A, Emmrich JV, Knauss S
The Impact of a Mobile Money–Based Intervention on Maternal and Neonatal Health Outcomes in Madagascar: Cluster-Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e70182
URL: https://publichealth.jmir.org/2025/1/e70182
doi:10.2196/70182
PMID:

©Lisa Bogler, Bítia Vieira, Harizaka Emmanuel Andriamasy, Zavaniarivo Rampanjato, Sebastian Vollmer, Till Bärnighausen, Andriamampianina Ralisimalala, Julius Valentin Emmrich, Samuel Knauss. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 15.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Microorganisms Linked to Health Care–Associated Infections: Modernization of Terminology Resources for Reporting to the National Healthcare Safety Network

Sheila Abner[1], PhD; Shuai Zheng[1], PhD; Caroline Macumber[2], MS; Shaun Shakib[2], PhD; Jonathan Edwards[1], MStat; Andrea Benin[1], MD

[1]Division of Healthcare Quality Promotion, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA, United States

[2]Clinical Architecture, Carmel, IN, United States

Corresponding Author:
Sheila Abner, PhD
Division of Healthcare Quality Promotion, National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, GA, United States

Abstract

The National Healthcare Safety Network (NHSN) of the Centers for Disease Control and Prevention (CDC) needed a modernized approach to manage resources containing standardized terminology that specify microorganism data submitted electronically for legacy reporting. Health care–associated infections (HAIs) reported to NHSN require the submission of data regarding specific microorganisms attributed to the patient's condition. Data on microorganisms submitted to the NHSN electronically must use the SNOMED CT terminology standard. Terminology artifacts that guide submission of microorganism data have been maintained in spreadsheets that have become increasingly challenging to manage. This case report details the initial use case for the implementation of off-the-shelf software within the NHSN to modernize the maintenance of terminology assets. Resources that guide reporting microorganisms for HAIs were used as a prototype to demonstrate how a software application can be practically implemented to streamline the maintenance of complex terminology assets. Mission-critical artifacts have been reconciled and consolidated into a single source of truth knowledgebase using an off-the-shelf software solution. This report shares progress and lessons learned regarding the modernization of NHSN's *Pathogen Codes* resource and its derivative artifacts. A model is now available that can be replicated across other NHSN legacy artifacts. Our experience can be applied to other public health use cases and information systems facing similar challenges with attachments to legacy terminology resources and systems.

(JMIR Public Health Surveill 2025;11:e66088) doi:10.2196/66088

KEYWORDS

SNOMED CT; terminology; microorganisms; health care–associated infections; public health surveillance

Introduction

The National Healthcare Safety Network (NHSN) of the Centers for Disease Control and Prevention (CDC) is the nation's surveillance program for health care–associated infections (HAIs) and other conditions, patient safety events, and antimicrobial use and resistance [1]. The NHSN collects and processes surveillance data for the purpose of promoting efforts to prevent adverse events in health care by enabling facilities and governance agencies to determine problematic areas and formulate corresponding corrective actions.

The NHSN was launched in 2005 as a web-based application, succeeding the smaller DOS-based National Nosocomial Infection Surveillance (NNIS) system that had been in operation since the 1980s [2]. A bellwether for the adoption of HL7 data standards and electronic public health reporting, the NHSN implemented its first clinical document architecture (CDA) data submissions in 2008. As of November 2024, there were more than 165,000 users and over 38,500 health care facilities across the health care continuum reporting to the NHSN. More than 59 CDA vendors have established the capacity to submit data electronically on behalf of facilities.

When the NHSN first enabled electronic reporting, specifications for value sets from standardized vocabularies were needed to guide CDA vendors. Spreadsheets became the de facto solution for managing those resources because terminology software was not available. Additional attributes of metadata and relationships to other artifacts were incorporated into the spreadsheets to meet business needs. Maintaining these assets became labor-intensive with the growth of the NHSN expanding beyond its original functions.

As part of the CDC's Data Modernization Initiative, the NHSN has adopted a new software platform for the maintenance of terminology assets. This paper highlights the NHSN's early

XSL•FO
RenderX

successes in modernizing resources that guide standardized data collection of microorganisms attributed to HAIs while preserving backward compatibility with legacy business processes.

Evaluation of Legacy State and Transition Approach

Terminology resource management processes and products were qualitatively evaluated before and after the adoption of an off-the-shelf solution. An assessment of the historical state was performed by terminologists and informaticists skilled in the laboratory data domain. Stakeholders—including data scientists, database administrators, system developers, in addition to epidemiologists and infection preventionists authoring surveillance protocols—were questioned in focus groups to determine pitfalls and pain points in legacy processes. Feedback and insight from those interviews guided the overarching business and technical requirements for the transition activities.

The deployment of the new platform has been approached with three major features. First, the required servers have been constructed in the cloud environment of the CDC, followed by integration and performance testing with other infrastructural systems and existing NHSN applications. Various data provisioning methods have also been explored and merged into the roadmap of the NHSN.

Second, a corresponding standard operating procedure (SOP) has been developed to guide usage of the platform and facilitate communication among collaborators. A Terminology Committee, including physician informaticists, has been assembled to centralize discussions of terminology-related topics and requests, providing convenient consultation services to all stakeholders. After requests are conceptualized, they are formally itemized into specific requirements and documented in standardized terminology asset management (TAM) request forms. The SOP improves the efficiency of communication and streamlines coauthoring and maintenance of terminology resources into the new platform.

Third, necessary training sessions have been scheduled for the software users and stakeholders. In addition to the basic operations of the software platform, fundamental informatics and terminology principles have been taught in the form of training videos, webinars, and workshops to educate the workforce of NHSN. Subgroups have also been defined to offer additional domain-specific trainings based on different roles in the workflow.

Historical State and Drivers for Implementing Improvements

The NHSN has developed an extensive collection of terminology assets, including 2 code systems and more than 80 value sets from various reference terminologies for electronic data specifications conformant to the HL7 CDA standard [3]. In addition to value sets in the value set authority center (VSAC), a publicly available authoring and provisioning tool hosted by the National Library of Medicine [4], CDA vendors are also dependent on the NHSN information data model (IDM), which combines business rules with standardized terminology content. The IDM has been maintained in spreadsheets that evolved into a multi-faceted resource and was well-positioned for modernization. The first proof of concept for such modernization focused on a high-priority table in the IDM, *Pathogen Codes* [5], that catalogs criteria for reporting microorganisms linked to all HAI events in the existing NHSN surveillance portfolio.

The NHSN retained approximately 1000 microorganism terms from its NNIS predecessor and established a methodology to create unique and human readable codes for analysis purposes. These short representations for pathogen concepts were challenging to curate after being coupled to SNOMED CT [6] concept identifiers required for the electronic reporting option. Although the naming conventions for NHSN codes became less intuitive as the number of terms increased to keep up with growth of NHSN over the years, these conventions remained mandatory for analyzing data.

By 2019, the *Pathogen Codes* resource had amassed more than 3500 rows of terms that included subsets of synonyms. Synonyms were particularly challenging to manage due to the ever-changing landscape of microbial nomenclature [7]. More than 35 columns of metadata were needed to characterize the reporting criteria across a growing number of complex HAI events. The sheer number of mixed-term types with manually curated mnemonic codes and business logic attributes managed in spreadsheets led to the need for comprehensive terminology software for the most effective maintenance.

Although the SNOMED CT U.S. Edition releases occur biannually (March and September), updates to the *Pathogen Codes* resource have coincided with annual NHSN application major releases in January. Because the NHSN development life cycles also need to accommodate the development cycles of CDA vendors, the offset timing of NHSN releases and SNOMED CT updates meant that it was challenging to maintain up-to-date SNOMED CT content in distributed NHSN resources.

Updating the *Pathogen Codes* resource was so labor-intensive that the task needed to be started on a version of SNOMED CT that preceded the NHSN application release by more than a year. The terminologist prepared enumerated lists of SNOMED CT concepts in VSAC, and exported files were manually compared to existing *Pathogen Codes* content to identify inactivated codes, name changes to term descriptions, and additions being considered. The resulting marked-up spreadsheet that highlighted proposed updates was shared with protocols and data analysis stakeholders to reach consensus on changes to microorganism membership and associated attributes. Following that reconciliation, the annotated spreadsheet was submitted to business analysts with enough time for the development team to implement the changes for the January NHSN application releases. Overall, the maintenance process relied heavily on manually processed spreadsheets and email communication which intensified the workload and prolonged the path to completion. The fragmented process caused challenges to version control and rigorous traceability.

In addition, supplemental spreadsheets derived from the master *Pathogen Codes* table were required to support the divergent

XSL•FO
RenderX

needs of the two different approaches available for data submission. Subsets of content needed to be parsed and simplified in a companion spreadsheet for users entering one event at a time into the web application. In contrast, almost 1000 additional terms needed to be curated separately to expand a subset of content (eg, more than 700 *E. coli* serotypes in SNOMED CT) to support users submitting data via CDA files.

Evolution of Technology Advancement

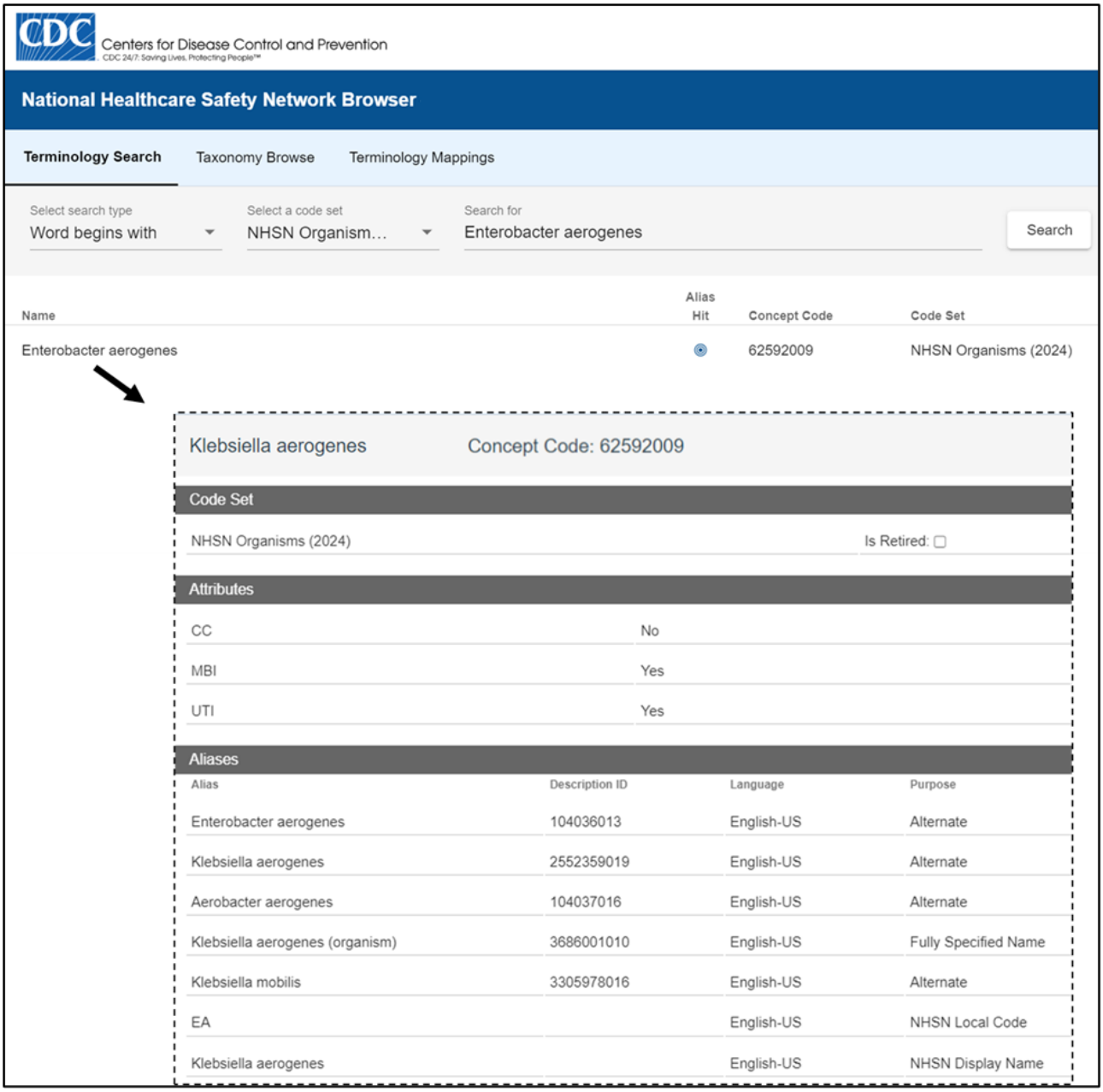
Migrating the *Pathogen Codes* resource from a spreadsheet to an off-the-shelf software product has modernized how NHSN is maintaining and provisioning this critical terminology asset. The SNOMED CT code system now serves as the core structure for *Pathogen Codes*, and the standard content is enriched with NHSN-specific metadata. Key principles relevant to this effort are derived from the *Desiderata for Controlled Medical Vocabularies* that have influenced the design of several modern code systems [8].

Identifying changes to SNOMED CT concepts and all associated NHSN-specific attributes in the *Pathogen Codes* resource is automated using the terminology software. Microorganism concepts that have moved to inactive status and updates to term descriptions due to changes in nomenclature are automatically flagged for review by subject matter experts from the protocols and data analysis teams with each new release of the code system. Maintenance is substantially more streamlined since the NHSN-specific metadata is now bound to SNOMED CT concept identifiers. Display names of microorganisms in dropdown lists in the web application are identical to the SNOMED CT Preferred Terms. All synonyms (ie, former and alternate names) for the microorganism concepts in the *Pathogen Codes* resource are automatically updated and are now searchable using a customized NHSN terminology browser (Figure 1). This simplified browser has eliminated the need to maintain a separate supplemental spreadsheet for users entering one event at a time into the web application. In addition to identifying synonyms, selected attributes that define reporting criteria for bloodstream infections and urinary tract infections (eg, the Common Commensals column in the *Pathogen Codes* asset) are displayed in the terminology browser for easy reference [9].

Since the NHSN codes are integral to analyzing data, change must be strategic to minimize disruptions among stakeholders. For maintenance purposes, existing NHSN mnemonic codes will persist as short alternate names (eg, EA for *Klebsiella aerogenes* that was formerly in the genus *Enterobacter*). Current analysis of electronically submitted data depends on translating standard codes from CDA files to NHSN codes. However, future-facing data aggregation and analysis will directly leverage the SNOMED CT ontology and non-semantic concept identifiers without the need for NHSN codes.

XSL•FO
RenderX

Figure 1. The National Healthcare Safety Network (NHSN) terminology browser for the Pathogen Codes SNOMED CT resource. CC: common commensals; MBI: mucosal barrier injury; UTI: urinary tract infection.



Overcoming Challenges From Consequences of Program Expansion

Due to the considerable growth and success of the NHSN, legacy approaches to maintaining terminology assets were no longer sustainable. These challenges placed a demand on adopting a more automated solution to manage all terminology-related resources.

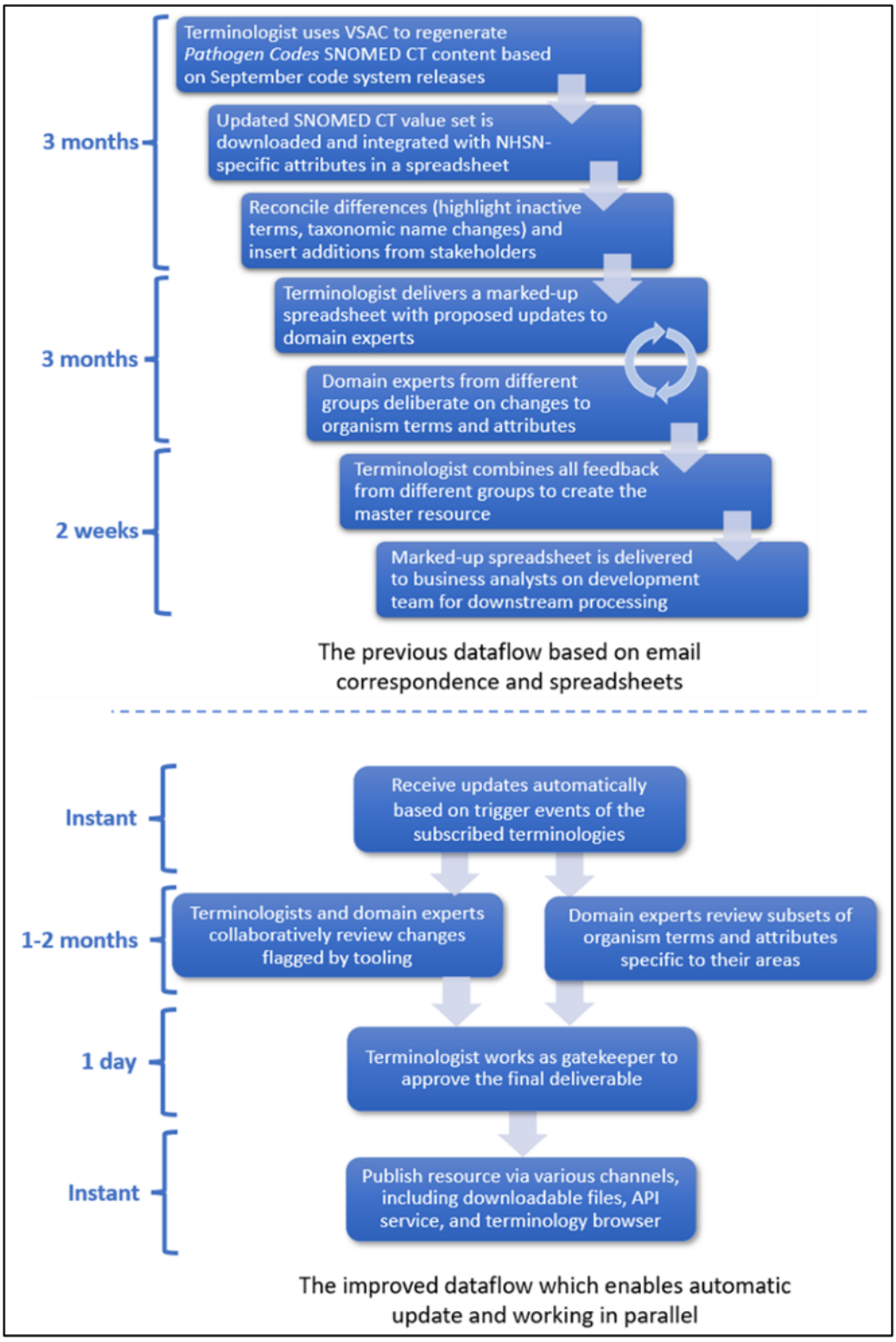
As a result of transitioning to the new terminology platform, reports can be configured to streamline export of customized assets derived from the master *Pathogen Codes* resource. This change will obviate the need to maintain separate companion spreadsheets for different use cases. In addition, the inherent version control and audit trail functionality will enable automated traceability.

Figure 2 illustrates the streamlined maintenance processes. Compared to the previous workflow, a major improvement is that reconciliation with new releases of SNOMED CT is essentially automated. After subscribing to a reference terminology in the platform, changes to NHSN vocabulary assets are applied seamlessly when triggered by new releases of the code system. Alleviating that burden of human intervention is amplified by features of the tooling that allow multiple domain specialists to work simultaneously on subsets of content related to their areas of expertise (eg, bloodstream infection or ventilator-associated events). Another feature that will standardize and streamline the maintenance process is the rich functionality for authoring rule-based value sets. This framework bridges the knowledge gap between domain experts and terminologists to apply rigorous inclusion and exclusion criteria for authoring and automatically refreshing value sets.

XSL•FO
RenderX

Overall, the maintenance timeframe has been reduced from more than 6 months to less than 2 months.

Figure 2. Streamlined data flow with the new approach.



In addition to an application that can automate and streamline the terminology maintenance procedures, the corresponding workflow design and education of the workforce are equally important. Enabling multiple collaborators to cohesively work in parallel on different aspects of a terminology asset not only fast tracks the overall processing time, but also reduces the risk of miscommunication. Meanwhile, training materials on essential terminology principles and methodologies should be offered to the workforce involved in defining requirements for terminology resources to transform routine processes. For example, a basic understanding of the approaches to building rule-based value sets is a fundamental step towards automating updates.

XSL•FO
RenderX

Preliminary progress with the NHSN terminology modernization efforts has been productive and promising. Our experience and lessons learned can be applied to similar use cases relying heavily on terminology resources that are subject to frequent updates. Further integration with other NHSN and CDC applications is a major future direction. A fully automated data pipe will be implemented based on the terminology services offered by the new platform, allowing systematic data provisioning to internal and external stakeholders. Meanwhile, all existing terminology resources will be migrated from the IDM to the new platform for improved maintenance and data sharing procedures. The NHSN is moving in the direction of maximizing electronic submission of data [10,11], but support for manual data entry in the web application and the current IDM will persist indefinitely to accommodate facilities with varying technical capacities and to support legacy surveillance reporting.

Summary

In conclusion, the preparations and directions outlined have drastically eased the adoption and deployment of the new technical solution. Migration of the *Pathogen Codes* asset to the new terminology platform has been a proving ground for updating additional resources in the IDM containing standardized terminology content coupled to the NHSN codes. Being able to tether business rules for HAI reporting criteria to the SNOMED CT concepts of interest and repurposing NHSN codes as short, alternate-term descriptions are pivotal gains from implementing the new approach. Directly using the framework of terminology standards extended with NHSN-specific attributes will streamline maintenance tasks and keep the content more current. Although this use case is restricted to the SNOMED CT organism hierarchy for microorganisms identified in laboratory test results, benefits of this work will scale to other clinical, medications, and laboratory data elements critical for the next generation of NHSN reporting.

Disclaimer

The conclusions, findings, and opinions expressed by authors contributing to this journal do not necessarily reflect the official position of the U.S. Department of Health and Human Services, the Public Health Service, the Centers for Disease Control and Prevention, or the authors' affiliated institutions.

Use of trade names is for identification only and does not imply endorsement by the Public Health Service or by the U.S. Department of Health and Human Services.

Conflicts of Interest

None declared.

References

1. National Healthcare Safety Network (NHSN). URL: https://www.cdc.gov/nhsn/index.html [accessed 2024-11-01]
2. Gaynes RP, Culver DH, Emori TG, et al. The national nosocomial infections surveillance system: plans for the 1990s and beyond. Am J Med 1991 Sep 16;91(3B):116S-120S. [doi: 10.1016/0002-9343(91)90355-2] [Medline: 1656746]
3. HL7 Clinical Document Architecture (CDA®) Standards. URL: https://www.hl7.org/implement/standards/product_section.cfm?section=10 [accessed 2024-11-01]
4. Value Set Authority Center (VSAC). URL: https://vsac.nlm.nih.gov/ [accessed 2024-11-01]
5. NHSN Implementation Toolkits & Resources. URL: https://www.cdc.gov/nhsn/cdaportal/toolkits.html [accessed 2024-11-01]
6. Systematized Nomenclature of Medicine-Clinical Terms (SNOMED CT®). URL: https://www.nlm.nih.gov/research/umls/rxnorm/sourcereleasedocs/snomedct_us.html [accessed 2024-11-01]
7. Prinzi AM, Moore NM. Change of plans: overview of bacterial taxonomy, recent changes of medical importance, and potential areas of impact. Open Forum Infect Dis 2023 Jul;10(7):ofad269. [doi: 10.1093/ofid/ofad269] [Medline: 37408854]
8. Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med 1998 Nov;37(4-5):394-403. [Medline: 9865037]
9. NHSN Terminology. URL: https://www.cdc.gov/nhsn/cdaportal/terminology/index.html [accessed 2024-11-01]
10. Classen DC, Rhee C, Dantes RB, Benin AL. Healthcare-associated infections and conditions in the era of digital measurement. Infect Control Hosp Epidemiol 2024 Jan;45(1):3-8. [doi: 10.1017/ice.2023.139] [Medline: 37747086]
11. Shehab N, Alschuler L, McILvenna S, et al. The national healthcare safety network's digital quality measures: CDC's automated measures for surveillance of patient safety. J Am Med Inform Assoc 2024 Apr 19;31(5):1199-1205. [doi: 10.1093/jamia/ocae064] [Medline: 38563821]

Abbreviations

CDA: clinical document architecture
CDC: Centers for Disease Control and Prevention
HAI: healthcare-associated infections

XSL•FO
RenderX

IDM: information data model
NHSN: National Healthcare Safety Network
NNIS: National Nosocomial Infection Surveillance
SNOMED CT: Systematized Nomenclature of Medicine – Clinical Terms
SOP: standard operating procedure
TAM: terminology asset management
VSAC: value set authority center

Edited by A Mavragani; submitted 03.09.24; peer-reviewed by B Braun, L Rodriguez; revised version received 03.12.24; accepted 04.12.24; published 19.08.25.

Please cite as:
Abner S, Zheng S, Macumber C, Shakib S, Edwards J, Benin A
Microorganisms Linked to Health Care–Associated Infections: Modernization of Terminology Resources for Reporting to the National Healthcare Safety Network
JMIR Public Health Surveill 2025;11:e66088
URL: https://publichealth.jmir.org/2025/1/e66088
doi:10.2196/66088

© Sheila Abner, Shuai Zheng, Caroline Macumber, Shaun Shakib, Jonathan Edwards, Andrea Benin. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Machine Learning Applications in Population and Public Health: Guidelines for Development, Testing, and Implementation

Andrew D Pinto[1,2,3,4], MSc, MD, CCFP; Sharon Birdi[1], MBT; Steve Durant[1], PhD; Roxana Rabet[1], MSc; Rahul Parekh[1,5], MD; Shehzad Ali[6,7,8], PhD; David Buckeridge[9], MSc, MD, PhD; Marzyeh Ghassemi[10], MSc, PhD; Jennifer Gibson[11], MA, PhD; Ava John-Baptiste[12], PhD; Jillian Macklin[1,13], MD, PhD; Melissa D McCradden[4,14,15], MHSc, PhD; Kwame McKenzie[16,17], MBBS; Parisa Naraei[18], PhD; Akwasi Owusu-Bempah[19], MA, PhD; Laura C Rosella[4,20,21,22], MHSc, PhD; James Shaw[23], MPT, PhD; Ross Upshur[3,4,11], MA, MSc, MD, CCFP; Sharmistha Mishra[24,25,26,27,28], MSc, MD, PhD

[1]Upstream Lab, MAP Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, Unity Health Toronto, 30 Bond Street, Toronto, ON, Canada

[10]Department of Electrical Engineering and Computer Science (EECS) and Institute for Medical Engineering and Science (IMES), Massachusetts Institute of Technology, Cambridge, MA, United States

[11]Joint Centre for Bioethics, University of Toronto, Toronto, ON, Canada

[12]Departments of Epidemiology and Biostatistics, Anesthesia and Perioperative Medicine, and Schulich Interfaculty Program in Public Health, Western University, London, ON, Canada

[13]Undergraduate Medical Education, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[14]Department of Bioethics, Hospital for Sick Children, Toronto, ON, Canada

[15]Australian Institute for Machine Learning, University of Adelaide, Adelaide, Australia

[16]Wellesley Institute, Toronto, ON, Canada

[17]Centre for Addiction and Mental Health, Toronto, ON, Canada

[18]Department of Computer Science, Toronto Metropolitan University, Toronto, ON, Canada

[19]Department of Sociology, Faculty of Arts and Sciences, University of Toronto, Toronto, ON, Canada

[2]Department of Family and Community Medicine, St. Michael's Hospital, Toronto, ON, Canada

[20]Institute for Better Health, Trillium Health Partners, Toronto, ON, Canada

[21]Department of Laboratory Medicine and Pathobiology, Temerty Faculty of Medicine, Toronto, ON, Canada

[22]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[23]Department of Physical Therapy, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[24]Division of Infectious Diseases, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[25]MAP Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, Unity Health Toronto, Toronto, ON, Canada

[26]Management and Evaluation, and Division of Epidemiology, Institute of Health Policy, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[27]Institute of Medical Science, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[28]Institute for Clinical Evaluative Sciences, Toronto, ON, Canada

[3]Department of Family and Community Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[4]Division of Clinical Public Health, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[5]Department of Family Medicine, College of Medicine, University of Saskatchewan, Saskatoon, SK, Canada

[6]Department of Epidemiology and Biostatistics, Western Centre for Public Health & Family Medicine, Western University, London, ON, Canada

[7]WHO Collaborating Centre for Knowledge Translation and Health Technology Assessment in Health Equity, Ottawa, ON, Canada

[8]Department of Health Sciences, University of York, York, United Kingdom

[9]Department of Epidemiology, Biostatistics and Occupational Health, School of Population and Global Health, McGill University, Montreal, QC, Canada

Corresponding Author:
Andrew D Pinto, MSc, MD, CCFP
Upstream Lab, MAP Centre for Urban Health Solutions, Li Ka Shing Knowledge Institute, Unity Health Toronto, 30 Bond Street, Toronto, ON, Canada

Abstract

Machine learning (ML), a subset of artificial intelligence, uses large datasets to identify patterns between potential predictors and outcomes. ML involves iterative learning from data and is increasingly used in population and public health. Examples include

XSL•FO
RenderX

early warning of infectious disease outbreaks, predicting the future burden of noncommunicable diseases, and assessing public health interventions. However, ML can inadvertently produce biased outputs related to the quality and quantity of data, who is engaged and helping direct the analysis, and how findings are interpreted. Specific guidelines for using ML in population and public health have not yet been created. We assembled a diverse team of experts in computer science, statistical modeling, clinical and population health epidemiology, health economics, ethics, sociology, and public health. Drawing on literature reviews and a modified Delphi process, we identified five key recommendations: (1) prioritize partnerships and interventions to support communities considered structurally disadvantaged; (2) use ML for dynamic situations, such as public health emergencies, while adhering to ethical standards; (3) conduct risk assessments and bias mitigation strategies aligned with identified risks; (4) ensure technical transparency and reproducibility by publicly sharing data sources and methodologies; and (5) foster multidisciplinary dialogue to discuss the potential harms of ML-related bias and raise awareness among the public and public health community. The proposed guidelines provide operational steps for stakeholders, ensuring that ML tools are not only effective but also ethically grounded and feasible in real-world scenarios.

(JMIR Public Health Surveill 2025;11:e68952) doi:10.2196/68952

KEYWORDS

population health; public health; machine learning; artificial intelligence; guideline; algorithmic bias; AI; health equity

Introduction

Machine learning (ML) is a form of artificial intelligence (AI) that is now used for a range of problems across many fields. ML involves a machine "learning" as it processes more data, improving predictive performance over time [1]. ML, as a set of tools, can be used for prediction, clustering, and causal inference. While prediction models are commonly used in public health to examine associations between potential predictors and outcomes, ML can also be used to identify groups with shared characteristics (clustering) and to identify potential causal associations between interventions and health outcomes [2,3]. However, these methods are often limited by the quality and quantity of the available data in public health [4].

As methods advance and as the availability of data increases, ML-based innovations are playing a central role in population and public health [5]. Key areas include the surveillance of infectious diseases [6]; predicting the burden of noncommunicable diseases (NCDs) [7]; and assessing public health interventions, including those focused on modifiable risk factors [8]. Awareness that ML could help improve population health is tempered by a growing understanding that it has risks and potentially negative consequences [9]. An important concern for the application of ML models is the perpetuation and amplification of biases reflecting patterns of societal inequality, when the accuracy of model predictions differs systematically across subpopulations, potentially leading to decisions that exacerbate health inequities [10]. This issue is rooted in differences in the amount and quality of data from different populations [11].

Bias in ML as it is applied to population and public health problems has been identified previously [12]; however, guidelines for ML applied to population and public health are limited. Without such guidelines, ML could exacerbate inequities and fail to adhere to methodological standards. For example, during the COVID-19 pandemic, ML models were used to predict infection rates and allocate health care resources. These models were found to be biased, particularly when data from underrepresented communities were either scarce or of lower quality, impacting the accuracy of predictions [13]. Public health organizations may lack the resources and expertise to thoroughly evaluate the ML models they use. While ethical frameworks for AI in health have emerged, they primarily focus on clinical settings. Public health involves unique applications, including population-level interventions, real-time surveillance, and communicable disease control, which require field-specific guidance. Our guidelines aim to address this gap by focusing on the distinct operational, ethical, and equity considerations involved in applying ML to public health. We aim to provide recommendations to those creating or revising ML tools for population and public health purposes, offer guidance to users of ML tools, and outline approaches on how to address bias. We centered our recommendations on the following questions: (1) What are best practices to identify and mitigate biases for those developing, testing, and implementing ML models in population health? (2) What are the priority areas for further research?

We used the GRADE (Grading of Recommendations Assessment, Development and Evaluation) [14], the National Institute for Health and Care Excellence (NICE) [15] approaches to guideline development, and other guidelines specific to epidemiology and health economics modeling [16,17]. We also drew insights from participatory frameworks for policy development governing ML [18] and from documentation on the governance of community data trusts, which provide data for many of the ML models we studied [19]. In addition, we considered international guidelines on AI ethics, such as the Montreal Declaration for Responsible Development of AI [20] and recommendations from the European Commission High-Level Expert Group on AI [21], along with the World Health Organization guidance on *Ethics & Governance of Artificial Intelligence for Health* [22].

Guideline Panel Composition and Management of Competing Interests

To identify experts for the Delphi process, we consulted with leading researchers and practitioners in the fields of ML, public health, and ethics. We also searched online databases and professional networks to identify potential experts. We aimed to include experts from diverse backgrounds, including

XSL•FO
RenderX

academia, industry, and government, with 17 chosen to be a part of the guideline panel (Multimedia Appendix 1). Our guideline panel includes academics in computer science, statistical modeling, epidemiology (clinical epidemiology, social epidemiology, infectious disease, and population health), health economics, ethics, sociology, primary care, public health, and the social determinants of health. Each member has contributed their disciplinary expertise and perspective regarding the social constructions influencing data, bias, and bias mitigation (Multimedia Appendix 1). Our panel members also bring diverse experiential knowledge, aligning with our guidelines' emphasis on recognizing the importance of lived experience in equitable policy development. Panel members did not report any direct conflicts of interest. To ensure transparency, members also agreed to disclose any less direct or potential competing interests, but no such conflicts were found.

Review of the Literature

We conducted a series of literature searches to retrieve existing guidelines related to population and public health, focusing on their intersections with equity, technical aspects, and knowledge mobilization. We searched a range of sources, including original research, reviews, commentaries, editorials, and gray literature (eg, government reports, policy documents, and organizational websites). In the course of this work, we drew on common elements found in other guidelines addressing bias in ML models [10,18,23,24]. These included practices such as engaging with communities impacted by biased models, prioritizing diversity in the teams developing and implementing models, considering which groups considered marginalized could benefit most from involvement in population and public health ML model development, ensuring robustness in the identification of groups considered vulnerable in data, systematically evaluating and reporting on bias, and conducting post-implementation evaluations to continue mitigating bias. In parallel, we conducted and reported three scoping reviews on ML applications in population and public health, specifically examining their use in addressing risk factors, NCDs, and communicable diseases. Two of these reviews have been published, providing an empirical foundation for our guidelines[25,26]. The first review found that bias mitigation was rarely addressed in ML studies focused on population health, particularly in the context of NCDs, with most efforts limited to addressing sex-related bias [25]. The second review, which examined ML applications for addressing risk factors for NCDs, found that although nine of the 20 studies mentioned algorithmic bias, these discussions were generally superficial and limited to traditional biases such as recall or misclassification [26]. To date, few systematic reviews have examined bias in ML applications for public health. Our findings are supported by another systematic review on the use of AI and ML in disaster response and public health emergencies, which similarly noted a lack of bias mitigation strategies despite the growing use of ML in these settings [27]. Our guidelines address this gap by offering operational recommendations that emphasize transparency, stakeholder engagement, and context-specific risk assessment [25]. Finally, we shared our evidence synthesis with the guideline committee alongside our draft recommendations.

Topic Selection and Development of Recommendations

We used a modified Delphi approach to develop our recommendations, consisting of two iterative rounds. An initial list of relevant topics for ML applications in population and public health was compiled by SD and SB, based on a review of gray literature, academic publications, current practice reports, and consultations with the study team.

In round 1, this list was refined during three virtual workshops held on November 24, 2022, December 9, 2022, and December 13, 2022. These sessions focused on ethics and equity, technical considerations, and knowledge mobilization. During the workshops, participants assessed draft recommendations using a modified Likert scale and provided written feedback. SB and SD analyzed the quantitative scores and qualitative input to identify consensus and areas of divergence.

In round 2, we circulated the revised draft recommendations to the panel in advance of a follow-up virtual meeting. Panel members discussed the content and submitted additional written feedback. We incorporated these inputs to finalize the recommendations, drawing on frameworks such as GRADE and other relevant guideline development methodologies.

The final recommendations were organized into five thematic categories: (1) equity, diversity, and inclusion; (2) public health emergencies, deidentified population data use, and consent; (3) due diligence during model conceptualization and early development; (4) technical transparency, consistency, and data management; and (5) knowledge mobilization to the public, ML experts, and public health professionals.

Recommendations

Recommendation 1: Prioritize Partnerships and Interventions That Support Communities That Are Disadvantaged by Social and Economic Policies

This includes activities such as algorithmic bias mitigation, capacity building, and fostering equitable representation. Partnerships should not only engage diverse collaborators with expertise in ethical AI and public health but also invest in building local capacity to support sustainable and responsible ML deployment in diverse settings.

ML models may be prone to various types of bias, which can significantly impact the health and well-being of communities made vulnerable by social and economic policies. ML approaches that rely on data from internet searches or social media may lack representativeness, tending to exhibit a bias toward processing data from individuals with higher socioeconomic status, from certain age groups, or those living in urban areas rather than remote or rural areas [6]. Similarly, ML models used in population health often encounter missing or nonrepresentative data when relying on electronic medical records or public health data, which may not fully capture the diversity of the communities they aim to serve [28,29].

XSL•FO
RenderX

Members of diverse community groups, including communities disadvantaged by social and economic policies, bring a variety of perspectives and lived experiences to the table. These perspectives and insights must be actively prioritized in decision-making to identify and mitigate algorithm-related harms effectively [30]. Such experiences are crucial for meaningful research and transformative change within institutions. Expert consultation, while valuable, should not replace the direct involvement of people with lived experience; instead, it should complement it [31]. Unfortunately, those who stand to benefit most from fair and responsible ML applications are also those most at risk of algorithmic harm due to systemic biases and exclusion [32].

Prioritizing equity in ML requires balancing ethical considerations with technical feasibility and real-world implementation challenges. While transparency is critical for trust and accountability, full transparency can be resource intensive and difficult to achieve in low- and middle-income countries (LMICs), where data infrastructure and regulatory oversight may be limited. Similarly, while equity-driven approaches aim to reduce algorithmic harms, they require sustained investment in local expertise, which may not always be immediately available. Feasibility must therefore be considered at each stage, ensuring that solutions are both ethically sound and realistically implementable in diverse public health settings.

In LMICs, where these guidelines are particularly relevant, implementing ML partnerships and equity-focused interventions is often constrained by financial limitations, workforce shortages, and fragmented health data systems. Ensuring that ML applications promote equity while remaining feasible requires balancing ethical considerations with practical constraints. Many public health institutions in LMICs operate with competing priorities, making it challenging to allocate resources solely toward algorithmic transparency, bias mitigation, or extensive auditing processes. Policymakers, health care organizations, and ML developers in these settings must adopt scalable, context-aware approaches that allow for phased implementation. Resource-limited settings can start with standard approaches to data definition and extraction, streamlined documentation, internal monitoring, and bias assessments targeted at high-risk applications where disparities are most pronounced. Furthermore, the literature from LMICs emphasizes the need for tailored implementation approaches. In Africa, for example, ongoing discussions have identified key priorities such as strengthening electricity and internet infrastructure [33], expanding the data science workforce through increased educational opportunities [33], leveraging smartphone-based AI apps [33], and developing AI frameworks that reflect regional needs [34]. In addition, fostering multisectoral partnerships and policy initiatives has been recommended as a way to incentivize and support responsible AI adoption [35].

The trade-offs between transparency, equity, and operational feasibility must also be carefully managed. While transparency fosters trust and accountability, excessive documentation requirements or mandatory external reviews may slow the deployment of ML-driven health interventions, particularly in rapidly evolving crises such as infectious disease outbreaks. In such cases, prioritizing community engagement, establishing advisory panels with representatives from populations considered marginalized, and implementing practical bias mitigation strategies such as integrating socioeconomic variables into model adjustments can ensure ethical safeguards while maintaining efficiency. Flexible regulatory frameworks that accommodate local resource constraints can allow LMICs to adopt ML responsibly without overwhelming already burdened health care infrastructures.

Partnering directly with communities can help address these gaps. This includes building trust, capacity building, and advancing representation [32]. Bias in ML models can stem from various sources, including data collection, label selection, and feature inclusion. Stakeholders from communities considered disadvantaged should be engaged not only to address data biases but also to guide decisions on selecting fair labels and features. For example, using health care expenditure as a proxy for need can introduce bias if spending patterns differ across communities due to economic disparities. Direct involvement of community members can also help identify potential biases in data sources and variables, as well as develop more equitable models. For instance, in the development of predictive models for health care access, community insights can reveal overlooked barriers, such as transport challenges or financial constraints. To mitigate this, ML users should prioritize diverse and representative datasets, monitor error rates and performance levels across different patient groups, and consider the downstream implications (eg, potential impacts on health care access and quality of public health interventions) [32].

True partnership goes beyond consultation; it requires ongoing collaboration and investment in local expertise. Many LMICs and resource-constrained settings lack appropriate safeguards, technical expertise, and infrastructure to implement ML tools effectively. Addressing these gaps requires more than just ethical guidelines. It demands direct investment in skills, resources, and institutional support. Practical strategies to foster capacity building include mentorship programs that connect ML experts in high-income countries with local researchers and policymakers to support training and facilitate knowledge exchange. For example, codeveloping open-access educational resources tailored to public health professionals can help increase AI literacy. Sharing technical resources, such as code and cloud-based computing capacity, can support implementation, provided these efforts are led by LMIC stakeholders. Additional strategies include investing in affordable and scalable ML tools designed for LMICs to ensure usability even with limited infrastructure and fostering partnerships between public health institutions, universities, and community organizations to bridge knowledge gaps. By prioritizing these efforts, ML tools can be developed not only with but also by the communities they aim to serve.

Ultimately, capacity building is a long-term commitment that requires sustained funding, institutional support, and local leadership. While external collaborations provide valuable expertise, true sustainability depends on empowering local communities to take ownership of ML initiatives. This ensures that ML applications in public health are not only effective and

XSL•FO
RenderX

transparent but also grounded in the realities and needs of the populations they serve [36].

Recommendation 2: Use ML in Public Health Emergencies and Other Dynamic, Fast-Paced Situations by Collecting, Analyzing, and Using Population-Wide Deidentified Data

ML can be useful in public health emergencies and other dynamic, fast-paced situations by collecting, analyzing, and using population-wide deidentified data. This process of information gathering and manipulation of deidentified data can be carried out without consent, provided that ethical safeguards are in place and risks related to privacy, misuse, and transparency are actively mitigated.

ML plays a crucial role in gathering, reporting, and analyzing information rapidly, which is vital for mitigating further health damage in dynamic, fast-paced situations such as public health emergencies. According to the United Nations' Sendai framework, reducing disaster risk involves understanding situational risk, strengthening governance, improving preparedness for an effective response, and allocating resources toward measures that can enhance resilience [37]. Prediction models and protocols, such as evacuation planning, have been used to alleviate the adverse effects of such emergencies. ML has greatly enhanced the ability to monitor information and make timely decisions during emergencies. This has enabled disease outbreak prediction, improved evacuation planning, and optimized the distribution of resources to areas in need [27].

During a public health emergency, the rapid collection of deidentified data without consent may be necessary to ensure timely responses. However, this urgency must be balanced with the potential risks of data misuse, privacy breaches, and lack of accountability, particularly in regions with weaker institutional safeguards. In LMICs and other resource-constrained settings, insufficient regulatory oversight may increase the risk of unauthorized access, data exploitation, or reidentification of individuals.

To minimize these risks while maintaining the efficiency of ML-driven public health responses, organizations and governments should adopt structured ethical frameworks that ensure transparency, accountability, and proportionality in data use. Principles such as FAIR (Findable, Accessible, Interoperable, and Reusable) and CARE (Collective Benefit, Authority to Control, Responsibility, and Ethics) offer structured guidelines for ethical data collection and governance and thus can be used in contexts where regulatory oversight is limited. The FAIR principles focus on making health data widely accessible and reusable, which is essential for rapid public health responses while maintaining data integrity and security. The CARE principles, on the other hand, ensure that data collection respects collective benefit and community authority, particularly in Indigenous communities and those considered marginalized, preventing exploitative data use [38]. In addition to regulatory measures, capacity-building efforts should prioritize the establishment of independent data governance bodies within LMICs to oversee ethical ML implementation. Encouraging multistakeholder collaborations, including engagement with civil society organizations and the communities affected, can further enhance accountability. By embedding these safeguards, public health responses can remain rapid and effective without compromising ethical considerations, even in regions with limited regulatory infrastructure.

Similarly, determining whether individuals or populations would be willing to disclose certain information may be complex. Most people and organizations recognize that some degree of data collection without consultation may be necessary. During the COVID-19 pandemic, researchers effectively used deidentified census data to identify high-risk neighborhoods without individual consent [39]. This approach can serve as a model for using public health data during emergencies while adhering to ethical standards. Ethical frameworks related to consent, privacy of information, and their implications for population and public health, as well as crisis management, can help evaluate the ethical feasibility of proposed solutions, even if they do not address all aspects of the current situation. In their development of an ethical data access framework, Lusignan et al [40] outline that structuring collective as well as organizational thinking and decision-making regarding what is and is not appropriate within a crisis early on makes for improved trust between the data providers and recipients, as expectations may be collaboratively set, negotiated, and ideally fulfilled or exceeded. To accelerate the use of deidentified data in public health responses, ethical frameworks should allow faster-access pathways for researchers who can demonstrate secure data handling and infrastructure. This could involve streamlined approvals for projects with proven privacy measures.

Ideally, this process should occur during the interpandemic phase, using inclusive structures and governance models with oversight powers. This process should determine whether the benefits to society, such as reducing long-term loss of life and health, outweigh the trade-offs involving public interest, privacy breaches, and ethical considerations. To ensure the enforcement of these guidelines in real-world scenarios, particularly where regulatory oversight is lacking, we propose several mechanisms. One approach is to integrate these guidelines into existing public health governance structures, ensuring alignment with established ethical and legal frameworks [41]. In addition, public health institutions and regulatory bodies could establish independent review panels to assess compliance with these ML guidelines [42]. Another enforcement mechanism could be the development of accreditation or certification programs for ML applications in public health, which would require adherence to ethical and transparency standards [43].

Recommendation 3: Ensure Fairness Across Populations and Mitigate Bias

The level of bias mitigation should be proportional to the context and potential impact of the model, ensuring that even low-risk applications maintain essential fairness safeguards. All aspects of the model's implementation should be assessed, including inherent, external, or incompletely known factors that could contribute to bias.

Risk assessment is central to any innovation, but it is especially important for ML, which is prone to certain liabilities [44]. The use of algorithms in health care can perpetuate racial biases due

XSL•FO
RenderX

to existing disparities in health care delivery. When designing algorithms to predict health outcomes based on genetic findings, bias may arise if there is limited or no research conducted in certain populations. For instance, using data from the Framingham Heart Study to predict cardiovascular risk in non-White populations has led to biased results, overestimating or underestimating the risk [45].

To clarify how risk relates to bias mitigation, risk should be defined based on the potential for harm or inequity, particularly to populations considered vulnerable, rather than solely on model complexity or scope. Risk assessment should consider the model's intended use, the sensitivity of the outcome (eg, health care access or resource allocation), and the potential for systemic harm if biases are present. Importantly, a lower risk designation should never justify a lax approach to bias mitigation. Instead, risk levels should guide the *type* of mitigation strategies deployed, with simpler models undergoing appropriate fairness checks and more complex, high-impact models requiring comprehensive audits and subgroup analyses.

Given the wide range of ML models, conducting situation-specific risk assessments, especially for populations that may be vulnerable due to social or economic policies, is essential. We need to develop protocols to ensure fairness across diverse populations, using metrics such as subgroup performance, false-positive or false-negative rates, and bias-specific checks (eg, checking for overrepresentation or underrepresentation of particular groups). Transparency should extend to data sources, model methodologies, and decisions made during development. Providing accessible documentation, including technical reports and simplified explanations, can help both technical and nontechnical stakeholders understand potential biases and their mitigation [32]. Transparency should highlight how different populations are represented in training data and how model outputs vary among them. For example, models predicting health care access should clearly show performance disparities between urban and rural populations.

Transparency, explainability, and interpretability of ML models are critical for stakeholder trust and decision-making. Transparency refers to how openly the model's design, data sources, and decision-making processes are shared, helping stakeholders understand how the model operates. Efforts should be made to improve the transparency of models, such as through clear explanations of how inputs influence outputs, visual aids, and plain language summaries. Explainability focuses on making specific predictions understandable, clarifying why a model produced a particular output through methods such as feature importance analysis or decision trees. Interpretability relates to how easily a human can understand the model's internal logic or decision-making process, which is often linked to model simplicity or the use of interpretable techniques such as linear regression or rule-based systems. Altogether, these concepts ensure that ML models are accessible, trustworthy, and accountable to both technical and nontechnical stakeholders.

For example, using interpretable models in clinical settings can help physicians understand and validate AI recommendations, enhancing adoption and reliability. ML is sometimes portrayed as a technology requiring advanced training to understand, leading it to both hopeful expectations and suspicion. It is important to clearly communicate the inherent and associated risks of the model and define what is meant by ML in its context [32]. When ML-related technologies and their risks are explained in plain language, public distrust tends to decrease [46].

Recommendation 4: Ensure Public Availability of Data Sources, Model Methodologies, and Technical Details, Along With Bias Mitigation Strategies, Used in the Contexts of Models to Promote Transparency, Reproducibility, and Trustworthiness Across ML Studies

Providing accessible information about the technical aspects of an ML study or solution, such as its data sources, population, characteristics, and model variables, along with detailed descriptions of its methodology and the deidentified datasets used, supports reproducibility, bias mitigation, and trustworthiness [32]. However, ensuring public availability of data must be balanced with the sensitivity of health data, privacy regulations, and proprietary constraints. To promote meaningful transparency, a structured approach is necessary, one that goes beyond simply sharing code and considers regulatory, ethical, and contextual limitations. Drawing on European Union legislation and literature in computer science, Kiselva and De Hert [47] suggest that transparency must be seen as a fundamental "way of thinking" and an all-encompassing concept that characterizes the process of developing and using AI.

Achieving transparency requires balancing openness with the risks of reidentification, legal restrictions, and proprietary interests. While transparency enhances trust in ML models, full public disclosure of health data is not always feasible due to privacy laws such as the Health Insurance Portability and Accountability Act (HIPAA) and General Data Protection Regulation (GDPR), institutional data-sharing policies, and concerns over commercial confidentiality.

In public health contexts, transparency should function as a system of accountability rather than unrestricted access to all datasets. ML documentation should be structured to ensure usability while maintaining compliance with privacy and proprietary safeguards. This includes providing methodological details and bias mitigation strategies, ensuring that plain language summaries are available for nontechnical stakeholders, and adopting tiered-access models where sensitive data can be reviewed under controlled conditions rather than being made fully public. The need for transparency should also be weighed against the practical barriers to data sharing. Health data often originate from multiple institutions with varying governance policies, making harmonization difficult. In addition, proprietary models developed by industry partners may involve intellectual property protections that restrict full disclosure of methodologies. Addressing these challenges requires the development of standardized documentation and data governance agreements that balance the need for public trust with confidentiality concerns.

To enhance data sharing, transparency, and bias mitigation, dataset development protocols should align with existing

XSL•FO
RenderX

frameworks to integrate new and prior information. Data used for model training, validation, or implementation should be of high quality (eg, completeness, source consistency, and linkage potential) and consistently accessible and updated, especially data related to employment, education, occupation, other socioeconomic status factors, and health inequalities [48]. In applications of linked data and ML in the health sciences (eg, to estimate population-level health indicators), bias can arise from nonstandard data collection methodologies. Developing standardized protocols for all data sources used in each project and including variance assessment and context-appropriate handling (eg, oversampling, imputation) in bias mitigation strategies are essential. For example, in response to bias concerns during the COVID-19 pandemic, ML models were adjusted to incorporate sociodemographic data to improve accuracy in predicting infection spread among populations considered marginalized [49]. Another successful example includes the use of algorithmic audits in health care AI applications, where independent researchers identified and mitigated racial biases in predictive models used for hospital readmission rates [50]. Transparency should include public documentation of these bias mitigation strategies while ensuring that sensitive details remain protected.

Assessing bias and performance across subpopulations is another essential element of responsible transparency. For public health surveillance, sensitivity may be prioritized to avoid missing cases, whereas models used for treatment decisions may emphasize specificity to minimize unnecessary interventions. Metrics such as area under the curve, positive predictive value, and user satisfaction should be evaluated across different subpopulations to detect and correct bias. It is equally important to identify variance in datasets, as high variance can exacerbate bias in ML models. Underfitting, which occurs when a model fails to capture important patterns, and overfitting, where a model becomes hypersensitive to minor fluctuations in data, should be addressed through continuous model tuning and hyperparameter optimization.

In addition to predeployment fairness checks, postdeployment monitoring should include regular performance evaluations across diverse subpopulations to detect bias drift, which can emerge as population characteristics change. Bias drift occurs when the model's predictive performance deteriorates or produces skewed results over time, often reflecting shifts in population demographics or health care access patterns. We also highlight the importance of model update protocols incorporating community feedback loops, participatory audits, and transparent reporting of adjustments to ensure that updates are responsive and equitable.

We recommend that organizations implement structured, iterative audit cycles to support these processes. These cycles should include predeployment fairness checks using simulated population data and postdeployment ethical audits conducted at regular intervals or following significant model updates. In addition, establishing rapid-response mechanisms to address emerging ethical concerns during deployment can help maintain accountability. Incorporating stakeholder feedback from impacted communities through participatory reviews ensures that evaluations remain inclusive and socially responsive.

Recommendation 5: Facilitate Regular, Multidisciplinary Discussions Involving ML Developers, Public Health Professionals, Ethicists, and Community Representatives to Identify Biases, Ensure Fair Implementation, and Increase Transparency

Provide plain language summaries and guidelines that are consistent in terminology to raise awareness among both the public and experts about ML-related bias and debiasing.

Efforts to raise public awareness about ML and its benefits should be balanced with accurate information about potential harms that may be associated with its implementation. In a study by Musbahi et al [51], the views of patients and the public about AI in health care were analyzed. The top 5 concerns included decreased human interaction, data security, obtaining consent for data use, errors in AI systems, and the potential irrelevance of AI in health care. Despite factors promoting the adoption of technology in health care settings, achieving sustainable implementation remains a challenge. Providing information to address individual concerns about the safety and effectiveness of ML models is essential. Public disclosure and scientific interrogation of potentially harmful occurrences or risks associated with ML, including unintentional biases, build trust with the public. Efforts to raise awareness should use accessible language and incorporate research best practices (eg, consistent terminology, structured abstracts, and established reporting guidelines) to make technical concepts easier to understand for a wider audience [46]. This approach not only raises awareness but also fosters trust among nontechnical stakeholders, such as community members and health care practitioners. Past multidisciplinary engagements have proven valuable in addressing bias and transparency concerns in ML applications. For instance, the World Health Organization convened multistakeholder panels to assess AI-based disease surveillance tools, leading to refined ethical guidelines [52]. Similarly, the Canadian AI for Public Health initiative facilitated workshops bringing together data scientists, epidemiologists, and policymakers to discuss bias mitigation strategies in ML-based public health interventions [53]. A structured approach to future engagements could include regular stakeholder summits, interdisciplinary task forces, and community-centered consultations.

Summary

ML is rapidly advancing and holds potential for uses to improve the health of individuals and communities. However, these efforts must prioritize equity. Model developers, statisticians, epidemiologists, public health professionals, policymakers, and funders must collaborate to ensure that ML implementations avoid prejudice and discrimination while also enhancing human capabilities, connections, and knowledge in health and disease contexts. This approach aligns with the ethical integration of AI in health [54].

Our recommendations align with existing frameworks. In the United States, our guidelines support the principles outlined in the Executive Order on Safe, Secure, and Trustworthy AI, which emphasizes the importance of equity, privacy safeguards, and

XSL•FO
RenderX

risk mitigation in health-related AI applications [55]. Similarly, our focus on transparency and bias mitigation aligns with the European Union's AI Act and GDPR, which collectively promote ethical AI deployment, data protection, and algorithmic accountability [56,57]. Our recommendations complement the UK AI Regulatory Principles, which emphasize fairness, accountability, and transparency, particularly in public sector applications [58]. Finally, our guidelines build upon principles from the Pan-Canadian Artificial Intelligence Strategy [59], which highlights equity, interdisciplinarity, and inclusivity as core pillars of responsible AI adoption in public health. These country-specific frameworks emphasize the importance of adapting ethical ML guidelines to regional regulatory and social contexts, while ensuring global alignment on core principles such as equity, transparency, and data privacy. Our recommendations add to these discussions by offering a public health–specific lens, focusing on bias mitigation in population-level models, fostering community partnerships, and emphasizing the impact of social determinants of health in ML implementation.

We had a diverse range of disciplines and approaches represented in our team, including AI, ML, computer science, population health, epidemiology, ethics, and AI applications in health. A limitation of this work was that we have adapted methodologies initially designed for developing clinical practice guidelines, as there is no clear guidance for developing ML applications in population and public health. In LMICs, there might be limited resources available for adequately monitoring and documenting public health interventions that use ML. While our recommendations aim to promote equity and enhance the integration of ML models in LMICs for population and public health purposes, they may not cover all possible use cases. Implementing these guidelines in diverse settings presents several challenges, including variations in data quality, resource availability, and stakeholder engagement. Addressing these challenges is essential to ensure that ML models enhance public health outcomes equitably.

Limitations

Some key limitations include difficulties in engaging communities considered disadvantaged, ensuring data representativeness, and maintaining transparency with limited resources. Operationalizing these recommendations will require adaptable protocols, increased local capacity, and collaborative efforts. For example, developing simplified, context-specific inclusivity checklists and leveraging international partnerships can help overcome resource gaps. Ultimately, ongoing evaluation and stakeholder input are crucial for refining these guidelines and ensuring their effectiveness across varied public health contexts.

Conclusions

Without adequate bias prevention and mitigation during ML model development and implementation, ML applications in population and public health contexts could worsen existing health disparities or even contribute to new ones. Similar to the development of health policy aimed at promoting equity, it is crucial to carefully assess ML innovations before, during, and after their development and deployment. This ensures that model design and delivery are equitable and based on data representing all groups. Achieving equity requires adhering to ethical principles of equity, transparency, and engagement, as well as multidisciplinary efforts involving both model developers and population health practitioners committed to bias prevention and mitigation standards. Monitoring the outcomes of such adherence, including the guidelines proposed here, can promote less biased use of ML to inform policy. Future work should include piloting these guidelines in a range of public health settings, including LMICs, and developing practical tools to assess their effectiveness. Evaluation efforts could involve tracking the diversity of stakeholders involved in model development and testing, measuring improvements in model accuracy for underrepresented groups, and assessing whether decisions based on these models contribute to more equitable health outcomes. Ongoing assessment and adaptation of these guidelines will be crucial for ensuring responsible and fair use of ML in public health. ML is not just a rapidly evolving technology but also a tool that can promote equity in population and public health if a commitment to mitigating bias is maintained by all stakeholders involved in its design and delivery.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Guideline team composition.
[DOCX File, 16 KB - publichealth_v11i1e68952_app1.docx]

References

1. Beam AL, Kohane IS. Big data and machine learning in health care. JAMA 2018 Apr 3;319(13):1317-1318. [doi: 10.1001/jama.2017.18391] [Medline: 29532063]
2. Friedman DJ, Starfield B. Models of population health: their value for US public health practice, policy, and research. Am J Public Health 2003 Mar;93(3):366-369. [doi: 10.2105/ajph.93.3.366] [Medline: 12604473]
3. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning: Springer; 2013.

XSL•FO
RenderX

4. Atasoy H, Greenwood BN, McCullough JS. The digitization of patient care: a review of the effects of electronic health records on health care quality and utilization. Annu Rev Public Health 2019 Apr 1;40:487-500. [doi: 10.1146/annurev-publhealth-040218-044206] [Medline: 30566385]
5. Lavigne M, Mussa F, Creatore MI, Hoffman SJ, Buckeridge DL. A population health perspective on artificial intelligence. Healthc Manage Forum 2019 Jul;32(4):173-177. [doi: 10.1177/0840470419848428] [Medline: 31106580]
6. Aiello AE, Renson A, Zivich PN. Social media- and internet-based disease surveillance for public health. Annu Rev Public Health 2020 Apr 2;41(1):101-118. [doi: 10.1146/annurev-publhealth-040119-094402] [Medline: 31905322]
7. Silva KD, Lee WK, Forbes A, Demmer RT, Barton C, Enticott J. Use and performance of machine learning models for type 2 diabetes prediction in community settings: a systematic review and meta-analysis. Int J Med Inform 2020 Nov;143:104268. [doi: 10.1016/j.ijmedinf.2020.104268] [Medline: 32950874]
8. Allem JP, Ferrara E, Uppu SP, Cruz TB, Unger JB. E-Cigarette surveillance with social media data: social bots, emerging topics, and trends. JMIR Public Health Surveill 2017 Dec 20;3(4):e98. [doi: 10.2196/publichealth.8641] [Medline: 29263018]
9. AI, machine learning and the potential impacts on the practice of family medicine. AMS. URL: https://www.ams-inc.on.ca/resource/cfpc-briefing-paper/ [accessed 2025-10-01]
10. Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. Practical guidance on artificial intelligence for health-care data. Lancet Digit Health 2019 Aug;1(4):e157-e159. [doi: 10.1016/S2589-7500(19)30084-6] [Medline: 33323184]
11. Wilder J, Saraswathula A, Hasselblad V, Muir A. A systematic review of race and ethnicity in Hepatitis C clinical trial enrollment. J Natl Med Assoc 2016 Feb;108(1):24-29. [doi: 10.1016/j.jnma.2015.12.004] [Medline: 26928485]
12. Buckeridge DL. Precision, equity, and public health and epidemiology informatics - a scoping review. Yearb Med Inform 2020 Aug;29(1):226-230. [doi: 10.1055/s-0040-1701989] [Medline: 32823320]
13. Röösli E, Rice B, Hernandez-Boussard T. Bias at warp speed: how AI may contribute to the disparities gap in the time of COVID-19. J Am Med Inform Assoc 2021 Jan 15;28(1):190-192. [doi: 10.1093/jamia/ocaa210] [Medline: 32805004]
14. Schünemann H, Brożek J, Guyatt G, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendations. Grading of Recommendations Assessment, Development and Evaluation. 2013. URL: https://gdt.gradepro.org/app/handbook/handbook.html [accessed 2025-10-01]
15. Developing NICE guidelines: the manual. National Institute for Health and Care Excellence. 2014. URL: https://www.nice.org.uk/media/default/about/what-we-do/our-programmes/developing-nice-guidelines-the-manual.pdf [accessed 2025-10-01]
16. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices—overview. Med Decis Making 2012 Sep;32(5):667-677. [doi: 10.1177/0272989X12454577] [Medline: 22990082]
17. Eddy DM, Hollingworth W, Caro JJ, et al. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--7. Value Health 2012;15(6):843-850. [doi: 10.1016/j.jval.2012.04.012] [Medline: 22999134]
18. Lee NT, Resnick P, Barton G. Algorithmic bias detection and mitigation: best practices and policies to reduce consumer harms. Brookings Institution. 2019. URL: https://www.brookings.edu/articles/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/ [accessed 2025-10-01]
19. Paprica PA, Sutherland E, Smith A, et al. Essential requirements for establishing and operating data trusts. Int J Popul Data Sci 2020;5(1). [doi: 10.23889/ijpds.v5i1.1353] [Medline: 33644412]
20. Montreal declaration on responsible AI. Montréal Declaration. 2018. URL: https://www.montrealdeclaration-responsibleai.com/ [accessed 2019-01-24]
21. Courtland R. Bias detectives: the researchers striving to make algorithms fair. Nature New Biol 2018 Jun;558(7710):357-360. [doi: 10.1038/d41586-018-05469-3] [Medline: 29925973]
22. Ethics and governance of artificial intelligence for health: WHO guidance. World Health Organization. URL: https://www.who.int/publications/i/item/9789240029200 [accessed 2023-04-05]
23. Wiens J, Price WN, Sjoding MW. Diagnosing bias in data-driven algorithms for healthcare. Nat Med 2020 Jan;26(1):25-26. [doi: 10.1038/s41591-019-0726-6] [Medline: 31932798]
24. Pfohl SR, Foryciarz A, Shah NH. An empirical characterization of fair machine learning for clinical risk prediction. J Biomed Inform 2021 Jan;113:103621. [doi: 10.1016/j.jbi.2020.103621] [Medline: 33220494]
25. Birdi S, Rabet R, Durant S, et al. Bias in machine learning applications to address non-communicable diseases at a population-level: a scoping review. BMC Public Health 2024;24(1):1-16. [doi: 10.1186/S12889-024-21081-9/FIGURES/2] [Medline: 39732655]
26. Shergill M, Durant S, Birdi S, et al. Machine learning used to study risk factors for chronic diseases: A scoping review. Can J Public Health 2025 Jun 11. [doi: 10.17269/s41997-025-01059-9] [Medline: 40498391]
27. Lu S, Christie GA, Nguyen TT, Freeman JD, Hsu EB. Applications of artificial intelligence and machine learning in disasters and public health emergencies. Disaster Med Public Health Prep 2022 Aug;16(4):1674-1681. [doi: 10.1017/dmp.2021.125] [Medline: 34134815]
28. Bozkurt S, Cahan EM, Seneviratne MG, et al. Reporting of demographic data and representativeness in machine learning models using electronic health records. J Am Med Inform Assoc 2020 Dec 9;27(12):1878-1884. [doi: 10.1093/jamia/ocaa164] [Medline: 32935131]

XSL•FO
RenderX

29. Larrazabal AJ, Nieto N, Peterson V, Milone DH, Ferrante E. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. Proc Natl Acad Sci U S A 2020 Jun 9;117(23):12592-12594. [doi: 10.1073/pnas.1919012117] [Medline: 32457147]
30. Wark K, Woodbury RB, LaBrie S, Trainor J, Freeman M, Avey JP. Engaging stakeholders in social determinants of health quality improvement efforts. Perm J 2022 Dec 19;26(4):28-38. [doi: 10.7812/TPP/22.035] [Medline: 36154895]
31. Jones N, Harrison J, Aguiar R, Munro L. Transforming research for transformative change in mental health: toward the future. In: Nelson G, Kloos B, Ornelas J, editors. Community Psychology and Community Mental Health: Oxford Academic Press; 2024:351-372. [doi: 10.1093/acprof:oso/9780199362424.001.0001]
32. Vokinger KN, Feuerriegel S, Kesselheim AS. Mitigating bias in machine learning for medicine. Commun Med (Lond) 2021 Aug 23;1(1):25. [doi: 10.1038/s43856-021-00028-w] [Medline: 34522916]
33. Owoyemi A, Owoyemi J, Osiyemi A, Boyd A. Artificial intelligence for healthcare in Africa. Front Digit Health 2020;2:6. [doi: 10.3389/fdgth.2020.00006] [Medline: 34713019]
34. Alaran MA, Lawal SK, Jiya MH, et al. Challenges and opportunities of artificial intelligence in African health space. Digit Health 2025;11:20552076241305915. [doi: 10.1177/20552076241305915] [Medline: 39839959]
35. Tanui CK, Ndembi N, Kebede Y, Tessema SK. Artificial intelligence to transform public health in Africa. Lancet Infect Dis 2024 Sep;24(9):e542. [doi: 10.1016/S1473-3099(24)00435-3] [Medline: 39053483]
36. Baxter MS, White A, Lahti M, Murto T, Evans J. Machine learning in a time of COVID-19 - can machine learning support community health workers (CHWs) in low and middle income countries (LMICs) in the new normal? J Glob Health 2021 Jan 16;11(3017):03017. [doi: 10.7189/jogh.11.03017] [Medline: 33643627]
37. Aitsi-Selmi A, Murray V. The Sendai framework: disaster risk reduction through a health lens. Bull World Health Organ 2015 Jun 1;93(6):362. [doi: 10.2471/BLT.15.157362] [Medline: 26240454]
38. Carroll SR, Herczog E, Hudson M, Russell K, Stall S. Operationalizing the CARE and FAIR principles for Indigenous data futures. Sci Data 2021 Apr 16;8(1):108. [doi: 10.1038/s41597-021-00892-0] [Medline: 33863927]
39. Xiao C, Zhou J, Huang J, et al. C-watcher: a framework for early detection of high-risk neighborhoods ahead of COVID-19 outbreak. Proc AAAI Conf Artif Intell 2021;35(6):4892-4900. [doi: 10.1609/aaai.v35i6.16622]
40. De Lusignan S, Liyanage H, Di Iorio CT, Chan T, Liaw ST. Using routinely collected health data for surveillance, quality improvement and research: framework and key questions to assess ethics, privacy and data access. J Innov Health Inform 2016 Jan 19;22(4):426-432. [doi: 10.14236/jhi.v22i4.845] [Medline: 26855276]
41. Floridi L, Cowls J, Beltrametti M, et al. AI4People-an ethical framework for a good AI society: opportunities, risks, principles, and recommendations. Minds Mach (Dordr) 2018;28(4):689-707. [doi: 10.1007/s11023-018-9482-5] [Medline: 30930541]
42. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, Ghassemi M. Ethical machine learning in healthcare. Annu Rev Biomed Data Sci 2021 Jul;4:123-144. [doi: 10.1146/annurev-biodatasci-092820-114757] [Medline: 34396058]
43. Simon G, Aliferis C. Reporting standards, certification/accreditation, and reproducibility. In: Simon GJ, Aliferis C, editors. Artificial Intelligence and Machine Learning in Health Care and Medical Sciences: Best Practices and Pitfalls: Springer; 2024:693-707. [doi: 10.1007/978-3-031-39355-6_17]
44. Char DS, Shah NH, Magnus D. Implementing machine learning in health care - addressing ethical challenges. N Engl J Med 2018 Mar 15;378(11):981-983. [doi: 10.1056/NEJMp1714229] [Medline: 29539284]
45. Gijsberts CM, Groenewegen KA, Hoefer IE, et al. Race/ethnic differences in the associations of the Framingham risk factors with carotid IMT and cardiovascular events. PLoS ONE 2015;10(7):e0132321. [doi: 10.1371/journal.pone.0132321] [Medline: 26134404]
46. Edgell C, Rosenberg A. Putting plain language summaries into perspective. Curr Med Res Opin 2022 Jun;38(6):871-874. [doi: 10.1080/03007995.2022.2058812] [Medline: 35400253]
47. Kiseleva A, De Hert P. Creating a European health data space. Obstacles in four key legal areas. SSRN Journal. Preprint posted online on May 18, 2021. [doi: 10.2139/ssrn.3846781]
48. Haneef R, Tijhuis M, Thiébaut R, et al. Methodological guidelines to estimate population-based health indicators using linked data and/or machine learning techniques. Arch Public Health 2022 Jan 4;80(1):9. [doi: 10.1186/s13690-021-00770-6] [Medline: 34983651]
49. Zhao AP, Li S, Cao Z, et al. AI for science: predicting infectious diseases. J Saf Sci Resil 2024 Jun;5(2):130-146. [doi: 10.1016/j.jnlssr.2024.02.002]
50. Wang HE, Weiner JP, Saria S, Kharrazi H. Evaluating algorithmic bias in 30-day hospital readmission models: retrospective analysis. J Med Internet Res 2024 Apr 18;26(1):e47125. [doi: 10.2196/47125] [Medline: 38422347]
51. Musbahi O, Syed L, Le Feuvre P, Cobb J, Jones G. Public patient views of artificial intelligence in healthcare: a nominal group technique study. Digit Health 2021;7:20552076211063682. [doi: 10.1177/20552076211063682] [Medline: 34950499]
52. Blouin Genest G. World Health Organization and disease surveillance: jeopardizing global public health? Health (London) 2015 Nov;19(6):595-614. [doi: 10.1177/1363459314561771] [Medline: 25504474]
53. Fisher S, Rosella LC. Priorities for successful use of artificial intelligence by public health organizations: a literature review. BMC Public Health 2022 Nov 22;22(1):2146. [doi: 10.1186/s12889-022-14422-z] [Medline: 36419010]

XSL•FO
RenderX

54. Russo F, Schliesser E, Wagemans J. Connecting ethics and epistemology of AI. AI & Soc 2024 Aug;39(4):1585-1603. [doi: 10.1007/s00146-022-01617-6]
55. Safe, secure, and trustworthy development and use of artificial intelligence. Federal Register. URL: https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence [accessed 2025-05-09]
56. Regulation (EU) 2024/1689 of the european parliament and of the council of 13 june 2024 laying down harmonised rules on artificial intelligence and amending regulations (EC) no 300/2008, (EU) no 167/2013, (EU) no 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (artificial intelligence act). European Union. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj [accessed 2025-05-09]
57. General Data Protection Regulation (GDPR) – legal text. Intersoft Consulting Services. URL: https://gdpr-info.eu/ [accessed 2025-05-09]
58. Implementing the UK's AI regulatory principles: initial guidance for regulators. Government of UK. 2024. URL: https://www.gov.uk/government/publications/implementing-the-uks-ai-regulatory-principles-initial-guidance-for-regulators [accessed 2025-10-01]
59. Pan-Canadian artificial intelligence strategy. Government of Canada. URL: https://ised-isde.canada.ca/site/ai-strategy/en [accessed 2025-10-01]

Abbreviations

AI: artificial intelligence
CARE: Collective Benefit, Authority to Control, Responsibility, and Ethics
FAIR: Findable, Accessible, Interoperable, and Reusable
GDPR: General Data Protection Regulation
GRADE: Grading of Recommendations Assessment, Development and Evaluation
HIPAA: Health Insurance Portability and Accountability Act
LMIC: low- and middle-income country
ML: machine learning
NCD: noncommunicable disease
NICE: National Institute for Health and Care Excellence

Edited by OL Neto; submitted 18.11.24; peer-reviewed by E Chuang, R Ladhania, S Oworah; revised version received 06.06.25; accepted 06.06.25; published 24.10.25.

Please cite as:
Pinto AD, Birdi S, Durant S, Rabet R, Parekh R, Ali S, Buckeridge D, Ghassemi M, Gibson J, John-Baptiste A, Macklin J, McCradden MD, McKenzie K, Naraei P, Owusu-Bempah A, Rosella LC, Shaw J, Upshur R, Mishra S
Machine Learning Applications in Population and Public Health: Guidelines for Development, Testing, and Implementation
JMIR Public Health Surveill 2025;11:e68952
URL: https://publichealth.jmir.org/2025/1/e68952
doi:10.2196/68952

© Andrew D Pinto, Sharon Birdi, Steve Durant, Roxana Rabet, Rahul Parekh, Shehzad Ali, David Buckeridge, Marzyeh Ghassemi, Jennifer Gibson, Ava John-Baptiste, Jillian Macklin, Melissa D McCradden, Kwame McKenzie, Parisa Naraei, Akwasi Owusu-Bempah, Laura C Rosella, James Shaw, Ross Upshur, Sharmistha Mishra. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Ethical Considerations for Wastewater Surveillance Conducted by the US Department of Defense

Hunter Jackson Smith[1,2], MD, MPH, MBE; Richard T Agans[3], PhD; William J Kowallis[4,5], PhD

[1]Global Emerging Infections Surveillance Branch, Armed Forces Health Surveillance Division, 11800 Tech Rd Suite #200, Silver Spring, MD, United States

[2]Johns Hopkins Berman Institute of Bioethics, Baltimore, MD, United States

[3]Applied Technology and Genomics Division, Defense Centers for Public Health, Dayton, OH, United States

[4]Department of Emerging Biological Threats, Defense Centers for Public Health, Aberdeen, MD, United States

[5]Department of Computational and Chemical Sciences, Carlow University, Pittsburgh, PA, United States

Corresponding Author:
Hunter Jackson Smith, MD, MPH, MBE
Global Emerging Infections Surveillance Branch, Armed Forces Health Surveillance Division, 11800 Tech Rd Suite #200, Silver Spring, MD, United States

Abstract

The US Department of Defense (DoD) is establishing its wastewater surveillance capacities to support national security objectives and promote the public health and medical readiness of US service members. Wastewater surveillance is an emerging technology that has traditionally been leveraged for detecting infectious diseases. However, its potential future applications could yield a vast and unpredictable amount of information that could be used for a wide variety of both health- and nonhealth-related purposes. The US military also serves an inimitable role for the country and its citizens, and exercises significant levels of control over its service members compared to civilian organizations. Further, its present and potential wastewater surveillance activities may reach far beyond just military installations. These factors raise unique ethical considerations that must be accounted for by leaders and policymakers to ensure the DoD implements a wastewater surveillance network in a manner that is both impactful in supporting public health and appropriate to the scope and population under surveillance. This paper explores important ethical features in conducting wastewater surveillance that are both specific to the DoD experience and applicable to wider public health initiatives.

(JMIR Public Health Surveill 2025;11:e67145) doi:10.2196/67145

KEYWORDS

wastewater; surveillance; ethics; military; Department of Defense

Introduction

Wastewater represents a promising source for obtaining valuable public health data to inform decision-making and response efforts [1]. Wastewater testing includes the tools and technologies used for various activities (eg, research, law enforcement, public health), while wastewater surveillance (WWS) is the activity of assessing biological, chemical, and analyte signatures for determining population-level trends to inform public health decision-making [2-5]. Traditionally, WWS has been leveraged to monitor for select infectious disease pathogens, such as poliovirus and more recently, SARS-CoV-2 during the COVID-19 pandemic [2-4,6]. However, wastewater testing has been used in recent years to detect numerous other targets, ranging from dietary patterns to the genetic makeup of communities [2,3,6-10]. Wastewater surveillance is an appealing public health instrument due to its scalability, cost-effectiveness compared to clinical monitoring, and the ability to detect surges in infectious disease cases, days before they become clinically relevant [11,12]. Recognizing the potential utility and value of WWS, the Centers for Disease Control and Prevention led the development of the National Wastewater Surveillance System. This system allows environmental scientists and public health departments across the country to track specific pathogens, including SARS-CoV-2, influenza A, respiratory syncytial virus, and mpox and report data into the National Wastewater Surveillance System to inform public health decision-making [13]. This system currently collects data from over 1000 sites across the United States, representing 35% of the US population, and it continues to add collaborating sites and other pathogens of public health concern to its surveillance list [13]. There are also plans by researchers and private corporations to further expand the detection capabilities of wastewater testing technologies to include illicit drug use, noncommunicable diseases, genetic information, and other targets. There are meaningful differences in the ethical considerations between wastewater testing and WWS, partly because wastewater testing has broader applications, such as research, law enforcement, and espionage [14,15]. However, this paper focuses on WWS, addressing wastewater testing only as it pertains to WWS as a public health activity.

XSL•FO
RenderX

Recently, the Department of Defense (DoD) has started prioritizing wastewater testing and surveillance, as highlighted in the emerging capabilities and technologies section of the 2023 Biodefense Posture Review [16]. It currently oversees multiple individual WWS projects within the United States and overseas, partners with other government agencies and private industries to explore and expand its wastewater testing technologies, and is building the capacity to develop and manage a DoD WWS network. It has already invested tens of millions of dollars toward advancing its WWS efforts. The DoD plays a distinctive role in ensuring social security and maintains a unique relationship with its ervice members (SM). A DoD WWS network has a potentially vast global scope, covering diverse civilian and military populations [17]. Given (1) the DoD's unique population and security considerations, (2) the rapid evolution and promising applications of wastewater testing technologies, and (3) the significant unknowns regarding the potential future uses of these technologies, there is an urgent need to address the ethical considerations inherent to implementing a WWS network for the DoD. These efforts are imperative to ensure it is executed appropriately and has suitable policies in place for management.

It is critical that environmental scientists, policymakers, and those engaged in WWS must be aware of the current and potential ethical issues related to wastewater testing and surveillance. Policy development, particularly for emerging technologies, must be focused on ethical implications and should aim to provide appropriate guidance and boundaries. This paper explores the ethical features of WWS and offers recommendations to address these concerns— that are important for any group planning to implement a WWS system. While many ethical considerations discussed in this paper are universal to any WWS system, their application within the DoD (whether in identity, scope, or importance) compared to other contexts is unique and deserves particular attention.

Targets of Wastewater Surveillance

The scientific evidence and methodology for wastewater sampling and testing must be established for each pathogen or target under consideration for surveillance. Pursuing large-scale sampling without a sufficient understanding of the target would be both ineffective and an unethical use of resources. For infectious diseases, this might involve pathogen-specific knowledge of (1) sampling and detection methods; (2) laboratory testing, genomic sequencing, and analysis; (3) relationship between detection levels in wastewater and clinical detection in humans; and (4) methods to mitigate the spread of the disease. A key component of conducting ethical public health activities includes ensuring the surveillance and responses to the surveillance findings are done in effective, scientifically sound, and evidence-based ways. The negative opportunity costs of inefficiently utilizing public health funds could harm other health areas where funds could have otherwise been allocated. Addressing this concern should include pilot testing of new pathogen targets or wastewater testing technologies as well as cost-effectiveness analyses prior to full integration into a WWS system [18]. For example, if a pilot project discovered that a new potential feature of the WWS system only marginally reduced infectious disease transmission at a significant cost, then those funds might be better directed to more effective public health activities.

Kass [19] describes an ethical framework for the assessment of public health programs and emphasizes that one of the key aspects of an ethically conducted program includes its effectiveness in achieving the stated goals. Relatedly, monitoring and program evaluation should be incorporated into the workflow of a public health program to ensure alignment with its goals [20]. For WWS, this could involve monitoring to determine whether WWS programs are effective at limiting infectious disease transmission among communities where they are being implemented, compared to those without such programs. Additionally, as WWS technologies and methodologies are developed and refined, efforts should be made to assess for concordance between WWS findings and clinical or on-ground surveillance data. Thus, an ethical WWS program must employ scientifically sound sampling and analysis methods and demonstrate effectiveness in achieving its goals, supported by clinical and epidemiologic evidence.

Currently, most WWS programs are focused on infectious diseases. However, there are other potential targets that have been explored beyond infectious pathogens (eg, noncommunicable diseases such as cardiovascular disease and cancer). While it is difficult to predict what technologies will become available in the future to assess WWS [21,22], many of the core ethical issues at present will remain relevant regardless of the future technological capability. The DoD leadership should be cognizant of preventing excessive mission creep (ie, broadening of the original objective) [23,24]. Using WWS for purposes beyond promoting and preserving health can be dangerous and cause ethically fraught scenarios. For example, while opioids and other substances of abuse (eg, methamphetamines, cocaine, nicotine) could be monitored via WWS, it raises equity concerns and risks of overpolicing [25]. Although the DoD already implements randomized drug screening for its SMs, WWS could introduce additional equity concerns. For instance, many higher-ranking personnel and officers typically live off-base, but the DoD would more accessibly conduct WWS in on-base areas where lower-ranking personnel primarily reside. This disparity might result in military leadership focusing punitive actions on lower-ranking enlisted personnel in an inequitable, unjustified manner, despite the possibility that higher-ranking individuals living off-base might use illicit substances in equal measures but remain undetected. This inequity would be particularly problematic considering that WWS captures all individuals using the sanitation system, including SMs, families, civilian employees, contractors, and visitors. In the future, WWS could also be used to detect medication-related metabolites (eg, reproductive health medications), human genetic content (eg, racial or ethnic composition of communities), and other potentially sensitive or politically contentious targets. It is important to thoughtfully consider what may become possible for WWS capabilities and to ensure that DoD leaders and policymakers remain vigilant in keeping its focus on appropriate health targets and public health–promoting outcomes.

XSL•FO
RenderX

Populations and Areas of Wastewater Surveillance

Where should the DoD focus its wastewater surveillance efforts? This question is practically, legally, and ethically important because there are few areas where WWS could be conducted with only DoD SMs present. Understanding the size and vulnerability of the populations targeted by WWS can influence the interpretation of and responses to the data produced by it [26]. Thus, exploring where the DoD should conduct WWS naturally leads to the question: on whom should the DoD focus its WWS efforts?

On-base DoD locations and DoD transport vessels are obvious areas to focus WWS efforts. Four primary location types can be considered for on-base WWS: (1) facilities, (2) clinics and hospitals, (3) recruit and trainee sites, and (4) housing. On-base DoD locations refer to areas both within the United States and abroad, including permanent international installations and temporary bases established during conflicts. On-base facilities (both medical and nonmedical) and recruit and trainee sites are the least ethically problematic locations, as these areas typically constitute force health protection–relevant activities and have the highest concentration of SMs. While DoD civilian and contractor personnel also work in these areas, conducting WWS at these locations is fully within the DoD's responsibility and legal jurisdiction to support and promote the health of DoD SMs as well as civilian and contractor personnel. Military transport vessels (eg, large aircraft and naval ships) may prove to be particularly important catchment points for sentinel surveillance, especially for tracking the transmission of infectious diseases [27-29].

On-base housing is another force health protection–relevant area, as these are locations where SMs and those closest to them spend most of their off-duty time. An SM's family often lives on-base with them, and the family's health can directly impact the health of SMs and their units. Further, infectious diseases present among SM families are also present among SMs; therefore, conducting surveillance of beneficiaries can better inform the epidemiologic picture of infectious disease prevalence and transmission within military units. Thus, there is a high probability of detecting pathogens of DoD significance in these areas. Moreover, the DoD has a precedent of assuming some fiduciary responsibility for promoting the health and well-being of SM families, through a variety of clinical and social support systems (eg, Tricare health care coverage, paternity/maternity leave policy, education options, free counseling services, housing support, wellness centers). It is also impossible to extricate on-base housing WWS sampling of SMs from family members who live with them. Therefore, it is within the DoD's purview to promote SM families' health, and if the WWS is conducted in a minimally invasive, deidentified, and ethically sound manner, WWS could be a valuable way to further support these health objectives.

The DoD also has a vested interest in WWS abroad; these international locations can be grouped into three categories: (1) joint military exercises, (2) ally military installations, and (3) foreign areas of mutual interest with the host nation. There should be two key underlying requirements for all international locations (beyond DoD installations located abroad): (1) the DoD must obtain the permission of the host nation in which they are guests to conduct WWS and (2) the WWS should provide some benefit to that host nation. Ensuring these criteria are fulfilled will promote an equitable relationship between nations and help avoid exploitative or colonialist interpretations of mutually beneficial public health activities. Another critical consideration for the DoD in conducting international WWS is that the host nation should be empowered to exercise as much sovereign control over the public health activities occurring in its own territory as it sees fit. The DoD and the host nation may both have interests in conducting WWS in certain locations in that country; however, the host nation should decide if and how WWS should proceed. The DoD's role should be supportive, which can manifest in a variety of ways depending on the context and needs of the host nation, including leading the WWS project, funding host nation researchers, providing methodologic input and subject matter expertise, or developing technical and analytical capabilities. Challenges may arise if an ally nation requests WWS support that the DoD is not positioned to support, either on principle or due to resource constraints. International scientific collaborations require agreement from all involved parties. If the DoD does not want to engage, then they are not required to do so, particularly if the DoD has ethical reservations regarding an expansion of the ally nation's surveillance activities to include controversial targets or populations. If the DoD is supportive of the request but is resource constrained, then it can offer support in other ways (eg, technical consultation, references to other potential collaborators). If the foreign nation's request involves surveillance targets that are of importance to them but not of interest to the DoD, then the DoD can consider offering low-cost support (eg, information, consultation) in the spirit of preserving and strengthening the relationship. Generally, approaching host nations from the perspective of support, providing the host nation opportunities to lead decision-making, and emphasizing the mutually beneficial aspects of the collaboration will support an equitable partnership.

Joint military exercises serve as collaborative training opportunities for two or more militaries to build cohesiveness, promote mission readiness, and grow camaraderie among allies. In such scenarios, it is acceptable for either the DoD or other militaries to lead WWS efforts, provided all the participating militaries whose wastewater is sampled endorse their permission for WWS to be conducted during the exercise, and for the data to be shared with the sampled parties that participated, if requested. This practice supports the participating military's autonomy by providing an opportunity for group informed consent, whereby a nation's military can provide permission or denial for participation of its SMs. For WWS at ally military health care facilities and installations, it is appropriate for the ally nation's military to lead, with the DoD providing support these efforts as defined by what is agreed upon between the two militaries. The flow of data should be bidirectional and beneficial for both parties for the endeavor to be equitable. For international civilian settings, such as health care facilities, the ally nation's relevant health organizations (eg, ministries of health, health research institutes) should serve as key partners,

XSL•FO
RenderX

and the DoD should support and coordinate with them in the conduct of the surveillance and sharing of data. The host nation's organization should make efforts to ensure awareness of WWS efforts within the communities involved, especially in specific locations such as health care facilities, and that the facility's leadership is aware, and approves of the WWS.

Management of Wastewater Surveillance Data

Once the WWS has been conducted, how should data be managed? The DoD faces important ethical considerations regarding the type of data to be collected and its management. A balance must be struck between allowing beneficial data access and ensuring data protection. WWS can generate a variety of data, including molecular and biological signatures, health and disease markers, biochemical indicators of illicit activities, and associated identifiable demographic metadata [17]. While possessing large amounts of data is typically beneficial for determining appropriate responses, course of actions, and preventive planning, protecting these data is crucial, especially as the DoD arguably experiences more frequent and powerful attacks on its information security infrastructure compared to other organizations. Therefore, managing the focus on which data to protect (eg, personally identifiable information vs aggregate data) and how to protect it (eg, cloud or servers vs hardline firewalls) will be crucial for ensuring minimized risk of data weaponization or leakage. Wastewater surveillance can produce highly granular data that can be used by public health researchers and officials for improving the quality of the communities they serve; however, this level of detail is not always essential when communicating with end users. Establishment of well-defined data-sharing agreements which incorporate ethical guidelines and treat WWS data with the same scrutiny as Controlled Unclassified Information or Classified determinations, wherein raw information sharing is conducted only as needed, will aid in reducing improper use of WWS information. Data security can be handled in a similar manner by escalating access and storage restrictions corresponding to data aggregation and specificity. For instance, aggregated, community-level detection data of public health targets could be stored on shared drives with minimal restrictions; however, individual molecular and genetic data could be stored behind hardline firewalls on dedicated IT systems. Anonymization and the use of beacons are additional potential methods that could help in ensuring individual human subjects cannot be identified by WWS; however, both methods have their inherent limitations [30]. Nonetheless, it is likely that data security approaches will not lend to one size fits all solutions and will be dictated by the surveillance effort, parties involved, and required outcomes. The DoD must take strong measures to ensure only authorized personnel access appropriate levels of WWS data.

As WWS data becomes more granular and specific, the ethical responsibility to ensure data security and appropriate use and access increases proportionally. This escalation of responsibility in proportion to the level of granularity of the data is derived from the increasing probability of it being personalized and potentially damaging to individuals or small groups. For example, depending on where the wastewater is sampled, it could be possible to pinpoint illicit drug use or the location of an individual to a specific household. In fact, analyzing local sewage for genetic markers was one of the methods potentially used by the United States to confirm Osama bin Laden's location prior to him being killed in 2011 [15]. It may also become possible in the future to obtain personal genetic information through WWS and to associate it with personally identifiable information; the DoD must be prepared to either dispose or strongly protect this data. There are significant risks if foreign adversaries gain access to this level of data, as well as real concerns regarding how this information could be misused domestically in politically motivated ways. One way to combat these concerns is to establish clear sampling strategies with requirements to destroy unused sample portions once the appropriate predetermined targets have been tested or detected.

There are also important considerations regarding the sharing of WWS data when it has broader implications beyond the DoD, and is relevant to other public health entities, such as US government groups (eg, local health departments, the Centers for Disease Control and Prevention), ally nations, or nongovernmental organizations (eg, the Bill and Melinda Gates Foundation). There are certain ethical responsibilities in the sharing of appropriate levels of data when it can also be of benefit to others, and the DoD should take steps to collaborate with these groups to promote health security. Thus, taking all these considerations into account, one overarching goal for a DoD WWS network is to ensure the data is maintained at an aggregate level as much as possible for monitoring of community and population-level trends rather than specific information on individuals or small groups.

Utilization and Response to Wastewater Surveillance Data

Public health surveillance data should not be collected unless there is a potential for a meaningful and actionable response. If surveillance findings are not translated to public health action or policy changes, the benefits accrued from the surveillance efforts are minimal [31]. Collecting data only for the sake of collection is an inappropriate use of resources; therefore, there should be focused and organized plans in place for using the WWS systems and information derived from them prior to their implementation. Additionally, there should be rigorous discussions about how military leaders and public health officials respond to such information as it becomes available to them. There are innumerable ways in which the leadership could respond to WWS data when it is received in near real-time. While this information offers an opportunity to ethically use that data, there is a risk of inappropriate response to it [32]. Notably, military commanders make the final decisions regarding public health actions or interventions, particularly those of consequence. It is likely that a medical or public health professional will not be making the decision, and the military commander may have no background in science, medicine, or public health. While making the intervention decisions in response to WWS data, these commanders and

XSL•FO
RenderX

their leadership teams should lean heavily upon the expert advice of medical and public health professionals to inform their decisions to help ensure they take the most appropriate actions. The key consideration here is that the responses—whether determined by military command or leadership from preventive medicine, public health, or clinical groups—should be appropriate, proportional, and context- and community-specific based on the information received through the WWS system.

General Ethical Considerations

Finally, there are unique ethical considerations for the DoD that civilian organizations do not necessarily face to the same extent. The military has substantial authority, control, and access over the lives of its SMs. It can dictate aspects ranging from what to wear, where to live, how to behave, and nearly every aspect of SMs' lives. Given the extent of this control and its durability and depth, the DoD has an elevated ethical responsibility toward the health and well-being of its SMs [33]. Further, there is a public health ethics principle, which argues that a public health program should use the least invasive or minimally burdensome (to the population being targeted by the program) methods to accomplish its goals [16,34]. Service members already forgo significant autonomy afforded to civilians; therefore, any public health program that might additionally strip away their autonomy must be seriously deliberated before implementation. Although WWS may not directly interfere with the lives of SMs while the data is collected, the actions related to WWS (eg, how leadership responds to the information, what data is collected, and how the data is managed) could infringe upon their privacy and autonomy. However, if conducted appropriately, the potential health benefits for the SMs would likely outweigh minor privacy or autonomy concerns. Additionally, the implications extend beyond military populations when considering the health of SMs. The mission of the DoD and SMs is to protect US citizens, interests, and national security. If SMs are unwell, experience disease outbreaks, or are targeted using bioweapons, it places the country at risk, as they will be less able to support their security objectives by failing to achieve optimal operational capabilities and mission readiness. Furthermore, SMs and military installations do not exist in isolation. If an infectious disease emerges among a military population in a certain area, it is likely there is a similar problem in the surrounding civilian community, or at least presents a risk to them. Therefore, these considerations add to the elevated responsibility of the DoD to support and promote the health of its SMs.

There are also several broad ethical considerations for how the DoD conducts WWS. There is a tension underlying many public health actions, between preserving personal freedom and autonomy of individuals and maximizing the public health benefits for the group. This tension is manifested in several ways, including the amount and types of data to be collected, proportionality and appropriateness of response to the data, level of invasiveness of the surveillance, populations, and areas selected for surveillance [35,36]. DoD leadership should balance protecting the freedom and autonomy of its SMs while also prioritizing the health of the force from a population perspective [33]. Additional resources providing ethical guidance for the operation of public health surveillance systems that are applicable for WWS are available from the American Public Health Association's Public Health Code of Ethics and the World Health Organization's Guidelines on Ethical Issues in Public Health Surveillance [34,37].

Justice is another key ethical concept, and when appraising a DoD WWS program, two relevant perspectives must be considered: (1) distributive justice and (2) procedural justice. Distributive justice is an ethical principle concerned with the fair distribution of benefits and burdens. In the DoD WWS context, it involves ensuring that the distribution of resources, outcomes, and interventions based on the surveillance findings is done in an equitable way (eg, distributing vaccines based on WWS findings of acute elevations in vaccine-preventable diseases). It also includes the fair distribution of resources for conducting WWS throughout the United States and international installations. Procedural justice is an ethical principle concerned with fair processes and methods. It includes features such as inclusion, transparency, neutrality, respect, and trust [38]. For DoD WWS, it might involve transparent communications and decision-making, community considerations, stakeholder participation, access to relevant data for the surveilled communities, among others.

There are a variety of other general ethical considerations that are applicable to WWS, which are briefly mentioned here and should be explored further. First, there are differences between conducting public health surveillance versus research and the varying regulatory, oversight, and ethical requirements. Second, there may be differences in what actions are permissible when considering baseline surveillance versus those needed during public health emergencies. Finally, there are inevitable future possibilities beyond our current capabilities or knowledge regarding what might become possible through conducting WWS. What might WWS be capable of in the future, and how do we anticipate and plan for that? Although there are myriad ethical considerations for the appropriate conduct of WWS by the DoD both currently and in the future, it is undoubtedly a burgeoning area of public health that the DoD should pursue and be engaged in a substantial, effective, and meaningful manner, as there are innumerable potential benefits to a DoD WWS network.

Conclusions

Wastewater surveillance offers tremendous promise and opportunities to advance public health by enhancing disease surveillance, informing public health and clinical decision-making, and improving public health response times to mitigate negative health impacts to populations [39,40]. The DoD should adopt WWS and position itself to take advantage of the current and potential future benefits. It should also recognize the practical and ethical differences between using wastewater testing for public health surveillance, research, or other potential military purposes. With over 1.3 million active-duty SMs, 1.1 million National Guard and Reserve members, and 700,000 civilian personnel, the DoD is the largest organization in the United States, and has responsibilities to their health and readiness [41]. Further, the Military Health

XSL•FO
RenderX

System serves a total of approximately 9.5 million beneficiaries, which includes both SMs and their families [42]. WWS has the potential to benefit not only these populations, but others in the United States and overseas as well. However, these opportunities also present a potential for misuse, which could be detrimental to those being surveilled, raising important ethical questions about how and why WWS should be conducted. The DoD also has elevated ethical obligations to the well-being of its SMs due to the breadth and depth of control it exercises over them. Although ethical issues may arise related to WWS, these issues can and should be addressed through proper deliberation and policy implementation, as the potential public health benefits could be substantial. The DoD has responsibilities to protecting the national security of its citizens broadly, which implies a crucial role in infectious disease surveillance for outbreaks as well as bioengineered and emerging pathogen threats. Therefore, the DoD should pursue wastewater surveillance in an ethically appropriate, scientifically robust, and thoughtfully planned and executed manner to guide public health decision-making, protect the well-being of its SMs, and support US government health security objectives.

Acknowledgments

The authors would like to express their gratitude to Col Brett Nishikawa, CDR Matthew Kasper, Lt. Col Mark Lehman, Dr. Lindsay Morton, Dr. Seth Faith, Dr. Anthony Fries, Dr. Blake Stamps, Ms. Genny Cook, Dr. Liza Dawson, MAJ Jennifer Jackiw, MAJ Matthew Angel, Ms. Heidi Taylor, Dr. Subrahmanyam Yerramilli, Dr. Georgia Millward, and 1LT Daniel Evans for their thoughtful reviews and feedback on the paper.

Disclaimer

The Defense Health Agency has reviewed this material, and there is no objection to its publication. The contents, views or opinions expressed in this publication are those of the authors and do not necessarily reflect official policy or position of the Armed Forces Health Surveillance Division, the Defense Centers for Public Health-Aberdeen, the Defense Centers for Public Health-Dayton, the Defense Health Agency, the Department of Defense, or Departments of the Army, Navy, or Air Force. Mention of trade names, commercial products, or organizations does not imply endorsement by the U.S. Government.

Conflicts of Interest

None declared.

References

1. Diamond MB, Keshaviah A, Bento AI, et al. Wastewater surveillance of pathogens can inform public health responses. Nat Med 2022 Oct;28(10):1992-1995. [doi: 10.1038/s41591-022-01940-x]
2. Huizer M, Ter Laak TL, de Voogt P, van Wezel AP. Wastewater-based epidemiology for illicit drugs: a critical review on global data. Water Res 2021 Dec 1;207:117789. [doi: 10.1016/j.watres.2021.117789] [Medline: 34731667]
3. Kilaru P, Hill D, Anderson K, et al. Wastewater surveillance for infectious disease: a systematic review. Am J Epidemiol 2023 Feb 1;192(2):305-322. [doi: 10.1093/aje/kwac175] [Medline: 36227259]
4. Luo J, Bello D, Pagsuyoin S. Long-term wastewater-based surveillance and impacts of the COVID-19 pandemic on drug use trends in a U.S. northeast rural town. Science of The Total Environment 2023 Jun;877:162806. [doi: 10.1016/j.scitotenv.2023.162806]
5. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division; Division on Earth and Life Studies, Board on Population Health and Public Health Practice, Water Science and Technology Board, Committee on Community Wastewater-based Infectious Disease Surveillance. Wastewater-Based Disease Surveillance for Public Health Action: Washington (DC): National Academies Press (US); 2023. [doi: 10.17226/26767] [Medline: 37184191]
6. Bowes DA, Driver EM, Savic S, et al. Integrated multiomic wastewater-based epidemiology can elucidate population-level dietary behaviour and inform public health nutrition assessments. Nat Food 2023 Mar;4(3):257-266. [doi: 10.1038/s43016-023-00717-w] [Medline: 37118274]
7. Clark M, Severn M. Wastewater surveillance for communicable diseases: emerging health technologies. In: Canadian Journal of Health Technologies 2, 2023, Vol. 3. [doi: 10.51731/cjht.2023.575] [Medline: 38502748]
8. Choi PM, Tscharke B, Samanipour S, et al. Social, demographic, and economic correlates of food and chemical consumption measured by wastewater-based epidemiology. Proc Natl Acad Sci U S A 2019 Oct 22;116(43):21864-21873. [doi: 10.1073/pnas.1910242116] [Medline: 31591193]
9. Evans ED, Dai C, Isazadeh S, Park S, Ratti C, Alm EJ. Longitudinal wastewater sampling in buildings reveals temporal dynamics of metabolites. PLoS Comput Biol 2020 Jun;16(6):e1008001. [doi: 10.1371/journal.pcbi.1008001] [Medline: 32598361]
10. Yang Z, Xu G, Reboud J, Kasprzyk-Hordern B, Cooper JM. Monitoring genetic population biomarkers for wastewater-based epidemiology. Anal Chem 2017 Sep 19;89(18):9941-9945. [doi: 10.1021/acs.analchem.7b02257] [Medline: 28814081]

XSL•FO
RenderX

11. Farkas K, Hillary LS, Malham SK, McDonald JE, Jones DL. Wastewater and public health: the potential of wastewater surveillance for monitoring COVID-19. Curr Opin Environ Sci Health 2020 Oct;17:14-20. [doi: 10.1016/j.coesh.2020.06.001] [Medline: 32835157]
12. Hart OE, Halden RU. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: feasibility, economy, opportunities and challenges. Sci Total Environ 2020 Aug 15;730:138875. [doi: 10.1016/j.scitotenv.2020.138875] [Medline: 32371231]
13. National wastewater surveillance system (NWSS). Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/nwss/wastewater-surveillance.html [accessed 2024-01-17]
14. National wastewater drug monitoring program reports. Australian Criminal Intelligence Commission. 2024. URL: https://www.acic.gov.au/publications/national-wastewater-drug-monitoring-program-reports [accessed 2024-10-15]
15. Allison G. How it went down. TIME. 2012. URL: https://time.com/archive/6641671/how-it-went-down [accessed 2024-10-15]
16. 2023 Biodefense Posture Review. US Department of Defense. 2023. URL: https://media.defense.gov/2023/Aug/17/2003282337/-1/1/1/2023_BIODEFENSE_POSTURE_REVIEW.PDF [accessed 2024-05-20]
17. Goldberg Z, Linder AG, Miller LN, Sorrell EM. Wastewater collection and sequencing as a proactive approach to utilizing threat agnostic biological defense. Health Secur 2024;22(1):11-15. [doi: 10.1089/hs.2023.0075] [Medline: 37856169]
18. Sanjak JS, McAuley EM, Raybern J, et al. Wastewater surveillance pilot at US military installations: cost model analysis. JMIR Public Health Surveill 2024 Sep 6;10:e54750. [doi: 10.2196/54750] [Medline: 39240545]
19. Kass NE. An ethics framework for public health. Am J Public Health 2001 Nov;91(11):1776-1782. [doi: 10.2105/ajph.91.11.1776] [Medline: 11684600]
20. German RR, Lee LM, Horan JM, et al. Updated guidelines for evaluating public health surveillance systems: recommendations from the Guidelines Working Group. MMWR Recomm Rep 2001 Jul 27;50(RR-13):1-35. [Medline: 18634202]
21. Amin V, Bowes DA, Halden RU. Systematic scoping review evaluating the potential of wastewater-based epidemiology for monitoring cardiovascular disease and cancer. Sci Total Environ 2023 Feb 1;858(Pt 3):160103. [doi: 10.1016/j.scitotenv.2022.160103] [Medline: 36370774]
22. Ladyzhets B. What toilets can reveal about COVID, cancer and other health threats. Nature New Biol 2024 Apr;628(8008):492-494. [doi: 10.1038/d41586-024-01092-7] [Medline: 38632479]
23. Doorn N. Wastewater research and surveillance: an ethical exploration. Environ Sci (Camb) 2022 Oct 27;8(11):2431-2438. [doi: 10.1039/d2ew00127f] [Medline: 36353217]
24. Rinde M, Science History Institute. The murky ethics of wastewater surveillance. Distillations Magazine. 2023 Feb. URL: https://sciencehistory.org/stories/magazine/the-murky-ethics-of-wastewater-surveillance [accessed 2024-04-23]
25. Hrudey SE, Silva DS, Shelley J, et al. Ethics guidance for environmental scientists engaged in surveillance of wastewater for SARS-CoV-2. Environ Sci Technol 2021 Jul 6;55(13):8484-8491. [doi: 10.1021/acs.est.1c00308] [Medline: 34101444]
26. Coffman MM, Guest JS, Wolfe MK, et al. Preventing scientific and ethical misuse of wastewater surveillance data. Environ Sci Technol 2021 Sep 7;55(17):11473-11475. [doi: 10.1021/acs.est.1c04325] [Medline: 34431294]
27. Ahmed W, Bertsch PM, Angel N, et al. Detection of SARS-CoV-2 RNA in commercial passenger aircraft and cruise ship wastewater: a surveillance tool for assessing the presence of COVID-19 infected travellers. J Travel Med 2020 Aug 20;27(5):116. [doi: 10.1093/jtm/taaa116] [Medline: 32662867]
28. Traveler-based genomic surveillance for early detection of new SARS-cov-2 variants. Centers for Disease Control and Prevention. 2023. URL: https://wwwnc.cdc.gov/travel/page/travel-genomic-surveillance [accessed 2024-03-05]
29. Li J, Hosegood I, Powell D, et al. A global aircraft-based wastewater genomic surveillance network for early warning of future pandemics. Lancet Glob Health 2023 May;11(5):e791-e795. [doi: 10.1016/S2214-109X(23)00129-8] [Medline: 37061316]
30. Jacobs D, McDaniel T, Varsani A, Halden RU, Forrest S, Lee H. Wastewater monitoring raises privacy and ethical considerations. IEEE Trans Technol Soc 2021;2(3):116-121. [doi: 10.1109/TTS.2021.3073886]
31. Lee LM, Thacker SB. Public health surveillance and knowing about health in the context of growing sources of health data. Am J Prev Med 2011 Dec;41(6):636-640. [doi: 10.1016/j.amepre.2011.08.015] [Medline: 22099242]
32. Gable L, Ram N, Ram JL. Legal and ethical implications of wastewater monitoring of SARS-CoV-2 for COVID-19 surveillance. J Law Biosci 2020;7(1):lsaa039. [doi: 10.1093/jlb/lsaa039] [Medline: 32793373]
33. Smith HJ. Ethical responsibilities of a military to the social determinants of health of its service members. Mil Med 2022 Aug 25;187(9-10):252-256. [doi: 10.1093/milmed/usab554] [Medline: 35389484]
34. Public health code of ethics. American Public Health Association. 2019. URL: https://www.apha.org/-/media/files/pdf/membergroups/ethics/code_of_ethics.ashx [accessed 2024-05-30]
35. Marckmann G, Schmidt H, Sofaer N, Strech D. Putting public health ethics into practice: a systematic framework. Front Public Health 2015;3:23. [doi: 10.3389/fpubh.2015.00023] [Medline: 25705615]
36. Schröder-Bäck P, Duncan P, Sherlaw W, Brall C, Czabanowska K. Teaching seven principles for public health ethics: towards a curriculum for a short course on ethics in public health programmes. BMC Med Ethics 2014 Oct 7;15:73. [doi: 10.1186/1472-6939-15-73] [Medline: 25288039]
37. WHO guidelines on ethical issues in public health surveillance. World Health Organization. 2017. URL: https://www.who.int/ethics/publications/public-health-surveillance/en [accessed 2024-05-30]

XSL•FO
RenderX

38. Bowes DA, Darling A, Driver EM, et al. Structured ethical review for wastewater-based testing in support of public health. Environ Sci Technol 2023 Sep 5;57(35):12969-12980. [doi: 10.1021/acs.est.3c04529] [Medline: 37611169]
39. Keshaviah A, Diamond MB, Wade MJ, Scarpino SV, Global Wastewater Action Group. Wastewater monitoring can anchor global disease surveillance systems. Lancet Glob Health 2023 Jun;11(6):e976-e981. [doi: 10.1016/S2214-109X(23)00170-5] [Medline: 37202030]
40. Levy JI, Andersen KG, Knight R, Karthikeyan S. Wastewater surveillance for public health. Science 2023 Jan 6;379(6627):26-27. [doi: 10.1126/science.ade2503] [Medline: 36603089]
41. The executive branch. White House Archives. URL: https://obamawhitehouse.archives.gov/1600/executive-branch [accessed 2024-02-15]
42. Patients by beneficiary category. Health.mil: The official website of the Military Health System. 2023. URL: https://www.health.mil/Military-Health-Topics/MHS-Toolkits/Media-Resources/Media-Center/Patient-Population-Statistics/Patients-by-Beneficiary-Category [accessed 2024-05-30]

Abbreviations

DoD: Department of Defense
SM: Service member
WWS: wastewater surveillance

Edited by A Mavragani; submitted 13.11.24; peer-reviewed by C Mansfeldt, M Gutierrez, W Hall; revised version received 13.12.24; accepted 15.12.24; published 06.02.25.

Please cite as:
Smith HJ, Agans RT, Kowallis WJ
Ethical Considerations for Wastewater Surveillance Conducted by the US Department of Defense
JMIR Public Health Surveill 2025;11:e67145
URL: https://publichealth.jmir.org/2025/1/e67145
doi:10.2196/67145

© Hunter Jackson Smith, Richard T Agans, William J Kowallis. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 6.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Call for Action: Lessons Learned From a Pilot to Share a Complex, Linked COVID-19 Cohort Dataset for Open Science

Clara Amid[1*], PhD; Martine Y van Roode[1*], MSc; Gabriele Rinck[2], PhD; Janko van Beek[1], PhD; Rory D de Vries[1], PhD; Gijsbert P van Nierop[1], PhD; Eric C M van Gorp[1], Prof Dr; Frank Tobian[3], MSc; Bas B Oude Munnink[1], PhD; Reina S Sikkema[1], PhD; Thomas Jaenisch[3,4], MD, PhD; Guy Cochrane[2], PhD; Marion P G Koopmans[1,5], DVM, Prof Dr

[1]Department of Viroscience, Erasmus Medical Center (Erasmus MC), Dr Molewaterplein 40, Rotterdam, Netherlands

[2]European Molecular Biology Laboratory (EMBL), European Bioinformatics Institute (EBI), Wellcome Genome Campus, Hinxton, Cambridgeshire, United Kingdom

[3]Heidelberg Institute of Global Health (HIGH), Heidelberg University Hospital (UKHD), Heidelberg, Germany

[4]Center for Global Health, Colorado School of Public Health, Aurora, CO, United States

[5]Pandemic and Disaster Preparedness Center (PDPC), Rotterdam, Netherlands

[*]these authors contributed equally

Corresponding Author:
Martine Y van Roode, MSc
Department of Viroscience, Erasmus Medical Center (Erasmus MC), Dr Molewaterplein 40, Rotterdam, Netherlands

Abstract

The COVID-19 pandemic proved how sharing of genomic sequences in a timely manner, as well as early detection and surveillance of variants and characterization of their clinical impacts, helped to inform public health responses. However, the area of (re)emerging infectious diseases and our global connectivity require interdisciplinary collaborations to happen at local, national and international levels and connecting data to understand the linkages between all factors involved. Here, we describe experiences and lessons learned from a COVID-19 pilot study aimed at developing a model for storage and sharing linked laboratory data and clinical-epidemiological data using European open science infrastructure. We provide insights into the barriers and complexities of internationally sharing linked, complex cohort datasets from opportunistic studies for connected data analyses. An analytical timeline of events, describing key actions and delays in the execution of the pilot, and a critical path, defining steps in the process of internationally sharing a linked cohort dataset are included. The pilot showed how building on existing infrastructure that had previously been developed within the European Nucleotide Archive at the European Molecular Biology Laboratory-European Bioinformatics Institute for pathogen genomics data sharing, allowed the rapid development of connected “data hubs.” These data hubs were required to link human clinical-epidemiological data under controlled access with open high dimensional laboratory data, under FAIR (Findable, Accessible, Interoperable, Reusable) principles. Based on our own experiences, we call for action and make recommendations to support and to improve data sharing for outbreak preparedness and response.

(JMIR Public Health Surveill 2025;11:e63996) doi:10.2196/63996

KEYWORDS

data sharing; data management; open science; COVID-19; emerging infectious disease; global health

Introduction

The COVID-19 pandemic proved in a real-case scenario how rapid FAIR (Findable, Accessible, Interoperable, Reusable) [1] sequence data sharing was used to inform public health response policies [2]. An international collaboration, the Reconciliation of Cohort data in Infectious Diseases (ReCoDID) [3], a 4-year project that began in January 2019, was funded by the European Commission’s Horizon 2020 programme with a focus on the global response to emerging pathogens. Building on existing infrastructures and partnerships, the consortium was funded to develop a sustainable model for the storage, curation, and analyses of complex datasets collected from infectious disease related cohorts to facilitate and speed up outbreak research. In 2020, a pilot study was developed within this project using COVID-19 data from Erasmus Medical Center (Erasmus MC) as a use-case. The incentive was that in an ongoing global COVID-19 pandemic, linking data for connected data analysis with the potential to better understand the disease would be of public interest and of public health importance. The aim of the pilot was to test the potential for sharing linked laboratory data and clinical-epidemiological (CE) data in the evolving European infrastructure for open science in a centralized model. The centralized model, as opposed to a decentralized or federated data sharing model, was chosen for this pilot because we wanted to build on existing infrastructure [4,5]. The model has the

XSL•FO
RenderX

potential of hosting and curation of large datasets, the benefit of standardized data input and output leading to interoperability, easing re-analysis by others necessary for improved prediction and preparedness on global, EU and national level, as was needed for a newly emerged outbreak. A decentralized or federated data sharing model has the benefit of not needing to move the data, but the consequences often manifest themselves in no sharing or limited sharing and linking of datasets, difficulty to apply federated data analyses across a wide range of data formats and sources, with predictions and preparedness happening more at local or regional level. Given the diversity of datasets and the complexity of sharing data from clinical research and in line with the rapidly evolving situation during the COVID-19 pandemic, we wanted to learn what it would take to release linked data from opportunistic studies, rather than pre-approved cohort data.

Here, we describe experiences and lessons learned from this pilot, providing insights into the complexities of internationally sharing linked, cohort datasets for connected data analyses.

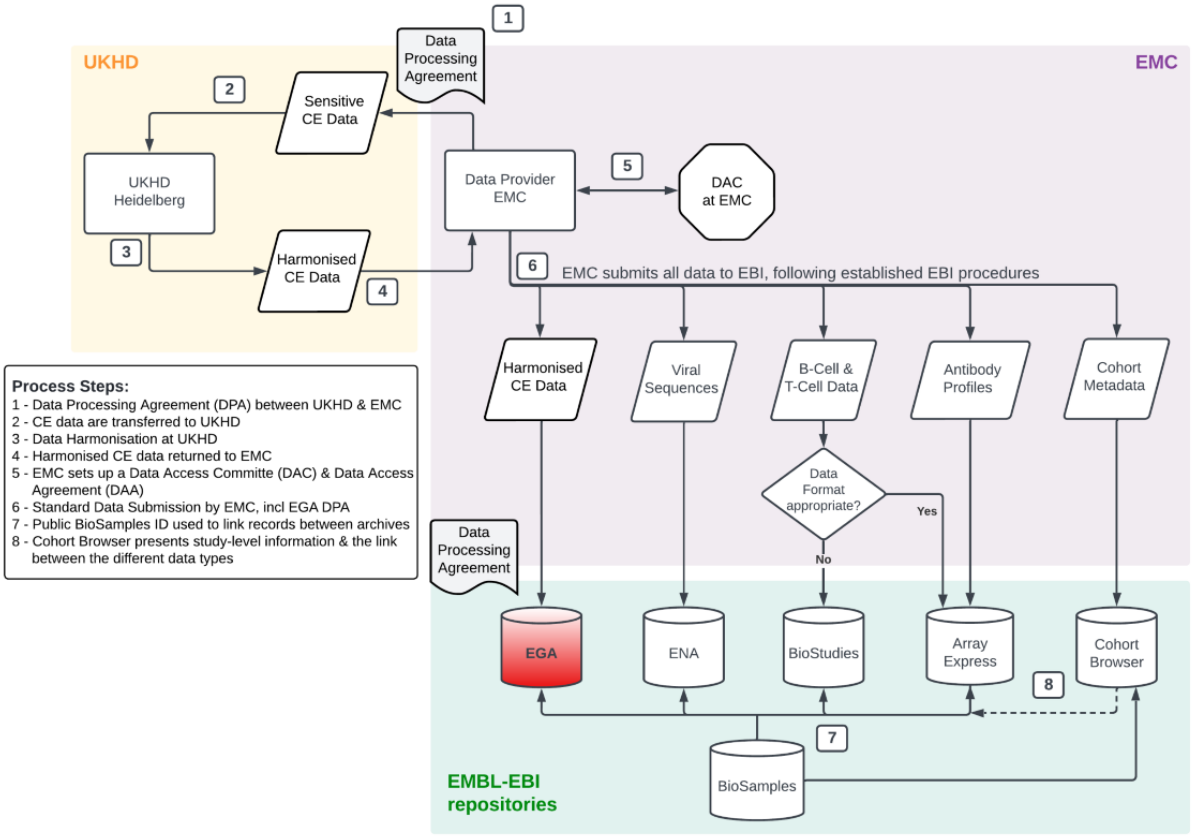
Pilot Set Up

The following sections have references to the timeline (Multimedia Appendix 1) in which the steps taken to set up and complete a pilot with a small but complex, linked dataset from opportunistic COVID-19 clinical and laboratory research studies as done in (academic) hospital settings are shown. We divided these steps into the following categories: data identification and acquisition, ethics and legal requirements (data protection), data harmonization, and data storage and delivery. Starting in November 2020, meetings took place with basic researchers and clinical scientists who had generated COVID-19 data as part of pandemic related research [6,7]. The drafted data map illustrated how complex the seemingly “simple” datasets (clinical, genomic, pathogen genomic, and high-dimensional laboratory) were. Different data sharing rules and governance for different parts of the datasets existed, and different identifiers and platforms for data capture and storage had been used. It became apparent that these procedures were not unique to COVID-19 pandemic and were common practice at the hospital. Finally, after a further meeting with the involved scientists to discuss datasets that could be linked in a pilot, a small cohort of 151 patients (intensive care unit and non–intensive care unit) who had tested SARS-CoV-2 positive (reverse transcription polymerase chain reaction) between May 2020 and May 2021 was selected for inclusion in the pilot, for whom several data types were partially available, including SARS-CoV-2 viral whole genomes, protein-microarray serological readouts, T- and B-cell data as well as CE data (data identification and acquisition).

Due to the extensive legal discussions around data sharing (ethics and legal requirements), the implementation of the long-term model of the ReCoDID data workflow was brought to a standstill (see “Enablers and Barriers Affecting the Execution of the Pilot”). Therefore, a slightly deviated dataflow model for the pilot was agreed on. The deviation was solely focused on CE data and meant that instead of Heidelberg University Hospital (UKHD) transferring harmonized CE data directly to the European Genome-Phenome Archive (European Molecular Biology Laboratory-European Bioinformatics Institute [EMBL-EBI]), the harmonized CE data were transferred back from UKHD to Erasmus MC, and Erasmus MC carried out the CE data submission to the European Genome-Phenome Archive, including a further data processing agreement (Figure 1). This change was necessary to circumvent one of the barriers (”Adherence to GDPR, and differences in its interpretation” in Table 1) and to move forward with the pilot and to facilitate the development of the connected datasets platform. Furthermore, the Erasmus MC’s legal and privacy offices and the institute’s COVID-19 clinical database team (established in June 2020) were contacted to discuss the aim of the pilot, the access procedures to the CE data, as well as all other steps and legal requirements for sharing the CE data internationally. For this purpose, a Data Protection Impact Assessment was carried out. A data access request for CE data though the Erasmus MC’s COVID-19 clinical database was also required (ethics and legal requirements).

XSL•FO
RenderX

Figure 1. Overview of Erasmus Medical Center (EMC)'s pilot dataflow, linkage and presentation for EMC's connected dataset pilot. The roles of the participating institutes within this pilot were as follows. EMC: mobilization and consolidation of the datasets (data identification and acquisition) by the EMC team at the source within EMC, participant-level and derived BioSample record registration to include metadata about biological samples, linking and submission of all data to repositories within the European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) infrastructure including deposition of clinical-epidemiological (CE) data under controlled access at European Genome-Phenome (EGA), depositions of viral sequences at the European Nucleotide Archive (ENA), B-cell and T-cell data, as well as antibody profiles at BioStudies and ArrayExpress repositories, respectively, and setting up a data access committee (DAC) and data access agreement (DAA) (data storage and delivery); University Hospital Heidelberg (UKHD): harmonization of the CE data and data dictionary at UKHD according to international standards (data harmonization); EMBL-EBI: provided the process for data linking on a participant level, including linking of the CE data to omics data types, using the hierarchical BioSamples database model and presentation of EMC's cohort study in the Reconciliation of Cohort data in Infectious Diseases Cohort Browser (data storage and delivery). All steps taken in this process (numbered arrows) are also numbered in the key entitled "Process Steps."



XSL•FO
RenderX

Table . Enablers and barriers that affected the Erasmus Medical Center (Erasmus MC) pilot's execution.

Description of enablers and barriers	Stakeholders	Category of actions
Enabler Description		
Active support from legal advisors or GDPR[a] specialists. Navigating the complexity of regulations governing data sharing, such as the GDPR, and to apply these accordingly, requires expert privacy and legal knowledge. The active support from legal or GDPR specialists was therefore pivotal in defining the process of dataflow within ReCoDID[b].	Academic hospitals or institutes	Ethics and legal requirements
Building forward on existing infrastructure for sharing (biomedical) data, allowed the rapid development of the connected data hubs to include controlled access sensitive data for/in this pilot.	Researchers; European Union and national funding agencies	Data storage and delivery
Building forward on existing collaborations, as developed in ReCoDID and COMPARE (Collaborative Management Platform for Detection and Analyses of (Re)emerging and Foodborne Outbreaks in Europe) [8], facilitated the steps in the pilot and helped to further build trust necessary for data sharing between participants. The collaboration between the data provider or controller (Erasmus MC) and the respective EMBL-EBI[c] repositories, as well as between the EMBL-EBI repositories, was needed in this pilot to drive further developments and to extend the existing infrastructures to host and present connected datasets.	Researchers; European Union and national funding agencies	All
Dedicated data scientist championing the pilot. Acquiring data from different sources and systems and collected by various specialist teams, linking such data, and ensuring the (legal) requirements are fulfilled before sharing such data, is cumbersome and time-consuming. A dedicated data scientist, supported by encouragement and advocacy of the principal investigator for data sharing amongst researchers, ensured the pilot kept moving ahead.	Academic hospitals or institutes	All
Barrier Description		
Adhering to GDPR, and differences in its interpretation. The data workflow gave rise to legal issues in the context of the GDPR, with regards to GDPR roles (eg, controller vs processor), potential liabilities, different interpretations of Article 49 (d) (Derogations for specific situations, important reasons of public interest), also in regard to the transfer of data to international organizations (here: EMBL-EBI), and resulted in long-lasting legal discussions to define the process of dataflow within ReCoDID.	European Union and national governments; supervisory authorities[d]	Ethics and legal requirements

XSL•FO
RenderX

Description of enablers and barriers	Stakeholders	Category of actions
Insufficient FAIRification[e] and sharing of initial data. Lack of both mandatory FAIRification and sharing rules and specialized staff in inter-pandemic times led to a situation during the pandemic, when time and resources were naturally scarce, often differently prioritized and competing priorities existed, where FAIRification and data sharing was not the main priority. This also led to a situation where harmonization of CE[f] data according to international standards (WHO-ISARIC[g] format) [9] was delayed or not completed.	Academic hospitals or institutes; researchers; European Union and national funding agencies	Data identification and acquisition; data storage and delivery
Insufficient interoperability of the data. Establishing linkages between CE data and corresponding high-dimensional lab data was challenging, and the availability of overlapping data types was limited. Various identifiers were used for CE data and derived high-dimensional laboratory data, and metadata was not fully harmonized, making it difficult to establish linkages. Furthermore, as high-dimensional data types were generated in separately funded projects, a limited amount of overlap in high-dimensional data types existed for the patients.	Academic hospitals or institutes; researchers; European Union and national funding agencies	Data identification and acquisition
Issues contributing to difficulty in CE data acquisition		

XSL•FO
RenderX

Description of enablers and barriers	Stakeholders	Category of actions
• Competing priorities: there was a plethora of clinical studies to be entered into the institute's COVID-19 clinical database. Consequently, there were difficulties in getting approval for prioritization of data entry and data access for this project. Finally, over time there was a reduced interest with decreasing burden of COVID-19 in completing the data entry of CE data. • Time consuming and resource intensive handling and storage of CE data: CE data was stored at different specialized hospital health record systems, that allowed only partial automation of data entry into the institute's COVID-19 clinical database, and records needed to be administered individually. During a time when clinical services were overburdened with COVID-19 cases this resulted in a backlog of CE data that needed to be entered into the institute's database. • Discontinued institutional funding and support for its COVID-19 clinical database: after one year (in 2021), the database was discontinued and subsequently placed under different management. This delay also meant that additional requirements, eg, related to data protection and ethics, had to be met before CE data could be accessed and shared.	Academic hospitals or institutes	Data identification and acquisition

[a]GDPR: General Data Protection Regulation.

[b]ReCoDID: Reconciliation of Cohort data in Infectious Diseases.

[c]EMBL-EBI: European Molecular Biology Laboratory-European Bioinformatics Institute.

[d]The supervisory authorities (barrier "Adhering to GDPR, and difference in its interpretation") refer to the independent public authorities in each Member State responsible for monitoring the application of the GDPR.

[e]FAIR: Findable, Accessible, Interoperable, Reusable.

[f]CE: clinical-epidemiological.

[g]WHO-ISARIC: World Health Organization-International Severe Acute Respiratory and emerging Infection Consortium.

Linking of Different Data Types

At the time of conducting this pilot, submission of SARS-CoV-2 viral genomes to repositories such as EMBL-EBI's European Nucleotide Archive (for raw sequences) and GISAID [10] (for consensus sequences) was done routinely, other laboratory datasets (protein-microarrays, T- and B-cell data) were not shared. To link SARS-CoV-2 genomes to the other datasets a check was done to verify whether viral genomes had been submitted from the 151 individuals. This exercise required iterative rounds of mapping 4 different identifiers. The internal mapping exercise further revealed the partiality of all other datasets for the individuals included in this cohort. Serum antibody profiles generated using protein-microarrays were available for 40 patients, T-cell phenotype and reactivity data for 28 patients and clonal antibody cross-reactivity data using B-cell profiling for 17 patients out of 151 selected patients in the cohort. For 7 patients out of 151 all data types were available.

Configuration of Data Infrastructure and Data Delivery

With the exception of the viral genomes that were routinely submitted to known repositories (refer to "Linking of different data types" section), for storage and linkage of the dataset it was needed to identify suitable respositories for all other laboratory datasets. This meant establishing where new pilot data types could be hosted, and how the EMBL-EBI system could be leveraged to link all pilot datasets within the EMBL-EBI system, so that the datasets could be eventually also "presented" through a Cohort browser [11]. Although the EMBL-EBI repositories were not fully designed to host one of our data types (T-cell) it was possible to submit the outputs with some adjustments to the format as proof of concept. This leaves room for discussion between the respective researchers and the EMBL-EBI to allow for a more routine application.

The upload of the last Erasmus MC pilot datasets was completed in November 2022. To link all datasets EMBL-EBI's BioSample database [12] was used with a hierarchical model allowing

XSL•FO
RenderX

linkage between individuals and their derived samples. With the completion of linking all data types on participant level and presentation of all datasets through the Cohort browser, the pilot was considered as completed (data storage and delivery) in December 2022 [13]. Study level metadata and links to the datasets in the corresponding repositories can be accessed through the Cohort browser and the BioSamples browser at EMBL-EBI [14].

Enablers and Barriers Affecting the Execution of the Pilot

Several enablers and barriers were identified that either facilitated or hampered the execution of this pilot, respectively (Table 1). The main category of actions affected by these barriers and enablers was also identified, as well as the main stakeholder(s) to influence the enablers and barriers.

Several barriers significantly delayed the sharing of a linked dataset in this pilot. The barriers associated with complying with ethics and legal requirements, and with data identification and acquisition, caused the most evidential delays. Some of these barriers tended to reinforce each other, particularly due to shared underlying issues, such as the siloed nature of clinical and complex laboratory data collection and storage. At the start of the pilot, CE data was in different hospital systems, and not yet integrated into the (newly established) institute's database for COVID-19 patient related CE data ("Issues contributing to difficulty in CE data acquisition" in Table 1). Derived high dimensional laboratory data was stored in different systems, separate from patient and diagnostic databases ("Insufficient FAIRification and sharing of initial data"). This was a result of data being gathered by multiple different specialist teams, often for separately funded projects, as well as the lack of linked data infrastructures capturing both patient and (raw) laboratory data, which is typical for hospital systems. Furthermore, the COVID-19 pandemic was characterized by a fragmented funding landscape. Many projects were funded, often focusing on a specific research question or field, thereby reinforcing the siloed nature of specialist teams' work. In addition, for each of these projects, patient sample materials were needed (eg, swabs, blood, and serum). Consequently, high-dimensional lab data were generated for only a subset of patients, based on availability of sample material and requirements of the specific research topic ("Insufficient interoperability of the data").

Interestingly, the newly established institute's COVID-19 clinical database could have in principle worked as an enabler for data sharing. It was designed with the aim for CE data to be collected, stored and retrieved from one central database, and potentially address the issue of CE data being siloed in different existing hospital IT systems and databases and improve the FAIR aspects of such data. Yet, a variety of issues, described in Table 1, caused significant delays in CE data acquisition and sharing.

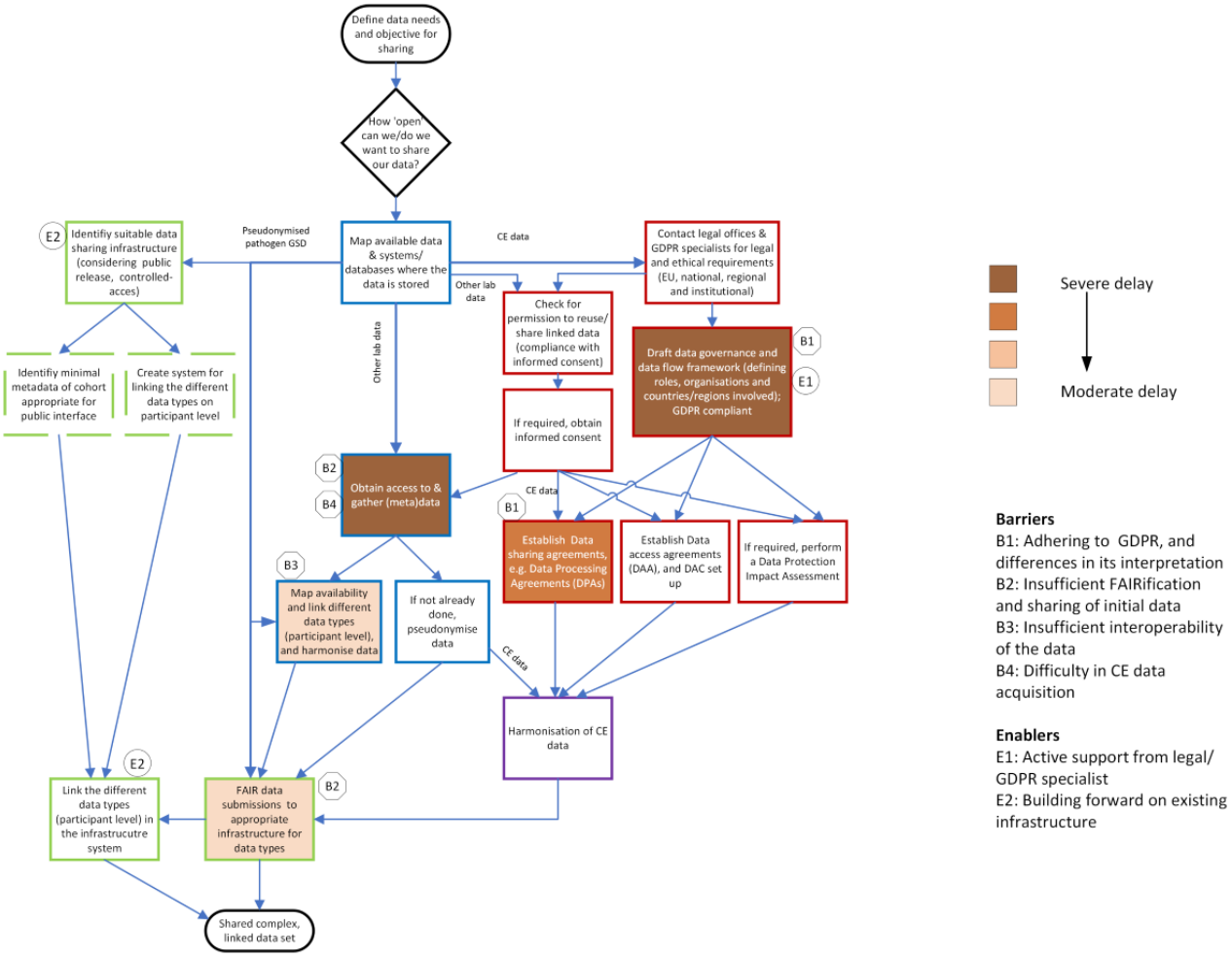
Several factors contributed to or facilitated the sharing of a linked dataset in this pilot, which showed the willingness of the involved teams to overcome the barriers in the sharing of a complex, linked dataset.

Critical Path for Sharing a Complex, Linked Dataset

The critical path analysis (Figure 2) defined the steps in the process of sharing a complex, linked cohort dataset, showing the necessary order of steps or tasks to complete the submission and sharing of such datasets through an appropriate infrastructure, based on our pilot experiences. These steps were classified into the same 4 categories of actions introduced above, data identification and acquisition (4 steps), ethics and legal requirements (7 steps), external data harmonization (1 step), and data storage and delivery (5 steps). Most of the steps in this process could be worked on concurrently, depending on the availability of resources (eg, time of legal staff; availability of appropriate, and standardized templates). As shown in Figure 2, pathogen genetic sequence data (GSD) required a minimum number of steps (1 step), followed by other laboratory data (6 steps), whereas CE data required at least 10 steps. This also corresponds to the time for completion of all steps for each data type in our pilot, with the least amount of time for pathogen GSD. Figure 2 also shows how the identified barriers and enablers affected the process for this pilot.

XSL•FO
RenderX

Figure 2. Critical path for sharing a complex, linked dataset. Two rounded boxes mark the start and endpoint; the diamond box represents a decision. Steps were classified into 4 categories of actions (color coding corresponds to timeline): data identification and acquisition (4 steps, shown in text boxes with blue margins), ethics and legal requirements (7 steps, shown in red), external data harmonization (1 step, shown in purple), and data storage and delivery (5 steps, shown in green). Dashed green boxes represent steps that were needed for this first pilot to build the system (European Molecular Biology Laboratory-European Bioinformatics Institute). Different types of data required different steps in this process, as shown for pseudonymized pathogen genetic sequence data (GSD), other laboratory datasets (other lab data), and clinical-epidemiological (CE) data. The identified barriers and enablers are indicated at the specific steps in the process that were mainly affected for our pilot; 2 enablers are not shown as these did not affect specific steps but acted as general facilitators of this process for our pilot. DAC: data access committee; EU: European Union; FAIR: Findable, Accessible, Interoperable, Reusable; GDPR: General Data Protection Regulation;



Mission (Im)possible: A Call for Action

The pilot demonstrated that sharing a linked cohort dataset using a centralized approach can succeed, but several crucial barriers (as described in Table 1) exist that cause long delays. Although these barriers were not unique to COVID-19 pandemic and have been extensively described previously [15-17], they persist today, as currently the burden to overcome or circumvent such barriers is placed on the individual researchers, projects, or institutes. However, the root causes of these barriers are systemic in nature. To alleviate these barriers, stakeholders at institutional, national and international or EU level (eg, governments, health care, research, journals, funders, and legal) need to work jointly toward building a model that not only supports but motivates data exchange by simplifying steps in the process of data sharing, taking privacy and legal aspects into account, and to reward researchers in new ways [18]. Below we describe some immediate actions to achieve this based on our experiences.

Support With Privacy Questions and General Data Protection Regulation

We recommend creating a roadmap amongst EU countries for how to deal with common issues, to avoid (recurrent) lengthy legal discussions. Harmonization of regional and national implementation of GDPR, and alleviation of the disproportionate burden in legal compliance activities for data sharing with international organizations, needs to be achieved [19]. Finally, additional support should be provided by the European Union to their funded projects, for example, by providing legal guidance or a top up for projects to hire legal experts.

Build and maintain FAIR Data Sharing During "Interpandemics"

Mandatory FAIRification of data in funded projects with funding for personnel to build and maintain FAIR data collections and FAIR sharing, and monitoring of and consequences for noncompliance. At the same time, more data repositories submitting to evaluation-based initiatives and accreditation

XSL•FO
RenderX

processes [20-22], could help with a cultural change toward more FAIR data sharing. Furthermore, FAIRification and sharing of data should become standard practice during the education of bachelor, master, and PhD students.

Decrease Fragmented Funding Landscape

Consideration for funding interdisciplinary projects to enable combined data sharing and, continued funding support for existing data infrastructures is key to sustaining the efforts of data FAIRification and sharing, both nationally and internationally. One clear example is the COVID-19 open data portal [23], an initiative launched at the explicit wish from the European Commission. It provides access to open data from a range of areas of expertise for use across sciences (biomedical and social sciences). A second example is the newly launched World Health Organization (WHO) pandemic hub, to develop "collaborative surveillance," a new area that seeks to bring together data collected by different areas of expertise. This includes the classical epidemiological data that are the core of public health surveillance, but also data on climate parameters, and data on factors that influence disease dynamics.

Willingness and Commitment Building (Culture Change)

We call for establishment of a common policy about data sharing allowing culture change at the EU or international level and for provision of resources to do so. One way of supporting culture change can be achieved by reflecting how researchers are recognized, rewarded, or assessed beyond publishing results in a journal. On short term, acknowledging data providers when reusing their datasets, and the use of citable data Digital Object Identifiers (DOI) pointing to their data could help (eg, as used for one of the datasets used in this pilot [24]). At institutional level, common policies about data sharing should also be established, and resources provided to implement these, to improve existing sharing norms and cultures.

Standardization and Interoperability

Institutes need to prioritize standardization or reconciliation of existing databases and systems to tackle data siloes. Overarching infrastructures at University Hospitals that would allow and enhance standardization, harmonization, and linking of different data types locally and nationally could also increase interoperability at EU and international level [25].

Discussion

Without action, barriers will persist, especially for sharing CE data, which hamper FAIR data sharing for infectious disease preparedness and response. The actions described here also contribute toward the commitment from closed to open research information, as put forward in the recent Barcelona Declaration [26]. The first and uppermost question needs to be answered at institutional and national level: Do people want to share data and what are they prepared to invest in terms of a collaborative effort to do so, if there was funding, GDPR and other support, for data sharing to move away from a felt "burden" to being a seen "benefit." On the path toward more open data sharing researchers' and other stakeholders' concerns about sharing "their" data need to be addressed [15,27,28].

Meanwhile, decentralized or federated approaches are often preferred, as these circumvent barriers [16,29,30] and are less resource- and time intensive for (clinical) researchers and institutes. Federated networks, where decentralized but interconnected nodes allow data to be queried or analyzed by other nodes in the network without the data leaving its location, have been proposed as a solution and are being worked on to address the siloing of health data and current barriers to data sharing [31]. However, common obstacles to broader uptake of federated networks include the absence of data standards or limited adherence to existing standards, the complexity of designing, implementing and deploying federated solutions that preserve privacy [32]. The challenges must be addressed before federated networks can be implemented more widely, especially across national borders [31], but the question remains if full federation will ever be possible. Other concerns with federated approaches, such as the forthcoming European Health Data Space [33], are about significant delays in the authorization by national health data access bodies for reuse or secondary use of health data [34], and introduction of paywalls [35], potentially leading to inequity.

There is a need to streamline the timeliness of steps for sharing CE and other data, and to ensure no additional barriers are created in sharing a linked cohort dataset. Ongoing discussions to include digital sequence information, including pathogen GSD, under the Nagoya protocol [36], or the WHO CA+ [37], could potentially result in new barriers due to additional steps to complete legal procedures before pathogen GSD could be shared. This may considerably delay the prompt sharing and access to pathogen GSD, which has proven crucial in early detection, diagnostics, and identifying variants during the COVID-19 pandemic [2], while its impact on improving access and benefit sharing mechanisms remains uncertain [38].

The multidisciplinary nature of infectious disease research and collaborations reflects the complexity of research and public health questions which need to be understood and answered around (re-) emerging diseases. While certain answers can be found in isolation within a specific area of research and small numbers of patients, larger complex questions require a joint outlook and a combination of data sources to be examined together [39].

Conclusions

Here we have shown the complexities in sharing a small but complex, linked dataset from opportunistic COVID-19 clinical and laboratory research studies as done in (academic) hospital settings. We also provided actionable elements to shorten the timeline to go through this process. Actions that are taken now to improve data sharing for outbreak preparedness and response will improve our ability to detect and respond to emerging threats in the future. We call upon governments, funders, global and regional organizations, scientists and their institutes, journals and industries to tackle known barriers hampering data sharing for infectious disease preparedness and response.

XSL•FO
RenderX

Acknowledgments

The work described here would not have been possible without many other people within the Erasmus Medical Center (Erasmus MC) and the Reconciliation of Cohort data in Infectious Diseases (ReCoDID) consortium; we are grateful for your dedication, effort, and for supporting the performance of this study. Specifically, we would like to thank the following people. At Erasmus MC: Pieter Fraaij and Jeroen van Kampen for support with the data landscape; Divyae Prasad and Marie-Catherine Bouquieaux for SARS-CoV-2 data submissions; Mickey Haasnoot and Laurens Groenendijk for support with the legal requirements; Tamara Van Huuksloot, Viola Woeckel, Anne-Lotte Masson and Kenneth Davids for support with the institutional database. At European Molecular Biology Laboratory-European Bioinformatics Institute: Nadim Rahman and Zahra Waheed for support with the European Nucleotide Archive; Dipayan Gupta for support with BioSample; and the ArrayExpress/BioStudies and European Genome-Phenome Archive teams. At Heidelberg University Hospital (UKHD; Germany): Fruzsina Molnar-Gabor for legal support concerning the General Data Protection Regulation. At McMaster University (Canada): Priscilla Cesar and Claudia Emerson for the overall RECoDID Data Governance framework. Financial support for this work was provided by the European Union's Horizon 2020 research and innovation program (grant agreements 825746 [ReCoDID] and 874735 [VEO]). The funders had no role in any aspect pertinent to the study, nor in the decision to submit the manuscript. A preprint of this manuscript has been published by Zenodo.

Authors' Contributions

Although the datasets themselves and their connected data analysis are not the focus of this manuscript, the pilot study, to share a complex linked COVID-19 cohort dataset, would not have been possible without the respective data providers. Data providers were BBOM, RDdV, GPvN, RSS, JvB, and ECMvG. FT carried out clinical-epidemiological data harmonization. Funding acquisition was done by TJ, MPGK, MYvR, and GC. CA and GR carried out the project administration (management and coordination responsibility for the pilot planning and execution at Erasmus Medical Center and European Molecular Biology Laboratory-European Bioinformatics Institute respectively). The project was conceptualized and the manuscript conceived by MYvR, CA, and MPGK. Writing-original draft was carried out by CA and MYvR, and Writing-review and editing by all authors. Figures were designed by MYvR and GR. Visualization (preparation, creation and presentation of work at consortium meetings) was done by CA and MYvR. CA and MYvR contributed equally to this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary material - timeline.
[DOCX File, 456 KB - publichealth_v11i1e63996_app1.docx]

References

1. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3:160018. [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
2. Oude Munnink BB, Nieuwenhuijse DF, Stein M, et al. Rapid SARS-CoV-2 whole-genome sequencing and analysis for informed public health decision-making in the Netherlands. N Med 2020 Sep;26(9):1405-1410. [doi: 10.1038/s41591-020-0997-y] [Medline: 32678356]
3. Reconciliation of Cohort data in infectious diseases. ReCoDID. URL: https://recodid.eu [accessed 2024-04-29]
4. Amid C, Pakseresht N, Silvester N, et al. The COMPARE data hubs. Database (Oxford) 2019 Jan 1;2019:baz136. [doi: 10.1093/database/baz136] [Medline: 31868882]
5. Rahman N, O'Cathail C, Zyoud A, Sokolov A, Oude Munnink B, Gruning B, et al. Mobilisation and analyses of publicly available SARS-CoV-2 data for pandemic responses. Microb Genom 2024;10(2). [doi: 10.1101/2023.04.19.537514]
6. Aguilar-Bretones M, Westerhuis BM, Raadsen MP, et al. Seasonal coronavirus-specific B cells with limited SARS-CoV-2 cross-reactivity dominate the IgG response in severe COVID-19. J Clin Invest 2021 Nov 1;131(21):e150613. [doi: 10.1172/JCI150613] [Medline: 34499051]
7. Lu L, Sikkema RS, Velkers FC, et al. Adaptation, spread and transmission of SARS-CoV-2 in farmed minks and associated humans in the Netherlands. Nat Commun 2021 Nov 23;12(1):6802. [doi: 10.1038/s41467-021-27096-9] [Medline: 34815406]
8. COMPARE. Collaborative management platform for detection and analyses for (re-)emerging and foodborne outbreaks in Europe. URL: https://www.compare-europe.eu/ [accessed 2024-01-17]
9. International severe acute respiratory and emerging infection consortium. ISARIC. URL: https://isaric.org/research-2/covid-19-clinical-research-resources/covid-19-crf/ [accessed 2024-01-09]
10. GISAID. URL: https://gisaid.org [accessed 2024-01-16]

XSL•FO
RenderX

11. Cohort Browser Infectious disease cohort studies shared at EMBL-EBI. URL: https://www.pathogensportal.org/cohorts?activeTab=Dashboard&acc=BSC0000005 [accessed 2024-04-29]
12. Courtot M, Gupta D, Liyanage I, Xu F, Burdett T. BioSamples database: FAIRer samples metadata to accelerate research data management. Nucleic Acids Res 2022 Jan 7;50(D1):D1500-D1507. [doi: 10.1093/nar/gkab1046] [Medline: 34747489]
13. Stroe O. First linked dataset on Pathogens Portal: EMBL-EBI. 2022. URL: https://www.ebi.ac.uk/about/news/updates-from-data-resources/pathogens-portal-linked-dataset/ [accessed 2022-12-20]
14. BioSamples, EMBL-EBI. Search results for ReCoDID COVID-19 pilot study. URL: https://www.ebi.ac.uk/biosamples/samples?text=ReCoDID%20COVID-19%20pilot%20study [accessed 2023-12-20]
15. Ribeiro CDS, Roode MY, Haringhuizen GB, Koopmans MP, Claassen E, Burgwal LHM. How ownership rights over microorganisms affect infectious disease control and innovation: A root-cause analysis of barriers to data sharing as experienced by key stakeholders. PLoS ONE 2018;13(5):e0195885. [doi: org/10.1371/journal.pone.0195885]
16. Bernier A, Molnár-Gábor F, Knoppers BM, et al. Reconciling the biomedical data commons and the GDPR: three lessons from the EUCAN ELSI collaboratory. Eur J Hum Genet 2024 Jan;32(1):69-76. [doi: 10.1038/s41431-023-01403-y] [Medline: 37322132]
17. Vlahou A, Hallinan D, Apweiler R, et al. Data sharing under the general data protection regulation: time to harmonize law and research ethics? Hypertension 2021 Apr;77(4):1029-1035. [doi: 10.1161/HYPERTENSIONAHA.120.16340] [Medline: 33583200]
18. European Commission, Directorate-General for Research and Innovation, Maxwell L. Maximising investments in health research – FAIR data for a coordinated COVID-19 response– Workshop report: Publications Office of the European Union. 2022 URL: https://data.europa.eu/doi/10.2777/726950
19. Science europe feedback on the implementation of the GDPR. Science Europe. 2024 Feb 16. URL: https://scienceeurope.org/media/5blavcwx/se_response_ec_consultation_gdpr_implementation_final.pdf
20. Lin D, Crabtree J, Dillo I, et al. The TRUST Principles for digital repositories. Sci Data 2020 May 14;7(1):144. [doi: 10.1038/s41597-020-0486-7] [Medline: 32409645]
21. ELIXIR core data resources: ELIXIR Europe. ELIXIR. URL: https://elixir-europe.org/platforms/data/core-data-resources/
22. Resources, GCB. Global Core Biodata Resources: Global Core Biodata Resources. URL: https://globalbiodata.org/what-we-do/global-core-biodata-resources
23. COVID-19 Data Portal - accelerating scientific research through data. URL: https://www.covid19dataportal.org/ [accessed 2024-11-25]
24. Oude Munnink B. ND, Sikkema R, Schapendonk C, et al. On behalf of the Dutch national COVID-19 response team SARS-CoV-2 whole-genome sequencing for informed public health decision making and outbreak tracking and tracing in the Netherlands. URL: https://www.ebi.ac.uk/biostudies/studies/S-BSST6622021. https://doi.org/10.6019/S-BSST662
25. Rinaldi E, Stellmach C, Rajkumar NMR, et al. Harmonization and standardization of data for a pan-European cohort on SARS- CoV-2 pandemic. NPJ Digit Med 2022 Jun 14;5(1):75. [doi: 10.1038/s41746-022-00620-x] [Medline: 35701537]
26. Barcelona declaration on open research information. 2024. URL: https://barcelona-declaration.org/ [accessed 2024-06-01]
27. Federer LM, Lu YL, Joubert DJ, Welsh J, Brandys B. Biomedical data sharing and reuse: attitudes and practices of clinical and scientific research staff. PLoS ONE 2015;10(6):e0129506. [doi: 10.1371/journal.pone.0129506] [Medline: 26107811]
28. van Roode MY, Dos S Ribeiro C, Farag E, et al. Six dilemmas for stakeholders inherently affecting data sharing during a zoonotic (re-)emerging infectious disease outbreak response. BMC Infect Dis 2024 Feb 12;24(1):185. [doi: 10.1186/s12879-024-09054-0] [Medline: 38347527]
29. Peñalvo JL, Mertens E, Ademović E, et al. Unravelling data for rapid evidence-based response to COVID-19: a summary of the unCoVer protocol. BMJ Open 2021 Nov 18;11(11):e055630. [doi: 10.1136/bmjopen-2021-055630] [Medline: 34794999]
30. Tacconelli E, Gorska A, Carrara E, et al. Challenges of data sharing in European Covid-19 projects: a learning opportunity for advancing pandemic preparedness and response. Lancet Reg Health Eur 2022 Oct;21:100467. [doi: 10.1016/j.lanepe.2022.100467] [Medline: 35942201]
31. Hallock H, Marshall SE, 't Hoen PAC, et al. Federated networks for distributed analysis of health data. Front Public Health 2021;9:712569. [doi: 10.3389/fpubh.2021.712569] [Medline: 34660512]
32. Casaletto J, Bernier A, McDougall R, Cline MS. Federated analysis for privacy-preserving data sharing: a technical and legal primer. Annu Rev Genomics Hum Genet 2023 Aug 25;24:347-368. [doi: 10.1146/annurev-genom-110122-084756] [Medline: 37253596]
33. European Parliamentary Research Service (EPRS). European health data space: European parliament. 2023 URL: https://www.europarl.europa.eu/RegData/etudes/ATAG/2023/754642/EPRS_ATA(2023)754642_EN.pdf
34. Lähteenmäki J, Vuorinen AL, Pajula J, et al. Integrating data from multiple Finnish biobanks and national health-care registers for retrospective studies: Practical experiences. Scand J Public Health 2022 Jun;50(4):482-489. [doi: 10.1177/14034948211004421] [Medline: 33845693]
35. FINDATA social and health data permit authority pricing: FINDATA social and health data permit authority. URL: https://findata.fi/en/pricing [accessed 2024-05-24]

XSL•FO
RenderX

36. Convention on Biological Diversity (United Nations) Nagoya protocol on access to genetic resources and the fair and equitable sharing of benefits arising from their utilization to the convention on biological diversity. URL: https://www.cbd.int/abs/doc/protocol/nagoya-protocol-en.pdf [accessed 2023-12-22]
37. World Health Organization (WHO). Bureau's text of the WHO convention, agreement, or other international instrument on pandemic prevention, preparedness and response (WHO CA+) for the consideration of the Intergovernmental Negotiating Body at its fifth meeting: World Health Organization (WHO). URL: https://apps.who.int/gb/inb/pdf_files/inb5/A_INB5_6-en.pdf [accessed 2023-06-02]
38. DSI Scientific Network Using digital sequence information (DSI) to design an mRNA vaccine: COVID-19 case study: DSI Scientific Network. 2023. URL: https://www.dsiscientificnetwork.org/wp-content/uploads/2023/04/DSI-covid-vaccines-factsheet.pdf
39. Eckhardt M, Hultquist JF, Kaake RM, Hüttenhain R, Krogan NJ. A systems approach to infectious disease. Nat Rev Genet 2020 Jun;21(6):339-354. [doi: 10.1038/s41576-020-0212-5] [Medline: 32060427]

Abbreviations

CE: clinical-epidemiological
EMBL-EBI: European Molecular Biology Laboratory-European Bioinformatics Institute
Erasmus MC: Erasmus Medical Center
FAIR: Findable, Accessible, Interoperable, Reusable
GDPR: General Data Protection Regulation
GSD: genetic sequence data
ReCoDID: Reconciliation of Cohort data in Infectious Diseases
UKHD: Universitätsklinikum Heidelberg (University Hospital Heidelberg)
WHO: World Health Organization

Edited by A Mavragani; submitted 05.07.24; peer-reviewed by LN Mayigane, S Pesälä; revised version received 28.11.24; accepted 01.12.24; published 11.02.25.

Please cite as:
Amid C, van Roode MY, Rinck G, van Beek J, de Vries RD, van Nierop GP, van Gorp ECM, Tobian F, Oude Munnink BB, Sikkema RS, Jaenisch T, Cochrane G, Koopmans MPG
A Call for Action: Lessons Learned From a Pilot to Share a Complex, Linked COVID-19 Cohort Dataset for Open Science
JMIR Public Health Surveill 2025;11:e63996
URL: https://publichealth.jmir.org/2025/1/e63996
doi:10.2196/63996

© Clara Amid, Martine Y van Roode, Gabriele Rinck, Janko van Beek, Rory D de Vries, Gijsbert P van Nierop, Eric C M van Gorp, Frank Tobian, Bas B Oude Munnink, Reina S Sikkema, Thomas Jaenisch, Guy Cochrane, Marion P G Koopmans. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 11.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Promise and Perils of Artificial Intelligence in Advancing Participatory Science and Health Equity in Public Health

Abby C King[1], PhD; Zakaria N Doueiri[2], BS; Ankita Kaulberg[2], BS, MBA; Lisa Goldman Rosas[3], MPH, PhD

[1]Departments of Epidemiology & Population Health and of Medicine (Stanford Prevention Research Center), Stanford University School of Medicine, 1701 Page Mill Road, Palo Alto, CA, United States

[2]Department of Epidemiology & Population Health, Stanford University School of Medicine, Palo Alto, CA, United States

[3]Departments of Epidemiology & Population Health and of Medicine (Division of Primary Care and Population Health), Stanford University School of Medicine, Stanford University, Palo Alto, CA, United States

Corresponding Author:
Abby C King, PhD
Departments of Epidemiology & Population Health and of Medicine (Stanford Prevention Research Center), Stanford University School of Medicine, 1701 Page Mill Road, Palo Alto, CA, United States

Abstract

Current societal trends reflect an increased mistrust in science and a lowered civic engagement that threaten to impair research that is foundational for ensuring public health and advancing health equity. One effective countermeasure to these trends lies in community-facing citizen science applications to increase public participation in scientific research, making this field an important target for artificial intelligence (AI) exploration. We highlight potentially promising citizen science AI applications that extend beyond individual use to the community level, including conversational large language models, text-to-image generative AI tools, descriptive analytics for analyzing integrated macro- and micro-level data, and predictive analytics. The novel adaptations of AI technologies for community-engaged participatory research also bring an array of potential risks. We highlight possible negative externalities and mitigations for some of the potential ethical and societal challenges in this field.

(JMIR Public Health Surveill 2025;11:e65699) doi:10.2196/65699

KEYWORDS

digital health; artificial intelligence; community-based participatory research; citizen science; health equity; societal trends; public health; viewpoint; policy makers; public participation; information technology; micro-level data; macro-level data; LLM; natural language processing; machine learning; language model; Our Voice

Introduction

> *"While the future might indeed be bright for AI, it wouldn't be so by accident. We will have to earn it, together."* [Dr. Fei-Fei Li] [Founding Director, Stanford Institute for Human-Centered AI] [From The Worlds I See: Curiosity, Exploration, and Discovery at the Dawn of AI 1]

The last several decades have witnessed a growing mistrust in science among both policy makers and the public at large on an unprecedented scale [2]. The reasons for this trend appear to be many and varied, including a lack of understanding about how science actually operates, partly at least due to an absence of participatory educational opportunities and direct engagement in meaningful science activities [3]. This, in turn, has been linked with poor self-ratings of scientific literacy among school-aged adolescents and adults alike [3]. A general confusion about what the public should reasonably expect from research in terms of scientific veracity or "truth" has been an additional contributor to the public's skepticism about science. These problems have been amplified to a significant extent by social media, which has often worked, either intentionally or unintentionally, to confuse or obfuscate the truth [4]. Unfortunately, such efforts arguably have been abetted by scientific institutions themselves as well as scientific journal outlets which, in their haste to grab the day's headlines, have at times stripped scientific results of their complexity, nuance, and context; or, shockingly, filled untold numbers of journals with "fake science" [5]. Well-documented historical injustices also have contributed to a lack of trust in academic and scientific institutions, particularly among communities of color. For example, the US Public Health Service Untreated Syphilis Study at Tuskegee has generated mistrust in health care institutions and research for generations [6]. More recently, the University of Arizona was found guilty of using DNA samples from the Havasupai Tribe for research without their consent [7]. Recent controversies and public health crises (eg, the COVID-19 pandemic) also have given rise to increased questioning of public health authorities and have sown doubts regarding scientific validity and reliability, which have exacerbated feelings of mistrust.

This alarming trend of mistrust in science threatens public health efforts in multiple ways, which include the following recent examples: (1) misinformation concerning vaccines, including

XSL•FO
RenderX

outright dismissal of the extensive multigenerational evidence base supporting the development and use of vaccines to prevent a myriad of dangerous and at times fatal diseases that have threatened the public's health; (2) skepticism surrounding the wide-ranging, multidisciplinary evidence base establishing the considerable effects of climate change that threaten the health of the planet and all of its living inhabitants; and (3) the growing distrust of health care systems as well as pharmaceutical companies (ie, "big Pharma"), which has become especially acute among marginalized groups, including minoritized and under-resourced communities [8].

The mistrust of science coincides with other pernicious societal trends reflecting disconnection and skepticism in other facets of daily life. These include growing concerns in the health and mental health fields around the current levels of civic disconnectedness and community disengagement among significant sections of the population. Recent data indicate, for example, that civic participation, social engagement, and community connection among segments of today's youth and young adults are declining, with this trend having been worsened by the COVID-19 pandemic [9].

The confluence of these trends with the emergence of an often-bewildering pace of advances in the worldwide IT sector has served to exacerbate feelings of confusion and mistrust among broad segments of the population. This has included, in particular, the unexpectedly disruptive force of the current artificial intelligence (AI) revolution across all sectors, including health care, education, and research. Threats regarding the spread of misinformation, increasing systematic biases against already marginalized populations, data privacy issues, and potential job displacement are among the leading concerns of the public, in addition to other impacts that the widespread acceptance and use of AI might have on various segments of society, from individuals through governments and policy makers [10].

The above issues have led to the following questions: (1) how can science in general be made more accessible to the public in ways that can reduce confusion and mistrust and increase the meaningfulness of scientific inquiry to directly benefit individuals and their communities, and (2) how can the potential of AI be harnessed to bolster the public's engagement in public health science while mitigating the distrust and confusion about both research and AI itself?

Increasing Public Participation in Scientific Research

Public participation in science research (PPSR) is one way researchers have been addressing the increasing mistrust among decision makers and the public toward science. A growing number of studies have included PPSR methods, such as citizen science, that have utilized educational outreach efforts to increase public knowledge and awareness of scientific principles and best practices [11]. Important, systematic efforts have been employed to actively involve a greater number and variety of community members in the scientific process [12]. Such participatory research paradigms can have substantial positive impacts for increasing transparency, knowledge, and trust in scientific inquiry, particularly when they involve generating actionable data of direct relevance to those involved [13].

Within the broader umbrella of PPSR frameworks, citizen science—generally defined as involving lay persons in the research process to advance science—is an increasingly popular approach to engage community members in research activities [14]. A strength behind citizen science is its underlying principle that scientific inquiry can often benefit from including diverse members of the community from varying educational, occupational, social, and cultural backgrounds. Various forms of citizen science have occurred over several centuries and have involved diverse fields, including the life, environmental, health, social, and behavioral sciences [15]. Citizen science can directly benefit research by providing extra pairs of "helping hands" in the data collection and problem-solving aspects of science [14]. Beyond democratizing data acquisition, citizen scientists also can contribute to the fuller research process, including problem definition, data interpretation and prioritization, data-driven solution generation, and results dissemination [16-18]. This more fully participatory "by the people" form of citizen science can advance scientific inquiry as well as promote meaningful knowledge and solutions for the local communities that are involved [13,16,17].

An example of this type of citizen science in the public health arena is the technology-enabled "Our Voice" Global Citizen Science Research Initiative [19], where residents from diverse and often under-resourced communities learn how to partner with researchers and community organizations in collecting and interpreting aggregated and anonymized contextual data (eg, photos, text, and audio). These data identify features of residents' local environments that help or hinder their health [16,17]. They then learn how to effectively communicate their data to decision makers and work with them to activate health-benefiting changes in their local environments [16,17]. Over nearly a decade, the multi-generational "Our Voice" Initiative has produced demonstrable community-driven physical and social environmental impacts across 100+ projects involving over 25 countries spanning 6 continents [16,17].

Leveraging the Potential of AI to Promote More Powerful Citizen Science for Advancing Public Health

While the current focus of health-oriented AI applications has been primarily at the individual patient level, the thoughtful and strategic application of AI at the community level offers potentially powerful tools to augment participatory science activities across the fuller research process. At the same time, it is important that both researchers and participants recognize the potential risks of AI and are committed to implementing multilayered mitigation tactics aimed at addressing them (see Table 1, Figure 1, and the subsequent section focused on risk mitigation).

XSL•FO
RenderX

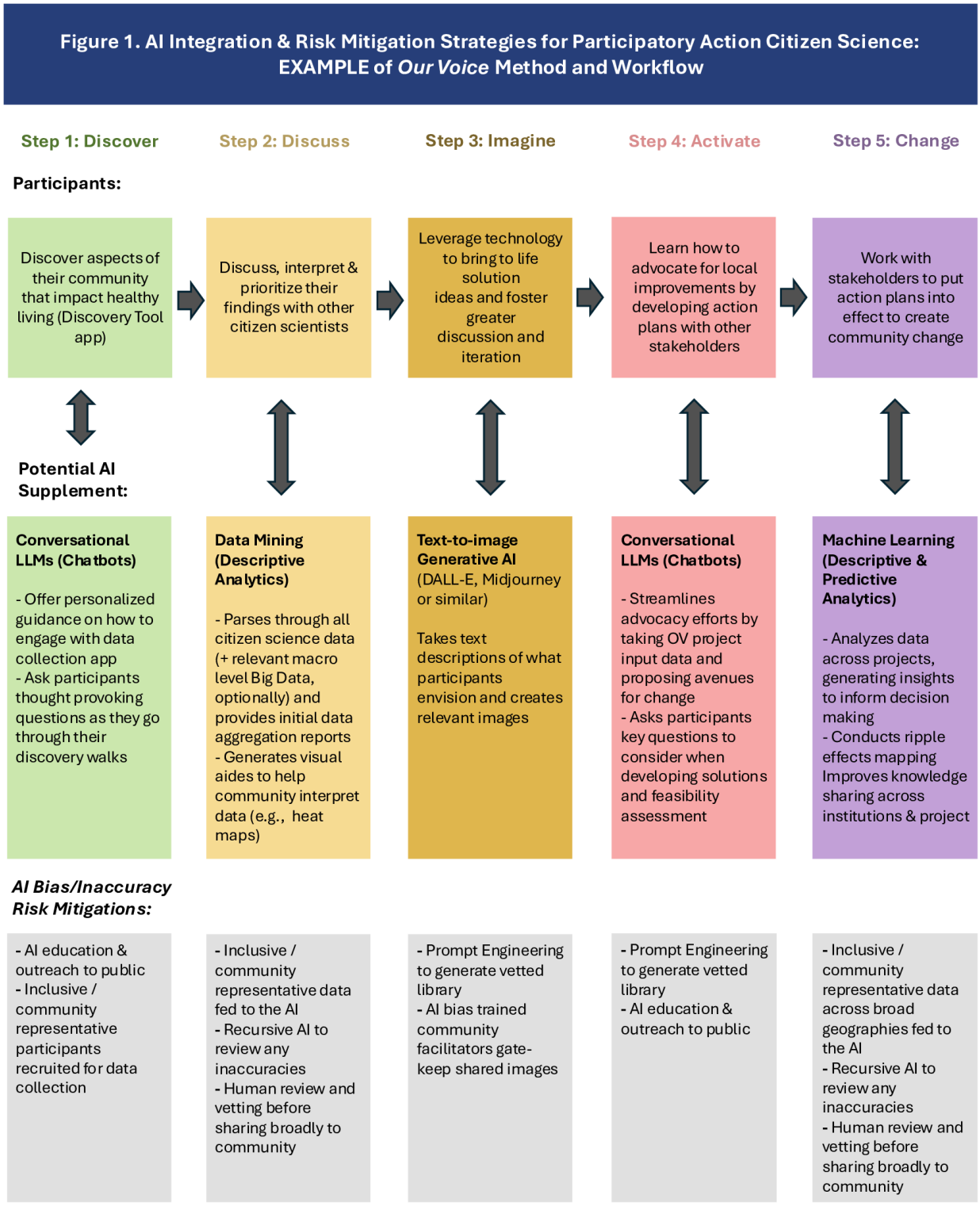
Table . Examples of potential artificial intelligence (AI) uses within community-engaged public health citizen science research, along with potential risks and risk mitigation strategies.

Potential use of AI within community-engaged research	Potential risks	Risk mitigation strategies
Conversational LLMs[a] (ChatGPT or similar) for onboarding participants, offering personalized guidance on engaging with technology, and asking thought-provoking questions	• Biased or culturally insensitive conversations emerge • AI "hallucinations" where factually inaccurate commentary is made	• Dedicated time spent in prompt engineering to determine language best used to minimize these biases in this context, and then shared as best practices to the public • Public education on AI, the training data that were used, and discussion of shortcomings
Text-to-image Generative AI tools (DALL-E, Midjourney, or similar) for helping the community visualize their data (eg, turning text descriptions into relevant images)	• Biased or culturally insensitive pictures generated and shared • Impossible or impractical to execute ideas presented	• Human (expert) review and gatekeeping before images shared broadly to the public • Prompt engineering and similar efforts aimed at community data collection are shared with participants in advance to enhance the potential relevance and feasibility of solutions
Data Mining (Descriptive Analytics) for analyzing macro- and microlevel data, generating integrated data reports, visualizations, etc.	• Inaccuracies or AI hallucinations injected into the dataset • Loss of control, research rigor, and cognitive competence by scientists who defer too much to AI	• Intentional efforts made to feed AI inclusive and community representative data • Recursive AI use: feed output back to AI to check for inaccuracies • Require scientists to review all data generated for accuracy and meet field standards for scientific literacy or competence
Data Mining (Predictive Analytics) for reviewing large datasets, providing data aggregation reports, and generating visual aids to help the community interpret data	• Data presented may result in wrong conclusions, resulting in further inequities • Loss of control, research rigor, and cognitive competence by scientists who defer too much to AI	• Human (expert) review and governance of all generated findings • Recursive AI use: feed output back to AI to check for inaccuracies • Require scientists to review all data and methods for accuracy and meet field standards for scientific literacy or competence • Empirical testing of AI to determine when it helps or hurts scientific skills and competencies

[a]LLM: large language model

XSL•FO
RenderX

Figure 1. Artificial intelligence (AI) integration and risk mitigation strategies for participatory action citizen science: example of "Our Voice" Method and workflow. LLM: large language model; OV: Our Voice.



Using the "Our Voice" multistep citizen science method as an example, potentially valuable AI applications can occur at each step (see Figure 1). The steps include resident-engaged data collection using a multilingual mobile app ("Discover" step); facilitated analysis, discussion, and prioritization of relevant local issues ("Discussion" step); leveraging of visualization technologies to bring potential solutions into clearer focus ("Imagine" step); communicating relevant community issues and brainstorming feasible solutions with local decision makers ("Activate" step); and continued partnering with stakeholders in bringing action plans to fruition ("Change" step). In addition to these integral steps, the above "Our Voice" method has been supplemented at times with an increasingly proactive and thoughtful approach to engage the community prior to

XSL•FO
RenderX

embarking on citizen science activities, as well as the thoughtful dissemination of findings and insights to researchers, decision makers, and the public (a "Share" step). The careful integration of relevant AI applications into these different citizen science processes may assist in supplementing and enriching the different step outputs, as well as helping to address at least some of the challenges with public trust, given the active role that community members play in this type of participatory science.

Some examples of potentially relevant AI tools that can be explored are given below for each step.

Discover (Data Collection) Step

A challenge to address in the discover (data collection) step is that the inclusion of AI in this step should not supersede the real, on-the-ground data collected by local citizen scientists. Any additional data presented by the AI at this step is intended to be an augmentation that goes through human review before it is included in project data capture.

Conversational large language models (LLMs) can be employed in a strategic and transparent way to personally guide the use of citizen science data collection apps such as the "Our Voice Discovery Tool" app, along with "real-time" dynamic prompts to more fully engage residents as they go through their data collection walks. In addition, utilizing computer vision, hearing, and other "real-world" perceptual sensing tools (eg, smell) may help to produce more robust and comprehensive data capture and classifications to augment the contextually focused citizen science data [20]. The goal of using such sensing tools along with LLMs would be to enrich citizen scientists' data collection activities to produce more nuanced and detailed data to analyze and combine with other relevant data sources (eg, geographic information systems and social network models).

Discuss Step

A challenge to address in the discuss step is that AI should not absolve researchers and participants from the task of ensuring analytical accuracy and rigor, nor replace efforts in interpreting the data and gleaning insights. As AI is known at times to experience "hallucinations" and commit errors when it comes to data analysis and interpretation, whatever is generated by it should be seen as a first-pass draft of something that helps researchers and community members enrich their own interpretations and iterations.

ChatGPT and similar conversational LLMs can be engaged during community meetings to support residents' group discussions during data interpretation and solution-building activities. For example, LLMs may be able to support thoughtful community-level discussion and advocacy, not as an omniscient "black box" but as an attentive community-level support system and Socratic thinking aide. Rather than attempting to "replace" community members' perspectives and lived experiences, LLMs may be employed to play the role of "devil's advocate" in a more neutral way, to help residents better understand the concerns of marginalized groups as well as policy makers. LLMs also may allow participants to better assess the feasibility of their data-driven ideas for positive change by supplementing knowledge or experience gaps that the average citizen scientist may have (eg, "Considering the financial constraints of X and timeframe to solve this issue by Y, which of these 5 solutions are the most feasible to address …..?"). In presenting participants with more detailed background information regarding the implementation of different data-driven solutions or strategies, residents can be further informed about the potential complexities that come with improving their communities. AI also could be leveraged to generate and describe multilevel insights by combining diverse sources of data, from citizen scientist–generated microscale data that capture local contexts and individual-level biosensing outputs (eg, heart rate and physiological stress responses) through macroscale population-based social determinants and health data. This, in turn, could result in the development of interactive maps from such mixed-methods data sources that arguably could foster faster insights and public health responses from policy makers. For example, the initial efforts to visually combine citizen scientist–generated microscale data describing residents' neighborhood-lived experiences with macroscale epidemiological data about their larger environmental contexts have yielded richer results than would have been obtained from either data source alone [21]. Leveraging AI in this manner to assist with combining and visualizing different forms of data could actually free up researchers and citizen scientists to engage in higher-order thinking and analysis that is required for creatively addressing some of society's most pressing issues. For instance, in California, low participation in community-supported public health initiatives that provide opportunities for affordable healthy food access is often driven by diverse factors that can vary by locale. The multifactorial issues driving low program uptake often are challenging for researchers, governmental agencies, and community-based organizations to understand sufficiently to be able to intervene effectively. It is possible that AI tools could help by aiding the visualization of how microlevel "lived experience" data vary with respect to macrolevel contextual data. Such multilevel visualizations would allow researchers and community partners to spend more of their time understanding and solving the problem, as opposed to gathering and synthesizing relevant data to describe the problem.

Imagine Step

A challenge to address in the imagine step is that AI should not have free reign in determining what the best solutions for the community are. Using culturally tailored inputs, AI's creativity can be used as a supplemental aide to the collective community's brainstorming while not superseding it.

Several AI tools may be useful in helping residents, decision makers, and researchers better imagine what scientifically generated health-enhancing changes can actually look like in their own communities and environments (eg, through generative AI text-to-image tools such as DALL-E and Midjourney). For example, Stanford's Center on Human-Centered Artificial Intelligence awarded a seed grant to "Our Voice" that will allow citizen scientists to take images collected during local "discovery walks" and transform them into reimagined futures through such generative AI tools (eg, an image of an unsafe intersection to get to school could be "morphed" into images of different solutions such as a crosswalk, pedestrian bridge, etc). The early testing of this form

XSL•FO
RenderX

of AI with school children living in the El Pozón neighborhood of Cartagena, Colombia—where the majority of inhabitants live below the poverty line and lack access to essential services—is showing how the thoughtful, culturally tailored use of text-to-image AI tools can generate meaningful and actionable solutions to local environmental problems (eg, stagnant water and uncollected garbage). These rapid visual renderings could in turn lead to enhanced and richer discussions of trade-offs and foster new ideas among community members and researchers.

Activate and Change Steps

A challenge to address in the activate and change steps is that, as noted above, an optimal goal of AI use is to supplement and enrich such conversations, as opposed to driving them. Sharing the above types of visualization tools and information with local decision makers may help to jump-start relevant solution-building efforts. Generative AI could be used to aid in the creation of presentation materials, help community members role-play conversations with local decision makers in preparation for scheduled meetings with them, and identify issues and suggest modifications to the action plan created during the community meetings.

Share Step

A challenge to address in the share step is that it is recommended that researchers avoid risking diminished scientific rigor and accuracy through outsourcing major data interpretation to AI. As noted earlier, people are in the best position to fully understand and evaluate what the data mean, including, in this case, the citizen scientists who have collected those data.

Machine learning tools also can be used to aid researchers in the aggregation of data across projects and expedite their ability to glean useful insights across different citizen science projects and research groups.

In addition to the above types of citizen science-oriented strategies, through utilizing AI-generated virtual reality simulations, citizen scientists along with other community members can experience physical, cultural, and social situations that they might not otherwise encounter in their daily lives [22]. Such AI-generated virtual reality simulations also can provide community members and decision makers with hands-on opportunities to interact with virtual characters representing differing cultural backgrounds that may lead to increased cultural sensitivity [22]. These AI-driven virtual or simulated environments can additionally allow community members to try out citizen science data gathering and similar tools before real-world use [20].

Additional Recommendations on Mitigating Potential Risks of AI

Amidst the potentially promising uses of AI for the public health and citizen science fields noted above, including its potential to expand the breadth and utility of the data being collected and interpreted, a number of ethical challenges and concerns have been raised, some of which have been touched on earlier and in other articles [23]. These include the different types of risks, as described in this paragraph. AI could diminish, distort, or replace data being collected in the field with output that is less accurate (ie, AI "hallucinations") and less representative of community member experiences and responses. AI output may be based on information and sources that are not culturally relevant or appropriate for a given community or context. Potential challenges emerge when AI is presented with contrasting ideas from citizen scientists, researchers, and other stakeholders and autonomously makes prioritization choices among them. The introduction of AI tools into the public health research process could actually serve to increase the "depersonalization" of the research process, leading to a greater rather than a lower distrust of science. Finally, major concerns have been raised that through potentially oversimplifying and directing aspects of the scientific process, LLMs and similar AI tools could significantly diminish scientists' data processing skills and cognitive competencies.

The publicly available use of ChatGPT and similar generative AI tools as "short cuts" for increasing time efficiencies at the potential expense of thought and deliberation has served to heighten many of these concerns.

In response to such threats, some potential solutions have begun to emerge (see Table 1), as described below.

Thoughtfulness and Transparency

A potential solution is being thoughtful and transparent in determining when IA (information augmentation to support human intelligence) versus AI (artificial intelligence, which is often used as a replacement for some human activities) is best suited. For example, in places where the radius for impact is large and missteps would be difficult to rectify or could cause grave harm, AI, which to date generally does not have community participation or checks for bias mitigation, ought to be used with great caution. In such instances, using technology instead to aggregate and synthesize data findings, and leaving the interpretation of those findings to people, could be wiser.

A mitigation example could be the data collection of citizen scientists' lived experiences in the context of community determinants of health that could be synthesized to share summary statistics, generate heat maps, or word clouds; however, any interpretation of their data, along with solution-building and implementation, would not be outsourced to AI without human supervision and community-level interpretations and buy-in.

Employing Explainable AI Concepts and Strategies

Another potential solution is the use of explainable AI (XAI) that involves processes and methods that allow human users to understand and better trust the results and output created by machine learning and similar algorithms. It can be used to describe an AI model, its expected uses and impacts, and potential biases (eg, based on sex or race).

A mitigation example is the use of a set of vetted chat prompts and outlined best practices that could be provided to users ("prompt engineering"), instead of assuming that the results of ChatGPT are correct and accurate. These best practices could include examples of wording to enter into ChatGPT to improve

XSL•FO
RenderX

results and ideally catch AI "hallucinations." Here is one such example: "For any and all responses, end each response with a bulleted list that includes (1) any assumptions you are making, (2) citations of any sources you have used to determine the response to my prompt, and (3) any biases or concerns you want to review if you had access to unlimited data or personnel." In addition, the responses generated by the AI system can be fed back to itself with the prompt "Check this response for any factual inaccuracies or state anything you cannot back up with reputable sources."

Launching Active Educational and Outreach Initiatives

There are growing concerns related to the potential of AI, including generative AI, to promote biased results and intensify health inequities through the lack of diverse data that are representative of different societal groups. Therefore, it is important to launch active educational and outreach initiatives to better inform the public across all walks of life about AI, its uses, ethical implications, and the precautions about which people should be aware. Educational efforts aimed specifically at diverse communities and populations may help to increase transparency concerning both the promise and limitations of this emerging field. It also may help ensure that key community norms and values are taken into account [24]. Such outreach-based multidisciplinary and community-driven educational efforts tailored to the needs of different populations may additionally help to empower currently under-represented groups to actively participate in data collection and information sharing to help mitigate such biases [24]. Involving community members in the development of AI methodologies from the beginning is important to ensure transparency and reduce mistrust among participants [24]. While such proactive educational efforts have been focused primarily on health care, they could benefit public health research and activities as well.

A mitigation example is described here. Zainab Garba-Sani and colleagues have developed the A.C.C.E.S.S. AI model, which lays a framework for involving communities in the AI-rollout process as follows: A = "Affirm your aims," C = "Consider your communities," C = "Cultivate your conversations," E = "Embrace your essentials," S = "Specify your scope," and S = "Scrutinize your spaces" [24]. Through training sessions and workshops with diverse groups, A.C.C.E.S.S. serves the dual purpose of not only educating people about AI and its potential applications but also gathering their input on how AI should be thoughtfully designed for their unique contexts. Another example of such an initiative is the nonprofit "AI4ALL" movement that aims to advance an increasingly human-centered, inclusive AI discipline [25]. Through proactive participatory approaches such as these, we can begin to demystify AI and ensure that diverse communities are brought to the table when implementing AI tools.

Increasing Efforts to Improve Data Upon Which AI is Modeled and Trained

As AI is merely holding up a mirror to the society or data upon which it is trained, the biases inherent in AI responses are likely due to the missing perspectives and incomplete data upon which it was built. Leveraging collaborations across institutions, social or civic interest groups, and diverse communities to improve the training data should help AI become more aware of differing perspectives and hopefully reduce if not completely minimize the bias. Outreach-based multidisciplinary and community-driven educational efforts as described below that are tailored to the needs of different populations may additionally help to empower currently under-represented groups to actively participate in data collection and information sharing to help mitigate such biases [24].

A mitigation example is the need for clear guidelines and incentives from scientific organizations, universities, journals, and granting organizations (eg, specific grant announcements from the NIH in this area) concerning expanding the population representativeness of the datasets upon which AI is based; this can send a clear message regarding the importance of this issue to the field as well as society at large.

Setting AI-Specific Ethical Standards and Having Expected AI-Influenced Outputs Verified by Experts

It is important to set AI-specific ethical standards across universities, other research organizations, scientific outlets, and funding sources that include the ethics review of study methods as well as have the expected AI-influenced outputs verified by experts trained in the relevant ethics fields. Publishers of scientific journals, books, and other scientific communication channels should be encouraged to provide specific guidelines related to the publication of research investigations that include AI. Demanding transparency of AI use in scientific research more generally is recommended, as a growing collection of policy makers, technology developers, and scientific organizations are currently doing. Setting up appropriate "guardrails" prior to AI use is similarly important to help ensure that flawed, offensive, or inaccurate information does not unexpectedly find its way into AI-driven conversations or other outputs.

A mitigation example is described here. As part of institutional review board policies and activities, some grant-making organizations have included, in addition to the review of a grant proposal's scientific merit, a separate ethics review involving experts in ethics and similar fields to identify areas of potential ethical concern and methods for the mitigation of such concerns. An example is the Stanford Human-Centered Artificial Intelligence Center's process of including a separate ethics review of AI grant proposals to help ensure that appropriate AI-relevant guardrails are in place prior to funding.

Regulating Transparent Data Ownership and Use

It is important to regulate transparent data ownership and use to ensure that access to citizen science data is not limited by commercial interests [20]; this helps to ensure that AI tools and resources are available to everyone, which can democratize its use and diminish public suspicion and skepticism. For example, many users of ChatGPT and similar platforms may not understand that their data currently may be sent to the parent company OpenAI to subsequently power future AI learning, and then be ultimately used by unnamed sources for a variety of undisclosed purposes.

A mitigation example is that, instead of relying on public GPTs whose data ownership and use are ambiguous, the Stanford

XSL•FO
RenderX

School of Medicine and other academic institutions have developed their own custom GPT that ensure data security, HIPAA compliance, and similar outcomes. Additional AI resources that are being developed at Stanford include a user-friendly "AI Playground" platform, built on open-source technologies, that allows the Stanford community to try out various AI models and tools as well as compare the utility of different tools in a head-to-head fashion.

Human Governance and Oversight Over AI Output

It is imperative when undertaking community-engaged forms of research that scientists work to ensure that AI-generated contributions to the research endeavor do not supersede or otherwise hamper the credibility of the contributions from community members and their representatives. While AI tools are often used to obtain meaning from large datasets, in the participatory public health research field, it is typically the residents themselves who can best provide meaning and context to the data being collected and interpreted.

A mitigation example is that the addition of community advisory experts, panels, and similar oversight bodies can help ensure that the use of AI enriches but does not overshadow either the scientific process itself or scientists and the public's participation in that process. It will be particularly valuable to include the expertise of ethicists familiar with AI who can identify potential areas of concern and help build appropriate safeguards.

Ensuring That Scientific Competencies are Enhanced Rather Than Diminished as a Consequence of the AI "Revolution"

With the rapid expansion and ease of use of ChatGPT and other LLM tools across a range of scientific and educational fields, there have been growing concerns related to the detrimental effects that the overuse of such tools can have on current and future generations of scientists and the public. These concerns include the real danger that regular substitution of LLM-generated written and oral communication in science, as well as across other facets of life, could significantly diminish cognitive competencies and the creative and innovative thinking that is the foundation of the scientific process itself. It is clear, at this early stage of generative AI use in participatory science as well as in other scientific areas, that carefully created guidelines for AI usage in all aspects of scientific education and training should be developed and put into place to ensure LLMs and other AI tools support, rather than replace, the facilitation and participation of community-engaged research endeavors.

A mitigation example is described here. A year-long deliberative process by Stanford University's Department of Epidemiology & Population Health has resulted in a clear set of early principles and guidelines for the use of generative AI tools in one of its major education programs. The guidelines have helped to diminish ambiguity among students and faculty while reinforcing the importance of human-centered educational activities in growing the cognitive competencies of vital importance to the advancement of science. In addition to clear guidelines for AI use and requirements that scientists at all levels of training and experience review their data and methods for accuracy, other suggestions include the further development of field-specific scientific standards for scientific literacy and competence independent of AI usage.

It is also important to ensure that scientific data collection and analysis activities themselves are not in some way diminished or curtailed through overdependence on AI tools. As noted earlier, the judicious use of certain AI tools in the participatory science area may help to enrich and expand the insights and impacts gained through participation by residents, decision makers, and scientists working together. However, it is critical that such AI tools undergo thorough empirical testing to determine when, how, and with whom they can enhance rather than impede scientific knowledge and insights. As one example, the "Our Voice" research team is conducting initial experiments to better determine how the use of "text-to-image" generative AI may affect the ability of citizen scientists to better visualize and understand the proposed health-promoting environmental changes drawn from their data. Among the populations for which such tools are being systematically tested are older US women and school children living in the El Pozón neighborhood of Cartagena, Colombia, which was highlighted earlier. In a second example, a recent "Our Voice" citizen science project conducted in Bogotá, Colombia has illustrated how resident-collected data that were augmented via virtual reality tools can ignite compelling discussions and solution building among participants, scientists, and local stakeholders [22].

The multidimensional mixed-methods data collection methods, sources, and analyses required by the types of participatory scientific methods described in this Viewpoint also may serve as a means for preventing the diminution of cognitive competencies. For example, we have found that the increased complexity of the multilevel, mixed-methods data acquired when microscale (eg, citizen scientists' "lived experiences") and macroscale data sources are integrated requires an increased interdisciplinary understanding of both these different forms of data and more intricate and multifaceted data analysis methods, which can serve to enhance cognitive competencies [21].

Discussion and Conclusions

Citizen science has been shown to be effective in positively engaging diverse groups of community members in meaningful, action-oriented research [17]. Such participatory methods can help to lessen both mistrust in science and civic disengagement, which are trends that threaten to impair advances in public health. Can the judicious applications of various AI tools serve to amplify the potential of participatory science for reducing these trends while broadening meaningful scientific advances in public health? Given that it is clear that the AI "genie" is out of the bottle and will likely have substantial impacts across virtually all aspects of science, it is critical that we meet the AI challenge head on. The inevitable integration of AI into participatory research methods must be thoughtfully managed in order to amplify, and not replace, the collective documentation of human lived experiences, and to ensure that the applications of such data to affect change truly reflect the priorities of local communities. We believe that the time is ripe to systematically evaluate the potential of AI beyond current individual-level applications by exploring its utility as a

XSL•FO
RenderX

community aid for advancing public health. At the same time, it is imperative that we continually keep in mind the potential risks accompanying its use, especially among underserved, under-resourced communities. We call on the research community to consider such challenges proactively and empirically over the coming years, finding innovative ways to mitigate the risks while testing the impacts of AI on trust in science, civic engagement, knowledge advancement, and community-driven change.

Acknowledgments

The authors thank Ann Banchoff, MSW, MPH and Steven Lovett, PhD for their valuable input related to this manuscript. We also thank Eduardo De La Vega Sr, PhD for his innovative early community-engaged research activities in Colombia using generative large language models, and Arjan Walia, MS for his leadership in the initial systematic evaluation of generative artificial intelligence models in community-engaged research with older US women. Dr. King received partial support from US Public Health Services grants 5R01AG07149002 (King) and 1R44AG071211 (M Fauci) from the National Institute on Aging. King, Doueiri, and Kaulberg received partial support from a seed grant provided by the Stanford Institute for Human-Centered Artificial Intelligence at Stanford University. The authors declare no potential conflicts of interest with respect to the research, authorship, or publication of this article. No copyrighted material, surveys, instruments, or tools are included in this manuscript.

Conflicts of Interest

None declared.

References

1. The Worlds I See: Curiosity, Exploration, and Discovery at the Dawn of AI: Macmillan Publishers; 2024:4.
2. Kennedy B, Tyson A. Americans' Trust in Scientists, Positive Views of Science Continue to Decline.: Pew Research Center; 2023. URL: https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines [accessed 2023-11-14]
3. Achterberg P, de Koster W, van der Waal J. A science confidence gap: Education, trust in scientific methods, and trust in scientific institutions in the United States, 2014. Public Underst Sci 2017 Aug;26(6):704-720. [doi: 10.1177/0963662515617367] [Medline: 26644011]
4. Epstein Z, Sirlin N, Arechar A, Pennycook G, Rand D. The social media context interferes with truth discernment. Sci Adv 2023 Mar 3;9(9):eabo6169. [doi: 10.1126/sciadv.abo6169] [Medline: 36867704]
5. Nasr N. Overcoming the discourse of science mistrust: how science education can be used to develop competent consumers and communicators of science information. Cult Stud Sci Educ 2021;16(2):345-356. [doi: 10.1007/s11422-021-10064-6] [Medline: 34178180]
6. Alsan M, Wanamaker M. Tuskegee and the health of black men. Q J Econ 2018 Feb;133(1):407-455. [doi: 10.1093/qje/qjx029] [Medline: 30505005]
7. Garrison NA. Genomic justice for native Americans: impact of the Havasupai case on genetic research. Sci Technol Human Values 2013;38(2):201-223. [doi: 10.1177/0162243912470009] [Medline: 28216801]
8. Bazargan M, Cobb S, Assari S. Discrimination and medical mistrust in a racially and ethnically diverse sample of California adults. Ann Fam Med 2021;19(1):4-15. [doi: 10.1370/afm.2632] [Medline: 33431385]
9. Kannan VD, Veazie PJ. US trends in social isolation, social engagement, and companionship nationally and by age, sex, race/ethnicity, family income, and work hours, 2003-2020. SSM Popul Health 2023 Mar;21:101331. [doi: 10.1016/j.ssmph.2022.101331] [Medline: 36618547]
10. Byrne M. The disruptive impacts of next generation generative artificial intelligence. Comput Inform Nurs 2023 Jul 1;41(7):479-481. [doi: 10.1097/CIN.0000000000001044] [Medline: 37417716]
11. Haywood BK, Besley JC. Education, outreach, and inclusive engagement: Towards integrated indicators of successful program outcomes in participatory science. Public Underst Sci 2014 Jan;23(1):92-106. [doi: 10.1177/0963662513494560] [Medline: 23887249]
12. Mintz E, Couch J. Biomedical citizen science at the National Institutes of Health. Citiz Sci 2022 Dec 15;7(1):16651505. [doi: 10.5334/cstp.543]
13. Skarlatidou A, Haklay M, Hoyte S, van Oudheusden M, Bishop IJ. How can bottom-up citizen science restore public trust in environmental governance and sciences? Recommendations from three case studies. Environ Sci Policy 2024 Oct;160:103854. [doi: 10.1016/j.envsci.2024.103854]
14. King AC, Winter SJ, Chrisinger BW, Hua J, Banchoff AW. Maximizing the promise of citizen science to advance health and prevent disease. Prev Med 2019 Feb;119:44-47. [doi: 10.1016/j.ypmed.2018.12.016] [Medline: 30593793]
15. Allf BC, Cooper CB, Larson LR, et al. Citizen science as an ecosystem of engagement: implications for learning and broadening participation. Bioscience 2022 Jul;72(7):651-663. [doi: 10.1093/biosci/biac035] [Medline: 35769502]
16. King AC, King DK, Banchoff A, et al. Employing participatory citizen science methods to promote age-friendly environments worldwide. Int J Environ Res Public Health 2020 Feb 27;17(5):1541. [doi: 10.3390/ijerph17051541] [Medline: 32121001]

XSL•FO
RenderX

17. King AC, Odunitan-Wayas FA, Chaudhury M, et al. Community-based approaches to reducing health inequities and fostering environmental justice through global youth-engaged citizen science. Int J Environ Res Public Health 2021 Jan 21;18(3):892. [doi: 10.3390/ijerph18030892] [Medline: 33494135]
18. English PB, Richardson MJ, Garzón-Galvis C. From crowdsourcing to extreme citizen science: participatory research for environmental health. Annu Rev Public Health 2018 Apr 1;39:335-350. [doi: 10.1146/annurev-publhealth-040617-013702] [Medline: 29608871]
19. King AC. Our Voice Global Citizen Science Research Initiative. 2024. URL: https://ourvoice.stanford.edu [accessed 2025-01-27]
20. Ceccaroni L, Bibby J, Roger E, et al. Opportunities and risks for citizen science in the age of artificial intelligence. Citiz Sci 2019 Nov 28;4(1):1-14. [doi: 10.5334/cstp.241]
21. Soltani S, Hinman JA, Blanco-Velazquez I, et al. Bringing micro to the macro: how citizen science data enrich geospatial visualizations to advance health equity. J Maps 2023;19(1):2216217. [doi: 10.1080/17445647.2023.2216217] [Medline: 37448978]
22. Guevara-Aladino P, Sarmiento OL, Rubio MA, et al. Urban care for unpaid caregivers: community voices in the care block program, in Bogotá, Colombia. J Urban Health 2024 Dec;101(6):1113-1127. [doi: 10.1007/s11524-024-00899-z] [Medline: 39316308]
23. Hsu YC, Verma H, Mauri A, Nourbakhsh I, Bozzon A. Empowering local communities using artificial intelligence. Patterns (N Y) 2022 Mar 11;3(3):100449. [doi: 10.1016/j.patter.2022.100449] [Medline: 35510187]
24. Garba-Sani Z, Farincacci-Roberts C, Essien A. AI: A new framework for advancing health equity in health care AI. Health Aff 2024. [doi: 10.1377/forefront.20240424.369302]
25. AI4ALL webpage. : Stanford University URL: https://ai-4-all.org [accessed 2024-11-11]

Abbreviations

AI: artificial intelligence
LLM: large language model
PPSR: public participation in science research

Edited by A Mavragani; submitted 22.08.24; peer-reviewed by FA Causio, R Feltrero; revised version received 21.11.24; accepted 28.11.24; published 14.02.25.

Please cite as:
King AC, Doueiri ZN, Kaulberg A, Goldman Rosas L
The Promise and Perils of Artificial Intelligence in Advancing Participatory Science and Health Equity in Public Health
JMIR Public Health Surveill 2025;11:e65699
URL: https://publichealth.jmir.org/2025/1/e65699
doi:10.2196/65699

© Abby C King, Zakaria N Doueiri, Ankita Kaulberg, Lisa Goldman Rosas. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Recent Use of Novel Data Streams During Foodborne Illness Cluster Investigations by the United States Food and Drug Administration: Qualitative Review

Michael C Bazaco[1], PhD; Christina K Carstens[1], PhD; Tiffany Greenlee[1], MPH; Tyann Blessington[1], PhD; Evelyn Pereira[1], JD; Sharon Seelman[1], MS; Stranjae Ivory[1], MPH; Temesgen Jemaneh[1], PhD; Margaret Kirchner[1], PhD; Alvin Crosby[1], MBA; Stelios Viazis[1], PhD; Sheila van Twuyver[2], BS; Michael Gwathmey[3], MS; Tanya Malais[3], BS; Oliver Ou[1], PhD; Stephanie Kenez[1], MPH; Nichole Nolan[1], MPH; Andrew Karasick[1], MD; Cecile Punzalan[1], MD; Colin Schwensohn[4], MPH; Laura Gieraltowski[4], PhD; Cary Chen Parker[1], MPH; Erin Jenkins[1], MPH; Stic Harris[1], DVM

[1]US Food and Drug Administration, College Park, MD, United States

[2]US Food and Drug Administration, Minneapolis, MN, United States

[3]US Food and Drug Administration, Silver Spring, MD, United States

[4]US Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Michael C Bazaco, PhD
US Food and Drug Administration, College Park, MD, United States

Abstract

Foodborne illness is a continuous public health risk. The recognition of signals indicating a cluster of foodborne illness is key to the detection, mitigation, and prevention of foodborne adverse event incidents and outbreaks. With increased internet availability and access, novel data streams (NDSs) for foodborne illness reports initiated by users outside of the traditional public health framework have emerged. These include, but are not limited to, social media websites, web-based product reviews posted to retailer websites, and private companies that host public-generated notices of foodborne illnesses. Information gathered by these platforms can help identify early signals of foodborne illness clusters or help inform ongoing public health investigations. Here we present an overview of NDSs and 3 investigations of foodborne illness incidents by the US Food and Drug Administration that included the use of NDSs at various stages. Each example demonstrates how these data were collected, integrated into traditional data sources, and used to inform the investigation. NDSs present a unique opportunity for public health agencies to identify clusters that may not have been identified otherwise, due to new or unique etiologies, as shown in the 3 examples. Clusters may also be identified earlier than they would have been through traditional sources. NDSs can further provide investigators supplemental information that may help confirm or rule out a source of illness. However, data collected from NDSs are often incomplete and lack critical details for investigators, such as product information (eg, lot numbers), clinical or medical details (eg, laboratory results of affected individuals), and contact information for report follow-up. In the future, public health agencies may wish to standardize an approach to maximize the potential of NDSs to catalyze and supplement adverse event investigations. Additionally, the collection of essential data elements by NDS platforms and data-sharing processes with public health agencies may aid in the investigation of foodborne illness clusters and inform subsequent public health and regulatory actions.

(JMIR Public Health Surveill 2025;11:e58797) doi:10.2196/58797

KEYWORDS

foodborne illness surveillance; novel data streams; outbreak investigations; novel data; foodborne illness; foodborne; illness; United States; public health; prevention; outbreaks; social media; product review; cluster; product information; surveillance; epidemiology

Introduction

An estimated 1 in 6 people experiences an episode of foodborne illness in the United States annually, placing an estimated US $77.7 billion burden on the US economy [1,2]. Most of these illnesses are considered sporadic; however, limitations in foodborne outbreak surveillance may hinder the detection of outbreak-associated cases. The identification and investigation of foodborne illness clusters requires a high degree of coordination across state and local jurisdictions and, in some cases, federal agencies [3].

XSL•FO
RenderX

At the federal level, outbreaks are investigated when they involve cases in multiple states, implicate imported products, or when federal support is requested. Most of these outbreaks that involve known pathogens are traditionally initiated through laboratory-based surveillance systems with the identification of illness clusters by molecular subtyping of clinical isolates or case reports from ill individuals with a common exposure and illness who seek medical care. These clusters can then be linked to a common vehicle based on supporting epidemiologic information, traceback findings, and laboratory testing results from food and environmental samples. This information is usually first received by state or local health departments (LHDs) before it is passed on to federal partners, when deemed necessary. Outbreaks that involve known pathogens may also be identified through matching the genetic similarity of historic and contemporary isolates from product or environmental samples to clinical isolates via molecular subtyping techniques, such as whole genome sequencing [4].

In situations where an etiology is unknown or has not yet been identified, consumer complaint systems provide a mechanism for consumers to self-report adverse health events, including foodborne illness. This allows public health entities to conduct disease surveillance to help mitigate this primary limitation of laboratory-based surveillance. Consumer complaint–based systems allow for the collection and storage of consumer reports that can be submitted in a range of formats, including in-person, phone, e-mail, and web form [5]. A survey of LHDs in the United States revealed that a majority (81%) use consumer complaint–based systems and attributed the detection of 69% of foodborne outbreaks in the past year to consumer complaints [6]. State health departments that have conducted self-assessments of their consumer-complaint systems, such as Minnesota and Rhode Island, have credited the detection of 79% and 80% of foodborne outbreaks to consumer complaints, respectively [7,8]. Of the surveyed LHDs in the United States, 75% reported the ability to receive consumer complaints via email, and 40% reported the availability of a web-based reporting form [6]. The Integrated Food Safety Centers for Excellence help states to continually improve these complaint-based surveillance systems as well [9]. In the United States, the various consumer complaint systems create a patchwork that helps to capture information on foodborne illnesses that may not be reported to medical professionals. However, holistically, this patchwork can pose challenges as data from the various consumer complaint systems are stored separately. Recognition of signals depends on a sufficient number of complaints to come to the same complaint system.

Complaint-based surveillance is also used at the federal level to identify potential clusters of foodborne illness. The US Food and Drug Administration (FDA) collects complaints or reports submitted from consumers, health care providers, industry members, public health officials, or other submitters via 2 web-based mechanisms using the MedWatch or Safety Reporting Portal as well as by phone, email, fax, and physical mail [10-13]. A complainant may submit similar reports via multiple mechanisms, and it may not be possible to distinguish if they are duplicate reports or are separate cases. A key element of identifying illness clusters via this approach is the availability of complete and reliable epidemiologic data, such as illness onset date and contact information, to link complainant reports. This information is also needed to narrow down potential etiology and food vehicle [5].

Disease surveillance and outbreak information received by the FDA from various internal and external sources, including federal, state, and local health partners, is evaluated by the FDA's Coordinated Outbreak Response & Evaluation (CORE) Network, which was established in 2011 and charged with coordinating efforts to find, stop, and prevent illnesses linked to FDA Center for Food Safety and Applied Nutrition (CFSAN) regulated products [14]. These products include all human foods not falling under the jurisdiction of the US Department of Agriculture Food Safety Inspection Service and account for 75% of the food supply in the United States [15]. Once an actionable signal has been identified, further investigational activities and public health actions (eg, product traceback, recalls, and communication) are coordinated across local, state, and federal partners, a process that has been described elsewhere [3].

Traditional investigative tools, such as laboratory-based surveillance, do face challenges including but not limited to underdiagnosis and underreporting of illnesses, varying illness reporting structures, the increased use of culture-independent diagnostic testing, and resource availability [16,17]. While some of these challenges are mitigated by public health consumer complaint systems described earlier, the increased accessibility and use of novel data streams (NDSs) provide new information that may further aid in the detection and investigation of foodborne illness clusters. NDSs include nontraditional sources of information such as social media, web-based customer reviews, multimedia news (comment sections from news stories about clusters of illnesses), and private participatory reporting platforms (PPRPs) that actively collect data on food poisoning.

The availability of crowdsourced web-based information provides an opportunity for public health entities to potentially augment existing traditional foodborne illness surveillance systems or identify clusters not detected by traditional tools. The evolution of cell phones to more readily allow users to access the internet has made these NDSs even more accessible as a tool for the submission of consumer complaints. These NDSs, in the context of foodborne illness, consist of reports of consumer information related to food products that are relayed digitally through social media, search histories, customer reviews (ie, product and restaurant reviews), and PPRPs [18]. Some PPRP sites have the express goal of collecting data on people who believe that they have gotten sick from food. Information from these sources may be shared directly with public health agencies or may be mined by agencies in response to a signal.

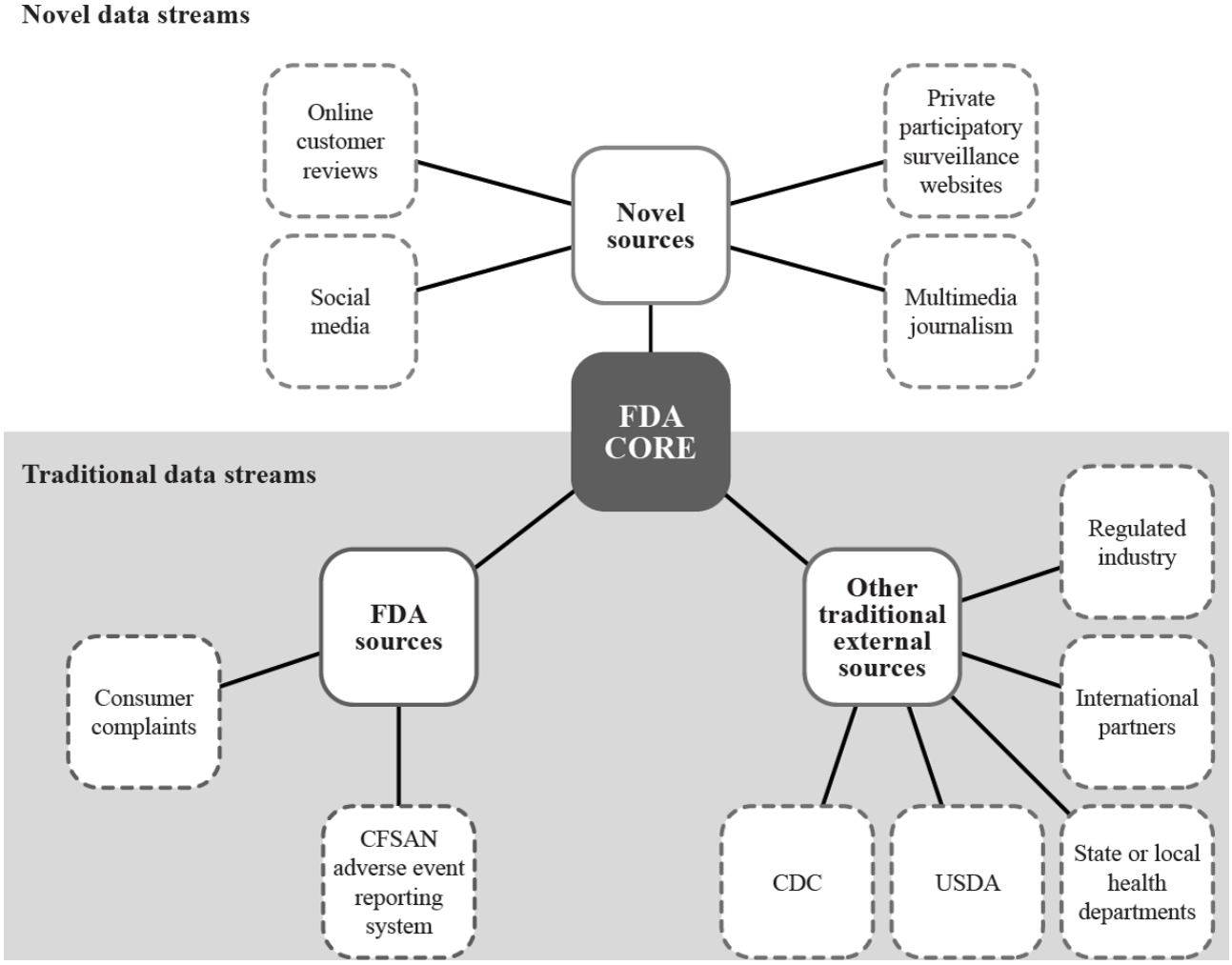
NDSs can be integrated with traditional public health surveillance programs to catalyze or supplement investigation, and several examples of this exist at the local and state health department level. For instance, the Chicago Department of Public Health has used social media posts generated within jurisdictional limits to address consumer complaints of food poisoning and to target health inspections [19]. The New York

XSL•FO
RenderX

City Department of Health and Mental Hygiene has similarly evaluated the use of customer-submitted restaurant reviews to detect outbreaks and target health inspections [20]. These are just a few documented examples of the use of NDSs by state and LHDs. No systematic use of these data sources has been documented at the federal level at this time.

Reports of foodborne illness clusters can come from a variety of traditional and novel data sources (Figure 1). In this review, we (1) highlight key details of 3 recent adverse event clusters that were investigated by the FDA and were catalyzed or supplemented by NDSs, (2) discuss challenges and limitations associated with the current use of NDSs by the FDA, and (3) propose opportunities for improvement and the future use of NDS applications during foodborne illness surveillance and investigational activities conducted by the FDA.

Figure 1. Examples of adverse event and illness cluster data sources evaluated by the US Food and Drug Administration Coordinated Outbreak Response & Evaluation (FDA CORE) Network. Multimedia journalism may also be considered a traditional data source in some cases. CFSAN: Center for Food Safety and Applied Nutrition; CDC: Centers for Disease Control and Prevention; USDA: US Department of Agriculture.



Three Case Studies

Overview

To illustrate the use of NDSs during FDA investigations of foodborne illness clusters, the following 3 case studies describe recent incidents linked to dry breakfast cereal, ketogenic shakes, and plant-based crumbles, respectively. Each of these incidents provides a unique look at how these data streams were used by the FDA to identify or supplement investigations. There may be duplicate reports included in the analyses of these case studies, as reports may have been submitted by the same complaint through multiple points in FDA and to a given PPRP or other form of NDS. The following case studies are meant to highlight examples of this use and are not meant to be comprehensive reviews of the entire incidents.

Dry Breakfast Cereal

On April 1, 2022, the New York Post published an article that highlighted an increase in illness reports mentioning a dry breakfast cereal product on a PPRP [21]. The article stated that at least 139 consumers reported vomiting and diarrhea after consuming the dry breakfast cereal of interest, which prompted an FDA investigation. After FDA CORE initiated an investigation, the PPRP shared complaint data with the FDA, which was reviewed by FDA CORE and used for an initial assessment of the incident. FDA's medical evaluation of 470 PPRP reports received from January 1, 2022, to April 5, 2022, identified symptom onset data for 349 (74%) reports, most of which (96%) suggested nonrapid symptom onset (exceeding 30 min since consumption). In some cases, due to the nature of the reporting tool, these data sometimes omitted necessary

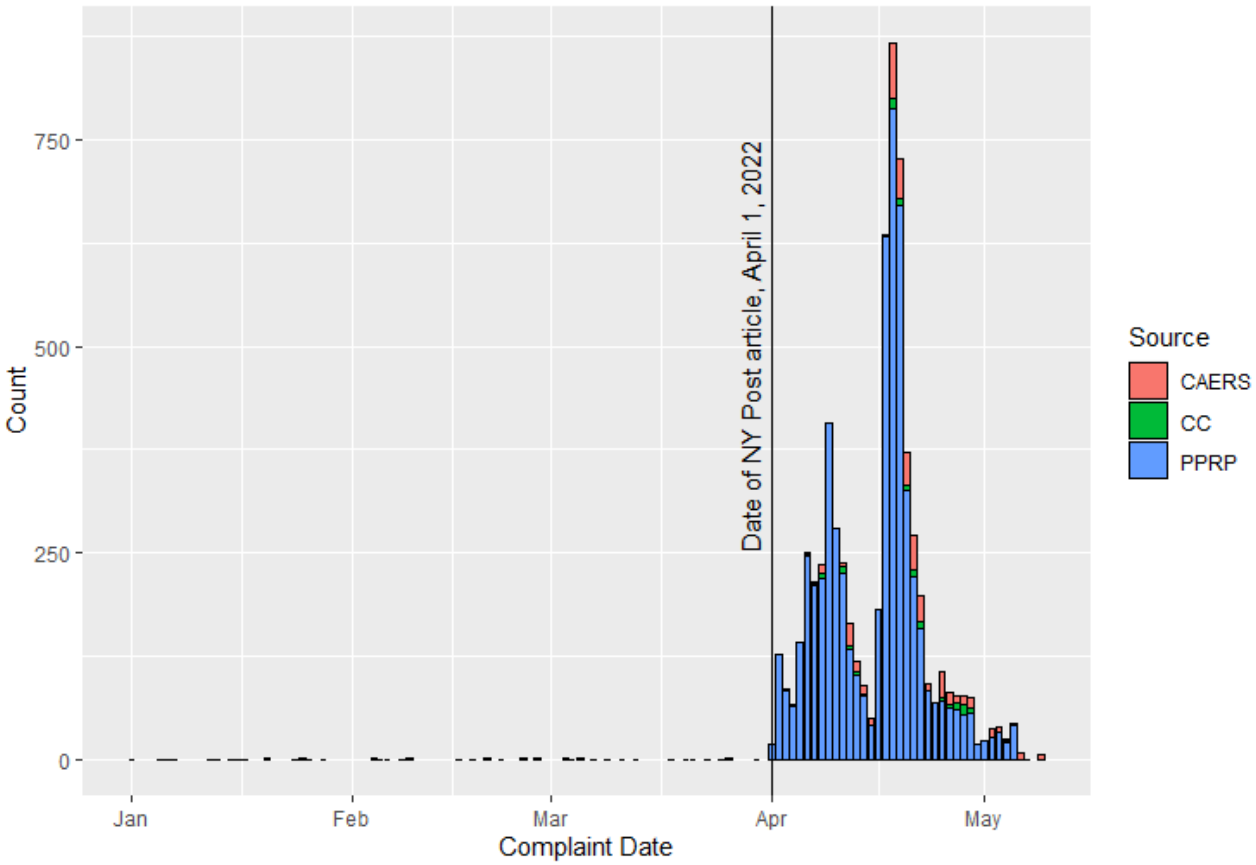
XSL•FO
RenderX

epidemiologic information such as the date the person first reported feeling ill, food exposure date, and illness duration.

Thus, the investigation instead focused on reports directly submitted to the FDA. FDA consumer complaint reports collected between June 1, 2021, and May 10, 2022, that mentioned exposure to the dry breakfast cereal product of interest and explicit gastrointestinal (GI) symptoms were reviewed. Only one of the 120 consumer complaint reports submitted between June 1, 2021, and May 10, 2022, was received in 2021. The remaining 119 were submitted after April 1, 2022, the day the NY Post article was published. Of these 120 consumer complaints, many reported GI symptoms.

Between June 1, 2021, and May 10, 2022, there were a total of 438 Center for Food Safety and Applied Nutrition Adverse Event Reporting System (CAERS) reports associated with the dry breakfast cereal product, 437 of which were submitted after April 1, 2022 (Figure 2). The FDA conducted an analysis of 406 CAERS reports that contained suspected cases with exposure to the breakfast cereal, explicit GI symptoms, and were received prior to April 30, 2022. Of 406 CAERS reports, only 160 had information on illness onset, most of which suggested symptom onset occurred 1‐24 hours after consumption. Overall, 558 FDA adverse event reports (120 consumer complaint and 438 CAERS reports) and 6272 PPRP reports of GI illness were associated with this incident.

Figure 2. Distribution of adverse event reports received by various reporting platforms concerning dry breakfast cereal (2022). CAERS: Center for Food Safety and Applied Nutrition Adverse Event Reporting System; CC: Food and Drug Administration consumer complaints; NY: New York; PPRP: private participatory reporting platform.



In response to illness reports related to the dry breakfast cereal product, the FDA conducted preventive control inspections at 2 manufacturing facilities that produced the dry breakfast cereal of interest and at 1 product supplier in order to document manufacturing operations and identify potential deficiencies. An FDA investigation was also conducted at the corporate offices of the dry breakfast cereal brand in order to collect records pertaining to manufacturing processes, facilities, products, suppliers, quality control practices, and consumer complaint data. Concurrently, the Michigan Department of Agriculture and Rural Development initiated a preventive control inspection at another supplier to the brand. No significant observations were found during these inspections. FDA chemical and microbial analyses were completed on product and environmental samples. FDA and state partners also analyzed clinical and consumer product samples. However, no causative agent for this incident was identified. Evidence from this investigation indicated that this was, in fact, not a microbial outbreak related to dry cereal.

Ketogenic Shakes

In March 2022, a PPRP shared with the FDA an analysis of complaints they received about Brand A protein or meal replacement shakes, a generic brand produced for Retailer A. The analysis demonstrated that there was an increase in complaints associated with Brand A when compared to other protein or meal replacement shake brands starting in January 2021. The PPRP provided all complaints to the FDA for Brand A protein or meal replacement shakes submitted between

XSL•FO
RenderX

January 2021 and March 2022. A review of this information found that 28 of 32 reports specifically identified the ketogenic shake from Brand A, and the FDA initiated an investigation into this product. During the investigation, 6 additional complaints were received, culminating in a total of 34 complaints specifically identifying Brand A ketogenic shakes reported to the PPRP between January 2, 2021, and March 30, 2022. These PPRP reports were challenging to use as they lacked complete epidemiologic information, and specific product details, such as lot numbers, were not consistently available. Further, some of the data that was provided was unstructured and required manual review and extraction.

Additionally, reviews for the ketogenic shake product were extracted on March 24, 2022, from Retailer A's website. This data also required manual review and data extraction. Between January 9, 2021, and February 2, 2022, 54 reviews contained text with complaints indicating an adverse event. Finally, FDA systems were queried for complaints related to the Brand A ketogenic shake over the previous 2 years. Between June 6, 2021, and April 8, 2022, 2 consumer complaints and 4 CAERS reports were received by the FDA. Complaints from all 3 data sources (FDA, consumer reviews, and PPRP complaints) were compiled, and a total of 93 were reviewed. Symptoms were varied but included mostly GI distress, and many reported a rapid onset of symptoms (<30 min). The number of reports received for Brand A was compared to the number for Brand B ketogenic shake to ascertain whether the amount of complaints was unusual. The Brand B ketogenic shake had nearly equivalent product labeling to the Brand A ketogenic shake, but the 2 brands did not share a manufacturer. Both brands were nationally distributed in the United States. The FDA CAERS database and the PPRP were also queried for complaints that specified Brand B during the same time frame. These Brand B queries returned no complaints from the PPRP and 1 complaint from FDA CAERS.

Closed retail samples of 2 flavors of Brand A ketogenic shakes and open and closed samples from a complainant's home were collected. Chemical and sensory analyses were completed on all product samples collected by the FDA, and no unusual findings were identified. No public communications or regulatory actions resulted from this investigation. However, details of the investigation were shared with Retailer A, who had already discontinued distribution of the product based on poor web-based reviews and who further, in response to FDA notification, made the remaining product in their stores unavailable for purchase. Ultimately, laboratory testing did not detect any specific undeclared ingredients or toxins as a causative agent for these adverse events, but the investigation did highlight potential consumer intolerances to some declared ingredients used in ketogenic food products.

Plant-Based Crumble Product

In June 2022, the FDA received consumer complaints and adverse event reports referencing GI illness suggesting abnormal liver function (eg, jaundice and dark urine) that may have been attributable to the consumption of a plant-based crumble product produced by Brand C. As of June 23, 2022, 20 CAERS reports and 13 consumer complaints had been evaluated. Although some complainants reported several Brand C products, all complaints specified the plant-based crumble product. Most reported GI and liver-associated symptoms. This incident became heavily publicized as reports pertaining to the adverse events were reported digitally. Additionally, consumers reported adverse health outcomes to various other sources, such as state reporting mechanisms, a PPRP, and directly to Brand C. As of October 18, 2022, this incident included 393 (177 consumer complaints and 216 CAERS reports) adverse event reports in 39 states with 133 hospitalizations and zero deaths. Estimated illness onset dates ranged from April 19, 2022, to September 4, 2022.

Closed and open samples of Brand C plant-based crumbles product were sampled from complainants' homes. Closed and undistributed product samples were also obtained from Brand C's warehouses. Both chemical and microbiological analyses were performed, but there were no significant findings. However, in 2024, the FDA made a public posting noting that tara flour, an ingredient used in this product, did not meet the Generally Recognized as Safe standard, so it was an unapproved food additive [22]. The FDA issued a web post on the investigation and posted the company recall on the FDA recalls page [23,24]. The FDA further performed several domestic inspections related to the incident and 1 foreign inspection, though none of these actions provided evidence of contamination.

During this investigation, the FDA also reached out to the PPRP to examine any reports they had received regarding this product. The website had received 24 complaints associated with Brand C products from April 4, 2022, to October 18, 2022, 7 of which mentioned the product of interest (plant-based crumbles). Of these 7 complaints, 4 reported GI symptoms and 3 reported various liver-associated ailments. Records from the PPRP were obtained to rule out other vehicles of interest, such as other Brand C products and products from a separate company, Brand D. Although 8 other products were mentioned, none were mentioned more than twice. Further, the PPRP had only received a single complaint for Brand D, which described nausea and vomiting. Based on these reports and the data from FDA systems, other Brand C and Brand D products were excluded as not likely vehicles of illness for this incident.

Discussion

Overview

The use of NDSs to identify adverse event signals or supplement FDA investigations of foodborne illness clusters may offer opportunities such as earlier detection of clusters and vehicle confirmation as seen in the case studies described earlier. However, data collection from NDSs also presents challenges that may reduce their utility during ongoing incident investigations. The advantages and challenges associated with the use of NDS data are discussed in the following section, along with a summary of future steps to better harmonize public health efforts.

XSL•FO
RenderX

Advantages

Traditional, laboratory-based methods of foodborne disease surveillance are largely reliant on the detection of foodborne illnesses via information collected at the clinical level, such as laboratory tests and health care provider reports [25]. These signals are dependent on ill persons seeking medical care, yet some research indicates that fewer than 15% of individuals with self-suspected foodborne illness seek medical attention or report to an LHD [25]. The most recent Foodborne Disease Active Surveillance Network (FoodNet) population survey supports this, noting that only 16.9% of respondents with an acute GI illness sought medical care [26]. Further, most episodes of foodborne illness are self-limiting, and only a small percentage of illnesses in the United States are estimated to result in hospitalization [2]. Once a case of foodborne illness is detected, associated epidemiologic information is gathered by trained epidemiologists, which ensures collected information is detailed and accurate. However, national mandatory reporting of foodborne illnesses is limited to a select set of diseases [27]. While this list includes most well-established foodborne pathogens, it is not as useful for illnesses without a known etiology. Notification of the public health system to new illnesses is also slow, as lag times often range from 2 to 4 weeks or more, depending on the etiology, time of year, and laboratory testing capability. As discussed, some of these limitations are mitigated by various consumer complaint-based systems that are not limited by specific etiology and do not require a person to seek medical care. This is particularly important for localized outbreak detection.

NDSs have the potential to address gaps that exist in traditional foodborne illness surveillance methods and empower consumers. For example, data from these systems can alert public health agencies to foodborne illnesses in situations where medical care is not obtained (eg, due to financial reasons, lack of perceived necessity). NDSs may also provide early signals for foodborne illness clusters that would otherwise be slowed due to expected lag times in case reporting or unforeseen delays. Common barriers to traditional reporting, such as cost, time, access, or awareness, may also be avoided by consumers who use NDSs. In the United States, mechanisms to report an adverse event directly to a public health entity vary across the state and federal levels, while NDS companies have varying user-friendly formats for accepting consumer complaints. Therefore, the generation of an adverse event report on a PPRP or a review on a product webpage may be more intuitive, simpler, and less time-consuming for the public in some situations.

The broad reach of NDSs strengthens their ability to catalyze public health investigations. The larger number of associated complaints can also contextualize reported symptoms, which is useful in situations where the pathogen or hazard is emerging and case symptomology is unknown. One example of this is the ketogenic shake investigation, in which complainants described a rapid onset of short-lived symptoms after consuming the product, making it more unlikely for complainants to seek care under these circumstances. Despite the low volume of related complaints reported to the FDA, which typically would not rise to the level of an investigation, an FDA investigation was triggered due to the number of complaints reporting a similar product or symptomology within product reviews and PPRP reports.

A signal received via NDS of a suspected illness related to an FDA-regulated product may allow the FDA to evaluate consumer risk using investigative tools such as traceback if the incoming data includes necessary details about the product. Traceback helps ascertain if illnesses are related to a common source, which leads to further investigation and, potentially, product actions. Details within NDS reports may help to target firm inspections, product sample collection, and product testing, all of which can inform compliance actions. Retailers may also use information from NDSs as an early alert to a poor-quality product or one that may be causing adverse reactions among consumers, such as in the example provided with ketogenic shakes. This unsolicited information allows retailers to internally initiate actions to protect consumers prior to a manufacturer-initiated product recall or the attention of regulatory and public health agencies. Traceback information obtained from NDSs may also help to rule out products not linked to adverse events. Although NDS data can inform industry and regulatory investigations, the source for NDS reports may be anonymous; therefore, the veracity of the data should be considered during decision-making.

Challenges

An overarching limitation of NDSs stems from the quality of available data. Consumer complaints made to the FDA use forms that incorporate standardized epidemiologic data elements, including those identified by the Council to Improve Foodborne Outbreak Response [27]. These forms also collect details used to assess complaint trends and detect potential food safety issues. NDSs do not reliably capture key data elements essential for public health investigations such as those outlined in Table 1, which includes an expanded list of data elements in addition to those put forth by the Council to Improve Foodborne Outbreak Response. As NDSs often obtain information for other purposes, data gaps are expected and can hinder public health investigations. For example, epidemiologic information is used by investigators to identify a common etiologic agent, which is critical in the absence of clinical validation. Without epidemiologic information, the cause of a food safety issue is more challenging to identify, and the scope of product testing is significantly more resource intensive. Product and exposure information are also necessary to identify a common source of illness and inform traceback investigations. A lack of actionable data inhibits public health response, which was illustrated during the 3 investigations described here.

Another limitation associated with NDS systems is that the data are often unstructured. Information from complainants is typically in the form of free-text entry (eg, social media posts, product reviews, web-based comments), which currently require manual review to extract key data elements. This process is resource-intensive, subject to human error, and has limited feasibility for large numbers of posts or comments. Automated data review may reduce the burden of manual review in the future, but specialized tools are expensive investments for ad hoc analyses, and varying data sources may necessitate different technical approaches. Similar to traditional outbreak data

XSL•FO
RenderX

sources, NDS reports may also be duplicated across reporting platforms, as individuals may submit complaints on multiple NDS platforms, multiple complaints on one platform, and directly to the FDA (Figure 1). Some NDS platforms do minimize report duplication with requirements for users to sign into an account, to have verified purchases to leave a review, or by tracking the IP address of the user. However, evaluation of complaint uniqueness for enumeration across platforms may require manual deduplication or simply not be possible.

Table . List of key data elements for actionable public health response for collection by novel data systems, expanded and modified from the Council to Improve Foodborne Outbreak Response [27].

Data type	Data element
Epidemiologic	• Illness onset date and time • Symptoms • Duration of symptoms • Care-seeking behavior • Suspected exposure of interest • Exposure date • Other recent exposures (past 7 days)
Product	• Photo of product packaging • Brand • Best buy date or best if used by date • Lot code • Product description on label • Purchase location and purchase date • Manufacturer, distributor information
Complainant's contact information	• Name • Phone number • Email address
Affected individual's contact information (if not the same as the complainant's)	• Name • Phone number • Email address
Affected individual's demographic information	• Age • Gender • State of affected individual's residence
Method of report	• Date of complaint • Type of platform • Method of complaint

Ultimately the structure, breadth, and utility of information available in an NDS are dependent on the purpose of the platform. Certain NDSs, such as web-based reviews and social media posts, share information publicly, though not explicitly for public health goals. Due to this, food safety complaints made on these platforms usually lack critical data elements but have the potential to provide contextual information that can catalyze or supplement an investigation. PPRPs do have a public health goal, and compared to official government complaint systems, they may be easier to find and use, which makes them attractive options to complainants. However, although these sites may be easier to use, the data captured may be incomplete or of lower quality than data collected from traditional complaint systems. Public health and regulatory officials must keep these limitations in mind when considering NDSs in an investigational context.

When a signal for an adverse event is received, public health investigators review the reported symptoms, in conjunction with laboratory diagnoses, to determine an etiology and define cluster criteria (ie, develop a case definition) [27]. With NDSs, limited epidemiologic information may result in nonspecific symptomology. Alternatively, investigators may rely on medical records and laboratory documentation to accumulate evidence in search of a possible etiology, but complainant contact information is not often included in NDS reports, which prevents investigators from obtaining follow-up information. Thus, reports from NDSs must be grouped based on available epidemiologic features, which results in a broad case definition, extensive record requests, and resource-intensive review by investigators. Therefore, biological plausibility and medical review should be considered when refining a case definition and evaluating likely exposures when using NDSs.

Evidence has shown that media attention, particularly TV news coverage, can lead to an increase in adverse event reporting during a health scare [28,29]. This would also affect reporting to NDSs, which could see greater impacts given their ease of reporting, particularly if an NDS uses tools such as social media to share information on the incident. This effect was illustrated during the dry cereal incident, as complaints to both NDSs and FDA increased dramatically after newspaper reporting (Figure 2). One possible further complication of media attention is the amplification of unreliable and spurious signals that leads to

XSL•FO
RenderX

stress on public health response entities, who already have limited resources.

Many actors are involved in multistate investigations, each with unique roles and challenges. One of these challenges is identifying proper medical and epidemiologic experts to evaluate data collected from PPRP or other NDS sources. CDC, for example, has investigational programs aligned to specific pathogens, so it can be difficult to identify which group would provide expertise during an investigation involving an unknown etiology.

As NDS platforms are not managed by established public health systems, information sharing with entities that have authority to take public health action is limited. Although some public health agencies may partner with third-party NDS platforms, subscribe to their services, or survey their data, NDS data may not be automatically escalated to a public health authority, which may not always be clear to the public when submitting their report. Also, some NDS companies may require public health agencies to pay to access complaint data. In addition, NDS platforms vary in the user information they collect, which can limit the ability to conduct follow-up with ill persons. It is also important to note that these NDS systems may divert ill people away from submitting reports to established public health reporting and surveillance systems. People who report their illnesses on these NDS platforms may assume that their reports will automatically be shared with authorities, which is not the case. In addition, submitting complaints to NDS systems may give people the false expectation that the information they provide will be actionable by public health agencies.

One reason that laboratory results can be confirmed relatively quickly in traditional outbreaks is that the etiological agent is known, and multiple laboratories are often able to perform and analyze testing results. Conversely, NDS-detected outbreaks often have an unknown etiological agent, which requires broad, instead of targeted, testing and often takes longer to return results.

Similar to limitations discussed regarding epidemiologic data, information collected from consumers through NDSs may lack critical elements needed for product traceback, such as product and purchase information. Although photos can be shared of a product, photos and product descriptions obtained from consumers do not consistently provide information on lot numbers, purchase dates, or purchase locations. Using a phone to submit complaints, while convenient, may limit the amount of information that people are able to submit. This information can aid regulatory authorities to conduct traceback, as some products may have the same name or label but have been produced or grown at different locations. Purchase information may also inform traceback investigations through identification of the manufacturer; however, these details are not consistently captured from NDSs or from traditional outbreak data sources. Ultimately, the lack of consistent product and purchase information may hinder the identification of a specific implicated product for further industry or regulatory action.

Opportunities

While NDSs have been helpful at the local level in identifying establishment-based illness clusters or targeting limited resources for inspections, foodborne investigations at the federal level focus on multistate outbreaks or vehicles that enter interstate commerce [7,20,28]. In these situations, etiologic agents are generally not related to ill workers or conditions at a retail establishment, but rather issues further up the supply chain. Similar to traditional outbreak data sources, NDSs reports may be of assistance during FDA investigations to identify signals related to branded or packaged products that are easily identified by the consumer, as opposed to loose items, such as produce, with no clear or easily identifiable brand or manufacturer. NDSs are also most likely to identify incidents that involve a rapid onset of symptoms, as with the ketogenic shakes investigation, or unusual symptoms, as in the plant-based crumble product investigation.

Signal detection and data collection through NDSs will likely become more prevalent. Therefore, it is incumbent upon public health agencies to determine the best ways to use this data, integrate it into established signal detection systems, and potentially collaborate with platform hosts. There is also a need to establish a process, once a signal is detected via NDS, where ill individuals can be interviewed with a standardized questionnaire that can be used to further evaluate potential exposures and etiologies. NDS providers should consider adding the noted critical data elements (Table 1) to existing data captured by their sites or provide public health officials with contact information to follow up with complainants. CDC and state health departments or LHDs can play an important role in coordinating patient interviews and epidemiological data collection, including medical records. In the future, the further development of tools for natural language processing, website scraping, and optical character recognition technology (software to extract text from an image) may also be useful when applied to NDSs for signal detection or to supplement ongoing outbreak investigations.

Conclusions

NDSs may enable quicker signal detection in certain situations compared to traditional surveillance methods for foodborne illness clusters attributable to unknown agents. Recently, NDSs have provided supplementary data during traditional investigations and sparked investigations into previously undetected signals, which may or may not have provided enough information for an actionable public health response. While PPRPs and product reviews are the focus in this paper, future signals may come from other NDSs, such as web-based social media platforms. To effectively integrate these data into public health surveillance and response, working protocols could be developed to help agencies better collect, surveil, and assess this data. NDSs with a focus on public health should also do their best to collect critical epidemiologic, product, and clinical data, enhancing their utility as a resource for public health agencies.

XSL•FO
RenderX

Acknowledgments

The response efforts discussed in this manuscript included numerous public health officials at local and state health and regulatory departments and public health laboratories in the United States, who serve as the backbone of any multi-state foodborne illness outbreak investigation. We also appreciate the assistance provided by US Food and Drug Administration (FDA) Emergency Response Coordinators who coordinated the response between FDA and state partners. The findings and conclusions of this report are those of the authors and do not necessarily represent the official position of the FDA and the Centers for Disease Control and Prevention (CDC).

Conflicts of Interest

None declared.

References

1. Scharff RL. Economic burden from health losses due to foodborne illness in the United States. J Food Prot 2012 Jan;75(1):123-131. [doi: 10.4315/0362-028X.JFP-11-058] [Medline: 22221364]
2. Scallan E, Griffin PM, Angulo FJ, Tauxe RV, Hoekstra RM. Foodborne illness acquired in the United States--unspecified agents. Emerg Infect Dis 2011 Jan;17(1):16-22. [doi: 10.3201/eid1701.091101p2] [Medline: 21192849]
3. Seelman S, Viazis S, Merriweather SP, Cloyd TC, Aldridge M, Irvin K. Integrating the Food and Drug Administration Office of the Coordinated Outbreak Response and Evaluation Network's foodborne illness outbreak surveillance and response activities with principles of the National Incident Management System. J Emerg Manag 2021;19(2):131-141. [doi: 10.5055/jem.0567] [Medline: 33954963]
4. Wellman A, Bazaco MC, Blessington T, et al. An overview of foodborne sample-initiated retrospective outbreak investigations and interagency collaboration in the United States. J Food Prot 2023 Jun;86(6):100089. [doi: 10.1016/j.jfp.2023.100089] [Medline: 37024093]
5. Kim TN, Edmundson AR, Hedberg CW. Foodborne Illness Complaint Systems Detect, and Restaurant Inspection Programs prevent restaurant-associated foodborne illness outbreaks. Foodborne Pathog Dis 2024 Feb;21(2):92-98. [doi: 10.1089/fpd.2023.0086] [Medline: 38010953]
6. Li J, Shah GH, Hedberg C. Complaint-based surveillance for foodborne illness in the United States: a survey of local health departments. J Food Prot 2011 Mar;74(3):432-437. [doi: 10.4315/0362-028X.JFP-10-353] [Medline: 21375880]
7. Li J, Smith K, Kaehler D, Everstine K, Rounds J, Hedberg C. Evaluation of a statewide foodborne illness complaint surveillance system in Minnesota, 2000 through 2006. J Food Prot 2010 Nov;73(11):2059-2064. [doi: 10.4315/0362-028x-73.11.2059] [Medline: 21219718]
8. Yousaf A, Viveiros B, Caron G. Rhode Island Department of Health Foodborne Illness Complaint System: a descriptive and performance analysis. J Food Prot 2019 Sep;82(9):1568-1574. [doi: 10.4315/0362-028X.JFP-19-135] [Medline: 31429595]
9. Integrated Food Safety Centers of Excellence. URL: https://foodsafetycoe.org [accessed 2025-02-25]
10. Punzalan C, Budnitz DS, Chirtel SJ, et al. Swallowing problems and dietary supplements: data from U.S. Food and Drug Administration Adverse Event Reports, 2006-2015. Ann Intern Med 2019 Nov 19;171(10):771-773. [doi: 10.7326/M19-0947] [Medline: 31426091]
11. MedWatch online voluntary reporting form. US Food and Drug Administration. 2024. URL: https://www.accessdata.fda.gov/scripts/medwatch/index.cfm [accessed 2025-02-25]
12. Safety Reporting Portal. 2024. URL: https://www.safetyreporting.hhs.gov/SRP2/en/Home.aspx?sid=09abe76b-57ab-49a2-9ff9-22cd52eb8e32 [accessed 2025-02-25]
13. Consumer complaint coordinators. US Food and Drug Administration. 2024. URL: https://www.fda.gov/safety/report-problem-fda/consumer-complaint-coordinators [accessed 2025-02-25]
14. About the CORE network. US Food and Drug Administration. 2024. URL: https://www.fda.gov/food/outbreaks-foodborne-illness/about-core-network [accessed 2025-02-25]
15. FDA at a glance: regulated products and facilities. US Food and Drug Administration. 2024. URL: https://www.fda.gov/media/115824/download#:~:text=FDA%20regulates%20about%2075%20percent%20of%20the%20U.S.%20food%20supply [accessed 2025-02-25]
16. Scallan E, Hoekstra RM, Angulo FJ, et al. Foodborne illness acquired in the United States--major pathogens. Emerg Infect Dis 2011 Jan;17(1):7-15. [doi: 10.3201/eid1701.p11101] [Medline: 21192848]
17. White AE, Tillman AR, Hedberg C, et al. Foodborne illness outbreaks reported to National Surveillance, United States, 2009-2018. Emerg Infect Dis 2022 Jun;28(6):1117-1127. [doi: 10.3201/eid2806.211555] [Medline: 35608555]
18. Deng X, Cao S, Horn AL. Emerging applications of machine learning in food safety. Annu Rev Food Sci Technol 2021 Mar 25;12:513-538. [doi: 10.1146/annurev-food-071720-024112] [Medline: 33472015]
19. Harris JK, Mansour R, Choucair B, et al. Health department use of social media to identify foodborne illness - Chicago, Illinois, 2013-2014. MMWR Morb Mortal Wkly Rep 2014 Aug 15;63(32):681-685. [Medline: 25121710]

XSL•FO
RenderX

20. Harrison C, Jorder M, Stern H, et al. Using online reviews by restaurant patrons to identify unreported cases of foodborne illness - New York City, 2012-2013. MMWR Morb Mortal Wkly Rep 2014 May 23;63(20):441-445. [Medline: 24848215]
21. Fickenscher L. Lucky Charms cereal causing vomiting, diarrhea: report. New York Post. URL: https://nypost.com/2022/04/01/lucky-charms-cereal-causing-vomiting-diarrhea-report/ [accessed 2025-02-25]
22. FDA update on the post-market assessment of tara flour. US Food and Drug Administration. 2024. URL: https://www.fda.gov/food/hfp-constituent-updates/fda-update-post-market-assessment-tara-flour [accessed 2025-02-25]
23. Daily harvest issues voluntary recall of french lentil + leek crumbles due to potential health risk. US Food and Drug Administration. 2022. URL: https://www.fda.gov/safety/recalls-market-withdrawals-safety-alerts/daily-harvest-issues-voluntary-recall-french-lentil-leek-crumbles-due-potential-health-risk [accessed 2025-02-25]
24. Investigation of adverse event reports: French lentil & leek crumbles (June 2022). US Food and Drug Administration. 2022. URL: https://www.fda.gov/food/outbreaks-foodborne-illness/investigation-adverse-event-reports-french-lentil-leek-crumbles-june-2022 [accessed 2025-02-25]
25. Fein SB, Lin CTJ, Levy AS. Foodborne illness: perceptions, experience, and preventive behaviors in the United States. J Food Prot 1995 Dec;58(12):1405-1411. [doi: 10.4315/0362-028X-58.12.1405] [Medline: 31159043]
26. Foodborne diseases active surveillance network - population survey. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/foodnet/surveys/population.html2024 [accessed 2025-02-25]
27. CIFOR Guidelines for Foodborne Disease Outbreak Response, 3rd edition: The Council to Improve Foodborne Outbreak Response (CIFOR); 2022. URL: https://cifor.us/downloads/clearinghouse/CIFOR-Guidelines-Complete-third-Ed.-FINAL.pdf [accessed 2025-02-25]
28. Faasse K, Gamble G, Cundy T, Petrie KJ. Impact of television coverage on the number and type of symptoms reported during a health scare: a retrospective pre-post observational study. BMJ Open 2012;2(4):e001607. [doi: 10.1136/bmjopen-2012-001607] [Medline: 22904334]
29. Philen RM, Kilbourne EM, McKinley TW, Parrish RG. Mass sociogenic illness by proxy: parentally reported epidemic in an elementary school. Lancet 1989 Dec 9;2(8676):1372-1376. [doi: 10.1016/s0140-6736(89)91976-4] [Medline: 2574312]

Abbreviations

CAERS: Center for Food Safety and Applied Nutrition Adverse Event Reporting System
CFSAN: Center for Food Safety and Applied Nutrition
CORE: Coordinated Outbreak Response & Evaluation
FDA: Food and Drug Administration
GI: gastrointestinal
LHD: local health department
NDS: novel data stream
PPRP: private participatory reporting platform

Edited by A Mavragani; submitted 26.03.24; peer-reviewed by AE White, C Hedberg, ES Walter, K Pogreba-Brown; revised version received 12.11.24; accepted 13.11.24; published 28.02.25.

Please cite as:
Bazaco MC, Carstens CK, Greenlee T, Blessington T, Pereira E, Seelman S, Ivory S, Jemaneh T, Kirchner M, Crosby A, Viazis S, van Twuyver S, Gwathmey M, Malais T, Ou O, Kenez S, Nolan N, Karasick A, Punzalan C, Schwensohn C, Gieraltowski L, Chen Parker C, Jenkins E, Harris S
Recent Use of Novel Data Streams During Foodborne Illness Cluster Investigations by the United States Food and Drug Administration: Qualitative Review
JMIR Public Health Surveill 2025;11:e58797
URL: https://publichealth.jmir.org/2025/1/e58797
doi:10.2196/58797

© Michael C Bazaco, Christina K Carstens, Tiffany Greenlee, Tyann Blessington, Evelyn Pereira, Sharon Seelman, Stranjae Ivory, Temesgen Jemaneh, Margaret Kirchner, Alvin Crosby, Stelios Viazis, Sheila van Twuyver, Michael Gwathmey, Tanya Malais, Oliver Ou, Stephanie Kenez, Nichole Nolan, Andrew Karasick, Cecile Punzalan, Colin Schwensohn, Laura Gieraltowski, Cary Chen Parker, Erin Jenkins, Stic Harris. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete

XSL•FO
RenderX

bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Scaling Up and Enhancing the Functionality of the Electronic Integrated Diseases Surveillance and Response System in Uganda, 2020-2022: Description of the Journey, Challenges, and Lessons Learned

Rodney Mugasha[1], MSc; Andrew Kwiringira[1], MPH; Vivian Ntono[1], MPH; Lydia Nakiire[1], MPH; Immaculate Ayebazibwe[2], MSc; Caroline Kyozira[3], MSc; Allan Niyonzima Muruta[3], MPH; Juliet Namugga Kasule[4], MSc; Dathan M Byonanebye[1], PhD; Judith Nanyondo[1], MPH; Richard Walwema[1], MBA; Francis Kakooza[1], PhD; Mohammed Lamorde[1], PhD

[1]Infectious Diseases Institute, College of Health Sciences, Makerere University, McKinnel Knowledge Center, 2 Hall Lane, Kampala, Uganda

[2]Health Information Systems Program Uganda, Kampala, Uganda

[3]Integrated Epidemiology, Surveillance, and Public Health Emergencies Department, Ministry of Health Uganda, Kampala, Uganda

[4]Division of Global Health Protection, Center for Global Health, Centers for Disease Control and Prevention, Kampala, Uganda

Corresponding Author:
Rodney Mugasha, MSc
Infectious Diseases Institute, College of Health Sciences, Makerere University, McKinnel Knowledge Center, 2 Hall Lane, Kampala, Uganda

Abstract

In 2017, Uganda implemented an electronic Integrated Disease Surveillance and Response System (eIDSR) to improve data completeness and reporting timelines. However, the eIDSR system had limited functionality and was implemented on a small scale. The Ministry of Health, with support from the Infectious Disease Institute, Makerere University, and Health Information Systems Program Uganda, upgraded the system functionality and scaled up its implementation. This study describes the process and impact of upgrading eIDSR functionality and expanding its implementation across additional districts. The Ministry of Health, through its Integrated Epidemiology, Surveillance & Public Health Emergency Department, coordinated the implementation of the eIDSR. User requirements were identified through consultations with national surveillance stakeholders. The feedback informed the design and development of the upgraded eIDSR functionalities. The eIDSR rollout followed a consultative workshop to create awareness of the system among stakeholders. A curriculum was developed, and a national training of trainers was conducted. These trainers cascaded the training to the district health teams, who later cascaded the training to health workers. The training adopted an on-site training approach, where a group of national or district trainers would train new users at their desks. The eIDSR system was upgraded to the District Health Information Software 2 (DHIS2) 2.35 platform featuring faster reading and writing tracker data, handling over 100 concurrent users and enhanced case-based surveillance features on Android and web platforms. From October 2020 to September 2022, the eIDSR was rolled out in 68% (100/146) of districts. Additionally, the system permitted prompt reporting of signals of epidemic-prone diseases. Improving the functionality and the expanded geographical scope of the eIDSR system enhanced disease surveillance. Stakeholder commitment and leveraging existing structures will be needed to scale up eIDSR.

(JMIR Public Health Surveill 2025;11:e59783) doi:10.2196/59783

KEYWORDS

electronic Integrated Disease Surveillance and Response; eIDSR; disease surveillance; training of trainers; Uganda; digital surveillance systems; health worker; eHealth; public health; digital health

Introduction

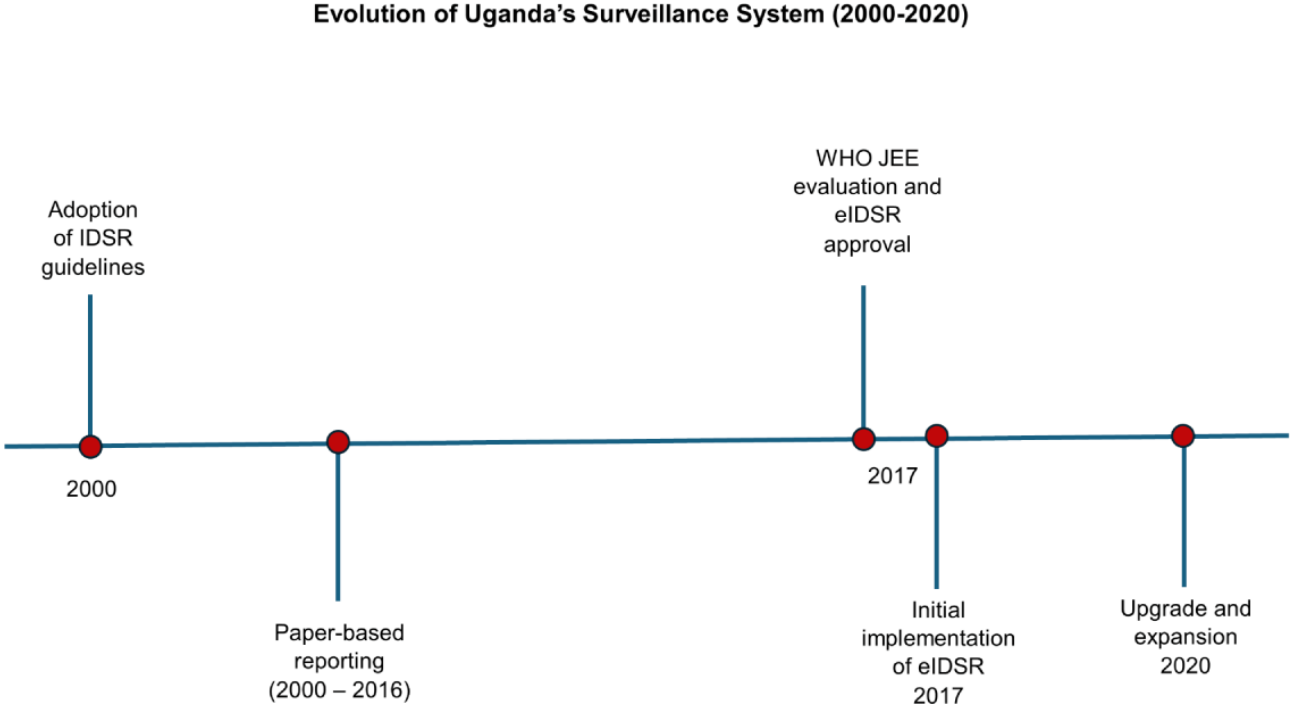
Uganda is highly susceptible to public health emergencies due to its proximity to the ecologically diverse and biologically rich tropical Congo Basin, recurrent epidemic-prone outbreaks, and refugee inflow [1]. In the last 5 years before 2024, Uganda has faced several significant outbreaks, including Ebola and Marburg virus diseases, cholera, measles, and Crimean-Congo hemorrhagic fever, highlighting the ongoing threat of infectious diseases [2]. Before 2000, surveillance systems were not robust enough to detect outbreaks early, leading to delayed responses and exacerbated public health impacts. For instance, the 2000 Ebola outbreak in Gulu, which resulted in 224 deaths,

XSL•FO
RenderX

underscored the inadequacies in early warning and response mechanisms at the time [3].

Uganda embraced the Integrated Disease Surveillance and Response (IDSR) technical guidelines in 2000 [2]. However, the IDSR system was largely paper based. Reports from health facilities were transmitted physically to the district-by-district surveillance focal persons. The process of transporting paper-based reports often took several days to weeks, depending on the distance between the facilities and district offices, the availability of transportation, and other logistical challenges. This manual process led to delayed data transmission and response [2,4]. The World Health Organization (WHO) Joint External Evaluation conducted in 2017 highlighted the fragmentation and lack of coordination between human health, animal health, and laboratory data systems, which hindered effective surveillance of zoonotic diseases [3]. Additionally, the system faced challenges with data quality, including errors and inconsistencies due to manual data entry, as well as inadequate infrastructure, such as unreliable internet and computer systems, which further limited the effectiveness of the surveillance system [3]. In light of these findings, the Ministry of Health (MOH) approved the implementation of the electronic Integrated Disease Surveillance and Response (eIDSR) System in 2017 (Figure 1) to improve the timeliness of reporting; enhance data accuracy and quality; facilitate integration of data across human health, animal health, and laboratory sectors; and improve the capacity for data analysis and decision-making in public health interventions [5].

Figure 1. Road map of Uganda's surveillance system since 2000. eIDSR: electronic Integrated Disease Surveillance and Response; IDSR: Integrated Disease Surveillance and Response; WHO JEE: World Health Organization Joint External Evaluation.



The Electronic Integrated Disease Surveillance and Response System

The eIDSR system is based on the electronic application of the principles of the IDSR to facilitate early detection, investigation, and response to public health events [6]. The eIDSR was developed and customized on the District Health Information Software 2 (DHIS2) open-source software platform (version 2.35; Health Information Systems Program [HISP], University of Oslo) [7]. Unlike the conventional paper-based system used in health facilities, which requires data to be recaptured electronically at the district level, the eIDSR facilitates the instantaneous reporting of notifiable diseases through its SMS, Android, and web platforms [8]. The Android platform had an offline data entry functionality that allowed case-based notification and registration. The SMS platform facilitated immediate notification through a toll-free code currently used by the Uganda MOH. The application could also save data offline and automatically submit it as soon as an internet connection was available [9]. While improving paper-based reporting, the eIDSR system still faced challenges in ensuring that the data collected were complete and accurate. The eIDSR system facilitated case reporting through electronic case investigation forms for 9 out of 33 priority diseases and conditions. Inconsistencies in data entry and gaps in reporting affected the reliability of the information, which in turn impacted decision-making. The interface was not sufficiently user-friendly, which sometimes led to errors in data entry and delays in reporting. Moreover, the need for advanced data visualization tools made it difficult for public health officials to quickly interpret data trends and make informed decisions.

The enhancements and upgrades to the eIDSR system were therefore justified by the need to address these specific challenges. These upgrades included the addition of electronic case investigation forms for other priority diseases and advanced reporting tools to improve data timeliness and quality by streamlining the data entry process. New data visualization features were also added to enable public health officials to

XSL•FO
RenderX

quickly interpret complex data through charts, graphs, and heat maps, facilitating more informed and timely public health actions. Additionally, enhancements were made to ensure better interoperability with other health information systems, thereby reducing data fragmentation and improving the overall efficiency of the surveillance system. The user interface was also redesigned to be more intuitive and user-friendly, reducing the likelihood of errors and encouraging more consistent use of the system by health care workers.

In October 2020, the Infectious Diseases Institute, Makerere University, in collaboration with the Uganda MOH and the HISP Uganda, upgraded the functionalities and scaled the eIDSR to 100 out of 146 districts. This study describes the process and impact of upgrading eIDSR functionality and expanding its implementation across additional districts.

Methods

Overview

A coordinated approach was used to systematically design the enhancement of functionalities and scale-up of the eIDSR. The process began with gathering user requirements from surveillance stakeholders. These requirements were then translated into technical upgrades. The MOH Integrated Epidemiology, Surveillance, and Public Health Emergencies (IES&PHE) Department led coordination. A cascade training model was then conducted to equip health care workers with the necessary skills.

Coordination

The MOH, through its IES&PHE Department, coordinated the implementation of the eIDSR. The IES&PHE Department provided leadership and oversight throughout the implementation process. By coordinating the efforts of various stakeholders, the department ensured that the system was aligned with the country's health priorities. The MOH coordinated the collection of user requirements to ensure that the eIDSR system addressed the specific needs of end-users, such as surveillance officers and health workers. Additionally, the MOH was responsible for convening national surveillance stakeholder meetings.

User Requirements Gathering

User requirements were identified through consultations with national surveillance stakeholders, including the Ministry of Agriculture, Animal Industry, and Fisheries (MAAIF), implementing partners, as well as regional and district surveillance officers. The requirements were gathered on several aspects: system functionality, such as its ability to capture and transmit data in real-time; data quality, accuracy, completeness, and timeliness of reporting; user adoption; and operational challenges, such as technical issues and resource constraints. The feedback gathered was instrumental in informing the design and development of the upgraded eIDSR functionalities, which aimed to address gaps in the existing system.

System Development and Enhancement

The technical team from the HISP Uganda then translated these requirements into operational enhancements. The eIDSR system was upgraded to the DHIS2 2.35 platform, featuring faster reading and writing tracker data and enhanced case-based surveillance features on Android and web platforms. Key changes included the integration of advanced reporting tools, real-time data visualization features, and enhanced interoperability with other health information systems, such as laboratory and animal health data systems. These upgrades were intended to streamline data collection allowing more concurrent users, improve accuracy, and enable more timely responses to public health threats (Figure 1).

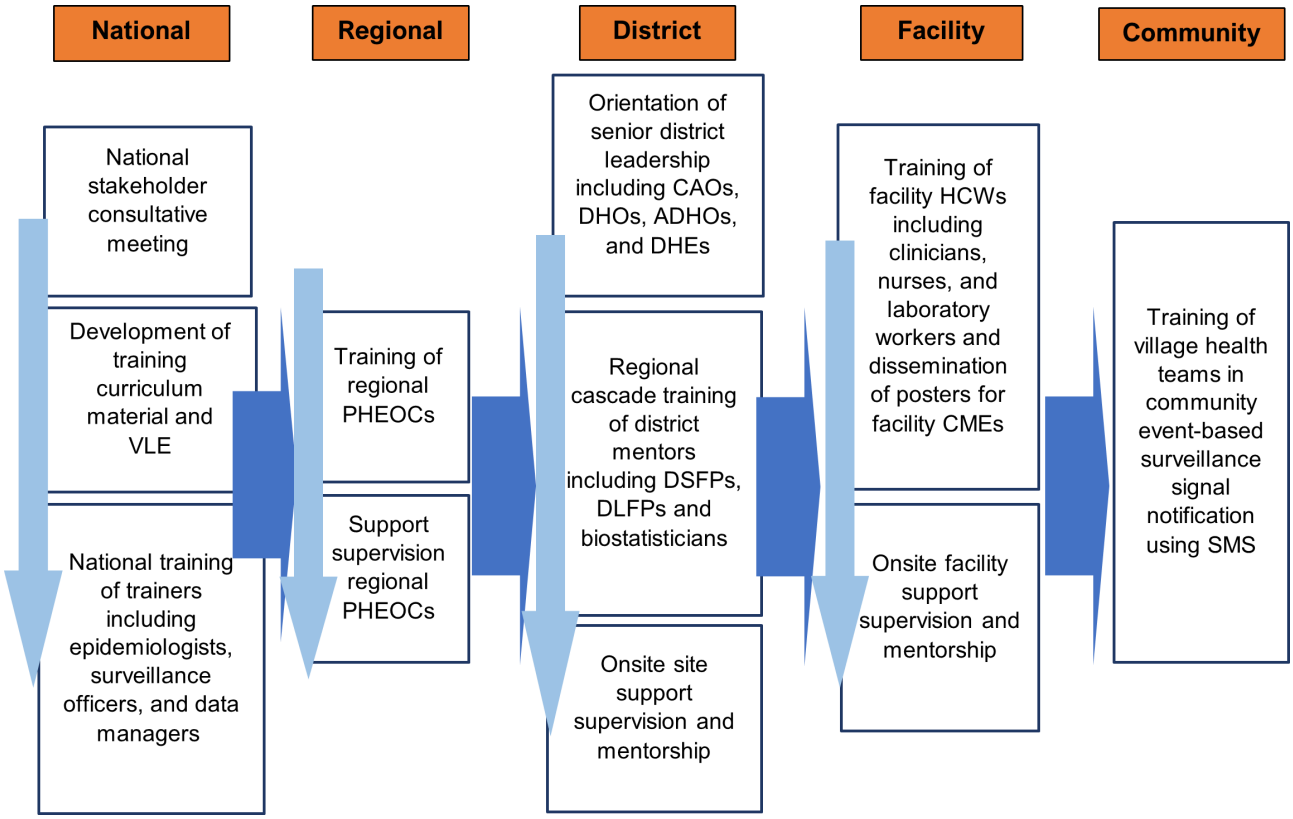
National surveillance stakeholder engagement meetings were regularly convened to obtain user feedback and ensure ongoing coordination and harmonization of the eIDSR rollout. These meetings brought together representatives from the MOH, MAAIF, and other stakeholders to review system performance, address challenges, and plan the expansion of eIDSR to additional districts.

Training

An eIDSR curriculum was developed through a collaborative process involving key stakeholders from the MOH, implementing partners, and subject matter experts in epidemiology, surveillance, and health information systems. The development process began with a needs assessment to identify the specific skills and knowledge gaps among health workers at various levels of the surveillance system. The curriculum was further refined through a series of consultative workshops, where feedback was gathered from national and regional trainers, district health teams, and other stakeholders. This iterative process ensured that the training content was relevant, practical, and aligned with the MOH's surveillance goals. Additionally, the curriculum incorporated IDSR best practices in disease surveillance and reporting, while being tailored to the local context of Uganda. Real-time simulations, case studies, and interactive sessions were added to make the training more engaging and to reinforce learning outcomes. The curriculum focused on alert notification, case notification, registration, and reporting using SMS, Android, and web platforms. The training curriculum also covered the fundamentals of surveillance, such as event-based surveillance and indicator-based surveillance, as well as priority diseases and events for immediate reporting. The rollout model adopted the MOH-recommended approach, where training was cascaded from the national to lower levels (Figure 2).

XSL•FO
RenderX

Figure 2. The electronic Integrated Disease Surveillance and Response (eIDSR) implementation approach in Uganda, 2020-2022. ADHO: assistant district health officer; CAO: chief administrative officer; CME: continuing medical education; DHE: district health educator; DHO: district health officer; DLFP: district laboratory focal person; DSFP: district surveillance focal person; HCW: health care worker; PHEOC: Public Health Emergency Operations Center; VLE: virtual learning environment.



At the national level, national training of trainers was conducted to establish a cohort of national trainers to cascade training to the regional level. This training targeted epidemiologists; health information analysts; laboratory, clinical, monitoring, and evaluation specialists from the MOH; and implementing partners. At the regional level, the national trainers oriented the district health teams on the eIDSR. The district health team comprised of biostatisticians and the Health Management Information System focal persons, district surveillance focal persons, district laboratory focal persons, district veterinary officers, and health subdistrict surveillance focal persons.

At the district level, regional trainers trained clinicians, nurses, laboratory and health information assistants, and facility- and community-based health workers (ie, village health teams) at the facilities. Health workers were selected from high-volume facilities; regional and district hospitals; level 3 health centers that provide a broader range of services including outpatient and inpatient care, basic laboratory services, maternity services, and minor surgeries; and level 4 health centers that provide comprehensive care to larger areas with emergency obstetric care, basic surgeries, laboratory services, and treatment for common diseases. District surveillance focal persons that were trained were given tablets, monthly airtime, and internet data to support case registration and reporting of priority diseases. The trainers subsequently provided post-training mentorship and support supervision to subnational health teams through mentorship to ensure that skills gained during the training were sustained.

System Upgrade and Stakeholder Engagement for eIDSR Enhancement

In 2020, the HISP Uganda developers improved the eIDSR system to the DHIS2 2.35 platform version following the user requirements gathering process. Improvements included the development of an additional 19 electronic case investigation forms for priority diseases, including rabies, anthrax, severe acute respiratory illnesses or influenza-like illnesses like COVID-19, and cholera. These diseases were identified due to their potential for outbreaks, high burden, and emerging threats. The existing eIDSR system lacked comprehensive tools to capture detailed case investigation data. The new forms were developed to address this gap by including essential variables such as exposure history, symptoms, and hospitalization details to enable effective disease tracking of case investigation details, exposure and travel history, symptoms, underlying conditions, and hospitalization.

The laboratory requests and results module were improved to capture laboratory requests for tests to be conducted on the specimen collected for each case. Variables captured include the identification, place, and time (date) of sample collection and the reference and final laboratory results. The improvements in the case outcome module also enabled more accurate tracking of health outcomes and final investigation results. The specimen tracking module was developed to link an existing system that facilitated tracking specimens across the laboratory hubs and laboratories until the completion of the case investigation. The module was developed to send an automated notification via

XSL•FO
RenderX

email and SMS to the reference laboratory and the national disease surveillance team.

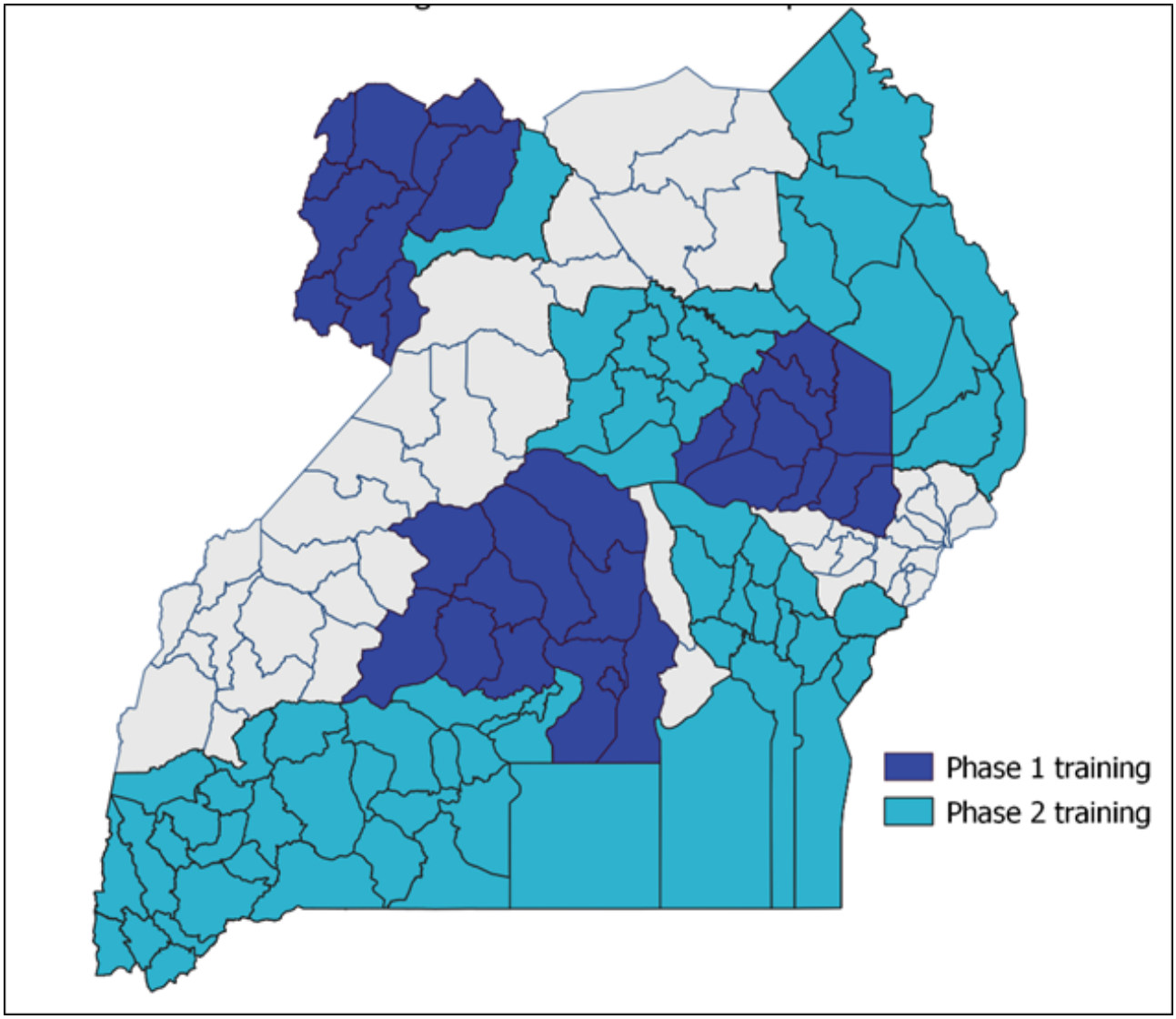
Additional upgrades to the system included enhanced data management and analytics capabilities that included the design of highly customizable forms to track individual-level data to facilitate case-based surveillance with data validation functionality for data quality assurance. Data analysis upgrades included robust charting capabilities to include epidemic curves to facilitate epidemiological data analysis. The upgrade included geographic information systems analysis using thematic mapping, allowing overlay of population, climate, and other layers using the Maps app and supporting geolocation of individual cases. In a bid to enhance situational awareness, a smart web-enabled television at the national level was procured to facilitate access to international news outlets and facilitate data visualization from the eIDSR.

The eIDSR System Adoption and Use

During the 2 years of implementation (October 2020 and September 2021), the eIDSR was scaled up to 68% (100/146) of districts and cities (Figure 3). Thirty national trainers and 191 of the targeted 200 district-level mentors were trained. The district mentors subsequently trained and mentored 2409 of 3295 targeted facility-level health care workers from 536 health facilities. The project also trained another cohort of national-level surveillance officers and field epidemiologists from the Public Health Emergency Operations Center on system navigation, data extraction, analysis, and use for action.

Since its rollout, the eIDSR has continued to be an essential data source for disease surveillance to facilitate the detection and reporting of public health events countrywide. Several events were detected through the eIDSR from SMS messages sent from the community between 2020 and 2022 (Table 1).

Figure 3. Coverage of the electronic Integrated Disease Surveillance and Response (eIDSR) phase 1 and phase 2 training in Uganda, 2020-2022.



XSL•FO
RenderX

Table . Outbreaks detected by the electronic Integrated Disease Surveillance and Response (eIDSR) in Uganda, 2020 - 2022.

Disease outbreak or event	Districts that notified through the eIDSR	Date (year)
Anthrax	Madi-Okollo	2022
Undiagnosed deaths	Kyotera	2021
Rift Valley fever	Kagadi	2020
Anthrax	Bududa and Manafwa	2022
Crimean-Congo hemorrhagic fever	Kagadi	2020
Undiagnosed illness (skin rash)	Pakwach	2020
Rabies	West Nile region: Arua, Adjumani, Moyo, and Yumbe	2020 - 2021
Suspected viral hemorrhagic fever: Rift Valley fever	Moyo, Obongi, Terego, Madi Okollo, and Yumbe	2020 - 2021
Anthrax	Arua and Madi-Okollo	2021 - 2022

Challenges in Implementing the eIDSR

The eIDSR system in Uganda faces several challenges that could threaten the system's sustainability. Data from system logs and user feedback indicate that insufficient technology infrastructure, including reliable internet connectivity and computer hardware, has led to delays in data submission from health care facilities.

Uganda faces significant challenges in its information and communication technology (ICT) infrastructure, impeding the successful implementation of the eIDSR in terms of real-time reporting. Inconsistent internet connectivity and limited access to computer hardware in some rural districts hindered the ability to report cases in real time similarly seen in other countries [10,11]. This was particularly problematic during peak times of disease outbreaks when timely data transmission was crucial for an effective response. It is challenging to operate ICT equipment due to limited access to broadband connectivity and electricity in some rural areas [12]. Another ICT-related challenge was the lack of smartphones among some district surveillance focal persons to register cases in the eIDSR mobile app. The lesson learned is that, for real-time reporting to be successful, ongoing investment in infrastructure, such as reliable internet and adequate hardware, is essential. Additionally, the importance of continuous monitoring and support to maintain data accuracy cannot be overstated.

Health care worker turnover and transfers have significantly impacted the successful implementation of the eIDSR. When trained staff leave, this creates skill shortages and disrupts continuity, causing a knowledge gap similarly seen in the first DHIS2 implementation [4]. This, coupled with the limited ICT skills among health care workers, especially in peripheral facilities, limits the adoption and effective use of the eIDSR. Continuous education and training of users is key to ensuring skill retention, and it is imperative to train many focal persons in each facility.

The inclusion of additional data variables for case investigations presented both an opportunity and a challenge. While the more detailed forms allowed for comprehensive data collection, health care workers found the increased data entry requirements burdensome, particularly in high-volume settings. This sometimes led to incomplete case investigations and delays in submitting reports. To address this, additional training and on-the-job support were provided, emphasizing the importance of thorough and timely data entry. The lesson learned is that there is a need to balance data comprehensiveness with the workload of health care workers to ensure that the system remains user-friendly and efficient.

Opportunities

Despite these challenges, the upgrades to the eIDSR system have provided several opportunities for enhancing disease surveillance and response. The improved system functionality, combined with the lessons learned from the challenges faced, have paved the way for future improvements. For instance, the introduction of mobile data collection tools and offline data entry options could further enhance real-time reporting capabilities, especially in remote areas. Additionally, increased stakeholder engagement and continuous feedback from users will be critical in refining the system and ensuring its long-term sustainability. The pivotal role played by the MOH throughout the implementation process, coupled with the active involvement of district health team trainers, creates a strong possibility of the government taking ownership of eIDSR implementation, independent of external project support.

Discussion

We successfully upgraded eIDSR system functionality and scaled up to more than half of the districts in the country within 2 years of implementation. We also established a cohort of surveillance staff at national and subnational levels who could train health workers across all levels of health care. The implementation of eIDSR was one of the key priority actions in real-time surveillance to address gaps that were recommended from the last WHO Joint External Evaluation conducted in 2017.

The WHO IDSR framework, which strongly emphasizes the necessity for electronic surveillance tools to allow a wide range of crucial surveillance functions, served as a guide for improving the functionality of the eIDSR system [6]. These functions encompass alert notification, case investigation and registration, data quality assurance, real-time reporting, and monitoring and evaluation [9]. The expanded goal of enhancing the national

XSL•FO
RenderX

electronic surveillance system has the potential to benefit significantly from the enhanced functionality of the eIDSR. This broader system intends to seamlessly integrate laboratory tests, notification processes, emergency responses, and reporting mechanisms with indicator-based surveillance, event-based surveillance, and case-based surveillance data across the human health sector. Notably, the eIDSR adheres to open standards and was carefully created on the DHIS2 platform, enabling smooth interchange and connectivity with other software applications and data sources [12]. This strategic alignment puts Uganda on a trajectory for developing interoperability with numerous electronic surveillance systems across multiple industries, in addition to laboratory systems.

A crucial element of the successful rollout of the eIDSR system was the establishment of the curriculum and training of health personnel. Learning is improved by creating a thorough curriculum that incorporates on-site mentoring improvement [13,14]. It was crucial to have onsite supervision and experienced national trainers with knowledge of the eIDSR, public health, and disease monitoring. Maintaining competency requires regular monitoring and evaluation, which includes refresher training. Documenting training records maintains accountability. Sustainability is promoted by incorporating eIDSR training into already-existing health care training programs, and community involvement increases understanding of the system's significance and the responsibilities that health care professionals play in disease monitoring, ultimately improving public health outcomes.

The commitment of the MOH and existing administrative structures provided an opportunity to implement the eIDSR. For example, the division of health information in the MOH is mandated to coordinate initiatives to digitize the health infrastructure [15]. The IES&PHE Department coordinates surveillance and response to public health emergencies. However, several factors documented by unpublished support supervision reports could threaten the sustainability of the system even after enhancement of the system; these include workforce challenges, such as weak ICT infrastructure at the district and health facility level and insufficient funding at the district to conduct surveillance activities. Additionally, Uganda does not have in place interoperable and interconnected electronic reporting for animal health [3]. Data use is lacking as we wait for outbreaks to act rather than use available data to predict specific disease outbreaks, and analysis also remains a challenge.

There are potential limitations to this project. First, the analysis was not intended to be a formal review of the eIDSR platform; we did not assess how the system performed. Thus, additional research is needed to examine to what extent the intervention contributed to improved reporting rates. Secondly, this project did not train village health teams due to inadequate funding. Our project did not give phones or tablets to facility health workers; therefore, we cannot make definitive claims about the effect of implementing the eIDSR on improving case reporting rates.

Conclusions

Improving the functionality and expanding the eIDSR system's geographical scope improved the efficiency of disease surveillance. The system's enhanced features, such as advanced reporting tools, improved data timeliness by streamlining the data entry process, and visualization features enabled quick interpretation of complex data, facilitating decision-making.

The wider geographical reach signifies a successful initial scale-up. To ensure the sustained success and scale of the eIDSR system, future programs should facilitate stakeholder commitment and leverage existing structures. Engaging internal and external stakeholders is crucial for securing long-term support, commitment, and endorsement at the national and subnational levels.

Acknowledgments

We acknowledge the US Centers for Disease Control and Prevention and the Infectious Diseases Institute, Makerere University for funding. Special appreciation is also extended to Health Information Systems Program Uganda for their invaluable technical assistance. Their expertise played a crucial role in the successful implementation and customization of the system. We acknowledge Uganda's Ministry of Health for its leadership and guidance throughout this project. Finally, we would like to express our gratitude to all individuals and organizations who contributed to this endeavor, directly or indirectly.

Conflicts of Interest

None declared.

References

1. Kayiwa J, Homsy J, Nelson LJ, et al. Establishing a public health emergency operations center in an outbreak-prone country: lessons learned in Uganda, January 2014 to December 2021. Health Secur 2022;20(5):394-407. [doi: 10.1089/hs.2022.0048] [Medline: 35984936]
2. Lukwago L, Nanyunja M, Ndayimirije N, et al. The implementation of Integrated Disease Surveillance and Response in Uganda: a review of progress and challenges between 2001 and 2007. Health Policy Plan 2013 Jan;28(1):30-40. [doi: 10.1093/heapol/czs022] [Medline: 22669899]
3. Joint external evaluation of IHR core capacities of the Republic of Uganda mission report june 26-30. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/WHO-WHE-CPI-REP-2017.49 [accessed 2025-04-02]

XSL•FO
RenderX

4. Kiberu VM, Matovu JKB, Makumbi F, Kyozira C, Mukooyo E, Wanyenze RK. Strengthening district-based health reporting through the district health management information software system: the Ugandan experience. BMC Med Inform Decis Mak 2014 May 13;14(1):40. [doi: 10.1186/1472-6947-14-40] [Medline: 24886567]
5. Uganda national eHealth strategy 2017 - 2021. Uganda Ministry of Health. URL: https://health.go.ug/sites/default/files/National%20e_Health%20Strategy_0.pdf [accessed 2024-10-25]
6. Technical guidelines for Integrated Disease Surveillance and Response (IDSR). World Health Organization. 2019. URL: https://www.who.int/publications/i/item/WHO-AF-WHE-CPI-01-2019 [accessed 2025-04-02]
7. DHIS2 overview - DHIS2. University of Oslo. URL: https://dhis2.org/overview/ [accessed 2024-10-04]
8. Joseph JJ, Mkali HR, Reaves EJ, et al. Improvements in malaria surveillance through the electronic Integrated Disease Surveillance and Response (eIDSR) system in mainland Tanzania, 2013-2021. Malar J 2022 Nov 8;21(1):321. [doi: 10.1186/s12936-022-04353-w] [Medline: 36348409]
9. Ibrahim LM, Okudo I, Stephen M, et al. Electronic reporting of integrated disease surveillance and response: lessons learned from northeast, Nigeria, 2019. BMC Public Health 2021 May 13;21(1):916. [doi: 10.1186/s12889-021-10957-9] [Medline: 33985451]
10. Youssef D, Yaghi A, Jouny A, Abou-Abbas L, Chammaa H, Ghosn N. Converting the existing disease surveillance from a paper-based to an electronic-based system using district health information system (DHIS-2) for real-time information: the Lebanese experience. BMC Health Serv Res 2022 Mar 25;22(1):395. [doi: 10.1186/s12913-022-07773-1] [Medline: 35337327]
11. Ibeneme S, Karamagi H, Muneene D, Goswami K, Chisaka N, Okeibunor J. Strengthening health systems using innovative digital health technologies in Africa. Front Digit Health 2022;4:854339. [doi: 10.3389/fdgth.2022.854339] [Medline: 35434700]
12. Adu-Gyamfi E, Nielsen P, Saebø JI. The dynamics of a global health information systems research and implementation project. Presented at: Proceedings of the 17th Scandinavian Conference on Health Informatics; Nov 12-13, 2019; Oslo, Norway URL: https://ep.liu.se/en/conference-article.aspx?series=ecp&issue=161&Article_No=13 [accessed 2025-04-02]
13. Maponga BA, Chirundu D, Shambira G, Gombe NT, Tshimanga M, Bangure D. Evaluation of the notifiable diseases surveillance system in Sanyati district, Zimbabwe, 2010-2011. Pan Afr Med J 2014;19:278. [doi: 10.11604/pamj.2014.19.278.5202] [Medline: 25870733]
14. Ngwa MC, Liang S, Mbam LM, et al. Cholera public health surveillance in the Republic of Cameroon-opportunities and challenges. Pan Afr Med J 2016;24:222. [doi: 10.11604/pamj.2016.24.222.8045] [Medline: 27800077]
15. Health information & digital health strategic plan. Uganda Ministry of Health. 2023 Jan 1. URL: https://library.health.go.ug/health-information-systems/digital-health/uganda-health-information-and-digital-health-strategic [accessed 2024-10-25]

Abbreviations

DHIS2: District Health Information Software 2
eIDSR: electronic Integrated Disease Surveillance and Response
HISP: Health Information Systems Program
ICT: information and communication technology
IDSR: Integrated Disease Surveillance and Response
IES&PHE: Integrated Epidemiology, Surveillance, and Public Health Emergencies
MAAIF: Ministryof Agriculture, Animal Industry, and Fisheries
MOH: Ministry of Health
WHO: World Health Organization

Edited by A Mavragani; submitted 02.05.24; peer-reviewed by C El-Hayek, K Goswami; revised version received 25.10.24; accepted 07.11.24; published 14.04.25.

Please cite as:
Mugasha R, Kwiringira A, Ntono V, Nakiire L, Ayebazibwe I, Kyozira C, Muruta AN, Kasule JN, Byonanebye DM, Nanyondo J, Walwema R, Kakooza F, Lamorde M
Scaling Up and Enhancing the Functionality of the Electronic Integrated Diseases Surveillance and Response System in Uganda, 2020-2022: Description of the Journey, Challenges, and Lessons Learned
JMIR Public Health Surveill 2025;11:e59783
URL: https://publichealth.jmir.org/2025/1/e59783
doi:10.2196/59783

XSL•FO
RenderX

© Rodney Mugasha, Andrew Kwiringira, Vivian Ntono, Lydia Nakiire, Immaculate Ayebazibwe, Caroline Kyozira, Allan Niyonzima Muruta, Juliet Namugga Kasule, Dathan M Byonanebye, Judith Nanyondo, Richard Walwema, Francis Kakooza, Mohammed Lamorde. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Collaborative Surveillance: Using a Minimum Set of Key Data Parameters for One Health Participatory Surveillance

Mark Smolinski[1*], MPH, MD; Nomita Divi[1*], MScPH; Onicio Leal Neto[1*], MScPH, PhD

Ending Pandemics Academy, Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, United States

[*]all authors contributed equally

Corresponding Author:
Mark Smolinski, MPH, MD
Ending Pandemics Academy
Mel and Enid Zuckerman College of Public Health
University of Arizona
1295 N Martin Ave
Tucson, 85724
United States
Phone: 1 520 626 4609
Email: drsmo@arizona.edu

Abstract

Early detection of a newly emerging or reemerging infectious disease is crucial to minimize the impact of such a threat on lives and livelihoods. With three of four pathogens capable of causing epidemics or pandemics arising first in animals and spreading to humans as zoonosis, a One Health approach to early detection is paramount. One Health participatory surveillance, defined as the bidirectional receiving and transmitting of data for action through direct engagement of the target population, is an effective form of collaborative surveillance to enhance global health security. Participatory surveillance systems can vary greatly when developed for a specific purpose or to meet a particular community's needs. Different geographies, languages, customs, beliefs, and practices often influence the breadth and depth of the data collected within each system. Imagine, however, if each of these varied systems could "speak" to each other, sharing their aggregated, deidentified data to create a comprehensive, real-time view of planetary health. The key is to collect the same information from users in each system, or at least a minimum set of key data parameters, to generate One Health surveillance greater than that of any individual system. To enable this vision, we propose a minimum set of key data parameters for One Health participatory surveillance that could be collected in any system through self-reporting by the public. This real-time collaborative surveillance could be the earliest indicator of a human, animal, or environmental health threat as it does not require interaction with a health care facility or provider where most disease surveillance traditionally occurs. One Health participatory surveillance that can detect major syndromes of potential emerging or reemerging pathogens through self-reporting on human, animal, or environmental health is a practical, scalable solution to identify and respond to rapidly spreading contagions.

(JMIR Public Health Surveill 2025;11:e77448) doi:10.2196/77448

KEYWORDS

One Health; participatory surveillance; collaborative surveillance; data parameters; pandemics; epidemics; global health security

Introduction

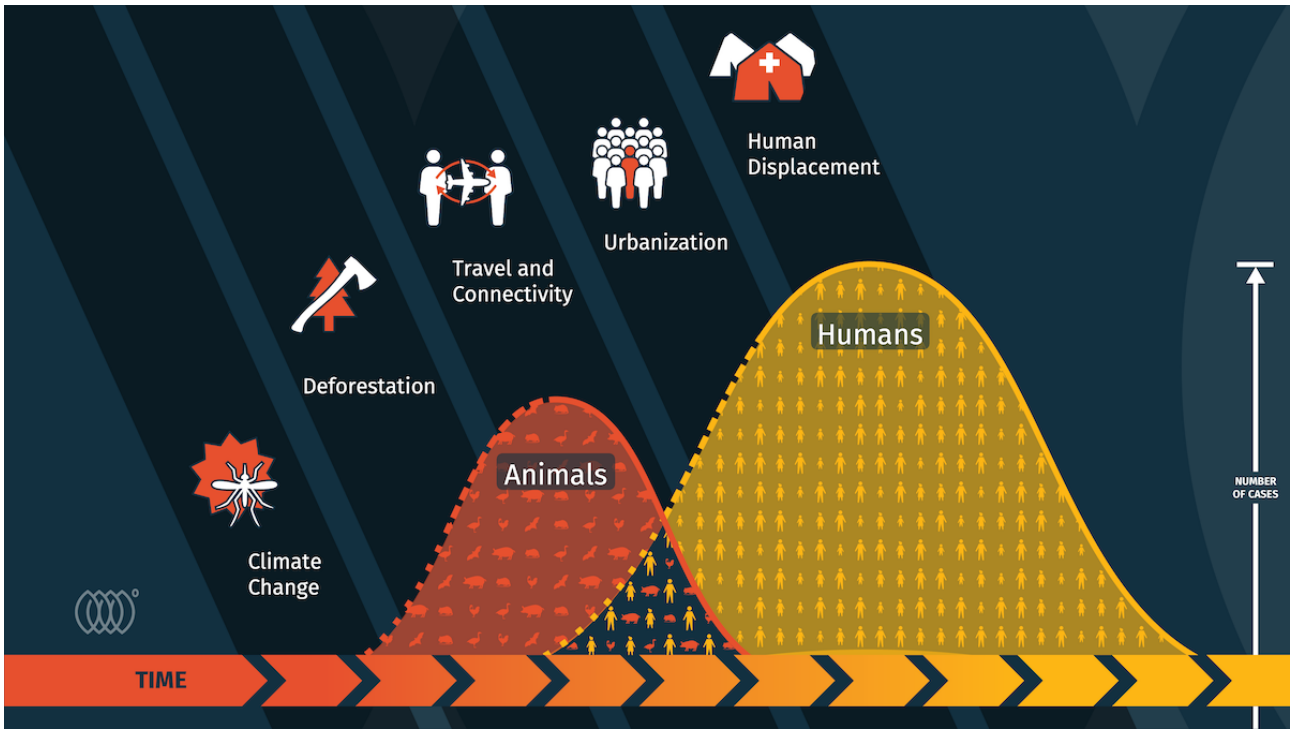
Of infectious diseases with epidemic or pandemic potential, 75% are zoonoses, meaning they spread to humans from wildlife or domestic animals [1,2]. Oftentimes, insect vectors such as mosquitoes, ticks, or fleas play a pivotal role in spreading disease from animals to humans [3,4]. Other times, infections spread from an infected animal to humans by way of direct contact with skin or saliva, or via the respiratory route [5]. In the case of hantavirus, for example, the inhalation of virus-contaminated particles when sweeping floors covered with dried urine and feces from deer mice led to acute respiratory infections in humans [6-8]. Regardless of the method of spillover, once a pathogen is established within the human population, numerous factors such as attitudes, behaviors, demographics, and politics can often exacerbate the further spread among humans [9,10].

Various additional dynamics lead to the emergence and spread of infectious diseases, including the speed and extent of international travel and trade, land use, climate change, lack of public health infrastructure, and access to health care, to name

XSL•FO
RenderX

just a few (Figure 1) [9,11]. Although not all drivers of disease emergence can be eliminated, cooperation and collaboration across public and private sectors, scientific disciplines, and geographic borders can significantly improve early detection and trigger an effective and timely response. An integrated approach to disease surveillance across species is paramount for the early detection and prompt response to emerging infections [12]. Prioritizing surveillance in planetary hot spots, those areas with the greatest likelihood of zoonotic outbreaks, is a compelling approach to reducing collective risk [13].

Figure 1. A One Health approach. Of every newly emerging infectious disease, 75% arise in animals and spread to humans. Animal outbreaks (shown in red) can quickly lead to human outbreaks (shown in yellow) when people are in close contact with livestock or wildlife. If we prioritize our efforts to find outbreaks faster in animals, it is possible to prevent human infection. Factors like climate change, urbanization, and increased travel and trade can exacerbate the spread and must also be considered.



One Health Participatory Surveillance

Directly engaging communities to report on the health of their family, themselves, and their animals, and any concerns regarding air or water quality, or other environmental health hazards is how we can safeguard the world from epidemics and pandemics. Nobody understands that something is “not quite right” better than the person who lives, works, or frequently spends time where these observations are occurring, and of course, nobody understands something is not right in how a person feels better than the individual themself.

Imagine if each of us took a couple of minutes each day, or even just once per week, to report on our health status and our family’s, or concerns within our immediate environmental surroundings. This could be done through email prompts, smartphone apps, brief SMS text messages, social media posts, or an interactive voice response system, or even seamlessly via passive data already captured by our phones, wearable devices, or other smart technologies. Ideally, if the majority of the community were doing the same, this form of collaborative surveillance [12] could detect the earliest signs of an emerging health threat. Everyone’s report would anonymously be connected on a visual platform to “see” that something is not right in a particular geographical region or simultaneously in several regions at a given point in time. This real-time collaborative surveillance could be the earliest indicator of a human, animal, or environmental health threat, as it does not require an interaction with a health care facility or provider, where most disease surveillance traditionally occurs.

Speed of detection is crucial. The faster a threat is identified and verified as true, the faster prompt actions can ensue to minimize its impact. Faster action saves lives and preserves livelihoods [14,15]. Once a threat gets a foothold in society, its impact grows exponentially in most cases. This is especially true for infectious disease threats once they enter the human population.

Collaborative surveillance can be achieved through participatory surveillance, defined as the process of receiving and transmitting data for action through direct engagement of the target population [16,17]. Participatory surveillance is the most efficient, cost-effective, and inclusive method to detect threats in real time—anywhere and everywhere across the world [18-21]. Each participatory surveillance system can be designed in any language, engaging users through local customs, practices, and beliefs, and adding features such as gamification to attract certain age groups or specific cohorts. Recruiting and retaining users is a significant obstacle that can be overcome with user-centric design principles and other nonfinancial incentives [22,23]. The more users that are engaged with the system, the more useful it becomes for early detection. The 2-way communication enabled through this approach can also

XSL•FO
RenderX

be instrumental in mounting a rapid and effective response. Numerous COVID-19–specific self-reporting systems were developed during the pandemic, validating the utility of participatory surveillance for the early detection and monitoring of emerging infections [24-26]. As more ecosystem dynamics and environmental data are integrated with human and animal disease surveillance efforts, the likelihood of accurately predicting and preventing zoonotic disease outbreaks will increase.

The most effective disease surveillance requires communities to harmoniously apply the best science, technology, and practices regarding human, animal, and environmental health in planetary hot spots for emerging infections and beyond. For the past decade, Ending Pandemics, a US-based nonprofit, has provided technical guidance and catalytic funding to support participatory surveillance in partnership with low- and middle-income countries on five continents [27]. A stellar example of One Health participatory surveillance was created in 2014 in Chiang Mai, Thailand, where community health volunteers are trained to report suspected illness in domestic livestock and poultry through submission of photos and simple data forms on mobile phones. Today, the system is receiving reports on food safety, contaminated water, and counterfeit drug sales in addition to reports of human or animal illness. In 2025, the system added wastewater surveillance as another mechanism for the early detection of an emerging zoonosis [28]. This One Health participatory approach was replicated in Tanzania in 2016, where village health volunteers report on human and animal illnesses among the pastoral herding communities whose livestock often comingle with the wildlife in the Serengeti region. Participatory surveillance has also been successfully used during mass gatherings in Brazil with the 2014 World Cup and 2016 Olympic Games [19]. The most successful participatory surveillance systems are those that are collaboratively developed, locally owned and resourced, and established as deep partnerships with the government from the time of inception.

Expanding Collaborative Surveillance

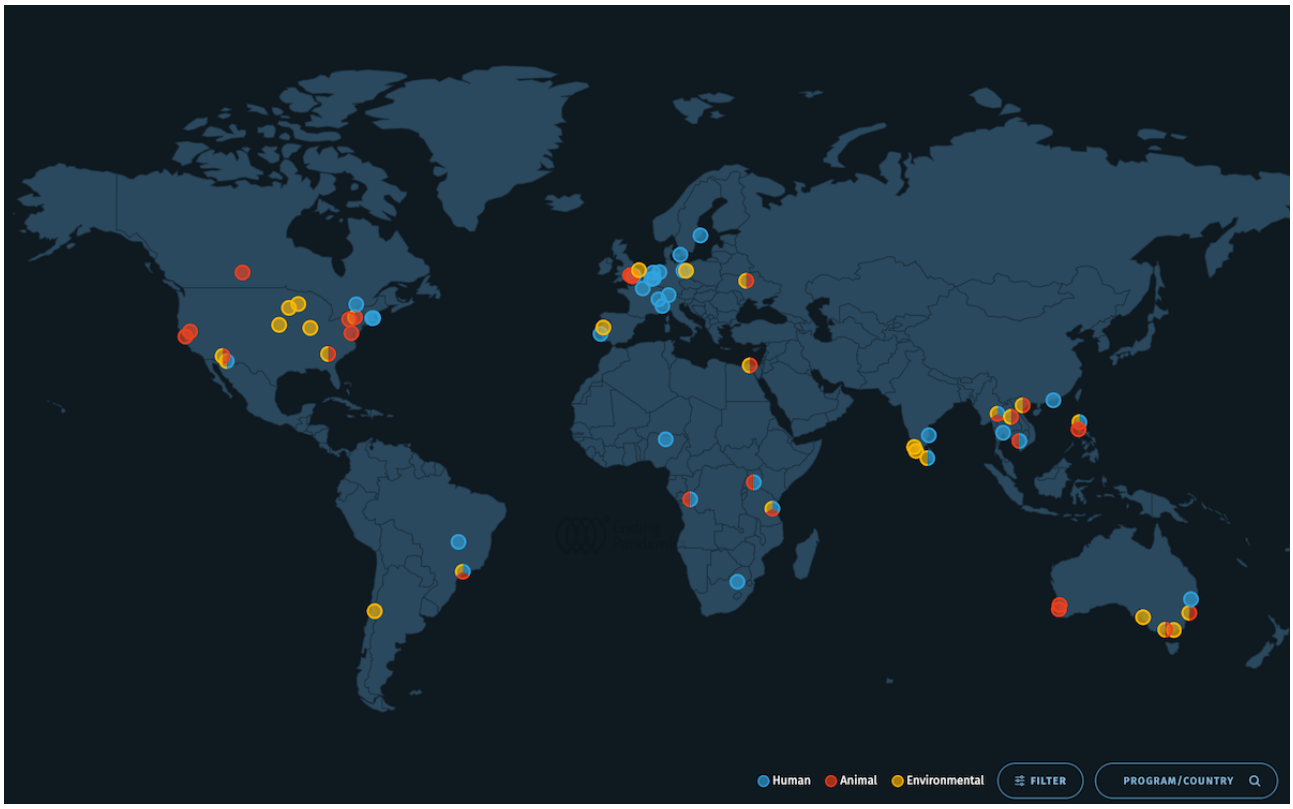
Although individual participatory surveillance systems exist, imagine if each of these varied systems could “speak” to each other, sharing their aggregated, deidentified data to expand the global view of the planet’s health. Each of these systems can be designed to engage specific population groups, churches, schools, workplaces, or any other social network. The key is to collect the same information from users in each system, or at least a minimum set of key data parameters, to generate One Health surveillance greater than that of any individual system.

Perhaps an analogy would be useful. If each surveillance system represents a spoke on a wheel, then each system has a pivotal role in holding the wheel’s rim in place and distributing pressure evenly. For the wheel to turn smoothly and bear weight without collapsing, every spoke must be made of the right material, cut to the correct length, and fastened securely. If even one spoke is too weak, too short, or poorly connected, the integrity of the entire wheel is compromised. The more spokes you have, the stronger and more resilient the wheel becomes. This redundancy adds durability and performance, allowing the wheel to handle greater challenges without faltering. If the wheel itself represents global disease surveillance, then each spoke signifies a vital part of this global system by capturing key information at the local, national, or regional levels.

To get a better idea of progress toward creating this global disease surveillance “wheel,” in 2022, Ending Pandemics conducted a systematic literature review and surveys to map the participatory surveillance “spokes” monitoring the health of animals, humans, or the environment across the world (Figure 2) [17,29]. The resulting map illustrates that the approach of directly engaging communities in self-reporting the health status of their members, animals, or environment is rapidly increasing in all areas of the world. Users of this map can learn key details about each system, including the number of reporters, the focus area of the system, and key contacts for the creators of the system. Although every system on the map represents bidirectional information sharing, some are sector specific and limit their reach to wildlife, livestock, the environment, or human health monitoring. Several systems, however, have achieved the One Health monitoring goal of incorporating data parameters that span all sectors. One Health participatory surveillance is the key to quickly finding and responding to emerging health threats and, in some cases, is contributing to the prediction and prevention of potential diseases with epidemic or pandemic potential. The ultimate success of One Health participatory surveillance will be realized when all systems can share data as mutual key data parameters are collected among the various systems.

XSL•FO
RenderX

Figure 2. One Health participatory surveillance interactive landscape map. This interactive map highlights participatory surveillance systems in people, animals, and ecosystems for detecting and responding to potential disease outbreaks. The systems' colors indicate the sector they focus on, with red indicating animal health, blue indicating human health, and yellow indicating environmental health.



One Health Participatory Surveillance Data Parameters Compendium

Every participatory surveillance system collects information as a set of data parameters. Ending Pandemics compiled a compendium of data parameters from participatory surveillance systems that operate in one or more of the One Health sectors: humans, animals, and the environment. Through a series of virtual workshops in 2021-2022 with experts from each sector, including health experts from the World Health Organization, the Food and Agriculture Organization of the United Nations, and the World Organization for Animal Health, a set of data parameters within each sector was compiled that were geared toward the public. Data parameters that seemed too "specialized" for public participation were excluded from the final set. Where available, information was gathered on how different parameters are collected within the individual systems.

The One Health Participatory Surveillance Data Parameters Compendium [30] is a snapshot into the world of participatory surveillance systems collecting data on human, animal, and environmental health. The compendium illustrates the breadth and depth of what is being used in participatory surveillance systems operating across the world. The compendium can serve the needs of those who may be contemplating the development of a participatory surveillance system, looking to expand a current system to include additional health sectors, or adapting data parameters to match other systems to foster collaborations and data sharing.

Establishing a Minimum Set of Key Data Parameters

Participatory surveillance systems are generally designed to collect information relevant to the community it is monitoring with a specific intent or purpose. Hence, some systems are focused directly on the public's health, and others might be tailored more to wildlife health or conservation, with the "reporter" having special training or expanded knowledge relative to the system's intent, such as a park ranger or veterinarian. Imagine, however, if we could collect a minimum set of key data parameters that monitor One Health across every participatory surveillance system, not necessarily all the parameters in each system, but a minimum set of key data parameters that can detect an emerging health threat of zoonotic origin by monitoring the health of animals, humans, and the environment.

Ending Pandemics used the compendium of data parameters collected across the known participatory surveillance systems and reduced this fairly massive set of data parameters to approximately 100 data parameters for consideration in a system designed for use by any member of a community (ie, the public). In 2022, the Fourth International Workshop on Participatory Surveillance (IWOPS) convened human, animal, and environmental health experts from 28 countries to further refine this set of 100 parameters into a potential minimum set of key data parameters for inclusion in any system [31]. The main goal of this intensive consultation process was the ability of this minimum set of key data parameters to detect any emerging

XSL•FO
RenderX

zoonosis, with the ultimate goal of potentially detecting environmental, wildlife, and domestic livestock and poultry threats before any human illness. If not detected before spillover, this minimum set of key data parameters can provide a fuller picture of an emerging infectious disease threat as the earliest human cases are identified through self-reporting. Our proposed minimum set of key data parameters consists of 42 data parameters across human, animal, and environmental health, along with 10 general data parameters to collect information on demographics, location, and timing of reports. We have also included the potential for photographs, diagnostic tests or laboratory confirmation, and digital biomarkers (eg, automated wearables, contactless physiological sensing, or other smart technologies) to be added by system designers where relevant and available (Figure 3). The addition of these latter data elements can add specificity to the system and help validate the self-reported syndromic data.

Figure 3. Minimal set of key data parameters for One Health participatory surveillance. This set of 42 parameters across human, animal, and environmental health, along with 10 general parameters and additional auxiliary parameters, can be incorporated into participatory systems to enable One Health surveillance to detect any emerging zoonosis.

The Value of a Minimum Set of Key Data Parameters

As stated earlier, the faster an emerging infection is detected, the quicker an effective response can ensue to limit loss of life and preserve livelihoods. Several major epidemics and pandemics throughout history were the result of a previously unrecognized pathogen [32]. Hence, relying on diagnostics or laboratory tests would not be useful at the earliest stages of such outbreaks. In fact, with the lack of readily available, widely dispersed diagnostics for even known pathogens, the clustering of similar symptoms in time and space is often the only means to detect an emerging disease. Syndromic surveillance, therefore, is often relied upon for the early detection of a newly emerging or reemerging infectious disease [33]. Participatory surveillance that can detect major syndromes of potential emerging or reemerging pathogens through self-reporting of illness is a practical way to help identify a rapidly spreading contagion.

In general, the minimum set of key data parameters should identify a newly emerging or reemerging pathogen based on the frequency of reports and the geographical location of such reports. We believe that respiratory diseases have the highest likelihood of wide-scale epidemics or pandemics, and the minimum set of key data parameters should pick up such a threat. Cough and shortness of breath, plus or minus sore throat, can be effective at identifying influenza, severe acute respiratory syndrome, Middle East respiratory syndrome, or a yet unknown respiratory pathogen. When combined with fever, malaise, body aches or pains, a cluster of such reports heightens concerns of a serious respiratory threat. It may even be possible to specify

XSL•FO
RenderX

a potential diagnosis with additional symptoms being reported. Likewise, data parameters on health-seeking behavior and absenteeism from school or work can provide further evidence of illness severity.

Additional self-reported symptoms can also identify other potential pathogens by collecting syndromic information on rash illness, diarrheal illness, or hemorrhagic fever. Key data parameters on exposure to domestic animals, wildlife (including rodents), and insect vectors like mosquitoes, ticks, and fleas can lead to a plethora of other zoonotic threats to be considered as the cause of an outbreak. Furthermore, several of the key data parameters in the minimum set can provide valuable information on sickness or die off in domestic animals or wildlife, or environmental concerns like flooding or water safety that might spur further investigation, leading to the prediction and prevention of human illness from zoonosis.

The minimum set of key data parameters was analyzed by expert human curation to validate that most emerging infectious disease threats with epidemic or pandemic potential could be identified through various combinations of terms. For example, the minimum set of key data parameters should be able to identify most emerging infections capable of causing epidemic or pandemic disease. The World Health Organization updated its list of priority pathogens capable of causing epidemic and pandemic diseases in 2024 [34] by convening over 300 scientists to consider the evidence on virus families and bacteria, as well as "Disease X." Disease X is included to indicate an unknown pathogen that could cause a serious epidemic or pandemic. With major syndromes being reported regularly, it is more than likely that Disease X would also be detected through One Health participatory surveillance.

The Value of Collecting Similar Data Parameters: Case of Global Flu View

Participatory surveillance systems for influenza-like illness have been well established in numerous countries, many operating for over a decade [35]. Ending Pandemics convened the developers of each of these systems at the Second IWOPS in 2013 [36]. The IWOPS community established a minimum set of data parameters among systems collecting self-reported information to monitor the annual risk of influenza. As a result, deidentified and aggregated data are shared on a digital platform, Global Flu View (GFV), to improve situational awareness and increase our understanding of how influenza emerges across the world [37].

GFV is a system orchestrator platform that converges and combines influenza-related data from 11 countries, harmonizing the way information is captured and reported by ensuring that datasets from different systems and regions remain consistent, comparable, and accurate. As a part of the platform's approach, GVF offers a ready-to-use tool kit for institutions interested in deploying or enhancing their own participatory surveillance systems. The tool kit significantly lowers the barriers to entry by providing a clear and concise guide on the key data parameters that have already been vetted across multiple regions. At the same time, GFV maintains a high degree of flexibility to accommodate local realities and linguistic nuances, allowing the naming and coding of parameters to be adapted into more culturally and contextually appropriate terms. Ultimately, this approach amplifies the quality and timeliness of influenza-like illness data worldwide, helping health authorities spot potential outbreaks sooner and collaborate more effectively in responding to public health threats.

GFV represents one example of what is possible when participatory surveillance systems can "speak" to each other regarding influenza-like illness data parameters. Embracing the entire minimum set of key data parameters encompassing One Health will expand the value of participatory surveillance for all epidemic and pandemic threats.

Limitations

This minimum set of key data parameters is not without limitations. The parameters are based on syndromic surveillance, with the goal of being highly sensitive to enable early detection of any potential health threat to the community. A primary goal of participatory surveillance is speed of detection; hence, the system may produce false positives. The occurrence of false positives, however, is offset by the potential to quickly uncover a rapidly spreading infection within the population. Measuring sensitivity generally requires comparison with "gold standard" surveillance data to determine the true frequency of the disease in the population [38]. To capture the sensitivity of a particular participatory surveillance system, a gold standard for early detection will need to match the purpose of this system. While the broad range of parameters in this minimum dataset provides high sensitivity, the parameters lack specificity. This reduced specificity may be overcome as participatory surveillance systems allow users to include diagnostic or laboratory tests, digital biomarkers, or other data derived from future advances in technology like wearables and environmental sensors. Furthermore, as the field of rapid diagnostics continues to expand in all areas of the world, especially the availability of home test kits, greater specificity in diagnosing the threat among the population reporting into the system will be possible. Further work is needed to determine the sensitivity and specificity of the minimum set of key data parameters for One Health participatory surveillance.

The minimum set of key data parameters is intended to rapidly detect any emerging zoonosis, whether signs and symptoms appear first in animals or humans. As such, the minimum set of key data parameters may not detect some endemic diseases or subclinical infections in humans and animals, nor are they intended to monitor the general health of the population. Participatory surveillance is intended to complement other surveillance efforts such as severe acute respiratory infection, sentinel sero-surveillance, sequential biologic sampling in high-risk communities, or health care–based surveillance systems. Integrating participatory surveillance among populations that are also part of the severe acute respiratory infection or other targeted surveillance systems can serve to expand the reach of monitoring within a given population.

XSL•FO
RenderX

Conclusion

Early detection is paramount when it comes to any emerging infectious disease. Participatory surveillance is an effective method of collaborative surveillance that allows anyone to participate in monitoring the health of their communities to detect the earliest signs of an emerging infectious disease threat. Furthermore, participatory surveillance allows for monitoring factors in the environment along with wildlife and domestic livestock surveillance for detecting zoonotic diseases. A minimum set of key data parameters for One Health participatory surveillance can enable an individual system anywhere in the world to connect to share information with other systems collecting similar data. This, of course, requires standardization of data collection or analysis to permit true comparisons among systems. As more systems become integrated into regional or global surveillance efforts, the greater the chances of detecting zoonotic diseases of epidemic or pandemic potential at their earliest stages of emergence. Persons or organizations using the minimum set of key data parameters to create a new participatory surveillance system or to expand a system currently in use can adapt these parameters to fit the needs of their specific community. For example, they can survey users within their system to help ascertain the applicability and clarity of the specific parameters and the potential need for revisions or additional terms. With further advances in environmental and animal health monitoring, it may be possible to predict and prevent human infections from occurring at all. When this potential is realized, the resulting greatly enhanced global health security will allow the prevention of major epidemics and pandemics.

Acknowledgments

We gratefully acknowledge the valuable contributions of all participants of the first to fourth International Workshops on Participatory Surveillance in advancing this work on participatory surveillance. We also wish to recognize the authors of the One Health Participatory Surveillance Data Parameters Compendium, whose work played a critical role in informing the development of the minimum set of key data parameters for One Health participatory surveillance. The collective efforts of these groups were instrumental in advancing this effort.

Conflicts of Interest

None declared.

References

1. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. Nature 2008 Mar 21;451(7181):990-993 [FREE Full text] [doi: 10.1038/nature06536] [Medline: 18288193]
2. Woolhouse M, Gowtage-Sequeria S. Host range and emerging and reemerging pathogens. Emerg Infect Dis 2005 Dec;11(12):1842-1847. [doi: 10.3201/eid1112.050997] [Medline: 16485468]
3. Gubler DJ. Vector-borne diseases. Rev Sci Tech 2009 Aug;28(2):583-588. [doi: 10.20506/rst.28.2.1904] [Medline: 20128467]
4. Chala B, Hamde F. Emerging and re-emerging vector-borne infectious diseases and the challenges for control: a review. Front Public Health 2021;9:715759 [FREE Full text] [doi: 10.3389/fpubh.2021.715759] [Medline: 34676194]
5. Rahman MT, Sobur MA, Islam MS, Ievy S, Hossain MJ, El Zowalaty ME, et al. Zoonotic diseases: etiology, impact, and control. Microorganisms 2020 Sep 12;8(9):1405 [FREE Full text] [doi: 10.3390/microorganisms8091405] [Medline: 32932606]
6. Engelthaler D, Mosley D, Cheek J, Levy CE, Komatsu KK, Ettestad P, et al. Climatic and environmental patterns associated with hantavirus pulmonary syndrome, Four Corners region, United States. Emerg Infect Dis 1999;5(1):87-94 [FREE Full text] [doi: 10.3201/eid0501.990110] [Medline: 10081675]
7. van Seventer JM, Hochberg NS. Principles of infectious diseases: transmission, diagnosis, prevention, and control. In: Quah SR, editor. International Encyclopedia of Public Health (Second Edition). Cambridge, MA: Academic Press; 2017:22-39.
8. Plowright R, Parrish C, McCallum H, Hudson PJ, Ko AI, Graham AL, et al. Pathways to zoonotic spillover. Nat Rev Microbiol 2017 Aug;15(8):502-510 [FREE Full text] [doi: 10.1038/nrmicro.2017.45] [Medline: 28555073]
9. Sabin N, Calliope A, Simpson S, Arima H, Ito H, Nishimura T, et al. Implications of human activities for (re)emerging infectious diseases, including COVID-19. J Physiol Anthropol 2020 Sep 25;39(1):29 [FREE Full text] [doi: 10.1186/s40101-020-00239-5] [Medline: 32977862]
10. Institute of Medicine. In: Smolinski MS, Hamburg MA, Lederberg J, editors. Microbial Threats to Health: Emergence, Detection, and Response. Washington, DC: The National Academies Press; 2003.
11. Findlater A, Bogoch I. Human mobility and the global spread of infectious diseases: a focus on air travel. Trends Parasitol 2018 Sep;34(9):772-783 [FREE Full text] [doi: 10.1016/j.pt.2018.07.004] [Medline: 30049602]
12. Archer B, Abdelmalik P, Cognat S, Grand PE, Mott JA, Pavlin BI, et al. Defining collaborative surveillance to improve decision making for public health emergencies and beyond. Lancet 2023 Jun 03;401(10391):1831-1834 [FREE Full text] [doi: 10.1016/S0140-6736(23)01009-7] [Medline: 37230104]
13. Allen T, Murray K, Zambrana-Torrelio C, Morse SS, Rondinini C, Di Marco M, et al. Global hotspots and correlates of emerging zoonotic diseases. Nat Commun 2017 Oct 24;8(1):1124. [doi: 10.1038/s41467-017-00923-8] [Medline: 29066781]

XSL•FO
RenderX

14. Smolinski M, Crawley A, Olsen J. Finding outbreaks faster. Health Secur 2017;15(2):215-220 [FREE Full text] [doi: 10.1089/hs.2016.0069] [Medline: 28384035]
15. COVID-19 could have been contained – new report shows how countries successfully beat infectious disease outbreaks before they became epidemics. Vital Strategies. 2021 Apr 27. URL: https://tinyurl.com/5ahyxy5f [accessed 2025-07-29]
16. Smolinski MS, Crawley AW, Olsen JM, Jayaraman T, Libel M. Participatory disease surveillance: engaging communities directly in reporting, monitoring, and responding to health threats. JMIR Public Health Surveill 2017 Oct 11;3(4):e62 [FREE Full text] [doi: 10.2196/publichealth.7540] [Medline: 29021131]
17. McNeil C, Verlander S, Divi N, Smolinski MS. The landscape of participatory surveillance systems across the One Health spectrum: systematic review. JMIR Public Health Surveill 2022 Aug 05;8(8):e38551 [FREE Full text] [doi: 10.2196/38551] [Medline: 35930345]
18. Smolinski M. Innovations in participatory disease surveillance. Int J Infect Dis 2016 Dec;53:10 [FREE Full text] [doi: 10.1016/j.ijid.2016.11.030]
19. Leal Neto O, Dimech GS, Libel M, de Souza WV, Cesse E, Smolinski MS, et al. Saúde na Copa: the world's first application of participatory surveillance for a mass gathering at FIFA World Cup 2014, Brazil. JMIR Public Health Surveill 2017 May 04;3(2):e26 [FREE Full text] [doi: 10.2196/publichealth.7313] [Medline: 28473308]
20. Leal Neto O, Cruz O, Albuquerque J, Nacarato de Sousa M, Smolinski MS, Pessoa Cesse EA, et al. Participatory surveillance based on crowdsourcing during the Rio 2016 Olympic Games using the Guardians of Health platform: descriptive study. JMIR Public Health Surveill 2020 Apr 07;6(2):e16119 [FREE Full text] [doi: 10.2196/16119] [Medline: 32254042]
21. Valerio MGP, Laher B, Phuka J, Lichand G, Paolotti D, Leal Neto O. Participatory disease surveillance for the early detection of cholera-like diarrheal disease outbreaks in rural villages in Malawi: prospective cohort study. JMIR Public Health Surveill 2024 Jul 16;10:e49539 [FREE Full text] [doi: 10.2196/49539] [Medline: 39012690]
22. Dalton C, Carlson S, Butler M, Cassano D, Clarke S, Fejsa J, et al. Insights from flutracking: thirteen tips to growing a web-based participatory surveillance system. JMIR Public Health Surveill 2017 Aug 17;3(3):e48 [FREE Full text] [doi: 10.2196/publichealth.7333] [Medline: 28818817]
23. Land-Zandstra AM, van Beusekom M, Koppeschaar C, van den Broek J. Motivation and learning impact of Dutch flu-trackers. J Sci Commun 2016;15(01):A04. [doi: 10.22323/2.15010204]
24. Wittwer S, Paolotti D, Lichand G, Leal Neto O. Participatory surveillance for COVID-19 trend detection in Brazil: cross-sectional study. JMIR Public Health Surveill 2023 Apr 26;9:e44517 [FREE Full text] [doi: 10.2196/44517] [Medline: 36888908]
25. Leal Neto O, Santos F, Lee J, Albuquerque J, Souza W. Prioritizing COVID-19 tests based on participatory surveillance and spatial scanning. Int J Med Inform 2020 Nov;143:104263 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104263] [Medline: 32877853]
26. Leal Neto O, Egger T, Schlegel M, Flury D, Sumer J, Albrich W, et al. Digital SARS-CoV-2 detection among hospital employees: participatory surveillance study. JMIR Public Health Surveill 2021 Nov 22;7(11):e33576 [FREE Full text] [doi: 10.2196/33576] [Medline: 34727046]
27. Divi N, Smolinski M. EpiHacks, a process for technologists and health experts to cocreate optimal solutions for disease prevention and control: user-centered design approach. J Med Internet Res 2021 Dec 15;23(12):e34286 [FREE Full text] [doi: 10.2196/34286] [Medline: 34807832]
28. Yano T, Phornwisetsirikun S, Susumpow P, Visrutaratna S, Chanachai K, Phetra P, et al. A participatory system for preventing pandemics of animal origins: pilot study of the Participatory One Health Disease Detection (PODD) system. JMIR Public Health Surveill 2018 Mar 21;4(1):e25 [FREE Full text] [doi: 10.2196/publichealth.7375] [Medline: 29563079]
29. One Health Participatory Surveillance. Ending Pandemics. URL: https://endingpandemics.org/participatory-surveillance-programs-map/ [accessed 2025-07-26]
30. McNeil C, Divi N, Bargeron Iv CT, Capobianco Dondona A, Ernst KC, Gupta AS, et al. Data parameters from participatory surveillance systems in human, animal, and environmental health from around the globe: descriptive analysis. JMIR Public Health Surveill 2025 Mar 26;11:e55356 [FREE Full text] [doi: 10.2196/55356] [Medline: 40138683]
31. Smolinski MS, Divi N. IWOPS: 4th International Workshop On Participatory Surveillance. Ending Pandemics. 2022. URL: https://endingpandemics.org/wp-content/uploads/2023/01/FINAL-IWOPS-IV-REPORT.pdf [accessed 2025-07-29]
32. Honeyman D, Gurdasani D, Notaras A, Akhtar Z, Edgeworth J, Moa A, et al. Global epidemiology of outbreaks of unknown cause identified by Open-Source Intelligence, 2020-2022. Emerg Infect Dis 2025 Mar;31(2):298-308. [doi: 10.3201/eid3102.240533] [Medline: 39983687]
33. Smith GE, Elliot AJ, Lake I, Edeghere O, Morbey R, Catchpole M, et al. Syndromic surveillance: two decades experience of sustainable systems - its people not just data!. Epidemiol Infect 2019 Jan;147:e101 [FREE Full text] [doi: 10.1017/S0950268819000074] [Medline: 30869042]
34. Ukoaka B, Okesanya O, Daniel F, Ahmed MM, Udam NG, Wagwula PM, et al. Updated WHO list of emerging pathogens for a potential future pandemic: implications for public health and global preparedness. Infez Med 2024;32(4):463-477 [FREE Full text] [doi: 10.53854/liim-3204-5] [Medline: 39660154]
35. Best practices for the design, implementation, reporting and analysis of participatory surveillance for influenza-like illness. World Health Organization. 2024 Jul 5. URL: https://www.who.int/publications/i/item/9789240095038 [accessed 2025-08-05]

XSL•FO
RenderX

36. Second International Workshop on Participatory Surveillance (IWOPS II). Ending Pandemics. URL: https://endingpandemics.org/projects/international-workshops-on-participatory-surveillance-iwops/iwops-ii/ [accessed 2025-07-29]
37. Leal Neto O, Paolotti D, Dalton C, Carlson S, Susumpow P, Parker M, et al. Enabling multicentric participatory disease surveillance for global health enhancement: viewpoint on Global Flu View. JMIR Public Health Surveill 2023 Sep 01;9:e46644 [FREE Full text] [doi: 10.2196/46644] [Medline: 37490846]
38. Buehler J, Hopkins R, Overhage J, Sosin D, Tong V, Working Group CDC. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC Working Group. MMWR Recomm Rep 2004 May 07;53(RR-5):1-11 [FREE Full text] [Medline: 15129191]

Abbreviations

GFV: Global Flu View
IWOPS: International Workshop on Participatory Surveillance

Edited by A Mavragani; submitted 15.05.25; peer-reviewed by T Yano, P Babar; comments to author 12.06.25; revised version received 07.07.25; accepted 21.07.25; published 08.08.25.

Please cite as:
Smolinski M, Divi N, Leal Neto O
Collaborative Surveillance: Using a Minimum Set of Key Data Parameters for One Health Participatory Surveillance
JMIR Public Health Surveill 2025;11:e77448
URL: https://publichealth.jmir.org/2025/1/e77448
doi:10.2196/77448
PMID:40779314

©Mark Smolinski, Nomita Divi, Onicio Leal Neto. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 08.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Injury Registry Data on Policymaking, Hospitalizations, and Mortality: Systematic Review

Ana Cláudia Medeiros-de-Souza[1], MSc; Luana Emanuelly Sinhori Lopes[1], MSc; Tayna Felicissimo Gomes de Souza Bandeira[1], MSc; Lucas Reis Correia[1,2], MSc; Naiza Nayla Bandeira de Sá[3,4], DSc; Bruno Zocca de Oliveira[1], MSc, MPP

[1]Hospital Israelita Albert Einstein, 755 Comendador Elias Jafet Street, L1 Floor, Room 134, São Paulo, Brazil

[2]Department of Preventive Medicine, University of São Paulo, São Paulo, Brazil

[3]Secretariat of Health and Environmental Surveillance, Department of Surveillance of Noncommunicable Diseases and Health Promotion, Ministry of Health, Brasília, Brazil

[4]Faculty of Nutrition, Federal University of Pará, Belém, Brazil

Corresponding Author:
Ana Cláudia Medeiros-de-Souza, MSc
Hospital Israelita Albert Einstein, 755 Comendador Elias Jafet Street, L1 Floor, Room 134, São Paulo, Brazil

Abstract

Background: The Brazilian project, launched in 2021, aims to establish a nationwide injury registry that systematically collects detailed information on incidents and individuals across the country, regardless of injury severity. The registry integrates information from prehospital and hospital care, various health systems lacking interoperability, and data from sectors such as firefighters and police. Its primary aim is to enhance health surveillance by providing timely, high-quality information that guides prevention strategies and informs policymaking. In addition, the project seeks to reduce morbidity and mortality associated with injuries.

Objective: This study aimed to investigate the effects of injury registry data on policymaking, hospitalization rates or duration, and mortality.

Methods: The systematic review followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, with a protocol registered in PROSPERO (International Prospective Register of Systematic Reviews, CRD42023481528). A total of 5 databases were searched in November 2023, with an update conducted in March 2024, incorporating reference lists from the studies included. Two reviewers independently screened records, extracted data, and assessed methodological quality using the Newcastle-Ottawa Scale, resolving disagreements with a third reviewer. Studies were eligible if they reported results related to the implementation and use of injury or trauma registry data for at least one outcome of interest, while those based on other sources were excluded. Synthesis of findings was presented in tables, and the observed results were reported as number or percentage differences.

Results: Out of 9100 studies retrieved, 3951 were excluded due to duplication, leaving 5149 for selection, with 15 full texts reviewed. Only 5 studies met the inclusion criteria, highlighting a notable scarcity of research on the effects or results of registry data on injury outcomes. It is important to note that the studies included reflect correlations rather than causalities, and there are currently no publications on impact. The findings suggest that injury and trauma registries have the potential to inform policymaking, which can lead to enhanced health outcomes. One study noted a 3-day reduction in intensive care unit stay (from 16 to 13 days; $P<.05$) and a 4% reduction in expected hospital mortality (from 17.5% to 21.5%) for patients with an Injury Severity Score ≥16, while another showed a 42% annual decrease in traffic injury hospital admissions (from 45 to 16). Significant methodological heterogeneity and the small number of studies limited the feasibility of a meta-analysis.

Conclusions: Establishing an injury registry in Brazil presents a significant opportunity to enhance health outcomes through informed policymaking. While it is crucial to set appropriate expectations regarding its effects on morbidity and mortality, particularly concerning the causality and transportability of the findings to the Brazilian context, its role in facilitating preventive measures and improving surveillance capabilities remains valuable.

(JMIR Public Health Surveill 2025;11:e67115) doi:10.2196/67115

KEYWORDS

injury registry; trauma registry; policymaking; health policy; wounds and injuries; prevention and control; injury; trauma; Brazil; Newcastle-Ottawa Scale; PRISMA; systematic review; registry; registry data; injury outcome; hospitalization; mortality; morbidity; health surveillance; quality information; epidemiology; hospital care

XSL•FO
RenderX

Introduction

Injuries resulting from various incidents, such as accidents, falls, drownings, burns, poisonings, and acts of violence directed at oneself or others, represent a global concern. Annually, approximately 4.4 million deaths are attributed to injuries, constituting nearly 8% of all deaths. Unintentional causes account for 3.16 million fatalities, while acts of violence result in 1.25 million deaths. According to the World Health Organization, among individuals aged 5 - 29 years old, 3 of the top 5 causes of death are injury-related: road traffic injuries, homicide, and suicide. In addition, falls lead to over 684,000 deaths each year, highlighting a growing and often overlooked public health issue [1].

In Brazil, injuries and violence have frequently ranked as the fourth leading cause of death. In 2022, the country recorded 152,945 fatalities attributed to injuries, which corresponds to approximately 75 deaths per 100,000 inhabitants, constituting nearly 10% of all deaths. Most of these fatalities were due to accidents (46.9%), with traffic accidents accounting for 22.8% and falls for 11.8%. Homicides represented 29.0% of the deaths, while suicides accounted for 10.8%. Studies highlight the significant impact of injuries on premature mortality and disability, particularly among young individuals aged 20 - 39 years old and males, emphasizing the urgent need to prioritize this issue in the country [2,3].

Conversely, nonfatal injuries also pose a significant public health challenge. Every year, millions of people worldwide experience nonfatal injuries, resulting in visits to emergency departments or hospitalizations [1]. In Brazil, in 2022, the number of hospitalizations due to injuries was nearly 8.8 times higher than the number of deaths, totaling 1,351,021 hospitalizations, which translates to over 600 injury-related hospitalizations per 100,000 Brazilians. Falls were the most prevalent cause of these hospitalizations, accounting for 34.4%, followed by traffic accidents at 17.9% [4].

The National Policy for the Reduction of Morbidity and Mortality from Accidents and Violence was established in response to the challenge posed by injuries as a public health issue in Brazil [5]. This policy has led to several key initiatives, including the launch of the National Policy for Emergency Care in 2003 [6] and the proposal for creating the National Network for the Prevention of Violence and Health Promotion in 2004 [7]. Two years later, in 2006, the Violence and Accident Surveillance System in Sentinel Services was implemented [8]. In 2010, the Life in Traffic Project was initiated, emphasizing intersectoral coordination and data integration for traffic accident surveillance and intervention across 54 Brazilian municipalities [9,10]. Furthermore, beginning in 2011, interpersonal and self-inflicted violence were included in the list of diseases and conditions subject to compulsory notification in public and private health services in the country [11].

An injury registry is a systematic and standardized database that records detailed information about events and individuals who have sustained injuries, regardless of their severity or whether they result from accidents, falls, violence, or other causes. These databases can encompass a wide range of injuries, from minor ones that may not require hospitalization to more serious injuries that do. Traumatic injuries that require acute care delivered to hospitalized patients, typically resulting from high-impact events such as car accidents, severe falls, or gunshot wounds, are usually included in trauma registries. Both types of registries play important roles in injury prevention, research, and health care management. The primary distinction lies in the scope and severity of injuries each type of registry deals with. They allow policy makers to efficiently target resources and implement evidence-based interventions [12-15].

Recently, in 2021, a project was initiated under the Institutional Development Support Program of the Unified Health System (Programa de Apoio ao Desenvolvimento Institucional do Sistema Único de Saúde [PROADI-SUS]), a strategy of the Brazilian Ministry of Health, executed by the Hospital Israelita Albert Einstein, to establish a national injury registry. The primary aim of this initiative is to enhance health surveillance capacity by providing timely, high-quality information to guide the development and evaluation of prevention strategies and policymaking related to accidents and violence. To achieve this, the project focuses on improving data integration and quality from prehospital and hospital care, as well as from various existing health information systems that currently lack interoperability. In addition, it seeks to incorporate data from multiple sectors involved in injuries and violence, including firefighters and police. The objective is to consolidate information from diverse systems into an integrated database that will be centrally accessible to health managers and researchers throughout Brazil by 2026. Ultimately, these efforts are designed to reduce morbidity and mortality associated with injuries [16,17].

The success of trauma systems has been reported in various localities, with their effectiveness well-established [14,18-20]. Estimates indicate a 15% reduction in the odds of mortality [21,22]. Furthermore, the implementation of trauma registries has been well-documented, and these registries are consistently recognized as essential tools for improving trauma systems and policymaking [13,23]. However, a significant knowledge gap remains regarding the effects of injury registries that encompass all injuries, regardless of their severity or whether hospitalization is required [23].

Considering the current efforts to implement an injury registry in Brazil, there is an urgent need to establish an evidence base in this context. Therefore, this systematic review aimed to investigate the effects and results of injury registries on policymaking and injury outcomes, particularly in terms of reducing hospitalizations and mortality. Hence, the research question guiding this review is as follows: “What is the effect of implementing an injury registry on policymaking, hospitalization rates or duration, and mortality due to injuries?”.

Methods

Protocol

The development of this study followed the recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement, ensuring adherence to

XSL•FO
RenderX

best practices in systematic review methodology [24,25]. The PRISMA 2020 checklist for a systematic review report is completed and provided in Checklist 1, while the checklist for the abstract is provided in Checklist 2. A protocol for this systematic review was registered in the PROSPERO (International Prospective Register of Systematic Reviews) database under registration number CRD42023481528. A comprehensive version of the protocol has been published [26].

Eligibility Criteria

The eligibility criteria for this systematic review were formulated according to the PICOS structure (Population, Intervention, Comparison, Outcomes, and Study), as shown in Table 1. A search was conducted for articles reporting results or effects related to the implementation and use of data from injury registries, including trauma registries, for at least one of the outcomes described in Table 1. These registries were established within cities, states, provinces, countries, hospitals, or other well-defined administrative boundaries. Nonrandomized intervention studies were considered for inclusion, which involve those that cannot use randomization to allocate individuals or clusters of individuals into intervention groups, including quasi-experimental studies. Analytical observational studies, such as cohort studies and before-and-after studies, were also considered for inclusion. Only nonrandomized intervention studies and analytical observational studies could adequately address the research question that guided this systematic review, as a randomized controlled trial would not be feasible in this context.

Table . Eligibility criteria presented in the population, intervention, comparison, outcomes, and study structure.

Acronym	Description	Criteria
P	Population	Cities, states, provinces, countries, hospitals, or other well-defined administrative boundaries.[a]
I	Intervention	Implementation and use of injury registries, including trauma registries.
C	Comparison	No restrictions on comparators
O	Outcomes	Policymaking (changes and design, including preventive measures, health care improvements, surveillance actions, among others), hospitalization rates or duration, and mortality.
S	Study	Nonrandomized intervention studies or observational studies

[a]The population consists of the locations where injury registries have been implemented, with no restrictions on the patients served by health services regarding age, gender, or other attributes.

Studies were excluded if they were based on sources other than an injury or trauma registry. Articles without full-text availability were also excluded after attempts to contact the corresponding authors to request the full text. In addition, opinion articles, editorials, letters to the editor, conference abstracts, systematic or literature reviews, and any other studies that lacked empirical research regarding the effects of injury registries. The studies excluded were documented in an attached table within the systematic review. There were no restrictions on publication date, country, or language.

Information Sources, Search, and Selection Process

The search was conducted in the MEDLINE databases via PubMed, Embase, Lilacs via the Virtual Health Library, Scopus, and Web of Science. To ensure comprehensive literature coverage, the reference lists of studies included and relevant systematic or literature reviews identified through the search were also reviewed. The bibliographic search was performed in November 2023, with an update in March 2024.

The implemented search strategies are presented in Multimedia Appendix 1. The search terms were defined based on the population, intervention, and outcomes, following the PICOS structure. These terms include controlled vocabulary specific to each database, as well as uncontrolled vocabulary terms, synonyms, natural language, related terms, expressions, and truncations, all aimed at broadening the scope of the search. The strategies were developed in collaboration with a librarian experienced in conducting systematic review searches. Initially, the MEDLINE strategy was created with input from the research team and subsequently reviewed by a specialist with expertise in the research topic. Once the MEDLINE strategy was finalized, it was adapted to the syntax and subject headings of the other databases.

The literature search results were uploaded to Rayyan, where duplicates were eliminated, and an initial screening of abstracts and titles was conducted to identify potential studies for inclusion. Finally, the full texts of these potential studies were assessed to determine their eligibility. Adhering to PRISMA recommendations, both steps were performed independently by 2 reviewers [ACMS and LESL]. Any disagreements not resolved through discussion between the reviewers were arbitrated by a third reviewer [BZO]. The complete list of articles excluded during the full-text assessment is detailed in Multimedia Appendix 2, along with the documented reasons for study exclusions.

Data Collection Process and Data Items

Studies that passed the eligibility stage proceeded to data extraction, which was performed using a predefined Excel spreadsheet by the same independent reviewers [ACMS and LESL]. The following information was extracted from each study: title, year of publication, authors, corresponding author

XSL•FO
RenderX

contact, language, abstract, objectives, population (inclusion criteria, size, and characteristics), methods (type of study, time horizon, study site, comparators, and statistical significance), outcome measures (policies implemented, including health care improvements and preventive measures, hospitalization rates or duration, and mortality), funding, and conflicts of interest. Regarding the registries, the information collected included the type of registry (whether it was an injury or trauma registry), registry age, and specific focus of the registry. When disagreements arose, the full-text articles were reviewed again by the reviewers.

Data Synthesis

Summaries of study characteristics were presented in a table, and data synthesis was explored based on outcomes. The policy changes and designs, including preventive measures for injuries, health care improvements, surveillance actions, stakeholder partnerships, among others, were also synthesized in a table. For each hospitalization rate or duration, and for mortality, the intervention results were reported through number or percentage differences. A meta-analysis could not be conducted due to considerable methodological heterogeneity among the included studies and the small number of studies reporting the same outcome; consequently, publication bias could also not be assessed.

Quality Assessment

Two reviewers [ACMS and LESL] independently assessed the methodological quality of studies using the Newcastle-Ottawa Scale (NOS) [27] for cohort studies and an adaptation of this tool for cross-sectional studies [28]. This tool was deemed the most appropriate considering the study designs of interest in this review. According to the instructions of the NOS tool, the assessment was conducted considering the study as a whole, regardless of the specific outcomes analyzed in this systematic review.

The NOS is organized around 3 main dimensions of observational studies: participant selection, comparability between groups, and outcome evaluation. Quality assessment in NOS uses a star system. For cohort studies, up to 9 stars can be awarded across the 3 domains: selection (up to 4 stars), comparability (up to 2 stars), and outcome (up to 3 stars) [27]. For the adapted tool for cross-sectional studies, up to 10 stars can be awarded across 3 domains: selection (up to 5 stars), comparability (up to 2 stars), and outcome (up to 3 stars) [28].

To standardize study evaluation, specific criteria were defined for all studies. In the selection dimension, the item "Representativeness of the exposed cohort" considered the country of the registry to assess population representativeness. For the item "Was follow-up long enough for outcomes to occur" in the outcome dimension, studies were awarded a star if they had a minimum follow-up period of 3 years. The bias assessment involved independent evaluation by pairs (ACM-d-S and LESL), with discrepancies resolved by a third reviewer (TFGdSB).

Results

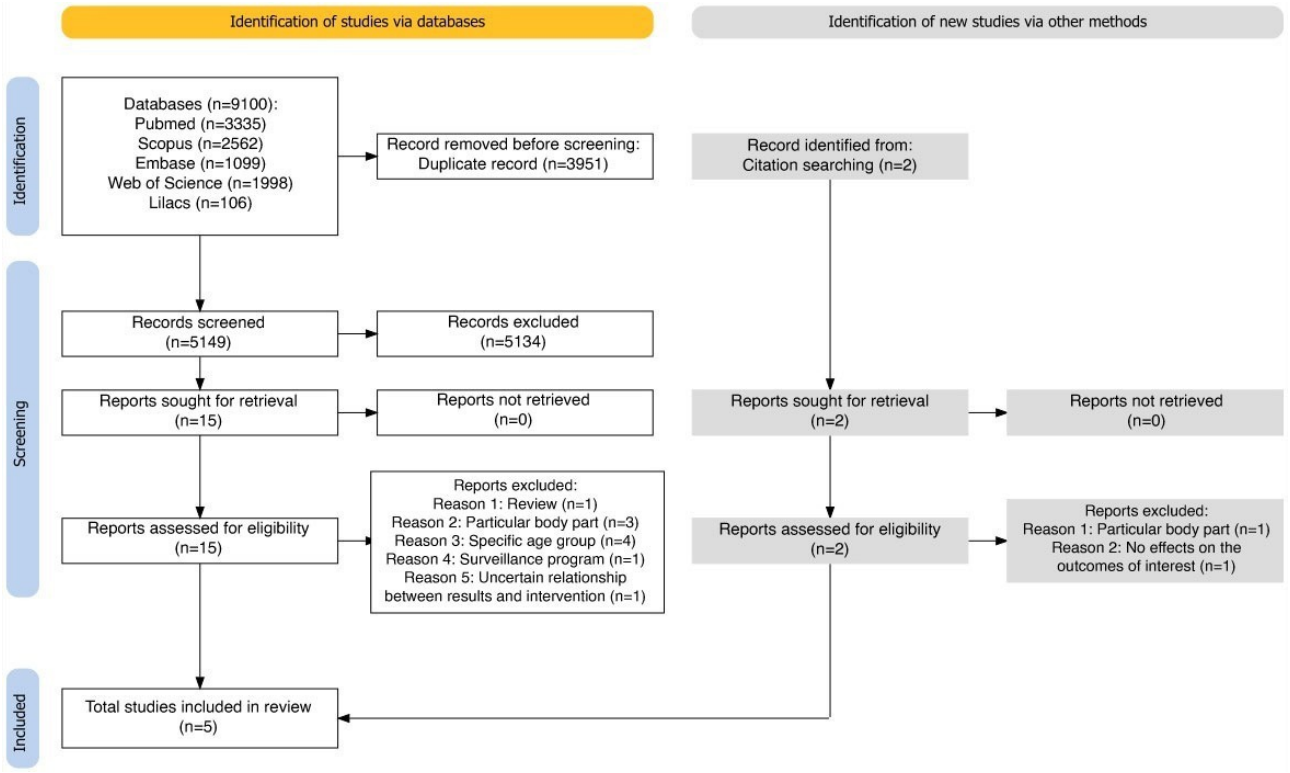
Study Selection

The database searches yielded 9100 potentially relevant records. After removing duplicates, 5149 unique records underwent screening based on their titles and abstracts. Subsequently, 15 full-text documents were examined, resulting in the inclusion of 5 papers [29-33]. In addition, a search was conducted in the references of the initial studies included. However, no additional articles meeting the inclusion criteria were identified through these searches (Figure 1).

A total of 12 studies were excluded from our review [34-45], with reasons for exclusion listed in Checklist 2. Among these exclusions, 1 study was disregarded due to its nature as a literature review [34]. In addition, 8 focused exclusively on a particular body part [38-41] or age group [36,42-44]. In one of them, the relationship between the reported results and the registry was not clear [35]. Another showed no associations on the outcomes of interest for this review [37]. Finally, 1 study presented results that appeared to extend beyond the implementation and use of data from registries; it reported findings from its surveillance program [45].

XSL•FO
RenderX

Figure 1. Flow diagram of the search and selection process of the studies included in the review.



Study Characteristics

The characteristics of each study included are provided in Tables 2 and 3. Of the 5 studies, 3 conducted in Germany involved data from trauma registries, with publication years ranging from 2004 to 2013 [29-31]. The remaining 2, one conducted in Cardiff, Wales, published in 2004 [32], and another in Harstad and Trondheim, Norway, published in 1995 [33], used data from injury registries.

Among the studies included, 4 used an observational cohort design, using a before-and-after approach [29-32]. One of these included a corresponding comparison group [29], while 2 others compared the obtained results with expected values as well [30,31]. Notably, 1 adopted a mixed design, incorporating a cross-sectional assessment involving hospitals to evaluate their efforts to improve care quality [31]. The fifth study classified itself as quasi-experimental; however, the outcome of interest for this review was assessed using a cohort [33].

Four of the studies described policymaking outcomes. Two, focusing on trauma registries, detailed health care improvements [29,31]. Meanwhile, 2 other centered on injury registries presented results on diverse approaches to injury management, including preventive measures against assault [32] and traffic injuries [33]. In addition, 2 studies reported associations of these policies on hospitalization [31,33], and 4 described the mortality outcomes [29-31,33].

XSL•FO
RenderX

Table . Characteristics of studies included.

First author, year (country)	Title	Registry type	Described outcomes
Helm M, 2013 (Germany) [29]	The Trauma Register DGU as the basis of medical quality management. Ten years' experience of a national trauma center exemplified by emergency room treatment	Trauma	Policymaking and mortality
Ruchholtz S, 2008 (Germany) [30]	Reduction in mortality of severely injured patients in Germany	Trauma	Mortality
Ruchholtz S, 2004 (Germany) [31]	External quality management in the clinical treatment of severely injured patients	Trauma	Policymaking, hospitalization, and mortality
Warburton AL, 2004 (Wales) [32]	Development, usage, and importance of accident and emergency department–derived assault data in violence management	Injury	Policymaking
Ytterstad B, 1995 (Norway) [33]	The Harstad Injury Prevention Study: evaluation of hospital-based injury recording and community-based intervention for traffic injury prevention	Injury	Policymaking, hospitalization, and mortality

XSL•FO
RenderX

Table . Details of studies included.

First author, year (country)	Study details				
	Study type	Time horizon	Population	Comparator	Funding
Helm M, 2013 (Germany) [29]	Observational (cohort)	Almost 10 years, divided into 4 periods: January 1998-December 1998; January 1999-December 2000; January 2001-December 2005; January 2006-March 2007.	Patients admitted via emergency rooms at a Level I trauma center (all patients: n=2239, and ISS[a] ≥16: n=1174)	First observed period and TR-DGU[b] hospitals data (ISS ≥16, national trauma centers from Germany)	Not explicitly indicated in the study
Ruchholtz S, 2008 (Germany) [30]	Observational (cohort)	7 years, from 1999 to 2005	Patients with ISS ≥16 admitted to 105 hospitals of the TR-DGU (n=11,013)	First observed period and values expected	not explicitly indicated in the study
Ruchholtz S, 2004 (Germany) [31]	Observational (cross-sectional+ cohort)	3 years, from 1999 to 2001	Hospitals transmitting data for ≥100 patients to TR-DGU (n=21) and patients with ISS ≥16, directly transferred from the accident site (n=2702)	First observed period and values expected	German Research Foundation and German Society for Trauma Surgery
Warburton AL, 2004 (Wales) [32]	Observational (cohort)	7 years, from April 1, 1995, to March 31, 2002, encompassing each annual period from April 1 to March 31.	All assault patients treated at the Cardiff Accident and Emergency (A&E) department	No comparator	Wales Office of Research and Development (WORD).
Ytterstad B, 1995 (Norway) [33]	Quasi-experimental+ cohort	7 years, from July 1, 1985, to June 30, 1993, was divided into three periods of 30 months each.	Individuals treated for traffic accident injuries in hospitals in two Norwegian cities, Harstad (n=22,970) and Trondheim (n=158,911).	Trondheim and first observed period	Norwegian Research Council for Science and the Humanities (NAVF) and Norwegian Directorate of Health/The Royal Norwegian Ministry of Social Affairs

[a]ISS: Injury Severity Score.

[b]TR-DGU: Trauma Register of the German Society of Trauma Surgery.

Outcomes

Policymaking

In total, 4 of the 5 studies included demonstrated the use of injury or trauma registry data as a policymaking tool. The key policies identified, including health care improvements and preventive measures, are synthesized in Table 4.

Two studies conducted in Germany involved data from trauma registries and demonstrated their use in continuously improving the care of hospitalized patients. One of these [29] shows the evolution of emergency room management in a national trauma center over a decade, marked by significant improvements across four distinct phases. These improvements resulted from ongoing monitoring of hospital data and comparisons with the Trauma Register of the German Society of Trauma Surgery (TR-DGU) to identify areas needing enhancement. Interventions included incorporating diagnostic exams and imaging procedures in the emergency room, refining personnel training protocols, and implementing time-efficient treatment algorithms. The other study [31] involved interviews with 21 hospitals regarding their efforts to improve care quality based on trauma registry data feedback. Optimization measures were introduced in the emergency room, surgical care, and intensive care. In the emergency room, changes included the installation of digital X-ray systems and spiral CT scanners. Organizational improvements focused on enhancing interdisciplinary coordination through quality circles and guidelines, as well as training and documentation. In surgical and intensive care, optimization targeted therapy areas, although specific details were not provided.

The 2 studies that described policymaking results related to injury registries presented preventive initiatives against assault and traffic injuries. For violence prevention, 1 study [32] detailed efforts by a multisector steering group, including assault awareness campaigns, targeted policing initiatives, and night transport management strategies. The use of data to guide judicial decisions, particularly for alcohol and entertainment license applications, was also highlighted. The authors concluded that using injury data from assault patients attending A&E departments in Cardiff, Wales, effectively targeted police

XSL•FO
RenderX

and local resources to reduce city-wide violence. The other study [33] also focused on traffic injury prevention through a multisector group's efforts. It described education initiatives targeting high-risk groups, traffic safety measures like installing stoplights and constructing separate pedestrian and cyclist roads, and enhanced traffic enforcement such as vehicle checks and speed limit enforcement. Legislative changes included assigning local health authorities the responsibility for accident prevention. The study emphasized using data to inform various stakeholders about traffic injury risks and prevention strategies. Conducted in Harstad, Norway, the data were actively used in a community-based injury prevention program.

Table . Key policies identified in studies.[a]

Policies	Studies
Trauma registry	
Continuous quality management in health care	3 [29,31]
Changes in emergency room management	2 [29,31]
Introducing exams and imaging diagnosis in the emergency room	2 [29,31]
Adjusting training and documentation	2 [29,31]
Other measures (eg, introducing therapy in the emergency room)	2 [29,31]
Changes in intensive care (organization, diagnosis, and therapeutic measures)	1 [31]
Changes in surgical care (organization, diagnosis, and therapeutic measures)	1 [31]
Injury registry	
Violence prevention by multiple sectors group	1 [32]
Assault awareness campaigns (eg, in schools and public libraries, and for bar and club managers)	1 [32]
Targeted policing initiatives (eg, a task force to hotspot areas, and targeting alcohol-related street crime)	1 [32]
Night transport management (eg, meetings to identify barriers and opportunities for education and staff protection)	1 [32]
Use of data to guide judicial decisions (eg, alcohol and entertainment license applications)	1 [32]
Traffic injury prevention by multiple sectors group	1 [33]
Traffic injury education initiatives and campaigns to high-risk groups (eg, in schools and driving schools, at health fairs and shopping centers, for parents of small children, for police and automobile societies, and through media campaigns)	1 [33]
Traffic safety measures (eg, installation of additional stoplights, constructing separate pedestrian and cyclist roads, planning safe pathways for children going to school, and loaning safety chairs for small children)	1 [33]
Enhanced traffic enforcement (eg, checks on vehicles, speed limit enforcement, and curfews for serving alcohol in bars and restaurants)	1 [33]
Legislative change (eg, a national law assigning local community health authorities' responsibility for accident injury prevention)	1 [33]
Use of data to inform educational institutions, city planners, private and public organizations, and the community	1 [33]

[a]Including health care improvements and preventive measures, among others.

Hospitalization

Although a meta-analysis for hospitalization results was planned, it could not be conducted due to the considerable methodological heterogeneity among the studies included and the small number of studies reporting this outcome (n=2). Since quantitative synthesis is not appropriate, the results observed by the authors are presented in narrative form below. Relevant findings identified in the studies can be found in Table 5.

XSL•FO
RenderX

Table . Hospitalization and mortality results identified in studies.

Outcome or studies	Follow-up	Results
Hospitalization		
Ruchholtz S, 2004 (Germany) [31]	1999 - 2001	• 3 days in the average length of stay in the ICU[a] among patients with an ISS[b] ≥16 (from 16 to 13 days, *P*<.05)
Ytterstad B, 1995 (Norway) [33]	1987 - 1992	• 42% per year reduction in admissions from traffic injuries (from 45 to 16 admissions)
Mortality		
Helm M, 2013 (Germany) [29]	1998 - 2007	• 6.0%[c] in mortality among all patients (from 10.8% to 4.8%, *P*<.05) • 20.1%[c] in mortality among patients with an ISS ≥16 (from 33.3% to 13.2%, *P*<.05) • 11.3%[c] greater reduction in mortality among ISS ≥16 patients compared with TR-DGU hospitals (−20.1% vs −8.8%)
Ruchholtz S, 2008 (Germany) [30]	1999 - 2005	• 4.1%[c] in observed mortality among patients with an ISS ≥16 (from 22.8% to 18.7%) • 3.4%[c] in expected mortality among patients with an ISS ≥16 (18.7% vs 22.1%)
Ruchholtz S, 2004 (Germany) [31]	1999 - 2001	• 4.0% in expected mortality among patients with an ISS ≥16 (17.5% vs 21.5%)
Ytterstad B, 1995 (Norway) [33]	1985 - 1993	• 66.7%[c] reduction in traffic-related deaths (from 6 to 2, *P*<.12)

[a]ICU: intensive care unit.

[b]ISS: Injury Severity Score.

[c]Values calculated by the systematic review authors based on study data.

Two studies described hospitalization results: 1 presented findings related to the implementation and use of a trauma registry, and the other presented results related to an injury registry. The first study [31], conducted in Germany, used continuous monitoring of hospital data to introduce optimization measures in the emergency room, surgical care, and intensive care. It compared the change in average values for the length of stay in the intensive care unit and total hospital stay over 3 years (1999 - 2001), involving data from 2702 patients with an Injury Severity Score (ISS)≥16 who were directly transferred from the accident site. The authors found no significant reduction in the total average length of hospital stay for patients (from 29 to 28 days). However, there was a significant reduction of 3 days in the average length of stay in the intensive care unit (from 16 to 13 days, *P*<.05).

In the second study [33], the injury registry was actively used in a community-based traffic injury prevention program conducted in Norway. The authors observed hospitalization results by comparing the 1987 sample of all admissions from traffic accidents at Harstad Hospital with the mean annual admissions during 1990 - 1992. As a result, a 42% per year reduction in admissions from traffic injuries was observed, decreasing from an average of 45 admissions per year to 26.

Mortality

The planned meta-analysis for the mortality could also not be conducted due to the considerable heterogeneity in terms of design and methods among the studies included that reported this outcome (n=4). Therefore, a quantitative synthesis was deemed inappropriate, and the findings from the studies are summarized in the text below. Relevant results identified in the studies can be found in Table 5.

Regarding the studies that focused on trauma registries, all were conducted in Germany and observed results on mortality. One study [29] implemented various interventions in the emergency room of a national trauma center through continuous monitoring of hospital data, including the refinement of personnel training protocols and the incorporation of diagnostic exams and imaging procedures. This study observed a significant reduction in hospital mortality among all patients, decreasing from 10.8% to 4.8% (−6.0%, *P*<.05), and specifically among patients with an ISS ≥16, decreasing from 33.3% to 13.2% (−20.1%, *P*<.05). The total observation period covered almost 10 years (January 1998 to March 2007), during which a total of 2239 trauma patients were admitted via the emergency room, including 1174 patients with ISS ≥16. Similar reductions were observed in hospitals part of the TR-DGU, where mortality rates among patients with an ISS ≥16 decreased by 8.8%, from 24.9% to

XSL•FO
RenderX

16.1%. However, the reduction in the studied hospital was notably more substantial, with an 11.3% greater decrease (−20.1% vs −8.8%).

Another study [30] evaluated patient data starting from 1999, which marked the beginning of receiving annual feedback in the form of reports comparing hospital data with TR-DGU results to identify areas for improvement and support continuous quality management. This study demonstrated that consistent feedback of treatment data to hospitals is associated with positive changes in survival prognosis. During the observation period (1999 to 2005), data from a total of 11,013 severely injured patients (ISS ≥16) in 105 hospitals were documented. Mortality significantly decreased by 4.1%, from 22.8% in 1999 to 18.7% in 2005, compared with an expected rate of 22.1% for that year, which represents a 3.4% reduction in expected mortality. The same author [31] had already noted that the observed mortality rate was lower than the expected rate (−4.0%, 17.5% vs 21.5%) after the introduction of optimization measures in the emergency room, surgical care, and intensive care in 21 hospitals that transmitted data to TR-DGU between 1999 and 2001. Despite reporting potential limitations due to the expected mortality calculation method being based on United States data and finding no significant change in overall mortality rates (16.6% vs 17.5%).

The Norway study [33] aimed to test the feasibility of an injury registry and evaluate its results within a community-based traffic injury prevention program over a follow-up period of 7.5 years, from July 1, 1985, to June 30, 1993, divided into 3 periods of 30 months each. The study observed a decrease in the number of traffic-related fatalities in Harstad, with a total of 10 resident deaths: 6 in period 1, 2 in period 2, and 2 in period 3 (P<.12). This represents a 66.7% reduction in traffic-related deaths from period 1 to period 3.

Quality Assessment

Among the 5 studies included in this review, 4 were rated as good quality [29-31,33], while only 1 study was assessed as fair quality [32]. The quality assessment results for the cohort studies are summarized in Multimedia Appendix 3, and the results for the cross-sectional studies are presented in Multimedia Appendix 4.

The only study classified as fair received just 2 stars in the assessment of the dimension of selection for exposed and unexposed cohorts. It was deemed not representative of the general population, specifically the country where the registry is located, as outlined in the methodology of this review. In addition, it lacked comparability of results with an unexposed cohort. While the NOS assigns higher scores to cohort studies that include nonexposed groups, similar to the other 2 studies in this review [30,31], this study did not aim to compare its results with a nonexposed cohort, as it was designed as a before-and-after study.

The analysis of evidence certainty using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) tool was not conducted due to the designs and objectives of the studies. The diverse methodologies used, along with their focus on practical implications rather than outcomes as such, limited the applicability of GRADE.

Discussion

Principal Findings

The findings of this review suggest that injury and trauma registries have the potential to inform policymaking, including preventive measures, health care improvements, and surveillance actions, which can lead to enhanced health outcomes. The studies indicate that trauma registries are primarily designed to provide information that can enhance the efficiency and quality of care for hospitalized patients, with findings including reductions in intensive care unit stay duration and the number of deaths [29-31]. Meanwhile, injury registries are used in community-based injury prevention to identify at-risk populations and health needs, as well as to plan and implement prevention policies that address local priorities [32,33]. One of the 2 studies on injury registries demonstrated reductions in traffic-related hospital admissions and deaths [33]. The other, while not reporting a decrease in these outcomes, noted a reduction in the overall number of assault cases attending the emergency room in the sixth and seventh years of the registry [32]. These studies also emphasized the importance of involving multiple sectors, beyond just health, in injury prevention efforts. All studies highlighted the value of registries for scientific research.

However, it is important to note that the vast majority of studies retrieved from our search used registry data to generate scientific knowledge rather than to assess the impact or effects of the registry itself. For instance, 1 study aimed to identify predictors of prehospital versus hospital mortality due to road traffic injuries in an Iranian population, using data from an integrated road traffic injury registry [46]. Another study examined the epidemiology of occupational injury hospitalizations in the United Arab Emirates using a trauma registry [47]. While some studies evaluated the results of implemented policies and measures, such as one that presented significant trends and developments in acute trauma care over 20 years [35], they did not intend or did not establish causality related to the implementation and use of registry data. Consequently, despite the large number of studies identified in our search, only a few advanced to the full reading phase, as many were excluded during the screening stage after reviewing their titles and abstracts.

An important observation from this review is the notable scarcity of studies dedicated to exploring the effects or results of the use of registry data on injury outcomes. In addition, a crucial point to highlight when interpreting the results is that even the few studies included in the review reflect correlations rather than causalities. This is particularly important as the studies did not adopt designs that allowed for the control of external factors, making it impossible to conclude that the observed results can be attributed solely to the implementation and use of injury or trauma registry data. Currently, there are no impact publications.

XSL•FO
RenderX

Limitations

The limitations of our review include the cautious search strategy, which may have excluded studies assessing surveillance information systems or other types that capture injury data, or that collected data solely for that specific research project. We were particularly careful not to subjectively define which studies originated from injury registries. Therefore, we focused on studies that explicitly identified their data sources as such. Our conservative approach focused on the terms "injury registry" and "trauma registry" to ensure the inclusion of studies containing structured registry data with direct relevance to our research question. For example, a study evaluating the impact of the Life in Traffic Project—implemented in 2010 across various Brazilian capitals to reduce traffic-related deaths and serious injuries, which is based on an integrated analysis of data from multiple sectors involved in road traffic injury issues—was not included in our systematic review [48]. The study used interrupted time series models to compare mortality rates before and after the project's implementation, revealing positive outcomes in some cities but concerning increases in others. It highlights the necessity of ongoing evaluations through data registries to enhance public policy analysis and monitoring. Despite its relevance, the tool implemented by this project did not function as an injury or trauma registry. This limitation has prompted the Brazilian Ministry of Health to organize a workshop aimed at establishing a national concept for accident and violence registries, which took place in October 2024.

Although the NOS was considered the most appropriate tool for analyzing the methodological quality of the studies included in this review, we still identified limitations in its application. Most of the studies included did not aim to compare their results with a nonexposed cohort, as they were designed as before-and-after studies. This led to one of them being classified as "fair" because the NOS assigns higher scores to cohort studies that include nonexposed groups. In addition, we had to use an adapted version of the tool for cross-sectional studies, and an analysis of evidence certainty using the GRADE tool was not conducted due to the designs and objectives of the studies. Furthermore, it is crucial to acknowledge the potential influence of publication bias, where studies yielding positive outcomes are more likely to be published, thereby shaping the landscape of the literature.

The studies included collectively provided insights into the results of trauma and injury registries on reducing hospitalizations and mortality; however, variations in methodology and context limited the ability to draw synthesized results through meta-analysis.

Comparison With Previous Work

To our knowledge, this is the first systematic review aimed at mapping the effects and results related to registry data, including those that encompass all types of injuries, regardless of their severity. Previous research has predominantly focused on this relationship with trauma systems, as evidenced by a meta-analysis that demonstrated a 15% reduction in mortality in favor of the presence of a trauma system [22], as well as by another systematic review that sought to identify which components contributed to its effectiveness [21]. In addition, we can mention a study that exclusively addressed the most common functions of trauma registry data in the formulation of health policies but did not examine its effects [13].

Conclusions

This study is especially timely, given the current efforts to develop an injury registry in Brazil. It addressed the existing knowledge gap related to the results of injury and trauma registries on policymaking and injury outcomes. Highlighting our acknowledgment that the registry alone does not directly influence morbidity and mortality; rather, these outcomes are indirect results of evidence-based policies and preventive measures implemented [13,34,49]. Besides, we emphasize that the findings of the studies included represent associations, making it essential to temper expectations regarding the causalities of their effects, as well as to exercise caution in the transportability of results from European studies to the Brazilian context.

Factors such as distinct social, economic, and cultural characteristics, access to health care, trauma burden, practices related to injury management and policies, along with the handling and use of data from the implemented injury registry, may influence the outcomes observed in studies conducted in Europe. Public health decision-makers should consider the differences in the distribution of potential effect modifiers when interpreting the results and their applicability in Brazil. Some studies discuss limitations and adjustments that can be applied through the combination of assumptions and statistical approaches to ensure the unbiased transfer of results obtained from different populations [50,51].

Furthermore, the tool being developed is expected to be low-cost, primarily based on the unification and sharing of existing information systems that currently lack interoperability. This approach represents one of the solutions identified in a study published in 2018 to address the challenges of implementing trauma registries in low- and middle-income countries [49]. Another study highlighted that trauma registries are rarely found in mid to low-income countries due to the lack of digital resources needed to maintain a costly registry. According to the author, until 2017, many countries with a significant trauma burden—including India, South Africa, Spain, China, and Brazil—had yet to develop nationwide trauma registries [52]. The national injury registry currently in development for Brazil aims not only to reduce costs but also to enhance the quality of data and analysis. The main objective is to strengthen epidemiological surveillance to support policies and preventive measures against accidents and violence.

In conclusion, the establishment of an injury registry in Brazil presents a significant opportunity to improve health outcomes through informed policymaking. While it is crucial to set appropriate expectations regarding its effects on morbidity and mortality, its role in facilitating preventive measures and enhancing surveillance capabilities remains valuable.

Future reviews should aim to broaden the scope of inquiry to include studies derived from other data sources, even if they were not explicitly identified as injury or trauma registries, but which still align with the concept of accident and violence

XSL•FO
RenderX

registries defined by Brazil after our study. In addition, it would be beneficial to conduct primary studies that examine the effects of these registries on the quality of surveillance. Without neglecting to mention that it is essential to apply methodologies that allow for the evaluation of the causal impact on health outcomes, particularly in the Brazilian context.

Acknowledgments

The authors express their sincere gratitude to Edna Terezinha Rother, the librarian responsible for developing the search strategies for the databases. This project was supported by Hospital Israelita Albert Einstein through the Programa de Apoio ao Desenvolvimento Institucional do Sistema Único de Saúde (PROADI-SUS 25000.159100/2023-77)[53], a strategy of the Brazilian Ministry of Health. The funding covered costs associated with the development and publication of this manuscript.

Authors' Contributions

ACMdS, LRC, and BZdO contributed to the conception of the study. ACMdS, TFGdSB, and LESL defined the design of the systematic review and the methods for data analysis. ACMdS and LESL drafted the manuscript, while BZdO, LRC, NNBdS, and TFGdSB provided feedback and revised the final version.

Conflicts of Interest

The authors are involved in PROADI-SUS TRAUMA, the project discussed in this article, which has the objective of establishing an injury registry for Brazil.

Multimedia Appendix 1
Search strategy.
[DOCX File, 26 KB - publichealth_v11i1e67115_app1.docx]

Multimedia Appendix 2
Full-text evaluated and studies excluded.
[DOCX File, 28 KB - publichealth_v11i1e67115_app2.docx]

Multimedia Appendix 3
Newcastle-Ottawa Scale quality assessment for cohort studies.
[DOCX File, 25 KB - publichealth_v11i1e67115_app3.docx]

Multimedia Appendix 4
Newcastle-Ottawa Scale quality assessment for cross-sectional studies.
[DOCX File, 22 KB - publichealth_v11i1e67115_app4.docx]

Checklist 1
PRISMA 2020 main checklist.
[DOCX File, 39 KB - publichealth_v11i1e67115_app5.docx]

Checklist 2
PRISMA abstract checklist.
[DOCX File, 30 KB - publichealth_v11i1e67115_app6.docx]

References

1. Injuries and violence. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/injuries-and-violence [accessed 2025-08-21]
2. Mortalidade – desde 1996 pela CID-10 [Web page in Portugese]. DATASUS - Ministério da Saúde. 2023. URL: https://datasus.saude.gov.br/mortalidade-desde-1996-pela-cid-10 [accessed 2025-08-21]
3. Malta DC, Minayo MDS, Soares AM Filho, et al. Mortality and years of life lost by interpersonal violence and self-harm: in Brazil and Brazilian states: analysis of the estimates of the Global Burden of Disease Study, 1990 and 2015. Rev Bras Epidemiol 2017 May;20Suppl 01(Suppl 01):142-156. [doi: 10.1590/1980-5497201700050012] [Medline: 28658379]
4. Morbidade hospitalar do SUS por causas externas - por local de residência [Web page in Portugese]. Ministério da Saúde. URL: http://tabnet.datasus.gov.br/cgi/deftohtm.exe?sih/cnv/fruf.def [accessed 2025-08-21]
5. Anexo 1 do anexo VII da portaria de consolidação nº 2, de 28 de setembro de 2017. Consolidação das normas sobre as políticas nacionais de saúde do sistema único de saúde [Report in Portugese]. : Brasília: Diário Oficial da República

XSL•FO
RenderX

Federativa do Brasil; 2017 URL: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2017/prc0002_03_10_2017.html#ANEXOVII [accessed 2025-08-21]
6. Portaria nº 1863, de 29 de setembro de 2003. institui a política nacional de atenção às urgências, a ser implantada em todas as unidades federadas, respeitadas as competências das três esferas de gestão [Report in Portugese]. : Brasília: Diário Oficial da República Federativa do Brasil; 2003 URL: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2003/prt1863_26_09_2003.html [accessed 2025-08-21]
7. Portaria n. : Diário Oficial da República Federativa do Brasil; 2004 URL: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2004/prt0936_19_05_2004.html
8. Portaria nº 1.356, de 23 de junho de 2006 institui incentivo aos estados, ao distrito federal e aos municípios para a vigilância de acidentes e violências em serviços sentinela, com recursos da secretaria de vigilância em saúde (svs) [Report in Portugese]. : Diário Oficial da República Federativa do Brasil; 2006 URL: https://www.saude.sp.gov.br/resources/cve-centro-de-vigilancia-epidemiologica/areas-de-vigilancia/doencas-cronicas-nao-transmissiveis/observatorio-promocao-a-saude/portarias/portaria_gm1356_2006.pdf [accessed 2025-08-21]
9. Portaria interministerial nº 2.268, de 10 de agosto de 2010 institui a comissão nacional interministerial para acompanhamento da implantação e implementação do projeto vida no trânsito [Report in Portugese]. : Diário Oficial da República Federativa do Brasil; 2010 URL: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2010/pri2268_10_08_2010_rep.html [accessed 2025-08-21]
10. Abulatif LI, Grupo Técnico de Parceiros do Projeto Vida no Trânsito. Data Integation Process: an information management model for multiple databases on traffic accidents in Brazil. Epidemiol Serv Saude 2018 Jun 28;27(2):e2017160. [doi: 10.5123/S1679-49742018000200018] [Medline: 29995103]
11. Portaria de consolidação nº 4, de 28 de setembro de 2017 [Report in Portugese]. : Diário Oficial da República Federativa do Brasil; 2017 URL: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2017/prc0004_03_10_2017.html [accessed 2025-08-21]
12. Moore L, Clark DE. The value of trauma registries. Injury 2008 Jun;39(6):686-695. [doi: 10.1016/j.injury.2008.02.023] [Medline: 18511052]
13. Mobinizadeh M, Berenjian F, Mohamadi E, et al. Trauma registry data as a policy-making tool: a systematic review on the research dimensions. Bull Emerg Trauma 2022 Apr;10(2):49-58. [doi: 10.30476/BEAT.2021.91755.1286] [Medline: 35434165]
14. O'Reilly GM, Cameron PA, Joshipura M. Global trauma registry mapping: a scoping review. Injury 2012 Jul;43(7):1148-1153. [doi: 10.1016/j.injury.2012.03.003] [Medline: 22483995]
15. Differences between an injury and a trauma. McCarthy + Co LLP. 2023. URL: https://mccarthy.ie/blog/personal-injury-claims/differences-between-an-injury-and-a-trauma/#:~:text=What [accessed 2025-08-21]
16. Unified Fast Data Access Technology for Accident Mitigation (TRAUMA) [Web page in Portugese]. PROADI. 2023. URL: https://hospitais.proadi-sus.org.br/projeto/tecnologia-de-rapido-acesso-de-dados-unificado-para-mitigacao-da-acidentalidade [accessed 2025-08-21]
17. Saúde vai integrar dados sobre acidentes, lesões e violência no país: Projeto Trauma deve ser instalado no Ministério da Saúde até 2026 [Web page in Portugese]. Agência Brasil. 2024. URL: https://agenciabrasil.ebc.com.br/saude/noticia/2024-03/saude-vai-integrar-dados-sobre-acidentes-lesoes-e-violencia-no-pais [accessed 2025-08-21]
18. St-Louis E, Paradis T, Landry T, Poenaru D. Factors contributing to successful trauma registry implementation in low- and middle-income countries: a systematic review. Injury 2018 Dec;49(12):2100-2110. [doi: 10.1016/j.injury.2018.10.007] [Medline: 30333086]
19. O'Reilly GM, Joshipura M, Cameron PA, Gruen R. Trauma registries in developing countries: a review of the published experience. Injury 2013 Jun;44(6):713-721. [doi: 10.1016/j.injury.2013.02.003] [Medline: 23473265]
20. Sleat GKJ, Ardolino AM, Willett KM. Outcome measures in major trauma care: a review of current international trauma registry practice. Emerg Med J 2011 Dec;28(12):1008-1012. [doi: 10.1136/emermed-2011-200326] [Medline: 21890862]
21. Moore L, Champion H, Tardif PA, et al. International injury care improvement initiative impact of trauma system structure on injury outcomes: a systematic review and meta-analysis world. World J Surg 2018 May(5):1327-1339. [doi: 10.1007/s00268-017-4292-0]
22. Celso B, Tepas J, Langland-Orban B, et al. A systematic review and meta-analysis comparing outcome of severely injured patients treated in trauma centers following the establishment of trauma systems. J Trauma 2006 Feb;60(2):371-378. [doi: 10.1097/01.ta.0000197916.99629.eb] [Medline: 16508498]
23. Hoepelman RJ, Driessen MLS, de Jongh MAC, et al. Concepts, utilization, and perspectives on the Dutch Nationwide Trauma registry: a position paper. Eur J Trauma Emerg Surg 2023 Aug;49(4):1619-1626. [doi: 10.1007/s00068-022-02206-4] [Medline: 36624221]
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
25. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160. [doi: 10.1136/bmj.n160] [Medline: 33781993]

XSL•FO
RenderX

26. Medeiros-de-Souza AC, Lopes LES, de Oliveira BZ, Rother ET, Correia LR. Effects of injury registry data on policy making, hospitalizations, and mortality: protocol for a systematic review and meta-analysis. JMIR Res Protoc 2024 Oct 30;13:e55029. [doi: 10.2196/55029] [Medline: 39475822]
27. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute. URL: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2025-09-09]
28. Modesti PA, Reboldi G, Cappuccio FP, et al. ESH working group on CV risk in low resource settings. panethnic differences in blood pressure in Europe: a systematic review and meta-analysis. PLoS ONE 2016 Jan 25(1):e0147601. [doi: 10.1371/journal.pone.0147601]
29. Helm M, Bitzl A, Klinger S, Lefering R, Lampl L, Kulla M. The TraumaRegister DGU® as the basis of medical quality management. Ten years experience of a national trauma centre exemplified by emergency room treatment. Unfallchirurg 2013 Jul;116(7):624-632. [doi: 10.1007/s00113-012-2251-7] [Medline: 22971955]
30. Ruchholtz S, Lefering R, Paffrath T, et al. Reduction in mortality of severely injured patients in Germany. Dtsch Arztebl Int 2008 Mar;105(13):225-231. [doi: 10.3238/arztebl.2008.0225] [Medline: 19629200]
31. Ruchholtz S, German Society of Trauma Surgery. External quality management in the clinical treatment of severely injured patients. Unfallchirurg 2004 Oct;107(10):835-843. [doi: 10.1007/s00113-004-0814-y] [Medline: 15300327]
32. Warburton AL, Shepherd JP. Development, utilisation, and importance of accident and emergency department derived assault data in violence management. Emerg Med J 2004 Jul;21(4):473-477. [Medline: 15208234]
33. Ytterstad B, Wasmuth HH. The Harstad Injury Prevention Study: evaluation of hospital-based injury recording and community-based intervention for traffic injury prevention. Accid Anal Prev 1995 Feb;27(1):111-123. [doi: 10.1016/0001-4575(94)00036-l] [Medline: 7718072]
34. Bouillon B, Lefering R, Paffrath T, Sturm J, Hoffmann R. Treatment of severely injured patients: Impact of the German Trauma Registry DGU®. Unfallchirurg 2016 Jun;119(6):469-474. [doi: 10.1007/s00113-016-0193-1] [Medline: 27240850]
35. TraumaRegister DGU. 20years of trauma documentation in Germany—Actual trends and developments. Injury 2014 Oct;45:S14-S19. [doi: 10.1016/j.injury.2014.08.012]
36. Weiner EJ, Tepas JJ. Application of electronic surveillance and global information system mapping to track the epidemiology of pediatric pedestrian injury. J Trauma 2009 Mar;66(3 Suppl):S10-S16. [doi: 10.1097/TA.0b013e3181937bc8] [Medline: 19276720]
37. Isles S, Christey G, Civil I, Hicks P. The New Zealand Major Trauma Registry: the foundation for a data-driven approach in a contemporary trauma system. N Z Med J 2017 Oct 6;130(1463):19-27. [Medline: 28981491]
38. Holstein JH, Stuby FM, Herath SC, Culemann U, Aghayev E, Pohlemann T. Einfluss des Beckenregisters der DGU auf die Versorgung von Beckenringfrakturen. Unfallchirurg 2016 Jun;119(6):475-481. [doi: 10.1007/s00113-016-0168-2]
39. Hauschild O, Strohm PC, Culemann U, et al. Mortality in patients with pelvic fractures: results from the German pelvic injury register. J Trauma 2008 Feb;64(2):449-455. [doi: 10.1097/TA.0b013e31815982b1] [Medline: 18301214]
40. Richard-Denis A, Gravel LF, Dionne A, et al. An evaluation of the representativeness of a national spinal cord injury registry: a population-based cohort study. Spinal Cord 2021 Oct;59(10):1072-1078. [doi: 10.1038/s41393-021-00622-8] [Medline: 33828247]
41. Noonan VK, Kwon BK, Soril L, et al. The Rick Hansen Spinal Cord Injury Registry (RHSCIR): a national patient-registry. Spinal Cord 2012 Jan;50(1):22-27. [doi: 10.1038/sc.2011.109] [Medline: 22042297]
42. Pressley JC, Barlow B, Durkin M, Jacko SA, Dominguez DR, Johnson L. A national program for injury prevention in children and adolescents: the injury free coalition for kids. J Urban Health 2005 Sep;82(3):389-402. [doi: 10.1093/jurban/jti078] [Medline: 15958785]
43. Ytterstad B. The Harstad injury prevention study: community based prevention of fall-fractures in the elderly evaluated by means of a hospital based injury recording system in Norway. J Epidemiol Community Health 1996 Oct;50(5):551-558. [doi: 10.1136/jech.50.5.551] [Medline: 8944864]
44. Ytterstad B, Smith GS, Coggan CA. Harstad injury prevention study: prevention of burns in young children by community based intervention. Inj Prev 1998 Sep;4(3):176-180. [doi: 10.1136/ip.4.3.176] [Medline: 9788086]
45. Razzak JA, Shamim MS, Mehmood A, Hussain SA, Ali MS, Jooma R. A successful model of road traffic injury surveillance in a developing country: process and lessons learnt. BMC Public Health 2012 May 16;12:357. [doi: 10.1186/1471-2458-12-357] [Medline: 22591600]
46. Homayoun SB, Milad JD, Mina G, Parvin S. Predictors of pre-hospital vs. hospital mortality due to road traffic injuries in an Iranian population: results from Tabriz integrated road traffic injury registry. BMC Emerg Med 2022 Mar 8;22(1):37. [doi: 10.1186/s12873-022-00593-w] [Medline: 35260101]
47. Barss P, Addley K, Grivna M, Stanculescu C, Abu-Zidan F. Occupational injury in the United Arab Emirates: epidemiology and prevention. Occup Med (Lond) 2009 Oct;59(7):493-498. [doi: 10.1093/occmed/kqp101] [Medline: 19640929]
48. Figueiredo Filho DB. Avaliação de impacto do projeto vida no trânsito [Report in Portugese]. : Escola Nacional de Administração Pública (Enap) URL: http://repositorio.enap.gov.br/handle/1/6962 [accessed 2025-08-21]

XSL•FO
RenderX

49. Bommakanti K, Feldhaus I, Motwani G, Dicker RA, Juillard C. Trauma registry implementation in low- and middle-income countries: challenges and opportunities. J Surg Res 2018 Mar;223:72-86. [doi: 10.1016/j.jss.2017.09.039] [Medline: 29433888]
50. Westreich D, Edwards JK, Lesko CR, Cole SR, Stuart EA. Target validity and the hierarchy of study designs. Am J Epidemiol 2019 Feb 1;188(2):438-443. [doi: 10.1093/aje/kwy228] [Medline: 30299451]
51. Pearl J, Bareinboim E. External validity: from do-calculus to transportability across populations. Statist Sci 2014 Nov;29(4):579-595. [doi: 10.1214/14-STS486]
52. Choi J, Carlos G, Nassar AK, Knowlton LM, Spain DA. The impact of trauma systems on patient outcomes. Curr Probl Surg 2021 Jan;58(1):100849. [doi: 10.1016/j.cpsurg.2020.100849] [Medline: 33431134]
53. Extrato de ajuste [Web page in Portugese]. Diário Oficial da União. URL: https://www.in.gov.br/web/dou/-/extrato-de-ajuste-533243037 [accessed 2025-08-21]

Abbreviations

GRADE: Grading of Recommendations Assessment, Development and Evaluation
ISS: Injury Severity Score
NOS: Newcastle-Ottawa Scale
PICOS: population, intervention, comparison, outcomes, and study
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROADI-SUS: Programa de Apoio ao Desenvolvimento Institucional do Sistema Único de Saúde
PROSPERO: Prospective Register of Systematic Reviews
TR-DGU: Trauma Register of the German Society of Trauma Surgery
WHO: World Health Organization

Edited by KS Sahu; submitted 02.10.24; peer-reviewed by Y Sarikhani, Z He; revised version received 28.04.25; accepted 21.05.25; published 10.09.25.

Please cite as:
Medeiros-de-Souza AC, Sinhori Lopes LE, Felicissimo Gomes de Souza Bandeira T, Reis Correia L, Bandeira de Sá NN, Zocca de Oliveira B
Effects of Injury Registry Data on Policymaking, Hospitalizations, and Mortality: Systematic Review
JMIR Public Health Surveill 2025;11:e67115
URL: https://publichealth.jmir.org/2025/1/e67115
doi:10.2196/67115

© Ana Cláudia Medeiros-de-Souza, Luana Emanuelly Sinhori Lopes, Tayna Felicissimo Gomes de Souza Bandeira, Lucas Reis Correia, Naiza Nayla Bandeira de Sá, Bruno Zocca de Oliveira. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Pathological Processes Among Content Creators on Social Media: Scoping Review

Sergio Martinez-Aguirre[1], MD; Javier Sanz-Valero[2], MPH, PhD; Elena Ronda-Pérez[3], MD, PhD

[1]Doctoral Program in Health Sciences, University of Alicante, Alicante, Spain

[2]National School of Occupational Medicine, Carlos III Health Institute, Monforte de Lemos Street, 5, Madrid, Spain

[3]Public Health Research Group, Department of Community Nursing, Preventive Medicine and Public Health and History of Science, University of Alicante, CIBER of Epidemiology and Public Health, Alicante, Spain

Corresponding Author:
Javier Sanz-Valero, MPH, PhD
National School of Occupational Medicine, Carlos III Health Institute, Monforte de Lemos Street, 5, Madrid, Spain

Abstract

Background: Content creators (CCs), like any other worker, are exposed to various occupational hazards that can affect their physical, mental, and social well-being, with psychosocial and ergonomic risks being particularly relevant. The combination of prolonged work hours, sedentary lifestyles, excessive public scrutiny, and often job insecurity and unpredictability (manifested as continuous connectivity and anticipation of sporadic tasks) presents a significant risk for the development of health issues.

Objective: This study reviews the scientific literature to identify the potential pathological processes affecting CCs on social media.

Methods: The scoping review method was used. Data were obtained from the following bibliographic databases: MEDLINE (via PubMed), Embase, Cochrane Library, PsycINFO, Scopus, Web of Science, and Virtual Health Library. The terms used as descriptors and in the title and abstract fields were “Content Creator” and “Pathologic Processes.” The search was conducted in May 2024. Agreement between authors for paper selection was measured using the Cohen κ coefficient. The documentary quality of the papers was assessed using the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) questionnaire, and the level of evidence and recommendation grade were determined according to the Scottish Intercollegiate Guidelines Network recommendations. Bias was evaluated using the Risk of Bias in Non-randomized Studies of Exposures (ROBINS-E) tool.

Results: Of the 1522 references retrieved, 6 papers were selected based on the inclusion and exclusion criteria. Of the 6 studies reviewed, 3 were exclusively focused on a single gender. The agreement on the relevance of the selected studies, calculated using the κ index, was 84.9% (P<.01). The study population ranged from a minimum of 6 to a maximum of 1544 participants. The STROBE scores ranged from 81.3% to 96.8%, with a median of 14.9% (IQR 2.1). According to the Scottish Intercollegiate Guidelines Network criteria, this review provided evidence level 2++ with a recommendation grade of B. ROBINS-E highlighted a higher number of biases in Domains 5, 6, and 7. All interventions were based on interviews, either conducted online or via email. Participant activities, as documented in the respective studies, comprised influencer roles (n=2), blogging (n=2), YouTube content creation (n=1), and live streaming (n=1). The design of the reviewed works comprised 4 qualitative studies and 2 mixed methods (qualitative and quantitative) studies. The reported health impacts were diverse, comprising burnout (n=2), anxiety (n=1), co-occurring anxiety and depression (n=1), eating disorder (n=1), chronic pain (n=1), and unspecified mental health issues (n=1). All studies highlighted the necessity for further investigation into potential pathological processes among CCs engaged in social media activities.

Conclusions: It was found that the most affected area was mental health, as observed in nearly all the reviewed studies. Despite the extensive documentation of mental health impact, it is necessary to identify the risk factors associated with the pathological processes of CCs to prevent the signs and symptoms identified in this literature review.

(JMIR Public Health Surveill 2025;11:e76708) doi:10.2196/76708

KEYWORDS

pathological processes; social media; content creator; psychological phenomena; postural balance

XSL•FO
RenderX

Introduction

The rise of social media is primarily supported by the creation of content that appeals to the public. These social media platforms have become not only a means of social entertainment but also a commercial medium where brands offer their products and services through individuals with a significant ability to influence a broad audience.

In this framework, the role of content creation (CC) stands out as a new form of employment. The emergence of this labor role, popularly known as the "influencer," led the European Union to enact the Digital Services Act, which introduced regulations for digital platforms and CCs to combat illegal content and misinformation [1]. This regulation prompted the Spanish government to introduce specific legislation that established the rights and obligations for those engaging in this profession [2].

Moreover, Spain does not currently possess specific legislation governing the working hours of CCs, nor is there any European regulation directly addressing the working time (broadcast duration) of CCs.

Thus, the exercise of this activity, as indicated by a study from the University of Valencia [3], is based on the economic performance derived from the advertising impact it has on social media. In other words, it connects commercial brands with their target audience through their ability to influence a community. Collaboration with CCs is now part of the global strategy of brands [4]. To achieve this, CCs must devote much of their day not only to content creation but also to promoting their work and engaging with sponsoring companies.

In the context of their professional activity, like any other worker, CCs are exposed to various occupational risks that can affect their physical, mental, and social well-being [5]. Particularly relevant are psychosocial and ergonomic risks. Psychosocial risks are related to factors arising from the interaction between work, individuals, and the social context: stress, pressure to maintain a constant presence on digital platforms, or public exposure can lead to anxiety or emotional disorders. Regarding ergonomic risks, these relate to the organization and conditions of the work environment: the use of data display screens, forced postures, or repetitive movements. Identifying, evaluating, and preventing these risks must be the foundation of this labor sector to avoid future health problems in CCs.

Although the negative health implications of social media engagement are recognized, there is limited understanding regarding the precise role social media plays in the development of associated pathological conditions [6]. Thus, this cocktail of marathon workdays, sedentary behavior, public overexposure, and, in many cases, job insecurity and uncertainty (being constantly connected and awaiting tasks without predictability) results in health problems, ranging from mild to severe [7].

Similarly, CCs have been associated with components of addiction (prominence, tolerance, mood modification, relapses, withdrawal, and conflict). However, Peng and Liao [8] have critiqued the ability to differentiate between problematic and highly active CCs. Recently, a novel media phenomenon termed "pathostreaming" has gained prominence. "Pathostreaming" refers to sharing misleading or harmful content (eg, obscenity and hate speech) online, alongside the propagation of hate speech across various digital platforms. The distinct characteristics of "pathostreaming" have led to the emergence of new media entities: "pathostreamers," "pathoinfluencers," and "pathousers." A critical area of ongoing research focuses on the mental health of these individuals and their influence on their followers [9].

There is no doubt that work through digital platforms is growing in importance; however, it also presents significant challenges for the health and safety of its workers. Specifically, all the risks involved in the activities will be present (transport, cleaning, workplace adaptation, home-based work, etc), but these are exacerbated by the unique characteristics of work on platforms [10,11].

CCs are often expected to produce quality work while managing their own stress. The pressure of the creative process, combined with the challenge of meeting deadlines, can lead to burnout. One of the main forms of digital content creation is the live streaming of video games. The study noted that video game players and those engaged in electronic sports were at a higher risk of developing numerous chronic diseases and experiencing increased mortality, highlighting a concerning reality in this research [12].

At the same time, participation on this Web 2.0 platform, with prolonged exposure as an idealized image, was correlated with depressive symptoms, self-esteem problems, anxiety, and body dissatisfaction [13]. Furthermore, this overexposure may occasionally result in the reception of hate speech. The experience of being targeted by hate speech is distressing for content creators (CCs); this sense of attack can diminish their self-perception and self-esteem, thereby compromising their well-being [14]. In this line, Fung et al [15] conducted a systematic review that highlighted the importance of social media in public health issues. They emphasized how the generated content influenced both the well-being of creators and viewers.

All of the above justifies the need for this review. In recent years, social media has experienced a remarkable surge, giving rise to new forms of employment. Among these, CCs stand out as a rapidly growing and increasingly influential group across various sectors. This professional field has evolved and diversified, attracting a significant portion of the active workforce, particularly among younger generations.

Therefore, the central aim of this study is to critically analyze the available scientific bibliography, seeking to understand the potential pathological processes derived from CC within the realm of social networks and the interventions undertaken, with the ultimate aim of establishing the nature of their causal link.

XSL•FO
RenderX

Methods

Design

This is a descriptive cross-sectional study and a critical analysis of studies retrieved through a systematic technique. The structure of this review followed the guidelines of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [16] and the methodological framework proposed by Arksey and O'Malley [17] for scoping reviews.

Protocol Registration

The protocol of this review was registered in the institutional repository REPISALUD of the Instituto de Salud Carlos III.

Study Population

In this review, a CC was defined as an individual responsible for generating and distributing diverse forms of material (including text, video, audio, and images) across digital platforms. Their primary aims include entertaining, informing, educating, or fostering a connection with an audience. This creative work can be performed on any web-based social media platform.

Data Collection Source

Data were obtained via direct consultation of the following bibliographic databases in the field of health sciences: MEDLINE (via PubMed), Embase, Cochrane Library, PsycINFO, Scopus, Web of Science, and Virtual Health Library.

Information Processing

To define the search terms, the Thesaurus of Descriptores de Ciencias de la Salud, developed by the Latin American and Caribbean Center on Health Sciences Information, was consulted along with its equivalence to Medical Subject Headings from the US National Library of Medicine.

From the hierarchical study of both thesauri and their indexing sheets (entry terms), the following search equations were deemed appropriate:

Equation 1: CCs on social media (Web 2.0).

"Content Creator*"[Title/Abstract] OR "Influencer*"[Title/Abstract] OR "Streamer*"[Title/Abstract] OR "Gamer*"[Title/Abstract] OR "Youtuber*"[Title/Abstract] OR "TikToker*"[Title/Abstract] OR "Instagramer*"[Title/Abstract] OR "Twitcher*"[Title/Abstract] OR "Blogger*"[Title/Abstract] OR "Vloggers*"[Title/Abstract] OR "Podcaster*"[Title/Abstract] OR "Content Strategist"[Title/Abstract]

Equation 2: pathological processes—abnormal forms and mechanisms involved in dysfunctions of tissues and organs.

"Pathologic Processes"[Mesh] OR "Pathologic Processes"[Title/Abstract] OR "Pathological Process"[Title/Abstract] OR "Disease*"[Title/Abstract] OR "Syndrome*"[Title/Abstract] OR "Symptom*"[Title/Abstract] OR "Illness"[Title/Abstract]

The final search equation was developed for MEDLINE (via PubMed) by combining the two proposed equations using Boolean logic (Equation 1 and Equation 2).

This strategy was then adapted to the characteristics of each of the other consulted databases, and the search was conducted from the earliest available date up to May 2024. In addition, a complementary search strategy was implemented to minimize the possibility of publication bias by manually searching the bibliographies of the papers selected for the review. Furthermore, experts in the study topic were contacted to determine the potential existence of gray literature (materials and research produced by organizations outside of traditional commercial or academic publications and disseminated through alternative channels). These experts were not involved in the delineation of the research scope, the refinement of the research question, the selection of studies, or the interpretation of the findings.

Final Selection of Papers

The selection of papers for review and critical analysis was informed by the criteria outlined in Textbox 1.

Textbox 1. Inclusion and exclusion criteria adopted for the review.

Inclusion criteria:

- Original paper published in a peer-reviewed journal.
- The language must be English, Spanish, or Portuguese.
- The complete paper is available for review.
- Participants aged 13 years and older (minimum legal age for engagement in the activity).

Exclusion criteria:

- Studies that do not address the objectives of the review (eg, content creator [CC] and pathological processes).
- Studies conducted exclusively with participants under 13 years [2].

The selection of relevant papers was performed by two authors of this review (SMA and JSV). To validate the inclusion of papers, the agreement rating between the authors had to be above 0.60 [18]. Any disagreements were resolved through consensus among all authors of the review.

Document Quality, Evidence Level, and Recommendation Grade

The structural validity of the papers was evaluated using the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [19], which contain a list of 22

XSL•FO
RenderX

essential aspects that must be described in each paper. For each selected paper, one point was assigned for each item present. If an item contained several subcategories, these were independently evaluated, with the same value given to each subcategory, and then an average was calculated (this average was the result for that item), ensuring that the total score for each item did not exceed one point.

To determine the level of evidence and its recommendation grade, the recommendations of the Scottish Intercollegiate Guidelines Network (SIGN) Grading Review Group [20] were used.

Bias Study

The risk of bias in exposure effect studies was assessed using the Risk of Bias in Non-randomized Studies of Exposures (ROBINS-E) tool [21]. This tool classifies biases into 7 dimensions, rating each one on four possible interpretations: low risk (little or no concern about bias); some concerns (there is some concern about bias in this domain, though it is unclear whether a risk exists); high (the study has some significant issues in this domain, leading to a high risk of bias); and very high (the study is highly problematic in this domain. The study's characteristics lead to a very high risk of bias.

Data Extraction

Data accuracy control was performed using double-entry tables that allowed for the detection of deviations and their correction through reconsulting the original data. Duplicate records were excluded using the cross-platform ZOTERO program (George Mason University), a reference management tool developed by the Center for History and New Media at George Mason University.

Variables Under Study

The papers were grouped based on the variables under study to systematize and facilitate the understanding of the results. The following data were extracted: first author, year of publication, design, population (CC activity and age), disease, platform or location, time frame, instrument, and health issue.

Data Analysis

The data related to information retrieval were presented in terms of frequency and percentage. The agreement for paper selection by the two authors was measured using Cohen κ coefficient, with an agreement considered adequate when the value was above 60% (good or very good agreement strength).

To determine the currency or obsolescence of the selected papers, the Burton-Kebler (BK) semiperiod was calculated—defined as the median age based on the temporal range analyzed—and the Price Index (PI), which represents the percentage of papers published within the past 5 years.

STROBE questionnaire scores were analyzed using the median, maximum, and minimum. The evolution of these scores, in relation to the years of publication, was obtained through regression analysis (R^2).

For statistical analyses, SPSS (version 29; IBM Corp) for Windows was used. The level of significance used was α=.05.

Ethical Considerations

All data were obtained from papers accepted for review. Therefore, in accordance with Law 14/2007 on biomedical research, ethical committee approval was not required due to the use of secondary data [22].

Results

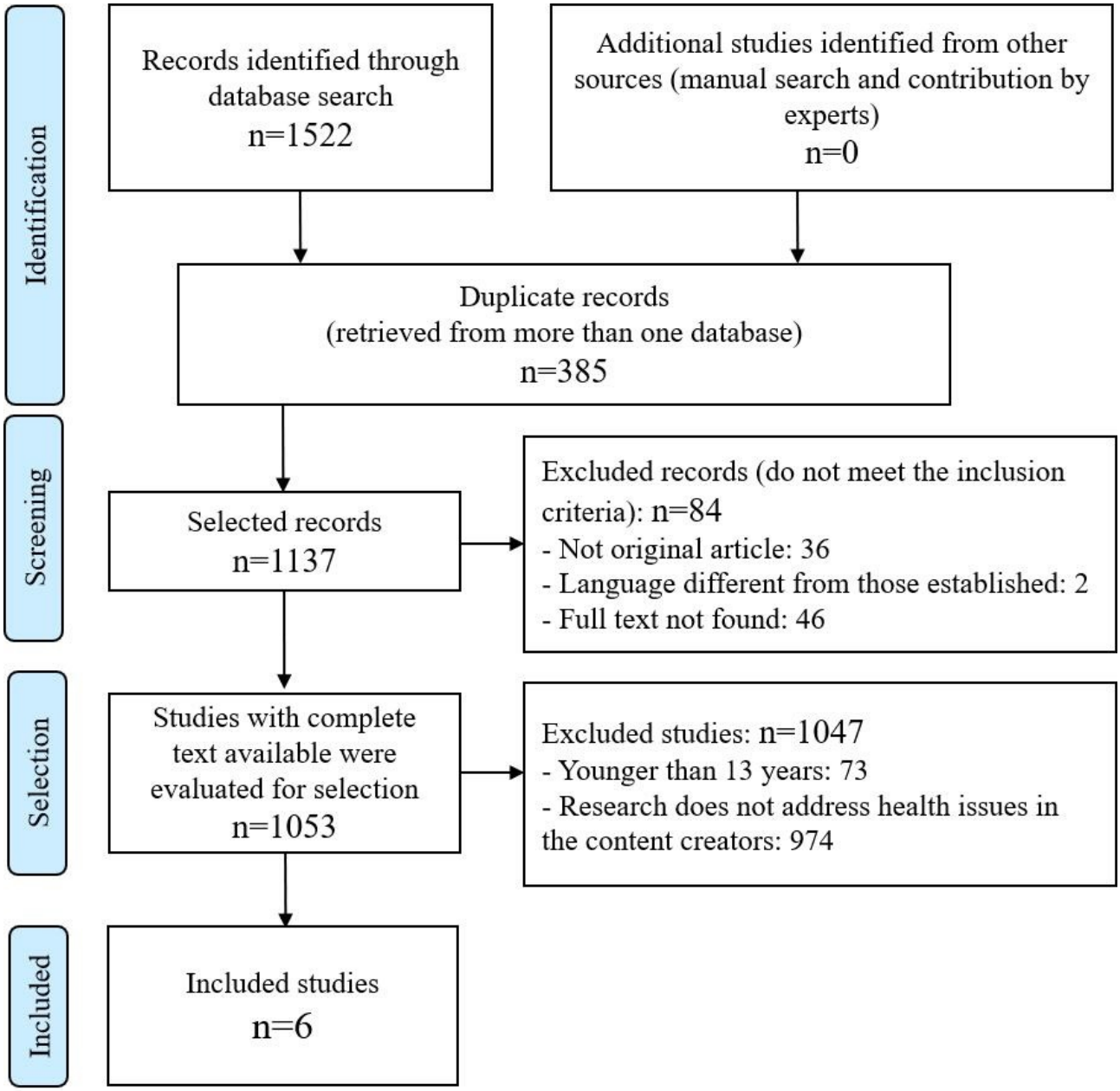
Overview

By applying the search criteria, a total of 1522 references were retrieved: 200 (13.1%) from MEDLINE (via PubMed), 61 (4.0%) from Embase, 148 (9.7%) from Cochrane Library, 95 (6.2%) from PsycINFO, 519 (34.1%) from Scopus, 344 (22.6%) from Web of Science, and 155 (10.2%) from Virtual Health Library. The consultation of bibliographies of selected papers and expert consultation did not yield any new documents. The entire paper selection process can be referenced.

After eliminating 385 duplicate records and applying the inclusion and exclusion criteria (Figure 1), 6 documents [23-28] were selected for review and critical analysis; refer to Table 1.

XSL•FO
RenderX

Figure 1. Identification and selection of studies in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.



XSL•FO
RenderX

Table . Selected studies for review on pathological processes among content creators on social networks.

Author	Design	Activity	Population (n; gender; age)	Disease	Platform or location	Time frame	Instrument	Health issue
Zsila et al [23]	Qualitative and quantitative	Influencer	• Study 1: n=1544; M/W=1336/208; 14-53 years • Studies 2 and 3: n=525; M/W=427/98; 18-56 years	Depression and anxiety	Twitch	• Study 1: spring 2021 • Study 2: 2021 - 2022 • Study 3: 2021 - 2022	• Study 1: online questionnaire on gaming disorder (DSM-5[b] criteria) • Study 2: 10-item Need for Belonging Scale. The participants' need for group dependency was evaluated. • Study 3: following the same approach as Study 2, but the focus was on aspirational identification and parasocial relationships	• Study 1: internet gaming disorder, inadequate sleep, and symptoms of depression and anxiety • Study 2: high scores were obtained, indicating the need for belonging • Study 3: investigates the type of contact with the streamer, parasocial relationship bonds, identification of desires, and the need for belonging as predictors of financial support to a favorite streamer
Niederauer and Maggi [24]	Qualitative	YouTuber	• n=6; M/W=6/0; 18-40 years	Mental health (burnout)	YouTube	• Not specified	• Semistructured interviews addressing subjective questions about the digital influencer on the YouTube platform	• Exhaustion, discomfort, pressure, distress or anxiety, and burnout

XSL•FO
RenderX

Author	Design	Activity	Population (n; gender; age)	Disease	Platform or location	Time frame	Instrument	Health issue
Gandhi et al [25]	Qualitative	Streamer	• n=25; M/W=16/9; 18-39 years	Mental health (unspecified)	Twitch	• 2019	• Semistructured interviews with an average duration of 59 minutes (SD: 25.5); $20 per interview	• Mental health issues (no specific symptoms are outlined); the study focuses on how streamers, viewers, or both address these issues during live streams
Tran et al [26]	Qualitative	Influencer	• n=9; M/W=0/9; 21-40 years	Burnout	YouTube	• Not specified	• 20-question interview distributed via email (Jacobsen Framework, theoretical model for the study of esthetics psychology)	• Burnout (feelings of guilt related to inconsistent work, delayed video posts, creative blocks, and failure to keep social media platforms updated) • Mood regulation and self-esteem • Anxiety
Yeshua-Katz and Martins [27]	Qualitative	Blogger	• n=33; M/W=0/33; 15-33 years	Eating disorder	Blog	• Not specified		• Eating disorder (anorexia nervosa) • Anxiety

XSL•FO
RenderX

Author	Design	Activity	Population (n; gender; age)	Disease	Platform or location	Time frame	Instrument	Health issue
							Semistructured interview with three sections of questions: • First section: bloggers' experiences with their eating disorder • Second section: focused on their reasons for blogging about their disorder • Third section: questions about the perceived benefits and disadvantages of pro-ana blogging	
Ressler et al [28]	Qualitative and quantitative	Blogger	• n=230; M/W=17/213; 18-75 years	Chronic pain	Blog	• May 5 to July 11, 2011	• Online open-ended survey with 34 items	• Chronic pain diseases • Therapeutic potential of blogging (the act of posting about the pathology they endure—in this case, chronic pain diseases—allows individuals to alleviate the emotional burden associated with it)

[a]M/W/O: man/woman/other (nonbinary).

XSL•FO
RenderX

[b]DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition.*

The agreement on the relevance of the selected studies between the evaluators, calculated using the κ index, was 84.9% (P<.01).

The total retrieved papers showed an obsolescence rate, according to the BK semiperiod, of 3 years, with a PI of 71.40% (7/10 papers were published in the last 4 years). Meanwhile, the papers selected for review showed a BK semiperiod of 3.5 years, with a PI of 66.7%.

When evaluating the document correction of the reviewed papers using the STROBE questionnaire, the scores ranged from a minimum of 13.0 (81.3% compliance) to a maximum of 15.5 (96.8% compliance), with a median of 14.9 (Table 2). A very low, nonsignificant direct linear trend over time was observed (R^2=0.02; P=.80).

Table . Analysis of the methodological quality of the studies according to the 22 evaluation points of the Strengthening the Reporting of Observational Studies in Epidemiology guidelines.

Paper	Score of the points in the questionnaire[a]																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	Value, n (%[b])
Zsila et al [23]	1	1	1	1	1	1	1	1	N/A	1	N/A	N/A	1	0.5	N/A	N/A	N/A	1	1	1	1	0.5	15.0 (93.8)
Nadler et al [24]	0.5	1	1	1	1	1	1	0	N/A	1	N/A	N/A	1	1	N/A	N/A	N/A	1	1	1	1	1	14.5 (90.6)
Gandhi et al [25]	1	1	1	1	1	1	1	1	N/A	1	N/A	N/A	1	1	N/A	N/A	N/A	1	1	1	1	0.5	15.5 (96.8)
Tran et al [26]	0.5	1	1	1	1	1	1	1	N/A	1	N/A	N/A	0.3	1	N/A	N/A	N/A	1	1	1	1	1	14.8 (92.5)
Yehuda-Katz et al [27]	0.5	1	1	0	1	1	1	1	N/A	1	N/A	N/A	0.5	1	N/A	N/A	N/A	1	1	1	1	0	13.0 (81.3)
Reshef et al [28]	0.5	1	1	1	1	1	1	1	N/A	1	N/A	N/A	1	1	N/A	N/A	N/A	1	1	1	1	1	15.5 (96.8)

[a]0: does not meet the item or any of its parts; 1: fully meets the item; 0 to 1: partially meets the item; and NA: does not apply.

[b]Percentage of compliance for the total items, excluding those that do not apply (N/A).

According to SIGN criteria, this review presented an evidence level of 2++ (systematic reviews with a high probability of a causal relationship) with a recommendation grade of B (a body of evidence including studies directly applicable to the target population and showing overall consistency of results).

The risk of bias in the analyzed studies and the evaluation of the domains included in the ROBINS-E tool are summarized in Table 3, where a higher number of biases were observed in older papers, particularly in Domains 5, 6, and 7. This indicates that these studies have difficulty in appropriately measuring the results.

XSL•FO
RenderX

Table . Evaluation of the reviewed papers according to the bias domains included in the Risk of Bias in Non-randomized Studies of Exposures tool.

Domain	Zsila et al [23]	Niederauer et al [24]	Gandhi et al [25]	Tran et al [26]	Yeshua-Katz et al [27]	Ressler et al [28]
Domain 1[a]	Some concerns	Low	Low	Low	Low	Low
Domain 2[b]	Low	Low	Some concerns	Some concerns	Low	Some concerns
Domain 3[c]	Low	Low	Low	Low	Some concerns	Low
Domain 4[d]	N/A[e]	N/A	N/A	N/A	N/A	N/A
Domain 5[f]	Some concerns	Very high	High	Very high	High	Some concerns
Domain 6[g]	High	High	High	Very high	High	High
Domain 7[h]	High	High	High	High	High	High

[a]Domain 1: due to confounding.

[b]Domain 2: arising from exposure measurement.

[c]Domain 3: participant selection in the study.

[d]Domain 4: due to interventions after exposure.

[e]N/A: not applicable.

[f]Domain 5: due to missing data.

[g]Domain 6: arising from outcome measurement.

[h]Domain 7: due to the selection of the reported outcome.

The study population ranged from a minimum of 6 participants [24] to a maximum of 1544 participants [23]. Three of the 6 studies focused on a single gender: Niederauer and Maggi [24] included only men, while Tran et al [26] and Yeshua-Katz and Martins [27] included only women. Gandhi et al [25] was the only study that identified gender differences. Two studies noted that the population initially presented eating disorders [27] and chronic pain [28], while the other studies did not report any specific pathology in the studied population.

The activities performed by the individuals included in the various studies were: influencers [23,26], YouTubers [24], streamers [25], and bloggers [27,28].

Although this review focused on adults, as it investigated CC as a professional activity, two studies included minors: the studies by Zsila et al [23] and Yeshua-Katz et al [27]. Only the study by Ressler et al [28] included individuals older than 65 years.

Tools Used

All tools were shared either online or via email. The study by Zsila et al [23] was mixed (qualitative-quantitative). It analyzed qualitative aspects and established quantitative scales or values through 3 sections. It examined the problematic gaming behavior pattern, characterized by an increasing priority given to gaming over other activities, to the point where gaming takes precedence over daily interests and activities (gaming disorder). It also assessed the sense of trust, intimacy, and acceptance within a group, whether family, friends, or work (Need for Belonging Scale), and considered the potential financial support based on the second point.

Niederauer and Maggi [24] conducted an exploratory qualitative study using semistructured interviews. The methodology involved content analysis of participants' responses to identify patterns and generate thematic categories.

Gandhi et al [25] used a qualitative methodology with semistructured interviews to explore how CCs address mental health topics.

Tran et al [26] conducted a qualitative study using an ad hoc interview to analyze guilt feelings related to potential inconsistent work, using the Jacobsen Framework (a theoretical model for studying aesthetics psychology) to understand the effects of intrinsic and extrinsic motivation on CC self-esteem.

Yeshua-Katz and Martins [27] also used a qualitative methodology with a focus on personal narratives and the meanings individuals assign to their experiences with illness. In addition, content analysis based on grounded theory was used to identify emerging themes.

Ressler et al [28] performed a mixed methods study. They conducted a quantitative part with closed-ended questions to collect data on participants' characteristics, using descriptive statistics. They also conducted a qualitative portion with open-ended questions to explore bloggers' personal experiences, motivations, and perceptions of blogging's impact on their emotional and social well-being, using grounded theory to identify recurring themes and patterns.

Main Results

The study by Zsila et al [23] revealed results indicating that the lack of group belonging, along with certain types of video games, led to the development of mental health symptoms, such as anxiety and depression, as well as disturbances in sleep.

Niederauer and Maggi [24] highlighted how the lack of satisfaction in content creation could lead to signs and symptoms of burnout. They observed that unsatisfactory CC on YouTube led to various health problems, including exhaustion, discomfort, pressure, distress, anxiety, and burnout.

XSL•FO
RenderX

In the study conducted by Gandhi et al [25], it was described that the type of video game influenced how mental health was addressed for both CCs and their audiences, without specifying concrete symptoms.

The results of the study by Tran et al [26] indicated signs or symptoms of burnout in the studied population. Feelings of guilt related to unconscious work, delayed video posts, artistic blocks, and failure to keep social media platforms up to date were observed. In addition, mood regulation, self-esteem, and anxiety were reflected.

Yeshua-Katz and Martins [27] highlighted the presence of bloggers with health concerns related to eating disorders, manifesting issues such as anorexia nervosa and anxiety.

Ressler et al [28] showed that content creation for bloggers dealing with chronic pain and illness helped reduce the feeling of isolation through online communication, presenting a therapeutic potential in the act of blogging.

Thus, both Yeshua-Katz and Martins [27] and Ressler et al [28] suggested that blogging could serve as an effective means to express a patient's clinical situation and even be considered a therapeutic activity.

All studies agreed on the need for further research into the potential pathological processes of CCs engaged in social media activities.

Discussion

Principal Findings

Considering the recommendations for the objectives of a systematic review [29], this review summarized relevant information related to the occupational health of CCs working on social media. In fact, there is a wide range of research on the effects of social media use on health, particularly in young people. However, very little is known about the health of social media CCs, who also tend to be quite young.

The reviewed papers presented low obsolescence, indicating that the topic is novel. This is further confirmed by the high percentage of papers included in this study that were published in the last 5 years, ensuring confidence in the findings being up to date.

The evaluation of the document accuracy of the studies included in this review, using the STROBE guidelines, did not reveal any temporal evolution. Usually, more recent papers show better results and are linked to the progressive implementation of quality questionnaires. In fact, older studies often did not follow these quality guidelines; for example, the first documents in STROBE date back to 2004, and their usage was gradual [30]. In light of the recent publication of the reviewed studies, it was anticipated that this circumstance would not arise and that the obtained results would be satisfactory.

The level of evidence and recommendation grade for this work, according to the SIGN criteria, were in line with expectations for reviewing observational studies. Despite aiming for a consistent relationship between intervention and outcome, it is important to remember the study designs from which the evidence is derived, acknowledging that some are more susceptible to biases than others and, therefore, justify decisions less strongly. It is well known that many health and occupational safety studies still do not rely on the highest possible level of evidence [31]. This could be due to the limitations of certain primary study designs, such as clinical trials, which are considered robust but may not be suitable for evaluating occupational health interventions, as they generally have long-term effects [32]. In this review, the pathological processes of CCs may not yet be well-studied causes.

The evaluation of the domains included in the ROBINS-E tool showed the main expected biases in the observational studies, but it should be noted that no studies with appropriate cause-and-effect relationships were available, as already indicated [21]. In fact, studies on occupational health, particularly those analyzing data from social media, have shown an uncontrolled process that leads to a biased focal point [33].

On the other hand, the population size was very small, except for the studies by Zsila et al [23] and Ressler et al [28], which could establish results directly derived from the intervention and, consequently, generalize their conclusions. This issue was mentioned in all studies, which unanimously agreed on the need for further research into the possible pathological processes of CCs engaged in social media activities.

Regarding the activity developed, it was observed over the years that there has been an evolution from blogging to the use of platforms that allow live streaming, such as YouTube or Twitch, as streamers can broadcast in real time and interact with their audience. Despite the differences each platform may present, there is a unique characteristic: the effort of CCs to attract audiences and establish parasocial relationships (ie, emotional connection with someone they do not know personally) [34]. Furthermore, people aged 15-24 years (the most active on social media) are precisely those who use platforms such as YouTube and Twitch the most [35]. These 2 platforms have contributed to the so-called "democratization of content generation," allowing the audience to have a proactive role in the process. The "Twitch Generation" values, above all, authenticity and interactivity, genuine, unfiltered content, and direct interactions with their favorite streamers [36].

The presence of minors as CCs in the study by Yeshua-Katz and Martins [27] should not come as a surprise, as most platforms allow content creation starting at the age of 13. However, we must not forget that these minors are performing adult-level work, with all the associated risks, particularly from a psychological perspective. Therefore, it is crucial to raise awareness among parents and guardians while also keeping in mind the responsibility of digital platforms in safeguarding minors' rights.

In this regard, the study by Gómez et al [37] examined how "kidfluencers" (minor CCs) created content and integrated "branded content" (content linked to a brand that connects the brand with the consumer) into their creations. Martínez Pastor et al [38] demonstrated that, despite the existence of advertising regulations in all the countries studied, compliance was minimal when it came to identifying advertising content as such. In this vein, Tur-Viñes et al [39] assumed that content creation by

XSL•FO
RenderX

minor CCs is carried out with the support of parents, and in some cases, parents even open their own channels to support the transmedia strategy initiated by their children.

While the study by Ressler et al [28] included a population up to 75 years old, no specific results were found for the 65 - 75 age range. However, other studies showed that the content production of senior CCs was characterized by a pronounced digital marketing strategy, highlighting professional routines such as regularity in posting frequency, post quality, and creativity, and a close relationship with followers [40].

Traditionally associated with a more passive role as content consumers, some older adults have been developing their digital skills and are now active CCs on social media. In some cases, they have even created a vast audience and a wide range of engaged, interested, and committed followers [41]. However, there is little evidence concerning these older digital CCs, and even less regarding the pathologies associated with this activity, as these individuals are often considered excluded from the labor market.

Methodological Considerations

Conducting interviews online or via email are methods commonly used for data collection, as they are extremely cost-effective and can be sent to a large audience simultaneously. However, the reliability of the data can sometimes be questioned, as the responses received may not always be authentic. Therefore, a method should be established to support the results that might be obtained. Nevertheless, there are studies that have considered online or email interviews as an effective and convenient method [42].

In studies with quantitative methodology [23,28], a validated tool should have been used, or the homogeneity, reliability, internal consistency, and validity should have been reported [43]. In addition, in studies with qualitative techniques, the trustworthiness—credibility, transferability, dependability, and confirmability—should have been discussed [44].

The need for validation of results is fundamental in web-based studies. Collecting high-quality datasets from social media users is problematic, not only due to the biases associated with data collection methods but also with respect to consent management and the selection of appropriate analytical techniques [45].

Health Impacts and Preventive Measures

As observed in this review, the impact on mental health was one of the primary issues resulting from the activities of CCs. This is due to the constant pressure, whether imposed externally or self-imposed, that these workers face in their daily activities to meet targets.

At the same time, they must contend with the need to please their viewers, often unable to fully express their emotions or feelings, either to their audience or, even less so, to their sponsors. This situation leads to the manifestation of symptoms that affect the mental health of CCs, such as disturbances in sleep, anxiety, emotional lability, excessive worry, irritability, loss of interest, obsessions, and fatigue. This set of symptoms creates the ideal breeding ground for the development of psychiatric illnesses such as generalized anxiety disorder, dysthymia, or depressive episodes. In the most extreme cases, these can even lead to the suicide of a CC [46,47].

In addition to the extreme burnout associated with these conditions, CCs may also experience physical exhaustion and a sense of depersonalization, resulting in burnout syndrome, as indicated in the studies by Niederauer and Maggi [24] and Tran et al [26].

However, as demonstrated in this review, it would be a mistake to think that the health consequences for CCs are limited solely to mental health disorders. The prolonged use of data display screens, visual fatigue, persistent forced postures, repetitive upper limb movements, and a sedentary lifestyle can lead to musculoskeletal disorders and systemic issues. These risks associated with the occupational activity of CCs are critical to understanding the various impacts that this professional activity can have on the health of these workers.

It is essential to address these consequences at their source through preventive actions to avert health problems. To prevent these health issues, CCs should be educated and informed about the potential consequences of their work activity, emphasizing the importance of a balance between work and rest. Promoting periodic digital disconnection and avoiding marathon workdays that hinder recovery or rest is key. In this regard, it is important to identify whether CCs, or certain profiles within this group, are exposed to shift work in their activities and, if so, to apply the corresponding preventive measures. To ensure the health of workers, periodic evaluations of their well-being are essential and should be conducted regularly to identify and prevent potential health issues.

Regarding the consequences of sedentary behavior in CCs' activities, workers must be made aware of the importance of taking regular breaks during their workday, alternating between sitting and standing, to reduce the negative effects of prolonged sedentary behavior.

To mitigate the risks associated with the use of data display screens, persistent forced postures, and repetitive upper limb movements, it is crucial for CCs to ensure adequate ergonomic conditions in their work environment that promote a healthy and appropriate setting for carrying out their activities.

Critical Reflection by Authors

It is essential to consider the importance of conducting research that analyzes the impact on the health of CCs from multiple perspectives, including age, the type of Web 2.0 platform used for content creation, the target audience, the social impact generated, and the consequences for health in each of these areas. This multidimensional approach would allow for a comprehensive and detailed understanding of the occupational risks associated with this activity for the health of those engaged in it. The foundation of occupational risk prevention lies in addressing the problem at its origin. Furthermore, appropriate corrective and preventive measures must be established to minimize the risk. In this regard, there is a noticeable gap in proposals related to such measures. As long as the active population continues their work without regular medical supervision and appropriate follow-up, health problems will

XSL•FO
RenderX

persist or develop, which, if left uncorrected, could lead to severe and irreversible damage to the CCs.

Limitations

A primary limitation of this review could be the low number of studies selected, as it concerns an emerging area of knowledge. It has been stated that systematic reviews should be based on studies with designs and selections that ensure greater scientific rigor. However, this analysis included all relevant studies retrieved on the topic under investigation.

The significant difference between the papers retrieved and the final selection is primarily due to the bibliographic databases Scopus and Web of Science. Although some studies were retrieved from other databases, the majority were irrelevant, which could be attributed to their lack of proper indexing (the search was conducted in text format, consulting titles, abstracts, and keywords) and the inability to limit the search by paper type. This high "documentary noise" has been previously observed in other reviews.

According to the United States Agency for Healthcare Research and Quality, the epidemiological designs of the studies selected in this review do not guarantee the full validity and reliability of the observations obtained [48]. However, the available evidence is likely the best given the difficulties of studying this area of research and based on the observations obtained in the various interventions.

Although the real limitations are due to the characteristics of each study, important lessons can be drawn from them to formulate appropriate actions for the development, implementation, and evaluation of future Web 2.0 applications. As Drucker et al [49] rightly stated, understanding these sources of bias is crucial for both authors and consumers of scientific literature as they conduct and read systematic reviews and incorporate their findings into clinical practice and policy formulation.

Conclusions

The studies revealed that the most observed area was mental health. Although the impact on mental health has been extensively documented, it is evident that there is a need to address other aspects of this occupational activity, such as physical health and its social impact, in future studies.

Content creation, therefore, is an emerging activity that is not without health issues, especially in the realm of mental health. It is crucial to continue conducting research that addresses these problems at their origin and determines preventive measures to mitigate their effects. Thus, more detailed studies from an occupational health perspective are necessary.

Acknowledgments

No artificial intelligence tools were used in the conduct of this review.

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Authors' Contributions

SMA and JSV were responsible for conceptualization, methodology, formal analysis, supervision, project administration, writing (original draft), and writing (review and editing). ERP was responsible for methodology, formal analysis, data curation, and writing (review and editing).

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR checklist.
[PDF File, 660 KB - publichealth_v11i1e76708_app1.pdf]

References

1. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and Amending Directive 2000/31/EC (Digital Services Act). EUR-Lex. 2022. URL: https://eur-lex.europa.EU/legal-content/EN/TXT/HTML/?uri=CELEX:32022R2065 [accessed 2025-08-25]
2. Real decreto 444/2024, de 30 de abril, por el que se regulan los requisitos a efectos de ser considerado usuario de especial relevancia de los servicios de intercambio de vídeos a través de plataforma, en desarrollo del artículo 94 de la ley 13/2022, de 7 de julio, general de comunicación audiovisual [Website in Spanish]. : Boletín Oficial del Estado (BOE); 2024 URL: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2024-8716 [accessed 2025-08-25]
3. Pătra L, Todolí A. Ser Influencer Hoy: Posibilidades y Obstáculos de Una Nueva Fuente de Empleo [Book in Spanish]: Universitat de València, Cátedra de Economía Colaborativa y Transformación Digital; 2022. URL: https://bit.ly/3RyjLPI [accessed 2024-06-18]

XSL•FO
RenderX

4. Alvarez Monzoncillo JM, editor. The Dynamics of Influencer Marketing: A Multidisciplinary Approach, 1st edition: Routledge; 2023.
5. Fernandes C, Pereira A. Exposure to psychosocial risk factors in the context of work: a systematic review. Rev Saude Publica 2016;50(24):24. [doi: 10.1590/S1518-8787.2016050006129] [Medline: 27253900]
6. Filip D, Van der Hallen R, Smeets G, Franken I, Prinzie P. Pathological personality domains and social media use in emerging adults: mediation by social media self-control failure. Emerg Adulthood 2025 Apr;13(2):291-307. [doi: 10.1177/21676968241264323]
7. Facultad de Ciencias Empresariales y del Trabajo [Website in Spanish]. 2018. URL: https://uvadoc.uva.es/bitstream/handle/10324/35992/TFG-O-1603.pdf?sequence=1&isAllowed=y [accessed 2024-06-18]
8. Peng P, Liao Y. Six addiction components of problematic social media use in relation to depression, anxiety, and stress symptoms: a latent profile analysis and network analysis. BMC Psychiatry 2023 May 8;23(1):321. [doi: 10.1186/s12888-023-04837-2] [Medline: 37158854]
9. Kułaga W, Szkoła Doktorska Nauk Społecznych Uniwersytetu Jagiellońskiego. The phenomenon of pathostreaming and pathological content in the face of media policy. Real Polit 2023;23(1):49-67. [doi: 10.15804/rop2023103]
10. A review on the future of work: online labour exchanges or crowdsourcing. European Agency for Safety and Health at Work (EU-OSHA). 2016. URL: https://oshwiki.osha.europa.eu/en/themes/review-future-work-online-labour-exchanges-or-crowdsourcing [accessed 2024-06-18]
11. Foresight on new and emerging occupational safety and health risks associated with digitalisation by 2025. European Agency for Safety and Health at Work (EU-OSHA). 2018. URL: https://osha.europa.eu/en/publications/foresight-new-and-emerging-occupational-safety-and-health-risks-associated [accessed 2024-06-18]
12. Ketelhut S, Martin-Niedecken AL, Zimmermann P, Nigg CR. Physical activity and health promotion in esports and gaming-discussing unique opportunities for an unprecedented cultural phenomenon. Front Sports Act Living 2021;3:693700. [doi: 10.3389/fspor.2021.693700] [Medline: 34604743]
13. Sherlock M, Wagstaff DL. Exploring the relationship between frequency of Instagram use, exposure to idealized images, and psychological well-being in women. Psychol Pop Media Cult 2019;8(4):482-490. [doi: 10.1037/ppm0000182]
14. Pluta A, Mazurek J, Wojciechowski J, Wolak T, Soral W, Bilewicz M. Exposure to hate speech deteriorates neurocognitive mechanisms of the ability to understand others' pain. Sci Rep 2023 Mar 13;13(1):4127. [doi: 10.1038/s41598-023-31146-1] [Medline: 36914701]
15. Fung ICH, Blankenship EB, Ahweyevu JO, et al. Public health implications of image-based social media: a systematic review of Instagram, Pinterest, Tumblr, and Flickr. Perm J 2020;24(18):18.307. [doi: 10.7812/TPP/18.307] [Medline: 31852039]
16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
17. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
18. Wanden-Berghe C, Sanz-Valero J. Systematic reviews in nutrition: standardized methodology. Br J Nutr 2012 Jun;107 Suppl 2:S3-S7. [doi: 10.1017/S0007114512001432] [Medline: 22591902]
19. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology STROBE statement: guidelines for reporting observational studies. Gac Sanit 2008;22(2):144-150. [doi: 10.1157/13119325] [Medline: 18420014]
20. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. BMJ 2001 Aug 11;323(7308):334-336. [doi: 10.1136/bmj.323.7308.334] [Medline: 11498496]
21. Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). Environ Int 2024 Apr;186:108602. [doi: 10.1016/j.envint.2024.108602] [Medline: 38555664]
22. Ley 14/2007, de 3 de julio, de investigación biomédica [Website in Spanish]. : Boletín Oficial del Estado; 2007 URL: http://boe.es/eli/es/l/2007/07/03/14 [accessed 2025-08-25]
23. Zsila Á, Shabahang R, Aruguete MS, Bőthe B, Gregor-Tóth P, Orosz G. Exploring the association between Twitch use and well-being. Psychol Pop Media 2023;13. [doi: 10.1037/ppm0000486]
24. Niederauer RC, Maggi A. Os desafios do algoritmo da plataforma para o trabalho do influenciador digital conhecido como YouTuber [Article in Portuguese]. Rev Eletron Comun Inf Inov Saúde 2022;16(4):820-836. [doi: 10.29397/reciis.v16i4.3379]
25. Gandhi R, Cook CL, LaMastra N, Uttarapong J, Wohn DY. An exploration of mental health discussions in live streaming gaming communities. Front Psychol 2021;12:575653. [doi: 10.3389/fpsyg.2021.575653] [Medline: 33796040]
26. Tran A, Rosales R, Copes L. Paint a better mood? Effects of makeup use on YouTube beauty influencers' self-esteem. SAGE Open 2020 Apr;10(2). [doi: 10.1177/2158244020933591]
27. Yeshua-Katz D, Martins N. Communicating stigma: the pro-ana paradox. Health Commun 2013;28(5):499-508. [doi: 10.1080/10410236.2012.699889] [Medline: 22873763]
28. Ressler PK, Bradshaw YS, Gualtieri L, Chui KKH. Communicating the experience of chronic pain and illness through blogging. J Med Internet Res 2012 Oct 23;14(5):e143. [doi: 10.2196/jmir.2002] [Medline: 23092747]

XSL•FO
RenderX

29. Hagger MS. What makes a 'good' review article? Some reflections and recommendations. Health Psychol Rev 2012 Sep 1;6(2):141-146. [doi: 10.1080/17437199.2012.705556]
30. Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. BMJ 2004 Jan 3;328(7430):39-41. [doi: 10.1136/bmj.328.7430.39] [Medline: 14703546]
31. Teufer B, Ebenberger A, Affengruber L, et al. Evidence-based occupational health and safety interventions: a comprehensive overview of reviews. BMJ Open 2019 Dec 11;9(12):e032528. [doi: 10.1136/bmjopen-2019-032528] [Medline: 31831544]
32. Climent-Ballester S, García-Salom P, Sanz-Valero J. Computer programs used in the field of hospital pharmacy for the management of dangerous drugs: systematic review of literature. Front Public Health 2023;11:1233264. [doi: 10.3389/fpubh.2023.1233264] [Medline: 37711235]
33. Palomo-Llinares R, Sánchez-Tormo J, Wanden-Berghe C, Sanz-Valero J. Occupational health applied infodemiological studies of nutritional diseases and disorders: scoping review with meta-analysis. Nutrients 2023 Aug 14;15(16):3575. [doi: 10.3390/nu15163575] [Medline: 37630765]
34. Lozano-Blasco R, Mira-Aladrén M, Gil-Lamata M. Social media influence on young people and children: analysis on Instagram, Twitter and YouTube. Comunicar 2023 Jan 1;31(74):125-137. [doi: 10.3916/C74-2023-10]
35. Buitrago Á, Torres Ortiz L. Influencers de ciencia en Twitch. Divulgación científica a través de vídeo-streaming en tiempos de COVID-19. Rev Teknokultura 2022;19(2):165-176. [doi: 10.5209/tekn.77941]
36. Gregoris F. Twitch: la plataforma que lo cambió todo [Article in Spanish]. Forbes. 2023. URL: https://forbes.es/empresas/332558/twitch-la-plataforma-que-lo-cambio-todo/ [accessed 2025-01-16]
37. Gómez EF, Vázquez JF, Fernández BF. Children as content creators on YouTube and Instagram: analysis of the formats used, parental presence and brand presence. Rev ICONO14 2022;20(1). [doi: 10.7195/ri14.v20i1.1762]
38. Martínez Pastor E, Cetina Presuel R, Serrano Maíllo I. Minors as content creators: a study of the effectiveness of advertising regulation in the United States, the United Kingdom and Spain. Medijske Stud 2023;13(26):48-69. [doi: 10.20901/ms.13.26.3]
39. Tur-Viñes V, Núñez-Gómez P, Martínez-Pastor E. YouTube, menores y cultura colaborativa. Revisión bibliográfica de la investigación académica. Hist Comun Soc ;24(1):331-351. [doi: 10.5209/hics.64498]
40. Martín Martín FM, Ballesteros Aguayo L, Pedrero García E, Cruz Díaz R. Producción y caracterización de los influencers sénior españoles en Instagram. Fonseca J Commun (26):103-120. [doi: 10.14201/fjc.31237]
41. Antunes AC, Miranda S, Gama A. The (not so) secret life of senior social media influencers. ICIEMC Proc 2022;3:30-39. [doi: 10.34624/ICIEMC.V0I3.29635]
42. Fritz RL, Vandermause R. Data collection via in-depth email interviewing: lessons from the field. Qual Health Res 2018 Aug;28(10):1640-1649. [doi: 10.1177/1049732316689067] [Medline: 29298576]
43. Bernaldo-de-Quirós M, Labrador-Méndez M, Sánchez-Iglesias I, Labrador FJ. Measurement instruments of online gaming disorder in adolescents and young people according to DSM-5 criteria: a systematic review. Adicciones 2020 Nov 17;32(4):291-302. [doi: 10.20882/adicciones.1277] [Medline: 31342078]
44. Rigour in qualitative research. Aten Primaria 1999;24(5):295-300 [FREE Full text]
45. Wongkoblap A, Vadillo MA, Curcin V. Researching mental health disorders in the era of social media: systematic review. J Med Internet Res 2017 Jun 29;19(6):e228. [doi: 10.2196/jmir.7215] [Medline: 28663166]
46. Pirkis J, Machlin A. Differing perspectives on what is important in media reporting of suicide. Br J Psychiatry 2013 Sep;203(3):168-169. [doi: 10.1192/bjp.bp.112.124396]
47. Niederkrotenthaler T, Braun M, Pirkis J, et al. Association between suicide reporting in the media and suicide: systematic review and meta-analysis. BMJ 2020 Mar 18;368:m575. [doi: 10.1136/bmj.m575] [Medline: 32188637]
48. Methods guide for effectiveness and comparative effectiveness reviews. : Agency for Healthcare Research and Quality (AHRQ); 2014 URL: https://effectivehealthcare.ahrq.gov/sites/default/files/pdf/cer-methods-guide_overview.pdf [accessed 2025-01-16]
49. Drucker AM, Fleming P, Chan AW. Research techniques made simple: assessing risk of bias in systematic reviews. J Invest Dermatol 2016 Nov;136(11):e109-e114. [doi: 10.1016/j.jid.2016.08.021] [Medline: 27772550]

Abbreviations

BK: Burton-Kebler
CC: content creator
PI: Price Index
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
ROBINS-E: Risk of Bias in Non-randomized Studies of Exposures
SIGN: Scottish Intercollegiate Guidelines Network
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Mavragani; submitted 29.04.25; peer-reviewed by J Martínez, LA Sabucedo; revised version received 16.07.25; accepted 17.07.25; published 05.09.25.

Please cite as:
Martinez-Aguirre S, Sanz-Valero J, Ronda-Pérez E
Pathological Processes Among Content Creators on Social Media: Scoping Review
JMIR Public Health Surveill 2025;11:e76708
URL: https://publichealth.jmir.org/2025/1/e76708
doi:10.2196/76708

© Sergio Martinez-Aguirre, Javier Sanz-Valero, Elena Ronda-Pérez. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 5.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Preparing for Potential Health and Safety Risks at the Olympic Games: Scoping Review

Shaotong Ren[1], MRes; Tiantian Li[1], MRes; Yongzhong Zhang[1], PhD; Song Bai[1], PhD; Zichen Zhou[2], PhD; Shengxin Li[3], PhD

[1]School of Disaster and Emergency Medicine, Tianjin University, 92 Weijin Road Nankai District, Tianjin, China

[2]College of Public Health, Zhengzhou University, Zhengzhou, China

[3]School of Management, Beijing Sport University, Beijing, China

Corresponding Author:
Yongzhong Zhang, PhD
School of Disaster and Emergency Medicine, Tianjin University, 92 Weijin Road Nankai District, Tianjin, China

Abstract

Background: The Olympic Games are an example of a mass gathering that involves a complex and large crowd composition, with a large number of illnesses and injuries occurring at previous Olympic Games, and the Olympic Games also becoming a target for terrorist attacks.

Objective: With the help of mass-gathering medicine as a guide, this study aims to critically summarize and analyze the state of illness, injury, and terrorism during the Olympic Games in order to reduce the incidence of illnesses and injuries in crowds and to offer lessons for the organization of major international sporting events such as the Olympics.

Methods: The procedure for this scoping review followed the 5-step methodological framework of Arksey and O'Malley. We searched electronic databases such as PubMed, Web of Science, and Scopus. We extracted, summarized, and categorized general information on each study, game characteristics, illness and injury profiles, terrorism characteristics, preventive measures, and surveillance paradigms.

Results: We conducted a database search and retrieved a total of 9587 studies on 2 occasions. After removing duplicates and screening, we included 120 studies. Only 12 studies on the Summer, Winter, and Paralympic Games published before 2000, and 108 studies from 2000 onward, comprise the 120 studies, marking an unprecedented number of studies in this field of research, particularly in recent times. Of the 120 studies, 80 were illness-related, 81 were injury-related, and 2 were terrorism-related. Nine studies explicitly assessed body parts, including shoulders, feet, and dentistry; 26 studies specifically investigated certain illnesses and injuries, such as COVID-19 disease, heat-related illnesses, and concussions. Of the 120 research studies, 18 specifically analyzed sports such as gymnastics and weight lifting, with 11 studies focusing especially on COVID-19 disease. The most studied games were the Tokyo 2021 Olympic or Paralympic Games, the London 2012 Olympic or Paralympic Games, and the Rio 2016 Olympic or Paralympic Games. The system of injury and illness surveillance in the Olympic Games goes through 3 stages of development: the first trial of information technology, the construction of networks, and the enhancement of intelligence.

Conclusions: A critical summary of studies of illness, injury, and terrorist attacks at previous Olympic Games is important for injury and terrorism prevention at major sporting events such as the Olympic Games. Surveillance methods require improvements in surveillance technology, data sharing, and privacy protection.

(JMIR Public Health Surveill 2025;11:e66829) doi:10.2196/66829

KEYWORDS

Olympic Games; illness; injury; terrorism; surveillance; PRISMA

Introduction

According to the World Health Organization definition, a mass gathering is "any organized or spontaneous event that attracts a sufficient number of people to strain the planning and response resources of the community, city, or country hosting the event." [1]. Mass gatherings include World Youth Day, the Olympic Games, and political rallies [2]. The Olympic Games are the world's largest and most influential sporting event, with a particularly large and complex population that involves many stakeholders. Since the XXVI Olympic Games in Atlanta in 1996, hundreds of thousands of athletes, volunteers, spectators, and other stakeholders have consistently participated in each Olympic Game. Such large-scale crowd gatherings can pose significant public health concerns for the well-being and safety of participants, host country inhabitants, and even the worldwide human population [1].

XSL•FO
RenderX

Previous studies of major health and safety issues at the Olympic Games have focused sporadically on illness, injury, and terrorist attacks [3,4]. The studies focused on groups of tens of thousands of people, such as athletes and Olympic personnel, and found that the most common illnesses were respiratory, gastrointestinal, and heat-related illnesses [5-7]. Meanwhile, numerous studies have explored the sites, types, mechanisms, and incidence of injuries in athletes from previous Olympic Games, as well as the competitions with the highest risk. In recent years, traditional high-risk and emerging uncertain infectious illnesses such as the global epidemic and variation of COVID-19 disease, the active global cholera epidemic [8], and the unknown future of mankind's Illness X [9] have posed a significant threat to major international sporting events such as the Olympic Games. Furthermore, terrorism concerns have historically drawn attention to previous Olympic Games as a security issue. Simultaneously, bioterrorism has the characteristics of easy transport, low cost, hidden dissemination, and wide influence [10], and it takes advantage of the global influence of the Olympic Games, which makes it very easy to become a potential high-risk threat.

Protecting the health of athletes is an important task for the International Olympic Committee (IOC) [11]. The IOC and host nations, among other stakeholders, implement several strategies to mitigate and manage infections, injuries, and acts of terrorism. Regarding illness and injury prevention, host countries implement immunization programs, enhance publicity and education, prioritize personal hygiene, devote attention to medical personnel training, and modify the competition rules [12-14]. Furthermore, surveillance systems such as the Web-based Injury and Illness Surveillance System and Event-based Surveillance are adopted for monitoring [15,16]. Regarding counterterrorism responses, actions encompass organizing drills, providing funding, and strengthening security. Examples include the deployment of a disaster medical assistance team and a highly sophisticated chemical-biological incident response unit during the 1997 Atlanta Olympic Games. [17,18]. Meanwhile, mass-gathering medicine is a specialized discipline [19] that focuses on the provision of medical care and health protection during large crowd gatherings, providing some guidelines for the prevention of illnesses, injuries, and terrorism arising from the Olympic Games.

Extensive studies have demonstrated the prevalence of illnesses and injuries during the Olympic Games, along with multiple instances of terrorism, posing a significant risk to the physical and emotional well-being and security of athletes and individuals participating in the Games. However, systematic studies on the prevention and control of illness, injury, and terrorism in the Olympic Games are rare, and analytical studies of the evidence on the factors contributing to these health and safety impacts are still very sporadic. To fill this gap, the objective of this scoping review is to provide an overview and critical summary of the available evidence on risk factors, preventive measures, and surveillance approaches to illness, injury, and terrorism at the Olympic Games, with a view to providing counterstrategies to help major national and international sporting events in the future reduce health and safety risks.

Methods

Design and Search Strategy

The procedure for the scoping review followed Arksey and O'Malley's 5-step methodological framework, including (1) identification of the study problem, (2) identification of relevant studies, (3) study selection, (4) extraction and graphing of data, and (5) collation, summarization, and reporting of results [20]. Results were presented in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) recommendations [21].

On April 1, 2024, we conducted a systematic search of electronic databases (PubMed, Web of Science, and Scopus). The search strategy was divided into two parts: (1) the sports event, such as "Olympic," "Paralympic," or "Winter Olympic Games"; and (2) terms that covered topics such as "injury," "illness," "infectious diseases," or "terrorism" (see Multimedia Appendix 1 for the detailed search strategy [22-87]).

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria applied in this study are detailed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Studies explicitly addressing the Olympic Games context.
- Study populations involving Olympic athletes or stakeholders (eg, spectators, volunteers, and officials).
- Focus on health risks (eg, illness and injury) or safety risks (eg, terrorism).
- All publication types with accessible full text.
- Papers written in English.

Exclusion criteria

- Studies not related to the Olympic Games.
- Case reports or studies on individual athletes without group-level analysis.
- Studies unrelated to illness, injury, and terrorism, or imaging detection studies.
- Publications with unavailable full text.
- Paper not written in English.

Data Extraction

General information was extracted (ie, year of publication and type of study). Data related to the characteristics of the Summer Paralympic Games, the Winter Paralympic Games (ie, date, location, type of study group, illnessnesses, injury, or terrorism studied, etc), was extracted, along with the key findings of each study, including the occurrence of illnesses, injuries, terrorist attacks, and so forth, as well as preventive measures (Multimedia Appendix 1).

Classification of Illnesses

The classification of illnesses in this study will be the IOC consensus statement: methods for recording and reporting of epidemiological data on injury and illness in Sport 2020 (including STROBE [STrengthening the Reporting of OBservational studies in Epidemiology] Extension for Sport Injury and Illness Surveillance) [88] were combined with the *International Classification of Diseases* [89], which was processed and fused to form an innovative classification of illnesses applicable to this review and eventually divided into 18 organ or system classifications (Multimedia Appendix 1).

Results

General Characteristics of Included Studies

We searched the database, and the first search included 8587 studies, which were supplemented for the Olympic infectious illness studies, and 1000 were searched again, and 120 studies were contained after removing duplicates and screening (Figure 1).

Only 12 of the 120 studies on the Summer, Winter, and Paralympic Games were published before 2000, while 108 appeared from 2000 onward, reflecting an unprecedented surge in this field of research, particularly in recent years. Eighty-one of the 120 studies focused on illness, 81 on injury, and 2 on terrorism. Of the 120 studies, 18 analyzed events such as weight lifting and gymnastics; 9 analyzed bodily parts including shoulders, feet, and dentistry; and 26 analyzed particular illnesses or injuries such as COVID-19 disease, heat-related illnesses, and concussions, including 11 on COVID-19 disease. Out of all the studies conducted, 79 focused on the Summer Games, 48 on the Winter Games, and 50 on the Paralympic Games. Tokyo 2021, London 2012, and Rio 2016 were the top 3 most studied games (Figure 2). Across all illness studies, the most common organ or system illness for each study based on illness organ or system classification was respiratory system illness (n=35), the second most common organ or system illnesses for each study were gastrointestinal system illness (n=17) and skin system illness (n=8), respectively, and the third most common organ or system illnesses for each study were gastrointestinal system illness (n=10) and skin system illness (n=7). Additionally, the 120 publications included 3 reviews on injuries in Paralympic sports, 1 review on injuries in Winter Paralympic sports, 1 review on injuries in Winter Olympic Games, and 1 review on terrorist attacks.

XSL•FO
RenderX

Figure 1. Screening process for studies eligible for inclusion. The first search yielded 8587 studies, and the supplementary search yielded 1000 studies.

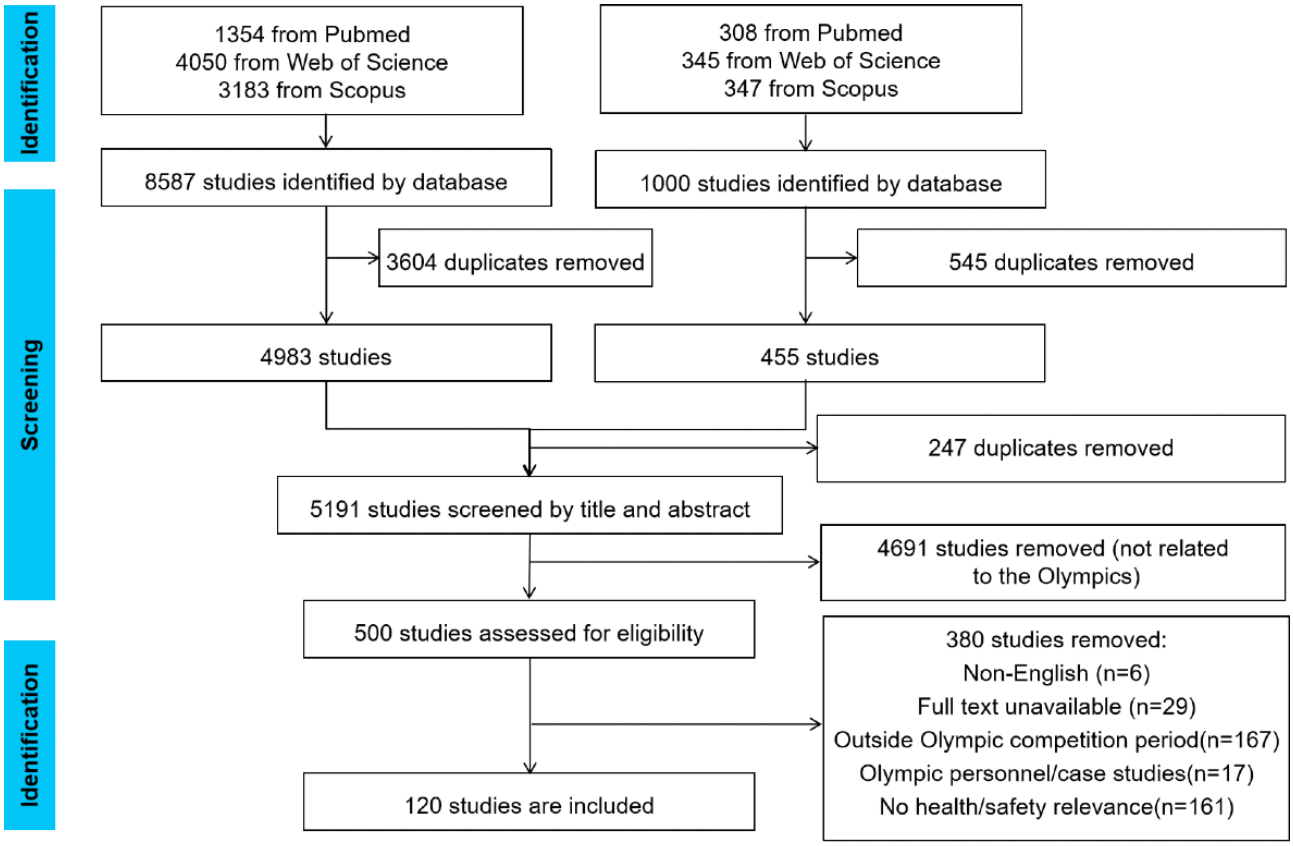
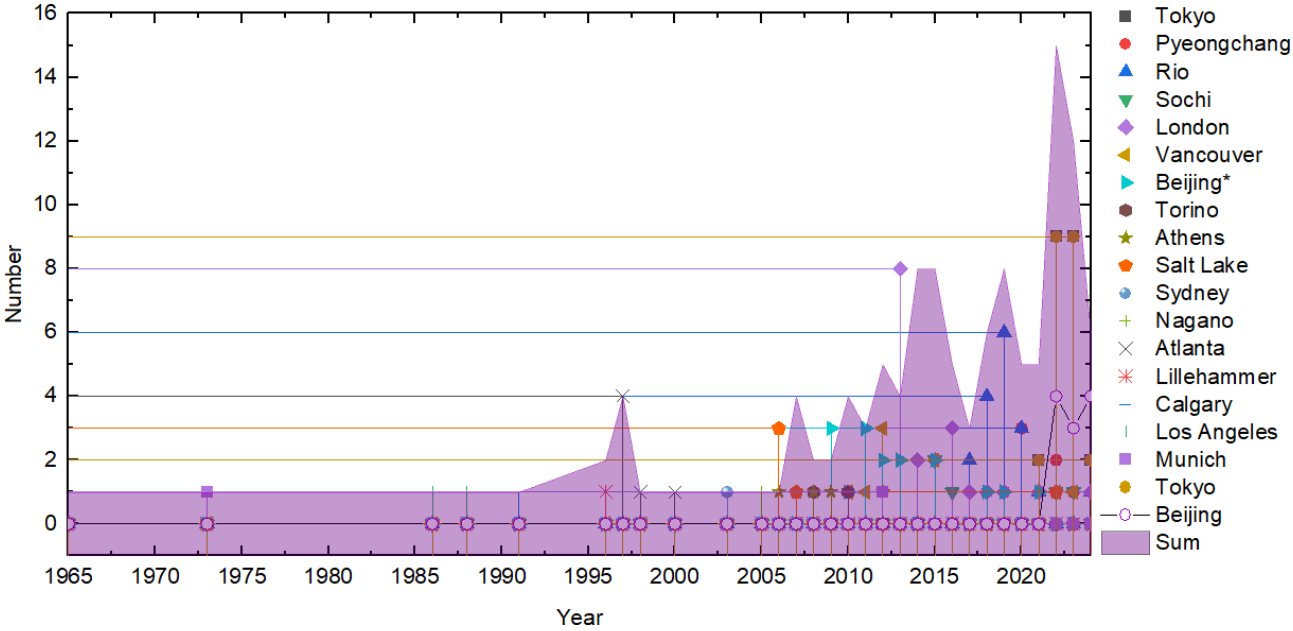


Figure 2. Trends in the number of studies per Olympics and the number of publications per year in the literature (shaded). Beijing refers to the 2022 Beijing Winter Olympics, and Beijing* refers to the 2008 Beijing Olympics.



Health and Safety Risks and Influencing Factors

Risk and Influencing Factors of Illnesses

Respiratory System Illness

Respiratory illnesses have been the most common illnesses at all Olympic Games, and the IOC Medical and Scientific Committee has noted that acute respiratory illnesses are often the most common reason for athletes to seek medical assistance at major sporting events [90]. The number of respiratory infections during the 1996 Summer Olympics in Atlanta [91] and the overall number of respiratory illnesses at the 2008 Beijing Olympics [92] reached a peak of 922 and 3944 cases, respectively. The system most affected by illness during the London 2012 Paralympic Games was the respiratory system [15]. COVID-19 disease has also caused some spread in events such as the Olympics. For example, the Tokyo Olympic Committee recorded 430 cases of infections related to the

XSL•FO
RenderX

Olympic Games [93]. The total number of positive cases monitored under closed-loop management during the Beijing Winter Olympics and Paralympics was 212 [94].

The primary determinants of respiratory illnesses are infection, allergies, environment, and various other related factors. Viruses and bacteria, such as measles, chicken pox, and influenza, primarily cause infections and widely affect the respiratory system. For example, respiratory illnesses were mostly caused by infections in the London 2012 Paralympics [95], Sochi 2014 Winter Olympics [96], Rio 2016 Olympic Games [97], and Pyongyang 2018 Winter Olympics [98]. The main risk factors for allergy and the environment are the mechanical and dehydrating pressures generated in the airways and the levels of air pollutants, irritants, and allergens inhaled by athletes during hyperventilated exercise conditions [99]. Symptoms of pollen allergy, such as worsening rhinoconjunctivitis and asthma, can be destructive to athletes expecting their best performance [100]. Prolonged exposure to mixed exhaust aerosols during exercise may result in reduced lung function and may promote vascular dysfunction [101]. During the Winter Olympics, physical activity in cold and dry air conditions causes inflammation and damage to the airways, possibly leading to a higher incidence of hyperresponsiveness to airways and asthma [102]. The long period of preparation and hard work during the 2008 Beijing Olympic and Paralympic Games led to the prevalence of other factors, such as fatigue, in staff respiratory illnesses [92].

Gastrointestinal System Illness

Illnesses of the gastrointestinal system are widely affecting athletes and those associated with the Olympic Games. Some studies have found that competitions, such as running, are the most demanding type of exercise for the gastrointestinal tract. More than 80% of athletes encounter gastrointestinal issues while exercising [103]. There were 120 cases of illnesses in the gastrointestinal system during the 2012 Summer Olympics in London, which accounted for 16% of all illnesses [104]. There were 131 cases of gastrointestinal system illnesses at the Rio 2016 Olympics [7]. At the 2018 PyeongChang Winter Olympics, 168 staff members were diagnosed with norovirus at the team's accommodation near PyeongChang, and 4 athletes were also detected with the virus [105].

The primary determinants of gastrointestinal system disorders encompass infection, food, environment, and other related factors. In terms of infection, about 90% of viral gastroenteritis outbreaks are caused by norovirus [106]. The majority of virus transmission occurrences happen through person-to-person contact, contact with infected surfaces or environments, ingestion of food, or ingestion of water [107]. Allergies, environmental pollutants, and vigorous physical activity also contribute to the development of gastrointestinal illnesses [95,108]. Between 30% and 50% of people who play power endurance sports may have 1 or more problems with their stomachs, such as diarrhea, nausea, vomiting, and heartburn [108]. It has also been shown that variables such as gastrointestinal symptoms at rest, age, gender, diet, and years of training are associated with gastrointestinal discomfort during exercise [109].

Skin System Illness

Olympic athletes are susceptible to traumatic, environmental, and infectious skin disorders [110]. At the 1996 Atlanta Olympic Village polyclinic, 261 cases of dermatological conditions were documented, including rashes, infections, and other skin-related issues [111]. During the Tokyo 2020 Paralympic Games, there were 73 cases of skin illnesses, accounting for 26.1% (73/280) of all illnesses [112]. Skin and subcutaneous tissue illnesses occurred in 91 cases during the London 2012 Paralympic Games and are the second most common illnesses [113]. During the PyeongChang 2018 Paralympic Winter Games, skin system disorders such as limb or prosthetic interface lesions, residual limb or skin subcutaneous pain, rashes, skin breakdown, and abscesses were observed [114]. Skin system illnesses were the third most common system illness during the Beijing 2022 Paralympic and Paralympic Winter Games [5,14].

The main influences on skin system illnesses are infections, high-intensity physical activity, and specific competitions [115]. Skin system illnesses are particularly common among Paralympic athletes. Causes of skin infections in athletes with spinal cord injuries, amputations or limb deficiencies, and so forth, include loss of sensation and prolonged exposure to pressure in wheelchair-bound athletes, high forces at the residual limb-socket interface, hot or humid environments, sweating during exercise, and possible bacterial contamination in the exercise environment [95]. Skin and subcutaneous tissue are the most commonly affected systems in wheelchair basketball, weight lifting, and seated volleyball [113]. Certain sports are particularly prone to frictional stresses caused by repetitive movements, such as jogging, racquetball, baseball, golf, or any sport that entails frequent contact with equipment [115].

Heat-Related Illness

Heat-related illnesses affect a significant number of athletes and Olympic professionals at the Summer Olympic Games. A total of 213 heat-related illnesses were reported at the 1984 Los Angeles Olympics [116]; 372 people were treated for heat-related illnesses at the 1996 Atlanta Olympics, with symptoms including heat cramps, dehydration, heat syncope, and heatstroke [117]; at the 2008 Beijing Olympics, during the opening ceremony alone, there were up to 381 cases of heat-related illnesses [118]; and 225 cases of heatstroke occurred during the 2021 Tokyo Olympics, of which 100 were treated in athlete clinics and 125 in spectator clinics [119]. Therefore, it is evident that a rise in temperatures and other environmental factors during the Summer Olympic Games will result in heat-related illnesses among a significant number of people.

Factors contributing to mass heat-related illnesses are the environment, the specific competition, and the climate of the host country, with climate being the dominant factor. Spectators in hot outdoor environments, among others, account for the majority of heat-related illness visits (939/1056, 88.9%), while members of the Olympic family are largely indoors in cooler temperatures, where heat-related illnesses are rare [91]. A study pointed out that marathon, race walking, cycling, and competitive swimming are the Olympic sports with the highest frequency of heatstroke illnesses [120]. The climate of Atlanta, which is expected to pose a significant risk for heat-related

XSL•FO
RenderX

illnesses [120], is consistent with this study, showing that the climate of the host country of an event plays a determining role in the occurrence of heat-related illnesses. As an example, Tokyo experiences a tropical monsoon climate, Beijing experiences a temperate monsoon climate, and Los Angeles experiences a Mediterranean climate (Figure 3). The shared features of the Olympic Games held in these countries are high temperatures, which directly contribute to the prevalence of heat-related illnesses.

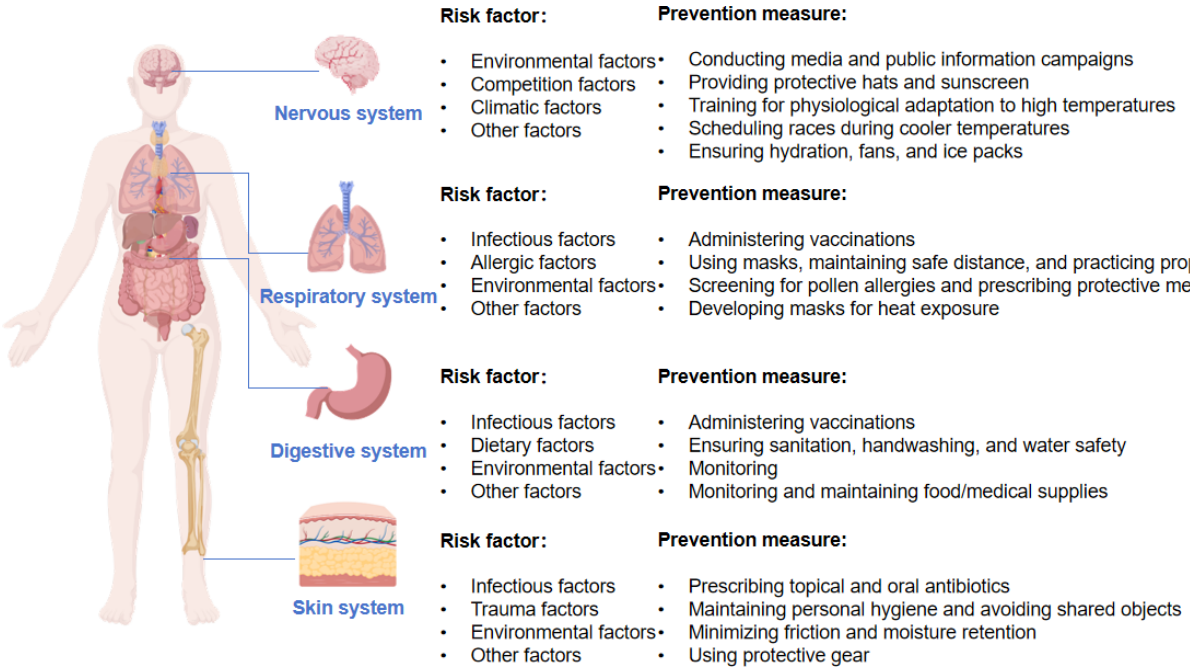
Figure 3. Geospatial location of the Olympic host cities included in the study.

Other Organ or System Illness

Furthermore, there was a higher prevalence of neurological illnesses, genitourinary illnesses, and circulation illnesses. A total of 38 neurological illnesses were recorded at the 2016 Rio Olympics, representing 6% of all illnesses [7]. At the opening ceremony of the 2008 Beijing Olympics, a total of 25 instances of urinary tract illnesses were observed [118]. The 2008 Beijing Olympics and Paralympics reported a total of 49 instances of cardiac illness and 34 instances of vascular illnesses [121]. In addition to this, psychological issues such as competition stress and mental illness in the Olympic Games, especially the Paralympic Games, should also be given attention [111,122]. Special consideration should be given to illnesses associated with colds in susceptible populations participating in the Winter Olympic and Paralympic Games [123]. There have been numerous fatalities at the Olympic Games, with 1 athlete in each of the snowboarding events in 1964 and 2010 losing their life [124,125], a cameraman dying of cardiac arrest during the 1996 Olympic Games [91], a doctor from the Alpine Ski Team accidentally colliding with a tracked snowmobile and dying instantly in 2012 [123], and an athlete at the 2016 Summer Paralympic Games in Rio who experienced a head injury during the competition (Figure 4) [126].

XSL•FO
RenderX

Figure 4. Factors affecting high-risk illness and preventive measures in the Olympic Games. Figure 4 is an original image created by the authors; however, the organ images were sourced from BioGDP [127].



Risk and Influencing Factors of Injuries

Information on injuries related to the Summer and Winter Olympic Games is given in Table 1. Alpine skiing, Nordic skiing, and sledge hockey are the 3 key sports in the Winter Paralympics, and their associated sports injuries are more prominent [122]. Compared with recreational sports, the Olympics are faster, higher, and more competitive, with athletes having greater difficulty completing competitions and often neglecting minor injuries for the sake of winning honors, which can lead to more serious injuries [128,129]. The most common areas of sports injuries in the Summer Olympics are focused on the lower limbs and lower back, which may be related to the fact that the lower limbs and lower back need to bear the weight and impact of the human body in high-speed running, change of direction, and jumping, as well as high confrontation and high-load states, and are prone to injuries [130]. The most common types of injuries in Winter Olympic Games may be related to the movements executed in these sports, many of which involve high-speed twisting and shear forces [128]. For instance, ice hockey, a sport included in the Winter Olympics, is a sport marked by fast movements and intense body contact, resulting in contusions and sprains in vulnerable anatomical regions such as the hands [131]. Paralympic athletes experience a higher frequency of injuries to their upper extremities, namely, the shoulder, wrist, and elbow. This is because wheelchair athletes heavily depend on their shoulders for all their activities. Additionally, the nature of the injuries experienced by some of these athletes results in a smaller number of athletes participating in competitions with fully functioning lower extremities compared with their upper extremities [132]. The effect of competition difficulty, protective equipment, and assistive devices on injury has been noted in Winter Paralympic injury research [122]. Compared with summer sports, winter sports have a higher incidence of contusions, fractures, and concussions due to impact and speed [133]. Additionally, the Olympic Games have introduced new sports, posing challenges to precompetition and in-competition prevention and treatment efforts [129].

Table . Information on injuries related to the Summer and Winter Olympic Games.

Games	Highest risk competitions	Most common areas of injury	Most common types of injuries
Summer Olympics [130]	BMX, taekwondo, football, rugby, mountain biking, boxing, handball, hockey, weight lifting, etc	Knee, thigh, ankle, lumbar spine, calf, and foot	Joint sprains or ligament ruptures, muscle strains, contusions or hematomas, skin lacerations, and tendinopathies
Winter Olympics [131]	Freestyle skiing, snowboarding, alpine skiing, bobsledding, and ice hockey	Knees, chest/waist/back, and wrists/hands/fingers	Contusions, hematomas, and bruises, followed by strains (including muscle breaks, tears, or tendon ruptures) and sprains (including dislocations, subluxations, and ligament ruptures)

XSL•FO
RenderX

Risk and Influencing Factors of Terrorist Attacks

The Olympic Games are highly susceptible to terrorist groups due to their prominent presence, the magnitude of the events, and the significant number of spectators. On September 5, 1972, a group of 11 Israeli athletes and coaches were abducted and subsequently murdered by members of the Black September organization at the Athletes' Village [18], and during the 1996 Olympics in Atlanta, the 100-year anniversary of the Olympic Park bombing killed 2 people and sent 111 victims to city hospitals [17]. Besides terrorist acts that took place immediately during the Olympic Games, there have been other terrorist occurrences that are either directly or indirectly associated with the Olympic Games or other significant athletic events. Following the news of London's successful attempt to host the 2012 Olympic Games on July 7, 2005, 4 terrorist assaults occurred in the city within 24 hours, resulting in the loss of 52 lives. On the eve of the 2010 World Cup final, an al-Qaeda spokesman declared on a North African jihadist website that the tournament would be targeted by suicide bombers using "undetectable explosives" [18]. In a study that noted at least 74 terrorist attacks against stadiums in the last 50 years, there have been 4 or more attacks in each subsequent year since 2011, with a peak of 11 in 2015 [134]. In recent years, although there have been no terrorist attacks at Olympic Games events, the Games may face other emergencies, such as demonstrations, bombings, and so forth, including bioterrorist attacks.

Illness, Injury, and Terrorism Risk Surveillance System

Injury and illness surveillance at major sporting events prior to the 1990s relied primarily on traditional paper records and manual counting [135]. Actually, this appears to be the situation with the monitoring of the Olympic Games. With the rise of the internet and information technology, injury surveillance systems have gone through the following 3 phases.

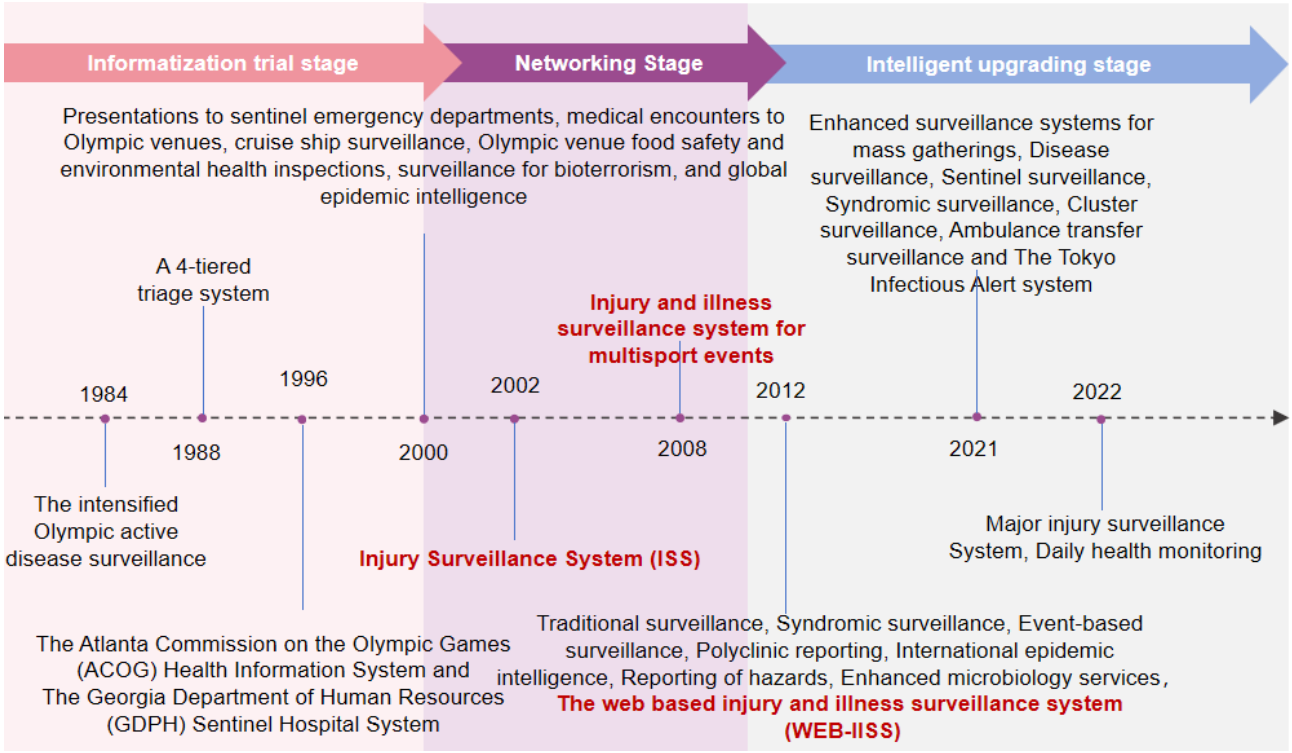
The stage of information technology attempts was roughly from the 1980s to the early 2000s. With the rise of computer and internet technology, the Olympic Games began to experiment with the application of information technology to injury and illness surveillance, recording athletes' health status through a preliminary surveillance network system or an electronic medical record system, and using databases for data storage and management. For example, the active illness surveillance system during the 1984 Summer Olympics in Los Angeles included a network of 46 hospitals, 90 doctors' offices, and so forth [136]. Two complementary public health surveillance systems, the Georgia Department of Human Resources Sentinel Hospital System and the Atlanta Commission for the Olympic Games Health Information System, were established for the 1996 Summer Olympic Games in Atlanta [117]. Early surveillance methods still do not have convenient and efficient features. They are achieved through the employment of workers and fixed weekly calls with the match venue to achieve the purpose of surveillance, such as 4 levels of detection of injury classification scale, in order for the researchers of the charts to be retrospectively reviewed in order to understand the patient's body-related situation [136,137] .

The networking phase was roughly from the beginning of the 21st century to around 2010. With the spread and development of internet technology, the Olympics began to establish a web-based injury and illness surveillance system. These technologies provide instantaneous transmission and exchange of data to enhance the effectiveness and precision of surveillance mechanisms. For the Sydney 2000 Olympic Games, public health planning covering disaster response planning, environmental hazard monitoring, and food safety strategies was developed, as well as surveillance systems covering both domestic and international parts of the country, including enhanced infectious illness surveillance, bioterrorism monitoring, global outbreak intelligence, and so forth [138,139]. The 2002 Salt Lake Winter Olympics developed a new public health surveillance system based on electronic medical records, with the goal of identifying unusual patterns of symptoms that could represent clusters of illness, with the hope of identifying these events of significance in a timely manner so as to allow an adequate public health response and management [140]. The International Paralympic Committee implemented the injury surveillance system and continued to do so at all subsequent Winter Paralympic Games [141]. The Athens 2004 Olympic Games used a team sports injury-reporting system, which reported information such as the injured player's jersey number, time of game, site of injury, and type of injury [142]. The 2008 Beijing Olympics used a medical intake system to record unusual patterns of illness or injury that required further investigation [92].

The intelligent enhancement stage is mainly from 2010 to the present. With the rapid development of big data, artificial intelligence, and other technologies, the Olympic Games' web-based injury and illness surveillance system has been continuously upgraded and improved, and the system is able to use data analysis technology to dig out the rules and trends behind the data, providing a scientific basis for the formulation of prevention and control measures (Figure 5). London 2012 Olympic and Paralympic Games used traditional monitoring, integrated monitoring, event-based surveillance, integrated clinic reporting, and so forth [16], with escalating scientific and technological methods. The IOC's injury and illness surveillance system for multisport events was used in the Olympic Games in Beijing 2008, London 2012, Sochi 2014, Rio 2016, and Pyeongchang 2018 [96-98,143]. The intelligence drive led to the first adoption of the Web-based Injury and Illness Surveillance System at the London 2012 Paralympic Games, and since then, the WEB-IISS has been adopted and continually improved and enhanced at the Sochi, Rio, PyeongChang, Tokyo, and other Paralympic Games [112,114,132,144,145]. Surveillance for the Tokyo 2021 Olympic Games also involves multiagency and multifaceted aspects, including Japan's Extreme Emergency and Post-Disaster Surveillance, Open-Source Epidemic Intelligence, and Multi-Surveillance System [119,146,147]. Injury surveillance for the Beijing 2022 Winter Olympic Games remains an integral part of the IOC's initiative to protect the health of athletes and record injuries and illnesses in a web-based tabular format, as well as data from the local organizing committee's medical (polyclinic) facilities and on-site medical support [14,119].

XSL•FO
RenderX

Figure 5. Main surveillance systems for injury and illness at previous Olympic Games.



Discussion

General Characteristics

There were only 12 studies on the Summer Olympics, Winter Olympics, and Paralympics before 2000, and a total of 108 studies published in 2000 and beyond, particularly in recent years when the amount of literature in this area of research reached an unprecedented level, reflecting the fact that in recent years the importance attached to health and safety issues at the Olympic Games has reached an unprecedented level. Furthermore, the 3 most extensively studied games also demonstrate the growing emphasis on health and safety in the games in recent years. The 120 studies involving comprehensive papers on injury and disease, single competitions, body parts, and single injuries or illnesses show that in addition to holistic studies of illness and injury, scholars are also focusing on high-risk competitions and major injury areas. It is noteworthy that only 2 studies were conducted on terrorism, reflecting, on the one hand, the theme of peace and development and the near absence of terrorism in recent years and, on the other hand, the lack of ability to think in peace and insufficient research on new types of bioterrorist attacks.

Related Health Risks

Illness is the most common risk factor in Olympic events. Illnesses of the respiratory system, the gastrointestinal system, and the skin system should pay particular attention to the infectious factors, which can be prevented and controlled in terms of the sources of infection, the route of transmission, and the susceptible population. In controlling the source of infection, attention is paid to screening for incoming illnesses, early diagnosis and isolation, and tracking, surveillance, and, if necessary, isolation and observation of close contacts of patients [16,148]. Emphasis should also be placed on dynamic surveillance and early detection of global illnesses, such as open-source epidemiological intelligence and multisource surveillance systems for infectious diseases [146,147]. In terms of cutting off the routes of transmission, it is advocated to wash hands frequently, wear masks, and avoid touching the mouth, nose, and eyes with hands, and so forth, so as to reduce the chances of pathogens spreading through hand-to-hand contact [6]. Maintaining cleanliness and hygiene in living and competition environments, with regular disinfection and ventilation, thereby reduces the survival time of pathogens in the environment and decreases pathogen concentrations in enclosed spaces [14]. Pay special attention to food and drinking water hygiene, avoiding food sold by street vendors, following water safety recommendations, and so forth [149]. With regard to the protection of susceptible groups, timely inoculation of appropriate vaccines and popularization of knowledge on the prevention and treatment of infectious illnesses, special attention is paid to the older adults, children, pregnant women, and other vulnerable and highly sensitive groups, and the necessary protective measures are taken, such as reducing the number of outings and avoiding crowded places [150].

Furthermore, illnesses affecting the respiratory and gastrointestinal systems, among others, can be attributed to a range of potential influencing factors such as environmental conditions (temperature, humidity, altitude, pollution, and pollen), diet, and exposure to diverse cultures, populations, and pathogens [151]. Particular attention should therefore be paid to environmental challenges related to heat, air quality, and water quality [152]. The management of respiratory system illnesses resulting from high altitude or cold conditions, air pollution, and pollen allergies includes the use of heat and moisture exchange respiratory device masks, allergy screening,

XSL•FO
RenderX

and the administration of preventive drugs for illness prevention and control [100,102]. Dietary and nutraceutical approaches to regulate the microbial, together with the avoidance of nonsteroidal anti-inflammatory medicines and high fructose meals, can effectively treat gastrointestinal system issues induced by intense aerobic activity [103,153]. For illnesses of the skin system, especially moisture, sweat, friction [110], UV exposure [154] of body parts of specific competition events, and (people with disabilities) special populations should be considered. As heat-related illnesses can cause a medical surge, athletes and spectators (especially children, older adults, and vulnerable groups with underlying medical conditions) should be cooled, protected from the sun, hydrated, and avoided from competing in the heat [155]. Attention is also drawn to the recommendations of the World Economic Forum's Global Commission on Biodiversity Cities on urban greening and the installation of white reflective roofs and continuously shaded pavements [156]. Furthermore, apart from preventing the aforementioned illnesses, our upcoming studies should also prioritize the mental health of professional athletes [122,128,157]. The issue of cardiac arrest should be taken seriously as deaths have occurred in previous competitions and reference can be made to the Resuscitation Council of Great Britain's best practice guidelines on cardiac arrest [158]. Vigilance against traditional high-risk [159] and emerging unpredictable infectious illnesses should also be an important measure in illness prevention.

The IOC strongly encourages National Olympic Committees and International Federations to protect the health of athletes and actively focus on injury prevention [160]. It has been suggested that changes in the incidence of injuries may be the result of changes in the composition of the Olympic program, changes in environmental factors (eg, venue or track design, rules of the game, or other factors), changes in the awareness of athletes and their medical staff of the importance of recognizing and reporting minor incidents, and the availability of training in the recording of injuries and illnesses [96,98]. For example, encouraging taekwondo athletes to use more head kicks and removing helmet protection for boxers have led to a high incidence of injuries in combat sports, with the removal of protective helmet protection tripling the overall risk of injuries [143]. Injuries are common in the Olympic Games, and risk factors for injury mechanisms in high-risk events, video analysis, and precise biomechanics factors for movement techniques are essential for better direct injury prevention strategies [128,160]. Difficult and high-speed races are highly prone to injuries, and among the measures to prevent injuries are changing the rules of the competition, strengthening protective measures, and optimizing the track. For example, at the Sochi and PyeongChang Winter Olympics, modifiable injury risk factors were identified: redesigning the track, increasing the number of training sessions, changing the start time of the competition, and so forth [161]. Meanwhile, in the training of medical personnel, emphasis should be placed on clarifying common types of injuries and emergency procedures, with targeted training on contusions, muscle breaks, and sprains, including those to the knee, thoracic spine, lumbar spine, back, and other parts of the body [131]. Some studies, however, have also found an inverse relationship between the size of the National Olympic Committee (NOC) and the risk of health problems, noting that athletes from the smallest NOCs have significantly higher prevalence rates than athletes from the largest NOCs [98]. This is because larger delegations, often from countries with well-developed exercise physiology and sports medicine communities, can provide more comprehensive medical care and closer medical follow-up during the Olympic flag period, and therefore the IOC and host countries should consider the allocation of medical resources as a whole, with the inclusion of multilingual triage nurses and doctors as a noteworthy area of concern in the allocation of medical resources [162].

The problem of terrorist attacks at the Olympic Games has a long history, but in recent years, strong initiatives such as huge investments, high-tech equipment for monitoring, and armed military forces for security have hardly been associated with rumors of terrorist attacks at the Olympic Games [98]. Medical staff instructed on bomb blast injuries, pharmacies stocked with therapeutic drugs, and prepared to hose down contaminated facilities prior to the Atlanta Olympics Games [111]. The Sydney Olympics also established a bioterrorism-monitoring system [139]. The Tokyo Olympics also organized a number of training events to prepare for possible mass casualties, such as terrorist attacks [163]. It appears that plans have been formed for acts of terrorism. The world's unprecedented changes are accelerating, international political disputes and military conflicts are erupting at various points, religious and ethnic conflicts still exist, and the global development and security situation is intricate and complex, which also poses certain risks to the security of the Olympic Games. Emerging biotechnology has also increased the cost, difficulty, and impact of bioterrorism attacks, with a particular focus on physical and mental health. Because of the Olympic Games' widespread global influence, bioterrorism is a prime target, so it is critical to increase surveillance and take preventative measures beforehand.

Illness and Injury Surveillance System

Surveillance systems in the Olympic Games have progressed from traditional paper records and manual statistics to information technology, networking, and intelligence, which can better facilitate the IOC's discovery of injury and illness prevalence, but there are still shortcomings such as lagging surveillance technology, limitations in the scope of surveillance (eg, lack of epidemiological data on athletes' mental health [164] and little attention to female athletes in the epidemiological data on injuries and illnesses [165]), data standards lack of uniformity, unequal distribution of monitoring resources in delegations, equipment safety, and radiation risks. In the future, technological innovation will integrate cutting-edge information technologies such as Internet of Things and blockchain to enhance monitoring intelligence and automation, enabling real-time surveillance and rapid response. Concurrently, data sharing and utilization efforts will strengthen cross-departmental and interinstitutional collaboration to dismantle information silos while promoting open data sharing and effective utilization. In privacy protection, reasonable data use will be advanced while ensuring data security through reinforced measures including encryption and access controls, guaranteeing safety throughout data transmission, storage, and usage. Finally, monitoring effectiveness will adopt an integrated

XSL•FO
RenderX

ecosystem approach to assess population-level infectious disease risks, forging synergies among medical and laboratory systems, validated tools, and essential human and financial resources [166].

Limitations of the Study

There are 3 key limitations to this study. First, most current illness classifications are by body system, which then makes it difficult to explore the causes and influences that trigger illness when analyzing the literature. Second, the study is also constrained in comparing different competitions as many papers have used different approaches to illness and injury data collection and research reporting. Third, this study does not include papers written in languages other than English, and information on Olympic-related illnesses, injuries, and terrorism may be published in the host country's native language.

Conclusions

From the perspective of mass-gathering medicine, this study investigates the occurrence of illnesses, injuries, and terrorism in the Olympic Games, analyzes the real reasons behind them, and then proposes targeted preventive measures. Through this study, we found that respiratory system illnesses, gastrointestinal system illnesses, skin system illnesses, and so on, are more common, and at the same time, heat-related illnesses also affect the Olympic personnel involved in a larger way; fighting, cycling, and ball and snow sports are prone to crowd casualties, and injuries are mainly in the knee, thigh, ankle, chest, waist, back, and so on, and the common types of injuries are contusions, strains, and so on; and the advancement of biotechnology and the volatile international situation have made bioterrorism attacks possible. In terms of health and safety risk prevention, the focus should be on vaccine-preventable illnesses, while health education should be strengthened, and so on; the rules of the competition should be revised, protective equipment should be upgraded, medical resources should be rationally allocated, and so on; and bioterrorist attacks should also be given high attention. In addition, injury and illness surveillance systems should pay attention to technological innovation, resource sharing, and information confidentiality.

Acknowledgments

The authors declare that they received financial support for the research, authorship, and publication of this paper. The National Key Research and Development Program of China (2021YFC2600504) and the National Natural Science Foundation of China (72474152) supported this research, which relied on the China National Key Research and Development Program's Research on Bioterrorism Prevention and Control Technology in Important Public Places and Research on the Elastic Control Mechanism of Biological Triage Process Based on BPR Theory.

Authors' Contributions

ST, SX, and YZ designed the study; ST, ZC, and BS were involved in study selection and data extraction; ST, TT, and YZ contributed to the writing of the manuscript. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy and detailed study characteristics.
[DOCX File, 130 KB - publichealth_v11i1e66829_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 115 KB - publichealth_v11i1e66829_app2.pdf]

References

1. Public health for mass gatherings: key considerations. World Health Organization. 2008. URL: https://www.who.int/publications/i/item/public-health-for-mass-gatherings-key-considerations [accessed 2025-06-22]
2. Memish ZA, Stephens GM, Steffen R, Ahmed QA. Emergence of medicine for mass gatherings: lessons from the Hajj. Lancet Infect Dis 2012 Jan;12(1):56-65. [doi: 10.1016/S1473-3099(11)70337-1] [Medline: 22192130]
3. Abubakar I, Gautret P, Brunette GW, et al. Global perspectives for prevention of infectious diseases associated with mass gatherings. Lancet Infect Dis 2012 Jan;12(1):66-74. [doi: 10.1016/S1473-3099(11)70246-8] [Medline: 22192131]
4. Steffen R, Bouchama A, Johansson A, et al. Non-communicable health risks during mass gatherings. Lancet Infect Dis 2012 Feb;12(2):142-149. [doi: 10.1016/S1473-3099(11)70293-6] [Medline: 22252147]
5. Derman W, Runciman P, Eken M, et al. Incidence of injury and illness at the Beijing 2022 Paralympic Winter Games held in a closed-loop environment: a prospective cohort study of 7332 athlete days. Br J Sports Med 2024 Jul 25;58(15):836-843. [doi: 10.1136/bjsports-2023-107525] [Medline: 38346775]

XSL•FO
RenderX

6. Soligard T, Palmer D, Steffen K, et al. New sports, COVID-19 and the heat: sports injuries and illnesses in the Tokyo 2020 Summer Olympics. Br J Sports Med 2023 Jan;57(1):46-54. [doi: 10.1136/bjsports-2022-106155]
7. Soligard T, Steffen K, Palmer D, et al. Sports injury and illness incidence in the Rio de Janeiro 2016 Olympic Summer Games: a prospective study of 11274 athletes from 207 countries. Br J Sports Med 2017 Sep;51(17):1265-1271. [doi: 10.1136/bjsports-2017-097956] [Medline: 28756389]
8. Memish ZA, Blumberg L, Al-Maani AS, et al. Moving cholera vaccines ahead of the epidemic curve. Lancet 2024 Jan;403(10422):127-129. [doi: 10.1016/S0140-6736(23)02244-4]
9. WHO to identify pathogens that could cause future outbreaks and pandemics. World Health Organization. 2022. URL: https://www.who.int/news/item/21-11-2022-who-to-identify-pathogens-that-could-cause-future-outbreaks-and-pandemics [accessed 2025-06-22]
10. Jansen HJ, Breeveld FJ, Stijnis C, Grobusch MP. Biological warfare, bioterrorism, and biocrime. Clin Microbiol Infect 2014 Jun;20(6):488-496. [doi: 10.1111/1469-0691.12699] [Medline: 24890710]
11. Steffen K, Soligard T, Engebretsen L. Health protection of the Olympic athlete. Br J Sports Med 2012 Jun;46(7):466-470. [doi: 10.1136/bjsports-2012-091168] [Medline: 22661696]
12. Xiang-lan Z, Bi-bo P. Management and nursing practice for the patients in the polyclinic of Beijing Olympics Village. Chin J Nurs 2009;44(1):21-23.
13. Han PD, Gao D, Liu J, et al. Medical services for sports injuries and illnesses in the Beijing 2022 Olympic Winter Games. World J Emerg Med 2022;13(6):459-466. [doi: 10.5847/wjem.j.1920-8642.2022.106] [Medline: 36636567]
14. Soligard T, Palmer D, Steffen K, et al. Olympic Games during nationwide lockdown: sports injuries and illnesses, including COVID-19, at the Beijing 2022 Winter Olympics. Br J Sports Med 2024 Jan;58(1):11-17. [doi: 10.1136/bjsports-2023-107412]
15. Derman W, Schwellnus M, Jordaan E, et al. Illness and injury in athletes during the competition period at the London 2012 Paralympic Games: development and implementation of a WEB-based surveillance system (WEB-IISS) for team medical staff. Br J Sports Med 2013 May;47(7):420-425. [doi: 10.1136/bjsports-2013-092375]
16. McCloskey B, Endericks T, Catchpole M, et al. London 2012 Olympic and Paralympic Games: public health surveillance and epidemiology. Lancet 2014 Jun 14;383(9934):2083-2089. [doi: 10.1016/S0140-6736(13)62342-9] [Medline: 24857700]
17. Brennan RJ, Keim ME, Sharp TW, et al. Medical and public health services at the 1996 Atlanta Olympic Games: an overview. Med J Aust 1997;167(11-12):595-598. [doi: 10.5694/j.1326-5377.1997.tb138908.x] [Medline: 9418799]
18. Sugden J. Watched by the Games: surveillance and security at the Olympics. Int Rev Sociol Sport 2012 Jun;47(3):414-429. [doi: 10.1177/1012690211433482]
19. Memish ZA, Zumla A, McCloskey B, et al. Mass gatherings medicine: international cooperation and progress. Lancet 2014 Jun;383(9934):2030-2032. [doi: 10.1016/S0140-6736(14)60225-7]
20. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2007 Feb 1;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
22. Jegathesan M. Pattern of injuries and illnesses in the Malaysian Olympic Team. Med J Malaysia 1973 Jun;27(4):248-252. [Medline: 4270780]
23. Centers for Disease Control and Prevention. Prevention and management of heat-related illness among spectators and staff during the Olympic Games—Atlanta. JAMA 1996 Jul;276(8):593-595. [doi: 10.1001/jama.1996.03540080017008]
24. Cantwell JD. Cardiovascular events in the 1996 Olympic Games. Am J Cardiol 1997 Nov 15;80(10):1394. [doi: 10.1016/s0002-9149(97)00698-x] [Medline: 9388129]
25. Keim ME, Williams D. Hospital use by Olympic athletes during the 1996 Atlanta Olympic Games. Med J Aust 1997;167(11-12):603-605. [doi: 10.5694/j.1326-5377.1997.tb138910.x] [Medline: 9418801]
26. Nyland J, Snouse SL, Anderson M, Kelly T, Sterling JC. Soft tissue injuries to USA Paralympians at the 1996 summer games. Arch Phys Med Rehabil 2000;81(3):368-373. [doi: 10.1053/apmr.2000.0810368]
27. Athanasopoulos S, Kapreli E, Tsakoniti A, et al. The 2004 Olympic Games: physiotherapy services in the Olympic Village polyclinic. Br J Sports Med 2007 Sep;41(9):603-609. [doi: 10.1136/bjsm.2007.035204] [Medline: 17502331]
28. Succo G, Crosetti E, Mattiazzo A, Riontino E, Torino MG. XX Olympic and IX Paralympic Winter Games: the ENT experience. Acta Otorhinolaryngol Ital 2006;28(3):101-109. [Medline: 18646571]
29. Junge A, Engebretsen L, Mountjoy ML, et al. Sports injuries during the Summer Olympic Games 2008. Am J Sports Med 2009 Nov;37(11):2165-2172. [doi: 10.1177/0363546509339357] [Medline: 19783812]
30. Piat SC, Minniti D, Traversi D, Gianino MM, Massazza G, Siliquini R. Torino 2006 Winter Olympic Games: highlight on health services organization. J Emerg Med 2010 Oct;39(4):454-461. [doi: 10.1016/j.jemermed.2009.08.028]
31. Hanstad DV, Rønsen O, Andersen SS, Steffen K, Engebretsen L. Fit for the fight? Illnesses in the Norwegian team in the Vancouver Olympic Games. Br J Sports Med 2011 Jun;45(7):571-575. [doi: 10.1136/bjsm.2010.081364] [Medline: 21393259]
32. Taunton J, Wilkinson M, Celebrini R, et al. Paralympic medical services for the 2010 Paralympic Winter Games. Clin J Sport Med 2012 Jan;22(1):10-20. [doi: 10.1097/JSM.0b013e318242ec68] [Medline: 22222594]

XSL•FO
RenderX

33. Webborn N, Willick S, Emery CA. The injury experience at the 2010 Winter Paralympic Games. Clin J Sport Med 2012 Jan;22(1):3-9. [doi: 10.1097/JSM.0b013e318243309f] [Medline: 22222596]
34. Gawronski W, Sobiecka J, Malesza J. Fit and healthy Paralympians—medical care guidelines for disabled athletes: a study of the injuries and illnesses incurred by the Polish Paralympic team in Beijing 2008 and London 2012. Br J Sports Med 2013 Sep;47(13):844-849. [doi: 10.1136/bjsports-2013-092298] [Medline: 23902777]
35. Vanhegan IS, Palmer-Green D, Soligard T, et al. The London 2012 Summer Olympic Games: an analysis of usage of the Olympic Village "Polyclinic" by competing athletes. Br J Sports Med 2013 May;47(7):415-419. [doi: 10.1136/bjsports-2013-092325] [Medline: 23467963]
36. Grant ME, Steffen K, Glasgow P, Phillips N, Booth L, Galligan M. The role of sports physiotherapy at the London 2012 Olympic Games. Br J Sports Med 2014 Jan;48(1):63-70. [doi: 10.1136/bjsports-2013-093169] [Medline: 24334506]
37. Webborn N, Emery C. Descriptive epidemiology of Paralympic sports injuries. PM R 2014 Aug;6(8 Suppl):S18-S22. [doi: 10.1016/j.pmrj.2014.06.003] [Medline: 25134748]
38. Palmer-Green D, Elliott N. Sports injury and illness epidemiology: Great Britain Olympic Team (TeamGB) surveillance during the Sochi 2014 Winter Olympic Games. Br J Sports Med 2015 Jan;49(1):25-29. [doi: 10.1136/bjsports-2014-094206] [Medline: 25425714]
39. Steffen K, Engebretsen L. Injury and illness surveillance among Olympic athletes: summary of the 2010 Winter, and the 2008 and 2012 summer Olympic Games. Sports Inj Prev ;2015:39-50. [doi: 10.1007/978-4-431-55318-2_3]
40. Blauwet CA, Lexell J, Derman W. Paralympic sports medicine. In: Handbook of Sports Medicine and Science: Training and Coaching the Paralympic Athlete: Wiley; 2016:75-95.
41. Derman W, Schwellnus MP, Jordaan E, et al. Sport, sex and age increase risk of illness at the Rio 2016 Summer Paralympic Games: a prospective cohort study of 51 198 athlete days. Br J Sports Med 2018 Jan;52(1):17-23. [doi: 10.1136/bjsports-2017-097962] [Medline: 29074477]
42. Janse Van Rensburg DC, Schwellnus M, Derman W, Webborn N. Illness among Paralympic athletes: epidemiology, risk markers, and preventative strategies. Phys Med Rehabil Clin N Am 2018 May;29(2):185-203. [doi: 10.1016/j.pmr.2018.01.003] [Medline: 29627085]
43. Kim DS, Lee YH, Bae KS, et al. PyeongChang 2018 Winter Olympic Games and athletes' usage of "polyclinic" medical services. BMJ Open Sport Exerc Med 2019;5(1):e000548. [doi: 10.1136/bmjsem-2019-000548] [Medline: 31548900]
44. Nabhan D, Windt J, Taylor D, Moreau W. Close encounters of the US kind: illness and injury among US athletes at the PyeongChang 2018 Winter Olympic Games. Br J Sports Med 2020 Aug;54(16):997-1002. [doi: 10.1136/bjsports-2018-100015] [Medline: 31375500]
45. Chang JY, You SJH, Grant ME, et al. Review of physiotherapy service for athletes of 2018 Olympic Winter games: consideration of preparation for two polyclinics. Phys Ther Sport 2021 May;49:106-111. [doi: 10.1016/j.ptsp.2021.02.012] [Medline: 33667775]
46. Park H, Chang JY, Hwang J, Lee YH. Sports injury and physiotherapy services in the 2018 PyeongChang Winter Paralympic Games: considerations and potential recommendations for future Paralympics. Phys Ther Korea 2022 Aug 20;29(3):241-248. [doi: 10.12674/ptk.2022.29.3.241]
47. Noble-Taylor KE, Adams WM, Cali MG, et al. Incidence of injuries among Team USA athletes competing in Tokyo 2020 Summer Olympic Games. Med Sci Sports Exerc 2022;54(9S):666. [doi: 10.1249/01.mss.0000883380.16252.a0]
48. Kuroki N, Yagishita K, Shimizu K, et al. Emergency department activities at the Athletes' Village during the Tokyo 2020 Olympic and Paralympic Games. Acute Med Surg 2023 Jan;10(1). [doi: 10.1002/ams2.905]
49. Adam F, Derman W, Schwellnus M, Boer PH, Jordaan E, Runciman P. Injury incidence according to athlete impairment type during the 2012 and 2016 Summer Paralympic Games: a combined analysis of 101 108 athlete days. Br J Sports Med 2024 Jul 25;58(15):852-859. [doi: 10.1136/bjsports-2023-107144] [Medline: 38286574]
50. Anderson T, Cali MG, Clark SC, et al. Team USA injury and illness incidence at the 2022 Beijing Winter Olympic and Paralympic Games. Br J Sports Med 2024 Jan;58(1):25-33. [doi: 10.1136/bjsports-2023-107185]
51. McCloskey B, Saito T, Shimada S, et al. The Tokyo 2020 and Beijing 2022 Olympic Games held during the COVID-19 pandemic: planning, outcomes, and lessons learnt. Lancet 2024 Feb 3;403(10425):493-502. [doi: 10.1016/S0140-6736(23)02635-1] [Medline: 38244561]
52. Sakanashi S, Tanaka H, Yokota H, et al. Injuries and illness of athletes at the Tokyo 2020 Olympic and Paralympic summer games visiting outside facilities. Sports Med Health Sci 2024 Mar;6(1):48-53. [doi: 10.1016/j.smhs.2024.01.003] [Medline: 38463667]
53. Grim C, Hotfiel T, Engelhardt M, Plewinski S, Spahl O, Wolfarth B. Sports injuries and illnesses of the German National Team during the 2016 Olympic Summer Games in Rio de Janeiro. Sportverletz Sportschaden 2017 Jan;31(1):25-30. [doi: 10.1055/s-0043-101967] [Medline: 28359131]
54. Soria V. Beyond London 2012: The quest for a security legacy. RUSI J 2011;156(2):36-43. [doi: 10.1080/03071847.2011.576473]
55. Kubosch EJ, Kosel J, Steffen K, et al. Upcoming Paralympic summer games in Rio: what did the German medical team learn from the London Games? J Sports Med Phys Fitness 2017 Nov;57(11):1486-1493. [doi: 10.23736/S0022-4707.16.06537-3] [Medline: 27377254]

XSL•FO
RenderX

56. Valtonen M, Waris M, Vuorinen T, et al. Common cold in Team Finland during 2018 Winter Olympic Games (PyeongChang): epidemiology, diagnosis including molecular point-of-care testing (POCT) and treatment. Br J Sports Med 2019 Sep;53(17):1093-1098. [doi: 10.1136/bjsports-2018-100487]
57. Yashio T, Murayama A, Kami M, Ozaki A, Tanimoto T, Rodriguez-Morales AJ. COVID-19 infection during the Olympic and Paralympic Games Tokyo 2020. Travel Med Infect Dis 2021;44:102205. [doi: 10.1016/j.tmaid.2021.102205] [Medline: 34785374]
58. Akashi H, Shimada S, Tamura T, Chinda E, Kokudo N. SARS-CoV-2 infections in close contacts of positive cases in the Olympic and Paralympic Village at the 2021 Tokyo Olympic and Paralympic Games. JAMA 2022 Mar 8;327(10):978-980. [doi: 10.1001/jama.2022.0818] [Medline: 35113161]
59. Shimizu K, Mossialos E, Shibuya K. What has the 2020 Tokyo Olympic and Paralympic Games taught global health on sporting mass gatherings under COVID-19 pandemic? Anaesth Crit Care Pain Med 2022 Feb;41(1):101001. [doi: 10.1016/j.accpm.2021.101001]
60. Sugawara M, Manabe Y, Yamasawa F, Hosokawa Y. Athlete medical services at the marathon and race walking events during Tokyo 2020 Olympics. Front Sports Act Living 2022;4:872475. [doi: 10.3389/fspor.2022.872475] [Medline: 35529419]
61. Valtonen M, Waris M, Luoto R, et al. Non-SARS-CoV-2 respiratory viruses in athletes at major winter sport events, 2021 and 2022. Emerg Infect Dis 2022 Oct;28(10):2096-2099. [doi: 10.3201/eid2810.220478] [Medline: 36148959]
62. Runciman P, Blauwet C, Kissick J, et al. "Heading" in the right direction: concussions reported at the Tokyo 2020 Paralympic Games. Br J Sports Med 2022 Dec 13;57(1):bjsports-2022-106550. [doi: 10.1136/bjsports-2022-106550] [Medline: 36588429]
63. Xiong R, Zhou J, Li W, et al. Effects of a closed-loop system against SARS-CoV-2 at the Beijing 2022 Olympic Winter Games: a descriptive and modeling study. World J Emerg Med 2023;14(6):471-476. [doi: 10.5847/wjem.j.1920-8642.2023.105] [Medline: 37969215]
64. Yao Y, Wang P, Zhang H. The impact of preventive strategies adopted during large events on the COVID-19 pandemic: a case study of the Tokyo Olympics to provide guidance for future large events. Int J Environ Res Public Health 2023 Jan 29;20(3):2408. [doi: 10.3390/ijerph20032408] [Medline: 36767780]
65. Wang RX, Wang ZM, Tian HY. Application of mathematical models of infectious diseases in the evaluation of COVID-19 transmission risk at mass gatherings. Zhonghua Yu Fang Yi Xue Za Zhi 2022 Aug 6;56(8):1055-1061. [doi: 10.3760/cma.j.cn112150-20220429-00433] [Medline: 35922231]
66. Ekeland A, Dimmen S, Lystad H, Aune AK. Completion rates and injuries in alpine races during the 1994 Olympic Winter Games. Scand J Med Sci Sports 1996 Oct;6(5):287-290. [doi: 10.1111/j.1600-0838.1996.tb00472.x] [Medline: 8960650]
67. Allen TL, Jolley SJ, Cooley VJ, et al. The epidemiology of illness and injury at the alpine venues during the Salt Lake City 2002 Winter Olympic Games. J Emerg Med 2006 Feb;30(2):197-202. [doi: 10.1016/j.jemermed.2005.07.006] [Medline: 16567258]
68. Grissom CK, Finnoff JT, Murdock DC, Culberson JT. Nordic venue medical services during the 2002 Winter Olympics. J Emerg Med 2006 Feb;30(2):203-210. [doi: 10.1016/j.jemermed.2005.08.005] [Medline: 16567259]
69. Shadgan B, Feldman BJ, Jafari S. Wrestling injuries during the 2008 Beijing Olympic Games. Am J Sports Med 2010 Sep;38(9):1870-1876. [doi: 10.1177/0363546510369291] [Medline: 20522826]
70. Pocecco E, Ruedl G, Stankovic N, et al. Injuries in judo: a systematic literature review including suggestions for prevention. Br J Sports Med 2013 Dec;47(18):1139-1143. [doi: 10.1136/bjsports-2013-092886] [Medline: 24255909]
71. Blauwet C, Cushman D, Derman W, Schwellnus M, Willick S. The epidemiology of injuries in Paralympic athletics (track and field) at the London 2012 Paralympic Games. Med Sci Sports Exerc 2015;47(5S):903. [doi: 10.1249/01.mss.0000479181.69154.9e]
72. Webborn N, Cushman D, Blauwet CA, et al. The epidemiology of injuries in football at the London 2012 Paralympic Games. PM R 2016 Jun;8(6):545-552. [doi: 10.1016/j.pmrj.2015.09.025] [Medline: 26454234]
73. Willick SE, Cushman DM, Blauwet CA, et al. The epidemiology of injuries in powerlifting at the London 2012 Paralympic Games: an analysis of 1411 athlete-days. Scand J Med Sci Sports 2016 Oct;26(10):1233-1238. [doi: 10.1111/sms.12554] [Medline: 26453890]
74. Derman W, Blauwet C, Webborn N, Schwellnus M, Vliet PD, Lazarovski D. Mitigating risk of injury in alpine skiing in the Pyeongchang 2018 Paralympic Winter Games: the time is now!. Br J Sports Med 2018 Apr;52(7):419-420. [doi: 10.1136/bjsports-2017-098864] [Medline: 29351955]
75. Edouard P, Steffen K, Junge A, Leglise M, Soligard T, Engebretsen L. Gymnastics injury incidence during the 2008, 2012 and 2016 Olympic Games: analysis of prospectively collected surveillance data from 963 registered gymnasts during Olympic Games. Br J Sports Med 2018 Apr;52(7):475-481. [doi: 10.1136/bjsports-2017-097972] [Medline: 29032364]
76. Thomas RE, Thomas BC. A systematic review of injuries in gymnastics. Phys Sportsmed 2019 Feb;47(1):96-121. [doi: 10.1080/00913847.2018.1527646] [Medline: 30246587]
77. Ashikaga K, Yoneyama K, Hirayama K, et al. Medical care provision at the venue of the weightlifting event of the Tokyo 2020 Olympic Games. Sport Sci Health 2022;18(3):847-852. [doi: 10.1007/s11332-021-00865-1] [Medline: 34840613]

XSL•FO
RenderX

78. Yanagawa Y, Muto T, Shakagori M, Terakado A, Kumai T. Activities of medical centers for athletes and spectators at cycling track events in the Tokyo 2020 Olympic and Paralympic Games. Acute Med Surg 2022 Jan;9(1). [doi: 10.1002/ams2.728]
79. Yagi M, Kasanami R, Tarumi Y, Dohi K. Medical care management based on disaster medicine for the triathlon events at the XXXII Olympiad and Tokyo 2020 Paralympic Games. Int J Environ Res Public Health 2023 Oct 7;20(19):6891. [doi: 10.3390/ijerph20196891] [Medline: 37835161]
80. Piccininni PM, Fasel R. Sports dentistry and the Olympic Games. J Calif Dent Assoc 2005 Jun;33(6):471-483. [Medline: 16060340]
81. Vougiouklakis G, Tzoutzas J, Farmakis ET, Farmakis EE, Antoniadou M, Mitsea A. Dental data of the Athens 2004 Olympic and Paralympic Games. Int J Sports Med 2008 Nov;29(11):927-933. [doi: 10.1055/s-2008-1038489] [Medline: 18418811]
82. Badekas T, Papadakis SA, Vergados N, et al. Foot and ankle injuries during the Athens 2004 Olympic Games. J Foot Ankle Res 2009 Apr 12;2:9. [doi: 10.1186/1757-1146-2-9] [Medline: 19361341]
83. Yang XJ, Schamach P, Dai JP, et al. Dental service in 2008 Summer Olympic Games. Br J Sports Med 2011 Apr;45(4):270-274. [doi: 10.1136/bjsm.2010.075283] [Medline: 21047841]
84. D'Ath PJ, Thomson WD, Wilson CM. Seeing you through London 2012: eye care at the Olympics. Br J Sports Med 2013 May;47(7):463-466. [doi: 10.1136/bjsports-2013-092152] [Medline: 23486936]
85. On MG, Oh JR, Jang YH, Kim DS. Epidemiologic study of shoulder injuries in the PyeongChang 2018 Winter Olympic Games. Clin Orthop Surg 2019 Jun;11(2):187-191. [doi: 10.4055/cios.2019.11.2.187] [Medline: 31156771]
86. Fernandes AG, Bastawrous A, Ferraz NN, et al. Eye clinic attendance at the Olympic and Paralympic Games Rio 2016 and its correlation to the WHO indicators on eye health. Br J Sports Med 2020 Oct 8:bjsports-2020-102706. [doi: 10.1136/bjsports-2020-102706] [Medline: 33032992]
87. Derman W, Runciman P, Eken M, et al. Incidence and burden of injury at the Tokyo 2020 Paralympic Games held during the COVID-19 pandemic: a prospective cohort study of 66 045 athlete days. Br J Sports Med 2023 Jan;57(1):63-70. [doi: 10.1136/bjsports-2022-106234]
88. Bahr R, Clarsen B, Derman W, et al. International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in Sport 2020 (including STROBE Extension for Sport Injury and Illness Surveillance (STROBE-SIIS)). Br J Sports Med 2020 Apr;54(7):372-389. [doi: 10.1136/bjsports-2019-101969] [Medline: 32071062]
89. International Statistical Classification of Diseases and Related Health Problems (ICD). World Health Organization. 2024. URL: https://www.who.int/standards/classifications/classification-of-diseases [accessed 2024-06-17]
90. Hull JH, Clemm H, Backer V, et al. Protecting respiratory health of athletes: an Olympic challenge. Lancet Respir Med 2024 Jul;12(7):509-510. [doi: 10.1016/S2213-2600(24)00183-8]
91. Wetterhall SF, Coulombier DM, Herndon JM, Zaza S, Cantwell JD. Medical care delivery at the 1996 Olympic Games. Centers for Disease Control and Prevention Olympics Surveillance Unit. JAMA 1998 May 13;279(18):1463-1468. [doi: 10.1001/jama.279.18.1463] [Medline: 9600481]
92. Lang XY, Lan L, Chen WN, et al. Disease distribution and medical resources during the Beijing 2008 Olympic and Paralympic Games. Chin Med J 2011;124(7):1031-1036. [doi: 10.3760/cma.j.issn.0366-6999.2011.07.016]
93. Torres JR. Are rapid antigen SARS-Cov-2 tests effective for mass screening of travelers at airports? The Olympic experience. J Travel Med 2021 Oct 11;28(7):taab135. [doi: 10.1093/jtm/taab135] [Medline: 34480181]
94. Wang X, Yan X, Pei S, Jia Z. Evaluation of effectiveness of COVID-19 prevention and control in a closed-loop management manner during Beijing 2022 Winter Olympics and Paralympics Games. Dis Surveill 2022;37(12):1594-1598. [doi: 10.3784/jbjc.202208060347]
95. Vanlandewijck YC, Thompson WR. Training and Coaching the Paralympic Athlete: Handbook of Sports Medicine and Science: Wiley; 2016:1-244. [doi: 10.1002/9781119045144]
96. Soligard T, Steffen K, Palmer-Green D, et al. Sports injuries and illnesses in the Sochi 2014 Olympic Winter Games. Br J Sports Med 2015 Apr;49(7):441-447. [doi: 10.1136/bjsports-2014-094538]
97. Yoon J, Bae M, Kang H, Kim T. Descriptive epidemiology of sports injury and illness during the Rio 2016 Olympic Games: a prospective cohort study for Korean team. Int J Sports Sci Coach 2018 Dec;13(6):939-946. [doi: 10.1177/1747954118768686]
98. Soligard T, Palmer D, Steffen K, et al. Sports injury and illness incidence in the PyeongChang 2018 Olympic Winter Games: a prospective study of 2914 athletes from 92 countries. Br J Sports Med 2019 Sep;53(17):1085-1092. [doi: 10.1136/bjsports-2018-100236] [Medline: 31235615]
99. Kippelen P, Fitch KD, Anderson SD, et al. Respiratory health of elite athletes—preventing airway injury: a critical review. Br J Sports Med 2012 Jun;46(7):471-476. [doi: 10.1136/bjsports-2012-091056] [Medline: 22522585]
100. Katelaris CH, Carrozzi FM, Burke TV, Byth K. A springtime Olympics demands special consideration for allergic athletes. J Allergy Clin Immunol 2000 Aug;106(2):260-266. [doi: 10.1067/mai.2000.108603] [Medline: 10932068]
101. Rundell KW. Effect of air pollution on athlete health and performance. Br J Sports Med 2012 May;46(6):407-412. [doi: 10.1136/bjsports-2011-090823]

XSL•FO
RenderX

102. Hanstock HG, Ainegren M, Stenfors N. Exercise in sub-zero temperatures and airway health: implications for athletes with special focus on heat-and-moisture-exchanging breathing devices. Front Sports Act Living 2020;2:34. [doi: 10.3389/fspor.2020.00034]
103. Bertuccioli A, Zonzini GB, Cazzaniga M, et al. Sports-related gastrointestinal disorders: from the microbiota to the possible role of nutraceuticals, a narrative analysis. Microorganisms 2024 Apr 16;12(4):804. [doi: 10.3390/microorganisms12040804] [Medline: 38674748]
104. Engebretsen L, Soligard T, Steffen K, et al. Sports injuries and illnesses during the London Summer Olympic Games 2012. Br J Sports Med 2013 May;47(7):407-414. [doi: 10.1136/bjsports-2013-092380] [Medline: 23515712]
105. Kim K, Jang JY, Moon G, et al. Experiences of the emergency department at the Pyeongchang polyclinic during the 2018 PyeongChang Winter Olympic Games. Yonsei Med J 2019 May;60(5):474-480. [doi: 10.3349/ymj.2019.60.5.474] [Medline: 31016910]
106. Kageyama T, Shinohara M, Uchida K, et al. Coexistence of multiple genotypes, including newly identified genotypes, in outbreaks of gastroenteritis due to Norovirus in Japan. J Clin Microbiol 2004 Jul;42(7):2988-2995. [doi: 10.1128/JCM.42.7.2988-2995.2004] [Medline: 15243049]
107. Jeong MH, Song YH, Ju SY, Kim SH, Kwak HS, An ES. Surveillance to prevent the spread of norovirus outbreak from asymptomatic food handlers during the PyeongChang 2018 Olympics. J Food Prot 2021 Oct 1;84(10):1819-1823. [doi: 10.4315/JFP-21-136] [Medline: 34115864]
108. Brouns F, Beckers E. Is the gut an athletic organ? Digestion, absorption and exercise. Sports Med 1993 Apr;15(4):242-257. [doi: 10.2165/00007256-199315040-00003] [Medline: 8460288]
109. Peters HPF, Bos M, Seebregts L, et al. Gastrointestinal symptoms in long-distance runners, cyclists, and triathletes: prevalence, medication, and etiology. Am J Gastroenterol 1999;94(6):1570-1581. [doi: 10.1111/j.1572-0241.1999.01147.x]
110. De Luca JF, Adams BB, Yosipovitch G. Skin manifestations of athletes competing in the Summer Olympics: what a sports medicine physician should know. Sports Med 2012 May 1;42(5):399-413. [doi: 10.2165/11599050-000000000-00000] [Medline: 22512412]
111. Eaton SB, Woodfin BA, Askew JL, et al. The polyclinic at the 1996 Atlanta Olympic Village. Med J Aust 1997;167(11-12):599-602. [doi: 10.5694/j.1326-5377.1997.tb138909.x] [Medline: 9418800]
112. Derman W, Runciman P, Eken M, et al. Incidence and burden of illness at the Tokyo 2020 Paralympic Games held during the COVID-19 pandemic: a prospective cohort study of 66 045 athlete days. Br J Sports Med 2023 Jan;57(1):55-62. [doi: 10.1136/bjsports-2022-106312]
113. Schwellnus M, Derman W, Jordaan E, et al. Factors associated with illness in athletes participating in the London 2012 Paralympic Games: a prospective cohort study involving 49,910 athlete-days. Br J Sports Med 2013 May;47(7):433-440. [doi: 10.1136/bjsports-2013-092371] [Medline: 23525473]
114. Derman W, Runciman P, Jordaan E, et al. Incidence rate and burden of illness at the Pyeongchang 2018 Paralympic Winter Games. Br J Sports Med 2019 Sep;53(17):1099-1104. [doi: 10.1136/bjsports-2018-100096] [Medline: 30636703]
115. Carr PC, Cropley TG. Sports dermatology: skin disease in athletes. Clin Sports Med 2019 Oct;38(4):597-618. [doi: 10.1016/j.csm.2019.06.001] [Medline: 31472769]
116. Baker WM, Simone BM, Niemann JT, Daly A. Special event medical care: the 1984 Los Angeles Summer Olympics experience. Ann Emerg Med 1986 Feb;15(2):185-190. [doi: 10.1016/s0196-0644(86)80019-1] [Medline: 3946862]
117. Centers for Disease Control and Prevention (CDC). Prevention and management of heat-related illness among spectators and staff during the Olympic Games—Atlanta, July 6-23, 1996. MMWR Morb Mortal Wkly Rep 1996 Jul 26;45(29):631-633. [Medline: 8965791]
118. Yu Q, Cheng Z. Analysis on the emergency medical services for the 2008 Beijing Olympic Games opening ceremony. Chin J Emerg Med 2009;18(4):355-360. [doi: 10.3760/cma.j.issn.1671-0282.2009.04.005]
119. Tanaka H, Tanaka S, Yokota H, et al. Acute in-competition medical care at the Tokyo 2020 Olympics: a retrospective analysis. Br J Sports Med 2023 Nov;57(21):1361-1370. [doi: 10.1136/bjsports-2022-105778] [Medline: 37055080]
120. Inoue H, Tanaka H, Sakanashi S, et al. Incidence and factor analysis for the heat-related illness on the Tokyo 2020 Olympic and Paralympic Games. BMJ Open Sport Exerc Med 2023;9(2):e001467. [doi: 10.1136/bmjsem-2022-001467] [Medline: 37051574]
121. He W, Xiang DY, Dai JP. Sonography in the 29th Olympic and Paralympic Games: a retrospective analysis. Clin Imaging 2011;35(2):143-147. [doi: 10.1016/j.clinimag.2009.12.023] [Medline: 21377054]
122. Wu F, Liu Y, Zhuang M. Lessons from the Winter Paralympic Games disclosing the epidemiology of winter sports injury in paralytic athletes: a meta-analysis. BMC Sports Sci Med Rehabil 2022 Mar 29;14(1):53. [doi: 10.1186/s13102-022-00446-w] [Medline: 35351197]
123. Blauwet CA, Cushman D, Emery C, et al. Risk of injuries in Paralympic track and field differs by impairment and event discipline: a prospective cohort study at the London 2012 Paralympic Games. Am J Sports Med 2016 Jun;44(6):1455-1462. [doi: 10.1177/0363546516629949] [Medline: 26920432]
124. Haid BC. Medical Service During the Winter Olympic Games...1964... Innsbruck, Austria. Anesth Analg 1965;44:240-244. [Medline: 14284268]

XSL•FO
RenderX

125. Blauwet CA, Lexell J, Derman J. Paralympic sports medicine. In: Training and Coaching the Paralympic Athlete: Handbook of Sports Medicine and Science: Wiley; 2016:75-95.
126. Derman W, Runciman P, Schwellnus M, et al. High precompetition injury rate dominates the injury profile at the Rio 2016 Summer Paralympic Games: a prospective cohort study of 51 198 athlete days. Br J Sports Med 2018 Jan;52(1):24-31. [doi: 10.1136/bjsports-2017-098039] [Medline: 29030389]
127. BioGDP. URL: https://biogdp.com/ [accessed 2025-08-14]
128. Wu Y, Dai R, Yan W, Ren S, Ao Y. Characteristics of sports injuries in athletes during the Winter Olympics: a systematic review and meta-analysis. Orthop J Sports Med 2023 Dec;11(12):23259671231209286. [doi: 10.1177/23259671231209286] [Medline: 38107844]
129. Kennedy MI, Moatshe G, Engebretsen L. Global forum: orthopaedic physicians in the Winter and Summer Olympic Games. J Bone Joint Surg Am 2020;102(11):e52. [doi: 10.2106/JBJS.19.01245]
130. Jihang C, Qing B, Haiqing X, et al. Analysis of the characteristics of sports injury in previous Summer Olympic Games in recent 15 years. Infect Dis Now 2024 Jun;54(4). [doi: 10.1016/j.idnow.2024.104889]
131. Zang W, Fang M, Zhang X, Xiao N, Wang S, Mu L. Exploring the epidemiology of injuries in athletes of the Olympic Winter Games: a systematic review and meta-analysis. J Sports Sci Med 2023 Dec;22(4):748-759. [doi: 10.52082/jssm.2023.748] [Medline: 38045742]
132. Willick SE, Webborn N, Emery C, et al. The epidemiology of injuries at the London 2012 Paralympic Games. Br J Sports Med 2013 May;47(7):426-432. [doi: 10.1136/bjsports-2013-092374]
133. Oh H, Johnson W, Syrop IP. Winter adaptive sports participation, injuries, and equipment. Sports Med Arthrosc Rev 2019 Jun;27(2):56-59. [doi: 10.1097/JSA.0000000000000236] [Medline: 31046009]
134. Rahman GR, Jasani GN, Liang SY. Terrorist attacks against sports venues: emerging trends and characteristics spanning 50 years. Prehosp Disaster Med 2023 Jun;38(3):366-370. [doi: 10.1017/S1049023X23000377] [Medline: 36938664]
135. Yan X, Fang Y, Li Y, Jia Z, Zhang B. Risks, epidemics, and prevention measures of infectious diseases in major sports events: scoping review. JMIR Public Health Surveill 2022 Dec 2;8(12):e40042. [doi: 10.2196/40042] [Medline: 36459401]
136. Weiss BP, Mascola L, Fannin SL. Public health at the 1984 Summer Olympics: the Los Angeles County experience. Am J Public Health 1988 Jun;78(6):686-688. [doi: 10.2105/ajph.78.6.686] [Medline: 3285705]
137. Thompson JM, Savoia G, Powell G, Challis E, Law P. Level of medical care required for mass gatherings: the XV Winter Olympic Games in Calgary, Canada. Ann Emerg Med 1991 Apr;20(4):385-390. [doi: 10.1016/S0196-0644(05)81660-9]
138. Thackway SV, Delpech VC, Jorm LR, McAnulty JM, Visotina M. Monitoring acute diseases during the Sydney 2000 Olympic and Paralympic Games. Med J Aust 2000 Sep 18;173(6):318-321. [doi: 10.5694/j.1326-5377.2000.tb125666.x] [Medline: 11061404]
139. Jorm LR, Thackway SV, Churches TR, Hills MW. Watching the Games: public health surveillance for the Sydney 2000 Olympic Games. J Epidemiol Community Health 2003 Feb;57(2):102-108. [doi: 10.1136/jech.57.2.102] [Medline: 12540684]
140. Gundlapalli AV, Olson J, Smith SP, et al. Hospital electronic medical record–based public health surveillance system deployed during the 2002 Winter Olympic Games. Am J Infect Control 2007 Apr;35(3):163-171. [doi: 10.1016/j.ajic.2006.08.003]
141. Webborn N, Willick S, Reeser JC. Injuries among disabled athletes during the 2002 Winter Paralympic Games. Med Sci Sports Exerc 2006 May;38(5):811-815. [doi: 10.1249/01.mss.0000218120.05244.da] [Medline: 16672831]
142. Junge A, Langevoort G, Pipe A, et al. Injuries in team sport tournaments during the 2004 Olympic Games. Am J Sports Med 2006 Apr;34(4):565-576. [doi: 10.1177/0363546505281807] [Medline: 16303876]
143. Lystad RP, Alevras A, Rudy I, Soligard T, Engebretsen L. Injury incidence, severity and profile in Olympic combat sports: a comparative analysis of 7712 athlete exposures from three consecutive Olympic Games. Br J Sports Med 2021 Oct;55(19):1077-1083. [doi: 10.1136/bjsports-2020-102958] [Medline: 33219113]
144. Derman W, Schwellnus MP, Jordaan E, et al. High incidence of injury at the Sochi 2014 Winter Paralympic Games: a prospective cohort study of 6564 athlete days. Br J Sports Med 2016 Sep;50(17):1069-1074. [doi: 10.1136/bjsports-2016-096214] [Medline: 27307272]
145. Ona Ayala KE, Li X, Huang P, et al. Injury epidemiology and preparedness in powerlifting at the Rio 2016 Paralympic Games: an analysis of 1410 athlete‐days. Transl Sports Med 2019 Nov;2(6):358-369. [doi: 10.1002/tsm2.107]
146. Sugishita Y, Somura Y, Abe N, et al. Multisource surveillance conducted by the Tokyo Metropolitan Government during the Tokyo 2020 Olympic and Paralympic Games. Western Pac Surveill Response J 2023;14(3):1-10. [doi: 10.5365/wpsar.2023.14.3.978] [Medline: 37955030]
147. Yanagawa M, Lorenzo JC, Fukusumi M, et al. Use of Epidemic Intelligence from Open Sources for global event-based surveillance of infectious diseases for the Tokyo 2020 Olympic and Paralympic Games. Western Pac Surveill Response J 2022;13(3):1-7. [doi: 10.5365/wpsar.2022.13.3.959] [Medline: 36688182]
148. Adams WM, Anderson T, Finnoff JT. Team USA COVID-19 Prevalence at Tokyo 2020 Summer Olympic and Paralympic Games. Curr Sports Med Rep 2023 Feb 1;22(2):49-51. [doi: 10.1249/JSR.0000000000001035] [Medline: 36757122]
149. Daly P, Gustafson R. Public health recommendations for athletes attending sporting events. Clin J Sport Med 2011;21(1):67-70. [doi: 10.1097/JSM.0b013e31820758dc]

XSL•FO
RenderX

150. Bougault V, Valorso R, Sarda-Esteve R, et al. Paris air quality monitoring for the 2024 Olympics and Paralympics: focus on air pollutants and pollen. Br J Sports Med 2024 Sep 4;58(17):973-982. [doi: 10.1136/bjsports-2024-108129] [Medline: 39054048]
151. Schwellnus MP, Derman WE, Jordaan E, et al. Elite athletes travelling to international destinations >5 time zone differences from their home country have a 2-3-fold increased risk of illness. Br J Sports Med 2012 Sep;46(11):816-821. [doi: 10.1136/bjsports-2012-091395] [Medline: 22875910]
152. Mougin L, Bougault V, Racinais S, et al. Environmental challenges facing athletes, stakeholders and spectators at Paris 2024 Olympic and Paralympic Games: an evidence-based review of mitigation strategies and recommendations. Br J Sports Med 2024 Jul 25;58(15):870-881. [doi: 10.1136/bjsports-2024-108281] [Medline: 38955507]
153. de Oliveira EP, Burini RC, Jeukendrup A. Gastrointestinal complaints during exercise: prevalence, etiology, and nutritional recommendations. Sports Med 2014 May;44(Suppl 1):S79-S85. [doi: 10.1007/s40279-014-0153-2] [Medline: 24791919]
154. Adams BB. Dermatologic disorders of the athlete. Sports Med 2002;32(5):309-321. [doi: 10.2165/00007256-200232050-00003]
155. Vanos JK, Kosaka E, Iida A, et al. Planning for spectator thermal comfort and health in the face of extreme heat: the Tokyo 2020 Olympic marathons. Sci Total Environ 2019 Mar 20;657:904-917. [doi: 10.1016/j.scitotenv.2018.11.447] [Medline: 30677956]
156. McCloskey B, Memish ZA, Sharma A, Traore T, Parker S, Zumla A. Confronting heat-related illnesses and deaths at mass gathering religious and sporting events. Lancet Planet Health 2024 Aug;8(8):e522-e523. [doi: 10.1016/S2542-5196(24)00161-X] [Medline: 39009004]
157. Post EG, Anderson T, Shilt JS, et al. Incidence of injury and illness among paediatric Team USA athletes competing in the 2020 Tokyo and 2022 Beijing Olympic and Paralympic Games. BMJ Open Sport Exerc Med 2023;9(4):e001730. [doi: 10.1136/bmjsem-2023-001730] [Medline: 38143720]
158. Smith CM, Moore F, Drezner JA, et al. Resuscitation on the field of play: a best-practice guideline from Resuscitation Council UK. Br J Sports Med 2024 Oct 17;58(19):1098-1106. [doi: 10.1136/bjsports-2024-108440] [Medline: 39122372]
159. Zumla A, Peiris M, Memish ZA, Perlman S. Anticipating a MERS-like coronavirus as a potential pandemic threat. Lancet 2024 May;403(10438):1729-1731. [doi: 10.1016/S0140-6736(24)00641-X]
160. Engebretsen L, Steffen K, Soligard T. Injury risk in the Olympic Games: the injury and illness surveillance during the XXIX 2008 Summer and the XXI 2010 Winter Olympic Games. In: Nuclear Medicine and Radiologic Imaging in Sports Injuries: Springer; 2015:1107-1121. [doi: 10.1007/978-3-662-46491-5_51]
161. Derman W, Runciman P, Jordaan E, et al. High incidence of injuries at the Pyeongchang 2018 Paralympic Winter Games: a prospective cohort study of 6804 athlete days. Br J Sports Med 2020 Jan;54(1):38-43. [doi: 10.1136/bjsports-2018-100170]
162. Nakamura S, Wada K, Yanagisawa N, Smith DR. Health risks and precautions for visitors to the Tokyo 2020 Olympic and Paralympic Games. Travel Med Infect Dis 2018 Mar;22:3-7. [doi: 10.1016/j.tmaid.2018.01.005]
163. Kobayashi K, Kimura A, Sasaki R, et al. Actual situation of handling Tokyo 2020 Games-related patients at a designated hospital during COVID-19 pandemic. Glob Health Med 2022 Aug 31;4(4):230-232. [doi: 10.35772/ghm.2022.01009] [Medline: 36119786]
164. Mountjoy M, Junge A, Bindra A, et al. Surveillance of athlete mental health symptoms and disorders: a supplement to the International Olympic Committee's consensus statement on injury and illness surveillance. Br J Sports Med 2023 Nov;57(21):1351-1360. [doi: 10.1136/bjsports-2022-106687] [Medline: 37468210]
165. Moore IS, Crossley KM, Bo K, et al. Female athlete health domains: a supplement to the International Olympic Committee consensus statement on methods for recording and reporting epidemiological data on injury and illness in sport. Br J Sports Med 2023 Sep;57(18):1164-1174. [doi: 10.1136/bjsports-2022-106620] [Medline: 37349084]
166. Gallien Y, Fournet N, Delamare H, Haroutunian L, Tarantola A. Epidemiological surveillance and infectious disease outbreaks during mass international summertime sports gatherings: a narrative review. Infect Dis Now 2024 Jun:104889. [doi: 10.1016/j.idnow.2024.104889]

Abbreviations

IOC: International Olympic Committee
NOC: National Olympic Committee
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 24.09.24; peer-reviewed by K Fukumoto, T Endericks; revised version received 26.05.25; accepted 29.08.25; published 03.09.25.

Please cite as:
Ren S, Li T, Zhang Y, Bai S, Zhou Z, Li S
Preparing for Potential Health and Safety Risks at the Olympic Games: Scoping Review
JMIR Public Health Surveill 2025;11:e66829
URL: https://publichealth.jmir.org/2025/1/e66829
doi:10.2196/66829

© Shaotong Ren, Tiantian Li, Yongzhong Zhang, Song Bai, Zichen Zhou, Shengxin Li. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 3.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association of Indoor Temperature Level and Mental Health Among Community-Dwelling Older Adults: Systematic Review

Min Kyung Park[1], PhD; Bada Kang[1], PhD; Dahye Hong[1,2], MPH; Seolah Yoon[3], MSN; Minsub Kim[1], BA; Yubeen Kim[1], BSN; Zhao Ni[4], PhD; Eleanor S McConnell[5,6], PhD

[1]Mo-Im Kim Nursing Research Institute, Yonsei University College of Nursing, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

[2]College of Nursing and Brain Korea 21 FOUR Project, Yonsei University, Seoul, Republic of Korea

[3]Department of Nursing, Yonsei University Graduate School, Seoul, Republic of Korea

[4]School of Nursing, Yale University, Orange, CT, United States

[5]School of Nursing, Duke University, Durham, NC, United States

[6]Geriatric Research, Education, and Clinical Center (GRECC), Durham VA Health Care System, Durham, NC, United States

Corresponding Author:
Bada Kang, PhD
Mo-Im Kim Nursing Research Institute, Yonsei University College of Nursing, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: The association between indoor temperature level and mental health is becoming increasingly important as climate change leads to extreme temperature fluctuations. Older adults are particularly vulnerable to indoor temperature changes because of their diminished ability to regulate body temperature and extended time spent indoors.

Objective: This study aims to examine the association between indoor temperature levels and mental health outcomes among community-dwelling older adults, aiming to provide essential evidence to support the development of interventions and policy strategies to improve their mental health.

Methods: In this systematic review, we conducted a comprehensive search of 7 electronic databases (MEDLINE, Embase, Cochrane Library, CINAHL, PsycINFO, Google Scholar, and ProQuest) on April 4, 2024, without restrictions on language or publication date. The National Institutes of Health quality assessment tool for observational cohorts and cross-sectional studies was used to evaluate the methodological quality of the included literature.

Results: Of the 2328 studies identified, 15 met the inclusion criteria. The majority (8/15, 53%) were conducted in Asia, followed by Europe (4/15, 27%) and 1 study each in Australia (7%), Egypt (7%), and the United States (7%). Mental health outcomes associated with indoor temperature exposure were categorized into four groups: (1) sleep problems, including insomnia; (2) emotional problems, such as emotional distress and negative mood; (3) social interaction problems, such as social exclusion and low social participation; and (4) other mental health issues, including anxiety, agitation, and annoyance. Sleep problems were the most frequently reported mental health outcome related to indoor temperature levels (9/15, 60%). Older adults living in substandard housing conditions, facing economic difficulties, and residing in urban areas were vulnerable to exposure to uncomfortable indoor temperatures because of housing-related risks, such as low energy efficiency, inadequate heating or cooling, and limited access to green spaces.

Conclusions: The findings highlight the need to develop evidence-based guidelines to improve mental health by managing indoor temperature levels in the community. Improving housing conditions through policy support could enhance the mental health of community-dwelling older adults.

Trial Registration: PROSPERO CRD42024536215; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024536215

(JMIR Public Health Surveill 2025;11:e78257) doi:10.2196/78257

KEYWORDS

emotional; housing; climate change; sleep-wake disorder; social participation

Introduction

Global temperature is rapidly increasing as a result of climate change [1]. The Lancet Countdown, an international collaboration monitoring the health impacts of climate change, warns that the unequal health impacts of climate change will worsen without immediate intervention [2,3]. As climate change leads to more frequent and intense extreme temperatures [1], interest is growing in studies on health outcomes associated

XSL•FO
RenderX

with temperature level [4]. To better adapt to environmental changes caused by climate change, communities need to understand how to protect themselves from exposure to uncomfortable temperatures [5].

Temperature is closely linked to mental health, with growing evidence demonstrating that exposure to extreme temperatures is associated with increased hospital admissions and emergency department visits related to adverse mental health outcomes such as affective disorders, insomnia, anxiety, depressive disorders, schizophrenia, and organic mental disorders [6-8]. Indoor temperature may be evaluated by both objective measurements and subjective perceptions of indoor comfort [9], which is crucial given that recent studies have found that exposures measured by a single approach are significantly associated with negative mental health outcomes.

Determining an appropriate indoor temperature is complex as it depends not only on environmental and structural factors, such as regional climate and housing conditions [10,11], but also on individual response to temperature levels. Aging further complicates this issue, as older adults experience a diminished ability to perceive and regulate body temperature, increasing their vulnerability to the adverse effects of exposure to both high and low temperatures [12]. In addition, older adults are more susceptible to inappropriate indoor temperature exposure, because they tend to spend more time indoors than do younger populations [9,13]. A recent study found that community-dwelling older adults spend up to 85% of their time indoors, whether at home or in other settings [14]. Accordingly, a growing number of studies have examined the association between indoor temperature levels and mental health in older adults [15-17]. However, without synthesizing this fragmented evidence, it remains difficult to determine which mental health outcomes are most frequently reported in association with indoor temperature or to identify which subgroups of older adults are most vulnerable. A systematic review can integrate these findings to enhance our understanding of the associations between indoor temperature, individual vulnerabilities, and mental health outcomes, essential for informing targeted interventions and policy development in the context of aging populations and climate change.

In this systematic review, we aim to investigate the association between indoor temperature levels and mental health—encompassing both clinical disorders and nonclinical outcomes—in community-dwelling older adults, addressing the following questions:

1. What characteristics of indoor temperature exposure are related to mental health outcomes in older adults?
2. Which mental health outcomes are related to indoor temperature levels in older adults?
3. What factors make older adults vulnerable to indoor temperature levels?

Methods

Design

This study is a systematic review. The reporting of this review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Checklist 1) [18]. The protocol was registered in the International Prospective Register of Systematic Reviews (protocol: CRD42024536215). Owing to the use of previously published studies and publicly available anonymized data, ethical approval was not required.

Search Strategy

A systematic search was conducted on 7 electronic databases with the assistance of experienced medical librarians: MEDLINE, Embase, Cochrane Library, CINAHL, PsycINFO, Google Scholar, and ProQuest. In addition, the bibliographies of the screened studies were manually searched. Three categories of concepts were used individually and in combination: (1) population (older adults), (2) exposure (indoor temperature), and (3) outcome (mental health). The database literature search was conducted on April 4, 2024, with no restrictions on publication dates or language. Non-English studies were screened using machine translation tools and cross-checked among the reviewers. The detailed search strategy for each database is provided in Table S1 in Multimedia Appendix 1.

Eligibility Criteria

The inclusion and exclusion criteria are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for study selection.

Inclusion criteria

- Population: older adults aged ≥60 years.
- Exposure: indoor temperature measurement
- Outcome: mental health outcomes related to indoor temperature [19,20], including *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5)-defined mental illnesses (eg, insomnia, mood disorders, and delirium) and nonclinical states (eg, annoyance, social isolation, and loneliness).
- Setting: community settings including homes and nonhospital residential environments (eg, geriatric and nursing homes).
- Study design: quantitative research, including cross-sectional and longitudinal studies.
- Language: no restrictions

Exclusion criteria

- Population: Studies not reporting separate results for older adults.
- Exposure: studies combining indoor and outdoor temperature without separate indoor effects.
- Outcome: studies without mental health–related outcomes.
- Setting: Hospital settings or nonresidential environments.
- Study design: protocol papers without data and studies lacking methodological reporting (eg, editorials and commentaries).
- Language: not applicable.

Study Selection

The results of the literature search were exported from the databases into EndNote 20 reference management software (Clarivate Analytics) for organization and duplicate removal. The titles and abstracts of the retrieved articles were independently screened by 5 authors (MKP, DH, SY, YK, and MK), with all studies undergoing independent double review. For potentially relevant articles, 3 authors (MKP, DH, and SY) conducted full-text assessments independently in a double-review format. At each stage of the review process, any disagreements or discrepancies were resolved through discussion among the 3 authors until consensus was achieved.

Data Extraction and Synthesis

Relevant data from eligible studies were extracted into a Microsoft Excel spreadsheet. The same 5 authors extracted data from the selected studies for analysis. One reviewer (MKP) extracted information from all 15 articles, and the remaining 4 reviewers (DH, SY, YK, and MK) independently extracted information from 3 to 4 articles per person. Discrepancies were reviewed by five authors and resolved in discussion with the sixth author (BK) after re-examining the data to ensure accuracy. The extracted data included the characteristics of the studies (country, design, setting, sample size, participant characteristics, vulnerability factors, and quality assessment), exposure (target season, temperature, and measure), and outcome (mental health outcome and measurement).

Predefined research questions guided the synthesis. We conducted a structured narrative synthesis organized around: (1) characteristics of indoor temperature exposure related to mental health outcomes in older adults, (2) mental health outcomes in older adults related to indoor temperature levels, and (3) factors associated with indoor temperature exposure at varying levels in older adults. Given the heterogeneity in study designs, exposure data formats, and mental health outcome measures, pooling of effect sizes was not possible; therefore, a meta-analysis was not conducted. Instead, we summarized the direction and consistency of associations for each study-outcome pair. To enhance clarity and comparability for the heterogeneous studies, we organized the findings by outcome domain (sleep problems, emotional problems, social interaction problems, and other mental health issues) and further by exposure type (hot vs cold indoor temperature). When studies reported both objective and subjective measurements, results were presented separately, and any discrepancies were noted to highlight methodological variability.

Quality Assessment in Individual Studies

The National Institutes of Health quality assessment tool for observational cohort and cross-sectional studies [21] was used for quality assessment. This tool evaluates the internal validity of a study by asking 14 questions. The quality of the studies was categorized based on the number of criteria they met: good (>9), fair (5–9), and poor (<5) [22]. Studies were independently scored by three authors (MKP, DH, and SY) using Microsoft Excel. Any discrepancies regarding quality appraisal were resolved in discussion with the fourth author (BK).

Results

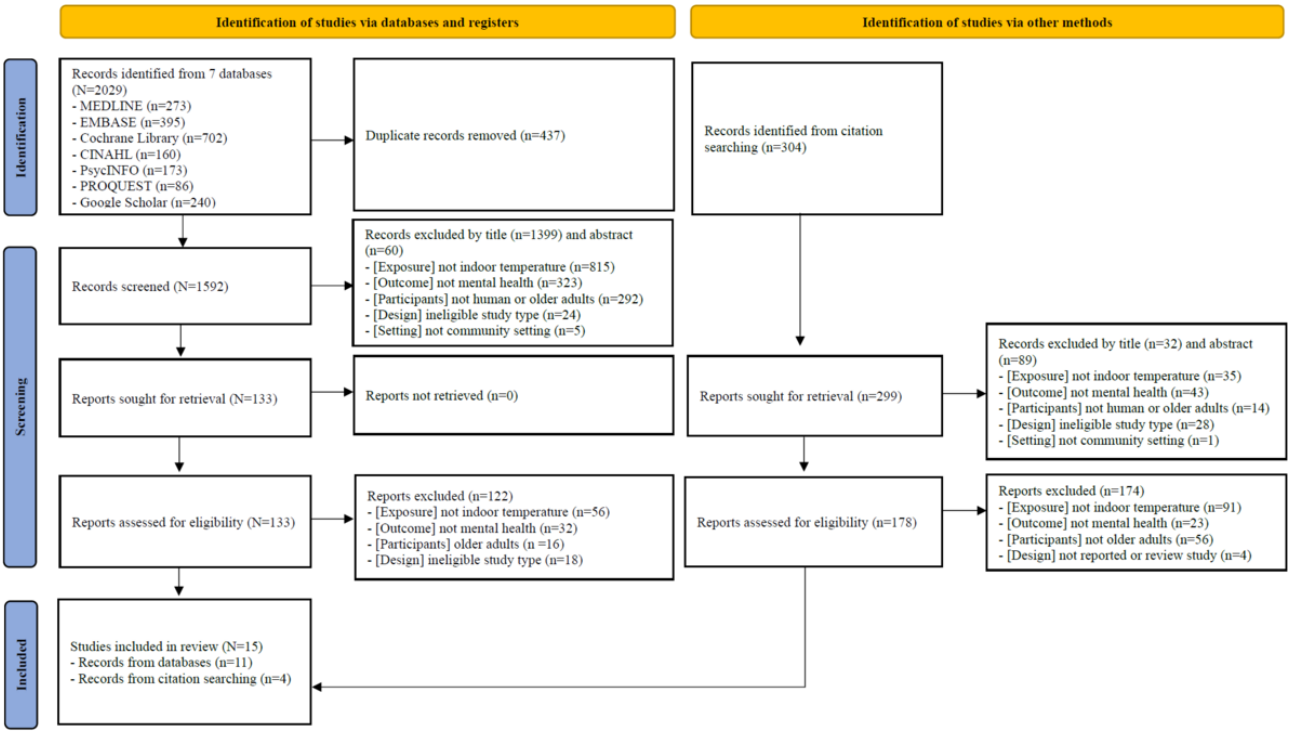
Study Selection

Figure 1 presents the screening and selection process in accordance with PRISMA 2020 guidelines, detailing the number of records included and excluded at each stage, as well as the specific reasons for exclusion during the eligibility assessment. A total of 2328 records were identified through electronic database searches (n=2029) and citation searching (n=299). After removing 437 duplicates, 1891 articles remained. A total of 1580 articles were excluded after title (n=1431) and abstract

XSL•FO
RenderX

(n=149) screening for the following reasons—ineligible exposure (not indoor temperature), ineligible outcome (not mental health), ineligible population (not human or not older adults), ineligible study design (eg, editorial or review), and noncommunity setting (eg, hospital)— leaving 311 articles for full-text assessment. Following a detailed eligibility assessment, 296 articles were excluded owing to ineligible exposure, outcome, population, or study design. Ultimately, 15 studies met the inclusion criteria and were included for the final synthesis.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.



Characteristics of Included Studies

Table 1 presents the characteristics of the included studies. Of the 15 included studies, all except 1 (7%) study [23] were published after 2010, with 10 published within the last 10 years. The majority of the studies (8/15, 53%) were conducted in Asia; the remainder were conducted in Europe (4/15, 27%) and Australia, Egypt, and the United States (1/15, 7% each). As shown in Figure 2, according to the World Bank classification [24], most studies (12/15, 80%) were conducted in high-income countries. From the 15 studies, 7 (47%) used a longitudinal design, whereas the remaining 8 (53%) used a cross-sectional design. Eleven (73%) studies focused on older adults residing at home, and the remaining 4 targeted older adults residing in nursing homes (2/15, 13%), geriatric homes (1/15, 7%), and sheltered facilities (1/15, 7%). The sample sizes ranged from 19 to 29,380. Three (20%) studies included participants with a mean age in the 60s, and 7 (47%) and 2 (13%) studies involved participants with a mean age in the 70s and 80s, respectively. The remaining 3 (20%) studies did not report mean age. Regarding participant characteristics, 3 (20%) studies examined healthy older adults [17,23,25] where 2 (13%) focused on older adults with dementia [26,27], and 1 (7%) examined low-income older adults [28].

XSL•FO
RenderX

Table . Characteristics of the included studies (N=15).

Authors (year)	Country (urban or rural)	Study design	Setting	Sample size	Age (y), mean (SD; range)[c]	Quality assessment
Ohnaka (1995) [23]	Japan (urban)[a]	Cross-sectional observational study	Home	20	73.0 (67–82)	Good quality
Okamoto-Mizuno and Tsuzuki (2010) [25]	Japan (urban)[a]	Longitudinal observational study	Home	19	65.8 (2.6; ≥62)	Good quality
Bakr et al (2012) [6]	Egypt (urban)[b]	Cross-sectional observational study	Geriatric home	184	68.02 (6.71; 60–97)	Poor quality
Cotter et al (2012) [29]	Ireland (mixed)[b]	Cross-sectional observational study	Home	722	72.5	Poor quality
Garre‐Olmo et al (2012) [26]	Spain (urban)[a]	Cross-sectional observational study	Nursing home	160	82.6 (11.6)	Fair quality
Saeki et al (2015) [30]	Japan (rural)[a]	Longitudinal observational study	Home	861	72.1 (7.1; ≥60)	Good quality
Ahrentzen et al (2016) [28]	USA (urban)[b]	Longitudinal observational study	Home	57	73 (62‐92)	Fair quality
Van Loenhout et al (2016) [31]	Netherlands (urban)[b]	Longitudinal observational study	Home	113	73.8 (7.5; ≥65)	Fair quality
Tartarini et al (2017) [27]	Australia (urban)[a]	Longitudinal observational study	Nursing home	21	(61–92)	Good quality
Kim et al (2020) [32]	Korea (rural)[b]	Longitudinal observational study	Home	104	79.6 (65–96)	Fair quality
Lindemann et al (2018) [33]	Germany (urban)[b]	Longitudinal observational study	Facilities of sheltered	81	80.9 (6.53; 63–93)	Good quality
Lee et al (2020) [34]	Korea (mixed)[a]	Cross-sectional observational study	Home	248	N/A[d] (≥60)	Fair quality
Kanno et al (2022) [16]	Japan (mixed)[a]	Cross-sectional observational study	Home	29,380	61.7 (11.5)	Fair quality
Yan et al (2022) [17]	China (urban)[a]	Cross-sectional observational study	Home	40	70.7 (5.0)	Good quality
Liu et al (2023) [15]	China (mixed)[b]	Cross-sectional observational study	Home	356	N/A (≥60)	Fair quality

[a]Participants' residence was categorized as urban, rural, or mixed based on regional information.

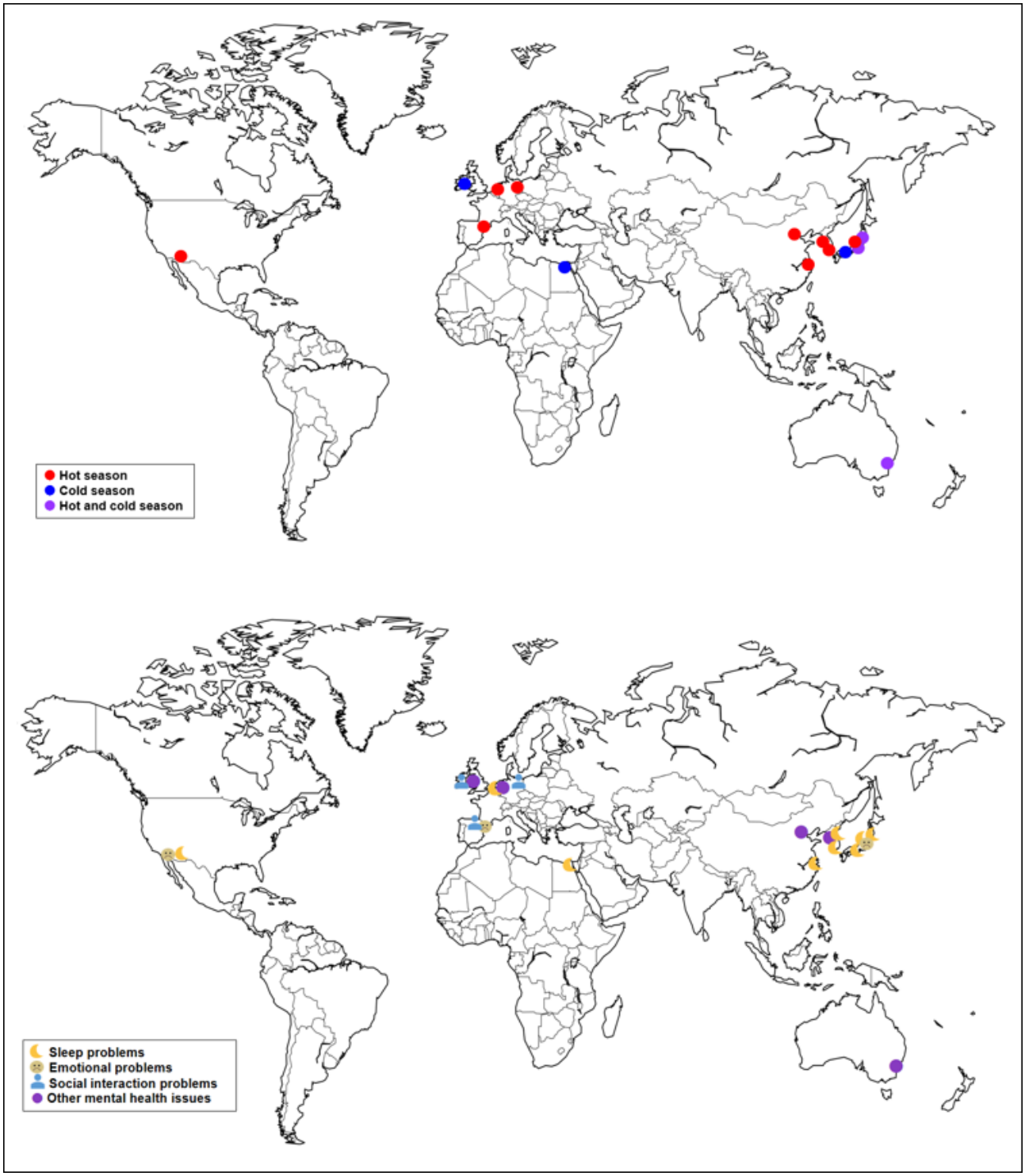
[b]In the literature, participants' characteristics are clearly classified as urban or rural, with over 90% sharing the same regional background.

[c]Age values are presented as reported; mean, SD, and/or range are provided only when available.

[d]N/A: not applicable.

XSL•FO
RenderX

Figure 2. Target season and mental health outcomes by region.



Quality Assessment in Individual Studies

The risk of bias assessment for the 15 studies is summarized in Table 1. The quality of the included studies varied. Six (40%) studies were rated as having "good quality." Seven (47%) studies were categorized as "fair quality" and 2 (13%) as "poor quality." Bias was primarily due to the unclear participation rate of eligible persons, a lack of justification for the sample size, and the inability to blind the exposure. Detailed results from the quality assessment of each study are available in Table S2 in Multimedia Appendix 2.

Characteristics of Indoor Temperature Exposure

The seasonal background of the studies was predominantly hot (9/15, %), with 3 (20%) targeting the cold season and 3 (20%) including both hot and cold seasons. The characteristics of indoor temperatures to which older adults were exposed are shown in Table 2. The results for indoor temperatures were presented as means in 10 studies, of which 3 (20%) also reported maximum and minimum temperatures. One (7%) study provided the set air conditioning temperature, while 5 (33%) studies did not specify the exact temperatures. The average indoor temperature during the hot season ranged from 22.7 °C to 30.5

XSL•FO
RenderX

°C, whereas the average temperature during the cold season ranged from 9.5 °C to 17.1 °C.

XSL•FO
RenderX

Table . Characteristics of indoor temperature (N=15).

Authors (year)	Season and time ▪ Target season (hot or cold) ▪ Month and year	Exposure measure ▪ Objective versus subjective ▪ Location ▪ Duration/interval ▪ Measurement tool	Exposure temperature ▪ Format of temperature reporting ▪ TMean[a] (SD[b]; TRange[c])	Vulnerability characteristics ▪ Vulnerable group ▪ Related indoor temperature
Ohnaka et al (1995) [23]	▪ Hot ▪ July–August (year N/A[d])	▪ Objective ▪ Home (bedroom) ▪ July to August or every 2 minutes throughout the night ▪ Data logger (XT-102, JNIS)	▪ TMean (SD; TRange) ▪ 26.26 °C (2.24 °C; 25 °C-28 °C)	▪ Younger age ▪ Higher room temperature
Okamoto-Mizuno and Tsuzuki (2010) [25]	▪ Cold and hot ▪ February, July–August, October–November (year N/A)	▪ Objective ▪ Home (bedroom) ▪ 5 days or every 1 minute. ▪ Thermistor and hygrometer probe (RS-12; Espec Mic Corporation)	▪ TMean (SD) ▪ Fall: 15.4 °C (0.25 °C) Winter: 9.5 °C (0.69 °C) Summer: 27.7 °C (0.63 °C)	• N/A
Bakr et al. (2012) [6]	▪ Cold ▪ October–December (2010)	▪ Subjective ▪ N/A ▪ N/A ▪ Self-reported questionnaire	▪ Response rate (%): feeling cold or hot ▪ N/A	• N/A
Cotter et al (2012) [29]	▪ Cold ▪ January–April (2011)	▪ Subjective ▪ Home ▪ N/A ▪ Self-reported questionnaire	▪ Response rate (%) ▪ N/A	▪ Poorer health condition (disability, arthritis, and fall in the previous 6 mo); poorer quality housing (mold, damp, and draughts in the home; no central heating system); difficulty paying for heating; social exclusion (feeling loneliness; low social activities or hobbies) ▪ Very cold homes
Garre‐Olmo et al (2012) [26]	▪ Hot ▪ April–July (2008)	▪ Objective ▪ Home (bedroom, dining room, and living room) ▪ 7 days or twice in the morning and twice in the afternoon ▪ DT-8820 environment meter	▪ TMean (SD), TMedian[e] ▪ 25.8 °C (1.3 °C)	• N/A
Saeki et al (2015) [30]	▪ Cold ▪ October–April (2010‐2013)	▪ Objective ▪ Home (living room and bedroom) ▪ 2 days or every 10 minutes ▪ Thermochron iButtons (DS1922L; Maxim Integrated)	▪ TMean ▪ Evening 17.1 °C (4.1 °C)	• N/A
Ahrentzen et al (2016) [28]	▪ Hot ▪ June–August (2010‐2012)	▪ Objective ▪ Home (kitchen, bedroom, and living area) ▪ 5 days or every 15 minutes ▪ Mobile Onset HOBO data loggers	▪ TMean (SD), TMedian, TMax[f], TMin[g], TRange ▪ Panel 1: 26.0 °C (1.31 °C; 24.2 °C‐28.0 °C) ▪ Panel 3: 25.4 °C (1.27 °C; 22.9 °C‐28.2 °C)	▪ Affordable housing (low energy efficiency housing) ▪ Higher indoor temperature (mean minimum and exceed 27 °C)

XSL•FO
RenderX

Authors (year)	Season and time ▪ Target season (hot or cold) ▪ Month and year	Exposure measure ▪ Objective versus subjective ▪ Location ▪ Duration/interval ▪ Measurement tool	Exposure temperature ▪ Format of temperature reporting ▪ TMean[a] (SD[b]; TRange[c])	Vulnerability characteristics ▪ Vulnerable group ▪ Related indoor temperature
Van Loenhout et al (2016) [31]	▪ Hot ▪ August (2012)	▪ Objective ▪ Living room and bedroom ▪ 21 days with high temperature and 1 cold reference week or every 30 minutes ▪ iButton Hygrochron temperature or humidity loggers (type DS1923) ▪ Subjective ▪ Living room and bedroom ▪ N/A ▪ Self-reported questionnaire	▪ TMean, TMax, TMin, TRange ▪ Living room: 25.4 °C (22.3 °C - 30.2 °C) Bedroom: 25.1 °C (20.8 °C - 29.3 °C)	• N/A
Tartarini et al (2017) [27]	▪ Cold and hot ▪ March–December (2015)	▪ Objective ▪ Nursing home (each room) ▪ 14 days or every 15 minutes ▪ iButton temperature loggers (Maxim Integrated)	▪ TMean, TRange ▪ 23.5 °C (16.2 °C - 33.6 °C)	• N/A
Kim et al (2020) [32]	▪ Hot ▪ August (2018)	▪ Objective ▪ Home ▪ 3 days or twice a day (morning and afternoon) ▪ Electronic hygrothermographs (AE-817CE, B&J)	▪ TMean, TMax, TMin, TRange, with or without air conditioners ▪ 30.5 °C (22.9 °C - 38.3 °C)	• N/A
Lindemann et al (2018) [33]	▪ Hot ▪ May–October (2015) (including heatwave period)	▪ Objective ▪ Facility (table in the shade in the room) ▪ From May to October and 2 heat waves in July and August or every 4 weeks ▪ Data logger (HL-1D, ROTRONIC Messgere GmbH)	▪ TMean, increased temperature ▪ 22.7 °C	▪ Lower gait speed; living in city or city-center (lower access to green spaces compared to living in garden cities or suburban) ▪ Stronger mean changes of participation per 10 °C increase of indoor temperature
Lee et al (2020) [34]	▪ Hot ▪ Heatwave, 2018 (survey conducted in September 2018)	▪ Objective ▪ Home and primary living spaces other than home (eg, workplace and school) ▪ N/A ▪ N/A	▪ Use air conditioner set ▪ Lower than 24 ℃	• N/A
Kanno et al (2022) [16]	▪ Cold and hot ▪ May (2013)–November (2018)	▪ Subjective ▪ Home (living room, bedroom, sanitary space, toilet, and corridor or stairs) ▪ N/A ▪ Self-reported questionnaire (CASBEE[h] Housing Health Checklist related to coldness or warmth)	▪ Evaluation scores ▪ N/A	▪ Younger age; with psychological distress; sleeping difficulties ▪ Poorer housing condition related to coldness or warmth

XSL•FO
RenderX

Authors (year)	Season and time ▪ Target season (hot or cold) ▪ Month and year	Exposure measure ▪ Objective versus subjective ▪ Location ▪ Duration/interval ▪ Measurement tool	Exposure temperature ▪ Format of temperature reporting ▪ TMean[a] (SD[b]; TRange[c])	Vulnerability characteristics ▪ Vulnerable group ▪ Related indoor temperature
Yan et al (2022) [17]	▪ Hot ▪ July–August (2018)	▪ Objective ▪ Bedroom ▪ 6 days or every 5 minutes ▪ Data logger (TR-76Ui logger)	▪ TMean, TRange ▪ 28.8 °C (26 °C - 32 °C)	• N/A
Liu et al (2023) [15]	▪ Hot ▪ October (2021)–March (2022)	▪ Subjective ▪ Home ▪ N/A ▪ Self-reported questionnaire (8 items about temperature and humidity)	▪ Evaluation score ▪ N/A	• N/A

[a]TMean: mean temperature.

[b]SD

[c]TRange: temperature range.

[d]N/A: not applicable.

[e]TMedian: median temperature.

[f]TMax: maximum temperature.

[g]TMin: minimum temperature.

[h]CASBEE: comprehensive assessment system for building environment efficiency.

The methods used to measure exposure to indoor temperatures varied significantly across the studies. From the 15 studies, 10 (67%) objectively assessed indoor temperature exposure using devices such as data loggers, while 4 (27%) studies used only self-report questionnaires for subjective evaluation. Among the studies conducted during the cold season, only 1 (7%) [30] objectively assessed indoor temperature exposure. The location of indoor temperature measurements also differed, with bedrooms being the most frequently measured area (8/15, 53%), followed by living rooms (5/15, 33%). Regarding the duration and frequency of temperature measurements, 10 (67%) studies provided this information. The duration of exposure measurements across the studies varied from a few days to several months. Continuous measurements at intervals of 1 to 30 minutes were conducted in 7 (47%) studies, twice a day in 2 (13%) studies, and every 4 weeks in 1 (7%) study.

Mental Health Outcomes Related to Indoor Temperature Exposure

Table 3 shows that mental health outcomes related to indoor temperature exposure were disproportionately represented by sleep-related outcomes, with 9 of the 15 (60%) included studies focusing on this domain. Sleep problems, including short sleep duration, sleep disturbances, sleep disorders, and insomnia, were negatively associated with both hot and cold indoor temperature exposure. This predominance of sleep-related findings underscores existing gaps in the literature, as relatively few studies have examined other mental health outcomes such as mood, cognition, or social-emotional well-being in relation to indoor temperature. Of the 15 total studies, only 4 (27%) used objective methods to focus exclusively on sleep problems [17,23,25,30], whereas the remaining 11 (73%) studies measured outcomes using subjective methods. In the study by Okamoto-Mizuno and Tsuzuki [25], the association between hot indoor temperature exposure and objective sleep disturbance was statistically significant—wake episode (r=0.46; P<.05), sleep efficiency index (r=−0.47; P<.05), and wake after sleep onset (r=0.49; P<.01)—whereas the association with subjective sleep disturbance was not (P<.05). This highlights the differences between subjective and objective outcomes. The findings regarding the association between hot or cold indoor temperature exposure and emotional problems were inconsistent. In the studies by Ahrentzen et al [28] and Kanno et al [16], significant differences in emotional distress (t=−2.085; P<.05) and psychological distress (P<.001) scores were observed based on indoor temperature exposure. By contrast, Garre-Olmo et al [26] did not find indoor temperature exposure to be a significant influence on signs of negative affective mood (β=−.10 - .20; P>.05).

XSL•FO
RenderX

Table . Mental health outcomes related to indoor temperature (N=15).

Authors (year)	Mental health outcomes				Outcome measure measurement tool
	Sleep problems	Emotional problems	Social interaction problems	Other mental health issues	
Ohnaka (1995) [23]	• Body movement during sleep (r=0.299; P<.01)	• N/A[a]	• N/A	• N/A	• Objective: using a static charge sensitive bed

Okamoto-Mizuno and Tsuzuki (2010) [25]	• Objective: sleep disturbance[b]
	  • wake episode (r=0.46; P<.05)
	  • sleep efficiency index (r=-0.47; P<.05)
	  • wake after sleep onset (r=0.49; P<.01)
	• Subjective
	  • sleep disturbance (r=not described; P>.05)
		• N/A	• N/A	• N/A	• Objective: using wrist actigraphy
					• Subjective: self-reported questionnaire

Bakr et al (2012) [6]	• Insomnia (P=.02)	• N/A	• N/A	• N/A	• Subjective: self-reported questionnaire (Athens Insomnia Scale)
Cotter et al (2012) [29]	• N/A	• N/A	• Social exclusion (frequency =21.4%; P=not applicable)	• Loneliness (frequency =26.4%; P=not applicable)	• Subjective: self-reported questionnaire
Garre‐Olmo et al (2012) [26]	• N/A	• Signs of negative affective mood (β=–.10‐0.20; P>.05)	• Behavioral signs of social interaction (β=–.03 to 0.18; P>.05)	• N/A	• Subjective: proxy-response questionnaire (Quality of Life in Late-Stage Dementia)
Saeki et al (2015) [30]	• Objective sleep onset latency (β=–.019; P=.02) • Subjective sleep onset latency (β=–.021; P<.01)	• N/A	• N/A	• N/A	• Objective: using actigraphy • Subjective: sleep diary
Ahrentzen et al (2016) [28]	• Sleep hours (t=2.150, P<.05)	• Emotional distress (t=−2.085; P<.05)	• N/A	• N/A	• Subjective: self-reported questionnaire
Van Loenhout et al (2016) [31]	• Sleep disturbance (β=1.24, P=.001)	• N/A	• N/A	• Annoyance (β=1.28 to 1.33; P<.001)	• Subjective: hourly diary

XSL•FO
RenderX

Authors (year)	Mental health outcomes				Outcome measure measurement tool
	Sleep problems	Emotional problems	Social interaction problems	Other mental health issues	
Tartarini et al (2017) [27]	• N/A	• N/A	• N/A	• Agitation (*t*=3.09; *P*=.002)	• Subjective: informant rating questionnaire (Cohen-Mansfield Agitation Inventory)
Kim et al (2020) [32]	• Sleep disturbance (frequency=58.7%; *P*=N/A) • Sleep hours (*t*=-0.680; *P*>.05)	• N/A	• N/A	• N/A	• Subjective: self-reported questionnaire
Lindemann et al (2018) [33]	• N/A	• N/A	• Social participation (β=–4.53; *P*<.05)	• N/A	• Subjective: self-reported questionnaire
Lee et al (2020) [34]	• Sleep disorder[c] (OR=0.68, *P*>.05)	• N/A	• N/A	• Mental disorder (including depression, anger, anxiety)[c](OR=1.61, *P*<.05)	• Subjective: self-reported questionnaire
Kanno et al (2022) [16]	• N/A	• ▪ Psychological distress (χ2=not described; *P*<.001)	• N/A	• N/A	• Subjective: self-reported questionnaire (Kessler Psychological Distress Scale)
Yan et al (2022) [17]		• N/A	• N/A	• N/A	• Objective: using a wrist-worn sleep tracker (Fitbit) • Subjective: self-reported questionnaire

XSL•FO
RenderX

Authors (year)	Mental health outcomes				Outcome measure measurement tool
	Sleep problems	Emotional problems	Social interaction problems	Other mental health issues	
	• Objective sleep quality • total sleep time (β=−86.59, *P*=.02) • sleep efficiency (β=−7.96, *P*=.01) • wake time (β=15.37, *P*=.04) • REM sleep (β=−32.97, *P*=.04) • light sleep (β=−38.90, *P*=.02) • deep sleep (β=−27.34, *P*=.27) • Subjective sleep quality • calmness of sleep (β=-1.05, *P*=.02) • ease of falling asleep (β=-1.25, *P*=.01) • ease of awakening (β=−0.23, *P*=.26) • freshness after awakening (β=-1.56, *P*=.02) • satisfaction with sleep (β=−2.26, *P*=.01)				
Liu et al (2023) [15]	• N/A	• N/A	• N/A	• Mental health (*r*=.234, *P* < .01) • vitality • role limitations due to emotional problems • social functioning • mental health	• Subjective: self-reported questionnaire

[a]N/A: not applicable.

[b]Results comparing bedroom temperature in summer and other seasons (fall and winter).

[c]Results for setting temperature <24 °C.

Similarly, the 3 (20%) studies showed inconsistent findings regarding the association between indoor temperature exposure and social interaction problems. In the study by Lindemann et al [33], exposure to hot indoor temperatures significantly reduced social life participation (β=−4.53; *P*<.05) and hindered older adults' participation in social activities, as reflected in their dissatisfaction with opportunities related to general activity, time management, the amount of activity, and participation in the community. By contrast, Garre-Olmo et al [26] found that indoor temperature was not a significant factor affecting the amount of social interaction. Cotter et al [29] identified a significant link between living in a cold home and social exclusion (21.4% in cold homes vs 17.4% in the total sample). In addition, significant associations were found between exposure to hot or cold indoor temperature and other mental health issues (5/15, 33%), including agitation (*t*=3.09; *P*=.002) [27], loneliness (26.4% in cold homes vs 12% in the total sample) [29], and annoyance (β=1.28 - 1.33; *P*<.001) [31].

XSL•FO
RenderX

Factors Associated With Vulnerability to Inappropriate Indoor Temperature Exposure

Five (33%) studies reported characteristics related to vulnerability to exposure to inappropriate indoor temperature. Residents living in poorly conditioned houses, particularly those without energy-efficiency upgrades, were found to be vulnerable to indoor temperature exposure, which can exacerbate health issues, particularly during hot weather [28]. Older adults living in cities or city centers experienced a greater increase in higher indoor temperature exposure than those living in garden cities or suburbs owing to lower access to green spaces [33]. Factors identified as increasing vulnerability to cold indoor temperature exposure included poorer health conditions (eg, disability, arthritis, and fall in the previous 6 months), poorer quality housing (eg, mold, damp, and draughts in the home; no central heating system), difficulty in paying for heating, and social exclusion (eg, feeling lonely and low social activities or hobbies) [29]. In addition, lower gait speed [33], younger age [16,23], and being female [16] were associated with greater exposure to hot or cold indoor temperatures.

Discussion

Principal Findings

To the best of our knowledge, this is the first systematic literature review investigating the association between exposure to varying indoor temperature levels and mental health among older adults. We found that exposure to both high and low indoor temperatures is associated with mental health problems in older adults. Specifically, sleep, emotional, and social interaction problems are frequently reported in association with indoor temperature levels.

Indoor Temperature Levels and Mental Health

Sleep disturbances were the most consistently reported mental health issue related to indoor temperature levels. Several studies suggested that both hot and cold indoor environments were associated with poorer sleep quality, including short sleep duration [28], insomnia [6], sleep disturbances [32], sleep disorders [34], and longer sleep-onset latency [30]. This finding is consistent with recent reviews, which have confirmed the negative association between extreme outdoor temperature exposure and sleep [35,36]. Van Loenhout et al [31] reported that more than 40% of older adults experienced severe sleep disturbances when indoor temperatures exceeded 25 °C, even during short heatwaves. The optimal indoor temperature for sleep is influenced by several factors, including regional climate and housing conditions [10,11], ventilation method used [10], and the type of bedding and sleepwear [11]. Therefore, it is important not only to maintain an appropriate bedroom temperature range but also to create a comfortable sleeping environment through adaptation strategies, such as appropriate bedding, sleepwear, and ventilation, to promote sleep health in older adults.

Indoor temperature levels were also reported to be associated with emotional problems, reduced social interaction, and symptoms such as agitation, loneliness, and annoyance in older adults. Regarding emotional problems, evaluations of housing coldness or warmth are significantly related to psychological distress [16]. In terms of social interactions, significant relationships have been identified between living in a cold home, social exclusion, and loneliness [29]. In addition, participation in social life decreases with increasing indoor temperatures [33]. Furthermore, the cumulative exposure to both hot and cold temperatures has been shown to correlate linearly with agitation [27], and increasing indoor temperatures are associated with heightened annoyance due to heat [31]. These findings are consistent with recent studies on outdoor temperature exposure, which have reported that high temperatures are associated with mood disturbances, reduced social participation due to decreased motivation, aggravation of preexisting mental health conditions, and reduced patience and tolerance [37]. However, in this analysis of indoor temperature exposure, the number of studies on outcomes other than sleep problems was three or fewer, and the methods of reporting results and statistical significance varied among studies, limiting the conclusions that can be drawn about the impact.

Meanwhile, we determined that there is a lack of studies investigating the impact of indoor temperature levels on depression, suicide, and anxiety, which have been highlighted as highly associated with outdoor temperatures [38,39]. To address these limitations, there is a need for longitudinal cohort studies and randomized intervention trials that can more robustly evaluate the causal relationships between indoor temperature exposure and a wider range of mental health outcomes in older adults.

Vulnerable Groups' Exposure to Inappropriate Indoor Temperature

Most of the characteristics related to inadequate indoor temperature exposure identified in this study were vulnerability-modifying factors that either amplified or mitigated the association between exposure and mental health outcomes. For example, this review demonstrates how poor housing and low economic status affect the mental health of older adults through indoor temperature exposures. Older adults living in affordable housing with low energy efficiency [28] or facing economic difficulties such as constraints in paying heating bills [29] are exposed to hotter indoor environments in summer and colder ones in winter. This exposure is related to various mental health outcomes such as emotional distress, reduced sleep length and quality, loneliness, and social exclusion. Older adults are considered the group most affected by poor housing conditions and are also the age group most likely to reside in such conditions [40]. Managing residential environments can be an effective strategy for improving the health outcomes of vulnerable populations by enhancing thermal comfort [41].

In addition, older adults living in cities or city centers with lower access to green spaces were more vulnerable to changes in social participation due to rising temperatures than those living in garden cities or suburban areas [33]. This might have been owing to the urban heat island effect. The urban heat island effect associated with urbanization can cause temperatures within cities to be significantly higher than those in surrounding rural areas, which is exacerbated when there is a lack of green spaces [42]. Previous studies have also emphasized that older

XSL•FO
RenderX

adults living in urban areas are at the greatest risk during heatwaves [43]. As the aging population increasingly moves to cities for health care access [44], addressing the challenges of urban heat by improving infrastructure and integrating green spaces will be essential to protecting the well-being of older adults.

On the contrary, certain preexisting conditions, such as disabilities or psychological distress, may act as potential confounders that influence both indoor temperature exposure and mental health outcomes. Previous studies have shown that people with disabilities experience mental distress 4.6 times more often than do adults without disabilities [45]. At the same time, people with disabilities have a higher risk of heat exposure-related emergency department admissions, particularly for mental disorders, compared with those without disabilities [46]. Therefore, future research should distinguish between vulnerability modifiers and potential confounders for accurate interpretation of the association between indoor temperature and mental health. This distinction will strengthen causal inference and support more precise identification of at-risk groups and intervention priorities.

Policy responses addressing the impact of inadequate indoor temperature exposure on vulnerable groups are limited. Although the World Health Organization's 2018 Housing and Health Guidelines [47] outline the importance of managing indoor temperatures for mental well-being, few policies focus directly on this issue. For example, while the US Department of Health and Human Services provides heating and cooling assistance for low-income households, these programs lack specific guidelines on optimal indoor temperatures for mental health [48]. Similarly, South Korea's Health Plan 2030 includes initiatives for creating age-friendly environments but does not focus on indoor temperature management [49]. Therefore, future policies should integrate evidence on the mental health impacts of indoor temperature exposures. From a policy perspective, legal integration could involve the incorporation of standardized indoor temperature thresholds into national building codes and long-term care facility regulations to safeguard the mental health of older adults. Embedding these standards within existing regulatory frameworks would promote compliance and sustainability, while establishing specific implementation guidelines could support nurses and facility managers in adhering to them. In community settings, incorporating indoor temperature assessments into home-visiting nursing protocols—where nurses systematically record environmental conditions alongside vital signs—could enhance surveillance of thermal environments and inform evidence-based interventions. It may be a useful strategy to prioritize older adults with poor health, those living in substandard housing, those with low economic status, and those experiencing social exclusion—identified in this review as being vulnerable to exposure to inappropriate indoor temperatures—as the target population for community home nursing services.

Measuring Indoor Temperature and Mental Health Outcomes

Objectively measuring indoor temperature exposure and identifying the appropriate temperature ranges and thresholds are the first steps to protecting older adults from extreme temperatures. Compared with younger generations, older adults are less sensitive to thermal stimuli and tend to feel less heat [50]. This means that older adults may feel thermally comfortable at temperatures that in fact could pose health risks. In our review, 6 (40%) studies [6,16,25,27,29,30] investigated sleep problems in cold environments, but only one [30] used objective temperature measurements, making it difficult to determine precise threshold values of safe temperature. The lack of objective data reduces the potential of leveraging behavioral measures to control indoor temperatures. Therefore, further research is needed to establish the optimal indoor temperature conditions for the mental health of older adults. This will enable health care providers to educate older adults on appropriate indoor temperature ranges and help them adjust their indoor temperature exposure.

This study highlights the importance of a balanced approach to using both subjective and objective measures of mental health outcomes to gain a comprehensive understanding of the influence of indoor temperature exposure. Our findings revealed that only sleep-related outcomes were objectively assessed in the literature, whereas other outcomes were measured solely using self-reported questionnaires. While subjective measures capture individual perceptions that can strongly influence behavior, they may introduce bias in older adults with mental health issues or memory limitations, potentially leading to inaccurate results when used alone. Conversely, a study conducted by Okamoto-Mizuno and Tsuzuki [25] revealed that relying solely on objective methods could result in underreporting of sleep disorders, as subjective sleep disturbances are not reflected in objective measurements. Similarly, previous studies have identified discrepancies between subjective and objective sleep assessments in older adults and recommended using both assessment methods for a more balanced evaluation [51]. Therefore, it is crucial to consider objective and standardized methods alongside subjective methods to measure mental health outcomes in older adults. Recently, methods that minimize recall bias by repeatedly measuring participants' experiences in real time using smartphones have been effectively applied to assess mental health outcomes in older adults [52,53]. Incorporating these innovative methods into research could improve the accuracy of measuring subjective mental health outcomes related to changes in indoor temperature exposure. In addition, linking mental health outcome data to corresponding indoor temperature measurements would provide more comprehensive insights.

Limitations

This study had several limitations. First, most included studies were conducted in high-income countries, with a notable number in Asia. This may limit the generalizability of the findings to other regions or lower-income settings where the impact of indoor temperature exposure on mental health might differ. Second, the heterogeneity of study designs, settings, and measurement methods in the included studies prevented the possibility of a meta-analysis. Future research using standardized methodologies is needed to establish more robust evidence. Third, the limited diversity of the outcomes in the included studies also limited our ability to fully grasp the wide-ranging

XSL•FO
RenderX

impact of indoor temperature exposure on the mental health of older adults, highlighting the need for future research that explores a broader range of mental health outcomes.

Implications

Despite the limitations, the findings of this study have significant implications for nursing care, education, research, and policy development. In nursing care, assessing indoor temperature exposure during home visits to older adults, particularly those facing economic hardships or living in substandard housing conditions, could be a valuable approach to improved recognition and mitigation of threats to mental health. Providing practical and specific guidance on maintaining optimal indoor temperatures may support the mental health of older adults.

From an educational perspective, this study highlights the importance of developing programs aimed at enhancing climate literacy among nurses and health care professionals. Such programs could help health care providers better understand the impact of indoor temperature fluctuations on the mental health of older adults within the context of climate change. In addition, these programs may improve health care providers' ability to guide older adults and their families on practical strategies for managing indoor temperatures.

In this research, the findings emphasize the value of establishing standardized study designs and measurement tools to examine the association between indoor temperature exposure and mental health. Consistent criteria for temperature measurement and mental health variables would facilitate robust meta-analyses and provide stronger evidence. In addition, future studies could benefit from adopting collaborative, multidisciplinary approaches involving environmental science, public health, and nursing. Conducting longitudinal studies that combine real-time temperature exposure monitoring with mental health assessments can provide stronger causal evidence.

Policies could focus on retrofitting older homes to enhance insulation and thermal regulation, particularly for low-income households. In addition, community-wide initiatives, such as expanding green spaces in urban areas, could mitigate heat island effects, aligning with public health objectives. Health care providers can contribute to these efforts by sharing evidence-based insights from clinical and community settings with policymakers.

Conclusion

In this systematic review, we identified an association between exposure to both hot and cold indoor temperatures and a range of mental health outcomes in community-dwelling older adults, with sleep disturbances being the most consistently reported. Evidence on other outcomes—such as depression and anxiety, which have been extensively studied in relation to outdoor temperature exposure—remains limited. Future studies using standardized and longitudinal designs are needed to establish optimal indoor temperature conditions and to inform health policies and practices that better protect the mental health of older adults.

Acknowledgments

The authors confirm that no generative artificial intelligence was used in the preparation of this manuscript.

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education [grant: RS-2020-NR049581].

Data Availability

All data generated or analyzed during this study are included in the article and supplementary materials.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies.
[DOCX File, 16 KB - publichealth_v11i1e78257_app1.docx]

Multimedia Appendix 2
Quality rating by the National Institutes of Health (NIH) quality assessment tool for observational cohort and cross-sectional studies.
[DOCX File, 19 KB - publichealth_v11i1e78257_app2.docx]

Checklist 1
PRISMA checklist.
[PDF File, 63 KB - publichealth_v11i1e78257_app3.pdf]

XSL•FO
RenderX

References

1. Intergovernmental Panel on Climate Change. The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change. In: Zhai P, Pirani A, Connors SL, editors. Summary for Policymakers Climate Change: Cambridge University Press; 2021. URL: https://www.ipcc.ch/report/ar6/wg1/ [accessed 2025-11-18]
2. Romanello M, Di Napoli C, Drummond P, et al. The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels. Lancet 2022 Nov 5;400(10363):1619-1654. [doi: 10.1016/S0140-6736(22)01540-9] [Medline: 36306815]
3. Watts N, Amann M, Arnell N, et al. The 2020 report of The Lancet Countdown on health and climate change: responding to converging crises. Lancet 2021 Jan 9;397(10269):129-170. [doi: 10.1016/S0140-6736(20)32290-X] [Medline: 33278353]
4. Tham S, Thompson R, Landeg O, Murray KA, Waite T. Indoor temperature and health: a global systematic review. Public Health 2020 Feb;179:9-17. [doi: 10.1016/j.puhe.2019.09.005] [Medline: 31707154]
5. Zhao J, Uhde E, Salthammer T, et al. Long-term prediction of the effects of climate change on indoor climate and air quality. Environ Res 2024 Feb 15;243:117804. [doi: 10.1016/j.envres.2023.117804] [Medline: 38042519]
6. Bakr IM, Elaziz KMA, Ezz N, Fahim HI. Insomnia in institutionalized older people in Cairo, Egypt: prevalence and risk factors associated. Eur Geriatr Med 2012 Apr;3(2):92-96. [doi: 10.1016/j.eurger.2012.02.002]
7. Basu R, Gavin L, Pearson D, Ebisu K, Malig B. Examining the association between apparent temperature and mental health–related emergency room visits in California. Am J Epidemiol 2018 Apr 1;187(4):726-735. [doi: 10.1093/aje/kwx295] [Medline: 29020264]
8. Zhang S, Yang Y, Xie X, et al. The effect of temperature on cause-specific mental disorders in three subtropical cities: a case-crossover study in China. Environ Int 2020 Oct;143:105938. [doi: 10.1016/j.envint.2020.105938] [Medline: 32688157]
9. Mu J, Kang J. Indoor environmental quality of residential elderly care facilities in Northeast China. Front Public Health 2022;10:860976. [doi: 10.3389/fpubh.2022.860976] [Medline: 35602153]
10. Sekhar SC, Goh SE. Thermal comfort and IAQ characteristics of naturally/mechanically ventilated and air-conditioned bedrooms in a hot and humid climate. Build Environ 2011 Oct;46(10):1905-1916. [doi: 10.1016/j.buildenv.2011.03.012]
11. Nicol F. Temperature and sleep. Energy Build 2019 Dec;204:109516. [doi: 10.1016/j.enbuild.2019.109516]
12. Son JY, Liu JC, Bell ML. Temperature-related mortality: a systematic review and investigation of effect modifiers. Environ Res Lett 2019 Jul;14(7):073004. [doi: 10.1088/1748-9326/ab1cdb] [Medline: 40837670]
13. Matz CJ, Stieb DM, Davis K, et al. Effects of age, season, gender and urban-rural status on time-activity: CanadianHuman Activity Pattern Survey 2 (CHAPS 2). Int J Environ Res Public Health 2014 Feb 19;11(2):2108-2124. [doi: 10.3390/ijerph110202108] [Medline: 24557523]
14. Sattari Z, Weitkamp G, Meijering L. What happens behind doors? Exploring everyday indoor activities when ageing in place. J Aging Stud 2023 Mar;64:101109. [doi: 10.1016/j.jaging.2023.101109] [Medline: 36868621]
15. Liu F, Li Y, Gao X, Du J. The association between perceived housing environment and health and satisfaction among the older adults during the COVID-19 pandemic: a cross-sectional survey in northern China. Buildings 2023;13(11):2875. [doi: 10.3390/buildings13112875]
16. Kanno I, Hasegawa K, Nakamura T, et al. Relationship between the housing coldness/warmth evaluation by CASBEE Housing Health Checklist and psychological distress based on TMM community-based cohort study: a cross-sectional analysis. Public Health (Fairfax) 2022 Jul;208:98-104. [doi: 10.1016/j.puhe.2022.05.003] [Medline: 35738131]
17. Yan Y, Lan L, Zhang H, et al. Association of bedroom environment with the sleep quality of elderly subjects in summer: a field measurement in Shanghai, China. Build Environ 2022 Jan;208:108572. [doi: 10.1016/j.buildenv.2021.108572] [Medline: 35602249]
18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
19. Thompson R, Hornigold R, Page L, Waite T. Associations between high ambient temperatures and heat waves with mental health outcomes: a systematic review. Public Health 2018 Aug;161:171-191. [doi: 10.1016/j.puhe.2018.06.008] [Medline: 30007545]
20. Wang J, Lloyd-Evans B, Giacco D, et al. Social isolation in mental health: a conceptual and methodological review. Soc Psychiatry Psychiatr Epidemiol 2017 Dec;52(12):1451-1461. [doi: 10.1007/s00127-017-1446-1] [Medline: 29080941]
21. The National Institutes of Health. Study quality assessment tools. National Heart, Lung, and Blood Institute. 2020. URL: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools [accessed 2025-01-03]
22. Montero-Vilchez T, Diaz-Calvillo P, Rodriguez-Pozo JA, et al. The burden of hidradenitis suppurativa signs and symptoms in quality of life: systematic review and meta-analysis. Int J Environ Res Public Health 2021 Jun 22;18(13):6709. [doi: 10.3390/ijerph18136709] [Medline: 34206415]
23. Ohnaka T, Tochihara Y, Kanda K. Body movements of the elderly during sleep and thermal conditions in bedrooms in summer. Appl Hum Sci 1995;14(2):89-93. [doi: 10.2114/ahs.14.89]
24. World bank country classifications by income level for 2024–2025. World Bank. 2024. URL: https://blogs.worldbank.org/en/opendata/world-bank-country-classifications-by-income-level-for-2024-2025 [accessed 2025-01-03]

XSL•FO
RenderX

25. Okamoto-Mizuno K, Tsuzuki K. Effects of season on sleep and skin temperature in the elderly. Int J Biometeorol 2010 Jul;54(4):401-409. [doi: 10.1007/s00484-009-0291-7] [Medline: 20041336]
26. Garre-Olmo J, López-Pousa S, Turon-Estrada A, Juvinyà D, Ballester D, Vilalta-Franch J. Environmental determinants of quality of life in nursing home residents with severe dementia. J Am Geriatr Soc 2012 Jul;60(7):1230-1236. [doi: 10.1111/j.1532-5415.2012.04040.x] [Medline: 22702541]
27. Tartarini F, Cooper P, Fleming R, Batterham M. Indoor air temperature and agitation of nursing home residents with dementia. Am J Alzheimers Dis Other Demen 2017 Aug;32(5):272-281. [doi: 10.1177/1533317517704898] [Medline: 28429641]
28. Ahrentzen S, Erickson J, Fonseca E. Thermal and health outcomes of energy efficiency retrofits of homes of older adults. Indoor Air 2016 Aug;26(4):582-593. [doi: 10.1111/ina.12239] [Medline: 26249033]
29. Cotter N, Monahan E, McAvoy H, Goodman P. Coping with the cold—exploring relationships between cold housing, health and social wellbeing in a sample of older people in Ireland. Qual Ageing Older Adults 2012 Mar 9;13(1):38-47. [doi: 10.1108/14717791211213607]
30. Saeki K, Obayashi K, Tone N, Kurumatani N. A warmer indoor environment in the evening and shorter sleep onset latency in winter: the HEIJO-KYO study. Physiol Behav 2015 Oct 1;149:29-34. [doi: 10.1016/j.physbeh.2015.05.022] [Medline: 26004170]
31. van Loenhout JAF, le Grand A, Duijm F, et al. The effect of high indoor temperatures on self-perceived health of elderly persons. Environ Res 2016 Apr;146:27-34. [doi: 10.1016/j.envres.2015.12.012] [Medline: 26710340]
32. Kim S, Kim SY, Oh J, et al. Effects of the 2018 heat wave on health in the elderly: implications for adaptation strategies to climate change. Environ Anal Health Toxicol 2020 Dec;35(4):e2020024-e2020020. [doi: 10.5620/eaht.2020024] [Medline: 33434424]
33. Lindemann U, Skelton DA, Oksa J, et al. Social participation and heat-related behavior in older adults during heat waves and on other days. Z Gerontol Geriatr 2018 Jul;51(5):543-549. [doi: 10.1007/s00391-017-1338-8] [Medline: 29110136]
34. Lee YH, Bae S, Hwang SS, et al. Association between air conditioning use and self-reported symptoms during the 2018 heat wave in Korea. J Prev Med Public Health 2020 Jan;53(1):15-25. [doi: 10.3961/jpmph.19.171] [Medline: 32023670]
35. Janssen H, Ford K, Gascoyne B, et al. Cold indoor temperatures and their association with health and well-being: a systematic literature review. Public Health 2023 Nov;224:185-194. [doi: 10.1016/j.puhe.2023.09.006] [Medline: 37820536]
36. Chevance G, Minor K, Vielma C, et al. A systematic review of ambient heat and sleep in a warming climate. Sleep Med Rev 2024 Jun;75:101915. [doi: 10.1016/j.smrv.2024.101915] [Medline: 38598988]
37. Rony MKK, Alamgir HM. High temperatures on mental health: recognizing the association and the need for proactive strategies—a perspective. Health Sci Rep 2023 Dec;6(12):e1729. [doi: 10.1002/hsr2.1729] [Medline: 38059052]
38. Cianconi P, Betrò S, Janiri L. The impact of climate change on mental health: a systematic descriptive review. Front Psychiatry 2020;11:74. [doi: 10.3389/fpsyt.2020.00074] [Medline: 32210846]
39. Hu J, He G, Meng R, et al. Temperature-related mortality in China from specific injury. Nat Commun 2023;14(1):37. [doi: 10.1038/s41467-022-35462-4]
40. Donald IP. Housing and health care for older people. Age Ageing 2009 Jul;38(4):364-367. [doi: 10.1093/ageing/afp060] [Medline: 19429922]
41. Thomson H, Thomas S, Sellstrom E, Petticrew M. Housing improvements for health and associated socio-economic outcomes. Cochrane Database of Systematic Reviews 2013;2013(3):1-348. [doi: 10.1002/14651858.CD008657.pub2]
42. Wang C, Ren Z, Dong Y, et al. Efficient cooling of cities at global scale using urban green space to mitigate urban heat island effects in different climatic regions. Urban Forestry & Urban Greening 2022 Aug;74:127635. [doi: 10.1016/j.ufug.2022.127635]
43. Wanka A, Arnberger A, Allex B, Eder R, Hutter HP, Wallner P. The challenges posed by climate change to successful ageing. Z Gerontol Geriatr 2014 Aug;47(6):468-474. [doi: 10.1007/s00391-014-0674-1] [Medline: 25119703]
44. Kenny GP, Flouris AD, Yagouti A, Notley SR. Towards establishing evidence-based guidelines on maximum indoor temperatures during hot weather in temperate continental climates. Temperature (Austin) 2019;6(1):11-36. [doi: 10.1080/23328940.2018.1456257] [Medline: 30906809]
45. Cree RA, Okoro CA, Zack MM, Carbone E. Frequent mental distress among adults, by disability status, disability type, and selected characteristics—United States, 2018. MMWR Morb Mortal Wkly Rep 2020 Sep 11;69(36):1238-1243. [doi: 10.15585/mmwr.mm6936a2] [Medline: 32914770]
46. Park J, Kim A, Kim Y, et al. Association between heat and hospital admissions in people with disabilities in South Korea: a nationwide, case-crossover study. Lancet Planet Health 2024 Apr;8(4):e217-e224. [doi: 10.1016/S2542-5196(24)00027-5] [Medline: 38580423]
47. WHO housing and health guidelines. World Health Organization. 2018. URL: https://www.who.int/publications/i/item/9789241550376 [accessed 2025-01-03]
48. Administration for Children and Families. U.S. Department of Health and Human Services. Low Income Home Energy Assistance Program (LIHEAP). 2024. URL: https://acf.gov/ocs/programs/liheap [accessed 2025-01-03]
49. The 5th National Health Promotion Plan [health plan 2030, 2021–2030]. Korea Health Promotion Institute. 2021. URL: https://www.khealth.or.kr/healthplan2020 [accessed 2025-01-03]

XSL•FO
RenderX

50. Dufour A, Candas V. Ageing and thermal responses during passive heat exposure: sweating and sensory aspects. Eur J Appl Physiol 2007;100:19-26. [doi: 10.1007/s00421-007-0396-9]
51. Hughes JM, Song Y, Fung CH, et al. Measuring sleep in vulnerable older adults: a comparison of subjective and objective sleep measures. Clin Gerontol 2018;41(2):145-157. [doi: 10.1080/07317115.2017.1408734] [Medline: 29283797]
52. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
53. Fingerman KL, Kim YK, Ng YT, Zhang S, Huo M, Birditt KS. Television viewing, physical activity, and loneliness in late life. Gerontologist 2022 Aug 12;62(7):1006-1017. [doi: 10.1093/geront/gnab120] [Medline: 34379115]

Edited by A Mavragani, T Sanchez; submitted 29.05.25; peer-reviewed by OC Balogun, PM Graffy; accepted 04.11.25; published 27.11.25.

Please cite as:
Park MK, Kang B, Hong D, Yoon S, Kim M, Kim Y, Ni Z, McConnell ES
Association of Indoor Temperature Level and Mental Health Among Community-Dwelling Older Adults: Systematic Review
JMIR Public Health Surveill 2025;11:e78257
URL: https://publichealth.jmir.org/2025/1/e78257
doi:10.2196/78257

© Min Kyung Park, Bada Kang, Dahye Hong, Seolah Yoon, Minsub Kim, Yubeen Kim, Zhao Ni, Eleanor S McConnell. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Safety and Efficacy of Respiratory Syncytial Virus Vaccination in Older Adults: Systematic Review and Meta-Analysis of Randomized Controlled Trials

Qingzhi Xiao[1*], BS; Ruizhe Yang[2*], MPH; Lei Zhang[3], PhD; Ye Tian[4], MD; Xu Wang[3], PhD; Wei Li[3], PhD

[1]School of Pediatrics, Xinjiang Medical University, Xinjiang, China

[2]Department of Public Health, Children's Hospital of Nanjing Medical University, Nanjing, China

[3]Clinical Medical Research Center, Children's Hospital of Nanjing Medical University, 72 Guangzhou Rd, Nanjing, China

[4]Department of Infectious Diseases, Children's Hospital of Nanjing Medical University, Nanjing, China

[*]these authors contributed equally

Corresponding Author:
Wei Li, PhD
Clinical Medical Research Center, Children's Hospital of Nanjing Medical University, 72 Guangzhou Rd, Nanjing, China

Abstract

Background: Although the respiratory syncytial virus (RSV) membrane protein F has become one of the major target proteins for RSV vaccine development, previous studies have identified safety issues with RSV vaccines in older adults. The causes of these issues remain unclear.

Objective: We aimed to evaluate the safety and efficacy of a novel RSV preF vaccine in preventing lower respiratory tract infections (LRTIs) and RSV-associated acute respiratory illnesses (RSV-ARIs) in older adults.

Methods: We conducted a systematic review of randomized controlled trials from 5 databases—PubMed, Embase, the Cochrane Library, Web of Science, and the Cochrane Central Register of Randomized Controlled Trials—published up to July 31, 2024. For categorical variables, we used the risk ratio; for continuous variables, we used the weighted mean difference or standardized mean difference. We extracted key study characteristics and assessed their quality using the Newcastle-Ottawa Scale. Sensitivity analyses were used to evaluate the impact of individual studies on pooled effects, and publication bias was examined with Egger and Begg tests.

Results: Meta-analysis of 5 studies (101,825 older adults) showed a reduced LRTI incidence in the vaccinated group, with heterogeneity resolved after excluding 1 study. Analysis of 4 studies (99,931 participants) confirmed a lower RSV-ARI incidence with no heterogeneity. Safety analysis (14 studies, 76,695 participants) showed higher adverse events (AEs) in the vaccinated group, mainly injection-site reactions and vaccine-related AEs, but no significant difference in serious AEs. Subgroup analyses identified potential sources of AE heterogeneity, and sensitivity analyses confirmed the efficacy but suggested variability in the injection-site reactions.

Conclusions: This meta-analysis showed a statistically significant difference in the incidence of LRTI and RSV-ARI between the vaccine and placebo groups in older adults. However, no significant differences were observed in overall AEs or serious AEs.

Trial Registration: PROSPERO CRD42024525116; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024525116.

(JMIR Public Health Surveill 2025;11:e74271) doi:10.2196/74271

KEYWORDS

respiratory syncytial virus vaccine; safety; efficacy; older population; adverse events

Introduction

Respiratory syncytial virus (RSV) is a pathogen that causes typical respiratory tract infections, with seasonal winter peaks in temperate climates [1]. Despite the relative antigenic stability of RSV and its fusion (F) protein, natural infection elicits only partial and short-lived immunity and does not consistently prevent subsequent infections [1-5]. RSV infections usually resolve without complications or sequelae in immunocompetent populations [6]. Among those at high risk, RSV may cause a potential disease burden comparable to that of influenza, including hospitalizations, intensive care unit admissions, and deaths [7-9].

In fact, in older adults aged ≥60 years, RSV can cause more serious respiratory illnesses, such as lower respiratory tract disease [10], especially in people with underlying medical conditions or those who are immunocompromised [11,12]. In older adults, RSV infections impose a significant disease burden

XSL•FO
RenderX

[13], which has been underestimated [14-16]. A recent systematic review based on data from high-income countries reported that the pooled estimates were 1.62% (95% CI 0.84 - 3.08) for the RSV-associated acute respiratory illnesses (RSV - ARI) attack rate, 0.15% (95% CI 0.09 - 0.22) for the hospitalization attack rate, and 7.13% (95% CI 5.40 - 9.36) for in - hospital case fatality rate [15]. According to 2019 global population data, the same review estimated that approximately 5 million cases of acute respiratory tract infection, 0.5 million hospitalizations, and 33,000 in-hospital deaths among older adults could be attributed to RSV in 2019 [15]. Based on the United Nations Department of Economic and Social Affairs population estimates for 2025, the number of RSV cases in older adults in high - income countries could be as high as 10.9 million, RSV hospitalizations as high as 0.8 million, and the number of deaths due to RSV as high as 74,000 [15].

The severity of RSV-associated disease in older adults has been ascribed to waning humoral and cellular immune responses (immunosenescence) that were induced by previous RSV infections [5,17-21]. Thus, vaccination approaches to overcome waning immunity could help avoid serious RSV-associated diseases in older adults [22]. An effective RSV vaccine in older adults would likely need to boost or induce potent and durable RSV-neutralizing antibody responses, as well as restore and elicit RSV-specific T-cell responses [23]. Multiple RSV vaccines targeting the prefusion F (preF) protein and administered through various platforms—including adjuvanted protein subunit, bivalent preF, adenoviral-vectored, and mRNA-based technologies—have recently been evaluated in randomized clinical trials involving older adults [24-26]. A first-in-human study in healthy adults aged 18 to 85 years has reported that the RSV preF vaccine was safe, well tolerated, and highly immunogenic [27,28]. Because immunosenescence may reduce vaccine responses in older adults, the addition of an adjuvant to vaccine formulations may enhance immune responses [29,30]. Moreover, in preclinical studies, combined regimens containing Ad26.RSV.preF and recombinant RSV preF proteins showed greater humoral and cellular immunogenicity and improved protection compared with either component alone [31].

Although the RSV membrane protein F has become one of the major target proteins for RSV vaccine development, previous studies have identified safety issues with RSV vaccines in older adults, but the causes of these issues remain unclear. We systematically evaluated safety outcomes—including overall adverse events (AEs), vaccine-related AEs, and serious AEs—using subgroup and sensitivity analyses to investigate potential heterogeneity. This review incorporates randomized controlled trials (RCTs) published up to July 31, 2024, thereby integrating the most recent large-scale evidence that was unavailable to earlier meta-analyses. In addition, we conducted a rigorous quality assessment using the Newcastle-Ottawa Scale, augmented by comprehensive sensitivity analyses and publication bias assessments, all of which strengthened the validity and reliability of our findings relative to those of prior reviews. By explicitly addressing key research gaps and synthesizing current high-quality evidence, this study provides a robust, up-to-date evaluation of the safety and efficacy of RSV vaccines in older adults, offering actionable insights for clinical practice and future research.

Methods

This meta-analysis was prospectively registered with PROSPERO (CRD42024525116) prior to commencing the study search. The analysis adhered strictly to the methodological standards established by Cochrane for intervention evaluations and is reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Search Strategy

A comprehensive and systematic search was performed across 5 major databases (PubMed, Embase, the Cochrane Library, Web of Science, and the Cochrane Central Register of Controlled Trials) from their inception to July 31, 2024. Search terms were used both individually and in combination using the advanced search functionalities of each database. Additionally, the reference lists of relevant reviews were examined to identify additional studies.

Outcome-related terms were included only when necessary to enhance specificity and minimize the risk of omitting relevant studies. Validated study design filters were applied to target-specific methodologies (eg, RCTs or observational studies), and searches were restricted to human subjects and relevant languages, as appropriate. In addition, the reference lists of all eligible studies and pertinent systematic reviews were manually screened to identify potentially eligible studies not captured by the electronic search. Search results were exported to a reference management system (EndNote version 11; Clarivate Analytics) for deduplication and record management. The complete search strategy, including the full search string, date of execution, and retrieval yield, was documented and reported in accordance with the PRISMA guidelines.

Selection Criteria

The inclusion criteria for the meta-analysis were as follows: (1) peer-reviewed original research articles, excluding conference abstracts, gray literature, and government reports; (2) studies published in English; (3) research specifically addressing vaccination in older adult populations; (4) administration of RSV vaccines; (5) RCTs comparing RSV vaccines with placebo; and (6) studies reporting on the safety and efficacy of RSV vaccines.

The exclusion criteria for the studies were as follows: (1) non-peer-reviewed studies, including systematic reviews, meta-analyses, conference abstracts, letters, editorial comments, case reports, unpublished articles, and articles not published in English; (2) vaccination of individuals younger than 60 years; (3) studies involving animals or preliminary clinical trials; (4) non-RCTs; and (5) studies that did not report on the efficacy and safety of RSV vaccines.

Study Screening and Data Extraction

Following the application of the inclusion criteria, we imported the selected studies into EndNote for automated deduplication. Subsequently, we conducted a thorough review of the titles and

XSL•FO
RenderX

abstracts of all identified studies to determine eligibility. The methodological quality of the included studies was evaluated using the Newcastle-Ottawa Scale for cohort studies (Table S3 in Multimedia Appendix 1). The literature was independently screened by 2 reviewers (WL and QX), and any discrepancies were resolved through discussion or, if necessary, by consulting a third reviewer. The extracted data encompassed the following elements: the first author's name, year of publication, country of study, study design, risk of bias assessment for RCTs, sample size, participant grouping, baseline characteristics, interventions, and outcome measures (Table S2 in Multimedia Appendix 1). The primary outcome measures encompassed medically attended lower respiratory tract illness, RSV-associated respiratory illness in older adults, AEs, and serious AEs within this population. We conducted a comprehensive review of the included studies, original texts, and supplementary materials to ensure that the relevant data were not overlooked.

Outcome Variables

According to the studies included in the meta-analysis, the primary outcome variable commonly assessed was the number of older adults with at least one serious AE or AE, including those deemed to be related to the vaccine. Additionally, some studies have reported the number of older adults with medically confirmed RSV-associated lower respiratory tract infections (LRTIs) or lower respiratory tract diseases and the number of older adults with medically confirmed RSV-associated acute respiratory infections, providing evidence for evaluating the efficacy of RSV vaccination. Details of the outcome variables, including both primary and secondary outcomes, are provided in Tables S4-S17 in Multimedia Appendix 1.

Statistical Analysis

The meta-analysis and subgroup analyses were conducted using Stata (version 17.1; StataCorp LLC) software. For categorical variables, the risk ratio was used as the effect measure, whereas for continuous variables, the weighted mean difference or standardized mean difference was used. Each effect size was expressed with a 95% CI, and the corresponding point estimate was provided. The heterogeneity of the literature was assessed using the χ^2 test. If $P>.1$ and $I^2 \leq 50\%$, a fixed-effect model was used. If $P \leq .1$ and $I^2 > 50\%$, potential sources of heterogeneity were investigated, and after excluding any evident clinical heterogeneity, a random-effects model was used to evaluate the heterogeneity. Sensitivity analyses were conducted for both models. The significance level for the meta-analysis was set at $\alpha=.05$, unless otherwise specified. A one-way sensitivity analysis was conducted to assess the impact of individual studies on pooled outcomes with significant heterogeneity. Publication bias was evaluated using funnel plots generated by Review Manager (version 5.4.1; Cochrane Collaboration) and Egger regression test in Stata 15.1 for outcomes with three or more studies. A $P<.05$ indicated statistical significance for publication bias.

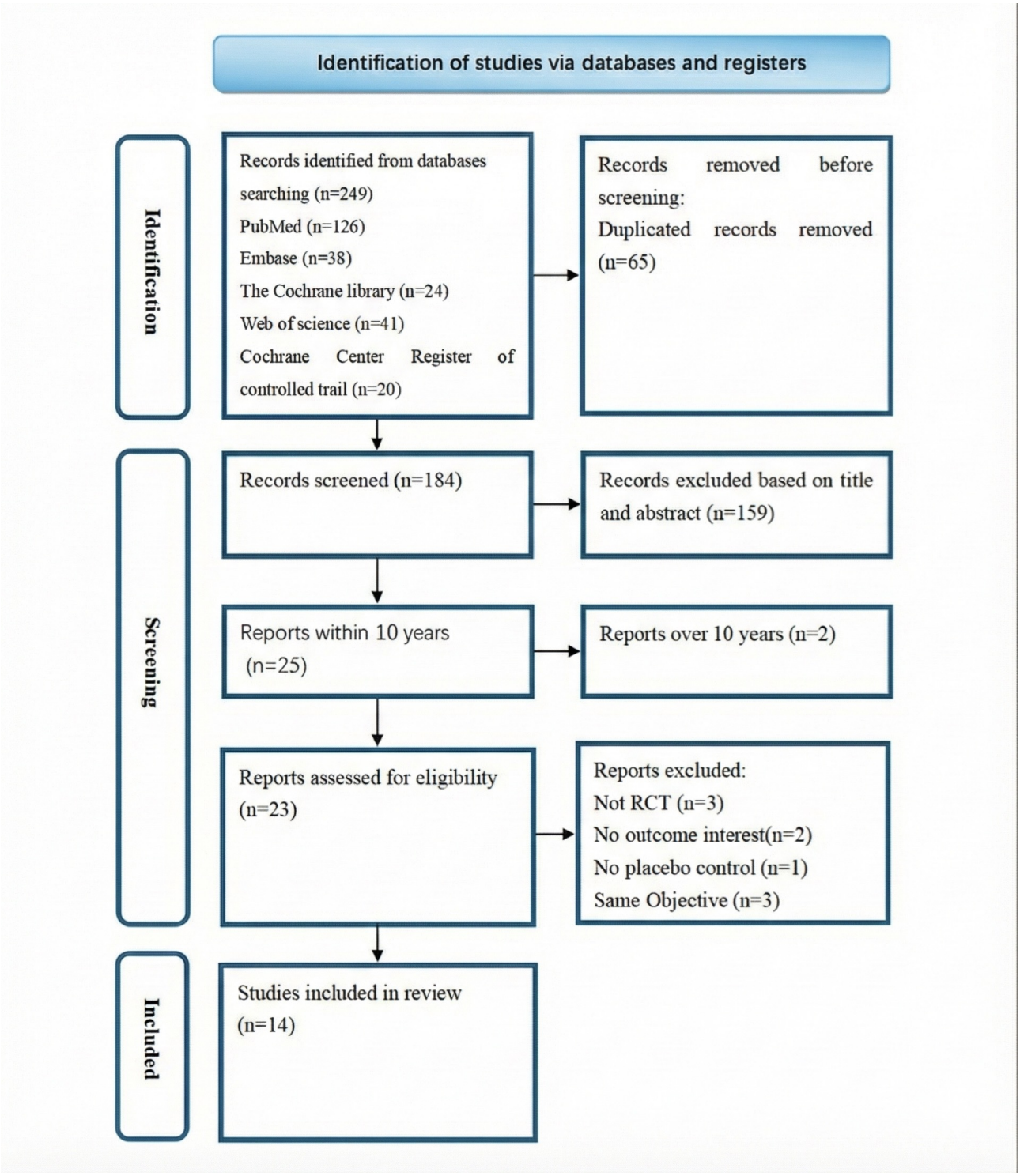
Results

Study Search and Characteristics Overview

A systematic and comprehensive literature search was conducted to identify all relevant studies pertinent to the scope of this review. Initial database queries across PubMed, Embase, the Cochrane Library, Web of Science, and the Cochrane Central Register of Controlled Trials retrieved a total of 249 records. After the removal of 65 (26.1%) duplicate entries, 184 (73.9%) unique records were screened for eligibility based on their titles and abstracts. Of these, 159 (86.4%) were excluded due to irrelevance to the research question, leaving 25 (13.6%) full-text articles for in-depth evaluation. One (4%) article was further excluded for being published more than 10 years ago, which may limit its applicability given advances in the field, resulting in 24 (96%) studies undergoing full-text eligibility assessment. Upon detailed appraisal, an additional 10 (41.7%) studies were excluded according to the predefined criteria: failure to employ an RCT design (3/10, 30%), absence of reported outcomes of interest (2/10, 20%), lack of a placebo control group (1/10, 10%), or significant overlap in the study population and objectives with another included study (3/10, 30%). Consequently, out of 24 studies, 14 (58.3%) fulfilled all prespecified inclusion criteria and were incorporated into the final qualitative synthesis. All included studies were assessed as high quality with low risk of bias (Table S2 in Multimedia Appendix 1). The systematic search and selection process is summarized in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) flow diagram of the systematic review and selection process for respiratory syncytial virus vaccine safety and efficacy in older adults (≥60 y). RCT: randomized controlled trial.

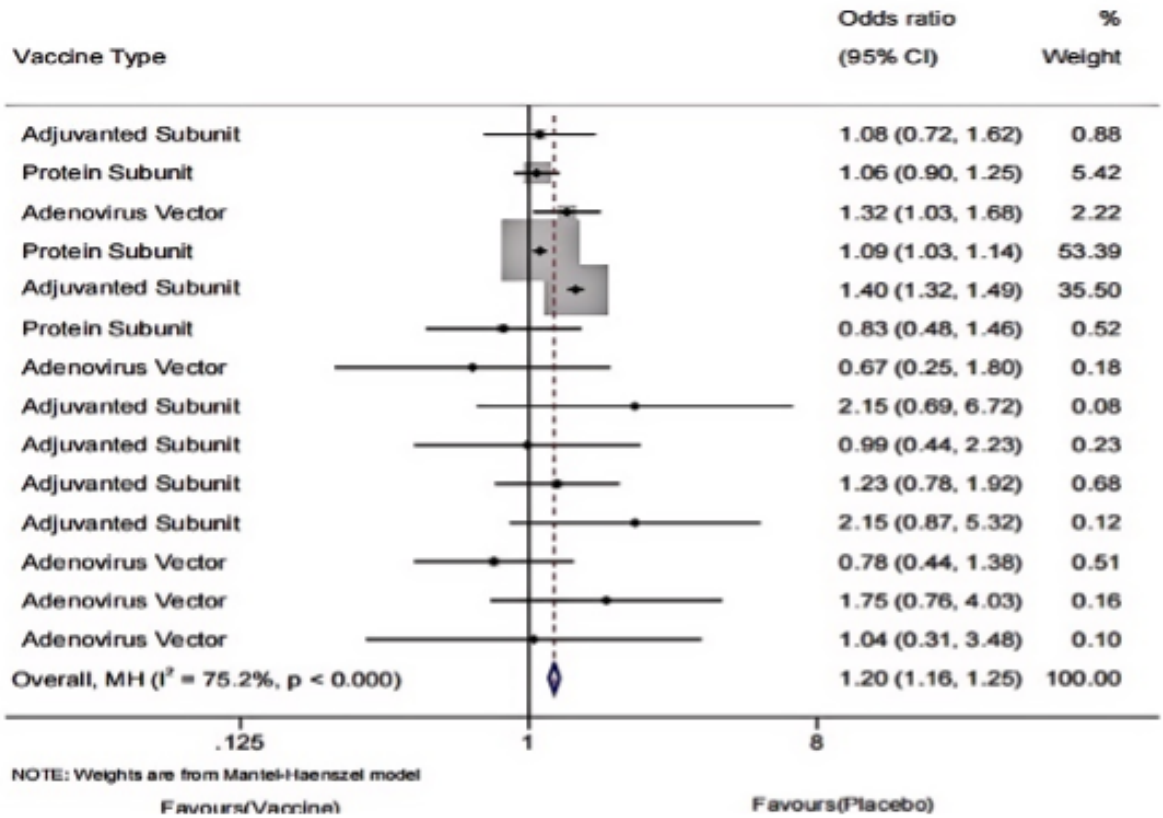


Safety of Vaccination

To assess the difference in overall AEs following vaccine or placebo injections, we included 14 studies reporting AE outcomes in 76,695 older adults, including 38,459 (50.1%) vaccinated and 38,236 (49.9%) in the placebo group. The meta-analysis showed a statistically significant higher incidence of AEs in the vaccinated group, primarily injection-site reactions (eg, swelling, pain, or redness), compared with the placebo group, and significant heterogeneity was also observed (Figure 2).

XSL•FO
RenderX

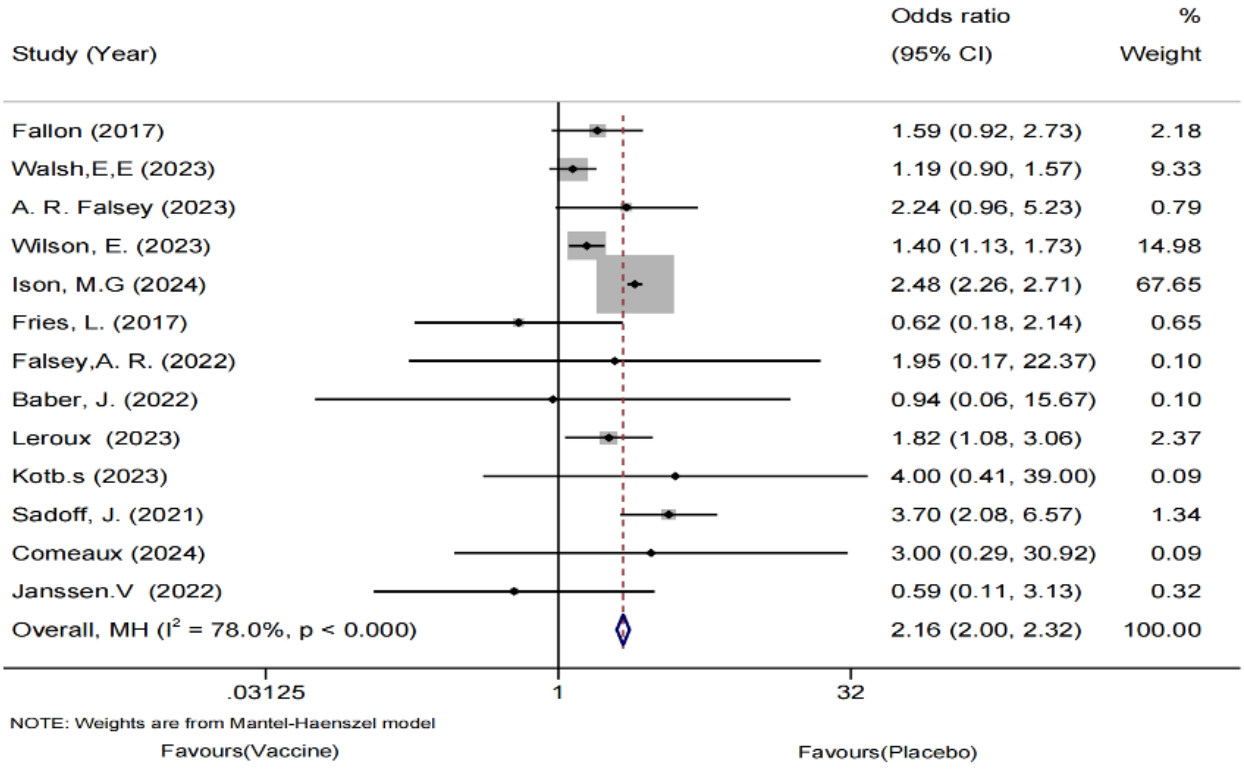
Figure 2. Forest plot of overall adverse event incidence following RSV vaccination versus placebo in older adults: 14 randomized controlled trials (n=76,695; 2017 - 2024).



Because Williams et al [32] did not report AEs, including those potentially related to vaccine components, their study was not included in the analysis in this section. A total of 13 studies (71,531 older adults) were included to analyze AEs attributable to the chemical components of the vaccine or placebo. The meta-analysis revealed a statistically significant higher incidence of AEs due to chemical components, including adjuvants and formulations containing Ad26.RSV.preF or recombinant RSV preF protein, compared with the placebo group. Significant heterogeneity was observed (Figure 3).

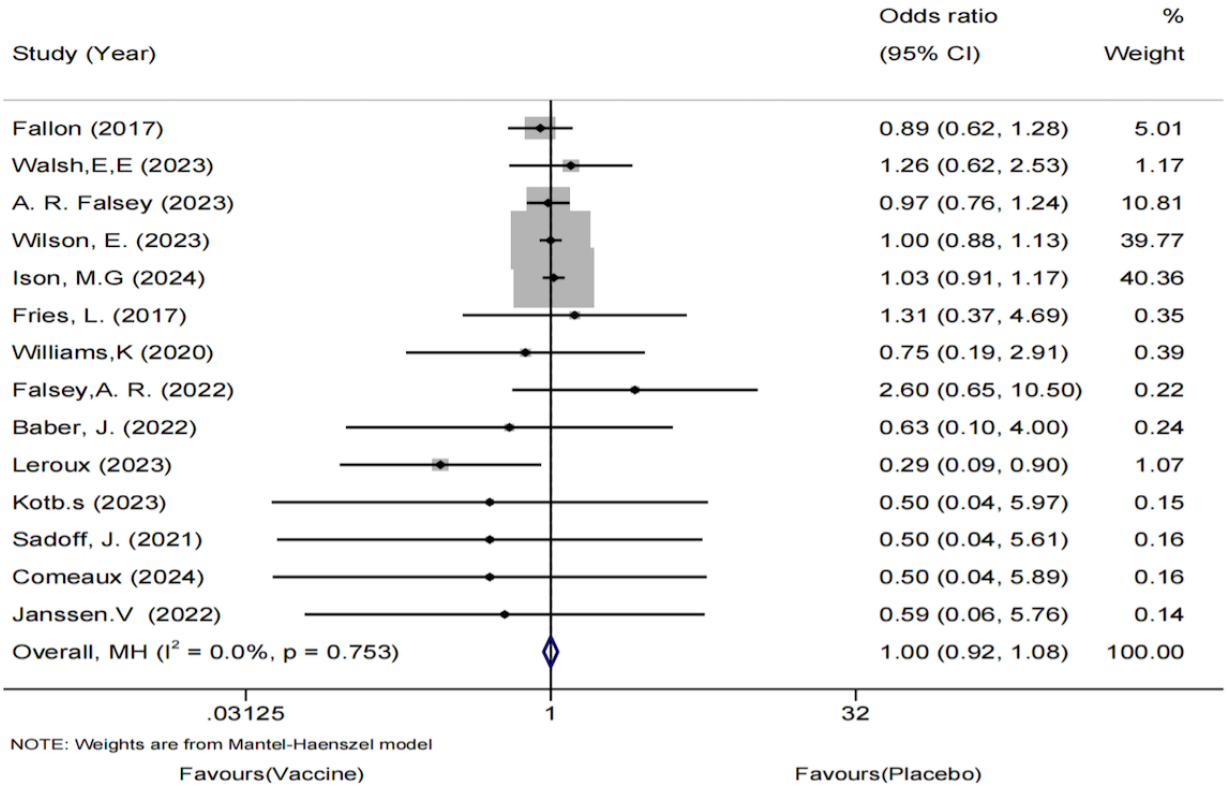
XSL•FO
RenderX

Figure 3. Forest plot of adverse events caused by vaccine chemical composition in older adults: 13 randomized controlled trials (n=71,531; 2017 - 2024) [26-28,33-34,35-36].



Serious AEs caused by vaccine or placebo injection were also reported in the included studies. A total of 14 studies involving a total of 76,695 older adults were included in the analysis. The meta-analysis showed no significant difference in the incidence of serious AEs between the vaccine and placebo groups (Figure 4). Moreover, no heterogeneity was detected, and publication bias was not observed based on funnel plots and Egger regression tests.

Figure 4. Forest plot of serious adverse events following respiratory syncytial virus vaccination versus placebo in older adults: 14 randomized controlled trials (n=76,695; 2017 - 2024) [26-28,32-34,35-36].



XSL•FO
RenderX

Efficacy of the RSV Vaccination

LRTI Analysis

Five studies involving 1,01,825 older adults were analyzed. The meta-analysis showed a significantly lower incidence of LRTIs in the vaccinated group (Figure 5); however, significant heterogeneity was observed. Rescreening revealed that Falloon et al's study [37] included participants with different baseline characteristics, many of whom had a history of respiratory infections. We excluded this study and re-conducted the pooled meta-analysis. The revised analysis showed a significant reduction in LRTIs in the vaccinated group compared with the placebo group (Figure 6), with significantly reduced heterogeneity. This finding suggests that the observed heterogeneity may have been influenced by the study by Falloon et al [37]. Additionally, no publication bias was detected using funnel plots or Egger regression test.

Figure 5. Forest plot of lower respiratory tract infection incidence following respiratory syncytial virus vaccination versus placebo in older adults: 5 randomized controlled trials (n=1,01,825; 2017 - 2024) [26,27,33,34,37].

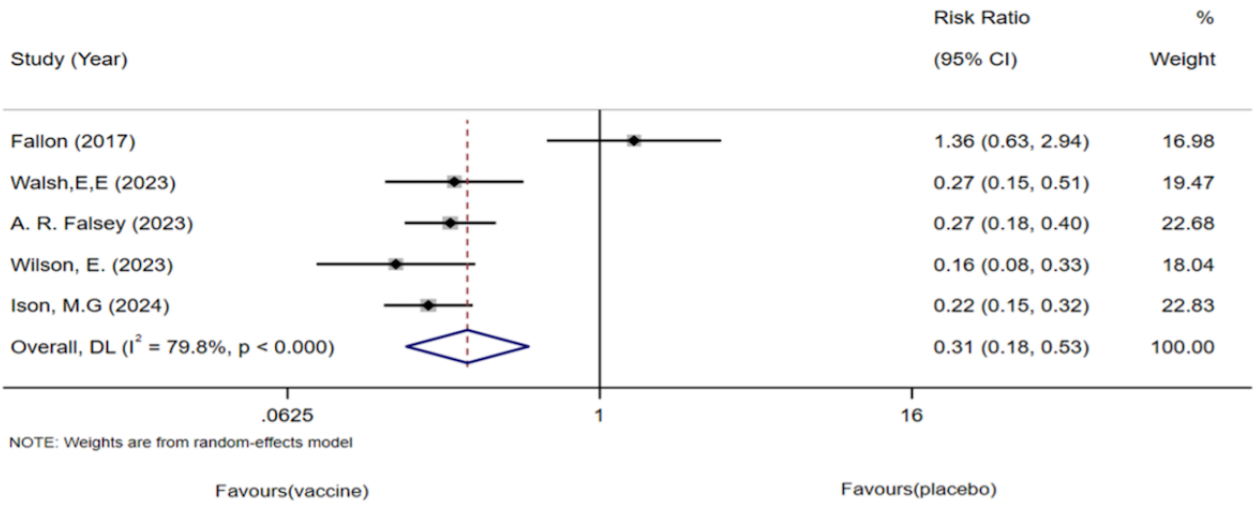
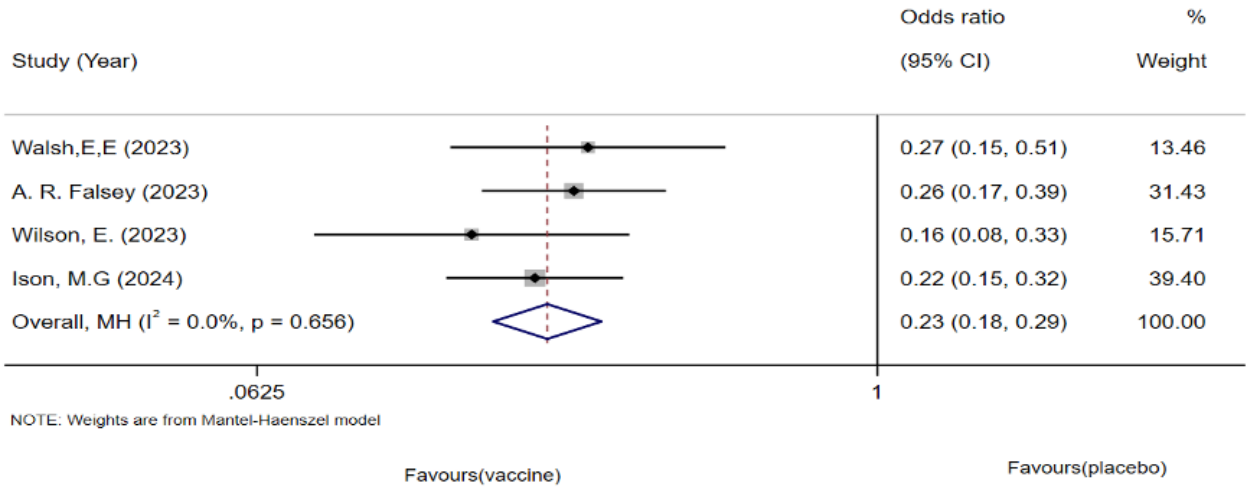


Figure 6. Forest plot of lower respiratory tract infection incidence following respiratory syncytial virus vaccination versus placebo after sensitivity analysis: 4 randomized controlled trials (n=99,931; 2023 - 2024) [26,27,33,34].



RSV-ARI Analysis

Because of the heterogeneity observed in the study by Falloon et al (Figure 7), we excluded it and analyzed 4 articles involving 99,931 older adults [37]. The pooled analysis showed a reduction in the incidence of RSV-ARI in the vaccinated group (Figure 8). No significant heterogeneity or publication bias was detected using funnel plots or Egger test.

XSL•FO
RenderX

Figure 7. Forest plot of respiratory syncytial virus (RSV)-associated acute respiratory illness incidence in older adults: 5 randomized controlled trials (n=1,01,825; 2017 - 2024) [26,27,33,34,37].

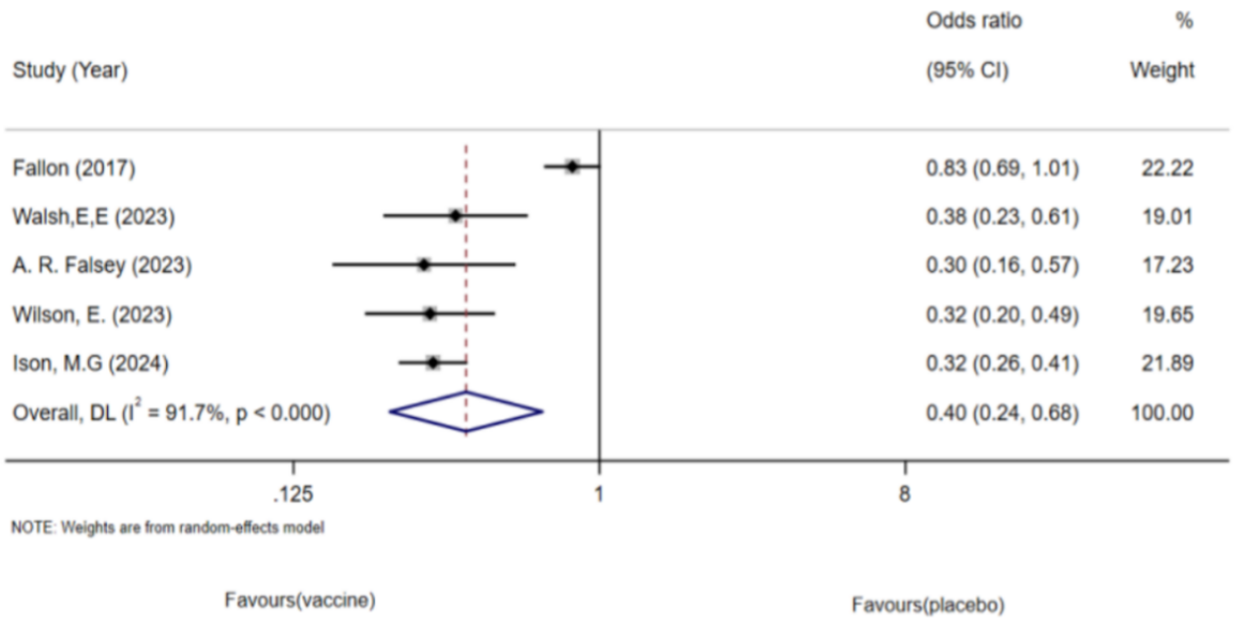
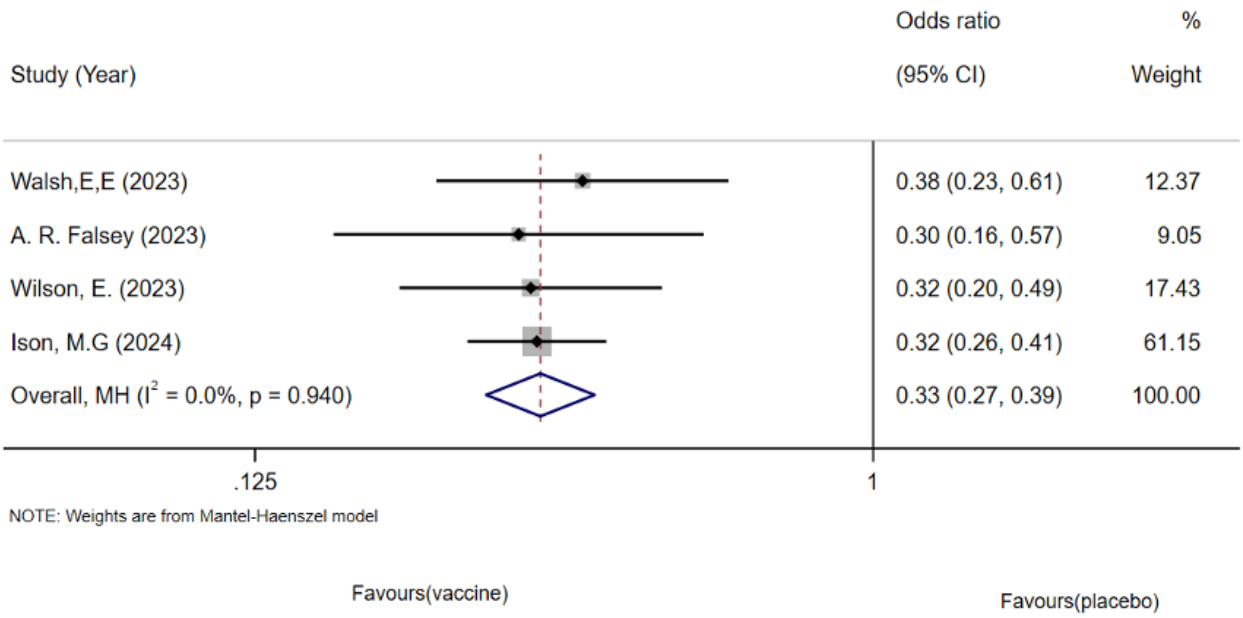


Figure 8. Forest plot of respiratory syncytial virus (RSV)-associated acute respiratory illness incidence after sensitivity analysis: 4 randomized controlled trials (n=99,931; 2023 - 2024) [26,27,33,34].



Subgroup Analysis

Overview

To investigate the cause of heterogeneity, we conducted subgroup analysis according to intervention mode, ethnic distribution, age at baseline, and the quality of studies. Stata (version 17; StataCorp LLC) was used for image rendering and analysis. The subgroup analysis in this study was only aimed at overall AEs following vaccination or placebo injection and AEs due to chemical compositions of the vaccine or placebo. Other subgroup analyses were not performed because fewer than 10 data pairs were available.

In addition, our classification of RSV vaccines was based on the target antigen rather than the manufacturing platform. mRESVIA (Moderna) encodes the RSV preF protein; although the antigen is produced in vivo through the mRNA platform, it is antigenically identical to the preF protein administered via recombinant vaccines. Therefore, mRESVIA was categorized under the preF group. Given that mRNA-based RSV vaccines only became commercially available in mid-2024, which was

XSL•FO
RenderX

beyond the primary evidence window of this review, mRESVIA was not evaluated as an independent subgroup.

Subgroup Analysis Based on Race Distribution

Given that many of the studies involved different countries, with most located in Europe or North America, we categorized the 14 included studies into 2 groups based on the percentage of White participants: those with more than 80% White participants and those with 80% or fewer. Among the studies reporting overall AE outcomes, group 1 exhibited significant heterogeneity, whereas group 2 showed a lack of heterogeneity (Figure S1 in Multimedia Appendix 2). Therefore, we conclude that racial differences may contribute to the heterogeneity observed in all 14 studies describing overall AE outcomes. However, when examining AEs due to chemical compositions, group 1 still demonstrated significant heterogeneity, whereas group 2 showed a notable improvement, with no heterogeneity (Figure S2 in Multimedia Appendix 2). These results indicate a possible association between racial composition and between-study heterogeneity. Nevertheless, given the limited number of studies in each subgroup, as well as the influence of intervention type and study quality identified through subgroup and meta-regression analyses, heterogeneity cannot be conclusively attributed to racial factors alone. Thus, race should be considered a potential, rather than a definitive, contributor to the observed heterogeneity.

Subgroup Analysis Based on Age Distribution

Despite the inclusion criteria specifying older adults aged 60 years and above, considerable age variability remained within this population. To address this, we stratified the 14 studies into 2 subgroups based on median age: those with a median age of less than 70 years and those with a median age of 70 years or more. For AEs, no significant heterogeneity was observed in either subgroup (Figure S3 in Multimedia Appendix 2). Consequently, it is difficult to ascertain whether age differences contributed to heterogeneity. However, subgroup analysis focusing on AEs attributable to chemical compositions revealed significant heterogeneity in the younger subgroup (median age <70 y), while the older subgroup (median age ≥70 y) suggested that age differences may also influence heterogeneity in AEs related to chemical compositions (Figure S4 in Multimedia Appendix 2).

Subgroup Analysis Based on Article Quality

When conducting a meta-analysis, variations in the quality of included studies can substantially influence the final pooled results. We categorized the studies based on previously conducted risk-of-bias assessments into 2 groups: high-quality studies (those with a low risk of bias) and lower-quality studies (those rated as medium or high risk of bias). For overall AE outcomes, heterogeneity was observed in the high-quality group, while no significant heterogeneity was noted in the lower-quality group (Figure S5 in Multimedia Appendix 2). This finding suggests that differences in study quality may be a contributing factor to heterogeneity among studies reporting overall AEs. Additionally, subgroup analysis of studies examining AEs related to vaccine components revealed heterogeneity in the high-quality group, whereas the lower-quality group exhibited the opposite trend (Figure S6 in Multimedia Appendix 2). These findings indicate that variations in study quality contributed to the heterogeneity observed in studies describing AEs due to chemical compositions.

Subgroup Analysis Based on Clinical Intervention Methods

In the included studies, interventions for participants were not only limited to the traditional RSV preF vaccine but also included the RSV preF vaccine with an adjuvant and the Ad26.RSV.preF vaccine. Participants were categorized into three groups based on the type of intervention: (1) RSV preF vaccine versus placebo, (2) RSV preF vaccine with adjuvant versus placebo, and (3) Ad26.RSV.preF vaccine versus placebo.

As shown in Figure S7 in Multimedia Appendix 2, significant heterogeneity was observed in the overall pooled results of the studies. Although each of the 3 groups exhibited low heterogeneity individually, meta-regression analysis revealed substantial heterogeneity between the groups. These findings suggest that the intervention method may be a potential source of the observed heterogeneity.

In the subgroup analysis of AE studies attributed to chemical compositions in vaccines (Figure S8 in Multimedia Appendix 2), significant heterogeneity was observed in the overall pooled results. Although each of the 3 groups demonstrated low within-group heterogeneity, substantial between-group heterogeneity was evident. Therefore, the intervention method can be considered a potential source of this heterogeneity.

Sensitivity Analysis

We conducted a one-way sensitivity analysis to evaluate the impact of each individual study on the risk ratio for LRTI, ARI, common AEs, and serious AEs in older adults by sequentially excluding each study (Figures S9-S13 in Multimedia Appendix 2). The sensitivity analysis revealed no statistically significant changes in LRTI, ARI, common AEs, or serious AEs among the included studies in older adults, indicating robustness of the overall results. However, in the analysis of solicited injection-site AEs, we observed that the sequential exclusion of the studies by Falsey et al [26], Walsh et al [27], and Ison et al [33] resulted in a shift from statistically significant to nonsignificant aggregated results. This finding suggests that the data on solicited injection-site AEs may lack robustness.

Discussion

This meta-analysis assessed the safety and efficacy of a novel RSV preF vaccine in preventing LRTIs and RSV-ARIs in older adults. A systematic review and meta-analysis of RCTs was conducted using 5 databases. The analysis indicated a low risk of publication bias, and the included studies exhibited high methodological quality [38]. Five studies showed a significant reduction in LRTI incidence, with low heterogeneity after removing 1 outlier. Four studies showed a significant decrease in RSV-ARI, with no heterogeneity. Safety data from 14 studies indicated more nonserious AEs in the vaccine group, mainly injection-site reactions, but no difference in serious AEs. Subgroup and sensitivity analyses supported efficacy

XSL•FO
RenderX

consistency and identified minor variability in local reactions. Overall, the RSV preF vaccine significantly reduced the risk of LRTI and RSV-ARI in older adults without increasing serious AEs. According to our findings, the administration of the RSV preF vaccine significantly reduced the risk of RSV-ARI [39].

These results align with previous findings from randomized, placebo-controlled, multicountry trials that enrolled diverse populations of adults aged ≥60 years, similar to our study participants [24,26,40]. These trials included adults with age-related medical conditions representative of those seen in the general older adult population [15,41,42]. Comparable trial designs and participant characteristics enhance the consistency of outcomes between those studies and ours. The vaccine effectively mitigated the leading causes of lower LRTIs and ARIs among older adults. Our findings indicate that the vaccine provided a significant enhanced level of protection against these infections in the target population. Furthermore, these results suggest that vaccination may be the most effective strategy for safeguarding older adults.

Our meta-analysis identified a statistically significant higher incidence of general AEs and AEs related to vaccine components in the vaccinated older adult population compared with the placebo group, with notable heterogeneity observed. Although our analysis suggests that this heterogeneity may stem from the quality of the included studies and the diverse chemical compositions among vaccine subtypes, the latter factor appears to have more significant negative implications in real-world settings. Epidemiological evidence demonstrates that older adults with chronic medical conditions are at an increased risk of developing symptomatic RSV-ARI and severe RSV disease. Additionally, these individuals are more likely to require medically attended visits and to experience progression to more severe disease [43-47]. In reality, our study did not examine the variations in health status among the enrolled older adults across the included studies owing to ambiguous family medical histories. This limitation hinders our ability to ascertain whether the statistical differences and heterogeneity in AEs are attributable to the underlying risk factors associated with subclinical or suboptimal health statuses. Despite previous studies demonstrating that one subtype of the RSV preF vaccine was well tolerated and exhibited an acceptable safety profile, the majority of solicited adverse events were transient and of mild to moderate severity [24,40]. Notwithstanding the enrollment criteria that limited participation to medically stable individuals, the overall positive benefit-risk profile of the RSV preF vaccine in older adults with comorbidities warrants further investigation [40].

On the basis of the subgroup analysis, clinical interventions were identified as a significant source of heterogeneity for both overall AEs and AEs associated with the chemical compositions in the vaccine. Previous studies have indicated that, among various types of RSV preF vaccines, subunit vaccines demonstrate significantly greater safety and a lower incidence of solicited injection-site and systemic adverse reactions [39]. To accurately characterize safety in mRNA-based RSV preF vaccines using original sources, the phase 2 or 3 randomized trial of Moderna mRNA-1345 (RSV preF mRNA) in adults aged ≥60 years reported serious AEs in 2.8% of participants in both the vaccine and placebo groups, with investigators noting no evident safety concerns and only a small number of serious AEs judged related to vaccination; thus, there was no excess of serious AEs associated with the mRNA-1345 vaccine compared with placebo [48,49]. Several solicited injection site AEs, including tenderness to touch and injection site pain, were reported more frequently in the mRNA-treated groups compared with the placebo group. However, most AEs experienced by mRNA-treated participants were transient and graded as grade 1 or 2 [50,51]. Additionally, myalgia, nausea, and chills were the most frequently reported symptoms among recipients of adenovirus-vectored vaccines. However, these symptoms were infrequently observed in our vaccination group [52]. These differences may potentially be attributed to variations in the magnitude of the immune response elicited by different vaccine platforms, a hypothesis that warrants confirmation through further immunological evaluations [53,54].

The potential for AEs among older adults could negatively influence public attitudes toward the RSV vaccine. Previous surveys have indicated that concerns about vaccine safety and side effects are the most common reasons for vaccine refusal [55]. Wang et al [55] demonstrated that older adults are more likely to have limited awareness and knowledge of RSV, which contributes to a higher propensity for vaccine refusal. Additionally, they found a negative correlation between advancing age and both vaccine confidence and vaccination rates [56,57]. This finding is consistent with the common reasons for refusing COVID-19 and monkeypox vaccines reported in previous studies [58,59]. Furthermore, administering the RSV vaccine concurrently with one or more additional vaccines during the same visit may increase the incidence of local or systemic reactogenicity [60]. Current studies have exclusively focused on the safety of the RSV preF vaccine when coadministered with various influenza vaccines in older adults. These studies have indicated that coadministration of the RSV preF vaccine with influenza vaccines results in an acceptable safety and tolerability profile [61-63]. However, data on the safety of the RSV preF vaccine in older adults when coadministered with other vaccines recommended for this age group—such as COVID-19, pneumococcal, adult tetanus, diphtheria, and pertussis vaccines, as well as the recombinant zoster vaccine—are lacking [60]. In other words, the prior vaccination history should also be considered a key factor contributing to variations within the vaccine group, necessitating more detailed clinical investigations.

Our study had several limitations. First, the number of studies included was limited, and the variations in vaccine types and doses across these studies may have led to heterogeneity in the results. This is particularly evident in the small number of experiments focusing on LRTIs and related ARIs, which prevented further subgroup analyses. Moreover, these factors may have impacted the overall study results and restricted more detailed analyses. Future large-scale studies are essential to further investigate and validate these findings. Most of the studies included in our analysis were conducted on a global scale, and the results may have been influenced by the ethnic and regional diversity inherent in these study populations. Therefore, it is necessary to conduct regional and ethnic

XSL•FO
RenderX

classification studies. Finally, it was challenging to accurately ascertain the health status of the survey participants at baseline. Any discrepancies in participant characteristics at baseline may have introduced heterogeneity into the results.

In conclusion, this meta-analysis provides strong evidence that RSV vaccination is a critical public health measure for older adults, offering substantial protection against RSV-related respiratory infections and a favorable safety profile, despite a higher frequency of mild-to-moderate AEs. The results showed that widespread RSV vaccination could significantly reduce the respiratory disease burden, lower health care use, and prevent thousands of hospitalizations and deaths annually. Consistent benefits across diverse populations and vaccine types support regulatory approval and inclusion in routine immunization programs for adults aged ≥60 years. Successful implementation will require addressing vaccine hesitancy with clear, evidence-based messaging, ensuring safe coadministration with other vaccines, and establishing surveillance systems to monitor real-world effectiveness and safety. Key research priorities include assessing long-term protection, optimal revaccination timing, evaluating performance in immunocompromised individuals, and analyzing cost-effectiveness for equitable resource use. As the global population ages and the burden of RSV increases, RSV vaccination should be considered not as an optional add-on but as an essential part of preventive care for older adults, with the potential to transform respiratory disease prevention.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to confidentiality restrictions but are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: XW and WL
Data curation: RY, QX, and YT
Formal analysis: RY
Funding acquisition: XW
Investigation: RY and QX
Methodology: QX and RY
Resources: WL and XW
Software: RY
Supervision: LZ
Validation: XW
Visualization: RY and QX
Writing—original draft: RY and QX
Writing—review and editing: XW.
All authors have reviewed and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Basic information and quality assessment of the included literature.
[DOCX File, 50 KB - publichealth_v11i1e74271_app1.docx]

Multimedia Appendix 2
Forest plots and sensitivity analyses (subgroup).
[DOCX File, 811 KB - publichealth_v11i1e74271_app2.docx]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File, 30 KB - publichealth_v11i1e74271_app3.docx]

References

1. Simoes EA. Respiratory syncytial virus infection. Lancet 1999 Sep 4;354(9181):847-852. [doi: 10.1016/S0140-6736(99)80040-3] [Medline: 10485741]
2. Ascough S, Paterson S, Chiu C. Induction and subversion of human protective immunity: contrasting influenza and respiratory syncytial virus. Front Immunol 2018;9:323. [doi: 10.3389/fimmu.2018.00323] [Medline: 29552008]

XSL•FO
RenderX

3. Hause AM, Henke DM, Avadhanula V, Shaw CA, Tapia LI, Piedra PA. Sequence variability of the respiratory syncytial virus (RSV) fusion gene among contemporary and historical genotypes of RSV/A and RSV/B. PLoS One 2017;12(4):e0175792. [doi: 10.1371/journal.pone.0175792] [Medline: 28414749]
4. Mas V, Nair H, Campbell H, Melero JA, Williams TC. Antigenic and sequence variability of the human respiratory syncytial virus F glycoprotein compared to related viruses in a comprehensive dataset. Vaccine (Auckl) 2018 Oct 29;36(45):6660-6673. [doi: 10.1016/j.vaccine.2018.09.056] [Medline: 30292456]
5. Russell CD, Unger SA, Walton M, Schwarze J. The human immune response to respiratory syncytial virus infection. Clin Microbiol Rev 2017 Apr;30(2):481-502. [doi: 10.1128/CMR.00090-16] [Medline: 28179378]
6. Falsey AR, Walsh EE. Respiratory syncytial virus infection in adults. Clin Microbiol Rev 2000 Jul;13(3):371-384. [doi: 10.1128/CMR.13.3.371] [Medline: 10885982]
7. Falsey AR, Walsh EE, House S, et al. Risk factors and medical resource utilization of respiratory syncytial virus, human metapneumovirus, and influenza-related hospitalizations in adults-a global study during the 2017-2019 epidemic seasons (Hospitalized Acute Respiratory Tract Infection [HARTI] study). Open Forum Infect Dis 2021 Nov;8(11):ofab491. [doi: 10.1093/ofid/ofab491] [Medline: 35559130]
8. Hartnett J, Donga P, Ispas G, et al. Risk factors and medical resource utilization in US adults hospitalized with influenza or respiratory syncytial virus in the Hospitalized Acute Respiratory Tract Infection study. Influenza Other Respir Viruses 2022 Sep;16(5):906-915. [doi: 10.1111/irv.12994] [Medline: 35474419]
9. Ackerson B, An J, Sy LS, Solano Z, Slezak J, Tseng HF. Cost of hospitalization associated with respiratory syncytial virus infection versus influenza infection in hospitalized older adults. J Infect Dis 2020 Aug 17;222(6):962-966. [doi: 10.1093/infdis/jiaa183] [Medline: 32300806]
10. Griffiths C, Drews SJ, Marchant DJ. Respiratory syncytial virus: infection, detection, and new options for prevention and treatment. Clin Microbiol Rev 2017 Jan;30(1):277-319. [doi: 10.1128/CMR.00010-16] [Medline: 27903593]
11. Falsey AR, Hennessey PA, Formica MA, Cox C, Walsh EE. Respiratory syncytial virus infection in elderly and high-risk adults. N Engl J Med 2005 Apr 28;352(17):1749-1759. [doi: 10.1056/NEJMoa043951] [Medline: 15858184]
12. Nam HH, Ison MG. Respiratory syncytial virus infection in adults. BMJ 2019 Sep 10;366:l5021. [doi: 10.1136/bmj.l5021] [Medline: 31506273]
13. Shi T, Denouel A, Tietjen AK, et al. Global disease burden estimates of respiratory syncytial virus-associated acute respiratory infection in older adults in 2015: a systematic review and meta-analysis. J Infect Dis 2020 Oct 7;222(Suppl 7):S577-S583. [doi: 10.1093/infdis/jiz059] [Medline: 30880339]
14. Kestler M, Muñoz P, Mateos M, Adrados D, Bouza E. Respiratory syncytial virus burden among adults during flu season: an underestimated pathology. J Hosp Infect 2018 Dec;100(4):463-468. [doi: 10.1016/j.jhin.2018.03.034] [Medline: 29614245]
15. Savic M, Penders Y, Shi T, Branche A, Pirçon JY. Respiratory syncytial virus disease burden in adults aged 60 years and older in high-income countries: a systematic literature review and meta-analysis. Influenza Other Respir Viruses 2023 Jan;17(1):e13031. [doi: 10.1111/irv.13031] [Medline: 36369772]
16. Schubert L, Steininger J, Lötsch F, et al. Surveillance of respiratory syncytial virus infections in adults, Austria, 2017 to 2019. Sci Rep 2021 Apr 26;11(1):8939. [doi: 10.1038/s41598-021-88537-5] [Medline: 33903713]
17. Cherukuri A, Patton K, Gasser RA Jr, et al. Adults 65 years old and older have reduced numbers of functional memory T cells to respiratory syncytial virus fusion protein. Clin Vaccine Immunol 2013 Feb;20(2):239-247. [doi: 10.1128/CVI.00580-12] [Medline: 23239796]
18. Falsey AR, Walsh EE. Relationship of serum antibody to risk of respiratory syncytial virus infection in elderly adults. J Infect Dis 1998 Feb;177(2):463-466. [doi: 10.1086/517376] [Medline: 9466538]
19. Malloy AMW, Falsey AR, Ruckwardt TJ. Consequences of immature and senescent immune responses for infection with respiratory syncytial virus. Curr Top Microbiol Immunol 2013;372:211-231. [doi: 10.1007/978-3-642-38919-1_11] [Medline: 24362692]
20. Piedra PA, Jewell AM, Cron SG, Atmar RL, Glezen WP. Correlates of immunity to respiratory syncytial virus (RSV) associated-hospitalization: establishment of minimum protective threshold levels of serum neutralizing antibodies. Vaccine (Auckl) 2003 Jul 28;21(24):3479-3482. [doi: 10.1016/s0264-410x(03)00355-4] [Medline: 12850364]
21. Walsh EE, Peterson DR, Falsey AR. Risk factors for severe respiratory syncytial virus infection in elderly persons. J Infect Dis 2004 Jan 15;189(2):233-238. [doi: 10.1086/380907] [Medline: 14722887]
22. Stephens LM, Varga SM. Considerations for a respiratory syncytial virus vaccine targeting an elderly population. Vaccines (Basel) 2021 Jun 9;9(6):624. [doi: 10.3390/vaccines9060624] [Medline: 34207770]
23. Graham BS, Modjarrad K, McLellan JS. Novel antigens for RSV vaccines. Curr Opin Immunol 2015 Aug;35:30-38. [doi: 10.1016/j.coi.2015.04.005] [Medline: 26070108]
24. Papi A, Ison MG, Langley JM, et al. Respiratory syncytial virus prefusion F protein vaccine in older adults. N Engl J Med 2023 Feb 16;388(7):595-608. [doi: 10.1056/NEJMoa2209604] [Medline: 36791160]
25. Walsh EE, Pérez Marc G, Zareba AM, et al. Efficacy and safety of a bivalent RSV prefusion F vaccine in older adults. N Engl J Med 2023 Apr 20;388(16):1465-1477. [doi: 10.1056/NEJMoa2213836] [Medline: 37018468]
26. Falsey AR, Williams K, Gymnopoulou E, et al. Efficacy and safety of an Ad26.RSV.preF-RSV preF protein vaccine in older adults. N Engl J Med 2023 Feb 16;388(7):609-620. [doi: 10.1056/NEJMoa2207566] [Medline: 36791161]

XSL•FO
RenderX

27. Walsh EE, Falsey AR, Scott DA, et al. A randomized phase 1/2 study of a respiratory syncytial virus prefusion F vaccine. J Infect Dis 2022 Apr 19;225(8):1357-1366. [doi: 10.1093/infdis/jiab612] [Medline: 34932102]
28. Falsey AR, Walsh EE, Scott DA, et al. Phase 1/2 randomized study of the immunogenicity, safety, and tolerability of a respiratory syncytial virus prefusion F vaccine in adults with concomitant inactivated influenza vaccine. J Infect Dis 2022 Jun 15;225(12):2056-2066. [doi: 10.1093/infdis/jiab611] [Medline: 34931667]
29. Crooke SN, Ovsyannikova IG, Poland GA, Kennedy RB. Immunosenescence and human vaccine immune responses. Immun Ageing 2019;16:25. [doi: 10.1186/s12979-019-0164-9] [Medline: 31528180]
30. Boraschi D, Italiani P. Immunosenescence and vaccine failure in the elderly: strategies for improving response. Immunol Lett 2014 Nov;162(1 Pt B):346-353. [doi: 10.1016/j.imlet.2014.06.006] [Medline: 24960535]
31. Saeland E, van der Fits L, Bolder R, et al. Combination Ad26.RSV.preF/preF protein vaccine induces superior protective immunity compared with individual vaccine components in preclinical models. NPJ Vaccines 2023 Mar 23;8(1):45. [doi: 10.1038/s41541-023-00637-7] [Medline: 36949051]
32. Williams K, Bastian AR, Feldman RA, et al. Phase 1 safety and immunogenicity study of a respiratory syncytial virus vaccine with an adenovirus 26 vector encoding prefusion F (Ad26.RSV.preF) in adults aged ≥60 years. J Infect Dis 2020 Aug 17;222(6):979-988. [doi: 10.1093/infdis/jiaa193] [Medline: 32320465]
33. Ison MG, Papi A, Athan E, et al. Efficacy and safety of respiratory syncytial virus (RSV) prefusion F protein vaccine (RSVPreF3 OA) in older adults over 2 RSV seasons. Clin Infect Dis 2024 Jun 14;78(6):1732-1744. [doi: 10.1093/cid/ciae010] [Medline: 38253338]
34. Wilson E, Goswami J, Baqui AH, et al. Efficacy and safety of an mRNA-based RSV PreF vaccine in older adults. N Engl J Med 2023 Dec 14;389(24):2233-2244. [doi: 10.1056/NEJMoa2307079] [Medline: 38091530]
35. Fries L, Shinde V, Stoddard JJ, et al. Immunogenicity and safety of a respiratory syncytial virus fusion protein (RSV F) nanoparticle vaccine in older adults. Immun Ageing 2017;14:8. [doi: 10.1186/s12979-017-0090-7] [Medline: 28413427]
36. A randomized, double-blind, placebo-controlled phase 3 efficacy study of an Ad26.RSV.preF-based vaccine in the prevention of lower respiratory tract disease caused by RSV in adults aged 60 years and older. Janssen Vaccines & Prevention B.V.. 2022. URL: https://cdn.clinicaltrials.gov/large-docs/83/NCT04908683/Prot_000.pdf [accessed 2025-11-14]
37. Falloon J, Yu J, Esser MT, et al. An adjuvanted, postfusion f protein-based vaccine did not prevent respiratory syncytial virus illness in older adults. J Infect Dis 2017 Dec 12;216(11):1362-1370. [doi: 10.1093/infdis/jix503] [Medline: 29029260]
38. Ma J, Chen L, Tang S, Shi Y. Efficacy and safety of respiratory syncytial virus vaccination during pregnancy to prevent lower respiratory tract illness in newborns and infants: a systematic review and meta-analysis of randomized controlled trials. Front Pediatr 2023;11:1260740. [doi: 10.3389/fped.2023.1260740] [Medline: 38357264]
39. Pang Y, Lu H, Cao D, et al. Efficacy, immunogenicity and safety of respiratory syncytial virus prefusion F vaccine: systematic review and meta-analysis. BMC Public Health 2024 May 6;24(1):1244. [doi: 10.1186/s12889-024-18748-8] [Medline: 38711074]
40. Feldman RG, Antonelli-Incalzi R, Steenackers K, et al. Respiratory syncytial virus prefusion F protein vaccine is efficacious in older adults with underlying medical conditions. Clin Infect Dis 2024 Jan 25;78(1):202-209. [doi: 10.1093/cid/ciad471] [Medline: 37698366]
41. Divo MJ, Martinez CH, Mannino DM. Ageing and the epidemiology of multimorbidity. Eur Respir J 2014 Oct;44(4):1055-1068. [doi: 10.1183/09031936.00059814] [Medline: 25142482]
42. Prince MJ, Wu F, Guo Y, et al. The burden of disease in older people and implications for health policy and practice. Lancet 2015 Feb 7;385(9967):549-562. [doi: 10.1016/S0140-6736(14)61347-7] [Medline: 25468153]
43. McLaughlin JM, Khan F, Begier E, Swerdlow DL, Jodar L, Falsey AR. Rates of medically attended RSV among US adults: a systematic review and meta-analysis. Open Forum Infect Dis 2022 Jul;9(7):ofac300. [doi: 10.1093/ofid/ofac300] [Medline: 35873302]
44. Beran J, Ramirez Villaescusa A, Devadiga R, et al. Respiratory syncytial virus acute respiratory infections in ≥ 65-year-old adults in long-term care facilities in the Czech Republic. Cent Eur J Public Health 2021 Sep;29(3):167-176. [doi: 10.21101/cejph.a6861] [Medline: 34623114]
45. Belongia EA, King JP, Kieke BA, et al. Clinical features, severity, and incidence of RSV illness during 12 consecutive seasons in a community cohort of adults ≥60 years old. Open Forum Infect Dis 2018 Dec;5(12):ofy316. [doi: 10.1093/ofid/ofy316] [Medline: 30619907]
46. Wyffels V, Kariburyo F, Gavart S, Fleischhackl R, Yuce H. A real-world analysis of patient characteristics and predictors of hospitalization among US medicare beneficiaries with respiratory syncytial virus infection. Adv Ther 2020 Mar;37(3):1203-1217. [doi: 10.1007/s12325-020-01230-3] [Medline: 32026380]
47. Ivey KS, Edwards KM, Talbot HK. Respiratory syncytial virus and associations with cardiovascular disease in adults. J Am Coll Cardiol 2018 Apr 10;71(14):1574-1583. [doi: 10.1016/j.jacc.2018.02.013] [Medline: 29622165]
48. Al Khames Aga QA, Alkhaffaf WH, Hatem TH, et al. Safety of COVID-19 vaccines. J Med Virol 2021 Dec;93(12):6588-6594. [doi: 10.1002/jmv.27214] [Medline: 34270094]
49. Ling Y, Zhong J, Luo J. Safety and effectiveness of SARS-CoV-2 vaccines: a systematic review and meta-analysis. J Med Virol 2021 Dec;93(12):6486-6495. [doi: 10.1002/jmv.27203] [Medline: 34264528]

XSL•FO
RenderX

50. Jackson LA, Anderson EJ, Rouphael NG, et al. An mRNA vaccine against SARS-CoV-2 - preliminary report. N Engl J Med 2020 Nov 12;383(20):1920-1931. [doi: 10.1056/NEJMoa2022483] [Medline: 32663912]
51. Feldman RA, Fuhr R, Smolenov I, et al. mRNA vaccines against H10N8 and H7N9 influenza viruses of pandemic potential are immunogenic and well tolerated in healthy adults in phase 1 randomized clinical trials. Vaccine (Auckl) 2019 May 31;37(25):3326-3334. [doi: 10.1016/j.vaccine.2019.04.074] [Medline: 31079849]
52. McDonald I, Murray SM, Reynolds CJ, Altmann DM, Boyton RJ. Comparative systematic review and meta-analysis of reactogenicity, immunogenicity and efficacy of vaccines against SARS-CoV-2. NPJ Vaccines 2021 May 13;6(1):74. [doi: 10.1038/s41541-021-00336-1] [Medline: 33986272]
53. Fang E, Liu X, Li M, et al. Advances in COVID-19 mRNA vaccine development. Signal Transduct Target Ther 2022 Mar 23;7(1):94. [doi: 10.1038/s41392-022-00950-y] [Medline: 35322018]
54. Pollard AJ, Bijker EM. A guide to vaccinology: from basic principles to new developments. Nat Rev Immunol 2021 Feb;21(2):83-100. [doi: 10.1038/s41577-020-00479-7] [Medline: 33353987]
55. Wang Q, Xiu S, Yang L, et al. Perceptions about respiratory syncytial virus (RSV) and attitudes toward the RSV vaccine among the general public in China: a cross-sectional survey. Hum Vaccin Immunother 2024 Dec 31;20(1):2310916. [doi: 10.1080/21645515.2024.2310916] [Medline: 38369712]
56. de Figueiredo A, Simas C, Karafillakis E, Paterson P, Larson HJ. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. Lancet 2020 Sep 26;396(10255):898-908. [doi: 10.1016/S0140-6736(20)31558-0] [Medline: 32919524]
57. Wang G, Yao Y, Wang Y, et al. Determinants of COVID-19 vaccination status and hesitancy among older adults in China. Nat Med 2023 Mar;29(3):623-631. [doi: 10.1038/s41591-023-02241-7] [Medline: 36720270]
58. Wang B, Peng X, Li Y, et al. Perceptions, precautions, and vaccine acceptance related to monkeypox in the public in China: a cross-sectional survey. J Infect Public Health 2023 Feb;16(2):163-170. [doi: 10.1016/j.jiph.2022.12.010] [Medline: 36535136]
59. Wang Q, Yang L, Jin H, Lin L. Vaccination against COVID-19: a systematic review and meta-analysis of acceptability and its predictors. Prev Med 2021 Sep;150:106694. [doi: 10.1016/j.ypmed.2021.106694] [Medline: 34171345]
60. Melgar M, Britton A, Roper LE, et al. Use of respiratory syncytial virus vaccines in older adults: recommendations of the advisory committee on immunization practices - United States, 2023. MMWR Morb Mortal Wkly Rep 2023 Jul 21;72(29):793-801. [doi: 10.15585/mmwr.mm7229a4] [Medline: 37471262]
61. Athan E, Baber J, Quan K, et al. Safety and immunogenicity of bivalent RSVpreF vaccine coadministered with seasonal inactivated influenza vaccine in older adults. Clin Infect Dis 2024 May 15;78(5):1360-1368. [doi: 10.1093/cid/ciad707] [Medline: 37992000]
62. Widagdo W, Bastian AR, Jastorff AM, et al. Concomitant administration of Ad26.RSV.preF/RSV preF protein vaccine and high-dose influenza vaccine in adults 65 years and older: a noninferiority trial. J Infect Dis 2024 Aug 16;230(2):e374-e383. [doi: 10.1093/infdis/jiad594] [Medline: 38134393]
63. Chandler R, Montenegro N, Llorach C, et al. Immunogenicity, reactogenicity, and safety of AS01E-adjuvanted RSV prefusion F protein-based candidate vaccine (RSVPreF3 OA) when co-administered with a seasonal quadrivalent influenza vaccine in older adults: results of a phase 3, open-label, randomized controlled trial. Clin Infect Dis 2024 Jan 8:ciad786. [doi: 10.1093/cid/ciad786] [Medline: 38189778]

Abbreviations

AE: adverse event
ARI: acute respiratory infection
LRTI: lower respiratory tract infections
preF: prefusion F
RCT: randomized controlled trial
RSV: respiratory syncytial virus
RSV-ARI: respiratory syncytial virus-associated acute respiratory illness
SAE: serious adverse event

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 21.03.25; peer-reviewed by MM Lino, M Raemet; revised version received 29.09.25; accepted 20.10.25; published 04.12.25.

Please cite as:
Xiao Q, Yang R, Zhang L, Tian Y, Wang X, Li W
Safety and Efficacy of Respiratory Syncytial Virus Vaccination in Older Adults: Systematic Review and Meta-Analysis of Randomized Controlled Trials
JMIR Public Health Surveill 2025;11:e74271
URL: https://publichealth.jmir.org/2025/1/e74271
doi:10.2196/74271

© Qingzhi Xiao, Ruizhe Yang, Lei Zhang, Ye Tian, Xu Wang, Wei Li. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 4.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Minimum Data Set and Metadata for Active Vaccine Safety Surveillance: Systematic Review

Mengdi Zhang[1,2], MD; Junting Yang[1,2], PhD; Yan Li[3], PhD; Yuan Li[3], MD; Tong Li[1,2], MD; Ziqi Dong[1,2], MD; Shuo Gong[1,2], MD; Yahui Wu[1,2], MD; Minrui Ren[3], MD; Chunxiang Fan[3], MD; Lina Zhang[3], MD; Yi Wang[4], MD; Yali Wang[4], PhD; Jingtian Ren[4], PhD; Feng Sun[1,2], PhD; Chuanyong Shen[4], PhD; Keli Li[3*], MD; Zhike Liu[1,2*], PhD; Siyan Zhan[1,2,5,6*], PhD

[1]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, 38 Xueyuan Road, Haidian District, Beijing, China

[2]Key Laboratory of Epidemiology of Major Diseases, Peking University, Ministry of Education, Beijing, China

[3]National Immunization Program, Chinese Center for Disease Control and Prevention, Beijing, China

[4]Center for Drug Reevaluation, National center for ADR Monitoring, Beijing, China

[5]Center for Intelligent Public Health, Institute for Artificial Intelligence, Peking University, Beijing, China

[6]Research Center of Clinical Epidemiology, Peking University Third Hospital, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Siyan Zhan, PhD
Department of Epidemiology and Biostatistics, School of Public Health, Peking University, 38 Xueyuan Road, Haidian District, Beijing, China

Abstract

Background: Active vaccine safety surveillance (AVSS) stands as a top priority for the World Health Organization (WHO), serving as a critical indicator of the fourth maturity level for national regulatory agencies.

Objective: This review aims to define the minimal data scope for association studies in vaccine safety, providing a reference framework for implementing AVSS systems worldwide, especially in low- and middle-income countries.

Methods: The study systematically searched PubMed, Embase, and Web of Science for cohort and case-control studies related to AVSS published between January 1, 2018, and September 7, 2022. Guided by the WHO and Council for International Organizations of Medical Sciences guidelines (CIOMS), we developed a 4D framework for Minimum Data Sets (MDSs), including "Vaccine," "Outcome," "Demographic Data," and "Covariate." Variables with a frequency of at least 5% were included in the MDS.

Results: Of the 123 included studies, 102 (82.9%) were cohort studies and 98 (79.7%) originated from high-income countries, covering populations across the entire life course. The MDS for COVID-19 vaccines identified 54 variables, while the MDS for maternal populations included 96 variables. WHO guidelines were found to align better with practical applications compared to CIOMS guidelines, though both require further optimization based on the MDS findings. However, metadata for these essential variables were inadequately described across the studies.

Conclusions: The proposed MDS provides clear guidance and concise requirements for AVSS data scope. Establishing a globally standardized MDS and comprehensive metadata based on these findings is essential to enhancing the global vaccine safety ecosystem.

Trial Registration: PROSPERO CRD42023449920; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023449920

(JMIR Public Health Surveill 2025;11:e63161) doi:10.2196/63161

KEYWORDS

vaccine safety; active surveillance; minimum data sets; metadata; association

Introduction

Vaccines are one of the most cost-effective interventions for preventing and controlling infectious diseases, and vaccination could reduce 2 - 3 million deaths annually, with an additional 1.5 million deaths that could be averted if the coverage expanded [1]. However, because vaccines are administered to healthy individuals, concerns about their safety can significantly impact public trust, leading to vaccine hesitancy. This hesitancy poses serious threats to global health, such as the resurgence of nearly eradicated diseases and challenges in achieving the Immunization Agenda 2030 (IA2030) goals [2]. To address

XSL•FO
RenderX

these challenges, robust systems for monitoring vaccine safety are essential. These systems must quickly and accurately evaluate vaccine risks and benefits while providing timely responses to safety concerns, which not only strengthens public confidence but also facilitates effective communication during vaccine rollouts [3,4]. Regulatory frameworks have increasingly recognized the importance of such measures. For instance, the US Food and Drug Administration (FDA) has mandated life-cycle safety management for vaccines since 2007, while China's Vaccine Administration Law enforces stringent safety requirements from development through postmarket surveillance [5,6]. From the perspective of evidence and monitoring approaches, the premarket clinical trials and postmarket surveillance should be essential measurements to ensure vaccine safety. While clinical trials are critical for assessing safety and efficacy before approval, proactive monitoring systems are indispensable for identifying rare or unforeseen safety issues that may not emerge during trials [7]. Active surveillance enables the timely detection of potential safety signals and provides robust evidence of vaccine benefits, complementing the limitations of passive surveillance [8].

Over the last decade, an increasing number of countries, particularly those in resource-rich regions, are attaching great importance to and accelerating the progress of active vaccine safety surveillance (AVSS) systems [9,10]. Several exemplary models were established and provided high-quality evidence for the critical decision-making of safety in immunization programs [11-16]. There are three types of data collection approaches and different functions among them. First, the Vaccine Safety Datalink (VSD) project collaborated with the Centers for Disease Control and Prevention in the United States and integrated a whole life-course cohort through data linkage, which covered 18 million individuals and included pregnant women, infants, children, and older people [17,18]. Second, the national pediatric hospital surveillance networks were established to enhance routine capacity and bridge the research gap to measure the impact and safety of vaccine programs [19]. Third, AusVaxSafety actively collects adverse events from national immunization sentinel sites through SMS text messaging or email reports [20]. In brief, these systems have gone through many years of investment and construction and now are stable, sustainable, and fast-response due to their large-scale and representative sample size.

The World Health Organization (WHO) has recently been urging and endeavoring to establish a global vaccine safety surveillance ecosystem that highlights the capacity of proactive vigilance in vaccine safety [21]. However, few studies were from low- and middle-income countries (LMICs) due to challenges in securing the necessary funding, technology, and human resources for these systems.

The use of electronic health data for AVSS requires addressing three key questions: (1) What information will researchers and policymakers need in the future for AVSS? (2) What data sources can identify this information? (3) What are the minimum data elements required? The Minimum Data Set (MDS) is a structured collection of essential data elements used for clinical and research purposes, facilitating standardized and efficient data collection. Many countries have adopted MDS frameworks to ensure data comparability, enhance public health decision-making, and support real-time safety monitoring [22]. In the context of AVSS, a well-defined MDS can help identify critical surveillance indicators, ensure data accessibility, and improve risk assessment methodologies [23]. Furthermore, metadata, which provides descriptive information about datasets—plays a key role in standardizing data classification, integration, and analysis across different surveillance systems [24]. Therefore, establishing a standardized MDS and metadata system is essential for optimizing AVSS, particularly in resource-limited settings.

Currently, 2 guidelines provide frameworks for data elements, approaches, and study designs of AVSS, but inconsistencies between these guidelines create challenges in standardizing data collection and usage [25,26]. Furthermore, no study has systematically evaluated or validated the essential data elements proposed by these frameworks, leaving their relevance and applicability uncertain. Robust AVSS systems must rely on credible and essential data to function effectively. To address this gap, our systematic review examines the data elements and metadata requirements for AVSS by synthesizing findings from association studies in vaccine safety. This review aims to define an MDS that ensures efficient AVSS implementation, particularly in resource-limited settings [27].

Methods

Search Strategy and Selection Criteria

We systematically searched 3 databases, PubMed, Embase, and Web of Science, from January 1, 2018, to September 7, 2022, to identify all the cohort studies and case-control studies of AVSS. The search strategy was registered on the PROSPERO (International Prospective Register of Systematic Reviews) database (CRD42023449920). The details are shown in Supplementary Materials (Multimedia Appendix 1).

There were two stages in the literature screening. In the first step, we selected studies with descriptions of AVSS or systems and excluded: (1) basic experiments (in vitro and animal) and basic research (pathology); (2) vaccine management and economics studies, such as policy, health economics, and sociomedical assessment, quality of life; (3) case reports, case series studies, nursing experiences, which primarily described the characteristics of individual patients or small groups; (4) vaccine effectiveness evaluation; (5) studies not in English; (6) studies belonging to other categories; and (7) uncertain studies, requiring full-text reading for clarification in the next step.

During the secondary step, we included only postmarketing association studies, such as cohort studies and case-control studies that were conducted with control groups from the same source population, regardless of whether databases were used. The following publications were excluded: (1) passive surveillance, (2) clinical trials, (3) theoretical or methodological studies for vaccine safety surveillance, (4) studies not classified as association studies, including correlational studies, signaling reports, descriptive studies, or those lacking a comparison group, (5) studies not related to vaccine intrinsic safety, (6) nonoriginal studies, such as reviews, conference abstracts, correspondence,

XSL•FO
RenderX

errata, or protocols, (7) studies not in English, or (8) studies without full text. Table 1 presents our inclusion and exclusion criteria.

Table . Inclusion or exclusion criteria.

Category	Inclusion criteria	Exclusion criteria
Article type	• Active surveillance • Original studies	• Passive surveillance • Nonoriginal studies (reviews, abstracts, etc)
Study type	• Association studies (cohort and case-control studies)	• Basic experiments (in vitro and animal) and basic research (pathology) • Vaccine management and economics studies • Case reports, case series studies, and nursing experiences • Vaccine effectiveness evaluation • Clinical trials • Theoretical or methodological studies on vaccine safety surveillance • Studies not classified as association studies • Studies unrelated to vaccine intrinsic safety
Language	• English	• Non-English studies

For the above steps, 2 trained researchers independently reviewed the title and abstract of each paper, if any disagreement occurred, a senior researcher would participate in the discussion and reach a final agreement.

Data Analysis

To formulate a standard, reasonable, and validated MDS for AVSS, 3 stages were undertaken to ensure its scientific and practical validity. First, a preliminary framework for the MDS was developed by referencing relevant variables from the Council for International Organizations of Medical Sciences (CIOMS) Guide to Active Vaccine Safety Surveillance and the *COVID-19 Vaccines: Safety Surveillance Manual* [25,26]. This initial framework comprised four dimensions: “Vaccine,” “Outcome,” “Demographic Data,” and “Covariate.” In addition, metadata including data sources, encoding, quality control, and assurance were also taken into consideration. Second, to improve the rationality and feasibility of the drafted framework, experts from diverse fields including immunization, epidemiology, clinical medicine, and databases were consulted to review the framework. A total of 30 eligible studies were randomly selected to refine the framework. Third, a standardized questionnaire for information extraction was developed based on the validated framework. Relevant variables were extracted from the methodology, results, and limitations sections of each study. During this process, any overlooked elements were identified and supplemented through discussion to finalize the framework. Furthermore, basic characteristics of the included studies (authors, publication year, country or region, study design, and sample size) were also collected.

The data extraction and management were completed parallelly by 2 trained researchers using Microsoft Excel 2019, with oversight from a senior researcher responsible for quality control. Subsequently, descriptive statistical analysis was performed using R 4.3.2 (R Core Team), and data visualization was carried out using both R 4.3.2 and OriginPro (Origin Laboratories) 2021.

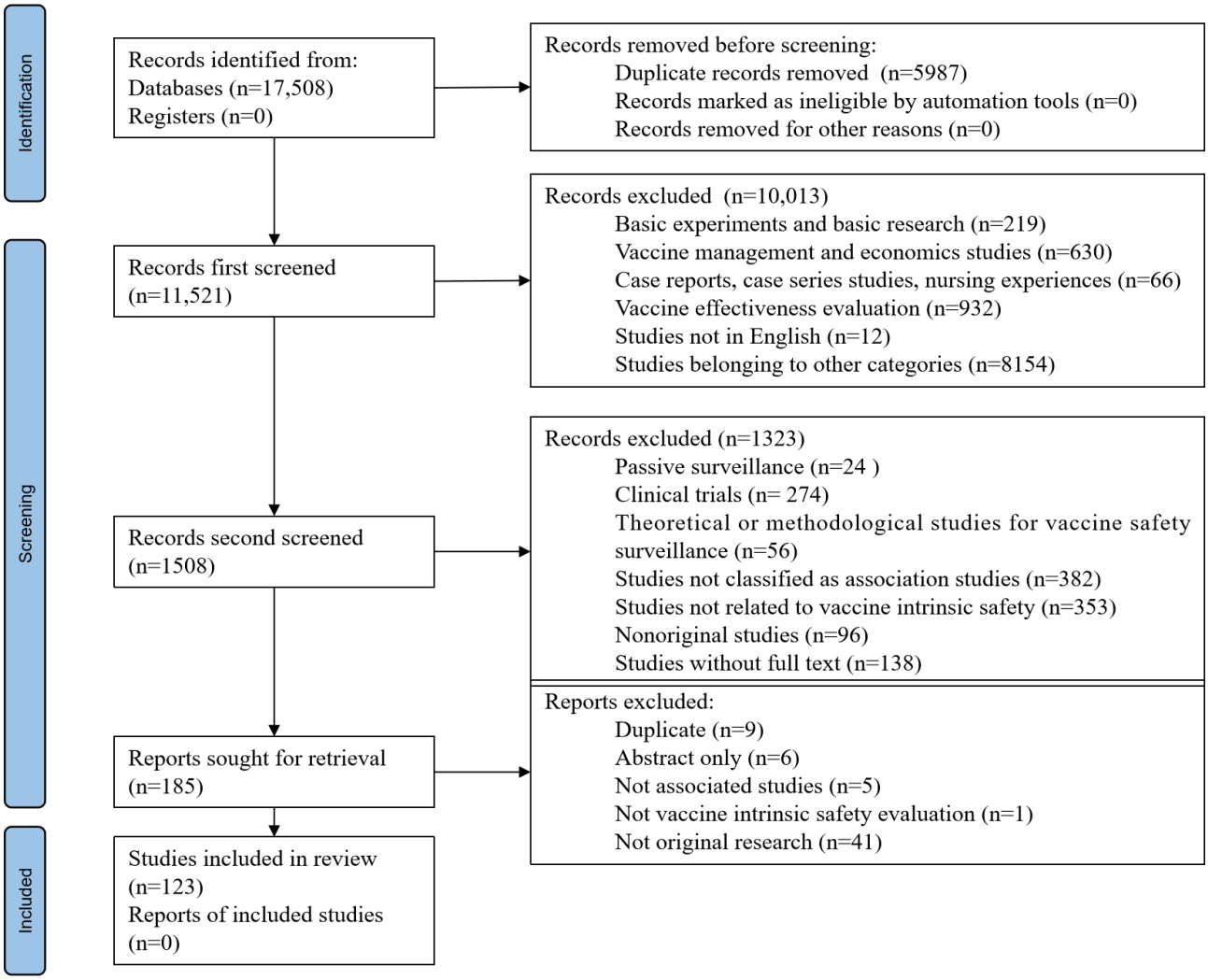
In the study, the key steps from study design, literature searching, framework formulation, information extraction, and analysis to interpretation have been supervised by external experts to strictly ensure high quality. Meanwhile, as for literature selection and data extraction, every participant was trained and qualified, and the information was double entry independently and parallelly, and any disagreement was ruled out by another senior researcher.

Results

Initially, 11,512 papers were identified after removing duplicated records. With 2 rounds of literature screening, 185 papers were evaluated in full text. Among them, 9 papers were excluded for duplication, 5 papers were excluded for not meeting the criteria for association studies, 1 paper was excluded due to its failure to evaluate vaccine intrinsic safety, 41 papers were excluded for not being original research, and 6 papers were excluded due to the lack of full-text availability. A total of 123 studies [14,15,17,28-147] were finally included. The inclusion and exclusion process is shown in Figure 1.

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic reviews and Meta-Analyses flow diagram.



The general characteristics of eligible studies are shown in Table 2, with more specific details provided in Table S1 in Multimedia Appendix 1. Of the eligible studies, 82.9% (102/123) were cohort studies, while case-control studies accounted for 17.1% (21/123). Regarding vaccines, 39.5% (49/124) investigated the association between the COVID-19 vaccine and adverse events, while 20.2% (25/124) focused on the influenza vaccine. As for the population, 20% (25/125) focused on children or adolescents, 14.4% (18/125) on pregnant women, 10.40% (13/125) on infants, and 4.8% (6/125) on the older people. In addition, 16.0% (20/125) of studies explored vaccine safety in patients with special diseases such as cancer and inflammatory bowel disease.

XSL•FO
RenderX

Table . General characteristics of selected studies.

Characteristics	Values, n (%)
Country or area	
High-income countries	98 (79.7)
Low- and middle-income countries	16 (13)
Multicountry settings	9 (7.3)
Study design	
Cohort study	102 (82.9)
Case-control study	21 (17.1)
Vaccine[a]	
COVID-19 vaccine	49 (39.5)
Influenza vaccine	25 (20.2)
Tdap[b]	9 (7.3)
HPV[c]	6 (4.8)
Zoster Vaccine	5 (4)
Others	30 (24.2)
Study population[a]	
General population	43 (34.4)
Children and Adolescents	25 (20)
Special patients[d]	20 (16)
Maternal	18 (14.4)
Infants	13 (10.4)
Older people	6 (4.8)

[a]Some studies involved multiple populations, multivaccines, or multioutcomes.

[b]Tdap: tetanus, diphtheria, acellular pertussis vaccine .

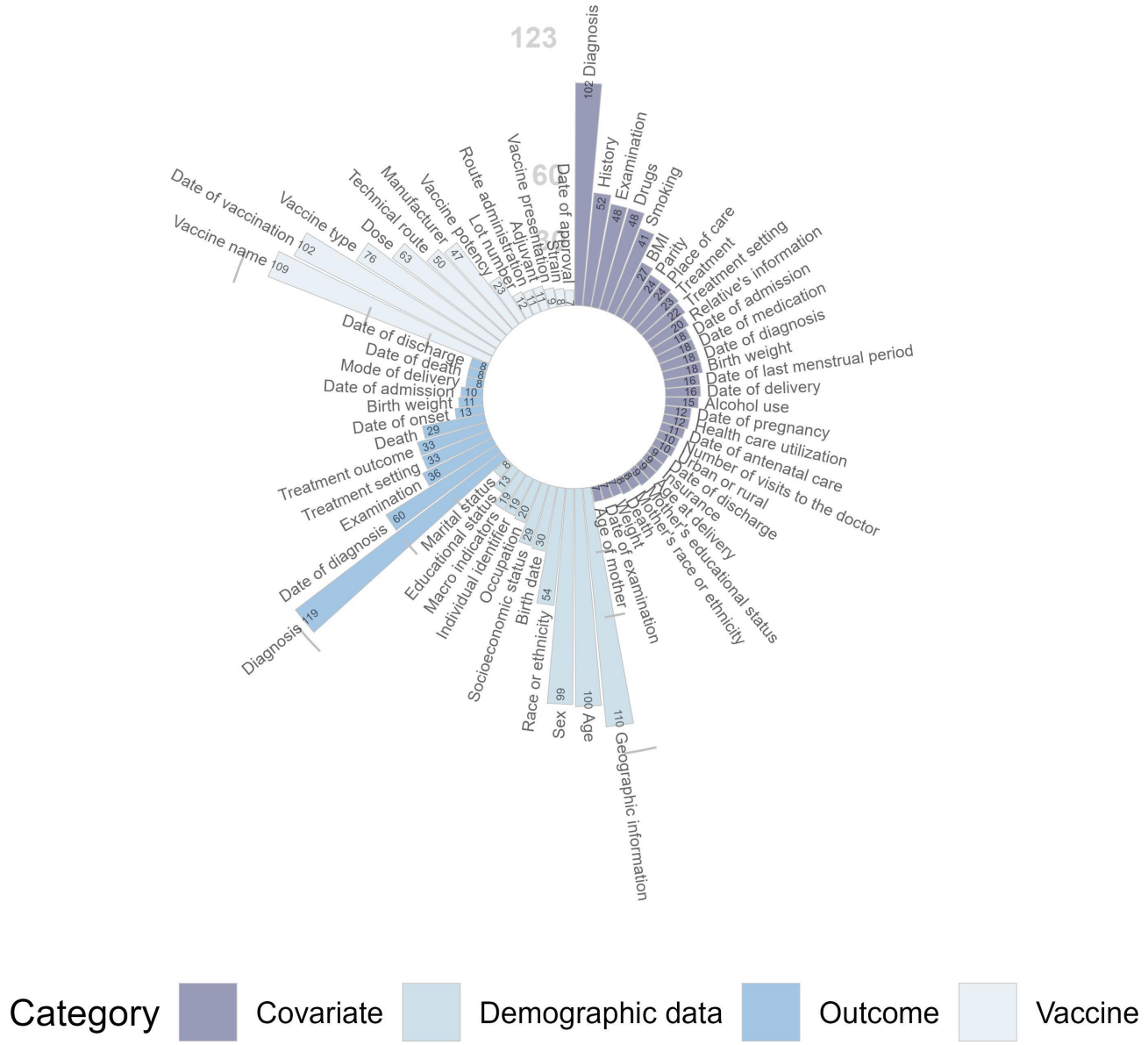
[c]HPV: human papillomavirus vaccine.

[d]Patients with inflammatory bowel disease, cancer, severe influenza, juvenile idiopathic arthritis, asthma, or other diseases.

Figure 2 illustrates the utilization of the summarized MDS across 123 included studies. Among the categories, “Diagnosis” within the “Outcome” and “Covariate” dimensions appeared most frequently, with 119 occurrences (96.8%) and 102 occurrences (82.9%), respectively. In the “Demographic Data” dimension, “Geographic Information” was the most frequently reported variable, appearing in 110 studies (89.4%). Similarly, “Vaccine Name” was the most commonly used variable in the “Vaccine” dimension, with 109 occurrences (88.6%). Table 3 shows the differences between the CIOMS and WHO guidelines and the application of both in the included studies. The MDS includes almost all variables recommended in the CIOMS and WHO guidelines, along with some additional variables. These extra variables include “Birth Weight,” “Examination,” and “Mode of Delivery” under the “Outcome” category; “Geographic Information” and “Race or Ethnicity” under the “Demographic Data” category; “Technical Route” and “Adjuvant” under the “Vaccine” category; and several additional variables within the “Covariate” category. The details for MDS are shown in Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 2. Frequency of Minimum Data Set variables.



XSL•FO
RenderX

Table . Comparison among the Council for International Organizations of Medical Sciences guidelines, and the World Health Organization guidelines.

Variables	CIOMS[a] guidelines	WHO[b] guidelines	Value, n (%)
Vaccine data			
Place of vaccination	✓	✓	1 (0.8)
Vaccine type	✓	✓[c]	77 (62.6)
Vaccine presentation, single or multiple dose	✓		9 (7.3)
Manufacturer	✓		47 (38.2)
Lot number (of vaccine and any diluents)	✓	✓[c]	12 (9.8)
Date of vaccination (and perhaps time)	✓	✓[c]	102 (82.9)
Vaccine injection site	✓	✓[c]	5 (4.1)
Number of dose	✓	✓[c]	63 (51.2)
Vaccine antigens		✓	7 (5.7)
Concomitant vaccines		✓	5 (4.1)
Route administration		✓	11 (8.9)
Health events or outcomes data			
Place of care	✓	✓	1 (0.8)
Diagnosis(es) or adverse event(s) or outcome	✓	✓[c]	119 (96.8)
Date (and time) of onset of (first) symptom of the event	✓	✓[c]	13 (10.6)
Serious		✓[c]	15 (12.2)
Demographic data			
Age at onset		✓[c]	100 (81.3)
Gender		✓[c]	99 (80.5)
Medical conditions		✓[c]	52 (42.3)
Medication		[c]✓	48 (39)

[a]CIOMS: Council for International Organizations of Medical Sciences

[b]WHO: World Health Organization

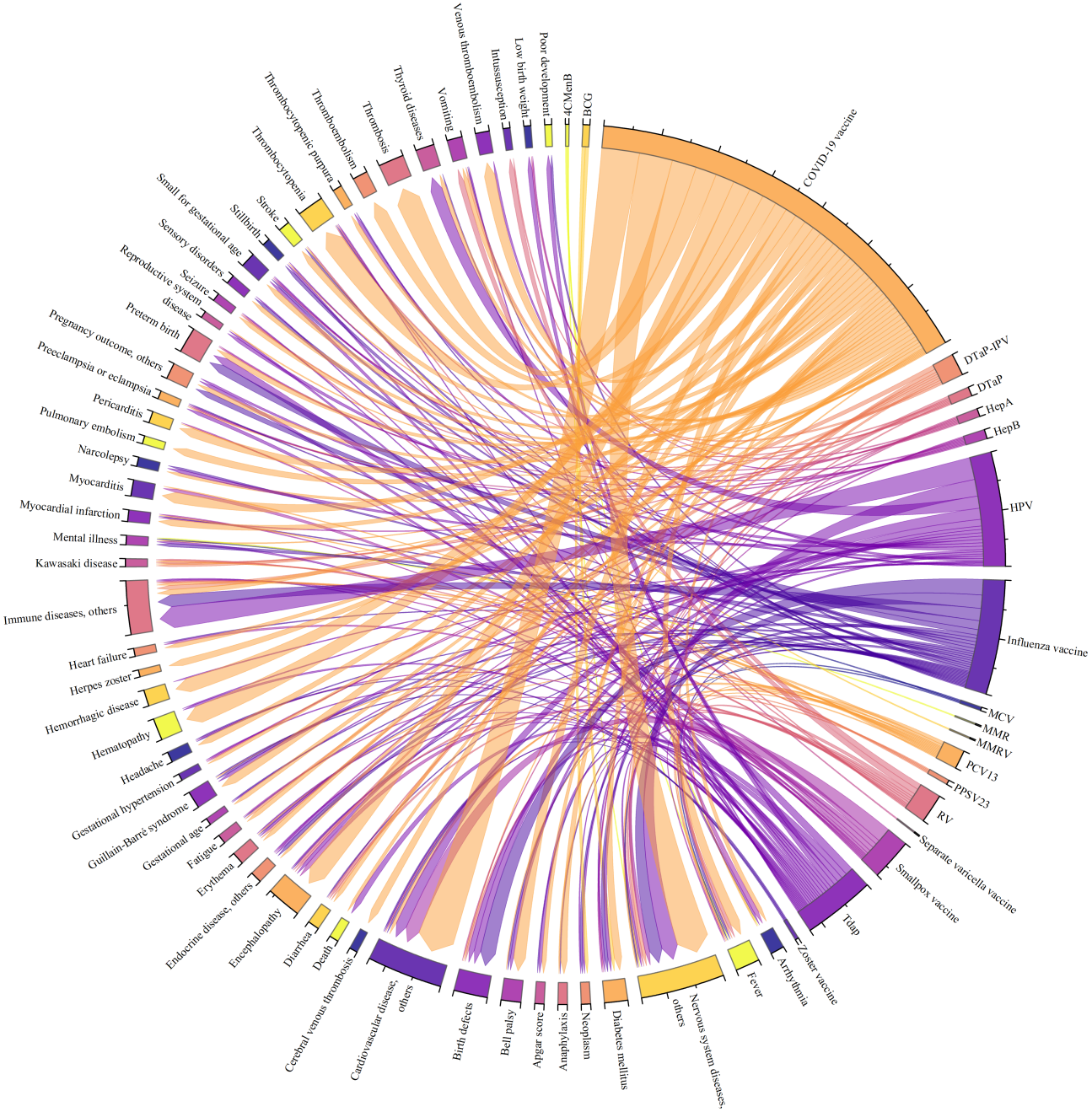
[c]The core dataset of the WHO guidelines.

Among the 49 studies focusing on COVID-19 vaccines, a total of 95 variables were identified, of which 54 variables appeared at least 3 times (frequency >5%) and were included in the MDS for COVID-19 vaccines. These variables were distributed as follows: 14 under the “Outcome” category, 10 under “Demographic Data,” 9 under “Vaccine,” and 21 under “Covariate.” In 18 studies targeting the pregnant population, 96 variables were identified, all of which were included in the MDS for this specific population (frequency >5%). These variables were categorized as follows: 20 under “Outcome,” 11 under “Demographic Data,” 11 under “Vaccine,” and 54 under “Covariate.” More details are shown in Tables S3 and S4 in Multimedia Appendix 1.

Figure 3 plotted the evaluation of the relationship between adverse events and vaccines. We only included diseases that appear more than five times, specifically cardiovascular disease, endocrine diseases, hemorrhagic diseases, nervous system diseases, mental illness, and immune diseases, and we consider the occurrences of outcomes to be counted twice if a publication studies two vaccines. A total of 1239 records specifying health outcomes and vaccine types were identified across 116 studies. The top 5 rare or severe adverse events, each occurring more than 20 times, were encephalopathy (n=30), birth defects (n=29), preterm birth (n=22), thrombocytopenia (n=22), and thrombosis (n=22). Adverse events with fewer than 5 (1%) occurrences were relatively concentrated in 3 groups of diseases: nervous system (n=74), cardiovascular system (n=65), and immune system (n=46). As for COVID-19 vaccines, the top 3 adverse events pairs were thrombosis (n=22), thrombocytopenia (n=19), and encephalopathy (n=17).

XSL•FO
RenderX

Figure 3. Frequency of "vaccine-outcome" pairs in included studies. 4CMenB: 4-Component group B meningococcal vaccine; BCG: Bacillus Calmette-Guerin vaccine; DTaP-IPV: Diphtheria, tetanus, acellular pertussis, and inactivated poliovirus combination vaccine; DTaP: Diphtheria, tetanus, acellular pertussis vaccine; HepA: Hepatitis A vaccine; HepB: Hepatitis B vaccine; HPV: Human papillomavirus vaccine; MCV: Measles and pertussis-containing vaccine; MMR: Measles, mumps, rubella vaccine; MMRV: Measles, mumps, rubella and varicella vaccine; PCV13: 13-Valent pneumococcal conjugate vaccine; PPSV-23: 23-Valent pneumococcal polysaccharide vaccine; RV: Rotavirus vaccine; Tdap: Diphtheria, tetanus, acellular pertussis vaccine.



A total of 17 studies were completed using the AVSS system, with 11 using the VSD project and the remaining 6 derived from various sources, including the Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC), Pharmacoepidemiological Research in Primary Care (Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria; BIFAP), Norway's emergency preparedness register for COVID-19 (Beredskapregisteret for COVID-19; Beredt C19), Post-Licensure Rapid Immunization Safety Monitoring (PRISM)–Sentinel, the Paediatric Active Enhanced Disease Surveillance network, and the Defense Medical Surveillance System. All these studies explicitly reported the standard codes for outcomes or events. Specifically, VSD used National Drug Codes and *International Classification of Diseases* (*ICD*) codes, PRISM-Sentinel used Current Procedural Terminology, BIFAP and Beredt C19 used *ICD* codes, and RCGP RSC applied 5-byte Read or Clinical Terms Version 3 codes. However, only 8 studies (9.3%) documented their quality control processes. Among these, 3 studies validated diagnoses in the database by assessing the positive predictive value, 2 confirmed data reliability by comparing original records with database entries, 2 used high-quality data that had undergone prior evaluation, and one maintained data quality through regular monitoring.

XSL•FO
RenderX

Discussion

Principal Findings

This study is the first to systematically clarify the minimum data requirements for AVSS by conducting a comprehensive review aligned with WHO and CIOMS guidelines, focusing on minimizing data range. The proposed MDS and Metadata framework for AVSS encompasses four key dimensions: "Vaccine," "Outcome," "Demographic Data," and "Covariate." Across these dimensions, 68 variables were identified: 13 in "Vaccine," 12 in "Outcome," 11 in "Demographic Data," and 32 in "Covariate." Certain variables were deemed essential for all vaccine safety association studies, including unique and anonymized individual identifiers; vaccine name and vaccination date from the "Vaccine" category; diagnosis and diagnostic date from the "Outcome" category; age and sex from the "Demographic Data" category; and history and drugs from the "Covariate" category. These aspects of data collection, quality control, and assurance must be transparently declared in AVSS protocols; however, regrettably, only a minority of studies addressed them adequately. Therefore, we strongly advocate for the establishment of a globally standardized and widely recognized MDS and Metadata framework. Such a standard would enable countries and regions, particularly LMICs, to rapidly develop AVSS systems or conduct related research. In addition to facilitating data comparability, this effort would significantly strengthen the global vaccine safety ecosystem.

The study design was carefully developed to ensure the reliability and representativeness of the findings. First, the study population covered the whole life course, and the number of studies from various groups, including the general population (43/125, 34.4%), maternal and infants (31/125, 24.8%), children (25/125, 20%), and special patients (20/125, 16%) was relatively balanced. Second, the investigated vaccines spanned a wide spectrum, comprising newly emergency-marketed COVID-19 vaccines, annually administered influenza vaccines, globally disseminated human papillomavirus vaccines, diphtheria, tetanus, acellular pertussis vaccines that often raise concerns, and more than 15 other vaccines in total. Finally, this review spanned over 23 countries and regions, with a notable proportion of 13% (n=16) of included studies from LMICs.

The 49 studies examining the safety of COVID-19 vaccines encompassed diverse populations, including the general population, infants, children and adolescents, pregnant women, and special patients. Despite this breadth, only 54 variables were used, likely reflecting the challenges of limited data availability during the rollout of an emergency vaccine. Previous studies have underscored several challenges in evaluating COVID-19 vaccine safety, including insufficient sample sizes and lower study quality [148-150]. Furthermore, misleading evidence has, at times, hindered efforts to increase vaccine coverage [151]. It is imperative for us to establish a stable AVSS system to provide reliable and sufficient data for the safety studies of emergency vaccines.

The maternal population, which includes 96 variables, reflects a broad spectrum of data, likely due to the inclusion of infant-related variables at this level. Pregnant women are frequently excluded from randomized controlled trials of drug safety because of their unique physiological characteristics [152]. Therefore, the real-world evidence from postmarketing surveillance plays a critical role in evaluating vaccine safety for this population. Furthermore, establishing continuous active surveillance systems facilitates the investigation of the long-term effects of prenatal exposure on offspring.

The standard MDS could provide precise guidance and concise extents of data requirements for the AVSS system. However, no relevant guidelines have yet been released. Although both the CIOMS and WHO guidelines briefly outline the data elements of AVSS, their contents lack consistency and empirical validation in real-world applications [153]. This study is the first to validate these concepts and develop a comprehensive framework. When compared with the common data models of established systems like the VSD and Sentinel, the MDS shares several core strengths. Like these systems, the MDS encompasses the entire lifecycle population, covers a wide array of vaccines, and accounts for a broad spectrum of events. Furthermore, the MDS in our study demonstrated the other two advantages. First, it focuses exclusively on vaccine safety surveillance, minimizing unnecessary data collection efforts and addressing the resource constraints often faced in such endeavors. Second, the MDS incorporates a wider range of data sources, including registration and survey data, which better reflect the diversity of real-world settings and enhance the representativeness of AVSS systems.

A well-defined MDS is crucial for ensuring data quality, comparability, and efficient monitoring in public health surveillance. Within an AVSS system, an effective MDS helps identify key indicators, enhances data accessibility, and strengthens risk assessment [22]. By standardizing data collection, it reduces fragmentation, ensures dataset consistency, and enables timely, automated signal detection, allowing policymakers to respond swiftly to emerging safety concerns. In resource-constrained regions, prioritizing key variables such as vaccination date, vaccine name, outcome diagnosis, and diagnosis date, among the most frequently reported in our study, ensures feasibility while maintaining surveillance effectiveness. From a public health perspective, our study aims to establish a robust AVSS that balances data accessibility and efficiency. By integrating elements from established frameworks such as CIOMS and WHO guidelines, we validate their applicability through a comprehensive literature review. Our findings offer a foundational reference for resource-limited regions, with future efforts focusing on refining MDS definitions through expert consultation and real-world application.

Metadata, which reflects data quality and applicability, is essential for ensuring that real-world data can generate real-world evidence. Unfortunately, only 14.6% of the 103 studies on databases provided Metadata information, indicating that most studies overlooked the importance of data quality control and assurance. This raises concerns about the overall quality of evidence in AVSS from previous studies. It's imperative to increase awareness and provide clear guidelines in the future. Certainly, the MDS and Metadata are essential components of the information infrastructure for scientific research on vaccine safety vigilance. Prioritizing their

XSL•FO
RenderX

importance ensures the establishment of high-quality evidence, enhances evidence transparency, and facilitates evidence translation, particularly in LMICs [154].

There were several limitations in our study. First, our review focused exclusively on association studies in vaccine safety from the past 5 years, which may only represent the latest progress in this field but not its full picture. Second, the variables mentioned in previous studies may not fully correspond to those actually used, potentially introducing some bias in the frequency description. Third, non-English studies were excluded from our analysis. However, given that AVSS studies mostly originate from high-income countries, the impact of this publication bias is likely minimal and can be disregarded. Fourth, our analysis was limited to comparative observational studies (eg, cohort studies and case-control studies) with explicit comparison groups, excluding self-controlled designs (eg, self-controlled case series, and case-crossover studies). While this exclusion represents a limitation, Self-Controlled Case Series (SCCS) studies typically require a smaller set of predefined variables, and the key variables used in SCCS designs are already encompassed within cohort and case-control studies. Therefore, although excluding SCCS studies may slightly reduce the frequency counts of certain variables, it does not affect the overall data scope of the identified MDS. Fifth, this study included variables with a frequency of ≥5% as part of the MDS. While variables with <5% frequency, although less commonly reported, may still hold significant importance, such as “place of vaccination” for detecting clustering of adverse events. For transparency and future reference, a complete list of variables identified in this study is provided in Table S5 in Multimedia Appendix 1.

Conclusion

In conclusion, following WHO and CIOMS guidance and adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Checklist 1), this study presents an initial framework of MDS and Metadata for AVSS. This MDS can provide precise guidance and concise requirements for the data scope of the AVSS system and studies. This framework is particularly valuable for resource-limited regions, providing a foundation for implementing AVSS systems effectively and efficiently. Therefore, it is crucial to establish a standardized MDS globally based on these findings to enhance the global vaccine safety ecosystem.

Acknowledgments

This study was supported by National Natural Science Foundation of China (grant number 82204175; 72361127500; 82330107), and the Special Project for Director, China Center for Evidence Based Traditional Chinese Medicine (grant number 2020YJSZX-2).No generative AI tools were utilized at any stage of this research or manuscript preparation

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

SZ and ZL were responsible for conceptualization, funding acquisition, and project administration. MZ conducted data curation, formal analysis, validation, visualization, and wrote the original draft. ZL and JY contributed to investigation and methodology. YL, YL, KL, MR, CF, LZ, YW, YW, CS, FS, and JR provided supervision. Validation was also performed by TL, ZD, SG, and YW. The manuscript was reviewed and edited by MZ, ZL, and SZ.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Basic information and extracted data elements of included studies.
[PDF File, 315 KB - publichealth_v11i1e63161_app1.pdf]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 checklist.
[PDF File, 101 KB - publichealth_v11i1e63161_app2.pdf]

References

1. Ryan J, Malinga T. Interventions for vaccine hesitancy. Curr Opin Immunol 2021 Aug;71:89-91. [doi: 10.1016/j.coi.2021.05.003] [Medline: 34271335]
2. Friedrich MJ. WHO’s top health threats for 2019. JAMA 2019 Mar 19;321(11):1041. [doi: 10.1001/jama.2019.1934] [Medline: 30874765]
3. The Lancet. Vaccine scandal and confidence crisis in China. Lancet 2018 Aug;392(10145):360. [doi: 10.1016/S0140-6736(18)31695-7]

XSL•FO
RenderX

4. Cheng AC, Buttery JP. Vaccine safety: what systems are required to ensure public confidence in vaccines? Med J Aust 2022 Aug 15;217(4):189-190. [doi: 10.5694/mja2.51662] [Medline: 35843626]
5. Ball R, Horne D, Izurieta H, et al. Statistical, epidemiological, and risk-assessment approaches to evaluating safety of vaccines throughout the life cycle at the Food and Drug Administration. Pediatrics 2011 May;127 Suppl 1:S31-S38. [doi: 10.1542/peds.2010-1722F] [Medline: 21502249]
6. Liu Z, Meng R, Yang Y, et al. Progress of active surveillance for vaccine safety in China. China CDC Wkly 2021 Jul 2;3(27):581-583. [doi: 10.46234/ccdcw2021.150] [Medline: 34594940]
7. Poland GA, Black S. Cryptic vaccine-associated adverse events: the critical need for a new vaccine safety surveillance paradigm to improve public trust in vaccines. Vaccine (Auckl) 2024 Mar 19;42(8):1860-1862. [doi: 10.1016/j.vaccine.2024.01.058] [Medline: 38296703]
8. Chen S, Yao L, Wang W, et al. Developing an effective and sustainable national immunisation programme in China: issues and challenges. Lancet Public Health 2022 Dec;7(12):e1064-e1072. [doi: 10.1016/S2468-2667(22)00171-2] [Medline: 36252582]
9. Ishiguro C, Mimura W, Murata F, et al. Development and application of a Japanese vaccine database for comparative assessments in the post-authorization phase: The Vaccine Effectiveness, Networking, and Universal Safety (VENUS) study. Vaccine (Auckl) 2022 Oct 6;40(42):6179-6186. [doi: 10.1016/j.vaccine.2022.08.069] [Medline: 36096969]
10. Liu Z, Zhang L, Yang Y, et al. Active surveillance of adverse events following human papillomavirus vaccination: feasibility pilot study based on the Regional Health Care Information Platform in the City of Ningbo, China. J Med Internet Res 2020 Jun 1;22(6):e17446. [doi: 10.2196/17446] [Medline: 32234696]
11. Blumenthal KG, Phadke NA, Bates DW. Safety surveillance of COVID-19 mRNA vaccines through the Vaccine Safety Datalink. JAMA 2021 Oct 12;326(14):1375-1377. [doi: 10.1001/jama.2021.14808] [Medline: 34477809]
12. Reddy SN, Nair NP, Tate JE, et al. Intussusception after Rotavirus Vaccine introduction in India. N Engl J Med 2020 Nov 12;383(20):1932-1940. [doi: 10.1056/NEJMoa2002276] [Medline: 33176083]
13. Zerbo O, Bartlett J, Fireman B, et al. Effectiveness of recombinant zoster vaccine against herpes zoster in a real-world setting. Ann Intern Med 2024 Feb;177(2):189-195. [doi: 10.7326/M23-2023] [Medline: 38190712]
14. Klein NP, Lewis N, Goddard K, et al. Surveillance for adverse events after COVID-19 mRNA vaccination. JAMA 2021 Oct 12;326(14):1390-1399. [doi: 10.1001/jama.2021.15072] [Medline: 34477808]
15. Hanson KE, Goddard K, Lewis N, et al. Incidence of Guillain-Barré syndrome after COVID-19 vaccination in the Vaccine Safety Datalink. JAMA Netw Open 2022 Apr 1;5(4):e228879. [doi: 10.1001/jamanetworkopen.2022.8879] [Medline: 35471572]
16. Deng L, Glover C, Dymock M, et al. The short term safety of COVID-19 vaccines in Australia: AusVaxSafety active surveillance, February - August 2021. Med J Aust 2022 Aug 15;217(4):195-202. [doi: 10.5694/mja2.51619] [Medline: 35781813]
17. Kharbanda EO, Vazquez-Benitez G, DeSilva MB, et al. Association of inadvertent 9-valent human papillomavirus vaccine in pregnancy with spontaneous abortion and adverse birth outcomes. JAMA Netw Open 2021 Apr 1;4(4):e214340. [doi: 10.1001/jamanetworkopen.2021.4340] [Medline: 33818618]
18. McNeil MM, Gee J, Weintraub ES, et al. The Vaccine Safety Datalink: successes and challenges monitoring vaccine safety. Vaccine (Auckl) 2014 Sep 22;32(42):5390-5398. [doi: 10.1016/j.vaccine.2014.07.073] [Medline: 25108215]
19. Top KA, Macartney K, Bettinger JA, et al. Active surveillance of acute paediatric hospitalisations demonstrates the impact of vaccination programmes and informs vaccine policy in Canada and Australia. Euro Surveill 2020 Jun;25(25):1900562. [doi: 10.2807/1560-7917.ES.2020.25.25.1900562] [Medline: 32613939]
20. Pillsbury AJ, Fathima P, Quinn HE, et al. Comparative postmarket safety profile of adjuvanted and high-dose influenza vaccines in individuals 65 years or older. JAMA Netw Open 2020 May 1;3(5):e204079. [doi: 10.1001/jamanetworkopen.2020.4079] [Medline: 32369177]
21. Chandler RE, Balakrishnan MR, Brasseur D, et al. Collaboration within the global vaccine safety surveillance ecosystem during the COVID-19 pandemic: lessons learnt and key recommendations from the COVAX Vaccine Safety Working Group. BMJ Glob Health 2024 Mar 7;9(3):e014544. [doi: 10.1136/bmjgh-2023-014544] [Medline: 38453518]
22. Bernardi FA, Mello de Oliveira B, Bettiol Yamada D, et al. The minimum data set for rare diseases: systematic review. J Med Internet Res 2023 Jul 27;25:e44641. [doi: 10.2196/44641] [Medline: 37498666]
23. Zahmatkeshan M, Farjam M, Mohammadzadeh N, et al. Design of infertility monitoring system: minimum data set approach. J Med Life 2019;12(1):56-64. [doi: 10.25122/jml-2018-0071] [Medline: 31123526]
24. Ulrich H, Kock-Schoppenhauer AK, Deppenwiese N, et al. Understanding the nature of metadata: systematic review. J Med Internet Res 2022 Jan 11;24(1):e25440. [doi: 10.2196/25440] [Medline: 35014967]
25. Covid-19 vaccines: safety surveillance manual. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240032781 [accessed 2022-09-22]
26. CIOMS guide to active vaccine safety surveillance. CIOMS. 2017. URL: https://cioms.ch/publications/product/cioms-guide-to-active-vaccine-safety-surveillance/ [accessed 2017-07-13]

XSL•FO
RenderX

27. Good practice guide for the use of the metadata catalogue of real-world data sources. EMA. 2022. URL: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/good-practice-guide-use-metadata-catalogue-real-world-data-sources_en.pdf [accessed 2022-09-01]
28. Cross JW, Joy M, McGee C, et al. Adverse events of interest vary by influenza vaccine type and brand: sentinel network study of eight seasons (2010–2018). Vaccine (Auckl) 2020 May;38(22):3869-3880. [doi: 10.1016/j.vaccine.2020.03.034]
29. Alfayadh NM, Gowdie PJ, Akikusa JD, et al. Vaccinations do not increase arthritis flares in juvenile idiopathic arthritis: a study of the relationship between routine childhood vaccinations on the Australian immunization schedule and arthritis activity in children with juvenile idiopathic arthritis. Int J Rheumatol 2020;2020:1078914. [doi: 10.1155/2020/1078914] [Medline: 32831849]
30. Rolfes MA, Vonglokham P, Khanthamaly V, et al. Measurement of birth outcomes in analyses of the impact of maternal influenza vaccination. Influenza Other Respir Viruses 2019 Nov;13(6):547-555. [doi: 10.1111/irv.12673] [Medline: 31424627]
31. Elding Larsson H, Lynch KF, Lönnrot M, et al. Pandemrix® vaccination is not associated with increased risk of islet autoimmunity or type 1 diabetes in the TEDDY study children. Diabetologia 2018 Jan;61(1):193-202. [doi: 10.1007/s00125-017-4448-3] [Medline: 28990147]
32. Corsenac P, Parent M, Mansaray H, et al. Early life Bacillus Calmette-Guerin vaccination and incidence of type 1, type 2, and latent autoimmune diabetes in adulthood. Diabetes Metab 2022 May;48(3):101337. [doi: 10.1016/j.diabet.2022.101337]
33. Caspard H, Steffey A, Mallory RM, et al. Evaluation of the safety of live attenuated influenza vaccine (LAIV) in children and adolescents with asthma and high-risk conditions: a population-based prospective cohort study conducted in England with the Clinical Practice Research Datalink. BMJ Open 2018 Dec 9;8(12):e023118. [doi: 10.1136/bmjopen-2018-023118] [Medline: 30530581]
34. Yamamoto-Hanada K, Pak K, Saito-Abe M, et al. Cumulative inactivated vaccine exposure and allergy development among children: a birth cohort from Japan. Environ Health Prev Med 2020 Jul 7;25(1):27. [doi: 10.1186/s12199-020-00864-7] [Medline: 32635895]
35. Salmon C, Conus F, et al. Association between Bacillus Calmette–Guérin vaccination and lymphoma: a population‐based birth cohort study. J Intern Med 2019 Nov;286(5):583-595. [doi: 10.1111/joim.12965] [Medline: 31361936]
36. Øland CB, Mogensen SW, Rodrigues A, et al. Reduced mortality after oral polio vaccination and increased mortality after diphtheria-tetanus-pertussis vaccination in children in a low-income setting. Clin Ther 2021 Jan;43(1):172-184. [doi: 10.1016/j.clinthera.2020.11.010] [Medline: 33277047]
37. Gögenur M, Fransgård T, Krause TG, et al. Association of influenza vaccine and risk of recurrence in patients undergoing curative surgery for colorectal cancer. Acta Oncol 2021 Nov;60(11):1507-1512. [doi: 10.1080/0284186X.2021.1967444] [Medline: 34459323]
38. Christiansen CF, Thomsen RW, Schmidt M, et al. Influenza vaccination and 1-year risk of myocardial infarction, stroke, heart failure, pneumonia, and mortality among intensive care unit survivors aged 65 years or older: a nationwide population-based cohort study. Intensive Care Med 2019 Jul;45(7):957-967. [doi: 10.1007/s00134-019-05648-4] [Medline: 31187170]
39. Orta OR, Hatch EE, Regan AK, et al. A prospective study of influenza vaccination and time to pregnancy. Vaccine (Auckl) 2020 Jun 2;38(27):4246-4251. [doi: 10.1016/j.vaccine.2020.04.054] [Medline: 32409134]
40. Hall C, Abramovitz LM, Bukowinski AT, et al. Safety of tetanus, diphtheria, and acellular pertussis vaccination among pregnant active duty U.S. military women. Vaccine (Auckl) 2020 Feb;38(8):1982-1988. [doi: 10.1016/j.vaccine.2020.01.009]
41. Proaños NJ, Reyes LF, Bastidas A, et al. Prior influenza vaccine is not a risk factor for bacterial coinfection in patients admitted to the ICU due to severe influenza. Med Intensiva (Engl Ed) 2022 Aug;46(8):436-445. [doi: 10.1016/j.medine.2021.05.009] [Medline: 35868720]
42. Knowlton KU, Knight S, Muhlestein JB, et al. A small but significantly greater incidence of inflammatory heart disease identified after vaccination for severe acute respiratory syndrome Coronavirus 2. Open Forum Infect Dis 2022 Mar;9(3):ofab663. [doi: 10.1093/ofid/ofab663] [Medline: 35141346]
43. Horvat JV, Sevilimedu V, Becker AS, et al. Frequency and outcomes of MRI-detected axillary adenopathy following COVID-19 vaccination. Eur Radiol 2022 Aug;32(8):5752-5758. [doi: 10.1007/s00330-022-08655-0] [Medline: 35247087]
44. Peretz-Machluf R, Hirsh-Yechezkel G, Zaslavsky-Paltiel I, et al. Obstetric and neonatal outcomes following COVID-19 vaccination in pregnancy. J Clin Med 2022 Apr 30;11(9):2540. [doi: 10.3390/jcm11092540] [Medline: 35566665]
45. Whiteley WN, Ip S, Cooper JA, et al. Association of COVID-19 vaccines ChAdOx1 and BNT162b2 with major venous, arterial, or thrombocytopenic events: a population-based cohort study of 46 million adults in England. PLoS Med 2022 Feb;19(2):e1003926. [doi: 10.1371/journal.pmed.1003926] [Medline: 35192597]
46. Toepfner N, von Meißner WCG, Strumann C, et al. Comparative safety of the BNT162b2 messenger RNA COVID-19 vaccine vs other approved vaccines in children younger than 5 years. JAMA Netw Open 2022 Oct 3;5(10):e2237140. [doi: 10.1001/jamanetworkopen.2022.37140] [Medline: 36255723]
47. Chou OHI, Zhou J, Lee TTL, et al. Comparisons of the risk of myopericarditis between COVID-19 patients and individuals receiving COVID-19 vaccines: a population-based study. Clin Res Cardiol 2022 Oct;111(10):1098-1103. [doi: 10.1007/s00392-022-02007-0] [Medline: 35333945]

XSL•FO
RenderX

48. Becerra-Culqui TA, Getahun D, Chiu V, et al. Prenatal influenza vaccination or influenza infection and autism spectrum disorder in offspring. Clin Infect Dis 2022 Sep 30;75(7):1140-1148. [doi: 10.1093/cid/ciac101] [Medline: 35174388]
49. Petousis-Harris H, Jiang Y, Yu L, et al. A retrospective cohort study of safety outcomes in New Zealand infants exposed to Tdap vaccine in utero. Vaccines (Basel) 2019 Oct 11;7(4):147. [doi: 10.3390/vaccines7040147] [Medline: 31614582]
50. Wesselink AK, Hatch EE, Rothman KJ, et al. A prospective cohort study of COVID-19 vaccination, SARS-CoV-2 infection, and fertility. Am J Epidemiol 2022 Jul 23;191(8):1383-1395. [doi: 10.1093/aje/kwac011] [Medline: 35051292]
51. Weaver KN, Zhang X, Dai X, et al. Impact of SARS-CoV-2 vaccination on inflammatory bowel disease activity and development of vaccine-related adverse events: results from PREVENT-COVID. Inflamm Bowel Dis 2022 Oct 3;28(10):1497-1505. [doi: 10.1093/ibd/izab302] [Medline: 34871388]
52. Tseng HF, Sy LS, Qian L, et al. Pneumococcal conjugate vaccine safety in elderly adults. Open Forum Infect Dis 2018 Jun;5(6):ofy100. [doi: 10.1093/ofid/ofy100] [Medline: 29977960]
53. Agger WA, Deviley JA, Borgert AJ, et al. Increased incidence of giant cell arteritis after introduction of a live varicella zoster virus vaccine. Open Forum Infect Dis 2021 Feb;8(2):ofaa647. [doi: 10.1093/ofid/ofaa647] [Medline: 33598502]
54. Tseng HF, Sy LS, Ackerson BK, et al. Safety of tetanus, diphtheria, acellular pertussis (Tdap) vaccination during pregnancy. Vaccine (Auckl) 2022 Jul 30;40(32):4503-4512. [doi: 10.1016/j.vaccine.2022.06.009] [Medline: 35717267]
55. Andrews NJ, Stowe J, Ramsay ME, et al. Risk of venous thrombotic events and thrombocytopenia in sequential time periods after ChAdOx1 and BNT162b2 COVID-19 vaccines: a national cohort study in England. Lancet Reg Health Eur 2022 Feb;13:100260. [doi: 10.1016/j.lanepe.2021.100260] [Medline: 34927118]
56. Corrao G, Rea F, Franchi M, et al. Balancing benefits and harms of COVID-19 vaccines: lessons from the ongoing mass vaccination campaign in Lombardy, Italy. Vaccines (Basel) 2022 Apr 15;10(4):623. [doi: 10.3390/vaccines10040623] [Medline: 35455374]
57. Hoffmann SS, Thiesson EM, Johansen JD, et al. Risk factors for granulomas in children following immunization with aluminium-adsorbed vaccines: a Danish population-based cohort study. Contact Derm 2022 Nov;87(5):430-438. [doi: 10.1111/cod.14180] [Medline: 35778959]
58. Baldolli A, Fournier A, Verdon R, et al. Reactogenicity among health care workers following a BNT162b2 or mRNA-1273 second dose after priming with a ChAdOx1 nCOV-19 vaccine. Clin Microbiol Infect 2022 Jun;28(6):885. [doi: 10.1016/j.cmi.2022.02.010] [Medline: 35182759]
59. Zerbo O, Modaressi S, Goddard K, et al. Safety of measles and pertussis-containing vaccines in children with autism spectrum disorders. Vaccine (Auckl) 2022 Apr 20;40(18):2568-2573. [doi: 10.1016/j.vaccine.2022.03.031] [Medline: 35315325]
60. Taquet M, Husain M, Geddes JR, et al. Cerebral venous thrombosis and portal vein thrombosis: a retrospective cohort study of 537,913 COVID-19 cases. EClinicalMedicine 2021 Sep;39:101061. [doi: 10.1016/j.eclinm.2021.101061] [Medline: 34368663]
61. Foo D, Sarna M, Pereira G, et al. Maternal influenza vaccination and child mortality: longitudinal, population-based linked cohort study. Vaccine (Auckl) 2022 Jun 15;40(27):3732-3736. [doi: 10.1016/j.vaccine.2022.05.030] [Medline: 35606236]
62. Ou MT, Boyarsky BJ, Motter JD, et al. Safety and reactogenicity of 2 doses of SARS-CoV-2 vaccination in solid organ transplant recipients. Transplantation 2021 Oct 1;105(10):2170-2174. [doi: 10.1097/TP.0000000000003780] [Medline: 33859151]
63. Martín-Merino E, Castillo-Cano B, Martín-Perez M, et al. Papillomavirus vaccination and Guillain-Barre Syndrome among girls: a cohort study in Spain. Vaccine (Auckl) 2021 Jul 13;39(31):4306-4313. [doi: 10.1016/j.vaccine.2021.06.029] [Medline: 34175126]
64. Pottegård A, Lund LC, Karlstad Ø, et al. Arterial events, venous thromboembolism, thrombocytopenia, and bleeding after vaccination with Oxford-AstraZeneca ChAdOx1-S in Denmark and Norway: population based cohort study. BMJ 2021 May 5;373:n1114. [doi: 10.1136/bmj.n1114] [Medline: 33952445]
65. Blakeway H, Prasad S, Kalafat E, et al. COVID-19 vaccination during pregnancy: coverage and safety. Am J Obstet Gynecol 2022 Feb;226(2):236. [doi: 10.1016/j.ajog.2021.08.007] [Medline: 34389291]
66. Dick A, Rosenbloom JI, Karavani G, et al. Safety of third SARS-CoV-2 vaccine (booster dose) during pregnancy. Am J Obstet Gynecol MFM 2022 Jul;4(4):100637. [doi: 10.1016/j.ajogmf.2022.100637] [Medline: 35398583]
67. Bartels LE, Ammitzbøll C, Andersen JB, et al. Local and systemic reactogenicity of COVID-19 vaccine BNT162b2 in patients with systemic lupus erythematosus and rheumatoid arthritis. Rheumatol Int 2021 Nov;41(11):1925-1931. [doi: 10.1007/s00296-021-04972-7] [Medline: 34476603]
68. Karlstad Ø, Hovi P, Husby A, et al. SARS-CoV-2 vaccination and myocarditis in a Nordic cohort study of 23 million residents. JAMA Cardiol 2022 Jun 1;7(6):600-612. [doi: 10.1001/jamacardio.2022.0583] [Medline: 35442390]
69. Speake HA, Pereira G, Regan AK. Risk of adverse maternal and foetal outcomes associated with inactivated influenza vaccination in first trimester of pregnancy. Paediatr Perinat Epidemiol 2021 Mar;35(2):196-205. [doi: 10.1111/ppe.12715] [Medline: 33155331]
70. Hviid A, Hansen JV, Frisch M, et al. Measles, mumps, rubella vaccination and autism: a nationwide cohort study. Ann Intern Med 2019 Apr 16;170(8):513-520. [doi: 10.7326/M18-2101] [Medline: 30831578]

XSL•FO
RenderX

71. Foo D, Sarna M, Pereira G, et al. Prenatal influenza vaccination and allergic and autoimmune diseases in childhood: a longitudinal, population-based linked cohort study. PLoS Med 2022 Apr;19(4):e1003963. [doi: 10.1371/journal.pmed.1003963] [Medline: 35381006]
72. Kent A, Beebeejaun K, Braccio S, et al. Safety of meningococcal group B vaccination in hospitalised premature infants. Arch Dis Child Fetal Neonatal Ed 2019 Mar;104(2):F171-F175. [doi: 10.1136/archdischild-2017-314152] [Medline: 29636385]
73. Groom HC, Irving SA, Koppolu P, et al. Uptake and safety of Hepatitis B vaccination during pregnancy: a Vaccine Safety Datalink study. Vaccine (Auckl) 2018 Oct 1;36(41):6111-6116. [doi: 10.1016/j.vaccine.2018.08.074] [Medline: 30194002]
74. Lee SM, Kim SJ, Chen J, et al. Post-marketing surveillance to assess the safety and tolerability of a combined diphtheria, tetanus, acellular pertussis and inactivated poliovirus vaccine (DTaP-IPV) in Korean children. Hum Vaccin Immunother 2019;15(5):1145-1153. [doi: 10.1080/21645515.2019.1572406] [Medline: 30668217]
75. Tu TM, Yi SJ, Koh JS, et al. Incidence of cerebral venous thrombosis following SARS-CoV-2 infection vs mRNA SARS-CoV-2 vaccination in Singapore. JAMA Netw Open 2022 Mar 1;5(3):e222940. [doi: 10.1001/jamanetworkopen.2022.2940] [Medline: 35297971]
76. Dick A, Rosenbloom JI, Gutman-Ido E, et al. Safety of SARS-CoV-2 vaccination during pregnancy- obstetric outcomes from a large cohort study. BMC Pregnancy Childbirth 2022 Feb 28;22(1). [doi: 10.1186/s12884-022-04505-5] [Medline: 35227233]
77. Becerra-Culqui TA, Getahun D, Chiu V, et al. Prenatal tetanus, diphtheria, acellular pertussis vaccination and autism spectrum disorder. Pediatrics 2018 Sep;142(3):e20180120. [doi: 10.1542/peds.2018-0120] [Medline: 30104424]
78. Baker MA, Baer B, Kulldorff M, et al. Kawasaki disease and 13-valent pneumococcal conjugate vaccination among young children: A self-controlled risk interval and cohort study with null results. PLoS Med 2019 Jul;16(7):e1002844. [doi: 10.1371/journal.pmed.1002844] [Medline: 31265459]
79. Laverty M, Crowcroft N, Bolotin S, et al. Health outcomes in young children following pertussis vaccination during pregnancy. Pediatrics 2021 May;147(5):e2020042507. [doi: 10.1542/peds.2020-042507] [Medline: 33875535]
80. Pawlowski C, Rincón-Hekking J, Awasthi S, et al. Cerebral venous sinus thrombosis is not significantly linked to COVID-19 vaccines or non-COVID vaccines in a large multi-state health system. J Stroke Cerebrovasc Dis 2021 Oct;30(10):105923. [doi: 10.1016/j.jstrokecerebrovasdis.2021.105923] [Medline: 34627592]
81. Shasha D, Bareket R, Sikron FH, et al. Real-world safety data for the Pfizer BNT162b2 SARS-CoV-2 vaccine: historical cohort study. Clin Microbiol Infect 2022 Jan;28(1):130-134. [doi: 10.1016/j.cmi.2021.09.018] [Medline: 34592420]
82. Rottenstreich M, Sela HY, Rotem R, et al. Covid-19 vaccination during the third trimester of pregnancy: rate of vaccination and maternal and neonatal outcomes, a multicentre retrospective cohort study. BJOG 2022 Jan;129(2):248-255. [doi: 10.1111/1471-0528.16941] [Medline: 34554630]
83. Hviid A, Svanström H, Scheller NM, et al. Human papillomavirus vaccination of adult women and risk of autoimmune and neurological diseases. J Intern Med 2018 Feb;283(2):154-165. [doi: 10.1111/joim.12694] [Medline: 29044769]
84. Bardenheier BH, Gravenstein S, Blackman C, et al. Adverse events following mRNA SARS-CoV-2 vaccination among U.S. nursing home residents. Vaccine (Auckl) 2021 Jun 29;39(29):3844-3851. [doi: 10.1016/j.vaccine.2021.05.088] [Medline: 34092431]
85. Scherrer JF, Salas J, Wiemken TL, et al. Impact of herpes zoster vaccination on incident dementia: a retrospective study in two patient cohorts. PLoS ONE 2021;16(11):e0257405. [doi: 10.1371/journal.pone.0257405] [Medline: 34788293]
86. Arora NK, Das MK, Poluru R, et al. A prospective cohort study on the safety of infant pentavalent (DTwP-HBV-Hib) and oral polio vaccines in two South Indian districts. Pediatr Infect Dis J 2020 May;39(5):389-396. [doi: 10.1097/INF.0000000000002594] [Medline: 32301918]
87. Bruxvoort K, Slezak J, Qian L, et al. Association between 2-dose vs 3-dose hepatitis b vaccine and acute myocardial infarction. JAMA 2022 Apr 5;327(13):1260-1268. [doi: 10.1001/jama.2022.2540] [Medline: 35333303]
88. Mohammed H, Roberts CT, Grzeskowiak LE, et al. Safety of maternal pertussis vaccination on pregnancy and birth outcomes: a prospective cohort study. Vaccine (Auckl) 2021 Jan 8;39(2):324-331. [doi: 10.1016/j.vaccine.2020.11.052] [Medline: 33280854]
89. Ludvigsson JF, Winell H, Sandin S, et al. Maternal influenza A(H1N1) immunization during pregnancy and risk for autism spectrum disorder in offspring: a cohort study. Ann Intern Med 2020 Oct 20;173(8):597-604. [doi: 10.7326/M20-0167] [Medline: 32866418]
90. Xiong X, Wong CKH, Au ICH, et al. Safety of inactivated and mRNA COVID-19 vaccination among patients treated for hypothyroidism: a population-based cohort study. Thyroid 2022 May;32(5):505-514. [doi: 10.1089/thy.2021.0684] [Medline: 35216517]
91. Scherrer JF, Salas J, Wiemken TL, et al. Lower risk for dementia following adult tetanus, diphtheria, and pertussis (Tdap) vaccination. J Gerontol A Biol Sci Med Sci 2021 Jul 13;76(8):1436-1443. [doi: 10.1093/gerona/glab115] [Medline: 33856020]
92. MacDonald SE, Dover DC, Hill MD, et al. Is varicella vaccination associated with pediatric arterial ischemic stroke? A population-based cohort study. Vaccine (Auckl) 2018 May 11;36(20):2764-2767. [doi: 10.1016/j.vaccine.2018.04.012] [Medline: 29655630]

XSL•FO
RenderX

93. Ohfuji S, Deguchi M, Tachibana D, et al. Safety of influenza vaccination on adverse birth outcomes among pregnant women: a prospective cohort study in Japan. Int J Infect Dis 2020 Apr;93:68-76. [doi: 10.1016/j.ijid.2020.01.033] [Medline: 31982621]
94. Wijn DH, Groeneveld GH, Vollaard AM, et al. Influenza vaccination in patients with lung cancer receiving anti-programmed death receptor 1 immunotherapy does not induce immune-related adverse events. Eur J Cancer 2018 Nov;104:182-187. [doi: 10.1016/j.ejca.2018.09.012] [Medline: 30368069]
95. Walsh LK, Donelle J, Dodds L, et al. Health outcomes of young children born to mothers who received 2009 pandemic H1N1 influenza vaccination during pregnancy: retrospective cohort study. BMJ 2019 Jul 10;366:l4151. [doi: 10.1136/bmj.l4151] [Medline: 31292120]
96. Andersson NW, Thiesson EM, Laursen MV, et al. Safety of heterologous primary and booster schedules with ChAdOx1-S and BNT162b2 or mRNA-1273 vaccines: nationwide cohort study. BMJ 2022 Jul 13;378:e070483. [doi: 10.1136/bmj-2022-070483] [Medline: 35831006]
97. Barda N, Dagan N, Ben-Shlomo Y, et al. Safety of the BNT162b2 mRNA Covid-19 vaccine in a nationwaide setting. N Engl J Med 2021 Sep 16;385(12):1078-1090. [doi: 10.1056/NEJMoa2110475] [Medline: 34432976]
98. Sheel M, Wood N, Macartney K, et al. Severity of rotavirus-vaccine-associated intussusception: prospective hospital-based surveillance, Australia, 2007-2018. Pediatr Infect Dis J 2022 Jun 1;41(6):507-513. [doi: 10.1097/INF.0000000000003521] [Medline: 35363642]
99. Decker MD, Garman PM, Hughes H, et al. Enhanced safety surveillance study of ACAM2000 smallpox vaccine among US military service members. Vaccine (Auckl) 2021 Sep 15;39(39):5541-5547. [doi: 10.1016/j.vaccine.2021.08.041] [Medline: 34454787]
100. da Cunha GK, de Matos MB, Trettim JP, et al. Thimerosal-containing vaccines and deficit in child development: population-based study in southern Brazil. Vaccine (Auckl) 2020 Feb 24;38(9):2216-2220. [doi: 10.1016/j.vaccine.2019.12.044] [Medline: 32014269]
101. Rider LG, Parks CG, Wilkerson J, et al. Baseline factors associated with self-reported disease flares following COVID-19 vaccination among adults with systemic rheumatic disease: results from the COVID-19 global rheumatology alliance vaccine survey. Rheumatology (Oxford) 2022 Jun 28;61(SI2):SI143-SI150. [doi: 10.1093/rheumatology/keac249] [Medline: 35460240]
102. Layton JB, Butler AM, Panozzo CA, et al. Rotavirus vaccination and short-term risk of adverse events in US infants. Paediatr Perinat Epidemiol 2018 Sep;32(5):448-457. [doi: 10.1111/ppe.12496] [Medline: 30048564]
103. Sarna M, Pereira GF, Foo D, et al. The risk of major structural birth defects associated with seasonal influenza vaccination during pregnancy: a population-based cohort study. Birth Defects Res 2022 Nov 15;114(19):1244-1256. [doi: 10.1002/bdr2.2049] [Medline: 35678518]
104. Hertel M, Heiland M, Nahles S, et al. Real-world evidence from over one million COVID-19 vaccinations is consistent with reactivation of the varicella-zoster virus. J Eur Acad Dermatol Venereol 2022 Aug;36(8):1342-1348. [doi: 10.1111/jdv.18184] [Medline: 35470920]
105. Bukhbinder AS, Ling Y, Hasan O, et al. Risk of Alzheimer's disease following influenza vaccination: a claims‐based cohort study using propensity score matching. J Alzheimers Dis 2022;88(3):1061-1074. [doi: 10.3233/JAD-220361] [Medline: 35723106]
106. van Dongen JAP, Rouers EDM, Schuurman R, et al. Rotavirus vaccine safety and effectiveness in infants with high-risk medical conditions. Pediatrics 2021 Dec 1;148(6):e2021051901. [doi: 10.1542/peds.2021-051901] [Medline: 34814164]
107. Peppa M, Thomas SL, Minassian C, et al. Seasonal influenza vaccination during pregnancy and the risk of major congenital malformations in live-born infants: a 2010–2016 historical cohort study. Clin Infect Dis 2021 Dec 6;73(11):e4296-e4304. [doi: 10.1093/cid/ciaa845] [Medline: 32572453]
108. Goldshtein I, Steinberg DM, Kuint J, et al. Association of BNT162b2 COVID-19 vaccination during pregnancy with neonatal and early infant outcomes. JAMA Pediatr 2022 May 1;176(5):470-477. [doi: 10.1001/jamapediatrics.2022.0001] [Medline: 35142809]
109. Groom HC, Smith N, Irving SA, et al. Uptake and safety of hepatitis A vaccination during pregnancy: a Vaccine Safety Datalink study. Vaccine (Auckl) 2019 Oct 16;37(44):6648-6655. [doi: 10.1016/j.vaccine.2019.09.043] [Medline: 31548013]
110. Cuschieri S, Borg M, Agius S, et al. Adverse reactions to Pfizer-BioNTech vaccination of healthcare workers at Malta's state hospital. Int J Clin Pract 2021 Oct;75(10):e14605. [doi: 10.1111/ijcp.14605] [Medline: 34228863]
111. Yoon D, Lee JH, Lee H, et al. Association between human papillomavirus vaccination and serious adverse events in South Korean adolescent girls: nationwide cohort study. BMJ 2021 Jan 29;372:m4931. [doi: 10.1136/bmj.m4931] [Medline: 33514507]
112. Goud R, Lufkin B, Duffy J, et al. Risk of Guillain-Barré syndrome following recombinant zoster vaccine in medicare beneficiaries. JAMA Intern Med 2021 Dec 1;181(12):1623-1630. [doi: 10.1001/jamainternmed.2021.6227] [Medline: 34724025]
113. Lai FTT, Huang L, Peng K, et al. Post-Covid-19-vaccination adverse events and healthcare utilization among individuals with or without previous SARS-CoV-2 infection. J Intern Med 2022 Jun;291(6):864-869. [doi: 10.1111/joim.13453] [Medline: 35043503]

XSL•FO
RenderX

114. Kang W, Shami JJP, Yan VKC, et al. Safety of two-dose COVID-19 vaccination (BNT162b2 and CoronaVac) in adults with cancer: a territory-wide cohort study. J Hematol Oncol 2022 May 19;15(1):66. [doi: 10.1186/s13045-022-01265-9] [Medline: 35590336]
115. Frisch M, Besson A, Clemmensen KKB, et al. Quadrivalent human papillomavirus vaccination in boys and risk of autoimmune diseases, neurological diseases and venous thromboembolism. Int J Epidemiol 2018 Apr 1;47(2):634-641. [doi: 10.1093/ije/dyx273] [Medline: 29425361]
116. Hertel M, Schmidt-Westhausen AM, Wendy S, et al. Onset of oral lichenoid lesions and oral lichen planus following COVID-19 vaccination: a retrospective analysis of about 300,000 vaccinated patients. Vaccines (Basel) 2022 Mar 20;10(3):480. [doi: 10.3390/vaccines10030480] [Medline: 35335112]
117. Hviid A, Myrup Thiesson E. Association between human papillomavirus vaccination and primary ovarian insufficiency in a nationwide cohort. JAMA network open 2021 Aug 2;4(8):e2120391. [doi: 10.1001/jamanetworkopen.2021.20391] [Medline: 34436612]
118. Wong CKH, Lau KTK, Xiong X, et al. Adverse events of special interest and mortality following vaccination with mRNA (BNT162b2) and inactivated (CoronaVac) SARS-CoV-2 vaccines in Hong Kong: a retrospective study. PLoS Med 2022 Jun;19(6):e1004018. [doi: 10.1371/journal.pmed.1004018] [Medline: 35727759]
119. Houghton DE, Wysokinski W, Casanegra AI, et al. Risk of venous thromboembolism after COVID-19 vaccination. J Thromb Haemost 2022 Jul;20(7):1638-1644. [doi: 10.1111/jth.15725] [Medline: 35398975]
120. Faix DJ, Gordon DM, Perry LN, et al. Prospective safety surveillance study of ACAM2000 smallpox vaccine in deploying military personnel. Vaccine (Auckl) 2020 Oct;38(46):7323-7330. [doi: 10.1016/j.vaccine.2020.09.037]
121. Hviid A, Hansen JV, Thiesson EM, et al. Association of AZD1222 and BNT162b2 COVID-19 vaccination with thromboembolic and thrombocytopenic events in frontline personnel: a retrospective cohort study. Ann Intern Med 2022 Apr;175(4):541-546. [doi: 10.7326/M21-2452] [Medline: 35103482]
122. Lai FTT, Chua GT, Chan EWW, et al. Adverse events of special interest following the use of BNT162b2 in adolescents: a population-based retrospective cohort study. Emerg Microbes Infect 2022 Dec;11(1):885-893. [doi: 10.1080/22221751.2022.2050952] [Medline: 35254219]
123. Becerra-Culqui TA, Getahun D, Chiu V, et al. The association of prenatal tetanus, diphtheria, and acellular pertussis (Tdap) vaccination with attention-deficit/hyperactivity disorder. Am J Epidemiol 2020 Oct 1;189(10):1163-1172. [doi: 10.1093/aje/kwaa074] [Medline: 32378703]
124. Simpson CR, Shi T, Vasileiou E, et al. First-dose ChAdOx1 and BNT162b2 COVID-19 vaccines and thrombocytopenic, thromboembolic and hemorrhagic events in Scotland. Nat Med 2021 Jul;27(7):1290-1297. [doi: 10.1038/s41591-021-01408-4] [Medline: 34108714]
125. Khan N, Trivedi C, Aberra F, et al. Safety of recombinant zoster vaccine in patients with inflammatory bowel disease. J Crohns Colitis 2022 Sep 8;16(9):1505-1507. [doi: 10.1093/ecco-jcc/jjac040] [Medline: 35350070]
126. Failing JJ, Ho TP, Yadav S, et al. Safety of influenza vaccine in patients with cancer receiving pembrolizumab. JCO Oncol Pract 2020 Jul;16(7):e573-e580. [doi: 10.1200/JOP.19.00495] [Medline: 32048920]
127. Kiely M, Billard MN, Toth E, et al. Investigation of an increase in large local reactions following vaccine schedule change to include DTaP-HB-IPV-Hib (Infanrix-hexa®) and MMRV (ProQuad®) at 18 months of age. Vaccine (Auckl) 2018 Oct;36(45):6688-6694. [doi: 10.1016/j.vaccine.2018.09.049]
128. Weibel D, Sturkenboom M, Black S, et al. Narcolepsy and adjuvanted pandemic influenza A (H1N1) 2009 vaccines - Multi-country assessment. Vaccine (Auckl) 2018 Oct 1;36(41):6202-6211. [doi: 10.1016/j.vaccine.2018.08.008] [Medline: 30122647]
129. Yokomichi H, Tanaka-Taya K, Koshida R, et al. Immune thrombocytopenic purpura risk by live, inactivated and simultaneous vaccinations among Japanese adults, children and infants: a matched case-control study. Int J Hematol 2020 Jul;112(1):105-114. [doi: 10.1007/s12185-020-02866-1] [Medline: 32253664]
130. Nayak MK, Kumar M, Nayak RK, et al. Prevalence of Intussusception after rotavirus vaccination: a hospital based study from Odisha, India. J Clin Diagn Res . [doi: 10.7860/JCDR/2019/39517.12616]
131. Wan EYF, Chui CSL, Ng VWS, et al. Messenger RNA Coronavirus Disease 2019 (COVID‐19) vaccination with BNT162b2 increased risk of Bell's palsy: a nested case‐control and self‐controlled case series study. Clin Infect Dis 2023 Feb 8;76(3):e291-e298. [doi: 10.1093/cid/ciac460] [Medline: 35675702]
132. Donahue JG, Kieke BA, King JP, et al. Inactivated influenza vaccine and spontaneous abortion in the Vaccine Safety Datalink in 2012-13, 2013-14, and 2014-15. Vaccine (Auckl) 2019 Oct 16;37(44):6673-6681. [doi: 10.1016/j.vaccine.2019.09.035] [Medline: 31540812]
133. Oberle D, Hoffelner M, Pavel J, et al. Retrospective multicenter matched case-control study on the risk factors for intussusception in infants less than 1 year of age with a special focus on rotavirus vaccines - the German Intussusception Study. Hum Vaccin Immunother 2020 Oct 2;16(10):2481-2494. [doi: 10.1080/21645515.2020.1726679] [Medline: 32271647]
134. Le Vu S, Bertrand M, Jabagi MJ, et al. Age and sex-specific risks of myocarditis and pericarditis following Covid-19 messenger RNA vaccines. Nat Commun 2022 Jun 25;13(1):3633. [doi: 10.1038/s41467-022-31401-5] [Medline: 35752614]

XSL•FO
RenderX

135. Wan EYF, Chui CSL, Wang Y, et al. Herpes zoster related hospitalization after inactivated (CoronaVac) and mRNA (BNT162b2) SARS-CoV-2 vaccination: a self-controlled case series and nested case-control study. Lancet Reg Health West Pac 2022 Apr;21:100393. [doi: 10.1016/j.lanwpc.2022.100393] [Medline: 35128500]
136. Kleinstern G, Larson MC, Ansell SM, et al. Vaccination history and risk of lymphoma and its major subtypes. Cancer Epidemiol Biomarkers Prev 2022 Feb;31(2):461-470. [doi: 10.1158/1055-9965.EPI-21-0383] [Medline: 34782394]
137. Abu-Rumeileh S, Mayer B, Still V, et al. Varicella zoster virus-induced neurological disease after COVID-19 vaccination: a retrospective monocentric study. J Neurol 2022 Apr;269(4):1751-1757. [doi: 10.1007/s00415-021-10849-3] [Medline: 34724572]
138. Geier DA, Kern JK, Geier MR. Premature puberty and thimerosal-containing hepatitis B vaccination: a case-control study in the Vaccine Safety Datalink. Toxics 2018 Nov 15;6(4):67. [doi: 10.3390/toxics6040067] [Medline: 30445743]
139. Shemer A, Pras E, Einan-Lifshitz A, et al. Association of COVID-19 vaccination and facial nerve palsy: a case-control study. JAMA Otolaryngol Head Neck Surg 2021 Aug 1;147(8):739-743. [doi: 10.1001/jamaoto.2021.1259] [Medline: 34165512]
140. Panagiotakopoulos L, McCarthy NL, Tepper NK, et al. Evaluating the association of stillbirths after maternal vaccination in the Vaccine Safety Datalink. Obstet Gynecol 2020 Dec;136(6):1086-1094. [doi: 10.1097/AOG.0000000000004166] [Medline: 33156197]
141. Wan EYF, Chui CSL, Lai FTT, et al. Bell's palsy following vaccination with mRNA (BNT162b2) and inactivated (CoronaVac) SARS-CoV-2 vaccines: a case series and nested case-control study. Lancet Infect Dis 2022 Jan;22(1):64-72. [doi: 10.1016/S1473-3099(21)00451-5] [Medline: 34411532]
142. Lai FTT, Li X, Peng K, et al. Carditis after COVID-19 vaccination with a messenger RNA vaccine and an inactivated virus vaccine: a case-control study. Ann Intern Med 2022 Mar;175(3):362-370. [doi: 10.7326/M21-3700] [Medline: 35073155]
143. Palmsten K, Suhl J, Conway KM, et al. Influenza vaccination during pregnancy and risk of selected major structural noncardiac birth defects, National Birth Defects Prevention Study 2006-2011. Pharmacoepidemiol Drug Saf 2022 Aug;31(8):851-862. [doi: 10.1002/pds.5435] [Medline: 35366035]
144. Murata K, Onoyama S, Yamamura K, et al. Kawasaki disease and vaccination: prospective case-control and case-crossover studies among infants in Japan. Vaccines (Basel) 2021 Jul 30;9(8):839. [doi: 10.3390/vaccines9080839] [Medline: 34451964]
145. Sing CW, Tang CTL, Chui CSL, et al. COVID-19 vaccines and risks of hematological abnormalities: nested case-control and self-controlled case series study. Am J Hematol 2022 Apr;97(4):470-480. [doi: 10.1002/ajh.26478] [Medline: 35080247]
146. Lophatananon A, Mekli K, Cant R, et al. Shingles, Zostavax vaccination and risk of developing dementia: a nested case-control study-results from the UK Biobank cohort. BMJ Open 2021 Oct 8;11(10):e045871. [doi: 10.1136/bmjopen-2020-045871] [Medline: 34625411]
147. Chen CC, Lin CH, Chiu CC, et al. Influenza vaccination and risk of stroke in women with chronic obstructive pulmonary disease: a nationwide, population-based, propensity-matched cohort study. Front Med (Lausanne) 2022;9:811021. [doi: 10.3389/fmed.2022.811021] [Medline: 35665329]
148. Wack S, Patton T, Ferris LK. COVID-19 vaccine safety and efficacy in patients with immune-mediated inflammatory disease: review of available evidence. J Am Acad Dermatol 2021 Nov;85(5):1274-1284. [doi: 10.1016/j.jaad.2021.07.054] [Medline: 34363909]
149. Piechotta V, Siemens W, Thielemann I, et al. Safety and effectiveness of vaccines against COVID-19 in children aged 5-11 years: a systematic review and meta-analysis. Lancet Child Adolesc Health 2023 Jun;7(6):379-391. [doi: 10.1016/S2352-4642(23)00078-0] [Medline: 37084750]
150. Watanabe A, Kani R, Iwagami M, et al. Assessment of efficacy and safety of mRNA COVID-19 vaccines in children aged 5 to 11 years: a systematic review and meta-analysis. JAMA Pediatr 2023 Apr 1;177(4):384-394. [doi: 10.1001/jamapediatrics.2022.6243] [Medline: 36689319]
151. Fendler A, de Vries EGE, GeurtsvanKessel CH, et al. COVID-19 vaccines in patients with cancer: immunogenicity, efficacy and safety. Nat Rev Clin Oncol 2022 Jun;19(6):385-401. [doi: 10.1038/s41571-022-00610-8] [Medline: 35277694]
152. Rahmati M, Yon DK, Lee SW, et al. Effects of COVID-19 vaccination during pregnancy on SARS-CoV-2 infection and maternal and neonatal outcomes: a systematic review and meta-analysis. Rev Med Virol 2023 May;33(3):e2434. [doi: 10.1002/rmv.2434] [Medline: 36896895]
153. The Lancet Global Health. Implementing implementation science in global health. Lancet Glob Health 2023 Dec;11(12):e1827. [doi: 10.1016/S2214-109X(23)00523-5] [Medline: 37973323]
154. Peterson HB, Dube Q, Lawn JE, et al. Achieving justice in implementation: the Lancet Commission on Evidence-Based Implementation in Global Health. Lancet 2023 Jul 15;402(10397):168-170. [doi: 10.1016/S0140-6736(23)00870-X] [Medline: 37354916]

Abbreviations

AVSS: active vaccine safety surveillance
Beredt C19: Beredskapregisteret for COVID-19
BIFAP: Base de datos para la Investigación Farmacoepidemiológica en Atención Primaria

XSL•FO
RenderX

CIOMS: Council for International Organizations of Medical Sciences guidelines
FDA: Food and Drug Administration
IA2030: Immunization Agenda 2030
***ICD*:** *International Classification of Diseases*
LMIC: low-income and middle-income country
MDS: Minimum Data Set
PRISM: Post-Licensure Rapid Immunization Safety Monitoring
PRISMA : Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCGP: Royal College of General Practitioners
RSC : Research and Surveillance Centre
SCCS: Self-Controlled Case Series
VSD: Vaccine Safety Datalink
WHO: World Health Organization

Edited by A Mavragani; submitted 12.06.24; peer-reviewed by AS Stergachis, J Knapp; revised version received 21.03.25; accepted 02.04.25; published 17.06.25.

Please cite as:
Zhang M, Yang J, Li Y, Li Y, Li T, Dong Z, Gong S, Wu Y, Ren M, Fan C, Zhang L, Wang Y, Wang Y, Ren J, Sun F, Shen C, Li K, Liu Z, Zhan S
Minimum Data Set and Metadata for Active Vaccine Safety Surveillance: Systematic Review
JMIR Public Health Surveill 2025;11:e63161
URL: https://publichealth.jmir.org/2025/1/e63161
doi:10.2196/63161

© Mengdi Zhang, Junting Yang, Yan Li, Yuan Li, Tong Li, Ziqi Dong, Shuo Gong, Yahui Wu, Minrui Ren, Chunxiang Fan, Lina Zhang, Yi Wang, Yali Wang, Jingtian Ren, Feng Sun, Chuanyong Shen, Keli Li, Zhike Liu, Siyan Zhan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Health Care Services Offered in the HIV Care Continuum via a Machine Learning–Based Topic Modeling Approach: Exploratory Literature Review

SangA Lee[1], PhD; Layoung Kim[1,2], PhD; Mi-So Shim[3], PhD; Gwang Suk Kim[1], PhD

[1]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

[2]College of Nursing and Brain Korea 21 FOUR Project, Yonsei University, Seoul, Republic of Korea

[3]College of Nursing, Keimyung University, Daegu, Republic of Korea

Corresponding Author:
Gwang Suk Kim, PhD
Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, 50-1 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: It remains unclear whether the existing health care services reflect the HIV care continuum, which underscores the need for integrated care beyond viral suppression.

Objective: This study aimed to analyze the literature on health care services for people living with HIV to enhance the understanding of trends and knowledge structures.

Methods: A literature review was conducted using BERTopic, an advanced machine learning-based topic modeling technique. We searched PubMed, CINAHL, EMBASE, and Cochrane databases for English-language studies published between 2013 and 2023. Analyses were performed twice: first, to gain a broad understanding of the literature, and second, to examine the specific details of health care services described.

Results: Among the 11,269 articles screened, 204 studies met the inclusion criteria. Within the HIV care continuum, most studies focused on the treatment retention stage, while studies focusing on the long-term stage were limited. A broad literature analysis identified five key topics, with “ART adherence” emerging as the most prominent topic. A more comprehensive analysis of health care services within the literature revealed 7 topics, reflecting diverse delivery methods and content in providing health care services for people living with HIV. The predominant topic, “ART adherence and counseling,” encompassed the largest number of studies, indicating the strongest emphasis in the field. Notably, the distribution of topics exhibited a distinct pattern: while health care service diversity was the highest in the earlier stages of the HIV care continuum, it became increasingly limited in the later stages.

Conclusions: This study provides valuable insights into current HIV care services and highlights areas for future research and intervention. Despite the shift toward lifelong HIV management, existing literature remains heavily focused on medication treatment, overlooking the multifaceted health care needs of people living with HIV. Research disparities, particularly concerning vulnerable populations, underscore the need for more inclusive studies and tailored health care services. Efforts should be intensified to bridge these gaps, ensuring inclusive and equitable health care services across diverse populations and fostering interdisciplinary collaboration to meet the evolving needs of people living with HIV, thereby enhancing the HIV care continuum for all.

(JMIR Public Health Surveill 2025;11:e65081) doi:10.2196/65081

KEYWORDS

HIV infections; continuity of patient care; health services; review; machine learning

Introduction

Similar to other chronic illnesses, people living with HIV require a lifelong disease-specific regimen to maintain their physical, psychological, and social well-being [1]. While individuals with sustained viral suppression do not risk sexually transmitting the virus to others [2], those with poor psychological health and lower socioeconomic status—measured by factors such as income, housing, and education—are particularly vulnerable to dropping out of care. This vulnerability is exacerbated by HIV-specific barriers, including stigma, discrimination in health care settings, and individual risk factors [3-5]. To address these challenges, health care services must be scaled up to provide targeted support [3,4]. Moreover, as these barriers cannot be

XSL•FO
RenderX

immediately eliminated, a more continuous and comprehensive approach to HIV treatment and life journey is necessary.

The global response to HIV treatment has focused on retaining people living with HIV in care to ensure their healthy lives and to end AIDS [6]. In 2013, the HIV care continuum was established to outline the steps for the continuity of treatment outcomes, namely: diagnosis, linkage to care, treatment initiation, treatment retention, and the achievement and maintenance of viral suppression [7,8]. Institutions such as the World Health Organization (WHO) [9], International Association of Providers of AIDS Care [10], and US Department of Health and Human Services [11] proposed various guidelines, including health care services for people living with HIV on the care continuum at the international and national levels. These guidelines have agreed on the importance of the HIV care continuum, which requires a broader scope of integrated HIV care to achieve holistic health beyond viral suppression [12].

In the context of continuity, the WHO highlights that integrated, people-centered care expands the capacity of patients to play a more active role in long-term survival [13]. Positioned as components of the HIV care continuum, health care services such as HIV testing, medication supply, antiretroviral therapy (ART) adherence monitoring, adverse event management, immunological and virological monitoring, and follow-up for missed visits have been explored in previous studies [14-16]. However, it remains unclear whether the services being offered to people living with HIV align with lifelong continuity of care. To provide effective health care services supporting the HIV care continuum, a comprehensive data review to gain an in-depth understanding of historical trends and key perspectives on the services being provided to people living with HIV must be conducted. There are global, regional, and national differences in coverage, HIV incidence, prevalence, and mortality [17]. Low-income countries have focused on high-priority services, such as child and maternal health, due to resource constraints and immediate health care needs; whereas high-income countries possess the infrastructure, funding, and policy frameworks necessary to support more advanced health care systems. In particular, high-income countries have prioritized the development of integrated care, a prominent policy theme aimed at addressing complex and interconnected health issues, including chronic diseases and the HIV care continuum [18]. Furthermore, high-income countries often serve as benchmarks for global policy development, making their trends and outcomes particularly valuable for informing evaluation and service provision [19]. Their well-established systems, combined with consistent access to resources and comprehensive data, enable the study of integrated care systems in a way that minimizes variability and reduces potential confounding factors associated with differences in economic conditions or institutional capacities. This focus on high-income countries ensures greater comparability and facilitates the derivation of meaningful, context-specific insights. Therefore, health care services provided in high-income countries [20] need to be analyzed first to review the HIV care continuum from the long-term perspective.

As various HIV health care service studies are being conducted across treatment stages, individuals, and methods, the literature needs to be comprehensively reviewed. Traditional methods, such as manual literature reviews and qualitative content analysis, while valuable for deriving specific insights from research criteria, face significant limitations in handling large and complex datasets. These methods are time-intensive, rely heavily on human expertise, and are subjective to judgment, making them prone to inconsistency as well as bias [21,22]. Furthermore, these methods are less effective at detecting overarching trends, correlations, and topic evolution over time, particularly in diverse and extensive datasets. Manual keyword extraction and topic grouping often fail to fully capture nuanced relationships between concepts, limiting the ability to generate a comprehensive and graphical perspective of the literature. To review, summarize, classify, and detect trends more efficiently in large-scale documents, text analysis techniques offer a valuable and effective solution [23,24].

To address these challenges, this study used Bidirectional Encoder Representations from Transformers (BERT) developed by Devlin [25] in Google artificial intelligence (AI) Language. Among various text analysis technical models, BERT is recognized as one of the most recent and advanced machine learning-based approaches to topic modeling. Unlike traditional text analysis methods, techniques like BERT offer several advantages. Its ability to pretrain deep bidirectional representations enables BERT to capture the contextual meanings of words and sentences by considering both left and right contexts simultaneously [25], resulting in more accurate and nuanced topic modeling. An example of traditional methods is the Latent Dirichlet Allocation (LDA), which relies on simpler probabilistic frameworks and often fails to account for complex semantic relationships [25,26]. Studies comparing LDA and BERT have demonstrated that BERT outperforms LDA in extracting deeper and distinct topics, identifying intricate topic relationships in datasets with complex linguistic structures [27,28]. BERT is particularly effective at processing large datasets, efficiently summarizing, classifying, and detecting trends across a wide range of documents. Considering its strengths, it has been increasingly used in research, particularly literature reviews analyzing numerous studies as well as social media analyses processing numerous tweets to extract insights on health care issues [29-32]. Its scalability reduces the time and resources required for manual analysis while ensuring consistency in topic extraction and classification. In addition, BERT excels in modeling topic correlations as well as their evolution over time, providing a dynamic and graphical perspective of the data that traditional methods struggle to achieve. Its ability to handle complex linguistic structures and low-resource tasks enables the identification of meaningful patterns in diverse contexts, further enhancing its utility for analyzing comprehensive datasets [33]. While BERT offers numerous strengths, it effectively addresses the limitations of traditional approaches by ensuring objectivity, scalability, and the ability to generate insights that would otherwise require extensive human effort and expertise. These capabilities make BERT particularly well-suited for reviewing the extensive and heterogeneous body of literature on HIV health care services. This study therefore used BERT to depict the bibliographic trends and knowledge structure of publications on health care

XSL•FO
RenderX

services to improve the HIV care continuum for people living with HIV in high-income countries.

Methods

Design

This study used a quantitative content analysis using BERTopic modeling to examine the keywords and topics in the literature on health care services provided to people living with HIV. The methodological process comprised phases: (1) search strategy and selection, (2) data preprocessing, and (3) data analysis, using topic modeling with BERTopic.

Search Strategy and Selection Process

Relevant literature was searched from PubMed, CINAHL, EMBASE, and Cochrane databases. Key search terms included Medical Subject Heading (MeSH) terms, trade words, and natural language words such as "care continuum," "HIV/AIDS," and "service." The search was limited to studies published between January 2013—when the WHO first introduced comprehensive guidelines on ART for all populations across the HIV care continuum—and December 2023 [34]. The inclusion criteria were studies: (1) targeting adults diagnosed with HIV, (2) investigating or providing health care services (eg, interventions, services, or programs) to improve the HIV care continuum, and (3) any study design conducted in high-income countries. The exclusion criteria were: (1) studies targeting specific subpopulations (eg, drug users, individuals with mental illnesses, or sex workers), (2) studies targeting HIV prevention services, (3) studies for which the full text was unavailable in English, (4) studies targeting public policies, (5) literature reviews, and (6) duplicates. The exclusion of studies on specific subpopulations aimed to focus the review on general HIV care rather than focusing on tailored interventions for specific groups. While these populations face unique challenges and require specialized interventions, their exclusion was conducted to provide a broader perspective on the HIV care continuum. This approach ensures that the analysis reflects the overall capacities and integration efforts of health care systems, offering insights applicable to the wider population rather than a subset of individuals.

In total, 14,377 articles were retrieved from the databases. After removing duplicates, 11,269 articles were screened for their titles, abstracts, and full texts to determine their eligibility. Multiple studies using the same health care service were assessed through text frequency analysis to detect keyword matches. Most exhibited few keyword matches, suggesting distinct studies. Subsequently, all studies, with the exception of secondary analyses, underwent a full-text review. Three researchers (SL, MS, and LK) independently screened all the articles and selected the final list of articles for analysis after discussion. Discrepancies were discussed with the principal investigator (GK) and consensus was reached as to whether to include the article. The dataset size selected was deemed appropriate for BERT, as it demonstrated efficacy in handling modest-sized datasets with a comparable scale in previous studies, often using 100 documents [28,35]. As with many bibliometric analysis studies exploring research trends [36], a quality assessment was unnecessary. All the data, including the metadata elements of the title, abstract, authors, and year, were downloaded for analysis.

Data Preprocessing

Before data analysis, a sequence of preprocessing steps were conducted to clean and prepare the data for further analysis. Articles without abstracts were excluded to ensure a comprehensive set of textual information. Additional columns were created to classify studies by country, HIV care continuum, and health care services provided. The countries from which each study collected data and the specific stages of the HIV care continuum on which each study focused were identified to derive more meaningful interpretations of the distribution of research topics. Specific stages of the HIV care continuum, including linkage to care, treatment initiation, treatment retention, and the long-term stage [7,8], were identified. The classification was finalized following a consensus reached through discussion between 2 researchers (SL and LK), with multiple stages being investigated in one study. After incorporating studies covering multiple stages, the final corpus comprised 242 documents for analysis. To ensure a more thorough investigation of health care services, detailed descriptions of services from the methods sections of each article were extracted. This process, conducted by 2 researchers (SL and LK), enhanced the accuracy of the dataset. The usage of details regarding the health care services described in the methods section helped create a more robust text corpus, enhancing the efficacy of the subsequent topic modeling analysis.

Stop words were then excluded, and a replacement dictionary was generated. Stop words, including "a," "the," and "of," refer to common words that appear frequently in the text but do not convey insights into the specific topic of a document. All English stop words were removed using the Natural Language Toolkit library in Python. Words that barely contained significant meanings for the research purpose and degraded or hindered the interpretability of results (eg, background, method, results, and trial registration) were added to the toolkit. The extended stop words were then excluded from the data. The replacement dictionary was used to combine several words or abbreviations with the same or similar meanings and designate a single representative word (eg, people living with HIV, HIV, and AIDS). During this stage, repeated word analyses were performed with regard to abstract, and the researchers reached consensus on the words to be registered in the dictionaries through discussion.

Data Analysis

BERTopic, a Python-based package, was used to extract the primary topics from the literature. It was implemented with the all-MiniLM-L6-v2 embedding model to ensure accurate and efficient document embedding, leveraging its compact yet powerful transformer architecture to generate high-quality sentence representations for topic modeling [37,38]. The BERTopic process consists of three steps: (1) document embedding, (2) document clustering, and (3) topic representation [39,40]. Document embedding, using the all-MiniLM-L6-v2, reads sentences in both directions, pretrains them in deep representations, and then generates word embeddings (ie, representations of valued vectors transformed from words to

XSL•FO
RenderX

numbers depending on the surrounding contexts). Document clustering applies Uniform Manifold Approximation and Projection to reduce the dimensions [41] and cluster semantically similar words through the Hierarchical Density-Based Spatial Clustering of Applications with the Noise algorithm [42]. Topic representation extracts significant words on class-based term frequency-inverse document frequency, which is a metric to measure the relevance of words [43].

To validate the findings, LDA—a widely used topic modeling method—was also applied. Since no definitive method exists to determine the number of topics, the research team conducted multiple discussions to assess topics generated by the BERT-based approach. The final number of topics was determined by considering evaluation metrics, intertopic distance maps, data characteristics, and interpretability of the results, in comparison to LDA [25,44,45]. The evaluation metrics used to assess topic quality were topic coherence and topic diversity, both ranging from 0 to 1. Topic coherence measures how frequently words co-occur in a dataset, with higher values indicating more interpretable topics, while topic diversity assesses the uniqueness of words within topics, with higher values indicating more distinct topics [46]. The intertopic distance map visualizes topics as circles in a 2D space, where the size of each circle represents the number of studies assigned to that topic, and the proximity between circles indicates the degree of similarity between topics [47]. In this study, the BERT-generated model demonstrated superior evaluation metrics and a more balanced distribution of topics in the intertopic distance maps, with minimal overlap. In addition, the BERT approach yielded more data-related, distinct, and plausible results, outperforming LDA topic modeling in both structure and longitudinal topic trends. In line with the common practice in prior studies using BERTopic, Topic -1—a nonthematic cluster of documents with weak thematic alignment and often containing words commonly found across different topics [48]—was excluded from the interpretation of results. The proportion of nonthematic documents assigned to Topic -1 varied among the studies, ranging from approximately zero to 54% [25,36]. In this study, the analyses were performed twice, with one aimed at broadening the understanding of the literature related to health care services for people living with HIV and the other focused on a specific exploration of the health care services described in the literature.

Ethical Considerations

The study was exempted from the approval process because it was a review and analysis of existing literature. No human participants were not directly involved, and no personal or private information was used.

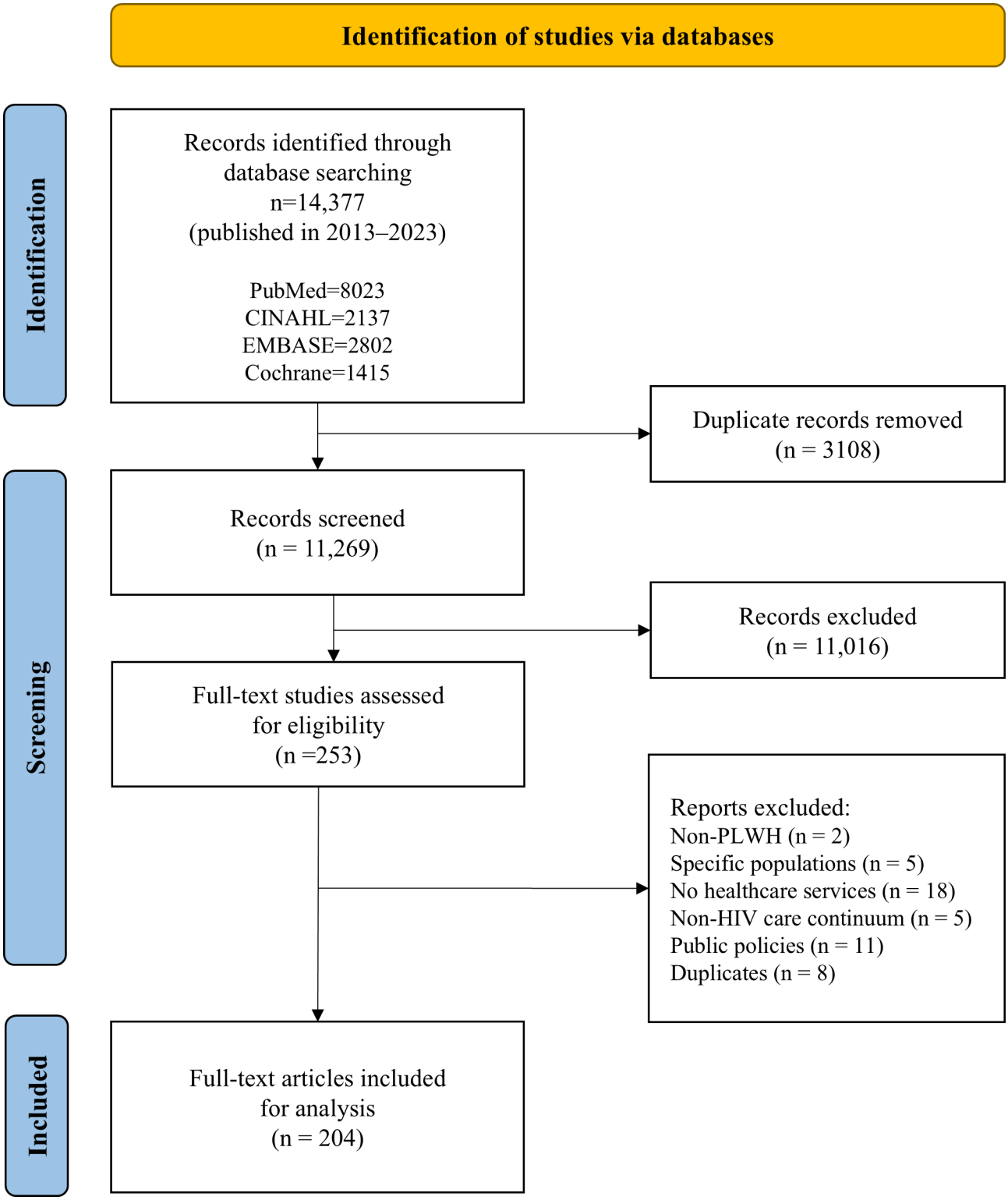
Results

Data Description

A total of 204 studies were included in the analysis (Figure 1). Over the past decade, an average of approximately 19 studies on health care services for people living with HIV were conducted annually, ranging from 10 to 31 studies (Figure 2). The number of publications increased significantly between 2018 and 2020, but declined in 2016 and from 2021 onwards. Table 1 summarizes the characteristics of the included studies. The majority of studies were conducted in the United States (n=174, 85.3%). Regionally, as defined by the WHO [49], the Region of the Americas (AMR) accounted for the highest representation, while the Eastern Mediterranean Region was represented by only one study. The average age of study participants was 40.52 (SD 10.67) years, with the longest recorded duration since their HIV diagnosis being 30 years. Among the HIV care continuum stages, the treatment retention stage comprised the highest number of studies, totaling 185, while the long-term stage had the fewest studies.

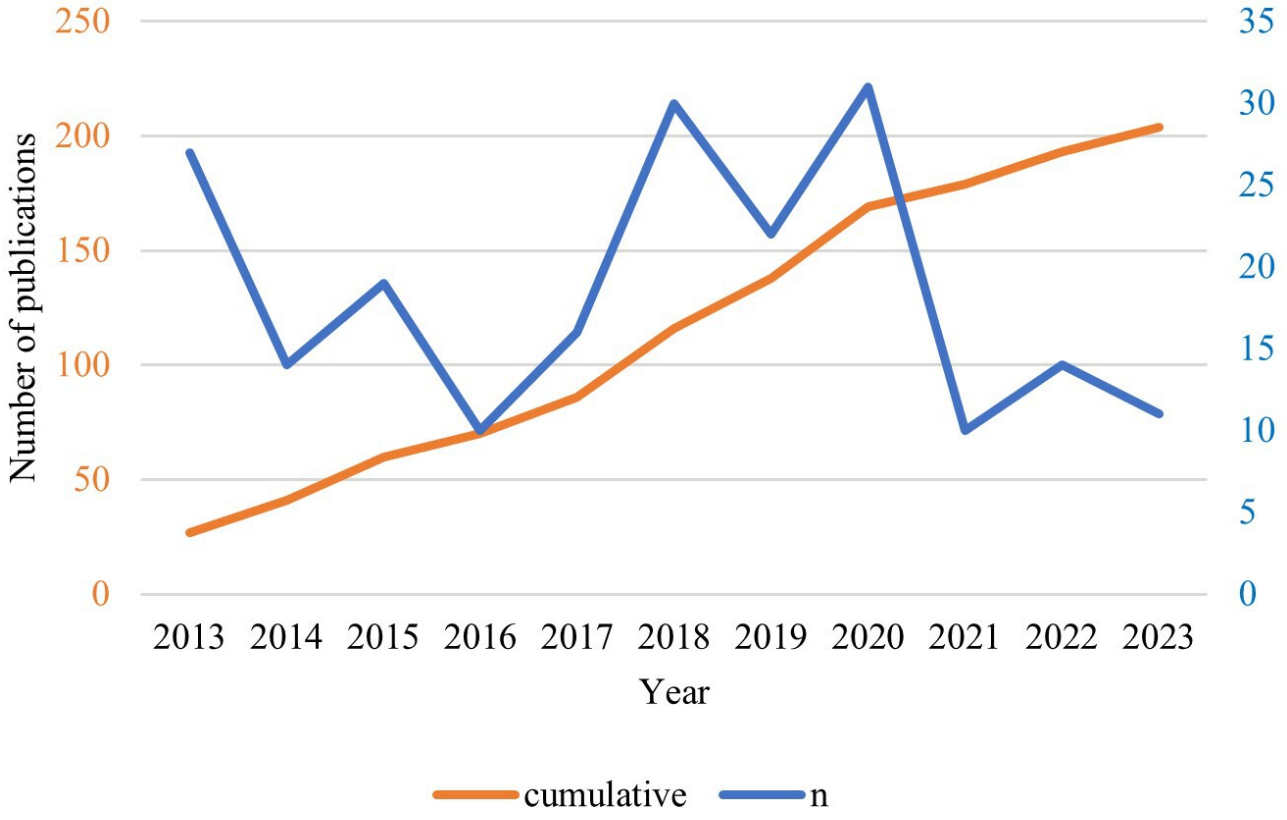
XSL•FO
RenderX

Figure 1. PRISMA flowchart for article selection.



XSL•FO
RenderX

Figure 2. Number of publications on health care services for people living with HIV from 2013 to 2023.



XSL•FO
RenderX

Table . Characteristics of the analyzed studies (N=204).

Characteristics	Values
Research design, n (%)	
Quantitative	164 (80.4)
Qualitative	22 (10.8)
Mixed methods	18 (8.8)
Sample	
Size, min-max	10 - 32,999
Age, years	18 - 85
Period of HIV diagnosis, years	0 - 30
Country, n (%)	
Region of the Americas, n (%)	
United States	174 (85.3)
Canada	6 (2.9)
European Region, n (%)	
Spain	8 (3.9)
United Kingdom	4 (2.0)
Denmark	1 (0.5)
Italy	1 (0.5)
Netherlands	1 (0.5)
Poland	1 (0.5)
Switzerland	1 (0.5)
Western Pacific Region, n (%)	
Australia	3 (1.5)
Korea	2 (1.0)
Singapore	1 (0.5)
Oman	1 (0.5)
HIV care continuum[a], %	
Linkage to care	17
Treatment initiation	29
Treatment retention	185
Long-term	11

[a]Records duplicates.

Topic Modeling on Health Care Services Research for People Living With HIV

BERTopic was initially applied to the abstract corpus to identify overarching trends and the knowledge structure within the included publications. Among the 242 documents, 191 were classified into 5 distinct topics, while the remaining 51 (21.1%) were identified as a nonthematic topic. This proportion is relatively small compared to previous studies, where the proportion ranged from 0% to 54% [25,36]. The most frequently occurring keywords within the nonthematic topic included “hiv,” “intervention,” “health,” “use,” and “improve,” which represent broad and general concepts rather than distinct thematic categories.

Table 2 presents the 5 distinct topics, labelled by the research team, alongside their respective keywords and the number of documents assigned to each topic. The 2 prominent topics—“ART adherence” and “Retention and engagement in care”—align with key stages of the HIV care continuum. The remaining 3 topics focus on specific populations: “Minorities-focused” consists of studies concentrating on race and ethnicities, while “Young population-focused,” emerged as an age-specific category. In addition, “Newly diagnosed men who have sex with men (MSM)” encompasses health care services tailored to a specific population at a specific stage.

XSL•FO
RenderX

Table . Topics and top keywords identified in health care services research for people living with HIV.

Topics	Top keywords (importance weights)	n
ART[a] adherence	adherence (0.0470), intervention (0.0274), hiv (0.0267), art (0.0255), use (0.0229), medication (0.0221), group (0.0184), people living with HIV (0.0172), improve (0.0164), self (0.0158)	120 (62.8%)
Retention and engagement in care	retention (0.0575), hiv (0.0505), intervention (0.0344), clinic (0.0253), health (0.0226), peer (0.0221), program (0.0209), improve (0.0208), adult (0.0191), engagement (0.0175)	29 (15.2%)
Minorities-focused	hiv (0.0451), component (0.0357), intervention (0.0292), black (0.0251), interview (0.0241), american (0.0237), african (0.0219), african-american (0.0219), use (0.0214), latino (0.0209)	16 (8.4%)
Young population-focused	intervention (0.0641), adherence (0.0586), youth (0.0504), adolescent (0.0344), art (0.0342), hiv (0.0336), condition (0.0292), accept (0.0261), medication (0.0246), group (0.0240)	14 (7.3%)
Newly diagnosed MSM[b]	test (0.0572), hiv (0.0496), ltc (0.0428), clinic (0.0387), client (0.0341), msm (0.0303), diagnose (0.0301), hiv-test (0.0250), case (0.0234), men (0.0216)	12 (6.3%)

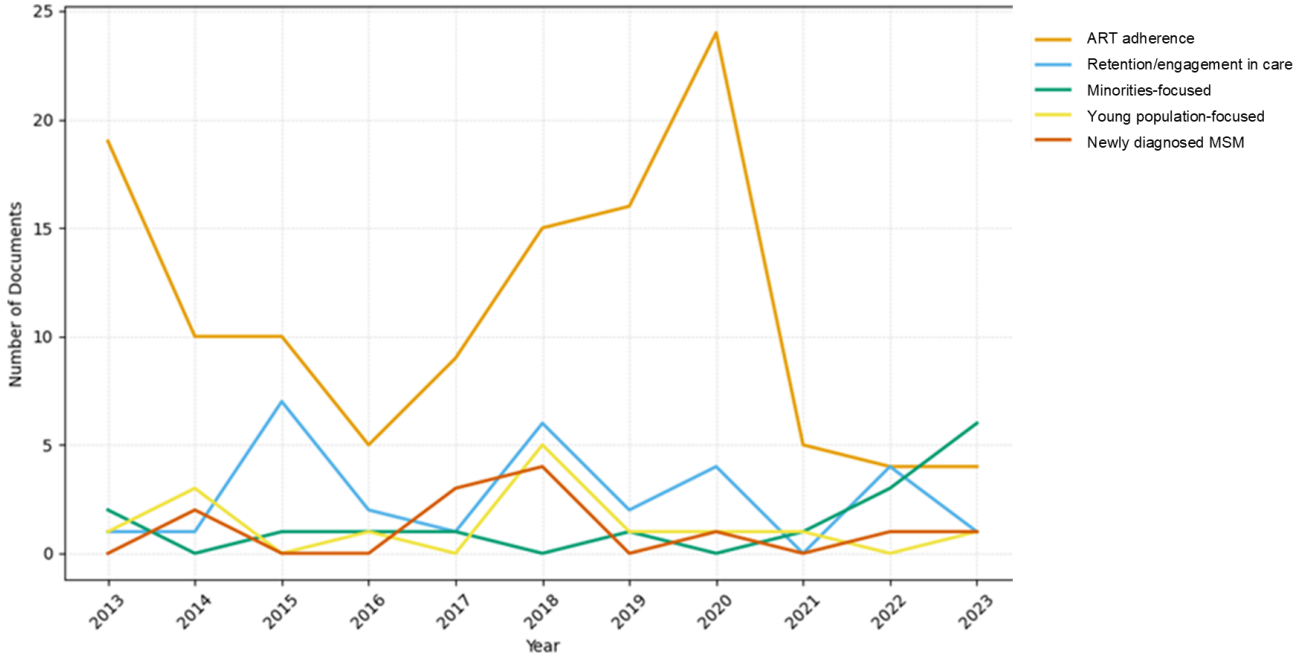
[a]ART: antiretroviral therapy.

[b]MSM: men who have sex with men.

When comparing the evaluation metrics of the models generated using LDA and BERTopic, the topic coherence scores were 0.2987 and 0.4264, respectively, while the topic diversity scores were 0.4200 and 0.6667, respectively, indicating superior performance of BERTopic in both aspects. In addition, the intertopic distance map of the BERT-generated model presented a more well-balanced distribution of topics with no overlap, whereas the LDA-generated model exhibited substantial topic overlap (Multimedia Appendix 1). Considering these comparisons, including topic interpretability, the BERTopic-generated model was determined to be the optimal choice for analysis.

Figure 3 depicts the distribution of topics over the years. The research focus on “ART adherence” peaked in 2020, while notable declines were observed during mid-2010 and after 2020. “Retention and engagement in care” exhibited fluctuating trends. Coinciding with the onset of the COVID-19 pandemic, studies involving people living with HIV tended to decrease post 2020, but research on “Minorities-focused” topic exhibited recent increases.

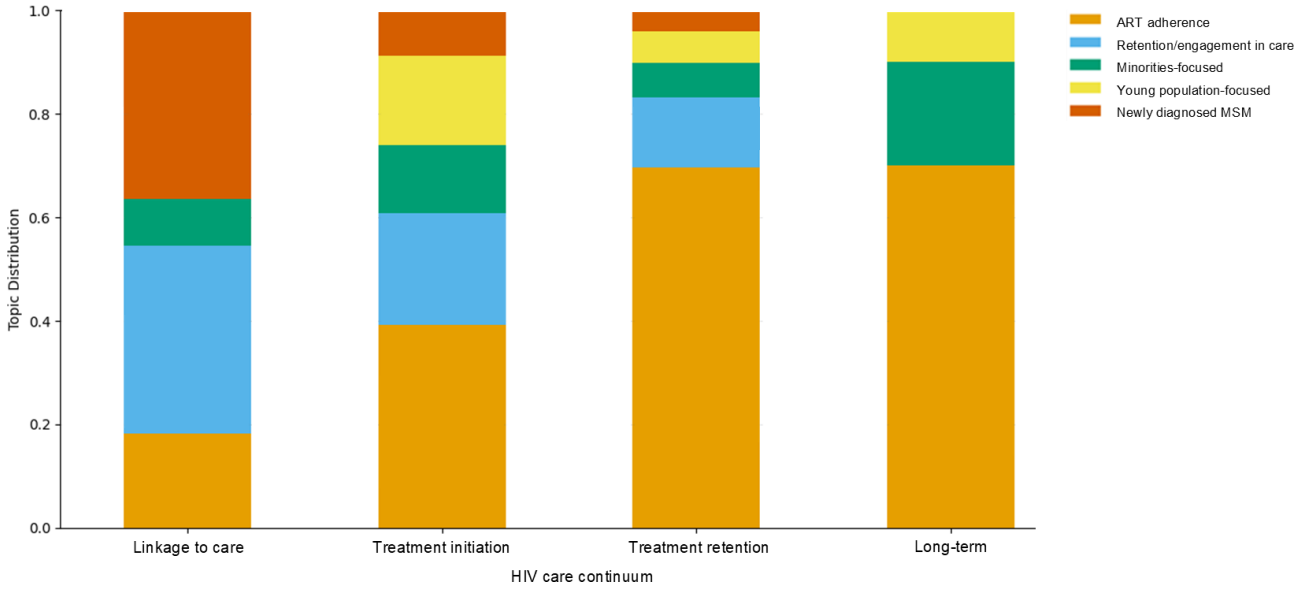
Figure 3. Topic distribution in health care services research for people living with HIV from 2013 to 2023. ART: antiretroviral therapy.



XSL•FO
RenderX

Among the 4 stages of the HIV care continuum, the treatment initiation and retention stages were investigated across all topics (Figure 4). In the first stage (linkage to care), research on young populations was notably lacking, whereas the "Newly diagnosed MSM" topic had the highest representation compared to other stages. In the treatment initiation stage, the "Young population-focused" topic began to emerge, with all topics being more evenly distributed. The "ART adherence" topic accounted for more than half of the studies in the treatment retention stage. Finally, the long-term stage had the fewest studies overall. Regarding the regional distribution of topics by country, only the AMR accounted for all topics, while studies from other regions were limited. Research conducted in the AMR predominantly focused on ART adherence.

Figure 4. Distribution of topics in health care services research for people living with HIV across the HIV care continuum. ART: antiretroviral therapy; MSM: men who have sex with men.



Topic Modeling on Health Care Services for People Living With HIV

The second analysis aimed to provide more comprehensive examination of health care services by using the detailed descriptions of health care services found in the methods section of the articles. When running BERTopic, 181 out of the total 242 documents were categorized into 7 topics, while the remaining 61 (25.2%) fell into a nonthematic category. The most frequent occurring keywords in the nonthematic topic included "care," "participant," "hiv," "patient," and "health," suggesting that these documents contained broad, general concepts rather than distinct thematic distinctions.

Table 3 presents the 7 topics, as labeled by the research team, along with their respective keywords and the number of documents assigned to each. Among these, "ART adherence and counseling" was the most prevalent topic, reflecting strategies to enhance ART adherence, as indicated by the keywords "counseling" and "message." Other topics explore more specific aspects of health care service delivery models, target populations, and service content. The "Peer support" and "Pharmacist involvement" topics highlight the engagement of both health care and non-health care professionals in studies providing health care services to people living with HIV. The "Pharmacist involvement" and "Youth-to-adults transition care" topics specify who delivers health care services and for whom, respectively. "Monetary incentives" pertain to studies using financial rewards as benefits for people living with HIV. The topic "Health promotion and care navigation" focuses on structured interventions designed to guide individuals in maintaining and improving health. Finally, the "Linkage to care" represents studies addressing a specific stage of the HIV care continuum. As reflected in their labels, the 7 topics collectively reflect a combination of delivery methods and content in the provision of health care services to people living with HIV.

XSL•FO
RenderX

Table . Topics and top keywords identified as health care services for people living with HIV.

Topics	Top keywords (importance weights)	n
ART[a] adherence and counseling	adherence (0.0312), participant (0.0283), session (0.0280), medication (0.0212), intervention (0.0209), use (0.0186), patient (0.0184), message (0.0150), counsel (0.0139), receive (0.0128)	81 (44.8%)
Peer support	care (0.0369), hiv (0.0344), peer (0.0246), health (0.0229), intervention (0.0211), support (0.0196), participant (0.0190), patient (0.0182), session (0.0176), barrier (0.0172)	40 (22.1%)
Youth-to-adults transition care	care (0.0572), hiv (0.0473), patient (0.0435), team (0.0329), clinic (0.0276), transition (0.0253), visit (0.0222), service (0.0213), provide (0.0200), adult (0.0183)	20 (11.0%)
Pharmacist involvement	art (0.0606), patient (0.0568), medication (0.0389), pharmacist (0.0382), hiv (0.0375), clinic (0.0304), day (0.0292), drug (0.0283), pharmd (0.0246), hiv-pharmd (0.0246)	17 (9.4%)
Monetary incentives	chance (0.0482), participant (0.0449), point (0.0440), chance-win (0.0437), win (0.0437), hiv (0.0416), component (0.0395), prize (0.0351), tm (0.0328), test (0.0320)	9 (5.0%)
Health promotion and care navigation	hiv-care (0.0662), care (0.0613), hiv (0.0611), health (0.0514), health-promotion (0.0493), promotion (0.0490), care-navigation (0.0480), message (0.0479), navigation (0.0435), participant (0.0434)	8 (4.4%)
Linkage to care	care (0.0684), hiv (0.0554), site (0.0545), people living with HIV (0.0407), hiv-care (0.0347), classify (0.0341), ltc (0.0336), ed (0.0312), program (0.0311), patient (0.0289)	6 (3.3%)

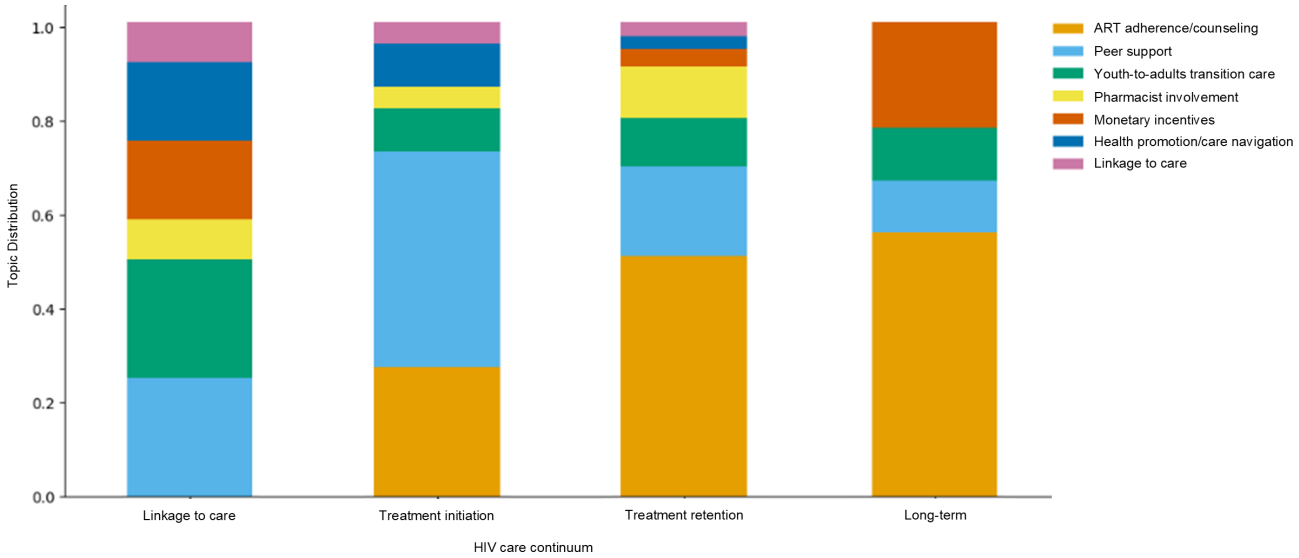
[a]ART: antiretroviral therapy.

When comparing the evaluation metrics of the models generated using LDA and BERTopic, the topic coherence scores were 0.3799 and 0.4458, respectively, while the topic diversity scores were 0.4429 and 0.6125, respectively, demonstrating BERTopic's superior performance in both measures. Furthermore, the intertopic distance map of the BERT-generated model exhibited a well-balanced distribution of topics with no overlap, whereas the LDA-generated model displayed overlapping topics (Supplementary Figure 1 in Multimedia Appendix 1). Considering these factors, including topic interpretability, the BERTopic-generated model outperformed the LDA-generated model and was ultimately selected as the preferred approach.

Figure 5 illustrates the distribution of health care service topics across each stage of the HIV care continuum, highlighting varying levels of research coverage. While research on the initial linkage to care and final long-term stages were similarly scarce, the distribution of health care service topics presented distinct patterns. At the linkage to care stage, all health care services were represented except those related to medication adherence. In the treatment initiation stage, services focusing on medication adherence began to emerge, with peer support being the most prominent. During the treatment retention stage, medication adherence-focused services accounted for more than half of the studies. In the long-term stage, services related to medication adherence and monetary incentives became more prominent, whereas services involving pharmacist involvement and health navigation, which were present in earlier stages, declined. Overall, the diversity of health care services was the highest in the earlier stages of the continuum but became less varied in later stages, where research was more limited.

XSL•FO
RenderX

Figure 5. Distribution of health care service topics for people living with HIV across the HIV care continuum. ART: antiretroviral therapy; MSM: men who have sex with men.



Discussion

Principal Findings

The HIV care continuum is a framework that describes the treatment journey of people living with HIV, encompassing the stages from diagnosis to linkage to care, engagement in care, and viral suppression through ART [7]. Key health care service components within this continuum include HIV testing, medication supply, treatment adherence monitoring, adverse event management, as well as virological monitoring [14-16]. However, like other chronic diseases, the HIV care continuum requires a shift toward an integrated care system centered on patient self-management, equipping individuals with the health care resources needed to manage their condition throughout life [50]. Moreover, as patients are increasingly expected to take active roles in managing their success of HIV treatment and long-term health [51], it is critical to move beyond the sole objective of viral suppression and emphasize overall quality of care for all people living with HIV [12]. Nonetheless, the extent to which current services align with the principles of lifelong continuity of care remains unclear. To address this gap, this study aimed to enhance the understanding of health care service trends and knowledge structures within the HIV care continuum by reviewing the literature using BERT. Given that approximately a decade has passed since the global recognition of the HIV care continuum, reviewing research in this field is essential for identifying both dominant areas and those lacking attention and support.

In this study, research on health care services within the HIV care continuum has primarily focused on the treatment retention stage, with a notable emphasis on ART adherence. While the importance of the HIV care continuum is widely recognized [12], research beyond disease-focused models remains underexplored in the field. Notably, despite longstanding concerns regarding the quality of life (QoL) of people living with HIV—including issues such as comorbidities, stigma, depression, discrimination, and sexual health [52]—these topics were not prominently represented in the extracted themes. The dominance of treatment-focused topics reveals persistent gaps in the literature, even amid global calls for more comprehensive, continuum-based HIV care. To broaden perspectives on life beyond treatment, future research should embrace emerging concepts such as “long-term success” [53] and “healthy living” [54] for people living with HIV, which emphasize a holistic approach addressing the multifaceted needs of people living with HIV. Optimizing health-related QoL and addressing age-related concerns is crucial to ensuring people living with HIV can lead fulfilling lives. This highlights the importance of multidisciplinary teams in delivering integrated, person-centered care and support for people living with HIV.

From a population perspective, there remains a notable lack of research on health care services tailored to minority populations. The topic of MSM emerged as a distinct research focus, reflecting the longstanding dominance of this demographic among people living with HIV. Consequently, research targeting MSM has been abundant. However, research addressing racial and ethnic minorities was consolidated into a single topic, and no distinct topic related to older adults emerged. This gap may be attributable to challenges in engaging these populations, particularly during the COVID-19 pandemic. Such academic trends inadequately address the global HIV epidemic. In high-income countries such as the US, HIV prevalence and incidence estimates are disproportionately higher among Black or African American and Hispanic or Latino individuals compared to White individuals. In addition, women are disproportionately affected among acquiring HIV through heterosexual contact, highlighting their vulnerability despite lower numerical representation [55]. The disease burden attributable to HIV infection among older adults, including prevalence, mortality, and QoL concerns, has increased [56]. These data underscore the critical need for comprehensive actions across the HIV care continuum and increased support tailored to younger and older populations, irrespective of race, ethnicity, and gender. Prioritizing and funding research aimed at understanding the specific needs of diverse and vulnerable populations is critical to fostering a more equitable health care landscape for HIV-infected individuals.

XSL•FO
RenderX

The topics identified on health care services aligned well with the needs of people living with HIV. Peer support and youth-to-adult transition services were prominent in the early stages of the HIV care continuum, when people living with HIV struggle with limited information following diagnosis and must decide whether to undergo ART. This period also represents a high-risk time for disengagement, particularly among young adults with HIV transitioning from pediatric to adult-oriented care [57,58]. ART adherence and counseling services were primarily implemented in the subsequent stages, supporting people living with HIV as they continued treatment and worked toward clinical stability [8]. While the health care services identified in the literature corresponded to different stages of the HIV care continuum, they were predominantly concentrated in the treatment retention stage, with fewer and less diverse services available in other stages. Therefore, more health care services should be developed and implemented across the HIV care continuum, particularly at the long-term stage. Specifically, more varied services to facilitate treatment initiation for people living with HIV at the initial stage are necessary, along with supportive services to promote healthy living and aging with HIV in the long term [59]. Furthermore, the cost-effectiveness, feasibility, sustainability, and confidentiality of these services must be further investigated, particularly for marginalized, younger, and older people living with HIV [60,61].

Pharmacists emerged as a distinct topic in health care services, whereas nurses and physicians were not separately identified. This may be attributed to the unique nature of HIV infection, which heavily relies on ART. Since the introduction of the first ART drug in 1987, ART has been continuously and actively developed and evaluated [62]. Furthermore, the WHO has continuously issued updated evidence-based ART guidelines since publishing its first consolidated guidelines in 2013 [33]. Recent trends have emphasized the cost-effectiveness and accessibility of ART, alongside advancements in potential cures and vaccines [62]. Given the dependency on ART and active trends in its advancement, pharmacies serve as crucial locations for the delivery of HIV treatment, testing, and prevention services, although the scope of services varies among countries. Pharmacists' attitudes and services that are favorable to people living with HIV are essential for enhancing the HIV care continuum [63]. Expanding the services offered by pharmacies and pharmacists as primary service providers for testing and medication, particularly for high-risk and vulnerable populations, should be key consideration [64].

Although physicians and nurses play fundamental roles in HIV care, addressing evolving needs for health and longevity requires extending care beyond primary care to optimal long-term management beyond treatment. This need is particularly pertinent in high-income countries with abundant resources [65,66]. Nurses' practices have been closely aligned with the needs of people living with HIV in HIV care, contributing to interdisciplinary efforts that facilitate the active development and implementation of interventions across the HIV care continuum [67]. However, their involvement in policymaking at local, national, and international levels remains limited [67]. Given their deep understanding of the needs of people living with HIV, nursing professionals should actively participate in decision-making processes to reshape practices and policies within the HIV care continuum. Furthermore, while not a distinct topic in this study, social welfare professionals also play a crucial role in the long-term management of HIV.

In addition, involving family members in health care services is recommended to optimize the HIV care continuum, particularly for key populations such as women and adolescents [10]. Numerous studies advocate a holistic patient-centered approach for people living with HIV, which involves not only pharmacists and peers, but also partners, family, and other individuals. Ensuring a comprehensive approach that encompasses a diverse range of health care and non-health care professionals over the long term is critical for promoting the HIV care continuum, particularly for vulnerable populations.

Overall, a decline in the number of studies on HIV services was observed, likely influenced by the impact of the COVID-19 pandemic. Notable drops in research output occurred between 2014 and 2016, and again in 2021, coinciding with global outbreaks of infectious diseases such as Ebola, Zika, and COVID-19 [68]. Similar to the upwards trend observed after the resolution of outbreaks such as Ebola and Zika, research efforts may increase shortly after the end of the COVID-19 pandemic. Therefore, research published from this perspective is important. As the vast majority of HIV research is conducted in the US, findings from this context may not be generalizable to other countries owing to significant sociocultural differences. Expanding research efforts across diverse countries and tailoring health care services to the specific needs of each country is crucial.

This study has several limitations. First, it focused solely on studies from high-income countries; studies from middle- and low-income countries may present different characteristics. Second, studies that did not explicitly address the HIV care continuum were excluded, necessitating careful interpretation of the findings. Finally, our trend analysis relied on linear models, potentially oversimplifying the intricate evolution of topics influenced by multiple factors.

Conclusions

Despite the evolving nature of HIV care toward lifelong management, the literature predominantly emphasizes medical treatment, highlighting significant gaps in addressing the multifaceted needs of people living with HIV. This study provides a novel contribution by analyzing health care services aimed at enhancing HIV care continuum for people living with HIV using BERT, an advanced machine learning-based topic modeling method. By examining research trends across different stages of the HIV care continuum, this study offers valuable insights into both overarching research trends and specific service-related themes. Notably, it extends beyond traditional treatment-focused disease models to encompass long-term management and holistic well-being considerations, underscoring the importance of interdisciplinary collaboration in meeting the evolving needs of people living with HIV. Overall, this research advances understanding of HIV care and provides critical directions for future interventions and studies in this field.

XSL•FO
RenderX

Acknowledgments

This work was supported by the National Research Foundation of Korea funded by the Korean government (Ministry of Science and ICT) [Grant no. 2020R1A2C101081713], the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education [Grant no. 2020R1A6A1A03041989], and a Faculty Research Grant from Yonsei University College of Nursing [Grant no. 6-2022-0034]. Layoung Kim received a scholarship from the Brain Korea 21 FOUR Project, funded by the National Research Foundation of Korea, Yonsei University College of Nursing.

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Figure 1. Comparison of inter-topic distance maps generated utilizing BERTopic and LDA.
[PNG File, 923 KB - publichealth_v11i1e65081_app1.png]

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) Checklist.
[DOCX File, 109 KB - publichealth_v11i1e65081_app2.docx]

References

1. Swendeman D, Ingram BL, Rotheram-Borus MJ. Common elements in self-management of HIV and other chronic illnesses: an integrative framework. AIDS Care 2009 Oct;21(10):1321-1334. [doi: 10.1080/09540120902803158] [Medline: 20024709]
2. The Lancet. U=U taking off in 2017. Lancet HIV 2017 Nov;4(11):e475. [doi: 10.1016/S2352-3018(17)30183-2] [Medline: 29096785]
3. Choi JP, Seo BK. HIV-related stigma reduction in the era of undetectable equals untransmittable: the South Korean perspective. Infect Chemother 2021 Dec;53(4):661-675. [doi: 10.3947/ic.2021.0127] [Medline: 34979602]
4. Safreed-Harmon K, Anderson J, Azzopardi-Muscat N, et al. Reorienting health systems to care for people with HIV beyond viral suppression. Lancet HIV 2019 Dec;6(12):e869-e877. [doi: 10.1016/S2352-3018(19)30334-0] [Medline: 31776099]
5. Moen M, Doede M, Johantgen M, et al. Nurse-led hospital-to-community care, clinical outcomes for people living with HIV and health-related social needs. J Adv Nurs 2023 May;79(5):1949-1958. [doi: 10.1111/jan.15485] [Medline: 36345144]
6. Joint United Nations Programme on HIV/AIDS. Prevailing against pandemics by putting people at the centre: world AIDS day report 2020. : Joint United Nations Programme on HIV/AIDS; 2020 URL: https://www.unaids.org/en/resources/documents/2020/prevailing-against-pandemics [accessed 2024-03-27]
7. Babalola S, Van Lith LM, Mallalieu EC, et al. A Framework for health communication across the HIV treatment continuum. J Acquir Immune Defic Syndr 2017 Jan 1;74(Suppl 1):S5-S14. [doi: 10.1097/QAI.0000000000001206] [Medline: 27930606]
8. Kim GS, Kim Y, Park MK, Lee SA, Lee Y, Shim MS. Barriers and facilitators on the HIV care continuum among adults living with HIV in high-income countries: a scoping review protocol. Syst Rev 2022 Oct 15;11(1):222. [doi: 10.1186/s13643-022-02097-x] [Medline: 36243736]
9. World Health Organization. HIV Treatment: Guidelines for Managing Advanced HIV Disease and Rapid Initiation of Antiretroviral Therapy: Policy Brief; 2017. URL: https://www.who.int/publications/i/item/9789241550062 [accessed 2024-03-27]
10. International Advisory Panel on HIV Care Continuum Optimization. IAPAC guidelines for optimizing the HIV care continuum for adults and adolescents. J Int Assoc Provide AIDS Care 2015;14(1):S3-S34. [doi: 10.3389/10.1177/2325957415613442] [Medline: 26527218]
11. The White House. National HIV/AIDS strategy for the united states 2022–2025. 2021 URL: https://files.hiv.gov/s3fs-public/NHAS-2022-2025.pdf [accessed 2024-03-27]
12. Horberg MA, Certa JM, Rubenstein KB, et al. Beyond the HIV care continuum and viral suppression: broadening the scope of quality metrics for total HIV patient care. AIDS Patient Care STDS 2020 Nov;34(11):461-469. [doi: 10.1089/apc.2020.0145] [Medline: 33147087]
13. World Health Organization. Continuity and Coordination of Care: A Practice Brief to Support Implementation of the WHO Framework on Integrated People-Centred Health Services (9241514035) 2018. URL: https://www.who.int/publications/i/item/9789241514033 [accessed 2024-03-27]

XSL•FO
RenderX

14. McNulty MC, Schneider JA. Care continuum entry interventions: seek and test strategies to engage persons most impacted by HIV within the United States. AIDS 2018 Feb 20;32(4):407-417. [doi: 10.1097/QAD.0000000000001733] [Medline: 29381558]
15. Mugavero MJ. Elements of the HIV care continuum: improving engagement and retention in care. Top Antivir Med 2016;24(3):115-119. [Medline: 27841983]
16. Pinto RM, Witte SS, Filippone PL, Baird KL, Whitman WR. Factors that influence linkages to HIV continuum of care services: implications for multi-level interventions. Int J Environ Res Public Health 2017 Nov 7;14(11):1355. [doi: 10.3390/ijerph14111355] [Medline: 29112126]
17. Frank TD, Carter A, Jahagirdar D. Global, regional, and national incidence, prevalence, and mortality of HIV, 1980-2017, and forecasts to 2030, for 195 countries and territories: a systematic analysis for the global burden of diseases, injuries, and risk factors study 2017. Lancet HIV 2019 Dec;6(12):e831-e859. [doi: 10.1016/S2352-3018(19)30196-1] [Medline: 31439534]
18. Mounier-Jack S, Mayhew SH, Mays N. Integrated care: learning between high-income, and low- and middle-income country health systems. Health Policy Plan 2017 Nov 1;32(suppl_4):iv6-iv12. [doi: 10.1093/heapol/czx039] [Medline: 29194541]
19. Pincombe M, Reese V, Dolan CB. The effectiveness of national-level containment and closure policies across income levels during the COVID-19 pandemic: an analysis of 113 countries. Health Policy Plan 2021 Aug 12;36(7):1152-1162. [doi: 10.1093/heapol/czab054] [Medline: 33942081]
20. Cooper V, Clatworthy J, Youssef E, et al. Which aspects of health care are most valued by people living with HIV in high-income countries? a systematic review. BMC Health Serv Res 2016 Nov 30;16(1):677. [doi: 10.1186/s12913-016-1914-4] [Medline: 27899106]
21. Park S, Park J. Identifying the knowledge structure and trends of outreach in public health care: a text network analysis and topic modeling. Int J Environ Res Public Health 2021 Sep 3;18(17):9309. [doi: 10.3390/ijerph18179309] [Medline: 34501897]
22. Subakti A, Murfi H, Hariadi N. The performance of BERT as data representation of text clustering. J Big Data 2022;9(1):15. [doi: 10.1186/s40537-022-00564-9] [Medline: 35194542]
23. Thakur K, Kumar V. Application of text mining techniques on scholarly research articles: methods and tools. New Rev Acad Librariansh 2022 Jul 3;28(3):279-302. [doi: 10.1080/13614533.2021.1918190]
24. Dancy-Scott N, Dutcher GA, Keselman A, et al. Trends in HIV terminology: text mining and data visualization assessment of international AIDS conference abstracts over 25 years. JMIR Public Health Surveill 2018 May 4;4(2):e50. [doi: 10.2196/publichealth.8552] [Medline: 29728344]
25. Devlin J, Chang MW, Lee K, Toutanova K. Bert: pretraining of deep bidirectional transformers for language understanding. arXiv. Preprint posted online on 2018. [doi: 10.48550/arXiv.1810.04805]
26. Xu WW, Tshimula JM, Dubé È, et al. Unmasking the Twitter discourses on masks during the COVID-19 pandemic: user cluster-based BERT topic modeling approach. JMIR Infodemiology 2022;2(2):e41198. [doi: 10.2196/41198] [Medline: 36536763]
27. El-Gayar O, Al-Ramahi M, Wahbeh A, Nasralah T, Elnoshokaty A. A comparative analysis of the interpretability of LDA and LLM for topic modeling: the case of healthcare apps. 2024 Presented at: AMCIS 2024 Proceedings; Aug 15-17, 2024; Salt Lake City, Utah URL: https://aisel.aisnet.org/amcis2024/health_it/health_it/22 [accessed 2025-06-19]
28. Egger R, Yu J. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. Front Sociol 2022;7:886498. [doi: 10.3389/fsoc.2022.886498] [Medline: 35602001]
29. Raman R, Pattnaik D, Hughes L, Nedungadi P. Unveiling the dynamics of AI applications: a review of reviews using scientometrics and BERTopic modeling. J Innov Knowl 2024 Jul;9(3):100517. [doi: 10.1016/j.jik.2024.100517]
30. Baird A, Xia Y, Cheng Y. Consumer perceptions of telehealth for mental health or substance abuse: a Twitter-based topic modeling analysis. JAMIA Open 2022 Jul;5(2):ooac028. [doi: 10.1093/jamiaopen/ooac028] [Medline: 35495736]
31. Chandra R, Saini R. Biden vs Trump: modeling US general elections using BERT language model. IEEE Access 2021;9:128494-128505. [doi: 10.1109/ACCESS.2021.3111035]
32. Karabacak M, Jagtiani P, Jain A, Panov F, Margetis K. Tracing topics and trends in drug-resistant epilepsy research using a natural language processing-based topic modeling approach. Epilepsia 2024 Apr;65(4):861-872. [doi: 10.1111/epi.17890] [Medline: 38314969]
33. Caskey J, McConnell IL, Oguss M, et al. Identifying COVID-19 outbreaks from contact-tracing interview forms for public health departments: development of a natural language processing pipeline. JMIR Public Health Surveill 2022 Mar 8;8(3):e36119. [doi: 10.2196/36119] [Medline: 35144241]
34. WHO. Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach. 2013. URL: https://www.who.int/publications/i/item/9789241505727 [accessed 2025-06-20]
35. Asmussen CB, Møller C. Smart literature review: a practical topic modelling approach to exploratory literature review. J Big Data 2019 Dec;6(1):1-18. [doi: 10.1186/s40537-019-0255-7]
36. Karabacak M, Schupper AJ, Carr MT, Hickman ZL, Margetis K. From text to insight: a natural language processing-based analysis of topics and trends in neurosurgery. Neurosurgery 2024 Apr 1;94(4):679-689. [doi: 10.1227/neu.0000000000002763] [Medline: 37988054]

XSL•FO
RenderX

37. Wilianto D, Suganda Girsang A. Automatic short answer grading onhigh school's e-learning using semantic similarity methods. TEM Journal 2023 Feb 27;12(1):297-302. [doi: 10.18421/TEM121-37]
38. Hatvani P, Yang ZG. Training embedding models for Hungarian. Presented at: 2024 IEEE 3rd Conference on Information Technology and Data Science (CITDS); Aug 26-28, 2024; Debrecen, Hungary. [doi: 10.1109/CITDS62610.2024.10791389]
39. Hidayat AA, Nirwantono R, Budiarto A, Pardamean B. BERT-based topic modeling approach for malaria research publication. Presented at: 2022 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS); Nov 16-17, 2022; Jakarta, Indonesia. [doi: 10.1109/ICIMCIS56303.2022.10017743]
40. Sangaraju VR, Bolla BK, Nayak DK. KhJ topic modelling on consumer financial protection bureau data: an approach using BERT based embeddings. arXiv. Preprint posted online on 2022. [doi: 10.48550/arXiv.2205.07259]
41. McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction. arXiv 2018. [doi: 10.48550/arXiv.1802.03426]
42. McInnes L, Healy J, Astels S. Hdbscan: Hierarchical density based clustering. J Open Source Softw 2017;2(11):205. [doi: 10.21105/joss.00205]
43. Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv. Preprint posted online on 2022. [doi: 10.48550/arXiv.2203.05794]
44. Zhao W, Chen JJ, Perkins R, et al. A heuristic approach to determine an appropriate number of topics in topic modeling. BMC Bioinformatics 2015;16(Suppl 13):S8. [doi: 10.1186/1471-2105-16-S13-S8] [Medline: 26424364]
45. Riaz A, Abdulkader O, Ikram MJ, Jan S. Exploring topic modelling: a comparative analysis of traditional and transformer-based approaches with emphasis on coherence and diversity. Int J Electrical Computer Engineer (IJECE) 2025;15(2):1933. [doi: 10.11591/ijece.v15i2.pp1933-1948]
46. Li S, Pan R, Luo H, Liu X, Zhao G. Adaptive cross-contextual word embedding for word polysemy with unsupervised topic modeling. Knowl Based Syst 2021 Apr;218:106827. [doi: 10.1016/j.knosys.2021.106827]
47. Murthy D, Keshari S, Arora S, et al. Categorizing E-cigarette-related tweets using BERT topic modeling. Emerg Trends Drugs Addict Health 2024 Dec;4:100160. [doi: 10.1016/j.etdah.2024.100160]
48. Adam R, Kogler ML. Bidirectional topic matching: quantifying thematic overlap between corpora through topic modelling. arXiv. Preprint posted online on 2024. [doi: 10.48550/arXiv.2412.18376]
49. WHO. Countries. URL: https://www.who.int/countries [accessed 2025-06-20]
50. Allegrante JP, Wells MT, Peterson JC. Interventions to support behavioral self-management of chronic diseases. Annu Rev Public Health 2019 Apr 1;40(1):127-146. [doi: 10.1146/annurev-publhealth-040218-044008] [Medline: 30601717]
51. Kay ES, Batey DS, Mugavero MJ. The HIV treatment cascade and care continuum: updates, goals, and recommendations for the future. AIDS Res Ther 2016 Dec;13(1):1-7. [doi: 10.1186/s12981-016-0120-0]
52. Zeluf-Andersson G, Eriksson LE, Schönnesson LN, Höijer J, Månehall P, Ekström AM. Beyond viral suppression: the quality of life of people living with HIV in Sweden. AIDS Care 2019 Apr;31(4):403-412. [doi: 10.1080/09540121.2018.1545990] [Medline: 30444140]
53. Lazarus JV, Wohl DA, Cascio M, et al. Long-term success for people living with HIV: a framework to guide practice. HIV Med 2023 Mar;24(supplement 2):8-19. [doi: 10.1111/hiv.13460] [Medline: 36920412]
54. Guaraldi G, Arends J, Buhk T, et al. "Moving fourth": a vision toward achieving healthy living with HIV beyond viral suppression. AIDS Rev 2019;21(3):135-142. [doi: 10.24875/AIDSRev.19000088] [Medline: 31532396]
55. Centers for Disease Control and Prevention. Estimated HIV incidence and prevalence in the united states. 2017–2021. : HIV Surveillance Supplemental Report; 2023 URL: https://stacks.cdc.gov/view/cdc/149080 [accessed 2025-06-23]
56. Du M, Liu M, Liu J. Global, regional, and national disease burden and attributable risk factors of HIV/AIDS in older adults aged 70 years and above: a trend analysis based on the Global Burden of Disease study 2019. Epidemiol Infect 2023 Dec 15;152:e2. [doi: 10.1017/S0950268823001954] [Medline: 38097398]
57. Bor J, Fischer C, Modi M, et al. Changing knowledge and attitudes towards HIV treatment-as-prevention and "undetectable = untransmittable": a systematic review. AIDS Behav 2021 Dec;25(12):4209-4224. [doi: 10.1007/s10461-021-03296-8] [Medline: 34036459]
58. Hussen SA, Chahroudi A, Boylan A, Camacho-Gonzalez AF, Hackett S, Chakraborty R. Transition of youth living with HIV from pediatric to adult-oriented healthcare: a review of the literature. Future Virol 2014 Oct;9(10):921-929. [doi: 10.2217/fvl.14.73]
59. Lazarus JV, Safreed-Harmon K, Kamarulzaman A, et al. Consensus statement on the role of health systems in advancing the long-term well-being of people living with HIV. Nat Commun 2021;12(1):4450. [doi: 10.1038/s41467-021-24673-w]
60. Duthely LM, Sanchez-Covarrubias AP. Digitized HIV/AIDS treatment adherence interventions: a review of recent SMS/texting mobile health applications and implications for theory and practice. Front Commun (Lausanne) 2020 Nov;5:530164. [doi: 10.3389/fcomm.2020.530164] [Medline: 33644162]
61. Simoni JM, Kutner BA, Horvath KJ. Opportunities and challenges of digital technology for HIV treatment and prevention. Curr HIV/AIDS Rep 2015 Dec;12(4):437-440. [doi: 10.1007/s11904-015-0289-1] [Medline: 26412082]
62. Haris M, Abbas R. Four decades of HIV: global trends, testing assays, treatment, and challenges. Zoonoses 2024;4(1):997. [doi: 10.15212/ZOONOSES-2023-0039]

XSL•FO
RenderX

63. McCree DH, Byrd KK, Johnston M, Gaines M, Weidle PJ. Roles for pharmacists in the "ending the HIV epidemic: a plan for America" initiative. Public Health Rep 2020;135(5):547-554. [doi: 10.1177/0033354920941184] [Medline: 32780671]
64. Crawford ND, Myers S, Young H, Klepser D, Tung E. The role of pharmacies in the HIV prevention and care continuums: a systematic review. AIDS Behav 2021 Jun;25(6):1819-1828. [doi: 10.1007/s10461-020-03111-w] [Medline: 33386509]
65. Brennan-Ing M, Mattas E. Aging with HIV: working to ensure equity and inclusion. Gerontology 2023;69(8):946-952. [doi: 10.1159/000530723] [Medline: 37071979]
66. Tam G, Wong NS, Lee SS. A survey of the involvement of primary care doctors in HIV prevention and care in a low-prevalence, high-income setting. BMC Fam Pract 2021 Jan 28;22(1):27. [doi: 10.1186/s12875-021-01376-1] [Medline: 33504326]
67. Bond KT, Chandler R, Chapman-Lambert C, et al. Applying a nursing perspective to address the challenges experienced by cisgender women in the HIV status neutral care continuum: a review of the literature. J Assoc Nurses AIDS Care 2021;32(3):283-305. [doi: 10.1097/JNC.0000000000000243] [Medline: 33929979]
68. WHO. Public health milestones through the years. 2024. URL: https://www.who.int/campaigns/75-years-of-improving-public-health/milestones#year-1945 [accessed 2025-06-20]

Abbreviations

AMR: Region of the Americas
ART: antiretroviral therapy
BERT: bidirectional encoder representations from transformers
LDA: latent dirichlet allocation
QoL: quality of life
TF-IDF: term frequency-inverse document frequency
WHO: World Health Organization

Edited by TH Kwan; submitted 04.08.24; peer-reviewed by DCDES Jarlais, R Saifi; revised version received 07.04.25; accepted 28.04.25; published 09.07.25.

Please cite as:
Lee S, Kim L, Shim MS, Kim GS
Identifying Health Care Services Offered in the HIV Care Continuum via a Machine Learning–Based Topic Modeling Approach: Exploratory Literature Review
JMIR Public Health Surveill 2025;11:e65081
URL: https://publichealth.jmir.org/2025/1/e65081
doi:10.2196/65081

© SangA Lee, Layoung Kim, Mi-So Shim, Gwang Suk Kim. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 9.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Symptom Trajectories and Clinical Subtypes in Post–COVID-19 Condition: Systematic Review and Clustering Analysis

Mingzhi Hu*, Dr med; Tian Song*, Dr med; Zhaoyuan Gong*, Dr med; Qianzi Che, Dr med; Jing Guo, Dr med; Lin Chen, Dr med; Haili Zhang, Dr med; Huizhen Li, Dr med; Ning Liang, Dr med; Guozhen Zhao, Dr med; Yanping Wang, Prof Dr Med; Nannan Shi, Prof Dr Med; Bin Liu, Dr med

Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, No 16, Nanxiao Street, Dongzhimen, Dongcheng District, Beijing, China

*these authors contributed equally

Corresponding Author:
Bin Liu, Dr med
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, No 16, Nanxiao Street, Dongzhimen, Dongcheng District, Beijing, China

Abstract

Background: Post–COVID-19 condition presents complex symptomatology involving multifaceted interactions, which has resulted in a current lack of comprehensive understanding of its disease trajectory. This knowledge gap significantly compromises the efficiency of symptom management and adversely affects patients' quality of life.

Objective: This study aims to comprehensively characterize the temporal evolution of post–COVID-19 condition by identifying core symptom clusters and clinical phenotypes, thereby enhancing understanding of the disease trajectory.

Methods: The PubMed, Web of Science, and Embase databases were searched from December 1, 2019, to March 1, 2024. Observational studies related to the prevalence of symptoms in post–COVID-19 condition had been included. We conducted a meta-analysis to synthesize symptom prevalence across different follow-up intervals following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines and used a network to explore interrelationships and co-occurrence patterns among symptoms, enabling the identification of core symptoms and changes over time. Clustering analysis was used to classify included studies into distinct clinical subtypes.

Results: This study analyzed 155 sets of macrolevel data from 108 clinical studies, encompassing 63,771 patients. Fatigue was the most prevalent symptom across all 4 follow-up points (52%, 48%, 46%, and 54%). Dyspnea peaked at the third and sixth follow-ups (36% and 31%) and then declined steadily (28% and 22%). Subgroup analysis revealed that Africa reported the fewest symptoms overall, yet showed high early incidences of fatigue (68%, 95% CI 50% - 85%) and dyspnea (56%, 95% CI 15% - 98%). The Americas placed greater emphasis on symptom evolution within the first postinfection year, with notably higher prevalence of anxiety (60%, 95% CI 54% - 66%) and depression (36%, 95% CI 16% - 55%). Asia and Europe documented the most comprehensive symptom profiles, with Asia reporting lower early dyspnea rates (29%, 95% CI 18% - 40%) and Europe exhibiting more complex multisystem involvement during long-term follow-up. Network analysis showed that core post–COVID-19 symptoms evolved from early respiratory-neurological manifestations to chronic multisystem symptoms dominated by dizziness. Clustering analysis further indicated a progressive convergence of 2 initially distinct post–COVID-19 subtypes, with the acute inflammatory type becoming less prominent and gradually transitioning into a more chronic, persistent pattern.

Conclusions: This study provides a comprehensive characterization of the dynamic evolution of post–COVID-19 condition symptoms and clinical subtypes, highlighting their multisystem involvement. The results reveal a progressive decline in respiratory symptoms over time, while neurological manifestations emerge as the most persistent and systemically impactful core symptoms. Our findings emphasize the need for region-specific surveillance and early warning systems informed by symptom progression patterns. By continuously monitoring the trajectories of symptom clusters, this approach offers valuable insights for identifying early warning signals and targeted intervention points in the management of postinfectious sequelae arising from future large-scale epidemics.

Trial Registration: PROSPERO CRD42024537825; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024537825

(JMIR Public Health Surveill 2025;11:e72221) doi:10.2196/72221

KEYWORDS

post–COVID-19 condition; temporal dynamic; symptoms; subtypes; systematic review

XSL•FO
RenderX

Introduction

Studies have shown that approximately 10% - 30% of patients continue to experience persistent symptoms after nucleic acid test results turn negative or after discharge from the hospital, significantly impacting their daily lives [1-3]. Furthermore, some patients may experience symptoms lasting more than 2 years [4,5]. The World Health Organization defines the persistent symptoms following the acute phase of COVID-19 as "post–COVID-19 condition," characterized by symptoms that persist or emerge at least 3 months after the initial SARS-CoV-2 infection and cannot be explained by other causes. As of March 2025, a total of 777 million people worldwide have been infected with COVID-19 [6]. Assuming a conservative estimate that 10% of recovered individuals will experience persistent symptoms, over 70 million people may be affected by the long-term sequelae of COVID-19. Post–COVID-19 condition presents with complex symptoms affecting multiple systems and organs throughout the body, with studies reporting over 200 types of symptoms [7]. Common symptoms include fatigue, postexertional malaise, cognitive impairment (brain fog), dyspnea, chest pain, as well as anxiety and depression [8,9]. A single patient may simultaneously exhibit symptoms across multiple systems, with these symptoms interacting and forming a vicious cycle. The heterogeneity and dynamic nature of these symptoms pose significant challenges to the treatment of post–COVID-19 condition. Over the past few years, numerous researchers have conducted studies to understand the pathophysiology, risk factors, and optimal management strategies for post–COVID-19 condition [10-12]. However, no definitive pharmacological treatments or interventions have been established to date, and current approaches primarily rely on comprehensive management or symptomatic relief. Against this backdrop of symptom-based care, existing studies have reported the prevalence of post–COVID-19 condition symptoms and proposed clinical subtypes, such as fatigue-dominant and musculoskeletal-dominant phenotypes. Nevertheless, most of these studies have relied on cross-sectional assessments, failing to capture the dynamic symptom patterns or phenotypic transitions over time, and have not prioritized the identification of core symptoms [13,14].

Since the concept of symptom networks was first introduced by Fried et al [15], this research paradigm has been widely applied in fields such as psychiatry, chronic disease management, and long-term oncology follow-up [16,17]. Symptom networks not only provide direct measurements such as symptom prevalence and severity but also uncover the complex interactions among symptoms through a range of network metrics, including node centrality, predictability, clustering, network connectivity, and transitivity. Core symptoms, identified via centrality and related indicators, often exert a disproportionate influence over the entire network. Targeting these symptoms may yield a "leverage effect," leading to broader improvements in symptom management. Clustering analyses can detect co-occurring symptom patterns, pointing toward potential symptom groups and shared pathophysiological mechanisms [18]. They provide a theoretical and methodological foundation for exploring underlying mechanisms, refining symptom phenotyping, and identifying key therapeutic targets. Constructing dynamic symptom networks, in particular, allows for the identification of changes in intersymptom relationships over time [19]. This approach can help uncover patterns of symptom co-occurrence and mutual influence, contributing to the early detection and diagnosis of complex conditions such as post–COVID-19 condition. By identifying core and bridge symptoms and delivering interventions that specifically target these pivotal nodes, it becomes possible to disrupt symptom interaction pathways, rendering associated symptoms less connected and reducing their impact. This network-based strategy facilitates a shift from the traditional model of treating multiple symptoms independently to a more precise approach that targets the most influential symptom nodes [20], thereby enhancing symptom control and improving quality of life.

Although systematic reviews have investigated the incidence of post–COVID-19 condition symptoms, their inclusion of primary studies was limited to observation periods of up to 1 year at most, thereby precluding assessment of longer-term symptom trajectories [21]. These reviews also failed to explore the dynamic symptom patterns or phenotypic transitions of post–COVID-19 condition over time. Furthermore, the lack of timely updates in these studies has left symptom evolution patterns over the past 3 years largely unexplored. Notably, none of these systematic reviews have specifically examined the dynamic changes in core symptoms and symptom clusters as their primary research focus. Building upon this foundation, this study seeks to conduct a systematic evaluation and dynamic analysis of symptom data derived from published clinical investigations, with the primary objectives of identifying core symptom clusters associated with post–COVID-19 condition, elucidating their characteristic phenotypic evolution patterns, addressing the critical knowledge gap concerning long-term symptom trajectories extending beyond 1 year postinfection, and generating updated epidemiological evidence regarding symptom prevalence rates in this emerging clinical entity. The findings will contribute to providing a scientific basis for public health surveillance and stratified management of post–COVID-19 condition, thereby aiding in the optimization of medical resource allocation and the formulation of rehabilitation policies.

Methods

Study Design

This study was designed as a systematic review and meta-analysis of observational studies investigating post–COVID-19 condition combined with network analysis and clustering to identify longitudinal symptom patterns and clinical subtypes. We conducted a comprehensive literature search of published clinical studies from December 2019 to March 2025, extracting symptom prevalence data across different follow-up intervals (3rd, 6th, 12th, and 24th month). This temporal stratification allowed for the examination of symptom dynamics over time. To standardize the classification of reported symptoms across studies, we categorized each symptom into 1 of 9 system domains, including respiratory, nervous, circulatory, digestive, musculoskeletal, endocrine, skin, urogenital, and

XSL•FO
RenderX

other. The classification framework was primarily based on the *International Classification of Diseases, 11th Revision* (*ICD-11*). For symptoms that could plausibly be assigned to multiple systems, such as fatigue, we determined the most appropriate category based on predominant pathophysiological relevance, as indicated by current literature. For example, fatigue was classified under neurological manifestations due to its established association with autonomic dysfunction and neuroinflammatory processes in post–COVID-19 condition. A complete list of symptoms and their corresponding system classifications is provided in Table S1 in Multimedia Appendix 1 to facilitate the reproducibility of the categorization process. The meta-analysis was used to estimate pooled symptom prevalence rates at each time point, enhancing comparability across studies. To further explore the interrelationships among symptoms and their evolution, we applied network analysis, which revealed the co-occurrence structures and central symptoms within each follow-up period. Subsequently, we performed clustering analysis to group population-level data into distinct clinical phenotypes based on symptom profiles, enabling the identification of population-level subgroups with similar post–COVID-19 condition presentations. This stepwise, multimethod analytical framework was designed to address the complexity and heterogeneity of post–COVID-19 condition symptoms and provide a comprehensive understanding of their temporal and structural patterns. We have registered this meta-analysis on PROSPERO (CRD42024537825). We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statements for reporting our systematic review (Checklist 1) [22].

Search Strategy

We searched the following databases from December 1, 2019, to March 1, 2025 (with the language restricted to English): PubMed, Embase, and Web of Science. MeSH (Medical Subject Headings) terms were combined with text words in the literature retrieval related to "post–COVID-19 condition," "symptom," and "follow-up"; for specific details on the search strategy, please refer to Multimedia Appendix 2. To find additional studies, we screened the relevant reference lists of the included studies and previous literature reviews including systematic reviews and meta-analyses. All search results were imported to an EndNote (version X9; Clarivate Analytics) tool to facilitate the screening of titles and abstracts.

Data Collection and Extraction

Two independent reviewers (HL and GZ) worked together to screen eligible studies. The complete eligibility criteria are presented in Textbox 1. In the initial screening, they read titles and abstracts to include studies containing 1 or more intervention models. All potentially relevant studies were retrieved for full-text screening. A full-text review of potentially eligible studies then was conducted, and the reasons for exclusion in the full-text screening were recorded in the EndNote database. After screening, data were extracted independently by the reviewers using a data extraction form (LC and NL). The extraction form consists of the following categories: characteristics of studies (eg, years, time of publication, and title), characteristics of participants, follow-up intervals, symptoms, and others. Regarding the extraction of symptom data at different follow-up times, we will follow these criteria: if a single study involves multiple follow-up points, we will sequentially extract the number of patients with post–COVID-19 condition and symptom data for each follow-up point that meets the inclusion criteria. Studies reporting follow-up times that do not align with the predetermined follow-up intervals (eg, 1 month or spanning multiple months or even a year) will be excluded. Additionally, studies must report the number of confirmed post–COVID-19 condition cases at each follow-up; those reporting only the baseline sample size will also be excluded. Discrepancies will be resolved through discussions until reaching consensus with another reviewer (BL).

Textbox 1. Eligibility criteria for study inclusion in the systematic evaluation of symptom incidence in patients with post–COVID-19 condition.

Inclusion criteria

- Document type: Original studies
- Study population: Adults (age≥18 years) after diagnosed with COVID-19
- Data availability: Providing specific details about the follow-up time (≥3 months), sample size, and symptom data of post–COVID-19 condition cases
- Study design: Observational study (cohort study, case-control study, cross-sectional study, and case series)
- Language: English publications

Exclusion criteria

- Document type: Reviews, basic experimental studies, and other unrelated papers
- Study population: Pediatric or adolescent populations
- Data availability: The study has a limited sample size (n≤30)
- Study design: Basic research, randomized controlled trials
- Language: Publications in other languages

XSL•FO
RenderX

Risk of Bias Assessment

The same investigators (LC and NL) assessed the risk of bias for all qualifying studies based on the Joanna Briggs Institute critical appraisal tool for analytic study and cross-sectional study [23]. The investigators independently assessed and ranked the risk of bias for every eligible study as yes, no, unclear, and not applicable.

Statistical Analysis

We conducted a meta-analysis (*metafor* package in RStudio, version 4.2.0; Posit PBC) using random effects models to estimate the prevalence of post–COVID-19 condition symptoms, aggregating data from multiple studies and incorporating study-specific effects into our analysis. The pooled effect results were presented along with their corresponding 95% CIs. To assess the proportion of variation attributable to between-study heterogeneity, analysis was conducted using I^2 to quantify heterogeneity on a scale from 0% to 100%. In parallel, we performed a subgroup analysis based on the regional classification of the patients. Statistical significance for differences in prevalence among regions was determined using a threshold of $P<.05$. An extensive network analysis was performed to investigate the evolution of symptoms over 4 distinct time intervals, providing a comprehensive overview of the interrelations of symptom prevalence. We used Spearman intragroup methods to conduct correlation analysis, with the primary aim of identifying symptom pairs that exhibited statistically significant correlations ($P<.05$). In addition, fuzzy clustering was used to cluster symptoms across various follow-up intervals, providing insight into potential differences in clinical phenotypes among patients. The network was visualized using Cytoscape software (version 3.7.2; NRNB), providing a comprehensive and detailed illustration of symptom interactions over the designated time intervals.

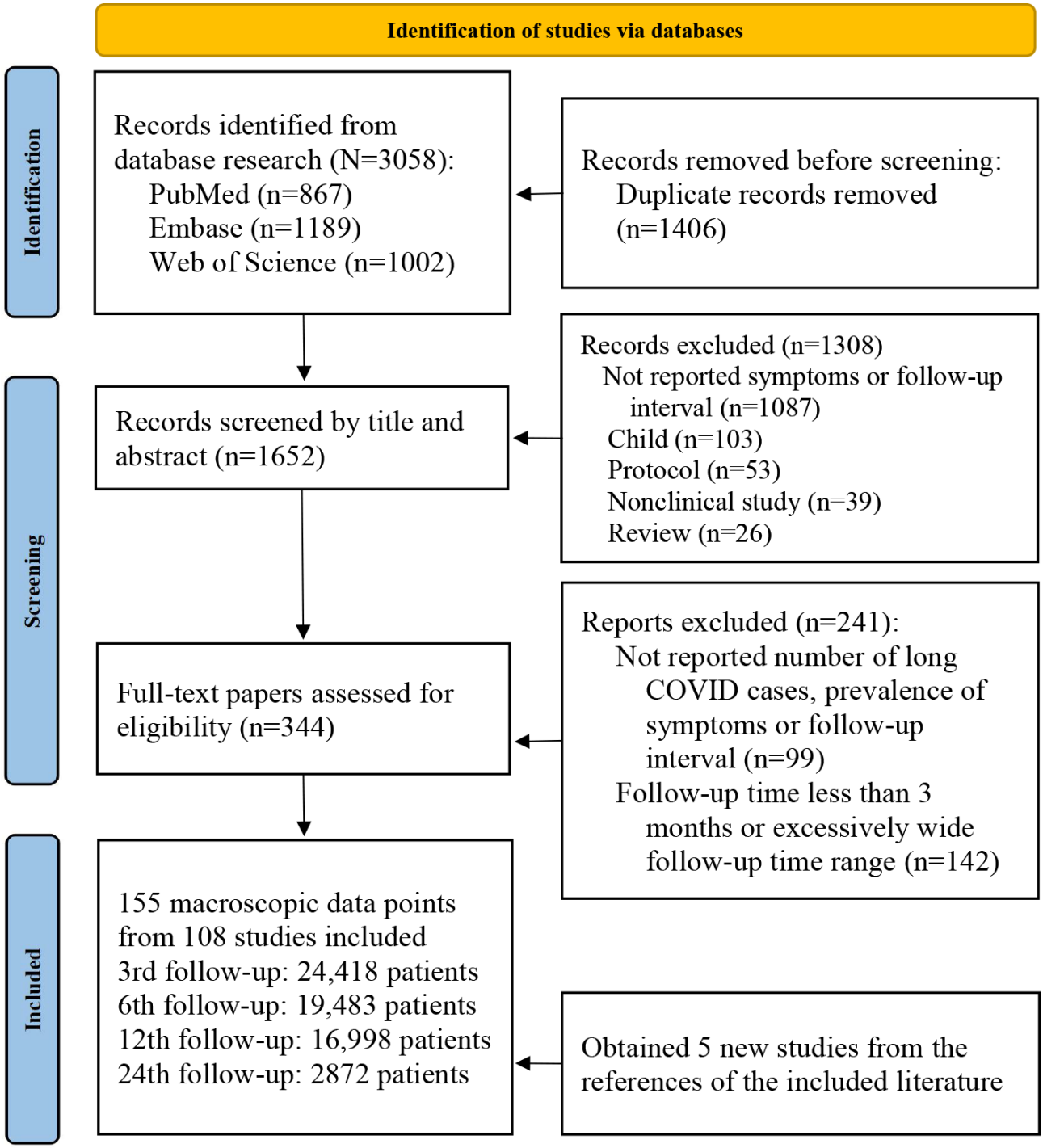
Results

Overview

A systematic review of post–COVID-19 condition identified 3058 records from database searches. After removing duplicates, title or abstract screening yielded 344 potentially eligible papers. The full-text review excluded 241 papers (unmet inclusion criteria), with 5 additional studies identified through reference screening. A total of 108 studies were included in the qualitative synthesis, encompassing 63,771 confirmed cases. These studies reported 64 distinct symptoms spanning 9 organ systems, including respiratory, neurological, cardiovascular, and musculoskeletal manifestations. Neurological symptoms were the most frequent (22/64), followed by respiratory (11/64) and digestive symptoms (9/64). The flowchart of the review process is shown in Figure 1. The studies were mainly conducted in Europe (55/108), Asia (38/108), America (13/108), and Africa (2/108). The number of patients and symptoms decreased over time, with neurological symptoms being the most prevalent. The characteristics of the studies are summarized in Table 1. The symptom classification system after standardization using *ICD-11* is presented in Table 2. The results of the quality assessment of the included studies are presented in Table S2 in Multimedia Appendix 1 [24-131].

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of study selection for the systematic review on symptom prevalence of post–COVID-19 condition (December 2019-March 2025).



XSL•FO
RenderX

Table . Summary of basic characteristics of included studies on post–COVID-19 condition, including study design, geographic location, symptom numbers, and follow-up period.

		All studies		3rd month		6th month		12th month		24th month	
		Studies, n	Patients, n	Studies, n	Patients, n	Studies, n	Patients, n	Studies, n	Patients, n	Studies, n	Patients, n
Overall[a]		108	63,771	43	24,418	51	19,483	50	16,998	11	2872
Study type											
	Prospective cohort study	35	17,894	15	11,505	17	2766	19	3027	5	596
	Cross-sectional study	23	20,364	10	7689	9	6259	6	5993	1	423
	Cohort study	30	18,122	9	4305	17	5795	17	6379	4	1643
	Prospective observational study	8	4835	3	194	4	3794	2	637	1	210
	Retrospective cohort study	7	1276	4	390	2	682	3	204	—[b]	—
	Retrospective observational study	4	1145	2	335	2	187	2	623	—	—
	Case control	1	135	—	—	—	—	1	135	—	—
Regions											
	Europe	55	34,433	19	12,886	24	8358	28	11,831	5	1358
	Asia	38	24,727	16	9023	18	9746	14	4484	5	1474
	America	13	3149	6	1189	8	1308	7	612	1	40
	Africa	2	1462	2	1320	1	71	1	71	—	—

[a]Since multiple follow-up dates were reported in 31 clinical studies, the total number of follow-up periods exceeds 99.

[b]Not available.

Table . Standardize and normalize symptoms using the *International Classification of Diseases, 11th Revision* (*ICD-11*) and classify all reported symptoms in the study into different systems.

Symptom-based disease system	Symptoms, n				
	All studies	3rd month	6th month	12th month	24th month
Nervous	23	22	18	19	16
Respiratory	11	10	9	8	9
Digestive	9	9	6	6	5
Circulatory	2	2	2	2	3
Musculoskeletal	5	4	3	4	3
Endocrine	3	3	2	1	—[a]
Skin	5	3	4	4	5
Urogenital	3	2	1	1	—
Other	3	3	3	3	2
Overall	64	60	50	50	45

[a]Not available.

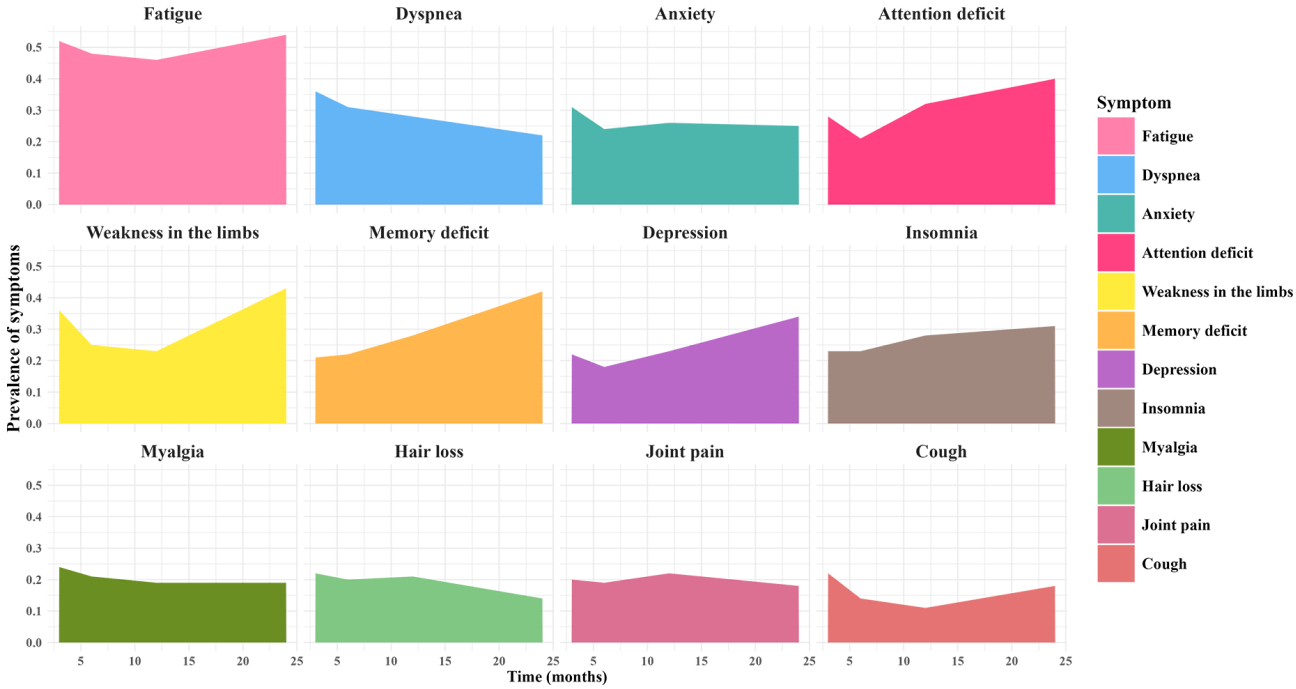
XSL•FO
RenderX

Prevalence of Clinical Symptoms

Meta-analysis of symptom incidence rates at different follow-up time points revealed distinct dynamic evolution characteristics. At the third-month follow-up, 43 studies reported on 60 clinical symptoms in a total of 24,418 patients. Fatigue (52%, 95% CI 45% - 59%) and dyspnea (36%, 95% CI 29% - 43%) were the most common symptoms, while neurological and psychiatric symptoms such as weakness in the limbs (36%, 95% CI 20% - 53%) and anxiety (31%, 95% CI 17% - 44%) had already begun to emerge. Additionally, there are 56 other symptoms (attention deficit, myalgia, insomnia, hair loss, and cough) with an incidence rate of less than 30% (Figure S1A in Multimedia Appendix 1). A total of 51 studies (n=19,483) reported the prevalence rates of 50 symptoms at the sixth-month follow-up. Although fatigue (48%, 95% CI 42% - 54%) remained the most prevalent symptom, dyspnea (31%, 95% CI 26% - 36%) showed a declining trend, and neurological and psychiatric symptoms (anxiety: 24% and insomnia: 23%) persisted. At the 12th-month follow-up, 50 studies with a combined total of 16,998 patients detailed the prevalence rates of 50 symptoms. Beyond the persistently high prevalence of fatigue (46%, 95% CI 41% - 52%), dyspnea has declined to the fourth most reported symptom, while neurological manifestations (attention deficit: 32%, 95% CI 22% - 43%; memory impairment: 28%, 95% CI 19% - 37%; and insomnia: 28%, 95% CI 21% - 35%) and cardiovascular symptoms (arrhythmia: 29%, 95% CI 12% - 46%) have emerged as more prominent clinical features. At the 2-year follow-up, the prevalence rates of 45 symptoms were reported across 11 studies involving 2872 patients. Fatigue remained the most prevalent symptom (54%, 95% CI 45% - 64%), with neurological manifestations (particularly memory deficit at 42%, 95% CI 28% - 56%) constituting the predominant clinical presentation. Notably, emerging symptoms including gastrointestinal bloating (45%, 95% CI 38% - 51%) and respiratory nasal congestion (38%, 95% CI 32% - 45%) have become increasingly apparent. Please refer to Figure S1 in Multimedia Appendix 1 for a detailed illustration of forest plots of symptom prevalence at different follow-ups.

Overall, fatigue consistently ranked as the most prevalent symptom, particularly prominent among patients with post–COVID-19 condition. Moreover, neurological symptoms demonstrated a progressive worsening trend. Most importantly, as the disease duration extended, the symptom spectrum shifted from a predominance of respiratory symptoms in the early stages to multisystem involvement, including cardiovascular, digestive, and neurological systems, ultimately forming a complex clinical presentation characterized by fatigue as the core feature and multisystem dysfunction at the 2-year follow-up. The prevalence rates of the 12 most frequently reported symptoms across the 4 follow-up time points are presented in Figure 2.

Figure 2. Trends in the prevalence of major post–COVID-19 condition symptoms across 4 follow-up time points (3rd, 6th, 12th, and 24th month) in studies conducted between December 2019 and March 2025, highlighting the progression of symptoms with fatigue as the core symptom and the shift from respiratory involvement to multisystem dysfunction.



Subgroup Analysis

To account for the comprehensive impact of temporal and spatial dynamics on symptom prevalence, we conducted a subgroup analysis based on the geographical locations of the studies. The results revealed notable differences in symptom prevalence across different regions. Overall, Africa reported the fewest symptoms, while Asia and Europe reported the highest number of symptoms. The Americas focused more on symptom incidence within 1 year. At the third-month follow-up, fatigue (68%, 95% CI 50% - 85%) and dyspnea (56%, 95% CI 15% - 98%) had the highest prevalence rates in Africa. However, insufficient attention to other symptoms has resulted in limited data availability, making further symptom comparisons challenging. The Americas exhibited a distinct neuromusculoskeletal symptom profile during short-term

XSL•FO
RenderX

follow-up, with significantly higher prevalence rates of anxiety (60%, 95% CI 54% - 66%) and depression (36%, 95% CI 16% - 55%) compared to other regions. Notably, joint pain (51%, 95% CI 33% - 68%) and weakness in the limbs (57%) were particularly prominent. Over time, both anxiety and joint pain maintained persistently high prevalence rates, reflecting a substantial long-term disease burden. During early follow-up periods, Asian cohorts demonstrated the highest reported symptom diversity (n=58 distinct manifestations) while exhibiting comparatively lower prevalence rates of dyspnea relative to other continental populations. This epidemiological pattern suggests a milder degree of respiratory involvement in Asian patient populations during the acute-postacute transition phase. Limb weakness was more commonly observed initially, but gradually diminished over time, evolving into a predominantly neurological symptom profile characterized by insomnia, depression, and memory impairment. The European region exhibited a more complex symptom spectrum during long-term follow-up, including an increasing prevalence of fatigue (66%, 95% CI 52% - 79%) and worsening cognitive dysfunction. Additionally, a range of new-onset symptoms emerged, such as bloating (45%) and other multisystem manifestations. Concrete details are provided in Table S3 in Multimedia Appendix 1.

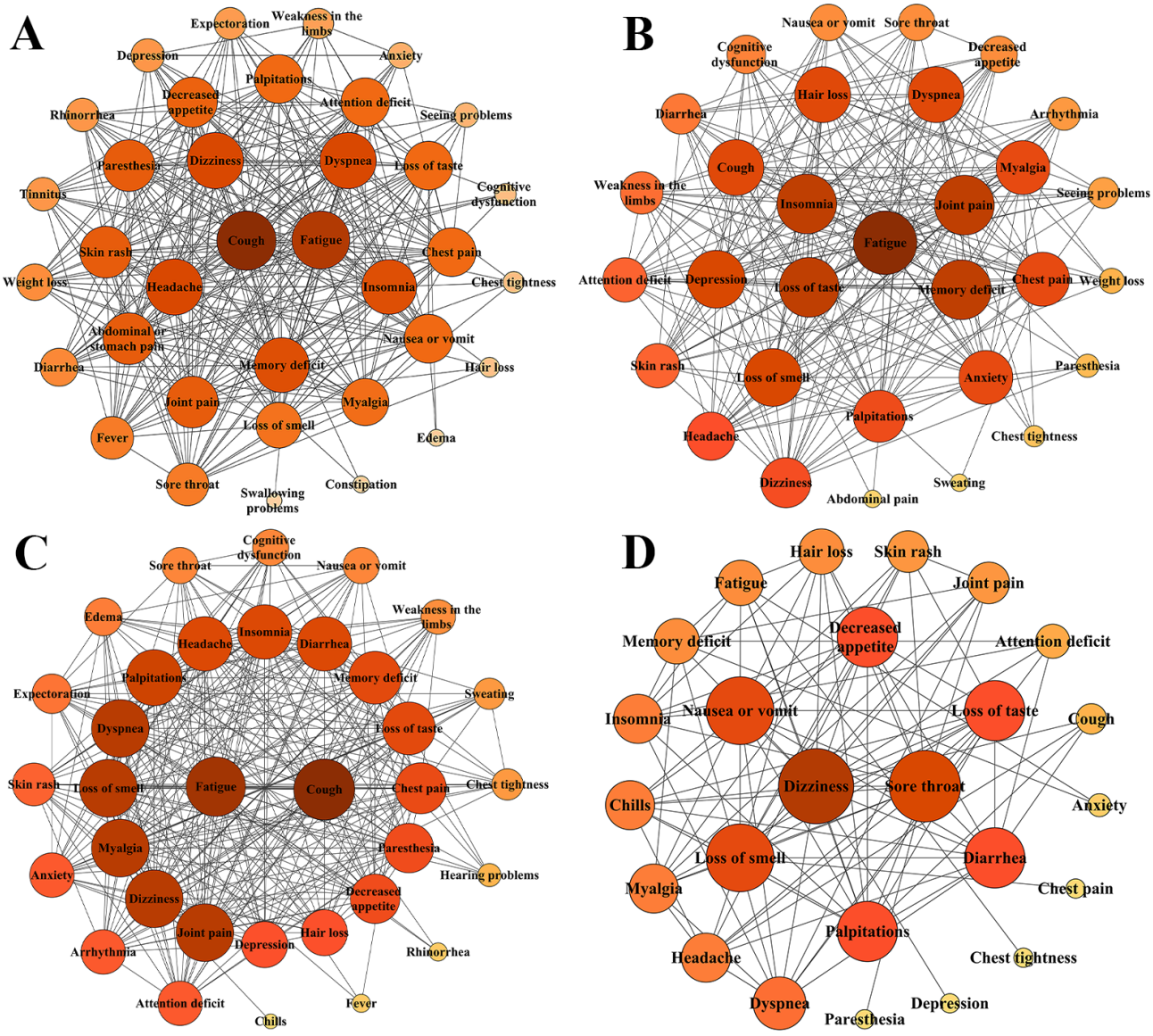
Association and Clustering Analysis of Symptoms

We conducted the Spearman rank correlation analysis of symptoms across different follow-up periods, identifying positively correlated symptom pairs with statistically significant differences. These correlations were visualized using Cytoscape to display the core symptom clusters for each follow-up period.

The third-month follow-up analysis revealed significant positive correlations between core symptoms (cough, fatigue, headache, dizziness, and dyspnea) and other persistent symptoms. Network analysis demonstrated a complex symptom interaction pattern during early follow-up, comprising 36 distinct symptoms (Figure 3A). Cough showed the highest closeness centrality (0.875), node degree (Degree=30), and significantly elevated stress centrality (424) compared to other symptoms, indicating its critical role in symptom propagation pathways. Notably, fatigue, dyspnea, dizziness, and headache demonstrated comparable closeness centrality and topological coefficients, suggesting early stage respiratory-neurological system interactions. The network analysis at the sixth-month follow-up identified a core symptom cluster consisting of fatigue, memory impairment, ageusia, arthralgia, and insomnia. Fatigue demonstrated significantly higher betweenness centrality and closeness centrality compared to other symptoms, with the highest node strength in the entire network (Degree=26), indicating its pivotal role as a central hub connecting multiple symptom subgroups. Comparative analysis revealed a progressive reduction in symptom complexity, with the number of nodes decreasing from 36 at 3 months to 31 at 6 months. Notably, respiratory-related acute symptoms (eg, dyspnea and chest pain) became marginalized in the network topology. The neurocognitive-somatic symptom cluster, including memory impairment, arthralgia, insomnia, and ageusia, showed converging topological coefficients (0.614-0.625), suggesting that these symptoms may share neural network connectivity, potentially reflecting common neuroinflammatory mechanisms. The 12th-month follow-up network analysis revealed a newly formed core symptom cluster consisting of fatigue, cough, dizziness, anosmia, myalgia, and arthralgia. Dizziness emerged as the most central symptom, exhibiting the highest betweenness centrality (0.067) and stress centrality (322) in the network, along with a closeness centrality of 0.846. Fatigue maintained its prominent position with equally high betweenness centrality (0.067) and significant stress centrality (322), confirming both symptoms as critical information hubs and key propagation nodes within the symptom network. Notably, myalgia and arthralgia demonstrated nearly identical topological coefficients (0.645), suggesting shared immunoinflammatory pathway involvement. The analysis further identified cough as having the highest closeness centrality (0.891) in the entire network while showing topological coefficient similarity with anosmia (0.61), potentially indicating common inflammatory pathway mechanisms underlying these symptom pairs. Comparative analysis revealed a significant reduction in network complexity at a 2-year follow-up, with the number of nodes decreasing to 25 and the complete disappearance of respiratory symptoms (eg dyspnea and chest pain) and acute inflammatory manifestations (fever). Dizziness maintained its network dominance, exhibiting persistently high betweenness centrality (0.189) and closeness centrality (0.75) while achieving the highest node degree (Degree=16), indicating substantial strengthening of its role as the primary information hub. In contrast, cough exited the core cluster, demonstrating marked reductions in both node degree (to 4) and stress centrality (to 14), reflecting its substantially diminished role in symptom propagation. Gastrointestinal and cardiovascular manifestations emerged as increasingly prominent over time, demonstrating that post–COVID-19 condition follows a characteristic evolutionary trajectory culminating in chronic, persistent multisystem involvement. Complete symptom network analysis results appear in Table S4 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 3. Symptom network diagrams of post–COVID-19 condition at different follow-up time points showing symptom interactions and core nodes. Longitudinal symptom network analysis revealed a transition from respiratory-neurological symptom clusters to chronic, multisystem symptoms over time. (A) 3rd-month follow-up, (B) 6th-month follow-up, (C) 12th-month follow-up, and (D) 24th-month follow-up; nodes represent individual symptoms and edges to illustrate the correlations between symptoms for each interval. Each node's color and size represent the degree of the symptom, with darker colors and larger size indicating a higher number of symptoms correlated with it.



Furthermore, we conducted a clustering analysis on longitudinal symptom prevalence data, which revealed 2 distinct clinical subtypes of post–COVID-19 condition. To avoid subjective severity-based labeling, we adopted descriptive terms based on symptom distribution patterns and classified the clusters as the "dispersed type" and the "concentrated type." The dispersed type (cluster 1) was characterized by a broader diversity of symptoms with generally lower prevalence rates across multiple organ systems. In contrast, the concentrated type (cluster 2) showed a narrower symptom spectrum but with higher prevalence rates, predominantly involving neurological and respiratory systems. This clustering analysis finding was most pronounced within 1 year of follow-up. At the third-month follow-up, notable differences in symptom distribution were evident (Multimedia Appendix 3): cluster 1 showed a wide (n=59) but lower-range distribution of symptom prevalence (ranging from 0.5% to 20%), whereas cluster 2 demonstrated a narrower (n=39) but higher-range distribution (from 20% to 50%), indicating a more focused and severe symptom profile. The different colors in the legend represent different systems, with symptoms of the same system shown in the same area. The letter "n" represents the number of macro data points, and the percentages indicate the average prevalence of symptoms within each cluster. These findings suggest 2 distinct trajectories of post–COVID-19 condition: one subtype characterized by chronic, multisystem, and persistent symptoms (cluster 1), and another with concentrated, inflammation-related symptoms of greater severity (cluster 2). The longitudinal analysis further revealed a progressive convergence in the number of reported symptoms between the 2 clusters over time, suggesting a potential evolutionary trend toward more chronic and diffuse symptomatology in the postacute phase of SARS-CoV-2 infection. Figure S2 in Multimedia Appendix 1 illustrates the 6th- and 12th-month follow-up findings.

XSL•FO
RenderX

Discussion

The Complex Network Dynamics of Long COVID Symptoms

The intricate interactions of post–COVID-19 condition symptoms are reflected in their dynamically evolving network characteristics. During the early follow-up phase (third month), respiratory symptoms (cough and dyspnea) showed significant correlations with neurological manifestations (headache and dizziness), suggesting potential shared pathways linking respiratory inflammation and neurological dysfunction, likely attributable to persistent multisystem impacts of residual viral components from the acute infection phase. Current evidence highlights autonomic dysfunction, particularly manifesting as postural orthostatic tachycardia syndrome–like symptoms (postural orthostatic tachycardia, dizziness, and gastrointestinal dysregulation), as a defining pathophysiological characteristic of post–COVID-19 condition. The proposed mechanisms include direct viral neurotropism, immune-mediated autonomic nerve injury, and chronic inflammatory modulation of autonomic regulatory circuits [132-134]. As the disease progressed, fatigue emerged as a persistent central node connecting multiple symptom clusters, with its centrality becoming particularly prominent at 6 - 12 months, demonstrating strong associations with neurocognitive symptoms (memory impairment and insomnia) and musculoskeletal symptoms (arthralgia and myalgia). This pattern may reflect chronic neuroinflammation or autoantibody-mediated damage mechanisms. Previous studies have similarly identified that post–COVID-19 condition is characterized by persistent immune dysregulation, manifesting as chronic inflammation and sustained immune activation following acute infection [135]. The underlying pathogenic mechanisms may include viral antigen persistence, autoimmune responses, and cytokine-mediated tissue damage, which collectively contribute to its multisystem clinical manifestations. Autoimmunity has been specifically implicated as a driver of symptoms such as joint pain and neurological manifestations, while prolonged immune activation may lead to progressive tissue injury and chronic symptomatology [136-139]. Of particular note, by the 2-year follow-up, the symptom network complexity had significantly diminished, with respiratory manifestations having largely resolved, while neurological, circulatory, and digestive system symptoms (including dizziness, nausea, vomiting, and palpitations) emerged as the predominant clinical features. Some researchers have proposed the hypothesis that the gut-brain axis constitutes a bidirectional neuro-immune-endocrine network, with vagal-mediated signaling and microbial metabolite production (particularly serotonin, predominantly gut-derived) critically modulating both enteric and central nervous system functions. Mounting evidence indicates that gut dysbiosis and impaired microbiome-immune crosstalk contribute to neuropsychiatric and systemic manifestations [140,141]. Another key hypothesis posits that vascular compromise leads to tissue hypoperfusion and hypoxia, particularly affecting high-metabolic organs (brain, heart, and skeletal muscle), potentially explaining fatigue, cognitive impairment, and myalgia [142-144]. Thus, it can be seen that the pathophysiology of long COVID is characterized by a multifactorial interplay of immune dysregulation, persistent vascular or organ injury, chronic inflammation, and neuroendocrine dysfunction, resulting in heterogeneous clinical manifestations [145-147]. This complex pathophysiology complicates mechanistic elucidation while accounting for the condition's diverse clinical phenotypes.

Regional Variations in Symptom Reporting and Underlying Factors

The study identified significant regional variations in symptom reporting, likely driven by multiple interacting factors. Disparities in health care resource distribution play a critical role—well-established follow-up systems in Europe and North America may capture a broader spectrum of symptoms, whereas Africa's constrained health care infrastructure and overburdened medical service systems may contribute to both significant underreporting of symptoms and elevated prevalence of primary clinical manifestations [148]. Genetic background differences also contribute; for instance, the high prevalence of arthralgia in the Americas may be linked to specific human leukocyte antigen genotypes [149,150]. Divergent vaccination strategies, particularly the varying protective effects of mRNA versus inactivated vaccines, might modulate immune responses and influence long-term symptom profiles. The spatiotemporal distribution of viral variants is another key factor, as the study period spanned both the Delta and Omicron waves. The Delta variant may induce long-term sequelae characterized by hyperinflammatory or neurotropic pathological patterns [151], whereas Omicron infection appears to confer a lower risk of developing post–COVID-19 condition [152]. Additionally, cultural factors, such as differing perceptions of mental health, may lead to underreporting of psychiatric symptoms (eg, anxiety) in Asian populations. These variations underscore the need for a globally standardized core symptom assessment tool and careful consideration of contextual factors when interpreting cross-regional findings.

Symptom Dynamics and Targeted Management

Fatigue, as a common nonspecific symptom, frequently overlaps with various other clinical manifestations, complicating the identification of condition-specific patterns. Notably, network analysis based on symptom prevalence suggests that cough and dyspnea may serve as key sentinel symptoms for early detection of post–COVID-19 condition. Within the context of long-term sequelae, symptoms such as dizziness, nausea or vomiting, and palpitations play a critical role in its recognition and diagnosis. The nonspecific nature of fatigue and its frequent overlap with diverse clinical manifestations complicate the identification of disease-specific patterns. Symptom prevalence–based network analysis highlights cough and dyspnea as potential sentinel symptoms for early recognition of post–COVID-19 condition, while neuro-circulatory-gastroenteric symptom clusters (particularly dizziness, nausea or vomiting, and palpitations) emerge as critical diagnostic markers in the chronic phase. This evolving understanding warrants dynamic, phase-adapted management approaches: early-stage interventions should prioritize respiratory rehabilitation and inflammatory response modulation, whereas chronic-phase care requires focused neurological symptom management integrated with

XSL•FO
RenderX

multidisciplinary strategies addressing concomitant digestive and circulatory system involvement. Regionally tailored public health strategies are essential, exemplified by musculoskeletal symptom–focused interventions in the Americas and enhanced cognitive rehabilitation frameworks in Europe, reflecting geographical variations in disease manifestation and health care system capacities.

Study Limitations and Future Directions

With regard to sampling constraints, the limited representation of data from Africa and certain parts of Southeast Asia may restrict the global generalizability of our conclusions, as region-specific symptom patterns or subtypes may remain unidentified. This regional imbalance could also introduce bias, as symptom prevalence and reporting practices are likely influenced by local health care infrastructure, follow-up systems, and cultural perceptions of illness. Moreover, the included studies largely lacked adjustment for socioeconomic determinants of health. Given that many countries have lifted containment measures, resumed economic activities, and reopened borders, this global shift toward adaptive coexistence with SARS-CoV-2 may lead to systematic underreporting of symptoms, especially mild or chronic manifestations. Additionally, the predominant reliance on passive follow-up methods may have missed clinically relevant but less severe cases, further contributing to data skewness. Another notable limitation is the absence of variant-specific data. The study period spans multiple waves of the pandemic, including those driven by the Delta and Omicron variants, which have been associated with distinct clinical trajectories. However, due to data availability constraints, we were unable to stratify symptom trajectories by viral strain, limiting the capacity to attribute observed patterns to specific viral characteristics. These limitations underscore the need for future studies to adopt a more standardized, comprehensive, and globally inclusive approach. We recommend prospective cohort studies with individual-level, longitudinal follow-up, stratified by region, socioeconomic status, and viral variant. Active surveillance in underrepresented regions, combined with systematic collection of social determinants and clinical metadata, is essential to generate more robust, equitable, and actionable insights into post–COVID-19 condition.

Conclusions

Building upon these findings, a comprehensive synthesis of post–COVID-19 condition research reveals a complex and evolving clinical landscape. Across varying follow-up periods encompassing diverse symptoms and affected organ systems, we demonstrate that the chronic and persistent effects of post–COVID-19 condition pose significant challenges for both patients and health care providers. Symptom prevalence exhibits temporal fluctuations, with neurological manifestations emerging as the disease's most prominent feature. In regions with limited research coverage, such as Africa, greater attention should be directed toward understanding specific symptom patterns and disease progression. Consequently, the establishment of a stratified management system based on disease staging and regional characteristics becomes imperative to address this heterogeneous condition. However, given the limitations of population-level data, our findings should be used as a guideline for understanding the broader trends of post–COVID-19 condition, but clinical decisions should be personalized, taking into account each patient's unique characteristics. Future studies incorporating individual-level data are needed to provide more granular insights and inform precision medicine strategies for post–COVID-19 care.

Acknowledgments

The study is supported by the China Academy of Chinese Medical Sciences key collaborative project of Innovation Fund, Analysis of Dynamic Temporal and Spatial Characteristics of TCM Syndrome of COVID-19 and Research on Its Biological Connotation (CI2022C004-01), National Key R&D Program of China (2023YFC3503400), 2023 Agency/National Special Fund Project of National Administration of Traditional Chinese Medicine (F0212), and the 2022 Qi Huang Young Scholar program of the National Administration of Traditional Chinese Medicine (NS; Z0841 and Z0865).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

MH, BL, and NS conceptualized the study. MH, BL, and ZG prepared the original draft of the manuscript. MH and TS conducted all statistical analyses. HZ developed the systematic review workflow and created Figure 1. ZG prepared Figure 3. HL, GZ, LC, NL, and BL collaboratively conducted literature screening and data extraction. QC and JG contributed to reviewing and editing the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

Conflicts that the editors consider relevant to the content of the manuscript have been disclosed. All authors have submitted the International Committee of Medical Journal Editors Form for Disclosure of Potential.

Multimedia Appendix 1
Detailed data and analyses on symptom characteristics and analytical findings in post–COVID-19 condition.

XSL•FO
RenderX

[DOC File, 2558 KB - publichealth_v11i1e72221_app1.doc]

Multimedia Appendix 2
Search strategy.
[DOC File, 29 KB - publichealth_v11i1e72221_app2.doc]

Multimedia Appendix 3
Clustering analysis of post–COVID-19 condition symptoms by prevalence at the third-month follow-up.
[PNG File, 413 KB - publichealth_v11i1e72221_app3.png]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOC File, 68 KB - publichealth_v11i1e72221_app4.doc]

References

1. Ballering AV, van Zon SKR, Olde Hartman TC, Rosmalen JGM, Lifelines Corona Research Initiative. Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. Lancet 2022 Aug 6;400(10350):452-461. [doi: 10.1016/S0140-6736(22)01214-4] [Medline: 35934007]
2. Ayoubkhani D, Bermingham C, Pouwels KB, et al. Trajectory of long covid symptoms after covid-19 vaccination: community based cohort study. BMJ 2022 May 18;377(e069676):e069676. [doi: 10.1136/bmj-2021-069676] [Medline: 35584816]
3. Qin S, Zhang Y, Li Y, et al. Long COVID facts and findings: a large-scale online survey in 74,075 Chinese participants. Lancet Reg Health West Pac 2024 Nov;52:101218. [doi: 10.1016/j.lanwpc.2024.101218] [Medline: 39881668]
4. Fernández-de-Las-Peñas C, Palacios-Ceña D, Gómez-Mayordomo V, et al. Prevalence of post-COVID-19 symptoms in hospitalized and non-hospitalized COVID-19 survivors: a systematic review and meta-analysis. Eur J Intern Med 2021 Oct;92:55-70. [doi: 10.1016/j.ejim.2021.06.009] [Medline: 34167876]
5. Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global prevalence of post-coronavirus disease 2019 (COVID-19) condition or long COVID: a meta-analysis and systematic review. J Infect Dis 2022 Nov 1;226(9):1593-1607. [doi: 10.1093/infdis/jiac136] [Medline: 35429399]
6. COVID-19 cases. WHO COVID-19 Dashboard. URL: https://data.who.int/dashboards/covid19/cases?n=c [accessed 2025-03-03]
7. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol 2023 Mar;21(3):133-146. [doi: 10.1038/s41579-022-00846-2] [Medline: 36639608]
8. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. Nat Med 2021 Apr;27(4):626-631. [doi: 10.1038/s41591-021-01292-y] [Medline: 33692530]
9. Lechner-Scott J, Levy M, Hawkes C, Yeh A, Giovannoni G. Long COVID or post COVID-19 syndrome. Mult Scler Relat Disord 2021 Oct;55:103268. [doi: 10.1016/j.msard.2021.103268] [Medline: 34601388]
10. Konkle SL, Magleby R, Bonacci RA, et al. Post-COVID condition risk factors and symptom clusters and associations with return to pre-COVID health-results from a 2021 multi-state survey. Clin Infect Dis 2024 Dec 20;PMID:39705503. [doi: 10.1093/cid/ciae632] [Medline: 39705503]
11. Gusev E, Sarapultsev A. Exploring the pathophysiology of long COVID: the central role of low-grade inflammation and multisystem involvement. Int J Mol Sci 2024 Jun 9;25(12):38928096. [doi: 10.3390/ijms25126389] [Medline: 38928096]
12. Peluso MJ, Deeks SG. Mechanisms of long COVID and the path toward therapeutics. Cell 2024 Oct 3;187(20):5500-5529. [doi: 10.1016/j.cell.2024.07.054] [Medline: 39326415]
13. Hu LY, Cai AQ, Li B, et al. Prevalence and risk factors for long COVID in China: a systematic review and meta-analysis of observational studies. J Infect Public Health 2025 Mar;18(3):102652. [doi: 10.1016/j.jiph.2025.102652] [Medline: 39826380]
14. Woodrow M, Carey C, Ziauddeen N, et al. Systematic review of the prevalence of long COVID. Open Forum Infect Dis 2023 Jul;10(7):ofad233. [doi: 10.1093/ofid/ofad233] [Medline: 37404951]
15. Fried EI, Boschloo L, van Borkulo CD, et al. Commentary: "Consistent superiority of selective serotonin reuptake inhibitors over placebo in reducing depressed mood in patients with major depression". Front Psychiatry 2015 Aug 21;6(117). [doi: 10.3389/fpsyt.2015.00117]
16. Kachadourian LK, Harpaz-Rotem I, Tsai J, Southwick SM, Pietrzak RH. Posttraumatic stress disorder symptoms, functioning, and suicidal ideation in US Military Veterans: a symptomics approach. Prim Care Companion CNS Disord 2019 Apr 25;21(2):18m02402. [doi: 10.4088/PCC.18m02402] [Medline: 31050228]
17. Rush AJ, Kovacs M, Beck AT, Weissenburger J, Hollon SD. Differential effects of cognitive therapy and pharmacotherapy on depressive symptoms. J Affect Disord 1981 Sep;3(3):221-229. [doi: 10.1016/0165-0327(81)90024-0] [Medline: 6456289]
18. Zhu Z, Zhao R, Hu Y. Symptom clusters in people living with HIV: a systematic review. J Pain Symptom Manage 2019 Jul;58(1):115-133. [doi: 10.1016/j.jpainsymman.2019.03.018] [Medline: 30951828]

XSL•FO
RenderX

19. Santos HP, Kossakowski JJ, Schwartz TA, Beeber L, Fried EI. Longitudinal network structure of depression symptoms and self-efficacy in low-income mothers. PLoS ONE 2018;13(1):e0191675. [doi: 10.1371/journal.pone.0191675] [Medline: 29360876]
20. Bekhuis E, Schoevers R, de Boer M, et al. Symptom-Specific effects of psychotherapy versus combined therapy in the treatment of mild to moderate depression: a network approach. Psychother Psychosom 2018;87(2):121-123. [doi: 10.1159/000486793] [Medline: 29495015]
21. Kuodi P, Gorelik Y, Gausi B, Bernstine T, Edelstein M. Characterization of post-COVID syndromes by symptom cluster and time period up to 12 months post-infection: a systematic review and meta-analysis. Int J Infect Dis 2023 Sep;134:1-7. [doi: 10.1016/j.ijid.2023.05.003] [Medline: 37150350]
22. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
23. Munn Z, Moola S, Lisy K, Riitano D, Tufanaru C. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. Int J Evid Based Healthc 2015 Sep;13(3):147-153. [doi: 10.1097/XEB.0000000000000054] [Medline: 26317388]
24. Kim Y, Bae S, Chang HH, Kim SW. Characteristics of long COVID and the impact of COVID-19 vaccination on long COVID 2 years following COVID-19 infection: prospective cohort study. Sci Rep 2024 Jan 9;14(1):854. [doi: 10.1038/s41598-023-50024-4] [Medline: 38191556]
25. Irisson-Mora I, Salgado-Cordero AM, Reyes-Varón E, et al. Comparison between the persistence of post COVID-19 symptoms on critical patients requiring invasive mechanical ventilation and non-critical patients. PLOS ONE 2022;17(8):e0273041. [doi: 10.1371/journal.pone.0273041] [Medline: 35994441]
26. Domènech-Montoliu S, Puig-Barberà J, Pac-Sa MR, et al. Complications post-COVID-19 and risk factors among patients after six months of a SARS-CoV-2 infection: a population-based prospective cohort study. Epidemiologia (Basel) 2022 Feb 10;3(1):49-67. [doi: 10.3390/epidemiologia3010006] [Medline: 36417267]
27. Catalán IP, Martí CR, Sota DDL, et al. Corticosteroids for COVID-19 symptoms and quality of life at 1 year from admission. J Med Virol 2022 Jan;94(1):205-210. [doi: 10.1002/jmv.27296] [Medline: 34436783]
28. Messin L, Puyraveau M, Benabdallah Y, et al. COVEVOL: natural evolution at 6 months of COVID-19. Viruses 2021 Oct 25;13(11):2151. [doi: 10.3390/v13112151] [Medline: 34834958]
29. Babicki M, Kapusta J, Pieniawska-Śmiech K, et al. Do COVID-19 vaccinations affect the most common post-COVID symptoms? Initial data from the STOP-COVID register-12-month follow-up. Viruses 2023 Jun 13;15(6):1370. [doi: 10.3390/v15061370] [Medline: 37376668]
30. Sykes DL, Van der Feltz-Cornelis CM, Holdsworth L, et al. Examining the relationship between inflammatory biomarkers during COVID-19 hospitalization and subsequent long-COVID symptoms: a longitudinal and retrospective study. Immun Inflamm Dis 2023 Oct;11(10):e1052. [doi: 10.1002/iid3.1052] [Medline: 37904690]
31. Serrano MN, Muñoz OM, Rueda C, Arboleda AC, Botero JD, Bustos MM. Factors associated with oxygen requirement and persistent symptoms 1 year after severe COVID-19 infection. J Int Med Res 2023 May;51(5):3000605231173317. [doi: 10.1177/03000605231173317] [Medline: 37170571]
32. Huang L, Li X, Gu X, et al. Health outcomes in people 2 years after surviving hospitalisation with COVID-19: a longitudinal cohort study. Lancet Respir Med 2022 Sep;10(9):863-876. [doi: 10.1016/S2213-2600(22)00126-6] [Medline: 35568052]
33. Villa A, Milito C, Deiana CM, et al. High prevalence of long COVID in common variable immunodeficiency: an Italian multicentric study. J Clin Immunol 2024 Feb 6;44(2):59. [doi: 10.1007/s10875-024-01656-2] [Medline: 38319477]
34. Amorim CEN, Gomes VLT, Cristelli MP, et al. High prevalence of long-COVID among kidney transplant recipients: a longitudinal cohort study. Transplantation 2022 Dec 1;106(12):2408-2415. [doi: 10.1097/TP.0000000000004359] [Medline: 36228200]
35. Becker C, Beck K, Zumbrunn S, et al. Long COVID 1 year after hospitalisation for COVID-19: a prospective bicentric cohort study. Swiss Med Wkly 2021 Oct 11;151(4142):w30091. [doi: 10.4414/smw.2021.w30091] [Medline: 34694106]
36. Rigoni M, Torri E, Nollo G, et al. “Long COVID” results after hospitalization for SARS-CoV-2 infection. Sci Rep 2022 Jun 10;12(1):9581. [doi: 10.1038/s41598-022-13077-5] [Medline: 35688830]
37. Blomberg B, Mohn KGI, Brokstad KA, et al. Long COVID in a prospective cohort of home-isolated patients. Nat Med 2021 Sep;27(9):1607-1613. [doi: 10.1038/s41591-021-01433-3] [Medline: 34163090]
38. Pérez-González A, Araújo-Ameijeiras A, Fernández-Villar A, et al. Long COVID in hospitalized and non-hospitalized patients in a large cohort in Northwest Spain, a prospective cohort study. Sci Rep 2022 Mar 1;12(1):3369. [doi: 10.1038/s41598-022-07414-x] [Medline: 35233035]
39. Yaksi N, Teker AG, Imre A. Long COVID in hospitalized COVID-19 patients: a retrospective cohort study. Iran J Public Health 2022 Jan;51(1):88-95. [doi: 10.18502/ijph.v51i1.8297] [Medline: 35223629]
40. Adler L, Gazit S, Pinto Y, et al. Long-COVID in patients with a history of mild or asymptomatic SARS-CoV-2 infection: a Nationwide Cohort Study. Scand J Prim Health Care 2022 Sep;40(3):342-349. [doi: 10.1080/02813432.2022.2139480] [Medline: 36314555]

XSL•FO
RenderX

41. Lombardo MDM, Foppiani A, Peretti GM, et al. Long-term coronavirus disease 2019 complications in inpatients and outpatients: a one-year follow-up cohort study. Open Forum Infect Dis 2021 Aug;8(8):ofab384. [doi: 10.1093/ofid/ofab384] [Medline: 34386546]
42. Kim Y, Kim SW, Chang HH, Kwon KT, Hwang S, Bae S. One year follow-up of COVID-19 related symptoms and patient quality of life: a prospective cohort study. Yonsei Med J 2022 Jun;63(6):499-510. [doi: 10.3349/ymj.2022.63.6.499] [Medline: 35619573]
43. Bhandari S, Rankawat G, Joshi S, Tiwaskar M, Lohmror A, Bhandari S. Post-COVID syndrome: the stranger ghost of culprit COVID-19. J Assoc Physicians India 2023 Feb;71(2):11-12. [doi: 10.5005/japi-11001-0193] [Medline: 37354471]
44. Fang X, Ming C, Cen Y, et al. Post-sequelae one year after hospital discharge among older COVID-19 patients: a multi-center prospective cohort study. J Infect 2022 Feb;84(2):179-186. [doi: 10.1016/j.jinf.2021.12.005] [Medline: 34902448]
45. van der Maaden T, Mutubuki EN, de Bruijn S, et al. Prevalence and severity of symptoms 3 months after infection with SARS-CoV-2 compared to test-negative and population controls in the Netherlands. J Infect Dis 2023 Apr 26;227(9):1059-1067. [doi: 10.1093/infdis/jiac474] [Medline: 36477364]
46. Hossain MA, Hossain KMA, Saunders K, et al. Prevalence of long COVID symptoms in Bangladesh: a prospective Inception Cohort Study of COVID-19 survivors. BMJ Glob Health 2021 Dec;6(12):e006838. [doi: 10.1136/bmjgh-2021-006838] [Medline: 34906986]
47. Gutiérrez-Canales LG, Muñoz-Corona C, Barrera-Chávez I, Viloria-Álvarez C, Macías AE, Guaní-Guerra E. Quality of life and persistence of symptoms in outpatients after recovery from COVID-19. Medicina (Kaunas) 2022 Dec 6;58(12):1795. [doi: 10.3390/medicina58121795] [Medline: 36556998]
48. Zhang Y, Han Z, Dai Y, et al. Symptom burden among older COVID-19 survivors two years after hospital discharge. Aging Dis 2023 Dec 1;14(6):2238-2248. [doi: 10.14336/AD.2023.0304] [Medline: 37199576]
49. Shrestha DS, Manandhar S, Chalise BS, et al. Symptoms 6 months following SARS-CoV-2 infection in Nepali women. PLOS ONE 2024;19(3):e0299141. [doi: 10.1371/journal.pone.0299141] [Medline: 38466665]
50. Perisse A, De Cacqueray F, Delarbre D, et al. Symptoms of long-COVID 1-year after a COVID-19 outbreak among sailors on a French aircraft carrier. Infect Dis Now 2023 Jun;53(4):104673. [doi: 10.1016/j.idnow.2023.104673] [Medline: 36775065]
51. Huang L, Yao Q, Gu X, et al. 1-year outcomes in hospital survivors with COVID-19: a longitudinal cohort study. Lancet 2021 Aug;398(10302):747-758. [doi: 10.1016/S0140-6736(21)01755-4]
52. Maestrini V, Birtolo LI, Francone M, et al. Cardiac involvement in consecutive unselected hospitalized COVID-19 population: in-hospital evaluation and one-year follow-up. Int J Cardiol 2021 Sep 15;339:235-242. [doi: 10.1016/j.ijcard.2021.06.056] [Medline: 34224766]
53. Li D, Liao X, Ma Z, et al. Clinical status of patients 1 year after hospital discharge following recovery from COVID-19: a prospective cohort study. Ann Intensive Care 2022 Jul 10;12(1):64. [doi: 10.1186/s13613-022-01034-4] [Medline: 35816225]
54. Zhao Y, Yang C, An X, et al. Follow-up study on COVID-19 survivors one year after discharge from hospital. Int J Infect Dis 2021 Nov;112:173-182. [doi: 10.1016/j.ijid.2021.09.017] [Medline: 34520845]
55. Peghin M, De Martino M, Palese A, et al. Post-COVID-19 syndrome 2 years after the first wave: the role of humoral response, vaccination and reinfection. Open Forum Infect Dis 2023 Jul;10(7):ofad364. [doi: 10.1093/ofid/ofad364] [Medline: 37520419]
56. Kirchberger I, Meisinger C, Warm TD, Hyhlik-Dürr A, Linseisen J, Goßlau Y. Post-COVID-19 syndrome in non-hospitalized individuals: healthcare situation 2 years after SARS-CoV-2 infection. Viruses 2023 Jun 5;15(6):1326. [doi: 10.3390/v15061326] [Medline: 37376625]
57. Zuschlag D, Grandt D, Custodis F, Braun C, Häuser W. Spontaneously reported persistent symptoms related to coronavirus disease 2019 one year after hospital discharge: a retrospective cohort single-center study. Schmerz 2022 Oct;36(5):315-325. [doi: 10.1007/s00482-022-00626-0] [Medline: 35217881]
58. Zhang X, Wang F, Shen Y, et al. Symptoms and health outcomes among survivors of COVID-19 infection 1 year after discharge from hospitals in Wuhan, China. JAMA Netw Open 2021 Sep 1;4(9):e2127403. [doi: 10.1001/jamanetworkopen.2021.27403] [Medline: 34586367]
59. Millet C, Narvaneni S, Roman S, et al. Symptoms persist in patients two years after COVID-19 infection: a prospective follow-up study. Clin Microbiol Infect 2022 Nov;28(11):1505-1507. [doi: 10.1016/j.cmi.2022.06.008] [Medline: 35724867]
60. Liu T, Wu D, Yan W, et al. Twelve-month systemic consequences of coronavirus disease 2019 (COVID-19) in patients discharged from hospital: a prospective cohort study in Wuhan, China. Clin Infect Dis 2022 Jun 10;74(11):1953-1965. [doi: 10.1093/cid/ciab703] [Medline: 34390330]
61. Wahlgren C, Forsberg G, Divanoglou A, et al. Two-year follow-up of patients with post-COVID-19 condition in Sweden: a prospective cohort study. Lancet Reg Health Eur 2023 May;28:100595. [doi: 10.1016/j.lanepe.2023.100595] [Medline: 36855599]
62. Nehme M, Vieux L, Kaiser L, et al. The longitudinal study of subjective wellbeing and absenteeism of healthcare workers considering post-COVID condition and the COVID-19 pandemic toll. Sci Rep 2023 Jul 4;13(1):10759. [doi: 10.1038/s41598-023-37568-1] [Medline: 37402726]

XSL•FO
RenderX

63. Berentschot JC, Bek LM, Heijenbrok-Kal MH, et al. Long-term health outcomes of COVID-19 in ICU- and non-ICU-treated patients up to 2 years after hospitalization: a longitudinal cohort study (CO-FLOW). J Intensive Care 2024 Nov 8;12(1):47. [doi: 10.1186/s40560-024-00748-w] [Medline: 39516956]
64. Rocha RPS, Andrade ADS, Melanda FN, Muraro AP. Síndrome pós-COVID-19 entre hospitalizados por COVID-19: estudo de coorte após 6 e 12 meses da alta hospitalar [Post-COVID-19 syndrome among hospitalized COVID-19 patients: a cohort study assessing patients 6 and 12 months after hospital discharge]. Cad Saúde Pública ;40(2):e00027423. [doi: 10.1590/0102-311xpt027423]
65. Zhang D, Chen C, Xie Y, et al. Prevalence and risk factors of long COVID-19 persisting for 2 years in Hainan Province: a population-based prospective study. Sci Rep 2025 Jan 2;15(1):369. [doi: 10.1038/s41598-024-84598-4] [Medline: 39747631]
66. Martino GP, Benfaremo D, Bitti G, et al. 6 and 12 month outcomes in patients following COVID-19-related hospitalization: a prospective monocentric study. Intern Emerg Med 2022 Sep;17(6):1641-1649. [doi: 10.1007/s11739-022-02979-x] [Medline: 35397762]
67. Huang C, Huang L, Wang Y, et al. RETRACTED: 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. Lancet 2021 Jan;397(10270):220-232. [doi: 10.1016/S0140-6736(20)32656-8]
68. Emad Eldin EM, Mohammed RS, Batanony M, Al-Sharawy LA. 12-month risk factor evaluation for persistent pulmonary symptoms in long COVID-19 patients. Egypt J Bronchol ;18(1). [doi: 10.1186/s43168-024-00265-2]
69. Nayani S, Castanares-Zapatero D, De Pauw R, et al. Classification of post COVID-19 condition symptoms: a longitudinal study in the Belgian population. BMJ Open 2023 Oct 6;13(10):e072726. [doi: 10.1136/bmjopen-2023-072726] [Medline: 37802617]
70. Evans RA, Leavy OC, Richardson M, et al. Clinical characteristics with inflammation profiling of long COVID and association with 1-year recovery following hospitalisation in the UK: a prospective observational study. Lancet Respir Med 2022 Aug;10(8):761-775. [doi: 10.1016/S2213-2600(22)00127-8]
71. Straudi S, Manfredini F, Baroni A, et al. Construct validity and responsiveness of the COVID-19 Yorkshire Rehabilitation Scale (C19-YRS) in a cohort of Italian hospitalized COVID-19 patients. Int J Environ Res Public Health 2022 May 30;19(11):6696. [doi: 10.3390/ijerph19116696] [Medline: 35682280]
72. Malesevic S, Sievi NA, Baumgartner P, et al. Impaired health-related quality of life in long-COVID syndrome after mild to moderate COVID-19. Sci Rep 2023 May 12;13(1):7717. [doi: 10.1038/s41598-023-34678-8] [Medline: 37173355]
73. Karaarslan F, Güneri FD, Kardeş S. Long COVID: rheumatologic/musculoskeletal symptoms in hospitalized COVID-19 survivors at 3 and 6 months. Clin Rheumatol 2022 Jan;41(1):289-296. [doi: 10.1007/s10067-021-05942-x] [Medline: 34713356]
74. Seang S, Itani O, Monsel G, et al. Long COVID-19 symptoms: clinical characteristics and recovery rate among non-severe outpatients over a six-month follow-up. Infect Dis Now 2022 May;52(3):165-169. [doi: 10.1016/j.idnow.2022.02.005] [Medline: 35158095]
75. Martinez AE, Banderet F, Labhardt ND, Battegay M. Long-term outcome after SARS-CoV-2 infection in healthcare workers: a single centre cohort study. Swiss Med Wkly 2021 Oct 11;151(4142):w30094. [doi: 10.4414/smw.2021.w30094] [Medline: 34694107]
76. da Costa E Silva GR, Moura W, Dos Santos KC, et al. Long-term symptoms after mild coronavirus disease in healthy healthcare professionals: a 12-month prospective cohort study. Int J Environ Res Public Health 2023 Jan 13;20(2):1483. [doi: 10.3390/ijerph20021483] [Medline: 36674238]
77. Comelli A, Viero G, Bettini G, et al. Patient-reported symptoms and sequelae 12 months after COVID-19 in hospitalized adults: a multicenter long-term follow-up study. Front Med (Lausanne) 2022;9:834354. [doi: 10.3389/fmed.2022.834354] [Medline: 35391879]
78. Och A, Tylicki P, Polewska K, et al. Persistent post-COVID-19 syndrome in hemodialyzed patients—a longitudinal cohort study from the North of Poland. J Clin Med 2021 Sep 28;10(19):4451. [doi: 10.3390/jcm10194451] [Medline: 34640471]
79. Seeßle J, Waterboer T, Hippchen T, et al. Persistent symptoms in adult patients 1 year after coronavirus disease 2019 (COVID-19): a prospective cohort study. Clin Infect Dis 2022 Apr 9;74(7):1191-1198. [doi: 10.1093/cid/ciab611] [Medline: 34223884]
80. Tok PSK, Kang KY, Ng SW, et al. Post COVID-19 condition among adults in Malaysia following the Omicron wave: a prospective cohort study. PLOS ONE 2024;19(1):e0296488. [doi: 10.1371/journal.pone.0296488] [Medline: 38181017]
81. Smith P, De Pauw R, Van Cauteren D, et al. Post COVID-19 condition and health-related quality of life: a longitudinal cohort study in the Belgian adult population. BMC Public Health 2023 Jul 27;23(1):1433. [doi: 10.1186/s12889-023-16336-w] [Medline: 37495947]
82. Monteiro IP, Appelt PA, Sisconetto AT, et al. Post-coronavirus disease 2019 functional impairments, limitations, and restrictions: a prospective cohort study based on the international classification of functioning, disability, and health. J Cent Nerv Syst Dis 2023;15:11795735231195759. [doi: 10.1177/11795735231195759] [Medline: 38025402]
83. Shah S, Bhattarai SR, Basnet K, et al. Post-COVID syndrome: a prospective study in a tertiary hospital of Nepal. PLOS ONE 2022;17(8):e0272636. [doi: 10.1371/journal.pone.0272636] [Medline: 35947595]

XSL•FO
RenderX

84. Augustin M, Schommers P, Stecher M, et al. Post-COVID syndrome in non-hospitalised patients with COVID-19: a longitudinal prospective cohort study. Lancet Reg Health Eur 2021 Jul;6:100122. [doi: 10.1016/j.lanepe.2021.100122] [Medline: 34027514]
85. Dryden M, Mudara C, Vika C, et al. Post-COVID-19 condition 3 months after hospitalisation with SARS-CoV-2 in South Africa: a prospective cohort study. Lancet Glob Health 2022 Sep;10(9):e1247-e1256. [doi: 10.1016/S2214-109X(22)00286-8] [Medline: 35961348]
86. Malinowska A, Muchlado M, Ślizień Z, et al. Post-COVID-19 syndrome and decrease in health-related quality of life in kidney transplant recipients after SARS-COV-2 infection—a cohort longitudinal study from the North of Poland. J Clin Med 2021 Nov 8;10(21):5205. [doi: 10.3390/jcm10215205] [Medline: 34768725]
87. Hyassat D, El-Khateeb M, Dahbour A, et al. Post-COVID-19 syndrome among healthcare workers in Jordan. East Mediterr Health J 2023 Apr 27;29(4):247-253. [doi: 10.26719/emhj.23.029] [Medline: 37246435]
88. Rodríguez Onieva A, Vallejo Basurte C, Fernández Bersabé A, et al. Clinical characterization of the persistent COVID-19 symptoms: a descriptive observational study in primary care. J Prim Care Community Health 2023 Jan;14. [doi: 10.1177/21501319231208283]
89. Gaspar P, Dias M, Parreira I, et al. Predictors of long-COVID-19 and its impact on quality of life: longitudinal analysis at 3, 6 and 9 months after discharge from a Portuguese centre. Acta Med Port 2023 Oct 2;36(10):647-660. [doi: 10.20344/amp.19047] [Medline: 36827994]
90. Guliani A, Tandon A, Chakroborty A, Gupta PP. Predictors of post COVID complications in patients admitted with moderate to severe COVID symptoms: a single center, prospective, observational study. Monaldi Arch Chest Dis 2022 Jul 20;93(2). [doi: 10.4081/monaldi.2022.2307] [Medline: 35872628]
91. Lapa J, Rosa D, Mendes JPL, Deusdará R, Romero GAS. Prevalence and associated factors of post-COVID-19 syndrome in a Brazilian cohort after 3 and 6 months of hospital discharge. Int J Environ Res Public Health 2023 Jan 2;20(1):848. [doi: 10.3390/ijerph20010848] [Medline: 36613170]
92. Kalak G, Jarjou'i A, Bohadana A, et al. Prevalence and persistence of symptoms in adult COVID-19 survivors 3 and 18 months after discharge from hospital or corona hotels. J Clin Med 2022 Dec 14;11(24):7413. [doi: 10.3390/jcm11247413] [Medline: 36556030]
93. Muz F, Metintaş S, Önsüz MF, et al. Prevalence of post-COVID-19 syndrome and related factors among university employees: a prospective cohort study. J Clin Pract Res 2023;45(3):258-264. [doi: 10.14744/etd.2023.56313]
94. Degli Antoni M, Maifredi G, Storti S, et al. Long-term symptoms after SARS-CoV-2 infection in a cohort of people living with HIV. Infection 2024 Dec;52(6):2339-2350. [doi: 10.1007/s15010-024-02288-9] [Medline: 38700658]
95. Emecen AN, Keskin S, Turunc O, et al. The presence of symptoms within 6 months after COVID-19: a single-center longitudinal study. Ir J Med Sci 2023 Apr;192(2):741-750. [doi: 10.1007/s11845-022-03072-0]
96. Méndez R, Balanzá-Martínez V, Luperdi SC, et al. Long-term neuropsychiatric outcomes in COVID-19 survivors: a 1-year longitudinal study. J Intern Med 2022 Feb;291(2):247-251. [doi: 10.1111/joim.13389] [Medline: 34569681]
97. Bellan M, Baricich A, Patrucco F, et al. Long-term sequelae are highly prevalent one year after hospitalization for severe COVID-19. Sci Rep 2021 Nov 22;11(1):22666. [doi: 10.1038/s41598-021-01215-4] [Medline: 34811387]
98. Rank A, Tzortzini A, Kling E, et al. One year after mild COVID-19: the majority of patients maintain specific immunity, but one in four still suffer from long-term symptoms. J Clin Med 2021 Jul 27;10(15):3305. [doi: 10.3390/jcm10153305] [Medline: 34362088]
99. Chai C, Feng X, Lu M, et al. One-year mortality and consequences of COVID-19 in cancer patients: a cohort study. IUBMB Life 2021 Oct;73(10):1244-1256. [doi: 10.1002/iub.2536] [Medline: 34318585]
100. Tortajada C, Navarro A, Andreu-Ballester JC, Mayor A, Añón S, Flores J. Prevalence and duration of symptoms among moderate and severe COVID-19 patients 12 months after discharge. Intern Emerg Med 2022 Apr;17(3):929-934. [doi: 10.1007/s11739-021-02895-6] [Medline: 35023003]
101. Fernández-de-Las-Peñas C, Guijarro C, Plaza-Canteli S, Hernández-Barrera V, Torres-Macho J. Prevalence of post-COVID-19 cough one year after SARS-CoV-2 infection: a multicenter study. Lung 2021 Jun;199(3):249-253. [doi: 10.1007/s00408-021-00450-w] [Medline: 33993321]
102. Boscolo‐Rizzo P, Guida F, Polesel J, et al. Sequelae in adults at 12 months after mild‐to‐moderate coronavirus disease 2019 (COVID‐19). Int Forum Allergy Rhinol 2021 Dec;11(12):1685-1688. [doi: 10.1002/alr.22832]
103. Tessitore E, Handgraaf S, Poncet A, et al. Symptoms and quality of life at 1-year follow up of patients discharged after an acute COVID-19 episode. Swiss Med Wkly 2021 Dec 6;151(4950):w30093. [doi: 10.4414/smw.2021.w30093] [Medline: 34909433]
104. Yang X, Hou C, Shen Y, et al. Two-year health outcomes in hospitalized COVID-19 survivors in China. JAMA Netw Open 2022 Sep 1;5(9):e2231790. [doi: 10.1001/jamanetworkopen.2022.31790] [Medline: 36107425]
105. Ho FF, Xu S, Kwong TMH, et al. Prevalence, patterns, and clinical severity of long COVID among Chinese medicine telemedicine service users: preliminary results from a cross-sectional study. Int J Environ Res Public Health 2023 Jan 19;20(3):1827. [doi: 10.3390/ijerph20031827] [Medline: 36767195]

XSL•FO
RenderX

106. Gutiérrez-Canales LG, Muñoz-Corona C, Barrera-Chávez I, et al. Persistence of COVID-19 symptoms and quality of life at three and twelve months after hospital discharge. Medicina (Kaunas) 2024 Jun 5;60(6):944. [doi: 10.3390/medicina60060944] [Medline: 38929561]
107. Petrakis V, Rafailidis P, Terzi I, et al. The prevalence of long COVID-19 syndrome in hospitalized patients with COVID-19 pneumonia. Infez Med 2024 Jun 1;32(2):202-212. [doi: 10.53854/liim-3202-8]
108. Newlands F, Rojas NK, Nugawela M, et al. A cross-sectional study of the health of emerging young adults in England following a COVID-19 infection. J Adolesc Health 2023 Jul;73(1):20-28. [doi: 10.1016/j.jadohealth.2023.01.026] [Medline: 37024311]
109. Ahmed SS, Adil PI, Rasheed NA, Hussein NR, Dhama K. A study of long COVID-19 in Duhok, Kurdistan region, Iraq. J Infect Dev Ctries 2023 Jun 30;17(6):805-811. [doi: 10.3855/jidc.17468] [Medline: 37406070]
110. Shang H, Chang T, Yang W, et al. Analysis of influencing factors on long COVID in COVID-19 patients infected with omicron variant three months after discharge: a cross-sectional study. BMC Infect Dis 2024 Jan 2;24(1):36. [doi: 10.1186/s12879-023-08947-w] [Medline: 38166694]
111. Daitch V, Yelin D, Awwad M, et al. Characteristics of long-COVID among older adults: a cross-sectional study. Int J Infect Dis 2022 Dec;125:287-293. [doi: 10.1016/j.ijid.2022.09.035] [Medline: 36191820]
112. Gasnier M, Choucha W, Radiguer F, et al. Comorbidity of long COVID and psychiatric disorders after a hospitalisation for COVID-19: a cross-sectional study. J Neurol Neurosurg Psychiatry 2022 Oct;93(10):1091-1098. [doi: 10.1136/jnnp-2021-328516]
113. Morioka S, Tsuzuki S, Maruki T, et al. Epidemiology of post-COVID conditions beyond 1 year: a cross-sectional study. Public Health (Fairfax) 2023 Mar;216:39-44. [doi: 10.1016/j.puhe.2023.01.008] [Medline: 36791649]
114. da Silva MRT, Costa AP, da Luz AA, et al. From acute phase to long COVID: a cross-sectional study of the epidemiological profile and clinical evaluation of SARS-CoV-2 infection in employees at a pediatric hospital. J Clin Med 2023 Jun 27;12(13):4325. [doi: 10.3390/jcm12134325] [Medline: 37445359]
115. Kumar NP, Thompson S, Ramchandra T, Waghmare R. Long COVID disease among symptomatic and asymptomatic COVID-19 patients: a multi-centric descriptive study. Clin Med (Northfield) 2022;9(6):1549-1555.
116. Bougea A, Georgakopoulou VE, Palkopoulou M, et al. New-onset non-motor symptoms in patients with Parkinson's disease and post-COVID-19 syndrome: a prospective cross-sectional study. Med Int (Lond) 2023;3(3):23. [doi: 10.3892/mi.2023.83] [Medline: 37214229]
117. Leon L, Perez-Sancristobal I, Madrid A, et al. Persistent post-discharge symptoms after COVID-19 in rheumatic and musculoskeletal diseases. Rheumatol Adv Pract 2022;6(1):rkac008. [doi: 10.1093/rap/rkac008] [Medline: 35233480]
118. Raj ASV, Jacob A, Ambu V, Wilson T, Renuka R. Post COVID-19 clinical manifestations and its risk factors among patients in a Northern District in Kerala, India. J Family Med Prim Care 2022;11(9):5312-5319. [doi: 10.4103/jfmpc.jfmpc_131_22]
119. Tran TK, Nguyen Truong S, Tran Thanh L, Le Hoang Gia N, Pham Trung H, Tran Dinh B. Post-COVID condition: a survey of patients recovered from COVID-19 in Central Vietnam. J Infect Dev Ctries 2023 Sep 30;17(9):1213-1220. [doi: 10.3855/jidc.18029] [Medline: 37824353]
120. Oliveira TD, Almeida HS, Pereira IB, et al. Post-COVID conditions: a Portuguese primary care study. RPMGF 2023 Aug 31;39(4):351-354. [doi: 10.32385/rpmgf.v39i4.13443]
121. Maamar M, Artime A, Pariente E, et al. Post-COVID-19 syndrome, low-grade inflammation and inflammatory markers: a cross-sectional study. Curr Med Res Opin 2022 Jun;38(6):901-909. [doi: 10.1080/03007995.2022.2042991] [Medline: 35166141]
122. Kosowan L, Sanchez-Ramirez DC, Katz A. Understanding symptoms suggestive of long COVID syndrome and healthcare use among community-based populations in Manitoba, Canada: an observational cross-sectional survey. BMJ Open 2024 Jan 12;14(1):e075301. [doi: 10.1136/bmjopen-2023-075301] [Medline: 38216180]
123. Maestre-Muñiz MM, Arias Á, Mata-Vázquez E, et al. Long-term outcomes of patients with coronavirus disease 2019 at one year after hospital discharge. J Clin Med 2021 Jun 30;10(13):2945. [doi: 10.3390/jcm10132945] [Medline: 34209085]
124. Fernández-de-Las-Peñas C, Rodríguez-Jiménez J, Cancela-Cilleruelo I, et al. Post-COVID-19 symptoms 2 years after SARS-CoV-2 infection among hospitalized vs nonhospitalized patients. JAMA Netw Open 2022 Nov 1;5(11):e2242106. [doi: 10.1001/jamanetworkopen.2022.42106] [Medline: 36378309]
125. Iqbal A, Iqbal K, Arshad Ali S, et al. The COVID-19 sequelae: a cross-sectional evaluation of post-recovery symptoms and the need for rehabilitation of COVID-19 survivors. Cureus 2021 Feb 2;13(2):e13080. [doi: 10.7759/cureus.13080] [Medline: 33680620]
126. Savith A, Meah A, Murthy RSS, Phal NB. Long-term health implications of coronavirus disease 2019: a prospective study on post-coronavirus disease 2019 symptoms. Ann Afr Med 2025;24(1):167-172. [doi: 10.4103/aam.aam_199_23]
127. de Bruijn S, Tulen AD, Rodenburg J, et al. Post-acute sequelae of COVID-19 3 to 12 months after infection: Delta vs Omicron. Int J Infect Dis 2025 Jan;150:107302. [doi: 10.1016/j.ijid.2024.107302] [Medline: 39549783]
128. Pérez Catalán I, Roig Martí C, Folgado Escudero S, et al. Presence of COVID-19 self-reported symptoms at 12 months in patients discharged from hospital in 2020-2021: a Spanish cross-sectional study. Sci Rep 2024 Nov 4;14(1):26575. [doi: 10.1038/s41598-024-78017-x] [Medline: 39496842]

XSL•FO
RenderX

129. Arnold DT, Hamilton FW, Milne A, et al. Patient outcomes after hospitalisation with COVID-19 and implications for follow-up: results from a prospective UK cohort. Thorax 2021 Apr;76(4):399-401. [doi: 10.1136/thoraxjnl-2020-216086] [Medline: 33273026]
130. Xiong Q, Xu M, Li J, et al. Clinical sequelae of COVID-19 survivors in Wuhan, China: a single-centre longitudinal study. Clin Microbiol Infect 2021 Jan;27(1):89-95. [doi: 10.1016/j.cmi.2020.09.023] [Medline: 32979574]
131. Peluso MJ, Kelly JD, Lu S, et al. Rapid implementation of a cohort for the study of post-acute sequelae of SARS-CoV-2 infection/COVID-19. medRxiv. Preprint posted online on Mar 12, 2021. [doi: 10.1101/2021.03.11.21252311] [Medline: 33758895]
132. Blitshteyn S, Whitelaw S. Postural orthostatic tachycardia syndrome (POTS) and other autonomic disorders after COVID-19 infection: a case series of 20 patients. Immunol Res 2021 Apr;69(2):205-211. [doi: 10.1007/s12026-021-09185-5] [Medline: 33786700]
133. Scott-Solomon E, Boehm E, Kuruvilla R. The sympathetic nervous system in development and disease. Nat Rev Neurosci 2021 Nov;22(11):685-702. [doi: 10.1038/s41583-021-00523-y] [Medline: 34599308]
134. Kwan AC, Ebinger JE, Wei J, et al. Apparent risks of postural orthostatic tachycardia syndrome diagnoses after COVID-19 vaccination and SARS-Cov-2 infection. Nat Cardiovasc Res 2022 Dec;1(12):1187-1194. [doi: 10.1038/s44161-022-00177-8] [Medline: 37303827]
135. Phetsouphanh C, Darley DR, Wilson DB, et al. Immunological dysfunction persists for 8 months following initial mild-to-moderate SARS-CoV-2 infection. Nat Immunol 2022 Feb;23(2):210-216. [doi: 10.1038/s41590-021-01113-x] [Medline: 35027728]
136. Puntmann VO, Carerj ML, Wieters I, et al. Outcomes of cardiovascular magnetic resonance imaging in patients recently recovered from coronavirus disease 2019 (COVID-19). JAMA Cardiol 2020 Nov 1;5(11):1265-1273. [doi: 10.1001/jamacardio.2020.3557] [Medline: 32730619]
137. Yin K, Peluso MJ, Luo X, et al. Long COVID manifests with T cell dysregulation, inflammation and an uncoordinated adaptive immune response to SARS-CoV-2. Nat Immunol 2024 Feb;25(2):218-225. [doi: 10.1038/s41590-023-01724-6] [Medline: 38212464]
138. Cervia C, Zurbuchen Y, Taeschler P, et al. Immunoglobulin signature predicts risk of post-acute COVID-19 syndrome. Nat Commun 2022 Jan 25;13(1):446. [doi: 10.1038/s41467-021-27797-1] [Medline: 35078982]
139. Choutka J, Jansari V, Hornig M, Iwasaki A. Unexplained post-acute infection syndromes. Nat Med 2022 May;28(5):911-923. [doi: 10.1038/s41591-022-01810-6] [Medline: 35585196]
140. Zollner A, Koch R, Jukic A, et al. Clearance of gut mucosal SARS-CoV-2 antigens and postacute COVID-19 after 2 years in patients with inflammatory bowel disease. Gastroenterology 2024 Aug;167(3):604-607. [doi: 10.1053/j.gastro.2024.04.008] [Medline: 38631418]
141. Hashimoto K. Detrimental effects of COVID-19 in the brain and therapeutic options for long COVID: the role of Epstein-Barr virus and the gut-brain axis. Mol Psychiatry 2023 Dec;28(12):4968-4976. [doi: 10.1038/s41380-023-02161-5] [Medline: 37402856]
142. Conway EM, Mackman N, Warren RQ, et al. Understanding COVID-19-associated coagulopathy. Nat Rev Immunol 2022 Oct;22(10):639-649. [doi: 10.1038/s41577-022-00762-9] [Medline: 35931818]
143. Appelman B, Charlton BT, Goulding RP, et al. Muscle abnormalities worsen after post-exertional malaise in long COVID. Nat Commun 2024 Jan 4;15(1):17. [doi: 10.1038/s41467-023-44432-3] [Medline: 38177128]
144. Libby P, Lüscher T. COVID-19 is, in the end, an endothelial disease. Eur Heart J 2020 Sep 1;41(32):3038-3044. [doi: 10.1093/eurheartj/ehaa623] [Medline: 32882706]
145. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. Nat Med 2021 Apr;27(4):601-615. [doi: 10.1038/s41591-021-01283-z]
146. Wise J. Covid-19: Increased risk of some neurological and psychiatric disorders remains two years after infection, study finds. BMJ 2022 Aug 17;378:o2048. [doi: 10.1136/bmj.o2048]
147. Taquet M, Sillett R, Zhu L, et al. Neurological and psychiatric risk trajectories after SARS-CoV-2 infection: an analysis of 2-year retrospective cohort studies including 1 284 437 patients. Lancet Psychiatry 2022 Oct;9(10):815-827. [doi: 10.1016/S2215-0366(22)00260-7] [Medline: 35987197]
148. Madhi SA, Nel J. Epidemiology of severe COVID-19 from South Africa. Lancet HIV 2021 Sep;8(9):e524-e526. [doi: 10.1016/S2352-3018(21)00183-1] [Medline: 34363788]
149. Sampaio-Barros PD, Bortoluzzo AB, Conde RA, Costallat LTL, Samara AM, Bértolo MB. Undifferentiated spondyloarthritis: a longterm followup. J Rheumatol 2010 Jun;37(6):1195-1199. [doi: 10.3899/jrheum.090625] [Medline: 20436080]
150. Geirsson AJ, Eyjolfsdottir H, Bjornsdottir G, Kristjansson K, Gudbjornsson B. Prevalence and clinical characteristics of ankylosing spondylitis in Iceland—a nationwide study. Clin Exp Rheumatol 2010;28(3):333-340. [Medline: 20406616]
151. Gennaro FD, Veronese N, Segala FV, et al. Protective role of vaccination on the development of long COVID: data from a large, multicenter, prospective cohort study. BMC Infect Dis 2024 Nov 18;24(1):1313. [doi: 10.1186/s12879-024-10226-1] [Medline: 39558292]

XSL•FO
RenderX

152. Bai F, Santoro A, Hedberg P, et al. The Omicron variant is associated with a reduced risk of the post COVID-19 condition and its main phenotypes compared to the wild-type virus: results from the EuCARE-POSTCOVID-19 Study. Viruses 2024 Sep 23;16(9):1500. [doi: 10.3390/v16091500] [Medline: 39339976]

Abbreviations

***ICD-11*:** *International Classification of Diseases, 11th Revision*
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by A Mavragani, T Sanchez; submitted 05.02.25; peer-reviewed by A Adeoye, D Fan; revised version received 09.05.25; accepted 19.05.25; published 18.07.25.

Please cite as:
Hu M, Song T, Gong Z, Che Q, Guo J, Chen L, Zhang H, Li H, Liang N, Zhao G, Wang Y, Shi N, Liu B
Symptom Trajectories and Clinical Subtypes in Post–COVID-19 Condition: Systematic Review and Clustering Analysis
JMIR Public Health Surveill 2025;11:e72221
URL: https://publichealth.jmir.org/2025/1/e72221
doi:10.2196/72221

© Mingzhi Hu, Tian Song, Zhaoyuan Gong, Qianzi Che, Jing Guo, Lin Chen, Haili Zhang, Huizhen Li, Ning Liang, Guozhen Zhao, Yanping Wang, Nannan Shi, Bin Liu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Leveraging Administrative Health Databases to Address Health Challenges in Farming Populations: Scoping Review and Bibliometric Analysis (1975-2024)

Pascal Petit[1], PhD; Nicolas Vuillerme[1,2], PhD, Prof Dr

[1]Laboratoire AGEIS, Université Grenoble Alpes, La Tronche Cedex, France

[2]Institut Universitaire de France, Paris, France

Corresponding Author:
Pascal Petit, PhD
Laboratoire AGEIS
Université Grenoble Alpes
Bâtiment Jean Roget, UFR de Médecine
Domaine de La Merci
La Tronche Cedex, 38706
France
Phone: 33 4 76 63 71 04
Email: pascal.petit@univ-grenoble-alpes.fr

Abstract

Background: Although agricultural health has gained importance, to date, much of the existing research relies on traditional epidemiological approaches that often face limitations related to sample size, geographic scope, temporal coverage, and the range of health events examined. To address these challenges, a complementary approach involves leveraging and reusing data beyond its original purpose. Administrative health databases (AHDs) are increasingly reused in population-based research and digital public health, especially for populations such as farmers, who face distinct environmental risks.

Objective: We aimed to explore the reuse of AHDs in addressing health issues within farming populations by summarizing the current landscape of AHD-based research and identifying key areas of interest, research gaps, and unmet needs.

Methods: We conducted a scoping review and bibliometric analysis using PubMed and Web of Science. Building upon previous reviews of AHD-based public health research, we conducted a comprehensive literature search using 72 terms related to the farming population and AHDs. To identify research hot spots, directions, and gaps, we used keyword frequency, co-occurrence, and thematic mapping. We also explored the bibliometric profile of the farming exposome by mapping keyword co-occurrences between environmental factors and health outcomes.

Results: Between 1975 and April 2024, 296 publications across 118 journals, predominantly from high-income countries, were identified. Nearly one-third of these publications were associated with well-established cohorts, such as Agriculture and Cancer and Agricultural Health Study. The most frequently used AHDs included disease registers (158/296, 53.4%), electronic health records (124/296, 41.9%), insurance claims (106/296, 35.8%), population registers (95/296, 32.1%), and hospital discharge databases (41/296, 13.9%). Fifty (16.9%) of 296 studies involved >1 million participants. Although a broad range of exposure proxies were used, most studies (254/296, 85.8%) relied on broad proxies, which failed to capture the specifics of farming tasks. Research on the farming exposome remains underexplored, with a predominant focus on the specific external exposome, particularly pesticide exposure. A limited range of health events have been examined, primarily cancer, mortality, and injuries.

Conclusions: The increasing use of AHDs holds major potential to advance public health research within farming populations. However, substantial research gaps persist, particularly in low-income regions and among underrepresented farming subgroups, such as women, children, and contingent workers. Emerging issues, including exposure to per- and polyfluoroalkyl substances, biological agents, microbiome, microplastics, and climate change, warrant further research. Major gaps also persist in understanding various health conditions, including cardiovascular, reproductive, ocular, sleep-related, age-related, and autoimmune diseases. Addressing these overlooked areas is essential for comprehending the health risks faced by farming communities and guiding public health policies. Within this context, promoting AHD-based research, in conjunction with other digital data sources (eg, mobile health, social health data, and wearables) and artificial intelligence approaches, represents a promising avenue for future exploration.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e62939) doi:10.2196/62939

KEYWORDS

farming population; digital public health; digital epidemiology; administrative health database; farming exposome; review; bibliometric analysis; data reuse

Introduction

Background

Public health research seeks to identify and understand the factors that influence population health to effectively prevent diseases and promote health and well-being for all [1,2]. A broad range of environmental determinants can impact health across the life span. One of the core areas of public health research, known as the exposome, investigates how cumulative environmental influences contribute to disease etiology and pathogenesis [3-18]. The exposome, which complements genomic research, refers to the comprehensive examination of all environmental exposures experienced throughout an individual's lifetime, including physical, chemical, biological, psychosocial, and behavioral factors, from conception to death [3-18]. The exposome classically includes 3 overlapping domains: the general external exposome (eg, climate and built environment); the specific external exposome (eg, chemical exposure, lifestyle, and occupations); and the internal exposome (eg, aging, oxidative stress, metabolism, and gut microbiome) [8,14,16,17,19,20]. Understanding the exposome is crucial for enabling both population-wide and precision prevention [3,21-23]. However, fully describing the exposome is challenging due to the vast diversity and the temporal and spatial variability of environmental factors [3]. Public health research in this area requires data on both risk factors and adverse health outcomes to progress effectively [3,14,24,25].

The volume of data collected has grown exponentially as the world becomes increasingly reliant on technology and digitization [26,27]. Data are omnipresent in our everyday lives, leading science toward data-driven research [27,28], in particular in the health field. The digital transformation in health care has enabled unprecedented data availability, collection, storage, and analysis capabilities, leading to a paradigm shift in health care systems, with entire care pathways becoming digitized [29,30]. Health-related data now represent approximately 6% of all digital data globally, a figure that continues to rise [31]. This explosion of data has transformed research, providing new opportunities, especially in public health, to enhance disease understanding and evaluate intervention effectiveness [27,28,32-35]. The integration of digital technologies and digital data in public health has led to the emergence of "digital public health," an evolving field focused on using digital data to achieve public health goals [33,36-41]. Public health research is moving from isolated data systems to more integrated, accessible, and reusable data resources [42]. Reusing data allows researchers to explore various health determinants, including environmental, occupational, behavioral, and organizational factors, fostering a holistic approach to disease prevention and health promotion strategies [14].

Within the digital public health framework, 2 main types of data are being used, namely primary and secondary data. Primary data are tailor-made, designed for a specific purpose, and often used once or repeatedly for the same goal [43-45]. Primary data are the cornerstone of traditional public health policy and decision-making. These data are derived from several types of studies [46-50], in particular observational cohorts (eg, the Framingham Cohort study [51,52]) [46-50,53,54], case-control studies [46-50], cross-sectional surveys (eg, the China Health and Retirement Longitudinal Study [55,56]) [46-50,53], and experimental studies [46-50]. Primary data have many advantages [46-50]. They are rich, of high quality, and are designed to answer specific research questions for public health and epidemiological purposes. Primary data are usually available at the individual level and are derived from studies that control for certain biases. By contrast, they are cumbersome, time-consuming, and costly to set up and maintain [53,54,57]. The representativeness of primary data is also limited in size, geographic scope, and temporal coverage and can erode with time [46-50,53]. Primary data are not free from bias, such as selection, healthy worker, recall, or prevarication biases [53,58].

Unlike traditional public health, digital public health does not rely solely on primary data but takes advantage of the myriad of existing digital data that have not been generated originally for research purposes (ie, secondary data) to overcome some limitations intrinsic to primary data and complement them [28,43,44,53,59-68]. Indeed, some data can have an additional impact when used beyond the context for which they were originally created [68,69]. Secondary data are collected for purposes other than public health or epidemiology and include contextual data (eg, air quality and climate data) [14,24,26,29,70-74], person-generated data (eg, social media, crowdsourcing, and mobile health) [2,24,26,31,43,61,62,73,75-86], synthetic data (eg, digital twin) [87-91], and administrative health databases (AHDs) [26,64,68,81,92-102].

AHD is a broad term encompassing a wide range of routinely collected data on individuals' health and sociodemographic information collected for registration, billing, record keeping, and other administrative purposes [26,64,81,93,95,98,100-102]. For this review, based on previous works [93,95,96,103-105], AHDs included population registers, claims databases, disease registers, electronic health or medical records, and hospital discharge databases that were collected at a local, regional, national, or international level (Table 1) [26,61,62,93,95,96,100,103-110].

XSL•FO
RenderX

Table 1. Definition and characteristics of administrative health databases included in this review.

	Population register	Claims database	Disease register	Electronic health or medical record	Hospital discharge database
Definition	Digital sociodemographic information on the residents of a country	Routinely collected digital information on individual data regarding reimbursement, records of health services, medical procedures, prescriptions, and medical diagnoses	A continuous and exhaustive digital collection of individual data regarding 1 or more health events in a geographically defined population	Systematized digital record of a patient's medical information collected in real time	Digital records of service use with information about patients, their care, and their stay in the hospital
Source	Local or national authorities	Insurance programs or schemes and health care providers	Health care institutions (eg, hospitals)	Hospitals, physicians, health care centers, and institutions	Hospitals
Population	All individuals residing in a country	All individuals covered by an insurance program or scheme	All individuals diagnosed with a specific health event in a population on a geographically defined scale	All patients using the health care system	All patients from a hospital
Purpose or finality	For the administrative purposes of government: to provide reliable information	To store financial and administrative information for medical insurers' and providers' use	For clinical and research purposes: to collect information about people diagnosed with a specific health event	For clinical and billing purposes: to document patients' clinical condition	For billing or accounting purposes
Health event	None	Health events covered by insurance or a health care provider	Specific health events (eg, cancer)	Health events requiring care that are reported in medical records	Health events from hospital admission

AHDs offer many advantages for research. Such data are collected as part of routine administrative processes, reducing additional costs for researchers. Therefore, AHDs offer relatively inexpensive access to a large number of individuals who can be tracked with time for several years, guaranteeing the representativeness of the populations studied [26,54,78,93,95,104,105,111-114]. Data recorded within AHDs are structured, coded in a standardized way, and less affected by participation and recall biases [54,58,95,106,113,115]. AHDs enable the study of rare events and populations underrepresented in studies using only primary data [95,111-113]. AHDs have limitations inherent to their nature, such as the absence of some confounding factors, the limited granularity of certain information, the data complexity, and confidentiality issues [73,78,93,95,115-125].

Rationale

AHDs are increasingly used in population-based health research due to their complementarity with traditional sources of public health and epidemiological data (ie, primary data) [42,59,64,87,93,95,96,126-128]. The reuse of AHDs, referring to their application beyond their original or intended purpose, holds major potential to advance public health and epidemiological research, offering insights that can guide public health decision-making [42,59,64,87,93,95,96,105,126-131]. Although several reviews have previously explored the general use of AHDs in research [42,95,107,129,132-135], others have focused on their application within specific countries [96,136], examined individual AHDs [108,137], or investigated their role in studying specific diseases and adverse health outcomes [44,93,104,138-142]. However, to the best of our knowledge, no study has synthesized how AHDs are reused for epidemiological and public health research within a specific population group.

To address this gap, we conducted a comprehensive scoping review and bibliometric analysis aimed at identifying how AHDs are used to address health issues in a specific population. We selected farming populations as an illustrative example because they present unique health and disease patterns [143-147]. Globally, approximately 27% of the workforce is engaged in occupational farming, and this group is exposed to numerous risk factors (ie, exposomes), including pesticides, biological agents, and limited access to health care [148]. These exposures put them at heightened risk for a wide range of adverse health outcomes [143,145,147,149]. Although agricultural safety and health have become a major public health issue in recent decades, most research on the health of farming populations has relied on traditional epidemiological and community-based studies, which often face limitations in terms of sample size, geographic scope, temporal coverage, and the range of health events examined [145,150,151].

In this context, AHDs offer valuable opportunities to enhance public health and epidemiological research in farming populations by providing broader insights, identifying at-risk subgroups, and informing health services and policy development [152]. The primary objectives of this scoping review were two-fold: (1) to summarize the current state of AHD-based research in farming populations by examining which types of AHDs are used and why, whether AHDs are integrated with other data sources, which farming populations have been studied, and what exposures and health outcomes have been explored and (2) to identify key areas of interest and potential research gaps and unmet needs in this field.

XSL•FO
RenderX

Methods

Overview

This scoping review was conducted and reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) and evidence maps guidelines (Table S1 in Multimedia Appendix 1) [153] following a single screening approach. The protocol of this study was not registered. A 7-step procedure was used: research question formulation, identifying relevant publications, title review, abstract review, full-text review, data extraction, and data analysis.

To formulate our research question, we followed the Joanna Briggs Institute guidelines, using the population, concept, and context criteria framework [154]. Our population included all individuals engaged in farming and all individuals exposed to farming-related exposures. The concepts included all possible public health and epidemiological research works that involved the study of a health outcome of interest. The context was the use, in any setting, of at least one of the AHDs, as defined in Table 1.

Search Strategy and Selection Criteria

To develop and validate the search strategy, previous reviews that examined the reuse of AHDs for population-based research were identified and refined [93,96,103,104]. Our initial search revealed that electronic health records (EHRs) are often interchangeably referred to as electronic medical records (EMRs). A distinction between EHR and EMR is sometimes made, with EMR describing patients' care from only 1 practice (eg, specific encounters in hospitals), which is contrary to EHR [105]. In that case, EMR serves as a data source for EHR. This distinction was not considered in this paper. In addition, to ensure comprehensiveness, the search terms were broadened by searching for their synonyms. For example, search terms such as "electronic health record," "digital health record," "electronic medical record," "EHR," or "EMR" were used as synonyms for electronic health or medical records. A total of 72 terms pertaining to 2 categories (farming and AHDs) were used (Textbox 1). The search terms were reflective of our research topic and question.

Textbox 1. Search terms.

Farming

- husbandry* OR agriculture* OR farming OR farm* OR agricultural* OR farmworker*

Administrative health databases (combined using AND)

- "health record" OR "health records" OR "digital record" OR "digital records" OR "health administrative register" OR "health administrative registry" OR "health register" OR "health registry" OR "medical register" OR "medical registry" OR "electronic health record" OR "electronic health records" OR "EHR" OR "EMR" OR "electronic medical record" OR "electronic medical records" OR "digital medical record" OR "digital medical records" OR "digital health record" OR "digital health records" OR "health administrative data" OR "health administrative database" OR " health administrative dataset" OR " health administrative datasets" OR "health administrative databases" OR "administrative health data" OR "administrative health database" OR "administrative health dataset" OR "administrative health datasets" OR "administrative health databases" OR "insurance data" OR "insurance database" OR "insurance databases" OR "insurance dataset" OR "insurance claim" OR "insurance claims" OR "cancer registry" OR "cancer register" OR "health insurance" OR "health surveillance program" OR "health surveillance programs" OR "Mutualite Sociale Agricole" OR "MSA" OR "health insurance system" OR "record-linkage" OR "population register" OR "population registry" OR "insurance scheme" OR "social security scheme" OR "hospital discharge" OR "administrative claim" OR "administrative claims" OR "medical claims" OR "medical claim" OR "electronic claim" OR "electronic claims" OR "mortality register" OR "mortality registry" OR "mortality record" OR "mortality records" OR "disease register" OR "disease registry" OR "illness register" OR "illness registry" OR "disorder register" OR "disorder registry"

To develop the eligibility criteria, an initial search of the literature was conducted on PubMed, with a review of the first 100 articles that used AHDs for public health and epidemiological research. In our pilot run, disease and morbidity registers were initially not considered as AHD because they were created for clinical and research purposes [47-50,53,76,155,156]. However, because disease registers contain some information derived from medical records, we decided to consider them as AHD for this review. The eligibility criteria are presented in Textbox 2. The search was restricted to original peer-reviewed records (all types were included) written in English or French but not constrained by the year of publication [93,106,157]. Publications that examined partly farming populations, with, for instance, studies reporting health risks for various sectors of activity, were included.

XSL•FO
RenderX

Textbox 2. Eligibility criteria for selection of publications.

Inclusion criteria for articles

- Data had to originate at least partly from the administrative health database (AHD)
- The study had to pertain at least partly to the farming population
- The study had to relate to public health or epidemiological research
- Original peer-reviewed publications
- Publications in English or French

Exclusion criteria for articles

- Publications not describing the use of an AHD
- Animal or in vitro studies
- Publications not in English or French

The final literature search was conducted on both PubMed and Web of Science Core Collection databases. Regarding the Web of Science Core Collection, a topic search was performed. To reduce the bias induced by daily database changes, all data collection (literature retrieval and data download) was conducted and completed on the same day, that is, April 15, 2024. Titles, abstracts, and full-text publications were screened based on pre-established inclusion and exclusion criteria. The inclusion criteria for each phase of the literature search are provided in Table S1 in Multimedia Appendix 1. When abstracts did not contain enough information about correspondence to inclusion or exclusion criteria, the article was considered for full-text review. Reference lists of included publications were not searched, although they might have also yielded new relevant studies.

Data Collection and Processing

A total of 29 metadata were extracted from each publication included in the scoping review (Table 2).

The data underwent rigorous manual validation, cleaning, and harmonization following a structured 5-step process. First, duplicate items (eg, keywords and institutions) were removed. Second, leading and trailing white spaces were eliminated. Third, items were standardized by converting text to lower case, with only the first letter capitalized. In the fourth step, items were harmonized to either singular or plural forms consistently. Finally, synonyms or terms with similar meanings (eg, “illness” and “disease”) were unified under a single term. For instance, “Pesticide,” “Pesticide exposure,” and “Pesticide use” were standardized to “Pesticide,” while “Pulmonary disease copd,” “Copd,” and “Chronic obstructive pulmonary disease” were unified as “COPD.” For cancer-related keywords, the International Classification of Diseases, eleventh revision, was used to consolidate varied terms (eg, “lung cancer,” “lung cancer risk,” “lung and bronchus cancer,” “lung tumor,” “lung tumour,” “lung neoplasm,” and “basal cell carcinoma of the lung”) into standard categories (eg, lung cancer). Quality appraisals were not performed because they were beyond the aim of this review [106,157].

XSL•FO
RenderX

Table 2. List of metadata of interest to collect from the literature search.

Metadata	Fictional example
Publication year	2024
Publication type	Article
Study name	Project X
Goal of the study	To study the association between farming and health outcome
Study type	Ecological study
Is the study nationwide?	Yes
Digital data used	Insurance claims
Goal of the digital data used	To identify farmers
Is active data used?	Yes
Active data used	Clinical examination
Farming exposure considered	Farming activity and pesticide compounds
Farming activities studied, n	10
Pesticide compounds studied, n	29
Population	Adults
Sex	Female
Participants included, n	100 to 1000
Country	France
Oldest data used (year)	1991
Most recent data used (year)	2020
Data follow-up period (years)	4
Years between the most recent data used and publication year, n	7
Disease or health events	Parkinson disease
Authors' names	Gauthier J
Authors' keywords	Pesticide
Authors' country	France
Authors' institution	Université Grenoble Alpes
Journal	Environmental Health Perspectives
Funding body	MIAI@Grenoble Alpes[a]
Citations, n	14

[a]MIAI@Grenoble Alpes: Multidisciplinary Institute in Artificial Intelligence at the Université Grenoble Alpes.

Data Analysis

To analyze the research directions (ie, hot spots and gaps) on the use of AHDs for public health and epidemiological research in farming populations, a bibliometric approach was conducted [158-160]. This analysis examined the number of publications, countries of publications, most active journals, institutions, authors, funding bodies, subject areas, citations of publications, and keywords of publications. Seven bibliometric metrics were computed, including the h-, g-, m-, and Y-indices; dominance factor; annual growth rate (AGR); and fractionalized frequency (Table S3 in Multimedia Appendix 1). The h-index attempts to measure both the productivity and citation impact of the published body of work of an entity (eg, author, institution, and journal) [161,162]. It refers to the total number of publications by a particular entity with at least the same number of citations. The m-index is calculated by dividing the h-index by the number of years of an entity's productive life (eg, researcher) [161]. The g-index of an entity corresponds to the largest number g such that the top g publications have at least $\geq g^2$ citations together [162]. The Y-index refers to the sum of both the total number of first-authored publications and the total number of corresponding author publications [163]. The dominance factor refers, for a particular researcher, to the proportion of multiauthored publications as a specific author's rank to the total number of multiauthored publications [164]. The fractionalized frequency intends to reflect an author's contribution. The AGR refers to the variable's change in

XSL•FO
RenderX

percentage as a year-over-year statistic [165]. The most up-to-date journals' impact factors and ranks were retrieved manually using the Journal Citation Report in April 2024.

Spearman correlations were calculated to examine the association between the number of publications and gross domestic product (GDP); population size [166]; and the total labor force, the number of researchers in research and development (per million people), fertilizer consumption (in both % of fertilizer production and kilograms per hectare of arable land), agricultural land (km^2), agricultural land (% of land area), land under cereal production (hectares), permanent cropland (% of land area), cereal production (metric tons), crop production index, food production index, livestock production index, cereal yield (kilogram per hectare), female individual employment in agriculture (% of female employment), male individual employment in agriculture (% of male employment), employment in agriculture (% of total employment); and agriculture, forestry, and fishing, value added (% of GDP). These country characteristics were obtained from the World Bank. The most recent country characteristic (eg, GDP) was used when available.

Research directions, including hot spots and gaps, were investigated with keyword frequency, co-occurrence (counting of paired keywords), and thematic mapping analyses. Thematic mapping and keyword co-occurrence network are 2 complementary but distinct approaches that serve different purposes and offer different insights. In summary, thematic mapping focuses more on the strategic positioning of research themes within a field, while keyword co-occurrence networks emphasize the relationships and connections between specific keywords in the literature [158,167]. Both methods complement each other and are usually used to provide a more comprehensive understanding of research landscapes. The co-occurrence of 2 keywords was defined by the frequency with which they appear together in publications and was quantified using association strength (AS) or equivalent index, calculated as , where c_{ij} is the number of publications in which keywords i and j co-occur, while c_i and c_j are the number of publications in which each keyword appears, respectively [158,167]. AS measures how close 2 keywords are to each other. An AS value of 1 indicates keywords always appear together, while 0 indicates they never co-occur. These keyword co-occurrences can be visualized using a co-occurrence network graph, where a vertex or node represents a keyword, the size of the node represents the keyword frequency, and the edge represents the association between 2 keywords [158,167]. On the basis of the keyword co-occurrence network graph, a community detection procedure can be used to identify groups of words highly associated with each other [158,167]. In other words, equivalent keywords based on AS can be grouped together to identify research themes [158,167]. A strategic diagram or thematic map is based on Callon centrality (x-axis) and Callon density (y-axis) [158,167]. Callon centrality measures the degree of interaction of a theme with other themes. It is defined as , where k is a keyword belonging to a theme and h is a keyword belonging to another theme [158,167]. Callon centrality can be interpreted as an indicator of the importance of a particular topic within the broader research landscape. Callon density measures the internal strength of a theme. It is defined as , where i and j are keywords belonging to the same theme and w is the total number of keywords in a theme [158,167]. Callon density serves as a metric for assessing the progression and maturation of that topic [158,167]. A strategic diagram is divided into 4 quadrants according to Callon centrality and density values, which correspond to 4 types of topics. Hot spots or hot topics are defined by both high density and high centrality values (upper-right quadrant), while basic topics are defined by high centrality but low density values (lower-right quadrant). Peripheral topics are defined by both low centrality and low density values (lower-left quadrant), while niche topics are defined by low centrality and high density values (upper-left quadrant) [158,167].

To focus on agricultural or farming exposome research, a bibliometric profile of the "farming exposome" was constructed, which restricts the exposome concept to environmental exposures specific to farming populations [152,168]. This bibliometric farming exposome picture examined co-occurrences between keywords related to potential risk factors and specific health events (eg, cancers and reproductive disorders).

The bibliometric analysis was conducted and reported according to the preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO; Table S4 in Multimedia Appendix 1) [169]. All analyses were performed using R software (version 4.3.2; R Foundation for Statistical Computing) for Windows 10 (Microsoft Corporation). The bibliometric analysis was performed using the *bibliometrix* R package (version 4.1.4) [170].

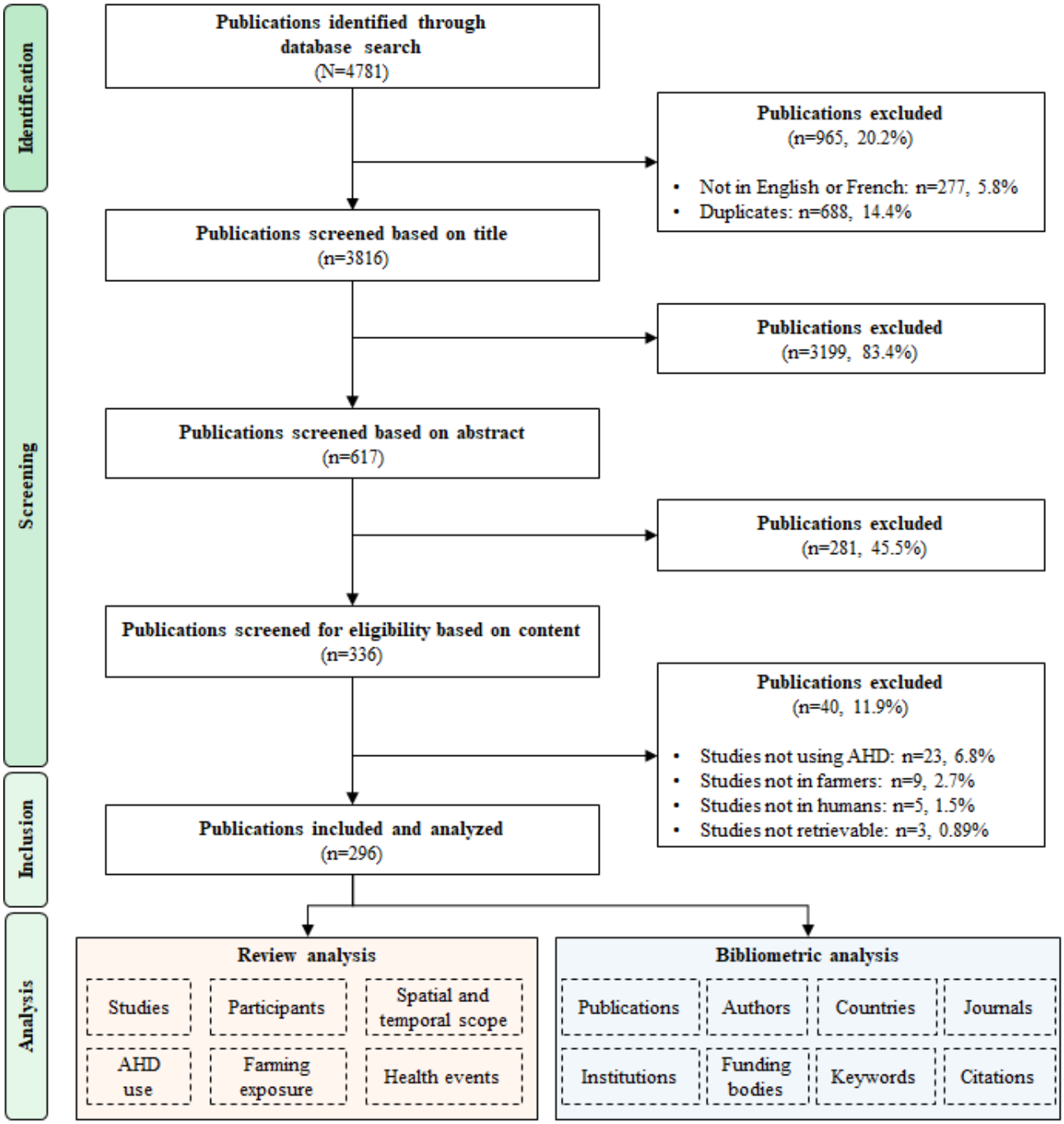
Results

Overview

After excluding 4485 irrelevant records, 296 publications were analyzed (Figure 1). The majority were articles (293/296, 98.9%), with a small number of reviews (2/296, 0.7%) and editorial materials (1/296, 0.3%; Table 3). Only one-third of the publications (107/296, 36.1%) were open access (Table 3 and Figure S1 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flowchart depicting the literature search and the evaluation process for finding relevant records. The search, conducted on April 15, 2024, in PubMed and Web of Science, had no date restrictions. AHD: administrative health database.



XSL•FO
RenderX

Table 3. Main characteristics of the included publications (N=296).

Description	Results
Timespan	1975 to 2024
Publication type, n (%)	
Article	293 (99)
Review	2 (0.7)
Editorial material	1 (0.3)
Open-access publications, n (%)	107 (36.1)
Document age (y), mean (SD)	14.2 (11.8)
Annual growth rate (%)	5.2
Publication citations	
Total, n[a]	9379
Average citations per publication	31.7
Average citations per year per publications	2.02
References, n	8814
Journals	
Total, n	118
Average number of publications per journal	1.86
Average number of citations per journal	79.5
Authors	
Total, n	1225
Single-author publications, n	4
Author appearances, n	1882
Average number of coauthors per publication	6.36
Average number of publications per author	0.24
International coauthorships (%)	24.3
Author's keywords, n	576
Author's country	
Total, n	34
Average number of publications per country	2.86
Average number of citations per country	436.0
Author's institution	
Total, n	338
Average number of publications per institution	3.11
Average number of citations per institution	101.3
Author's funding body	
Total, n	181
Average number of publications per funding body	2.48
Average number of citations per funding body	77.7

[a]Total, n indicates that the respective parameter has been cited n number of times, as in 296 publications have been cited 9379 times.

The average publication age was 14.2 (SD 11.8) years, ranging from the oldest in 1975 [171] to the most recent in April 2024 [152]. From 1975 onward, there has been a steady increase in publications using AHDs to address health issues in farming populations, with an AGR of 5.2%. Notably, almost one-third of these articles (91/296, 30.7%) were published in the last 5 years, highlighting the rising interest in AHD-based public health research in this population (Figure S2 in Multimedia

XSL•FO
RenderX

Appendix 1). Collectively, the publications received 9379 citations, averaging 31.7 citations per publication. Figure S3 in Multimedia Appendix 1 presents the historical direct citation network. The body of work involved 1225 authors from 338 institutions, with 1882 author appearances and an average of 6 authors per paper (Table 3). Four (1.4%) out of the 296 publications were single-author publications. On average, each paper cited 30 references.

Studies were led by authors from 34 countries, predominantly high-income nations, with 24.3% (72/296) of studies involving multicountry collaborations (Figure S4 in Multimedia Appendix 1). US-based authors contributed the most publications (91/296, 30.7%), followed by authors based in France (71/296, 24%) and Finland (35/296, 11.8%). US authors also had the most citations (3495/9379, 37.2%), with France and Finland ranking second and third, respectively.

Of 296 publications, the 25 (8.4%) most cited ones, appearing in 17 different journals, received between 83 (83/9379, 0.9%) and 485 (485/9379, 5.2%) citations (Table S5 in Multimedia Appendix 1) [150,172-196]. Of these 25 publications, 10 (40%) were published before 2000, another 10 (40%) between 2000 and 2010, and 5 (20%) after 2010. Most of these studies focused on cancer risk (16/25, 64%), while others investigated neurodegenerative disorders (5/25, 20%); respiratory conditions (2/25, 8%); and multiple health outcomes, such as sleep disorders, mental health disorders, and musculoskeletal disorders (2/25, 8%).

Table S6 in Multimedia Appendix 1 provides details on the most productive countries, prolific authors, active journals, institutions, and funding bodies.

Study Characteristics

Table 4 provides an overview of the included publications. Longitudinal study designs were the most common, including retrospective cohorts (129/296, 43.6%) and prospective cohorts (56/296, 18.9%). Case-control studies (62/296, 20.9%), cross-sectional studies (39/296, 13.2%), and ecological studies (17/296, 5.7%) were less common (Multimedia Appendix 2). A few studies (10/296, 3.4%) used multiple study designs [188,194,197-204].

The median follow-up period was 9.5 (IQR 5-17) years. On average, there was a 7-year gap (90% CI 3-14) between the most recent data used and the year of publication, with considerable variation depending on the publication year (Figure 2). The oldest data were from 1801 [205], and the most recent data were from 2022 [206]. Notably, one-third of the data used (98/296, 33.1%) were from before 2000, while nearly three-quarters (214/296, 72.3%) were from before 2015 (Figure S2 in Multimedia Appendix 1). Of 296 studies, only 10 (3.4%) used data from the last 5 years (from 2020), while 80 (27%) used data from the last 10 years (from 2015).

Studies were conducted in all continents, but most participants were from Europe (249/296, 84.1%), followed by North America (85/296, 28.7%), Asia (24/296, 8.1%), Oceania (17/296, 5.7%), Africa (4/296, 1.4%), and Central and South America (4/296, 1.4%). France (70/296, 23.6%) and the United States (67/296, 22.6%) were the most represented countries, followed by Finland (36/296, 12.2%), Sweden (32/296, 10.8%), Denmark (28/296, 9.5%), and Norway (26/296, 8.8%; Figure 3 and Figure S5 in Multimedia Appendix 1). Most studies had a regional or local scope (177/296, 59.8%), in particular, traditional epidemiological studies, such as Agriculture and Cancer (AGRICAN) [207] and Agricultural Health Study (AHS) [172], which used AHDs to either identify potential individuals for inclusion or enrich their cohorts.

XSL•FO
RenderX

Table 4. Characteristics of the included studies (1975 to 2024; N=296).

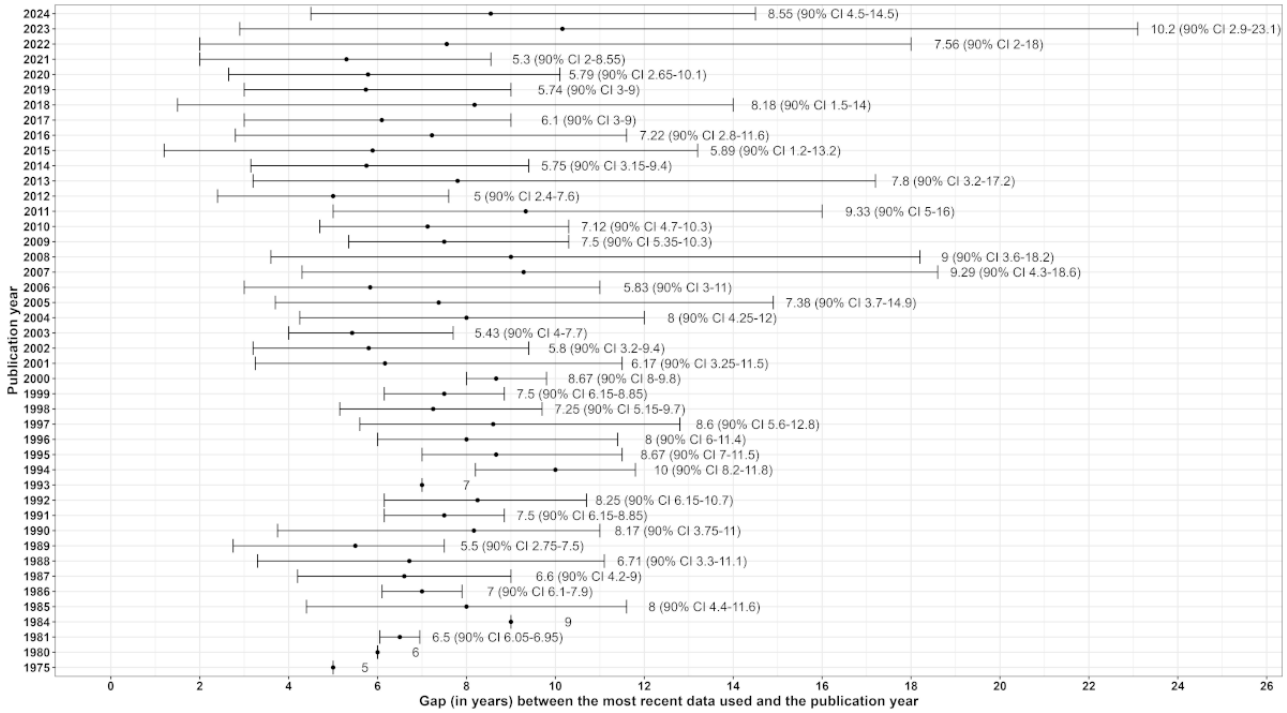
Characteristic	Values
Research goal, n (%)	
Study the association between farming and a health event	156 (52.7)
Study the association between individual characteristics and a health event	131 (44.3)
Other research goals	9 (3)
Study design, n (%)	
Retrospective cohort	129 (43.6)
Case-control study	62 (20.9)
Prospective cohort	56 (18.9)
Cross-sectional study	39 (13.2)
Ecological study	17 (5.7)
Multiple designs	10 (3.4)
Review	2 (0.7)
Perspective	1 (0.3)
Geographic scope, n (%)	
Nationwide	117 (39.5)
Regional or local	176 (59.5)
Temporal scope (y)	
Follow-up period, median (IQR)	9.50 (5-17)
Follow-up period, mean (SD)	12.8 (14.0)
Gap between the latest data used and publication year, median (IQR)	7.21 (5-9)
Gap between the latest data used and publication year, mean (SD)	7.21 (4.67)
Population, n (%)	
Adult	265 (89.5)
Adult and child	19 (6.4)
Child	8 (2.7)
Not reported	1 (0.3)
Sex, n (%)	
Female	130 (43.9)
Male	169 (57.1)
Female and male	188 (63.5)
Not specified	108 (36.5)
Participants, n (%)	
>1,000,000	50 (16.9)
100,001 to 1,000,000	53 (17.9)
10,001 to 100,000	65 (22)
1001 to 10,000	67 (22.6)
101 to 1000	47 (15.9)
10 to 100	8 (2.7)
Not reported	3 (1)
AHD [a] type, n (%)	
Disease register	158 (53.4)
Electronic health or medical record	124 (41.9)

XSL•FO
RenderX

Characteristic	Values
Insurance claim	106 (35.8)
Population register	95 (32.1)
Hospital discharge databases	41 (13.9)
AHD use, n (%)	
Obtain information on sociodemographics	272 (91.9)
Obtain information on a health event	269 (90.9)
Identify a farmer	147 (49.7)
Identify an individual	140 (47.3)
Obtain information on occupations	117 (39.5)
Exposure assessment	57 (19.3)
Obtain information on a farming activity	43 (14.5)
Other uses	14 (4.7)

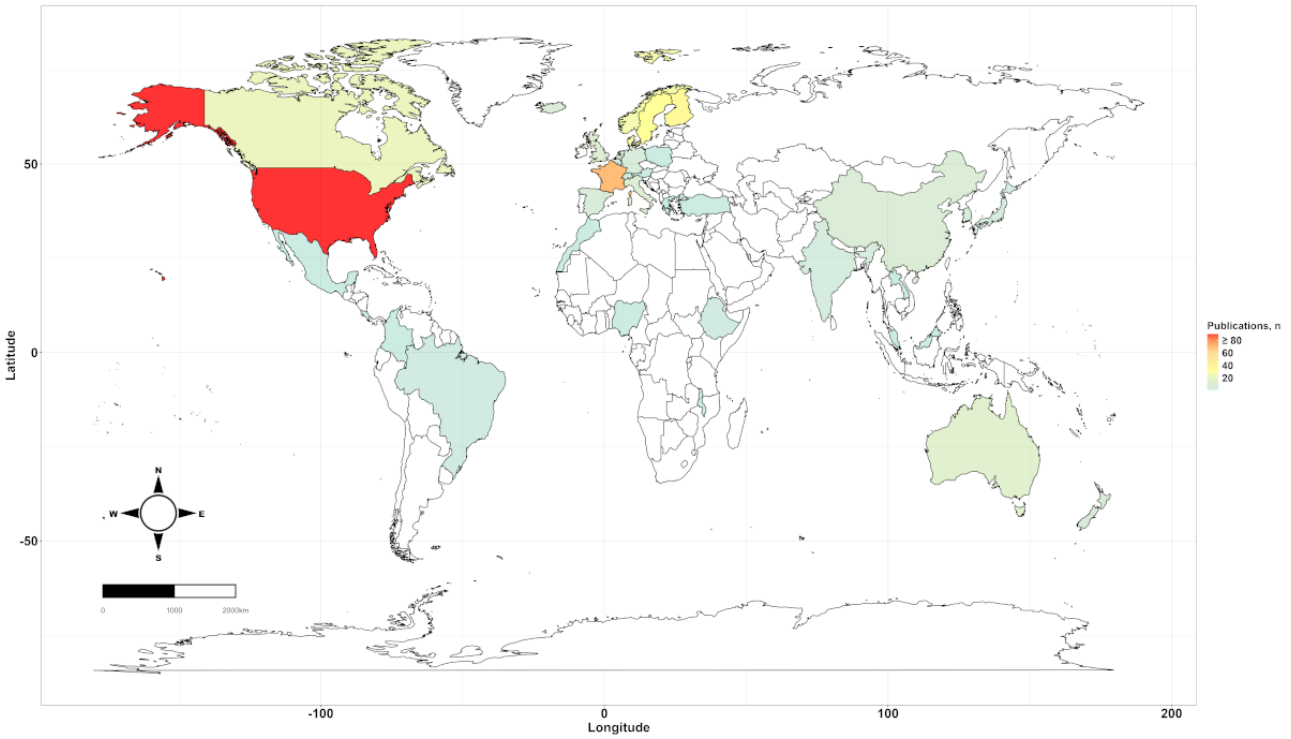
[a]AHD: administrative health database.

Figure 2. Number of years between the most recent data used and publication for all included articles (1975-2024). Points refer to the average number of years or gap between the most recent data used and publication (x-axis) for each publication year (y-axis). Error bars refer to the 90% CI of the number of years between the most recent data used and publication.



XSL•FO
RenderX

Figure 3. World map of the number of publications per country of the farming population studied between 1975 and 2024.



Most studies included 1001 to 10,000 participants (67/296, 22.6%), followed by studies with 10,001 to 100,000 participants (65/296, 22%) and 100,001 to 1 million participants (53/296, 17.9%; Table 4). Larger studies (>1 million participants) accounted for 16.9% (50/296) of the included publications. Smaller studies, with 100 to 1000 participants, were less common (47/296, 15.9%), and very few (8/296, 2.7%) had <100 participants. Most studies included adult participants (284/296, 95.9%). Of 296 studies, 169 (57.1%) examined male individuals, 130 (43.9%) examined female individuals, and 188 (63.5%) examined both sexes, but 108 (36.5%) did not specify the participants' sex.

More than half of the studies (156/296, 52.7%) aimed to explore the relationship between farming activities (eg, dairy farming) and health events, while 131 studies (44.3%) focused on individual characteristics, such as occupation, age, sex, and socioeconomic status (Figure S6 in Multimedia Appendix 1). Among those studies examining individual characteristics, farming was often considered broadly and compared to other occupations (95/131, 72.5%). Conversely, in studies investigating health outcomes specifically related to farming activities, agriculture was treated as a broad category in only 27.6% (43/156) of the cases. Most studies (277/296, 93.6%) used the general population or other nonfarming groups as the reference category without differentiating farmers by job role (eg, farm managers vs farm workers). Descriptive statistics and multivariable regression were the most commonly used methods. Notably, only 2 studies (2/296, 0.7%) incorporated artificial intelligence (AI) in their analysis [208,209].

Few studies investigated health outcomes in farmers' family members or nonfarmers exposed to farming. Of 296 studies, only 3 (1%) focused on health events in farmers' partners [177,210,211], 5 (1.7%) on farmers' children [179,212-215], and 6 (2%) on nonfarmers exposed to farming-related risks [209,216-221]. There were 11 (3.7%) studies that explored health risks in migrant workers.

Some publications reported findings from the same cohorts (Figure S7 in Multimedia Appendix 1). The 10 most prolific cohorts included France-based AGRICAN (18/296, 6.1%) [207], the US-based AHS (17/296, 5.7%) [172], Nordic Occupational Cancer Study (NOCCA; 12/296, 4.1%) from Nordic countries (Finland, Denmark, Norway, Sweden, and Iceland) [189], France-based Tracking and Monitoring Occupational Risks in Agriculture (TRACTOR; 7/296, 2.4%) [222], and Cancer in the Norwegian Agricultural Population (7/296, 2.4%) [182] cohorts. Other notable cohorts included the US-based United Farm Workers of America (6/296, 2.0%) [223], France-based BALISTIC (5/296, 1.7%) [224], the international (29 countries) consortium agricultural cohort (AGRICOH; 4/296, 1.4%) [150,225], AIRBAg (4/296, 1.4%) from France [226], and the US-based National Agricultural Workers Survey (3/296, 1%) [227]. Among these top 10 cohorts, only NOCCA, United Farm Workers of America, and TRACTOR exclusively used AHDs.

AHD Use

There was high heterogeneity in the coding systems used and the granularity of the information available regarding health events (outcomes), population, and exposure determinants, depending on the AHD and study considered. Regardless of the publication reviewed, AHDs and other datasets were never reported as adhering to the findable, accessible, interoperable, and reusable (FAIR) data principles [228-230]. In addition, none of them could be considered as FAIR data because, with a few exceptions [222], most AHDs were not precisely described, and data availability statements were rare. Furthermore, mainly due to privacy concerns, AHDs were not available for open and free access.

XSL•FO
RenderX

The most commonly used AHDs were disease registers, used in more than half of the studies (158/296, 53.4%), followed by electronic health or medical records (124/296, 41.9%), insurance claims (106/296, 35.8%), population registers (95/296, 32.1%), and hospital discharge databases (41/296, 13.9%; Table 4). Among disease registers, cancer (120/158, 75.9) and mortality registers (75/158, 47.5%) were the most frequently used (Figure S8 in Multimedia Appendix 1 and Multimedia Appendix 2). Nearly one-third of the studies (91/296, 30.7%) relied on a single AHD, with disease registers being the most common (38/91, 42%), followed by insurance claims (29/91, 32%), electronic health or medical records (18/91, 20%), population registers (5/91, 5%), and hospital discharge databases (1/91, 1%). Other types of digital data were used less frequently, including pesticide registration records (13/296, 4.4%), job-exposure matrices (JEMs; 12/296, 4.1%), crop-exposure matrices (11/296, 3.7%), pesticide use records (8/296, 2.7%), climate data (7/296, 2.4%), and air quality data (2/296, 0.7%). While contextual data were sometimes used (9/296, 3.0%), person-generated data, smart agriculture data, and omics were never used.

The AHDs and other digital data were primarily used to obtain sociodemographic information (272/296, 91.9%) and health event data (269/296, 90.9%). They were also used to identify farmers (147/296, 49.7%) or individuals (140/296, 47.3%), gather occupational information (117/296, 39.5%), assess exposure (57/296, 19.3%), obtain data on farming activities (43/296, 14.5%), or track climate conditions (7/296, 2.4%).

Nearly two-thirds of the studies (181/296, 61.1%) relied exclusively on digital data (AHDs or other), while more than one-third (112/296, 37.8%) incorporated self-reported information/active data (requiring active participant involvement) as part of epidemiological cohorts. A total of 111 (37.5%) out of 296 studies used participant-completed questionnaires (paper or electronic) to gather sociodemographic data and confounding factors (98/296, 33.1%), assess exposure (96/296, 32.4%), or collect health information (83/296, 28%). Some information was obtained through interviews (44/296, 14.9%) or clinical examinations (32/296, 10.8%). Biological monitoring (24/296, 8.1%) and airborne monitoring (2/296, 0.7%) were sometimes used, whereas no study reported dermal monitoring (Figure S9 in Multimedia Appendix 1).

Among all the AHDs used, the Mutualité Sociale Agricole (MSA) is a singularity. To the best of our knowledge, it is the only AHD specifically dedicated to the entire farming population of a country. Indeed, MSA is the French national insurance scheme that covers the entire farming workforce (5% of the overall French population) [115,128]. MSA was used in 60 studies (60/296, 20.3%). These studies were often part of cohorts with multiple publications, such as AGRICAN (18/60, 30%), TRACTOR (7/60, 12%), BALISTIC (5/60, 8%), AIRBAg (4/60, 7%), Aging Multidisciplinary Investigation (2/60, 3%) [151], BM3R (2/60, 3%) [231], FERMA (risk factors of the rural environment and allergic and respiratory disease; 1/60, 2%) [232], and Phytoner (1/60, 2%) [233]. Of these, TRACTOR was the only cohort using exclusively MSA data [222].

Farming Exposure

A variety of exposure proxies were used to assess farming-related exposure. The most common proxy was a job title, which generally referred to whether the individual was a farmer (184/296, 62.2%). Other proxies included specific farming activities, such as dairy or crop farming (111/296, 37.5%), general pesticide exposure (yes or no; 62/296, 20.9%), and exposure to specific pesticide compounds (eg, glyphosate or paraquat) or pesticide classes (eg, insecticides; 51/296, 17.2%; Figure S10 in Multimedia Appendix 1 and Multimedia Appendix 2). The number of farming activities studied ranged from just 1 [226] to 78 [222], with an average of 8 farming activities per study. Similarly, the number of pesticide compounds assessed ranged from 1 [234] to 943 [235], with an average of 42 pesticides per study. Only 1 study investigated the mixture effect of exposure to multiple pesticide combinations on human health [236]. Investigations into other chemical exposures were also limited, with only 2 papers each addressing silica exposure [237,238] and air pollution [194,217] (2/296, 0.7%). Notably, no studies examined exposure to per- and polyfluoroalkyl substances or micro- and nanoplastics. Research on the broader farming exposome was rare (5/296, 1.7%) and typically used farming activities as proxies [152].

Of 296 studies, few explored exposure to physical agents, with 5 studies (1.7%) focusing on radiation [187,218,239-241]. No studies investigated the effects of climate change on farmers' health. Exposure to biological agents was rarely studied as well, with just 3 (1%) out of 296 papers addressing mycotoxins [241-243]. Finally, only 3 studies (1%) examined psychological factors related to farming exposure [244-246].

Health Events

The most frequently studied health events were cancer (142/296, 48%), followed by mortality (44/296, 14.9%), injuries (38/296, 12.8%), workplace accidents (32/296, 10.8%), respiratory disorders (30/296, 10.1%), neurodegenerative diseases (28/296, 9.5%), and mental health issues (26/296, 8.8%; Figure S11 in Multimedia Appendix 1 and Multimedia Appendix 2). Less frequently studied conditions included cardiovascular diseases (16/296, 5.4%), autoimmune disorders (11/296, 3.7%), musculoskeletal disorders (11/296, 3.7%), reproductive disorders (3/296, 1.0%), sleep disorders (1/296, 0.3%), and frailty (1/296, 0.3%). Notably, no studies explored the farming microbiome.

Among cancers, lung cancer was the most commonly investigated cancer (43/142, 30.3%), followed by prostate cancer (38/142, 26.8%), leukemia (37/142, 26.1%), colorectal cancer (35/142, 34.6%), multiple myeloma (35/142, 34.6%), non-Hodgkin lymphoma (35/142, 34.6%), bladder cancer (31/142, 21.8%), and brain cancer (31/142, 21.8%; Figure S12 in Multimedia Appendix 1). Respiratory disorders were primarily focused on asthma (15/30, 50%) and COPD (chronic obstructive pulmonary disease; 14/30, 47%). Parkinson disease was the most studied neurodegenerative condition (16/28, 57%), followed by multiple sclerosis (6/28, 21%). Fewer publications examined Alzheimer disease (2/28, 7%) and amyotrophic lateral sclerosis (2/28, 7%; Figure S13 in Multimedia Appendix 1). In the mental health field, suicide (12/26, 46%) and depression

XSL•FO
RenderX

(8/26, 31%) were the most investigated issues (Figure S14 in Multimedia Appendix 1).

Keyword Analysis

Overview

Following an initial extraction of 1259 authors' keywords, manual harmonization was performed. Duplicate keywords were removed through singular or plural standardization (130/1259, 10.3%) and synonym unification and grouping of cancer-related terms (553/1259, 43.9%), yielding a final set of 576 (45.8%) harmonized keywords, which were all used in subsequent analyses to prevent selection bias.

On average, each publication included 8.90 keywords (90% CI 0-17), although 35 (11.8%) out of 296 publications lacked any keywords, in line with the journal guidelines. Keyword analysis confirmed prior findings regarding farming exposure and health outcomes. It also provided deeper insights into emerging research hot spots, directions, and gaps.

Of the total 576 keywords, 301 (52.3%) appeared only once, while 68 (11.8%) were mentioned at least 10 times. More frequently used keywords included 39 that appeared at least 20 times (39/576, 6.8%) and 11 that featured in at least 50 publications (11/576, 1.9%). The 50 most frequently used keywords were mentioned in at least 17 (5.7%) out of 296 publications, while the top 10 appeared in at least 51 publications (17.2%; Table S7 in Multimedia Appendix 1). The most frequently cited keyword was "cancer" (150/296, 50.7%), followed by "mortality" (96/296, 32.4%), "pesticide" (88/296, 29.7%), "occupation" (82/296, 27.7%), "farmer" (77/296, 26.0%), "agriculture" (74/296, 25%), "exposure" (57/296, 19.3%), and "epidemiology" (57/296, 19.3%).

In terms of overall citations, "cancer" (5766/9379, 61.5%), "pesticide" (3569/9379, 38.1%), and "mortality" (3097/9379, 33%) were the most cited keywords. During the past decade, the frequency of the top 5 keywords has drastically increased (Figure S15 in Multimedia Appendix 1). Notably, keywords such as "cancer," "mortality," "occupation," "pesticide," "agriculture," and "farmer" have been consistently present in publications spanning at least 30 years (not necessarily consecutively; Figure S16 in Multimedia Appendix 1). In the last decade, emerging keywords, such as "big data," "administrative health database," "dust," and "BMI," have gained prominence (Figure S17 in Multimedia Appendix 1).

Keyword Co-Occurrence

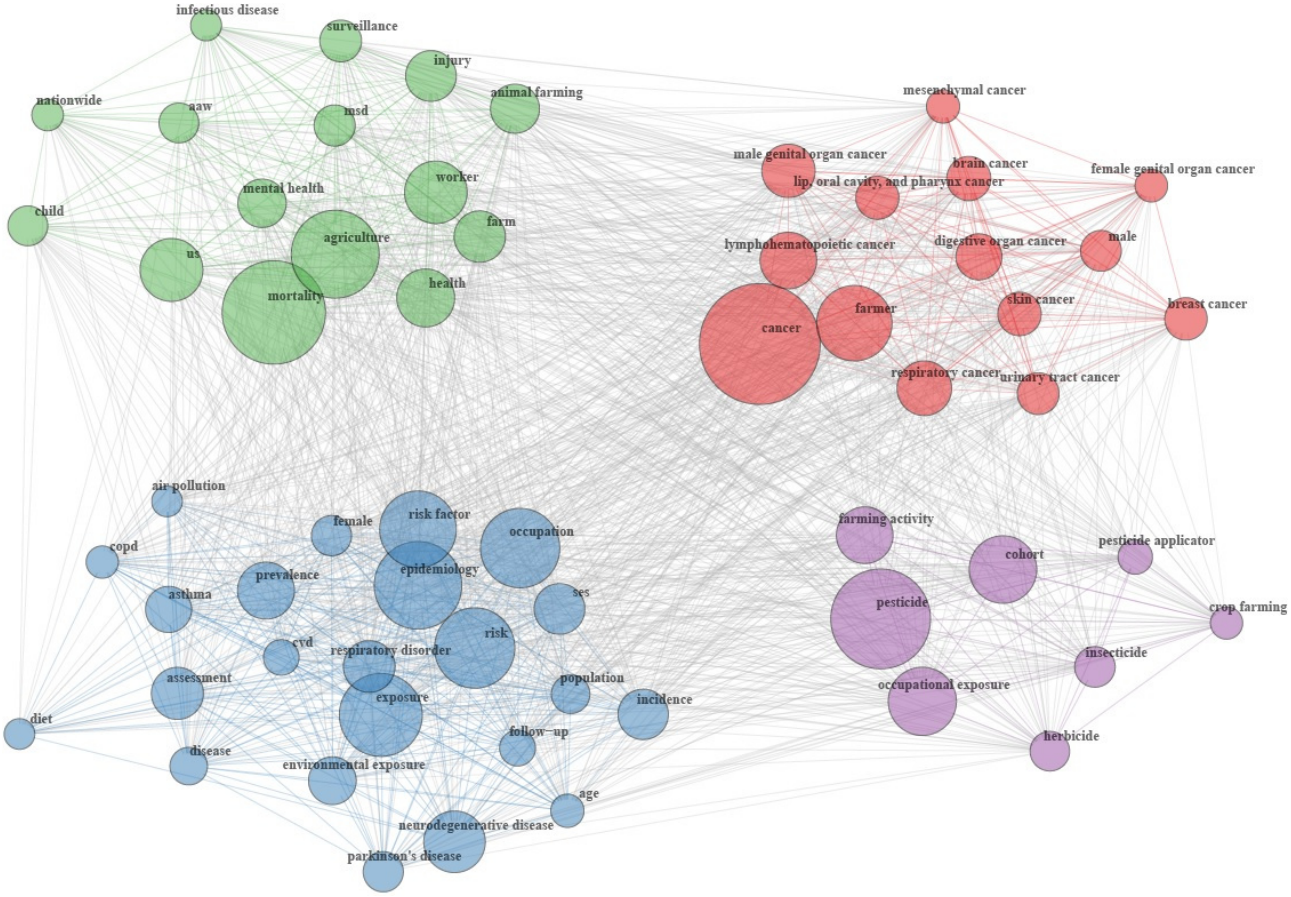
A keyword co-occurrence network illustrating the frequency of keyword co-mentions in publications was constructed, thereby revealing relationships and conceptual connections (Figure 4). In this network, nodes or vertices represent keywords, with their sizes indicating frequency, while edges denote co-occurrences. The network's density and arrangement reveal topic interconnectivity, with larger vertices representing more frequently mentioned keywords. The network visualization helps identify clusters of related topics and highlights core research areas.

Using a community detection algorithm (spin-glass model with simulated annealing), 4 distinct clusters or communities of keywords were identified. Each cluster groups keywords that are often mentioned together, with stronger internal associations and weaker connections to keywords in other clusters.

The most frequently used keywords for each cluster were "cancer" (red cluster), "pesticide" (purple cluster), "mortality" (green cluster), and "exposure" (blue cluster). The red cluster highlights associations between various types of cancer, reflecting the fact that studies investigating cancer risks often examine multiple types of cancer. The green cluster links "mortality" with terms such as "mental health," "injury," and "animal farming," explained by the association between workplace accidents, mental health issues (eg, suicide), animal farming, and mortality. In the purple cluster, "pesticide" connects with "occupational exposure" and "farming activity," emphasizing that pesticide exposure is primarily studied in occupational settings across different types of farming. The blue cluster connects "exposure" to terms such as "neurodegenerative disease," "respiratory disorder," "cardiovascular disorder," "risk factor," "air pollution," "age," and "diet," indicating the study of various risk factors in relation to several health events. These clusters highlight current research hot spots that focus on 4 main interconnected themes: the associations between risk factors, pesticide exposure, farming activities, and a range of diseases.

XSL•FO
RenderX

Figure 4. Keyword co-occurrence network of the 296 articles published between 1975 and 2024. Each vertex or node represents a keyword, while edges represent the co-occurrence between keywords. Two keywords are connected when they co-occur in the same publication, and the size of each vertex indicates the frequency of a keyword: larger vertices represent more frequently mentioned keywords. Keywords with the same color (cluster) represent a research area. AAW: workplace accident; COPD: chronic obstructive pulmonary disease; CVD: cardiovascular disorder; MSD: musculoskeletal disorder; SES: socioeconomic status.



Thematic Mapping: Research Hot Spots

Figure 5 presents a thematic map that illustrates current research directions. Thematic mapping visualizes the relationship between research themes or topics, enabling the identification of directions, emerging areas, and gaps in the literature. The result is a strategic diagram that shows how themes relate to each other and their relevance within a specific field. The graph is divided into 4 quadrants, categorizing topics based on their relevance (x-axis, Callon centrality) and maturity (y-axis, Callon density) within the broader research landscape. Each circle represents a theme or topic (ie, a cluster of equivalent keywords), with the circle size corresponding to the frequency of the keywords associated with that theme.

The upper-right quadrant represents “hot topics,” which are both highly relevant and mature in the research landscape. Four key hot topics drive AHD-based public health research in farming populations. These include 1 topic focused on cancer research; another on respiratory disorders; and a third encompassing neurodegenerative diseases, workplace accidents, injuries, and mental health issues. The final hot topic involves large-scale studies in France and Europe using big data and insurance claims.

The lower-right quadrant contains “basic topics,” which are relevant but not yet mature in the research landscape. Only 1 such theme emerged: research related to pesticide exposure, mortality, and farming.

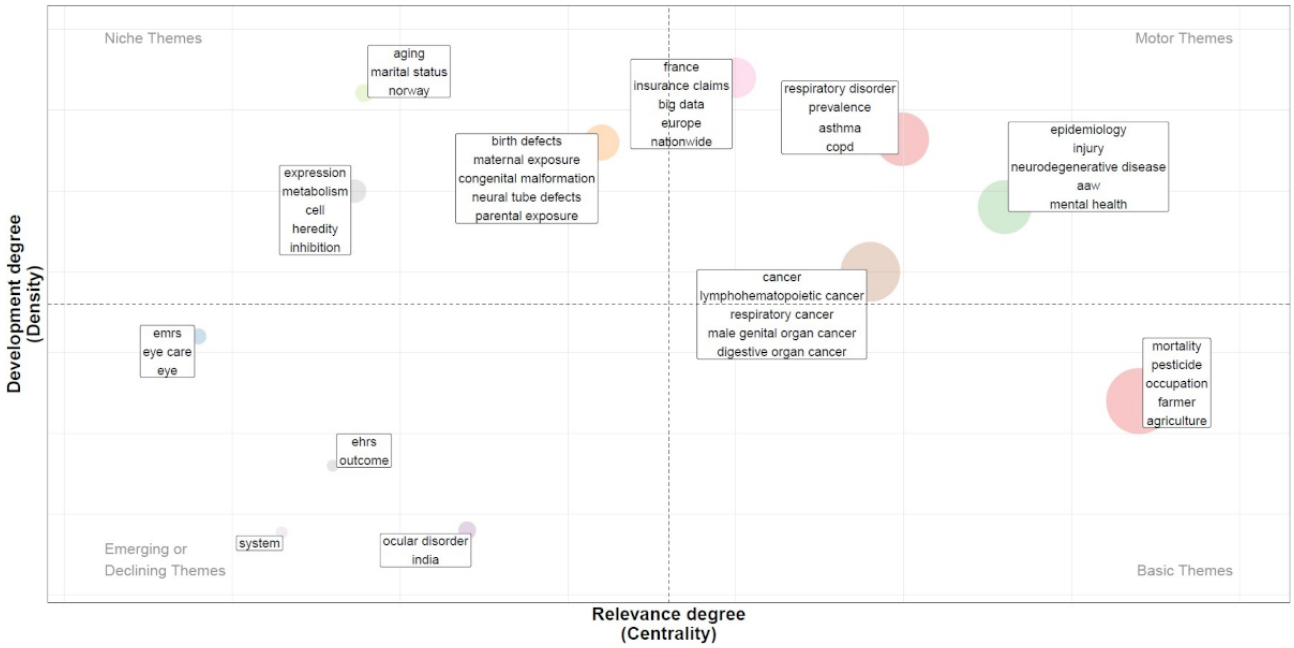
In the upper-left quadrant, “niche themes” refer to mature research topics that have not yet achieved full relevance. Three niche themes were identified: the first involves studies examining aging and research conducted in Norway; the second focuses on reproductive disorders and parental exposure, a theme poised to potentially evolve into a hot topic; and the final niche theme covers genetics and metabolism.

Finally, the lower-left quadrant contains “peripheral topics,” which represent either emerging or declining themes with low relevance and maturity. Four peripheral topics were observed, of which 2 (50%) were primarily centered on research on ocular disorders, 1 (25%) on the use of electronic health or medical records, and 1 (25%) on studies conducted in India.

This thematic map helps highlight both well-established and emerging areas of research, as well as gaps that may be ripe for future investigation.

XSL•FO
RenderX

Figure 5. Thematic mapping: research hot spots based on keywords from the 296 articles published between 1975 and 2024. The graph shows how themes relate to each other and their relevance within a specific field. This graph is divided into 4 quadrants, categorizing topics based on their relevance (x-axis and Callon centrality) and maturity (y-axis and Callon density) within the broader research landscape. Each circle represents a theme or topic (ie, a cluster of equivalent keywords), with the circle size corresponding to the frequency of the keywords associated with that theme. AAW: accident at work; COPD: chronic obstructive pulmonary disease; EHR: electronic health record; EMR: electronic medical record.



Bibliometric Farming Exposome

To identify research directions and gaps in the farming exposome literature, a bibliometric keyword co-occurrence analysis was conducted to explore the farming exposome by examining the co-occurrence between keywords associated with potential risk factors and specific health outcomes. This analysis was restricted to exposome-related and health event–related keywords. Of 576 keywords, 130 (22.6%) were related to the exposome, among which 93 (16.1%) were related to the specific external exposome (eg, pesticide), 19 (3.3%) to the general external exposome (eg, climate), and 18 (3.1%) to the internal exposome (eg, oxidative stress). Furthermore, there were 70 (12.2%) health event–related keywords (eg, brain cancer).

The results of this analysis are provided in Tables 5 and 6 and Multimedia Appendix 3, with each cell representing the percentage of occurrences of an exposome-related keyword (eg, air pollution) in all publications mentioning a specific health event–related keyword (eg, Alzheimer disease). For example, a value of 33.3 indicates that an exposome-related keyword appeared in 33.3% of all publications mentioning a specified health event–related keyword. To facilitate interpretation and ease the reading of Tables 5 and 6, exposome-related keywords were categorized into 19 groups (eg, chemical agent) and health event–related keywords into 20 groups.

XSL•FO
RenderX

Table 5. Co-occurrence between keywords related to internal exposomes and health event categories among the articles published between 1975 and 2024. Each cell refers to the number of times (%) a keyword related to an exposome category (eg, chemical agent) was mentioned among all publications in which a keyword related to a health event category (eg, cancer) appeared (N=296). Please note that the absolute value for each row is provided in parentheses with the row header and remains the same for all the parameters in that row.

Health event, n	Internal exposome (%)										
	Age	Sex	BMI	BP[a]	Heredity	Ethnicity	Hormone	Menopause	Metabolism	OS[b]	Inflammation
Cardiovascular disease (n=17)	5.88	11.8	5.88	5.88	0	5.88	0	0	0	0	0
Work-related disease (n=5)	0	20	0	0	0	0	0	0	0	0	0
Autoimmune disease (IBD[c], RA[d], vasculitis, and NR[e]; n=9)	0	0	11.1	0	0	0	0	0	0	0	0
Cancer (n=150)	3.62	27.5	2.17	0	0.73	1.45	3.62	0.73	2.9	0.73	0
Dental health (n=2)	0	0	0	0	0	0	0	0	0	0	0
Ocular disorder (n=6)	16.7	0	0	0	0	0	0	0	0	0	0
Frailty (n=2)	50	0	50	0	0	0	0	0	0	0	0
Anemia (n=1)	100	100	0	0	0	0	0	0	100	0	0
Infectious disease (malaria, Lyme disease, tuberculosis, toxoplasmosis, and NR; n=14)	0	7.14	0	0	0	7.14	0	0	0	0	0
Injury (including workplace accident and disability; n=40)	4.26	8.51	2.13	0	0	2.13	0	2.13	0	0	0
Chronic kidney disease (n=3)	0	0	0	0	0	0	0	0	0	0	0
Mental health disorder (depression, suicide, and NR; n=25)	9.52	4.76	4.76	0	0	0	0	0	0	0	0
Metabolic disorder (diabetes, dysthyroidism, and NR; n=9)	11.1	0	11.1	11.1	0	11.1	0	0	0	0	0
Mortality (n=75)	4.17	17.7	0	0	1.04	0	1.04	2.08	0	0	0
Musculoskeletal disorder (arthritis, low-back pain, and NR; n=14)	7.14	14.3	7.14	0	0	0	0	0	0	0	0
Neurodegenerative disease (AD[f], ALS[g], MND[h], MS[i], PD[j], and NR; n=33)	6.06	0	0	0	3.03	6.06	0	0	12.1	6.06	0
Sensory impairment (n=1)	0	0	0	0	0	0	0	0	0	0	0
Reproductive disorder (birth defects, infertility, spontaneous abortion, and NR; n=24)	0	28.6	14.3	0	0	7.14	7.14	0	7.14	0	0
Respiratory disorder (allergy, asthma, COPD[k], pneumonia, sarcoidosis, and NR; n=39)	5.88	11.8	8.82	2.94	0	5.88	0	0	5.88	0	2.94
Skin disorder (dermatitis and NR; n=2)	0	50	0	0	0	0	0	0	0	0	0

[a]BP: blood pressure.

[b]OS: oxidative stress.

XSL•FO
RenderX

[c]IBD: inflammatory bowel disease.
[d]RA: rheumatoid arthritis.
[e]NR: not reported.
[f]AD: Alzheimer disease.
[g]ALS: amyotrophic lateral sclerosis.
[h]MND: motor neuron disease
[i]MS: multiple sclerosis.
[j]PD: Parkinson disease.
[k]COPD: chronic obstructive pulmonary disease.

Distinct keyword exposome profiles were developed for each health event–related keyword (Figures S18-S43 in Multimedia Appendix 1), as illustrated in Figure 6 for mental health disorders. Most exposome-related keywords associated with keywords related to mental health disorders pertained to the type of occupations as well as chemical, lifestyle, socioeconomic, and psychological factors. Cancer-related keywords were associated mostly with keywords related to the internal (sex) and specific external exposome (chemical agents, lifestyle, and type of occupations). Autoimmune disease–related keywords co-occurred mostly with external exposome–related keywords (chemical agents, lifestyle, type of occupations, and socioeconomic factors). Neurodegenerative disease–related keywords were associated mostly with keywords related to the specific external exposome (lifestyle, chemical agents, and type of occupations). Reproductive disorders co-occurred mostly with internal (sex and BMI) and specific external exposome–related keywords (chemical agents and type of occupations). Keywords related to both musculoskeletal disorder and injury were associated with keywords from all 3 exposome components, in particular sex, type of occupations, lifestyle, biomechanical factors, chemical agents, and psychological factors. Infectious disease–related keywords co-occurred with specific external exposome–related keywords (biological agents and type of occupations), while respiratory disorder–related keywords were associated mostly with internal (sex) and specific external exposome–related keywords (lifestyle, chemical and biological agents, and type of occupations). Cardiovascular disorder–related keywords were associated with keywords from all 3 exposome components, in particular the sex, type of occupations, lifestyle, chemical agents, and socioeconomic and psychological factors.

XSL•FO
RenderX

Table 6. Co-occurrence between keywords related to specific external and general external exposomes and health event categories among the articles published between 1975 and 2024. Each cell refers to the number of times (%) a keyword related to an exposome category (eg, chemical agent) was mentioned among all publications in which a keyword related to a health event category (eg, cancer) appeared (N=296). Please note that the absolute value for each row is provided in parentheses with the row header and remains the same for all the parameters in that row.

Health event, n	Specific external exposome (%)					General external exposome (%)		
	Lifestyle	CA[a]	BA[b]	BF[c]	Occupation	PA[d]	SF[e]	PF[f]
Cardiovascular disease (n=17)	11.8	29.4	0	11.8	23.5	0	29.4	17.6
Work-related disease (n=5)	0	20	20	20	40	0	20	0
Autoimmune disease (IBD[g], RA[h], vasculitis, and NR[i]; n=9)	22.2	44.4	0	0	55.6	0	22.2	0
Cancer (n=150)	17.4	51.4	7.25	1.45	58.7	5.8	8.7	0
Dental health (n=2)	50	0	0	0	0	0	50	0
Ocular disorder (n=6)	16.7	16.7	0	0	0	0	0	16.7
Frailty (n=2)	100	50	0	0	0	0	0	50
Anemia (n=1)	0	0	0	0	0	0	0	0
Infectious disease (malaria, Lyme disease, tuberculosis, toxoplasmosis, and NR; n=14)	21.4	21.4	50	7.14	28.6	0	14.3	7.14
Injury (including workplace accident and disability; n=40)	12.8	14.9	2.13	14.9	23.4	2.13	4.26	6.38
Chronic kidney disease (n=3)	0	100	66.7	0	0	33.3	0	0
Mental health disorder (depression, suicide, and NR; n=25)	28.6	23.8	0	4.76	33.3	4.76	23.8	28.6
Metabolic disorder (diabetes, dysthyroidism, and NR; n=9)	33.3	44.4	0	11.1	44.4	11.1	0	22.2
Mortality (n=75)	17.7	55.2	8.33	4.17	54.2	5.21	17.7	2.08
Musculoskeletal disorder (arthritis, low-back pain, and NR; n=14)	14.3	21.4	7.14	14.3	42.9	7.14	0	21.4
Neurodegenerative disease (AD[j], ALS[k], MND[l], MS[m], PD[n], and NR; n=33)	21.2	48.5	6.06	0	36.4	0	9.09	3.03
Sensory impairment (n=1)	0	0	0	0	100	0	0	0
Reproductive disorder (birth defects, infertility, spontaneous abortion, and NR; n=24)	7.14	35.7	7.14	0	64.3	0	7.14	0
Respiratory disorder (allergy, asthma, COPD[o], pneumonia, sarcoidosis, and NR; n=39)	47.1	35.3	11.8	2.94	35.3	0	8.82	5.88
Skin disorder (dermatitis and NR; n=2)	50	50	0	0	50	0	0	0

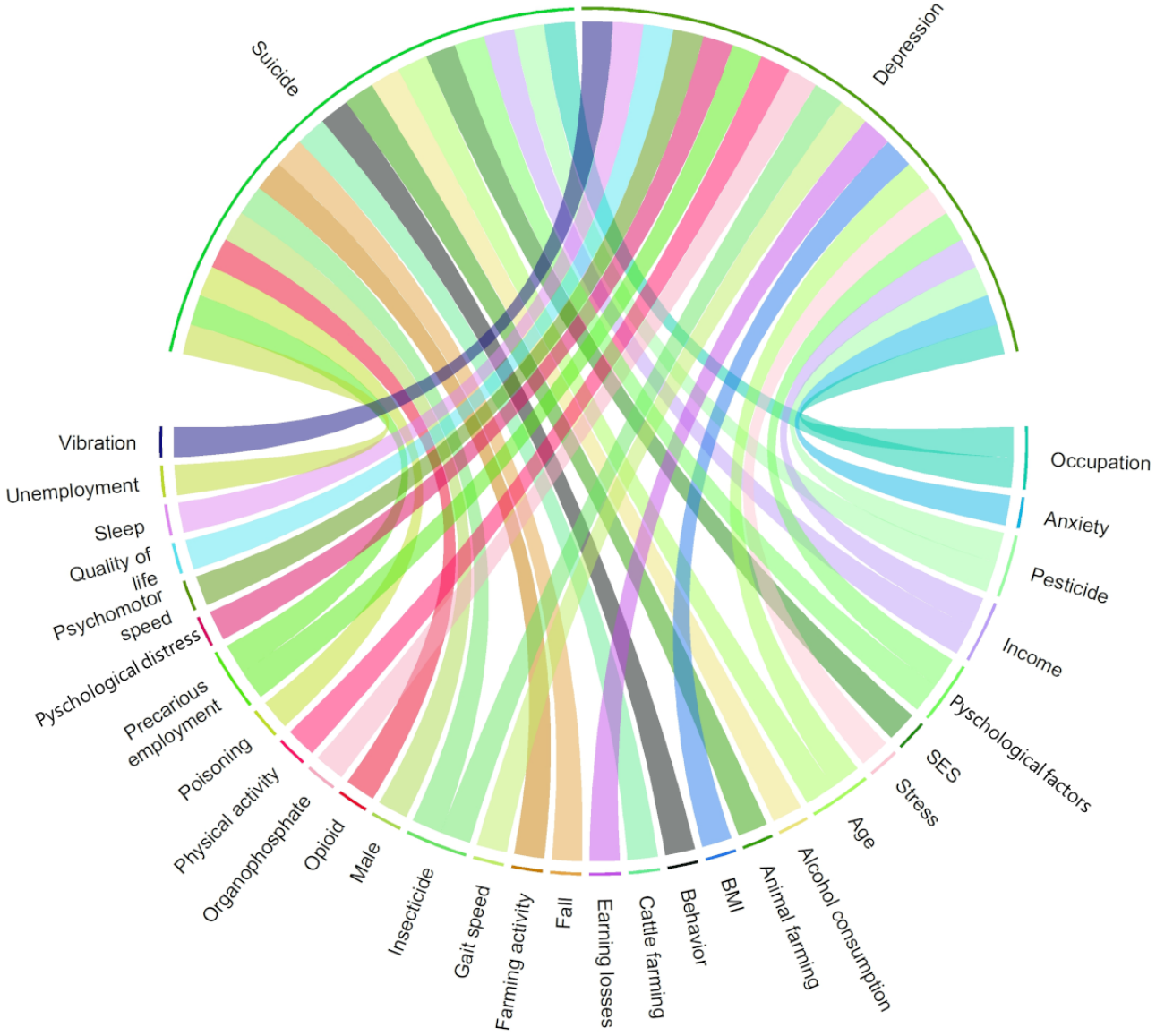
[a]CA: chemical agent.

[b]BA: biological agent.

XSL•FO
RenderX

[c]BF: biomechanical factor.
[d]PA: physical agent.
[e]SF: socioeconomic factor.
[f]PF: psychological factor.
[g]IBD: inflammatory bowel disease.
[h]RA: rheumatoid arthritis.
[i]NR: not reported.
[j]AD: Alzheimer disease.
[k]ALS: amyotrophic lateral sclerosis.
[l]MND: motor neuron disease
[m]MS: multiple sclerosis.
[n]PD: Parkinson disease.
[o]COPD: chronic obstructive pulmonary disease.

Figure 6. Chord diagram of keyword co-occurrence between potential risk factors and mental health disorder keywords from the 296 articles published between 1975 and 2024. Disease-related keywords are displayed on the top half of the chord diagram, while exposome-related keywords are displayed on the bottom half. Each chord or link indicates that an exposome-related keyword was mentioned with a disease-related keyword (co-occurrence) at least once in the same publication. The chord color differs from one exposome-related keyword to another. SES: socioeconomic status.



XSL•FO
RenderX

Discussion

Principal Findings

This review provides the first comprehensive and objective synthesis of research on the use of AHDs to address health issues in farming populations. It identifies major contributors, key publications, and existing research gaps while also suggesting future directions for leveraging AHDs to study health issues in farming populations. Overall, findings indicate that only a small part of the exposome and a limited range of health events have been examined within farming populations through the reuse of AHDs.

Current Directions in AHD Use for Public Health Research in Farming Populations

Research using AHDs in farming populations has been predominantly conducted in developed countries [150,225], with the United States [172,223,227]; France [207,222,224]; Canada [145,186,195]; and Scandinavian nations [182,189], including Denmark, Finland, Norway, and Sweden, leading the field. This dominance is linked to considerable funding from these regions and international collaborations. Scandinavian countries are particularly advanced in AHD use, offering databases that are highly complete, accessible, and well-integrated into public health research. AHDs from Denmark, Sweden, Canada, and France also provide comprehensive data on a patient's digital trajectory within their respective health systems [93,98,104,108,113,115,128,247]. France stands out further, with an AHD dedicated specifically to the entire farming population (MSA). This may explain the frequency of large-scale and long-term studies from these countries, some of which included >100,000 participants. However, many studies still had a regional focus, partly due to the use of AHDs by traditional epidemiological studies such as AGRICAN [207] and AHS [172], which rely on limited resources [47-50,53,54,57]. These studies often used AHDs to identify farming populations for inclusion or to supplement cohort data. The international AGRICOH consortium, initiated by the US National Cancer Institute and the International Agency for Research on Cancer, includes 11 (38%) of the 29 cohorts identified in this review [150,225]. However, several cohorts in AGRICOH were not identified, potentially due to lack of publications, language barriers, or limited use of AHDs. There were many publications associated with these well-established cohorts, for which many of the most prolific authors were working [172,182,189,207,222-227].

The most frequently used AHDs in farming population health research were disease registers, followed by electronic health or medical records and insurance claims. More than two-thirds of the studies used disease registers, in particular, cancer and mortality registers. This is not surprising because disease registers are created for clinical and research purposes with a continuous, exhaustive, and optimal digital collection of individual data regarding ≥1 health event in a geographically defined population [53,105]. The coding systems and the granularity of information related to health outcomes, populations, and exposure determinants varied widely across studies. Most studies (291/296, 98.3%) used AHDs to collect sociodemographic and health event information [222].

There was no consensus on the best methods or proxies to assess farming exposure. A variety of exposure proxies and determinants were used across studies, with indirect methods being the most common. Many studies (237/296, 80.1%) dichotomized proxy, for example, classifying individuals as farmers or nonfarmers or as pesticide-exposed versus nonexposed. In nearly two-thirds of the included studies, job title (ie, "being a farmer") served as the primary exposure proxy. About one-third of studies took into account specific farming activities (eg, dairy farming and crop farming) to reflect the diversity of farming practices. This approach is a valuable proxy for agricultural exposure, offering a broader representation of the farming exposome, which involves multiple stressors beyond just pesticides [147,152,188,201,248,249]. Farming activity information was often derived from digitalized data, such as agricultural censuses or self-reported data from mandatory insurance enrollments [152]. Many studies (111/296, 37.5%) combined AHDs with self-reported data (eg, questionnaires) [172,207], which allowed for more comprehensive data collection but tended to restrict the scope to regional studies due to resource constraints. These studies typically yielded high-quality data, with more potential confounders considered compared to studies relying solely on AHDs. Most studies (68/111, 61.3%) using self-reported data focused on single exposures, mainly pesticides, with only 1 study addressing multiexposure to pesticides [236]. Biological monitoring and airborne chemical sampling were rarely conducted, likely due to practical and financial constraints and the short half-lives of most pesticides [250]. Dermal chemical monitoring has not been reported, even though it is the main exposure route for pesticides [251]. The high number of studies investigating exposure to pesticides may be explained by the fact that AHS focuses on pesticide applicators and their spouses [172] and because many pesticides have adverse health effects on humans, such as neurotoxicity or endocrine disruption [252-256]. Beyond pesticides, farmers face exposure to other chemicals [257], such as air pollution; micro- and nanoplastics [258-263]; biological agents (eg, endotoxins and zoonoses) [264-266]; physical agents (eg, UV radiation, noise, and vibrations) [187,267]; biomechanical factors (eg, repetitive movements, heavy load, and working posture) [198,268,269]; and psychosocial stressors [270-274], which have been less studied than pesticides. Despite these multiple exposures, the broader farming exposome remains understudied.

In addition to AHDs, some studies integrated other secondary data, such as climate data [187,275], air quality data [194,217], JEMs [212,276,277], or crop-exposure matrices [278]. JEMs provide exposure level estimates for various chemicals and stressors based on job categories [250,279]. Although JEMs can provide valuable exposure data, they often lack the specificity of individual-level data, making it difficult to account for task-specific risks, temporal variations, and the inclusion of specific worker subgroups such as female individuals [250,279-282]. The lack of a universal standard for JEMs further complicates their application, which may explain why many

XSL•FO
RenderX

studies still rely on self-reported data for more accurate exposure assessment despite the risk of recall bias [250,279-282].

The health outcomes studied were predominantly cancer [145,150,151,171-175,177-179,182,183,185,189,190,196,207,210-212,223], mortality [173,186,194,195,202,205], workplace injuries [198,208], respiratory disorders (eg, asthma and COPD) [151,180,181,213,224,226], neurodegenerative diseases (eg, Parkinson disease) [111,151,176,184,187,188,193,197,201,248,249,283], and mental health issues [151,244,273], which represent focal points within the research field. This is not surprising because many well-established cohorts centered on cancer research, in particular AGRICAN [207], Cancer in the Norwegian Agricultural Population [182], and NOCCA [189]. In addition, arsenic and inorganic arsenic compounds are classified as carcinogenic to humans by the International Agency for Research on Cancer [252,253,284], while malathion, glyphosate, diazinon, dichlorodiphenyltrichloroethane, and occupational exposures in spraying and application of nonarsenical insecticides are classified as probably carcinogenic to humans (group 2A), and several other pesticides are ranked as possibly carcinogenic to humans (group 2B), such as tetrachlorvinphos and parathion. Regarding mortality, it is often cancer and suicide mortalities that are investigated [186,195,202,244]. Furthermore, several pesticides are neurotoxic [252,255], but existing studies focused mainly on Parkinson disease and multiple sclerosis, with a paucity of data on Alzheimer disease and other neurodegenerative diseases [147,283]. In contrast, certain areas, such as cardiovascular diseases [151,194,203,227,285,286], autoimmune disorders [168,219,237], musculoskeletal disorders [192,204,245,287,288], reproductive disorders [242,289,290], sleep disorders [191], aging-related conditions [151,291], hearing impairment [267,292], and the microbiome [8,293-298], remain underexplored, despite their potential relevance to farming populations.

Challenges of Reusing AHD for Public Health Research in Farming Populations

Each AHD presents unique advantages and limitations. For example, large sample sizes and a large number of available health events are strengths, while generalizability and the absence of key confounders are challenges [64,93,95,105,115]. Access to AHDs is frequently restricted by a variety of factors, including governance and technical barriers, such as language, data structure, interoperability, and coding systems. Additional challenges stem from the type of AHD (eg, insurance claims or cancer registers), inadequate documentation (eg, absence of a data dictionary), limited accessibility due to costs or conditions, and jurisdictional and legal constraints [30,62,64,81,113,115]. Identifying the optimal AHD for a given research question is also complex, especially when considering the heterogeneity in coding systems and country-specific data structures. In countries such as Scandinavia, Canada, and France, individual identifiers facilitate data linkage across multiple AHDs, enhancing research opportunities [93,104,113,115,128,247]. However, many AHDs are not ready for research and require significant processing, cleaning, and understanding before they can be analyzed [93,105,113,222,299]. Another major challenge is the long lag between data access, analysis, and research publication. On average, studies used data that were 7 years old at the time of publication, largely due to delays in data access, administrative approvals, and the need to prepare complex datasets for analysis [223,300]. For instance, the TRACTOR project took 2 years to clean and prepare its dataset for research use [222]. These delays are compounded by the time required to conduct statistical analyses and prepare manuscripts for publication, as well as review and publication times (delay from submission to acceptance and from acceptance to publication). Another limitation of AHDs is the lack of detailed exposure data. AHDs rarely include exposure information due to their administrative focus, requiring researchers to supplement with additional data sources, such as JEMs or self-reported data. When exposure information is recorded in AHDs, it is often too generalized, typically only reflecting broad job classifications, such as farming, without specifying detailed activities or stressors. There are some exceptions, such as MSA data, which capture a wider range of specific farming activities (eg, dairy farming) [222]. However, exposure to specific stressors (eg, chemical compounds) is largely absent from AHDs.

The reference populations used in farming studies vary, which precludes direct comparisons and limits the generalizability of the findings. For example, AGRICAN used the general population as a reference [207], while TRACTOR used a farming population [152,168]. Furthermore, studies differ in their focus on specific farming populations, such as the entire agricultural workforce [207], farm managers [152], or pesticide applicators [172], which may lead to distinct exposure profiles that influence health outcomes because these farming populations have different socioeconomic status, experiences, and behaviors. Hence, to avoid or lessen bias, some studies focused on specific farming populations [152,172]. Moreover, the scope of farming populations included in studies is often limited, omitting subgroups, such as farm families, nearby residents, or consumers exposed to agricultural products, which limits the broader application of the findings. In addition, farming practices can vary significantly between countries and studies, and there is no international standardized classification for farming activities. In some cases, farming categories are derived from legal or administrative sources, as seen in the MSA data [152,168,283]. This lack of standardization limits the comparability and generalizability of findings across studies. In addition, the generalizability of the findings to other countries when using AHDs may be limited because of the differences in health care systems [93].

There are several well-known limitations of AHDs that complicate the investigation of health outcomes [301]. Health events captured in AHDs are typically limited to those requiring medical attention, which may not reflect the true incidence of diseases. In addition, the level of detail varies across diseases, even within the same AHD [92]. Although diagnostic accuracy is generally higher in disease registers, these are often geographically limited and cover only a subset of health conditions. For example, in France, cancer registers only cover 23 (24%) out of 96 administrative regions [155]. In addition, certain conditions, such as depression, are not covered by any registers. Identifying health outcomes in AHDs often requires complex algorithms that combine data from multiple sources,

XSL•FO
RenderX

such as drug reimbursements, disease declarations, or medical procedures [104,105,128,152,302-304]. In addition, inconsistencies in case definitions and algorithms across studies and countries hinder the ability to compare and pool risk estimates [104,302,304]. Some AHDs also lack critical clinical information, such as laboratory results and genetic data [115,128,305], and the recorded diagnosis or treatment date may not correspond to the actual onset of the disease. Furthermore, diagnosis codes are not always indicative of a confirmed diagnosis.

Emerging Opportunities and Research Needs

While AHDs are well-used in certain countries, there are underexplored opportunities in regions such as the United Kingdom, where AHDs exist but are underused for research [113]. For low- and middle-income countries, the development and access to AHDs remain limited, and international support is needed to expand this research infrastructure. As already reported by Habib et al [306], there is a notable lack of sex-specific data, even though occupational exposures and health outcomes can vary significantly between sexes due to genetic, physiological, psychological, and behavioral factors [307-311]. Future research should address these disparities to provide a more comprehensive understanding of health risks in farming communities. Although there are inherent delays in using AHDs due to the time required for data generation, consolidation, and access, we advocate for the continued publication of studies, even those using older data. Historical data remain vital for better understanding long-term health patterns, particularly for diseases such as cancer, where tumor initiation can span decades [312,313]. Editors should encourage the publication of studies using older datasets, especially when addressing long-term health outcomes (eg, cancer and neurodegenerative diseases) or when recent data are not available [312].

None of the analyzed AHDs fully adhere to the FAIR principles, possibly because most were developed before the establishment of these principles in 2016 [228-230]. Moreover, the assessment of FAIR compliance of AHDs relied solely on information presented in publications, which may not provide a comprehensive evaluation. Nevertheless, there is a critical need to advocate for the integration of FAIR principles within AHDs to enhance public health research [228-230]. Currently, the landscape is favorable for data reuse, particularly with initiatives such as the forthcoming European Health Data Space [314-316]. Data reuse extends beyond mere access; it encompasses data discovery, a fundamental aspect of FAIR principles that involves recognizing the existence of databases [228-230]. To facilitate this, the creation of data catalogs is essential [228]. Numerous data repositories, such as Re3Data [317], Zenodo [318], CANUE [319], Figshare [320], “Epidémiologie – France” [321], data.gouv [322], Dataverse [320], or Data Europa [320], already exist. In addition, specialized multidisciplinary open-access and peer-reviewed journals such as *Scientific Data* and *Data in Brief* publish datasets [318]. A dataset search can also be conducted using the Google (Google LLC) platform [323]. However, the documentation and access conditions for datasets can highly vary across inventories, complicating the selection process for researchers. The absence of indicators or scores for data reusability further hampers efforts to identify the most suitable datasets for specific research questions [45,69]. To our knowledge, no comprehensive catalog of AHDs currently exists to date. A web-based inventory of AHDs, modeled after existing resources, such as OccupationalCohorts.net [324], OccupationalExposureTools.net [325], and Toxicological and Exposure Database Inventory [10], could greatly enhance research endeavors. The motivation for analyzing AHDs often stems from the data they contain. Consequently, as data availability increases, researchers will be better positioned to formulate research questions and engage in a parallel process of “datagraphy” or “datagraphic search” akin to traditional bibliographic research. The objective of datagraphy would be to determine which datasets are best suited for addressing specific research questions, highlighting the need for accessible catalogs to support this goal.

There is also an opportunity to integrate other secondary data, such as person-generated data (eg, mobile health, social media, digital footprints, and wearable sensors); contextual data (eg, climate and air quality); and smart agriculture data [2,83,84,326-328]. These datasets, largely untapped in farming population research, could provide new insights into health outcomes and environmental exposures [101].

Nationwide studies using big data were a hot spot. AI, such as machine learning, is particularly useful for analyzing big data and holds substantial promise for future research [329-331], particularly for identifying predictors of health outcomes in farming populations [332]. To date, AI has been underused, with only 2 studies using it, 1 identifying occupational injuries in agriculture [208] and 1 reviewing the development of chronic kidney disease risk prediction models [209]. Incorporating AI, along with cohort enrichment and interdisciplinary expert interpretation, could open new avenues for research.

Many studies continue to examine agriculture as a broad category, highlighting the need for more detailed investigations into specific farming activities and tasks [147,152,188,201,248,249]. Our analysis reveals major research gaps in understanding environmental and occupational exposures among farming populations, particularly with regard to emerging concerns such as per- and polyfluoroalkyl substances, biological agents, micro- and nanoplastics, and the impact of climate change. Climate change is a critical issue for agriculture, as it may drive shifts in pests, diseases, and farming practices [274,333-339]. Parental exposure appears to be a theme that will soon become a hot topic. Furthermore, research is needed to explore the farming exposome, particularly focusing on the “mixture effects” of multiple simultaneous exposures [340,341]. Omics data, which have not been used in farming population studies to date, represent a promising avenue for future research because genetics and metabolism were found to be a niche theme. Omics data refer to the large-scale datasets generated from various omics technologies that analyze biological molecules (eg, genomics, transcriptomics, proteomics, and metabolomics), which provide comprehensive insights into different biological layers and processes [11,342,343].

To enhance the characterization of farming exposome research using keyword analysis, there is a pressing need for standardized keyword reporting. We advocate for the development of a

XSL•FO
RenderX

standardized approach to reporting keywords in scientific journals, including defining a minimum set of information (eg, study type, health outcome, population studied, data sources, and positive, negative, or null associations) and adopting a list of standardized terms. Although challenging, this approach would improve literature searches, make data more comparable and FAIR [228-230], and lead to more efficient, frugal (less time and energy spent to identify relevant information), and accurate synthesis of the scientific literature, such as in reviews and bibliometric analyses.

The prominence of topics such as cancer, neurodegenerative diseases, mortality, injuries, and mental health issues underscores the need for targeted prevention strategies. The thematic map analysis indicates that reproductive disorders (eg, birth defects, endometriosis, and infertility) are on the verge of becoming a central research focus. Emerging and understudied health conditions, including ocular disorders, autoimmune diseases (eg, inflammatory bowel disease and rheumatoid arthritis), sleep disorders (eg, sleep apnea), cardiovascular diseases, and musculoskeletal conditions (eg, low-back pain), warrant increased attention and further research. Aging-related health issues, such as frailty, also represent promising avenues for future research, particularly given the growing aging population and associated health care challenges [24].

Limitations

The findings of this review should be considered in view of their limitations. Because of time and resource constraints, a single screening approach was used, with only 1 author (PP) conducting the review and bibliometric analysis. While single screening is an efficient use of time and resources, there is a higher risk of missing relevant studies than when using dual screening [344,345]. However, when completed by an experienced reviewer familiar with the research topic, the proportion of missed studies is limited and estimated to be around 3% [344]. Therefore, we cannot exclude the possibility that some studies may have been missed. Nevertheless, we are confident that none of these methodological limitations would change the overall conclusions of this work. Our restriction on articles published in English and French may have inadvertently excluded potentially relevant publications. We cannot exclude the possibility that publications using AHDs for addressing health issues in farming populations may have been missed if there was no mention of AHD in the publications' titles and abstracts. However, it is important to mention that our search strategy was similar to recent reviews that specifically examined the use of AHDs for population-based research [93,96,103,104]. We further broadened our search by including synonyms to improve the comprehensiveness of our literature search. Some details and specificities on the AHDs and other digital data used may be limited because only information reported in each study was used. Shortcomings inherent to bibliometric analysis cannot be excluded. Some authors may have duplicate names, and namesakes could exist. This limitation could not be prevented as a unique author identifier (eg, Open Researcher and Contributor ID number) was not available. Self-citation could not be identified.

While our keyword analysis helped map the farming exposome in AHD-based public health research, this profile is incomplete and potentially biased. Because our review focused on AHD-based studies, we likely missed relevant epidemiological studies not using AHDs, leaving gaps in our understanding of the complete farming exposome across public health. In addition, the variability in keyword reporting practices across journals introduced bias into our keyword analysis. Some journals limit the number of keywords, and the lack of standardized keyword ontologies adds further variability. To mitigate this bias, we manually harmonized the keywords (eg, the use of 1 unique term for a given entity). While this approach is time-consuming, it allows for a more accurate analysis. For instance, if this approach was not performed, the same entity could be designated by various terms that would have been considered separate entities or terms, potentially resulting in underestimates (eg, in the number of publications). Despite these challenges, the findings from our scoping review were consistent with the keyword analysis.

Notwithstanding the aforementioned limitations, most of which are inherent to all scoping reviews and bibliometric analyses [93,96,103,104,158-160], we are confident that our findings can provide a comprehensive picture of what has been published until now (the current state of research and general directions) regarding the use of AHDs for addressing health issues in farming populations. This study may lay the groundwork for researchers to quickly identify research priorities and emerging research directions investigating health issues in farming populations using AHDs.

Conclusions

Technological advancements have greatly increased the volume of research data available, positioning AHDs as critical resources for population-based public health studies [41]. Our review underscores the broad public health implications of AHDs, providing actionable insights for researchers, physicians, and policy makers (Textbox 3). Addressing the identified research gaps is crucial to comprehensively understanding health risks in farming populations.

The insights derived from AHDs can inform meaningful recommendations for policy makers and guide future research directions, ultimately aiding health services and health policy development. Our findings underscore the necessity of comprehensive, interdisciplinary approaches to better understand and mitigate the health risks encountered by farming populations. Such efforts will improve data comparability and research quality while also supporting the formulation of targeted prevention strategies. This, in turn, will enhance health outcomes for farming populations and promote the sustainability of agriculture in an increasingly dynamic environment. The findings from this review offer insights that are not only relevant to farming populations but also potentially generalizable to other populations.

XSL•FO
RenderX

Textbox 3. Take-home messages.

Farming population

Research focusing on low- and middle-income countries, as well as on underrepresented subgroups within farming communities (eg, women, children, and contingent workers), remains insufficiently developed. These areas warrant further investigation to ensure more comprehensive insights.

Administrative health database (AHD) use

The use of AHDs in public health research among farming populations is expanding, offering major potential to enhance epidemiological studies and inform public health decisions. Promoting AHD-based research alongside the integration of other secondary data and artificial intelligence approaches represents a promising direction for future exploration. There is also a need to promote findable, accessible, interoperable, and reusable principles. Creating an AHD catalog or inventory could be a solution that would allow researchers to conduct a "datagraphic search" akin to traditional bibliographic research.

Farming exposure

Published studies on farming-related exposures often rely on broad proxies, such as job titles, neglecting to capture the nuances of specific agricultural tasks. While pesticide exposure remains a predominant research focus, emerging concerns, such as per- and polyfluoroalkyl substances, biological agents, micro or nanoplastics, and the effects of climate change, require urgent attention. The farming exposome remains underexplored despite its potential to uncover important associations between risk factors and a diverse range of health outcomes.

Health outcomes

Cancer, respiratory diseases, neurodegenerative disorders, and mental health issues are among the most frequently studied health outcomes in farming populations. However, significant gaps exist in understanding other critical conditions, such as cardiovascular diseases, reproductive disorders, ocular conditions, autoimmune diseases, musculoskeletal disorders, age-related health issues, and microbiome impacts. Addressing these overlooked areas is essential for a more complete understanding of the health risks faced by farming communities.

Acknowledgments

This work was partly funded by MIAI@Grenoble Alpes (Multidisciplinary Institute in Artificial Intelligence at Grenoble Alpes; grant ANR-19-P3IA-0003, 2019) and the French National Research Agency within the framework of the Investissements d'avenir program (grants ANR - 10 - AIRT - 05, 2012; ANR - 15 - IDEX - 02, 2016). The funding sources had no role in the study design and conduct of the study; collection, management, analysis, and interpretation of data; writing of the report; preparation, review, or approval of the paper; or decision to submit the paper for publication. The authors were not precluded from accessing data in the study, and all authors took final responsibility for the decision to submit for publication.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request and also in the multimedia appendices.

Authors' Contributions

PP was involved in the conceptualization, methodology, software selection, validation, formal analysis, investigation, data curation, writing the original draft and reviewing and editing, visualization, project administration, and funding acquisition of the study. NV was involved in the conceptualization, reviewing and editing, resource collection, and funding acquisition of the study.
All authors meet all 4 criteria for authorship in the International Committee of Medical Journal Editors recommendations. All authors affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any potential discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File , 8778 KB - publichealth_v11i1e62939_app1.docx]

Multimedia Appendix 2
Publications included and analyzed.
[XLSX File (Microsoft Excel File), 80 KB - publichealth_v11i1e62939_app2.xlsx]

XSL•FO
RenderX

Multimedia Appendix 3
Co-occurrence between exposome-related keywords and health event-related keywords.
[XLSX File (Microsoft Excel File), 39 KB - publichealth_v11i1e62939_app3.xlsx]

Multimedia Appendix 4
PRISMA-ScR checklist.
[PDF File (Adobe PDF File), 186 KB - publichealth_v11i1e62939_app4.pdf]

Multimedia Appendix 5
BIBLIO checklist for reporting the bibliometric reviews of the biomedical literature.
[PDF File (Adobe PDF File), 270 KB - publichealth_v11i1e62939_app5.pdf]

References

1. Azari R, Borisch B. What is public health? A scoping review. Arch Public Health 2023 May 10;81(1):86 [FREE Full text] [doi: 10.1186/s13690-023-01091-6] [Medline: 37165370]
2. Velmovitsky PE, Kirolos M, Alencar P, Leatherdale S, Cowan D, Morita PP. Leveraging mHealth technologies for public health. JMIR Public Health Surveill 2024 Sep 12;10:e49719 [FREE Full text] [doi: 10.2196/49719] [Medline: 39265164]
3. Barouki R, Lévi Y. Rapport 24-02. L'exposome, une contribution majeure pour la prévention. Bull Acad Natl Med 2024 May;208(5):523-535. [doi: 10.1016/j.banm.2024.03.004]
4. Barouki R, Audouze K, Becker C, Blaha L, Coumoul X, Karakitsios S, et al. The exposome and toxicology: a win-win collaboration. Toxicol Sci 2022 Feb 28;186(1):1-11 [FREE Full text] [doi: 10.1093/toxsci/kfab149] [Medline: 34878125]
5. Bennett B, Workman T, Smith MN, Griffith WC, Thompson B, Faustman EM. Characterizing the neurodevelopmental pesticide exposome in a children's agricultural cohort. Int J Environ Res Public Health 2020 Feb 25;17(5):1479 [FREE Full text] [doi: 10.3390/ijerph17051479] [Medline: 32106530]
6. Goodrich JA, Wang H, Jia Q, Stratakis N, Zhao Y, Maitre L, et al. Integrating Multi-Omics with environmental data for precision health: a novel analytic framework and case study on prenatal mercury induced childhood fatty liver disease. Environ Int 2024 Aug;190:108930 [FREE Full text] [doi: 10.1016/j.envint.2024.108930] [Medline: 39128376]
7. Haddad N, Andrianou XD, Makris KC. A scoping review on the characteristics of human exposome studies. Curr Pollution Rep 2019 Nov 28;5(4):378-393. [doi: 10.1007/S40726-019-00130-7]
8. Merra G, Gualtieri P, La Placa G, Frank G, Della Morte D, De Lorenzo A, et al. The relationship between exposome and microbiome. Microorganisms 2024 Jul 08;12(7):1386 [FREE Full text] [doi: 10.3390/microorganisms12071386] [Medline: 39065154]
9. Pandics T, Major D, Fazekas-Pongor V, Szarvas Z, Peterfi A, Mukli P, et al. Exposome and unhealthy aging: environmental drivers from air pollution to occupational exposures. Geroscience 2023 Dec 09;45(6):3381-3408 [FREE Full text] [doi: 10.1007/s11357-023-00913-3] [Medline: 37688657]
10. Petit P. Toxicological and exposure database inventory: a review. Int J Hyg Environ Health 2022 Sep;246:114055. [doi: 10.1016/j.ijheh.2022.114055] [Medline: 36279787]
11. Price EJ, Vitale CM, Miller GW, David A, Barouki R, Audouze K, et al. Merging the exposome into an integrated framework for "omics" sciences. iScience 2022 Mar 18;25(3):103976 [FREE Full text] [doi: 10.1016/j.isci.2022.103976] [Medline: 35310334]
12. Santos S, Maitre L, Warembourg C, Agier L, Richiardi L, Basagaña X, et al. Applying the exposome concept in birth cohort research: a review of statistical approaches. Eur J Epidemiol 2020 Mar 27;35(3):193-204 [FREE Full text] [doi: 10.1007/s10654-020-00625-4] [Medline: 32221742]
13. Siroux V, Agier L, Slama R. The exposome concept: a challenge and a potential driver for environmental health research. Eur Respir Rev 2016 Jun 31;25(140):124-129 [FREE Full text] [doi: 10.1183/16000617.0034-2016] [Medline: 27246588]
14. Tagliaferro S, Maio S, Pirona F, Stanisci I, Sarno G, Silvi P, EarlyFOOD, HEALS EXHES, et al. Assessing external exposome by implementing an environmental data management system using open data. Sci Rep 2024 Jul 26;14(1):17142 [FREE Full text] [doi: 10.1038/s41598-024-62924-0] [Medline: 39060268]
15. Vineis P, Barouki R. The exposome as the science of social-to-biological transitions. Environ Int 2022 Jul;165:107312 [FREE Full text] [doi: 10.1016/j.envint.2022.107312] [Medline: 35635963]
16. Wei X, Huang Z, Jiang L, Li Y, Zhang X, Leng Y, et al. Charting the landscape of the environmental exposome. Imeta 2022 Dec 02;1(4):e50. [doi: 10.1002/imt2.50] [Medline: 38867899]
17. Wild CP. The exposome: from concept to utility. Int J Epidemiol 2012 Feb;41(1):24-32. [doi: 10.1093/ije/dyr236] [Medline: 22296988]
18. Zhang P, Carlsten C, Chaleckis R, Hanhineva K, Huang M, Isobe T, et al. Defining the scope of exposome studies and research needs from a multidisciplinary perspective. Environ Sci Technol Lett 2021 Oct 12;8(10):839-852 [FREE Full text] [doi: 10.1021/acs.estlett.1c00648] [Medline: 34660833]

XSL•FO
RenderX

19. Ibanez A, Melloni L, Świeboda P, Hynes W, Ikiz B, Ayadi R, et al. Neuroecological links of the exposome and one health. Neuron 2024 Jun 19;112(12):1905-1910 [FREE Full text] [doi: 10.1016/j.neuron.2024.04.016] [Medline: 38723637]
20. Liang KH, Colombijn JM, Verhaar MC, Ghannoum M, Timmermans EJ, Vernooij RW. The general external exposome and the development or progression of chronic kidney disease: a systematic review and meta-analyses. Environ Pollut 2024 Oct 01;358:124509 [FREE Full text] [doi: 10.1016/j.envpol.2024.124509] [Medline: 38968981]
21. Canfell OJ, Littlewood R, Burton-Jones A, Sullivan C. Digital health and precision prevention: shifting from disease-centred care to consumer-centred health. Aust Health Rev 2022 Jun 10;46(3):279-283. [doi: 10.1071/AH21063] [Medline: 34882538]
22. Mess F, Blaschke S, Schick TS, Friedrich J. Precision prevention in worksite health-a scoping review on research trends and gaps. PLoS One 2024 Jun 10;19(6):e0304951 [FREE Full text] [doi: 10.1371/journal.pone.0304951] [Medline: 38857277]
23. Wang Y, Zhu M, Ma H, Shen H. Polygenic risk scores: the future of cancer risk prediction, screening, and precision prevention. Med Rev (2021) 2021 Dec;1(2):129-149 [FREE Full text] [doi: 10.1515/mr-2021-0025] [Medline: 37724297]
24. Kalia V, Belsky DW, Baccarelli AA, Miller GW. An exposomic framework to uncover environmental drivers of aging. Exposome 2022;2(1):osac002 [FREE Full text] [doi: 10.1093/exposome/osac002] [Medline: 35295547]
25. Chung MK, House JS, Akhtari FS, Makris KC, Langston MA, Islam KT, Members of the Exposomics Consortium. Decoding the exposome: data science methodologies and implications in exposome-wide association studies (ExWASs). Exposome 2024;4(1):osae001 [FREE Full text] [doi: 10.1093/exposome/osae001] [Medline: 38344436]
26. Andreu-Perez J, Poon CC, Merrifield RD, Wong ST, Yang GZ. Big data for health. IEEE J Biomed Health Inform 2015 Jul;19(4):1193-1208 [FREE Full text] [doi: 10.1109/JBHI.2015.2450362] [Medline: 26173222]
27. Li C, Huang M. Environmental sustainability in the age of big data: opportunities and challenges for business and industry. Environ Sci Pollut Res Int 2023 Dec 03;30(56):119001-119015. [doi: 10.1007/s11356-023-30301-5] [Medline: 37919492]
28. Pastorino R, De Vito C, Migliara G, Glocker K, Binenbaum I, Ricciardi W, et al. Benefits and challenges of Big Data in healthcare: an overview of the European initiatives. Eur J Public Health 2019 Oct 01;29(Supplement_3):23-27 [FREE Full text] [doi: 10.1093/eurpub/ckz168] [Medline: 31738444]
29. Flahault A, Geissbuhler A, Guessous I, Guérin P, Bolon I, Salathé M, et al. Precision global health in the digital age. Swiss Med Wkly 2017 Apr 07;147(1314):w14423 [FREE Full text] [doi: 10.4414/smw.2017.14423] [Medline: 28421566]
30. Manaouil C, Chamot S, Petit P. Le médecin confronté à l'IA (Intelligence artificielle): Éthique et responsabilité. Méd Droit 2024 Jun;2024(186):50-66. [doi: 10.1016/j.meddro.2024.02.001]
31. Samuel G, Lucassen AM. The environmental sustainability of data-driven health research: a scoping review. Digit Health 2022 Jul 05;8:20552076221111297 [FREE Full text] [doi: 10.1177/20552076221111297] [Medline: 35847526]
32. Adam NR, Wieder R, Ghosh D. Data science, learning, and applications to biomedical and health sciences. Ann N Y Acad Sci 2017 Jan 25;1387(1):5-11. [doi: 10.1111/nyas.13309] [Medline: 28122121]
33. Iyamu I, Xu AX, Gómez-Ramírez O, Ablona A, Chang H, Mckee G, et al. Defining digital public health and the role of digitization, digitalization, and digital transformation: scoping review. JMIR Public Health Surveill 2021 Nov 26;7(11):e30399 [FREE Full text] [doi: 10.2196/30399] [Medline: 34842555]
34. Mooney SJ, Pejaver V. Big data in public health: terminology, machine learning, and privacy. Annu Rev Public Health 2018 Apr 01;39:95-112 [FREE Full text] [doi: 10.1146/annurev-publhealth-040617-014208] [Medline: 29261408]
35. Wesson P, Hswen Y, Valdes G, Stojanovski K, Handley MA. Risks and opportunities to ensure equity in the application of big data research in public health. Annu Rev Public Health 2022 Apr 05;43(1):59-78 [FREE Full text] [doi: 10.1146/annurev-publhealth-051920-110928] [Medline: 34871504]
36. Iyamu I, Gómez-Ramírez O, Xu AX, Chang H, Watt S, Mckee G, et al. Challenges in the development of digital public health interventions and mapped solutions: findings from a scoping review. Digit Health 2022;8:20552076221102255. [doi: 10.1177/20552076221102255] [Medline: 35656283]
37. Iyamu I, McKee G, Haag D, Gilbert M. Defining the role of digital public health in the evolving digital health landscape: policy and practice implications in Canada. Health Promot Chronic Dis Prev Can 2024 Feb;44(2):66-69 [FREE Full text] [doi: 10.24095/hpcdp.44.2.04] [Medline: 38353941]
38. Khan S, Kandukuri DK, Subramaniyan EU, MohanaSundaram A. Harnessing the untapped potential of digital twin technology in digital public health interventions. Explor Digit Health Technol 2023 Nov 1;5(4):11-16. [doi: 10.37349/edht.2023.00003]
39. Maaß L, Pan CC, Freye M. Mapping digital public health interventions among existing digital technologies and internet-based interventions to maintain and improve population health in practice: protocol for a scoping review. JMIR Res Protoc 2022 Mar 31;11(3):e33404 [FREE Full text] [doi: 10.2196/33404] [Medline: 35357321]
40. Odone A, Buttigieg S, Ricciardi W, Azzopardi-Muscat N, Staines A. Public health digitalization in Europe. Eur J Public Health 2019 Oct 01;29(Supplement_3):28-35 [FREE Full text] [doi: 10.1093/eurpub/ckz161] [Medline: 31738441]
41. Wienert J, Jahnel T, Maaß L. What are digital public health interventions? First steps toward a definition and an intervention classification framework. J Med Internet Res 2022 Jun 28;24(6):e31921 [FREE Full text] [doi: 10.2196/31921] [Medline: 35763320]
42. Burgun A, Bernal-Delgado E, Kuchinke W, van Staa T, Cunningham J, Lettieri E, et al. Health data for public health: towards new ways of combining data sources to support research efforts in Europe. Yearb Med Inform 2017 Aug 19;26(01):235-240. [doi: 10.1055/S-0037-1606509]

XSL•FO
RenderX

43. Canfell OJ, Davidson K, Sullivan C, Eakin E, Burton-Jones A. Data sources for precision public health of obesity: a scoping review, evidence map and use case in Queensland, Australia. BMC Public Health 2022 Mar 24;22(1):584 [FREE Full text] [doi: 10.1186/s12889-022-12939-x] [Medline: 35331189]
44. Canfell OJ, Davidson K, Woods L, Sullivan C, Cocoros NM, Klompas M, et al. Precision public health for non-communicable diseases: an emerging strategic roadmap and multinational use cases. Front. Public Health 2022 Apr 8;10:854525 [FREE Full text] [doi: 10.3389/fpubh.2022.854525] [Medline: 35462850]
45. Sielemann K, Hafner A, Pucker B. The reuse of public datasets in the life sciences: potential risks and rewards. PeerJ 2020;8:e9954 [FREE Full text] [doi: 10.7717/peerj.9954] [Medline: 33024631]
46. Deal JA, Betz J, Lin FR, Reed NS. Interpreting results from epidemiologic studies. Semin Hear 2021 Feb 15;42(1):3-9 [FREE Full text] [doi: 10.1055/s-0041-1725996] [Medline: 33883787]
47. DiPietro NA. Methods in epidemiology: observational study designs. Pharmacotherapy 2010 Oct 06;30(10):973-984. [doi: 10.1592/phco.30.10.973] [Medline: 20874034]
48. Cancer epidemiology: principles and methods. International Agency for Research on Cancer. 1999. URL: https://publications.iarc.fr/_publications/media/download/4106/a72983d2524d630f6ad0539f747e4780447f1ae3.pdf [accessed 2024-04-29]
49. Munnangi S, Boktor SW. Epidemiology of study design. StatPearls. 2014. URL: https://www.ncbi.nlm.nih.gov/books/NBK470342/ [accessed 2024-04-29]
50. Röhrig B, du Prel JB, Wachtlin D, Blettner M. Types of study in medical research: part 3 of a series on evaluation of scientific publications. Dtsch Arztebl Int 2009 Apr;106(15):262-268 [FREE Full text] [doi: 10.3238/arztebl.2009.0262] [Medline: 19547627]
51. Himali JJ, Baril AA, Cavuoto MG, Yiallourou S, Wiedner CD, Himali D, et al. Association between slow-wave sleep loss and incident dementia. JAMA Neurol 2023 Dec 01;80(12):1326-1333. [doi: 10.1001/jamaneurol.2023.3889] [Medline: 37902739]
52. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. Lancet 2014 Mar 15;383(9921):999-1008 [FREE Full text] [doi: 10.1016/S0140-6736(13)61752-3] [Medline: 24084292]
53. Salines G, De LC. Les cohortes : intérêt, role et position de l'InVS. In: Saint-Maurice (Fra) : Institut de veille sanitaire, janvier 2010, 20 p. Princeton, NJ: Princeton University Press; 2018.
54. Zins M, Goldberg M. L'apport des bases de données d'origine administrative aux cohortes épidémiologiques : l'exemple de la cohorte Constances. Statistique et Société 2014;2(2):43-48 [FREE Full text]
55. Chen W, Wang X, Chen J, You C, Ma L, Zhang W, et al. Household air pollution, adherence to a healthy lifestyle, and risk of cardiometabolic multimorbidity: results from the China health and retirement longitudinal study. Sci Total Environ 2023 Jan 10;855:158896. [doi: 10.1016/j.scitotenv.2022.158896] [Medline: 36150596]
56. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China Health and Retirement Longitudinal Study (CHARLS). Int J Epidemiol 2014 Feb 12;43(1):61-68 [FREE Full text] [doi: 10.1093/ije/dys203] [Medline: 23243115]
57. Kesse-Guyot E, Assmann K, Andreeva V, Castetbon K, Méjean C, Touvier M, et al. Metadata correction of: lessons learned from methodological validation research in e-epidemiology. JMIR Public Health Surveill 2016 Oct 31;2(2):e160 [FREE Full text] [doi: 10.2196/publichealth.6876] [Medline: 27797322]
58. Quantin C. Bases de données médicoadministratives et épidémiologie : intérêt, limites. Revue d'Épidémiologie et de Santé Publique 2008 Sep;56(5):S254. [doi: 10.1016/j.respe.2008.06.011]
59. Babik KR, Shockey TM, Moore LL, Wurzelbacher SJ. Standardizing industrial hygiene data collection forms used by workers' compensation insurers. J Occup Environ Hyg 2018 Sep 04;15(9):676-685 [FREE Full text] [doi: 10.1080/15459624.2018.1490022] [Medline: 29985777]
60. Bosward R, Braunack-Mayer A, Frost E, Carter S. Mapping precision public health definitions, terminology and applications: a scoping review protocol. BMJ Open 2022 Feb 23;12(2):e058069 [FREE Full text] [doi: 10.1136/bmjopen-2021-058069] [Medline: 35197357]
61. Canfell OJ, Kodiyattu Z, Eakin E, Burton-Jones A, Wong I, Macaulay C, et al. Real-world data for precision public health of noncommunicable diseases: a scoping review. BMC Public Health 2022 Nov 24;22(1):2166 [FREE Full text] [doi: 10.1186/s12889-022-14452-7] [Medline: 36434553]
62. Chao K, Sarker MN, Ali I, Firdaus RB, Azman A, Shaed MM. Big data-driven public health policy making: potential for the healthcare industry. Heliyon 2023 Sep;9(9):e19681 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e19681] [Medline: 37809720]
63. Dolley S. Big Data's role in precision public health. Front Public Health 2018 Mar 7;6:68 [FREE Full text] [doi: 10.3389/fpubh.2018.00068] [Medline: 29594091]
64. Doyle CM, Lix LM, Hemmelgarn BR, Paterson JM, Renoux C. Data variability across Canadian administrative health databases: differences in content, coding, and completeness. Pharmacoepidemiol Drug Saf 2020 Jan 10;29 Suppl 1(S1):68-77. [doi: 10.1002/pds.4889] [Medline: 31507029]
65. Gallini A, Moisan F, Maura G, Carcaillon-Bentata L, Leray E, Haesebaert J, et al. [Identification of neurodegenerative diseases in administrative databases in France: a systematic review of the literature]. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S183-S197. [doi: 10.1016/j.respe.2017.01.115] [Medline: 28341166]

XSL•FO
RenderX

66. Khoury MJ, Iademarco MF, Riley WT. Precision public health for the era of precision medicine. Am J Prev Med 2016 Mar;50(3):398-401 [FREE Full text] [doi: 10.1016/j.amepre.2015.08.031] [Medline: 26547538]
67. McCullough LE, Maliniak ML, Amin AB, Baker JM, Baliashvili D, Barberio J, et al. Epidemiology beyond its limits. Sci Adv 2022 Jun 10;8(23):eabn3328 [FREE Full text] [doi: 10.1126/sciadv.abn3328] [Medline: 35675391]
68. Näher AF, Vorisek CN, Klopfenstein SA, Lehne M, Thun S, Alsalamah S, et al. Secondary data for global health digitalisation. Lancet Digit Health 2023 Feb;5(2):e93-101 [FREE Full text] [doi: 10.1016/S2589-7500(22)00195-9] [Medline: 36707190]
69. Koesten L, Vougiouklis P, Simperl E, Groth P. Dataset reuse: toward translating principles to practice. Patterns (N Y) 2020 Nov 13;1(8):100136 [FREE Full text] [doi: 10.1016/j.patter.2020.100136] [Medline: 33294873]
70. Hu H, Galea S, Rosella L, Henry D. Big data and population health: focusing on the health impacts of the social, physical, and economic environment. Epidemiology 2017 Nov;28(6):759-762. [doi: 10.1097/EDE.0000000000000711] [Medline: 28682850]
71. Shaw RJ, Harron KL, Pescarini JM, Pinto Junior EP, Allik M, Siroky AN, et al. Biases arising from linked administrative data for epidemiological research: a conceptual framework from registration to analyses. Eur J Epidemiol 2022 Dec 05;37(12):1215-1224 [FREE Full text] [doi: 10.1007/s10654-022-00934-w] [Medline: 36333542]
72. Subramanian M, Wojtusciszyn A, Favre L, Boughorbel S, Shan J, Letaief KB, et al. Precision medicine in the era of artificial intelligence: implications in chronic disease management. J Transl Med 2020 Dec 09;18(1):472 [FREE Full text] [doi: 10.1186/s12967-020-02658-5] [Medline: 33298113]
73. Velmovitsky PE, Bevilacqua T, Alencar P, Cowan D, Morita PP. Convergence of precision medicine and public health into precision public health: toward a big data perspective. Front Public Health 2021 Apr 6;9:561873 [FREE Full text] [doi: 10.3389/fpubh.2021.561873] [Medline: 33889555]
74. Zhang H, Hu H, Diller M, Hogan WR, Prosperi M, Guo Y, et al. Semantic standards of external exposome data. Environ Res 2021 Jun;197:111185 [FREE Full text] [doi: 10.1016/j.envres.2021.111185] [Medline: 33901445]
75. Daniels K, Bonnechère B. Harnessing digital health interventions to bridge the gap in prevention for older adults. Front Public Health 2023 Jan 8;11:1281923 [FREE Full text] [doi: 10.3389/fpubh.2023.1281923] [Medline: 38259780]
76. Liu F, Panagiotakos D. Real-world data: a brief review of the methods, applications, challenges and opportunities. BMC Med Res Methodol 2022 Nov 05;22(1):287 [FREE Full text] [doi: 10.1186/s12874-022-01768-6] [Medline: 36335315]
77. Maaß L, Freye M, Pan CC, Dassow HH, Niess J, Jahnel T. The definitions of health apps and medical apps from the perspective of public health and law: qualitative analysis of an interdisciplinary literature overview. JMIR Mhealth Uhealth 2022 Oct 31;10(10):e37980 [FREE Full text] [doi: 10.2196/37980] [Medline: 36315221]
78. Mooney SJ, Garber MD. Sampling and sampling frames in big data epidemiology. Curr Epidemiol Rep 2019 Mar 2;6(1):14-22 [FREE Full text] [doi: 10.1007/s40471-019-0179-y] [Medline: 31360626]
79. Pearson AL, Tribby C, Brown CD, Yang JA, Pfeiffer K, Jankowska MM. Systematic review of best practices for GPS data usage, processing, and linkage in health, exposure science and environmental context research. BMJ Open 2024 Feb 01;14(2):e077036 [FREE Full text] [doi: 10.1136/bmjopen-2023-077036] [Medline: 38307539]
80. Prosperi M, Min JS, Bian J, Modave F. Big data hurdles in precision medicine and precision public health. BMC Med Inform Decis Mak 2018 Dec 29;18(1):139 [FREE Full text] [doi: 10.1186/s12911-018-0719-2] [Medline: 30594159]
81. Safran C. Update on data reuse in health care. Yearb Med Inform 2017 Aug;26(1):24-27 [FREE Full text] [doi: 10.15265/IY-2017-013] [Medline: 29063535]
82. Severo M, Feredj A, Romele A. Soft data and public policy: can social media offer alternatives to official statistics in urban policymaking? Policy Internet 2016 Jul 19;8(3):354-372. [doi: 10.1002/poi3.127]
83. Shakeri Hossein Abad Z, Butler GP, Thompson W, Lee J. Physical activity, sedentary behavior, and sleep on Twitter: multicountry and fully labeled public data set for digital public health surveillance research. JMIR Public Health Surveill 2022 Feb 14;8(2):e32355 [FREE Full text] [doi: 10.2196/32355] [Medline: 35156938]
84. Singh T, Roberts K, Cohen T, Cobb N, Wang J, Fujimoto K, et al. Social media as a research tool (SMaaRT) for risky behavior analytics: methodological review. JMIR Public Health Surveill 2020 Nov 30;6(4):e21660 [FREE Full text] [doi: 10.2196/21660] [Medline: 33252345]
85. Zins M, Cuggia M, Goldberg M. Les données de santé en France - Abondantes mais complexes. Med Sci (Paris) 2021 Feb 16;37(2):179-184 [FREE Full text] [doi: 10.1051/medsci/2021001] [Medline: 33591261]
86. Motahari-Nezhad H, Fgaier M, Mahdi Abid M, Péntek M, Gulácsi L, Zrubka Z. Digital biomarker-based studies: scoping review of systematic reviews. JMIR Mhealth Uhealth 2022 Oct 24;10(10):e35722 [FREE Full text] [doi: 10.2196/35722] [Medline: 36279171]
87. Ayilara OF, Platt RW, Dahl M, Coulombe J, Ginestet PG, Chateau D, et al. Generating synthetic data from administrative health records for drug safety and effectiveness studies. Int J Popul Data Sci 2023 Nov 27;8(1):2176 [FREE Full text] [doi: 10.23889/ijpds.v8i1.2176] [Medline: 38414538]
88. Fadel M, Petot J, Gourraud PA, Descatha A. Flexibility of a large blindly synthetized avatar database for occupational research: example from the CONSTANCES cohort for stroke and knee pain. PLoS One 2024 Jul 31;19(7):e0308063 [FREE Full text] [doi: 10.1371/journal.pone.0308063] [Medline: 39083487]
89. Giuffrè M, Shung DL. Harnessing the power of synthetic data in healthcare: innovation, application, and privacy. NPJ Digit Med 2023 Oct 09;6(1):186 [FREE Full text] [doi: 10.1038/s41746-023-00927-3] [Medline: 37813960]

XSL•FO
RenderX

90. Kamel Boulos MN, Zhang P. Digital twins: from personalised medicine to precision public health. J Pers Med 2021 Jul 29;11(8):745 [FREE Full text] [doi: 10.3390/jpm11080745] [Medline: 34442389]
91. Vallevik VB, Babic A, Marshall SE, Elvatun S, Brøgger HB, Alagaratnam S, et al. Can I trust my fake data - a comprehensive quality assessment framework for synthetic tabular data in healthcare. Int J Med Inform 2024 May;185:105413 [FREE Full text] [doi: 10.1016/j.ijmedinf.2024.105413] [Medline: 38493547]
92. Chamot S, Manaouil C, Fantoni S, Frimat P, Al-Salameh A, Desailloud R, et al. Système national des données de santé (SNDS) : perspectives en santé au travail et environnementale. Arch Mal Prof Environ 2023 Nov;84(6):101875. [doi: 10.1016/j.admp.2023.101875]
93. Cohen S, Gilutz H, Marelli AJ, Iserin L, Benis A, Bonnet D, et al. Administrative health databases for addressing emerging issues in adults with CHD: a systematic review. Cardiol Young 2018 Jun 29;28(6):844-853. [doi: 10.1017/S1047951118000446] [Medline: 29704902]
94. Fraser HS, Mugisha M, Bacher I, Ngenzi JL, Seebregts C, Umubyeyi A, et al. Factors influencing data quality in electronic health record systems in 50 health facilities in Rwanda and the role of clinical alerts: cross-sectional observational study. JMIR Public Health Surveill 2024 Jul 03;10:e49127 [FREE Full text] [doi: 10.2196/49127] [Medline: 38959048]
95. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. J Epidemiol Community Health 2014 Mar;68(3):283-287. [doi: 10.1136/jech-2013-202744] [Medline: 24248997]
96. Hinds A, Lix LM, Smith M, Quan H, Sanmartin C. Quality of administrative health databases in Canada: a scoping review. Can J Public Health 2016 Jun 27;107(1):e56-e61 [FREE Full text] [doi: 10.17269/cjph.107.5244] [Medline: 27348111]
97. Howell CR, Zhang L, Clay OJ, Dutton G, Horton T, Mugavero MJ, et al. Social determinants of health phenotypes and cardiometabolic condition prevalence among patients in a large academic health system: latent class analysis. JMIR Public Health Surveill 2024 Aug 07;10:e53371 [FREE Full text] [doi: 10.2196/53371] [Medline: 39113389]
98. Jones KH, Heys S, Tingay KS, Jackson P, Dibben C. The good, the bad, the clunky: improving the use of administrative data for research. Int J Popul Data Sci 2018 Dec 04;4(1):587 [FREE Full text] [doi: 10.23889/ijpds.v4i1.587] [Medline: 32935024]
99. Namulanda G, Qualters J, Vaidyanathan A, Roberts E, Richardson M, Fraser A, et al. Electronic health record case studies to advance environmental public health tracking. J Biomed Inform 2018 Mar;79:98-104 [FREE Full text] [doi: 10.1016/j.jbi.2018.02.012] [Medline: 29476967]
100. Nicholls SG, Quach P, von Elm E, Guttmann A, Moher D, Petersen I, et al. The REporting of studies conducted using observational routinely-collected health data (RECORD) statement: methods for arriving at consensus and developing reporting guidelines. PLoS One 2015;10(5):e0125620 [FREE Full text] [doi: 10.1371/journal.pone.0125620] [Medline: 25965407]
101. Silverwood R, Rajah N, Calderwood L, De Stavola B, Harron K, Ploubidis G. Examining the quality and population representativeness of linked survey and administrative data: guidance and illustration using linked 1958 National Child Development Study and Hospital Episode Statistics data. Int J Popul Data Sci 2024 Jan 09;9(1):2137 [FREE Full text] [doi: 10.23889/ijpds.v9i1.2137] [Medline: 38425790]
102. Zuiderwijk A, Türk BO, Brazier F. Identifying the most important facilitators of open research data sharing and reuse in Epidemiology: a mixed-methods study. PLoS One 2024 Feb 8;19(2):e0297969 [FREE Full text] [doi: 10.1371/journal.pone.0297969] [Medline: 38330007]
103. Hobensack M, Song J, Scharp D, Bowles KH, Topaz M. Machine learning applied to electronic health record data in home healthcare: a scoping review. Int J Med Inform 2023 Feb;170:104978 [FREE Full text] [doi: 10.1016/j.ijmedinf.2022.104978] [Medline: 36592572]
104. Tassi MF, le Meur N, Stéfic K, Grammatico-Guillon L. Performance of French medico-administrative databases in epidemiology of infectious diseases: a scoping review. Front Public Health 2023 May 12;11:1161550 [FREE Full text] [doi: 10.3389/fpubh.2023.1161550] [Medline: 37250067]
105. Yoshida S, Takeuchi M, Tanaka-Mizuno S, Mizuno K, Nakashima M, Fukasawa T, et al. Clinical epidemiology and pharmacoepidemiology studies with real-world databases. Proc Jpn Acad Ser B Phys Biol Sci 2022;98(10):517-528 [FREE Full text] [doi: 10.2183/pjab.98.026] [Medline: 36504194]
106. Howren A, Aviña-Zubieta JA, Puyat JH, Esdaile JM, Da Costa D, De Vera MA. Defining depression and anxiety in individuals with rheumatic diseases using administrative health databases: a systematic review. Arthritis Care Res (Hoboken) 2020 Feb 09;72(2):243-255. [doi: 10.1002/acr.24048] [Medline: 31421021]
107. Kim MK, Rouphael C, McMichael J, Welch N, Dasarathy S. Challenges in and opportunities for electronic health record-based data analysis and interpretation. Gut Liver 2024 Mar 15;18(2):201-208 [FREE Full text] [doi: 10.5009/gnl230272] [Medline: 37905424]
108. Lee S, Xu Y, D Apos Souza AG, Martin EA, Doktorchik C, Zhang Z, et al. Unlocking the potential of electronic health records for health research. Int J Popul Data Sci 2020 Jan 30;5(1):1123 [FREE Full text] [doi: 10.23889/ijpds.v5i1.1123] [Medline: 32935049]
109. Pletscher F, Mändli Lerch K, Glinz D. Willingness to share anonymised routinely collected clinical health data in Switzerland: a cross-sectional survey. Swiss Med Wkly 2022 Jun 20;152(2324):w30182 [FREE Full text] [doi: 10.4414/smw.2022.w30182] [Medline: 35752970]

XSL•FO
RenderX

110. Yuan H, Ali MS, Brouwer ES, Girman CJ, Guo JJ, Lund JL, ISPE Comparative Effectiveness Research Special Interest Group. Real-world evidence: what it is and what it can tell us according to the International Society for Pharmacoepidemiology (ISPE) Comparative Effectiveness Research (CER) Special Interest Group (SIG). Clin Pharmacol Ther 2018 Aug 07;104(2):239-241. [doi: 10.1002/cpt.1086] [Medline: 29733448]
111. Andrew A, Zhou J, Gui J, Harrison A, Shi X, Li M, et al. Pesticides applied to crops and amyotrophic lateral sclerosis risk in the U.S. Neurotoxicology 2021 Dec;87:128-135 [FREE Full text] [doi: 10.1016/j.neuro.2021.09.004] [Medline: 34562505]
112. Chazard E, Ficheur G, Caron A, Lamer A, Labreuche J, Cuggia M, et al. Secondary use of healthcare structured data: the challenge of domain-knowledge based extraction of features. Stud Health Technol Inform 2018;255:15-19. [Medline: 30306898]
113. Grath-Lone LM, Jay MA, Blackburn R, Gordon E, Zylbersztejn A, Wiljaars L, et al. What makes administrative data "research-ready"? A systematic review and thematic analysis of published literature. Int J Popul Data Sci 2022;7(1):1718 [FREE Full text] [doi: 10.23889/ijpds.v6i1.1718] [Medline: 35520099]
114. Hashimoto RE, Brodt ED, Skelly AC, Dettori JR. Administrative database studies: goldmine or goose chase? Evid Based Spine Care J 2014 Oct 24;5(2):74-76 [FREE Full text] [doi: 10.1055/s-0034-1390027] [Medline: 25278880]
115. Bezin J, Duong M, Lassalle R, Droz C, Pariente A, Blin P, et al. The national healthcare system claims databases in France, SNIIRAM and EGB: powerful tools for pharmacoepidemiology. Pharmacoepidemiol Drug Saf 2017 Aug 24;26(8):954-962. [doi: 10.1002/pds.4233] [Medline: 28544284]
116. Al-Sahab B, Leviton A, Loddenkemper T, Paneth N, Zhang B. Biases in electronic health records data for generating real-world evidence: an overview. J Healthc Inform Res 2024 Mar 14;8(1):121-139. [doi: 10.1007/s41666-023-00153-2] [Medline: 38273982]
117. Bastarache L, Brown JS, Cimino JJ, Dorr DA, Embi PJ, Payne PR, et al. Developing real-world evidence from real-world data: transforming raw data into analytical datasets. Learn Health Syst 2022 Jan 14;6(1):e10293 [FREE Full text] [doi: 10.1002/lrh2.10293] [Medline: 35036557]
118. Fang J, Kapral MK, Richards J, Robertson A, Stamplecoski M, Silver FL. The Registry of Canadian Stroke Network : an evolving methodology. Acta Neurol Taiwan 2011 Jun;20(2):77-84. [Medline: 21739386]
119. Fillinger S, de la Garza L, Peltzer A, Kohlbacher O, Nahnsen S. Challenges of big data integration in the life sciences. Anal Bioanal Chem 2019 Oct 28;411(26):6791-6800. [doi: 10.1007/s00216-019-02074-9] [Medline: 31463515]
120. Fonteneau L, Le Meur N, Cohen-Akenine A, Pessel C, Brouard C, Delon F, et al. [The use of administrative health databases in infectious disease epidemiology and public health]. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S174-S182 [FREE Full text] [doi: 10.1016/j.respe.2017.03.131] [Medline: 28624133]
121. Silenou BC, Avalos M, Helmer C, Berr C, Pariente A, Jacqmin-Gadda H. Health administrative data enrichment using cohort information: comparative evaluation of methods by simulation and application to real data. PLoS One 2019 Jan 31;14(1):e0211118 [FREE Full text] [doi: 10.1371/journal.pone.0211118] [Medline: 30703112]
122. Wang SV, Schneeweiss S. Data checks before registering study protocols for health care database analyses. JAMA 2024 May 07;331(17):1445-1446. [doi: 10.1001/jama.2024.2988] [Medline: 38587830]
123. Yamana H, Yamasaki K, Jo T, Yatsuhashi H, Yasunaga H. A descriptive analysis of acute viral hepatitis using a database with electronic medical records and claims data. Ann Clin Epidemiol 2023;5(4):107-112 [FREE Full text] [doi: 10.37737/ace.23014] [Medline: 38504953]
124. Gianfrancesco MA, Goldstein ND. A narrative review on the validity of electronic health record-based research in epidemiology. BMC Med Res Methodol 2021 Oct 27;21(1):234 [FREE Full text] [doi: 10.1186/s12874-021-01416-5] [Medline: 34706667]
125. Denisov NS, Kamenskikh EM, Fedorova OS. Trends in population-based studies: molecular and digital epidemiology (review). Sovrem Tekhnologii Med 2022 Aug 29;14(4):60-70 [FREE Full text] [doi: 10.17691/stm2022.14.4.07] [Medline: 37179982]
126. Hamm NC, Hamad AF, Wall-Wieler E, Roos LL, Plana-Ripoll O, Lix LM. Multigenerational health research using population-based linked databases: an international review. Int J Popul Data Sci 2021 Oct 07;6(1):1686 [FREE Full text] [doi: 10.23889/ijpds.v6i1.1686] [Medline: 34734126]
127. Poulain R, Tarek MF, Beheshti R. Improving fairness in AI models on electronic health records: the case for federated learning methods. FAccT 23 (2023) 2023 Jun;2023:1599-1608 [FREE Full text] [doi: 10.1145/3593013.3594102] [Medline: 37990734]
128. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: from the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S149-S167. [doi: 10.1016/j.respe.2017.05.004] [Medline: 28756037]
129. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. Annu Rev Public Health 2016;37:61-81 [FREE Full text] [doi: 10.1146/annurev-publhealth-032315-021353] [Medline: 26667605]

XSL•FO
RenderX

130. Chan JM, Carroll MW, Smyth M, Hamilton Z, Evans D, McGrail K, et al. Comparing health administrative and clinical registry data: trends in incidence and prevalence of pediatric inflammatory bowel disease in British Columbia. Clin Epidemiol 2021;13:81-90 [FREE Full text] [doi: 10.2147/CLEP.S292546] [Medline: 33603489]
131. Chowdhury T, Hemmelgarn B. Evidence-based decision-making 6: utilization of administrative databases for health services research. In: Parfrey PS, Barrett BJ, editors. Clinical Epidemiology: Practice and Methods. Cham, Switzerland: Springer; 2015:469-484.
132. Chishtie J, Sapiro N, Wiebe N, Rabatach L, Lorenzetti D, Leung AA, et al. Use of epic electronic health record system for health care research: scoping review. J Med Internet Res 2023 Dec 15;25:e51003 [FREE Full text] [doi: 10.2196/51003] [Medline: 38100185]
133. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health 2011;32:91-108. [doi: 10.1146/annurev-publhealth-031210-100700] [Medline: 21219160]
134. Mazzali C, Duca P. Use of administrative data in healthcare research. Intern Emerg Med 2015 Jun 25;10(4):517-524. [doi: 10.1007/s11739-015-1213-9] [Medline: 25711312]
135. McPheeters ML, Sathe NA, Jerome RN, Carnahan RM. Methods for systematic reviews of administrative database studies capturing health outcomes of interest. Vaccine 2013 Dec 30;31 Suppl 10:K2-K6. [doi: 10.1016/j.vaccine.2013.06.048] [Medline: 24331070]
136. Sato S, Yasunaga H. A review of studies using Japanese nationwide administrative claims databases. Ann Clin Epidemiol 2023;5(2):58-64 [FREE Full text] [doi: 10.37737/ace.23008] [Medline: 38505730]
137. Mues KE, Liede A, Liu J, Wetmore J, Zaha R, Bradbury BD, et al. Use of the Medicare database in epidemiologic and health services research: a valuable source of real-world evidence on the older and disabled populations in the US. Clin Epidemiol 2017;9:267-277 [FREE Full text] [doi: 10.2147/CLEP.S105613] [Medline: 28533698]
138. Bernstein CN. Large registry epidemiology in IBD. Inflamm Bowel Dis 2017 Nov;23(11):1941-1949. [doi: 10.1097/MIB.0000000000001279] [Medline: 28991858]
139. Bloom PP, Tapper EB. The use of administrative data to investigate the population burden of hepatic encephalopathy. J Clin Med 2020 Nov 10;9(11):3620 [FREE Full text] [doi: 10.3390/jcm9113620] [Medline: 33182743]
140. Hudson M, Tascilar K, Suissa S. Comparative effectiveness research with administrative health data in rheumatoid arthritis. Nat Rev Rheumatol 2016 Jun 15;12(6):358-366. [doi: 10.1038/nrrheum.2016.34] [Medline: 27080692]
141. Mitchell E, Waring T, Ahern E, O'Donovan D, O'Reilly D, Bradley DT. Predictors and consequences of homelessness in whole-population observational studies that used administrative data: a systematic review. BMC Public Health 2023 Aug 24;23(1):1610 [FREE Full text] [doi: 10.1186/s12889-023-16503-z] [Medline: 37612701]
142. Ulrich EH, So G, Zappitelli M, Chanchlani R. A review on the application and limitations of administrative health care data for the study of acute kidney injury epidemiology and outcomes in children. Front Pediatr 2021 Oct 27;9:742888 [FREE Full text] [doi: 10.3389/fped.2021.742888] [Medline: 34778133]
143. Curl CL, Spivak M, Phinney R, Montrose L. Synthetic pesticides and health in vulnerable populations: agricultural workers. Curr Environ Health Rep 2020 Mar 20;7(1):13-29 [FREE Full text] [doi: 10.1007/s40572-020-00266-5] [Medline: 31960353]
144. Demos K, Sazakli E, Jelastopulu E, Charokopos N, Ellul J, Leotsinidis M. Does farming have an effect on health status? A comparison study in west Greece. Int J Environ Res Public Health 2013 Feb 26;10(3):776-792 [FREE Full text] [doi: 10.3390/ijerph10030776] [Medline: 23442558]
145. Kachuri L, Harris MA, MacLeod JS, Tjepkema M, Peters PA, Demers PA. Cancer risks in a population-based study of 70,570 agricultural workers: results from the Canadian census health and Environment cohort (CanCHEC). BMC Cancer 2017 May 19;17(1):343 [FREE Full text] [doi: 10.1186/s12885-017-3346-x] [Medline: 28525996]
146. Ramos AK, Girdžiūtė L, Starič J, Rautianinen RH. Identifying "vulnerable agricultural populations" at risk for occupational injuries and illnesses: a European perspective. J Agromedicine 2021 Jul 27;26(3):340-345. [doi: 10.1080/1059924X.2020.1771498] [Medline: 32458756]
147. Sturm ET, Castro C, Mendez-Colmenares A, Duffy J, Burzynska AZ, Stallones L, et al. Risk factors for brain health in agricultural work: a systematic review. Int J Environ Res Public Health 2022 Mar 13;19(6):3373 [FREE Full text] [doi: 10.3390/ijerph19063373] [Medline: 35329061]
148. Guindo Y, Parent ME, Richard H, Luce D, Barul C. Expert-based assessment of chemical and physical exposures, and organizational factors, in past agricultural jobs. Environ Res 2024 Dec 15;263(Pt 3):120238 [FREE Full text] [doi: 10.1016/j.envres.2024.120238] [Medline: 39461702]
149. Komarek AM, De Pinto A, Smith VH. A review of types of risks in agriculture: what we know and what we need to know. Agric Syst 2020 Feb;178:102738. [doi: 10.1016/j.agsy.2019.102738]
150. Leon ME, Schinasi LH, Lebailly P, Beane Freeman LE, Nordby KC, Ferro G, et al. Pesticide use and risk of non-Hodgkin lymphoid malignancies in agricultural cohorts from France, Norway and the USA: a pooled analysis from the AGRICOH consortium. Int J Epidemiol 2019 Oct 01;48(5):1519-1535 [FREE Full text] [doi: 10.1093/ije/dyz017] [Medline: 30880337]
151. Pérès K, Matharan F, Allard M, Amieva H, Baldi I, Barberger-Gateau P, et al. Health and aging in elderly farmers: the AMI cohort. BMC Public Health 2012 Jul 27;12(1):558 [FREE Full text] [doi: 10.1186/1471-2458-12-558] [Medline: 22839764]

XSL•FO
RenderX

152. Petit P, Chamot S, Al-Salameh A, Cancé C, Desailloud R, Bonneterre V. Farming activity and risk of treated thyroid disorders: insights from the TRACTOR project, a nationwide cohort study. Environ Res 2024 May 15;249:118458 [FREE Full text] [doi: 10.1016/j.envres.2024.118458] [Medline: 38365059]
153. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
154. The Joanna Briggs Institute Reviewers' Manual 2015 methodology for JBI scoping reviews. The Joanna Briggs Institute. 2015. URL: https://reben.com.br/revista/wp-content/uploads/2020/10/Scoping.pdf [accessed 2024-04-03]
155. Beydon M, Pinto S, De Rycke Y, Fautrel B, Mariette X, Seror R, et al. Risk of cancer for patients with rheumatoid arthritis versus general population: a national claims database cohort study. Lancet Reg Health Eur 2023 Dec;35:100768 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100768] [Medline: 37954220]
156. Moysan M. Des archives à la base de données : valeur(s) des données en épidémiologie. Hist Méd Santé 2022 Dec 15(22):151-164 [FREE Full text] [doi: 10.4000/hms.6274]
157. Syed R, Eden R, Makasi T, Chukwudi I, Mamudu A, Kamalpour M, et al. Digital health data quality issues: systematic review. J Med Internet Res 2023 Mar 31;25:e42615 [FREE Full text] [doi: 10.2196/42615] [Medline: 37000497]
158. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F. An approach for detecting, quantifying, and visualizing the evolution of a research field: a practical application to the Fuzzy Sets Theory field. J Informetr 2011 Jan;5(1):146-166. [doi: 10.1016/j.joi.2010.10.002]
159. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res 2021 Sep;133:285-296. [doi: 10.1016/j.jbusres.2021.04.070]
160. Moral-Muñoz JA, Herrera-Viedma E, Santisteban-Espejo A, Cobo MJ. Software tools for conducting bibliometric analysis in science: an up-to-date review. Prof Inf J 2020 Jan 19;29(1) [FREE Full text] [doi: 10.3145/epi.2020.ene.03]
161. Bienert IR, Oliveira RC, Andrade PB, Caramori CA. Bibliometric indexes, databases and impact factors in cardiology. Rev Bras Cir Cardiovasc 2015;30(2):254-259 [FREE Full text] [doi: 10.5935/1678-9741.20150019] [Medline: 26107458]
162. Bihari A, Tripathi S, Deepak A. A review on h-index and its alternative indices. J Inf Sci 2021 May 31;49(3):624-665. [doi: 10.1177/01655515211014478]
163. Shao Y, Chien TW, Jang FL. The use of radar plots with the Yk-index to identify which authors contributed the most to the journal of Medicine in 2020 and 2021: a bibliometric analysis. Medicine (Baltimore) 2022 Nov 11;101(45):e31033 [FREE Full text] [doi: 10.1097/MD.0000000000031033] [Medline: 36397440]
164. Kumar S, Kumar S. Trends of collaborative research in journals of oilseeds research (India), 1993-2004. Indian J Agric Library Inf Services 2008;24:90.
165. Wu T, Duan Y, Zhang T, Tian W, Liu H, Deng Y. Research trends in the application of artificial intelligence in oncology: a bibliometric and network visualization study. Front Biosci (Landmark Ed) 2022 Aug 31;27(9):254 [FREE Full text] [doi: 10.31083/j.fbl2709254] [Medline: 36224012]
166. Wang B, Zhao P. Worldwide research productivity in the field of back pain: a bibliometric analysis. Medicine (Baltimore) 2018 Oct;97(40):e11566 [FREE Full text] [doi: 10.1097/MD.0000000000011566] [Medline: 30290588]
167. Aria M, Cuccurullo C, D'Aniello L, Misuraca M, Spano M. Thematic analysis as a new culturomic tool: the social media coverage on COVID-19 pandemic in Italy. Sustainability 2022 Mar 20;14(6):3643. [doi: 10.3390/su14063643]
168. Petit P, Leroyer A, Chamot S, Fumery M, Bonneterre V. Farming activities and risk of inflammatory bowel disease: a French nationwide population-based cohort study. J Crohns Colitis 2024 Sep 03;18(9):1415-1429. [doi: 10.1093/ecco-jcc/jjae050] [Medline: 38605515]
169. Montazeri A, Mohammadi S, M Hesari P, Ghaemi M, Riazi H, Sheikhi-Mobarakeh Z. Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): a minimum requirements. Syst Rev 2023 Dec 15;12(1):239 [FREE Full text] [doi: 10.1186/s13643-023-02410-2] [Medline: 38102710]
170. Aria M, Cuccurullo C. bibliometrix : an R-tool for comprehensive science mapping analysis. J Informetr 2017 Nov;11(4):959-975. [doi: 10.1016/j.joi.2017.08.007]
171. Spitzer WO, Hill GB, Chambers LW, Helliwell BE, Murphy HB. The occupation of fishing as a risk factor in cancer of the lip. N Engl J Med 1975 Aug 28;293(9):419-424. [doi: 10.1056/nejm197508282930903]
172. Alavanja MC, Samanic C, Dosemeci M, Lubin J, Tarone R, Lynch CF, et al. Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. Am J Epidemiol 2003 May 01;157(9):800-814. [doi: 10.1093/aje/kwg040] [Medline: 12727674]
173. Alavanja MC, Dosemeci M, Samanic C, Lubin J, Lynch CF, Knott C, et al. Pesticides and lung cancer risk in the agricultural health study cohort. Am J Epidemiol 2004 Nov 01;160(9):876-885. [doi: 10.1093/aje/kwh290] [Medline: 15496540]
174. Andersen A, Barlow L, Engeland A, Kjaerheim K, Lynge E, Pukkala E. Work-related cancer in the Nordic countries. Scand J Work Environ Health 1999;25 Suppl 2:1-116 [FREE Full text] [Medline: 10507118]
175. Andreotti G, Koutros S, Hofmann JN, Sandler DP, Lubin JH, Lynch CF, et al. Glyphosate use and cancer incidence in the agricultural health study. J Natl Cancer Inst 2018 May 01;110(5):509-516 [FREE Full text] [doi: 10.1093/jnci/djx233] [Medline: 29136183]

XSL•FO
RenderX

176. Elbaz A, Clavel J, Rathouz PJ, Moisan F, Galanaud J, Delemotte B, et al. Professional exposure to pesticides and Parkinson disease. Ann Neurol 2009 Oct;66(4):494-504. [doi: 10.1002/ana.21717] [Medline: 19847896]
177. Engel LS, Hill DA, Hoppin JA, Lubin JH, Lynch CF, Pierce J, et al. Pesticide use and breast cancer risk among farmers' wives in the agricultural health study. Am J Epidemiol 2005 Jan 15;161(2):121-135. [doi: 10.1093/aje/kwi022] [Medline: 15632262]
178. Eriksson M, Karlsson M. Occupational and other environmental factors and multiple myeloma: a population based case-control study. Br J Ind Med 1992 Feb 01;49(2):95-103 [FREE Full text] [doi: 10.1136/oem.49.2.95] [Medline: 1536825]
179. Hemminki K, Saloniemi I, Salonen T, Partanen T, Vainio H. Childhood cancer and parental occupation in Finland. J Epidemiol Community Health 1981 Mar 01;35(1):11-15 [FREE Full text] [doi: 10.1136/jech.35.1.11] [Medline: 7264527]
180. Karjalainen A, Kurppa K, Martikainen R, Klaukka T, Karjalainen J. Work is related to a substantial portion of adult-onset asthma incidence in the Finnish population. Am J Respir Crit Care Med 2001 Aug 15;164(4):565-568. [doi: 10.1164/ajrccm.164.4.2012146] [Medline: 11520716]
181. Karjalainen A, Martikainen R, Klaukka T, Saarinen K, Uitti J. Risk of asthma among Finnish patients with occupational rhinitis. Chest 2003 Jan;123(1):283-288. [doi: 10.1378/chest.123.1.283] [Medline: 12527633]
182. Kristensen P, Andersen A, Irgens LM, Laake P, Bye AS. Incidence and risk factors of cancer among men and women in Norwegian agriculture. Scand J Work Environ Health 1996 Feb;22(1):14-26 [FREE Full text] [doi: 10.5271/sjweh.104] [Medline: 8685669]
183. Lee WJ, Sandler DP, Blair A, Samanic C, Cross AJ, Alavanja MC. Pesticide use and colorectal cancer risk in the Agricultural Health Study. Int J Cancer 2007 Jul 15;121(2):339-346 [FREE Full text] [doi: 10.1002/ijc.22635] [Medline: 17390374]
184. Levecque C, Elbaz A, Clavel J, Richard F, Vidal JS, Amouyel P, et al. Association between Parkinson's disease and polymorphisms in the nNOS and iNOS genes in a community-based case-control study. Hum Mol Genet 2003 Jan 01;12(1):79-86. [doi: 10.1093/hmg/ddg009] [Medline: 12490535]
185. Mills PK, Yang R. Prostate cancer risk in California farm workers. J Occup Environ Med 2003 Mar;45(3):249-258. [doi: 10.1097/01.jom.0000058339.05741.0c] [Medline: 12661182]
186. Morrison H, Savitz D, Semenciw R, Hulka B, Mao Y, Morison D, et al. Farming and prostate cancer mortality. Am J Epidemiol 1993 Feb 01;137(3):270-280. [doi: 10.1093/oxfordjournals.aje.a116674] [Medline: 8452135]
187. Orton SM, Wald L, Confavreux C, Vukusic S, Krohn J, Ramagopalan S, et al. Association of UV radiation with multiple sclerosis prevalence and sex ratio in France. Neurology 2011 Feb;76(5):425-431. [doi: 10.1212/wnl.0b013e31820a0a9f]
188. Pouchieu C, Piel C, Carles C, Gruber A, Helmer C, Tual S, et al. Pesticide use in agriculture and Parkinson's disease in the AGRICAN cohort study. Int J Epidemiol 2018 Feb 01;47(1):299-310. [doi: 10.1093/ije/dyx225] [Medline: 29136149]
189. Pukkala E, Martinsen JI, Lynge E, Gunnarsdottir HK, Sparén P, Tryggvadottir L, et al. Occupation and cancer - follow-up of 15 million people in five Nordic countries. Acta Oncol 2009 Jul 21;48(5):646-790. [doi: 10.1080/02841860902913546] [Medline: 19925375]
190. Reif J, Pearce N, Fraser J. Cancer risks in New Zealand farmers. Int J Epidemiol 1989 Dec;18(4):768-774. [doi: 10.1093/ije/18.4.768] [Medline: 2621012]
191. Stokes L, Stark A, Marshall E, Narang A. Neurotoxicity among pesticide applicators exposed to organophosphates. Occup Environ Med 1995 Oct 01;52(10):648-653 [FREE Full text] [doi: 10.1136/oem.52.10.648] [Medline: 7489054]
192. Vingård E, Alfredsson L, Goldie I, Hogstedt C. Occupation and osteoarthrosis of the hip and knee: a register-based cohort study. Int J Epidemiol 1991 Dec;20(4):1025-1031. [doi: 10.1093/ije/20.4.1025] [Medline: 1800399]
193. Vukusic S, Van Bockstael V, Gosselin S, Confavreux C. Regional variations in the prevalence of multiple sclerosis in French farmers. J Neurol Neurosurg Psychiatry 2007 Jul 18;78(7):707-709 [FREE Full text] [doi: 10.1136/jnnp.2006.101196] [Medline: 17299020]
194. Weichenthal S, Villeneuve PJ, Burnett RT, van Donkelaar A, Martin RV, Jones RR, et al. Long-term exposure to fine particulate matter: association with nonaccidental and cardiovascular mortality in the agricultural health study cohort. Environ Health Perspect 2014 Jun;122(6):609-615 [FREE Full text] [doi: 10.1289/ehp.1307277] [Medline: 24633320]
195. Wigle DT, Semenciw RM, Wilkins K, Riedel D, Ritter L, Morrison HI, et al. Mortality study of Canadian male farm operators: non-Hodgkin's lymphoma mortality and agricultural practices in Saskatchewan. J Natl Cancer Inst 1990 Apr 04;82(7):575-582. [doi: 10.1093/jnci/82.7.575] [Medline: 2313734]
196. Wiklund K, Dich J, Holm LE, Eklund G. Risk of cancer in pesticide applicators in Swedish agriculture. Br J Ind Med 1989 Nov 01;46(11):809-814 [FREE Full text] [doi: 10.1136/oem.46.11.809] [Medline: 2590647]
197. Horwitz H, Ahlgren B, Nærum E. Effect of occupation on risk of developing MS: an insurance cohort study. BMJ Open 2013 Jun 20;3(6):e002894 [FREE Full text] [doi: 10.1136/bmjopen-2013-002894] [Medline: 23794592]
198. Karttunen JP, Rautiainen RH, Quendler E. Gender division of farm work and occupational injuries. J Agric Saf Health 2019 Jul 16;25(3):117-127. [doi: 10.13031/jash.13177] [Medline: 32429612]
199. Keller JE, Howe HL. Risk factors for lung cancer among nonsmoking Illinois residents. Environ Res 1993 Jan;60(1):1-11. [doi: 10.1006/enrs.1993.1001] [Medline: 8432260]
200. Marant Micallef C, Shield KD, Vignat J, Baldi I, Charbotel B, Fervers B, et al. Cancers in France in 2015 attributable to occupational exposures. Int J Hyg Environ Health 2019 Jan;222(1):22-29. [doi: 10.1016/j.ijheh.2018.07.015] [Medline: 30174219]

XSL•FO
RenderX

201. Moisan F, Spinosi J, Dupupet JL, Delabre L, Mazurie J, Goldberg M, et al. The relation between type of farming and prevalence of Parkinson's disease among agricultural workers in five French districts. Mov Disord 2011 Feb 01;26(2):271-279 [FREE Full text] [doi: 10.1002/mds.23370] [Medline: 21412834]
202. Toch-Marquardt M, Menvielle G, Eikemo TA, Kulhánová I, Kulik MC, Bopp M, Euro-GBD-SE consortium. Occupational class inequalities in all-cause and cause-specific mortality among middle-aged men in 14 European populations during the early 2000s. PLoS One 2014 Sep 30;9(9):e108072 [FREE Full text] [doi: 10.1371/journal.pone.0108072] [Medline: 25268702]
203. Tüchsen F, Andersen O, Costa G, Filakti H, Marmot MG. Occupation and ischemic heart disease in the European Community: a comparative study of occupations at potential high risk. Am J Ind Med 1996 Oct;30(4):407-414. [doi: 10.1002/(SICI)1097-0274(199610)30:4<407::AID-AJIM5>3.0.CO;2-S] [Medline: 8892545]
204. Wang CC, Lin GL, Lin YJ, Chen WL, Wu WT. Occupational health surveillance and detection of emerging occupational diseases among Taiwan farmers, through analysis of national-based farmers' and medico-administrative databases. Am J Ind Med 2023 Jan 17;66(1):85-93. [doi: 10.1002/ajim.23443] [Medline: 36398405]
205. Langholz PL, Sommerseth HL. Old-age mortality and social class in northern Norway in the first half of the twentieth century. Scand Econ Hist Rev 2023 Sep 13;72(2):153-170. [doi: 10.1080/03585522.2023.2255599]
206. Das AV, Bagga B, Joseph J. Clinical profile and demographic distribution of keratitis: an electronic medical record-driven data analytics from an eye care network in India. Ocul Immunol Inflamm 2024 Jul 21;32(5):722-726. [doi: 10.1080/09273948.2023.2199331] [Medline: 37084281]
207. Lemarchand C, Tual S, Levêque-Morlais N, Perrier S, Belot A, Velten M, AGRICAN group. Cancer incidence in the AGRICAN cohort study (2005-2011). Cancer Epidemiol 2017 Aug;49:175-185 [FREE Full text] [doi: 10.1016/j.canep.2017.06.003] [Medline: 28709062]
208. Scott E, Hirabayashi L, Levenstein A, Krupa N, Jenkins P. The development of a machine learning algorithm to identify occupational injuries in agriculture using pre-hospital care reports. Health Inf Sci Syst 2021 Dec 29;9(1):31 [FREE Full text] [doi: 10.1007/s13755-021-00161-9] [Medline: 34422257]
209. Wilke RA, Qamar M, Lupu RA, Gu S, Zhao J. Chronic kidney disease in agricultural communities. Am J Med 2019 Oct;132(10):e727-e732 [FREE Full text] [doi: 10.1016/j.amjmed.2019.03.036] [Medline: 30998912]
210. Engel LS, Werder E, Satagopan J, Blair A, Hoppin JA, Koutros S, et al. Insecticide use and breast cancer risk among farmers' wives in the agricultural health study. Environ Health Perspect 2017 Sep 06;125(9):097002 [FREE Full text] [doi: 10.1289/EHP1295] [Medline: 28934092]
211. Louis LM, Lerro CC, Friesen MC, Andreotti G, Koutros S, Sandler DP, et al. A prospective study of cancer risk among agricultural health study farm spouses associated with personal use of organochlorine insecticides. Environ Health 2017 Sep 06;16(1):95 [FREE Full text] [doi: 10.1186/s12940-017-0298-1] [Medline: 28874165]
212. Elbaek Pedersen J, Hansen J. Risk of breast cancer in daughters of agricultural workers in Denmark. Environ Res 2024 Jan 01;240(Pt 1):117374. [doi: 10.1016/j.envres.2023.117374] [Medline: 37866542]
213. Kasten-Arias C, Hodes T, Marino M, Kaufmann J, Lucas JA, Estela Vasquez Guzman C, et al. Healthcare utilization for asthma exacerbation among children of migrant and seasonal farmworkers. Prev Med Rep 2024 Feb;38:102598 [FREE Full text] [doi: 10.1016/j.pmedr.2024.102598] [Medline: 38283959]
214. Mengist Dessie A, Fenta Feleke S, Getaye Workie S, Getinet Abebe T, Mossu Chanie Y, Kassa Yalew A. Prevalence of skin disease and its associated factors among primary schoolchildren: a cross-sectional study from a Northern Ethiopian town. Clin Cosmet Investig Dermatol 2022;15:791-801 [FREE Full text] [doi: 10.2147/CCID.S361051] [Medline: 35521561]
215. Mills PK, Zahm SH. Organophosphate pesticide residues in urine of farmworkers and their children in Fresno County, California. Am J Ind Med 2001;40(5):571-577. [doi: 10.1002/ajim.10007.abs]
216. Andrew N, Beare R, Ravipati T, Parker E, Snowdon D, Naude K, et al. Developing a linked electronic health record derived data platform to support research into healthy ageing. Int J Popul Data Sci 2023 Jun 12;8(1):2129 [FREE Full text] [doi: 10.23889/ijpds.v8i1.2129] [Medline: 37670961]
217. Contiero P, Borgini A, Bertoldi M, Abita A, Cuffari G, Tomao P, et al. An epidemiological study to investigate links between atmospheric pollution from farming and SARS-CoV-2 mortality. Int J Environ Res Public Health 2022 Apr 12;19(8):4637 [FREE Full text] [doi: 10.3390/ijerph19084637] [Medline: 35457503]
218. Lotterman A, Baliatsas C, de Rooij MM, Huss A, Jacobs J, Dückers M, et al. Increased risk of pneumonia amongst residents living near goat farms in different livestock-dense regions in the Netherlands. PLoS One 2023 Jul 5;18(7):e0286972 [FREE Full text] [doi: 10.1371/journal.pone.0286972] [Medline: 37405987]
219. Nerich V, Monnet E, Weill A, Vallier N, Vanbockstael V, Auleley G, et al. Fine-scale geographic variations of inflammatory bowel disease in France: correlation with socioeconomic and house equipment variables. Inflamm Bowel Dis 2010 May;16(5):813-821. [doi: 10.1002/ibd.21122] [Medline: 19902547]
220. Nilles JD, Lim D, Boyer MP, Wilson BD, Betar RA, Showalter HA, et al. The occurrence of bone and joint cancers and their association with rural living and radon exposure in Iowa. Environ Geochem Health 2023 Mar 05;45(3):925-940 [FREE Full text] [doi: 10.1007/s10653-022-01261-5] [Medline: 35381949]

XSL•FO
RenderX

221. Yzermans CJ, Moleman YP, Spreeuwenberg P, Nielen MM, Dückers ML, Smit LA, et al. Risk of pneumonia in the vicinity of goat farms: a comparative assessment of temporal variation based on longitudinal health data. Pneumonia (Nathan) 2023 Sep 05;15(1):13 [FREE Full text] [doi: 10.1186/s41479-023-00115-7] [Medline: 37667350]
222. Petit P, Bosson-Rieutort D, Maugard C, Gondard E, Ozenfant D, Joubert N, et al. The TRACTOR project: TRACking and MoniToring occupational risks in agriculture using French insurance health data (MSA). Ann Work Expo Health 2022 Mar 15;66(3):402-411. [doi: 10.1093/annweh/wxab083] [Medline: 34562080]
223. Mills PK, Kwong S. Cancer incidence in the United Farmworkers of America (UFW), 1987–1997. Am J Ind Med 2001;40(5):596-603. [doi: 10.1002/ajim.1125.abs]
224. Guillien A, Laurent L, Soumagne T, Puyraveau M, Laplante J, Andujar P, et al. Anxiety and depression among dairy farmers: the impact of COPD. Int J Chron Obstruct Pulmon Dis 2018;13:1-9 [FREE Full text] [doi: 10.2147/COPD.S143883] [Medline: 29296078]
225. A consortium of agricultural cohort studies. AGRICOH. URL: https://agricoh.iarc.fr/ [accessed 2024-05-17]
226. Jouneau S, Pinault M, Gouyet T, Brinchault G, Guillot S, Viel J, GERPRA. [AIRBAg study: preliminary results after one year of screening for COPD in dairy farmers]. Rev Mal Respir 2015 Mar;32(3):240-248. [doi: 10.1016/j.rmr.2014.03.014] [Medline: 25847201]
227. Soto S, Yoder AM, Nuño T, Aceves B, Sepulveda R, Rosales CB. Health conditions among farmworkers in the southwest: an analysis of the National Agricultural Workers survey. Front Public Health 2022 Nov 3;10:962085 [FREE Full text] [doi: 10.3389/fpubh.2022.962085] [Medline: 36407981]
228. Contaxis N, Clark J, Dellureficio A, Gonzales S, Mannheimer S, Oxley PR, et al. Ten simple rules for improving research data discovery. PLoS Comput Biol 2022 Feb 10;18(2):e1009768 [FREE Full text] [doi: 10.1371/journal.pcbi.1009768] [Medline: 35143479]
229. Vesteghem C, Brøndum RF, Sønderkær M, Sommer M, Schmitz A, Bødker JS, et al. Implementing the FAIR Data Principles in precision oncology: review of supporting initiatives. Brief Bioinform 2020 May 21;21(3):936-945 [FREE Full text] [doi: 10.1093/bib/bbz044] [Medline: 31263868]
230. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3(1):160018 [FREE Full text] [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
231. Roux P, Guillien A, Soumagne T, Ritter O, Laplante J, Travers C, et al. Smoking habits in French farmers: a cross-sectional study. BMC Public Health 2017 Feb 04;17(1):166 [FREE Full text] [doi: 10.1186/s12889-017-4030-4] [Medline: 28160776]
232. Audi C, Baïz N, Maesano CN, Ramousse O, Reboulleau D, Magnan A, et al. Serum cytokine levels related to exposure to volatile organic compounds and PM2.5 in dwellings and workplaces in French farmers - a mechanism to explain nonsmoking COPD. Int J Chron Obstruct Pulmon Dis 2017;12:1363-1374 [FREE Full text] [doi: 10.2147/COPD.S117866] [Medline: 28503065]
233. Baldi I, Gruber A, Rondeau V, Lebailly P, Brochard P, Fabrigoule C. Neurobehavioral effects of long-term exposure to pesticides: results from the 4-year follow-up of the PHYTONER study. Occup Environ Med 2011 Feb 22;68(2):108-115. [doi: 10.1136/oem.2009.047811] [Medline: 21097948]
234. Christensen CH, Platz EA, Andreotti G, Blair A, Hoppin JA, Koutros S, et al. Coumaphos exposure and incident cancer among male participants in the Agricultural Health Study (AHS). Environ Health Perspect 2010 Jan;118(1):92-96 [FREE Full text] [doi: 10.1289/ehp.0800446] [Medline: 20056581]
235. Achard P, Maugard C, Cancé C, Spinosi J, Ozenfant D, Maître A, et al. Medico-administrative data combined with agricultural practices data to retrospectively estimate pesticide use by agricultural workers. J Expo Sci Environ Epidemiol 2020 Jul 04;30(4):743-755. [doi: 10.1038/s41370-019-0166-x] [Medline: 31484997]
236. Hohenadel K, Harris SA, McLaughlin JR, Spinelli JJ, Pahwa P, Dosman JA, et al. Exposure to multiple pesticides and risk of non-Hodgkin lymphoma in men from six Canadian provinces. Int J Environ Res Public Health 2011 Jun 21;8(6):2320-2330 [FREE Full text] [doi: 10.3390/ijerph8062320] [Medline: 21776232]
237. Chung EY, Risi D, Holt JL, Lonergan M, Kotwal S, Yong K, et al. Retrospective study on the epidemiology of antineutrophil cytoplasmic autoantibodies-associated vasculitis in two Australian health districts. Intern Med J 2022 Apr 13;52(4):605-613. [doi: 10.1111/imj.15098] [Medline: 33040456]
238. Li X, Sundquist J, Sundquist K. Socioeconomic and occupational risk factors for rheumatoid arthritis: a nationwide study based on hospitalizations in Sweden. J Rheumatol 2008 Jun;35(6):986-991. [Medline: 18464310]
239. Apalla Z, Lallas A, Sotiriou E, Lazaridou E, Vakirlis E, Trakatelli M, et al. Farmers develop more aggressive histologic subtypes of basal cell carcinoma. Experience from a tertiary hospital in Northern Greece. J Eur Acad Dermatol Venereol 2016 Apr 17;30 Suppl 3(S3):17-20. [doi: 10.1111/jdv.13605] [Medline: 26995018]
240. Lope V, Pollán M, Gustavsson P, Plato N, Pérez-Gómez B, Aragonés N, et al. Occupation and thyroid cancer risk in Sweden. J Occup Environ Med 2005 Sep;47(9):948-957. [doi: 10.1097/01.jom.0000169564.21523.5d] [Medline: 16155480]
241. Nordby KC, Andersen A, Kristensen P. Incidence of lip cancer in the male Norwegian agricultural population. Cancer Causes Control 2004 Aug;15(6):619-626. [doi: 10.1023/b:caco.0000036169.90864.e2]

XSL•FO
RenderX

242. Kristensen P, Andersen A, Irgens LM. Hormone-dependent cancer and adverse reproductive outcomes in farmers' families--effects of climatic conditions favoring fungal growth in grain. Scand J Work Environ Health 2000 Aug;26(4):331-337 [FREE Full text] [doi: 10.5271/sjweh.550] [Medline: 10994799]
243. Døssing M, Petersen KT, Vyberg M, Olsen JH. Liver cancer among employees in Denmark. Am J Ind Med 1997 Sep;32(3):248-254. [doi: 10.1002/(sici)1097-0274(199709)32:3<248::aid-ajim10>3.0.co;2-v] [Medline: 9219654]
244. Klingelschmidt J, Chastang JF, Khireddine-Medouni I, Chérié-Challine L, Niedhammer I. [Occupational factors associated with suicide among French employees from the special agricultural social security scheme (MSA) working between 2007 and 2013]. Rev Epidemiol Sante Publique 2020 Feb;68(1):1-8. [doi: 10.1016/j.respe.2019.09.008] [Medline: 31843361]
245. Manninen P, Heliövaara M, Riihimäki H, Mäkelä P. Does psychological distress predict disability? Int J Epidemiol 1997 Oct;26(5):1063-1070. [doi: 10.1093/ije/26.5.1063] [Medline: 9363529]
246. Zöller B, Li X, Sundquist J, Sundquist K. Socioeconomic and occupational risk factors for venous thromboembolism in Sweden: a nationwide epidemiological study. Thromb Res 2012 May;129(5):577-582. [doi: 10.1016/j.thromres.2011.07.050] [Medline: 21868069]
247. Lucyk K, Lu M, Sajobi T, Quan H. Administrative health data in Canada: lessons from history. BMC Med Inform Decis Mak 2015 Aug 19;15(1):69 [FREE Full text] [doi: 10.1186/s12911-015-0196-9] [Medline: 26286712]
248. Kab S, Spinosi J, Chaperon L, Dugravot A, Singh-Manoux A, Moisan F, et al. Agricultural activities and the incidence of Parkinson's disease in the general French population. Eur J Epidemiol 2017 Mar 9;32(3):203-216. [doi: 10.1007/s10654-017-0229-z] [Medline: 28185034]
249. Perrin L, Spinosi J, Chaperon L, Kab S, Moisan F, Ebaz A. Pesticides expenditures by farming type and incidence of Parkinson disease in farmers: a French nationwide study. Environ Res 2021 Jun;197:111161. [doi: 10.1016/j.envres.2021.111161] [Medline: 33887276]
250. Carles C, Bouvier G, Lebailly P, Baldi I. Use of job-exposure matrices to estimate occupational exposure to pesticides: a review. J Expo Sci Environ Epidemiol 2017 Mar;27(2):125-140. [doi: 10.1038/jes.2016.25] [Medline: 27189257]
251. Tudi M, Li H, Li H, Wang L, Lyu J, Yang L, et al. Exposure routes and health risks associated with pesticide application. Toxics 2022 Jun 19;10(6):335 [FREE Full text] [doi: 10.3390/toxics10060335] [Medline: 35736943]
252. Ahmad MF, Ahmad FA, Alsayegh AA, Zeyaullah M, AlShahrani AM, Muzammil K, et al. Pesticides impacts on human health and the environment with their mechanisms of action and possible countermeasures. Heliyon 2024 Apr 15;10(7):e29128 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e29128] [Medline: 38623208]
253. Cavalier H, Trasande L, Porta M. Exposures to pesticides and risk of cancer: evaluation of recent epidemiological evidence in humans and paths forward. Int J Cancer 2023 Mar 01;152(5):879-912 [FREE Full text] [doi: 10.1002/ijc.34300] [Medline: 36134639]
254. Cresto N, Forner-Piquer I, Baig A, Chatterjee M, Perroy J, Goracci J, et al. Pesticides at brain borders: impact on the blood-brain barrier, neuroinflammation, and neurological risk trajectories. Chemosphere 2023 May;324:138251. [doi: 10.1016/j.chemosphere.2023.138251] [Medline: 36878369]
255. Richardson JR, Fitsanakis V, Westerink RH, Kanthasamy AG. Neurotoxicity of pesticides. Acta Neuropathol 2019 Sep 13;138(3):343-362 [FREE Full text] [doi: 10.1007/s00401-019-02033-9] [Medline: 31197504]
256. Sun J, Fang R, Wang H, Xu D, Yang J, Huang X, et al. A review of environmental metabolism disrupting chemicals and effect biomarkers associating disease risks: where exposomics meets metabolomics. Environ Int 2022 Jan;158:106941 [FREE Full text] [doi: 10.1016/j.envint.2021.106941] [Medline: 34689039]
257. Blair A, Freeman LB. Epidemiologic studies in agricultural populations: observations and future directions. J Agromedicine 2009 May 07;14(2):125-131 [FREE Full text] [doi: 10.1080/10599240902779436] [Medline: 19437268]
258. Carvalho JG, Augusto HC, Ferraz R, Delerue-Matos C, Fernandes VC. Micro(nano)plastic and related chemicals: emerging contaminants in environment, food and health impacts. Toxics 2024 Oct 20;12(10):762 [FREE Full text] [doi: 10.3390/toxics12100762] [Medline: 39453182]
259. Clarke K, Manrique A, Sabo-Attwood T, Coker ES. A narrative review of occupational air pollution and respiratory health in farmworkers. Int J Environ Res Public Health 2021 Apr 13;18(8):4097 [FREE Full text] [doi: 10.3390/ijerph18084097] [Medline: 33924663]
260. Domingo NG, Balasubramanian S, Thakrar SK, Clark MA, Adams PJ, Marshall JD, et al. Air quality-related health damages of food. Proc Natl Acad Sci U S A 2021 May 18;118(20):e2013637118 [FREE Full text] [doi: 10.1073/pnas.2013637118] [Medline: 33972419]
261. Hao Y, Sun H, Zeng X, Dong G, Kronzucker HJ, Min J, et al. Smallholder vegetable farming produces more soil microplastics pollution than large-scale farming. Environ Pollut 2023 Jan 15;317:120805. [doi: 10.1016/j.envpol.2022.120805] [Medline: 36470457]
262. Ng EL, Huerta Lwanga E, Eldridge SM, Johnston P, Hu H, Geissen V, et al. An overview of microplastic and nanoplastic pollution in agroecosystems. Sci Total Environ 2018 Jun 15;627:1377-1388. [doi: 10.1016/j.scitotenv.2018.01.341] [Medline: 30857101]
263. van Kersen W, Bossers A, de Steenhuijsen Piters WA, de Rooij MM, Bonten M, Fluit AC, et al. Air pollution from livestock farms and the oropharyngeal microbiome of COPD patients and controls. Environ Int 2022 Nov;169:107497 [FREE Full text] [doi: 10.1016/j.envint.2022.107497] [Medline: 36088872]

XSL•FO
RenderX

264. Ganter M. Zoonotic risks from small ruminants. Vet Microbiol 2015 Dec 14;181(1-2):53-65. [doi: 10.1016/j.vetmic.2015.07.015] [Medline: 26275853]
265. Golomazou E, Mamedova S, Eslahi AV, Karanis P. Cryptosporidium and agriculture: a review. Sci Total Environ 2024 Mar 15;916:170057. [doi: 10.1016/j.scitotenv.2024.170057] [Medline: 38242460]
266. McDaniel CJ, Cardwell DM, Moeller Jr RB, Gray GC. Humans and cattle: a review of bovine zoonoses. Vector Borne Zoonotic Dis 2014 Jan;14(1):1-19 [FREE Full text] [doi: 10.1089/vbz.2012.1164] [Medline: 24341911]
267. Arora K, Xu L, Bhagianadh D. Dementia and cognitive decline in older adulthood: are agricultural workers at greater risk? J Gerontol B Psychol Sci Soc Sci 2021 Sep 13;76(8):1629-1643 [FREE Full text] [doi: 10.1093/geronb/gbab005] [Medline: 33406265]
268. Douphrate DI, Rosecrance JC, Reynolds SJ, Stallones L, Gilkey DP. Tractor-related injuries: an analysis of workers' compensation data. J Agromedicine 2009 May 07;14(2):198-205. [doi: 10.1080/10599240902773215] [Medline: 19437278]
269. Dos Anjos Magri C, Garófallo Garcia R, Binotto E, Duarte da Silva Lima N, de Alencar Nääs I, Sgavioli S, et al. Occupational risk factors in health of broiler-farm workers: a systematic review. Arch Environ Occup Health 2021 Oct 15;76(8):482-493. [doi: 10.1080/19338244.2020.1832036] [Medline: 33054688]
270. Barbosa Junior M, Sokulski CC, Salvador R, Pinheiro E, de Francisco AC, Trojan F. What kills the agricultural worker? A systematic review on suicide. Rural Remote Health 2021 Aug;21(3):6067 [FREE Full text] [doi: 10.22605/RRH6067] [Medline: 34384223]
271. Garrett-Wright D, Malin C, Jones MS. Mental health in farming communities. J Psychosoc Nurs Ment Health Serv 2023 Oct;61(10):39-43. [doi: 10.3928/02793695-20230424-02] [Medline: 37134283]
272. Lunner Kolstrup C, Kallioniemi M, Lundqvist P, Kymäläinen HR, Stallones L, Brumby S. International perspectives on psychosocial working conditions, mental health, and stress of dairy farm operators. J Agromedicine 2013 Jul 03;18(3):244-255. [doi: 10.1080/1059924X.2013.796903] [Medline: 23844791]
273. Petit P, Gandon G, Dubuc M, Vuillerme N, Bonneterre V. Agricultural activities and risk of treatment for depressive disorders among the entire French agricultural workforce: the TRACTOR project, a nationwide retrospective cohort study. Lancet Reg Health Eur 2023 Aug;31:100674 [FREE Full text] [doi: 10.1016/j.lanepe.2023.100674] [Medline: 37408876]
274. Daghagh Yazd S, Wheeler SA, Zuo A. Key risk factors affecting farmers' mental health: a systematic review. Int J Environ Res Public Health 2019 Dec 02;16(23):4849 [FREE Full text] [doi: 10.3390/ijerph16234849] [Medline: 31810320]
275. Nordby K, Andersen A, Irgens LM, Kristensen P. Indicators of mancozeb exposure in relation to thyroid cancer and neural tube defects in farmers' families. Scand J Work Environ Health 2005 Apr;31(2):89-96 [FREE Full text] [doi: 10.5271/sjweh.855] [Medline: 15864902]
276. Kauppinen T, Uuksulainen S, Saalo A, Mäkinen I, Pukkala E. Use of the Finnish Information System on Occupational Exposure (FINJEM) in epidemiologic, surveillance, and other applications. Ann Occup Hyg 2014 Apr;58(3):380-396. [doi: 10.1093/annhyg/met074] [Medline: 24401793]
277. Shim YK, Mlynarek SP, van Wijngaarden E. Parental exposure to pesticides and childhood brain cancer: U.S. Atlantic coast childhood brain cancer study. Environ Health Perspect 2009 Jun;117(6):1002-1006 [FREE Full text] [doi: 10.1289/ehp.0800209] [Medline: 19590697]
278. Baldi I, Carles C, Blanc-Lapierre A, Fabbro-Peray P, Druet-Cabanac M, Boutet-Robinet E, PESTIMAT Group. A French crop-exposure matrix for use in epidemiological studies on pesticides: PESTIMAT. J Expo Sci Environ Epidemiol 2017 Jan 23;27(1):56-63. [doi: 10.1038/jes.2015.72] [Medline: 26696463]
279. Oltramare C, Mediouni Z, Shoman Y, Hopf NB, Graczyk H, Berthet A. Determinants of pesticide exposure in occupational studies: a meta-analysis. Toxics 2023 Jul 18;11(7):623 [FREE Full text] [doi: 10.3390/toxics11070623] [Medline: 37505588]
280. Burstyn I, Lavoué J, Van Tongeren M. Aggregation of exposure level and probability into a single metric in job-exposure matrices creates bias. Ann Occup Hyg 2012 Nov;56(9):1038-1050. [doi: 10.1093/annhyg/mes031] [Medline: 22986426]
281. Fadel M, Evanoff BA, Andersen JH, d'Errico A, Dale AM, Leclerc A, et al. Not just a research method: if used with caution, can job-exposure matrices be a useful tool in the practice of occupational medicine and public health? Scand J Work Environ Health 2020 Sep 01;46(5):552-553 [FREE Full text] [doi: 10.5271/sjweh.3900] [Medline: 32367143]
282. Peters S, Vienneau D, Sampri A, Turner MC, Castaño-Vinyals G, Bugge M, et al. Occupational exposure assessment tools in Europe: a comprehensive inventory overview. Ann Work Expo Health 2022 Jun 06;66(5):671-686 [FREE Full text] [doi: 10.1093/annweh/wxab110] [Medline: 34935027]
283. Petit P, Gondard E, Gandon G, Moreaud O, Sauvée M, Bonneterre V. Agricultural activities and risk of Alzheimer's disease: the TRACTOR project, a nationwide retrospective cohort study. Eur J Epidemiol 2024 Mar 10;39(3):271-287 [FREE Full text] [doi: 10.1007/s10654-023-01079-0] [Medline: 38195954]
284. List of classifications. International Agency for Research on Cancer. URL: https://monographs.iarc.who.int/list-of-classifications [accessed 2024-10-30]
285. Desai B, Sahni S, Jordan H, Sahni R, Reinbeau R, Nguyen A, et al. Risk of cardiovascular disease in male farmers over the age of 45: a review of literature. Cureus 2022 May;14(5):e24642 [FREE Full text] [doi: 10.7759/cureus.24642] [Medline: 35663675]

XSL•FO
RenderX

286. Kim H, Jung W, Jung S, Cho S, Jung I, Song H, et al. Is farming a risk occupation for cardio-cerebrovascular diseases? A scoping review on cardio-cerebrovascular disease risk in farmers. J Prev Med Public Health 2024 Aug 23;57(6):521-529 [FREE Full text] [doi: 10.3961/jpmph.24.302] [Medline: 39438011]
287. Hulkkonen S, Shiri R, Auvinen J, Miettunen J, Karppinen J, Ryhänen J. Risk factors of hospitalization for carpal tunnel syndrome among the general working population. Scand J Work Environ Health 2020 Jan 01;46(1):43-49 [FREE Full text] [doi: 10.5271/sjweh.3835] [Medline: 31168633]
288. Kim J, Youn K, Park J. Risk factors for musculoskeletal disorders in Korean farmers: survey on occupational diseases in 2020 and 2022. Healthcare (Basel) 2024 Oct 11;12(20):2026 [FREE Full text] [doi: 10.3390/healthcare12202026] [Medline: 39451441]
289. Hemminki K, Niemi ML, Saloniemi I, Vainio H, Hemminki E. Spontaneous abortions by occupation and social class in Finland. Int J Epidemiol 1980 Jun;9(2):149-153. [doi: 10.1093/ije/9.2.149] [Medline: 7409966]
290. Strand K, Wergeland E, Bjerkedal T. Fertility patterns according to occupational grouping in Norway, 1989. Scand J Soc Med 1996 Mar 01;24(1):50-54. [doi: 10.1177/140349489602400108] [Medline: 8740876]
291. Liu Y, Huang L, Hu F, Zhang X. Investigating frailty, polypharmacy, malnutrition, chronic conditions, and quality of life in older adults: large population-based study. JMIR Public Health Surveill 2024 Oct 11;10:e50617 [FREE Full text] [doi: 10.2196/50617] [Medline: 39145920]
292. Lie A, Skogstad M, Johannessen HA, Tynes T, Mehlum IS, Nordby K, et al. Occupational noise exposure and hearing: a systematic review. Int Arch Occup Environ Health 2016 Apr 7;89(3):351-372 [FREE Full text] [doi: 10.1007/s00420-015-1083-5] [Medline: 26249711]
293. Kraemer JG, Aebi S, Oppliger A, Hilty M. The Indoor-Air Microbiota of pig farms drives the composition of the pig farmers' nasal microbiota in a season-dependent and farm-specific manner. Appl Environ Microbiol 2019 May 01;85(9):e03038 [FREE Full text] [doi: 10.1128/AEM.03038-18] [Medline: 30824439]
294. Mucci N, Tommasi E, Chiarelli A, Lulli LG, Traversini V, Galea RP, et al. WORKbiota: a systematic review about the effects of occupational exposure on microbiota and workers' health. Int J Environ Res Public Health 2022 Jan 18;19(3):1043 [FREE Full text] [doi: 10.3390/ijerph19031043] [Medline: 35162072]
295. Shukla SK, Ye Z, Sandberg S, Reyes I, Fritsche TR, Keifer M. The nasal microbiota of dairy farmers is more complex than oral microbiota, reflects occupational exposure, and provides competition for staphylococci. PLoS One 2017 Aug 29;12(8):e0183898 [FREE Full text] [doi: 10.1371/journal.pone.0183898] [Medline: 28850578]
296. Sun J, Liao XP, D'Souza AW, Boolchandani M, Li S, Cheng K, et al. Environmental remodeling of human gut microbiota and antibiotic resistome in livestock farms. Nat Commun 2020 Mar 18;11(1):1427 [FREE Full text] [doi: 10.1038/s41467-020-15222-y] [Medline: 32188862]
297. Vilela C, Araújo B, Soares-Guedes C, Caridade-Silva R, Martins-Macedo J, Teixeira C, et al. From the gut to the brain: is microbiota a new paradigm in Parkinson's disease treatment? Cells 2024 Apr 30;13(9):770 [FREE Full text] [doi: 10.3390/cells13090770] [Medline: 38727306]
298. Gois MF, Fernández-Pato A, Huss A, Gacesa R, Wijmenga C, Weersma RK, et al. Impact of occupational pesticide exposure on the human gut microbiome. Front Microbiol 2023 Aug 10;14:1223120 [FREE Full text] [doi: 10.3389/fmicb.2023.1223120] [Medline: 37637104]
299. Christen P, Schnell R. Thirty-three myths and misconceptions about population data: from data capture and processing to linkage. Int J Popul Data Sci 2023 Jan 31;8(1):2115 [FREE Full text] [doi: 10.23889/ijpds.v8i1.2115] [Medline: 37636835]
300. Dattani N, Hardelid P, Davey J, Gilbert R. Accessing electronic administrative health data for research takes time. Arch Dis Child 2013 May 28;98(5):391-392. [doi: 10.1136/archdischild-2013-303730] [Medline: 23448863]
301. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. J Med Internet Res 2018 May 29;20(5):e185 [FREE Full text] [doi: 10.2196/jmir.9134] [Medline: 29844010]
302. Bousquet P, Caillet P, Coeuret-Pellicer M, Goulard H, Kudjawu Y, Le Bihan C, et al. [Using cancer case identification algorithms in medico-administrative databases: literature review and first results from the REDSIAM Tumors group based on breast, colon, and lung cancer]. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S236-S242. [doi: 10.1016/j.respe.2017.04.057] [Medline: 28576379]
303. Goldberg M, Carton M, Doussin A, Fagot-Campagna A, Heyndrickx E, Lemaitre M, et al. [The REDSIAM network]. Rev Epidemiol Sante Publique 2017 Oct;65 Suppl 4:S144-S148. [doi: 10.1016/j.respe.2017.06.001] [Medline: 28844426]
304. Lipscombe LL, Hwee J, Webster L, Shah BR, Booth GL, Tu K. Identifying diabetes cases from administrative data: a population-based validation study. BMC Health Serv Res 2018 May 02;18(1):316 [FREE Full text] [doi: 10.1186/s12913-018-3148-0] [Medline: 29720153]
305. Caputo M, Pecere A, Sarro A, Mele C, Ucciero A, Pagano L, et al. Incidence and prevalence of hyperthyroidism: a population-based study in the Piedmont Region, Italy. Endocrine 2020 Jul 13;69(1):107-112. [doi: 10.1007/s12020-020-02222-7] [Medline: 32056093]
306. Habib RR, Hojeij S, Elzein K. Gender in occupational health research of farmworkers: a systematic review. Am J Ind Med 2014 Dec 16;57(12):1344-1367 [FREE Full text] [doi: 10.1002/ajim.22375] [Medline: 25227724]

XSL•FO
RenderX

307. Biswas A, Harbin S, Irvin E, Johnston H, Begum M, Tiong M, et al. Sex and gender differences in occupational hazard exposures: a scoping review of the recent literature. Curr Environ Health Rep 2021 Dec 27;8(4):267-280 [FREE Full text] [doi: 10.1007/s40572-021-00330-8] [Medline: 34839446]
308. Biswas A, Tiong M, Irvin E, Zhai G, Sinkins M, Johnston H, et al. Gender and sex differences in occupation-specific infectious diseases: a systematic review. Occup Environ Med 2024 Aug 29;81(8):425-432 [FREE Full text] [doi: 10.1136/oemed-2024-109451] [Medline: 39168602]
309. Comfort N, Re DB. Sex-specific neurotoxic effects of organophosphate pesticides across the life course. Curr Environ Health Rep 2017 Dec 23;4(4):392-404 [FREE Full text] [doi: 10.1007/s40572-017-0171-y] [Medline: 29063415]
310. Wang W, Jin J, He R, Gong H. Gender differences in pesticide use knowledge, risk awareness and practices in Chinese farmers. Sci Total Environ 2017 Jul 15;590-591:22-28. [doi: 10.1016/j.scitotenv.2017.03.053] [Medline: 28285132]
311. Xu L, Xu X, Wu X, Kuang H, Xu C. Sex-dependent environmental health risk analysis of flupyradifurone. Environ Sci Technol 2022 Feb 01;56(3):1841-1853. [doi: 10.1021/acs.est.1c07726] [Medline: 35041393]
312. Ketchen Jr DJ, Roccapriore AY, Connelly BL. Using old data: when is it appropriate? J Manag 2023 Jun 04;49(8):2541-2548. [doi: 10.1177/01492063231177785]
313. Vallot C. Epigenomic tumor evolution under the spotlight: the promises of single-cell approaches. C R Biol 2022 May 11;345(1):11-16 [FREE Full text] [doi: 10.5802/crbiol.75] [Medline: 35787617]
314. Horgan D, Hajduch M, Vrana M, Soderberg J, Hughes N, Omar MI, et al. European health data space-an opportunity now to grasp the future of data-driven healthcare. Healthcare (Basel) 2022 Aug 26;10(9):1629 [FREE Full text] [doi: 10.3390/healthcare10091629] [Medline: 36141241]
315. Martani A, Geneviève LD, Wangmo T, Maurer J, Crameri K, Erard F, et al. Sensing the (digital) pulse. Future steps for improving the secondary use of data for research in Switzerland. Digit Health 2023 Apr 20;9:20552076231169826 [FREE Full text] [doi: 10.1177/20552076231169826] [Medline: 37113255]
316. Raab R, Küderle A, Zakreuskaya A, Stern AD, Klucken J, Kaissis G, et al. Federated electronic health records for the European Health Data Space. Lancet Digit Health 2023 Nov;5(11):e840-e847 [FREE Full text] [doi: 10.1016/S2589-7500(23)00156-5] [Medline: 37741765]
317. Pampel H, Weisweiler NL, Strecker D, Witt M, Vierkant P, Elger K, et al. re3data - indexing the global research data repository landscape since 2012. Sci Data 2023 Aug 29;10(1):571 [FREE Full text] [doi: 10.1038/s41597-023-02462-y] [Medline: 37644024]
318. Suhr B, Dungl J, Stocker A. Search, reuse and sharing of research data in materials science and engineering-a qualitative interview study. PLoS One 2020 Sep 15;15(9):e0239216 [FREE Full text] [doi: 10.1371/journal.pone.0239216] [Medline: 32931508]
319. Noaeen M, Doiron C, Syer J, Brook J. Advancing population health through open environmental data platforms. Curr Top Behav Neurosci 2024;68:297-323. [doi: 10.1007/7854_2024_512] [Medline: 39112811]
320. Dul J, van Raaij E, Caputo A. Advancing scientific inquiry through data reuse: necessary condition analysis with archival data. Strateg Change 2023 Nov 02;33(1):35-40. [doi: 10.1002/jsc.2562]
321. Catalogue des bases de données individuelles en santé en France. INSERM. 2020. URL: https://epidemiologie-france.aviesan.fr/ [accessed 2024-10-30]
322. Plateforme ouverte des données publiques françaises. République Française. URL: https://www.data.gouv.fr/fr/ [accessed 2024-10-30]
323. Dataset search. Google. URL: https://datasetsearch.research.google.com/ [accessed 2024-10-30]
324. Kogevinas M, Schlünssen V, Mehlum IS, Turner MC. The OMEGA-NET international inventory of occupational cohorts. Ann Work Expo Health 2020 Jul 01;64(6):565-568 [FREE Full text] [doi: 10.1093/annweh/wxaa039] [Medline: 32556221]
325. Peters S. Although a valuable method in occupational epidemiology, job-exposure -matrices are no magic fix. Scand J Work Environ Health 2020 May 01;46(3):231-234 [FREE Full text] [doi: 10.5271/sjweh.3894] [Medline: 32356897]
326. Nittas V, Puhan MA, von Wyl V. Toward a working definition of eCohort studies in health research: narrative literature review. JMIR Public Health Surveill 2021 Jan 21;7(1):e24588 [FREE Full text] [doi: 10.2196/24588] [Medline: 33475521]
327. Singh AV, Bansod G, Mahajan M, Dietrich P, Singh SP, Rav K, et al. Digital transformation in toxicology: improving communication and efficiency in risk assessment. ACS Omega 2023 Jun 20;8(24):21377-21390. [doi: 10.1021/acsomega.3c00596] [Medline: 37360489]
328. Soussi A, Zero E, Sacile R, Trinchero D, Fossa M. Smart sensors and smart data for precision agriculture: a review. Sensors (Basel) 2024 Apr 21;24(8):2647 [FREE Full text] [doi: 10.3390/s24082647] [Medline: 38676264]
329. Bukhtiyarova O, Abderrazak A, Chiu Y, Sparano S, Simard M, Sirois C. Major areas of interest of artificial intelligence research applied to health care administrative data: a scoping review. Front Pharmacol 2022 Jul 18;13:944516 [FREE Full text] [doi: 10.3389/fphar.2022.944516] [Medline: 35924057]
330. Caruana A, Bandara M, Musial K, Catchpoole D, Kennedy PJ. Machine learning for administrative health records: a systematic review of techniques and applications. Artif Intell Med 2023 Oct;144:102642. [doi: 10.1016/j.artmed.2023.102642] [Medline: 37783537]

XSL•FO
RenderX

331. Leal Neto O, Von Wyl V. Digital transformation of public health for noncommunicable diseases: narrative viewpoint of challenges and opportunities. JMIR Public Health Surveill 2024 Jan 25;10:e49575 [FREE Full text] [doi: 10.2196/49575] [Medline: 38271097]
332. Wang J, Kharrat FG, Gariépy G, Gagné C, Pelletier J, Massamba VK, et al. Predicting the population risk of suicide using routinely collected health administrative data in Quebec, Canada: model-based synthetic estimation study. JMIR Public Health Surveill 2024 Jun 28;10:e52773 [FREE Full text] [doi: 10.2196/52773] [Medline: 38941610]
333. Ben-Ari T, Boé J, Ciais P, Lecerf R, Van der Velde M, Makowski D. Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France. Nat Commun 2018 Apr 24;9(1):1627 [FREE Full text] [doi: 10.1038/s41467-018-04087-x] [Medline: 29691405]
334. Bhawra J, Elsahli N, Patel J. Applying digital technology to understand human experiences of climate change impacts on food security and mental health: scoping review. JMIR Public Health Surveill 2024 Jul 23;10:e54064 [FREE Full text] [doi: 10.2196/54064] [Medline: 39042453]
335. Comi M, Becot F, Bendixsen C. Automation, climate change, and the future of farm work: cross-disciplinary lessons for studying dynamic changes in agricultural health and safety. Int J Environ Res Public Health 2023 Mar 08;20(6):4778 [FREE Full text] [doi: 10.3390/ijerph20064778] [Medline: 36981685]
336. Ebi KL, Vanos J, Baldwin JW, Bell JE, Hondula DM, Errett NA, et al. Extreme weather and climate change: population health and health system implications. Annu Rev Public Health 2021 Apr 01;42:293-315 [FREE Full text] [doi: 10.1146/annurev-publhealth-012420-105026] [Medline: 33406378]
337. El Khayat M, Halwani DA, Hneiny L, Alameddine I, Haidar MA, Habib RR. Impacts of climate change and heat stress on farmworkers' health: a scoping review. Front Public Health 2022 Feb 8;10:782811 [FREE Full text] [doi: 10.3389/fpubh.2022.782811] [Medline: 35211437]
338. Shahzad A, Ullah S, Dar AA, Sardar MF, Mehmood T, Tufail MA, et al. Nexus on climate change: agriculture and possible solution to cope future climate change stresses. Environ Sci Pollut Res 2021 Jan 29;28(12):14211-14232. [doi: 10.1007/S11356-021-12649-8]
339. Gibb K, Beckman S, Vergara XP, Heinzerling A, Harrison R. Extreme heat and occupational health risks. Annu Rev Public Health 2024 May;45(1):315-335 [FREE Full text] [doi: 10.1146/annurev-publhealth-060222-034715] [Medline: 38166501]
340. Lefèvre-Arbogast S, Chaker J, Mercier F, Barouki R, Coumoul X, Miller GW, et al. Assessing the contribution of the chemical exposome to neurodegenerative disease. Nat Neurosci 2024 May 29;27(5):812-821 [FREE Full text] [doi: 10.1038/s41593-024-01627-1] [Medline: 38684891]
341. Shan L, Heusinkveld HJ, Paul KC, Hughes S, Darweesh SK, Bloem BR, et al. Towards improved screening of toxins for Parkinson's risk. NPJ Parkinsons Dis 2023 Dec 19;9(1):169 [FREE Full text] [doi: 10.1038/s41531-023-00615-9] [Medline: 38114496]
342. Abdelzaher H, Tawfik SM, Nour A, Abdelkader S, Elbalkiny ST, Abdelkader M, et al. Climate change, human health, and the exposome: utilizing OMIC technologies to navigate an era of uncertainty. Front Public Health 2022 Sep 21;10:973000 [FREE Full text] [doi: 10.3389/fpubh.2022.973000] [Medline: 36211706]
343. DeBord DG, Carreón T, Lentz TJ, Middendorf PJ, Hoover MD, Schulte PA. Use of the "exposome" in the practice of epidemiology: a primer on -Omic technologies. Am J Epidemiol 2016 Aug 15;184(4):302-314 [FREE Full text] [doi: 10.1093/aje/kwv325] [Medline: 27519539]
344. Waffenschmidt S, Knelangen M, Sieben W, Bühn S, Pieper D. Single screening versus conventional double screening for study selection in systematic reviews: a methodological systematic review. BMC Med Res Methodol 2019 Jun 28;19(1):132 [FREE Full text] [doi: 10.1186/s12874-019-0782-0] [Medline: 31253092]
345. Waffenschmidt S, Sieben W, Jakubeit T, Knelangen M, Overesch I, Bühn S, et al. Increasing the efficiency of study selection for systematic reviews using prioritization tools and a single-screening approach. Syst Rev 2023 Sep 14;12(1):161 [FREE Full text] [doi: 10.1186/s13643-023-02334-x] [Medline: 37705060]

Abbreviations

AGR: annual growth rate
AGRICAN: Agriculture and Cancer
AGRICOH: agricultural cohort
AHD: administrative health database
AHS: Agricultural Health Study
AI: artificial intelligence
AMI: Aging Multidisciplinary Investigation
ANR: French National Research Agency (Agence Nationale de la Recherche in French)
AS: association strength
BIBLIO: preliminary guideline for reporting bibliometric reviews of the biomedical literature
COPD: chronic obstructive pulmonary disease
EHR: electronic health record

XSL•FO
RenderX

EMR: electronic medical record
FAIR: findable, accessible, interoperable, and reusable
FERMA: risk factors of the rural environment and the allergic and respiratory disease
GDP: gross domestic product
JEM: job-exposure matrix
MIAI: Multidisciplinary Institute in Artificial Intelligence
MSA: Mutualité Sociale Agricole
NOCCA: Nordic Occupational Cancer Study
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
TRACTOR: Tracking and Monitoring Occupational Risks in Agriculture

Edited by A Mavragani; submitted 05.06.24; peer-reviewed by G Wu, F Clerc; comments to author 30.09.24; revised version received 08.10.24; accepted 07.11.24; published 09.01.25.

Please cite as:
Petit P, Vuillerme N
Leveraging Administrative Health Databases to Address Health Challenges in Farming Populations: Scoping Review and Bibliometric Analysis (1975-2024)
JMIR Public Health Surveill 2025;11:e62939
URL: https://publichealth.jmir.org/2025/1/e62939
doi:10.2196/62939
PMID:

©Pascal Petit, Nicolas Vuillerme. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 09.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

When Infodemic Meets Epidemic: Systematic Literature Review

Chaimae Asaad[1,2], MSc; Imane Khaouja[1], PhD; Mounir Ghogho[1,3], PhD; Karim Baïna[2], PhD

[1]TICLab, College of Engineering and Architecture, International University of Rabat, Salé, Morocco

[2]ENSIAS, Alqualsadi, Rabat IT Center, Mohammed V University, Rabat, Morocco

[3]University of Leeds, Leeds, United Kingdom

Corresponding Author:
Chaimae Asaad, MSc
TICLab
College of Engineering and Architecture
International University of Rabat
Shore Rocade, Rocade S, Rabat
Salé, 11103
Morocco
Phone: 212 (5)30112063
Email: chaimae.asaad@uir.ac.ma

Abstract

Background: Epidemics and outbreaks present arduous challenges, requiring both individual and communal efforts. The significant medical, emotional, and financial burden associated with epidemics creates feelings of distrust, fear, and loss of control, making vulnerable populations prone to exploitation and manipulation through misinformation, rumors, and conspiracies. The use of social media sites has increased in the last decade. As a result, significant amounts of public data can be leveraged for biosurveillance. Social media sites can also provide a platform to quickly and efficiently reach a sizable percentage of the population; therefore, they have a potential role in various aspects of epidemic mitigation.

Objective: This systematic literature review aimed to provide a methodical overview of the integration of social media in 3 epidemic-related contexts: epidemic monitoring, misinformation detection, and the relationship with mental health. The aim is to understand how social media has been used efficiently in these contexts, and which gaps need further research efforts.

Methods: Three research questions, related to epidemic monitoring, misinformation, and mental health, were conceptualized for this review. In the first PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) stage, 13,522 publications were collected from several digital libraries (PubMed, IEEE Xplore, ScienceDirect, SpringerLink, MDPI, ACM, and ACL) and gray literature sources (arXiv and ProQuest), spanning from 2010 to 2022. A total of 242 (1.79%) papers were selected for inclusion and were synthesized to identify themes, methods, epidemics studied, and social media sites used.

Results: Five main themes were identified in the literature, as follows: epidemic forecasting and surveillance, public opinion understanding, fake news identification and characterization, mental health assessment, and association of social media use with psychological outcomes. Social media data were found to be an efficient tool to gauge public response, monitor discourse, identify misleading and fake news, and estimate the mental health toll of epidemics. Findings uncovered a need for more robust applications of lessons learned from epidemic "postmortem documentation." A vast gap exists between retrospective analysis of epidemic management and result integration in prospective studies.

Conclusions: Harnessing the full potential of social media in epidemic-related tasks requires streamlining the results of epidemic forecasting, public opinion understanding, and misinformation detection, all while keeping abreast of potential mental health implications. Proactive prevention has thus become vital for epidemic curtailment and containment.

(JMIR Public Health Surveill 2025;11:e55642) doi:10.2196/55642

KEYWORDS

epidemics; social media; epidemic surveillance; misinformation; mental health

XSL•FO
RenderX

Introduction

Background

The global community braved the COVID-19 crisis, with multiple emerging variants, more than 6 million deaths, and 764 million cases being registered [1]. COVID-19 was dubbed "an individual and collective traumatic event," and has "directly or indirectly affected every individual in the world" [2]. Four years later, the world is still grappling with the emotional and socioeconomic aftermath of this crisis [3].

However, COVID-19 has not been the first crisis of its kind to affect global public health. Multiple epidemics have spanned the last 2 decades, causing varying degrees of instability and disease burden [4]. An epidemic is defined as "the occurrence in a community or region of cases of an illness, specific health-related behavior, or other health-related events clearly in excess of normal expectancy" [5]. When an epidemic "occurs worldwide or over a very wide area, crosses international boundaries, and affects a large number of people," it qualifies as a pandemic [5].

Epidemics are often linked to major feelings of uncertainty and loss. The 2014 Ebola outbreak caused rampant fear behaviors in West Africa [6]. The SARS outbreak has created a range of psychiatric conditions, including posttraumatic stress disorder, depressive disorders, and other anxiety spectrum disorders, such as panic, agoraphobia, and social phobia [7]. COVID-19 was associated with major stigma and psychological pressure, further aggravating feelings of guilt, shame, regret, sadness, self-pity, anger, internalized emotions, overwhelmed feelings, negative self-talk, unrealistic expectations, and perceived sense of failure [2]. During epidemics and outbreaks, mistrust of governments and health workers, misinformation, rumors, and conspiracies [8] present challenges to containment and can have a negative impact on mitigation efforts [9-11]. The particular vulnerability surrounding epidemics could render social media users highly suggestible and at risk for fake news acceptance and dissemination [12]. The substantial financial and medical burden imposed by outbreaks and epidemics, in addition to the substantial challenges arising in their progression and aftermath, further complicates the mental health toll they take on the affected population and on vulnerable communities [13].

The control strategies put in place in public health crises to contain the spread of infection are highly dependent on the transmission method and rate [14]. For instance, during COVID-19, various containment measures were adopted, including school closures, shut-downs of nonessential businesses, bans on mass gatherings, travel restrictions, border closures, and curfews [14]. These measures, although necessary for mitigation, can worsen emotional states, contribute to the exacerbation of preexisting socioeconomic inequalities in mental health [15], and lead to unhealthy coping mechanisms, such as problematic internet use, social media addiction, and emotional overeating [16-18].

During epidemics, social media platforms fulfill various functions ranging from informational support to emotional and peer support [19]. They are often a solemn companion offering a tool for connection, a space to grieve, and an instrument of outrage [19]. It is not surprising that the use of social media platforms massively increased during the COVID-19 pandemic [20], rendering them almost essential, ubiquitous, and a catalyst for change, for better and for worse [21].

Social media platforms offer significant amounts of data that can be leveraged for biosurveillance and syndromic surveillance of epidemics and outbreaks [22]. Biosurveillance provides early warning and situational awareness of events using diverse data streams [22]. Efforts directed at facilitating both the early detection and forecasting of disease outbreaks have been increasing in the past 2 decades [22]. Through the analysis of a variety of data sources, syndromic surveillance aims to discern individual and population health indicators before confirmed diagnoses are made [23] using trackable or exhibited behavioral patterns, symptoms, signs, or laboratory findings [23].

Understanding how social media shapes our experiences and preparedness during epidemics, and characterizing the roles it can fulfill, could allow for an improved apprehension of how to efficiently harness this resource for prevention efforts or alleviation of burden of disease [24].

Literature reviews have shown interest in understanding the roles social media fulfills during times of crisis, especially in the last decade [12,25-27]. Social media roles related to the facilitation of public health management, prevention of misinformation, and management of public health behavior and response were found to be of utmost priority [24], and social media topics related to surveillance and monitoring of public attitudes and perceptions, as well as mental health, misinformation, and fake news, were found to be the most well-developed research topics [28]. These 3 particular facets of social media's intersection with epidemics have not been approached in existing reviews; therefore, a gap remains for the research questions (RQs) proposed in this systematic literature review.

This Review

This review aimed to examine the "epidemic-social media" relationship and delineate its various aspects, as well as identify the methods used in harnessing social media in epidemics, with a particular focus on monitoring and surveillance, misinformation, and mental health. In light of the current state of global public health, it is vital to understand how a tool as influential as social media can shape the population's response in times of crisis and how it can be leveraged.

This systematic literature review outlines 3 RQs as follows: (1) How is social media harnessed for epidemic monitoring and management? (RQ1); (2) How is social media used for capturing and managing misinformation during epidemics? (RQ2); and (3) How is social media related to mental health during epidemics? (RQ3).

The remainder of this paper is organized as follows. Methods pertaining to the search strategy and extraction process are detailed in the Methods section. Results of the systematic review are synthesized in the Results section. Discussion of the major issues and practical implications as well as identified directions

XSL•FO
RenderX

for future research are presented in the Discussion section. Conclusions are summarized in the Conclusion section.

Methods

Overview

This systematic review builds upon the preferred reporting items outlined in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [29] (Multimedia Appendix 1).

Proposed RQs

The RQs proposed in this systematic literature review examine the epidemic-social media relationship from different perspectives. The first RQ aims to identify potential uses of social media in the context of epidemic management and mitigation. The second RQ examines potential methods used in the context of social media misinformation as it relates to epidemics. Furthermore, the third RQ aims to discern potential aspects of the relationship between social media and public mental health during epidemics.

Search Strategy

A systematic literature search was undertaken at the beginning of June 2021. A collaborative planning and task allocation process was developed and updated at each stage of the study. The systematic search was conducted across multiple digital libraries—PubMed, IEEE Xplore, ACM Digital Library, ScienceDirect, MDPI, ACL, SpringerLink, arXiv, and ProQuest. Gray literature sources (arXiv and ProQuest) were used to complement the search and reduce publication bias as they provide a venue for authors to share studies with null or negative results that might otherwise not be disseminated.

The RQs were used as a guideline to identify search keywords. The search terms used included “social media” and “epidemics,” with variations depending on the RQ’s objectives and the database searched. For RQ1, the search results of the query (“social media” AND “epidemics”) were complemented by the results of the query (“social media” AND “epidemics” AND “monitoring” AND “tracking”). The combination of these 2 queries allowed for result-filtering without overlimiting the output. The query (“social media” AND “epidemics” AND “fake news”) was used for RQ2. A combination of the queries (“social media” AND “epidemics” AND “mental health” AND “support system”) and (“social media” AND epidemic AND “mental health” AND addiction) was used for RQ3.

These queries were adapted to each database based on its settings. All searches used the parameters *full-text* or *all metadata* in the queries. All searches covered the time range 2010 to 2022.

Table 1 details the number of publications (without duplication) retrieved for screening from each database for each RQ.

Table 1. Output of search strategy for research questions (RQs) 1, 2, and 3.

Database	RQ1	RQ2	RQ3
IEEE Xplore	259	27	54
ScienceDirect	1180	371	240
SpringerLink	2189	367	2188
ACL	90	20	121
ACM Digital Library	672	923	1795
MDPI	178	113	70
arXiv	1544	5	0
ProQuest	226	26	127
PubMed	54	149	217

Study Selection and Data Extraction Strategy

At the initial screening stage, 3 authors assessed the titles and abstracts against the inclusion criteria. Publications included after this screening stage were then retrieved in full-text version, and subsequently screened in the eligibility stage. Three of the authors read the full-text articles independently to ascertain their relevance with regard to the search terms and the research aims. All disagreements on the included articles were resolved by consensus.

To organize the screening process, Rayyan [30], a web application facilitating the collaborative review process and screening process for systematic literature reviews, was used by the authors to import all articles initially collected and screen them following a “blind on” setting, where decisions and labels of any collaborator were not visible to others. Publications with inclusion disagreements were then identified after dropping the “blind on” setting and resolved among authors.

The inclusion and exclusion criteria specified the aims of the review and were agreed upon by all authors (Textbox 1). For a publication to be selected, it needed to address the RQs and be published within the time range. The publication was excluded if it was not a journal paper, conference proceedings paper, or peer-reviewed workshop or symposium paper. Long abstracts and posters were excluded. Publications related to the HIV or tuberculosis epidemic were excluded to preserve the homogeneity of the review. Tuberculosis is a bacterial infection with a high burden of disease, especially in developing countries, while HIV is the virus responsible for AIDS [1]. Both tuberculosis and HIV or AIDS are classified as ongoing worldwide public health issues by the World Health

XSL•FO
RenderX

Organization (WHO) and the Centers for Disease Control and Prevention [1]. Given the particularities of both tuberculosis and HIV or AIDS and the high volume of literature review publications related to them [31], the authors agreed to consider both beyond the scope of this review.

Textbox 1. Inclusion and exclusion criteria for the study selection process.

Inclusion criteria for studies

- Within the scope of one of the research questions
- Published between 2010 and 2022
- Relates to an epidemic or pandemic within the last 2 decades
- Includes the use of a social media site
- Is a journal, conference, or workshop paper

Exclusion criteria for studies

- Tuberculosis, HIV, or, noninfectious diseases
- Online forums or traditional media
- Book, e-book, letter to editor, magazine, abstracts, case reports, comments, reviews, or poster

In the data extraction stage, the final list of papers was analyzed to answer the RQs and extract pertinent information. The final stage of the PRISMA guidelines [29] was considered in this phase. The following data were extracted from selected papers: authors, publication year, epidemic studied, social media site used, theme, identified method, and key findings. All the related data were extracted independently by 2 investigators. When necessary, differences were resolved by discussing, examining, and negotiating with a third investigator.

Quality Assessment

The quality of the included studies in this review was appraised using a set checklist of quality criteria. Papers that did not fulfill at least 4 out of the 5 quality criteria were excluded. The checklist was defined as follows:

1. Are the study objectives clearly defined?
2. Are the methods clearly defined and applied?
3. Are the methods applied successfully and correctly?
4. Are accuracy values and efficiency and confidence levels reported?
5. Are limitations clearly reported and adequately represented?
6. Do the contributions outweigh the limitations of the study?

The quality criteria were formulated based on our understanding of the current state of research in this field and the research gap this systematic review is attempting to fill. The papers were assessed for their ability to answer the RQs and enrich the literature while fulfilling quality standards.

Bias was evaluated in this systematic literature review from 2 aspects. First, the risk of bias based on inclusion was limited through the use of multiple reviewers. Second, publication bias was limited by including gray literature which reports negative and null results. To enhance the quality of this review, the authors monitored the planned review tasks and ensured continuous progress monitoring. Collaborative worksheets were created to keep track of scheduled tasks and deadlines, and to note pertinent observations. Validation of the extracted data from selected papers was conducted by the authors and peer-reviewing was maintained at every stage of the systematic review process.

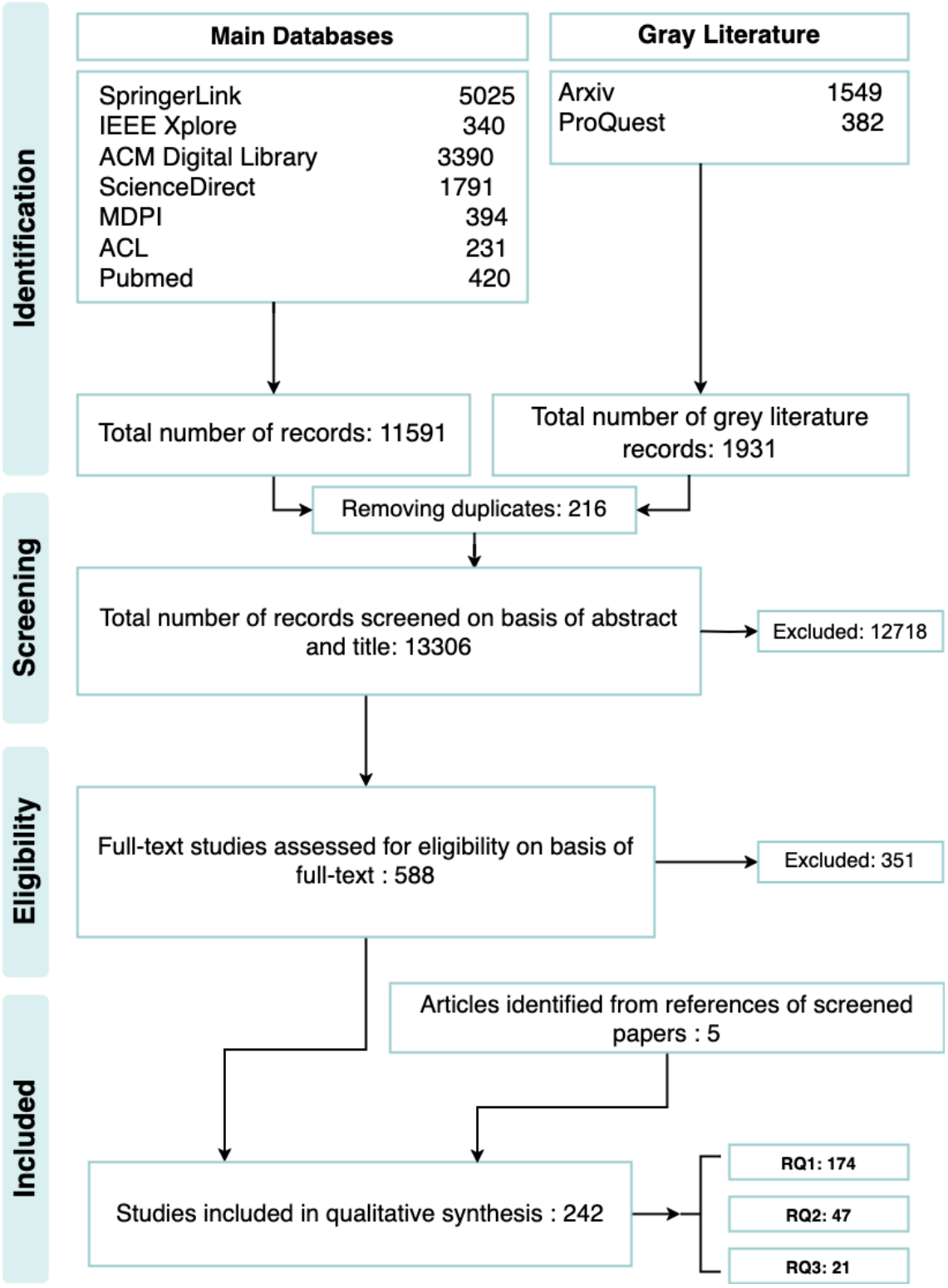
Results

Characteristics of the Selected Papers

The search process resulted in a total of 13,522 articles distributed over both the main and gray databases used. After the removal of duplicates, 13,306 (98.4%) titles remained. Of these, 12,718 (95.58%) studies were excluded after the title and abstract screening, as they did not fulfill the inclusion criteria. A flow diagram of the results of literature collection, screening, eligibility, and inclusion is presented in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for the selection of articles of the literature reviewed. RQ: research question.

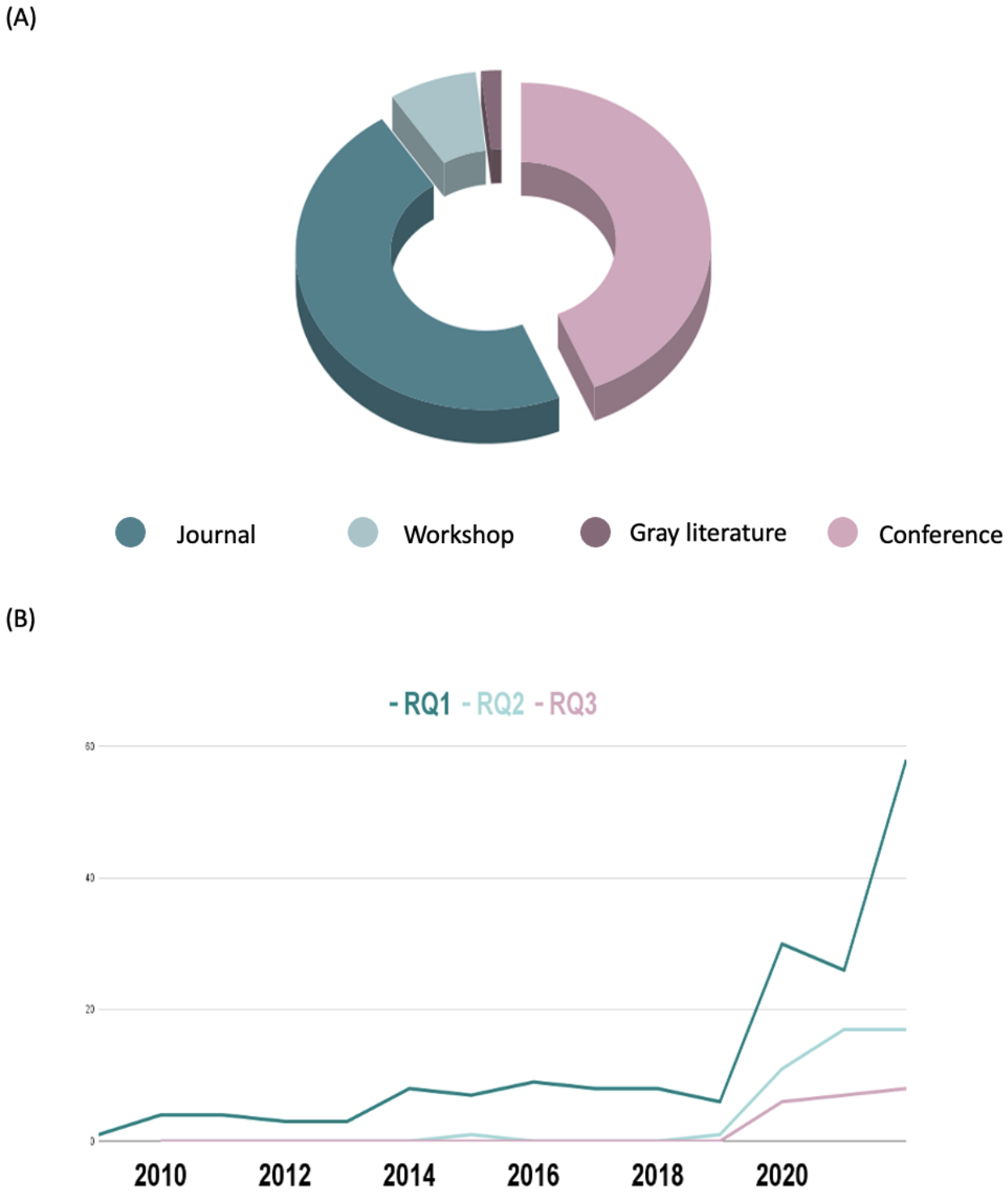


Of the 588 studies that were full-text screened, 351 (59.7%) did not meet the inclusion criteria. 5 (1.4%) papers were identified from reference lists of included papers. A total of 242 (67.9%) studies were selected for inclusion in this review as summarized in subsection Answers to RQs.

The papers included in the review were distributed as follows: 47.1% (114/242) were journal papers, 43.8% (106/242) were publications of conference proceedings, 7.4% (18/242) were workshop and symposium publications, while 1.7% (4/242) were gray literature (Figure 2A).

XSL•FO
RenderX

Figure 2. Distribution of selected papers by (A) type and (B) year. RQ: research question.



The publications spanned from 2010 to 2022. As can be seen in Figure 2B, the number of publications peaked in 2020 and continued to increase for all RQs. All the selected papers that answeredRQ3 spanned the from 2020 to 2021. A similar distribution was seen in papers that answeredRQ2, where selected papers were from 2015, 2019, 2020, and 2021. RQ1, which studies the aspects of epidemic management and mitigation using social media, included the highest number of papers and spanned the entire decade.

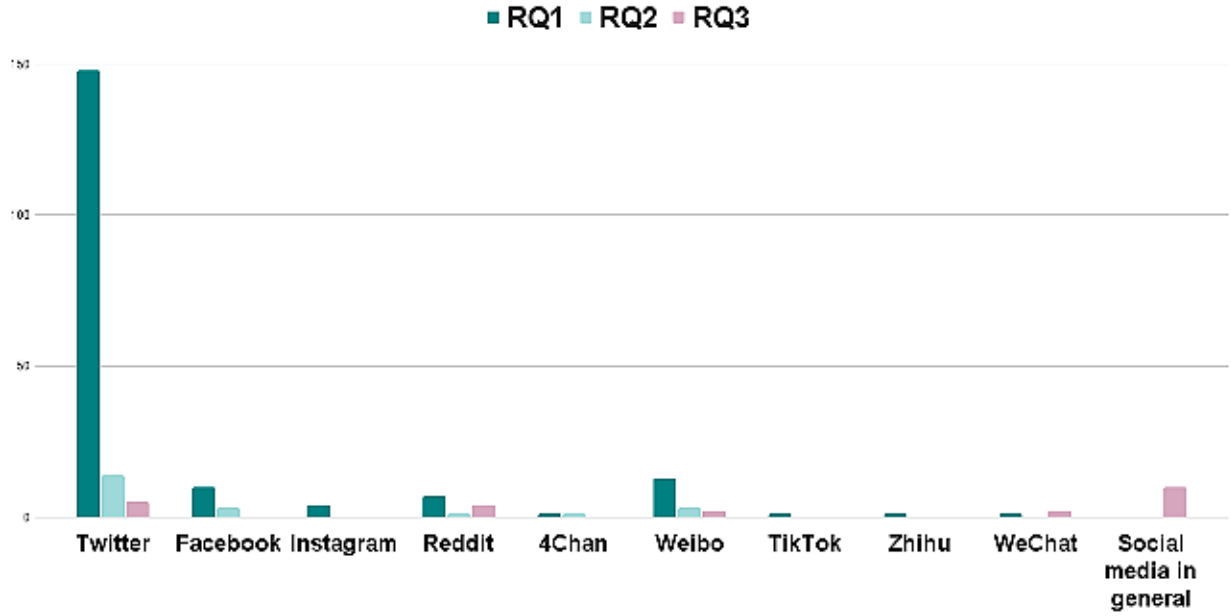
Social Media Platforms Used

Several social media platforms were used in the literature selected for this systematic review. X (formerly Twitter) is one of the most widely used platforms for sharing "microblogs." These short messages are called tweets and can take up to 280 characters. In contrast, Weibo, is a popular platform to share and discuss individual information and life activities as well as celebrity news in China. As can be seen in Figure 3, X followed by Weibo seems to be the platform of choice for most works aiming to study epidemic monitoring and mitigation through

XSL•FO
RenderX

social media (RQ1) and epidemic-related misinformation on social media (RQ2). For epidemic and social media–related mental health aspects, most works seem to take a generalist approach rather than a platform-specific one. Compared with other social media sites, such as Facebook and Instagram, which predominantly include heterogeneous posts, X offers a more concise "microblog" format.

Figure 3. Number of selected publications using each social media platform included in the systematic literature review. RQ: research question.



Epidemics Studied

The selected literature discussed multiple epidemics (Figures 4 and 5), including vital hemorrhagic fevers and influenza-like illness (ILI).

Dengue fever and Zika fever are mosquito-borne diseases caused by the dengue virus and Zika virus, respectively, and spread by several species of female mosquitoes of the Aedes genus [1]. The disease is now endemic in more than 100 countries with potential risk in other areas [1,32]. The WHO declared the Zika outbreak of 2016 and the Ebola outbreak in 2019 as public health emergencies of international concern (PHEICs) [1].

ILI is a nonspecific respiratory illness characterized by fever, fatigue, cough, and other symptoms. Cases of ILI can be caused either by influenza strains or by other viruses, such as coronaviruses. Influenza remains a global and year-round disease burden and causes illnesses that range in severity and sometimes lead to hospitalization and death. Seasonal influenza epidemics are mainly caused by influenza A and B viruses [1]. The influenza A virus subtype strain H1N1, commonly referred to as the swine flu, disproportionately affects children and younger people. H1N1 was declared a PHEIC in 2009 and then designated a pandemic [1]. Coronaviruses include SARS, MERS (Middle East respiratory syndrome), which can be contracted through direct or indirect contact with infected animals [1], as well as COVID-19 caused by the SARS-CoV-2 virus. The latter was designated a PHEIC and a pandemic by the WHO. As of April 26, 2023, the official death toll from COVID-19 reached 6,915,268 [1].

The highest number of selected publications for all RQs related to COVID-19, followed by influenza (Figure 4). This trend is due, in part, to the volume of the COVID-19 research output [33].

XSL•FO
RenderX

Figure 4. Number of selected publications pertaining to each epidemic included in the systematic literature review. MERS: Middle East respiratory virus; RQ: research question.

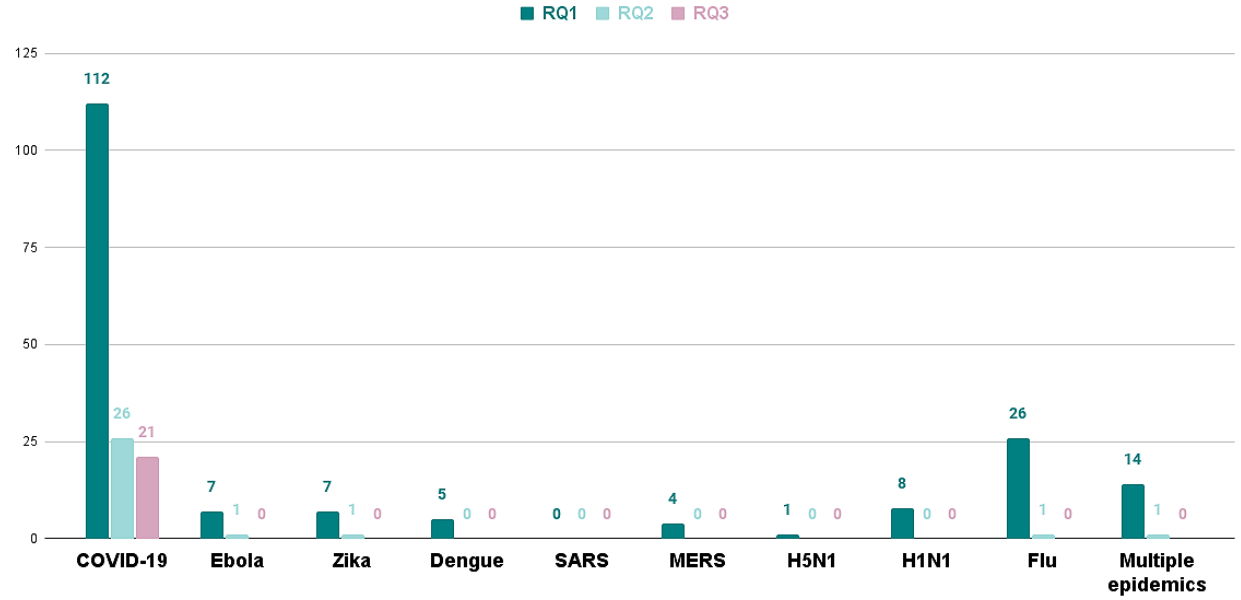
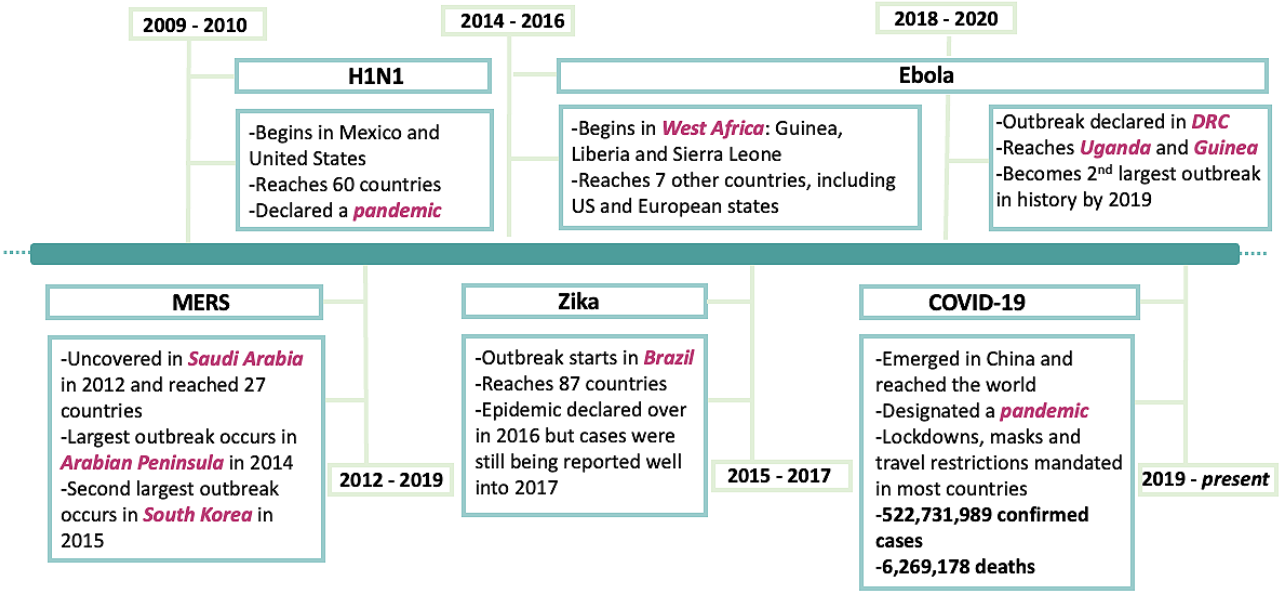


Figure 5. Timeline of the epidemics and pandemics spanning the last decade and included in the systematic literature review. SARS is pre-2009 and dengue fever has caused multiple outbreaks. Both are not illustrated in the timeline but are included in the systematic literature review. DRC: Democratic Republic of the Congo; MERS: Middle East respiratory syndrome.



Answers to RQs

Overview

A thematic analysis of the selected literature was conducted with the aim of identifying the main themes of each RQ. Themes were identified following the objectives of the paper and its results. For each theme, papers were organized by method, social media platform used, and the epidemic studied. Methods were grouped categorically. For instance, content analysis includes automated, linguistic, thematic, qualitative, or quantitative analysis, while dictionary-based classification entails a lexicon-based classification. Machine learning (ML) classification includes conventional ML models, while deep learning (DL) entails methods based on artificial neural networks with representation learning. Although it must be acknowledged that overlaps exist, the categorization used in this paper is based on the most distinctive and predominant use or theoretical approach of each method. This categorization is meant to facilitate a structured analysis and discussion of the literature by grouping papers according to their primary methodological approach, thus allowing a clearer comparison and contrast of their contributions, strengths, and limitations.

XSL•FO
RenderX

RQ1. Social Media for Epidemic Monitoring and Management

Overview

Social media platforms offer significant amounts of data, which can be potentially useful in biosurveillance and syndromic surveillance of epidemics and outbreaks.

Two main themes were identified in the selected papers that addressed how social media could be used in epidemic management, namely, (1) *epidemic surveillance and forecasting*, and (2) *public opinion understanding*.

Epidemic Surveillance and Forecasting

Several works proposed a dictionary-based classification of X for the surveillance of COVID-19 [34,35], dengue fever [36,37], Ebola [38,39], H1N1 [40,41], influenza [42-50], Zika [51,52], MERS [53], and a combination of epidemics [53-55]. Similar epidemic surveillance applications using dictionary-based classification were conducted using data from Weibo for Ebola [56] and influenza [57,58], Reddit for Zika [59], and Facebook for MERS and other epidemics [53].

Reported results indicated that epidemic surveillance can be achieved using varying strategies. For instance, social distancing–related tweets can be grouped into categories, such as implementation, purpose, social disruption, and adaptation, and used to quantify the spatiotemporal prevalence and evolution of COVID-19 social distancing on X [34]. Similarly, official social media channels of information and health organizations, such as the Centers for Disease Control and Prevention, WHO, and National Institutes of Health (NIH) can be monitored, and their X data can be classified to recognize “alarming” news and “concerning” news [35]. Dengue fever reported surveillance strategies to include systems aggregating social media data with weather and flood information [36], and using volume, location, time, and public perception as spatiotemporal dimensions [37].

Results also reported keyword-based data extraction and classification as a strategy for the creation of an Ebola monitoring platforming in China using Weibo data [56] and in Africa using X data [39], and for multiple epidemics [53-55].

Regression analysis was reported to be used for tracking and forecasting influenza [46,47,49,57] and Zika [52], and the Markov switching model was used for real-time early-stage influenza detection with emotion factors for epidemic and nonepidemic segmentation [58]. Statistical analysis was used to study the relationship between human activities collected from Sina Weibo and morbidity patterns and at-risk areas during COVID-19 in China [60].

Correlation analysis reported that X, in addition to other sources, could not provide an Ebola alert more than a week before the WHO and that X’s message volume was correlated more with news article volume than with the number of Ebola cases [38].

Additional dictionary-based surveillance methods include quantitative analysis, filtering, and normalization of X data for H1N1 [40,41] and Zika [51]; mathematical modeling of influenza trends using geo-tagged X streams [42]; time series for X symptom reporting matching ILI [43]; keyword analysis for Zika risk assessment [59,61] as well as influenza risk surveillance [44] and condition aggravation [45], and upcoming influenza spike detection [50], sentiment analysis [48].

Different methodologies using conventional ML were reported to be used for dengue-related event monitoring [62] and lazy associative classification [63]; influenza detection [64-66]; influenza activity monitoring [67]; seasonal influenza trend prediction [68,69]; ILI prevalence prediction [70] and awareness or infection classification [71]; location-specific influenza state detection [58,72,73]; avian influenza outbreak detection [74]; disease-related category classification [75-77] for Ebola, MERS, and dengue fever; guideline-related category classification for ILIs [78]; supervised text classification [79]; topic classification for symptomatic manifestation and prevention of mosquito-borne diseases [80]; infectious disease analytics [81,82] and COVID-19 case forecasting [83], and X-enabled contact tracing [84] and early detection [85].

Conventional ML models used for epidemic surveillance and monitoring include support vector machine (SVM), naive Bayes (NB), and logistic regression (LR).

Several DL techniques were applied for epidemic monitoring [86-90]; fine-tuning of semisupervised model with unlabeled COVID-19 dataset [91]; disease-infected individual detection in tweets using bidirectional encoder representations from transformers (BERT)–based model and disease-infection region identification using spatial analysis [92]; classification of Zika- and Ebola-related tweets [93], COVID-19 related tweets [94], and influenza-related information [95,96]; H1N1 outbreak forecasting and individual-level disease progressing using semisupervised multilayer perceptron and a online stochastic training algorithm [97]; and correlation of X reports on H1N1 infectious disease control using gray wolf optimizer and least square method [98]. Results reported that mathematical modeling can be used to understand the influence of X on the spread of H1N1 [98,99] and information entropy to quantify the impact of social network information [100].

Results showed that social network theory and social network analysis can be used for the prediction of infected groups and early detection of contagious outbreaks in social media [101-104], and that topic modeling techniques, such as latent Dirichlet allocation (LDA) can be used for epidemic intelligence [105-110] to detect major epidemic-related events [111], monitor information spread [112], and rank epidemic-related tweets [113].

Understanding Public Opinion

Several methods were used in the selected literature to extract and analyze public opinions expressed on social media. These methods were based on content analysis of social media data, linguistic analysis, qualitative analysis, lexicon-based analysis, sentiment analysis, valence aware dictionary and sentiment reasoner–based sentiment analysis, topic modeling, conventional ML models, and DL models.

Social media content analysis was used to analyze public discourse around H1N1- [114] and Zika-related risks [61], inspect social media coverage related to influenza vaccinations [115] and COVID-19 vaccinations [116-127], measure public

XSL•FO
RenderX

health concerns [128], identify stances toward policies, such as social distancing and face masks [129], identify emotional composition of online discourse before and after COVID-19 [130], and inspect the presence and escalation of negative sentiments toward China [131]. Latent semantic analysis and LDA were used to mine opinions on X related to the hashtag #IndiaFightsCorona [132]. Topic detection and sentiment analysis were performed for opinion mining, concern exploration, and public opinion analysis in the context of epidemics [133-142], and for pattern analysis [143-145]. Social media content analysis was also used for tracking information spread [146], narratives and information voids [147], monitoring engagement [148-152] and emotional response [153-161], requests for medical assistance [162], health behavior changes [163], governmental response [164,165], and physicians' opinions [166].

Public reaction tracking and investigation were performed using SVM and NB for topic and sentiment analysis [127,167-175]; SVM, NB, and random forest (RF) for social media content classification (eg, caution, advice, notifications, donations, etc) [176]; crisis analysis [177,178]; clustering for topic extraction [179]; and LR for prevention category tweet classification [180]. ML was used to analyze public discourse against masks [181], extract insights on policy response [182], and understand expressions of help-seeking during COVID-19 [183].

BERT-based models were used for public sentiment assessment of data related to COVID-19 available in X [184-186]. Multilingual COVID-19 emotion prediction was performed using a fine-tuned BERT "BERTmoticon" [187], while bias and user opinion were identified using a GPT [188]. A language model for Arabic Moroccan dialect was used for topic modeling, emotion recognition, and polarity analysis [189]. long short-term memory (LSTM), BERT, and enhanced language representation with informative entities were used to analyze the evolution of sentiments in the face of the public health crisis due to COVID-19 [190]. Bi-LSTM with an attention mechanism was used for sentiment analysis of COVID-19–related tweets [191]. Term-frequency analysis was adopted to build an emerging topic graph [192], while the k-means algorithm, LR, SVM, and NB were used to identify COVID-19–related topics [193]. An extra tree and convolutional neural network-based ensemble model was reported to have outperformed conventional ML models in a sentiment classification task [194]. French COVID-19 tweet classification was performed using FlauBERT [195], while opinion monitoring was achieved using a combination of LSTM and global vectors for word representations [196]. Convolutional neural network was used for COVID-19 personal health mention detection [197].

Findings of analyses performed in the context of Ebola, Zika, and influenza revealed that social media posts from health organizations were highly effective when incorporating visuals and that public response was more affected by these communications when they acknowledged the concerns and fear of the community [198]. In the context of Ebola, findings highlighted that online blame was directed toward the affected populations as well as figures with whom social media users had preexisting political frustrations [199].

Analysis of X discussions in relation to COVID-19 revealed the presence of negative sentiments and an association between the words "coronavirus" and "China" [200]; a gradual increase in calls for social distancing, quarantining, and working from home among social media users [201]; a growing number of anger expressions directed at individuals refusing sanitary protocols; and the frequent use of the words "family," "life," "health," and "death" [201]. Analysis of X hashtags also revealed categories, such as quarantine, panic buying, school closures, lockdowns, frustration, and hope [201], as well as mentions of mental health issues and gratitude for essential workers [201]. Other categories and themes identified or used for manual annotation of topics discussed on social media include resource provision, employment and strategies [87], statistics, prevention, hygiene, diagnosis, politics, world news [202], conspiracy, economy, mortality, origin, and outbreak [203].

Findings also indicated increased levels of connectivity and agency coordination during the early-stage response to COVID-19 [87]. Disregarding COVID-19–imposed sanitary and government recommendations was potentially linked to uncertainty in times of crisis, overwhelm by "noise" presented on social media, and varying socioeconomic factors [204].

Results revealed that social media analytics were an efficient approach to capture the attitudes and perceptions of the public during COVID-19 as mentioned in studies by Yigitcanlar et al [205] and Xia et al [206]. Fear and collectivism were identified as predictors of people's preventive intention in the context of COVID-19 [207]. "Sadness" appeared to spike after the WHO declared COVID-19 as a pandemic, while "anger" and "disgust" spiked after the death toll surpassed the hundred thousand in the United States [187].

Tables 2 and 3 summarize the methods, epidemics, and social media used in studies pertaining to epidemic forecasting and prediction and understanding of public opinion.

XSL•FO
RenderX

Table 2. Summary of methodologies used in studies addressing the first part of research question 1 (epidemic surveillance and forecasting).

Method, epidemic studied, and social media used	References
Dictionary-based classification	
COVID-19	
X (formerly Twitter)	[34,35]
Sina Weibo	[60]
Dengue fever	
X	[36,37]
Ebola	
X	[38,39]
Weibo	[56]
H1N1 or swine flu	
X	[40,41]
Influenza or flu	
X	[42-50]
Sina Weibo, Tancent Weibo	[57,58]
Zika	
X	[51,52]
Reddit	[59]
MERS[a]	
X	[53]
Facebook	[53]
Multiple epidemics	
X	[53-55]
Facebook	[53]
ML[b] **classification**	
Dengue fever	
X	[62,63]
Influenza or flu	
X	[64-73]
Sina Weibo, Tancent Weibo	[58,73]
Facebook	[68]
H1N1 or swine flu	
X	[78]
H5N1 or avian influenza	
X	[74]
MERS	
X	[75-77]
Ebola	
X	[93]
Zika	
X	[93]
COVID-19	

XSL•FO
RenderX

Method, epidemic studied, and social media used	References
X	[83-85]
Multiple epidemics	
X	[75-77,79-82]
DL[c] classification	
COVID-19	
X	[35,86,88-91,94]
Ebola	
X	[93]
Zika	
X	[93]
Influenza or flu	
X	[95]
Multiple epidemics	
X	[92]
Mathematical modeling	
COVID-19	
WeChat	[100]
H1N1 or swine flu	
X	[97-99]
Social network analysis	
COVID-19	
X	[101,102]
Influenza or flu	
Facebook	[104]
Multiple epidemics	
X	[103]
Topic modeling	
COVID-19	
X	[88,102,105,106]
Dengue fever	
X	[111]
Ebola	
X	[112]
Influenza or flu	
X	[107]
Weibo	[108]
Zika	
X	[109]
Multiple epidemics	
X	[113]

[a]MERS: Middle East respiratory syndrome.

[b]ML: machine learning.

[c]DL: deep learning.

XSL•FO
RenderX

Table 3. Summary of methodologies used in studies addressing the second part of research question 1 (understanding public opinion).

Method, epidemic studied and social media used	References
Content analysis	
COVID-19	
X	[87,123,126,141,144,146,148,151,156,157,159,160,200,201]
Instagram	[202]
Reddit	[155,157,204]
TikTok	[122]
Weibo	[151,162]
Facebook	[126,149]
Ebola	
X	[198,199]
Facebook	[199]
Instagram	[198]
Zika	
Reddit	[61]
Influenza or flu	
X	[115]
Facebook	[115]
H1N1 or swine flu	
X	[115]
Facebook	[115]
Dictionary-based classification	
COVID-19	
X	[128,129,147,153,158,165,166,203,205,206]
Facebook	[147]
Instagram	[147]
Reddit	[147]
Weibo	[129]
H1N1 or swine flu	
X	[114]
ML[a] classification	
COVID-19	
X	[127,132,164-167,177-179,181]
Weibo	[176]
Facebook	[182]
Instagram	[182]
Zika	
X	[180]
DL[b] classification	
COVID-19	
X	[118-120,130,184-189,191,194-197]
Weibo	[190]

XSL•FO
RenderX

Method, epidemic studied and social media used	References
Topic modeling	
COVID-19	
X	[116,133-135,138,143,154,163,189,192,193]
Reddit	[139]
Weibo	[136,137]
Zhihu	[183]
Social network analysis	
COVID-19	
X	[117,121,124,125,140,142,145,150,152]
Facebook	[161]

[a]ML: machine learning.

[b]DL: deep learning.

RQ2. Social Media for Misinformation Management During Epidemics

Overview

Misinformation, or "fake news," has become a social phenomenon and has received increased attention in the past few years. Although the term, "fake news" has been around since the 1890s [208], the emergence and exponential rise in popularity of social media platforms has brought the term to the "front page." Fake news can fall into multiple categories depending on the intent and form it takes [208]. For instance, fake news can be false information and rumor fabrication (eg, celebrity gossip), hoaxes (eg, doomsday 2012), conspiracy theories (Q-Anon), and satire (eg, The Onion). The intent can range from deception for the purposes of monetary or personal gain to satirizing real news.

One main theme was identified in the selected papers that addressed how social media could be used in misinformation management during epidemics, namely, misinformation detection and characterization. Three subsequent subthemes were identified based on the scope of selected literature, namely: fake news identification, fake news characterization, and information distortion and conspiracy theories.

Misinformation Detection and Characterization

Overview

The selected literature focused on the inspection of news or claims shared on social media, with the aim of classifying them based on trustworthiness. Several methods were used to analyze social media content and detect misleading information, such as expert annotation, DL models, and social network analysis. While some papers focused on technical approaches to the detection of fake news, other studies tried to identify various characteristics related to the source or propagation of fake news.

Fake News Identification

Several works performed fake news identification using DL models [209-211] with conventional ML models for comparison or as baselines. A modified 3-layer-each LSTM and gated recurrent unit were used along with 6 conventional ML classification models (decision trees, LR, k-nearest neighbors, RF, SVM, and NB) on a "Covid-19 fake news Twitter dataset" [212] to identify fake news [210]. Findings reported that the best test results were obtained by LSTM (2 layers), with an accuracy of 98.6%, a precision of 98.55%, a recall of 98.6%, and an F_1-score of 98.5% [210]. Similarly, a multilayer perceptron, LR, decision trees, RF, NB, SVM, and gradient boosting were used for COVID-19 fake news detection in tweets and concluded that RF outperformed other models with an accuracy of 78%, a recall of 100%, a precision of 85%, and an F_1-score of 83% [211]. Expert annotated tweets were used to evaluate the performance of a BERT-based misinformation detection system [213]. Findings suggest that knowledge about the domain vocabulary helps domain-adapted models in predicting the correct stance, as it did for retrieval.

Detecting misleading and fake news was also performed by several studies using methods based on pretrained transformer models, bi-LSTM networks, artificial neural networks, convolutional neural networks, deep transfer learning [214-220], and using hybrid methodologies [221-227].

A semisupervised probabilistic graphical model that aimed to jointly learn the interactions between user trustworthiness, content reliability, and post credibility for influenza posts' credibility analysis outperformed baseline models (RF and Bayesian network) with an accuracy of 71.7% on data from Sina Weibo [209]. LR was performed on a small dataset of Facebook comments to detect fake news [228]. Several ML models, including gradient boosting classifier, LR, RF classifier, and decision tree classification were used in multiple works for fake news classification on social media [229-233].

Other works seeking to curtail the misinformation of COVID-19–related news and support reliable information dissemination used manual analysis through fact-checkers as well as consensus to verify the veracity and correctness of selected tweets and social media posts [234,235]. This is illustrated in a use case analyzing Facebook and X content in both English and Amharic [234] and an Ebola study [235].

XSL•FO
RenderX

Fake News Characterization

A manual annotation of tweet sources following 5 categories (academic, government, media, health professional, and public) allowed for the creation of a gold standard dataset for training a LR model based on 6 million Arabic tweets related to infectious viruses, such as MERS and COVID-19 [236]. Rumor detection using a top-down strategy consisting of extracting posts associated with previously identified rumors reported an 84.03% accuracy for the LR classifier [236]. Higher precision was obtained at the expense of higher runtime using ML models [232]. Similarly, topic modeling based on the k-means algorithm was used to identify sources of COVID-19–related rumors [193]. An entropy-based method was used to investigate the potential control of COVID-19 rumors [237] and content analysis was used to evaluate rumor dissemination and official responses during COVID-19 [238].

Semantic correlations between textual content and attached images were mined using a pretrained convolutional neural network to learn image representations and use them to enhance textual representations and train a fake news detector [239].

Content analysis showed that fake news from multiple sources could be classified using a taxonomy of health and non–health-related types and reported that the response of the public health system was debilitated by the propagation of fake news [240]. Roots of misinformation were categorized as politically related, false medical information, celebrity and pop culture related, religious belief related, and fraud and criminality related [241]. A comparison of fake news sources between China, Iran, and the United States showed that fake science is the main “root” of misinformation in China, while counterexpertise, that is, the rejection of mainstream academic expertise, politically motivated and governmentally sourced misinformation is the most prevalent source of fake news in the United States. In Iran, discourse about COVID-19 was found to be politically manipulated by the government, while official religious figures hindered the dissemination of accurate information [241]. Statistical analysis found bias of sentiment in fake news, as well as biases of gender of the user and media use with respect to real news [242].

Bot detection using BERT was performed as a potential strategy to improve fake news detection [243]. Findings imply that the ratio of real news to fake news is very similar between human accounts and bot accounts, and bot detection could not improve the performance of the fake news detection model [243].

Findings of an information mutation study using A Lite BERT reported that misinformation propagation could potentially be exacerbated by user commentary and found a positive association between information mutation and spreading outcome [244].

The findings of a propagation analysis showed that false claims propagate faster than partially false claims and that tweets containing misinformation are more often concerned with discrediting other information on social media [245].

An investigation leveraging neural networks and quantitative content analysis that aimed to reveal the conditions that lead audiences to accept and disseminate a fake claim as it relates to the Zika virus showed that Zika tweets, including threat cues and protection cues, are positively associated with the likelihood of sharing fake news [246]. In addition, findings of a descriptive analysis showed that the quality of news sources varies considerably with regard to information on COVID-19 [247], Results of a computational analysis indicated that the COVID-19 infodemic is highly characteristic of community structure, shaped by ideological orientation, typology of fake news, and geographic areas of reference [248]. Data from X indicated that content could be labeled according to political affiliation, media source, and type of source (political, satire, mainstream media, science, conspiracy or junk science, clickbait, and fake or hoax) [248].

Information Distortion and Conspiracy Theories

Information distortion in X cascades was found to be linked to oversimplification, distortion of logical links, omission of facts, and a shift in the medical topic to political and business disputes [249]. Risk amplification by information dramatization appeared to be linked to controversial topics as well as social and cultural influences [250].

Manual content and semantic analysis and topic modeling (LDA) techniques of tweet content were conducted through an examination of key term distribution, context, and medical terminology verification [249]. In a COVID-19 5G conspiracy use case, LDA and social network analysis were used to identify several topics from dataset of tweets [251] related to “5G conspiracy” and “5G threat” and discuss topics, including 5G towers, radiation effects, network, and radiation [252,253]. Emerging COVID-19–related conspiracy theories were detected by estimating narrative networks with an underlying graphical model and using a collection of data from Reddit subreddits and 4Chan threads related to the pandemic [254]. Findings identified multiple central conspiracy theories illustrated by examples, such as incorporating the COVID-19 conspiracy into Q-Anon conspiracy, #scamdemic and #plandemic [255], 5G as the cause of COVID-19 [252,253], antivax conspiracy, Bill Gates, #filmyourhospital conspiracy [256], and Pizzagate conspiracy [254]. Table 4 summarizes the methods, epidemics, and social media used in studies pertaining to misinformation management and detection.

XSL•FO
RenderX

Table 4. Summary of methods used in papers addressing research question 2–misinformation identification and characterization.

Method, epidemic studied, and social media used	References
ML[a] classification	
COVID-19	
X	[193,210,211,221-227,229-232]
Facebook	[228,230]
Sina Weibo	[233]
Multiple epidemics	
X	[236]
DL[b] classification	
COVID-19	
X	[210,213-218,220-227,239]
Facebook	[156]
Instagram	[217]
Weibo	[219,239]
Topic modeling	
COVID-19	
X	[237,249,252,255]
Social network analysis	
COVID-19	
X	[248,252,253,256]
Reddit, 4Chan	[254]
Probabilistic graph modeling	
Influenza	
Weibo	[209]
Manual content analysis	
COVID-19	
X	[234,241,248,249]
Facebook	[234,241,247]
Weibo	[241,250]
Instagram	[241]
Ebola	
X	[235]
Quantitative content analysis	
COVID-19	
X	[213,242,245]
Weibo	[250]
Zika	
X	[246]

[a]ML: machine learning.

[b]DL: deep learning.

XSL•FO
RenderX

RQ3. Social Media's Relationship With Mental Health During Epidemics

Overview

During the implementation of restrictive measures requiring limited social contact, social media can become one of the few methods to safely engage with others, rendering it the sole support system of vulnerable populations. Mental health deterioration can manifest in expressions shared on the internet and be used to gauge the toll epidemics and subsequent containment strategies could potentially take on individuals.

Two main themes were identified in the selected papers addressing how social media can be integrated in aspects of public mental health management during epidemics, namely, (1) social media as a tool to gauge the mental health toll of epidemics, and (2) impact of social media consumption during epidemics on mental health.

Mental Health Assessment Using Social Media

Assessment of mental health state was performed using conventional ML [257-259], DL [260-262], and topic modeling techniques [263,264]. Psychological profiles of Weibo users were predicted using ML and online ecological recognition with emotional measures and cognitive indicators, such as anxiety, depression, Oxford happiness, social risk judgment, and life satisfaction [257]. LSTM was used to estimate the rate of depression in the population during the COVID-19 pandemic using Reddit data [260]. Topic modeling, expert intervention, and X data were used to evaluate the possible effects of critical factors related to COVID-19 on the mental well-being of the population in a psychological vulnerability study [263].

Findings revealed that negative emotional indicators of psychological traits increased in anxiety and depression after COVID-19 was declared an epidemic or pandemic [257,262], while life satisfaction and happiness decreased [257]. A 53% average increase in depression rate of Reddit users was noted in selected months after the pandemic [260], and negative psychological vulnerability manifested in negative emotions toward social distancing and hospitalization [263]. Financial burden was found to increase the odds of depressive nonsuicidal thoughts for individuals who suffered job loss during COVID-19 [264]. Results indicated the beginning of recovery following the immediate mental health impact of the COVID-19 pandemic [259].

Table 5 summarizes the methods, epidemics, and social media used in studies pertaining to the use of social media as a tool to gauge the mental health toll of epidemics.

Table 5. Summary of methods used in papers addressing the first part of research question 3 (mental health assessment using social media).

Method and epidemic studied	Social media used	References
ML[a] classification		
COVID-19		
	Weibo	[257]
	X	[258,259]
	Reddit	[259]
DL[b] classification		
COVID-19		
	Reddit	[260,261]
	X	[262]
Topic modeling		
COVID-19		
	X	[263,264]
	Reddit	[264]

[a]ML: machine learning.

[b]DL: deep learning.

Association of Social Media Consumption and Mental Health

Multiple papers conducted cross-sectional studies and statistical analysis to investigate the association between social media consumption and mental health complications during epidemics, particularly during COVID-19. Several studies relied on regression analysis, online surveys, the Generalized Anxiety Disorder Scale, and the Patient Health Questionnaire.

Findings revealed that frequent Sina Weibo use was associated with higher anxiety, depression, and a combination of both [265], and compulsive WeChat use was associated with social media fatigue, emotional stress, and social anxiety [266]. Frequent use of WeChat during COVID-19 was also associated with depression and secondary trauma and was found to be a significant predictor of both [19], while close contact with individuals with COVID-19, along with spending ≥2 hours daily on COVID-19–related news on WeChat was associated with probable anxiety and depression in community-based adults [267]. The association between social media consumption and anxiety and depression was found to be statistically significant

XSL•FO
RenderX

[265,268,269] and positively associated with emotional overeating and anxiety in individuals with neuroticism [18].

The association between the mental health of students receiving higher education and social media use during COVID-19 confinement was analyzed, and results indicated that students in the 18 to 24 years age group, who were not in a relationship and who had lower academic results, presented the highest levels of addiction to social media [16]. Significant positive associations were found between relatedness, need, frustration, and social media addiction, as well as between social media addiction, depressive symptoms, and loneliness [17]. Excessive social media use was also found to fully mediate the relationship between COVID-19–related life concerns and schizotypal traits [270].

Appropriate guidance of adolescents in the use of social networking sites was found to have a potential impact on the mitigation of negative emotions during the COVID-19 pandemic [271].

On the positive side, social media use was found to be rewarding for Wuhan's residents through information sharing and emotional and peer support [19]. Social media breaks were reported to have the potential to promote well-being during the COVID-19 pandemic [19]. In addition, positive mental health and mindfulness appeared to serve as protective factors, and positive mental health was found to be a mediator between the COVID-19 burden and addictive social media use [272].

Table 6 summarizes the methods, epidemics, and social media used in studies pertaining to the association of social media use with mental health issues during epidemics.

Table 6. Summary of methods used in papers addressing the second part of research question 3 (association of social media consumption with mental health).

Method, epidemic studied, and social media used	References
Statistical analysis	
COVID-19	
WeChat	[19,266]
Sina Weibo	[265]
Social media in general	[16-18,267-273]

Discussion

Principal Findings

This systematic literature review conceptualized 3 RQs to investigate if, when, and how social media can be harnessed for successful epidemic management and mitigation, effective curtailment of fake news propagation, and a refined understanding of social media's relationship with mental health during epidemics. It presented a systematic categorization and summary of methods, social media sites, and epidemics broached in the 242 selected works and identified potential research directions and practical implications related to the RQs.

Papers selected pertaining to RQ1 comprised the highest number of papers and included publications from all years of the decade, illustrating continuous and ongoing efforts by the scientific community to harness social media's potential for improved containment measures during epidemics.

COVID-19 was found to be the epidemic most studied in selected papers. This is due to the rapid increase of COVID-19–related publications since the first year of the pandemic. The frequency of publication and the volume of the academic output contributed to the creation of the COVID-19 Open Research Dataset [33]. A similar rising trend was seen in RQ2. This can be explained by the emergence of the "fake news" phenomena on social media and its particular increase in times of crisis. The selected publications answering RQ3 were published from 2020 to 2022. Papers that pertained to RQ3 were much lesser in number than those that pertained to RQ1 and RQ2. Given the mental health aspect of this particular RQ, a potential inference can be made suggesting a very recent interest in mental health as it relates to social media and epidemics. X was found to be the most used social media site in the selected literature, potentially suggesting its attractiveness to works conducting linguistic analysis and classification tasks. This can also be due to the differences in the popularity of social media sites by geographic location and key demographics. The availability of application programming interfaces to crawl data is also a major factor in choosing specific social media platforms as data sources.

General Discussion

The systematic literature review presented in this paper differs from existing reviews and aims to cover a different gap in the literature. Existing works have taken an interest in a broader range of crises, including noninfectious diseases and health risk behaviors [12], disasters in general [25], and new and reemerging infectious diseases [26]. Focus was directed toward effectively targeting vulnerable populations to test interventions and improve health outcomes [12], collective behavior [25], and generalized perspectives on emergency situations [27]. Differences other than scope include data sources, time range, and volume of literature. The review presented in this paper covered a broader time range, included gray literature, and reviewed a sizable volume of research papers.

The review's findings indicated that social media was found to be an effective way to understand the public's reactions and engagement during epidemics [205]. Monitoring topics of discussion during epidemics allowed for insights on whether aspects of epidemic management needed improvement, whether the public agrees with government decisions, and which emotions are linked to the onset of epidemics and mitigation

XSL•FO
RenderX

protocols [198,204-206]. Analysis of opinions related to aspects, such as COVID-19 vaccinations were proposed and could be used to give feedback to governments and health organizations to implement better suited protocols [116,122,124-126] for mitigation, and to identify topics of misinformation, and therefore offer clarifications or conduct further awareness efforts to combat rumors and conspiracies [254]. Results also indicated that social media can be used in case forecasting [83], X-enabled contact tracing [84], early detection [85], tracking adherence to preventive guidelines, such as wearing masks and social distancing [205,206], and monitoring symptomatic self-expressions of infection [80]. Misinformation detection on social media was performed as a classification task, manually using experts and fact checkers, and using artificial intelligence techniques; however, presented several challenges. Misinformation often used language styles of academics and health professionals to deceive the public [236] and propagated faster when it included higher levels of threat due to the collective stress reaction it generated [246]. "Troll" accounts were found to play the second most prominent role is misinformation spread and present a "substantial cause for concern" [248]. Other challenges of misinformation detection related to limitations of studies due to the use of small batches of data [252], false positives [228], and a "politicization" of neutral health emergency crises [235].

Although epidemics were found to cause negative emotions and mental health issues [260,262,263], many expressions of positive emotions were noted [257], reflecting group cohesiveness rather than pure personal emotions. Group threats contributed to the manifestation of more beneficial behaviors and social solidarity [269]. Viewing heroic acts, speeches from experts, and knowledge of the disease and prevention methods were associated with more positive effects and less expressions of depression [269]. Media content, including useful information for self-protection was found to be potentially helpful to people during epidemics and may enhance active coping, prevention behaviors, and instill a sense of control [269]. The use of social media during epidemics, although linked with manifestations of anxiety and depression, appeared to benefit Wuhan residents and was perceived as an important activity during lockdown [19]. Balancing social media use to obtain ample informational as well as emotional and peer support, while avoiding the potential mental health toll, is a difficult task for users, especially without the availability of alternative and easily accessible sources of health information [19].

Using social media data for mental health assessment has its challenges and limitations. It can add a population or demographic bias to results, given that some social media sites are predominantly used by younger people or are more or less popular depending on the country [257,263]. Depending on the social media site (eg, Reddit), the user pool skews younger, and thus could be more prone to depression [260]. Moreover, some analyses are based on a weekly basis, with a relatively large granularity, which has certain influences on reflecting the changing trend of social mentality in a timely manner [257]. The qualitative nature of the results obtained and interpreted by domain experts limits the generalization of the findings and requires more corroborating results. Consequently, findings may need additional data to be strengthened [260,263]. As for works pertaining to the association of social media consumption with psychological outcomes, a causal link has not been established due to the cross-sectional nature of the contributions. Studies reflected a single point in time for participants, therefore, further longitudinal studies are necessary. In addition, the surveys were conducted on the web, and consequently, respondent bias is possible [265]. The recruitment of all participants from the same country and from one social media platform can introduce bias to studies [266,268], in addition to potential gender biases and sample representativeness [18,19], and recall bias related to self-reporting [269]. The results could not exclude the possibility of residual confounding caused by unmeasured factors.

A change can be seen in the evolution of research themes over time and through different epidemics. A sizable number of works focused on the influenza epidemic surveillance using lexicon-based and dictionary-based classifications, as well as classical ML techniques. This volume of literature could potentially be linked to the influenza prediction "wave" that preceded, paralleled, and followed the dereliction of the "Google flu trend" after its failure to predict major outbreaks [274]. Although various methods were used, ML and DL techniques were most frequently used for COVID-19 surveillance. Scientific contributions evolved with the emergence of more epidemics. COVID-19 appeared to have benefited from the digitization of literature as well as the development and improvements taking place in the fields of natural language processing, ML, DL, social network analysis, and topic modeling. The global nature of the COVID-19 crisis generated an influx of publications and contributions. The theme of misinformation management has also evolved with epidemics and with the proliferation of social media fake news, bots, troll accounts, and widely propagated conspiracy theories. COVID-19 has been the subject of multiple controversies and conspiracies, which encouraged scientific efforts to study potential curtailment methods. As for the mental health aspect, all publications pertaining to the scope of RQ3 were related to COVID-19, and it appeared that previous epidemics were not subject to social media association analysis. This could be due to the fear linked to COVID-19 and the challenging nature of sanitary measures such as global lockdowns and social distancing, which led to an increase in social media reliance. It could also be due to the decade's zeitgeist which brought online mental health discussions and awareness front and center.

Identified Issues

One of the major issues identified was the lack of preemptive measures building on the results of previous studies and aiming to implement social media–enabled processes in real time or near real time. Lessons learned are not efficiently integrated in crisis mitigation measures nor used as building blocks for optimized proactive prevention. A synergy between government health agencies, research communities, and the public would allow for the success of social-media public health initiatives. Such collaborative efforts require effective and trustworthy interactions. This highlights an additional issue related to the relative inefficiency of social media campaigns. Populations need to be targeted for both informative purposes and for active

XSL•FO
RenderX

emotional support. Understanding public opinion is useful to gauge sentiments and reactions, and therefore it is important to remedy the gap for applications integrating extracted opinions in targeted epidemic management.

Because of the medical and financial burden of epidemics, mental health concerns are often ignored by both governments and the public. As a result, the manifestation of several mental health–related symptoms becomes more prevalent as epidemics progress. In the case of the Ebola outbreak in 2014, symptoms of posttraumatic stress disorder and anxiety-depression were more prevalent even after a year of the Ebola response [199]. When limited resources are geared for epidemic containment, the health care system focuses majorly on emergency services. Therefore, individuals with substance abuse and dependency disorders may see deterioration in their mental health [13]. During community crises, event-related information is often sought in an effort to retain a sense of control in the face of fear and uncertainty and their psychological manifestations. When misleading misinformation is propagated on social media, perceptions of risk are distorted, leading to extreme public panic, stigmatization, and marginalization [13]. Psychological interventions and psychosocial support would have a direct impact on the improvement of public mental health during epidemics.

Directions for Future Research

We identified several issues and gaps in the literature related to the RQs of this systematic literature review and suggest potential paths for future research.

Given the recognized impact of epidemics on mental health and the prevalent use of social media platforms during times of crisis, it is necessary to explore the aspects of social media leading to mental health deterioration during epidemics. Potential factors range from increased consumption levels of social media, social media addiction, emotional fatigue due to overwhelm, and consumption of “sad” content. Investigating which aspects of social media use are responsible for worsening states of mental health and mental health disorders would allow a targeted approach to curbing this negative impact during times of crisis. As for health-related fake news, it is important to understand what makes citizens prone to engaging in fake news sharing. Specifically, features identifying both an individual’s and a group’s susceptibility to believe and share misinformation need to be determined and categorized. Levels of education, geographic and demographic profiles, cultural influences, and psychological vulnerability are potential features requiring further investigation in their association with fake news dissemination on social media and within communities.

Epidemics are rapidly changing phenomena requiring fast interventions and decision-making. Although postcrisis analysis is imperative for an improved understanding of lessons learned, proactive epidemic management is vital and would have the most impact on mitigation efforts. Integrating artificial intelligence techniques into this proactive surveillance could further optimize this process.

In addition, misinformation propagation has a significant impact on the success of interventions given that both the components of exaggerated fear and apathy linked to misinformation can hinder management efforts. However, the investigation of misinformation needs to be extended to include potential links between misinformation and mental health deterioration.

Practical Implications

This work has several potential practical implications pertaining to different entities.

Implications for governing entities include the development of an efficient misinformation correction strategy to fight incorrect information, rumors, and conspiracy theories related to epidemics; the development of clear communication channels for knowledge dissemination to build trust with the public; the development of interventions to limit the impact of epidemics on stress responses (anxiety, depression) due to distorted risk perceptions; the bolstering of public awareness efforts on sanitary measures and proactive protection; and the insurance of the supply of medical staff available to treat patients, as well as psychological support staff to assist patients and their families in navigating the ramifications of infection and loss of loved ones.

Implications for social media platforms include taking a leadership position in the management of epidemic-related fake news by implementing built-in fact-checking processes and assisting health agencies and scientific entities in disseminating factual information about the disease, its symptoms, its potential risk, and efficient sanitary measures for the public to adopt.

Implications for the public include improving community resilience during epidemics using social media groups and assisting in combating misinformation.

Limitations

The results of this review should be considered in light of several limitations. The data sources used in this review did not cover all existing scientific databases, and therefore, cannot generalize findings to the entirety of the literature. The scope of the review focused on specific aspects of the epidemic-social media relationship, and so does not provide a general overview. Although the process of data extraction and analysis was undertaken with extreme diligence, there can be potential for bias. Despite our recognition of the inherent limitations of any search strategy, we have ensured our commitment to the rigor and transparency of the systematic review process.

Conclusions

Given the collective experience of epidemics, responses by communities can often provide insight into the degree of adherence toward preventive measures as well as mitigation protocols. In an effort to control the spread of epidemics, governments, public health institutions, and health care professionals generally issue guidelines for the public through online portals, news sources, and in the past decade, social media. Online “chatter” can indicate the public’s response to these guidelines, and their sentiments toward the epidemic itself or specific topics related to it, such as vaccinations, treatments, mortality rates, etc. Mitigation efforts require collaborative strategies and public involvement; therefore, gaining insight

XSL•FO
RenderX

into public opinion and response can prove vital in the success or failure of such efforts.

It is evident that epidemic preparedness and mitigation protocols need to be adjusted to deal with the special challenges that accompany the technological revolution taking place, especially in light of the considerable impact of the ongoing infodemic. In addition, it is vital to have effective ways to exploit the full potential of social media without risking the toll it could potentially take on users' mental health. The systematic literature review presented in this paper covers several key aspects of the relationship between epidemics and social media, especially with respect to fake news and mental health. Methods used to answer RQs are categorized. The findings of this review could shed light on broader implications related to data quality concerns and privacy considerations in epidemic surveillance, thus highlighting the lack of works proposing ethical, legal, and technical frameworks to accompany scientific efforts. Learning from past crises and integrating a digital and social media-enabled infrastructure into public health protocols could make a difference in future preparedness levels.

Acknowledgments

The authors thank the anonymous reviewers for their remarks and improvements to the paper. This work was carried out as part of the eCovid and CovInov projects which are funded by the Centre National de la Recherche Scientifique et Technique and the Rabat-Salé-Kénitra region.All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Data Availability

The datasets generated during and analyzed during this study are available in the Github repository [275].

Authors' Contributions

CA and MG conceived the study. CA and IK designed the experiments. CA, IK, and MG carried out the research. CA and IK prepared the first draft of the manuscript. MG and KB contributed to the experimental design and preparation of the manuscript. All authors were involved in the revision of the draft manuscript and have agreed to the final content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File , 33 KB - publichealth_v11i1e55642_app1.docx]

References

1. Health topics. World Health Organization. URL: https://www.who.int/health-topics [accessed 2024-12-24]
2. Mukhtar S. Psychological health during the coronavirus disease 2019 pandemic outbreak. Int J Soc Psychiatry 2020 Aug;66(5):512-516 [FREE Full text] [doi: 10.1177/0020764020925835] [Medline: 32434402]
3. Duarte Alonso A, Kok SK, Bressan A, O'Shea M, Sakellarios N, Koresis A, et al. COVID-19, aftermath, impacts, and hospitality firms: an international perspective. Int J Hosp Manag 2020 Oct;91:102654 [FREE Full text] [doi: 10.1016/j.ijhm.2020.102654] [Medline: 32863526]
4. Major epidemics of the modern era. Council on Foreign Relations. URL: https://www.cfr.org/timeline/major-epidemics-modern-era [accessed 2024-12-24]
5. Porta M. A Dictionary of Epidemiology. Oxford, UK: Oxford University Press; 2008.
6. Shultz JM, Baingana F, Neria Y. The 2014 Ebola outbreak and mental health: current status and recommended response. JAMA 2015 Feb 10;313(6):567-568. [doi: 10.1001/jama.2014.17934] [Medline: 25532102]
7. Mak IW, Chu CM, Pan PC, Yiu MG, Chan VL. Long-term psychiatric morbidities among SARS survivors. Gen Hosp Psychiatry 2009 Jul;31(4):318-326 [FREE Full text] [doi: 10.1016/j.genhosppsych.2009.03.001] [Medline: 19555791]
8. Saroj A, Pal S. Use of social media in crisis management: a survey. Int J Disaster Risk Reduct 2020 Sep;48:101584. [doi: 10.1016/j.ijdrr.2020.101584]
9. Lin TH, Chang MC, Chang CC, Chou YH. Government-sponsored disinformation and the severity of respiratory infection epidemics including COVID-19: a global analysis, 2001-2020. Soc Sci Med 2022 Mar;296:114744 [FREE Full text] [doi: 10.1016/j.socscimed.2022.114744] [Medline: 35124544]
10. Pavela Banai I, Banai B, Miklošić I. Beliefs in COVID-19 conspiracy theories, compliance with the preventive measures, and trust in government medical officials. Curr Psychol 2022 May 26;41(10):7448-7458 [FREE Full text] [doi: 10.1007/s12144-021-01898-y] [Medline: 34075284]

XSL•FO
RenderX

11. Pfeffer B, Goreis A, Reichmann A, Bauda I, Klinger D, Bock MM, et al. Coping styles mediating the relationship between perceived chronic stress and conspiracy beliefs about COVID-19. Curr Psychol 2022 Aug 16;43(15):1-9 [FREE Full text] [doi: 10.1007/s12144-022-03625-7] [Medline: 35990199]
12. Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EH, Olsen JM, et al. Using social media for actionable disease surveillance and outbreak management: a systematic literature review. PLoS One 2015;10(10):e0139701 [FREE Full text] [doi: 10.1371/journal.pone.0139701] [Medline: 26437454]
13. Roy A, Singh AK, Mishra S, Chinnadurai A, Mitra A, Bakshi O. Mental health implications of COVID-19 pandemic and its response in India. Int J Soc Psychiatry 2021 Aug;67(5):587-600 [FREE Full text] [doi: 10.1177/0020764020950769] [Medline: 32873106]
14. Kaimann D, Tanneberg I. What containment strategy leads us through the pandemic crisis? An empirical analysis of the measures against the COVID-19 pandemic. PLoS One 2021 Jun 21;16(6):e0253237 [FREE Full text] [doi: 10.1371/journal.pone.0253237] [Medline: 34153058]
15. Serrano-Alarcón M, Kentikelenis A, Mckee M, Stuckler D. Impact of COVID-19 lockdowns on mental health: evidence from a quasi-natural experiment in England and Scotland. Health Econ 2022 Feb 12;31(2):284-296 [FREE Full text] [doi: 10.1002/hec.4453] [Medline: 34773325]
16. Oliveira AP, Nobre JR, Luis H, Luis LS, Pinho LG, Albacar-Riobóo N, et al. Social media use and its association with mental health and internet addiction among Portuguese higher education students during COVID-19 confinement. Int J Environ Res Public Health 2022 Dec 30;20(1):664 [FREE Full text] [doi: 10.3390/ijerph20010664] [Medline: 36612991]
17. Cheng C, Lau YC. Social media addiction during COVID-19-mandated physical distancing: relatedness needs as motives. Int J Environ Res Public Health 2022 Apr 12;19(8):4621 [FREE Full text] [doi: 10.3390/ijerph19084621] [Medline: 35457489]
18. Gao Y, Ao H, Hu X, Wang X, Huang D, Huang W, et al. Social media exposure during COVID-19 lockdowns could lead to emotional overeating via anxiety: the moderating role of neuroticism. Appl Psychol Health Well Being 2022 Feb;14(1):64-80 [FREE Full text] [doi: 10.1111/aphw.12291] [Medline: 34265167]
19. Zhong B, Huang Y, Liu Q. Mental health toll from the coronavirus: social media usage reveals Wuhan residents' depression and secondary trauma in the COVID-19 outbreak. Comput Human Behav 2021 Jan;114:106524 [FREE Full text] [doi: 10.1016/j.chb.2020.106524] [Medline: 32836728]
20. Social media and COVID-19: a global study of digital crisis interaction among Gen Z and Millennials. World Health Organization. 2021 Dec 1. URL: https://www.who.int/news-room/feature-stories/detail/social-media-covid-19-a-global-study-of-digital-crisis-interaction-among-gen-z-and-millennials [accessed 2024-12-24]
21. Kapoor KK, Tamilmani K, Rana NP, Patil P, Dwivedi YK, Nerur S. Advances in social media research: past, present and future. Inf Syst Front 2017 Nov 6;20(3):531-558. [doi: 10.1007/s10796-017-9810-y]
22. Walters RA, Harlan PA, Nelson NP, Hartley DM, Voeller JG. Data sources for biosurveillance. In: Wiley Handbook of Science and Technology for Homeland Security. Hoboken, NJ: John Wiley & Sons; 2010.
23. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F, et al. Implementing syndromic surveillance: a practical guide informed by the early experience. J Am Med Inform Assoc 2004 Nov 21;11(2):141-150 [FREE Full text] [doi: 10.1197/jamia.M1356] [Medline: 14633933]
24. Selerio EJ, Caladcad JA, Catamco MR, Capinpin EM, Ocampo L. Emergency preparedness during the COVID-19 pandemic: modelling the roles of social media with fuzzy DEMATEL and analytic network process. Socioecon Plann Sci 2022 Aug;82:101217 [FREE Full text] [doi: 10.1016/j.seps.2021.101217] [Medline: 35001981]
25. Eismann K, Posegga O, Fischbach K. Collective behaviour, social media, and disasters: a systematic literature review. In: Proceedings of the European Conference on Information Systems. 2016 Presented at: ECIS 2016; June 12-15, 2016; Istanbul, Turkey.
26. Tang L, Bie B, Park SE, Zhi D. Social media and outbreaks of emerging infectious diseases: a systematic review of literature. Am J Infect Control 2018 Sep;46(9):962-972 [FREE Full text] [doi: 10.1016/j.ajic.2018.02.010] [Medline: 29628293]
27. Abdulhamid NG, Ayoung DA, Kashefi A, Sigweni B. A survey of social media use in emergency situations: a literature review. Inf Dev 2020 Mar 26;37(2):274-291. [doi: 10.1177/0266666920913894]
28. Michailidis PD. Visualizing social media research in the age of COVID-19. Information 2022 Aug 03;13(8):372. [doi: 10.3390/info13080372]
29. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. [The PRISMA 2020 statement: an updated guideline for reporting systematic reviewsDeclaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas]. Rev Panam Salud Publica 2022;46:e112 [FREE Full text] [doi: 10.26633/RPSP.2022.112] [Medline: 36601438]
30. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
31. Taggart T, Grewe ME, Conserve DF, Gliwa C, Roman Isler M. Social media and HIV: a systematic review of uses of social media in HIV communication. J Med Internet Res 2015 Nov 02;17(11):e248 [FREE Full text] [doi: 10.2196/jmir.4387] [Medline: 26525289]
32. Guzman MG, Harris E. Dengue. The Lancet 2015 Jan;385(9966):453-465. [doi: 10.1016/s0140-6736(14)60572-9]

XSL•FO
RenderX

33. Wang LL, Lo K, Chandrasekhar Y, Reas R, Yang J, Eide D, et al. CORD-19: the COVID-19 open research dataset. arXiv Preprint posted online on April 22, 2020 [FREE Full text]
34. Kwon J, Grady C, Feliciano JT, Fodeh SJ. Defining facets of social distancing during the COVID-19 pandemic: Twitter analysis. J Biomed Inform 2020 Nov;111:103601 [FREE Full text] [doi: 10.1016/j.jbi.2020.103601] [Medline: 33065264]
35. Rao HR, Vemprala N, Akello P, Valecha R. Retweets of officials' alarming vs reassuring messages during the COVID-19 pandemic: implications for crisis management. Int J Inf Manage 2020 Dec;55:102187 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2020.102187] [Medline: 32836644]
36. Othman MK, Danuri MS. Proposed conceptual framework of Dengue Active Surveillance System (DASS) in Malaysia. In: Proceedings of the International Conference on Information and Communication Technology. 2016 Presented at: ICICTM 2016; May 16-17, 2016; Kuala Lumpur, Malaysia. [doi: 10.1109/icictm.2016.7890783]
37. Gomide J, Veloso A, Meira WJ, Almeida V, Benevenuto F, Ferraz F, et al. Dengue surveillance based on a computational model of spatio-temporal locality of Twitter. In: Proceedings of the 3rd International Web Science Conference. 2011 Presented at: WebSci '11; June 15-17, 2011; Koblenz, Germany. [doi: 10.1145/2527031.2527049]
38. Yom-Tov E. Ebola data from the internet: an opportunity for syndromic surveillance or a news event? In: Proceedings of the 5th International Conference on Digital Health 2015. 2015 Presented at: DH '15; May 18-20, 2015; Florence, Italy. [doi: 10.1145/2750511.2750512]
39. Marc T, Kodzo A. How useful are social networks for analyzing epidemics? The example of Twitter for monitoring the 2018–2019 Ebola epidemic in Africa. In: Proceedings of the International Multi-Conference on: “Organization of Knowledge and Advanced Technologies”. 2020 Presented at: OCTA 2020; February 6-8, 2020; Tunis, Tunisia. [doi: 10.1109/octa49274.2020.9151663]
40. Szomszor M, Kostkova P, de Quincey E. #Swineflu: Twitter predicts Swine Flu outbreak in 2009. In: Proceedings of the Third International Conference on Electronic Healthcare. 2010 Presented at: eHealth 2010; December 13-15, 2010; Casablanca, Morocco. [doi: 10.1007/978-3-642-23635-8_3]
41. Quincey D, Kostkova P. Early warning and outbreak detection using social networking websites: the potential of Twitter. In: Proceedings of the Second International ICST Conference on Electronic Healthcare. 2009 Presented at: eHealth 2009; September 23-25, 2009; Istanbul, Turkey. [doi: 10.1007/978-3-642-11745-9_4]
42. Chon J, Raymond R, Wang H, Wang F. Modeling flu trends with real-time geo-tagged Twitter data streams. In: Proceedings of the 10th International Conference on Wireless Algorithms, Systems, and Applications. 2015 Presented at: WASA 2015; August 10-12, 2015; Qufu, China. [doi: 10.1007/978-3-319-21837-3_7]
43. Stilo G, Velardi P, Tozzi AE, Gesualdo F. Predicting flu epidemics using Twitter and historical data. In: Proceedings of the International Conference on Brain Informatics and Health. 2014 Presented at: BIH 2014; August 11-14, 2014; Warsaw, Poland. [doi: 10.1007/978-3-319-09891-3_16]
44. Hwang MH, Wang S, Cao G, Padmanabhan A, Zhang Z. Spatiotemporal transformation of social media geostreams: a case study of Twitter for flu risk analysis. In: Proceedings of the 4th ACM SIGSPATIAL International Workshop on GeoStreaming. 2013 Presented at: IWGS '13; November 5, 2013; Orlando, FL. [doi: 10.1145/2534303.2534310]
45. Talvis K, Chorianopoulos K, Kermanidis KL. Real-time monitoring of flu epidemics through linguistic and statistical analysis of Twitter messages. In: Proceedings of the 9th International Workshop on Semantic and Social Media Adaptation and Personalization. 2014 Presented at: SMAP 2014; November 6-7, 2014; Corfu, Greece. [doi: 10.1109/smap.2014.38]
46. Lampos V, Zou B, Cox IJ. Enhancing feature selection using word embeddings: the case of flu surveillance. In: Proceedings of the 26th International Conference on World Wide Web. 2017 Presented at: WWW '17; April 3-7, 2017; Perth, Australia. [doi: 10.1145/3038912.3052622]
47. Lampos V, Bie TD, Cristianini N. Flu detector-tracking epidemics on Twitter. In: Proceedings of the Machine Learning and Knowledge Discovery in Databases. 2010 Presented at: ECML PKDD 2010; September 20-24, 2010; Barcelona, Spain. [doi: 10.1007/978-3-642-15939-8_42]
48. Byrd K, Mansurov A, Baysal O. Mining Twitter data for influenza detection and surveillance. In: Proceedings of the International Workshop on Software Engineering in Healthcare Systems. 2016 Presented at: SEHS '16; May 14-22, 2016; Austin, Texas. [doi: 10.1145/2897683.2897693]
49. Comito C, Forestiero A, Pizzuti C. Improving influenza forecasting with web-based social data. In: Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2018 Presented at: ASONAM 2018; August 28-31, 2018; Barcelona, Spain. [doi: 10.1109/asonam.2018.8508563]
50. Kostkova P, Szomszor M, St. Louis C. #swineflu: the use of Twitter as an early warning and risk communication tool in the 2009 swine flu pandemic. ACM Trans Manage Inf Syst 2014 Jul;5(2):1-25. [doi: 10.1145/2597892]
51. Abouzahra M, Tan J. Twitter vs. Zika—the role of social media in epidemic outbreaks surveillance. Health Policy Technol 2021 Mar;10(1):174-181. [doi: 10.1016/j.hlpt.2020.10.014]
52. Masri S, Jia J, Li C, Zhou G, Lee MC, Yan G, et al. Use of Twitter data to improve Zika virus surveillance in the United States during the 2016 epidemic. BMC Public Health 2019 Jun 14;19(1):761 [FREE Full text] [doi: 10.1186/s12889-019-7103-8] [Medline: 31200692]

XSL•FO
RenderX

53. Chaudhary S, Naaz S. Use of big data in computational epidemiology for public health surveillance. In: Proceedings of the International Conference on Computing and Communication Technologies for Smart Nation. 2017 Presented at: IC3TSN 2017; October 12-14, 2017; Gurgaon, India. [doi: 10.1109/ic3tsn.2017.8284467]
54. Romano S, Di Martino S, Kanhabua N, Mazzeo A, Nejdl W. Challenges in detecting epidemic outbreaks from social networks. In: Proceedings of the 30th International Conference on Advanced Information Networking and Applications Workshops. 2016 Presented at: WAINA 2016; March 23-25, 2016; Crans-Montana, Switzerland. [doi: 10.1109/waina.2016.111]
55. Ji X, Chun SA, Geller J. Epidemic outbreak and spread detection system based on Twitter data. In: Proceedings of the First International Conference on Health Information Science. 2012 Presented at: HIS 2012; April 8-10, 2012; Beijing, China. [doi: 10.1007/978-3-642-29361-0_19]
56. Feng S, Hossain L. Weibo surveillance of public awareness to Ebola disaster in China. In: Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016. 2016 Presented at: IntelliSys 2016; September 21-22, 2016; London, UK. [doi: 10.1007/978-3-319-56991-8_58]
57. Zhang Q, Gioannini C, Paolotti D, Perra N, Perrotta D, Quaggiotto M, et al. Social data mining and seasonal influenza forecasts: the FluOutlook platform. In: Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases. 2015 Presented at: ECML PKDD 2015; September 7-11, 2015; Porto, Portugal. [doi: 10.1007/978-3-319-23461-8_21]
58. Sun X, Ye J, Ren F. Real time early-stage influenza detection with emotion factors from Sina Microblog. In: Proceedings of the Fifth Workshop on South and Southeast Asian Natural Language Processing. 2014 Presented at: WSSANLP@COLING 2014; August 23, 2014; Dublin, Ireland. [doi: 10.3115/v1/w14-5511]
59. Gui X, Kou Y, Pine KH, Chen Y. Managing uncertainty: using social media for risk assessment during a public health crisis. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 6-11, 2017; Denver, CO. [doi: 10.1145/3025453.3025891]
60. Yuan M, Liu T, Yang C. Exploring the relationship among human activities, COVID-19 morbidity, and at-risk areas using location-based social media data: knowledge about the early pandemic stage in Wuhan. Int J Environ Res Public Health 2022 May 27;19(11):6523 [FREE Full text] [doi: 10.3390/ijerph19116523] [Medline: 35682104]
61. Gui X, Kou Y, Pine K, Ladaw E, Kim H, Suzuki-Gill E, et al. Multidimensional risk communication: public discourse on risks during an emerging epidemic. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI '18; April 21-26, 2018; Montreal, QC URL: https://doi.org/10.1145/3173574.3173788 [doi: 10.1145/3173574.3173788]
62. Souza RC, de Brito DE, Cardoso RL, de Oliveira DM, Meira WJ, Pappa GL. An evolutionary methodology for handling data scarcity and noise in monitoring real events from social media data. In: Proceedings of the 14th Ibero-American Conference on AI. 2014 Presented at: IBERAMIA 2014; November 24-27, 2014; Santiago de Chile, Chile. [doi: 10.1007/978-3-319-12027-0_24]
63. Souza RC, Assunção RM, de Oliveira DM, de Brito DE, Meira WJ. Infection hot spot mining from social media trajectories. In: Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases. 2016 Presented at: ECML PKDD 2016; September 19-23, 2016; Riva del Garda, Italy. [doi: 10.1007/978-3-319-46227-1_46]
64. Zhang F, Luo J, Li C, Wang X, Zhao Z. Detecting and analyzing influenza epidemics with social media in China. In: Proceedings of the 18th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. 2014 Presented at: PAKDD 2014; May 13-16, 2014; Tainan, Taiwan. [doi: 10.1007/978-3-319-06608-0_8]
65. Culotta A. Towards detecting influenza epidemics by analyzing Twitter messages. In: Proceedings of the First Workshop on Social Media Analytics. 2010 Presented at: SOMA '10; July 25-28, 2010; Washington, DC. [doi: 10.1145/1964858.1964874]
66. Wakamiya S, Kawai Y, Aramaki E. After the boom no one tweets: microblog-based influenza detection incorporating indirect information. In: Proceedings of the Sixth International Conference on Emerging Databases: Technologies, Applications, and Theory. 2016 Presented at: EDB '16; October 17-19, 2016; Jeju, Republic of Korea. [doi: 10.1145/3007818.3007822]
67. Hassan Zadeh A, Zolbanin HM, Sharda R, Delen D. Social media for nowcasting flu activity: spatio-temporal big data analysis. Inf Syst Front 2019 Jan 5;21(4):743-760. [doi: 10.1007/s10796-018-9893-0]
68. Achrekar H, Gandhe A, Lazarus R, Yu SH, Liu B. Online social networks flu trend tracker: a novel sensory approach to predict flu trends. In: Proceedings of the 5th International Joint Conference on Biomedical Engineering Systems and Technologies. 2012 Presented at: BIOSTEC 2012; February 1-4, 2012; Vilamoura, Portugal. [doi: 10.1007/978-3-642-38256-7_24]
69. Aramaki E, Maskawa S, Morita M. Twitter catches the flu: detecting influenza epidemics using Twitter. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2011 Presented at: EMNLP '11; July 27-31, 2011; Edinburgh, UK.
70. Zhang K, Arablouei R, Jurdak R. Predicting prevalence of influenza-like illness from geo-tagged tweets. In: Proceedings of the 26th International Conference on World Wide Web Companion. 2017 Presented at: WWW '17 Companion; April 3-7, 2017; Perth, Australia. [doi: 10.1145/3041021.3051150]

XSL•FO
RenderX

71. Lamb A, Paul MJ, Dredze M. Separating fact from fear: tracking flu infections on Twitter. In: Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2013 Presented at: NAACL HLT 2013; June 9-15, 2013; Atlanta, GA.
72. Sun X, Ye J, Ren F. Detecting influenza states based on hybrid model with personal emotional factors from social networks. Neurocomputing 2016 Oct;210:257-268. [doi: 10.1016/j.neucom.2016.01.107]
73. Sun X, Ye J, Ren F. Hybrid model based influenza detection with sentiment analysis from social networks. In: Proceedings of the 4th National Conference on Social Media Processing. 2015 Presented at: SMP 2015; November 16-17, 2015; Guangzhou, China. [doi: 10.1007/978-981-10-0080-5_5]
74. Yousefinaghani S, Dara R, Poljak Z, Bernardo TM, Sharif S. The assessment of Twitter's potential for outbreak detection: avian influenza case study. Sci Rep 2019 Dec 03;9(1):18147 [FREE Full text] [doi: 10.1038/s41598-019-54388-4] [Medline: 31796768]
75. Rudra K, Sharma A, Ganguly N, Imran M. Classifying and summarizing information from microblogs during epidemics. Inf Syst Front 2018 Mar 20;20(5):933-948 [FREE Full text] [doi: 10.1007/s10796-018-9844-9] [Medline: 32214879]
76. Rudra K, Sharma A, Ganguly N, Imran M. Classifying information from microblogs during epidemics. In: Proceedings of the 2017 International Conference on Digital Health. 2017 Presented at: DH '17; July 2-5, 2017; London, UK. [doi: 10.1145/3079452.3079491]
77. Ghosh S, Rudra K, Ghosh S, Ganguly N, Podder S, Balani N, et al. Identifying multi-dimensional information from microblogs during epidemics. In: Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. 2019 Presented at: CODS-COMAD '19; January 3-5, 2019; Kolkata, India. [doi: 10.1145/3297001.3297030]
78. Collier N, Son NT, Nguyen NM. OMG U got flu? Analysis of shared health messages for bio-surveillance. J Biomed Sem 2011;2(Suppl 5):S9. [doi: 10.1186/2041-1480-2-s5-s9]
79. Sousa L, de Mello R, Cedrim D, Garcia A, Missier P, Uchôa A, et al. VazaDengue: an information system for preventing and combating mosquito-borne diseases with social networks. Inf Syst 2018 Jun;75:26-42. [doi: 10.1016/j.is.2018.02.003]
80. Jain VK, Kumar S. Effective surveillance and predictive mapping of mosquito-borne diseases using social media. J Comput Sci 2018 Mar;25:406-415. [doi: 10.1016/j.jocs.2017.07.003]
81. Hong Y, Sinnott RO. A social media platform for infectious disease analytics. In: Proceedings of the 18th International Conference on Computational Science and Its Applications. 2018 Presented at: ICCSA 2018; July 2-5, 2018; Melbourne, Australia. [doi: 10.1007/978-3-319-95162-1_36]
82. Swain S, Seeja KR. Analysis of epidemic outbreak in Delhi using social media data. In: Proceedings of the Second International Conference on Information, Communication and Computing Technology. 2017 Presented at: ICICCT 2017; May 13, 2017; New Delhi, India. [doi: 10.1007/978-981-10-6544-6_3]
83. Comito C. Sensing social media to forecast COVID-19 cases. In: Proceedings of the IEEE Symposium on Computers and Communications. 2022 Presented at: ISCC 2022; June 30-July 3, 2022; Rhodes, Greece. [doi: 10.1109/iscc55528.2022.9913033]
84. Perera D, Bamunusinghe J. Contact tracing of covid-19 patients using tweets. In: Proceedings of the 2nd International Conference on Advanced Research in Computing. 2022 Presented at: ICARC 2022; February 23-24, 2022; Belihuloya, Sri Lanka. [doi: 10.1109/icarc54489.2022.9753859]
85. Lazebnik T, Bunimovich-Mendrazitsky S, Ashkenazi S, Levner E, Benis A. Early detection and control of the next epidemic wave using health communications: development of an artificial intelligence-based tool and its validation on COVID-19 data from the US. Int J Environ Res Public Health 2022 Nov 30;19(23):16023 [FREE Full text] [doi: 10.3390/ijerph192316023] [Medline: 36498096]
86. Mundotiya RK, Baruah R, Srivastava B, Singh AK. NLPRL at WNUT-2020 task 2: ELMo-based system for identification of COVID-19 tweets. In: Proceedings of the Sixth Workshop on Noisy User-Generated Text. 2020 Presented at: W-NUT 2020; November 19, 2020; Online. [doi: 10.18653/v1/2020.wnut-1.60]
87. Wang Y, Hao H, Platt LS. Examining risk and crisis communications of government agencies and stakeholders during early-stages of COVID-19 on Twitter. Comput Human Behav 2021 Jan;114:106568 [FREE Full text] [doi: 10.1016/j.chb.2020.106568] [Medline: 32982038]
88. Cornelius J, Ellendorff T, Furrer L, Rinaldi F. COVID-19 twitter monitor: aggregating and visualizing COVID-19 related trends in social media. In: Proceedings of the Fifth Social Media Mining for Health Applications Workshop & Shared Task. 2020 Presented at: SMM4H@COLING 2020; December 12, 2020; Barcelona, Spain. [doi: 10.18653/v1/w19-3208]
89. Klein AZ, Magge A, O'Connor K, Flores Amaro JI, Weissenbacher D, Gonzalez Hernandez G. Toward using Twitter for tracking COVID-19: a natural language processing pipeline and exploratory data set. J Med Internet Res 2021 Jan 22;23(1):e25314 [FREE Full text] [doi: 10.2196/25314] [Medline: 33449904]
90. Zhou Y, Jiang JY, Chen X, Wang W. #StayHome or #Marathon?: social media enhanced pandemic surveillance on spatial-temporal dynamic graphs. In: Proceedings of the 30th ACM International Conference on Information & Knowledge Management. 2021 Presented at: CIKM '21; November 1-5, 2021; Virtual Event. [doi: 10.1145/3459637.3482222]

XSL•FO
RenderX

91. Lwowski B, Najafirad P. COVID-19 surveillance through Twitter using self-supervised and few shot learning. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Presented at: COVID-19@EMNLP 2020; November 19, 2020; Virtual Event. [doi: 10.18653/v1/2020.nlpcovid19-2.9]
92. Amin S, Uddin MI, Zeb MA, Alarood AA, Mahmoud M, Alkinani MH. Detecting dengue/flu infections based on tweets using LSTM and word embedding. IEEE Access 2020;8:189054-189068. [doi: 10.1109/access.2020.3031174]
93. Khatua A, Khatua A, Cambria E. A tale of two epidemics: contextual Word2Vec for classifying Twitter streams during outbreaks. Inf Process Manage 2019 Jan;56(1):247-257. [doi: 10.1016/j.ipm.2018.10.010]
94. Yu X, Xie Z, Mashhadi A, Hong L. Multi-task models for multi-faceted classification of pandemic information on social media. In: Proceedings of the 14th ACM Web Science Conference 2022. 2022 Presented at: WebSci '22; June 26-29, 2022; Barcelona, Spain. [doi: 10.1145/3501247.3531552]
95. Wang CK, Singh O, Tang ZL, Dai HJ. Using a recurrent neural network model for classification of tweets conveyed influenza-related information. In: Proceedings of the International Workshop on Digital Disease Detection using Social Media 2017. 2017 Presented at: DDDSM-2017; November 27, 2017; Taipei, Taiwan.
96. Kia MA, Khaksefidi FE. Twitter flu trend: a hybrid deep neural network for tweet analysis. In: Proceedings of the 42nd SGAI International Conference on Artificial Intelligence. 2022 Presented at: AI 2022; December 13-15, 2022; Cambridge, UK. [doi: 10.1007/978-3-031-21441-7_3]
97. Zhao L, Chen J, Chen F, Jin F, Wang W, Lu CT, et al. Online flu epidemiological deep modeling on disease contact network. Geoinformatica 2019 Jul 25;24(2):443-475. [doi: 10.1007/s10707-019-00376-9]
98. Huo HF, Jing SL, Wang XY, Xiang H. Modeling and analysis of a H1N1 model with relapse and effect of Twitter. Physica A 2020 Dec;560:125136. [doi: 10.1016/j.physa.2020.125136]
99. Huo HF, Zhang XM. Modeling the influence of Twitter in reducing and increasing the spread of influenza epidemics. Springerplus 2016 Jan 27;5(1):88 [FREE Full text] [doi: 10.1186/s40064-016-1689-4] [Medline: 26848428]
100. Nie Q, Liu Y, Zhang D, Jiang H. Dynamical SEIR model with information entropy using COVID-19 as a case study. IEEE Trans Comput Soc Syst 2021 Aug;8(4):946-954. [doi: 10.1109/tcss.2020.3046712]
101. Spurlock K, Elgazzar H. Predicting COVID-19 infection groups using social networks and machine learning algorithms. In: Proceedings of the 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference. 2020 Presented at: UEMCO; October 28-31, 2020; New York, NY. [doi: 10.1109/uemcon51285.2020.9298093]
102. Haupt MR, Jinich-Diamant A, Li J, Nali M, Mackey TK. Characterizing Twitter user topics and communication network dynamics of the “liberate” movement during COVID-19 using unsupervised machine learning and social network analysis. Online Soc Netw Media 2021 Jan;21:100114. [doi: 10.1016/j.osnem.2020.100114]
103. Bagavathi A, Krishnan S. Social sensors early detection of contagious outbreaks in social media. In: Ahram T, editor. Advances in Artificial Intelligence, Software and Systems Engineering. Cham, Switzerland: Springer; 2019.
104. Martín G, Marinescu MC, Singh DE, Carretero J. Leveraging social networks for understanding the evolution of epidemics. BMC Syst Biol 2011 Dec 23;5(S3):S14. [doi: 10.1186/1752-0509-5-s3-s14]
105. Zheng H, Goh DH, Lee EW, Lee CS, Theng YL. Uncovering topics related to COVID-19 pandemic on Twitter. In: Proceedings of the 22nd International Conference on Asia-Pacific Digital Libraries. 2020 Presented at: ICADL 2020; November 30-December 1, 2020; Kyoto, Japan. [doi: 10.1007/978-3-030-64452-9_28]
106. Shin M, Han S, Park S, Cha M. A risk communication event detection model via contrastive learning. In: Proceedings of the 3rd NLP4IF Workshop on NLP for Internet Freedom. 2020 Presented at: NLP4IF 2020; December 12, 2020; Barcelona, Spain.
107. Chen L, Hossain KS, Butler P, Ramakrishnan N, Prakash BA. Flu gone viral: syndromic surveillance of flu on Twitter using temporal topic models. In: Proceedings of the IEEE International Conference on Data Mining. 2014 Presented at: ICDM 2014; December 14-17, 2014; Shenzhen, China. [doi: 10.1109/icdm.2014.137]
108. Fu Q, Hu C, Xu W, He X, Zhang T. Detect and analyze flu outlier events via social network. In: Proceedings of the Web Technologies and Applications. 2014 Presented at: APWeb 2014; September 5, 2014; Changsha, China. [doi: 10.1007/978-3-319-11119-3_13]
109. Nolasco D, Oliveira J. Subevents detection through topic modeling in social media posts. Future Gener Comput Syst 2019 Apr;93:290-303. [doi: 10.1016/j.future.2018.09.008]
110. Zheng H, Goh DH, Lee CS, Lee EW, Theng YL. Uncovering temporal differences in COVID-19 tweets. Proc Assoc Inf Sci Technol 2020 Oct 22;57(1):e233 [FREE Full text] [doi: 10.1002/pra2.233] [Medline: 33173810]
111. Missier P, Romanovsky A, Miu T, Pal A, Daniilakis M, Garcia A, et al. Tracking dengue epidemics using Twitter content classification and topic modelling. In: Proceedings of the Current Trends in Web Engineering. 2016 Presented at: ICWE 2016; June 6-9, 2016; Lugano, Switzerland. [doi: 10.1007/978-3-319-46963-8_7]
112. Odlum M, Yoon S. What can we learn about the Ebola outbreak from tweets? Am J Infect Control 2015 Jun;43(6):563-571 [FREE Full text] [doi: 10.1016/j.ajic.2015.02.023] [Medline: 26042846]
113. Diaz-Aviles E, Stewart A. Tracking Twitter for epidemic intelligence: case study: EHEC/HUS outbreak in Germany, 2011. In: Proceedings of the 4th Annual ACM Web Science Conference. 2012 Presented at: WebSci '12; June 22-24, 2012; Evanston, Illinois. [doi: 10.1145/2380718.2380730]

XSL•FO
RenderX

114. Chew C, Eysenbach G. Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. PLoS One 2010 Nov 29;5(11):e14118 [FREE Full text] [doi: 10.1371/journal.pone.0014118] [Medline: 21124761]
115. Lehmann BA, Ruiter RA, Kok G. A qualitative study of the coverage of influenza vaccination on Dutch news sites and social media websites. BMC Public Health 2013 Jun 05;13:547 [FREE Full text] [doi: 10.1186/1471-2458-13-547] [Medline: 23738769]
116. Kobayashi R, Takedomi Y, Nakayama Y, Suda T, Uno T, Hashimoto T, et al. Evolution of public opinion on COVID-19 vaccination in Japan: large-scale Twitter data analysis. J Med Internet Res 2022 Dec 22;24(12):e41928 [FREE Full text] [doi: 10.2196/41928] [Medline: 36343186]
117. Martin S, Vanderslott S. "Any idea how fast 'It's just a mask!' can turn into 'It's just a vaccine!'": from mask mandates to vaccine mandates during the COVID-19 pandemic. Vaccine 2022 Dec 05;40(51):7488-7499 [FREE Full text] [doi: 10.1016/j.vaccine.2021.10.031] [Medline: 34823912]
118. Zhang C, Xu S, Li Z, Liu G, Dai D, Dong C. The evolution and disparities of online attitudes toward COVID-19 vaccines: year-long longitudinal and cross-sectional study. J Med Internet Res 2022 Jan 21;24(1):e32394 [FREE Full text] [doi: 10.2196/32394] [Medline: 34878410]
119. Argyris YA, Zhang N, Bashyal B, Tan PN. Using deep learning to identify linguistic features that facilitate or inhibit the propagation of anti- and pro-vaccine content on social media. In: Proceedings of the IEEE International Conference on Digital Health. 2022 Presented at: ICDH 2022; July 10-16, 2022; Barcelona, Spain. [doi: 10.1109/icdh55609.2022.00025]
120. Islam T, Goldwasser D. Understanding COVID-19 vaccine campaign on Facebook using minimal supervision. In: Proceedings of the IEEE International Conference on Big Data. 2022 Presented at: BigData 2022; December 17-20, 2022; Osaka, Japan. [doi: 10.1109/bigdata55660.2022.10021123]
121. Durmaz N, Hengirmen E. The dramatic increase in anti-vaccine discourses during the COVID-19 pandemic: a social network analysis of Twitter. Hum Vaccin Immunother 2022 Dec 31;18(1):2025008 [FREE Full text] [doi: 10.1080/21645515.2021.2025008] [Medline: 35113767]
122. Basch CH, Meleo-Erwin Z, Fera J, Jaime C, Basch CE. A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok. Hum Vaccin Immunother 2021 Aug 03;17(8):2373-2377 [FREE Full text] [doi: 10.1080/21645515.2021.1894896] [Medline: 33764283]
123. Kariyapperuma KR, Banujan K, Wijeratna PM, Kumara BT. Classification of covid19 vaccine-related tweets using deep learning. In: Proceedings of the International Conference on Data Analytics for Business and Industry. 2022 Presented at: ICDABI 2022; October 25-26, 2022; Sakhir, Bahrain. [doi: 10.1109/icdabi56818.2022.10041615]
124. Stella M, Vitevitch MS, Botta F. Cognitive networks extract insights on COVID-19 vaccines from English and Italian popular tweets: anticipation, logistics, conspiracy and loss of trust. Big Data Cogn Comput 2022 May 12;6(2):52. [doi: 10.3390/bdcc6020052]
125. Jabalameli S, Xu Y, Shetty S. Spatial and sentiment analysis of public opinion toward COVID-19 pandemic using twitter data: at the early stage of vaccination. Int J Disaster Risk Reduct 2022 Oct 01;80:103204 [FREE Full text] [doi: 10.1016/j.ijdrr.2022.103204] [Medline: 35935613]
126. Feizollah A, Anuar NB, Mehdi R, Firdaus A, Sulaiman A. Understanding COVID-19 halal vaccination discourse on Facebook and Twitter using aspect-based sentiment analysis and text emotion analysis. Int J Environ Res Public Health 2022 May 21;19(10):6269 [FREE Full text] [doi: 10.3390/ijerph19106269] [Medline: 35627806]
127. Khan K, Yadav S. Sentiment analysis on covid-19 vaccine using Twitter data: a NLP approach. In: Proceedings of the IEEE 9th Region 10 Humanitarian Technology Conference. 2021 Presented at: R10-HTC 2021; September 30-October 2, 2021; Bangalore, India. [doi: 10.1109/r10-htc53172.2021.9641515]
128. Ji X, Chun SA, Wei Z, Geller J. Twitter sentiment classification for measuring public health concerns. Soc Netw Anal Min 2015;5(1):13 [FREE Full text] [doi: 10.1007/s13278-015-0253-5] [Medline: 32226558]
129. Wang S, Schraagen M, Sang ET, Dastani M. Public sentiment on governmental COVID-19 measures in Dutch social media. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Presented at: EMNLP 2020; November 16-20, 2020; Online. [doi: 10.18653/v1/2020.nlpcovid19-2.17]
130. Zhunis A, Lima G, Song H, Han J, Cha M. Emotion bubbles: emotional composition of online discourse before and after the COVID-19 outbreak. In: Proceedings of the ACM Web Conference 2022. 2022 Presented at: WWW '22; April 25-29, 2022; Virtual event. [doi: 10.1145/3485447.3512132]
131. Muñoz LM, Ramirez MF, Camargo JE. A data-driven method for measuring the negative impact of sentiment towards China in the context of COVID-19. In: Proceedings of the Third International Conference on Applied Informatics. 2020 Presented at: ICAI 2020; October 29-31, 2020; Ota, Nigeria. [doi: 10.1007/978-3-030-61702-8_15]
132. Shah CS, Sebastian MP. Sentiment analysis and topic modelling of Indian government's Twitter handle #IndiaFightsCorona. In: Proceedings of the International Conference on Transfer and Diffusion of IT. 2020 Presented at: TDIT 2020; December 18-19, 2020; Tiruchirappalli, India. [doi: 10.1007/978-3-030-64861-9_30]
133. Garcia K, Berton L. Topic detection and sentiment analysis in Twitter content related to COVID-19 from Brazil and the USA. Appl Soft Comput 2021 Mar;101:107057 [FREE Full text] [doi: 10.1016/j.asoc.2020.107057] [Medline: 33519326]

XSL•FO
RenderX

134. Massaro M, Tamburro P, La Torre M, Dal Mas F, Thomas R, Cobianchi L, et al. Non-pharmaceutical interventions and the infodemic on Twitter: lessons learned from Italy during the covid-19 pandemic. J Med Syst 2021 Mar 06;45(4):50 [FREE Full text] [doi: 10.1007/s10916-021-01726-7] [Medline: 33675427]
135. Shi J, Li W, Yang Y, Yao N, Bai Q, Yongchareon S, et al. Automated concern exploration in pandemic situations - COVID-19 as a use case. In: Proceedings of the 17th Pacific Rim Knowledge Acquisition Workshop. 2020 Presented at: PKAW 2020; January 7-8, 2021; Yokohama, Japan. [doi: 10.1007/978-3-030-69886-7_15]
136. Chen L, Huang X, Zhang H, Niu B. Covid-19 public opinion analysis based on LDA topic modeling and data visualization. In: Proceedings of the Third International Conference on Machine Learning for Cyber Security. 2020 Presented at: ML4CS 2020; October 8-10, 2020; Guangzhou, China. [doi: 10.1007/978-3-030-62460-6_20]
137. Zhu B, Zheng X, Liu H, Li J, Wang P. Analysis of spatiotemporal characteristics of big data on social media sentiment with COVID-19 epidemic topics. Chaos Solitons Fractals 2020 Nov;140:110123 [FREE Full text] [doi: 10.1016/j.chaos.2020.110123] [Medline: 32834635]
138. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top concerns of tweeters during the COVID-19 pandemic: infoveillance study. J Med Internet Res 2020 Apr 21;22(4):e19016 [FREE Full text] [doi: 10.2196/19016] [Medline: 32287039]
139. Whitfield C, Liu Y, Anwar M. Surveillance of COVID-19 pandemic using social media: a reddit study in North Carolina. In: Proceedings of the 12th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics. 2021 Presented at: BCB '21; August 1-4, 2021; Gainesville, FL. [doi: 10.1145/3459930.3469550]
140. Kirabo L, Namara M, Mcneese N. The power of the blue tick (): Ugandans' experiences and engagement on Twitter at the onset of the COVID-19 pandemic. In: Proceedings of the 3rd African Human-Computer Interaction Conference: Inclusiveness and Empowerment. 2021 Presented at: AfriCHI '21; March 8-12, 2021; Maputo, Mozambique. [doi: 10.1145/3448696.3448746]
141. Thakur N, Han CY. An exploratory study of tweets about the SARS-CoV-2 Omicron variant: insights from sentiment analysis, language interpretation, source tracking, type classification, and embedded URL detection. COVID 2022 Jul 25;2(8):1026-1049. [doi: 10.3390/covid2080076]
142. Avasthi S, Chauhan R, Acharjya DP. Information extraction and sentiment analysis to gain insight into the COVID-19 crisis. In: Proceedings of the International Conference on Innovative Computing and Communications. 2021 Presented at: ICICC 2021; February 20-21, 2021; Delhi, India. [doi: 10.1007/978-981-16-2594-7_28]
143. Sabareesha SS, Bhattacharjee S, Shetty RD. Pattern analysis of COVID-19 based on geotagged social media data with sociodemographic factors. In: Proceedings of the 27th International Conference on Automation and Computing. 2022 Presented at: ICAC 2022; September 1-3, 2022; Bristol, UK. [doi: 10.1109/icac55051.2022.9911119]
144. Hoque MU, Lee K, Beyer JL, Curran SR, Gonser KS, Lam NS, et al. Analyzing tweeting patterns and public engagement on Twitter during the recognition period of the COVID-19 pandemic: a study of two U.S. States. IEEE Access 2022;10:72879-72894. [doi: 10.1109/access.2022.3189670]
145. Koukaras P, Tjortjis C, Rousidis D. Mining association rules from COVID-19 related Twitter data to discover word patterns, topics and inferences. Inf Syst 2022 Nov;109:102054 [FREE Full text] [doi: 10.1016/j.is.2022.102054] [Medline: 36569358]
146. Comito C. How COVID-19 information spread in U.S.? The role of Twitter as early indicator of epidemics. IEEE Trans Serv Comput 2022 May 1;15(3):1193-1205. [doi: 10.1109/tsc.2021.3091281]
147. Purnat TD, Vacca P, Burzo S, Zecchin T, Wright A, Briand S, et al. WHO digital intelligence analysis for tracking narratives and information voids in the COVID-19 infodemic. Stud Health Technol Inform 2021 May 27;281:989-993. [doi: 10.3233/SHTI210326] [Medline: 34042821]
148. Lorenzoni V, Andreozzi G, Bazzani A, Casigliani V, Pirri S, Tavoschi L, et al. How Italy tweeted about COVID-19: detecting reactions to the pandemic from social media. Int J Environ Res Public Health 2022 Jun 24;19(13):7785 [FREE Full text] [doi: 10.3390/ijerph19137785] [Medline: 35805444]
149. Pang PC, Jiang W, Pu G, Chan KS, Lau Y. Social media engagement in two governmental schemes during the COVID-19 pandemic in Macao. Int J Environ Res Public Health 2022 Jul 23;19(15):8976 [FREE Full text] [doi: 10.3390/ijerph19158976] [Medline: 35897346]
150. Abramova O, Batzel K, Modesti D. Collective response to the health crisis among German Twitter users: a structural topic modeling approach. Int J Inf Manage Data Insights 2022 Nov;2(2):100126. [doi: 10.1016/j.jjimei.2022.100126]
151. Wang T, Brooks I, Bashir M. Public reaction on social media during COVID-19: a comparison between Twitter and Weibo. In: Proceedings of the 2021 Intelligent Computing Conference. 2021 Presented at: IC 2021; July 15-16, 2021; Online. [doi: 10.1007/978-3-030-80119-9_38]
152. Bogović PK, Meštrović A, Martinčić-Ipšić S. Topic modeling for tracking COVID-19 communication on Twitter. In: Proceedings of the 28th International Conference on Information and Software Technologies. 2022 Presented at: ICIST 2022; October 13-15, 2022; Kaunas, Lithuania.
153. León-Sandoval E, Zareei M, Barbosa-Santillán LI, Falcón Morales LE, Pareja Lora A, Ochoa Ruiz G. Monitoring the emotional response to the COVID-19 pandemic using sentiment analysis: a case study in Mexico. Comput Intell Neurosci 2022;2022:4914665 [FREE Full text] [doi: 10.1155/2022/4914665] [Medline: 35634092]

XSL•FO
RenderX

154. Perera S, Perera I, Ahangama S. Exploring Twitter messages during the COVID-19 pandemic in Sri Lanka: topic modelling and emotion analysis. In: Proceedings of the 2nd International Conference on Advanced Research in Computing. 2022 Presented at: ICARC 2022; February 23-24, 2022; Belihuloya, Sri Lanka. [doi: 10.1109/icarc54489.2022.9753866]
155. Abro HU, Shah ZS, Abbasi H. Analysis of COVID-19 effects on wellbeing - study of Reddit posts using natural language processing techniques. In: Proceedings of the International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering. 2022 Presented at: ETECTE 2022; December 2-4, 2022; Lahore, Pakistan. [doi: 10.1109/etecte55893.2022.10007300]
156. Sreeraag G, Shynu PG. A comparative analysis of tweets from the South Indian states based on COVID-19 Omicron wave. In: Proceedings of the International Conference on Computing, Communication, Security and Intelligent Systems. 2022 Presented at: IC3SIS 2022; June 23-25, 2022; Kochi, India. [doi: 10.1109/ic3sis54991.2022.9885333]
157. Fukuda S, Nanba H, Shoji H. Changes in interests and emotional responses to news coverage of coronavirus disease 2019 case numbers over time. In: Proceedings of the Joint 12th International Conference on Soft Computing and Intelligent Systems and 23rd International Symposium on Advanced Intelligent Systems. 2022 Presented at: SCIS&ISIS 2022; November 29-December 2, 2022; Ise, Japan. [doi: 10.1109/scisisis55246.2022.10001983]
158. Nandan M, Mitra S, Dey S. Tweet classification and sentiment analysis of Covid 19 epidemic by applying hybrid based techniques. In: Proceedings of the IEEE VLSI Device Circuit and System. 2022 Presented at: VLSI DCS 2022; February 26-27, 2022; Kolkata, India. [doi: 10.1109/vlsidcs53788.2022.9811438]
159. Qazi A, Qazi J, Naseer K, Zeeshan M, Khan B, Dey S. Sentiment analysis of nationwide lockdown amid COVID 19: evidence from Pakistan. In: Proceedings of the IEEE 7th International Conference on Information Technology and Digital Applications. 2022 Presented at: ICITDA 2022; November 4-5, 2022; Yogyakarta, Indonesia. [doi: 10.1109/icitda55840.2022.9971397]
160. Pitroda H. Opinion mining approach: to understand public reactions towards reopening of schools and colleges/universities in India. In: Proceedings of the 2nd International Conference on Innovative Practices in Technology and Management. 2022 Presented at: ICIPTM 2022; February 23-25, 2022; Gautam Buddha Nagar, India. [doi: 10.1109/iciptm54933.2022.9754208]
161. Chang A, Xian X, Liu MT, Zhao X. Health communication through positive and solidarity messages amid the COVID-19 pandemic: automated content analysis of Facebook uses. Int J Environ Res Public Health 2022 May 19;19(10):6159 [FREE Full text] [doi: 10.3390/ijerph19106159] [Medline: 35627696]
162. Yang W, Wu Z, Mok NY, Ma X. How to save lives with microblogs? Lessons from the usage of Weibo for requests for medical assistance during COVID-19. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022 Presented at: CHI '22; April 29-May 5, 2022; New Orleans, LA. [doi: 10.1145/3491102.3517591]
163. Maylawati DS, Ramdhani MA. Indonesian citizens' health behavior in a pandemics: Twitter conversation analysis using latent Dirichlet allocation. In: Proceedings of the 8th International Conference on Wireless and Telematics. 2022 Presented at: ICWT 2022; July 21-22, 2022; Yogyakarta, Indonesia. [doi: 10.1109/icwt55831.2022.9935502]
164. Li L, Aldosery A, Vitiugin F, Nathan N, Novillo-Ortiz D, Castillo C, et al. The response of governments and public health agencies to COVID-19 pandemics on social media: a multi-country analysis of Twitter discourse. Front Public Health 2021 Sep 28;9:716333 [FREE Full text] [doi: 10.3389/fpubh.2021.716333] [Medline: 34650948]
165. Neves JC, de França TC, Bastos MP, de Carvalho PV, Gomes JO. Analysis of government agencies and stakeholders' Twitter communications during the first surge of COVID-19 in Brazil. Work 2022 Oct 19;73(s1):S81-S93. [doi: 10.3233/WOR-211213] [Medline: 36214022]
166. Wahbeh A, Nasralah T, Al-Ramahi M, El-Gayar O. Mining physicians' opinions on social media to obtain insights into COVID-19: mixed methods analysis. JMIR Public Health Surveill 2020 Jun 18;6(2):e19276 [FREE Full text] [doi: 10.2196/19276] [Medline: 32421686]
167. Addawood A, Alsuwailem A, Alohali A, Alajaji D, Alturki M, Alsuhaibani J, et al. Tracking and understanding public reaction during COVID-19: Saudi Arabia as a use case. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Presented at: COVID-19@EMNLP 2020; November 16-20, 2020; Online. [doi: 10.18653/v1/2020.nlpcovid19-2.24]
168. Comito C. COVID-19 concerns in US: topic detection in Twitter. In: Proceedings of the 25th International Database Engineering & Applications Symposium. 2021 Presented at: IDEAS '21; July 14-16, 2021; Montreal, QC. [doi: 10.1145/3472163.3472169]
169. Kumari KR, Ratna T, Gayathri T. Machine learning technique with spider monkey optimization for COVID-19 sentiment analysis. In: Proceedings of the International Conference on Computing, Communication and Power Technology. 2022 Presented at: IC3P 2022; January 7-8, 2022; Visakhapatnam, India. [doi: 10.1109/ic3p52835.2022.00070]
170. Paramarta V, Mailangkay A, Amalia H, Chrismas D. Public sentiment analysis of Indonesian tweets about COVID-19 vaccination using different machine learning approaches. In: Proceedings of the Seventh International Conference on Informatics and Computing. 2022 Presented at: ICIC 2022; December 8-9, 2022; Bali, Indonesia. [doi: 10.1109/icic56845.2022.10006896]

XSL•FO
RenderX

171. Bhatia S, Alhaider M, Alarjani M. Sentiment analysis for Arabic tweets on covid-19 using computational techniques. In: Proceedings of the 12th International Conference on Cloud Computing, Data Science & Engineering. 2022 Presented at: Confluence 2022; January 27-28, 2022; Noida, India. [doi: 10.1109/confluence52989.2022.9734188]
172. Arias F, Guerra-Adames A, Zambrano ME, Quintero-Guerra E, Tejedor-Flores N. Analyzing Spanish-language public sentiment in the context of a pandemic and social unrest: the Panama case. Int J Environ Res Public Health 2022 Aug 19;19(16):10328 [FREE Full text] [doi: 10.3390/ijerph191610328] [Medline: 36011965]
173. Elsaka T, Afyouni I, Hashem I, Al Aghbari Z. Spatio-temporal sentiment mining of COVID-19 Arabic social media. ISPRS Int J Geo Inf 2022 Sep 02;11(9):476. [doi: 10.3390/ijgi11090476]
174. Rahmanti AR, Chien CH, Nursetyo AA, Husnayain A, Wiratama BS, Fuad A, et al. Social media sentiment analysis to monitor the performance of vaccination coverage during the early phase of the national COVID-19 vaccine rollout. Comput Methods Programs Biomed 2022 Jun;221:106838 [FREE Full text] [doi: 10.1016/j.cmpb.2022.106838] [Medline: 35567863]
175. Kanakaraddi S, Chikaraddi AK, Aivalli N, Maniyar J, Singh N. Sentiment analysis of Covid-19 tweets using machine learning and natural language processing. In: Pandian AP, Narayanan M, Palanisamy R, Senjyu T, editors. Proceedings of Third International Conference on Intelligent Computing, Information and Control Systems. Singapore, Singapore: Springer; 2022.
176. Li L, Zhang Q, Wang X, Zhang J, Wang T, Gao TL, et al. Characterizing the propagation of situational information in social media during COVID-19 epidemic: a case study on Weibo. IEEE Trans Comput Soc Syst 2020 Apr;7(2):556-562. [doi: 10.1109/tcss.2020.2980007]
177. Guo R, Xu K. A large-scale analysis of COVID-19 Twitter dataset in a new phase of the pandemic. In: Proceedings of the IEEE 12th International Conference on Electronics Information and Emergency Communication. 2022 Presented at: ICEIEC 2022; July 15-17, 2022; Beijing, China. [doi: 10.1109/iceiec54567.2022.9835047]
178. Buvanasri AK, Meenakshi R, Karthika S. Applications of open source intelligence in crisis analysis—a COVID-19 case study. In: Proceedings of the ICT Systems and Sustainability. 2021 Presented at: ICT4SD 2021; July 22-23, 2021; Goa, India. [doi: 10.1007/978-981-16-5987-4_32]
179. Satu MS, Khan MI, Mahmud M, Uddin S, Summers MA, Quinn JM, et al. TClustVID: a novel machine learning classification model to investigate topics and sentiment in COVID-19 tweets. Knowl Based Syst 2021 Aug 17;226:107126 [FREE Full text] [doi: 10.1016/j.knosys.2021.107126] [Medline: 33972817]
180. Mandal S, Rath M, Wang Y, Patra BG. Predicting Zika prevention techniques discussed on Twitter: an exploratory study. In: Proceedings of the 2018 Conference on Human Information Interaction & Retrieval. 2018 Presented at: CHIIR '18; March 11-15, 2018; New Brunswick, NJ. [doi: 10.1145/3176349.3176874]
181. Al-Ramahi M, Elnoshokaty A, El-Gayar O, Nasralah T, Wahbeh A. Public discourse against masks in the COVID-19 era: infodemiology study of Twitter data. JMIR Public Health Surveill 2021 Apr 05;7(4):e26780 [FREE Full text] [doi: 10.2196/26780] [Medline: 33720841]
182. Liang G, Zhao J, Lau HY, Leung CW. Using social media to analyze public concerns and policy responses to COVID-19 in Hong Kong. ACM Trans Manag Inf Syst 2021 Sep 08;12(4):1-20. [doi: 10.1145/3460124]
183. Li L, Hua L, Gao F. Int J Environ Res Public Health 2022 Dec 31;20(1):780 [FREE Full text] [doi: 10.3390/ijerph20010780] [Medline: 36613100]
184. Mahor K, Manjhvar AK. Public sentiment assessment of coronavirus-specific tweets using a transformer-based BERT classifier. In: Proceedings of the International Conference on Edge Computing and Applications. 2022 Presented at: ICECAA 2022; October 13-15, 2022; Tamilnadu, India. [doi: 10.1109/icecaa55415.2022.9936448]
185. Rifat N, Ahsan M, Gomes R, Chowdhury M. COVID-19 sentiment analysis applying BERT. In: Proceedings of the IEEE International Conference on Electro Information Technology. 2022 Presented at: eIT 2022; May 19-21, 2022; Mankato, MN. [doi: 10.1109/eit53891.2022.9813777]
186. Dewangan A, Sharma A, Badholia A. COVID-19 tweet analysis using deep convolutional neural network (DCNN)-BERT. In: Proceedings of the 4th International Conference on Inventive Research in Computing Applications. 2022 Presented at: ICIRCA 2022; September 21-23, 2022; Coimbatore, India. [doi: 10.1109/icirca54612.2022.9985742]
187. Stoikos S, Izbicki M. Multilingual emoticon prediction of tweets about COVID-19. In: Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media. 2020 Presented at: PEOPLES 2020; December 13, 2020; Online.
188. Feldman P, Tiwari S, Cheah CS, Foulds JR, Pan S. Analyzing COVID-19 tweets with transformer-based language models. arXiv Preprint posted online on April 20, 2021 [FREE Full text]
189. Ghanem A, Asaad C, Hafidi H, Moukafih Y, Guermah B, Sbihi N, et al. Real-time infoveillance of moroccan social media users' sentiments towards the COVID-19 pandemic and its management. Int J Environ Res Public Health 2021 Nov 19;18(22):12172 [FREE Full text] [doi: 10.3390/ijerph182212172] [Medline: 34831927]
190. Lyu X, Chen Z, Wu D, Wang W. Sentiment analysis on Chinese Weibo regarding COVID-19. In: Proceedings of the 9th CCF International Conference on Natural Language Processing and Chinese Computing. 2020 Presented at: NLPCC 2020; October 14-18, 2020; Zhengzhou, China. [doi: 10.1007/978-3-030-60450-9_56]

XSL•FO
RenderX

191. Pan Z, Zhou S, Sun N. Twitter sentiment analysis of COVID-19 vaccine based on BiLSTM with attention mechanism. In: Proceedings of the 4th International Conference on Advances in Computer Technology, Information Science and Communications. 2022 Presented at: CTISC 2022; April 22-24, 2022; Suzhou, China. [doi: 10.1109/ctisc54888.2022.9849814]
192. De Santis E, Martino A, Rizzi A. An infoveillance system for detecting and tracking relevant topics from Italian tweets during the COVID-19 event. IEEE Access 2020;8:132527-132538. [doi: 10.1109/access.2020.3010033]
193. Alsudias L, Rayson P. COVID-19 and Arabic Twitter: how can Arab world governments and public health organizations learn from social media? In: Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020. 2020 Presented at: NLP COVID-19 ACL 2020; July 5-10, 2020; Online.
194. Umer M, Sadiq S, Karamti H, Abdulmajid Eshmawi A, Nappi M, Usman Sana M, et al. ETCNN: extra tree and convolutional neural network-based ensemble model for COVID-19 tweets sentiment classification. Pattern Recognit Lett 2022 Dec;164:224-231 [FREE Full text] [doi: 10.1016/j.patrec.2022.11.012] [Medline: 36407854]
195. Malo S, Bayala TR, Kinda Z. French COVID-19 tweets classification using FlauBERT layers. In: Proceedings of the Congress on Intelligent Systems. 2021 Presented at: CIS 2021; September 4-5, 2021; Bengaluru, India. [doi: 10.1007/978-981-16-9416-5_59]
196. Miao L, Last M, Litvak M. Tracking social media during the COVID-19 pandemic: the case study of lockdown in New York State. Expert Syst Appl 2022 Jan;187:115797 [FREE Full text] [doi: 10.1016/j.eswa.2021.115797] [Medline: 34566273]
197. Luo L, Wang Y, Liu H. COVID-19 personal health mention detection from tweets using dual convolutional neural network. Expert Syst Appl 2022 Aug 15;200:117139 [FREE Full text] [doi: 10.1016/j.eswa.2022.117139] [Medline: 35399189]
198. Guidry JP, Jin Y, Orr CA, Messner M, Meganck S. Ebola on Instagram and Twitter: how health organizations address the health crisis in their social media engagement. Public Relat Rev 2017 Sep;43(3):477-486. [doi: 10.1016/j.pubrev.2017.04.009]
199. Roy M, Moreau N, Rousseau C, Mercier A, Wilson A, Atlani-Duault L. Ebola and localized blame on social media: analysis of Twitter and Facebook conversations during the 2014-2015 Ebola epidemic. Cult Med Psychiatry 2020 Mar 18;44(1):56-79 [FREE Full text] [doi: 10.1007/s11013-019-09635-8] [Medline: 31214902]
200. Chetty N, Alathur S, Kumar V. 2019-nCoV disease control and rehabilitation: insights from Twitter analytics. In: Proceedings of the 5th International Conference on Computing, Communication and Security. 2020 Presented at: ICCCS 2020; October 14-16, 2020; Patna, India. [doi: 10.1109/icccs49678.2020.9277142]
201. Shanthakumar SG, Seetharam A, Ramesh A. Understanding the socio-economic disruption in the United States during COVID-19's early days. arXiv Preprint posted online on April 11, 2020 [FREE Full text]
202. Niknam F, Samadbeik M, Fatehi F, Shirdel M, Rezazadeh M, Bastani P. COVID-19 on Instagram: a content analysis of selected accounts. Health Policy Technol 2021 Mar;10(1):165-173 [FREE Full text] [doi: 10.1016/j.hlpt.2020.10.016] [Medline: 33723502]
203. Essam BA, Abdo MS. How do Arab tweeters perceive the COVID-19 pandemic? J Psycholinguist Res 2021 Jun 14;50(3):507-521 [FREE Full text] [doi: 10.1007/s10936-020-09715-6] [Medline: 32797330]
204. Ölcer S, Yilmaz-Aslan Y, Brzoska P. Lay perspectives on social distancing and other official recommendations and regulations in the time of COVID-19: a qualitative study of social media posts. BMC Public Health 2020 Jun 19;20(1):963 [FREE Full text] [doi: 10.1186/s12889-020-09079-5] [Medline: 32560716]
205. Yigitcanlar T, Kankanamge N, Preston A, Gill PS, Rezayee M, Ostadnia M, et al. How can social media analytics assist authorities in pandemic-related policy decisions? Insights from Australian states and territories. Health Inf Sci Syst 2020 Dec 15;8(1):37 [FREE Full text] [doi: 10.1007/s13755-020-00121-9] [Medline: 33078073]
206. Xia H, An W, Li J, Zhang ZJ. Outlier knowledge management for extreme public health events: understanding public opinions about COVID-19 based on microblog data. Socioecon Plann Sci 2022 Mar;80:100941 [FREE Full text] [doi: 10.1016/j.seps.2020.100941] [Medline: 32921839]
207. Huang F, Ding H, Liu Z, Wu P, Zhu M, Li A, et al. How fear and collectivism influence public's preventive intention towards COVID-19 infection: a study based on big data from the social media. BMC Public Health 2020 Nov 16;20(1):1707 [FREE Full text] [doi: 10.1186/s12889-020-09674-6] [Medline: 33198699]
208. Machete P, Turpin M. The use of critical thinking to identify fake news: a systematic literature review. In: Responsible Design, Implementation and Use of Information and Communication Technology. Cham, Switzerland: Springer International Publishing; 2020.
209. Guo Q, Huang W, Huang K, Liu X. Information credibility: a probabilistic graphical model for identifying credible influenza posts on social media. In: Proceedings of the International Conference on Smart Health. 2015 Presented at: ICSH 2015; November 17-18, 2015; Phoenix, AZ. [doi: 10.1007/978-3-319-29175-8_12]
210. Abdelminaam D, Ismail F, Taha M, Taha A, Houssein E, Nabil A. CoAID-DEEP: an optimized intelligent framework for automated detecting COVID-19 misleading information on Twitter. IEEE Access 2021;9:27840-27867. [doi: 10.1109/access.2021.3058066]
211. Madani Y, Erritali M, Bouikhalene B. Using artificial intelligence techniques for detecting Covid-19 epidemic fake news in Moroccan tweets. Results Phys 2021 Jun;25:104266. [doi: 10.1016/j.rinp.2021.104266]
212. Cui L, Lee D. Coaid: Covid-19 healthcare misinformation dataset. arXiv Preprint posted online on May 22, 2020 [FREE Full text]

XSL•FO
RenderX

213. Hossain T, Logan RLIV, Ugarte A, Matsubara Y, Young S, Singh S. COVIDLies: detecting COVID-19 misinformation on social media. In: Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020. 2020 Presented at: COVID-19@EMNLP 2020; November 16-20, 2020; Online. [doi: 10.18653/v1/2020.nlpcovid19-2.11]
214. Malla S, Alphonse PJ. Fake or real news about COVID-19? Pretrained transformer model to detect potential misleading news. Eur Phys J Spec Top 2022 Jan 13;231(18-20):3347-3356 [FREE Full text] [doi: 10.1140/epjs/s11734-022-00436-6] [Medline: 35039760]
215. Alajramy L, Jarrar R. Using artificial neural networks to identify COVID-19 misinformation. In: Proceedings of the 4th Multidisciplinary International Symposium on Disinformation in Open Online Media. 2022 Presented at: MISDOOM 2022; October 11-12, 2022; Boise, ID. [doi: 10.1007/978-3-031-18253-2_2]
216. Samantaray S, Kumar A. Bi-directional long short-term memory network for fake news detection from social media. In: Proceedings of the Intelligent and Cloud Computing. 2021 Presented at: ICICC 2021; February 20-21, 2021; Delhi, India. [doi: 10.1007/978-981-16-9873-6_42]
217. Ghayoomi M, Mousavian M. Deep transfer learning for COVID-19 fake news detection in Persian. Expert Syst 2022 Apr 03;39(8):e13008 [FREE Full text] [doi: 10.1111/exsy.13008] [Medline: 35599852]
218. Raj C, Meel P. ARCNN framework for multimodal infodemic detection. Neural Netw 2022 Feb;146:36-68 [FREE Full text] [doi: 10.1016/j.neunet.2021.11.006] [Medline: 34839091]
219. Tang C, Ma K, Cui B, Ji K, Abraham A. Long text feature extraction network with data augmentation. Appl Intell (Dordr) 2022 Apr 04;52(15):17652-17667 [FREE Full text] [doi: 10.1007/s10489-022-03185-0] [Medline: 35400845]
220. Zuo C, Banerjee R, Chaleshtori FH, Shirazi H, Ray I. Seeing should probably not be believing: the role of deceptive support in COVID-19 misinformation on Twitter. J Data Inf Quality 2022 Dec 28;15(1):1-26. [doi: 10.1145/3546914]
221. Biradar S, Saumya S, Chauhan A. Combating the infodemic: COVID-19 induced fake news recognition in social media networks. Complex Intell Systems 2023 Feb 18;9(3):2879-2891 [FREE Full text] [doi: 10.1007/s40747-022-00672-2] [Medline: 35194546]
222. Jeyasudha J, Seth P, Usha G, Tanna P. Fake information analysis and detection on pandemic in Twitter. SN Comput Sci 2022 Aug 24;3(6):456 [FREE Full text] [doi: 10.1007/s42979-022-01363-y] [Medline: 36035506]
223. Alouffi B, Alharbi A, Sahal R, Saleh H. An optimized hybrid deep learning model to detect COVID-19 misleading information. Comput Intell Neurosci 2021 Nov 15;2021(1):9615034 [FREE Full text] [doi: 10.1155/2021/9615034] [Medline: 34790233]
224. Hussna A, Trisha II, Karim MS, Alam MG. COVID-19 fake news prediction on social media data. In: Proceedings of the IEEE Region 10 Symposium. 2021 Presented at: TENSYMP 2021; August 23-25, 2021; Jeju, Republic of Korea. [doi: 10.1109/tensymp52854.2021.9550957]
225. Ajao O, Garg A, Da Costa-Abreu M. Exploring content-based and meta-data analysis for detecting fake news infodemic: a case study on COVID-19. In: Proceedings of the 12th International Conference on Pattern Recognition Systems. 2022 Presented at: ICPRS 2022; June 7-10, 2022; Saint-Etienne, France. [doi: 10.1109/icprs54038.2022.9854058]
226. Alhakami H, Alhakami W, Baz A, Faizan M, Khan MW, Agrawal A. Evaluating intelligent methods for detecting COVID-19 fake news on social media platforms. Electronics 2022 Aug 03;11(15):2417. [doi: 10.3390/electronics11152417]
227. Ali M, Murtza I, Ejaz A. Pandemic rumor identification on social networking sites: a case study of COVID-19. In: Proceedings of the 2021 5th International Conference on Natural Language Processing and Information Retrieval. 2021 Presented at: NLPIR '21; December 17-20, 2021; Sanya, China. [doi: 10.1145/3508230.3508246]
228. Maakoul O, Boucht S, Hachimi KE, Azzouzi S. Towards evaluating the COVID'19 related fake news problem: case of Morocco. In: Proceedings of the IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science. 2020 Presented at: ICECOCS 2020; December 2-3, 2020; Kenitra, Morocco. [doi: 10.1109/icecocs50124.2020.9314517]
229. Mehta V, Mishra RK. Machine learning based fake news detection on Covid-19 tweets data. In: Proceedings of International Conference on Computational Intelligence and Data Engineering. 2021 Presented at: ICCIDE 2021; August 13-14, 2021; Andhra Pradesh, India. [doi: 10.1007/978-981-16-7182-1_8]
230. Khan S, Hakak S, Deepa N, Prabadevi B, Dev K, Trelova S. Detecting COVID-19-related fake news using feature extraction. Front Public Health 2021 Jan 4;9:788074 [FREE Full text] [doi: 10.3389/fpubh.2021.788074] [Medline: 35059379]
231. Jouyandeh F, Rahmatikargar S, Rahmatikargar B, Zadeh PM. Fake news and COVID-19 vaccination: a comparative study. In: Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2021 Presented at: ASONAM '21; November 8-11, 2021; Virtual Event, Netherlands. [doi: 10.1145/3487351.3490960]
232. Bamiro B, Assayad I. Data-based automatic Covid-19 rumors detection in social networks. In: Proceedings of the Networking, Intelligent Systems and Security. 2021 Presented at: NISS 2021; April 1-2, 2021; Kenitra, Morocco. [doi: 10.1007/978-981-16-3637-0_57]
233. Zhang Y, Guo B, Ding Y, Liu J, Qiu C, Liu S, et al. Investigation of the determinants for misinformation correction effectiveness on social media during COVID-19 pandemic. Inf Process Manag 2022 May;59(3):102935 [FREE Full text] [doi: 10.1016/j.ipm.2022.102935] [Medline: 35400028]
234. Belay EG, Beyene M, Alemu T, Negash A, Tesema T, Mohammed A, et al. Towards curtailing infodemic in the era of COVID-19: a contextualized solution for Ethiopia. In: Proceedings of the HCI International 2020 – Late Breaking Papers:

XSL•FO
RenderX

Interaction, Knowledge and Social Media. 2020 Presented at: HCII 2020; July 19-24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-60152-2_17]
235. Sell TK, Hosangadi D, Trotochaud M. Misinformation and the US Ebola communication crisis: analyzing the veracity and content of social media messages related to a fear-inducing infectious disease outbreak. BMC Public Health 2020 May 07;20(1):550 [FREE Full text] [doi: 10.1186/s12889-020-08697-3] [Medline: 32375715]
236. Alsudias L, Rayson P. Classifying information sources in Arabic Twitter to support online monitoring of infectious diseases. In: Proceedings of the 3rd Workshop on Arabic Corpus Linguistics. 2019 Presented at: WACL-3 2019; July 22, 2019; Cardiff, UK.
237. Jain L. An entropy-based method to control COVID-19 rumors in online social networks using opinion leaders. Technol Soc 2022 Aug;70:102048 [FREE Full text] [doi: 10.1016/j.techsoc.2022.102048] [Medline: 35765463]
238. Chen B, Chen X, Pan J, Liu K, Xie B, Wang W, et al. Dissemination and refutation of rumors during the COVID-19 outbreak in China: infodemiology study. J Med Internet Res 2021 Feb 15;23(2):e22427 [FREE Full text] [doi: 10.2196/22427] [Medline: 33493124]
239. Zeng J, Zhang Y, Ma X. Fake news detection for epidemic emergencies via deep correlations between text and images. Sustain Cities Soc 2021 Mar;66:102652 [FREE Full text] [doi: 10.1016/j.scs.2020.102652] [Medline: 36570569]
240. Atehortua NA, Patino S. COVID-19, a tale of two pandemics: novel coronavirus and fake news messaging. Health Promot Int 2021 Apr 15;36(2):524-534 [FREE Full text] [doi: 10.1093/heapro/daaa140] [Medline: 33450022]
241. Madraki G, Grasso I, Otala JM, Liu Y, Matthews J. Characterizing and comparing COVID-19 misinformation across languages, countries and platforms. In: Proceedings of the Companion Proceedings of the Web Conference 2021. 2021 Presented at: WWW '21; April 19-23, 2021; Ljubljana, Slovenia. [doi: 10.1145/3442442.3452304]
242. Raj C, Meel P. People lie, actions Don't! Modeling infodemic proliferation predictors among social media users. Technol Soc 2022 Feb;68:101930. [doi: 10.1016/j.techsoc.2022.101930]
243. Heidari M, Zad S, Hajibabaee P, Malekzadeh M, HekmatiAthar S, Uzuner O. BERT model for fake news detection based on social bot activities in the COVID-19 pandemic. In: Proceedings of the IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference. 2021 Presented at: UEMCON 2021; December 1-4, 2021; New York, NY. [doi: 10.1109/uemcon53757.2021.9666618]
244. Yan M, Lin YR, Chung WT. Are mutated misinformation more contagious? A case study of COVID-19 misinformation on Twitter. In: Proceedings of the 14th ACM Web Science Conference 2022. 2022 Presented at: WebSci '22; June 26-29, 2022; Barcelona, Spain. [doi: 10.1145/3501247.3531562]
245. Shahi GK, Dirkson A, Majchrzak TA. An exploratory study of COVID-19 misinformation on Twitter. Online Soc Netw Media 2021 Mar;22:100104 [FREE Full text] [doi: 10.1016/j.osnem.2020.100104] [Medline: 33623836]
246. Valecha R, Srinivasan SK, Volety T, Kwon KH, Agrawal M, Rao HR. Fake news sharing: an investigation of threat and coping cues in the context of the Zika virus. Digital Threats 2021 Apr 15;2(2):1-16. [doi: 10.1145/3410025]
247. Singh L, Bode L, Budak C, Kawintiranon K, Padden C, Vraga E. Understanding high- and low-quality URL sharing on COVID-19 Twitter streams. J Comput Soc Sci 2020 Nov 27;3(2):343-366 [FREE Full text] [doi: 10.1007/s42001-020-00093-6] [Medline: 33263092]
248. Sacco PL, Gallotti R, Pilati F, Castaldo N, De Domenico M. Emergence of knowledge communities and information centralization during the COVID-19 pandemic. Soc Sci Med 2021 Sep;285:114215 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114215] [Medline: 34364158]
249. Ermakova L, Nurbakova D, Ovchinnikova I. Covid or not Covid? Topic shift in information cascades on Twitter. In: Proceedings of the 3rd International Workshop on Rumours and Deception in Social Media. 2020 Presented at: RDSM 2020; December 12, 2020; Barcelona, Spain.
250. Fan Y. The social amplification of risk on Weibo-take the COVID-19 epidemic in China as an example. In: Proceedings of the 2020 Conference on Artificial Intelligence and Healthcare. 2020 Presented at: CAIH2020; October 23-25, 2020; Taiyuan, China. [doi: 10.1145/3433996.3434013]
251. Chen E, Lerman K, Ferrara E. Tracking social media discourse about the COVID-19 pandemic: development of a public coronavirus Twitter data set. JMIR Public Health Surveill 2020 May 29;6(2):e19273 [FREE Full text] [doi: 10.2196/19273] [Medline: 32427106]
252. Bahja M, Safdar GA. Unlink the link between COVID-19 and 5G networks: an NLP and SNA based approach. IEEE Access 2020;8:209127-209137. [doi: 10.1109/access.2020.3039168]
253. Ahmed W, Vidal-Alaball J, Downing J, López Seguí F. COVID-19 and the 5G conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020 May 06;22(5):e19458 [FREE Full text] [doi: 10.2196/19458] [Medline: 32352383]
254. Shahsavari S, Holur P, Wang T, Tangherlini TR, Roychowdhury V. Conspiracy in the time of corona: automatic detection of emerging COVID-19 conspiracy theories in social media and the news. J Comput Soc Sci 2020;3(2):279-317 [FREE Full text] [doi: 10.1007/s42001-020-00086-5] [Medline: 33134595]
255. Lanier HD, Diaz MI, Saleh SN, Lehmann CU, Medford RJ. Analyzing COVID-19 disinformation on Twitter using the hashtags #scamdemic and #plandemic: retrospective study. PLoS One 2022 Jun 22;17(6):e0268409 [FREE Full text] [doi: 10.1371/journal.pone.0268409] [Medline: 35731785]

XSL•FO
RenderX

256. Ahmed W, López Seguí F, Vidal-Alaball J, Katz MS. COVID-19 and the "film your hospital" conspiracy theory: social network analysis of Twitter data. J Med Internet Res 2020 Oct 05;22(10):e22374 [FREE Full text] [doi: 10.2196/22374] [Medline: 32936771]
257. Li S, Wang Y, Xue J, Zhao N, Zhu T. The impact of COVID-19 epidemic declaration on psychological consequences: a study on active Weibo users. Int J Environ Res Public Health 2020 Mar 19;17(6):2032 [FREE Full text] [doi: 10.3390/ijerph17062032] [Medline: 32204411]
258. Fatimah N, Budi I, Santoso AB, Putra PK. Analysis of mental health during the Covid-19 pandemic in Indonesia using Twitter data. In: Proceedings of the 8th International Conference on Advanced Informatics: Concepts, Theory and Applications. 2021 Presented at: ICAICTA 2021; September 29-30, 2021; Bandung, Indonesia. [doi: 10.1109/icaicta53211.2021.9640265]
259. Davis BD, McKnight DE, Teodorescu D, Quan-Haase A, Chunara R, Fyshe A, et al. Quantifying depression-related language on social media during the COVID-19 pandemic. Int J Popul Data Sci 2020 Mar 30;5(4):1716 [FREE Full text] [doi: 10.23889/ijpds.v5i4.1716] [Medline: 35516163]
260. Wolohan JT. Estimating the effect of COVID-19 on mental health: linguistic indicators of depression during a global pandemic. In: Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020. 2020 Presented at: NLP COVID-19 ACL 2020; July 5-10, 2020; Online.
261. Chen JQ, Qi K, Zhang A, Shalaginov M, Zeng TH. COVID-19 impact on mental health analysis based on Reddit comments. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine. 2022 Presented at: BIBM 2022; December 6-8, 2022; Las Vegas, NV. [doi: 10.1109/bibm55620.2022.9995512]
262. Nandy S, Kumar V. My mind is a prison: a boosted deep learning approach to detect the rise in depression since COVID-19 using a stacked bi-LSTM CatBoost model. In: Proceedings IEEE International Conference on Big Data. 2021 Presented at: BigData 2021; December 15-18, 2021; Orlando, FL. [doi: 10.1109/bigdata52589.2021.9671482]
263. Viviani M, Crocamo C, Mazzola M, Bartoli F, Carrà G, Pasi G. Assessing vulnerability to psychological distress during the COVID-19 pandemic through the analysis of microblogging content. Future Gener Comput Syst 2021 Dec;125:446-459 [FREE Full text] [doi: 10.1016/j.future.2021.06.044] [Medline: 34934256]
264. Kumar R, Mukherjee S, Choi TM, Dhamotharan L. Mining voices from self-expressed messages on social-media: diagnostics of mental distress during COVID-19. Decis Support Syst 2022 Nov;162:113792 [FREE Full text] [doi: 10.1016/j.dss.2022.113792] [Medline: 35542965]
265. Gao J, Zheng P, Jia Y, Chen H, Mao Y, Chen S, et al. Mental health problems and social media exposure during COVID-19 outbreak. PLoS One 2020 Apr 16;15(4):e0231924 [FREE Full text] [doi: 10.1371/journal.pone.0231924] [Medline: 32298385]
266. Pang H. How compulsive WeChat use and information overload affect social media fatigue and well-being during the COVID-19 pandemic? A stressor-strain-outcome perspective. Telemat Inform 2021 Nov;64:101690 [FREE Full text] [doi: 10.1016/j.tele.2021.101690] [Medline: 36567817]
267. Ni MY, Yang L, Leung CM, Li N, Yao XI, Wang Y, et al. Mental health, risk factors, and social media use during the COVID-19 epidemic and cordon sanitaire among the community and health professionals in Wuhan, China: cross-sectional survey. JMIR Ment Health 2020 May 12;7(5):e19009 [FREE Full text] [doi: 10.2196/19009] [Medline: 32365044]
268. Neill R, Blair C, Best P, McGlinchey E, Armour C. Media consumption and mental health during COVID-19 lockdown: a UK cross-sectional study across England, Wales, Scotland and Northern. PsyArXiv Preprints Preprint posted online July 1, 2020 [FREE Full text] [doi: 10.31234/osf.io/5d9fc]
269. Chao M, Xue D, Liu T, Yang H, Hall BJ. Media use and acute psychological outcomes during COVID-19 outbreak in China. J Anxiety Disord 2020 Aug;74:102248 [FREE Full text] [doi: 10.1016/j.janxdis.2020.102248] [Medline: 32505918]
270. Daimer S, Mihatsch LL, Neufeld SA, Murray GK, Knolle F. Investigating the relationship of COVID-19 related stress and media consumption with schizotypy, depression, and anxiety in cross-sectional surveys repeated throughout the pandemic in Germany and the UK. Elife 2022 Jul 04;11:e75893 [FREE Full text] [doi: 10.7554/eLife.75893] [Medline: 35781372]
271. Yu L, Du M. Social networking use, mental health, and quality of life of Hong Kong adolescents during the COVID-19 pandemic. Front Public Health 2022 Oct 31;10:1040169 [FREE Full text] [doi: 10.3389/fpubh.2022.1040169] [Medline: 36388293]
272. Brailovskaia J, Margraf J. Positive mental health and mindfulness as protective factors against addictive social media use during the COVID-19 outbreak. PLoS One 2022 Nov 30;17(11):e0277631 [FREE Full text] [doi: 10.1371/journal.pone.0277631] [Medline: 36449483]
273. Zhao N, Zhou G. Social media use and mental health during the COVID-19 pandemic: moderator role of disaster stressor and mediator role of negative affect. Appl Psychol Health Well Being 2020 Dec;12(4):1019-1038 [FREE Full text] [doi: 10.1111/aphw.12226] [Medline: 32945123]
274. Lazer D, Kennedy R, King G, Vespignani A. Big data. The parable of Google Flu: traps in big data analysis. Science 2014 Mar 14;343(6176):1203-1205. [doi: 10.1126/science.1248506] [Medline: 24626916]
275. ChaiAsaad / SLR_EpidemicsAndSocialMedia. GitHub. URL: https://github.com/ChaiAsaad/SLR_EpidemicsAndSocialMedia [accessed 2024-12-24]

XSL•FO
RenderX

Abbreviations

BERT: bidirectional encoder representations from transformers
DL: deep learning
ILI: influenza-like illness
LDA: latent Dirichlet allocation
LR: logistic regression
LSTM: long short-term memory
ML: machine learning
NB: naive Bayes
PHEIC: public health emergency of international concern
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RF: random forest
RQ: research question
SVM: support vector machine
WHO: World Health Organization

Edited by A Mavragani; submitted 19.12.23; peer-reviewed by Y Wei, I Gryech, S Gordon; comments to author 06.02.24; revised version received 25.03.24; accepted 22.05.24; published 03.02.25.

Please cite as:
Asaad C, Khaouja I, Ghogho M, Baïna K
When Infodemic Meets Epidemic: Systematic Literature Review
JMIR Public Health Surveill 2025;11:e55642
URL: https://publichealth.jmir.org/2025/1/e55642
doi:10.2196/55642
PMID:

©Chaimae Asaad, Imane Khaouja, Mounir Ghogho, Karim Baïna. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 03.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Exploring the Use of Digital Educational Tools for Sexual and Reproductive Health in Sub-Saharan Africa: Systematic Review

Ramatu Hajia Abdul Hamid Alhassan[1], MPH; Catherine L Haggerty[1], MPH, PhD; Abimbola Fapohunda[2], MS, MPH, DrPH; Nabeeha Jabir Affan[1], MPH; Martina Anto-Ocrah[1,3], MPH, PhD

[1]Department of Epidemiology, University of Pittsburgh School of Public Health, Pittsburgh, PA, United States

[2]Department of Behavioral and Community Health Sciences, University of Pittsburgh School of Public Health, Pittsburgh, PA, United States

[3]Department of Medicine, Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

Corresponding Author:
Martina Anto-Ocrah, MPH, PhD
Department of Medicine
Division of General Internal Medicine
University of Pittsburgh School of Medicine
230 McKee Place
Pittsburgh, PA, 15213
United States
Phone: 1 000 000 0000
Email: maa509@pitt.edu

Abstract

Background: Adolescents, particularly those in Sub-Saharan Africa, experience major challenges in getting accurate and comprehensive sexual and reproductive health (SRH) information because of sociocultural norms, stigma, and limited SRH educational resources. Digital educational tools, leveraging the widespread use of mobile phones and internet connectivity, present a promising avenue to overcome these barriers and enhance SRH education among adolescents in Sub-Saharan Africa.

Objective: We conducted a systematic review to describe (1) the geographic and demographic distributions (designated objectives 1a and 1b, respectively, given their interrelatedness) and (2) the types and relevant impacts of digital educational tools (objective 2).

Methods: We followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, using databases, such as Ovid-MEDLINE, Google Scholar, PubMed, and ERIC, to conduct literature searches. The selection criteria focused on studies that specifically addressed digital educational tools used to assess or deliver SRH education, their implementation, and their effectiveness among the adolescent population in Sub-Saharan Africa. We used the JBI critical appraisal tools for the quality assessment of papers included in the review.

Results: The review identified 22 studies across Sub-Saharan Africa that met the inclusion criteria. The 22 studies spanned populations in West, Central, East, and South Africa, with an emphasis on youth and adolescents aged 10-24 years, reflecting the critical importance of reaching these age groups with effective, accessible, and engaging health education (objectives 1a and 1b). There was a diverse range of digital tools used, including social media platforms, mobile apps, and gamified learning experiences, for a broad age range of adolescent youth. These methods were generally successful in engaging adolescents by providing them with accessible and relevant SRH information (objective 2). However, challenges, such as the digital divide, the cultural sensitivity of the material, and the necessity for a thorough examination of the long-term influence of these tools on behavior modification, were noted.

Conclusions: Digital educational tools provide great potential to improve SRH education among adolescents in Sub-Saharan Africa. These technologies can help enhance relevant health outcomes and accessibility by delivering information that is easy to understand, interesting, and tailored to their needs. Future research should focus on addressing the identified challenges, including bridging the digital divide, ensuring cultural and contextual relevance of content, and assessing the long-term impact of digital SRH education on adolescent behavior and health outcomes. Policymakers and educators are encouraged to integrate digital tools into SRH educational strategies that target adolescents in order to improve the SRH of this age group and contribute to improving public health in Sub-Saharan Africa.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e63309) doi:10.2196/63309

KEYWORDS

digital health; adolescents; Africa; sexual health; reproductive health; human-centered design

Introduction

According to the World Health Organization, there are more than 1.2 billion adolescents aged 10 to 19 years with sexual and reproductive health (SRH) problems, such as sexually transmitted infections (STIs) and unplanned pregnancies, many of whom face significant barriers to accessing SRH-related education. With the increasing rates of SRH problems and mental health issues among adolescents, there is a pressing global need for effective educational tools that address the SRH needs of adolescents [1,2]. There are global disparities in adolescents' access to SRH education, with those in low- and middle-income countries (LMICs) facing the greatest barriers to accessing knowledge and resources due to sociocultural norms, beliefs, practices, and stigma, and yet, these parts of the world have the greatest burden of poor adolescent SRH outcomes [3]. For example, over 80% of all adolescents living with HIV are in Sub-Saharan Africa [4], which also has the highest prevalence of adolescent pregnancy globally [5]. Melesse et al [6] conducted a review of demographic and health survey data from 33 African countries and reported that despite improvements in SRH among adolescents in Sub-Saharan Africa, young people in this part of the world continue to have disproportionately high rates of unwanted pregnancies and STIs (including HIV/AIDS). These persistently high rates of poor health outcomes have been attributed, at least in part, to limited access to SRH education, more formally referred to as comprehensive sexuality education, in many Sub-Saharan African countries [7]. Challenges impacting access to sex education in Sub-Saharan Africa include cultural and societal norms that serve as barriers to youths' access to information, especially taboos regarding the open discussion of topics such as family planning, menstruation, and contraception [8]. Additionally, comprehensive sexuality education programs have been constrained by inconsistent funding and poor accountability [9]. It is crucial to promote SRH education to combat the challenges faced by adolescents and foster a more inclusive and equitable society as these young people transition into adulthood. Additionally, universal access to SRH is an essential component of sustainable development goals [10].

Digital educational tools have the potential to effectively provide SRH education for adolescents in Sub-Saharan Africa. Such technologies use the widespread availability of mobile phones and internet infrastructure to innovatively address the unique challenges faced by the youth in accessing current information and resources [11,12]. The use of social media platforms, websites, games, and mobile apps has become prevalent in digital education. These tools can provide instructional content with the potential to stimulate teenagers' interest in learning and promote favorable attitudes and knowledge that may lead to desirable SRH outcomes. The integration of digital tools into SRH education offers a promising avenue to overcome health disparities [13,14]. Mobile health (mHealth), which is the practice of medicine and public health that uses mobile devices, has exploded over the last few years, with mobile phone penetration reaching 90% in some LMICs [15]. The increased usage of mobile phones among younger populations in LMICs presents a valuable opportunity to use mHealth, for example, as a tool to overcome barriers to accessing SRH information and service [16]. Digital technology has the potential to close the information gap for adolescents in LMICs, ensuring equitable access to essential health knowledge for all adolescents irrespective of their geographical location or socioeconomic status.

We have conducted a systematic review to describe: (1) the geographic and demographic distribution (designated objectives 1a and 1b, respectively, given their interrelatedness) and (2) the types and formats of digital educational tools available for delivering SRH education to adolescents in Sub-Saharan Africa and, as appropriate, the impacts of these digital innovations on target populations (objective 2). A systematic review is appropriate for this work as it allows us to comprehensively synthesize the current state of the literature on such an underresearched topic, identify gaps in knowledge, and provide insights on moving the field forward [17]. The public health significance of this review lies in its potential to inform policy and practice by highlighting successful digital interventions that can enhance SRH education for adolescents. Given the increasing prevalence of internet and mobile device usage in Sub-Saharan Africa, identifying effective educational strategies is crucial for addressing the high rates of unintended pregnancies, STIs, and other reproductive health challenges faced by young people in the region. Ultimately, our findings can be used to support the development of targeted evidence-based digital interventions that can significantly improve health outcomes for adolescents in this part of the world.

Methods

Definitions of Digital Tools and Adolescents

We have defined the terms "digital tools" and "adolescents" as presented below.

Digital Tools

Digital tools are technological resources or applications used in the provision of information, support, and resources related to SRH. These tools use digital technology to enhance access, outreach, and effectiveness in the delivery of SRH education and services. They are available through a variety of digital platforms, such as websites, mobile apps, and social media platforms [11]. The use of such tools has become increasingly prevalent in recent years as they offer a convenient and efficient means of accessing information and resources related to SRH. These tools have gained considerable popularity owing to their ease of use and convenience in terms of obtaining information related to SRH.

XSL•FO
RenderX

Adolescents

Adolescents, who are individuals in the intermediate stage of development between childhood and adulthood, commonly referred to as teenagers or young adults, experience substantial transformations in physical structure, thinking, interpersonal relationships, and emotions [18]. As adolescents face major educational and health challenges, especially in the context of SRH-related information, our review is focused on this age group. We used the World Health Organization's definition of "adolescents" as individuals aged 10 to 19 years, "youth" as those aged 15 to 24 years, and "young people" as those aged 10 to 24 years [19]. Therefore, if any studies included these age groups, we erred on the side of caution and included them to ensure we were being as comprehensive as possible in capturing the breadth and scope of the research in this field.

Database and Search Dates

This systematic review adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) criteria [20] to ensure that all relevant primary materials match the review requirements [21]. Additionally, we used the JBI critical appraisal tools for the quality assessment of all papers included in the review [22] (Multimedia Appendix 1). A systematic search of peer-reviewed articles was conducted in November 2023 using Ovid-MEDLINE, with the assistance of a University of Pittsburgh Health Sciences Library System (HSLS) professional, to identify original research articles published between 2013 and 2024. The first Ovid-MEDLINE search was developed using a mix of Medical Subject Heading (MeSH) phrases, keywords, and Boolean operators (eg, "adolescent* or teen* or youth"; "digital health education/ or health promotion/ or sex education/"; "Africa or Angola or Benin or Botswana or Burkina Faso or Burundi or Cabo Verde or Cameroon or Central African Republic or Chad or Comoros or Congo or Cote d'Ivoire or Djibouti or Eritrea or Eswatini or Ethiopia or Gabon or Gambia or Ghana or Accra or Kumasi or Tamale or Sekondi-Takoradi or Guinea or Kenya or Lesotho or Liberia or Madagascar or Malawi or Mali or Mauritania or Mozambique or Namibia or Niger or Nigeria or Rwanda or Sao Tome or Principe or Senegal or Seychelles or Sierra Leone or Somalia or Sudan or Tanzania or Togo or Uganda or Zambia or Zimbabwe.ti,ab,kf.") (Multimedia Appendix 1). The search results were managed and tracked in SciWheel.

Study Selection

We restricted the search to studies conducted in humans and written in the English language. Additional articles were identified by cross-referencing bibliography lists. In alignment with the PRISMA standards, the methodology encompasses 4 phases: identification of sources, assessment of eligibility, detailed screening, and final selection for inclusion [21]. Initially, potential sources were selected by examining their titles and abstracts, ensuring relevance to the research question. This meticulous process facilitated the systematic evaluation and filtering of records, governed by rigorously defined inclusion and exclusion criteria. Articles that advanced beyond the preliminary screening phase were subjected to a comprehensive review of their full texts to determine their eligibility for inclusion in the systematic review. Articles were first reviewed by the first author (RHAHA) in collaboration with the co-author NJA and under the supervision of the second author (CLH) to ensure they met the eligibility criteria. The JBI critical appraisal tools for the quality assessment of all papers included in the review were further used to ensure the papers met the necessary standards for research and publication (internally valid, appropriately randomized, participants appropriately consented, etc). The eligibility criteria are presented in Textbox 1.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Region: Focuses on the Sub-Saharan African region or has significant relevance to Sub-Saharan African populations
- Target population: Adolescents and youth
- Area of focus: Addresses sexual and reproductive health, including sexually transmitted infection education, mental health, sexuality, and positive sexuality
- Mode of delivery: Digital tools, including apps, websites, gaming, e-learning modules, social media, etc
- Timeline: Published within the last 10 years (2013-2024) to ensure relevance
- Language: In English to ensure comprehension and interpretation by research teams who are primarily English speaking

Exclusion criteria

- Region: Excluded Sub-Saharan African populations or did not have a Sub-Saharan Africa focus
- Target population: Outside of the age range of adolescents and young people
- Area of focus: Does not specifically address sexual and reproductive health
- Mode of delivery: Excludes digital tools
- Timeline: Outside of the specified dates
- Language: Not in English

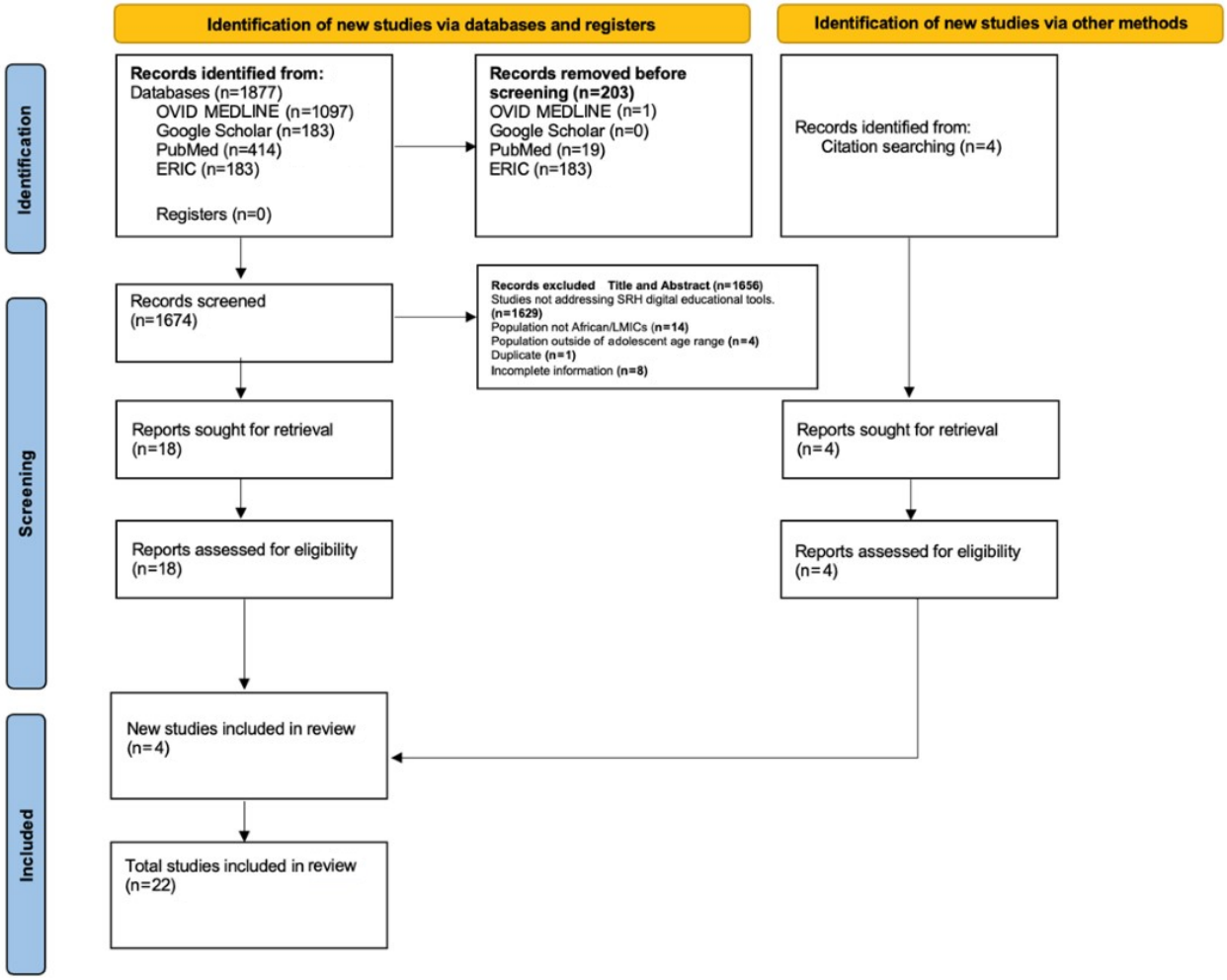
XSL•FO
RenderX

Results

Included Studies

Figure 1 outlines the systematic and comprehensive search and selection process employed in this systematic review, focusing on digital educational tools for adolescent SRH in Sub-Saharan Africa. Initially, a total of 1877 records were identified through meticulous searches across major databases, including OVID-MEDLINE, Google Scholar, PubMed, and ERIC. An additional 4 records were identified through citation searching, augmenting the pool of potential studies. Before screening, 203 records were removed due to duplication, leaving 1674 records for detailed screening. The screening process, grounded in rigorously defined inclusion and exclusion criteria, led to the exclusion of 1656 records for reasons such as not directly addressing SRH digital educational tools, not focusing on the Sub-Saharan African region or LMIC context, targeting populations outside the adolescent age range and young people, and duplications. After the exclusion process, 18 reports were retrieved and assessed for eligibility, and 4 additional studies were identified and included in the review alongside the initially identified studies, bringing the total number of studies included in the review to 22 (Figure 1; Table 1) [22].

Figure 1. Flowchart for our systematic review of digital educational tools for adolescent sexual and reproductive health (SRH) in Sub-Saharan Africa. LMIC: low- and middle-income country.



XSL•FO
RenderX

Table 1. Summary of studies on the availability of digital sexual and reproductive health education for adolescents in Sub-Saharan Africa.

Study (authors and year)	Study population/region/country	Study design	Digital tool type	Findings
Dulli et al [23], 2020	349 youth aged 15-24 years from Nigeria	Randomized controlled trial. Participants were randomized into 2 groups: those receiving the SMART Connections intervention and those receiving standard care services.	Social media	Primary results: The intervention did not significantly improve retention (50% vs 45% retention in treatment vs control groups) or social support but did significantly improve HIV-related knowledge (t=–2.96; P=.003).
Fakoya et al [24], 2022	111 adolescents aged 15-19 years from Nigeria, Tanzania, and Ethiopia	Application of youth-led participatory action research (PAR) approaches within human-centered design (HCD)	Youth-engaged version of HCD as part of the Adolescents 360 initiative	Primary results: Successful engagement of youth as project partners and action researchers. Identified opportunities to improve program empathy and responsiveness. Challenges in recruiting “extreme users” due to high competencies needed in HCD. Empathy and design standards during prototyping helped in decision-making. Real-world testing of services and products emphasized the importance of continuing youth-adult partnership.
Feroz et al [16], 2021	Young people (adolescents and youth) aged 10-24 years in low- and middle-income countries (LMICs), specifically Sub-Saharan Africa and Asia	Systematic review	Mobile phones used for mHealth interventions aimed at improving sexual and reproductive health (SRH) outcomes	Primarily used for client education and behavior change communication (n=14, 93%), followed by financial transactions and incentives, and data collection and reporting. Studies evaluated the effect mHealth interventions had on access to SRH services (n=9) and SRH outcomes (n=6). mHealth interventions improved access to SRH services and outcomes by addressing barriers such as provider prejudice, stigmatization, discrimination, and issues of privacy and confidentiality. Included decreased technological literacy, inferior network coverage, and lower linguistic competency.
Gonsalves et al [25], 2018	705 young people aged 13-24 years from Kenya and Peru	Protocol of an open, 3-arm, individually randomized trial. Participants will be randomized into one of three arms: (1) intervention arm receiving SRH information (using ARMADILLO) via mobile phones, (2) control arm receiving no intervention, and (3) a third arm with variations by site (Kenya and Peru).	Mobile phone sexual health information	If proven to be effective, interventions like ARMADILLO can bridge an important gap in achieving universal access to SRH information and education for an otherwise difficult-to-reach group. Outcomes will be assessed via questionnaires administered at baseline, intervention end, and 8 weeks after intervention end. Statistical analysis will include comparisons of proportions (chi-square tests) and means (t-tests) between arms and difference-in-difference techniques.
Haruna et al [26], 2019	348 secondary school participants (students) aged 11-15 years from 3 schools were recruited to participate in this study from Tanzania	Design-based research (DBR)	Digital gamified learning platform named “My Future Begins Today”	A paired t-test revealed a statistically significant improvement in sexual health literacy scores from pretest (mean 26.40, SD 7.29) to posttest (mean 74.12, SD 16.21) (t_{347}=52.230; P<.000), with an average increase of 47.72 points. A 1-way ANOVA showed no significant differences among the 3 groups’ pretest scores, but significant differences were found in posttest scores ($F_{2,345}$=210.43; P<.001), with follow-up Tukey post hoc tests indicating better performance by the gamified learning platform group compared to the traditional teaching group. Positive stakeholder and participant feedback on the gamified approach. Effective use of participatory design in developing educational tools.

XSL•FO
RenderX

Study (authors and year)	Study population/region/country	Study design	Digital tool type	Findings
Nalwanga et al [27], 2021	1086 Kyambogo University students aged 18-30 years from Uganda	Cross-sectional analysis of data from an endline survey of a randomized controlled trial (RCT) and data from use of a mobile phone app over a 6-month period.	Mobile phone app	The mobile phone app demonstrated predominantly positive (responsiveness, nondistracting in-app advertisements, and ease of use) attributes. 86% (n=464) of students who received a recommendation to download the app accepted and downloaded it. Of these students, 81% (n=374) used the app to access SRH information, goods, and services over 6 months. In terms of responsiveness, 55.1% (n=206) of students stated that clinics were responsive (*P*=.11). 75.4% (n=282) said the in-app advertisements were not distracting (*P*=.23). 44.4% (n=166) found that in-app instructions were very easy to use (*P*=.50).
Ippoliti et al [28], 2017	Adolescents and youth aged 10-24 years from Africa (67%), Eurasia (26%), and Latin America (13%)	Global landscape analysis. The authors issued a global call for project resources in 2014 and reviewed the submissions to confirm they met specific inclusion criteria.	Mobile phones (including SMS text messaging and mobile phone apps)	Most projects (n=12, 70%) relied on text messaging to transmit SRH information. The majority of projects were based in Africa (67%), followed by Eurasia (26%) and Latin America (13%). Mobile phones were effectively used to increase the reach of SRH information and services, especially in conservative societies where SRH topics are stigmatized. 70% of the projects relied on text messaging to transmit SRH information. Several projects have been adapted and scaled to other countries, demonstrating the scalability of mHealth interventions.
L'Engle et al [29], 2016	Review targeted adolescents aged 10-24 years and included studies that provided results from mobile phone interventions designed to improve adolescent SRH	Systematic review	Mobile phone interventions, primarily text messaging	Most programs (n=23, 82%) used text messaging. An average of 41% of essential mHealth criteria were met. An average of 82% of methodological reporting criteria were met. Evidence suggests improvements in health promotion campaigns, sexually transmitted infection (STI) screening and follow-up, and medication adherence.
Nolan et al [30], 2020	6000 youth aged 12-19 years across 8 districts in Rwanda	3-arm cluster-randomized. Arms: (1) CyberRwanda self-service, (2) CyberRwanda facilitated, and (3) control schools, which will receive the standard services that are available in the community.	CyberRwanda, a digital health intervention consisting of interactive stories, questions and answers, videos, and an online shop for health products	The study protocol describes the design and intended impact evaluation of the CyberRwanda program using a 3-arm cluster randomized noninferiority trial, but it does not report on the outcomes or effectiveness of the intervention. The primary outcomes are measurements of uptake of a modern method of contraception, initiation of childbearing, and HIV testing, all measured at the participant level. Data analysis will consist of a generalized linear mixed model.
Onukwugha et al [31], 2022	Review includes individuals aged 10-19 years, with consideration for interventions focusing on young people aged 10-24 years in Sub-Saharan Africa	Systematic review	mHealth interventions	mHealth interventions were effective in improving adolescents' uptake of SRH services across a wide range of services, with the strongest evidence for contraceptive use. Interventions with 2-way interactive functions and more behavior change techniques embedded were more effective. Limited data suggested that interventions were inexpensive, but cost-effectiveness was not evaluated.
Patel et al [32], 2022	Adolescents in LMICs (countries not specified by the authors)	Environmental scan is used to identify current SRH mobile apps available in the iOS App Store and Android Play Store.	Mobile apps	2165 mobile apps were initially identified, with only 8 apps meeting the inclusion criteria and assessed using the Mobile App Rating Scale (MARS) tool. The functionality subdomain scored the highest at 4.6. The information subdomain scored the lowest at 2.5. None of the assessed apps contained information on the MARS items related to the evidence base and goals. "Too Shy to Ask" had the highest individual app mean score of 4.1, while "e-SRHR" scored the lowest at 2.3.

XSL•FO
RenderX

Study (authors and year)	Study population/region/country	Study design	Digital tool type	Findings
Pfeiffer et al [33], 2014	60 adolescents aged 15-19 years in Dar es Salaam and Mtwara, Southern Tanzania	Mixed methods study	Social media, with a particular focus on Facebook	Adolescents access the internet mainly through mobile phones (68%). Facebook is the most popular internet site among the youth (73%). Adolescents expressed interest in receiving reproductive and sexual health messages through social media, with 92% of respondents saying Facebook should be used. Youth role models, such as music stars and actors, are influential in delivering health messages.
Rogers et al [34], 2019	Adolescents aged 10-19 years living in Zambia	Content analysis of the TuneMe website	Mobile-optimized website (TuneMe)	The TuneMe platform provided extensive information on SRH and HIV. Topics on relationships (21%) as well as sexual rights and sexual citizenship (11%) were also covered, albeit to a lesser extent. Areas, such as pleasure, violence, diversity, and gender, received significantly less attention. Content was presented in culturally relevant contexts but sometimes portrayed mixed or problematic views on gender norms. 89% of the TuneMe content falls within the IPPF Framework components for CSE, with 37% falling in the SRH and HIV primary category. The Gunning Fog Index score indicated that a minimum of an 8th grade reading level would be required to fully understand TuneMe content.
Rokicki et al [35], 2017	756 female adolescents aged 14-24 years enrolled in secondary schools in Accra, Ghana	Cluster randomized controlled trial. Unidirectional intervention (n=12 schools): participants received text messages with reproductive health information; interactive intervention (n=12 schools): participants engaged in text messaging reproductive health quizzes; control (n=14 schools): did not receive reproductive health–related text messages.	Text messaging program	The unidirectional intervention increased reproductive health knowledge by 11 percentage points (95% CI 7-15). The interactive intervention increased knowledge by 24 percentage points (95% CI 19-28) from a control baseline of 26%. No significant changes in reproductive health outcomes overall. Both interventions lowered the odds of self-reported pregnancy for sexually active participants (unidirectional: OR 0.14, 95% CI 0.03-0.71; interactive: OR 0.15, 95% CI 0.03-0.86). Results showed that text-messaging programs may help improve reproductive health knowledge among adolescents.
Soehnchen et al [13], 2023	77 young adults aged 18-35 years in the resource-poor regions of Kenya	Survey based on the Unified Theory of Acceptance and Use of Technology (UTAUT)	Web-based prototype designed to provide essential sexual health information	77 responses were included in the analysis. High acceptance and usability of the digital tool among the target population. Perceived usefulness, attitude toward health care–integrated evidence technology, and usability had significant positive impacts on acceptance and intention to use. Having the resources and knowledge necessary for the usage of a digital tool had a significant negative impact. The chi-square test statistic was 532 (P<.001), meaning it was highly significant compared to the original model. A System Usability Scale (SUS) score of 67.3 indicated the tool's usability as "okay."
Soehnchen et al [8], 2023	17 young female individuals aged 15-25 years from Kenya	Qualitative research design, with a sample size of 12 pilot phase interviews and 5 expert interviews.	Web-based sexual health education app	Stigmatization around sexual health in Kenya leads to myths and a lack of information. Sexual health education is not part of the Kenyan school curriculum, leading to insufficient knowledge about safe contraception, menstruation, and female genital mutilation. A digital app could support and provide education and information for universal equal access. Barriers to using a digital sexual health education app include conservative cultural background, classic text communication, and social affiliation influence.

XSL•FO
RenderX

Study (authors and year)	Study population/region/country	Study design	Digital tool type	Findings
Frankline et al [36], 2020	398 female adolescents aged 10-19 years from Cameroon	Single-centered randomized controlled single-blinded trial. Intervention group: received weekly educational 1-way text messages on sexoreproductive health (199 participants); control group: did not receive the text messages (199 participants).	Mobile phone text messaging (SMS)	Significant increase in mean knowledge, attitude, and practice scores from baseline to the end of the study (from 6.03, 4.01, and 3.45 to 7.99, 5.83, and 4.99, respectively). Statistical significance in the overall improvement in adolescents' perception of sexoreproductive health as a result of the intervention (F=15.12; P=.02 for knowledge; F=60.21; P=.001 for attitudes; F=57; P=.01 for practices). Majority (65.3%) of participants were satisfied with the SMS text messaging service.
Akande et al [37], 2024	1280 in-school adolescents aged 15-17 years from Nigeria	Cluster randomized controlled trial. Intervention group: received the mHealth-based intervention; control group: did not receive the mHealth intervention.	mHealth-based sex education program	Significant improvements in SRH knowledge (F=2117.252; P<.001) and attitudes (F=148.493; P<.001) in the intervention group. No significant change in risky sexual behavior scores in the intervention group (P=.57). Females had higher odds of having good SRH knowledge (aOR 2.5, 95% CI 1.04-6.13). Males had lower odds of practicing protective sexual behavior (aOR 0.3, 95% CI 0.15-0.55), and higher class levels (SS2: aOR 5.2; SS3: aOR 6.2) were associated with more protective behavior.
Rokicki and Fink [15], 2017	756 female adolescents and young women aged 14-24 years from Ghana	Randomized controlled trial. Unidirectional intervention (text messages with reproductive health information), interactive intervention (engaged adolescents in text messaging reproductive health quizzes), and control (no intervention).	Interactive mobile phone quiz (mHealth intervention)	The mHealth intervention was effective in engaging adolescent girls across sociodemographic strata. Participants showed increased health knowledge regarding SRH. Knowledge scores increased from 26% to 32% in the control group, 30% to 45% in the unidirectional group, and 31% to 60% in the interactive group when looking from baseline to the 3-month follow-up. The interactive intervention was more effective than the unidirectional intervention, increasing knowledge scores by 13 percentage points (95% CI 8-18). The intervention reached adolescents who may be at higher risk of poor SRH outcomes.
Haruna et al [38], 2018	120 lower secondary school students aged 11-15 years from Tanzania	Randomized controlled trial. Groups: (1) game-based learning (GBL), (2) gamification, (3) control group (traditional teaching).	GBL and gamification platforms	The Adolescent Sexual Health Literacy Test (ASHLT) showed a significant increase in scores from pretest (mean 29.26, SD 8.689) to posttest (mean 70.36, SD 18.201) (t_{119}=–23.787; P=.001). GBL and gamification methods significantly improved students' motivation, attitude, knowledge, and engagement (MAKE) compared to traditional teaching methods, as evidenced by higher scores in these areas and supported by positive feedback from focus group discussions. Pairwise comparisons revealed significant differences between traditional and experimental methods, with no substantial differences between GBL and gamification, indicating that both experimental approaches were more effective in enhancing sexual health education.
Alhassan et al [39], 2019	250 adolescents and young people aged 18-24 years at the University of Ghana, Legon in Accra, Ghana	Cross-sectional analytical study	Mobile phones (including smartphones)	99% of participants owned mobile phones, with 58% being smartphone users. Male young adults and young adults who owned a smartphone were more likely to use mobile phones for STI education and prevention (P=.000 and P=.01, respectively). The study suggests high mobile phone penetration among young adults, with a belief in the efficacy of mobile phone programs for STI education and prevention.
Ippoliti et al [40], 2021	Adolescents aged 12-19 years from Rwanda	HCD approach	CyberRwanda, a web-based digital health platform	Over 1000 Rwandan youth, caregivers, teachers, health care providers, and government stakeholders engaged. Revealed participants' beliefs, behavioral preferences, and experiences related to family planning and reproductive health. Designed for urban and periurban young people and pharmacy staff. Evaluations will be conducted across 60 schools and 9 youth centers in 8 districts in Rwanda (n=6082 at baseline).

XSL•FO
RenderX

Objectives 1a and 1b: Geographic and Demographic Focus

As shown in Figure 2, the geographical scope of research on digital interventions for SRH education spans diverse settings in Sub-Saharan Africa, each offering unique insights into the challenges and opportunities inherent in deploying technology-based solutions within various cultural and infrastructural contexts. The studies span populations in West, Central, East, and South Africa. This geographical diversity is not merely a backdrop for these studies but a critical factor that shapes the design, implementation, and impact of digital health interventions across the continent; it underscores the importance of contextually adapted solutions. As depicted in Table 1, the studies included youth and adolescents of various ages, integrating variations of human-centered designs (HCDs) that resonated with their lived experiences and preferences. In Nigeria, Dulli et al [23], Fakoya et al [24], and Akande et al [37] highlighted the country's engagement with digital platforms to address SRH among youth and adolescents. Feroz et al [16] and Ippoliti and L'Engle [28] also included studies from Nigeria. Nigeria, with its vast population and significant digital penetration, represents a critical context for understanding how social media and participatory design can be harnessed to meet young people's health education needs. The focus on Nigeria underscores the potential of digital tools to transcend traditional barriers to health education, leveraging the widespread use of social media and mobile phones among the youth. Besides Nigeria, this study included some other countries in West Africa, such as Ghana [16,31,35,39], Senegal, and Mali [28].

Figure 2. Map of Africa showing the countries included in the reviewed studies (green).

The reviewed studies extend beyond West Africa to include Kenya, as illustrated by Gonsalves et al [25], who explored the impact of mobile phone–based sexual health information among young people, as well as Soehnchen et al [13], who assessed acceptance of a digital tool for delivering sexual health education. Some other authors, namely Feroz et al [16], Ippoliti and L'Engle [28], Onukwugha et al [31], and Soehnchen et al [8], also included studies that were performed in Kenya in their reviews. Kenya's inclusion in this body of research is indicative of its role as a hub for technological innovation and digital health initiatives in East Africa. Nalwanga et al [27] performed a cross-sectional analysis to assess the use of a mobile phone app by university students in Uganda, another East African country. Haruna et al [26] focused on secondary school

XSL•FO
RenderX

participants in Tanzania, using a gamified learning platform to engage students in SRH education. This emphasis on educational settings highlights the potential of integrating digital tools into formal education systems, showing how gamification can enhance learning outcomes and engagement among school-aged youth. Data from Tanzania, Rwanda, and Ethiopia, which are also East African countries, were included in studies by Fakoya et al [24], Feroz et al [16], Haruna et al [26,38], Ippoliti and L'Engle [28], Ippoliti et al [40], Onukwugha et al [31], Nolan et al [30], and Pfeiffer et al [33]. The cross-regional approach of Gonsalves et al [25], encompassing participants from both Kenya and Peru, further emphasizes the global relevance of digital health interventions and the importance of considering diverse cultural and infrastructural landscapes in their design and deployment.

This review also included some countries in southern Africa, including South Africa, Mozambique, and Zambia. For example, Rogers et al [34] performed a content analysis of TuneMe, a mobile-optimized website that offers extensive information on SRH and HIV. Importantly, the content was presented in culturally relevant contexts. In addition to that study, Feroz et al [16], Ippoliti and L'Engle [28], and Onukwugha et al [31] included data from southern African countries in their systematic reviews.

Democratic Republic of Congo [16] and Cameroon [36] were the 2 Central African countries included in this review. Collectively, these studies paint a picture of a continent actively engaging with digital technologies to overcome the challenges in delivering SRH education. The wide geographical reach of this research, ranging from the populous country of Nigeria to the tech-savvy environments of Kenya, reflects the varied landscapes in which these digital interventions are implemented. Each setting provides unique insights into the interplay between cultural norms, technological accessibility, and health education needs, contributing to a richer understanding of how digital tools can be adapted and scaled across the Sub-Saharan African context.

Target Demographics

Among the studies reviewed, there was a notable emphasis on youth and adolescents, spanning ages 10 to 24 years, which reflected the critical importance of reaching these age groups with effective, accessible, and engaging health education. Although the focus of this study is the adolescent age group, they are often included in broader age groups, such as young people and youth, as shown in Table 1. For instance, Dulli et al [23] concentrated on youth aged 15 to 24 years in Nigeria. The authors performed a randomized controlled trial using social media as a platform to deliver sexual health interventions (SMART Connections). Similar populations were targeted by Gonsalves et al [25] in Kenya and Rokicki et al [35] in Ghana. Similarly, Fakoya et al [24] engaged adolescents aged 15 to 19 years from Nigeria, Tanzania, and Ethiopia, incorporating youth-led participatory action research with an HCD to create interventions that resonate with the lived experiences and preferences of young individuals. The authors directly engaged the target demographic in the study design process. In Rwanda, Nolan et al [30] analyzed data from 6000 youth aged 12 to 19 years in a cluster-randomized trial and Rogers et al [34] studied a similar age group (10 to 19 years) in Zambia.

Further broadening the demographic spectrum, Feroz et al [16] extended their review to young people aged 10 to 24 years in Sub-Saharan Africa and Asia. The Sub-Saharan African countries in that study included Tanzania, Nigeria, Ghana, Uganda, and Kenya. The authors performed a comprehensive review of mHealth interventions, identified a wide range of options, and elucidated their applications as well as barriers and facilitators for adoption.

The inclusive age range spanning from early adolescence to young adulthood acknowledges the diverse needs and challenges faced by individuals as they navigate through different stages of the early life course, emphasizing the importance of tailoring interventions to meet these varying requirements. The more expansive approach of Gonsalves et al [25], by including young people aged 13 to 24 years in Kenya and Peru, illustrates the global relevance of digital sexual health interventions and the need to consider cross-cultural applicability and customization in their design and implementation. This cross-regional study underscores the universal challenges related to sexual health education, while acknowledging the specific nuances that different cultural contexts bring to the fore. Finally, Haruna et al [26] involved secondary school participants from Dar es Salaam, Tanzania in the development of a digital gamified learning platform. This engagement not only emphasizes the educational context as a critical setting for sexual health education but also showcases the potential of gamification as a strategy to enhance learning outcomes and retention among school-aged youth.

Collectively, the studies reflect a strategic and inclusive approach, which is aimed at bridging gaps in access and engagement. By examining the implementation and impact of these digital tools across various African settings, researchers and practitioners can gain valuable insights into the factors that drive the success and scalability of health education interventions, thus paving the way for more inclusive and effective public health strategies.

Objective 2: Digital Tool Types and Effectiveness

The landscape of digital tools deployed for SRH education in various African contexts not only signifies a shift toward more accessible and engaging formats for health education but also reflects the adaptability of health practitioners and researchers to leverage technology to meet the needs of target populations.

With regard to design, Fakoya et al [24] introduced a youth-engaged version of HCD, a problem-solving participatory process that centers the needs, perspectives, and experiences of target populations, when developing solutions to complex challenges. The authors used HCD as part of Adolescents 360, a transdisciplinary initiative to increase the use of modern contraception among 15- to 19-year-old girls in Nigeria, Ethiopia, and Tanzania. The youth-led participatory approach involved 111 "youth designers" trained in HCD methods to inform the design and implementation of the interventions. The authors had successfully engaged youth as project partners and action researchers involved in an iterative cycle of program

XSL•FO
RenderX

research, design, and evaluation. However, they noted challenges in recruiting "extreme users" due to the high competencies needed in HCD and emphasized the need for real-world testing (not just prototyping to enhance adolescents' experiences) and the importance of planned and dedicated financial and human resources. An HCD approach was also deployed in Rwanda by Ippoliti et al [40] for their web-based digital health platform CyberRwanda. The involvement of not only youth but also caregivers, teachers, health care providers, and government stakeholders revealed participants' beliefs, behavioral preferences, and experiences related to family planning and reproductive health. The authors noted that the HCD process, although iterative, resulted in significant pivots to the design and implementation of the digital platform and resulted in a superior digital health intervention with and for Rwandan youth.

For deployment and implementation, the ubiquity and accessibility of mobile phones make them an ideal platform for reaching a wide audience, particularly in regions where mobile technology may be more readily available than traditional health care services. In fact, Alhassan et al [39] report that 99% of 250 surveyed adolescents and young people aged 18 to 24 years owned mobile phones and 58% were smartphone users. Young male adults and those who owned a smartphone were more likely to use mobile phones for STI education and prevention. In their systematic review of the use of mobile phone technology in LMICs for mHealth SRH interventions, Feroz et al [16] noted that most of the included mHealth intervention studies (n=14, 93%) focused on behavioral change and patient education using a variety of modalities, including SMS text messaging, video clips, images, and voice communication. mHealth tools improved young people's SRH knowledge, ensured safer sexual behavior, maximized reach and access to family planning information, and improved several HIV outcomes. Barriers to mHealth uptake for SRH included high cost of service, infrastructural or network quality, request for sociodemographic information that could break anonymity, and sociocultural beliefs and norms. Ippoliti and L'Engle [28] also echoed these findings in their global landscape analysis of the use of mobile phones for SRH content among adolescents and youth aged 10 to 24 years in Africa, Eurasia, and Latin America. The authors reported that mobile phones (particularly text messaging) were effectively used to increase the reach of SRH information and services, especially in conservative societies where SRH topics are stigmatized. Nalwanga et al [27] assessed the use of a mobile phone app to increase access to SRH information, goods, and services among university students in Uganda and reported high use of the frequently asked questions portal (71%) and high product use by both sexes (condoms for males [77% use]; sanitary pads for females [94% use]), and the most popularly accessed service was HIV testing and counseling (60% use). Participants appreciated the responsiveness, nondistracting in-app advertisements, and "ease of use" attributes of the mobile phone app. One interesting aspect of this study was the inclusion of a modest co-payment (paid by users), which, interestingly, was not a barrier to usage. Contrary to what has been reported by others, the youth in this study did not find the co-pay to be cost prohibitive. In fact, its incorporation ensured the utility of client payments in future iterations of the app after the pilot period. In a cluster randomized controlled trial of 756 adolescent female individuals aged 14 to 24 years enrolled in secondary schools in Accra, Ghana, Rokicki et al [35] compared the following 3 arms: unidirectional intervention (n=12 schools), where participants received text messages with reproductive health information; interactive intervention (n=12 schools), where participants engaged in text messaging reproductive health quizzes; and control (n=14 schools), where students did not receive reproductive health messaging. They found that the unidirectional intervention increased reproductive health knowledge by 11 percentage points (95% CI 7-15); however, the inclusion of quizzes in the interactive intervention increased knowledge by 24 percentage points (95% CI 19-28) from a control baseline of 26%. Both interventions lowered the odds of self-reported pregnancy for sexually active participants (unidirectional: OR 0.14, 95% CI 0.03-0.71; interactive: OR 0.15, 95% CI 0.03-0.86). These studies highlight the versatility of mobile phones in delivering tailored health education and interventions ranging from SMS text messaging–based information dissemination to more interactive app-based learning experiences and quizzes to reinforce knowledge retention and behavioral change. They can act as a one-stop shop, offering not only educational and clinical counseling services but also condoms, contraceptives, and menstrual products, which give users a holistic experience that can be anonymized to ensure privacy and protection.

The digital gamified learning platform developed in Tanzania by Haruna et al [26,38] represents an innovative leap in engaging and educating secondary school participants in the African setting. The authors used game-based and participatory designs to evaluate the ability of the game to motivate students, improve their attitudes, increase their acquisition of knowledge, and engage them in learning. A paired *t*-test revealed a statistically significant improvement in sexual health literacy scores from pretest (mean 26.40, SD 7.29) to posttest (mean 74.12, SD 16.21) (t_{347}=52.230; P<.000), with an average increase of 47.72 points. Additional analyses with ANOVA testing showed no significant differences in pretest scores among the 3 groups, but significant differences were found in posttest scores ($F_{2,345}$=210.43; P<.001), with follow-up Tukey post-hoc tests indicating better performance in the gamified learning platform groups than in the traditional teaching group. These findings suggest that gamified learning platforms can enhance motivation, attitude, knowledge acquisition, and engagement in sexual health education more effectively than traditional methods [26].

The research team took the intervention further in a 3-armed randomized clinical trial to compare the effectiveness of game-based learning and gamification to a control approach (traditional teaching method) [38]. The results showed that the average posttest scores were significantly higher for game-based learning (mean 79.94, SD 11.169) and gamification (mean 79.23, SD 9.186) than for the control approach (mean 51.93, SD 18.705) ($F_{2,117}$=54.75; P=.001). Additionally, statistically significant differences (P≤.05) were found for the constructs of motivation, attitude, knowledge, and engagement (MAKE) for the 2 intervention groups compared to the control arm. Gamification not only enhances learning outcomes but also introduces an element of fun and interaction that can

XSL•FO
RenderX

significantly increase participant engagement and knowledge retention [38]. Such gamified platforms offer promising avenues for making health education more appealing and effective for younger people.

Web-based platforms may eliminate the costs associated with the implementation of gamified learning platforms. In a content analysis of the TuneMe website, designed to deliver SRH information to Zambian adolescents aged 10 to 19 years, Rogers et al [34] reported that the website comprehensively covered topics of reproductive anatomy and STIs, including HIV, as well as relationships and sexual rights. However, the challenge of presenting SRH education in culturally relevant contexts was apparent when topics of sexual pleasure and gendered norms, behaviors, expectations, and violence were introduced, demonstrating nuanced beliefs around sexual rights.

Finally, social media emerged as a significant platform in the digital toolbox, with Dulli et al [23] exploring its utility among youth aged 15 to 24 years in Nigeria. The authors performed a randomized controlled trial using social media to deliver sexual health interventions for youth living with HIV (SMART Connections). The design and content of SMART Connections were as informed through workshops conducted with stakeholders and youth living with HIV. The content was delivered anonymously through Facebook over approximately 22 weeks to the intervention and control groups, with daily activities to promote engagement. The primary outcome was retention in HIV treatment, while secondary outcomes included antiretroviral therapy (ART) adherence, HIV knowledge, and social support. Although the intervention did not significantly improve the primary outcome of retention (50% and 45% retention in the treatment and control groups, respectively), there was a significant improvement in HIV-related knowledge for the SMART Connections group compared to controls (t=–2.96; P=.003). Intervention group participants overwhelmingly reported that the intervention was useful, that they enjoyed taking part, and that they would recommend it to peers. Similarly, Pfeiffer et al [33] found that Facebook was the most used social media platform for sexual health promotion in Southern Tanzania for youth aged 15 to 19 years. Adolescents were most interested in SRH messages through humorous posts, links, and clips, as well as through youth role models like music stars and actors in a manner that reflected up-to-date trends of modern youth culture. Both studies highlighted the integral role of social media in the daily lives of youth, positioning it as a critical channel for delivering sexual health education.

Collectively, these digital tools, ranging from social media and participatory design to mobile phone apps and gamification, reflect a dynamic and multifaceted digital ecosystem for SRH education. Each tool, with its unique features and applications, contributes to a broader strategy aimed at enhancing the accessibility, engagement, and effectiveness of health interventions in different African settings. These diversified digital tools not only address the varied needs and preferences of target demographics but also showcase the potential of technology in transforming health education and empowerment.

Discussion

In this research endeavor, we conducted a systematic literature review to achieve the following two objectives: (1) describe the geographic and demographic distributions of digital educational tools for SRH education for adolescents in Sub-Saharan Africa (objectives 1a and 1b) and (2) examine the types, formats, and impacts of these digital tools on target populations (objective 2). For the first objective, we identified 22 studies conducted across West, Central, East, and South Africa that focused on digital educational tools for SRH education targeting youth and adolescents aged 10 to 24 years. These studies offer valuable insights into the types of digital tools currently being employed in SRH education and highlight the regions where such tools are being most actively implemented. For the second objective, we found that a variety of digital tools, including social media, mobile apps, and gamified learning, were used to engage adolescents with accessible SRH information. These methods were generally effective, although challenges, such as the digital divide, cultural sensitivity, and the need for assessing long-term behavior change, were noted.

The 22 studies included in this review covered a broad spectrum of digital interventions, ranging from mobile apps to social media campaigns, and reflected a growing interest in digital platforms as a means of delivering SRH education to young people in Sub-Saharan Africa. The geographical distribution of the studies indicates that digital tools are being used across various cultural contexts, with notable contributions from West (eg, Nigeria and Mali), East (eg, Kenya and Tanzania), Central (eg, Cameroon and Democratic Republic of Congo), and Southern Africa (eg, Mozambique and South Africa). This regional diversity is significant as it shows that digital SRH interventions are becoming increasingly accessible across different subregions of Africa, each with unique sociocultural challenges and opportunities. Our findings suggest that digital tools can overcome geographic barriers, offering a means to engage with youth who may otherwise be excluded from formal SRH education programs due to geographical isolation, economic constraints, or cultural taboos around discussing SRH topics [41,42]. These findings are consistent with the broader literature on digital health interventions in the African region, which has highlighted the promising role of digital tools in expanding access to SRH education [31,43]. For example, in Nigeria, mobile apps focused on HIV prevention and sexual health have been shown to improve knowledge and self-efficacy among adolescents [44]. In Kenya, WhatsApp-based health campaigns have successfully increased awareness of sexual health issues among young people [45]. Conducting cross-country comparisons through systematic reviews, as performed in this work, provides valuable insights into the best practices that can be adapted and scaled across diverse African contexts. For example, WhatsApp is widely used around the world, and rather than reinventing the wheel, could developers in Nigeria adapt and implement the WhatsApp-based intervention designed for Kenyan youth by Chory et al [45] for a similar population in Nigeria? What would be the short- and long-term impacts of such an approach? Moreover, would policymakers be more inclined to support a “repackaged”

XSL•FO
RenderX

intervention that is cost-effective, evidence-based, and already proven in another African context? Future research should explore these cross-country comparisons in greater depth, examining the potential for scaling successful digital interventions across African nations. By redesigning these tools to reflect the cultural, linguistic, and societal contexts of each country, researchers can assess the feasibility and cost-effectiveness of adapting existing interventions, potentially saving time and resources compared to developing entirely new solutions.

Our second objective revealed that while a variety of digital tools were used to engage adolescents with accessible SRH information, mobile apps and gamified learning emerged as particularly popular and effective approaches. Mobile apps allow adolescents to access information in a flexible, anonymous, and user-friendly format, which helps overcome the barriers posed by traditional methods of SRH education [46]. This flexibility is particularly important in contexts where discussing sexual health is stigmatized, as it empowers youth to seek information discreetly and at their own pace. Additionally, gamified learning, by integrating interactive elements, not only captures adolescents' attention but also reinforces learning through engaging experiential methods, which may result in better knowledge retention and a deeper understanding of SRH topics [31]. Gamified learning, in particular, aligns with the preferences of a digitally native generation, where interactive and visually stimulating content is highly favored. Social media, including platforms like Facebook, WhatsApp, and Instagram, were also used to reach adolescents, leveraging peer influence and interactive communication strategies. This social aspect of learning enhances engagement by making the information feel more relevant and personal for youth. However, a major issue noted across the studies was the "digital divide," which is used to describe the unequal access to technology and internet connectivity that can limit the effectiveness of digital SRH education programs [47]. While mobile phones are widespread in Sub-Saharan Africa, reliable internet access remains a significant barrier, particularly in rural and underdeveloped regions [45]. This digital divide may result in inequitable access to SRH resources, especially for marginalized youth [45]. Innovative solutions, such as an artificial intelligence–powered phone, have the potential to bridge the digital divide [48]. This system delivers information and advice through voice calls, eliminating the need for a smartphone or internet connection and making information accessible to individuals without reliable internet or those who cannot afford smartphones and data plans [48]. By using artificial intelligence to understand natural language queries and provide tailored information from a knowledge base, these technologies can significantly improve access to SRH education in underserved areas with limited technological infrastructure [48]. Future research should explore effective solutions to overcome the digital divide, such as offline digital tools as described, low-cost data plans, and internet infrastructure development in rural and underserved areas.

Cultural sensitivity also emerged as a critical issue. In many African contexts, cultural norms and values significantly shape how SRH topics are perceived and discussed, and failure to account for these factors can limit the effectiveness of digital SRH interventions. For instance, discussions around issues like sexual consent, gender equality, and contraception can be seen as a taboo in some regions, which may lead to resistance or disengagement from the target audience [41]. Thus, tailoring content to reflect local values, languages, and cultural norms around sexuality is critical, as failing to consider these factors could undermine the effectiveness of SRH interventions. Involving adolescents in the creation and assessment of these tools can provide essential insights into their preferences, needs, and challenges, which can significantly enhance the acceptability and effectiveness of digital interventions [49]. Moreover, including parents could create opportunities to address cultural taboos and foster more open discussions between parents and adolescents on SRH [42]. Such involvement could lead to long-term benefits, as parent-adolescent communication is a key factor in reducing adverse sexual behaviors and improving overall sexual health outcomes [42]. According to Agbeve et al [42], African parents who hold sex-positive attitudes are more inclined to engage in discussions with their adolescents that help them make informed decisions regarding their sexual health and overall well-being. When parents are actively involved in the design and implementation of digital interventions, they can provide support, guidance, and reinforcement of the messages delivered through digital platforms [50,51]. Parental involvement can help create a more comprehensive approach to sexual education, ensuring that teens not only receive the digital information but also have a trusted adult to turn to for advice and clarification [50,51]. By fostering open communication between parents and teens, digital interventions can be better tailored to the specific needs of the family, ultimately promoting healthier decision-making and outcomes for young people.

Future studies could also explore avenues for incorporating comprehensive SRH education into the formal education system, expanding the scope of topics to include healthy relationships, positive sexuality, mental health, gender equality, and consent, as was done by Fakoya et al [24]. Integrating these topics into the curriculum would not only equip adolescents with essential knowledge about SRH but also foster a more holistic understanding of well-being and personal development. By addressing issues like mental health and gender equality alongside traditional SRH topics, this approach could help challenge harmful stereotypes, promote healthier attitudes toward sexuality, and empower young people to make informed decisions, leading to positive long-term social and health outcomes.

Finally, there was a noted gap regarding the impact of these digital tools on long-term adolescent sexual behavioral change. The majority of studies measured short-term outcomes, such as knowledge acquisition, while long-term studies assessing outcomes, such as improved contraceptive use and reductions in risky sexual behavior, remained limited. The anonymity and privacy afforded by digital platforms can empower adolescents to seek information and services without fear of stigma, promoting more proactive and informed health behaviors that can be tracked over time [46]. Tracking tools can be incorporated in the design phase of the intervention development

XSL•FO
RenderX

and can have input from adolescent end users. Long-term data support the sustainability of digital interventions, providing valuable information that can aid in the buy-in of investors, software developers, public health enthusiasts, and policymakers. Regardless, the implications of our findings are multifaceted and show that digital SRH tools are a promising avenue for enhancing adolescent sexual health education across the African region.

The use of a systematic literature review approach in this work has some limitations that are worth mentioning. The main limitations of our work are related to the methodology. There is a risk of selection bias, meaning that some relevant studies may have been omitted or that the included studies may not fully represent all available research on the topic. However, we believe that our extensive use of multiple databases (Ovid-MEDLINE, Google Scholar, PubMed, and ERIC) mitigated the risk of selection bias inherent in systematic reviews. A meta-analysis could further inform the selection bias risk and the heterogeneity of the included studies, and quantify the associations reported.

In conclusion, although digital SRH tools are a promising avenue for enhancing adolescent sexual health education in Sub-Saharan Africa, there is a need for infrastructure improvements to ensure the tools are equitably accessible and a need for understanding long-term impacts and sustainability. Policymakers, educators, and public health researchers must collaborate to ensure that digital innovations in SRH education are grounded in evidence, culturally sensitive, and widely accessible to all adolescents, regardless of socioeconomic status or geographical location. This will address the sustainable development goals of reducing the global disparities in SRH health (goal 10) and achieving good health and well-being across nations (goal 3).

Acknowledgments

MAO was funded by National Institutes of Health (NIH) National Institute of Neurological Disorders and Stroke (NINDS) K01 Award# K01NS121199.

Data Availability

All processes are outlined in the Methods section. Additional information beyond what is reported in the manuscript (eg, template data collection forms, data extracted from the included studies, data used for all analyses, analytic codes, etc) is available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Line-by-line search to retrieve titles and abstracts.
[DOCX File , 17 KB - publichealth_v11i1e63309_app1.docx]

Multimedia Appendix 2
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 173 KB - publichealth_v11i1e63309_app2.pdf]

References

1. Melcher J, Lavoie J, Hays R, D'Mello R, Rauseo-Ricupero N, Camacho E, et al. Digital phenotyping of student mental health during COVID-19: an observational study of 100 college students. J Am Coll Health 2023 Apr;71(3):736-748. [doi: 10.1080/07448481.2021.1905650] [Medline: 33769927]
2. Wadham E, Green C, Debattista J, Somerset S, Sav A. New digital media interventions for sexual health promotion among young people: a systematic review. Sex Health 2019 Apr;16(2):101-123. [doi: 10.1071/SH18127] [Medline: 30819326]
3. Chandra-Mouli V, Neal S, Moller A. Adolescent sexual and reproductive health for all in sub-Saharan Africa: a spotlight on inequalities. Reprod Health 2021 Jun 17;18(Suppl 1):118 [FREE Full text] [doi: 10.1186/s12978-021-01145-4] [Medline: 34134737]
4. Armstrong A, Nagata JM, Vicari M, Irvine C, Cluver L, Sohn AH, et al. A Global Research Agenda for Adolescents Living With HIV. J Acquir Immune Defic Syndr 2018 Aug 15;78 Suppl 1(1):S16-S21 [FREE Full text] [doi: 10.1097/QAI.0000000000001744] [Medline: 29994915]
5. Kassa GM, Arowojolu AO, Odukogbe AA, Yalew AW. Prevalence and determinants of adolescent pregnancy in Africa: a systematic review and Meta-analysis. Reprod Health 2018 Nov 29;15(1):195 [FREE Full text] [doi: 10.1186/s12978-018-0640-2] [Medline: 30497509]

XSL•FO
RenderX

6. Melesse DY, Mutua MK, Choudhury A, Wado YD, Faye CM, Neal S, et al. Adolescent sexual and reproductive health in sub-Saharan Africa: who is left behind? BMJ Glob Health 2020;5(1):e002231 [FREE Full text] [doi: 10.1136/bmjgh-2019-002231] [Medline: 32133182]
7. Wangamati CK. Comprehensive sexuality education in sub-Saharan Africa: adaptation and implementation challenges in universal access for children and adolescents. Sex Reprod Health Matters 2020 Dec;28(2):1851346 [FREE Full text] [doi: 10.1080/26410397.2020.1851346] [Medline: 33295853]
8. Soehnchen C, Rietz A, Weirauch V, Meister S, Henningsen M. Creating an Intercultural User-Centric Design for a Digital Sexual Health Education App for Young Women in Resource-Poor Regions of Kenya: Qualitative Self-Extended Double Diamond Model for Requirements Engineering Analysis. JMIR Form Res 2023 Nov 03;7:e50304 [FREE Full text] [doi: 10.2196/50304] [Medline: 37921860]
9. Achen D, Fernandes D, Kemigisha E, Rukundo GZ, Nyakato VN, Coene G. Trends and Challenges in Comprehensive Sex Education (CSE) Research in Sub-Saharan Africa: a Narrative Review. Curr Sex Health Rep 2023 May 06:1-9 [FREE Full text] [doi: 10.1007/s11930-023-00362-1] [Medline: 37362203]
10. Otu A, Ukpeh I, Okuzu O, Yaya S. Leveraging mobile health applications to improve sexual and reproductive health services in Nigeria: implications for practice and policy. Reprod Health 2021 Jan 23;18(1):21 [FREE Full text] [doi: 10.1186/s12978-021-01069-z] [Medline: 33485356]
11. Hubert C, Estrada F, Campero L, Heredia-Pi IB, Villalobos A, Suárez-López L, et al. Designing Digital Tools Capable of Reaching Disadvantaged Adolescents and Improving their Sexual Health: A Mexican Experience. Journal of Health Care for the Poor and Underserved 2021;32(2S):62-84. [doi: 10.1353/hpu.2021.0051]
12. Simuyaba M, Hensen B, Phiri M, Mwansa C, Mwenge L, Kabumbu M, et al. Engaging young people in the design of a sexual reproductive health intervention: Lessons learnt from the Yathu Yathu ("For us, by us") formative study in Zambia. BMC Health Serv Res 2021 Jul 29;21(1):753 [FREE Full text] [doi: 10.1186/s12913-021-06696-7] [Medline: 34325696]
13. Soehnchen C, Weirauch V, Schmook R, Henningsen M, Meister S. An acceptance analysis of a sexual health education digital tool in resource-poor regions of Kenya: an UTAUT based survey study. BMC Womens Health 2023 Dec 19;23(1):676 [FREE Full text] [doi: 10.1186/s12905-023-02839-6] [Medline: 38114976]
14. Kaitano B. Technology and sexual health education in sub-Saharan Africa. Council on Business & Society Insights. 2023 Feb 02. URL: https://cobsinsights.org/2023/02/02/technology-and-sexual-health-education-in-sub-saharan-africa/ [accessed 2024-03-18]
15. Rokicki S, Fink G. Assessing the reach and effectiveness of mHealth: evidence from a reproductive health program for adolescent girls in Ghana. BMC Public Health 2017 Dec 20;17(1):969 [FREE Full text] [doi: 10.1186/s12889-017-4939-7] [Medline: 29262823]
16. Feroz AS, Ali NA, Khoja A, Asad A, Saleem S. Using mobile phones to improve young people sexual and reproductive health in low and middle-income countries: a systematic review to identify barriers, facilitators, and range of mHealth solutions. Reprod Health 2021 Jan 16;18(1):9 [FREE Full text] [doi: 10.1186/s12978-020-01059-7] [Medline: 33453723]
17. Sukhera J. Narrative reviews: Flexible, rigorous, and practical. J Grad Med Educ 2022 Aug;14(4):414-417 [FREE Full text] [doi: 10.4300/JGME-D-22-00480.1] [Medline: 35991099]
18. Van Ryn RM, Nyrongo L. STI education for adolescents. SPARK (Scholarly Publications and Repository of Knowledge). 2023. URL: https://spark.siue.edu/dnpprojects/263 [accessed 2024-03-18]
19. Youth-centred digital health interventions: a framework for planning, developing and implementing solutions with and for young people. World Health Organization. 2020 Oct 26. URL: https://iris.who.int/handle/10665/336223 [accessed 2024-03-18]
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol 2009 Oct;62(10):1006-1012 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.06.005] [Medline: 19631508]
21. Salameh J, Bossuyt PM, McGrath TA, Thombs BD, Hyde CJ, Macaskill P, et al. Preferred reporting items for systematic review and meta-analysis of diagnostic test accuracy studies (PRISMA-DTA): explanation, elaboration, and checklist. BMJ 2020 Aug 14;370:m2632 [FREE Full text] [doi: 10.1136/bmj.m2632] [Medline: 32816740]
22. Critical appraisal tools. JBI. URL: https://jbi.global/critical-appraisal-tools [accessed 2024-06-04]
23. Dulli L, Ridgeway K, Packer C, Murray KR, Mumuni T, Plourde KF, et al. A Social Media-Based Support Group for Youth Living With HIV in Nigeria (SMART Connections): Randomized Controlled Trial. J Med Internet Res 2020 Jun 02;22(6):e18343 [FREE Full text] [doi: 10.2196/18343] [Medline: 32484444]
24. Fakoya I, Cole C, Larkin C, Punton M, Brown E, Ballonoff Suleiman A. Enhancing Human-Centered Design With Youth-Led Participatory Action Research Approaches for Adolescent Sexual and Reproductive Health Programming. Health Promot Pract 2022 Jan;23(1):25-31. [doi: 10.1177/15248399211003544] [Medline: 33858267]
25. Gonsalves L, Hindin MJ, Bayer A, Carcamo CP, Gichangi P, Habib N, et al. Protocol of an open, three-arm, individually randomized trial assessing the effect of delivering sexual and reproductive health information to young people (aged 13-24) in Kenya and Peru via mobile phones: adolescent/youth reproductive mobile access and delivery initiative for love and life outcomes (ARMADILLO) study stage 2. Reprod Health 2018 Jul 11;15(1):126 [FREE Full text] [doi: 10.1186/s12978-018-0568-6] [Medline: 29996854]

XSL•FO
RenderX

26. Haruna H, Zainuddin Z, Mellecker RR, Chu SK, Hu X. An iterative process for developing digital gamified sexual health education for adolescent students in low-tech settings. ILS 2019 Nov 11;120(11/12):723-742. [doi: 10.1108/ils-07-2019-0066]
27. Nalwanga R, Nuwamanya E, Nuwasiima A, Babigumira JU, Asiimwe FT, Babigumira JB. Utilization of a mobile phone application to increase access to sexual and reproductive health information, goods, and services among university students in Uganda. Reprod Health 2021 May 17;18(1):95 [FREE Full text] [doi: 10.1186/s12978-020-01037-z] [Medline: 34001178]
28. Ippoliti NB, L'Engle K. Meet us on the phone: mobile phone programs for adolescent sexual and reproductive health in low-to-middle income countries. Reprod Health 2017 Jan 17;14(1):11 [FREE Full text] [doi: 10.1186/s12978-016-0276-z] [Medline: 28095855]
29. L'Engle KL, Mangone ER, Parcesepe AM, Agarwal S, Ippoliti NB. Mobile Phone Interventions for Adolescent Sexual and Reproductive Health: A Systematic Review. Pediatrics 2016 Sep 01;138(3):49. [doi: 10.1542/peds.2016-0884] [Medline: 27553221]
30. Nolan C, Packel L, Hope R, Levine J, Baringer L, Gatare E, et al. Design and impact evaluation of a digital reproductive health program in Rwanda using a cluster randomized design: study protocol. BMC Public Health 2020 Nov 13;20(1):1701 [FREE Full text] [doi: 10.1186/s12889-020-09746-7] [Medline: 33187485]
31. Onukwugha FI, Smith L, Kaseje D, Wafula C, Kaseje M, Orton B, et al. The effectiveness and characteristics of mHealth interventions to increase adolescent's use of Sexual and Reproductive Health services in Sub-Saharan Africa: A systematic review. PLoS One 2022;17(1):e0261973 [FREE Full text] [doi: 10.1371/journal.pone.0261973] [Medline: 35061757]
32. Patel A, Louie-Poon S, Kauser S, Lassi Z, Meherali S. Environmental scan of mobile apps for promoting sexual and reproductive health of adolescents in low- and middle-income countries. Front Public Health 2022;10:993795 [FREE Full text] [doi: 10.3389/fpubh.2022.993795] [Medline: 36504952]
33. Pfeiffer C, Kleeb M, Mbelwa A, Ahorlu C. The use of social media among adolescents in Dar es Salaam and Mtwara, Tanzania. Reprod Health Matters 2014 May;22(43):178-186 [FREE Full text] [doi: 10.1016/S0968-8080(14)43756-X] [Medline: 24908469]
34. Rogers E, Hemal K, Tembo Z, Mukanu M, Mbizvo M, Bellows B. Comprehensive Sexuality Education for Adolescents in Zambia via the Mobile-Optimized Website TuneMe: A Content Analysis. American Journal of Sexuality Education 2019 Jul 08;15(1):82-98. [doi: 10.1080/15546128.2019.1635546]
35. Rokicki S, Cohen J, Salomon JA, Fink G. Impact of a text-messaging program on adolescent reproductive health: a cluster-randomized trial in Ghana. Am J Public Health 2017 Feb;107(2):298-305. [doi: 10.2105/AJPH.2016.303562] [Medline: 27997236]
36. Frankline S, Dickson S, Omer T, Joseph B. A protocol for a randomized controlled trial on mobile phone text messaging to improve sexo-reproductive health in Cameroon. J Public Health Epidemiol 2020 Jul 31;12(3):179-185 [FREE Full text] [doi: 10.5897/jphe2019.1141]
37. Akande OW, Muzigaba M, Igumbor EU, Elimian K, Bolarinwa OA, Musa OI, et al. The effectiveness of an m-Health intervention on the sexual and reproductive health of in-school adolescents: a cluster randomized controlled trial in Nigeria. Reprod Health 2024 Jan 13;21(1):6 [FREE Full text] [doi: 10.1186/s12978-023-01735-4] [Medline: 38218840]
38. Haruna H, Hu X, Chu SKW, Mellecker RR, Gabriel G, Ndekao PS. Improving sexual health education programs for adolescent students through game-based learning and gamification. Int J Environ Res Public Health 2018 Sep 17;15(9):2027 [FREE Full text] [doi: 10.3390/ijerph15092027] [Medline: 30227642]
39. Alhassan RK, Abdul-Fatawu A, Adzimah-Yeboah B, Nyaledzigbor W, Agana S, Mwini-Nyaledzigbor PP. Determinants of use of mobile phones for sexually transmitted infections (STIs) education and prevention among adolescents and young adult population in Ghana: implications of public health policy and interventions design. Reprod Health 2019 Aug 09;16(1):120 [FREE Full text] [doi: 10.1186/s12978-019-0763-0] [Medline: 31399123]
40. Ippoliti N, Sekamana M, Baringer L, Hope R. Using human-centered design to develop, launch, and evaluate a national digital health platform to improve reproductive health for Rwandan youth. Glob Health Sci Pract 2021 Nov 29;9(Suppl 2):S244-S260 [FREE Full text] [doi: 10.9745/GHSP-D-21-00220] [Medline: 34845048]
41. Ackah J, Esia-Donkoh K, Amponsah A, Agbemavi W, Bediako V, Tettey G. Use of digital media by adolescents for sexual and reproductive health and rights communication in sub-Saharan Africa: a protocol for systematic review. Syst Rev 2024 May 14;13(1):130-167 [FREE Full text] [doi: 10.1186/s13643-024-02534-z] [Medline: 38745293]
42. Agbeve AS, Fiaveh DY, Anto-Ocrah M. Parent-adolescent sexuality communication in the African context: a scoping review of the literature. Sex Med Rev 2022 Oct;10(4):554-566 [FREE Full text] [doi: 10.1016/j.sxmr.2022.04.001] [Medline: 36210094]
43. Chipako I, Singhal S, Hollingsworth B. Impact of sexual and reproductive health interventions among young people in sub-Saharan Africa: a scoping review. Front Glob Womens Health 2024;5(2):1344135-1344218 [FREE Full text] [doi: 10.3389/fgwh.2024.1344135] [Medline: 38699461]
44. Oladele D, Iwelunmor J, Gbajabiamila T, Obiezu-Umeh C, Okwuzu J, Nwaozuru U, et al. The 4 youth by youth mHealth photo verification app for HIV self-testing in Nigeria: Qualitative analysis of user experiences. JMIR Form Res 2021 Nov 17;5(11):e25824-e25831 [FREE Full text] [doi: 10.2196/25824] [Medline: 34787579]

XSL•FO
RenderX

45. Chory A, Callen G, Nyandiko W, Njoroge T, Ashimosi C, Aluoch J, et al. A pilot study of a mobile intervention to support mental health and adherence among adolescents living with HIV in Western Kenya. AIDS Behav 2022 Jan;26(1):232-242 [FREE Full text] [doi: 10.1007/s10461-021-03376-9] [Medline: 34292429]
46. Haruna H, Okoye K, Zainuddin Z, Hu X, Chu S, Hosseini S. Gamifying sexual education for adolescents in a low-tech setting: Quasi-experimental design study. JMIR Serious Games 2021 Oct 12;9(4):e19614 [FREE Full text] [doi: 10.2196/19614] [Medline: 34636739]
47. Dowling R, Howell E, Dasco M, Schwartzman J. Digital adolescent sexual and reproductive health in low- and middle-income countries: A scoping review. Youth 2025 Feb 06;5(1):15-60. [doi: 10.3390/youth5010015]
48. No smartphone or internet? No problem; AI-backed phone has the answers. Reuters. 2024. URL: https://www.reuters.com/technology/no-smartphone-or-internet-no-problem-ai-backed-phone-has-answers-2024-03-15/ [accessed 2024-04-24]
49. Decker MJ, Harrison S, Price M, Gutmann-Gonzalez A, Yarger J, Tenney R. Educators' perspectives on integrating technology into sexual health education: implementation study. JMIR Hum Factors 2022 Jan 12;9(1):e31381 [FREE Full text] [doi: 10.2196/31381] [Medline: 35019842]
50. Aventin Á, Gough A, McShane T, Gillespie K, O'Hare L, Young H, et al. Engaging parents in digital sexual and reproductive health education: evidence from the JACK trial. Reprod Health 2020 Aug 27;17(1):132-765 [FREE Full text] [doi: 10.1186/s12978-020-00975-y] [Medline: 32854734]
51. Santa Maria D, Markham C, Bluethmann S, Mullen P. Parent-based adolescent sexual health interventions and effect on communication outcomes: a systematic review and meta-analyses. Perspect Sex Reprod Health 2015 Mar;47(1):37-50 [FREE Full text] [doi: 10.1363/47e2415] [Medline: 25639664]

Abbreviations

HCD: human-centered design
LMIC: low- and middle-income country
SRH: sexual and reproductive health
STI: sexually transmitted infection

Edited by A Mavragani; submitted 16.06.24; peer-reviewed by N Levitz, MDM Fernández Álvarez; comments to author 06.09.24; revised version received 09.10.24; accepted 24.01.25; published 26.02.25.

Please cite as:
Abdul Hamid Alhassan RH, Haggerty CL, Fapohunda A, Affan NJ, Anto-Ocrah M
Exploring the Use of Digital Educational Tools for Sexual and Reproductive Health in Sub-Saharan Africa: Systematic Review
JMIR Public Health Surveill 2025;11:e63309
URL: https://publichealth.jmir.org/2025/1/e63309
doi:10.2196/63309
PMID:40009849

©Ramatu Hajia Abdul Hamid Alhassan, Catherine L Haggerty, Abimbola Fapohunda, Nabeeha Jabir Affan, Martina Anto-Ocrah. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Uptake of the Second Dose of the Measles Vaccine and Its Determinants Among Children Aged Less Than 5 Years: Systematic Review and Meta-Analysis

Molalign Aligaz Adisu[1], MSc; Tesfaye Engdaw Habtie[2], MSc; Tegene Atamenta Kitaw[2], MSc; Alemu Birara Zemariam[1], MSc

[1]Department of Pediatrics and Child Health, College of Health Science, Woldia University, Woldia, Ethiopia

[2]Department of Nursing, College of Health Science, Woldia University, Woldia, Ethiopia

Corresponding Author:
Molalign Aligaz Adisu, MSc
Department of Pediatrics and Child Health
College of Health Science
Woldia University
RH4X+GM, A2
Woldia, 00000
Ethiopia
Phone: 251 0942767333
Email: molalignaligaz@gmail.com

Abstract

Background: Ethiopia is faced with poor measles-containing vaccine second dose (MCV2) coverage, leading to recurrent outbreaks.

Objective: This meta-analysis and systematic review aimed at combining evidence of MCV2 uptake in Ethiopia and its determinants to inform interventions for increased vaccination uptake and control of public health challenges.

Methods: This review examined observational quantitative research on measles second dose vaccination among children in Ethiopia using databases such as PubMed Central, Cochrane Library, Web of Science, Scopus, MEDLINE, and Embase. The quality of the included studies was assessed using the Newcastle-Ottawa Scale. Heterogeneity was evaluated using I^2 statistics and the Cochran *Q* test, and the analysis used a random-effects model. Publication bias was assessed through funnel plots, the Egger test, and nonparametric trim-and-fill tests.

Results: The overall MCV2 uptake among children aged <5 years in Ethiopia was only 34.4% (95% CI 18.8%-49.9%). Significant determinants of MCV2 uptake included high level of maternal education (adjusted odds ratio [AOR] 3.31, 95% CI 1.32-5.30), attendance to antenatal care follow-ups (AOR 2.02, 95% CI 1.12-2.92), use of postnatal care services (AOR 3.03, 95% CI 1.77-4.28), reduced waiting times at vaccination sites (AOR 2.56, 95% CI 1.98-3.13), good awareness of measles vaccination (AOR 2.17, 95% CI 1.59-2.74), and positive perceptions of the vaccine (AOR 3.58, 95% CI 1.97-6.30).

Conclusions: This study found that the uptake of MCV2 among children aged <5 years in Ethiopia was 34%, which is far below the global and national goal of 95%. Key factors contributing to low coverage include mothers' educational levels, use of antenatal and postnatal care services, waiting times at vaccination sites, and mothers' awareness of the measles vaccine. Improving community-based education programs, increasing access to antenatal and postnatal care services, reducing waiting times, and raising awareness about immunization all contribute to increasing vaccine uptake.

Trial Registration: PROSPERO CRD42024619031; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024619031

(JMIR Public Health Surveill 2025;11:e77195) doi:10.2196/77195

KEYWORDS

second dose; measles; vaccine; uptake; children; Ethiopia; meta-analysis

XSL•FO
RenderX

Introduction

Background

Measles, a virulent viral disease caused by the measles morbillivirus, continues to exert a significant global health burden despite decades of vaccination efforts. The disease is characterized by fever, maculopapular rash, cough, coryza, and conjunctivitis. Alongside these symptoms, severe complications such as encephalitis, severe pneumonia, and even death may occur. These complications tend to be worse in young children, malnourished individuals, and people with a weakened immune system [1,2]. The highly contagious nature of measles is further highlighted by its basic reproductive number (R0), which is estimated to be between 12 and 18. This elevates the threshold population immunity required to prevent sustained transmission of the virus to >95% immunization [3].

The lack of global progress made toward measles control in the early 21st century is concerning, especially with the recent regressions observed. One study revealed the increasing rate of measles infections worldwide, projecting that 10.3 million people were infected in 2023, which is a shocking 20% increase from the previous year's estimate. Moreover, the number of countries with outbreaks also increased from 36 to 57 [4].

This surge critically underscores the precarious state of global immunity, a condition made even more fragile by disruptions to routine immunization services—such as those observed during the COVID-19 pandemic—and compounded by prevailing vaccine hesitancy across many populations [5,6]. In fact, insufficient coverage of immunization is the main impetus for these increasing outbreaks. More than 22 million children worldwide missed their first dose of the measles vaccine in 2023 alone—this creates growing susceptible cohorts that fuel epidemics [5].

The burden of measles falls on low- and middle-income countries. That is where health systems become even more strained, in many cases due to infrastructural, financial, and geographical barriers preventing access to strong equitable routine immunization services [7]. Measles can certainly be a disaster for morbidity and mortality within vulnerable populations but also severely strain already weak health care infrastructures by taking away badly needed resources from other important frontlines of public health, as well as increasing existing inequities in these vulnerable populations. The big picture goal of global measles elimination—as so unmistakably articulated in the World Health Organization (WHO) Immunization Agenda 2030—is decidedly imperiled at present by exactly such gaps in coverage, thus demanding recommitment and inspiring novelty in contextualized strategies directed toward achieving the lofty coverage rates necessary for eradication [8-10].

Prevention of measles and its complications is best achieved through vaccination. There are several live, attenuated measles vaccines that range from monovalent vaccines to combinations under the "measles-containing vaccine" name that also provide immunity against rubella, mumps, and varicella. Both doses of the measles vaccine currently in use worldwide are safe and effective. The first dose of the measles-containing vaccine (MCV1) was introduced in 1963, and by the 1980s, the WHO recommended that all countries implement routine MCV1 vaccination in their childhood immunization schedules. Ethiopia followed suit and introduced the MCV1 in the late 1980s [11,12].

Even following the introduction of the MCV1, which provides immunity to 85% of vaccine recipients, up to 15% of individuals do not develop long-term immunity following a single dose. MCV1 in some cases will also not generate an adequate immune response to provide immunity, a condition that comes to be known as primary vaccine failure. Identifying this gap, the WHO recommended in 2009 that countries introduce a second dose of the measles-containing vaccine (MCV2) into the national immunization schedules after they achieved 80% MCV1 uptake for 3 consecutive years. However, in 2017, the WHO revised this and recommended the introduction of MCV2 into national immunization schedules regardless of the level of uptake of MCV1 [13,14]. In response, Ethiopia introduced the MCV2 in February 2019 [15].

To maximize the effectiveness of vaccination, WHO guidelines stipulate that, in regions where the risk of measles mortality among infants is high, the MCV1 should be administered at 9 months of age followed by the MCV2 at the ages of 15 to 18 months. In lower-risk countries, the MCV1 may be administered at 12 months, with MCV2 timing adjusted to achieve the highest uptake by children [13]. Given the high risk of measles mortality in Ethiopia, both doses are administered at 9 and 15 months, respectively [16-18].

Even with national efforts and strategic frameworks adopted by the Ethiopian government as part of measles elimination efforts by 2030—strengthening routine immunization services, supplementary immunization activities, and surveillance—there remains a profound burden of measles in Ethiopia. Recent data clearly show an alarming reality that underscores gaps left behind in immunization coverage [19]. The United Nations Children's Fund and Centers for Disease Control and Prevention reports place Ethiopia consistently among those countries that have a large number of unvaccinated children [20,21]. Furthermore, it has always been among the topmost countries reporting measles cases, with both periodic and significant outbreaks across its regions [21].

Recent studies on measles epidemics in Ethiopia have described the entire picture of systemic and operational factors that contribute such challenges. A detailed review of outbreak investigations carried out in 2024 highlighted critical issues such as staff absenteeism, vaccines being out of stock at different levels of the supply chain, and shortage of operational funding leading to high dropout rates in vaccination and compromised service delivery [22]. Studies carried out in 2023 and 2024 further detailed such barriers from a qualitative perspective by introducing community-based factors—rural residence; low educational level of mothers among their sociodemographic characteristics; perceived safety of vaccines; inadequate health facilities within accessible distance; and inadequate community mobilization that impedes attainment of adequate coverage, particularly for the second dose [23,24]. Consequently, MCV2

XSL•FO
RenderX

uptake in Ethiopia persistently remains below both national targets and the critical 95% threshold necessary for achieving population-level herd immunity, contributing directly to the ongoing susceptibility of communities and the perpetuation of the disease [18,25]. Therefore, addressing this *last mile* gap in MCV2 coverage is identified as an urgent public health priority.

Objectives

Even though individual studies have looked at different aspects of vaccine acceptance and uptake across various regions of Ethiopia, there is still a lack of clear, comprehensive, and quantitative insights into the specific factors influencing MCV2 uptake at the national level. This systematic review and meta-analysis aimed to bring together all available data from Ethiopia on MCV2 uptake and its key determinants. By thoroughly synthesizing the existing evidence, we aimed to deepen the understanding of this pressing public health issue and support the development of more effective, evidence-based strategies to protect the health of Ethiopian children—and, ultimately, help achieve global and national measles elimination goals.

Methods

Study Design and Reporting

This systematic review and meta-analysis aimed to provide an estimate of the overall MCV2 uptake and determinants among children aged <5 years in Ethiopia. This review used the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. In addition, this review has been registered with PROSPERO with protocol CRD42024619031.

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria for this review can be found in Textbox 1.

Textbox 1. Inclusion and exclusion criteria for the systematic review and meta-analysis on the uptake of the second dose of the measles vaccine and its determinants among children aged <5 years in Ethiopia (criteria defined for study selection and rationale; 2025).

Inclusion criteria

- Setting: research conducted only in Ethiopia
- Population: studies involving children
- Language: published in English
- Study design: observational studies (longitudinal, cohort, cross-sectional, or case-control)
- Outcome variables: clearly defined outcome variables
- Quality: studies of sufficient quality
- Access: accessible through electronic resources

Exclusion criteria

- Setting: studies conducted outside of Ethiopia
- Population: studies focusing on adults or other populations
- Language: published in languages other than English
- Study design: experimental studies or reviews
- Outcome variables: undefined or unclear outcome variables
- Quality: low-quality studies
- Access: studies for which necessary information could not be obtained due to unavailability of the full text

Search Strategy and Source of Information

The search strategy was built using an adapted population, exposure, outcomes, study design, and setting framework to generate the MeSH (Medical Subject Headings) terms necessary to search for appropriate studies in the databases as follows:

1. Population: “children” and “pediatrics population”
2. Exposure: “second-dose measles vaccine” and “second-dose measles-containing vaccine”
3. Outcome: “uptake,” “utilization,” “coverage,” “associated factors,” “predictors,” “barriers,” or “determinants”
4. Study design: “observational studies”
5. Setting (context): “Ethiopia”

For a systematic review with primary studies as part of these studies, review questions focused on Ethiopia based on the aforementioned format were as follows: (1) what is the aggregate uptake of MCV2 among children aged <5 years in Ethiopia? (2) What associated factor explains the uptake of the MCV2 among children in Ethiopia?

A systematic search was conducted in various electronic databases: PubMed Central, Cochrane Library, Web of Science, Scopus, MEDLINE, Embase, and Google Scholar. The search used specific free-text terms and Boolean operator strings to locate studies pertinent to the aforementioned questions: (“Second-dose measles vaccine” OR “Second-doses measles-containing vaccine” OR “two-doses of measles vaccine” OR “measles vaccination”) AND (“children” OR “pediatrics”)

XSL•FO
RenderX

AND (“uptake” OR “utilization” OR “coverage” OR “acceptance” OR “adherence”) AND (“associated factors” OR “predictors” OR “barriers” OR “determinants” OR “influencers” OR “challenges”) AND (“Ethiopia” OR “Ethiopian context” OR “Ethiopian children”). After retrieving accessible articles, all the results were sorted, and duplicates were removed.

Study Selection and Quality Assessment

Two investigators, TEH and TAK, conducted a preliminary screening of studies by reviewing titles and abstracts to identify potentially relevant research before retrieving the full texts. If additional clarification was required to determine eligibility, they reached out to the corresponding author (MAA). Any disagreements between the investigators were resolved through discussion to achieve a consensus on inclusion or exclusion. All retrieved articles were imported into EndNote (version 20; Clarivate Analytics), where duplicate entries were eliminated. Next, the researchers reviewed the articles independently based on their structures and purposes, searching titles, abstracts, and full texts for studies that met the predetermined inclusion criteria. The articles for this review were selected and analyzed jointly by both reviewers.

To ensure the quality of each study, the adapted Newcastle-Ottawa Scale was used, which was reviewed by 3 independent assessors. The assessors undertook an inclusive review of all articles that met the inclusion criteria and study purpose. Differences between the reviewers were resolved through consensus, and a fourth reviewer was involved in adjudication as necessary.

The criteria for the assessment addressed a series of key areas: representativeness of the sample, adequacy of the sample size, response rates, respondent and nonrespondent characteristics, quality of the intended measurement tools for exposure or risk factors, comparability of outcome groups based on study design, controlling for major confounding variables, and measurement of outcome and use of appropriate statistical tests. Each of the studies was scored out of 10 points, with scores of >7 indicating low risk of bias for inclusion. All studies included in the analysis had a low risk of bias, with scores ranging from 9 to 10 (the full Newcastle-Ottawa Scale results can be found in Table S1 in Multimedia Appendix 1 [26-36]).

Data Extraction

The data extraction form was developed using a Microsoft Excel spreadsheet. Two independent reviewers (MAA and ABZ) systematically extracted data from the full-text articles. The data extraction form included key variables such as the first author’s name, year of publication, study region, study design, sample size, age of the participants, sample distribution by sex, uptake or coverage, and the number of children vaccinated with MCV2.

For each primary outcome, the uptake of MCV2 was recorded, whereas adjusted odds ratios (AORs) for each associated factor (secondary outcome) were also noted along with 95% CIs. Any discrepancies between the data extractors were addressed through discussion and consensus with another author.

Statistical Analysis

After extracting the data, a meta-analysis was conducted using Stata (version 17; StataCorp). Uptake estimates were calculated along with their corresponding SEs using the formulas $p = r/n$ and $SE = \sqrt{p(1-p)/n}$, where p represents the proportion, r is the total number of children vaccinated with MCV2, and n is the sample size. The results of the meta-analysis are presented as the pooled uptake of MCV2 accompanied by 95% CIs. A significance threshold was established at P values of <.05.

To investigate the factors influencing the uptake of MCV2, Stata (version 17) was used. A random-effects model was applied as heterogeneity between studies was high, as confirmed through an I^2 statistic of >75% and a P value of <.05. The I^2 index and Cochran Q test were used for testing heterogeneity. The I^2 statistic ranges from 0% to 100%, with 0% indicating no heterogeneity and 100% indicating high heterogeneity.

The effect of individual studies on the overall prevalence estimate was tested using sensitivity analysis. Publication bias was also tested using funnel plots for symmetry assessment, as well as the Egger test and nonparametric trim-and-fill tests. Although both these tests were used, the Egger test is commonly deemed superior for detecting publication bias because it provides actual effect sizes along with their accuracy. Contrary to this, the funnel plot is a less objective measure of potential asymmetry and may hint at publication bias.

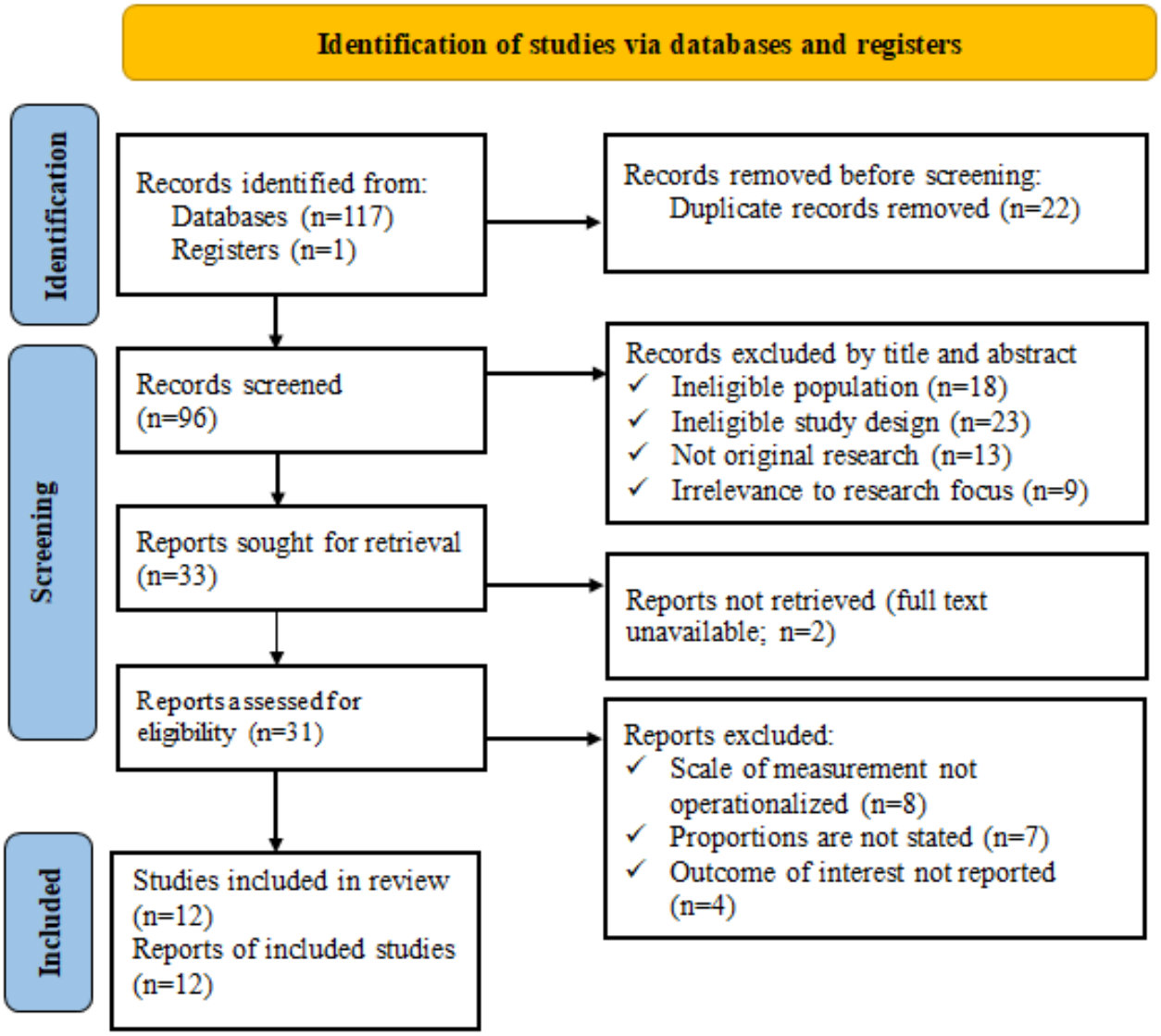
Results

Literature Search

This systematic review and meta-analysis followed the PRISMA guidelines (Multimedia Appendix 2). A comprehensive search of PubMed Central, Cochrane Library, Web of Science, Scopus, MEDLINE, and Embase initially yielded 118 articles addressing MCV2 uptake and its predictors among Ethiopian children. Of these 118 articles, after removing 22 (18.6%) duplicate records, 96 (81.4%) underwent title and abstract screening. Of these 96 articles, 63 (66%) were excluded for the following reasons: ineligible population (n=18, 29%), ineligible study design (n=23, 37%), nonoriginal research (n=13, 21%), or irrelevance to the research focus (n=9, 14%). The remaining 33 full-text articles were assessed for retrieval, with 2 (6%) records unavailable. Of the 31 articles evaluated in full text, 19 (61%) were excluded due to non-operationalized measurement scales (n=8, 42%), unstated proportions (n=7, 37%), or unreported outcomes of interest (n=4, 21%). Ultimately, 12 studies met all the inclusion criteria and were included in the final analysis (Figure 1).

XSL•FO
RenderX

Figure 1. Identification of studies via databases and registers.



Study Characteristics

The total review comprised 12 studies on 7847 children from all regions of the country. Specifically, the proportion of studies by region was 50% (6/12) of the studies conducted in Amhara, 25% (3/12) in Oromia, 8% (1/12) in Somalia, and 17% (2/12) from all regions. A total of 92% (11/12) of the studies were cross-sectional, and 8% (1/12) were case-control studies. All studies (12/12, 100%) were conducted during the period from 2019 to 2023. The average number of participants per study was 654 (SD 254; range 372-1172). Almost half (49%) of the participants were female (Table 1).

XSL•FO
RenderX

Table 1. Descriptive summary of the studies analyzed in the meta-analysis on measles vaccine second dose uptake and its determinants among children aged <5 years in Ethiopia (including authors, study design, participant ages, sample size, locations, and key findings).

Study	Year conducted	Region	Design	Participant age (y)	Sample size, n	Male participants, n	Female participants, n	Prevalence, %	Outcome
Taffie et al [26]	2022	Amhara	CS[a]	24-35	418	192	226	41.39	173
Tadesse et al [27]	2022	Oromia	CS	15-36	372	__[b]	—	42.5	158
Adisu et al [28]	2023	Amhara	CS	24-36	621	319	302	75.68	470
Demewoz et al [29]	2020	Amhara	CS	24-35	837	406	431	48.1	403
Woyessa et al [23]	2021	Oromia	CS	24-35	1172	558	613	48	563
Goshu Muluneh et al [30]	2019	All regions	CS	12-36	965	545	420	12.36	119
Ibrahim et al [31]	2023	Somalia	CS	15-36	429	215	214	21.4	92
Nurgi et al [32]	2023	Oromia	CC[c]	18-24	446	—	—	75.3	336
Teshale and Amare [33]	2019	All regions	CS	24-35	800	424	376	9.84	79
Zeleke et al [34]	2022	Amhara	CS	15-23	633	332	301	53.08	336
Mulatu et al [35]	2022	Amhara	CS	24-35	732	390	342	63.3	463
Asichalew [36]	2021	Amhara	CS	24-35	422	199	223	53	224

[a]CS: cross-sectional.

[b]Not available.

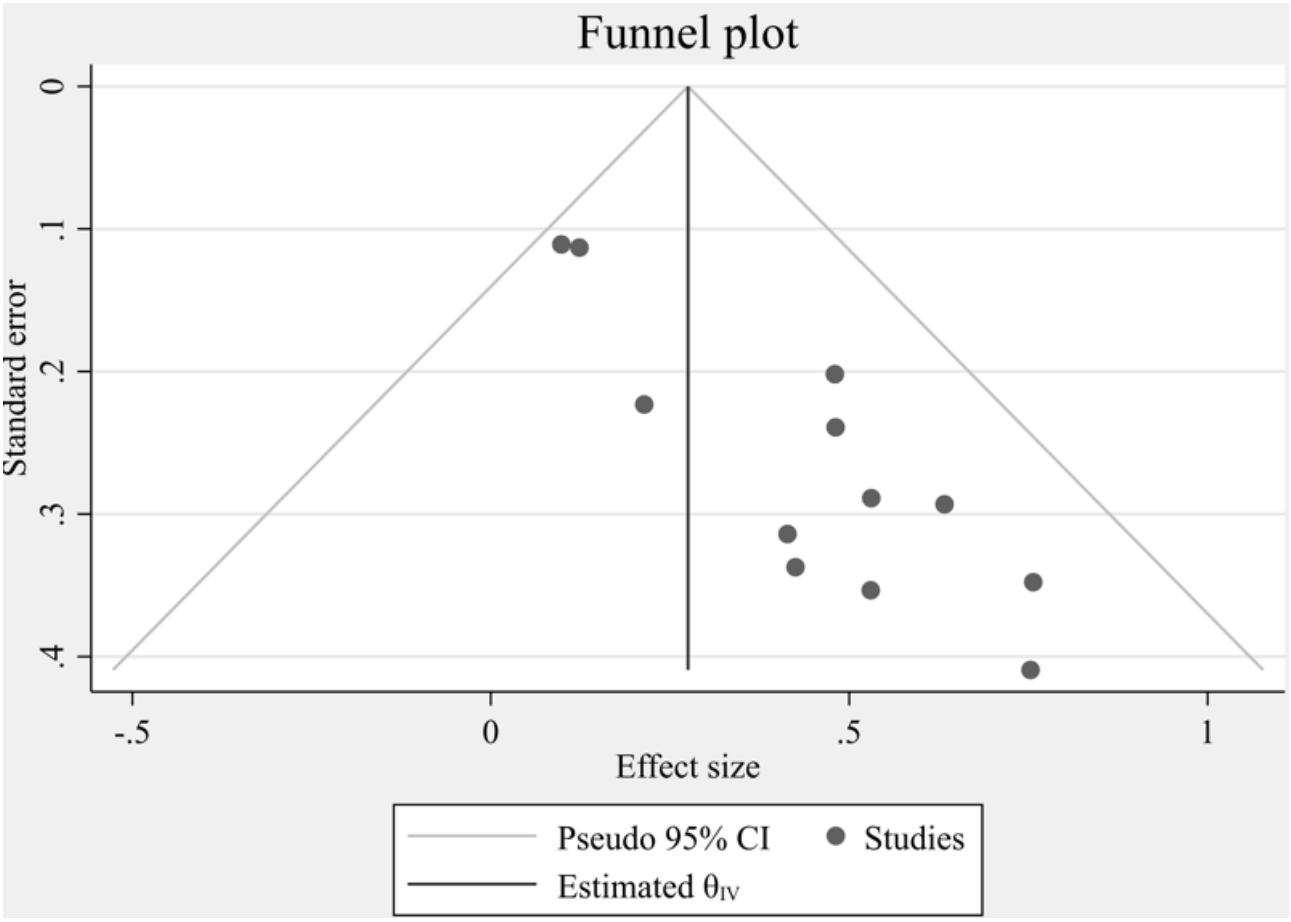
[c]CC: case-control.

Publication Bias

The publication bias was estimated through funnel plot analysis and the Egger regression test. The funnel plot revealed an asymmetrical shape, intuitively suggesting the presence of publication bias (Figure 2). The Egger regression test also yielded a *P* value of .001, indicating statistical evidence of publication bias. This agreement between subjective and objective analysis strengthens the likelihood of bias in the published studies.

XSL•FO
RenderX

Figure 2. Funnel plot illustrating publication bias for second-dose measles vaccine uptake among children younger than 5 years in Ethiopia (analysis conducted in 2025), reflecting study distribution and variability.



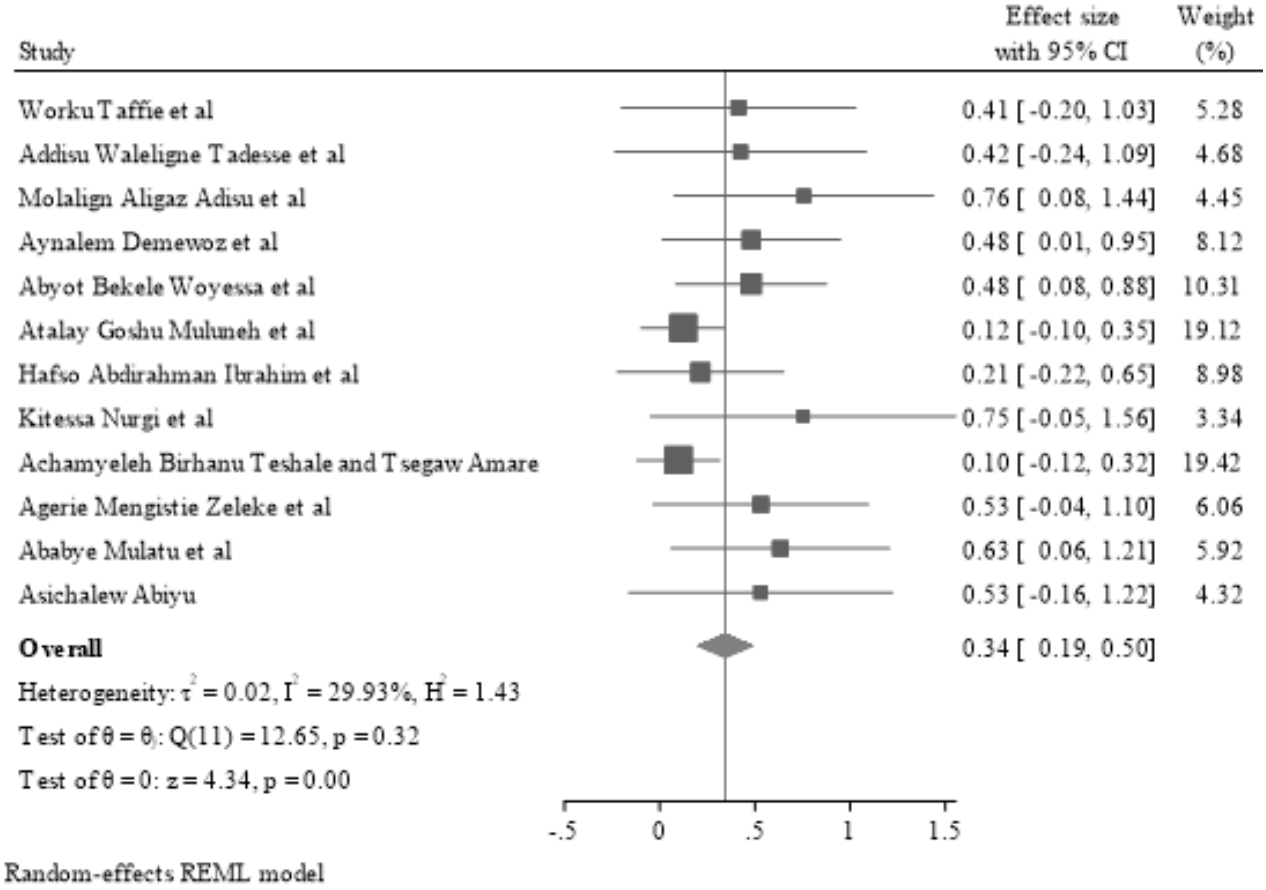
To explain the potential issue of publication bias indicated by the asymmetrical funnel plot, a nonparametric trim-and-fill analysis was conducted to calculate the overall effect. The trim-and-fill analysis flagged 18 studies and imputed 6 studies to calculate the potential bias. The observed effect size was 34.4 (95% CI 18.8-49.9). Following imputation of the studies, the effect size was 20.6 and was well within the CI of the observed effect. This means that the imputed studies had zero or negligible contribution to the overall effect size. The trim-and-fill analysis was used to include any missing studies due to publication bias and provide a better estimate of the true effect size. The tiny difference between the observed and adjusted effect sizes informed us that publication bias was likely not a significant problem in this dataset.

Pooled Uptake of MCV2 Among Children in Ethiopia

The pooled uptake of MCV2 among children in Ethiopia was 34.4% (95% CI 18.8%-49.9%), and there was low heterogeneity between the studies (P=.32; τ^2=0.02; I^2=29.9%; H=1.41; Figure 3 [26-36]).

XSL•FO
RenderX

Figure 3. Forest plot depicting the pooled uptake of second-dose measles vaccine among children younger than 5 years in Ethiopia.

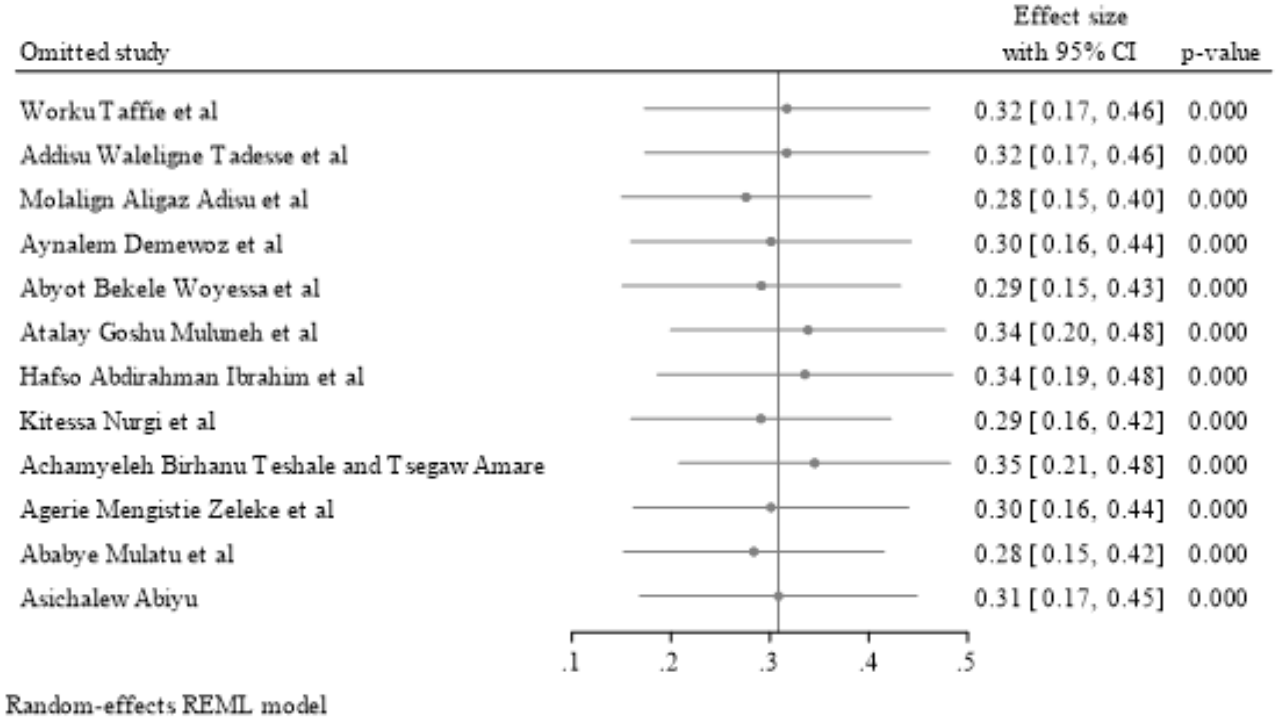


Sensitivity Analysis

Sensitivity analysis was conducted to evaluate the influence of outlying or potentially influential studies on the overall uptake estimate of the MCV2 among children. Through a random-effects model, the results of this analysis showed that there were no influential studies or outliers detected because all the point estimates fell within the limits of the 95% CI. This finding indicates that the overall estimate of uptake was robust and not determined to a large degree by a single study, thus providing additional strength for the validity of the pooled results. The full sensitivity analysis results are presented in Figure 4 [26-36].

XSL•FO
RenderX

Figure 4. Results of the sensitivity analysis on the uptake of second-dose measles vaccine among under-five children in Ethiopia.



Factors Associated With the Uptake of MCV2 in Ethiopia

The pooled odds ratio for mothers with the highest educational attainment as indicated in the studies was 3.31 times greater for vaccinating their children with MCV2 than for mothers with low or no education (AOR 3.31, 95% CI 1.32-5.30). Moreover, mothers who attended antenatal care (ANC) follow-ups were over twice as likely to vaccinate their children with MCV2 compared to those who did not attend ANC follow-ups (AOR 2.02, 95% CI 1.12-2.92). Similarly, mothers who used postnatal care (PNC) services were 3 times more likely to vaccinate their children with MCV2 than mothers who did not use PNC services (AOR 3.03, 95% CI 1.77-4.28).

In addition, mothers who had a waiting time of <30 minutes at vaccination centers were 2.5 times more likely to vaccinate their children than those with waiting times of >30 minutes (AOR 2.56, 95% CI 1.98-3.13). Moreover, mothers who were highly aware of measles vaccination were twice as likely to vaccinate their children with the second dose as mothers who were lowly aware (AOR 2.17, 95% CI 1.59-2.74). Similarly, the odds of being vaccinated with the MCV2 were 3.5 times higher (AOR 3.58, 95% CI 1.97-6.30) among children of mothers who had a positive perception of the measles vaccination compared to those of mothers with a poor perception (Table 2).

Table 2. Pooled estimates of determinants affecting the uptake of the second dose of the measles vaccine among children aged <5 years in Ethiopia (analyzed in 2025).

Associated factors	Pooled effect size (95% CI)	I^2 (%)	Egger test *P* value	Studies pooled
Mothers who attained the highest level of education	3.31 (1.32-5.30)	0	—[a]	[23,31]
Mothers who attended ANC[b] service follow-up	2.02 (1.12-2.92)	0	.40	[28,34]
Mothers who attended PNC[c] service follow-up	3.03 (1.77-4.28)	18.64	.19	[26,28,31,32]
Waiting time of <30 min at the vaccination site	2.56 (1.98-3.13)	0	.62	[27,34]
Mothers who had good awareness	2.17 (1.59-2.74)	0	.01	[26-29,31]
Mothers who had a positive perception	4.30 (2.63-5.97)	0	.35	[28,29]

[a]Not available.

[b]ANC: antenatal care.

[c]PNC: postnatal care.

XSL•FO
RenderX

Discussion

Principal Findings

Pooled uptake of MCV2 among children aged <5 years in Ethiopia was 34.4% (95% CI 18.8%-49.9%). This result is in line with the statistics from Nigeria, Madagascar, and Afghanistan, where low uptake rates have been documented [37-39]. However, this figure is significantly lower than the WHO target of 95% for MCV2 uptake, as well as the national immunization targets set by Ethiopia and other countries [38,40-42]. A WHO report showed that there are countries even below this level of uptake, such as the Democratic Republic of the Congo, Uganda, and Angola [38], which poses a substantial risk to public health in these regions. Several factors contribute to the low vaccination rates in Ethiopia. Ongoing strife and instability in certain regions have significantly hampered health care provision and access to vaccine services. Internal displacement and conflict have led to destruction of health infrastructure, which is making it increasingly difficult for families to access core immunization services. In addition, socioeconomic obstacles for several families, such as illiteracy and poverty, further compound the problem, leading to poor awareness and prioritization of vaccination. The consequences of the very low uptake are extensive. A high level of vaccination uptake should be reached if measles epidemics are to be avoided and herd immunity is to be achieved. All this emphasizes once more the imperative of urgent specific interventions aimed at improving vaccine uptake. Some likely options include increasing public awareness regarding the importance of the second dose, enhanced health service access, and targeting obstacles keeping families from vaccinating their children. Without significant improvements in these areas, the goal of measles elimination in Ethiopia and similar contexts will remain unattainable.

The pooled odds ratio indicated that mothers with the highest educational attainment were 3.31 times more likely to vaccinate their children with MCV2 compared to mothers with low or no education. Similar observations have been recorded in other sub-Saharan countries [41,43]. This is because those mothers who are better educated to some level will understand the benefit of vaccination, appreciate the hazard of vaccine-preventable diseases, and be in a position to avail their children care. This underscores the value of targeted educational interventions to increase awareness and knowledge of the importance of immunization, particularly among more poorly educated communities. Through investment in education, public health interventions will be able to improve levels of vaccination uptake and, ultimately, child health outcomes among high-risk groups.

The findings indicate that maternal attendance to ANC and PNC services significantly influences whether children are immunized with MCV2. Specifically, mothers who attended ANC follow-ups had twice the chances of vaccinating their children compared to those who did not attend ANC follow-ups. Similarly, mothers using PNC services were 3 times more likely to vaccinate their children with MCV2 than those who did not receive PNC services. Similar results have been reported in sub-Saharan African and South Asian countries [41,44]. This is because women attending ANC and PNC are most likely to know more about health habits, including the significance of vaccinations. During these visits to health care facilities, there are opportunities for health care workers to inform mothers of the advantages of vaccination, clarify queries, and provide support for appropriate timing of vaccination. All these findings demonstrate the importance of maternal health care uptake in the initiation of immunization.

The combined outcome revealed that mothers who waited for <30 minutes at the vaccine center were 2.5 times more likely to vaccinate their children compared to those waiting for >30 minutes. This is consistent with findings reported in other studies [45]. This is because shorter waiting times enhance perceived convenience in attending vaccination sessions. When mothers enjoy a smooth experience, they tend to perceive vaccination as possible and, hence, have a high compliance. In contrast, longer waiting times provide opportunities for mothers to become exhausted and stressed, especially with respect to children, who are most likely to become restless. Reduced waiting times create a smoother experience, allowing mothers to proceed with vaccinations easily. Apart from this, time is a valuable resource for most families, with women typically having a work schedule as well as caring duties. Thus, it is preferable to minimize waits to counter the opportunity cost of attending vaccinations. This provides evidence of the significant impacts of waiting times on acceptance of vaccinations, suggesting the role of the efficiency of the vaccination process in shaping immunization uptake.

The combined findings indicate that maternal perception and awareness of measles vaccination exert a significant influence on children vaccination uptake. Mothers with high awareness of measles vaccination were more than twice as likely to vaccinate their children with the second dose compared with mothers with low awareness. In addition, the odds of vaccinating children with the second dose were 3.5 times higher among mothers with a positive perception of measles vaccination. This outcome is consistent with those of a study conducted in sub-Saharan Africa [43,46]. This is because educated mothers who comprehend the value and advantages of measles vaccination are most likely to place greater emphasis on immunization among their children. Effective awareness helps in clearing misconceptions and myths that could lead to vaccine hesitation [46,47]. Moreover, when mothers believe in the efficacy and safety of vaccines, they are more likely to adhere to immunization schedules. Such beliefs may be based on personal experience, social norms, or successful health promotion. These findings highlight the exceptional significance of enhancing maternal attitudes and perceptions regarding measles vaccination. Vaccine benefit knowledge can significantly boost rates of vaccination among children. Public health interventions should be grounded in educational activities and community mobilization to establish an enabling environment for immunization.

This study has several limitations that warrant acknowledgment. First, the small sample size and cross-sectional design restrict both statistical power and the ability to infer causality or temporal relationships concerning vaccine acceptance. A

XSL•FO
RenderX

significant limitation stems from the potential publication bias observed among the included studies (Egger test: *P*=.001), which, despite mitigation efforts such as trim-and-fill analysis, could still influence the pooled estimates. This is compounded by the reliance on self-reported data within many of the included studies, introducing potential social desirability and recall biases, particularly for subjective measures such as awareness. Furthermore, there was a notable geographical imbalance, with a disproportionate number of included studies originating from the Amhara region and underrepresentation of other diverse Ethiopian contexts. This limits the generalizability of our findings across the nation's varied cultural and health care landscapes. Finally, the analysis did not fully account for the impact of temporal factors and policy changes, especially given the introduction of MCV2 in Ethiopia in 2019 and the varying time frames of the included studies. Important covariates such as cultural beliefs; accessibility challenges (eg, rural vs urban disparities); and conflict-related disruptions, which are known to influence vaccination rates in Ethiopia, were also not explicitly explored. Future research should aim to address these gaps to provide a more comprehensive understanding of vaccine uptake determinants.

Conclusions and Recommendations

The pooled MCV2 uptake among Ethiopian children aged <5 years was 34%, which is far below the national target of 95%. The level of education of mothers, use of ANC and PNC services, waiting time at vaccination sites, and mothers' awareness and perceptions of measles vaccination were statistically significant predictors of uptake of MCV2. To increase the uptake of MCV2 in Ethiopia, it is necessary to target interventions. Interventions should target the strengthening of community-based educational programs, mobile clinics for remote areas, enhanced access to ANC and PNC services, reduction of waiting times at the point of vaccine delivery, and raising awareness about the benefits of measles vaccination. These multistakeholder initiatives are crucial in achieving the national immunization objective.

Acknowledgments

The authors acknowledge the authors of the included studies for their original contributions.

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author (MAA) on reasonable request.

Authors' Contributions

MAA contributed to conceptualization, data curation, formal analysis, methodology, software, visualization, writing—original draft, and writing—review and editing. TEH contributed to data curation, formal analysis, methodology, writing—original draft, and writing—review and editing. TAK contributed to formal analysis, methodology, writing—original draft, and writing—review and editing. ABZ contributed to conceptualization, formal analysis, methodology, writing—original draft, and writing—review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Newcastle-Ottawa Scale scores for included studies.
[PDF File (Adobe PDF File), 214 KB - publichealth_v11i1e77195_app1.pdf]

Multimedia Appendix 2
PRISMA 2020 Checklist.
[DOCX File , 36 KB - publichealth_v11i1e77195_app2.docx]

References

1. Bankamp B, Kim G, Hart D, Beck A, Ben Mamou M, Penedos A, et al. Global update on measles molecular epidemiology. Vaccines (Basel) 2024 Jul 22;12(7):810 [FREE Full text] [doi: 10.3390/vaccines12070810] [Medline: 39066448]
2. Dattani S. Measles leaves children vulnerable to other diseases for years. Our World in Data. 2025 Jun 16. URL: https://ourworldindata.org/measles-increases-disease-risk [accessed 2025-08-08]
3. Guerra FM, Bolotin S, Lim G, Heffernan J, Deeks SL, Li Y, et al. The basic reproduction number (R) of measles: a systematic review. Lancet Infect Dis 2017 Dec;17(12):e420-e428. [doi: 10.1016/S1473-3099(17)30307-9] [Medline: 28757186]

XSL•FO
RenderX

4. Minta AA, Ferrari M, Antoni S, Lambert B, Sayi TS, Hsu CH, et al. Progress toward measles elimination - worldwide, 2000-2023. MMWR Morb Mortal Wkly Rep 2024 Nov 14;73(45):1036-1042 [FREE Full text] [doi: 10.15585/mmwr.mm7345a4] [Medline: 39541251]
5. Goje O, Kapoor A. Meeting the challenge of vaccine hesitancy. Cleve Clin J Med 2024 Sep 04;91(9 suppl 1):S50-S56 [FREE Full text] [doi: 10.3949/ccjm.91.s1.08] [Medline: 39231603]
6. Minta AA, Ferrari M, Antoni S, Portnoy A, Sbarra A, Lambert B, et al. Progress toward measles elimination - worldwide, 2000-2022. MMWR Morb Mortal Wkly Rep 2023 Nov 17;72(46):1262-1268 [FREE Full text] [doi: 10.15585/mmwr.mm7246a3] [Medline: 37971951]
7. Bantie B, Atnafu Gebeyehu N, Adella GA, Ambaw Kassie G, Mengstie MA, Abebe EC, et al. Mapping geographical inequalities of incomplete immunization in Ethiopia: a spatial with multilevel analysis. Front Public Health 2024 Jun 7;12:1339539 [FREE Full text] [doi: 10.3389/fpubh.2024.1339539] [Medline: 38912271]
8. Okesanya OJ, Olatunji G, Olaleke NO, Mercy MO, Ilesanmi AO, Kayode HH, et al. Advancing immunization in Africa: overcoming challenges to achieve the 2030 global immunization targets. Adolesc Health Med Ther 2024 Oct 19;15:83-91 [FREE Full text] [doi: 10.2147/AHMT.s494099]
9. Dixon MG, Ferrari M, Antoni S, Li X, Portnoy A, Lambert B, et al. Progress toward regional measles elimination - worldwide, 2000-2020. MMWR Morb Mortal Wkly Rep 2021 Nov 12;70(45):1563-1569 [FREE Full text] [doi: 10.15585/mmwr.mm7045a1] [Medline: 34758014]
10. Routine immunization catch-up vaccination guidelines. Federal Ministry of Health, Ethiopia. 2022. URL: https://publications.jsi.com/JSIInternet/Inc/Common/_download_pub.cfm?id=25790&lid=3 [accessed 2025-08-15]
11. Berche P. History of measles. Presse Med 2022 Sep;51(3):104149 [FREE Full text] [doi: 10.1016/j.lpm.2022.104149] [Medline: 36414136]
12. Patel MK, Goodson JL, Alexander JPJ, Kretsinger K, Sodha SV, Steulet C, et al. Progress toward regional measles elimination - worldwide, 2000-2019. MMWR Morb Mortal Wkly Rep 2020 Nov 13;69(45):1700-1705 [FREE Full text] [doi: 10.15585/mmwr.mm6945a6] [Medline: 33180759]
13. A guide to introducing a second dose of measles vaccine into routine immunization schedules. World Health Organization. 2013 Aug 19. URL: https://www.who.int/publications/i/item/WHO-IVB-13.03 [accessed 2025-08-08]
14. Biellik RJ, Davis R. The new World Health Organization recommendation on the 2-dose measles vaccine schedule and the way forward in African region. Pan Afr Med J 2017 Jun 21;27(Suppl 3):14 [FREE Full text] [doi: 10.11604/pamj.supp.2017.27.3.11611] [Medline: 29296149]
15. Ethiopia launches measles vaccine second dose (MCV2) introduction: over 3.3 million children will receive the vaccine annually. World Health Organization. 2019 Feb 11. URL: https://reliefweb.int/report/ethiopia/ethiopia-launches-measles-vaccine-second-dose-mcv2-introduction-over-33-million#:~:text=Addis [accessed 2025-08-08]
16. National Immunization Program Comprehensive Strategic Plan 2021-2025. Ministry of Health Ethiopia. 2021 Jul. URL: https://www.collegesidekick.com/study-docs/2499137 [accessed 2025-08-09]
17. Trend of coverage of measles - containing vaccine second dose (MCV2) in Africa. World Health Organization. URL: https://tinyurl.com/mrx7x6sf [accessed 2025-08-08]
18. Adisu MA. Timeliness of the second dose of measles-containing vaccine uptake and its determinants among children aged 24-36 months in Gondar City, Northwest Ethiopia, 2023: community-based cross-sectional study design. J Virus Erad 2025 Mar 17;11(2):100594 [FREE Full text] [doi: 10.1016/j.jve.2025.100594] [Medline: 40226158]
19. Ethiopia conducts a measles vaccination campaign to respond to an outbreak in three regions. World Health Organization. 2022 Apr 21. URL: https://www.afro.who.int/countries/ethiopia/news/ethiopia-conducts-measles-vaccination-campaign-respond-outbreak-three-regions [accessed 2025-08-08]
20. More than 870,000 children in Ethiopia miss out on lifesaving measles, diphtheria and tetanus vaccines in 2018. UNICEF. 2019 Jul 16. URL: https://www.unicef.org/ethiopia/press-releases/more-870000-children-ethiopia-miss-out-lifesaving-measles-diphtheria-and-tetanus [accessed 2025-08-09]
21. Global measles outbreaks. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/global-measles-vaccination/data-research/global-measles-outbreaks/index.html [accessed 2025-08-09]
22. Alemu MA, Dema ST, Anteneh HA, Guade HT, Gebrekidan AH, Mersha AM, et al. Insights from a measles outbreak root cause analysis in Ethiopia - 2024. Pan Afr Med J 2025;51(1):8 [FREE Full text]
23. Woyessa AB, Shah MP, Azmeraye BM, Pan J, Lisanwork L, Yimer G, et al. Factors associated with uptake of routine measles-containing vaccine doses among young children, Oromia Regional State, Ethiopia, 2021. Vaccines (Basel) 2024 Jul 11;12(7):762 [FREE Full text] [doi: 10.3390/vaccines12070762] [Medline: 39066400]
24. Solomon K, Aksnes BN, Woyessa AB, Sadi CG, Matanock AM, Shah MP, et al. Qualitative insights on barriers to receiving a second dose of measles-containing vaccine (MCV2), Oromia region of Ethiopia. Vaccines (Basel) 2024 Jun 22;12(7):702 [FREE Full text] [doi: 10.3390/vaccines12070702] [Medline: 39066340]
25. WHO/UNICEF estimates of national immunization coverage. World Health Organization. URL: https://www.who.int/teams/immunization-vaccines-and-biologicals/immunization-analysis-and-insights/global-monitoring/immunization-coverage/who-unicef-estimates-of-national-immunization-coverage [accessed 2025-08-08]

XSL•FO
RenderX

26. Taffie W, Temesgen H, Ashebir W, Mekonen H. Measles second dose vaccine uptake and its associated factors among children aged 24-35 months in Northwest Ethiopia, 2022. Sci Rep 2024 May 14;14(1):11059 [FREE Full text] [doi: 10.1038/s41598-024-61048-9] [Medline: 38744908]
27. Tadesse AW, Sahlu D, Benayew M. Second-dose measles vaccination and associated factors among under-five children in urban areas of North Shoa Zone, Central Ethiopia, 2022. Front Public Health 2022 Dec 9;10:1029740 [FREE Full text] [doi: 10.3389/fpubh.2022.1029740] [Medline: 36568740]
28. Adisu MA, Bogale WA, Alemu TG. Second dose of measles-containing vaccine coverage and associated factors among children aged 24-36 months in Gondar city, Central Gondar, Northwest Ethiopia, 2023. Front Public Health 2024 May 2;12:1364865 [FREE Full text] [doi: 10.3389/fpubh.2024.1364865] [Medline: 38756896]
29. Demewoz A, Wubie M, Mengie MG, Kassegn EM, Jara D, Aschale A, et al. Second dose measles vaccination utilization and associated factors in Jabitehnan District, Northwest Ethiopia. Dose Response 2023 Mar 10;21(1):15593258231164042 [FREE Full text] [doi: 10.1177/15593258231164042] [Medline: 36923301]
30. Goshu Muluneh A, Woldemariam Merid M, Tigabu B, Getie Ferede M, Molla Kassa G, Animut Y. Less than one-fifth of Ethiopian children were vaccinated for measles second dose; evidence from the Ethiopian mini demographic and health survey 2019. Vaccine X 2022 Sep 14;12:100217 [FREE Full text] [doi: 10.1016/j.jvacx.2022.100217] [Medline: 36148266]
31. Ibrahim HA, Wariyo A, Asefa EM, Cheru A, Abebe Lonsako A, Dirirsa G. Measles second dose vaccine uptake and associated factors among under-five children in Jigjiga City, Somali Region, Eastern Ethiopia: a community-based cross-sectional study. Front Public Health 2024 Jul 30;12:1395802 [FREE Full text] [doi: 10.3389/fpubh.2024.1395802] [Medline: 39139669]
32. Nurgi K, Ahmed S, Ganfure G, Bekele GG. Determinants of measles second dose vaccination dropout among children aged 18-24 months in Ejere woreda, central Ethiopia; unmatched case-control study. Front Pediatr 2024 Sep 18;12:1432762 [FREE Full text] [doi: 10.3389/fped.2024.1432762] [Medline: 39359739]
33. Teshale AB, Amare T. Exploring spatial variations and the individual and contextual factors of uptake of measles-containing second dose vaccine among children aged 24 to 35 months in Ethiopia. PLoS One 2023 Jan 4;18(1):e0280083 [FREE Full text] [doi: 10.1371/journal.pone.0280083] [Medline: 36598928]
34. Zeleke AM, Ferede YA, Assfa GW, Lakew ZH. Measles second dose vaccine uptake and determinants among 15–23 month old children at Bahir Dar, Northwest Ethiopia. J Midwifery Reprod Health 2025 Jan;13(1):4510-4662. [doi: 10.22038/JMRH.2024.71216.2089]
35. Mulatu A, Tsega Y, Cherie N, Kasaye MD, Mekonen AM. Uptake of second dose measles containing vaccine and associated factors among children aged 24-35 months in central Ethiopia: a community based cross-sectional study. BMC Public Health 2025 Apr 21;25(1):1470 [FREE Full text] [doi: 10.1186/s12889-025-22750-z] [Medline: 40259257]
36. Asichalew A. Utilization of measles conjugated vaccine 2 and associated factors among children aged 24-35 months in north Mecha district, west Gojjam zone, north west Ethiopia. Bahir Dar University. URL: http://ir.bdu.edu.et/bitstream/handle/123456789/13817/ASICHALEW%20ABIYU.pdf?sequence=1 [accessed 2025-08-09]
37. Sato R. Trends of the second measles vaccine (MCV2) over time after its launch as part of routine immunization in Nigeria: a brief research report. Front Public Health 2024;12:1392996 [FREE Full text] [doi: 10.3389/fpubh.2024.1392996] [Medline: 39529712]
38. Measles-containing-vaccine second-dose (MCV2) immunization coverage by the locally recommended age (%) (WUENIC). World Health Organization. URL: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/measles-containing-vaccine-second-dose-(mcv2)-i [accessed 2025-08-08]
39. Immunization dashboard, Afghanistan Reported cases of vaccine-preventable diseases (VPDs). World Health Organization. URL: https://immunizationdata.who.int/dashboard/regions/eastern-mediterranean-region/AFG [accessed 2025-08-15]
40. Measles and rubella strategic framework: 2021-2030. World Health Organization. 2020 Nov 08. URL: https://www.who.int/publications/i/item/measles-and-rubella-strategic-framework-2021-2030 [accessed 2025-08-08]
41. Chilot D, Belay DG, Shitu K, Gela YY, Getnet M, Mulat B, et al. Measles second dose vaccine utilization and associated factors among children aged 24-35 months in Sub-Saharan Africa, a multi-level analysis from recent DHS surveys. BMC Public Health 2022 Nov 12;22(1):2070 [FREE Full text] [doi: 10.1186/s12889-022-14478-x] [Medline: 36371164]
42. Sugishita Y, Kurita J, Akagi T, Sugawara T, Ohkusa Y. Determinants of vaccination coverage for the second dose of measles-rubella vaccine in Tokyo, Japan. Tohoku J Exp Med 2019 Dec;249(4):265-273 [FREE Full text] [doi: 10.1620/tjem.249.265] [Medline: 31852852]
43. Melis T, Mose A, Fikadu Y, Haile K, Habte A, Jofiro G. Predictors for low coverage of uptake of second dose of measles vaccine among children in sub-Saharan Africa, 2023: a systematic review and meta-analysis. J Pharm Policy Pract 2024 Dec 07;17(1):2285507 [FREE Full text] [doi: 10.1080/20523211.2023.2285507] [Medline: 38205190]
44. Mahazabin M, Tabassum N, Syfullah SM, Majumder UK, Islam MA. Socio-demographic factors affecting the first and second dose of measles vaccination status among under-five children: perspectives from South Asian countries. Prev Med Rep 2024 Jul 31;45:102839 [FREE Full text] [doi: 10.1016/j.pmedr.2024.102839] [Medline: 39188972]
45. Magodi R, Mmbaga EJ, Massaga J, Lyimo D, Mphuru A, Abade A. Factors associated with non-uptake of measles-rubella vaccine second dose among children under five years in Mtwara district council, Tanzania, 2017. Pan Afr Med J 2019 May 29;33:67 [FREE Full text] [doi: 10.11604/pamj.2019.33.67.17055] [Medline: 31448029]

XSL•FO
RenderX

46. Makokha FM. Uptake of second dose of measles vaccine among children in Kakamega County, Kenya. Jomo Kenyatta University of Agriculture and Technology. 2016. URL: http://ir.jkuat.ac.ke/bitstream/handle/123456789/2396/MAKOKHA%20FREDRICK%20MIKE,%20MSC%20APPLIED%20EPIDEMIOLOGY,%20%202016..pdf?sequence=1&isAllowed=y [accessed 2025-08-09]
47. Oluseye OM, Jimoh NA, Ogunleye CA. Knowledge and attitude of mothers towards measles and measles, mumps and rubella (MMR) vaccine in IDI- ABA community Abeokuta, Nigeria. Open J Med Res 2021 Nov 30;2(2):136-148. [doi: 10.52417/ojmr.v2i2.295]

Abbreviations

ANC: antenatal care
AOR: adjusted odds ratio
MCV1: first dose of the measles-containing vaccine
MCV2: second dose of the measles-containing vaccine
MeSH: Medical Subject Headings
PNC: postnatal care
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 09.05.25; peer-reviewed by A Dessie, A Girma; comments to author 06.07.25; revised version received 17.07.25; accepted 22.07.25; published 27.08.25.

Please cite as:
Adisu MA, Habtie TE, Kitaw TA, Zemariam AB
Uptake of the Second Dose of the Measles Vaccine and Its Determinants Among Children Aged Less Than 5 Years: Systematic Review and Meta-Analysis
JMIR Public Health Surveill 2025;11:e77195
URL: https://publichealth.jmir.org/2025/1/e77195
doi:10.2196/77195
PMID:40864500

©Molalign Aligaz Adisu, Tesfaye Engdaw Habtie, Tegene Atamenta Kitaw, Alemu Birara Zemariam. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Risk Factors for Typhoid Fever: Systematic Review

Portia Boakye Okyere[1], MPhil; Sampson Twumasi-Ankrah[2], PhD; Sam Newton[1], PhD; Samuel Nkansah Darko[3], PhD; Michael Owusu Ansah[4], MD; Eric Darko[5], MSc; Francis Agyapong[5], MD; Hyon Jin Jeon[6], MPH; Yaw Adu-Sarkodie[5], PhD; Florian Marks[6,7,8,9], PhD; Ellis Owusu-Dabo[1], PhD

[1]School of Public Health, College of Health, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

[2]Department of Statistics and Actuarial Science, College of Science, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

[3]Department of Molecular Medicine, College of Health, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

[4]Community Health, College of Health, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

[5]Department of Clinical Microbiology, College of Health, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana

[6]International Vaccine Institute, Seoul, South Korea, Republic of Korea

[7]Madagascar Institute for Vaccine Research (MIVR), University of Antananarivo, Antananarivo, Madagascar

[8]Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

[9]Heidelberg Institute of Global Health, University of Heidelberg, Heidelberg, Germany

Corresponding Author:
Portia Boakye Okyere, MPhil
School of Public Health
College of Health
Kwame Nkrumah University of Science and Technology
KNUST-SPH, SMS PMB
Kumasi, 233
Ghana
Phone: 233 266075051
Email: portiaboky@gmail.com

Abstract

Background: Typhoid fever, a significant global health problem, demonstrates a multifaceted transmission pattern. Knowledge of the factors driving its transmission is critical for developing effective control strategies and optimizing resource allocation.

Objective: This review aimed to comprehensively synthesize evidence on risk factors associated with typhoid fever transmission from 1928 to 2024.

Methods: We searched PubMed, Scopus, Google Scholar, and Semantic Scholar databases using keywords related to risk, contributors, determinants, and causes of typhoid fever. We followed a registered protocol to support our search and triangulated the results.

Results: Overall, we retrieved 1614 articles, of which 219 were reviewed. Of these, 109 addressed multiple, non–mutually exclusive typhoid fever risk factors. Unsurprisingly, of the total articles reviewed on risk factors, approximately 70.6% (77/109) originated from the Asian continent (51/109, 46.8%) and the African continent (26/109, 23.9%). Half of the articles (55/109, 50.5%) focused on risk factors related to demographic and socioeconomic transmission, while 44% (48/109) of the articles examined foodborne transmission. Additional risk factors included water, sanitation, and hygiene practices: waterborne transmissions (45/109, 41.3%) and sanitation and hygiene practices (34/109, 31.2%), travel-related risk (19/109, 17.4%), antimicrobial use (14/109, 12.8%), climate-related factors (15/109, 13.8%), environment-related factors (9/109, 8.3%), typhoid carriers (11/109, 10.1%), and host-related risk factors (6/109, 5.5%).

Conclusions: This review identifies demographic and socioeconomic factors as key drivers of typhoid transmission, underscoring the need for targeted interventions. Strengthening street food regulation in urban Asia and investing in water infrastructure in rural Africa can significantly mitigate risk. Integrating water, sanitation, and hygiene interventions with typhoid vaccines can reduce immediate exposure while enhancing long-term immunity. Prioritizing these strategies in schools and high-risk communities is essential for sustainable typhoid control. Future research should focus on longitudinal studies to assess risk factor causality and vaccine impact, guiding more effective public health interventions.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e67544) doi:10.2196/67544

KEYWORDS

typhoid fever; risk factors; foodborne; waterborne; systematic review

Introduction

Background

Typhoid fever is a potentially fatal febrile systemic disease caused by *Salmonella enterica* serotype Typhi (*Salmonella* Typhi or *S* Typhi), a rod-shaped gram-negative bacterium belonging to the Enterobacteriaceae family. *S* Typhi exists exclusively in humans and causes illnesses (typhoid fever) that resemble many other febrile diseases [1]. In this study, the terms "typhoid fever" and "typhoid" are used interchangeably. A description of the infection was reviewed by Cunha [2], clearly separating it from other febrile illnesses and associating its clinical manifestation with significant pathological abnormalities in the spleen, mesenteric lymph nodes, and intestines. Nonetheless, the mainstay of diagnosis is a microbial culture, usually with blood or bone marrow samples. Although bone marrow culture is highly sensitive, it is both invasive and technically unfeasible in most settings. As a result, the disease is usually diagnosed with blood culture, despite its limited sensitivity of approximately 40% to 80%, partly due to antibiotic exposures before the patient visits the health facility [3].

The disease transmission is by the fecal-oral route and can take 2 main forms: direct transmission, where food and water in the immediate environment are contaminated through poor hygiene and sanitation practices, either by transient or chronic carriers; and indirect transmission, where the broader environment becomes contaminated when sewage pollutes water supplies, raw human feces or untreated sewage is used as fertilizer for crops, or piped water is inadequately treated [4].

Typhoid fever is reported to affect people of all ages, although children are more susceptible than other age groups [5]. Before 2000, the global burden of typhoid fever was estimated at 16 million illnesses and 600 thousand deaths annually [6]. In 2000, approximately 21.7 million illnesses and 216,000 deaths occurred [7]. By 2010, annual estimates indicated approximately 26.9 million cases and 200,000 fatalities [8]. However, a more recent estimate from 2017 to 2024 suggests a decline in the annual incidence of typhoid cases [9,10]. Despite this decrease, typhoid fever remains a significant public health concern, particularly in areas with limited access to clean water and sanitation. Typhoid fever can be prevented and controlled concurrently with vaccinations and advancements in food safety, water quality, hygiene, and sanitation [11]. Three main generations of typhoid vaccines are presently approved for use by the World Health Organization (WHO): typhoid conjugate vaccines (TCVs), live attenuated Ty21a, and the unconjugated Vi polysaccharide vaccines [4]. The WHO strongly recommends using TCVs for all ages due to their superior immunological properties, suitability for use in younger children, and predicted longer period of protection above 2 years, which was a major limitation for using the Vi polysaccharide. However, to inform the choice of vaccination in a country, evidence is needed on both the scope of the problem and the risk factors contributing to disease transmission [11]. Despite notable progress in typhoid control, the disease remains a significant cause of morbidity and mortality to which billions of people worldwide are continuously exposed, particularly in Asia and sub-Saharan Africa.

Typhoid fever susceptibility involves multiple factors, each contributing through distinct transmission pathways. In endemic countries, knowledge of typhoid fever risk factors is critical for developing effective control strategies and allocating resources. Several epidemiological and modeled studies [12-14] have explored location- and time-specific risk factors for typhoid fever. In addition, various review studies have been undertaken to comprehensively understand and address the risk of typhoid across different transmission routes. For instance, Lee et al [15] used geospatial modeling to develop a typhoid risk index based on factors such as water sources, toilet facilities, and population density, providing insight into the geographical distribution of typhoid risk in impoverished countries. Similarly, Kim et al [16] investigated the relationship between observed incidence rates and geospatial covariates, such as access to improved water and sanitation, as well as broader health and environmental conditions influencing the transmission of *S* Typhi. Furthermore, Brockett et al [17] systematically reviewed case-control studies to uncover associations between water, sanitation, and hygiene (WASH) practices, food exposures, and typhoid fever. Similarly, Mogasale et al [18] conducted a meta-analysis spanning 1990 to 2013 to estimate the risk of typhoid associated with inadequate access to safe water. In addition, Wang et al [19], in a systematic review and meta-analysis, described the patterns of salmonellosis outbreaks in China from 1970 to 2023. Other reviews have examined specific aspects of typhoid transmission. For instance, the study by Ma et al [20] reviewed human genetic variants affecting susceptibility to enteric fever infection, while Levantesi et al [21] assessed the contribution of natural freshwater and drinking water as routes of *Salmonella* contamination from 2000 to 2010.

Objectives

While these studies provide valuable insights into different aspects of typhoid risk, a holistic synthesis of socioeconomic, environmental, and other factors remains lacking. Furthermore, previous reviews have often been limited in temporal scope or focused on specific transmission pathways. To address this gap, this study examined typhoid risk factors across a broader time frame (1928-2024) and incorporated a multidimensional perspective on transmission dynamics. By systematically integrating evidence from diverse sources, we aimed to provide a more comprehensive understanding of typhoid fever risk factors, which can inform targeted interventions to reduce typhoid incidence worldwide.

XSL•FO
RenderX

Methods

Search Strategy

We searched PubMed, Scopus, Google Scholar, and Semantic Scholar databases for articles published on risk factors for typhoid fever. The search was conducted in June 2023, and titles and abstracts from databases were downloaded and saved. Each database was searched using the following terms and keywords: risk factors, factors, contributors, determinants, causes, predictors, susceptibility factors, factors of exposures, predisposing factor, typhoid fever, typhoid, *Salmonella* typhoid, *Salmonella* Typhi, *S* Typhi, typhoid disease, typhoidal salmonellosis, and typhoidal *Salmonella*, and searches excluded terms related to perforation, complication, virulence, severity, and nontyphoidal infections. We placed no restrictions on the publication year, but the language was restricted to English. We followed a protocol adapted from the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, which was registered with the Open Science Framework in January 2024 [22] to structure our search. See Multimedia Appendix 1 for the PRISMA checklist. This study used published articles, and as such, permission from the institutional review board was not required.

Study Screening and Selection Criteria

We screened and selected studies on typhoid fever risk factors based on the inclusion and exclusion criteria summarized in Textbox 1.

Textbox 1. Inclusion and exclusion criteria for the desk review search strategy on typhoid fever risk factors.

Inclusion criteria

- Studies with confirmed S Typhi infections or outbreaks identified through the following: culture of bodily fluids or stool; polymerase chain reaction; Widal or other serological methods
- Epidemiological or modeled studies of any design

Exclusion criteria

- Studies on severe infections (complications and mortality)
- Studies that classified typhoid fever solely based on clinical indicators (ie, signs and symptoms) or with unclear diagnostic methods
- Studies involving nonhuman participants (animals, water, and farm produce)
- Articles whose full text was not available in English or those inaccessible

Further details on inclusion and exclusion criteria can be found in the review protocol [22].

The titles and abstracts retrieved from each database were imported into EndNote X8.2 (build 13302), merged into a single reference list, and duplicates were eliminated. The deduplicated list was then uploaded to a web-based systematic review tool, Rayyan (Qatar Computing Research Institute) [23] for title and abstract screening. All included citations were exported into Microsoft Excel (version 16.16.27) for full-text retrieval and screening. Each subsequent process, including title and abstract review, full-text review, and data extraction, was performed using predefined screening guidelines outlined in the screener instruction section of the review protocol [22] to ensure consistency. One author (PBO) performed the initial screening with supervision from coauthors (EO-D, SN, and ST-A). Discrepancies in the study selection were resolved through discussion among authors (PBO, EO-D, SN, and ST-A), with unresolved cases adjudicated by EO-D. Additional relevant articles were identified through other sources (expert input) and included after being assessed using the same eligibility criteria. Data were then extracted into Microsoft Excel and a shared Google Sheets spreadsheet (Google LLC; Multimedia Appendix 2). All authors reviewed the final dataset for completeness and accuracy.

Data Extraction

Electronic searches were performed using the internet to locate all eligible articles, and all relevant data relating to the research question were manually extracted into Microsoft Excel after reading the full text. The extracted data included specific risk factors for typhoid fever in all eligible articles. In addition, data on the route of transmission, sources of infection, year of publication, data collection period, town or district, country and continent of the study, study setting (outbreak or endemic), diagnostics method, study type, number of *S* Typhi cases, total participants enrolled, ages of participants, study design, and citations were extracted (Multimedia Appendix 2). We grouped the ages of participants into 3 categories based on inclusion age and age ranges: “children” were ≤15 years, “adults” were >15 years, and “mixed ages” were both children and adult participants. Information on typhoid fever susceptibility was grouped according to their transmission routes: waterborne and foodborne transmissions, host risk factors, vaccination, travel-related risk, health education, occupational risk, population growth and overcrowding, sanitation and sewage systems, climate and meteorological factors, antimicrobial resistance factors were extracted.

Results

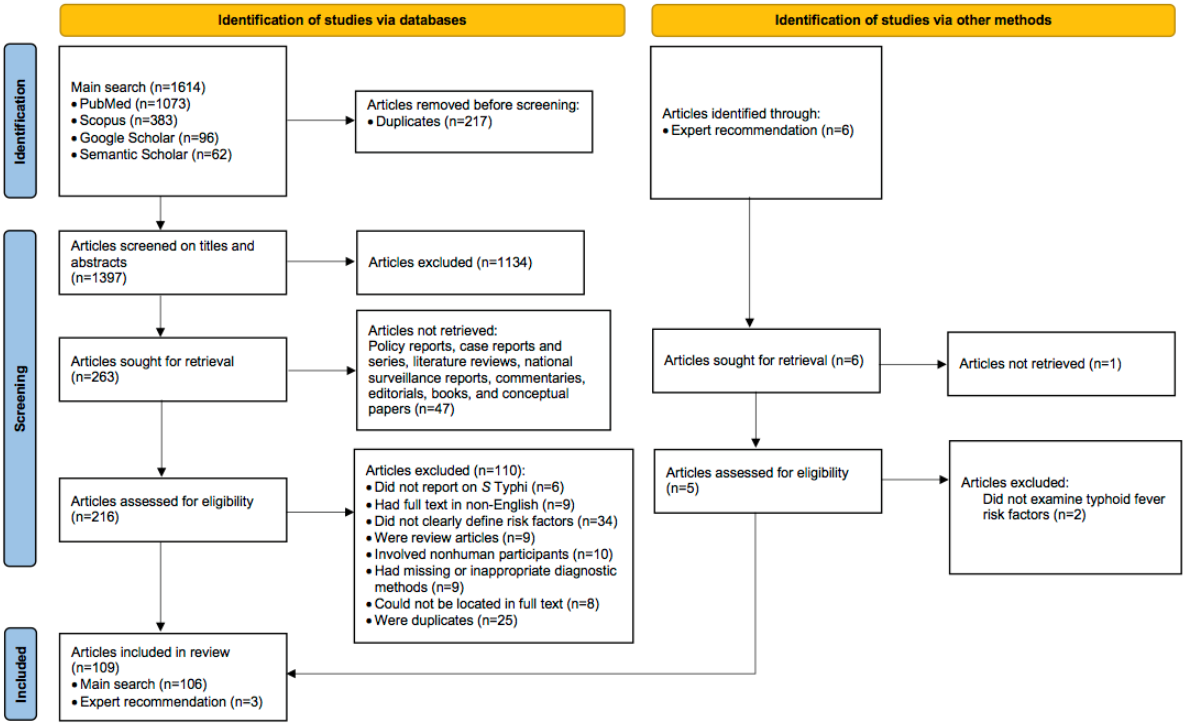
Overview

Our search strategy initially identified 1614 articles published between 1928 and 2023. After removing 217 duplicates, 1397 titles and abstracts remained for screening. Of these, 1181 were excluded, with the majority (687 articles) not examining typhoid fever risk factors. After a full-text review of 219 articles from the main databases (216 articles) and expert recommendations (3 articles). Other reasons for exclusion included 25 duplicates

XSL•FO
RenderX

of the same study published by different authors in different journals; 10 studies involving nonhuman participants such as farm-produced and water samples; 17 articles with non-English full texts or unavailable full texts; 6 articles not specifically related to *S* Typhi; and 9 articles with missing or inappropriate diagnosis based on recall typhoid fever episodes, unclear diagnosis, and clinical indicators (signs and symptoms). Furthermore, to avoid content duplication, 9 review articles were excluded, as shown in Figure 1. Finally, 109 published articles were included in this study [12-14,24-128] (Multimedia Appendix 2).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of search strategies and article selection of risk factors for typhoid fever (2024).



Study Characteristics

Among the 109 eligible articles, data were extracted from publications spanning 1972 to 2024 (Multimedia Appendix 2), covering 6 continents: 24% (26/109) from Africa, 46.8% (51/109) from Asia, 0.9% (1/109) from South America, 8.3% (9/109) from North America, 10.1% (11/109) from Europe, and 6.4% (7/109) from Australia. Furthermore, 2.8% (3/109) of the articles collected data from mixed continents, while one article (1/109, 0.9%) relied on the GeoSentinel Surveillance Network database, lacking specific location details (Table S1 in Multimedia Appendix 3). Regarding transmission routes, multiple non–mutually exclusive typhoid risk factors were identified across the eligible articles (Table 1) for both common and specific risk factors. Waterborne transmission was reported in 41.3% (45/109) of the articles, while 44% (48/109) of the articles focused on foodborne transmission. Demographic and socioeconomic factors were identified in more than 50.5% (55/109) of the studies, with hygiene and sanitation discussed in 31.2% (34/109) of the studies. Additional risk factors included travel-related exposure (19/109, 17.4%), climatic influences (15/109, 13.8%), and antimicrobial use (14/109, 12.8%; Table S2 in Multimedia Appendix 3). Across 103 papers, 253,951 typhoid fever cases were reported, with a median (IQR) of 110 (51-283). The diagnosis was confirmed predominantly through culture in 84.4% (92/109) of the articles, with approximately 9% (10/109) of articles using Widal and other serological tests and 6.4% (7/109) of the articles using polymerase chain reaction and other sequencing techniques. Age distribution data were available in 103 articles. Most of the articles (83/103, 80.6%) included participants of mixed ages, while approximately (14/103, 13.6%) of the articles focused exclusively on those aged ≤15 years and 5.8% (6/103) of the articles focused on those aged >15 years (Table S1 in Multimedia Appendix 3).

XSL•FO
RenderX

Table 1. Common and specific risk factors for typhoid fever.

Common risk factors	Specific risk factors
Waterborne transmission	
Water sources	Household sources of drinking water [24,58], having multiple drinking water sources [68], using water from other sources than the municipal water networks for bathing, brushing teeth or drinking [112], primary water sources with unpleasant smell [37], obtaining water from an outside tap [46], unsafe water source [53,115], obtaining water from municipal pipe for drinking [107], obtaining water from a river or stream [12,115]
Water supply	Water supplied by an outdated gravity-fed network [53], intermittent water availability [12], inadequate safe water supply [14], defective water systems [113], household water supply from public wells and boreholes or merchants [98], water from a community [34], water from government overhead tanks [43]
Contaminated or unsafe water	Drinking unsafe or contaminated water [28,33,63,79,95], using substandard water [83], *Escherichia coli* in stored drinking water [79], water sold in small plastic bags [80], use of ice cubes from a street vendor [29,123]
Untreated water	Drinking water from a well [96], untreated household drinking water [37,46-48,70,120], use of untreated public water after rains [32], drinking water from untreated open sources [12,27,89], drinking water at the work site [34], use of bore water [107], accidental ingestion of contaminated river water during swimming or bathing [28], cooking and cleaning with river water or an open dug well [68]
Water storage	Storing water in plastic containers without a lid [98], not storing water for drinking in a narrow-mouthed container [38], not using tipped containers to draw water [38], water sold in small plastic bags [80]
Foodborne transmission	
Street stall or restaurant food	Eating out from commercial food stalls, restaurants, or mobile food vendors in the street or outside home [26,34,45,91,101,122], eating outdoors at least once a week [29,30,117], eating food from a roadside cabin during the summer [34], mothers eating food from street vendors [70], eating cold beverages outside home [45], not dining at a tea-house [47], consuming French fries with sauce and poppadum from street vendors [107], eating commercially available foods or drinks [120], eating food at community market [97]
Poor food hygiene	Eating unwashed farm produce [12], unwashed guavas [38], not washing vegetables before eating [47]
Food handlers	Consuming food items from vendors [108,122], male food handlers [40], consuming food with the mother's or caregiver's assistance [30], drinking orange juice with hand contact [50], poor or unhygienic food handling practices or procedures [14,100], eating food prepared at home [97]
Salads or other contaminated food	Eating contaminated foods [63], including cucumber salad [94], potato salad [114], lettuce salad [31], and raw salads (onion, cucumber, and tomato) [103]
Uncooked or raw food	Eating raw traditional foods such as cig kofte [31], raw onions and cabbage [38], raw milk and meat [58], uncooked shellfish [118], papaya [37], and unwashed guava [38]
Frozen food	Eating ice cream [30,34,49], ice cubes in beverages [29], frozen tropical fruit (mamey) shakes [84], and fresh ice cream during the hot season [13]
Milk products	Eating butter and yogurt [38], fresh yogurt made from cow and sheep in the summer [13], and homemade cheese [13]
Local or traditional food	consuming locally made beverages [80], eating locally prepared popsicles [103], and consuming locally prepared flavored beverages [103]
Other	Eating food brought by relatives from endemic areas [91], sharing food from the same plate [123], and sharing food plates [53]
Demographic, socioeconomic, hygiene, and sanitation factors	
Demographic background	Age [25,58,87,126], older age [39,86], younger age [33,42,59,82,108,109], sex of the individual [102,126], being female [108,121] or male [87,101], presence of preschool children in the household [103], young adult [121], demographic status [56], and low BMI [128]
Socioeconomic status	Poor or low socioeconomic status [25,56,76,83,104], attending a gathering [96], per capita disposable income of all residents, and per capita gross domestic product [102], unemployment or part-time work [29,128], being part of a nuclear family [49], wealth index [90], being a student [47], and attending school or daycare [68]
Education level	No or low educational level [27,48,120], educational level [58], years of schooling [90], students in conventional institutions of higher learning [102], and illiteracy rate [49,104,105]
Occupational risk	Mishandling of *S* Typhi samples by clinical microbiology laboratory staff [72]; medical and laboratory personnel and sewage workers occupationally exposed to *Salmonella* bacteria [91]; household member growing crops [68]; farmers [87,128]; rearing chicken or goats [92]; job-related cause [63]
Population growth and overcrowding	Rising or bigger household size [104,127], living in a crowded household [31], crowding poor living conditions [63], and increased population density [33,113]
Housing system or condition	Poor housing conditions [79]

XSL•FO
RenderX

Common risk factors	Specific risk factors
Hygiene and behavioral factors	Scarcity of soap near a hand washing facility [127], nonuse of soap for handwashing [12,49,69,108], nonuse of medicated soap [127], nonavailability of soap to wash hands after toilet use [68,104], a habit of not washing hands before cooking or after defecating [53], infrequent hand washing after latrine use [12,97], poor hand washing practices [105], occasionally or never washing hand with water and soap [26,122], never or rarely washing hands before preparing or handling food, and eating or feeding [29,30,120]
Water, sanitation, and hygiene (WASH) practices	Poor WASH practices [74], not living in a better WASH household [75,106]
Sanitation and sewage systems	Use of pit latrine [27,92], open defecation [92], improper disposal of solid waste [27], burst sewer pipes at home [96], living in houses with open sewers [29], visible urine or feces [53], poor sanitary practice [14,63,95,98,113], having home latrines [38], no toilets in the residence [108,115], poor excreta disposal [79], having unimproved or malfunctioning sanitation infrastructure [12], unsterilized water from the hospital disposal and residential sewage used to irrigate vegetable farmlands [110], inadequate public sewerage system [90], and poor toilet drainage soil [79].
Other	
Antimicrobial use	Frequent use of antimicrobials or history of antimicrobial use [34,47,55,62,63,76,81,101], chloramphenicol-resistant *S* Typhi strain [44,79], ceftriaxone-resistant *S* Typhi strain [119], multidrug resistant or extensively drug-resistant *S* Typhi strains [121], circulation of virulent *S* Typhi strain (H58-lineage) [60,103]
Host risk factors	Polymorphism in intronic variable number tandem repeat of IL-4 [35], presence of serum anti–*Helicobacter pylori* immunoglobulin G antibodies [49,109], history of chronic underlying disease [69], HIV infections [71], and haplotype of tumor necrosis factor locus from single nucleotide polymorphisms [124]
Typhoid carriers	Recent or close contact with a patient with confirmed or active typhoid fever [32,48,63,88,101,105], hospitalization of household member with febrile illness [68], history of typhoid fever infections [69], having typhoid carrier at home [65], recent typhoid fever case in the household [108], and having a housekeeper (a boy or girl) [120]
Vaccination	No or lack of vaccination [63,75,106], vaccine hesitancy [120], vaccine ineffectiveness [94], and poor vaccination coverage [103]
Health education	Lack of knowledge regarding typhoid fever contact [127] and poor awareness of typhoid fever disease [97]
Travel-related risk	Longer duration of stay in the endemic area [39]; returning from or visiting endemic countries [76,86,116,126]; visiting friends and relatives in endemic areas [39,73]; travel destination [64]; travel outside the United States, Sweden, or United Kingdom (international travel) [41,59,93,129]; Asian travelers [52]; children visiting friends and relatives in endemic places (particularly South Asia) [61,99]; recent travel to endemic areas [85,125]; transient male workers [82]; living in a metropolitan area [86]; urbanization [90]; and number of foreign tourists received (tourism) [102]
Environmental conditions or factors	Living in geographically lower elevation areas [42], neighbors to a typhoid fever case [54], potentially floodable areas [57], proximity to major rivers and creeks [57], housing (external condition) [79], a lack of agricultural land [115], hydrological catchment areas [119], residing closer to waterbodies, residing near typhoid study treatment centers [104], anthropogenic alteration of land cover and hydrology [78], and environmental factors [56]
Climate or meteorological factors	Seasonal variation or fluctuations [25,82,102,110], high temperatures during summer [103], rainfall [57,115], temperature and precipitation [111,115,126], high vapor pressure [115], rainy and harmattan seasons [36], extreme weather conditions [77], higher or hot temperatures [66,86], flooding [87], and wind speed [90]

Discussion

Principal Findings

This study synthesizes typhoid fever risk factors from 1928 to 2024; however, the included studies span from 1972 to 2024, as few earlier publications met the inclusion criteria. We have identified demographic and socioeconomic factors as the predominant pathways for typhoid fever transmission, with additional contributions from waterborne and foodborne routes, hygiene and sanitation, travel-related exposures, antimicrobial use, and typhoid carriers. These findings highlight the multifaceted nature of typhoid fever transmission and underscore the importance of targeted interventions. This evidence is crucial for clinicians, public health experts, and policy makers in designing effective control strategies and optimizing resource allocation, particularly in endemic regions.

Approximately 70.6% (77/109) of the reviewed articles collected data from Asia and Africa, where most developing countries are located. This aligns with the global burden of typhoid fever estimation, which shows that Asian and African countries bear the greatest burden [130]. The low number of articles in Europe, North America, and Australia can be attributed to the introduction of control programs such as the treatment of municipal water, pasteurization of dairy products, and strict food safety regulations [1,131]. Similarly, the lack of studies from South America can be attributed to the decline in typhoid burden, given the economic transition with improved water and sanitation in the area [1]. Despite the lower disease burden in these underrepresented regions (Europe, North and South America, and Australia), the risk factors identified in this study, such as socioeconomic, foodborne and waterborne transmission, and antimicrobial use, remain relevant to these regions. Emerging threats, including climate change, urbanization, and

XSL•FO
RenderX

increased global travel, may contribute to the reemergence of typhoid in areas where it was previously controlled. In addition, the rise of antibiotic-resistant strains presents an ongoing challenge, underscoring the need for public health preparedness in these regions.

Waterborne Transmission

Water sources greatly impact the spread of typhoid fever, with protected wells and piped water classified as safe, while rivers, streams, and other unprotected sources are deemed unsafe [132]. This study identified several specific risk factors associated with typhoid fever and water sources, including having multiple drinking water sources [68], use of nonmunicipal water sources for various purposes [29,68,107], and main water sources with a foul smell [37]. Households with access to multiple water sources may be at increased risk of typhoid fever, particularly when they face challenges with potable water access such as limited supply hours, high tariffs, low-pressure, and long distances to collection points. These challenges often impede their ability to meet daily needs. Consequently, some households, particularly those located close to open dug wells, rivers, or streams, may use these alternative sources for bathing, cooking, or even drinking [115]. Although such sources offer convenience, they often lack chlorination and may be contaminated with fecal matter, thereby posing a risk for typhoid fever transmission.

A water supply system defect can facilitate the transmission of typhoid fever. Our study discovered a typhoid fever outbreak associated with a gravity-fed network [53], indicating a probable spread through an outdated mains system. This was attributed to low water pressure, insufficient chlorination, and fecal infiltration [12].

Furthermore, we uncovered an intermittent piped water supply as a risk factor for typhoid transmission. Intermittent piped water supply is a common challenge in many developing countries, potentially compromising water supply quality by allowing contaminants in nonpressurized pipes and creating negative pressure conditions that enable pathogens such as *S* Typhi to infiltrate the system [12,133]. Such interruptions often compel households to rely on water storage, which may introduce additional contamination risks, and to seek alternative sources that may not be microbiologically safe [12,29]. Sources such as government overhead tanks [43], community or public taps, and protected wells [34,115], which are often considered safe, can become contaminated due to environmental exposures, including industrial activities, sewage discharges, agricultural runoff, and animal waste [28,80]. Under these conditions, the presence of *Escherichia coli* (*E coli*) in drinking water is commonly regarded as an indicator of fecal contamination and, by extension, a potential risk factor for typhoid fever due to the possible presence of enteric pathogens [79]. However, the relationship between *E coli* and typhoid fever risk remains inconclusive. For instance, Karkey et al [134] observed a link between high *E coli* concentrations and the presence of *S* Typhi nucleic acids in drinking water, suggesting that *E coli* contamination may serve as a proxy for typhoid transmission risk. In contrast, Luby et al [34] found no significant difference in the levels of *E coli* in water samples between households with typhoid cases and control households, challenging the consistency of this connection. These discrepancies may stem from environmental conditions, regional differences in water treatment practices, and microbial competition. For instance, in settings with inadequate chlorination, *E coli* presence may serve as a proxy for recent fecal contamination, thereby increasing typhoid risk. Conversely, in areas with intermittent water supply and biofilm formation within pipes, *S* Typhi may persist independently of *E coli*, complicating its use as a universal indicator.

In addition, water storage practices emerged as a risk factor for typhoid fever in our study. Evidence suggests that the microbiological quality of water often deteriorates after collection, particularly during storage and handling [98]. The risk of contamination is influenced by the type of vessel used for storing or drawing drinking water. Wide-mouthed containers, in particular, are associated with a higher risk of infection due to their larger open surface area, which increases exposure to environmental contaminants and the likelihood of fecal contamination, compared to narrow-tipped or covered containers [38,98].

Moreover, the use of untreated water, both inside and outside the home, poses significant risks for typhoid fever transmission [27,32,48,89]. Although municipal water systems may incorporate filtration and chlorination, cross-contamination with *S* Typhi through wastewater intrusion remains a concern [46]. In such contexts, household-level water disinfection may be necessary to enhance water quality [37]. However, some households opt out of water treatment due to the perceived reliability of their water sources [98]. While water from piped or otherwise protected sources may contain insufficient bacteria to cause typhoid fever, untreated water from unprotected sources may carry high *S* Typhi levels, sufficient to cause clinical disease [28,34,47]. All the aforementioned risk factors are listed in Textbox 1.

Foodborne Transmission

Food serves as a highly efficient medium for the growth of *S* Typhi compared to water [34]. This study revealed specific risk factors indicating potential foodborne transmission of typhoid fever, with street stalls (restaurants) emerging as a major factor in this category. Street food stalls are typically small, with outdoor seating and without refrigerators or easy access to potable water or adequate facilities for washing food and utensils [29,34,45]. Therefore, persons who frequently consume food from such establishments are at increased risk of developing typhoid fever [29]. This increased risk may stem from using untreated or tap water stored or served in contaminated containers for food preparation and drinking purposes [45]. In addition, poor hygiene practices among street food vendors, including irregular handwashing [122], and the potential exposure to carriers of *S* Typhi further contribute to the risk of infection [135]. In addition, many of these food preparers and handlers in the street eateries lack adequate knowledge of safe food handling practices necessary to avert *S* Typhi infection transmission. Often, they operate without licenses or registration from food safety authorities; as a result, they are neither trained nor subject to regular inspections. Consequently, they may

XSL•FO
RenderX

unknowingly share food and drinks using poorly cleaned cups and utensils among multiple clients [97,120]. Furthermore, this study discovered a relationship between the consumption of frozen foods, including ice cream [13,30,49], fruit shakes [84], and iced beverages [29], and typhoid fever. A potential source of contamination lies in the ice used by street vendors, who often purchase large blocks of ice produced from untreated water, typically intended for industrial use, such as fisheries, rather than for human consumption. Despite this, the ice is commonly served in drinks for customers. Notably, research has shown that *S* Typhi can survive in ice for extended periods, underscoring its potential role as a vehicle for transmission [29,45,122]. Moreover, iced drinks may be further contaminated by street vendors who are asymptomatic carriers of *S* Typhi during the distribution chain [29]. In contrast, dining at tea houses decreases the risk of typhoid fever, as customers are typically served boiled water and tea, practices that limit exposure to *S* Typhi through thermal inactivation of the pathogen [47]. In addition, consumption of uncooked or raw foods such as onions, milk, meat, shellfish, papaya, cabbage, and other traditional raw ingredients poses a considerable risk, as these items may harbor *S* Typhi if not properly handled or sourced [31,37,38,58,118]. For instance, failure to wash fruits and vegetables before consumption increases the risk of infection due to surface contamination [12,38,47]. Cross-contamination during meat handling is another concern; for example, using the same knives and cutting boards for both infected and uninfected meat in butcheries and restaurants. Similarly, if contaminated water is used to wash carcasses or clean food-contact surfaces, *S* Typhi may be introduced during processing [58]. Moreover, inadequate hygiene practices during milk processing may contribute to the contamination of dairy products, including butter, yogurt, and homemade cheese [13,38,58]. Although dairy animals do not harbor *S* Typhi [136], improperly handled dairy products can serve as effective growth media for the pathogen [58].

Demographic and Socioeconomic Factors

This study identified mixed findings regarding sex-based susceptibility to typhoid fever. While a study suggests that men face greater exposure due to occupational differences, mobility patterns, dietary factors, or a lack of immunity [82], other studies indicate that women, influenced by their physiological status, hormonal imbalance, and gender-specific activities, are more susceptible [120,121]. Building on this, we propose the hypothesis that occupational exposure may explain the higher risk in men, whereas caregiving roles could contribute to increased susceptibility among women. Nonetheless, a study by Rasul et al [137] concludes that typhoid fever incidence is independent of gender, affecting men and women equally across all age groups. Age serves as a significant factor in typhoid transmission, with both young children and older adults identified as vulnerable groups [121,138]. Young children, characterized by their underdeveloped immune systems [33,42,87] and a limited understanding of disease transmission [58,87], face heightened risks of infection. While younger adults are predisposed to infections due to their adventurous lifestyle or unsanitary activities, such as eating junk food, and an increased number of social gatherings [25,42,58], older adults are more likely to be *S* Typhi resistant due to continual immune boosting [138]. Conversely, older adults may experience susceptibility due to a waning immune system or increased exposure to occupational and environmental risks. These exposures may include involvement in farming-related water contact activities [58,68], rearing chicken or goats [68], handling *S* Typhi specimens in clinical settings [72,91], and working in sewage management [91]. Although *S* Typhi is a human-adapted pathogen and not naturally harbored by animals, the association with poultry and livestock rearing may act as a confounder, reflecting underlying poor sanitary and hygienic conditions within the household rather than direct transmission from animals [92].

Our investigation further revealed that socioeconomic status correlates with an increased likelihood of *S* Typhi infection [96]. While studies suggest that typhoid fever is more common in low-income countries and is connected to poor public health and low socioeconomic indicators [25,83], one study [90] identified a protective effect of the wealth index. This suggests that residing in affluent districts within low- or middle-income countries may significantly mitigate the risk. This may be attributed to the per capita gross domestic product and individual disposable income within a community [102]. Wealthier households are more likely to afford preventive measures such as clean drinking water, improved sanitation, and timely medical consultations, thereby reducing the risk of infection. Conversely, individuals in lower socioeconomic brackets often face health care barriers, leading to underdiagnosis and delayed treatment. The protective effect of a higher wealth index underscores the need for targeted interventions. Public health strategies should prioritize health care accessibility and sanitation improvements in economically disadvantaged areas. Investments in water and sanitation infrastructure, subsidized vaccination programs, and awareness campaigns tailored to low-income communities could help reduce the disproportionate burden of typhoid fever. In addition, strengthening diagnostic capacity in resource-limited settings can improve case detection, ensure timely treatment, and curb typhoid transmission while enhancing health outcomes.

In the context of education, our research highlights the lack of certificate education as a significant risk factor influencing the perception of typhoid fever. Individuals who have never attended school tend to possess limited knowledge about the disease and its modes of transmission, highlighting a strong link between lower educational attainment and reduced awareness [102,120]. Previous studies [90,104,105] emphasize that formal education increases understanding, with uneducated individuals more likely to contract typhoid fever. Interestingly, being a student [47] or attending school (daycare) [68] may also pose risks, likely due to certain exposures in educational settings. While essential health knowledge, such as WHO-recommended practices of handwashing with soap, can be acquired outside formal education, a general lack of awareness significantly increases the risk of infection. Individuals who do not recognize the risks are less likely to take preventive actions [97,127]. Another significant factor for typhoid transmission, besides knowledge, is awareness of the presence of a patient with typhoid at home [120]. Individuals who are unsure of the presence of a patient with chronic or current typhoid at home

XSL•FO
RenderX

are more likely to have typhoid or a recurrence than those with full awareness. This may be attributed to the continued shedding of *S* Typhi in the stool and urine of infected individuals, even after initial antimicrobial treatment. Up to 10% of patients may continue shedding the bacteria for as long as 3 months, with some proceeding to become long-term asymptomatic carriers [1]. These transient or chronic carriers can be sources of infection within households [139,140]. In contrast, household members who are well informed of such cases are more likely to acquire knowledge about the disease, its transmission routes, and effective preventive measures. Recent contact with patients with typhoid has also been observed as a potential risk factor, further emphasizing the role of household-level awareness in reducing transmission [48,88]. In many communities, traditional practices of visiting the sick can increase interpersonal contact and inadvertently raise the risk of exposure. Close contacts are often residents of the same area and may share communal water sources, suggesting that transmission could still occur via water contamination (broader environment) rather than direct person-to-person spread. Therefore, health education initiatives should consider addressing the risks associated with visiting infected individuals, alongside broader messaging on water hygiene and disease prevention.

Hygiene and Sanitation Risk Factors

This study further identified risk factors that underscore the significant impact of hygiene and sanitation on the spread of typhoid fever. Poor handwashing practice is a critical risk factor, given the crucial role hands play in transmitting *S* Typhi through the fecal-oral cycle [26,30,120,123]. While handwashing with soap and clean water effectively removes pathogens [29], inadequacies in technique, such as rinsing without soap [53,122] or neglecting handwashing after defecation [12,97], can increase the risk of bacterial spread [106]. Furthermore, using medicated soaps is an added advantage because it is more effective in eliminating bacteria from hands compared to regular soaps [127].

In addition, the condition of the sewerage system in the house has an important impact on typhoid fever incidence. According to Prasad et al [12], people lacking access to improved sanitation facilities or with damaged improved sewerage systems are particularly vulnerable to infections. In many cases, household toilets are built without professional expertise, often on permeable soil, and in flood-prone areas, increasing the likelihood of leakage and pollution of surface water and crops with human waste [79]. Strengthening the construction and maintenance of sanitary excreta disposal facilities, alongside effective solid waste management, is essential for preventing typhoid fever transmission. Studies have shown that poor sanitation, including improper disposal of solid waste and excreta in residential settings, is directly correlated with higher typhoid prevalence [27,79]. Inadequate waste disposal infrastructure, such as pit latrines, open defecation sites, burst sewer pipes, and the presence of visible feces or urine, has been consistently identified as a significant risk factor [29,53,92,96]. Notably, preventing human excreta from entering the domestic arena has a greater impact on interrupting typhoid transmission than behaviors preventing pathogens in the environment from being ingested by humans (eg, hand washing).

Finally, the discharge of unsterilized water from hospitals and residential areas into rainwater canal systems, often used for irrigating farmlands, contributes to the contamination of vegetables cultivated in these areas [110]. These contaminated crops are frequently consumed without thorough washing, thereby increasing the risk of typhoid fever. Particularly, the risk is pronounced during the rainy season, when heavily polluted irrigation water is more commonly used, and runoff from farms mixed with garbage is more likely to spread into residential zones, further endangering public health.

Other Risk Factors for Typhoid Fever

Other typhoid risk factors identified in this study include antimicrobial exposure, host-related factors, vaccination status, travel history, and environmental or climate conditions. Among these, antimicrobial use has the greatest impact on *S* Typhi infection. Several studies have shown that prior or recent use of antibiotics, particularly within 4 weeks before disease onset, is associated with an increased risk of typhoid fever, particularly in cases involving multidrug resistant or extensively drug-resistant strains [63,81,101]. Antimicrobial exposure can induce prolonged alterations in gut flora and compromise the barrier against bacterial colonization, thereby reducing the threshold of *S* Typhi required for infection [47]. Studies by Yousafzai et al [119], Srinivasan et al [103], and Kamal et al [81], further highlight antimicrobial resistance as a major contributor to typhoid fever, with certain resistant strains capable of causing epidemics. This resistance is largely caused by the routine presence of *S* Typhi in the human intestine and the indiscriminate use of antibiotics [44,63]. Consequently, drug-resistant *S* Typhi strains, often carrying multiple virulence factors, are becoming increasingly prevalent worldwide. Notably, this study also identified specific risk factors associated with *S* Typhi–resistant strains harboring virulence genes, including those within the H58 lineage, which enhance their ability to infect and interact with host cells [60,62,103,121].

Host-Related Factors

Host genetic factors influence susceptibility to infectious diseases in humans. This study referenced research by Manal et al [35], which explored the relationship between genetic polymorphisms and typhoid fever risk. Their findings suggested that individuals carrying the 2R3R heterozygote of the intronic variable number tandem repeat in the *IL4* gene may have a genetic predisposition to typhoid fever. However, a study by Dunstan et al [124] reported that a specific haplotype within the tumor necrosis factor gene locus offers protection from typhoid. These associations may be explained by the influence of genetic variation on immune response pathways. Another study [49] found a link between serum anti *H pylori* immunoglobulin G (IgG) levels and an increased risk of typhoid fever. Serum IgG antibodies indicate either prior or active *H pylori* infection, as these antibodies can persist even after infection clearance [49,141,142]. A possible explanation for this association lies in the role of the gastric acid barrier as a crucial defense mechanism against ingested pathogens such as *S* Typhi. *H pylori* infection has been associated with hypochlorhydria, a condition characterized by reduced stomach acid production, which weakens this protective barrier [49,109].

XSL•FO
RenderX

This impairment may facilitate the survival and subsequent colonization of *S* Typhi in the gastrointestinal tract, thereby increasing susceptibility to typhoid fever.

In addition, this review identified a significant association between typhoid fever and the presence of chronic underlying conditions [69]. A plausible explanation is that chronic illnesses can weaken the immune system, impairing the body's ability to clear *S* Typhi and increasing the risk of persistent or severe infection. For instance, although *S* Typhi is not widely associated with AIDS in developed countries, studies from endemic areas suggest a different pattern. We uncovered a study by Gotuzzo et al [71] that reported an increased risk of typhoid in patients infected with HIV from typhoid-endemic areas. In addition, the study noted that a large proportion of HIV-positive participants were men who have sex with men, raising the possibility that direct fecal-oral transmission may contribute to increased incidence in this subgroup. However, further research is needed to clarify the specific transmission dynamics within this population and to distinguish the role of immunosuppression from that of behavioral factors.

Travel-Related Risk

Typhoid fever, once prevalent in industrialized countries, is now effectively controlled [91,93]; however, imported infections remain a significant public health concern [39,52,85]. The risk of infection among travelers varies depending on factors such as age, destination, duration, and purpose of travel [99]. Travelers visiting friends or relatives are in a high-risk category for typhoid fever [39,61,73]. As they are much less likely than other travelers to seek pretravel counseling, they may visit more rural, remote areas and engage closely with local people as well as eat high-risk foods and beverages [51]. Children and young adolescents who are visiting friends or relatives are also at high risk of contracting typhoid fever due to a lack of immunity or the possibility of traveling under unhygienic conditions [61,82,99]. We also discovered that traveling to endemic locations increases the risk of contracting typhoid [52,59,86,93]. According to Lin et al [86], more than half of all travelers with typhoid returning to developed countries have visited Asia or Africa, where the disease is widespread. This trend likely reflects increased exposure associated with travel to these endemic areas, particularly during extended stays. In contrast, short-term visitors to endemic areas face a comparatively lower risk of infection [39]. Furthermore, increasing global mobility driven by economic globalization has facilitated the movement of travelers for business, tourism, or labor migration, thereby contributing to disease spread. This growing influx of individuals, often without adequate vaccination or awareness of preventive measures, contributes to the continued transmission and global spread of typhoid fever [82,102].

Vaccination

Vaccination is essential for the control of typhoid fever in endemic and epidemic settings as well as among travelers moving between nonendemic and endemic areas. The WHO recommends the programmatic use of typhoid fever vaccines in endemic areas [4]. We retrieved 2 studies [75,106] that demonstrated a reduced risk of typhoid fever among individuals who received effective typhoid vaccines and resided in households with improved water. Conversely, poor vaccination coverage, particularly when combined with inadequate WASH infrastructure, can exacerbate typhoid transmission in a given area [63,103]. Moreover, the effectiveness of vaccination may also be compromised by factors, including defective vaccine batches, incorrect immunization procedures, or the interval since vaccination. Evidence from this study suggests that individuals vaccinated more than 3 years prior may face a heightened risk of infection, likely due to waning immunity [94]. This is particularly relevant for polysaccharide vaccines, which have an estimated cumulative efficacy of approximately 55% over 3 years, with the strongest protection occurring within the first 2 years after immunization.

Environmental, Seasonal, and Climate Factors

Typhoid fever transmission exhibits distinct seasonal patterns influenced by environmental and climatic factors such as temperature, humidity, and precipitation. This study identified diverse peak periods across different endemic regions. For example, Taiwan experiences a surge in cases during the fall (September-November) and winter seasons (December-February) [86], whereas in India, the peak occurs in June in Allahabad [25], and during the monsoon season (July-November) in Ahmedabad [77]. A study by Corner et al [56] discovered that approximately half of the yearly typhoid cases in the Dhaka Metropolitan Area, reaching up to 11 per 100,000 individuals, occurred during summer and fall (July-October). Similarly, Srinivasan et al [103] found a positive correlation between summer temperatures (June-August) and increased typhoid cases. These seasonal fluctuations may stem from a complex interplay of climatic conditions, hygiene practices, and local cultural dynamics [86,143]. In warmer climates or during summer, elevated temperatures enhance the proliferation of *S* Typhi in contaminated foods [103]. Conversely, in colder and more humid conditions, *S* Typhi survive longer in water and soil, thereby increasing the likelihood of environmental transmission [144]. In addition, heavy rainfall during the rainy season can trigger flooding and sewage overflows, leading to contamination of water sources and an increase in typhoid incidences [25,66]. This risk is particularly high in low-lying or flood-prone areas where surface water is commonly used for drinking, cooking, and cleaning [42,56,57]. Settlements in hydrologically vulnerable terrains, such as river floodplains, face particularly heightened risks during these periods due to increased sewage runoff and widespread contamination of water and food supplies [67,78].

Study Limitations

This study has some limitations. First, publication bias may be present, as studies with significant findings are more likely to be published, potentially underrepresenting negative or null results and skewing risk factor assessments. Second, the lack of formal risk-of-bias assessments may also affect the reliability of findings. In addition, excluding gray literature and non-English studies may have limited the scope, as multilingual and unpublished data could provide further insights. Furthermore, although the review highlights sex-based differences in susceptibility to typhoid fever, the underlying mechanisms remain unclear due to mixed findings, making the

XSL•FO
RenderX

proposed explanation, such as occupational exposure in men versus caregiving roles in women, speculative.

Conclusions

This review combines current knowledge of typhoid fever risk factors and identifies critical areas for targeted intervention. While food and water have been traditionally recognized as the primary transmission pathways for typhoid fever, this review identifies demographic and socioeconomic factors as predominant drivers for transmission. This shift in understanding underscores the need to move beyond conventional mitigation strategies and adopt targeted interventions that address demographic and socioeconomic disparities, which may play a more significant role in typhoid transmission than previously acknowledged. Considering all identified risk factors, mitigation strategies should be prioritized based on regional transmission dynamics and resource availability. In urban Asia, street food regulation should take precedence, while in rural Africa, investment in water infrastructure is equally critical. A combined approach integrating WASH and vaccination programs, particularly in schools and high-risk communities, remains a key priority for long-term disease control in endemic regions. In addition, addressing environmental and climate-related risks, such as flooding and poor waste management, should be incorporated into prevention strategies.

While this review provides a broad synthesis, key knowledge gaps remain. The absence of longitudinal studies limits the understanding of causality and the temporal dynamics of typhoid risk factors. Methodologically, while this review allows a flexible and wide-range synthesis, it lacks the rigor of systematic reviews and meta-analyses, particularly in terms of risk-of-bias assessments and standardized inclusion criteria.

Future studies should focus on longitudinal and case-control methodologies to establish causality in typhoid risk factors, particularly regarding sex, occupation, and disease susceptibility. In addition, structured bias evaluations and quantitative meta-analyses should be incorporated where possible to improve the reliability and comparability of findings. Future research should also investigate the impact of demographic, socioeconomic, and climate variability on typhoid transmission dynamics, given their increasing relevance. Moreover, TCVs are an important tool in reducing *S* Typhi transmission and should be prioritized for introduction in endemic areas to strengthen prevention and control efforts.

By addressing these gaps and priority key areas, future research can strengthen the empirical foundations for typhoid control strategies, enabling policy makers and public health professionals to develop more targeted, evidence-based interventions for effective typhoid fever prevention and control.

Acknowledgments

The authors would like to thank the EOD project team at Kwame Nkrumah University of Science, in collaboration with the International Vaccine Institute, for their support. The authors received no specific funding for this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[DOCX File , 270 KB - publichealth_v11i1e67544_app1.docx]

Multimedia Appendix 2
Data extraction entities for the systematic review of typhoid fever risk factors.
[XLSX File (Microsoft Excel File), 108 KB - publichealth_v11i1e67544_app2.xlsx]

Multimedia Appendix 3
Summary of study characteristics and query strings for article search.
[DOCX File , 25 KB - publichealth_v11i1e67544_app3.docx]

References

1. Crump JA, Sjölund-Karlsson M, Gordon MA, Parry CM. Epidemiology, clinical presentation, laboratory diagnosis, antimicrobial resistance, and antimicrobial management of invasive salmonella infections. Clin Microbiol Rev 2015 Oct;28(4):901-937 [FREE Full text] [doi: 10.1128/CMR.00002-15] [Medline: 26180063]
2. Cunha BA. Osler on typhoid fever: differentiating typhoid from typhus and malaria. Infect Dis Clin North Am 2004 Mar;18(1):111-125. [doi: 10.1016/S0891-5520(03)00094-1] [Medline: 15081508]
3. Parry CM, Wijedoru L, Arjyal A, Baker S. The utility of diagnostic tests for enteric fever in endemic locations. Expert Rev Anti Infect Ther 2011 Jun 10;9(6):711-725. [doi: 10.1586/eri.11.47] [Medline: 21692675]
4. Typhoid vaccines: WHO position paper - March 2018. World Health Organization. 2018 Mar 30. URL: https://www.who.int/publications/i/item/typhoid-vaccines-who-position-paper-march-2018 [accessed 2024-07-28]

XSL•FO
RenderX

5. Marks F, Im J, Park SE, Pak GD, Jeon HJ, Wandji Nana LR, et al. Incidence of typhoid fever in Burkina Faso, Democratic Republic of the Congo, Ethiopia, Ghana, Madagascar, and Nigeria (the Severe Typhoid in Africa programme): a population-based study. Lancet Glob Health 2024 Apr;12(4):e599-e610. [doi: 10.1016/S2214-109x(24)00007-x]
6. Stuart BM, Pullen RL. Typhoid; clinical analysis of 360 cases. Arch Intern Med (Chic) 1946 Dec 01;78(6):629-661. [doi: 10.1001/archinte.1946.00220060002001] [Medline: 20278487]
7. Crump JA, Luby SP, Mintz ED. The global burden of typhoid fever. Bull World Health Organ 2004 May;82(5):346-353 [FREE Full text] [Medline: 15298225]
8. Buckle GC, Walker CL, Black RE. Typhoid fever and paratyphoid fever: systematic review to estimate global morbidity and mortality for 2010. J Glob Health 2012 Jun;2(1):1-9 [FREE Full text] [doi: 10.7189/jogh.01.010401]
9. GBD 2017 Typhoid and Paratyphoid Collaborators. The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Infect Dis 2019 Apr;19(4):369-381 [FREE Full text] [doi: 10.1016/S1473-3099(18)30685-6] [Medline: 30792131]
10. Wang H, Zhang P, Zhao Q, Ma W. Global burden, trends and inequalities for typhoid and paratyphoid fever among children younger than 15 years over the past 30 years. J Travel Med 2024 Dec 10;31(8):e140. [doi: 10.1093/jtm/taae140] [Medline: 39450993]
11. World Health Organization. Typhoid vaccines: WHO position paper, March 2018 - recommendations. Vaccine 2019 Jan 07;37(2):214-216. [doi: 10.1016/j.vaccine.2018.04.022] [Medline: 29661581]
12. Prasad N, Jenkins AP, Naucukidi L, Rosa V, Sahu-Khan A, Kama M, et al. Epidemiology and risk factors for typhoid fever in Central Division, Fiji, 2014-2017: a case-control study. PLoS Negl Trop Dis 2018 Jun 8;12(6):e0006571 [FREE Full text] [doi: 10.1371/journal.pntd.0006571] [Medline: 29883448]
13. Zubair OA, Mohammad MY. Prevalent sources of enteric fevers among women in Mosul City. Ann Trop Med Public Health 2020;23(13B) [FREE Full text] [doi: 10.36295/ASRO.2020.231355]
14. Meftahuddin T. Review of the trends and causes of food borne outbreaks in Malaysia from 1988 to 1997. Med J Malaysia 2002 Mar;57(1):70-79 [FREE Full text] [Medline: 14569721]
15. Lee JS, Mogasale VV, Mogasale V, Lee K. Geographical distribution of typhoid risk factors in low and middle income countries. BMC Infect Dis 2016 Dec 05;16(1):732 [FREE Full text] [doi: 10.1186/s12879-016-2074-1] [Medline: 27919235]
16. Kim JH, Choi J, Kim C, Pak GD, Parajulee P, Haselbeck A, et al. Mapping the incidence rate of typhoid fever in sub-Saharan Africa. PLoS Negl Trop Dis 2024 Feb 26;18(2):e0011902. [doi: 10.1371/journal.pntd.0011902] [Medline: 38408128]
17. Brockett S, Wolfe MK, Hamot A, Appiah GD, Mintz ED, Lantagne D. Associations among water, sanitation, and hygiene, and food exposures and typhoid fever in case-control studies: a systematic review and meta-analysis. Am J Trop Med Hyg 2020 Sep;103(3):1020-1031 [FREE Full text] [doi: 10.4269/ajtmh.19-0479] [Medline: 32700668]
18. Mogasale VV, Ramani E, Mogasale V, Park JY, Wierzba TF. Estimating typhoid fever risk associated with lack of access to safe water: a systematic literature review. J Environ Public Health 2018;2018:9589208 [FREE Full text] [doi: 10.1155/2018/9589208] [Medline: 30174699]
19. Wang Z, Zhou H, Liu Y, Huang C, Chen J, Siddique A, et al. Nationwide trends and features of human salmonellosis outbreaks in China. Emerg Microbes Infect 2024 Dec 18;13(1):2372364 [FREE Full text] [doi: 10.1080/22221751.2024.2372364] [Medline: 38923510]
20. Ma PY, Tan JE, Hee EW, Yong DW, Heng YS, Low WX, et al. Human genetic variation influences enteric fever progression. Cells 2021 Feb 06;10(2):345 [FREE Full text] [doi: 10.3390/cells10020345] [Medline: 33562108]
21. Levantesi C, Bonadonna L, Briancesco R, Grohmann E, Toze S, Tandoi V. Salmonella in surface and drinking water: occurrence and water-mediated transmission. Food Res Int 2012 Mar 27;45(2):587-602 [FREE Full text] [doi: 10.1016/j.foodres.2011.06.037]
22. Okyere PB, Twumasi-Ankrah S, Newton S, Owusu-Dabo E. Protocol for a desk review on risk factors for typhoid fever. Open Science Framework Registries. 2024. URL: https://osf.io/87jv9 [accessed 2024-09-07]
23. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
24. Akwa TE, Nguimbous SP. Investigation of typhoid fever and their associated risk factors in children attending “Deo Gratias” hospital in Douala, Littoral, Cameroon. Eur J Med Educ Technol 2021;14(2):em2107. [doi: 10.30935/ejmets/10910]
25. Alajeely AA, Lawrence R, Jeyakumar E, Maurice NG. Risk factors of typhoid fever amongst patients in the Allahabad region, India. Int J Curr Res 2014;6(11) [FREE Full text]
26. Alba S, Bakker MI, Hatta M, Scheelbeek PF, Dwiyanti R, Usman R, et al. Risk factors of typhoid infection in the Indonesian archipelago. PLoS One 2016 Jun 9;11(6):e0155286 [FREE Full text] [doi: 10.1371/journal.pone.0155286] [Medline: 27281311]
27. Eba K, Bekele D. Prevalence of typhoid fever and its risk factors in Lalo Assabi District, West Wollega, Oromiya, Ethiopia. J Bacteriol Parasitol 2019;10:365 [FREE Full text]
28. Anita S, Amir KM, Fadzilah K, Ahamad J, Noorhaida U, Marina K, et al. Risk factors for typhoid outbreak in Sungai Congkak Recreational Park, Selangor 2009. Med J Malaysia 2012 Feb;67(1):12-16 [FREE Full text] [Medline: 22582542]

XSL•FO
RenderX

29. Gasem MH, Dolmans WM, Keuter MM, Djokomoeljanto RR. Poor food hygiene and housing as risk factors for typhoid fever in Semarang, Indonesia. Trop Med Int Health 2001 Jun 21;6(6):484-490 [FREE Full text] [doi: 10.1046/j.1365-3156.2001.00734.x] [Medline: 11422963]
30. Halder S, Rahman ME, Sarker MM, Mone FH, Roy K, Tajkia G. Identification of risk factors for typhoid fever in children admitted in a tertiary care hospital. J Health Med Sci 2021 Jun;4(2):158-162. [doi: 10.31014/aior.1994.04.02.168]
31. Hosoglu S, Celen MK, Geyik MF, Akalin S, Ayaz C, Acemoglu H, et al. Risk factors for typhoid fever among adult patients in Diyarbakir, Turkey. Epidemiol Infect 2006 Jun;134(3):612-616. [doi: 10.1017/S0950268805005583] [Medline: 16288686]
32. Javed N, Abbasi S, Tahir M, Bashir F. Identifying the risk factors for typhoid fever among the residents of rural Islamabad. Pak J Med Res 2017;56(2) [FREE Full text]
33. Khan MI, Ochiai RL, Soofi SB, Von-seidlein L, Khan MJ, Sahito SM, et al. Risk factors associated with typhoid fever in children aged 2–16 years in Karachi, Pakistan. Epidemiol Infect 2011 May 31;140(4):665-672 [FREE Full text] [doi: 10.1017/s0950268811000938]
34. Luby SP, Faizan MK, Fisher-Hoch SP, Syed A, Mintz ED, Bhutta ZA, et al. Risk factors for typhoid fever in an endemic setting, Karachi, Pakistan. Epidemiol Infect 1998 Mar;120(2):129-138. [doi: 10.1017/s0950268897008558] [Medline: 9593481]
35. Fadl MA, Aydarous MA, Mao C, Yasmeen A. An association of VNTR polymorphism in intron3 of IL-4 gene with susceptibility to typhoid fever in Khartoum State, Sudan. Kuwait J Sci 2016;43(1) [FREE Full text]
36. Rabasa AI, Mava Y, Pius S, Timothy SY, Baba UA. Typhoid fever in children: clinical presentation and risk factors. Niger J Paediatr 2012 Dec 04;40(1):2013 [FREE Full text] [doi: 10.4314/njp.v40i1.11]
37. Ram PK, Naheed A, Brooks WA, Hossain MA, Mintz ED, Breiman RF, et al. Risk factors for typhoid fever in a slum in Dhaka, Bangladesh. Epidemiol Infect 2007 Apr 08;135(3):458-465. [doi: 10.1017/S0950268806007114] [Medline: 16893490]
38. Sharma PK, Ramakrishnan R, Hutin Y, Manickam P, Gupte MD. Risk factors for typhoid in Darjeeling, West Bengal, India: evidence for practical action. Trop Med Int Health 2009 Jun 23;14(6):696-702 [FREE Full text] [doi: 10.1111/j.1365-3156.2009.02283.x] [Medline: 19392739]
39. Jelinek T, Nothdurft HD, von Sonnenburg F, Löscher T. Risk factors for typhoid fever in travelers. J Travel Med 1996 Dec 01;3(4):200-203 [FREE Full text] [doi: 10.1111/j.1708-8305.1996.tb00746.x] [Medline: 9815456]
40. Abera B, Yitayew G, Amare H. Salmonella serotypeTyphi, Shigella, and intestinal parasites among food handlers at Bahir Dar University, Ethiopia. J Infect Dev Ctries 2016 Feb 28;10(2):121-126 [FREE Full text] [doi: 10.3855/jidc.6890] [Medline: 26927451]
41. Ackers ML, Puhr ND, Tauxe RV, Mintz ED. Laboratory-based surveillance of Salmonella serotype Typhi infections in the United States: antimicrobial resistance on the rise. JAMA 2000 May 24;283(20):2668-2673. [doi: 10.1001/jama.283.20.2668] [Medline: 10819949]
42. Akullian A, Ng'eno E, Matheson AI, Cosmas L, Macharia D, Fields B, et al. Environmental transmission of typhoid fever in an urban slum. PLoS Negl Trop Dis 2015 Dec 3;9(12):e0004212 [FREE Full text] [doi: 10.1371/journal.pntd.0004212] [Medline: 26633656]
43. Anand PK, Ramakrishnan R. Investigation of the outbreak of typhoid in a village of Thar Desert Rajasthan, India. Indian J Med Res 2010 Jun;131:799-803 [FREE Full text] [Medline: 20571169]
44. Anderson ES, Smith HR. Chloramphenicol resistance in the typhoid bacillus. Br Med J 1972 Aug 05;3(5822):329-331 [FREE Full text] [doi: 10.1136/bmj.3.5822.329] [Medline: 4558307]
45. Batool R, Qureshi S, Yousafzai M, Kazi MT, Ali M, Qamar FN. Risk factors associated with extensively drug-resistant typhoid in an outbreak setting of Lyari Town Karachi, Pakistan. Am J Trop Med Hyg 2022 Mar 28;106(5):1379-1383 [FREE Full text] [doi: 10.4269/ajtmh.21-1323] [Medline: 35344928]
46. Centers for Disease Control and Prevention (CDC). Epidemic typhoid fever--Dushanbe, Tajikistan, 1997. MMWR Morb Mortal Wkly Rep 1998 Sep 18;47(36):752-756 [FREE Full text] [Medline: 9756457]
47. Srikantiah P, Vafokulov S, Luby SP, Ishmail T, Earhart K, Khodjaev N, et al. Epidemiology and risk factors for endemic typhoid fever in Uzbekistan. Trop Med Int Health 2007 Jul 25;12(7):838-847 [FREE Full text] [doi: 10.1111/j.1365-3156.2007.01853.x] [Medline: 17596250]
48. Tran HH, Bjune G, Nguyen BM, Rottingen JA, Grais RF, Guerin PJ. Risk factors associated with typhoid fever in Son La province, northern Vietnam. Trans R Soc Trop Med Hyg 2005 Nov;99(11):819-826. [doi: 10.1016/j.trstmh.2005.05.007] [Medline: 16099488]
49. Bhan MK, Bahl R, Sazawal S, Sinha A, Kumar R, Mahalanabis D, et al. Association between Helicobacter pylori infection and increased risk of typhoid fever. J Infect Dis 2002 Dec 15;186(12):1857-1860. [doi: 10.1086/345762] [Medline: 12447776]
50. Birkhead GS, Morse DL, Levine WC, Fudala JK, Kondracki SF, Chang H, et al. Typhoid fever at a resort hotel in New York: a large outbreak with an unusual vehicle. J Infect Dis 1993 May 01;167(5):1228-1232. [doi: 10.1093/infdis/167.5.1228] [Medline: 8486960]
51. Boggild AK, Castelli F, Gautret P, Torresi J, von Sonnenburg F, Barnett ED, et al. Vaccine preventable diseases in returned international travelers: results from the GeoSentinel Surveillance Network. Vaccine 2010 Oct 28;28(46):7389-7395. [doi: 10.1016/j.vaccine.2010.09.009] [Medline: 20851081]

XSL•FO
RenderX

52. Braddick MR, Sharp JC. Enteric fever in Scotland 1975-1990. Public Health 1993 May;107(3):193-198. [doi: 10.1016/s0033-3506(05)80441-x] [Medline: 8511239]
53. Brainard J, D'hondt R, Ali E, Van den Bergh R, De Weggheleire A, Baudot Y, et al. Typhoid fever outbreak in the Democratic Republic of Congo: case control and ecological study. PLoS Negl Trop Dis 2018 Oct 3;12(10):e0006795 [FREE Full text] [doi: 10.1371/journal.pntd.0006795] [Medline: 30281598]
54. Chao DL, Park JK, Marks F, Ochiai RL, Longini IM, Halloran ME. The contribution of neighbours to an individual's risk of typhoid outcome. Epidemiol Infect 2015 Dec 04;143(16):3520-3527 [FREE Full text] [doi: 10.1017/S0950268815000692] [Medline: 25936682]
55. Clark ST, Cronin K, Corbeil AJ, Patel SN. A ten-year retrospective survey of antimicrobial susceptibility patterns among salmonella enterica subsp. enterica Serovar Typhi Isolates in Ontario, Canada. Microbiol Spectr 2023 Feb 14;11(1):e0482822 [FREE Full text] [doi: 10.1128/spectrum.04828-22] [Medline: 36622222]
56. Corner RJ, Dewan AM, Hashizume M. Modelling typhoid risk in Dhaka metropolitan area of Bangladesh: the role of socio-economic and environmental factors. Int J Health Geogr 2013 Mar 16;12:13 [FREE Full text] [doi: 10.1186/1476-072X-12-13] [Medline: 23497202]
57. de Alwis R, Watson C, Nikolay B, Lowry JH, Thieu NT, Van TT, et al. Role of environmental factors in shaping spatial distribution of Salmonella enterica Serovar Typhi, Fiji. Emerg Infect Dis 2018 Feb;24(2):284-293 [FREE Full text] [doi: 10.3201/eid2402.170704] [Medline: 29350150]
58. Deksissa T, Gebremedhin EZ. A cross-sectional study of enteric fever among febrile patients at Ambo hospital: prevalence, risk factors, comparison of Widal test and stool culture and antimicrobials susceptibility pattern of isolates. BMC Infect Dis 2019 Mar 27;19(1):288 [FREE Full text] [doi: 10.1186/s12879-019-3917-3] [Medline: 30917795]
59. Ekdahl K, de Jong B, Andersson Y. Risk of travel-associated typhoid and paratyphoid fevers in various regions. J Travel Med 2005;12(4):197-204 [FREE Full text] [doi: 10.2310/7060.2005.12405] [Medline: 16086894]
60. Elemfareji OI, Thong KL. Comparative virulotyping of Salmonella typhi and Salmonella enteritidis. Indian J Microbiol 2013 Dec 23;53(4):410-417 [FREE Full text] [doi: 10.1007/s12088-013-0407-y] [Medline: 24426144]
61. Farmakiotis D, Varughese J, Sue P, Andrews P, Brimmage M, Dobroszycki J, et al. Typhoid fever in an inner city hospital: a 5-year retrospective review. J Travel Med 2013 Jan 01;20(1):17-21 [FREE Full text] [doi: 10.1111/j.1708-8305.2012.00665.x] [Medline: 23279226]
62. Feasey NA, Masesa C, Jassi C, Faragher EB, Mallewa J, Mallewa M, et al. Three epidemics of invasive multidrug-resistant Salmonella bloodstream infection in Blantyre, Malawi, 1998-2014. Clin Infect Dis 2015 Nov 01;61 Suppl 4(Suppl 4):S363-S371 [FREE Full text] [doi: 10.1093/cid/civ691] [Medline: 26449953]
63. Fida S, Mansoor H, Saif S, Iqbal J, Khan AQ. Clinical perspectives of multiple and extensively drug-resistant typhoid; result from a tertiary care hospital from Pakistan. J Infect Dev Ctries 2021 Apr 30;15(4):530-537. [doi: 10.3855/jidc.13539] [Medline: 33956653]
64. Forster DP, Leder K. Typhoid fever in travellers: estimating the risk of acquisition by country. J Travel Med 2021 Dec 29;28(8):taab150 [FREE Full text] [doi: 10.1093/jtm/taab150] [Medline: 34619766]
65. Forsyth JR, Bennett NM, Hogben S, Hutchinson EM, Rouch G, Tan A, et al. The year of the Salmonella seekers--1977. Aust N Z J Public Health 2003 Aug;27(4):385-389 [FREE Full text] [doi: 10.1111/j.1467-842x.2003.tb00414.x] [Medline: 14705299]
66. Gao Q, Liu Z, Xiang J, Zhang Y, Tong MX, Wang S, et al. Impact of temperature and rainfall on typhoid/paratyphoid fever in Taizhou, China: effect estimation and vulnerable group identification. Am J Trop Med Hyg 2021 Dec 06;106(2):532-542 [FREE Full text] [doi: 10.4269/ajtmh.20-1457] [Medline: 34872055]
67. Gauld JS, Olgemoeller F, Heinz E, Nkhata R, Bilima S, Wailan AM, et al. Spatial and genomic data to characterize endemic typhoid transmission. Clin Infect Dis 2022 Jun 10;74(11):1993-2000 [FREE Full text] [doi: 10.1093/cid/ciab745] [Medline: 34463736]
68. Gauld JS, Olgemoeller F, Nkhata R, Li C, Chirambo A, Morse T, et al. Domestic river water use and risk of typhoid fever: results from a case-control study in Blantyre, Malawi. Clin Infect Dis 2020 Mar 17;70(7):1278-1284 [FREE Full text] [doi: 10.1093/cid/ciz405] [Medline: 31144715]
69. Geteneh A, Tadesse S, Biset S, Girma L, Fissiha P. Rapid stool antigenic test for typhoid fever among suspected cases, Northeast, Ethiopia. Sci Rep 2023 Jan 12;13(1):649 [FREE Full text] [doi: 10.1038/s41598-023-27909-5] [Medline: 36635427]
70. Giri S, Mohan VR, Srinivasan M, Kumar N, Kumar V, Dhanapal P, et al. Case-control study of household and environmental transmission of typhoid fever in India. J Infect Dis 2021 Nov 23;224(Supple 5):S584-S592 [FREE Full text] [doi: 10.1093/infdis/jiab378] [Medline: 35238355]
71. Gotuzzo E, Frisancho O, Sanchez J, Liendo G, Carrillo C, Black RE, et al. Association between the acquired immunodeficiency syndrome and infection with Salmonella typhi or Salmonella paratyphi in an endemic typhoid area. Arch Intern Med 1991 Feb;151(2):381-382. [Medline: 1899554]
72. Grist NR, Emslie JA. Infections in British clinical laboratories, 1988-1989. J Clin Pathol 1991 Aug;44(8):667-669 [FREE Full text] [doi: 10.1136/jcp.44.8.667] [Medline: 1890201]

XSL•FO
RenderX

73. Gunaratnam P, Tobin S, Seale H, McAnulty JM. Infectious diseases in returned travellers, NSW, 2010-2011. N S W Public Health Bull 2014 Jun;24(4):171-175. [doi: 10.1071/NB13005] [Medline: 24939227]
74. Im J, Islam MT, Ahmmed F, Kim DR, Islam Khan A, Zaman K, et al. Can existing improvements of water, sanitation, and hygiene (WASH) in urban slums reduce the burden of typhoid fever in these settings? Clin Infect Dis 2021 Jun 01;72(11):e720-e726. [doi: 10.1093/cid/ciaa1429] [Medline: 32964216]
75. Im J, Khanam F, Ahmmed F, Kim DR, Kang S, Tadesse BT, et al. Prevention of typhoid fever by existing improvements in household water, sanitation, and hygiene, and the use of the Vi polysaccharide typhoid vaccine in poor urban slums: results from a cluster-randomized trial. Am J Trop Med Hyg 2022 Mar 07;106(4):1149-1155 [FREE Full text] [doi: 10.4269/ajtmh.21-1034] [Medline: 35385827]
76. Issack MI. Epidemiology of typhoid fever in Mauritius. J Travel Med 2005;12(5):270-274 [FREE Full text] [doi: 10.2310/7060.2005.12506] [Medline: 16256051]
77. Iyer V, Sharma A, Nair D, Solanki B, Umrigar P, Murtugudde R, et al. Role of extreme weather events and El Niño Southern Oscillation on incidence of Enteric Fever in Ahmedabad and Surat, Gujarat, India. Environ Res 2021 May;196:110417. [doi: 10.1016/j.envres.2020.110417] [Medline: 33217433]
78. Jenkins AP, Jupiter S, Mueller U, Jenney A, Vosaki G, Rosa V, et al. Health at the sub-catchment scale: typhoid and its environmental determinants in Central Division, Fiji. Ecohealth 2016 Dec 24;13(4):633-651. [doi: 10.1007/s10393-016-1152-6] [Medline: 27557784]
79. Jenkins AP, Jupiter SD, Jenney A, Naucukidi A, Prasad N, Vosaki G, et al. Environmental foundations of typhoid fever in the Fijian residential setting. Int J Environ Res Public Health 2019 Jul 06;16(13):2407 [FREE Full text] [doi: 10.3390/ijerph16132407] [Medline: 31284613]
80. Kabwama SN, Bulage L, Nsubuga F, Pande G, Oguttu DW, Mafigiri R, et al. A large and persistent outbreak of typhoid fever caused by consuming contaminated water and street-vended beverages: Kampala, Uganda, January - June 2015. BMC Public Health 2017 Jan 05;17(1):23 [FREE Full text] [doi: 10.1186/s12889-016-4002-0] [Medline: 28056940]
81. Kamal R, Ching C, Zaman MH, Sultan F, Abbas S, Khan E, et al. Identification of multiple variant extensively drug-resistant typhoid infections across Pakistan. Am J Trop Med Hyg 2023 Feb 01;108(2):278-284 [FREE Full text] [doi: 10.4269/ajtmh.22-0071] [Medline: 36623485]
82. Karkey A, Arjyal A, Anders KL, Boni MF, Dongol S, Koirala S, et al. The burden and characteristics of enteric fever at a healthcare facility in a densely populated area of Kathmandu. PLoS One 2010 Nov 15;5(11):e13988 [FREE Full text] [doi: 10.1371/journal.pone.0013988] [Medline: 21085575]
83. Karkey A, Thompson CN, Tran Vu Thieu N, Dongol S, Le Thi Phuong T, Voong Vinh P, et al. Differential epidemiology of Salmonella Typhi and Paratyphi A in Kathmandu, Nepal: a matched case control investigation in a highly endemic enteric fever setting. PLoS Negl Trop Dis 2013 Aug 22;7(8):e2391 [FREE Full text] [doi: 10.1371/journal.pntd.0002391] [Medline: 23991240]
84. Katz DJ, Cruz MA, Trepka MJ, Suarez JA, Fiorella PD, Hammond RM. An outbreak of typhoid fever in Florida associated with an imported frozen fruit. J Infect Dis 2002 Jul 15;186(2):234-239. [doi: 10.1086/341208] [Medline: 12134260]
85. Keller A, Frey M, Schmid H, Steffen R, Walker T, Schlagenhauf P. Imported typhoid fever in Switzerland, 1993 to 2004. J Travel Med 2008 Jul 01;15(4):248-251 [FREE Full text] [doi: 10.1111/j.1708-8305.2008.00216.x] [Medline: 18666925]
86. Lin FH, Chen BC, Chou YC, Hsieh CJ, Yu CP. Incidence and risk factors for notifiable typhoid and paratyphoid in Taiwan during the period 2011-2020. Healthcare (Basel) 2021 Oct 01;9(10):1316 [FREE Full text] [doi: 10.3390/healthcare9101316] [Medline: 34682996]
87. Liu Z, Lao J, Zhang Y, Liu Y, Zhang J, Wang H, et al. Association between floods and typhoid fever in Yongzhou, China: effects and vulnerable groups. Environ Res 2018 Nov;167:718-724. [doi: 10.1016/j.envres.2018.08.030] [Medline: 30241731]
88. Luxemburger C, Chau MC, Mai NL, Wain J, Tran TH, Simpson JA, et al. Risk factors for typhoid fever in the Mekong delta, southern Viet Nam: a case-control study. Trans R Soc Trop Med Hyg 2001 Jan;95(1):19-23. [doi: 10.1016/s0035-9203(01)90318-9] [Medline: 11280056]
89. Makungo UB, Ramutshila TE, Mabotja MC, Thomas J, Lekalakala-Mokaba R, Smith AM, et al. Epidemiological investigation of a typhoid fever outbreak in Sekhukhune District, Limpopo province, South Africa - 2017. S Afr J Infect Dis 2020 Nov 16;35(1):107 [FREE Full text] [doi: 10.4102/sajid.v35i1.107] [Medline: 34485467]
90. Masinaei M, Eshrati B, Yaseri M. Spatial and spatiotemporal patterns of typhoid fever and investigation of their relationship with potential risk factors in Iran, 2012-2017. Int J Hyg Environ Health 2020 Mar;224:113432. [doi: 10.1016/j.ijheh.2019.113432] [Medline: 31978729]
91. Mathieu JJ, Henning KJ, Bell E, Frieden TR. Typhoid fever in New York City, 1980 through 1990. Arch Intern Med 1994 Aug 08;154(15):1713-1718. [Medline: 8042888]
92. Mbae C, Mwangi M, Gitau N, Irungu T, Muendo F, Wakio Z, et al. Factors associated with occurrence of salmonellosis among children living in Mukuru slum, an urban informal settlement in Kenya. BMC Infect Dis 2020 Jun 17;20(1):422 [FREE Full text] [doi: 10.1186/s12879-020-05134-z] [Medline: 32552753]
93. Mermin JH, Townes JM, Gerber M, Dolan N, Mintz ED, Tauxe RV. Typhoid fever in the United States, 1985-1994: changing risks of international travel and increasing antimicrobial resistance. Arch Intern Med 1998 Mar 23;158(6):633-638. [doi: 10.1001/archinte.158.6.633] [Medline: 9521228]

XSL•FO
RenderX

94. Michel R, Garnotel E, Spiegel A, Morillon M, Saliou P, Boutin J. Outbreak of typhoid fever in vaccinated members of the French Armed Forces in the Ivory Coast. Eur J Epidemiol 2005 Jul;20(7):635-642. [doi: 10.1007/s10654-005-7454-6] [Medline: 16119438]
95. Mohan VR, Srinivasan M, Sinha B, Shrivastava A, Kanungo S, Natarajan Sindhu K, et al. Geographically weighted regression modeling of spatial clustering and determinants of focal typhoid fever incidence. J Infect Dis 2021 Nov 23;224(Supple 5):S601-S611 [FREE Full text] [doi: 10.1093/infdis/jiab379] [Medline: 35238357]
96. Muti M, Gombe N, Tshimanga M, Takundwa L, Bangure D, Mungofa S, et al. Typhoid outbreak investigation in Dzivaresekwa, suburb of Harare City, Zimbabwe, 2011. Pan Afr Med J 2014;18:309 [FREE Full text] [doi: 10.11604/pamj.2014.18.309.4288] [Medline: 25469202]
97. Nyamusore J, Nahimana MR, Ngoc CT, Olu O, Isiaka A, Ndahindwa V, et al. Risk factors for transmission of Salmonella Typhi in Mahama refugee camp, Rwanda: a matched case-control study. Pan Afr Med J 2018;29:148. [doi: 10.11604/pamj.2018.29.148.12070] [Medline: 30050612]
98. Okpasuo O, Aguzie IO, Joy AT, Okafor FC. Risk assessment of waterborne infections in Enugu State, Nigeria: implications of household water choices, knowledge, and practices. AIMS Public Health 2020;7(3):634-649 [FREE Full text] [doi: 10.3934/publichealth.2020050] [Medline: 32968683]
99. Pommelet V, Mariani P, Basmaci R, Tourdjman M, Morin L, Gaschignard J, et al. Enteric fever among children: 50 cases in a French tertiary care centre. J Travel Med 2018 Jan 01;25(1):059. [doi: 10.1093/jtm/tay059] [Medline: 30060197]
100. Pradier C, Keita-Perse O, Bernard E, Gisbert C, Vezolles MJ, Armengaud A, et al. Outbreak of typhoid fever on the French Riviera. Eur J Clin Microbiol Infect Dis 2000 Jun 4;19(6):464-467. [doi: 10.1007/s100960000300] [Medline: 10947223]
101. Qamar FN, Yousafzai MT, Khalid M, Kazi AM, Lohana H, Karim S, et al. Outbreak investigation of ceftriaxone-resistant Salmonella enterica serotype Typhi and its risk factors among the general population in Hyderabad, Pakistan: a matched case-control study. Lancet Infect Dis 2018 Dec;18(12):1368-1376. [doi: 10.1016/S1473-3099(18)30483-3] [Medline: 30507460]
102. Ren X, Zhang S, Luo P, Zhao J, Kuang W, Ni H, et al. Spatial heterogeneity of socio-economic determinants of typhoid/paratyphoid fever in one province in central China from 2015 to 2019. BMC Public Health 2023 May 22;23(1):927 [FREE Full text] [doi: 10.1186/s12889-023-15738-0] [Medline: 37217879]
103. Srinivasan M, Sindhu K, Kumar JS, Ramasamy RK, Pragasam AK, Aasaithampi P, et al. Outbreak of typhoid fever in children of urban Vellore: a report from the surveillance for enteric fever in India cohort. Am J Trop Med Hyg 2022 Jul 13;107(1):82-85. [doi: 10.4269/ajtmh.21-0593] [Medline: 35895361]
104. Sur D, Ali M, von Seidlein L, Manna B, Deen JL, Acosta CJ, et al. Comparisons of predictors for typhoid and paratyphoid fever in Kolkata, India. BMC Public Health 2007 Oct 12;7(1):289 [FREE Full text] [doi: 10.1186/1471-2458-7-289] [Medline: 17935611]
105. Sutiono AB, Qiantori A, Suwa H, Ohta T. Characteristics and risk factors for typhoid fever after the tsunami, earthquake and under normal conditions in Indonesia. BMC Res Notes 2010 Apr 17;3:106 [FREE Full text] [doi: 10.1186/1756-0500-3-106] [Medline: 20398409]
106. Tadesse BT, Khanam F, Ahmmed F, Im J, Islam MT, Kim DR, et al. Prevention of typhoid by Vi conjugate vaccine and achievable improvements in household water, sanitation, and hygiene: evidence from a cluster-randomized trial in Dhaka, Bangladesh. Clin Infect Dis 2022 Nov 14;75(10):1681-1687 [FREE Full text] [doi: 10.1093/cid/ciac289] [Medline: 35412603]
107. Vighio A, Syed MA, Hussain I, Zia SM, Fatima M, Masood N, et al. Risk factors of extensively drug resistant typhoid fever among children in Karachi: case-control study. JMIR Public Health Surveill 2021 May 11;7(5):e27276 [FREE Full text] [doi: 10.2196/27276] [Medline: 33973861]
108. Vollaard AM, Ali S, van Asten HA, Ismid IS, Widjaja S, Visser LG, et al. Risk factors for transmission of foodborne illness in restaurants and street vendors in Jakarta, Indonesia. Epidemiol Infect 2004 Oct;132(5):863-872. [doi: 10.1017/s0950268804002742] [Medline: 15473149]
109. Vollaard AM, Verspaget HW, Ali S, Visser LG, Veenendaal RA, Van Asten HA, et al. Helicobacter pylori infection and typhoid fever in Jakarta, Indonesia. Epidemiol Infect 2006 Feb 22;134(1):163-170. [doi: 10.1017/S0950268805004875] [Medline: 16409664]
110. Wang JF, Wang Y, Zhang J, Christakos G, Sun JL, Liu X, et al. Spatiotemporal transmission and determinants of typhoid and paratyphoid fever in Hongta District, Yunnan Province, China. PLoS Negl Trop Dis 2013;7(3):e2112 [FREE Full text] [doi: 10.1371/journal.pntd.0002112] [Medline: 23516653]
111. Wang LX, Li XJ, Fang LQ, Wang DC, Cao WC, Kan B. Association between the incidence of typhoid and paratyphoid fever and meteorological variables in Guizhou, China. Chin Med J (Engl) 2012 Feb;125(3):455-460. [Medline: 22490402]
112. Afifi S, Earhart K, Azab MA, Youssef FG, El Sakka H, Wasfy M, et al. Hospital-based surveillance for acute febrile illness in Egypt: a focus on community-acquired bloodstream infections. Am J Trop Med Hyg 2005 Aug;73(2):392-399. [Medline: 16103611]
113. Bendib A, Dridi H, Kalla M, Baziz N. Spatial analysis of typhoid fever vulnerability in the city of Batna (eastern Algeria). Environ Risques Santé 2016;15(3):228-237 [FREE Full text]
114. Coté TR, Convery H, Robinson D, Ries A, Barrett T, Frank L, et al. Typhoid fever in the park: epidemiology of an outbreak at a cultural interface. J Community Health 1995 Dec;20(6):451-458. [doi: 10.1007/BF02277062] [Medline: 8568020]

XSL•FO
RenderX

115. Kelly-Hope LA, Alonso WJ, Thiem VD, Anh DD, Canh DG, Lee H, et al. Geographical distribution and risk factors associated with enteric diseases in Vietnam. Am J Trop Med Hyg 2007 Apr;76(4):706-712. [Medline: 17426175]
116. Khatami A, Khan F, Macartney KK. Enteric fever in children in Western Sydney, Australia, 2003-2015. Pediatr Infect Dis J 2017 Dec;36(12):1124-1128. [doi: 10.1097/INF.0000000000001606] [Medline: 28445250]
117. Salama RI, Emara MH, Mostafa HM, Abd-Elsalam S, Alnabawy S, Elshweikh SA, et al. Helicobacter pylori infection and risk of salmonella infection. Medicine (Baltimore) 2019 Feb;98(6):e14335 [FREE Full text] [doi: 10.1097/MD.0000000000014335] [Medline: 30732157]
118. Stroffolini T, Manzillo G, De Sena R, Manzillo E, Pagliano P, Zaccarelli M, et al. Typhoid fever in the Neapolitan area: a case-control study. Eur J Epidemiol 1992 Jul;8(4):539-542. [doi: 10.1007/BF00146373] [Medline: 1397222]
119. Yousafzai MT, Qamar FN, Shakoor S, Saleem K, Lohana H, Karim S, et al. Ceftriaxone-resistant salmonella typhi outbreak in Hyderabad City of Sindh, Pakistan: high time for the introduction of typhoid conjugate vaccine. Clin Infect Dis 2019 Feb 15;68(Suppl 1):S16-S21 [FREE Full text] [doi: 10.1093/cid/ciy877] [Medline: 30767003]
120. Abdullahi Tawfiq U, Shohaimi S, Mohd Nadzir MN, Amin Nordin S, Ab Rahman AH, Salari N. Epidemiology and risk factors for typhoid fever in Gombe metropolis, Gombe State, Nigeria. Malays J Public Health Med 2022;22(2):110-121 [FREE Full text] [doi: 10.37268/mjphm/vol.22/no.2/art.1656]
121. Titus I, Iliyasu MY, Sahal MR, Umar RD, Wali MM, Ismai'l S, et al. Phenotypic characterization of virulence factors in multidrug-resistant Salmonella Typhi serovars isolated from clinical specimens in Bauchi metropolis. South Asian J Res Microbiol 2023;15(3):10-20 [FREE Full text] [doi: 10.9734/sajrm/2023/v15i3287]
122. Velema JP, van Wijnen G, Bult P, van Naerssen T, Jota S. Typhoid fever in Ujung Pandang, Indonesia--high-risk groups and high-risk behaviours. Trop Med Int Health 1997 Nov 03;2(11):1088-1094 [FREE Full text] [doi: 10.1046/j.1365-3156.1997.d01-179.x] [Medline: 9391512]
123. Vollaard AM, Ali S, van Asten HA, Widjaja S, Visser LG, Surjadi C, et al. Risk factors for typhoid and paratyphoid fever in Jakarta, Indonesia. JAMA 2004 Jun 02;291(21):2607-2615. [doi: 10.1001/jama.291.21.2607] [Medline: 15173152]
124. Dunstan SJ, Nguyen TH, Rockett K, Forton J, Morris AP, Diakite M, et al. A TNF region haplotype offers protection from typhoid fever in Vietnamese patients. Hum Genet 2007 Aug 15;122(1):51-61 [FREE Full text] [doi: 10.1007/s00439-007-0372-9] [Medline: 17503085]
125. Steinberg EB, Bishop R, Haber P, Dempsey AF, Hoekstra RM, Nelson JM, et al. Typhoid fever in travelers: who should be targeted for prevention? Clin Infect Dis 2004 Jul 15;39(2):186-191. [doi: 10.1086/421945] [Medline: 15307027]
126. Feng Y, Pan H, Zheng B, Li F, Teng L, Jiang Z, et al. An integrated nationwide genomics study reveals transmission modes of typhoid fever in China. mBio 2023 Oct 31;14(5):e0133323 [FREE Full text] [doi: 10.1128/mbio.01333-23] [Medline: 37800953]
127. Siddiqui FJ, Haider SR, Bhutta ZA. Risk factors for typhoid fever in children in squatter settlements of Karachi: a nested case-control study. J Infect Public Health 2008;1(2):113-120 [FREE Full text] [doi: 10.1016/j.jiph.2008.10.003] [Medline: 20701852]
128. Muche G, Tesfaw A, Bayou FD. Prevalence of typhoid fever and its associated factors among febrile patients visiting Arerti Primary Hospital, Amhara Region, north east Ethiopia. Front Public Health 2024 Aug 16;12:1357131 [FREE Full text] [doi: 10.3389/fpubh.2024.1357131] [Medline: 39220452]
129. Herdman MT, Karo B, Dave J, Katwa P, Freedman J, Do Nascimento V, et al. Increasingly limited options for the treatment of enteric fever in travellers returning to England, 2014-2019: a cross-sectional analytical study. J Med Microbiol 2021 Aug;70(8):1359 [FREE Full text] [doi: 10.1099/jmm.0.001359] [Medline: 34351258]
130. Antillón M, Warren JL, Crawford FW, Weinberger DM, Kürüm E, Pak GD, et al. The burden of typhoid fever in low- and middle-income countries: a meta-regression approach. PLoS Negl Trop Dis 2017 Feb 27;11(2):e0005376 [FREE Full text] [doi: 10.1371/journal.pntd.0005376] [Medline: 28241011]
131. Anonymous. Typhoid in the large cities of the United States in 1919. JAMA 1920 Mar 06;74(10):672-675 [FREE Full text] [doi: 10.1001/jama.1920.26210100002011]
132. Progress on drinking water and sanitation: Joint Monitoring Programme update 2012. World Health Organization. 2012 Mar 06. URL: https://www.who.int/publications/i/item/9789280646320 [accessed 2024-02-08]
133. Ercumen A, Arnold BF, Kumpel E, Burt Z, Ray I, Nelson K, et al. Upgrading a piped water supply from intermittent to continuous delivery and association with waterborne illness: a matched cohort study in urban India. PLoS Med 2015 Oct 27;12(10):e1001892 [FREE Full text] [doi: 10.1371/journal.pmed.1001892] [Medline: 26505897]
134. Karkey A, Jombart T, Walker AW, Thompson CN, Torres A, Dongol S, et al. The ecological dynamics of fecal contamination and Salmonella Typhi and Salmonella Paratyphi A in municipal Kathmandu drinking water. PLoS Negl Trop Dis 2016 Jan 6;10(1):e0004346 [FREE Full text] [doi: 10.1371/journal.pntd.0004346] [Medline: 26735696]
135. Feglo PK, Frimpong EH, Essel-Ahun M. Salmonellae carrier status of food vendors in Kumasi, Ghana. East Afr Med J 2004 Jul 22;81(7):358-361. [doi: 10.4314/eamj.v81i7.9191] [Medline: 15490708]
136. Parry CM. Epidemiological and clinical aspects of human typhoid fever. In: Maskell D, Mastroeni P, editors. Salmonella Infections: Clinical, Immunological and Molecular Aspects. Cambridge, UK: Cambridge University Press; 2006.
137. Rasul F, Sughra K, Mushtaq A, Zeeshan N, Mehmood S, Rashid U. Surveillance report on typhoid fever epidemiology and risk factor assessment in district Gujrat, Punjab, Pakistan. Biomed Res 2017;28(8):1-6 [FREE Full text]

XSL•FO
RenderX

138. Parry CM, Karunanayake L, Coulter JB, Beeching NJ. Test for quinolone resistance in typhoid fever. BMJ 2006 Jul 29;333(7561):260-261 [FREE Full text] [doi: 10.1136/bmj.333.7561.260-b] [Medline: 16873871]
139. Gunn JS, Marshall JM, Baker S, Dongol S, Charles RC, Ryan ET. Salmonella chronic carriage: epidemiology, diagnosis, and gallbladder persistence. Trends Microbiol 2014 Nov;22(11):648-655 [FREE Full text] [doi: 10.1016/j.tim.2014.06.007] [Medline: 25065707]
140. Gupta V, Kaur M, Datta P, Chander J. Chronic urinary carrier state due to Salmonella Typhi causing urinary tract infection in an immunocompetent healthy woman. Trop Doct 2018 Jul 07;48(3):236-238. [doi: 10.1177/0049475517752205] [Medline: 29307275]
141. Clemens J, Albert MJ, Rao M, Qadri F, Huda S, Kay B, et al. Impact of infection by Helicobacter pylori on the risk and severity of endemic cholera. J Infect Dis 1995 Jun 01;171(6):1653-1656. [doi: 10.1093/infdis/171.6.1653] [Medline: 7769312]
142. Passaro DJ, Taylor DN, Meza R, Cabrera L, Gilman RH, Parsonnet J. Acute Helicobacter pylori infection is followed by an increase in diarrheal disease among Peruvian children. Pediatrics 2001 Nov;108(5):E87. [doi: 10.1542/peds.108.5.e87] [Medline: 11694671]
143. Saad NJ, Lynch VD, Antillón M, Yang C, Crump JA, Pitzer VE. Seasonal dynamics of typhoid and paratyphoid fever. Sci Rep 2018 May 02;8(1):6870 [FREE Full text] [doi: 10.1038/s41598-018-25234-w] [Medline: 29720736]
144. Butler T, Mahmoud AA, Warren KS. Algorithms in the diagnosis and management of exotic diseases. XXIII. Typhoid fever. J Infect Dis 1977 Jun 01;135(6):1017-1020. [doi: 10.1093/infdis/135.6.1017] [Medline: 864284]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
TCV: typhoid conjugate vaccine
WASH: water, sanitation, and hygiene
WHO: World Health Organization

Edited by A Mavragani; submitted 14.10.24; peer-reviewed by I Friesema, M Yue; comments to author 13.02.25; revised version received 07.05.25; accepted 04.06.25; published 28.08.25.

Please cite as:
Boakye Okyere P, Twumasi-Ankrah S, Newton S, Nkansah Darko S, Owusu Ansah M, Darko E, Agyapong F, Jeon HJ, Adu-Sarkodie Y, Marks F, Owusu-Dabo E
Risk Factors for Typhoid Fever: Systematic Review
JMIR Public Health Surveill 2025;11:e67544
URL: https://publichealth.jmir.org/2025/1/e67544
doi:10.2196/67544
PMID:40875987

©Portia Boakye Okyere, Sampson Twumasi-Ankrah, Sam Newton, Samuel Nkansah Darko, Michael Owusu Ansah, Eric Darko, Francis Agyapong, Hyon Jin Jeon, Yaw Adu-Sarkodie, Florian Marks, Ellis Owusu-Dabo. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Health Equity of Hypertension Management Between Local Residents and Internal Migrants in Shenzhen, China: Cross-Sectional Study

Jinyu He[1,2], MD; Yanjun Li[1], BA; Huatang Zeng[1,3], MSc; Haoran Sun[1,2], MD; Liqun Wu[3], MSc; Zhenzhen Zhu[3], MSc; Ning Zhang[1,2], PhD; Wannian Liang[1,2], PhD

[1]Vanke School of Public Health, Tsinghua University, Beijing, China

[2]Institute of Healthy China, Tsinghua University, Haidian District, Beijing, China

[3]Public Hygiene and Health Commission of Shenzhen Municipality, Shenzhen Health Development Research and Data Management Center, Shenzhen, China

Corresponding Author:
Ning Zhang, PhD
Vanke School of Public Health, Tsinghua University, Beijing, China

Abstract

Background: With hypertension emerging as a global public health concern, taking measures to alleviate its burden is urgently needed. The hypertension management program (HMP) in China is a standardized policy to help people with hypertension to improve their health levels and reduce health inequalities. However, studies focusing on details regarding participation in this program remain scarce.

Objective: This study aims to investigate the participation rates in HMPs in China and examine the health disparities in hypertension management between local residents and internal migrants in Shenzhen.

Methods: This study used the medical record of Shenzhen, Guangdong, China. We included adults with new-onset diagnosis of hypertension after 2017 and focused on patients who have a record in the community health center. We described the basic characteristics of people diagnosed with hypertension, including age, gender, marital status, occupation, education level, and health insurance type. Enrolled rate, follow-up rate, and adherence rate were used to measure the engagement with this program at the city, district, and community levels.

Results: Of the 1,160,214 hypertensive patients, 29.70% (344,526/1,160,214) were local residents, while 70.3% (815,688/1,160,214) were internal migrants. In addition, 55.36% (642,250/1,160,214) were enrolled in the HMP. Of those, 57.52% (369,439/642,25) were followed up. In addition, 62.59% (231,217/369,439) of followed up individuals were adherents. Internal migrants demonstrated a significantly higher participation in the HMP, especially for the enrolled rate (local residents: 156,085/344,526, 45.30% vs internal migrants: 486,165/815,688, 59.60%) and adherence rate (local residents: 44,044/84,635, 52.04% vs internal migrants: 187,173/284,804, 65.72%). Apart from that, local, single, and younger individuals had lower rates compared to their counterparts. There also existed within-district and within-community variation among enrolled, follow-up, and adherence rates.

Conclusions: Based on our research, individuals with different demographic and socioeconomic characteristics and in different regions had different enrolled, follow-up, and adherence rates. Internal migrants in Shenzhen showed a significantly higher participation in the HMP. Given these findings, there exists the potential to enhance the outreach and engagement of local, single, and younger populations through targeted promotional strategies.

(JMIR Public Health Surveill 2025;11:e65548) doi:10.2196/65548

KEYWORDS

health equity; hypertension management; immigrant population; hypertension; China; global health; public health; health disparities; medical record; community health care; native population; immigrant; socioeconomic; disease burden; well-being; chronic disease; community health center

XSL•FO
RenderX

Introduction

Background

The United Nations' Sustainable Development Goals are important global issues. The goals focus on promoting health and well-being, as well as reducing inequalities, highlighting the need for all individuals, regardless of their circumstances, to be able to access quality health care and essential services for optimal health outcomes.

Hypertension presents a significant global health challenge, emerging as a primary contributor to the global disease burden. Analysis from the Global Burden of Disease Study reveals that high systolic blood pressure stands as the foremost risk factor for mortality worldwide, accounting for 10.8 million cardiovascular disease deaths and 11.3 million deaths overall in 2021 [1]. The World Health Organization's latest report highlights a surge in hypertension prevalence, with figures doubling from 650 million individuals in 1990 to an alarming 1.3 billion in 2019 [2]. In China, about 270 million people have hypertension, while research showed that, among the 256 million individuals aged 30-79 years with hypertension, only 52% were aware of their condition, 39% were receiving treatment, and a scant 16% had their blood pressure under control [2,3]. In China, hypertension is one of the earliest chronic diseases undergoing widespread public health intervention. The government has bolstered efforts to combat hypertension through initiatives such as the Basic Public Health Services policy, China Medium- and Long-term Plan for the Prevention and Treatment of Chronic Diseases, National Demonstration Areas for Comprehensive Chronic Disease Prevention and Control, and the National Healthy Lifestyle Action. Chinese strategies underscore the inclusion of hypertension management within basic public health services at the grassroots level, ensuring regular dissemination of national guidelines on hypertension management and treatment. These guidelines provide comprehensive details on the management processes and protocols for diagnosed patients and vulnerable individuals [4]. According to the National Standard for Basic Public Health Services [5], the hypertension management program (HMP) is one of most crucial programs within the Basic Public Health Services. It requires community health centers (CHCs) to establish dedicated medical records for hypertensive patients and provide regular follow-up services without additional fees beyond outpatient costs. According to performance assessments, there must be at least 1 follow-up visit record every 3 months.

Significant health inequities exist in the health policy context; health-related behaviors and beliefs; quality of care; utilization of health care services; and health status [6]. Research has highlighted health inequities between local residents and internal migrants in China. Many internal migrants still face significant barriers in accessing local public health services, particularly in establishing health records and receiving health education. This is more common in economically developed eastern regions, where internal migrants use public health services less often than those in central regions, indicating regional disparities [7]. Furthermore, internal migrants often have poorer health compared to local residents due to socioeconomic and cultural factors, with social integration and income inequality being key determinants. In 2017, internal migrants in Qingdao had relatively better health, while those in Shenzhen faced the greatest health challenges [8].

In terms of hypertension management, treatment, and control within the population [9], previous research has explored the potential impacts of multidimensional factors on the health inequity of hypertension and cardiovascular disease [10]; the factors can be summarized as individual, socioeconomic [11,12], and environmental. A nationally representative cross-sectional study in China revealed that age, BMI, and economic status contributed to the inequitable situation of hypertension prevalence [13]. Another study showed that economic status and educational level are associated with socioeconomic inequalities of health services utilization among patients with hypertension in the Pearl River Delta of China [14]. In addition, due to economic status, the hypertension burden appeared higher in rural areas compared to urban regions in China [15].

As China's first Special Economic Zone, Shenzhen has undergone sweeping changes, including a significant demographic shift due to a large influx of internal migrants from other parts of Mainland China. According to The Shenzhen Statistical Yearbook 2022, the total population of Shenzhen was 17,681,600, of which the registered household population was 5,563,900 (31.5%) and the nonregistered population was 12,117,700 (68.5%). Despite this, Shenzhen, similar to many other regions in China, prioritizes the provision of health care services to its local residents [16]. For example, older adults with local household registrations or local medical insurance are prioritized for free influenza vaccination, while internal migrants without local insurance have to pay out of pocket [17]. Local people aged ≥60 years are eligible for free health checks, whereas internal migrants are required to be ≥65 years old [18].

As a survey indicated that the prevalence rate of hypertension among the residents of Shenzhen is 20.74% [5], and the incidence of hypertension in Shenzhen is increasing, the HMP is accessible to all patients regardless of household registration status, with the aim of reducing health disparities; however, as of yet, there has been no research comparing the health inequity in hypertension management between internal migrants and local residents. In the context of an internal migrant–dense city, understanding the unique challenges and needs of both local residents and internal migrants is essential.

This Study

The primary objective of this baseline report was to describe the characteristics of hypertensive patients and those who have received HMP services at a CHC, as well as the current status of the specific services they receive. The secondary objective was to reveal the health inequity for utilization of hypertension management between local residents and internal migrants in Shenzhen.

Methods

Study Design and Data Source

This study used a cross-sectional design. The data of the study were obtained from CHCs and the government-owned hospitals

XSL•FO
RenderX

in Shenzhen, Guangdong, China. All data were extracted with SQL queries from original datasets at the Shenzhen Health Development Research and Data Management Center. The data included the patients' basic information as well as the actual hypertension record and follow-up service record from January 1, 2017, to September 30, 2023. We defined participants without Shenzhen household registration who had been living in Shenzhen for 6 months or more as internal migrants. The HMP is a key component of primary care services for hypertensive patients. Therefore, this study uses the enrollment rate, follow-up rate, and adherence rate of this program to assess the health equity in utilization of primary care services between local residents and internal migrants.

Eligibility Criteria of Resident Records

We included the records of the following residents: (1) those who were documented in either the health information system of a CHC or medical record of government-owned hospitals in Shenzhen, Guangdong, China from January 1, 2017, to September 30 2023; (2) those who were diagnosed with hypertension according to these information systems (patients with gestational hypertension, based on the International Classification of Diseases Tenth Revision code, were excluded); (3) those whose earliest diagnosis date was after January 1, 2017, and (4) those who were aged ≥18 years, which is the age threshold for blood pressure measurements when people visit medical institutions as per the hypertension management policy in Shenzhen.

Sociodemographic Variables

A total of 6 individual-level sociodemographic factors were included in this study: gender (male or female), age (<45 years, 45 - 65 years, and >65 years), marital status (single, married, widowed, or divorced), occupation (technical personnel/office staff, business service industry, agriculture and forestry, industry, other, unemployed), education level (Bachelor's degree or above, college diploma, high school or technical school, junior high school, elementary school or below), and health insurance type (urban employees, urban residents, others, fully self-funded).

Key Variables

The participation in the HMP in this research was assessed using 3 measures. First, "enrollment" in the program was defined as having a hypertensive record, which includes the diagnosis date, blood pressure at the time of diagnosis, lifestyle, specific information about the physicians and institutions providing the services, and patient contact information. This indicates that the patient has visited the CHC and the doctor had the opportunity to interact with them. Enrollment in the HMP is the initial step for doctors when encountering a hypertensive patient for the first time according to the policy. Second, follow-up in this study was defined as patients having at least 2 follow-up records at different times. Third, regular follow-up was defined as patients having an average of more than 4 follow-up visits per year, in accordance with the national policy, which mandates that hypertensive patients should be followed up at least 4 times per year. We used 3 corresponding ratios to evaluate engagement. The enrolled rate refers to the proportion of individuals enrolled out of all patients who have ever been diagnosed with hypertension. The follow-up rate indicates the number of individuals who are followed up out of those enrolled. The adherence rate represents the proportion of individuals who undergo regular follow-up out of the total number of those who are followed up.

Statistical Analysis

Individuals were assessed by household register (internal migrants or local residents). Indicators were estimated for participation in the HMP; for the whole city, this was stratified by sociodemographic factors (resident status, gender, age category, occupation, education level, and insurance type). The city was also stratified by district and community. Categorical variables and prevalence rates were expressed as absolute frequencies and percentages. Missing values were categorized as 1 group for each variable. Logistic regression was used to evaluate the main characteristics associated with participation in the HMP among all patients and internal migrants and local residents as a sensitivity analysis. A 2-sample *z*-test for proportions was used to evaluate whether the observed differences in proportions between internal migrants and local residents were statistically significant. The map figure was drawn using ArcMap 10.8 (Esri) and the other figures were drawn using OriginPro (version 2020b; OriginLab). Statistical analyses were conducted using R Studio (version 4.2.2; Posit).

Ethical Considerations

The data used in this study were anonymized and encrypted, ensuring no access to personally identifiable information. Our data access was authorized through an agreement with government departments. In addition, our study is 1 component of a larger study that was approved by The Tsinghua University Science and Technology Ethics Committee (20230065).

Results

Demographic Characteristics

The total number of diagnosed individuals recorded in the electronic health record was 1,160,214, with 344,526 local residents (29.7%) and 815,688 internal migrants (70.3%). In the overall population, 58.24% (n=675,692) were male and 41.76% (n=484,522) were female. More demographic characteristics of the diagnosed population are detailed in Table 1. Males constituted a higher proportion across all age groups and resident categories, with a noticeable decline in older age groups. Higher educational attainment was more prevalent in younger age groups. Internal migrants had higher proportions of individuals with lower educational levels (elementary school or below). Insurance coverage varied, with urban employee insurance being more common in younger and middle-aged groups, while urban resident insurance was more prevalent in the oldest age group.

XSL•FO
RenderX

Table . Characteristics of residents with hypertension from the electronic health record in Shenzhen by age group (January 1, 2017, to September 30, 2023).

	Total (n=1,160,214), n (%)				Local population (n=344,526), n (%)				Migrant population (n=815,688), n (%)			
Characteristics	18 - 45 years (n=279,179)	46 - 65 years (n=609,238)	>65 years (n=271,797)	Subtotal	18 - 45 years (n=72,361)	46 - 65 years (n=175,887)	>65 years (n=96,278)	Subtotal	18 - 45 years (n=206,818)	46 - 65 years (n=433,351)	>65 years (n=175,519)	Subtotal
Sex												
Male	195,225 (69.9)	356,461 (58.5)	124,006 (45.6)	675,692 (58.2)	47,125 (65.1)	102,826 (58.5)	46,050 (47.8)	196,001 (56.9)	148,100 (71.6)	253,635 (58.5)	77,956 (44.4)	479,691 (58.8)
Female	83,954 (30.1)	252,777 (41.5)	147,791 (54.4)	484,522 (41.8)	25,236 (34.9)	73,061 (41.5)	50,228 (52.2)	148,525 (43.1)	587,18 (28.4)	179,716 (41.5)	97,563 (55.6)	335,997 (41.2)
Marital status												
Single	62,445 (22.4)	35,202 (5.8)	11,072 (4.1)	108,719 (9.4)	18,165 (25.1)	16,762 (9.5)	7141 (7.4)	42,068 (12.2)	44280 (21.4)	18,440 (4.3)	3931 (2.2)	66,651 (8.2)
Married	213,769 (76.6)	567,735 (93.2)	249,868 (91.9)	1,031,372 (88.9)	53,405 (73.8)	156,995 (89.3)	84,909 (88.2)	295,309 (85.7)	160,364 (77.5)	410,740 (94.8)	164,959 (94.0)	736,063 (90.2)
Widowed or divorced	1860 (0.7)	4738 (0.8)	10,293 (3.8)	16,891 (1.5)	358 (0.5)	1553 (0.9)	3935 (4.1)	5846 (1.7)	1502 (0.7)	3185 (0.7)	6358 (3.6)	11,045 (1.4)
Missing	1105 (0.4)	1563 (0.3)	564 (0.2)	3,232 (0.3)	433 (0.6)	577(0.3)	293 (0.3)	1303 (0.4)	672 (0.3)	986 (0.2)	271 (0.2)	1929 (0.2)
Occupation												
Technical personnel or office staff	58,423 (20.9)	86,191 (14.1)	34,917 (12.8)	179,531 (15.5)	25,601 (35.4)	49,862 (28.3)	21,889 (22.7)	97,352 (28.3)	32,822 (15.9)	36,329 (8.4)	13,028 (7.4)	82,179 (10.1)
Business service industry	68,022 (24.4)	134,895 (22.1)	20,419 (7.5)	223,336 (19.2)	18,089 (25.0)	41,752 (23.7)	11,077 (11.5)	70,918 (20.6)	49,933 (24.1)	93,143 (21.5)	9342 (5.3)	152,418 (18.7)
Agriculture and forestry	2617 (0.9)	13,573 (2.2)	23,599 (8.7)	39,789 (3.4)	432 (0.6)	2035 (1.2)	4213 (4.4)	6680 (1.9)	2185 (1.1)	11,538 (2.7)	19,386 (11.0)	33,109 (4.1)
Industry	70,811 (25.4)	114,431 (18.8)	9548 (3.5)	194,790 (16.8)	3024 (4.2)	6877 (3.9)	3988 (4.1)	13,889 (4.0)	67,787 (32.8)	107,554 (24.8)	5560 (3.2)	180,901 (22.2)
Other	37,624 (13.5)	78,876 (12.9)	26,787 (9.9)	143,287 (12.4)	16,709 (23.1)	38,393 (21.8)	16,096 (16.7)	71,198 (20.7)	20,915 (10.1)	40,483 (9.3)	10,691 (6.1)	72,089 (8.8)
Unemployed	41,271 (14.8)	180,768 (29.7)	156,430 (57.6)	378,469 (32.6)	8463 (11.7)	36,932 (21.0)	38,987 (40.5)	84,382 (24.5)	32,808 (15.9)	143,836 (33.2)	117,443 (66.9)	294,087 (36.1)
Missing	411 (0.1)	504 (0.1)	97 (0.0)	1012 (0.1)	43 (0.1)	36 (0.0)	28 (0.0)	107 (0.0)	368 (0.2)	468 (0.1)	69 (0.0)	905 (0.1)
Education level												
Bachelor's degree or above	42,576 (15.3)	43,995 (7.2)	10,930 (4.0)	97,501 (8.4)	25,497 (35.2)	33,534 (19.1)	7483 (7.8)	66,514 (19.3)	17,079 (8.3)	10,461 (2.4)	3447 (2.0)	30,987 (3.8)
College diploma	46,343 (16.6)	51,669 (8.5)	15,667 (5.8)	113,679 (9.8)	17,230 (23.8)	31,359 (17.8)	9634 (10.0)	58,223 (16.9)	29,113 (14.1)	20,310 (4.7)	6033 (3.4)	55,456 (6.8)

XSL•FO
RenderX

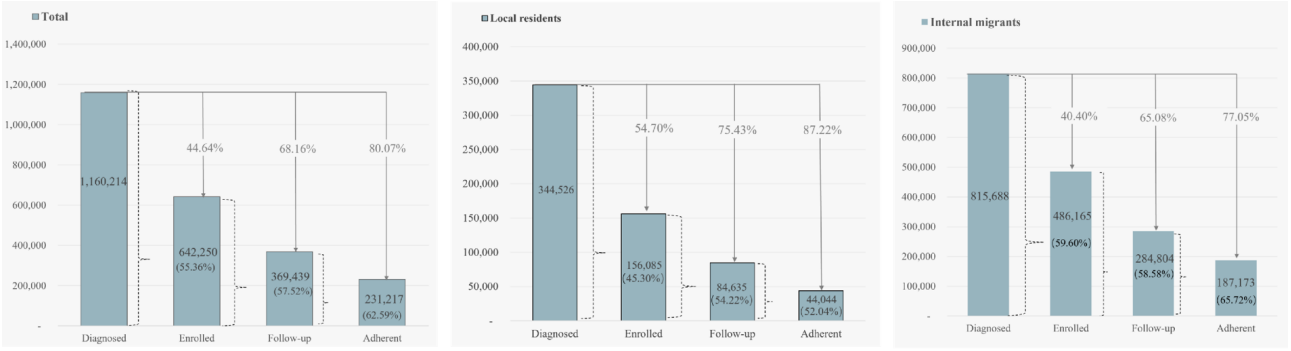
	Total (n=1,160,214), n (%)				Local population (n=344,526), n (%)				Migrant population (n=815,688), n (%)			
High school or technical school	84,781 (30.4)	164,361 (27.0)	52,107 (19.2)	301,249 (26.0)	13,239 (18.3)	53,872 (30.6)	24,805 (25.8)	91,916 (26.7)	71,542 (34.6)	110,489 (25.5)	27,302 (15.6)	209,333 (25.7)
Junior high school	64,310 (23.0)	194,959 (32.0)	65,548 (24.1)	324,817 (28.0)	3196 (4.4)	23,497 (13.4)	20,121 (20.9)	46,814 (13.6)	61,114 (29.5)	171,462 (39.6)	45,427 (25.9)	278,003 (34.1)
Elementary school or below	5205 (1.9)	91,713 (15.1)	108,233 (39.8)	205,151 (17.7)	372 (0.5)	7699 (4.4)	22,928 (23.8)	30,999 (9.0)	4833 (2.3)	84014 (19.4)	85,305 (48.6)	174,152 (21.4)
Missing	35,964 (12.9)	62,541 (10.3)	19,312 (7.1)	117,817 (10.2)	12,827 (17.7)	25,926 (14.7)	11,307 (11.7)	50,060 (14.5)	23,137 (11.2)	36,615 (8.4)	8005 (4.6)	67,757 (8.3)
Insurance type												
Urban employees	167,374 (60.0)	277,068 (45.5)	33,538 (12.3)	477,980 (41.2)	43,049 (59.5)	96,562 (54.9)	27,879 (29.0)	167,490 (48.6)	124,325 (60.1)	180,506 (41.7)	5659 (3.2)	310,490 (38.1)
Urban residents	27,268 (9.8)	74,175 (12.2)	40,742 (15.0)	142,185 (12.3)	11,510 (15.9)	44,050 (25.0)	35,570 (36.9)	91,130 (26.5)	15,758 (7.6)	30,125 (7.0)	5172 (2.9)	51,055 (6.3)
Other	65,806 (23.6)	222,946 (36.6)	184,410 (67.8)	473,162 (40.8)	8336 (11.5)	17,126 (9.7)	24,302 (25.2)	49,764 (14.4)	57,470 (27.8)	205,820 (47.5)	160,108 (91,.2)	423,398 (51.9)
Fully self-funded	1085 (0.4)	2681 (0.4)	1406 (0.5)	5172 (0.4)	458 (0.6)	749 (0.4)	308 (0.3)	1515 (0.4)	627 (0.3)	1932 (0.4)	1098 (0.6)	3657 (0.5)
Missing	17,646 (6.3)	32,368 (5.3)	11,701 (4.3)	61,715 (5.3)	9008 (12.4)	17,400 (9.9)	8219 (8.5)	34,627 (10.1)	8638 (4.2)	14,968 (3.5)	3482 (2.0)	27,088 (3.3)

City-Level HMP Participants

In Total

Among all diagnosed individuals, 55.36% (n=642,250) were enrolled in the HMP. We found that 57.52% (n=369,429) of those followed up. Furthermore, 62.59% (n=231,217) of the individuals who were followed up demonstrated adherence. Compared to local residents, internal migrants have a significantly higher participation rate (P<.001) in the HMP, as shown in Figure 1 (Table S1 in Multimedia Appendix 1).

Figure 1. City-level hypertension management program participants by household registration.



Enrollment Rate

Among all diagnosed individuals (n=1,160,214), 58.1% (n=642,250) were enrolled in the HMP. As is shown in Table 2, the enrollment rates were similar between different sexes, with males at 55.27% (n=373,452) and females at 55.48% (n=268,798). For internal residents, the enrollment rate was higher for females, with a rate of 60.15% (202,115/335,997) for females and 59.22% (284,050/479,691) for males. The opposite was seen for local migrants. Industrial workers had the highest enrollment rate, followed by those in agriculture and forestry, then the unemployed, with the lowest rate among those in the “other” category, and technical and administrative personnel just above that. For the local population, the business service industry was above the “other” group. The lower the educational level, the higher the enrollment rate. Older individuals had higher enrollment rates; the enrollment rate was only 39.27% (n=109,623) among those aged 18 - 44 years, but it increased to 64.14% (n=174,335) among those older than 65 years. Divorced and widowed individuals represented the

XSL•FO
RenderX

smallest proportion of the diagnosed population, but they had the highest enrollment rate at 67.52% (n=11,405), followed by married individuals at 59.16% (n=610,208), with single individuals having the lowest rate at 18.41% (n=20,018). The trends in enrolled rates across different age groups, marital statuses, and educational levels were similar between local residents and internal migrants. The logistic regression results indicated that all the research variables examined were significant factors influencing the enrollment rate (Table S2 in Multimedia Appendix 2).

XSL•FO
RenderX

Table . Participation rates in hypertension management programs by household registration in Shenzhen (January 1, 2017, to September 30, 2023).

Indicator	Enrolled individuals, n (%)[a]			Follow-up individuals, n (%)[a]			Adherent individuals, n (%)[a]		
Characteristics	Total (n=642,250)	Local residents (n=156,085)	Internal migrants (n=486,165)	Total (n=369,439)	Local residents (n=84,635)	Internal migrants (n=284,804)	Total (n=231,217)	Local residents (n=44,044)	Internal migrants (n=187,173)
Gender									
Male	373,452 (55.27)	89,402 (45.61)	284,050 (59.22)	220,066 (58.93)	48,848 (54.64)	171,218 (60.28)	140,515 (63.85)	26,207 (53.65)	114,308 (66.76)
Female	268,798 (55.48)	66,683 (44.9)	202,115 (60.15)	149,373 (55.57)	35,787 (53.67)	113,586 (56.2)	90,702 (60.72)	17,837 (49.84)	72,865 (64.15)
Age group (years)									
18 - 44	109,623 (39.27)	18,782 (25.96)	90,841 (43.92)	65,685 (59.92)	9565 (50.93)	56,120 (61.78)	45,436 (69.17)	5950 (62.21)	39,486 (70.36)
45 - 64	358,292 (58.81)	83,580 (47.52)	274,712 (63.39)	206,746 (57.7)	45,131 (54)	161,615 (58.83)	127,871 (61.85)	23,454 (51.97)	104,417 (64.61)
≥65	174,335 (64.14)	53,723 (55.8)	120,612 (68.72)	97,008 (55.64)	29,939 (55.73)	67,069 (55.61)	57,910 (59.7)	14,640 (48.9)	43,270 (64.52)
Marital status									
Single	20,018 (18.41)	5240 (12.46)	14,778 (22.17)	10,057 (50.24)	2428 (46.34)	7629 (51.62)	7166 (71.25)	1451 (59.76)	5715 (74.91)
Married	610,208 (59.16)	147,202 (49.85)	463,006 (62.90)	352,829 (57.82)	80,168 (54.46)	272,661 (58.89)	220,041 (62.36)	41,532 (51.81)	178,509 (65.47)
Widowed or divorced	11,405 (67.52)	3412 (58.36)	7993 (72.37)	6426 (56.34)	2001 (58.65)	4425 (55.36)	3923 (61.05)	1039 (51.92)	2884 (65.18)
Missing	619 (19.15)	231 (17.73)	388 (20.11)	127 (20.52)	38 (16.45)	89 (22.94)	87 (68.5)	22 (57.89)	65 (73.03)
Occupation									
Technical personnel or office staff	91,870 (51.17)	47,737 (49.04)	44,133 (53.7)	52,918 (57.6)	26,918 (56.39)	26,000 (58.91)	29,624 (55.98)	13,365 (49.65)	16,259 (62.53)
Business service industry	122,120 (54.68)	34,758 (49.01)	87,362 (57.32)	70,627 (57.83)	18,956 (54.54)	51,671 (59.15)	43,445 (61.51)	9998 (52.74)	33,447 (64.73)
Agriculture and forestry	24,092 (60.55)	3577 (53.55)	20,515 (61.96)	13,895 (57.67)	2015 (56.33)	11,880 (57.91)	8382 (60.32)	1043 (51.76)	7339 (61.78)
Industry	134,668 (69.13)	7855 (56.56)	126,813 (70.1)	88,760 (65.91)	4568 (58.15)	84,192 (66.39)	58,049 (65.4)	2405 (52.65)	55,644 (66.09)
Other	44,554 (31.09)	18,754 (26.34)	25,800 (35.79)	21,457 (48.16)	9216 (49.14)	12,241 (47.45)	12,915 (60.19)	4779 (51.86)	8136 (66.47)
Unemployed	224,840 (59.41)	43,397 (51.43)	181,443 (61.7)	121,778 (54.16)	22,962 (52.91)	98,816 (54.46)	78,798 (64.71)	12,454 (54.24)	66,344 (67.14)
Missing	106 (10.47)	7 (6.54)	99 (10.94)	4 (3.77)	0 (0)	4 (4.04)	4 (100)	0 (0)	4 (100)
Education level									
Bachelor's degree or above	40,803 (41.85)	28,613 (43.02)	12,190 (39.34)	21,729 (53.25)	15,248 (53.29)	6481 (53.17)	11,926 (54.89)	7798 (51.14)	4128 (63.69)
College diploma	53,614 (47.16)	27,111 (46.56)	26,503 (47.79)	29,547 (55.11)	14,640 (54)	14,907 (56.25)	17,369 (58.78)	7610 (51.98)	9759 (65.47)
High school or technical school	169,218 (56.17)	47,939 (52.16)	121,279 (57.94)	98,505 (58.21)	26,429 (55.13)	72,076 (59.43)	59,936 (60.85)	13,514 (51.13)	46,422 (64.41)

XSL•FO
RenderX

Indicator	Enrolled individuals, n (%)[a]			Follow-up individuals, n (%)[a]			Adherent individuals, n (%)[a]		
Characteristics	Total (n=642,250)	Local residents (n=156,085)	Internal migrants (n=486,165)	Total (n=369,439)	Local residents (n=84,635)	Internal migrants (n=284,804)	Total (n=231,217)	Local residents (n=44,044)	Internal migrants (n=187,173)
Junior high school	212,807 (65.52)	26,355 (56.3)	186,452 (67.07)	127,916 (60.11)	14,595 (55.38)	113,321 (60.78)	82,575 (64.55)	7776 (53.28)	74,799 (66.01)
Elementary school or below	144,562 (70.47)	18,481 (59.62)	126,081 (72.4)	83,165 (57.53)	10,490 (56.76)	72,675 (57.64)	54,244 (65.22)	5547 (52.88)	48,697 (67.01)
Missing	21,246 (18.03)	7586 (15.15)	13,660 (20.16)	8577 (40.37)	3233 (42.62)	5344 (39.12)	5167 (60.24)	1799 (55.64)	3368 (63.02)
Insurance type									
Urban employees	270,312 (56.55)	80,523 (48.08)	189,789 (61.13)	166,048 (61.43)	44,582 (55.37)	121,466 (64)	99,579 (59.97)	22,840 (51.23)	76,739 (63.18)
Urban residents	81,890 (57.59)	51,233 (56.22)	30,657 (60.05)	46,235 (56.46)	28,688 (56)	17,547 (57.24)	25,561 (55.28)	14,726 (51.33)	10,835 (61.75)
Fully self-funded	280,172 (59.21)	20,301 (40.79)	259,871 (61.38)	153,701 (54.86)	9797 (48.26)	143,904 (55.38)	103,928 (67.62)	5586 (57.02)	98,342 (68.34)
Other	2519 (48.70)	456 (30.1)	2063 (56.41)	1352 (53.67)	217 (47.59)	1135 (55.02)	878 (64.94)	130 (59.91)	748 (65.9)
Missing	7357 (11.92)	3572 (10.32)	3785 (13.97)	2103 (28.59)	1351 (37.82)	752 (19.87)	1271 (60.44)	762 (56.4)	509 (67.69)

[a]All percentages are calculated in relation to a broader population. For instance, the proportion of males within the enrolled individuals stands at 55.27% and is calculated as n/N × 100. Here, "n" represents the number of males within the enrolled individuals subset, while "N" denotes the number of males within the entire population. For N values, please refer to the corresponding "subtotal" columns in Table 1.

Follow-Up Rate

Of patients who have been enrolled in the hypertension program, about 57.52% (369,439/642,250) were followed up. As is shown in Table 2, the follow-up rate was higher among internal migrants (284,804/486,165, 58.58%), while it was 54.22% (84,635/156,085) among the local population. The proportions of those who follow up were higher among men (male: 220,066/373,452, 58.93%; female: 149,373/268,798, 55.57%) and industrial workers (88,760/134,668, 65.91%). For other characteristics, there were significant differences in follow-up rates between local residents and internal migrants. For the local population, follow-up rates increased with age: 50.93% (9565/18,782) for those aged 18‐44 years and 55.73% (29,939/53,723) for those over 65 years. However, for internal migrants, the trend was the opposite: the follow-up rate for those aged 18‐44 years was 61.78% (56,120/90,841), while it was 55.61% (67,069/120,612) for those over 65 years. Among local residents, the highest follow-up rate was observed among divorced and widowed individuals, at 58.65% (2001/3412). Conversely, for internal migrants, married individuals had the highest follow-up rate, at 58.89% (272,661/463,006). Single individuals exhibited the lowest follow-up rates in both populations, with 46.34% (2428/5240) of local residents and 51.62% (7629/14,778) of internal migrants following up. Among local residents, the highest follow-up rate was observed in individuals with an educational level of elementary school or below and those with urban resident insurance. Conversely, among internal migrants, the highest follow-up rate was found in individuals with a junior high school education and those with urban employee insurance. The logistic regression results indicated that all the research variables examined were significant factors influencing the follow-up rate (Table S2 in Multimedia Appendix 2).

Adherence Rate

A total of 62.59% (231,217/369,439) of patients who followed up continued to do so at least 4 times per year. As is shown in Table 2, similar to the enrolled rate and follow-up rate, the adherence rate was higher among internal migrants (187,173/284,804, 65.72%) compared to local residents (44,044/84,635, 52.04%). Among the followed-up patients, men (140,515/220,066, 63.85%), single individuals (7166/10,057, 71.25%), and people aged 18‐44 years (45,436/65,685, 69.17%) had the highest adherence rates. Individuals with urban employee and urban resident insurance did not have the highest adherence rates. These trends were consistent between both local and internal migrants. However, among local residents, those engaged in commercial and industrial occupations had the highest adherence rates, at 52.74% (9998/18,956) and 52.65% (2405/4568), respectively. Conversely, in the internal migrants group, the unemployed and those classified under other categories had the highest adherence rates, at 67.14% (66,344/98,816) and 66.47% (8136/12,241), respectively. Regarding educational level, among local residents, those with a junior high school level of education had the highest adherence rate at 53.28% (7776/14,595), whereas among internal migrants, those with an educational level of elementary school or below had the highest adherence rate. The logistic regression results indicated that all the research variables examined were significant factors influencing the adherence rate (Table S2 in Multimedia Appendix 2).

XSL•FO
RenderX

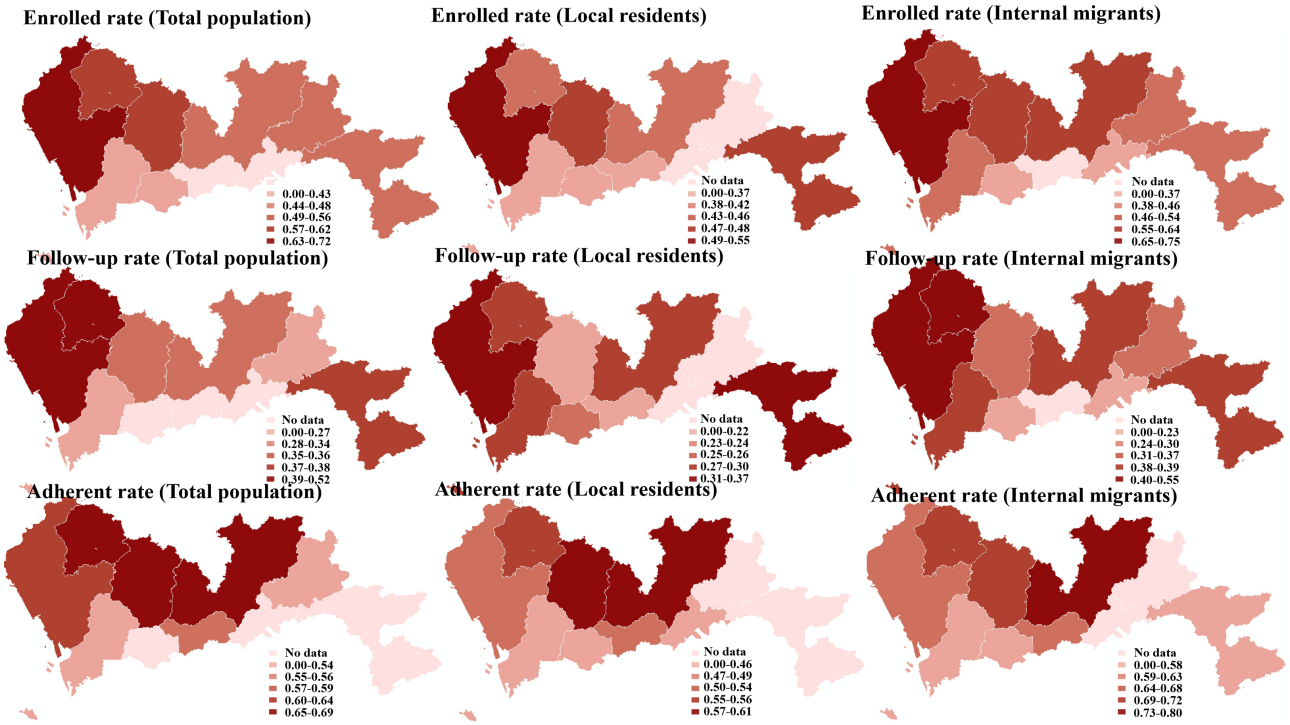
District-Level Hypertension Management

As shown in Figure 2, the highest enrolled rate was detected in Baoan district (180,301/250,036, 72.11%), Guangming district (33,308/54,063, 61.61%), and Longhua district (74,115/124,878, 59.35%), while the lowest enrolled rate was found in Luohu district (52,413/132,023, 39.70%), Yantian district (8609/19,924, 43.21%), and Futian district (59,962/132,337, 45.31%). It is worth noting that, apart from Luohu district, the enrolled rate of internal migrants was higher than that of local residents.

Regarding follow up rate, Baoan district, Guangming district, and Dapeng New Area showed the highest rates, which were 52.21% (131,171/251,237), 45.86% (24,915/54,328), and 36.72% (4301/11,713), while the lowest rate was found in the same places as the enrolled rate. The enrolled rate of internal migrants was higher than that of local residents in all districts.

Considering the adherence rate, Longhua district (31,779/45,917, 69.21%), Longgang district (54,088/79,965, 67.64%), and Guangming district (17,949/26,658, 67.33%) were ranked as the first 3. Similarly, the adherence rate of internal migrants was higher than that of local residents in all districts.

Figure 2. District-level hypertension management.

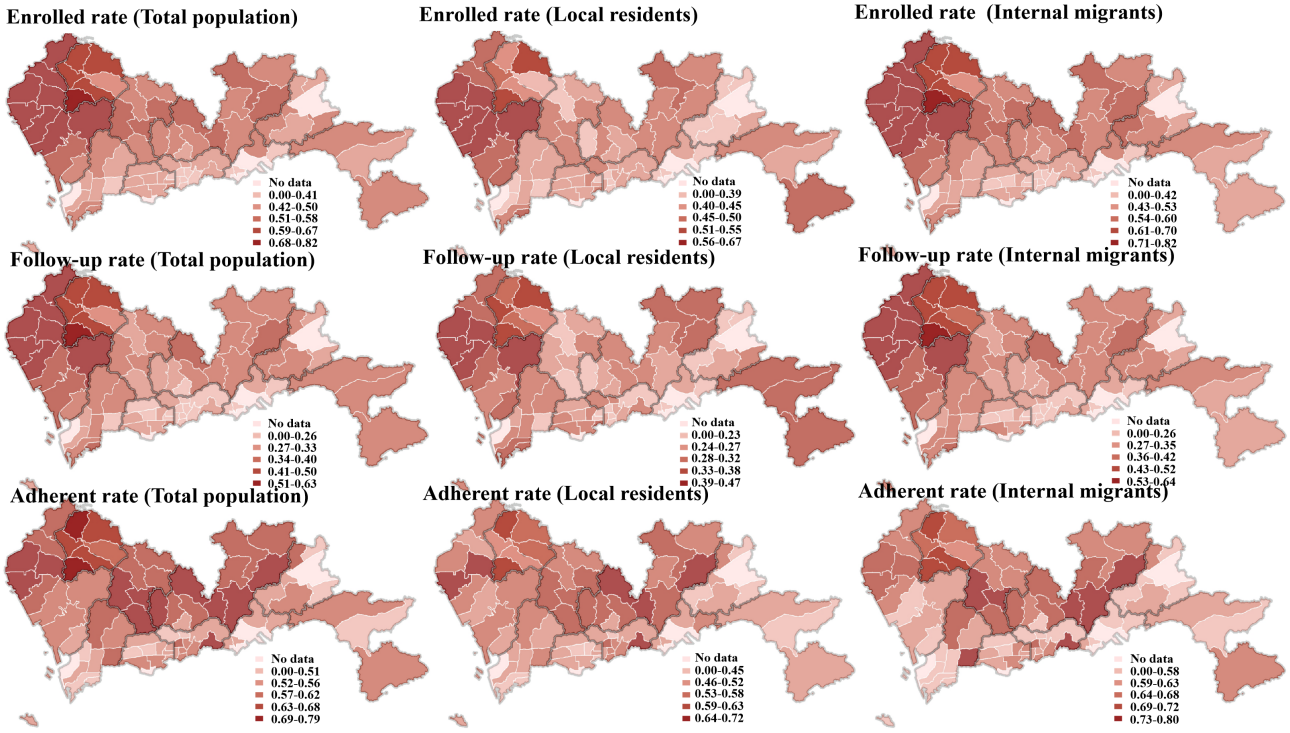


Community-Level Hypertension Management

There was considerable within-city and within-district variation in the community level of hypertension management. Generally speaking, the community-level enrolled rate, follow-up rate, and adherence rate were descending from east to west and from north to south. Although only the adherence rate was slightly different, some communities in the southern region showed a high rate, as shown in Figure 3.

XSL•FO
RenderX

Figure 3. Community-level hypertension management.



There also exists within-district variation such that a 24.30% gap between different communities was found in the enrollment rate of Guangming district. A 23.59% gap was detected in the follow-up rate of Baoan district. Similarly, a 23.59% gap was found in the adherence rate in Luohu district. In addition, rates of internal migrants were much higher than those of local residents in all districts.

Discussion

Principal Results

Regarding the crucial role of health service supply from primary health centers in achieving health equity in the domain of hypertension prevention, we summarized the enrolled rate, follow-up rate, and adherence rate of the Shenzhen population that participated in the HMP based on the data collected from CHCs. Among all diagnosed individuals, 55.36% (642,250/1,160,214) were enrolled in the HMP. Of those, 57.52% (369,439/642,250) followed up, and 62.59% (231,217/369,439) were adherents. Furthermore, we compared the health inequity in hypertension between internal migrants and local residents in Shenzhen.

Comparison With Prior Work

Overall, the data revealed that men exhibited higher rates of both follow-up and adherence compared to women, which is not consistent with the conclusion of some studies [19,20] related to hypertension treatment adherence. However, the medication adherence rates of men and women taking one medication were 61.04% and 55.86% in Beijing, respectively [21]. Despite the fact that the majority of individuals over the age of 65 years were enrolled in the program, this age group demonstrated the lowest rates of follow-up and adherence, although the rates were relatively high in Beijing and Shanghai [22]. Conversely, individuals aged between 18 and 44 years, who have the lowest enrollment rate, showed the highest rates of follow-up and adherence. It is not similar to common research results in which the oldest patients are the most adherent [19,20]. This phenomenon may be related to the fact that older individuals often have more complex health conditions, which CHCs cannot address well at present, leading them to generally prefer seeking treatment at higher-level medical institutions [23,24]. The observed low initial engagement of younger individuals with primary health care facilities may be attributed to a general mistrust or lower acceptance of these services among this demographic; the result is similar to that of a study conducted in Hong Kong [25]. Despite the low proportion of young patients who initially seek care at primary health centers following a hypertension diagnosis, those who do engage exhibit the highest follow-up and regular follow-up rates across all age groups. This indicates a high level of acceptance and satisfaction with primary health care management among younger patients once they have accessed these services.

Regarding marital status, the enrollment rate among widowed and divorced individuals was the highest, exceeding that of single individuals by more than 3 times. It is possible that this group, having experienced significant life events, became more concerned about their health. However, without a partner to provide encouragement and support [26], they may find it difficult to maintain consistent health behaviors and this group had higher mortality [27]. As a result, despite their high enrollment rates, their adherence tended to be relatively low. Interestingly, however, single individuals had the highest adherence rate. Perhaps this is because single individuals who enrolled in the HMP tend to pay more attention to their own health and are more likely to take action to keep healthy, leading to higher adherence; this is similar to a finding in the United States that never-married individuals reported better health habits and had higher rates of protective factors compared to

XSL•FO
RenderX

married or divorced/separated individuals [28]. Previous studies showed that, in general, blue-collar employees appeared less likely to participate in health promotion programs compared to white-collar workers [29], and they may not have enough time to participate in the program [30,31]. Conversely, in our research, industrial workers stood out with the highest rates of enrollment, follow-up, and adherence across all categories; this may be related to the fact that in Shenzhen, many primary health care centers operate until 9 PM, providing this group with sufficient time to access the programs. Furthermore, the data indicated that, regardless of age, individuals with lower educational levels tended to have higher rates for all 3 metrics. Previous research has also found that individuals with higher education levels may be less likely to participate in cancer screening or achieve favorable screening outcomes [32], suggesting the impact of education on health behaviors is complex.

Importantly, internal migrants demonstrated a significantly higher participation in the HMP, especially in terms of project inclusion and regular follow-up. The participation rate of internal migrants was more than 10% higher than that of local residents. This result is inconsistent with previous global and national studies on the health service utilization of internal migrants. Previous research has indicated that internal migrants face long-standing disadvantages due to various institutional, social, cultural, and other exclusionary factors [33]. These disadvantages often extend to health and health care access, limiting their access to health resources [34,35]. A review focusing on migrants' and refugees' health status and health care in Europe highlighted the issue of underuse of primary health care services among these populations [36]. A study [37] based on the 2017 National Migrants Dynamic Monitoring Special Survey in China revealed that, compared to the registered population, the floating population is at a disadvantage in utilizing basic public health services. Most studies [38,39] focusing on hypertension management in China showed that the rate of hypertension management at primary health care centers among the floating population was less than 30% and as low as 2.3% in Hunan province [40], but our study showed that this situation does not exist in Shenzhen. In fact, among the hypertensive population in Shenzhen, on the contrary, internal migrants are better able to utilize primary health care services than locals. The Shenzhen municipal government has invested significantly in ensuring that all residents can access health care services within a 15-minute radius, making it highly convenient for the entire population. Additionally, the city provides free blood pressure screening at CHCs for all individuals aged ≥18 years, which promotes accessibility and equity in hypertension management services [41]. Certain districts in Shenzhen supply free medication to hypertensive patients enrolled in the program [42]. For most internal migrants, free health care services hold strong appeal due to their economic constraints. Some internal migrants came to Shenzhen to help their son/daughter look after children [43], so they or their son/daughter may care more about their health and they usually have enough time to go to a CHC. Insurance is considered an important factor for the utilization of health services for internal migrants and immediate reimbursement of medical insurance can significantly increase inpatient utilization, promoting health [44]. The statistical model showed that insurance type is an important factor influencing the 3 hypertension management rates studied. However, there were notable differences between local residents and internal migrants. For local residents, those with urban resident insurance were more likely to participate in the HMP compared to those with urban employee insurance. In contrast, for internal migrants, those with other insurance types were less likely to participate compared to those with urban employee insurance. It takes less money to get treatment and medicine in CHCs compared with upper health care facilities, especially for people who do not have native health insurance or cannot get immediate reimbursement.

Our analysis revealed several intriguing trends regarding HMP engagement among different demographic groups in Shenzhen. Previous studies conducted in Beijing and Shanghai have also found variations in hypertension management across different demographic factors, including age, gender, education level, workload, marital status, residential area, and annual household income [45,46]. Among individuals aged 18 - 44 years, the enrolled rate of internal migrants was nearly 18% higher than that of local residents. This discrepancy may be attributed to the comparatively better economic conditions of local residents, who do not face the housing issues often encountered by migrants [47,48]. They may just go to upper hospitals when they think necessary, while internal migrants are more likely to utilize CHCs as the first choice for their routine health problems [49,50]. Additionally, the presence of their parents may lead locals to prioritize their own health less than migrants, who often lack familial support [51] and thus might be more vigilant about their health. Interestingly, in the divorced or widowed demographic, the follow-up rate among local residents surpassed that of internal migrants. Moreover, among local residents, the follow-up rate for this group was higher than that of married individuals. This finding contrasts with conventional research, which typically suggests that marriage promotes greater health awareness and engagement with health care services while the effects of bereavement usually lead to less health care utilization [52]. For married migrant individuals, having a spouse provides them with companionship [53] and significant support from their partner, whereas local residents can receive support from other relatives. This may cause internal migrants, compared to local residents, to have a more limited support system, making the protective effects of marriage on their health more pronounced. In contrast, local residents who are divorced or widowed may have access to a broader range of social support. The adherence rate showed an increasing divergence between local residents and internal migrants as age groups progress. Internal migrants tended to exhibit greater concern for their health after relocating to a new environment, which may explain this trend. Despite the fact that, among those with an undergraduate education or higher, the enrolled rate and follow-up rate for local residents were higher than those for internal migrants, the adherence rate for local residents remained lower than that of internal migrants. This suggests that, while highly educated local residents may initially engage with health care services, they may not maintain the same level of adherence as their migrant counterparts.

XSL•FO
RenderX

Some research showed that there exists vast disparities in health insurance coverage and health service utilization between floating populations and local residents; most of the floating population has no timely access to primary or some other kinds of health care service [39], which cause the poor control of chronic diseases. In our study, it appears that health insurance played a less significant role in the disparity of participation in the HMP between local residents and internal migrants.

From a district and community level, hypertensive patients were experiencing health inequity in hypertension treatment and prevention. The potential reasons included population, economic reasons, and policy differences. The populations of Longgang, Longhua, and Nanshan in Shenzhen are large and concentrated; in addition, the health care services are mature and systematic. Hence, the HMP is better than in other regions such as Yantian and Pingshan districts. Similarly, the wealthy regions had higher enrolled, follow-up, and adherence rates compared to the relatively poor regions, namely Yantian and Pingshan districts. These finding were consistent with previous research. A large national survey during 2004‐2018 found a provincial difference in awareness, treatment, and control is partly correlated with per capita gross domestic product [54]. Richer provinces are more successful in controlling hypertension. Globally, the proportions of hypertension awareness, treatment, and blood pressure control are low [55], particularly in low- and middle-income countries, and few comprehensive assessments of the economic impact of hypertension exist [56,57]. In terms of policy differences, the hypertensive patients who participate in the HMPs of CHCs could get free prescription medicines, blood pressure measurement services, and blood glucose measurement services once per quarter. The policy is different in different districts in Shenzhen, which may have caused the observed difference in enrolled, follow-up, and adherence rate. Furthermore, when it comes to health equity between internal migrants and native citizens in Shenzhen, the difference in enrolled, follow-up, and adherence rates within internal migrants was higher than that within local residents. This may be caused by different health care service supply situations toward internal migrants within different districts and communities [58]. The same problems were found in Ireland, Portugal, Spain, and the United States [59].

Limitations and Strengths

This study represents a pioneering effort to examine the participation of internal migrants and local residents in HMPs, thereby highlighting the equity of internal migrants' utilization of primary health care services under current Chinese hypertension management policies. By focusing on the specific involvement of diagnosed hypertension patients in HMPs, this research offers valuable insights into the accessibility and fairness of these services for internal migrants. The data utilized in this study were derived entirely from the primary health care system, ensuring its authenticity and reliability. However, this study has several limitations. It exclusively addressed the participation of hypertension patients in primary HMPs, which means that it did not account for patients who may receive regular treatment at other medical institutions without ever visiting CHCs. Consequently, the findings may not fully represent the hypertension management engagement of all patients citywide. Furthermore, the study was centered on participation rates within the HMP and did not capture the ultimate blood pressure control outcomes of the subjects. As a result, it is not possible to ascertain the differences in blood pressure control outcomes that may arise from varying levels of management participation.

Future research should consider integrating data from multiple health care providers and tracking long-term health outcomes to provide a more comprehensive understanding of hypertension management across diverse patient populations.

Conclusions

Based on our research, of the 1,160,214 hypertensive patients, 642,250 were enrolled in the HMP, 48,848 were included in the followed up population, and 34,186 were adherents. Notably, internal migrants demonstrated a significantly higher participation in the HMP. Furthermore, populations with different demographic and socioeconomic characteristics and in different regions showed different rates. Given these findings, there is the potential to enhance the outreach and engagement of native, younger, and single populations through targeted promotional strategies.

Acknowledgments

This project was supported by the Sanming Project of Medicine in Shenzhen (grant SZSM202111001). We would like to express our gratitude to the Shenzhen Government and the Shenzhen Health Development Research and Data Management Center for their invaluable support in providing data, funding, and an enabling environment for this project. We acknowledge the use of generative ChatGPT for grammatical checks.

Data Availability

The datasets generated during and/or analyzed during this study are not publicly available due to restrictions imposed by the local government. The data originate from the medical system and are managed by the local government. Access to the encrypted data stored on the cloud platform is limited to research institutions that have established formal collaborations with the local government.

XSL•FO
RenderX

Authors' Contributions

WL and NZ are the corresponding authors of this study, responsible for designing, guiding, organizing, and planning this study. JH was responsible for data analysis and manuscript writing. YL was responsible for data cleaning and provided helpful comments on several drafts of the manuscript. HZ, LW, HS, and ZZ provided advice and revised several drafts of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Two-proportion *z*-test for enrollment rates, follow-up rates, and adherence rates between local residents and internal migrants.
[XLSX File, 20 KB - publichealth_v11i1e65548_app1.xlsx]

Multimedia Appendix 2
Multivariable logistic regression analysis of the enrollment, follow up, and adherence of different groups in Shenzhen.
[XLSX File, 22 KB - publichealth_v11i1e65548_app2.xlsx]

References

1. Kyu HH, Abate D, Abate KH. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1859-1922. [doi: 10.1016/S0140-6736(18)32335-3] [Medline: 30415748]
2. Global report on hypertension: the race against a silent killer. : World Health Organization; 2023 URL: https://www.who.int/publications/i/item/9789240081062 [accessed 2024-12-16]
3. Feng XL, Pang M, Beard J. Health system strengthening and hypertension awareness, treatment and control: data from the China Health and Retirement Longitudinal Study. Bull World Health Organ 2014 Jan 1;92(1):29-41. [doi: 10.2471/BLT.13.124495] [Medline: 24391298]
4. Yin R, Yin L, Li L, et al. Hypertension in China: burdens, guidelines and policy responses: a state-of-the-art review. J Hum Hypertens 2022 Feb;36(2):126-134. [doi: 10.1038/s41371-021-00570-z] [Medline: 34215840]
5. Notice of the National Health and Family Planning Commission on the issuance of the "National Basic Public Health Service Standards (Third Edition)" [Website in Chinese]. National Health Commission of the PRC. 2017. URL: http://www.nhc.gov.cn/jws/s3578/201703/d20c37e23e1f4c7db7b8e25f34473e1b.shtml [accessed 2024-12-16]
6. Dover DC, Belon AP. The health equity measurement framework: a comprehensive model to measure social inequities in health. Int J Equity Health 2019 Feb 19;18(1):36. [doi: 10.1186/s12939-019-0935-0] [Medline: 30782161]
7. Zhang J, Lin S, Liang D, Qian Y, Zhang D, Hou Z. Public health services utilization and its determinants among internal migrants in China: evidence from a nationally representative survey. Int J Environ Res Public Health 2017 Sep 1;14(9):1002. [doi: 10.3390/ijerph14091002] [Medline: 28862682]
8. Lin Y, Zhang Q, Chen W, Ling L. The social income inequality, social integration and health status of internal migrants in China. Int J Equity Health 2017 Aug 4;16(1):139. [doi: 10.1186/s12939-017-0640-9] [Medline: 28778201]
9. Wang J, Tan F, Wang Z, et al. Understanding gaps in the hypertension and diabetes care cascade: systematic scoping review. JMIR Public Health Surveill 2024 Feb 16;10:e51802. [doi: 10.2196/51802] [Medline: 38149840]
10. Chaturvedi A, Zhu A, Gadela NV, Prabhakaran D, Jafar TH. Social determinants of health and disparities in hypertension and cardiovascular diseases. Hypertension 2024 Mar;81(3):387-399. [doi: 10.1161/HYPERTENSIONAHA.123.21354] [Medline: 38152897]
11. Doyle SK, Chang AM, Levy P, Rising KL. Achieving health equity in hypertension management through addressing the social determinants of health. Curr Hypertens Rep 2019 Jun 12;21(8):58. [doi: 10.1007/s11906-019-0962-7] [Medline: 31190099]
12. Jin Y, Jing M, Zhang L, Song S, Ma X. Internet access and hypertension management among the elderly population: a nationally representative cross-sectional survey in China. J Med Internet Res 2019 Jan 31;21(1):e11280. [doi: 10.2196/11280] [Medline: 30702439]
13. Zhang Z, Shi G, Xing Y, et al. Examining the potential impacts of intensive blood pressure treatment on the socioeconomic inequity in hypertension prevalence in China: a nationally representative cross-sectional study. Hypertens Res 2023 Dec;46(12):2746-2753. [doi: 10.1038/s41440-023-01441-5] [Medline: 37789112]
14. Liu Y, Liu N, Cheng M, et al. The changes in socioeconomic inequalities and inequities in health services utilization among patients with hypertension in Pearl River Delta of China, 2015 and 2019. BMC Public Health 2021 May 12;21(1):903. [doi: 10.1186/s12889-021-10879-6] [Medline: 33980187]
15. Cao D, Zhou Z, Si Y, et al. Prevalence and income-related equity in hypertension in rural China from 1991 to 2011: differences between self-reported and tested measures. BMC Health Serv Res 2019 Jul 1;19(1):437. [doi: 10.1186/s12913-019-4289-5] [Medline: 31262290]

XSL•FO
RenderX

16. Wang C. Local governance in China: realizing the citizens' rights of migrants. China Popul Dev Stud 2018 Mar;2(1):11-31. [doi: 10.1007/s42379-018-0002-3]
17. Who needs to get the flu vaccine? Do you need to get it again if you have had it before? [Website in Chinese]. Public Hygiene and Health Commission of Shenzhen Municipality. 2022. URL: https://wjw.sz.gov.cn/gzcy/ywzs/ggws/content/post_10268931.html [accessed 2024-12-16]
18. Free medical check-ups for Shenzhen hukou residents aged 60 or above [Website in Chinese]. Shenzhen News Network. 2020. URL: https://www.sznews.com/news/content/2020-06/06/content_23223870.htm [accessed 2024-12-16]
19. Khanam MA, Lindeboom W, Koehlmoos TLP, Alam DS, Niessen L, Milton AH. Hypertension: adherence to treatment in rural Bangladesh--findings from a population-based study. Glob Health Action 2014;7:25028. [doi: 10.3402/gha.v7.25028] [Medline: 25361723]
20. Gavrilova A, Bandere D, Rutkovska I, et al. Knowledge about disease, medication therapy, and related medication adherence levels among patients with hypertension. Medicina (Kaunas) 2019 Oct 28;55(11):715. [doi: 10.3390/medicina55110715] [Medline: 31661904]
21. Lan L, Hai P, Luo J, Li R, Wang Y. Medical behaviours and medication adherence of older hypertensive patients in different medical insurance programs in Beijing, China: a cross-sectional study. BMC Geriatr 2023 Dec 20;23(1):878. [doi: 10.1186/s12877-023-04476-y] [Medline: 38124122]
22. Yin M, Augustin B, Fu Z, Yan M, Fu A, Yin P. Geographic distributions in hypertension diagnosis, measurement, prevalence, awareness, treatment and control rates among middle-aged and older adults in China. Sci Rep 2016 Nov 14;6:37020. [doi: 10.1038/srep37020] [Medline: 27841326]
23. Yili Z, Xinyi H, Baoyu Q, Chuanrui S, Xiaoyang W, Xu W, et al. Multimorbidity in elderly population: challenges and coping strategies. Chin Gen Pract 2022;25(35):4363-4368. [doi: 10.12114/j.issn.1007-9572.2022.0514]
24. Na D, Jia-tang Z, Peng-fei W, Yuxia K, Xia T. Chronic diseases and comorbidity in the elderly in Beijing. Chin Gen Pract 2018;21(3):265-268. [doi: 10.3969/j.issn.1007-9572.2017.00.217]
25. Lim MK, Ha SCN, Luk KH, Yip WK, Tsang CSH, Wong MCS. Update on the Hong Kong Reference Framework for Hypertension Care for Adults in Primary Care Settings-review of evidence on the definition of high blood pressure and goal of therapy. Hong Kong Med J 2019 Feb;25(1):64-67. [doi: 10.12809/hkmj187701] [Medline: 30648830]
26. Scafato E, Galluzzo L, Gandin C, et al. Marital and cohabitation status as predictors of mortality: a 10-year follow-up of an Italian elderly cohort. Soc Sci Med 2008 Nov;67(9):1456-1464. [doi: 10.1016/j.socscimed.2008.06.026] [Medline: 18675500]
27. Dhindsa DS, Khambhati J, Schultz WM, Tahhan AS, Quyyumi AA. Marital status and outcomes in patients with cardiovascular disease. Trends Cardiovasc Med 2020 May;30(4):215-220. [doi: 10.1016/j.tcm.2019.05.012] [Medline: 31204239]
28. Kaplan RM, Kronick RG. Marital status and longevity in the United States population. J Epidemiol Community Health 2006 Sep;60(9):760-765. [doi: 10.1136/jech.2005.037606] [Medline: 16905719]
29. Glasgow RE, McCaul KD, Fisher KJ. Participation in worksite health promotion: a critique of the literature and recommendations for future practice. Health Educ Q 1993;20(3):391-408. [doi: 10.1177/109019819302000309] [Medline: 8307762]
30. Sorensen G, Linnan L, Hunt MK. Worksite-based research and initiatives to increase fruit and vegetable consumption. Prev Med 2004 Sep;39 Suppl 2:S94-100. [doi: 10.1016/j.ypmed.2003.12.020] [Medline: 15313078]
31. Wandel M, Roos G. Work, food and physical activity. A qualitative study of coping strategies among men in three occupations. Appetite 2005 Feb;44(1):93-102. [doi: 10.1016/j.appet.2004.08.002] [Medline: 15604036]
32. Korn AR, Walsh-Bailey C, Correa-Mendez M, et al. Social determinants of health and US cancer screening interventions: a systematic review. CA Cancer J Clin 2023;73(5):461-479. [doi: 10.3322/caac.21801] [Medline: 37329257]
33. Zhang J, Ye Y, Wang J. Social exclusion and psychopathologies of Chinese floating populations. J Affect Disord Rep 2021 Dec;6:100263. [doi: 10.1016/j.jadr.2021.100263]
34. Boardman J. Social exclusion and mental health – how people with mental health problems are disadvantaged: an overview. Ment Health Soc Incl 2011 Aug 15;15(3):112-121. [doi: 10.1108/20428301111165690]
35. Tang S, Long C, Wang R, Liu Q, Feng D, Feng Z. Improving the utilization of essential public health services by Chinese elderly migrants: strategies and policy implication. J Glob Health 2020 Jun;10(1):010807. [doi: 10.7189/jogh.10.010807] [Medline: 32257170]
36. Lebano A, Hamed S, Bradby H, et al. Migrants' and refugees' health status and healthcare in Europe: a scoping literature review. BMC Public Health 2020 Jun 30;20(1):1039. [doi: 10.1186/s12889-020-08749-8] [Medline: 32605605]
37. Song Y, Zhang G. Utilization of public health services and its influencing factors among migrant people with hypertension or diabetes in China. Chin J Public Health 2021;37(2):198-202. [doi: 10.11847/zgggws1123492]
38. Kui JZ. Analysis of hypertension morbidity among floating population of Sichuan in 2012. Chin J Prev Control Chronic Dis 2015;23(1):27-30. [doi: CNKI:SUN:ZMXB.0.2015-01-009]
39. Su L, Sun L, Xu L. Review on the prevalence, risk factors and disease management of hypertension among floating population in China during 1990-2016. Glob Health Res Policy 2018;3:24. [doi: 10.1186/s41256-018-0076-9] [Medline: 30123839]

XSL•FO
RenderX

40. Donghui J, Qiaohua X, Huilin L, et al. Management of hypertension and diabetes mellitus among floating population in Hunan. Pract Prev Med 2015;22(10):1211-1212. [doi: 10.3969/j.issn.1006-3110.2015.010.017]
41. Promote health management and prevent and control chronic diseases [Website in Chinese]. Public Hygiene and Health Commission of Shenzhen Municipality. 2024. URL: https://wjw.sz.gov.cn/ztzl/szsjmjkbps2022/jkszjs/content/post_11294466.html [accessed 2024-12-16]
42. Get free high blood pressure medicine! Come to these 12 community health centers in Shenzhen and get it before the end of the year [Website in Chinese]. Public Hygiene and Health Commission of Shenzhen Municipality. 2019. URL: https://wjw.sz.gov.cn/wzx/content/post_3196923.html [accessed 2024-12-16]
43. Yong T, Yipin W. The quality of life of the floating elderly and the establishment of social support network: a survey of the floating elderly in Shenzhen. Mgmt Admin 2014(4):132-134. [doi: 10.16517/j.cnki.cn12-1034/f.2014.04.011]
44. He W. Does the immediate reimbursement of medical insurance reduce the socioeconomic inequality in health among the floating population? Evidence from China. Int J Equity Health 2023 May 17;22(1):96. [doi: 10.1186/s12939-023-01913-7] [Medline: 37198632]
45. Jiang B, Liu H, Ru X, Zhang H, Wu S, Wang W. Hypertension detection, management, control and associated factors among residents accessing community health services in Beijing. Sci Rep 2014 May 2;4:4845. [doi: 10.1038/srep04845] [Medline: 24784167]
46. Zhang X, Zhang Y, Xiao X, Ma X, He J. The relation between health insurance and management of hypertension in Shanghai, China: a cross-sectional study. BMC Public Health 2016 Sep 10;16(1):959. [doi: 10.1186/s12889-016-3627-3] [Medline: 27613593]
47. Huang Y, Tao R. Housing migrants in Chinese cities: current status and policy design. Environ Plann C Gov Policy 2015 Jun;33(3):640-660. [doi: 10.1068/c12120]
48. Wang Y. Housing cost burden, homeownership, and self-rated health among migrant workers in Chinese cities: the confounding effect of residence duration. Cities 2023 Feb;133:104128. [doi: 10.1016/j.cities.2022.104128] [Medline: 37425220]
49. Li H, Chung RYN, Wei X, et al. Comparison of perceived quality amongst migrant and local patients using primary health care delivered by community health centres in Shenzhen, China. BMC Fam Pract 2014 Apr 29;15:76. [doi: 10.1186/1471-2296-15-76] [Medline: 24779564]
50. Wei X, Pearson S, Zhang Z, Qin J, Gerein N, Walley J. Comparing knowledge and use of health services of migrants from rural and urban areas in Kunming City, China. J Biosoc Sci 2010 Nov;42(6):743-756. [doi: 10.1017/S0021932010000313] [Medline: 20584352]
51. Bradby H, Liabo K, Ingold A, Roberts H. Visibility, resilience, vulnerability in young migrants. Health (Lond) 2019 Sep;23(5):533-550. [doi: 10.1177/1363459317739441] [Medline: 29090633]
52. Simeonova E. Marriage, bereavement and mortality: the role of health care utilization. J Health Econ 2013 Jan;32(1):33-50. [doi: 10.1016/j.jhealeco.2012.10.010] [Medline: 23202255]
53. Yang Y, Chen B, Huang P, Wang Y, Zhang L, Cai F. Prevalence and influencing factors of depressive symptoms among rural-to-urban migrant workers in China: a systematic review and meta-analysis. J Affect Disord 2022 Jun;307:11-19. [doi: 10.1016/j.jad.2022.03.061]
54. Wang W, Zhang M, Xu CD, et al. Hypertension prevalence, awareness, treatment, and control and their associated socioeconomic factors in China: a spatial analysis of a national representative survey. Biomed Environ Sci 2021 Dec 20;34(12):937-951. [doi: 10.3967/bes2021.130] [Medline: 34981717]
55. Mills KT, Stefanescu A, He J. The global epidemiology of hypertension. Nat Rev Nephrol 2020 Apr;16(4):223-237. [doi: 10.1038/s41581-019-0244-2] [Medline: 32024986]
56. Nguyen TN, Chow CK. Global and national high blood pressure burden and control. Lancet 2021 Sep 11;398(10304):932-933. [doi: 10.1016/S0140-6736(21)01688-3] [Medline: 34450082]
57. Kario K, Okura A, Hoshide S, Mogi M. The WHO Global report 2023 on hypertension warning the emerging hypertension burden in globe and its treatment strategy. Hypertens Res 2024 May;47(5):1099-1102. [doi: 10.1038/s41440-024-01622-w]
58. Zhou B, Carrillo-Larco RM, Danaei G. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021 Sep 11;398(10304):957-980. [doi: 10.1016/S0140-6736(21)01330-1] [Medline: 34450083]
59. Ledoux C, Pilot E, Diaz E, Krafft T. Migrants' access to healthcare services within the European Union: a content analysis of policy documents in Ireland, Portugal and Spain. Glob Health 2018 Jun 15;14(1):57. [doi: 10.1186/s12992-018-0373-6] [Medline: 29903036]

Abbreviations

CHC: Community Health Center
HMP: hypertension management program

XSL•FO
RenderX

Edited by A Mavragani; submitted 19.08.24; peer-reviewed by L Zhang, O Olasunkanmi; revised version received 22.10.24; accepted 29.10.24; published 10.02.25.

Please cite as:
He J, Li Y, Zeng H, Sun H, Wu L, Zhu Z, Zhang N, Liang W
Health Equity of Hypertension Management Between Local Residents and Internal Migrants in Shenzhen, China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e65548
URL: https://publichealth.jmir.org/2025/1/e65548
doi:10.2196/65548

© Jinyu He, Yanjun Li, Huatang Zeng, Haoran Sun, Liqun Wu, Zhenzhen Zhu, Ning Zhang, Wannian Liang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Accessibility of eHealth Before and During the COVID-19 Pandemic Among People With and People Without Impairment: Repeated Cross-Sectional Survey

Linda Pettersson[1,2,3], MD; Stefan Johansson[4,5], PhD; Ingrid Demmelmaier[2], PhD; Lena von Koch[6,7], PhD; Jan Gulliksen[4], Prof Dr; Per-Olof Hedvall[5], PhD; Karl Gummesson[1], PhD; Catharina Gustavsson[1,2,8], PhD

[1]Center for Clinical Research Dalarna, Uppsala University, Nissers väg 3, Falun, Sweden

[2]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden

[3]Primary Healthcare Center Mora, Region Dalarna, Mora, Sweden

[4]School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden

[5]Certec, Department of Design Sciences, Lund University, Lund, Sweden

[6]Division of Family Medicine and Primary Care, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[7]Theme Heart & Vascular and Neuro, Karolinska University Hospital, Stockholm, Sweden

[8]School of Health and Welfare, Dalarna University, Falun, Sweden

Corresponding Author:
Linda Pettersson, MD
Center for Clinical Research Dalarna, Uppsala University, Nissers väg 3, Falun, Sweden

Abstract

Background: The adoption of eHealth accelerated during the COVID-19 pandemic. Inequalities in the adoption of eHealth during the COVID-19 pandemic have been reported, but there are few such studies among people with impairment.

Objectives: This study aimed to investigate self-reported use and difficulty in the use of eHealth before the COVID-19 pandemic compared to during late social distancing restrictions in Sweden, among people with and without impairment, as well as between different types of impairment.

Methods: A cross-sectional survey was distributed twice by snowball sampling to people with self-reported impairment and a general population matched by age, gender, and county. Use and difficulty in the use of six eHealth services were compared between groups using chi-square test and logistic regression with year interaction terms, reported as odds ratio adjusted (aOR) for gender and age with 95% CI.

Results: The surveys included 1631 (in 2019) and 1410 (in 2021) participants with impairment, and 1084 (in 2019) and 1223 (in 2021) participants without. Participants with impairment, compared to those without impairment, reported less use and more difficulty in booking health care appointments online, digital identification, and the Swedish national web portal for health information and eHealth services (1177.se), both before and during the pandemic (P=.003 or lower). Video health care appointments were the exception to this disability digital divide in eHealth as video appointment adoption was the most likely among participants with attention, executive, and memory impairments (interaction term aOR 2.10, 95% CI 1.30 - 3.39). Nonuse and difficulty in the use of eHealth were consistently associated with language impairments and intellectual impairments. For example, language impairments were inversely associated with use of the logged-in eHealth services in 1177.se in 2021 (aOR 0.49, 95% CI 0.36 - 0.67) and were associated with difficulty in the use of 1177.se in 2019 (aOR 2.24, 95% CI 1.50 - 3.36) and the logged-in eHealth services in 1177.se in 2021 (aOR 1.89, 95% CI 1.32 - 2.70). Intellectual impairments were inversely associated with the use of the logged-in eHealth services in 1177.se in 2021 (aOR 0.19, 95% CI 0.13 - 0.27).

Conclusions: This repeated cross-sectional survey study, including participants with diverse types of impairment and a control group without impairment, reveals persisting disability digital divides, despite an accelerated adoption of eHealth across the pandemic. eHealth services were not accessible to some groups of people who were identified as being at risk of severe disease during the COVID-19 pandemic. This implies that all people could not use eHealth as a measure of infection protection.

(JMIR Public Health Surveill 2025;11:e64707) doi:10.2196/64707

KEYWORDS

eHealth; impairment; accessibility; digital inclusion; universal design; disability; digital divide; electronic health; COVID-19; pandemic; cross-sectional study; Sweden; online booking; digital identification; web portal; health information; control group; public health; digital health; digital literacy; health informatics; mobile phone

XSL•FO
RenderX

Introduction

Adoption of eHealth accelerated during the COVID-19 pandemic [1]. Before the pandemic, we identified a disability digital divide in eHealth, that is, people with impairment reported less use and more difficulty in the use of eHealth, particularly with some types of impairment (language impairments and intellectual impairments) [2]. Sustainable development [3], legislation, and human rights principles [4] demand accessible eHealth for all. During the pandemic, people with impairment have reported less use of COVID-19 digital services than people without impairment [5], and have reported difficulty in the use of eHealth [6]. Likewise, eHealth websites have failed to pass accessibility assessments [7]. Even though there are multiple studies and even reviews of associations between eHealth and other sociodemographic factors or somatic diseases, there are very few studies of how people with impairment perceive using eHealth. The three survey studies conducted prior to ours did not sample psychiatric or intellectual impairments, and they had no comparison to before the pandemic [5,6,8]. Further, the two of them that investigated the eHealth modality telehealth reported aggregated results of video health care appointments with telephone contacts.

The emergency phase of the COVID-19 pandemic in the spring of 2020 was characterized by uncertainty [9], and exceptional episodes of social distancing restrictions, including rapid digitalization [1]. Reprioritizing was a necessity, in order to provide acute health care in severe cases of COVID-19, triage with health care guidance, and public health assignments. Like the rest of the general population, health care workers were instructed to self-quarantine when having respiratory tract symptoms, which also limited health care availability [10]. To provide health care services in a social distancing manner, health care facilities were reorganized, separating dedicated areas for suspected COVID-19 cases [11] and the population was encouraged to redirect contacts to eHealth instead of face-to-face contacts [1]. Groups at risk for severe COVID-19 were identified and were recommended intensified social distancing restrictions [12], which included people with intellectual impairments, Down's syndrome, schizophrenia, bipolar disorder, and stroke. The Swedish restrictions were reimposed in a second wave from autumn 2020 to summer 2021 and a third wave from January 2022 to September 2022. This was a time of realization and concern about prolonged postpandemic scenarios [9].

This study aimed to investigate self-reported use and difficulty in the use of eHealth before the COVID-19 pandemic compared to during late social distancing restrictions in Sweden among people with and without impairment, as well as between different types of impairment.

Methods

Study Design

This study had a repeated cross-sectional comparative design, using data from the "Swedes with impairment and the internet 2019" (SMFOI19) survey and the "Swedes with impairment and the internet 2021" (SMFOI21) survey, to people with diverse types of impairment and matched controls.

Study Setting

The publicly financed Swedish national web portal for health information and eHealth services (1177.se) is open to the whole population and is the most commonly used eHealth service in Sweden [13]. To access the personal eHealth services in 1177.se, it is required that the user have digital identification. Digital identification is also needed to access other private eHealth services. In the past decade, private eHealth services that provide video health care appointments have been launched [14]. Subsequently, this has encouraged the public health care regions to also provide video health care appointments.

Participants

Participants were people with self-reported impairment and people from the general population matched to the sample of people with impairment by gender, age, and county of residence. Impairment status was self-reported by one questionnaire item with 43 checkboxes of diagnoses and activity limitations and a free-text response option for reporting "Other impairment" (Multimedia Appendix 1). Participants who received the survey as matched controls, but responded that they had impairment, were reallocated to be analyzed as participants with impairment.

Procedures

The first version of this survey was developed in 2017 to mirror two nationwide Swedish surveys [15,16], out of which the latter is connected to Eurostat [17]. In the SMFOI19 survey, questions on eHealth services were added [2], and SMFOI21 also included questions on video health care appointments. SMFOI19 comprised 47 questions and SMFOI21 comprised 43 questions, on information and communication technology, impairments, and background characteristics. The wording of the questionnaires was developed in collaboration with the Begripsam group, all of whom have lived experience of impairment [18]. Several response modalities were provided (web, pen, phone, or on-site interview) and adaptations were provided upon request (sign language, reading support, speech therapist, or pictograms). Snowball sampling was undertaken from June to October 2019 (SMFOI19) and from May to August 2021 (SMFOI21), respectively [19]. The snowball sampling was distributed by disability networks, as well as by the participants themselves. Then, in February 2020 (SMFOI19) and in March 2022 (SMFOI21), the surveys were posted to six randomly selected matched controls per participant with impairment. Addresses were provided by the Swedish state personal address register [20].

Data Collection

This study used SMFOI19 and SMFOI21 questionnaire items on eHealth, impairments, and background characteristics.

Outcomes

Table 1 provides an overview of outcome variables. In summary, outcomes were measured on the use of eHealth services (ie, whether the services had been used and whether avoiding or preferring use) and difficulty in the use of eHealth services (ie, perceived difficulty or ease of use).

XSL•FO
RenderX

Table . Outcome variables measuring "use of eHealth" and "difficulty in the use of eHealth."

Outcome variable followed by response options	In the SMFOI19[a] survey	In the SMFOI21[b] survey
Use of booking health care appointments online		
Checkbox items		
"I use booking medical appointments online" or "I use booking dental appointments online"	✓	✓
Use of video health care appointments		
Multiple-choice item		
"Yes, I used video health care appointments prior to the COVID-19 pandemic" or "Yes, I used video health care appointments during the COVID-19 pandemic"		✓
Comparisons were made respectively with the response option "No, I have not used video health care appointments," whereas "Do not know" was treated as missing value		
Use of private eHealth services (which are services that primarily provide video health care appointments, chat, and drug prescriptions)		
Checkbox items		
Use of five Swedish private eHealth services or "Other private eHealth service"		✓
Use of the Swedish national web portal for health information and eHealth services (1177.se)		
Checkbox items		
"I use the health information website in the 1177.se"		✓
"I use the logged-in eHealth services in the 1177.se"		✓
Use of digital identification		
Checkbox items		
"I use the digital identification app BankID" or "I use other digital identification than BankID"	✓	✓
Avoid booking health care appointments online		
Multiple-choice item		
"If possible, I avoid booking medical appointments online" or "If possible, I avoid booking dental appointments online"	✓	✓
Comparison was made with the response option "I try to book all appointments online," whereas "Not applicable" was treated as missing value.		
Avoid video health care appointments		
Multiple-choice item		
"If possible, I avoid video health care appointments"		✓
Comparison was made with the response option "I try to get all my appointments as video health care appointments," whereas "Not applicable" was treated as missing value.		
Difficulty in the use of the Swedish national web portal for health information and eHealth services (1177.se)		
Multiple-choice items		

XSL•FO
RenderX

Outcome variable followed by response options	In the SMFOI19[a] survey	In the SMFOI21[b] survey
Two items in SMFOI21 "The health information website in 1177.se is difficult to use" or		✓
"The logged-in eHealth services in 1177.se is difficult to use." For comparison between SMFOI19 and SMFOI21, they were combined.		✓
One single item in SMFOI19 "The Swedish national web portal for health information and eHealth services (1177.se) is difficult to use"	✓	
Comparison was made with the response option "It is easy to use," whereas "Not applicable" was treated as missing value.		
Difficulty in the use of digital identification		
Multiple-choice item		
"It is difficult to use the digital identification app Mobile BankID" or "It is difficult to use other digital identification than BankID"	✓	✓
Comparison was made with the response option "It is easy to use," whereas "Not applicable" was treated as missing value.		

[a]SMFOI19: Swedes with impairment and the internet 2019.

[b]SMFOI21: Swedes with impairment and the internet 2021.

Data on the use of eHealth were collected in both SMFOI19 and SMFOI21 on booking health care appointments online and digital identification. In SMFOI21, use of eHealth was also collected on video health care appointments "before the pandemic" and "during the pandemic," private eHealth services, and the Swedish national web portal for health information and eHealth services (1177.se).

Data on the use of eHealth were also collected by questions on whether the respondents "Avoided" or "Preferred" booking health care appointments online in both SMFOI19 and SMFOI21, and video health care appointments in SMFOI21.

Data on difficulties in the use of eHealth were collected in both SMFOI19 and SMFOI21, on the use of the 1177.se and use of digital identification.

Independent Variables

Age and gender were selected to be confounding variables based on empirical studies of known demographic associations with the study outcomes [21,22]. Gender was analyzed as "female" and "male" whereas the response options "Other gender" and "Prefer not to answer" were handled as missing values in the analysis. Age was categorized for analysis into four categories: 16 - 29, 30 - 44, 45 - 69, and 70 years of age and older.

Data on impairment were collected by one question with 43 checkbox multiple response options on activity limitations (eg, difficulties in understanding) and diagnoses (eg, intellectual disability), and a free-text option for "Other impairment" (Multimedia Appendix 1). For the analysis, we applied a conceptual model of purposeful subgrouping of impairments, which we had developed in a previous study [2]. The categorization of types of impairment was based on our competence in human-computer interaction, digital accessibility, and medical science, and with the International Classification of Functioning, Disability, and Health (ICF)-classification as a framework [23], as well as empirical knowledge of comorbidity and functioning [24]. The grouping of impairments was done independently by three of the researchers (LP, SJ, and CG). There was almost complete interrater agreement and consensus in a conceptual model. Our conceptual model comprises 11 types of impairments by merging ICF-Body functions and ICF-Activities, as well as the 7 global and 11 specific ICF-Cognitive functions. The robustness of the conceptual model was assessed in a cluster analysis [25] using SPSS (version 26.0; IBM Corp).

Background Characteristics

Data on participant characteristics were collected to describe the study samples in regard to educational level, native language, occupation, income, urban or rural status, accommodation, cohabitant status, professional support in everyday life, and access to digital devices.

Data Analysis

We compared 11 outcomes between participants with and without impairment, as well as between different types of impairment. Six of the outcomes were compared between the survey years (2021 and 2019; Table 1). The statistical software R (version 4.0.3; R Core Team) was used for the analyses. Pair-wise deletion was applied. The significance level was set to ≤.05. Chi-square test and logistic regression with robust standard errors (Huber-White) adjusted for gender and age were used. Multicollinearity was assessed in relation to a predetermined cut-off. Sensitivity analyses indicated the robustness of the estimates, that is, models were fitted with and without independent variables with wide CIs. Results from the following analyses are visualized in figures and presented as numeric values in Multimedia Appendix 2. Among all participants overall, SMFOI21 was compared to SMFOI19, by

XSL•FO
RenderX

logistic regression models, adjusted for types of impairment, in addition to adjusting for confounders. Comparisons between participants with impairment to those without were made by chi-square test. To identify types of impairment associated with nonuse and difficulty in the use of eHealth, year-stratified confounder-adjusted logistic regression was used. Comparisons between types of impairment in SMFOI21 to SMFOI19 were performed by the input of multiplicative interaction terms of types of impairment with survey year [26].

Ethical Considerations

This study was approved by the Swedish Ethical Review Authority (registration 2022-00184-01). The participants provided informed consent prior to participation. No compensation was provided for participation. This study used anonymized data.

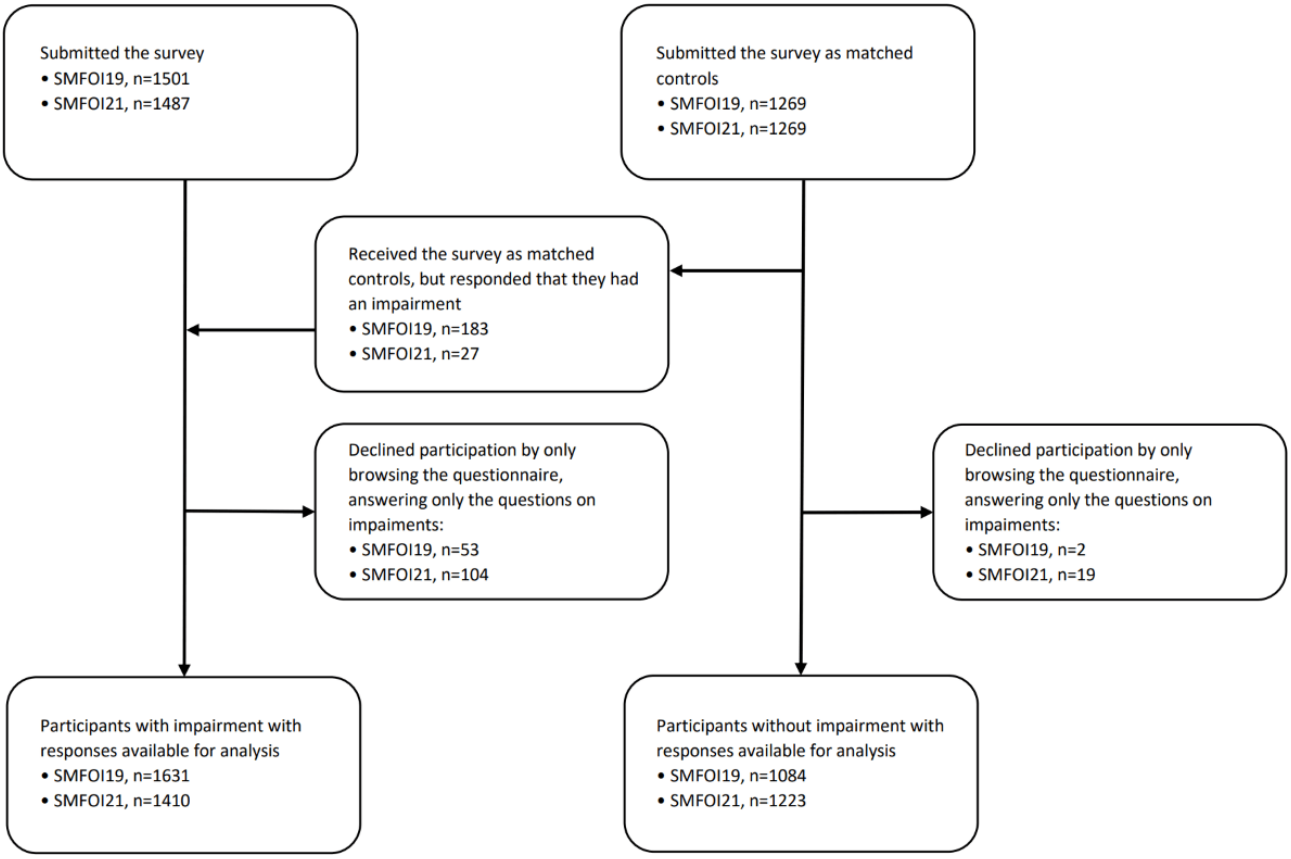
Results

The flowchart of participation (Figure 1) includes 1631 (in SMFOI19) and 1410 (in SMFOI21) participants with impairment, along with 1084 (in SMFOI19) and 1223 (in SMFOI21) participants without impairment.

Compared to participants without impairment, more of the participants with impairment were educated in a special school, were outside the labor market or not working, had lower income, lived in rental apartments or supported accommodation, were living alone, had professional support in everyday life, and lacked access to digital devices (Table 2).

In both survey years, participants with impairment had a median of 2 (IQR 1-6 in SMFOI19 and 1-5 in SMFOI21) impairments. The types of impairment were distributed in approximately similar proportions in the two survey years (Figure 2).

Figure 1. Flowchart of participation in the SMFOI19 and SMFOI21 surveys. SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.



XSL•FO
RenderX

Table . Background characteristics of participants with and without impairment.

	Participants with impairment in the SMFOI19[a] survey (n=1631)	Participants without impairment in the SMFOI19 survey (n=1084)	Participants with impairment in the SMFOI21[b] survey (n=1410)	Participants without impairment in the SMFOI21 survey (n=1223)
Gender, n (%)	n=1356	n=1060	n=1213	n=1193
Women	937 (69)	780 (74)	797 (66)	824 (69)
Men	419 (31)	280 (26)	416 (34)	369 (31)
Age (years), n (%)	n=1388	n=1069	n=1257	n=1200
16 - 29	176 (13)	58 (5)	186 (15)	48 (4)
30 - 44	304 (22)	156 (15)	269 (21)	173 (14)
45 - 69	773 (56)	716 (67)	599 (48)	681 (57)
≥70	135 (10)	139 (13)	203 (16)	298 (25)
Education, n (%)	n=1339	n=1040	n=1222	n=1190
Compulsory school	134 (10)	74 (7)	118 (10)	79 (7)
Special education school	152 (11)	2 (0)	113 (9)	10 (1)
Upper secondary school, Vocational school, or Folk high school (adult education)	567 (42)	442 (43)	488 (40)	447 (38)
University	486 (36)	522 (50)	503 (41)	654 (55)
Native language, n (%)	n=1361	n=1067	n=1253	n=1196
Swedish	1252 (92)	994 (93)	1136 (91)	1065 (89)
Other than Swedish	109 (8)	73 (7)	117 (9)	131 (11)
Occupation, n (%)	n=1367	n=1060	n=1244	n=1194
Retired	273 (17)	304 (28)	327 (26)	478 (40)
Outside of the labor market (disability-related early retirement, daily activity center, or temporary disability allowance)	519 (38)	18 (2)	344 (28)	14 (1)
Working	492 (30)	745 (69)	342 (27)	647 (54)
Student	136 (8)	36 (3)	111 (9)	26 (2)
On the labor market, but not working (unemployed, parental-leave, or sick-leave)	132 (10)	22 (2)	120 (10)	29 (2)
Monthly income (in Swedish Krona, SEK), n (%)	n=1157	n=866	n=1058	n=931
<5000	53 (5)	10 (1)	53 (5)	8 (1)
5000 - 24,999	714 (62)	228 (26)	632 (60)	278 (30)
≥25,000	390 (34)	628 (73)	373 (35)	645 (69)
Urban or rural status, n (%)	n=1368	n=1068	n=1239	n=1188
Rural	191 (14)	176 (16)	169 (14)	168 (14)
Suburban or town	325 (24)	252 (24)	284 (23)	290 (24)
Urban	852 (62)	640 (60)	786 (63)	730 (61)
Accommodation, n (%)	n=1365	n=1059	n=1233	n=1194
Supported accommodation (group living, service apartment, or other supported accommodation)	104 (8)	1 (0)	70 (6)	1 (0)

XSL•FO
RenderX

	Participants with impairment in the SMFOI19[a] survey (n=1631)	Participants without impairment in the SMFOI19 survey (n=1084)	Participants with impairment in the SMFOI21[b] survey (n=1410)	Participants without impairment in the SMFOI21 survey (n=1223)
Group living	52 (4)	1 (0)	38 (3)	0 (0)
Service apartment	43 (3)	0 (0)	23 (2)	1 (0)
Other supported accommodation	9 (1)	0 (0)	9 (1)	0 (0)
Apartment, condominium, house, or homeless	1261 (92)	1058 (100)	1163 (94)	1193 (100)
Apartment	470 (34)	168 (16)	407 (33)	225 (19)
Condominium	285 (21)	239 (22)	282 (23)	267 (22)
House	506 (37)	651 (61)	468 (38)	701 (59)
Homeless	0 (0)	0 (0)	6 (0)	0 (0)
Cohabitant status, n (%)	n=1365	n=1066	n=1249	n=1197
Living alone	549 (40)	183 (17)	487 (39)	231 (19)
Cohabiting	816 (60)	883 (83)	762 (61)	966 (81)
Professional support in everyday life, n (%)	n=1381	n=1063	n=1236	n=1190
Have professional support in everyday life	497 (36)	6 (1)	398 (32)	19 (2)
Home based support by municipal care services	112 (8)	2 (0)	91 (7)	4 (0)
Personal assistants	90 (7)	0 (0)	61 (5)	0 (0)
Supported-living staff, support persons or similar	214 (15)	0 (0)	166 (13)	0 (0)
Appointed guardian	111 (8)	0 (0)	71 (6)	0 (0)
Relative	81 (6)	4 (0)	152 (12)	6 (1)
Other support	54 (4)	4 (0)	55 (4)	7 (1)
No support	884 (64)	1057 (99)	838 (68)	1171 (98)
Access to digital devices, n (%)	n=1456	n=1067	n=1316	n=1213
Lack access to at least computer and one portable device	241 (17)	93 (9)	205 (16)	117 (10)
No device	29 (2)	7 (1)	21 (2)	19 (2)
Only computer at home	75 (5)	24 (2)	66 (5)	22 (2)
Only smartphone	78 (5)	19 (2)	44 (3)	30 (2)
Only tablet	17 (1)	1 (0)	13 (1)	1 (0)
Smartphone and tablet	42 (3)	39 (4)	61 (5)	45 (4)
Have access to at least computer and one portable device	1215 (83)	974 (91)	1111 (84)	1096 (90)
Computer and smartphone	418 (29)	264 (25)	395 (30)	362 (30)
Computer and tablet	65 (4)	25 (2)	60 (5)	19 (2)
Computer, smartphone, and tablet	732 (51)	688 (64)	656 (50)	715 (59)
Number of reported impairments	n=1631	n=1084	n=1410	n=1223

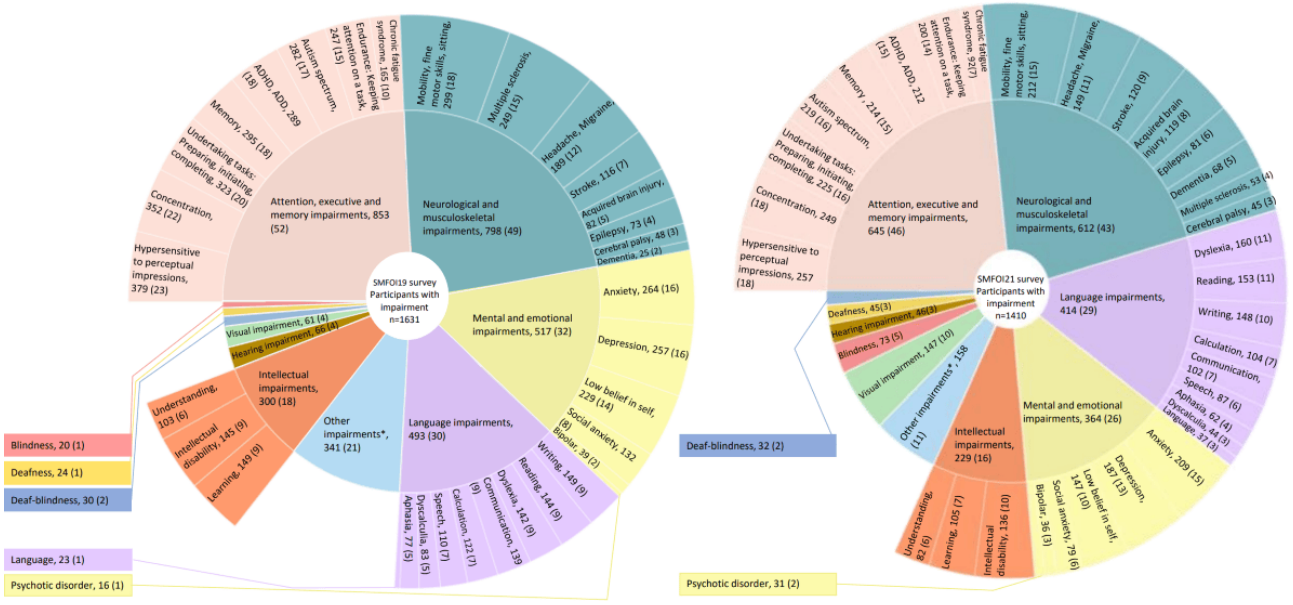
XSL•FO
RenderX

	Participants with impairment in the SMFOI19[a] survey (n=1631)	Participants without impairment in the SMFOI19 survey (n=1084)	Participants with impairment in the SMFOI21[b] survey (n=1410)	Participants without impairment in the SMFOI21 survey (n=1223)
Median (IQR), Maximum value	2 (1-6), 20	0 (0-0), 0	2 (1-5), 35	0 (0-0), 0

[a]SMFOI19: Swedes with impairment and the internet 2019.

[b]SMFOI21: Swedes with impairment and the internet 2021.

Figure 2. Description of participants' types of impairment (subgrouped in the inner circle and disaggregated to the level of the questionnaire in the outer circle). The labels are followed by number and proportion of all participants with impairment, n (%). Adds up to over 100% as multiple responses were allowed. The impairment question construction covered both activity limitations (eg, reading difficulties) and diagnoses (eg, dyslexia; Multimedia Appendix 1). *Free-text "other impairments" were in falling frequency: pain, genetic syndromes, paresis, and posttraumatic stress disorder, followed by infrequent responses. ADD: attention deficit disorder; ADHD: attention-deficit/hyperactivity disorder; SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.



Comparisons of use between the survey years, as well as between participants with and without impairment, are reported in Figures 3a and 4a. Among all participants, all eHealth services were used by more participants in SMFOI21 compared to SMFOI19: booking health care appointments online (adjusted odds ratio [aOR] 2.73, 95% CI 2.39 - 3.10), video health care appointments (aOR 3.34, 95% CI 2.81 - 3.96), and digital identification (aOR 1.74, 95% CI 1.43 - 2.10; Figures 3a and 4a). In both survey years, all but two of the eHealth services were used by fewer participants with impairment compared to participants without impairment: booking health care appointments online (P<.001), the information pages in 1177.se (P<.001), the logged-in eHealth services in 1177.se (P<.001), and digital identification (P<.001; Figures 3a and 4a). The exceptions were video health care appointments, which were used by 9% in both groups (P=.93) before the pandemic, and 29% with impairment versus 19% without impairment (P<.001) during the pandemic, and private eHealth services (P=.10; Figure 3a).

XSL•FO
RenderX

Figure 3. Use of booking health care appointments online, video health care appointments, and private eHealth services. The group with impairment (a) is disaggregated into the subgrouped types of impairment from Figure 2 (b-d). The numeric values visualized in this figure are presented in Tables S1 and S3-S5 in Multimedia Appendix 2. *aOR: adjusted odds ratio from logistic regression comparing use of eHealth among all participants overall between before and during the pandemic (reference group is before the pandemic), adjusted for type of impairment (reference is participants without impairment), gender (reference female), and age (reference <30 years of age); †2-sided *P* value from chi-square test comparing participants with and without impairment. ‡Intercept from year-stratified logistic regression (reference group is participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). §*P* value from logistic regression interaction terms of year with types of impairment (reference is before the pandemic and participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age).

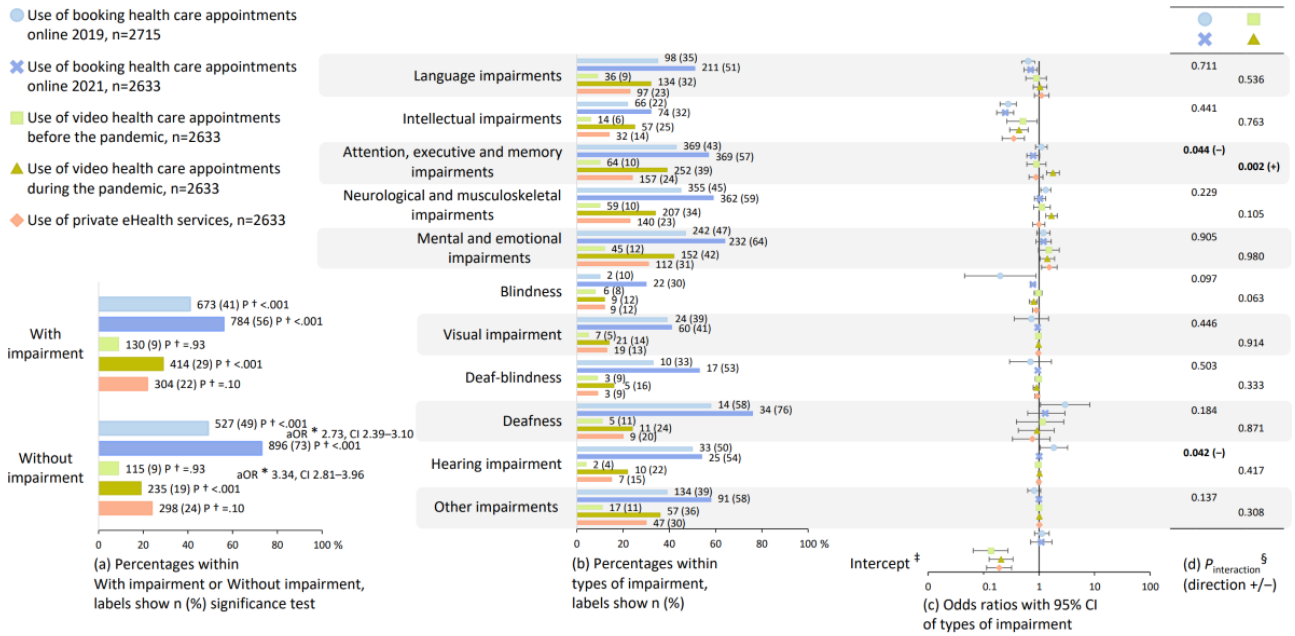
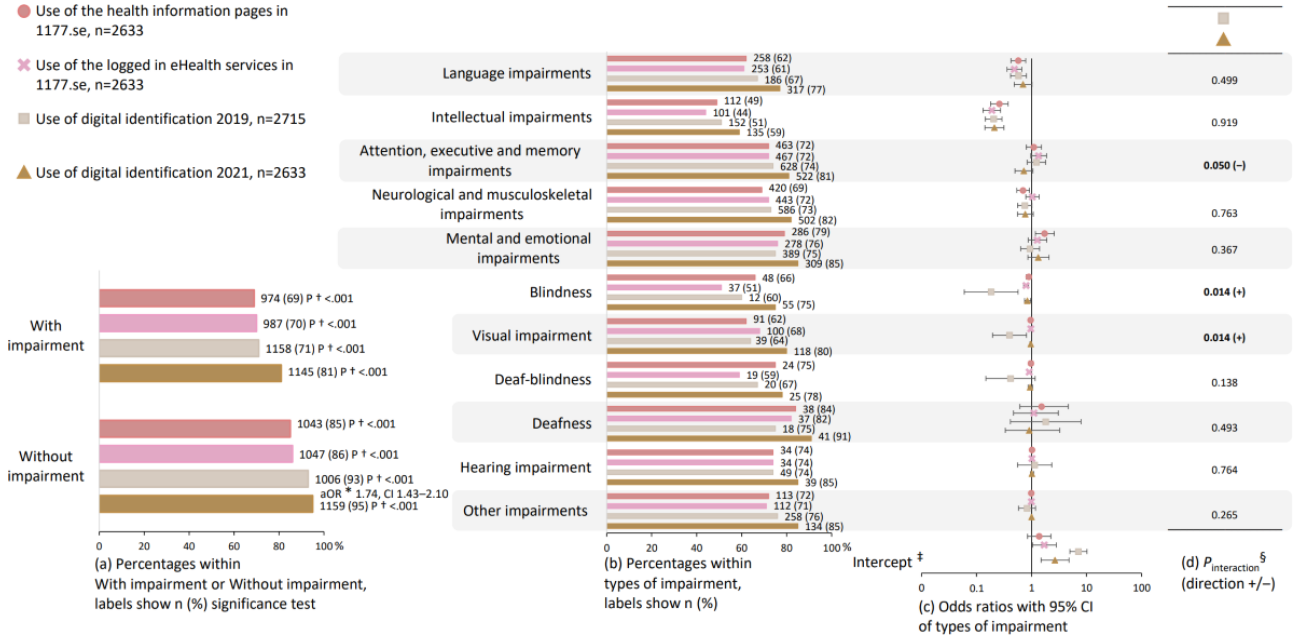


Figure 4. Use of the Swedish national web portal for health information and eHealth services (1177.se) and digital identification. The group with impairment (a) is disaggregated into the subgrouped types of impairment from Figure 2 (b–d). The numeric values visualized in this figure are presented in Tables S1, S6, and S7 in Multimedia Appendix 2. *aOR: adjusted odds ratio from logistic regression comparing use of eHealth among all participants overall between the SMFOI19 and SMFOI21 surveys (reference group is SMFOI19), adjusted for type of impairment (reference is participants without impairment), gender (reference female), and age (reference <30 years of age); †2-sided *P* value from chi-square test comparing participants with and without impairment. ‡Intercept from year-stratified logistic regression (reference group is participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). §*P* value from logistic regression interaction terms of year with types of impairment (reference is SMFOI19 and participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.



Associations between use and types of impairment are reported in Figures 3b, c, 4b, c, and further, changes in the associations between the survey years are reported in Figures 3d and 4d. Booking health care appointments online was inversely associated with language impairments in SMFOI19 (aOR 0.64, 95% CI 0.49 - 0.83) and in SMFOI21 (aOR 0.70, 95% CI

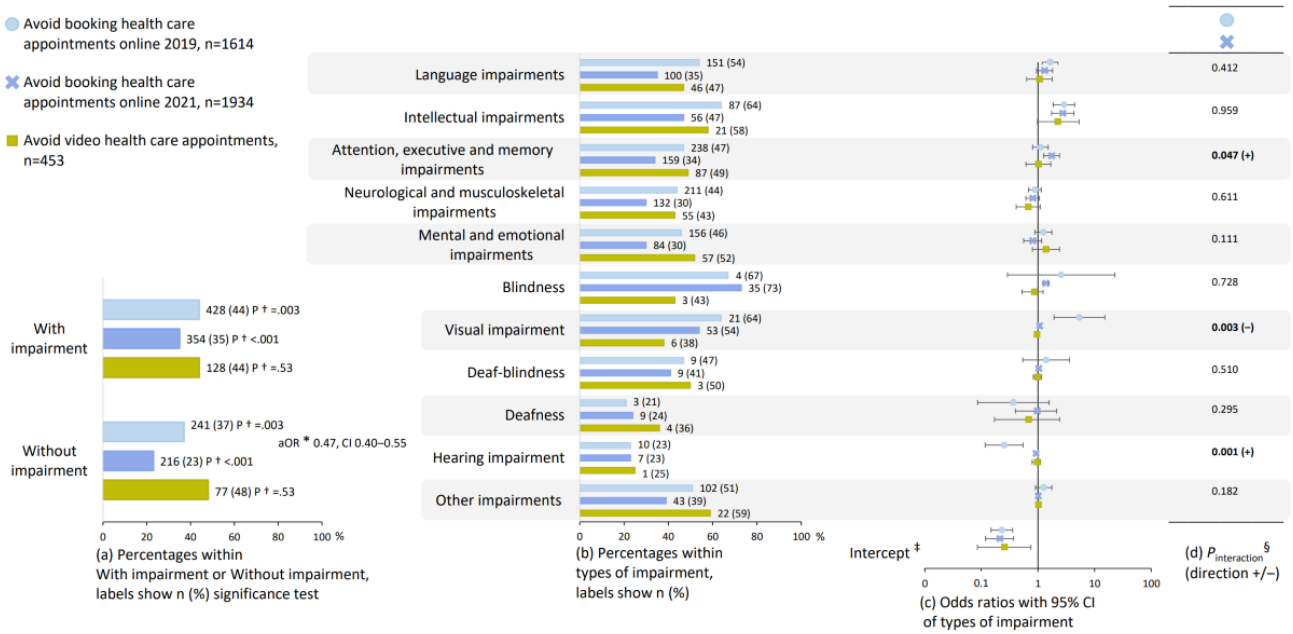
XSL•FO
RenderX

0.54 - 0.93), with intellectual impairments in SMFOI19 (aOR 0.28, 95% CI 0.20 - 0.39) and in SMFOI21 (aOR 0.25, 95% CI 0.17 - 0.35), and with blindness in SMFOI19 (aOR 0.20, 95% CI 0.05 - 0.88) and in SMFOI21 (aOR 0.78, 95% CI 0.70 - 0.86; Figure 3c). The use of video health care appointments was inversely associated with intellectual impairments before the pandemic (aOR 0.51, 95% CI 0.26 - 0.93) and during the pandemic (aOR 0.44, 95% CI 0.30 - 0.64; Figure 3c). Comparing during the pandemic to before, the use of video health care appointments was the most likely to increase among participants with attention, executive, and memory impairments (interaction term aOR 2.10, 95% CI 1.30 - 3.39; Figure 3d). The use of private eHealth services was associated with mental and emotional impairments (aOR 1.53, 95% CI 1.11 - 2.11) and inversely associated with intellectual impairments (aOR 0.35, 95% CI 0.22 - 0.55; Figure 3c). Use of the 1177.se was inversely associated with three types of impairments: language impairments (aOR 0.58, 95% CI 0.43 - 0.78 and aOR 0.49, 95% CI 0.36 - 0.67), intellectual impairments (aOR 0.26, 95% CI 0.18 - 0.38 and aOR 0.19, 95% CI 0.13 - 0.27), and blindness (aOR 0.89, 95% CI 0.80 - 1.00 and aOR 0.80, 95% 0.72 - 0.88; Figure 4c). The use of digital identification was inversely associated with intellectual impairments in SMFOI19 (aOR 0.21, 95% CI 0.15 - 0.29), and in SMFOI21 (aOR 0.21, 95% CI 0.14 - 0.32), and blindness in SMFOI19 (aOR 0.19, 95% CI 0.06 - 0.57) and SMFOI21 (aOR 0.84, 95% CI 0.75 - 0.96; Figure 4c).

Comparisons of avoiding use between the survey years, as well as between participants with and without impairment, are reported in Figure 5a. Among all participants, avoiding booking health care appointments online was less frequent in SMFOI21 compared to SMFOI19 (aOR 0.47, 95% CI 0.40 - 0.55). Booking health care appointments online was avoided by more participants with impairment compared to participants without impairment in SMFOI19 (44% vs 37%; *P*=.003) and SMFOI21 (35% vs 23%; *P*<.001; Figure 5a). Regarding avoiding video health care appointments, no difference was found between participants with and without impairment (44% vs 48%; *P*=.53; Figure 5).

Associations between avoiding use and types of impairment are reported in Figure 5b,c, and further, changes in the associations between the survey years are reported in Figure 5d. Avoiding booking health care appointments online was associated with intellectual impairments in SMFOI19 (aOR 2.88, 95% CI 1.86 - 4.45) and SMFOI21 (aOR 2.74, 95% CI 1.74 - 4.33), and visual impairment in SMFOI19 (aOR 5.40, 95% CI 1.92 - 15.18) and SMFOI21 (aOR 1.07, 95% CI 1.04 - 1.10; Figure 5c). Regarding avoiding video health care appointments, no associations were found with types of impairment (Figure 5c).

Figure 5. Avoid booking health care appointments online and video health care appointments. The group with impairment (a) is disaggregated into the subgrouped types of impairment from Figure 2 (b-d). The numeric values visualized in this figure are presented in Tables S2, S8, and S9 in Multimedia Appendix 2. *aOR: adjusted odds ratio from logistic regression comparing avoidance of eHealth among all participants overall between the SMFOI19 and SMFOI21 surveys (reference group is 2019), adjusted for type of impairment (reference is participants without impairment), gender (reference female), and age (reference <30 years of age); †2-sided *P* value from chi-square test comparing participants with and without impairment. ‡Intercept from year-stratified logistic regression (reference group is participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). §*P* value from logistic regression interaction terms of year with types of impairment (reference is SMFOI19 and participants without impairment), adjusted for gender (reference female), and age (reference <30 years of age). SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.

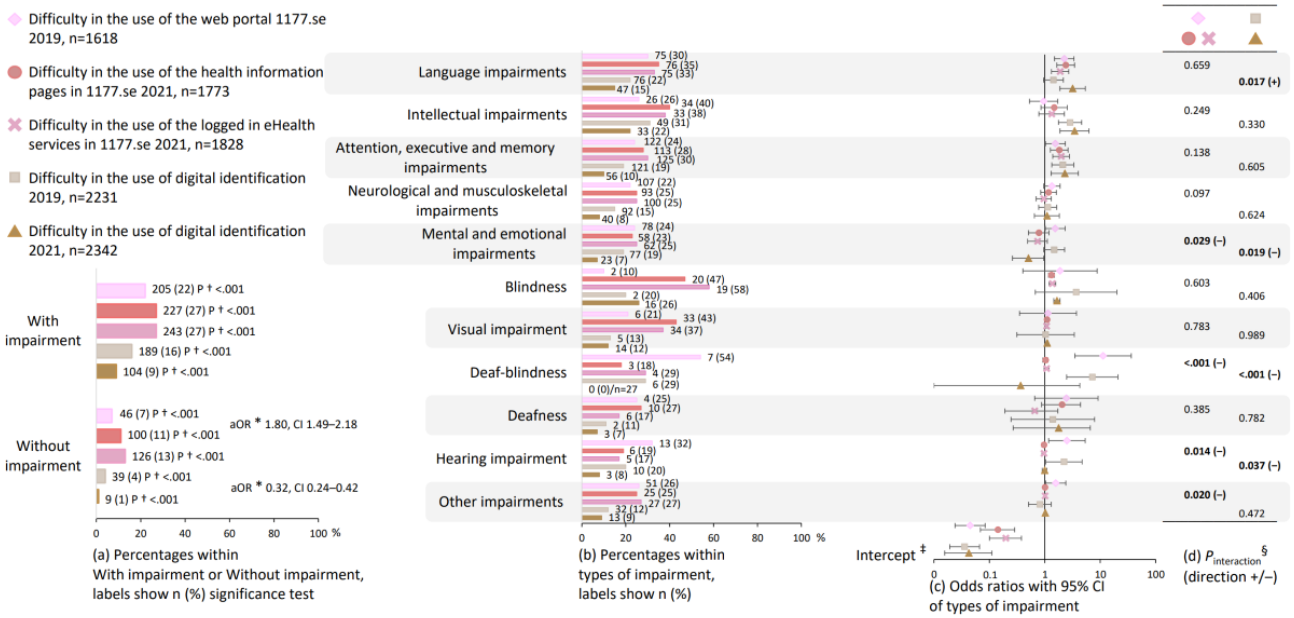


Comparisons of difficulty in the use of eHealth services between the survey years, as well as between participants with and without impairment, are reported in Figure 6a. Among all participants, difficulty in the use of the 1177.se was more frequent in SMFOI21 compared to SMFOI19 (aOR 1.80, 95% CI 1.49 - 2.18; Figure 6a). Difficulty in the use of digital identification was less frequent in SMFOI21 compared to SMFOI19 (aOR 0.32, 95% CI 0.24 - 0.42; Figure 6a). More

XSL•FO
RenderX

participants with impairment than participants without impairment reported difficulty in the use of the 1177.se in SMFOI19 (22% vs 7%; *P*<.001), the information pages in 1177.se in SMFOI21 (27% vs 11%; *P*<.001), the logged-in eHealth services in 1177.se in SMFOI21 (27% vs 13%; *P*<.001), and digital identification in SMFOI19 (16% vs 4%; *P*<.001), and in SMFOI21 (9% vs 1%; *P*<.001; Figure 6a).

Figure 6. Difficulty in the use of the Swedish national web portal for health information and eHealth services (1177.se) and digital identification. The group with impairment (a) is disaggregated into the subgrouped types of impairment from Figure 2 (b–d). The numeric values visualized in this figure are presented in Tables S2, S10, and S11 in Multimedia Appendix 2. *aOR: adjusted odds ratio from logistic regression comparing difficulty in the use of eHealth among all participants overall between the SMFOI19 and SMFOI21 surveys (reference group is 2019), adjusted for type of impairment (reference is participants without impairment), gender (reference female) and age (reference <30 years of age); †2-sided *P* value from chi-square test comparing participants with and without impairment. ‡Intercept from year-stratified logistic regression (reference group is participants without impairment), adjusted for gender (reference female) and age (reference <30 years of age). §*P* value from logistic regression interaction terms of year with types of impairment (reference is SMFOI19 and participants without impairment), adjusted for gender (reference female) and age (reference <30 years of age). SMFOI19: Swedes with impairment and the internet 2019; SMFOI21: Swedes with impairment and the internet 2021.



Associations between difficulty in the use of eHealth services and types of impairment are reported in Figure 6b,c, and further, changes in the associations between the survey years are reported in Figure 6d. Difficulty in the use of the 1177.se was associated with two types of impairments: language impairments and attention, executive, and memory impairments (Figure 6c). Language impairments were associated with difficulty in the use of the 1177.se in SMFOI19 (aOR 2.24, 95% CI 1.50 - 3.36), the 1177.se information pages (aOR 2.39, 95% CI 1.65 - 3.44), and the logged-in eHealth services (aOR 1.89, 95% CI 1.32 - 2.70) in SMFOI21. Attention, executive, and memory impairments were associated with difficulty in the use of the 1177.se in SMFOI19 (aOR 1.55, 95% CI 1.03 - 2.34), the 1177.se information pages (aOR 1.82, 95% CI 1.26 - 2.62), and the logged-in eHealth services (aOR 1.97, 95% CI 1.39 - 2.78) in SMFOI21. Difficulty in the use of digital identification was associated with intellectual impairments in SMFOI19 (aOR 2.86, 95% CI 1.77 - 4.62) and in SMFOI21 (aOR 3.44, 95% CI 1.87 - 6.26), and attention, executive, and memory impairments in SMFOI19 (aOR 2.11, 95% CI 1.33 - 3.34) and in SMFOI21 (aOR 2.30, 95% CI 1.30 - 4.02; Figure 6c).

Discussion

Principal Findings

In a rapid digitalization in health care, this first comparison across the COVID-19 pandemic revealed disability digital divides in eHealth that remained proportional. In this comparison of people with and without impairment, the former were disfavored in regard to eHealth. Namely, they reported less use of booking health care appointments online, the 1177.se and digital identification, which is in line with most studies before the pandemic as discussed in our previous article [2] and one during the pandemic [5]. Further, they reported substantial difficulty in the use of eHealth, which is in agreement with a survey conducted during the pandemic [6]. Nonuse and difficulty in the use of eHealth were consistently associated with language impairments and intellectual impairments. Our findings show that the highest adoption of video health care appointments was among participants with attention, executive, and memory impairments, which is in line with other surveys during the pandemic [8,27]. The latter of the cited studies concluded that video health care appointments increased among mental health service users when contacts were redirected to eHealth instead of face-to-face [27]. Accordingly, it is plausible that the adoption of video health care appointments among participants with attention, executive, and memory impairments in this study was due to disrupted face-to-face health care provision.

The Swedish national web portal for health information and eHealth services, 1177.se, was more frequently reported as difficult to use in SMFOI21 than in SMFOI19, both among participants with and without impairment, despite an accelerated adoption of eHealth across the pandemic. This may be

XSL•FO
RenderX

interpreted as a sign of forced digitalization, which implies that eHealth followed a previously shown societal pattern of nonaccessible social distancing restrictions [28]. People with intellectual impairments were afflicted on two fronts. They were distinguished as a risk group for severe COVID-19 [12] and thereby recommended intensified social distancing restrictions, which meant that they were encouraged to redirect contacts to eHealth instead of face-to-face contacts with health care if possible, while at the same time, they were facing substantial difficulties in the use of eHealth. Cognitive accessibility is insufficiently covered in eHealth studies [29,30] and guidelines [30]. Known adaptations for cognitive accessibility are intuitive structure, important elements before page scroll, large icons, visualizations, and third-party participation [31,32]. However, there is a lack of conformance to what is known about accessibility [7]. Moreover, in this study, people with language impairments were particularly afflicted by difficulties in the use of digital services, which has previously been shown by multiple studies [10,33], but not all [5]. Text and numerical tasks can be adapted by standards for understandable text, autocorrect, data import, speech recognition, audio, and enabling enough duration [31,32]. Barriers to the use of video appointments can be due to reduced nonverbal communication [10] and natural pauses [33].

Implications

Until this disability digital divide in eHealth is bridged, practitioners should provide eHealth integrated with staff support, or maintain conventional services as alternatives to digitalized options. In line with policy [34], we underscore that research and development must conform to legislated accessibility standards [4,31,32,35]. But in addition, we suggest that user participation in the design of eHealth [36] and adherence to principles for universal design [37] can bring additional benefits, beyond the accessibility standards. Such benefits exceed accessibility, by, for instance, readability, clarity, interoperability, and assistive functions, which improve usability for all users [32,35].

Strengths and Limitations

This is, to our knowledge, the first study comparing the accessibility of eHealth across the COVID-19 pandemic among people with diverse types of impairment relevant to the risk of disablement by poorly designed digital services. The pandemic provoked exceptional conditions, which can provide insights for developing resilient health care systems, in which people with high health care needs are included. This comparison between two cross-sectional surveys was deemed a suitable study design to study an unanticipated historical event, in the absence of an available cohort study.

Snowball sampling [19] is a major reason for achieving representation of seldom-heard populations. However, it is plausible that this sampling method engaged more people with higher digital literacy to respond to the survey. Further, self-assessment of digital literacy can be associated with overestimation of skills [38]. Thus, in summation, while this study did demonstrate the disability digital divide in eHealth, it may have underestimated its magnitude. In the two study samples, participants' characteristics and types of impairment demonstrate similar proportions, which makes it plausible that misclassification did not obstruct drawing inferences on comparisons between samples. To allow representation of individuals with multimorbidity in the analysis, we assessed that multicollinearity did not exceed the predetermined cut-off, and that outcomes were not associated with the number of reported impairments. The SMFOI21 survey sampled more participants in the oldest group (70 years or older) and less participants of 45 - 69 years of age, as compared to the SMFOI19 survey. This was handled as we adjusted for age in the statistical analysis. While we adjusted for gender and age, there were socioeconomic differences between people with and without impairment, for which we did not control. The dependent variables were investigated by different types of response options "I use"; "I avoid" or "I prefer"; as well as "is difficult to use" or "is easy to use" in order to capture different aspects of the phenomenon. We consider that avoiding is a dimension of using, in that it measures a type of "not using." However, it is possible that avoidance also can be associated with difficulty of use, but may also be influenced by other mechanisms, for example, self-efficacy, normative beliefs, or social support. This study focused on the association between types of impairment and eHealth services for patient-provider contact, meaning that other types of eHealth and other factors plausibly mediating the use of eHealth were not investigated, that is, eHealth literacy, socioeconomic factors, or lacking access to digital devices.

Conclusions

Despite an accelerated adoption of eHealth across the COVID-19 pandemic, our results show no relevant difference in the disability digital divide in eHealth between the survey years. In both before and during the pandemic, there were substantial nonuse and difficulty in the use of eHealth among people with impairment. Nonuse and difficulty in the use of eHealth were consistently associated with language impairments and intellectual impairments. This implies that groups recognized as being at risk of severe COVID-19 could not use eHealth, while emphasis was put on eHealth as a measure of infection protection in the social distancing policy. Accessibility standards can improve services. Until then, our results indicate that eHealth in its current state should be provided with optional support, or only be provided as a complement to conventional contacts.

Acknowledgments

Our research partners, the Begripsam group, all of whom have lived experience of impairment, contributed to the study design and development of the questionnaire of the first SMFOI survey (2017) and participated in further elaboration and refinement of the following surveys. Riccardo Lo Martire, statistician at the Center for Clinical Research Dalarna Sweden, provided consultation.

XSL•FO
RenderX

The study was funded by the Center for Clinical Research Dalarna, Sweden, and FORTE - The Swedish Research Council for Health, Working Life and Welfare (grant 2021-01875). The funders had no part in the undertaking of the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Question about impairments.
[DOCX File, 30 KB - publichealth_v11i1e64707_app1.docx]

Multimedia Appendix 2
Tables of all logistic regression models of this study. Unadjusted and adjusted model per outcome.
[DOCX File, 91 KB - publichealth_v11i1e64707_app2.docx]

References

1. Tu K, Sarkadi Kristiansson R, Gronsbell J, et al. Changes in primary care visits arising from the COVID-19 pandemic: an international comparative study by the International Consortium of Primary Care Big Data Researchers (INTRePID). BMJ Open 2022 May 9;12(5):e059130. [doi: 10.1136/bmjopen-2021-059130] [Medline: 35534063]
2. Pettersson L, Johansson S, Demmelmaier I, Gustavsson C. Disability digital divide: survey of accessibility of eHealth services as perceived by people with and without impairment. BMC Public Health 2023 Jan 27;23(1):181. [doi: 10.1186/s12889-023-15094-z] [Medline: 36707791]
3. UN General Assembly. Transforming our world: the 2030 agenda for sustainable development. A/RES/70/1. 2015 URL: http://digitallibrary.un.org/record/1654217 [accessed 2022-04-08]
4. Directive (EU) 2016/2102 of the European Parliament and of the Council of 26 October 2016 on the accessibility of the websites and mobile applications of public sector bodies. 2016. URL: https://eur-lex.europa.eu/eli/dir/2016/2102/oj [accessed 2024-01-31]
5. Park EY. Effect of COVID-19 on internet usage of people with disabilities: a secondary data analysis. Int J Environ Res Public Health 2022 Jun 25;19(13):7813. [doi: 10.3390/ijerph19137813] [Medline: 35805471]
6. Lawford BJ, Hinman RS, Morello R, Oliver K, Spittle A, Bennell KL. Perceptions about the efficacy and acceptability of telephone and video-delivered allied health care for adults with disabilities during the COVID-19 pandemic: a cross-sectional national survey. Arch Phys Med Rehabil 2022 Jul;103(7):1368-1378. [doi: 10.1016/j.apmr.2021.12.017] [Medline: 35041838]
7. Jonsson M, Gustavsson C, Gulliksen J, Johansson S. How have public healthcare providers in Sweden conformed to the European Union's Web Accessibility Directive regarding accessibility statements on their websites? Univ Access Inf Soc 2025 Mar;24(1):449-462. [doi: 10.1007/s10209-023-01063-1]
8. Xie Z, Hong YR, Jo A, Marlow NM. Telehealth utilization during the covid-19 pandemic among people with functional disabilities in the U.S.: a preliminary analysis of national health interview survey 2020 data. AJPM Focus 2023 Dec;2(4):100149. [doi: 10.1016/j.focus.2023.100149] [Medline: 37920402]
9. Gulliksen J, Lilliesköld J, Stenbom S. The 'New' new normal—digitalization and hybridization of work and education before, during and after the COVID-19 pandemic. Interact Comput 2023 Dec 31;35(5):691-706. [doi: 10.1093/iwc/iwac034]
10. Matenge S, Sturgiss E, Desborough J, Hall Dykgraaf S, Dut G, Kidd M. Ensuring the continuation of routine primary care during the COVID-19 pandemic: a review of the international literature. Fam Pract 2022 Jul 19;39(4):747-761. [doi: 10.1093/fampra/cmab115] [Medline: 34611708]
11. Fernemark H, Skagerström J, Seing I, Hårdstedt M, Schildmeijer K, Nilsen P. Working conditions in primary healthcare during the COVID-19 pandemic: an interview study with physicians in Sweden. BMJ Open 2022 Feb 8;12(2):e055035. [doi: 10.1136/bmjopen-2021-055035] [Medline: 35135771]
12. The National Board of Health and Welfare. Final report on risk for severe COVID-19 (S2021/00825) [in swedish: slutredovisning av regeringsuppdrag om risk för särskilt allvarlig sjukdomsutveckling vid COVID-19]. 2021 URL: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/dokument-webb/ovrigt/slutredovisning-riskgrupper-covid-19.pdf [accessed 2023-06-15]
13. Inera Company. 2024. URL: https://www.inera.se/tjanster/statistik-for-ineras-tjanster/statistik-for-1177.se/ [accessed 2024-05-23]
14. Ekman B, Ellegård L. Digitaliseringen av svensk vård och omsorg. In: The Digitalization of Swedish Healthcare: SNS Förlag; 2023:120.
15. Andersson J, Blomdahl F. The swedes and the internet 2022. The Swedish Internet Foundation (IIS). 2022 Oct 26 URL: https://svenskarnaochinternet.se/english/earlier-years-reports/ [accessed 2023-04-14]
16. Statistics Sweden (SCB). Use of computers and the internet by private persons 2016. 2016 Nov 24 URL: https://www.scb.se/hitta-statistik/statistik-efter-amne/forskning-och-det-digitala-samhallet/det-digitala-samhallet/

XSL•FO
RenderX

befolkningens-it-anvandning/pong/publikationer/privatpersoners-anvandning-av-datorer-och-internet-2016/ [accessed 2019-01-05]
17. Eurostat. Digital economy and society—database ICT usage in households and by individuals. URL: http://ec.europa.eu/eurostat/web/digital-economy-and-society/database [accessed 2024-01-31]
18. Begripsam. The Begripsam Company and the Begripsam Group Sweden. URL: https://www.begripsam.se/english [accessed 2024-01-31]
19. Lavrakas P. Encyclopedia of Survey Research Methods: SAGE Publications Ltd; 2008.
20. The Swedish state personal address register (SPAR). The Swedish Tax Agency. URL: https://www.statenspersonadressregister.se/master/start/english-summary [accessed 2024-01-31]
21. Reiners F, Sturm J, Bouw LJW, Wouters EJM. Sociodemographic factors influencing the use of eHealth in people with chronic diseases. Int J Environ Res Public Health 2019 Feb 21;16(4):645. [doi: 10.3390/ijerph16040645] [Medline: 30795623]
22. Mangin D, Parascandalo J, Khudoyarova O, Agarwal G, Bismah V, Orr S. Multimorbidity, eHealth and implications for equity: a cross-sectional survey of patient perspectives on eHealth. BMJ Open 2019 Feb 12;9(2):e023731. [doi: 10.1136/bmjopen-2018-023731] [Medline: 30760515]
23. Ustün TB, Chatterji S, Bickenbach J, Kostanjsek N, Schneider M. The International Classification of Functioning, Disability and Health: a new tool for understanding disability and health. Disabil Rehabil 2003;25(11-12):565-571. [doi: 10.1080/0963828031000137063] [Medline: 12959329]
24. Grigorenko EL, Compton DL, Fuchs LS, Wagner RK, Willcutt EG, Fletcher JM. Understanding, educating, and supporting children with specific learning disabilities: 50 years of science and practice. Am Psychol 2020 Jan;75(1):37-51. [doi: 10.1037/amp0000452] [Medline: 31081650]
25. MacQueen J. Some methods for classification and analysis of multivariate observations. In: Le Cam LM, Neyman J, editors. Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics: University of California Press:281-297 URL: http://projecteuclid.org/euclid.bsmsp/1200512992 [accessed 2025-04-07]
26. VanderWeele TJ, Knol MJ. A tutorial on interaction. Epidemiol Methods 2014 May 27;3(1):33-72. [doi: 10.1515/em-2013-0005]
27. Ivancic L, Bond DM, Nassar N. Impact of the COVID-19 pandemic on access and use of health services by middle-aged and older Australians. Aust Health Rev 2023 Feb;47(1):100-109. [doi: 10.1071/AH22183] [Medline: 36395815]
28. Chadwick D, Ågren KA, Caton S, et al. Digital inclusion and participation of people with intellectual disabilities during COVID‐19: a rapid review and international bricolage. Policy Practice Intel Disabi 2022 Sep;19(3):242-256. [doi: 10.1111/jppi.12410]
29. Jonsson M, Johansson S, Hussain D, Gulliksen J, Gustavsson C. Development and evaluation of eHealth services regarding accessibility: scoping literature review. J Med Internet Res 2023 Aug 17;25:e45118. [doi: 10.2196/45118] [Medline: 37590050]
30. European Commission: Directorate-General for Communications Networks, Content and Technology, Funka ICF, Kjellstrand S, Laurin S, et al. Inclusive web-accessibility for persons with cognitive disabilities—pilot project study—final report. European Commission. 2022 URL: https://data.europa.eu/doi/10.2759/3048 [accessed 2022-04-06]
31. ISO/TS 82304-2:2021. Health software—part 2: health and wellness apps—quality and reliability. URL: https://www.iso.org/standard/78182.html [accessed 2022-03-07]
32. WHO-ITU. Global standard for accessibility of telehealth services. Geneva: World Health Organization and International Telecommunication Union Licence: CC BY-NC-SA 3.0 IGO. 2022 URL: https://www.who.int/publications/i/item/9789240050464 [accessed 2022-12-01]
33. Ward K, Vagholkar S, Sakur F, Khatri NN, Lau AYS. Visit types in primary care with telehealth use during the COVID-19 pandemic: systematic review. JMIR Med Inform 2022 Nov 28;10(11):e40469. [doi: 10.2196/40469] [Medline: 36265039]
34. World Health Organization. Global strategy on digital health 2020-2025. Geneva: World Health Organization. Licence: CC BY-NC-SA 3.0 IGO. 2021 URL: https://www.who.int/publications/i/item/9789240020924 [accessed 2022-04-07]
35. ISO 21801-1:2020. Cognitive accessibility—part 1: general guidelines. URL: https://www.iso.org/standard/71711.html [accessed 2022-09-27]
36. World Wide Web Consortium (W3C). Involving users in evaluating web accessibility. URL: https://www.w3.org/WAI/test-evaluate/involving-users [accessed 2024-06-15]
37. Persson H, Åhman H, Yngling AA, Gulliksen J. Universal design, inclusive design, accessible design, design for all: different concepts—one goal? On the concept of accessibility—historical, methodological and philosophical aspects. Univers Access Inf Soc 2015 Nov;14(4):505-526. [doi: 10.1007/s10209-014-0358-z]
38. Merritt K, Smith D, Di Renzo Jr JC. An investigation of self-reported computer literacy: is it reliable? Issues Inf Syst 2005;6(1):289-295. [doi: 10.48009/1_iis_2005_289-295]

Abbreviations

aOR: adjusted odds ratio

XSL•FO
RenderX

ICF: International Classification of Functioning, Disability, and Health
SMFOI19: Swedes with impairment and the internet 2019
SMFOI21: Swedes with impairment and the internet 2021

Edited by AS Bhagavathula; submitted 30.07.24; peer-reviewed by M Koks-Leensen; revised version received 09.01.25; accepted 18.02.25; published 28.03.25.

Please cite as:
Pettersson L, Johansson S, Demmelmaier I, von Koch L, Gulliksen J, Hedvall PO, Gummesson K, Gustavsson C
Accessibility of eHealth Before and During the COVID-19 Pandemic Among People With and People Without Impairment: Repeated Cross-Sectional Survey
JMIR Public Health Surveill 2025;11:e64707
URL: https://publichealth.jmir.org/2025/1/e64707
doi:10.2196/64707

© Linda Pettersson, Stefan Johansson, Ingrid Demmelmaier, Lena von Koch, Jan Gulliksen, Per-Olof Hedvall, Karl Gummesson, Catharina Gustavsson. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Facilitated Telemedicine as a Patient-Centered, Sociotechnical Intervention to Integrate Hepatitis C Treatment Into Opioid Treatment Programs and Overcome the Digital Divide Among Underserved Populations: Qualitative Study

Andrew H Talal[1], MD, MPH; Arpan Dharia[1], MD; Marianthi Markatou[2], PhD; Lawrence S Brown Jr[3,4], MD, MPH; Kenneth E Bossert[1,5], BS; Zakiya Grubbs[6], MPH; Raktim Mukhopadhyay[2], MS; Boatemaa Ntiri-Reid[6], JD, MPH; Elisabeth J Houtsmuller[7,8], PhD

[1]Division of Gastroenterology, Hepatology and Nutrition, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, 875 Ellicott Street, Buffalo, NY, United States

[2]Department of Biostatistics, School of Public Health and Health Professions, University at Buffalo, Buffalo, NY, United States

[3]START Treatment and Recovery Centers, Brooklyn, NY, United States

[4]Addiction Medicine and Public Health, Weill Cornell Medicine, New York, NY, United States

[5]Drug Abuse Research and Treatment, Community Action Organization of Western New York, Buffalo, NY, United States

[6]National Alliance of State and Territorial AIDS Directors, Washington, DC, United States

[7]Patient-Centered Outcomes Research Institute, Washington, DC, United States

[8]Research Triangle Institute, Research Triangle Park, NC, United States

Corresponding Author:
Andrew H Talal, MD, MPH
Division of Gastroenterology, Hepatology and Nutrition, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, 875 Ellicott Street, Buffalo, NY, United States

Abstract

Background: People with opioid use disorder (OUD) have the highest rates of hepatitis C virus (HCV) infection. Despite the availability of curative HCV medication, people with OUD have limited health care access largely due to stigma. In a recent, pragmatic, randomized controlled trial (RCT), we compared a facilitated telemedicine intervention for HCV treatment integrated into opioid treatment programs (OTPs) with off-site referral. Facilitated telemedicine is bidirectional videoconferencing between a remote provider and a patient, supported by a case manager who facilitates the telemedicine encounter. The case manager schedules telemedicine visits, provides appointment reminders, and operates the digital equipment. Among 602 participants in the RCT, 90% (n=262) were cured through facilitated telemedicine and 39% (n=123) were cured through off-site referral. In this work, a multidisciplinary group of investigators, who directed the RCT, conducted a workshop, "Advancing Viral Hepatitis Screening and Treatment in Opioid Treatment Settings – Models & Resources," at the American Association for Treatment of Opioid Dependence Conference in May 2024 to disseminate knowledge of facilitated telemedicine, including implementation considerations. We highlighted facilitated telemedicine as a patient-centered, sociotechnical, pragmatic health care delivery model for underserved populations.

Objective: This study aimed to identify lessons learned to successfully overcome challenges of facilitated telemedicine implementation for HCV treatment integrated into OTPs.

Methods: We partnered with the National Alliance of State and Territorial AIDS Directors in planning the workshop. The workshop consisted of 7 presentations on topics related to facilitated telemedicine implementation. The workshop was recorded and transcribed by Zoom (Zoom Communications). The transcripts served as data for the thematic analysis. The transcripts were interpreted to elucidate patterns of meanings and nuances derived from each presentation. In an iterative process, preliminary findings were compared and coalesced into themes. Verbatim quotes from the workshop were highlighted to support the themes.

Results: We developed 3 themes. First, patient-centered care promotes HCV treatment for underserved populations through facilitated telemedicine. Case managers leveraged the destigmatizing environment of the OTP to build trust with patients, promoting an HCV cure through facilitated telemedicine. Second, sociotechnical approaches expand health care access for people with OUD. To be effective, facilitated telemedicine integrates 2 necessary components, a social aspect and a technical aspect. Third, facilitated telemedicine supports pragmatic research emphasizing people with OUD. Pragmatic research of facilitated telemedicine is needed to assess sustainability and scaling considerations beyond OTPs. Overall, we found that facilitated telemedicine overcame the digital divide, promoting access to digital technology, internet provision, and digital literacy.

XSL•FO
RenderX

Conclusions: Facilitated telemedicine incorporates both a technical and a social component. The technical component largely addresses geographical challenges, while the social component addresses temporal (ie, care coordination) issues, promotes trust, and largely assuages patients' concerns related to HCV treatment. The patient-centered, sociotechnical intervention can address the digital divide, thereby increasing health care access.

Trial Registration: ClinicalTrials.gov NCT02933970; https://clinicaltrials.gov/study/NCT02933970

(JMIR Public Health Surveill 2025;11:e68854) doi:10.2196/68854

KEYWORDS

telemedicine; patient-centered; sociotechnical; hepatitis; opioid; health equity; digital divide; underserved; pragmatic.

Introduction

People with opioid use disorder (OUD) have the highest hepatitis C virus (HCV) incidence and prevalence due to viral transmission through injection drug use [1]. If left untreated, HCV infection can lead to cirrhosis, liver failure, and death. The development of almost ubiquitously curative, direct-acting antivirals (DAAs) for HCV in 2 to 3 months with minimal side effects has profoundly changed the therapeutic landscape since 2013 [2]. Unfortunately, however, people with OUD have had limited access to DAAs, largely due to stigma in conventional health care settings, competing priorities, unawareness of their HCV status, and service unavailability [3,4]. A recent United States nationwide study documented that only 35% of eligible individuals received DAAs from 2013 to 2022 [5]. In addition, almost 40% of people with HCV are unaware of their infection status [6]. Nonetheless, HCV treatment is a health equity priority [7]. Thus, an HCV therapeutic divide exists with insufficient penetration of highly curative DAAs among the population most in need. Furthermore, people with OUD suffer from a digital divide, defined as the unequal access to digital technology, internet provision, and gaps in digital literacy [8].

Opioid treatment programs (OTPs) provide medical and behavioral treatment for OUD in a destigmatizing, supportive, and trusting environment [9]. Integrating HCV treatment into the OTP environment can promote therapeutic interactions, patient engagement, and retention in health care. To potentially overcome the HCV therapeutic divide among people with OUD, we developed a facilitated telemedicine intervention fully integrated into OTPs using the sociotechnical system framework [10]. Recently published facilitated telemedicine results demonstrated an HCV cure in almost all treated individuals, accompanied by high patient satisfaction with health care delivery [11-13] (details of facilitated telemedicine procedures and effectiveness results are provided in Multimedia Appendix 1). Facilitated telemedicine is bidirectional videoconferencing between a remote provider and patient, supported by a case manager. The case manager established trusting relationships with patients to promote personalized HCV care approaches. The intervention can overcome geographical and temporal challenges, as well as the therapeutic and digital divides. Since facilitated telemedicine encounters are embedded in OTPs facilitated by a case manager, patients had access to digital technology, internet provision, and assistance with digital literacy. Facilitated telemedicine not only increases DAA access among people with OUD, but it can also improve service provision by OTPs.

We obtained input from a variety of stakeholders involved in the implementation and conduct of facilitated telemedicine in OTPs. We conducted a workshop entitled "Advancing Viral Hepatitis Screening and Treatment in Opioid Treatment Settings – Models & Resources." The objective of the workshop was to disseminate successful approaches that overcame facilitated telemedicine implementation challenges. We also desired to inform workshop participants how facilitated telemedicine promotes patient-centered health care delivery for HCV treatment integrated into OTPs. We highlighted facilitated telemedicine implementation considerations that addressed the digital divide. We also highlighted specific considerations when conducting pragmatic trials, emphasizing underserved populations. We anticipate that the lessons learned could be beneficial in the expansion of facilitated telemedicine to address HCV-OUD treatment integration and health equity among people with OUD.

Methods

Study Design, Setting, and Sampling

We conducted a thematic analysis of the transcripts of a 3.5-hour workshop that addressed facilitated telemedicine implementation challenges. Our study's aim was to identify facilitated telemedicine considerations that overcame implementation challenges to promote the delivery of patient-centered HCV care. The primary study outcome was the results derived from the thematic analysis. The workshop transcripts served as data for the thematic analysis.

The workshop was sponsored by the University at Buffalo (UB) and took place as a preconference workshop at the 2024 American Association for the Treatment of Opioid Dependence (AATOD) Conference [14]. The workshop occurred on May 19, 2024, in Las Vegas, Nevada. Planning for the workshop began in 2018 when 3 of the workshop speakers attended the AATOD annual meeting; the workshop planning followed an iterative process (Figure 1). The data analysis occurred in July and August 2024.

XSL•FO
RenderX

Figure 1. Workshop development flow diagram. The objective of the workshop was to disseminate knowledge of facilitated telemedicine for hepatitis C virus infection treatment integrated into opioid treatment programs to attendees of the 2024 American Association for the Treatment of Opioid Dependence conference. AATOD: American Association for the Treatment of Opioid Dependence; NASTAD: National Alliance of State & Territorial AIDS Directors; Q&A: question and answer.

XSL•FO
RenderX

AATOD works with federal and state agencies to promote the growth and development of opioid treatment services. The association has a long history of promoting approaches to increase HCV management within OTPs, and the workshop concept originated at the 2018 annual meeting, where 3 workshop participants attended. Our workshop aligned with the AATOD conference theme of "Treating Opioid Use Disorder: So Much More Than Medication" [15]. In alignment with the conference objectives, we disseminated the outcomes of facilitated telemedicine among people with OUD. Our workshop attendees included OTP providers, OTP administrators, patient representatives, front-line staff, and state officials involved in the treatment of OUD.

Study Intervention

The workshop activities were approved by the UB Institutional Review Board. The principal investigator (AHT) of the facilitated telemedicine trial established a collaboration with the National Alliance of State and Territorial AIDS Directors (NASTAD) in planning the workshop. The focus of the alliance is to advance the health of people impacted by HIV, viral hepatitis, and intersecting epidemics through public health advocacy, capacity building, and social justice [16]. NASTAD represents public health officials who administer HIV and viral hepatitis programs in all 50 states. NASTAD representatives participated in developing the workshop's objectives and maintaining the relevance of the workshop through partnering with state and territorial health departments.

Workshop Speakers

The planning committee invited workshop speakers based on their roles and a high degree of engagement in the trial. The workshop was hosted by 2 planners, 1 moderator, and 7 speakers. The planners were AHT (also a speaker) and BN-R. AHT is a physician-scientist with more than 20 years of experience in treating HCV, especially among people with OUD. He is a Professor of Medicine at UB and the principal investigator on the randomized controlled trial entitled "Integrated Hepatitis C–Opioid Use Disorder Care Through Facilitated Telemedicine A Randomized Trial" [13]. AHT serves on the New York State HCV Elimination Task Force and its HCV Guidelines Committee. BN-R is the Senior Director of Syndemic Approaches at NASTAD. She leads the integration of HIV, viral hepatitis, and harm reduction. The moderator was ZG, Senior Hepatitis Manager at NASTAD. She supports prevention and surveillance capacity building through NASTAD's technical assistance center for viral hepatitis.

The speakers were AHT, EJH, MM, LSB, KEB, AD, and RM. EJH is a Program Director at Research Triangle Institute. Previously, she was an Associate Director at the Patient-Centered Outcomes Research Institute (PCORI) and the trial's program officer. PCORI is an independent, nonprofit organization that funds patient-centered comparative clinical effectiveness research with emphasis on stakeholder engagement, and dissemination and implementation of successful interventions. MM is a Distinguished Professor of the State University of New York and Associate Chair for Research and Healthcare Informatics, Department of Biostatistics at UB. She was the lead biostatistician for the trial. LSB is the former chief executive officer at START Treatment & Recovery Centers, treating patients with substance use and co-occurring medical and mental health disorders. Founded in 1969, START is the oldest and largest minority-run substance use treatment program in the United States, serving over 3000 patients through 7 medication-assisted programs in New York City. KEB is the former program administrator at Drug Abuse Research and Treatment of the Community Action Organization of Western New York. Founded in 1971, Drug Abuse Research and Treatment is a free-standing community organization that provides integrated services to people with OUD throughout Western New York. KEB was also an administrative surveyor for the Commission on Accreditation of Rehabilitation Facilities. AD is a physician and the Director of Liver Services at UB. RM is a research assistant in MM's laboratory group and a doctoral student in computational and data-enabled sciences and engineering at UB.

Workshop Presentations Overview

The workshop was developed through planning meetings between AHT, AD, and NASTAD's representative, BN-R. The planning committee developed an agenda with topics related to facilitated telemedicine and invited stakeholders who participated in the study by email. We used purposive sampling to select speakers who were the most appropriate to discuss each topic, as illustrated in Table 1. Furthermore, 2 individuals, representing an expert in substance use disorder and a patient-participant with lived experience, were unable to participate. These topics were represented by LSB, who has 40 years of experience in substance use treatment, and AD, who has served as a patient representative and program director during the trial. We used investigator and data source triangulation to understand when we had achieved sample saturation [17]. The workshop consisted of 7 presentations, and all invited speakers shared their presentations before the workshop. The dissemination of the trial's findings to an AATOD audience permitted scalability opportunities to a national and international audience substantially larger than New York State, where the trial occurred.

XSL•FO
RenderX

Table . Principal workshop topics by speaker.

Topic and subtopics	Participant, name (degree)
Research	
Facilitated telemedicine	AHT (MD, MPH)
Substance use	EJH (PhD)
Social determinants of health	MM (PhD)
Data acquisition and the use of large language models	MM (PhD) and RM (MS)[a]
Perspectives and experiences	
Patients	AD (MD)
OTP[b] leadership	LSB (MD, MPH)
OTP administration	KEB (BS)
Stakeholders	
Public health advocacy	BN-R (JD, MPH) and ZG (MPH)
Experience with OUD[c]	AATOD[d] audience

[a]Speaker.

[b]OTP: opioid treatment program.

[c]OUD: opioid use disorder.

[d]AATOD: American Association for Treatment of Opioid Dependence.

AHT opened with an overview of the trial, "Facilitated Telemedicine for Underserved Populations: A Stepped Wedge Trial of Hepatitis C Treatment." AHT discussed the background of HCV treatment access and health care delivery. He introduced the facilitated telemedicine model, the trial design, primary and secondary outcomes, and the trial's strengths, limitations, and lessons learned.

EJH followed with "Funding Projects in People with OUD." EJH explained that comparative effectiveness research funded by PCORI must be patient-centered, stakeholder-driven, and pragmatic, and that the facilitated telemedicine trial met all 3 criteria. She described the recent increase in patient-centered approaches for people with OUD. EJH highlighted several patient-centered parts of the final rule of the Substance Abuse and Mental Health Services Administration that was released in April 2024 [18]. The Substance Abuse and Mental Health Services Administration final rule expands treatment access by permitting telehealth evaluation, replacing stringent dosing requirements with clinical judgment, removing daily visit requirements, and meeting OUD diagnostic criteria. EJH cautioned that critical barriers remain, such as internet access in rural and underserved areas and for those with limited digital literacy.

MM presented on "Improvements in Substance Use and Social Determinants of Health (SDOH)." She discussed how substance use and SDOH can be evaluated as predictors for HCV treatment initiation and cure. MM explained that integrated HCV and OUD treatment resulted in improvements in substance use, education, employment, and criminal justice status. However, since treatment for HCV and OUD is not well-integrated into health care systems, managing these conditions will require addressing social and structural inequalities.

LSB discussed "Drivers for HCV-Related Telemedicine Research." He described the critical relationship between public policy and clinical care as he was originally from a neighborhood in which substance use was, and unfortunately still is, part of the landscape. LSB provided an overview of START Treatment and Recovery Centers in New York City. He explained that facilitated telemedicine aligned very closely with START's corporate philosophy of participating in medical and behavioral research and advocacy to address unmet patient needs.

KEB presented "The Evolution of Incorporating Facilitated Telemedicine in an Opioid Treatment Program: An Administrator's Perspective." KEB defined the notions of the "phenomenon of the OTP" and "full-service garage." He described the "phenomenon of the OTP" as a natural incubator of trust because the OTP provided structure, purpose, a safe environment, and socialization to people with OUD. The "full-service garage" implies that the OTP offers a panoply of health care and social services beyond medication-assisted treatment. KEB emphasized the interactions of facilitated telemedicine with both concepts.

AD discussed "Facilitated Telemedicine: Overcoming Challenges for Hepatitis C Treatment." He presented patient testimonials of facilitated telemedicine derived from qualitative interviews [3]. Patients were enthusiastic about health care delivery through facilitated telemedicine, and an HCV cure promoted self-confidence and improved whole health. AD also reviewed the CREATE (Culture, Respect, Educate, Advantage, Trust, Endorse) framework, an engagement approach for facilitated telemedicine [9].

RM concluded the workshop with "From Words to Wellness: Large Language Models Applied to the Identification of SDOH." He explained that SDOH data are limited in people with HCV,

XSL•FO
RenderX

and he described approaches used to derive SDOH data from trial participants [19]. RM concluded his presentation with a discussion of issues of quality and the application of models for the identification of SDOH derived from people with OUD.

Once the speakers had agreed to present at the workshop, they worked iteratively with the planning committee to obtain feedback on the presentations. When the presentations were finalized, each speaker uploaded the presentation to the AATOD website. Workshop speakers met in person before the workshop to review the presentations' content and to promote cohesiveness and flow. Speakers reviewed logistical considerations immediately before the workshop, including timing, recording, and moderation by NASTAD staff. Furthermore, 2 question-and-answer sessions were held during the workshop that were moderated by ZG. During the introductory remarks, the moderator discussed the objectives of the conference, the context of HCV with a special emphasis on the importance of HCV elimination, and specific areas of discussion for the audience. To minimize recollection bias, the question-and-answer sessions occurred immediately after the workshop presentations. As the workshop presenters had extensive experience working collaboratively to achieve the trial's goals, respondent bias was minimized. On the other hand, however, the moderator was not previously known to the workshop presenters and represented a public health advocacy perspective, which encouraged discussion of a diverse array of perspectives. To ensure question comprehension, the moderator repeated the question before eliciting a response. The moderator also sought responses from multiple workshop presenters as appropriate. After completion of the workshop, the planning committee met to discuss topics for manuscript development. The primary focus of the planning committee was the thematic analysis of the workshop presentation transcripts.

Thematic Analysis

The 3.5-hour workshop was recorded and transcribed by Zoom (Zoom Communications). All workshop speakers disseminated their successful approaches that overcame facilitated telemedicine implementation challenges. After the final set of transcripts was obtained, they were reviewed for completeness independently by 2 investigators (AHT and AD), with inconsistencies verified by listening to the recordings. While the transcripts did have identifiable quotations, those that were used in the analysis were deidentified. The workshop transcripts served as the raw data and the primary source for the thematic analysis. Since all workshop presenters had worked collaboratively for 8 years on evidence generation as part of the trial, the workshop adhered to the principle of investigator triangulation. The development of multiple manuscripts covering a variety of different perspectives from the trial adhered to the use of data triangulation [17].

Using a constructivist paradigm, we performed a thematic analysis to identify facilitated telemedicine considerations that overcame implementation challenges and delivered patient-centered care. In total, 2 investigators, AD and AHT, followed the 6-step framework for thematic analysis described by Braun and Clarke [20,21]. AD and AHT have published several qualitative studies in Q1 journals [3,9,22-27] and are well-versed in the approaches to minimize bias in qualitative research. They became familiar with the data by independently reviewing the transcripts for accuracy and formatting. They independently and inductively coded the transcript data based on its relevance to facilitated telemedicine considerations that overcame implementation challenges and delivered patient-centered care. Subsequently, they independently grouped the initial codes into general categories. In a recursive process, they discussed code generation to identify and interpret repeated patterns of meaning and nuances derived from each presentation. They met frequently to discuss these patterns of meaning until a consensus was reached and lower-level themes were developed with supporting quotations [28]. In the case of disagreements in the textual interpretations, they reviewed the transcripts until a consensus was reached.

Once a draft of concepts had been derived and agreed upon, the lower-level themes were discussed and reviewed with all coauthors in a series of meetings. Using a consensus and iterative approach, lower-level themes were defined and evolved into higher-level themes. Verbatim quotes from the workshop presentations were highlighted to support the themes. All coauthors reviewed the final themes and supporting quotes for agreement. All coauthors were active participants in the data-driven thematic analysis. We acknowledge that the experiences and interactions of the coauthors may have influenced the thematic analysis in this constructivist paradigm. As a concise report of qualitative analytical methods used, the Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist is provided (Multimedia Appendix 1).

Because of time constraints, several aspects related to care delivery through facilitated telemedicine could not be addressed at the workshop and are discussed in detail in separate papers. Specifically, the aspects that led to successful case management are discussed in a study by Talal et al [27]. The considerations for effective case conferencing at the patient and system levels are discussed in [22,29], respectively. Finally, the issues related to care delivery in studies by Ventuneac et al [23], Dickerson et al [24], and Talal et al [unpublished data, 2025] describe the considerations for scaling and maintaining facilitated telemedicine, the characteristics that led to full integration of facilitated telemedicine into OTP workflows,, and implementation characteristics required for effective treatment outcomes and high levels of patient satisfaction, respectively. In addition, the issues of data ownership, sharing, and privacy were not discussed at the conference, as these topics had been outlined in the trial protocol approved by the institutional review board.

Ethical Considerations

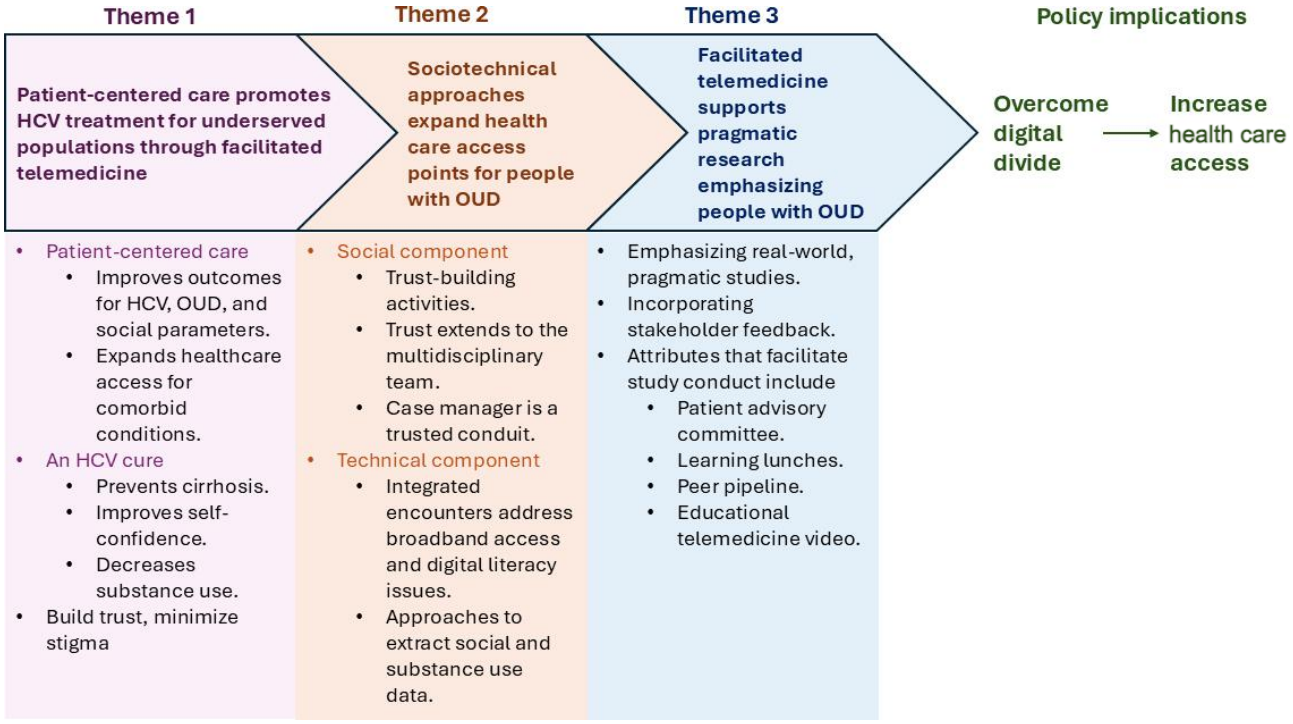
The study was approved by the Institutional Review Board at UB (MODCR00009273) (Multimedia Appendix 2). All workshop speakers consented to participate. All quotations were deidentified. Confidentiality concerns prohibit us from sharing data transcripts. We provided reimbursement for travel, lodging, and meals. Artificial intelligence was not used in any portion of this study.

XSL•FO
RenderX

Results

We developed 3 themes from the textual analysis of the workshop presentations (Figure 2).

Figure 2. Three themes and important attributes identified through thematic analysis of workshop presentations at the 2024 American Association for Treatment of Opioid Dependence Conference. The objective of the workshop was to disseminate knowledge of facilitated telemedicine for hepatitis C virus treatment. We highlighted facilitated telemedicine as a patient-centered, sociotechnical, pragmatic health care delivery model for underserved populations. We found that facilitated telemedicine integrated into opioid treatment programs overcame the digital divide. HCV: hepatitis C virus; OUD: opioid use disorder.



Theme 1: Patient-Centered Care Promotes Hepatitis C Treatment for Underserved Populations Through Facilitated Telemedicine

Patient-centered care for HCV and OUD promotes improvements not only in medical conditions but also in social factors. Patient-centered care emphasizes promoting individual health care choices and empowering patients to actively participate in health care decision-making [30] and is delivered by a multidisciplinary team.

> *[The] approach is to treat the whole person, and not just their substance use disorder… They have co-occurring alcohol use disorder… HIV, hepatitis C, diabetes, hypertension, and psychiatric disorders.*
>
> *[HCV and OUD treatment] promotes improvements in substance use and social functioning, [specifically] education, employment, and involvement with the legal system.*
>
> *The patients were actually helped to go back to school and complete whatever degrees they had [to complete].*

Consistent with the strong potential benefits of integrated OUD and HCV care, an OTP administrator cited his desire for participation in facilitated telemedicine as the "full-service garage" concept, providing essential medical care within the OTP for a large repertoire of comorbidities.

> *The full-service garage sentiment is the leverage of the trust equity and utilizing facilitated telemedicine provided a hub and spoke… If OTPs reach out to community providers, it becomes the home base for consumers and facilitated telemedicine becomes the opportunity to engage in essential services without having to leave the clinic.*

The value of integrated care was illustrated through obtaining an HCV cure, which represented a major achievement for people with OUD.

> *For patients that successfully completed [HCV treatment], it was the first major win for them. They realized that they had an active role in their recovery… Having a personal win that they could own was profound and had long term effects.*

An HCV cure removed the risk of cirrhosis and improved substance use.

> *He [trial PI] would actually use 'cure', and that's a 4-letter word we don't use in the [methadone] program.*
>
> *The message to take home is that once the patient initiates the treatment, you have a high probability that the patient is actually going to stick with it and obtain an HCV cure.*
>
> *Patients valued the HCV cure because it promoted self-confidence, and it enabled them to improve their health and lives.*

XSL•FO
RenderX

Among cured participants, we observed significant decreases in substance use, as assessed through the drug abuse screening test [31], with minimal HCV reinfections [13]. An HCV cure enabled the treatment team to reinforce adherence to OUD treatment.

> *When the patients progress through the study, the counselors and the treatment team will often utilize that [HCV cure] as a pivot with the patient… It was a wonderful tool that the treatment team used to support a patient's recovery.*

To engage people with OUD in research, we pursued approaches to build trust, such as integrating facilitated telemedicine encounters into the supportive, destigmatizing, and trusting environment of the OTP [32,33]. The foundation of facilitated telemedicine is patient-centeredness. Patients endorsed facilitated telemedicine because they realized the profound effects of an HCV cure.

> *The OTP became an extension of their family… A place of trust.*

People with OUD frequently encounter stigma [34] and feel ostracized in conventional health care settings. Stigma is a major impediment to health care access.

> *They describe themselves as addicts, dirty or clean. Patients blame themselves for contracting a lethal disease [i.e., HCV]. They face stigma in conventional healthcare settings outside of the OTP.*
>
> *We had tried multiple attempts to refer out [for HCV]… It was really met dismally. For some, it was logistics. They had a hard time getting to the other location. But for a lot of them, it was stigma.*

The approaches we used to maintain patient-centeredness are described in Multimedia Appendix 1. The growth of patient-centeredness in the treatment of substance use disorders has been gaining considerable momentum [35].

> *The vision of the agency [OTP] is dedicated to transforming the perception of addiction and behavioral health disorders by bringing dignity and respect to the lives, families, and communities we serve.*

Theme 2: Sociotechnical Approaches Expand Health Care Access for People With OUD

The sociotechnical system framework, which we used to evaluate the facilitated telemedicine intervention, is specifically designed to address challenges in telemedicine implementation in adaptive, complex health care systems, such as OTPs [29]. The facilitation aspect is the social component of facilitated telemedicine and was critical when designing digital health solutions for people with OUD.

> *[We had] to think about cultural and trust issues that may diminish its [telemedicine's] acceptability.*
>
> *Case managers were communicating information about the study, enrollment, retention, and all things hepatitis C beyond the study.*
>
> *[We had to] gain trust in the OTP to participate in patient-centered telemedicine encounters.*

An OTP leader explained the benefits of facilitated telemedicine and his reason for pursuing the intervention. He described the 4 pillars that guided his leadership strategy.

> *Our corporate philosophy is recognizing and overcoming OTP-related challenges like continuing relationships and addressing patient needs.*
>
> *Patient care, education, advocacy, and research... Telemedicine offers a number of opportunities to enhance value-based clinical care.*

The social structure within the OTP promoted trial enrollment and retention. Trust extended to the multidisciplinary team.

> *Early on, the treatment team did a lot of priming of the pump regarding recruitment… I thought it was important to provide education about the study to everyone because the trust equity that the patients have certainly was to the doctors, nurses, and counselors. But sometimes, their strongest relationships were with security staff or a business office person that they saw every day.*

During the question-and-answer periods, several audience members asked about the requirements to replicate the facilitated telemedicine intervention, including social parameters.

> *Remember the eligibility requirements [for the trial]. You had to be in [OUD] treatment for some period of time [6 months] so that means that there was a relationship that was developed between that patient and their counselor, as well as the employees of our program, who are related to the study. There was a relationship that had started to flourish at the time prior to [trial] enrollment.*

The other aspect of facilitated telemedicine is the technical component. We mitigated digital divide issues, such as technology access, digital health literacy, and broadband access, by fully embedding telemedicine encounters into the OTP. During the COVID-19 pandemic and lockdown, facilitated telemedicine became an approach to maintain patient engagement.

> *If one is thinking about adding a telehealth solution, one needs to consider the digital divide. Specific concerns would be access to the internet, digital health literacy, and medical-behavioral integration with virtual care.*
>
> *COVID-19 tethered the consumer away from the clinic… Telemedicine became the umbilicus; it brought the patient back.*

We developed workflows to extract patient-level data on social factors to support facilitated telemedicine implementation as well as for identifying social factors that might be associated with HCV treatment uptake and cure. These tools enabled cross-domain collaboration as well as played crucial roles in ensuring reproducibility and in optimizing clinical and social data acquisition and their integration [19]. An interest in the

XSL•FO
RenderX

use of artificial intelligence as a component of OTP operations was discussed during the question-and-answer period.

> *Recently, large language models have shown the capacity to process vast amounts of data to generate insights... We are interested in how these models can aid in the identification of social factors… to explore relationships between social factors and pursuit of HCV [treatment].*

These approaches are permissive for the direct inclusion of social factors and substance use variables in the patient's treatment plan.

> *[The analytical goal is] to emphasize social factors that can be integrated into the treatment approach.*

Facilitated telemedicine successfully manages the digital divide by addressing both social and technical challenges faced by people with OUD, as substantiated in Table 2.

Table . Sociotechnical system framework for facilitated telemedicine supported by workshop examples. Table adopted from study by Talal et al [29].

Sociotechnical system framework constructs	Definition	Workshop examples	Sociotechnical component
People	Clinicians, patients, administrators, researchers, or software developers.	"Case managers were communicating information about the study, enrollment, retention, and all things hepatitis C beyond the study."	Social
		"[We had to] gain trust in the OTP[a] to participate in patient-centered telemedicine encounters."	
Internal organizational features	Internal organizational policies, procedures, and institutional culture	"Our corporate philosophy is recognizing and overcoming OTP-related challenges like continuing relationships and addressing patient needs... Telemedicine offers a number of opportunities to enhance value-based clinical care."	Social
Hardware and software computing infrastructure	Technical requirements, (eg, hardware, equipment, devices, broadband)	"If one is thinking about adding a telehealth solution, one needs to consider the digital divide."	Technical
Human-computer interface	Comfort level and overall user-friendliness of the interface	"COVID-19 tethered the consumer away from the clinic… Telemedicine became the umbilicus; it brought the patient back."	Sociotechnical
Workflow and communication	Logistics surrounding the proper functioning of the system and ensuring cohesive communication	"I thought it was important to provide education about the study to everyone because the trust equity that the patients have certainly was to the doctors, nurses, and counselors."	Sociotechnical

[a]OTP: opioid treatment program.

Theme 3: Facilitated Telemedicine Supports Pragmatic Research Emphasizing People With OUD

The tremendous efficacy of DAAs promoted support for pragmatic effectiveness studies for HCV. Simultaneously, the trial sponsor emphasized pragmatism in research.

> *There was real interest in hepatitis C treatment and that opened the channels for funding OUD studies.*
>
> *The studies had to be done in the real world, and they had to include the people who are implementing the treatment and the patient-centered aspects of telehealth.*

Simultaneously, however, various stakeholders voiced concerns about engaging people with OUD in research.

> *The concerns regarding funding studies in people with OUD very much reflected the ones that payers and providers had… Grave concerns about treating this particular population because of its reputation for being unstable, that they would not complete treatment, would get reinfected, or would disappear.*

Historically, clinical research was largely focused on academia and was considered complete when the study results were published.

> *Studies were developed by researchers and then evaluated by other researchers… The problem with academic research was that if there was nobody who wanted to implement what was found effective, then nothing happened with it.*

Pragmatic studies achieve their full potential when they have long-lasting impact. In our case, 4 of 8 sites have continued to use facilitated telemedicine for evaluation of complex cases.

We used a variety of techniques, including the patient advisory committee, peer pipeline, and learning lunches, to understand and address facilitated telemedicine implementation challenges

XSL•FO
RenderX

within the OTP. The patient advisory committee facilitated participant understanding of facilitated telemedicine through recommendations to augment digital literacy. For example, we developed an educational video explaining facilitated telemedicine that we showed to participants during recruitment discussions.

> *We had a patient advisory committee that… formed a very cohesive bond and served as a mechanism for discussion and feedback on study-related interventions even before we implemented them.*

Alongside the patient advisory committee, there was a peer pipeline where patients had the opportunity to endorse trial participation to their peers.

> *We had a peer pipeline… Patients spoke to other patients that promoted subsequent engagement and maintained retention.*
>
> *They [patients] became the strongest advocates for enrollment into the study.*

We discovered that learning lunches were a particularly important mechanism to provide staff education, to facilitate OTP staff and research team interactions, and to receive feedback for potential course correction of trial procedures [9].

> *Learning lunches were open for the entire multidisciplinary team… Patient care providers, billing department and our security staff because many of them were seeing the patients every day… Having that opportunity to work through some of those knots was very instrumental to the [trial's] success.*

In addition to OTP leadership, patients, and trial sponsors, stakeholder engagement extended to the clinical and nonclinical staff. It was important to regularly communicate with OTP staff as they managed OTP workflows. Ongoing stakeholder engagement, combined with program funding, was synergistic toward facilitated telemedicine implementation, as explained by OTP leadership.

> *I've come to appreciate the need to have continuing relationships… Those relationships are not just research, but they are also economic. When you can get funding from a research institute to support what you do, it makes a big difference in day-to-day operations.*

HCV treatment access is a health equity priority [7]. Therefore, funding and policy should be evidence-based. Future research funding should focus on the scaling and sustainability of facilitated telemedicine to promote health care access.

> *Public policy is not always driven by science or clinical practice.*
>
> *What we need are some creative solutions to the access problem. Telehealth takes care of part of it, but we also need to think about where telehealth doesn't work… We still have the great digital divide. There are people who have no access to the internet... You have to have some familiarity with the technology… Not everybody has digital literacy.*

Discussion

Principal Results

Facilitated telemedicine, through its transcendence of geographical and temporal boundaries, resulted in remarkably high HCV cure rates when integrated into 12 OTPs throughout New York State. Implementing the intervention, however, required convincing stakeholders of the benefits of facilitated telemedicine. We conducted a pragmatic randomized controlled trial in people with OUD that satisfactorily addressed the digital divide in an underserved population and that successfully implemented a novel health care delivery modality using a complex design. We found that open and honest stakeholder discussions were usually sufficient to overcome initial resistance regarding integrating facilitated telemedicine into OTPs. Conducting clinical research in OTPs and recruiting people with OUD were novel concepts when the trial was proposed because neither typically participated in research. An initial challenge was convincing stakeholders to endorse a trial directed toward the setting and population. An additional challenge was whether true uncertainty (ie, equipoise) existed between the interventions being tested in the clinical trial. Since it was difficult to establish whether true uncertainty existed, we used the stepped wedge design [36]. The concept of using the stepped wedge design was also considered a potential complexity given its stringent enrollment and timing requirements. Patient engagement in a trusting and destigmatizing environment by staff, who had the patients' best interests as their priority, mitigated several of these concerns. From these perspectives, the OTP environment enabled us to fully integrate HCV treatment through facilitated telemedicine into its workflows. After the trial's conclusion, we have continued facilitated telemedicine for evaluation of complex cases in 4 sites. In 8 sites, facilitated telemedicine served as a bridge to onsite HCV treatment by OTP staff. From our investigation, we have learned that facilitated telemedicine can increase health care access using a value-based care approach.

Digital Divide

The digital divide remains a challenge to the scalability of facilitated telemedicine. The digital divide refers to the unequal access to digital technology, internet provision, and digital literacy that worsens inequality around access to information and resources [8]. In the design of facilitated telemedicine, we considered social and technical aspects, which were facilitated by a case manager. During telemedicine encounters and videoconferencing for patient advisory committee meetings, the case manager facilitated equipment operation. The case manager also addressed patients' questions and concerns and was an information conduit for any HCV-related issues. The importance of considering the social aspects of digital connectivity was demonstrated most poignantly during the COVID-19 pandemic and lockdown, when patient advisory committee meetings occurred weekly (instead of quarterly) through videoconferencing. The committee meetings became an extremely important approach to maintain patient engagement. In resource-limited settings, OTP staff can assume the role of HCV champion. Therefore, facilitated telemedicine has the potential to reach underserved populations (ie, migrants,

XSL•FO
RenderX

incarcerated individuals, and the homeless) and address the digital divide through specific approaches as illustrated (Table 3). The importance of the social aspects when increasing digital health care access to underserved populations has been investigated among people who use drugs in Athens, Greece [26,37]. Participants emphasized the importance of trust-building, including several suggestions to enhance trust through digital means, during focus group discussions. More importantly, social interaction and relationship building (ie, empathy, trust, support, and transparency) contributed to the patient-centeredness of facilitated telemedicine. Therefore, telemedicine without the "facilitation" aspect would be devoid of the social component that is crucial for high-quality, patient-centered care [24], especially for populations that often experience stigma from various sources. In this context, the social component was largely delivered by the case managers who facilitated the telemedicine encounters, educated patients, and served as patient advocates. Other facilitated telemedicine approaches have employed staff or peers to facilitate the telemedicine encounters [38-40]. The sociotechnical approach also has the potential to present the OTP as a "full-service garage" that can address OUD-related infectious diseases, mental health issues, common comorbidities, and social factors. The novelty of facilitated telemedicine remains the integration of the social and technical aspects for underserved populations.

Table . Lessons learned through the implementation of facilitated telemedicine to overcome the digital divide.

Topic	Lessons learned and recommendations
Medical	• Integrate HCV[a] and opioid use disorder treatment • Address medical comorbidities
Social	• Minimize stigma against underserved populations • Maintain patient-centeredness • Promote trust • Address competing priorities • Provide patient education • Employ case manager as patient advocate and patient navigator
Technical	• Facilitate telemedicine encounters • Promote access to technology • Ensure sufficient broadband strength • Negate digital literacy challenges
Lessons for pragmatic research	• Promote learning lunches and establish patient advisory committees • Reinforce the importance of HCV cure by OTP[b] staff • Disseminate information through peer pipeline • Engage stakeholders • Emphasize underserved population • Consider telemedicine sustainability through reimbursement

[a]HCV: hepatitis C virus.

[b]OTP: opioid treatment program.

Pragmatic research related to facilitated telemedicine implementation, conventionally conducted as explanatory clinical trials, sought to evaluate an intervention while controlling for multiple factors [41]. Within the past decade, increasing value has been placed on pragmatic research, including stakeholder engagement, implementation evaluation in real-world settings, and sustainability of effective interventions following study completion (Table 4).

XSL•FO
RenderX

Table . Comparison between explanatory and pragmatic research related to facilitated telemedicine implementation. Table adopted from studies by Le-Rademacher et al [41] and Krist et al [42].

Domain	Explanatory research	Pragmatic research	Facilitated telemedicine examples
Impact	To gain scientific knowledge	To make real-world changes	Facilitated telemedicine overcoming digital divide
Patient relevance	Patients are typically solely research participants	Usually patient-centered research matters to patients and caregivers; patient involvement in all study phases	Patients readily view the advantages of facilitated telemedicine
Stakeholder involvement	Largely investigator focused	Multiple stakeholders engaged in all study phases	OTP[a] staff referred patients for facilitated telemedicine
Research design	Extensive exclusion and inclusion criteria often lead to homogenous populations and settings	Real-world settings, populations, contextual assessments; usually parallel or stepped-wedge design usually with a high degree of heterogeneity.	Research in opioid treatment programs with usual workflows, development of a statewide telemedicine network.
Measures	Validated measures that minimize bias; focus on internal consistency	Practical to real-world and low-resource settings; matter to patients	HCV[b] cure, satisfaction with healthcare delivery through telemedicine
Costs	Infrequently collected	Assess cost of intervention	Telemedicine reimbursement
Data sources	Data collection from patients as part of the trial.	Data collected at patient and site level; can use existing datasets (ie, health records, administrative data, and patient lists)	Remote access to electronic health records, case managers faxing laboratory reports
Analyses	Can be per protocol	Intention-to-treat, relevant to multiple stakeholders; goal is understanding implementation process	Patient satisfaction with health care delivery through facilitated telemedicine.
Intervention adherence and follow-up requirements	Close monitoring and follow-up	Can be as usual care	Case managers ensured adherence to study and clinical procedures.
Sustainability	Emphasis on efficacy under ideal circumstances	Emphasis on effectiveness in real-world settings	HCV management of complex cases through telemedicine at 4 sites, on-site HCV treatment at 8 sites after study completion.

[a]OTP: opioid treatment program.

[b]HCV: hepatitis C virus.

Health care delivery approaches, such as facilitated telemedicine, are most effectively investigated through pragmatic as opposed to explanatory trials. The typical goal of explanatory research is to evaluate the efficacy of an intervention in a controlled setting, whereas pragmatic research evaluates its effectiveness in routine clinical practice settings to maximize applicability, generalizability, and transferability [43]. Explanatory research typically measures clinical symptoms or biological markers, whereas pragmatic research focuses on multiple outcomes, many of which are patient-centered. Pragmatic research emphasizes stakeholder involvement, real-world settings and populations, and outcomes that are relevant to patients and stakeholders. Therefore, data collection is typically more laborious and is required from patients, sites, and communities [41]. In addition, since sites where pragmatic studies are conducted are often part of a health care system, they typically share personnel, policies, and resources. Thus, the assumption that study sites are independent of each other, typically observed in explanatory studies, may be violated in pragmatic trials. Proper analysis of pragmatic trial data may require complex statistical models that can account for correlated responses between clusters. In the trial, we were fortunate to have a negligible correlation between our study sites [36]. Furthermore, we developed a comprehensive dataset that permitted the analysis of substance use and social factors in relation to HCV treatment initiation and cure [44].

Strengths and Limitations

Strengths of the workshop include multidisciplinary speakers who directed a multiyear trial of facilitated telemedicine. The presentations have been substantiated with extensive staff and patient interviews [3,22-24]. Another strength was partnering with national organizations (ie, AATOD, NASTAD, and PCORI) that represent patients with HCV and OUD and their caregivers. The workshop provided the opportunity to disseminate the findings of the trial to a national audience that specializes in OUD care. A minor limitation is the nonrecognition of all spoken words by the recording and transcription program; less than approximately 1% of spoken words were not picked up by the recording. In addition, a patient-participant was unable to attend the workshop. Fortunately, the patient's voice was communicated through a presentation about patient perspectives derived from a recent publication [3]. Given that the age of 18 years or older was a

XSL•FO
RenderX

trial inclusion criterion, we are unable to generalize the findings from this investigation to younger-aged individuals.

Conclusions

We developed facilitated telemedicine as a patient-centered intervention designed to expand DAA access among people with OUD. Because people with OUD may have difficulty crossing the digital divide, we used facilitated telemedicine to engage and retain the population in HCV care. A multidisciplinary group of investigators hosted a workshop to disseminate knowledge of facilitated telemedicine for HCV care integrated into OTPs. From the workshop presentations, we identified 3 themes. First, patient-centered care promotes HCV treatment for underserved populations through facilitated telemedicine (theme 1). Second, sociotechnical approaches expand health care access for people with OUD (theme 2). While the technical aspect of facilitated telemedicine was conducive to overcoming geographical and temporal obstacles, we had to consider the social aspect in the delivery of HCV care through digital approaches to people with OUD. These were primarily addressed by integrating HCV telemedicine encounters into the destigmatizing, supportive, and trusting OTP environment and by facilitation through a case manager.

Third, facilitated telemedicine supports pragmatic research emphasizing people with OUD (theme 3). Initial resistance of OTP leadership and frontline staff regarding facilitated telemedicine was effectively mitigated through engagement of a patient advisory committee, a peer pipeline operating within the OTP, and learning lunches. These approaches facilitated participant and OTP staff input on approaches to improve facilitated telemedicine. Addressing the digital divide requires developing patient-centered, sociotechnical innovations for health care delivery and equity directed toward underserved populations.

Acknowledgments

We would like to acknowledge the American Association for the Treatment of Opioid Dependence, its 2024 conference contributors and participants. We would like to thank Michael E Cain, MD, for his review of the manuscript.

This work was supported by a Patient-Centered Outcomes Research Institute (PCORI) Award (IHS-1507- - 31640) and partially supported by the Troup Fund of the Kaleida Health Foundation (to AHT). The statements in this work are solely the responsibility of the authors and do not necessarily represent the views of PCORI, its Board of Governors or Methodology Committee. Study funders were not involved in data collection, analysis, or manuscript preparation.

LSB was affiliated with the START Treatment and Recovery Centers, Brooklyn, New York, at the time of the study and is currently affiliated with the Weill Cornell Medical College.

EJH was affiliated with the PCORI, Washington, DC, at the time of the study and is currently affiliated with Research Triangle Park.

Data Availability

The datasets generated or analyzed during this study are not publicly available due confidentiality concerns.

Authors' Contributions

Conceptualization was contributed by AHT, BNR, and EJH. Investigation was contributed by all authors. Formal analysis and writing—original draft preparation were contributed by AHT, AD, and EJH. Data curation was contributed by AHT and AD. Writing—review and editing was contributed by all authors. Visualization was contributed by AHT, AD, MM, and EJH. Supervision and funding acquisition were handled by AHT. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

AHT reported receiving nonfinancial support from Abbott Laboratories and grants from Gilead Sciences, Novo Nordisk, AstraZeneca, and Salix. AHT has served as a committee/advisor for Gilead, AbbVie, Novo Nordisk, and Madrigal. He also serves as Medical Director for Empath Medical, Inc. MM reported receiving grants from PCORI and the Kaleida Health Foundation. No other conflicts of interest were reported. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Multimedia Appendix 1
Supplementary materials.
[DOCX File, 40 KB - publichealth_v11i1e68854_app1.docx]

Multimedia Appendix 2
Study protocol.
[DOCX File, 214 KB - publichealth_v11i1e68854_app2.docx]

XSL•FO
RenderX

References

1. Artenie A, Stone J, Fraser H, et al. Incidence of HIV and hepatitis C virus among people who inject drugs, and associations with age and sex or gender: a global systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2023 Jun;8(6):533-552. [doi: 10.1016/S2468-1253(23)00018-3] [Medline: 36996853]
2. Ghany MG, Morgan TR, AASLD-IDSA Hepatitis C Guidance Panel. Hepatitis C guidance 2019 update: American association for the study of liver diseases-infectious diseases society of America recommendations for testing, managing, and treating hepatitis C virus infection. Hepatology 2020 Feb;71(2):686-721. [doi: 10.1002/hep.31060] [Medline: 31816111]
3. Talal AH, Jaanimägi U, Dharia A, Dickerson SS. Facilitated telemedicine for hepatitis C virus: addressing challenges for improving health and life for people with opioid use disorder. Health Expect 2023 Dec;26(6):2594-2607. [doi: 10.1111/hex.13854] [Medline: 37641398]
4. Eckhardt B, Mateu-Gelabert P, Aponte-Melendez Y, et al. Accessible hepatitis C care for people who inject drugs: a randomized clinical trial. JAMA Intern Med 2022 May 1;182(5):494-502. [doi: 10.1001/jamainternmed.2022.0170] [Medline: 35285851]
5. Wester C, Osinubi A, Kaufman HW, et al. Hepatitis C virus clearance cascade - United States, 2013-2022. MMWR Morb Mortal Wkly Rep 2023 Jun 30;72(26):716-720. [doi: 10.15585/mmwr.mm7226a3] [Medline: 37384551]
6. Gnanapandithan K, Ghali MP. Self-awareness of hepatitis C infection in the United States: a cross-sectional study based on the National Health Nutrition and Examination Survey. PLoS ONE ;18(10):e0293315. [doi: 10.1371/journal.pone.0293315] [Medline: 10597475]
7. Valdiserri RO, Koh HK, Ward JW. Overcome health inequities to eliminate viral hepatitis. JAMA 2023 May 16;329(19):1637-1638. [doi: 10.1001/jama.2023.5381] [Medline: 37067806]
8. Saeed SA, Masters RM. Disparities in health care and the digital divide. Curr Psychiatry Rep 2021 Jul 23;23(9):61. [doi: 10.1007/s11920-021-01274-4] [Medline: 34297202]
9. Talal AH, George SJ, Talal LA, et al. Engaging people who use drugs in clinical research: integrating facilitated telemedicine for HCV into substance use treatment. Res Involv Engagem 2023 Aug 2;9(1):63. [doi: 10.1186/s40900-023-00474-x] [Medline: 37533127]
10. Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. Qual Saf Health Care 2010 Oct;19 Suppl 3(Suppl 3):i68-i74. [doi: 10.1136/qshc.2010.042085] [Medline: 20959322]
11. Talal AH, McLeod A, Andrews P, et al. Patient reaction to telemedicine for clinical management of hepatitis C virus integrated into an opioid treatment program. Telemed J E Health 2019 Sep;25(9):791-801. [doi: 10.1089/tmj.2018.0161] [Medline: 30325701]
12. Talal AH, Andrews P, Mcleod A, et al. Integrated, co-located, telemedicine-based treatment approaches for hepatitis C virus management in opioid use disorder patients on methadone. Clin Infect Dis 2019 Jul 2;69(2):323-331. [doi: 10.1093/cid/ciy899] [Medline: 30329042]
13. Talal AH, Markatou M, Liu A, et al. Integrated hepatitis C-opioid use disorder care through facilitated telemedicine: a randomized trial. JAMA 2024 Apr 23;331(16):1369-1378. [doi: 10.1001/jama.2024.2452] [Medline: 38568601]
14. Advancing viral hepatitis screening and treatment in opioid treatment settings – models & resources. Presented at: AATOD (American Association for the Treatment of Opioid Dependence); 2024 URL: https://aatod.eventscribe.net/fsPopup.asp?efp=TVRNWUhDSVgyMTIwNQ&PresentationID=1383218&rnd=0.4921656&mode=presInfo [accessed 2025-07-07]
15. Treating opioid use disorder: so much more than medication. 2024 Presented at: AATOD (American Association for the Treatment of Opioid Dependence) 2024 Conference URL: https://aatod.eventscribe.net/index.asp
16. NASTAD: About Us. National Association of State and Territorial AIDS Directors. URL: https://nastad.org/about [accessed 2025-10-25]
17. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
18. Medications for the treatment of opioid use disorder. Substance Abuse and Mental Health Services Administration. URL: https://public-inspection.federalregister.gov/2022-27193.pdf [accessed 2024-03-03]
19. Markatou M, Kennedy O, Brachmann M, Mukhopadhyay R, Dharia A, Talal AH. Social determinants of health derived from people with opioid use disorder: Improving data collection, integration and use with cross-domain collaboration and reproducible, data-centric, notebook-style workflows. Front Med (Lausanne) 2023;10:1076794. [doi: 10.3389/fmed.2023.1076794] [Medline: 36936205]
20. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
21. Terry G, Hayfield N, Clarke V, Braun V. Thematic analysis. In: SAGE Research Methods: The SAGE Handbook of Qualitative Research in Psychology: SAGE Publications Ltd; 2017. [doi: 10.4135/9781526405555.n2]
22. Talal AH, Sofikitou EM, Wang K, Dickerson S, Jaanimägi U, Markatou M. High satisfaction with patient-centered telemedicine for hepatitis C virus delivered to substance users: a mixed-methods study. Telemed J E Health 2023 Mar;29(3):395-407. [doi: 10.1089/tmj.2022.0189] [Medline: 35925809]

XSL•FO
RenderX

23. Ventuneac A, Dickerson SS, Dharia A, George SJ, Talal AH. Scaling and sustaining facilitated telemedicine to expand treatment access among underserved populations: a qualitative study. Telemed J E Health 2023 Dec;29(12):1862-1869. [doi: 10.1089/tmj.2022.0534] [Medline: 37252770]
24. Dickerson SS, George SJ, Ventuneac A, Dharia A, Talal AH. Care integration for hepatitis C virus treatment through facilitated telemedicine within opioid treatment programs: qualitative study. J Med Internet Res 2024 Jun 12;26:e53049. [doi: 10.2196/53049] [Medline: 38865703]
25. Gonzalez CJ, Kapadia SN, Niederdeppe J, et al. The state of hepatitis C elimination from the front lines: a qualitative study of provider-perceived gaps to treatment initiation. J GEN INTERN MED 2024 Sep;39(12):2268-2276. [doi: 10.1007/s11606-024-08807-6] [Medline: 38782810]
26. Papalamprakopoulou Z, Ntagianta E, Triantafyllou V, et al. Telehealth to increase healthcare access; perspectives of people who use drugs. BMC Med Inform Decis Mak 2024 Oct 19;24(1):306. [doi: 10.1186/s12911-024-02718-6] [Medline: 39425182]
27. Talal AH, Jaanimägi U, Davis K, et al. Facilitating engagement of persons with opioid use disorder in treatment for hepatitis C virus infection via telemedicine: stories of onsite case managers. J Subst Abuse Treat 2021 Aug;127:108421. [doi: 10.1016/j.jsat.2021.108421] [Medline: 34134875]
28. Sundler AJ, Lindberg E, Nilsson C, Palmér L. Qualitative thematic analysis based on descriptive phenomenology. Nurs Open 2019 Jul;6(3):733-739. [doi: 10.1002/nop2.275] [Medline: 31367394]
29. Talal AH, Sofikitou EM, Jaanimägi U, Zeremski M, Tobin JN, Markatou M. A framework for patient-centered telemedicine: application and lessons learned from vulnerable populations. J Biomed Inform 2020 Dec;112:103622. [doi: 10.1016/j.jbi.2020.103622] [Medline: 33186707]
30. Reynolds A. Patient-centered care. Radiol Technol 2009;81(2):133-147. [Medline: 19901351]
31. Skinner HA. The drug abuse screening test. Addict Behav 1982;7(4):363-371. [doi: 10.1016/0306-4603(82)90005-3] [Medline: 7183189]
32. Earnshaw V, Smith L, Copenhaver M. Drug addiction stigma in the context of methadone maintenance therapy: an investigation into understudied sources of stigma. Int J Ment Health Addict 2013 Feb 1;11(1):110-122. [doi: 10.1007/s11469-012-9402-5] [Medline: 23956702]
33. Islam MM. Missed opportunities for hepatitis C testing and other opportunistic health care. Am J Public Health 2013 Dec;103(12):e6. [doi: 10.2105/AJPH.2013.301611] [Medline: 24134357]
34. Harris M, Guy D, Picchio CA, White TM, Rhodes T, Lazarus JV. Conceptualising hepatitis C stigma: a thematic synthesis of qualitative research. Int J Drug Policy 2021 Oct;96:103320. [doi: 10.1016/j.drugpo.2021.103320] [Medline: 34261587]
35. Person-Centered Care Guidance for OASAS Certified Programs. New York State Office of Addiction Services and Supports. URL: https://oasas.ny.gov/system/files/documents/2020/01/oasasperson-centeredcareguidance.pdf [accessed 2021-09-09]
36. Talal AH, Markatou M, Sofikitou EM, et al. Patient-centered HCV care via telemedicine for individuals on medication for opioid use disorder: Telemedicine for Evaluation, Adherence and Medication for Hepatitis C (TEAM-C). Contemp Clin Trials 2022 Jan;112:106632. [doi: 10.1016/j.cct.2021.106632] [Medline: 34813962]
37. Papalamprakopoulou Z, Ntagianta E, Triantafyllou V, et al. Breaking the vicious cycle of delayed healthcare seeking for people who use drugs. Harm Reduct J 2025 Mar 5;22(1):27. [doi: 10.1186/s12954-025-01166-3] [Medline: 40045378]
38. Spencer H, Leichtling G, Babiarz J, et al. Peer-assisted telemedicine for hepatitis C (PATHS): process evaluation results from a State Opioid Response-funded program. J Subst Use Addict Treat 2024 Dec;167:209510. [doi: 10.1016/j.josat.2024.209510] [Medline: 39243982]
39. Morris MD, McDonell C, Luetkemeyer AF, Thawley R, McKinney J, Price JC. Community-based point-of-diagnosis hepatitis C treatment for marginalized populations: a nonrandomized controlled trial. JAMA Netw Open 2023 Oct 2;6(10):e2338792. [doi: 10.1001/jamanetworkopen.2023.38792] [Medline: 37862013]
40. Kim RG, McDonell C, McKinney J, Catalli L, Price JC, Morris MD. Staff-facilitated telemedicine care delivery for treatment of hepatitis C infection among people who inject drugs. Healthcare (Basel) 2024 Mar 25;12(7):715. [doi: 10.3390/healthcare12070715] [Medline: 38610138]
41. Le-Rademacher J, Gunn H, Yao X, Schaid DJ. Clinical trials overview: from explanatory to pragmatic clinical trials. Mayo Clin Proc 2023 Aug;98(8):1241-1253. [doi: 10.1016/j.mayocp.2023.04.013] [Medline: 37536808]
42. Krist AH, Glenn BA, Glasgow RE. Designing a valid randomized pragmatic primary care implementation trial: the my own health report (MOHR) project. Implement Sci 2023;8(73). [doi: 10.1186/1748-5908-8-73] [Medline: 3694031]
43. Ford I, Norrie J. Pragmatic trials. N Engl J Med 2016 Aug 4;375(5):454-463. [doi: 10.1056/NEJMra1510059] [Medline: 27518663]
44. Talal AH, Mukhopadhyay R, Veronesi V, Dharia A, Saraceno G, Markatou M. WED-479 Improved social outcomes after integrated hepatitis C and opioid use disorder treatment. J Hepatol 2024 Jun;80:S680. [doi: 10.1016/S0168-8278(24)01945-7]

Abbreviations

AATOD: American Association for the Treatment of Opioid Dependence
COREQ: Consolidated Criteria for Reporting Qualitative Research

XSL•FO
RenderX

DAA: direct-acting antiviral
HCV: hepatitis C virus
NASTAD: National Alliance of State and Territorial AIDS Directors
OTP: opioid treatment program
OUD: opioid use disorder
PCORI: Patient-Centered Outcomes Research Institute
RCT: randomized controlled trial
SDOH: social determinants of health
UB: University at Buffalo

Edited by A Mavragani; submitted 21.11.24; peer-reviewed by D Walker, S Gonzalez; revised version received 24.05.25; accepted 25.05.25; published 16.07.25.

Please cite as:
Talal AH, Dharia A, Markatou M, Brown Jr LS, Bossert KE, Grubbs Z, Mukhopadhyay R, Ntiri-Reid B, Houtsmuller EJ
Facilitated Telemedicine as a Patient-Centered, Sociotechnical Intervention to Integrate Hepatitis C Treatment Into Opioid Treatment Programs and Overcome the Digital Divide Among Underserved Populations: Qualitative Study
JMIR Public Health Surveill 2025;11:e68854
URL: https://publichealth.jmir.org/2025/1/e68854
doi:10.2196/68854

© Andrew H Talal, Arpan Dharia, Marianthi Markatou, Lawrence S Brown Jr, Kenneth E Bossert, Zakiya Grubbs, Raktim Mukhopadhyay, Boatemaa Ntiri-Reid, Elisabeth J Houtsmuller. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Changes in Epidemiological Characteristics of Varicella and Breakthrough Cases in Ningbo, China, From 2010 to 2023: Surveillance Study

Xingqiang Pan, MPH; Yan Zhang, MPH; Xuefei Zhao, MPH; Dandan Zhang, MPH

Ningbo Center for Disease Control and Prevention, No. 1166 Fanjiangan Road, Haishu District, Ningbo, China

Corresponding Author:
Dandan Zhang, MPH
Ningbo Center for Disease Control and Prevention, No. 1166 Fanjiangan Road, Haishu District, Ningbo, China

Abstract

Background: Varicella is a prevalent respiratory infectious disease. Continuous monitoring is essential to understand evolving epidemiological patterns, particularly given the impact of vaccination and recent nonpharmacological interventions.

Objective: This study aims to monitor the epidemiological characteristics of varicella and the changes in breakthrough cases to inform adjustments in immunization strategies and enhance prevention efforts.

Methods: From 2010 to 2023, varicella incidence was monitored using active (2010-2011) and passive (2012-2023) surveillance methods. Data were obtained from the Chinese Center for Disease Prevention and Control's information system and Ningbo's Immunization Information System. The study period was divided into four intervals to analyze trends. A birth cohort (2009-2013) was established to examine breakthrough cases. A recurrent neural network model was constructed for deep learning analysis of incidence trends and the impact of nonpharmaceutical interventions.

Results: Between 2010 and 2023, a total of 70,163 varicella cases were reported in Ningbo. Seasonal distribution indicated two incidence troughs before 2020 and only one from 2020 to 2023. The predominant age of onset was 10 - 14 years, accounting for 23.93% (16,795/70,163) of cases. From 2010 to 2013, the highest incidence was among children aged 5 - 9 years; from 2014 to 2019, it shifted to those aged 10 - 14 years; and from 2020 to 2023, it was primarily among individuals aged 15 - 19 years. Following the introduction of a second vaccine dose (2014 - 2019), incidence among children younger than 10 years of age decreased, notably by 59.54% in those aged 1 - 4 years. Conversely, incidence among individuals aged 10 years and older increased, particularly by 123.78% in the 15 - 19 years age group, with a significant upward trend (P_{trend}<.001). From 2020 to 2023, although incidence rates increased across age groups 15 years and older, the rise was modest. The average annual incidence rate of breakthrough cases after one vaccine dose was 83.40/100,000 (range, 51.21 - 119.50/100,000), significantly higher than the 24.80/100,000 (range, 17.67 - 32.90/100,000) observed after two doses. However, the incidence of breakthrough cases after the first dose declined following the implementation of the 2-dose program. The median time from vaccination to breakthrough case occurrence was 27 (IQR 17.50 - 48) months. The recurrent neural network model demonstrated high accuracy (mean squared error, 49.96) and indicated that implementation of emergency response and community lockdown measures in early 2020 correlated with a divergence between predicted and actual case numbers, suggesting an impact of nonpharmaceutical interventions on varicella transmission.

Conclusions: The significant shifts in varicella epidemiology between 2010 and 2023 highlight the importance of continuous monitoring and proactive immunization adjustments. We recommend enhanced varicella surveillance focusing on adult populations, and a targeted increase in 2-dose vaccine coverage, particularly in high-risk environments such as high schools and universities.

(JMIR Public Health Surveill 2025;11:e71691) doi:10.2196/71691

KEYWORDS

varicella; breakthrough case; epidemiology; disease surveillance; infectious; monitoring; vaccine; nonpharmacological interventions; machine learning

Introduction

Varicella, a highly contagious disease caused by the varicella-zoster virus, is primarily transmitted through direct contact with the rash of varicella or herpes zoster, as well as by inhaling aerosolized droplets from the respiratory secretions of patients with varicella [1]. While varicella is typically mild and self-limiting in healthy children, it can be life-threatening for immunosuppressed patients, pregnant women, and older people. In high-risk groups, varicella infection may result in severe complications such as pneumonia, encephalitis, hepatitis, organ failure, and acute cerebellar ataxia [2]. According to the World

XSL•FO
RenderX

Health Organization, the annual global burden of varicella is estimated at around 140 million cases, leading to 4,200,000 harsh complications necessitating hospitalization, and resulting in 4200 deaths [3]. Notably, China reported a total of 3,047,715 varicella cases between 2016 and 2019, with an average annual incidence rate of 55.05 per 100,000 [4]. The incidence rate significantly increased from 35.50 per 100,000 in 2016 to 70.14 per 100,000 in 2019, surpassing other legally notifiable infectious diseases like measles and rubella. Since varicella is not a notifiable infectious disease in China, the actual incidence rate is probably higher than the reported rate. This underscores the pressing need for effective prevention strategies to tackle the significant public health impact of varicella infection.

The introduction of the varicella vaccine has been crucial in disease control, significantly reducing the incidence, hospitalizations, and mortality rates associated with varicella [5,6]. The vaccine was first licensed for use in the United States in 1995. In 2006, a 2-dose schedule was recommended for children, with the first dose given at 12 - 15 months and the second dose between 4 and 6 years of age [6]. This immunization strategy has also been adopted by Germany, Australia, and Canada. In contrast, China introduced the varicella vaccine in 1998 but did not include it in the national immunization program, opting for a single-dose strategy targeting children aged 12 - 18 months. Starting in 2014, Ningbo implemented a 2-dose vaccination strategy, with the first dose administered at 12 - 15 months and the second dose at age 3 years and older. Vaccination in Ningbo is voluntary and self-funded, achieving a vaccination rate exceeding 90% for one dose and over 70% for two doses [7,8]. The use of the varicella vaccine in children has led to changes in the epidemiological characteristics of varicella, including a shift toward older age groups [3]. Nonpharmaceutical preventive measures for COVID-19 may also impact the epidemiological features of varicella. Studies indicated that the effectiveness of a single varicella vaccine dose diminishes over time, contrasting with the sustained effectiveness of the 2-dose vaccine [9]. Even if vaccinated, breakthrough cases of varicella can still occur. Although usually less severe, these cases can contribute to the spread of varicella within a community.

The World Health Organization emphasizes the need for thorough varicella surveillance to understand the disease's spread and to guide public health efforts in prevention and control. Since 2009, targeted surveillance efforts have been implemented, establishing specific varicella surveillance points in Ningbo, China, to monitor disease incidence and spread. In 2012, the Ningbo Municipal Bureau of Health required all health care facilities to report varicella cases to the China Information System for Disease Control and Prevention (CISDCP) within 24 hours, and this measure enhanced the accuracy and promptness of data collection. This study aims to use both active and passive surveillance data to analyze the epidemiological features of varicella and breakthrough cases.

Methods

Setting

Ningbo, a subprovincial city in Zhejiang province, East China, is one of the five state-planning cities in the country. Before 2016, Ningbo comprised 6 urban districts (Haishu, Jiangbei, Jiangdong, Zhenhai, Beilun, and Yinzhou), 3 satellite county-level cities (Fenghua, Yuyao, and Cixi), and 2 rural counties (Ninghai and Xiangshan). Following a regional restructuring in 2016, Ningbo now includes 6 urban districts: Haishu, Jiangbei, Zhenhai, Beilun, Yinzhou, and Fenghua; 2 county-level cities: Yuyao and Cixi; and 2 counties: Ninghai and Xiangshan. The city spans 9816.2 km² of land and 8355.8 km² of sea. As of the end of 2022, Ningbo had a population of 9.618 million permanent residents. Since 2005, the Ningbo Center for Disease Control and Prevention has operated the Ningbo Immunization Information System, a computerized system that maintains immunization data and demographic information on vaccines for the Ningbo region [10].

Active Monitoring Only, 2010-2011

The varicella active surveillance project [10] was conducted in Yuyao City, Yinzhou District, and Ninghai County during 2010 and 2011, through a collaboration between the Yuyao Center for Disease Prevention and Control (CDC), Yinzhou CDC, Ninghai CDC, and Ningbo CDC. Yuyao City is located in the western part of Ningbo, on the southern bank of the Yangtze River Delta. Yinzhou is a central district of Ningbo, covering an area of 1380 square kilometers. Ninghai County is situated to the south of Ningbo and spans approximately 1931 square kilometers. The populations of Yuyao, Yinzhou, and Ninghai County were 830,000, 730,000, and 580,000, respectively.

Since mid-2009, local health providers and physicians in the varicella surveillance areas have been required to electronically report any diagnosed varicella case to CISDCP within 24 hours. The epidemiologic information recorded for each case includes age, gender, current home address, household registration, date of symptoms onset, date of diagnosis, vaccine history, and other relevant details. Public health doctors in the surveillance area were also responsible for actively identifying varicella cases in hospitals, community health service centers, schools, and kindergartens on a weekly basis. If an unreported varicella case was found, public health doctors would report it through CISDCP.

Passive Surveillance, 2012-2023

Since 2012, under the requirement of the health administration department in Ningbo City, all medical institutions in the city have been required to report varicella cases through CISDCP. The CDC conducted quarterly investigations on underreporting to improve the quality of reporting.

Varicella Vaccine Immunization Strategy Adjustment in 2014

Ningbo City implemented a 2-dose varicella immunization strategy. Children receive voluntary and self-paid varicella vaccination. The recommended timing for the first dose is at 1

XSL•FO
RenderX

year of age, and the recommended timing for the second dose is between 3 and 4 years of age.

Nonpharmaceutical Intervention Measures

Starting from 2020, various nonpharmaceutical intervention (NPI) measures have been implemented by the government to control the COVID-19 epidemic. These measures include maintaining social distancing, wearing masks, school closures, and community lockdowns, which have had a significant impact on the transmission of other respiratory infectious diseases.

Diagnosis of Varicella and Breakthrough Cases

Varicella diagnostic criteria [11] are characterized by typical clinical manifestations, namely the presence of characteristic rashes that progress from macules to papules and then to vesicles before crusting. It is usually accompanied by fever and systemic symptoms such as malaise. If there are no typical clinical manifestations, the diagnosis needs to be made in conjunction with epidemiological history. The incidence of varicella in Ningbo is determined by dividing the number of reported cases by the corresponding population size. Population data for each age group is sourced from the Ningbo Municipal Public Security Bureau.

A birth cohort was established for children born between 2009 and 2013 who received one dose of varicella vaccine at the surveillance sites. Their varicella incidence was monitored through CISDCP. Varicella vaccine administration information was obtained through Ningbo's Immunization Information System.

Breakthrough varicella is defined as the occurrence of clinically diagnosed varicella in individuals who present with characteristic symptoms or have a documented epidemiological exposure history (as previously described), with symptom onset occurring more than 42 days after vaccination [12]. Furthermore, the rate of breakthrough infections has been delineated as the ratio of breakthrough infections occurring in children who have undergone varicella vaccination.

Data Extraction and Analysis

The varicella cases recorded between the years 2010 and 2023 were sourced from the comprehensive infectious disease surveillance and CISDCP. According to the temporal onset of illness, the cases were stratified into four phases for the purposes of descriptive analysis: 2010 - 2011, 2012 - 2013, 2014 - 2019, and 2020 - 2023. In the context of regional distribution analysis, and owing to specific regional modifications in Ningbo, the interval from 2014 to 2019 was further subdivided into two distinct phases: 2014 - 2015 and 2016 - 2019. Chi-square statistical tests were used to examine the variations in age and gender across the different phases, while trend chi-square tests were used to assess the evolving trends of various age cohorts within each phase.

To evaluate the impact of NPIs on the incidence of chickenpox, we developed an advanced deep-learning framework using recurrent neural networks. The model was trained on a dataset of chickenpox cases recorded between January 1, 2012, and December 31, 2018, and validated using data from January 1, 2019, to December 31, 2019. Predictions generated by the model were compared with actual incidence data from January 1 to June 30, corresponding to the implementation of Level 1 Emergency Response on January 22 and community lockdown measures introduced on February 2. The training process used a sliding window technique to forecast future case numbers, and model performance on the validation dataset was evaluated using metrics such as mean squared error, root mean squared error, and mean absolute error. The final model was then used to predict chickenpox incidence for the period from January 1, 2020, to June 30, 2020. These predictions were analyzed to assess the influence of NPIs by comparing the forecasted incidence with the observed data for the same timeframe.

The Python programming language facilitated the processes of data cleansing, analytical procedures, and data visualization. The 2-sided P values were documented, adhering to a significance threshold of $P<.05$. All analytical procedures were conducted using Python software (version 3.10; Python Software Foundation) [13].

Ethical Considerations

This study used secondary surveillance data on varicella cases obtained from the CISDCP. The study protocol was reviewed and approved by the Biological and Medical Ethics Committee of Ningbo Center for Disease Control and Prevention (approval: 202312), ensuring adherence to ethical guidelines for public health research. The original data collection involved informed consent procedures as per national regulations; the Ningbo CDC Ethics Committee specifically granted approval for this secondary analysis without requiring additional consent from individual patients. All data were fully anonymized and deidentified prior to access and analysis, protecting patient privacy and confidentiality. No personal identifiers were present in the dataset provided to the researchers. No compensation was provided to participants as this study involved the analysis of routinely collected surveillance data. The manuscript and supplementary materials do not contain any images of individual participants. We confirm that all data handling and analysis were conducted in compliance with relevant ethical standards and regulations for public health surveillance.

Results

Pathogenesis Overview

Between 2010 and 2023, a total of 70,163 varicella cases were documented in Ningbo. Within the designated monitoring region, 4753 cases were reported during 2010 - 2011. In the birth cohort from 2009 to 2013, there were 1404 varicella cases, including 912 classified as breakthrough cases. The overall incidence rate for breakthrough cases was 47,149 per 100,000 individuals. The frequency of breakthrough cases increased over time postvaccination, peaking in 2022 at 6307 per 100,000. The annual incidence of varicella rose from 63.89 per 100,000 in 2010 to 98.42 per 100,000 in 2013. It remained relatively stable between 2014 and 2017, ranging from 85.58 to 99.60 per 100,000, before increasing to 121.75 per 100,000 in 2018. A significant decline began in 2020, with the incidence rate dropping from 71.79 per 100,000 to 6389 per 100,000 (Table 1).

XSL•FO
RenderX

Table . The prevalence of varicella in Ningbo, 2010 - 2023

Year	Number of cases	Incidence rates (1/100,000)	
		Incidence of rate	BVR[a] of the birth cohorts
2010	2689	43.63	0.58
2011	3568	57.89	10.52
2012	4949	83.32	24.55
2013	5646	96.48	35.07
2014	5676	90.05	44.42
2015	5698	98.42	32.73
2016	5039	85.59	23.97
2017	5748	99.61	24.55
2018	7026	121.76	35.66
2019	6861	118.90	29.23
2020	4307	71.79	22.80
2021	4759	79.32	30.39
2022	4265	69.30	30.39
2023	3932	63.89	33.32

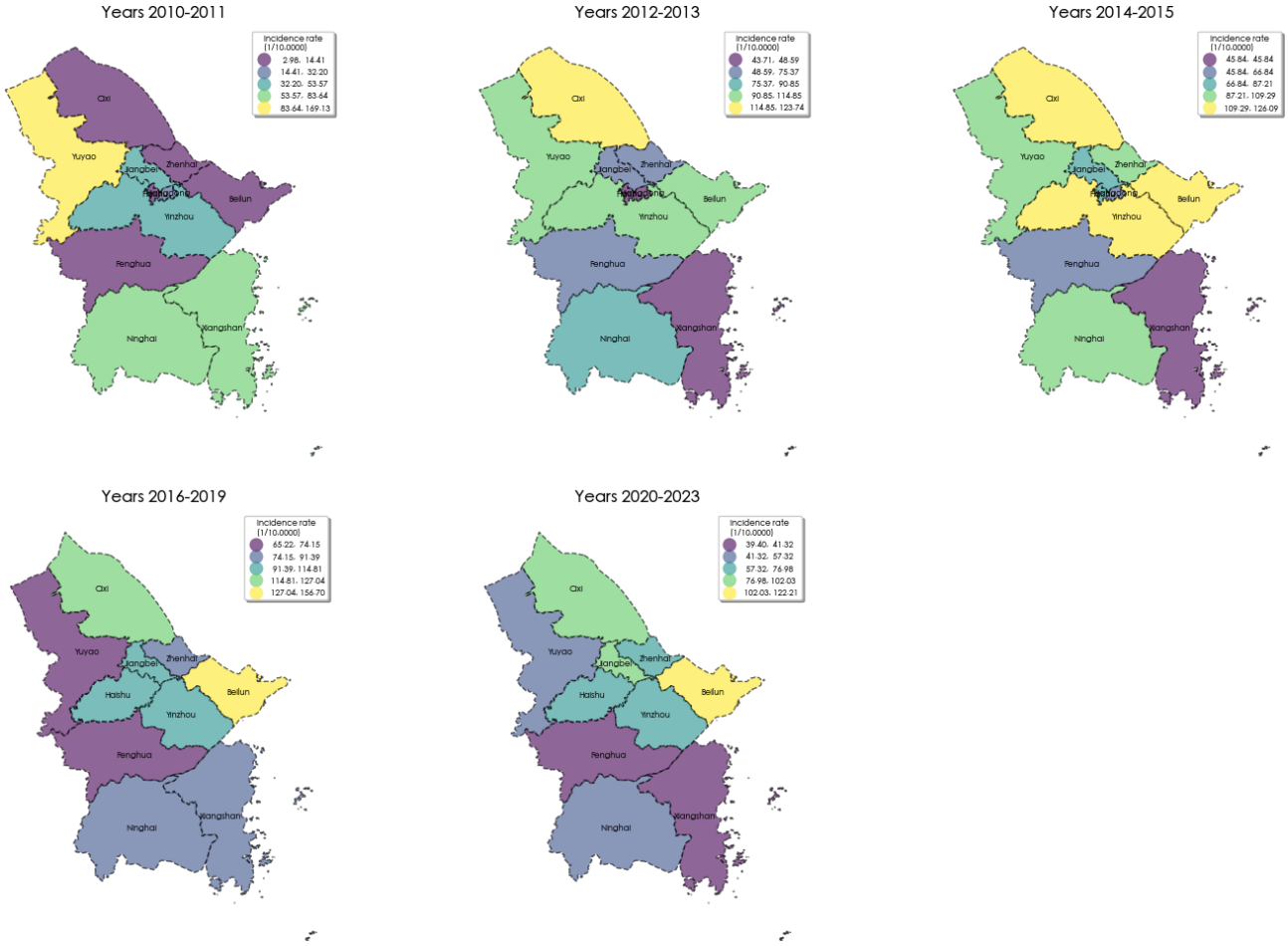
[a]BVR: breakthrough varicella incidence rate.

Spatial Distribution

Between 2010 and 2023, the municipalities with the highest average annual incidence rates of varicella were Beilun District, Yinzhou District, and Cixi City, with rates of 113.44, 105.95, and 99.70 per 100,000, respectively. Figure 1 illustrates the trajectory of these rates across each county and city district. During the active surveillance period, Yuyao City recorded the highest average annual incidence rate of 169.13 per 100,000 in 2010 - 2011. In the subsequent periods of 2012 - 2013 and 2014 - 2015, Cixi City reported the highest rates at 123.73 and 126.09 per 100,000, respectively. From 2016 to 2019 and 2020 to 2023, Beilun District had the highest rates at 156.70 and 122.21 per 100,000, respectively.

XSL•FO
RenderX

Figure 1. The spatial distribution of varicella in Ningbo, 2010 - 2023.



Temporal Distribution

Between the years 2010 and 2023, there were documented instances of varicella cases on a monthly basis in the city of Ningbo. Table 2 illustrates that during the periods 2010 - 2011, 2012 - 2013, and 2014 - 2019, two significant peaks in cases occurred: from May to July and from October to January of the following year. The incidence rates for May to July were 29.84% (1867/6257), 22.38% (2371/10,595), and 25.27% (9108/36,048), while the rates for October to January were 47.16% (2951/6257), 47.16% (2951/6257), and 45.51% (16,406/36,048), respectively. The months with the lowest incidence rates were February and September, with September accounting for 4.12% (258/6257), 4.95% (524/10,595), and 4.49% (1620/36,048), and February accounting for 2.70% (169/6257), 4.91% (520/10,595), and 5.90% (2127/36,048) across the respective years. From 2020 to 2023, the incidence rates in August and September showed no significant variation, while the number of cases from October to January remained high, comprising 45.32% (7823/17,263) of the total cases.

XSL•FO
RenderX

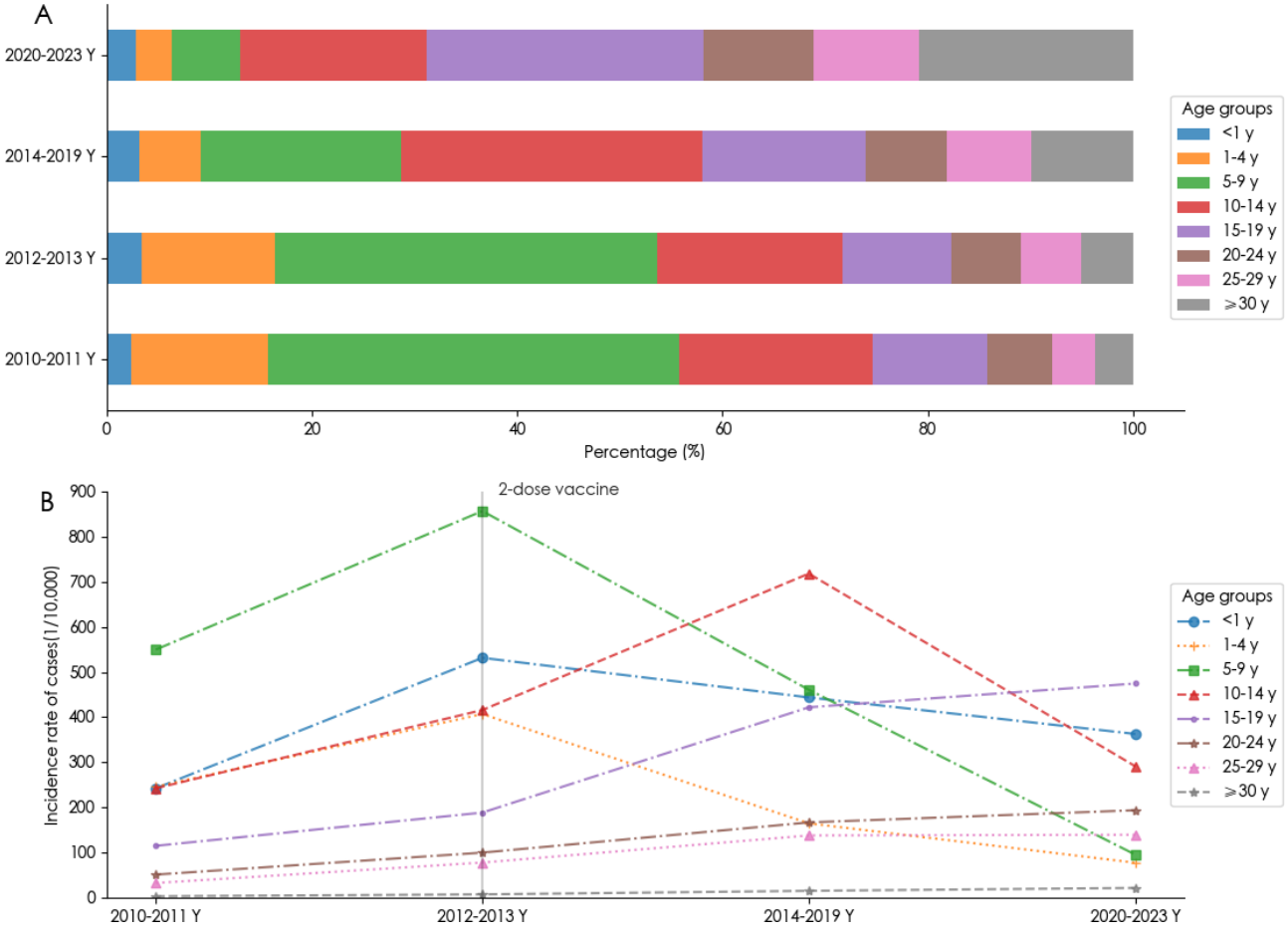
Table . Month of varicella onset by four phases in Ningbo.

Month of disease onset	Number of cases				
	Years 2010 - 2011	Years 2012 - 2013	Years 2014 - 2019	Years 2020 - 2023	Total
1	614	1072	3061	1252	5999
2	258	524	1620	752	3154
3	299	513	2049	899	3760
4	510	695	2506	1061	4772
5	601	816	3340	1456	6213
6	809	793	2899	1268	5769
7	457	762	2869	1394	5482
8	203	548	2232	1284	4267
9	169	520	2127	1326	4142
10	418	996	3287	1808	6509
11	817	1556	4969	2659	10,001
12	1102	1800	5089	2104	10,095

Population Distribution

From 2010 to 2023, the age groups with the highest number of cases were primarily 10 - 14, 5 - 9, and 15 - 19 years, comprising 23.93% (16,795/70,163), 20.86% (14,639/70,163), and 17.42% (12,224/70,163) of the total cases, respectively. Notably, between 2010 and 2011, the 5 - 9 age group had the highest proportion, accounting for 40.10% (2509/6257) and 37.31% (3953/10,595) of the total cases, respectively. In contrast, from 2014 to 2019, the 10 - 14 age group had the highest proportion, reaching 29.34% (10,577/36,048). From 2020 to 2023, the 15 - 19 age group had the highest proportion, accounting for 26.97% (4656/17,263) of the total cases (Figure 2A).

Figure 2. The age distribution of varicella in Ningbo. (A) The composition ratio of age by the phases. (B) The incidence rates by age during the phases.



XSL•FO
RenderX

From an analytical perspective on incidence rates, the age groups younger than 1 year and 5 - 9 years demonstrated a declining trend from 2012 - 2013 to 2020 - 2023, with reductions of 31.70% and 80.96%, respectively. In contrast, the 10 - 14 year age group experienced a significant increase of 72.68% from 2012 - 2013 to 2014 - 2019, followed by a decrease of 29.97% by 2020 - 2023. The age groups 15 - 19 years, 20 - 24 years, 25 - 29 years, and 30 years and older showed consistent upward trends from 2010 - 2011 to 2020 - 2023, with increases of 312.99%, 278.79%, 327.99%, and 575.62%, respectively. These trends were statistically significant, as determined by the trend chi-square test (all P<.001; and Table 3) (Figure 2B).

Table . The incidence rates of varicella by age and gender in Ningbo.

Characteristics	Years 2010 - 2011	Years 2012 - 2013	Years 2014 - 2019	Years 2020-2023	Difference between years 2010 - 2011 and 2020 - 2023		Difference between years 2012 - 2013 and 2014 - 2019		Difference between years 2012 - 2013 and 2020-2023		Difference between years 2014 - 2019 and 2020-2023	
					%	*P* value	%	*P* value	%	*P* value	%	*P* value
Gender												
Male	61.86	101.79	109.15	76.31	64.56	<.001	7.23	<.001	–25.04	<.001	–30.09	<.001
Female	46.60	81.27	91.30	62.76	74.40	<.001	12.34	<.001	–22.77	<.001	–31.26	<.001
Age groups (years)												
<1	242.39	531.46	443.73	362.81	49.68	<.001	–16.51	<.001	–31.73	<.001	–18.24	<.001
1 - 4	245.95	407.03	164.68	77.50	–68.49	<.001	–59.54	<.001	–80.96	<.001	–52.94	<.001
5 - 9	549.74	856.84	460.81	95.44	–82.64	<.001	–46.22	<.001	–88.86	<.001	–79.29	<.001
10 - 14	242.44	415.85	718.10	291.22	20.12	<.001	72.68	<.001	–29.97	<.001	–59.45	<.001
15 - 19[a]	114.91	188.53	421.90	474.57	312.99	<.001	123.78	<.001	151.72	<.001	12.48	<.001
20 - 24[a]	51.16	99.83	166.79	193.79	278.79	<.001	67.07	<.001	94.12	<.001	16.19	<.001
25 - 29[a]	32.58	77.82	137.91	139.44	327.99	<.001	77.22	<.001	79.18	<.001	1.11	<.001
≥30[a]	3.20	7.32	15.16	21.62	575.62	<.001	107.10	<.001	195.36	<.001	42.61	<.001
Total	50.76	90.53	104.11	70.13	78.35	<.001	15.0	<.001	–22.53	<.001	–32.64	<.001

[a]Chi-square test for trend: <.001.

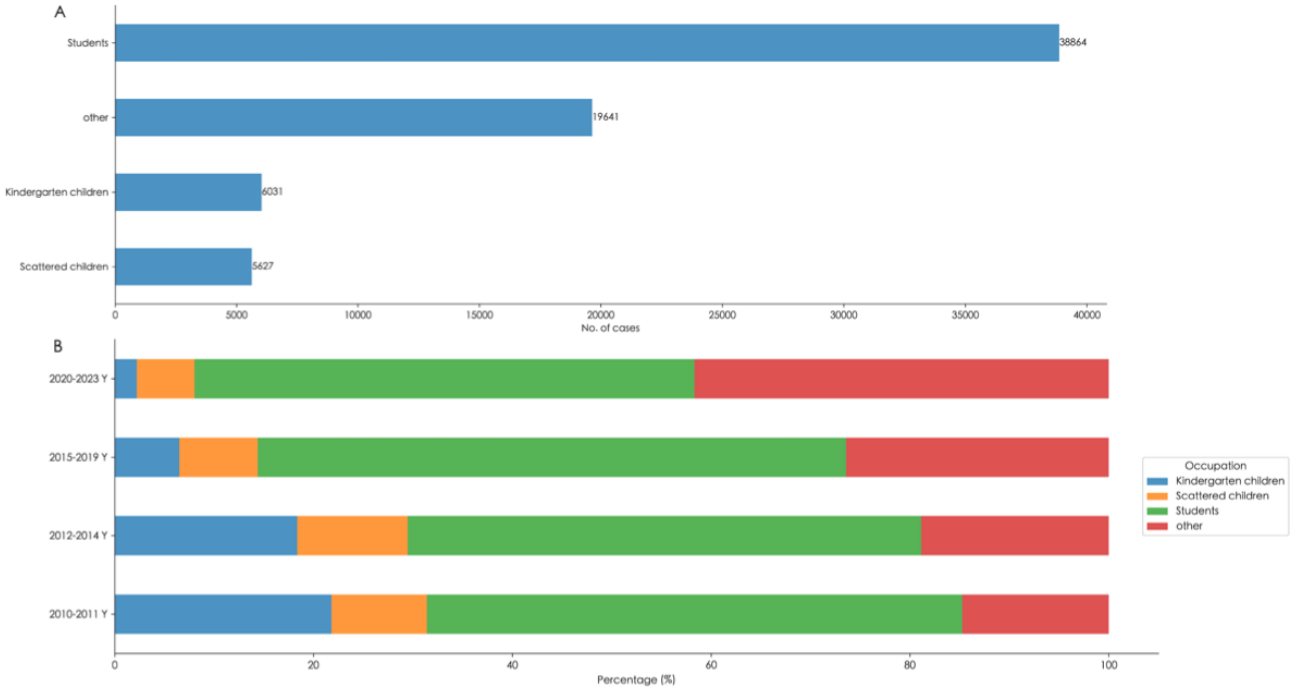
Among the total cases, male participants constituted 54.90% (38,519/70,163), while female participants accounted for 45.10% (31,644/70,163), resulting in a male-to-female ratio of 1.22:1. From 2010 - 2011 to 2020 - 2023, the incidence rate for male increased by 146.72%, whereas the rate for female rose by 169.37% (Table 3).

Occupational Distribution

Among the cases, there were 38,864 students (comprising 55.39%), 6031 children in childcare (representing 8.60%), 5627 children in disparate circumstances (accounting for 8.02%), and 19,641 instances classified as other nonchildren or students, which collectively constituted 28%. Figure 3 shows the distribution of cases by occupational classification. The proportion of children in childcare significantly declined (P<.001), from 21.80% in 2010 - 2011 to 2.21% in 2020 - 2023. Similarly, the percentage of children in disparate circumstances decreased from 9.61% in 2010 - 2011 to 5.82% in 2023 (P<.001). In contrast, the proportion of students remained relatively stable, with percentages of 53.84%, 51.67%, 59.19%, and 50.30% during the periods 2010 - 2011, 2012 - 2013, 2014 - 2019, and 2020 - 2023, respectively. Conversely, the proportion of other nonchildren or students increased significantly, from 14.75% in 2010 - 2011 to 41.67% in 2023 (P<.001).

XSL•FO
RenderX

Figure 3. The occupational distribution of varicella by the phases in Ningbo.

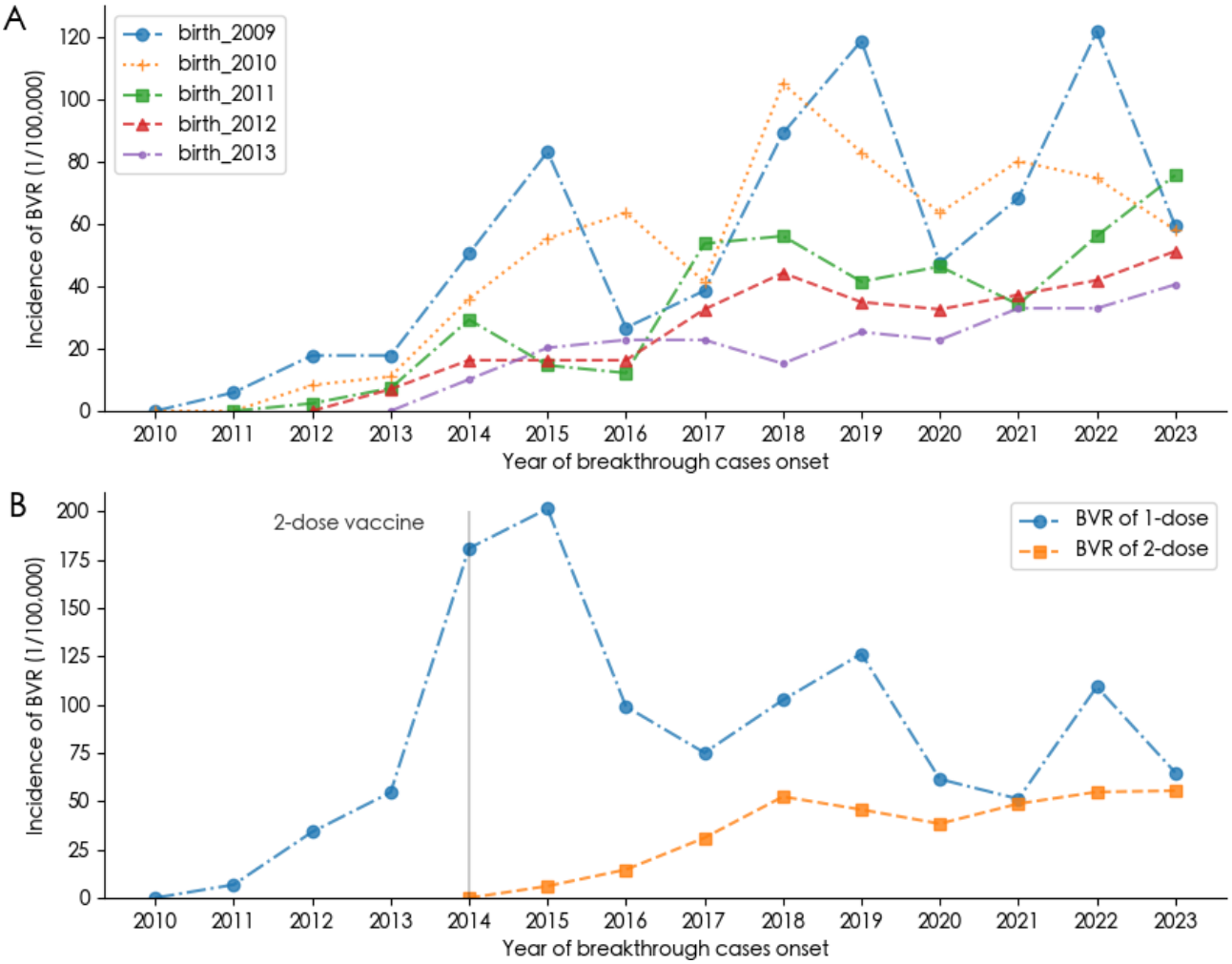


The Incidence and Number of Varicella Breakthrough Cases

Figure 4A presents the annual incidence of breakthrough varicella infection among different birth cohorts from 2009 to 2013, spanning the years 2010 to 2023. The mean annual breakthrough varicella incidence rate has shown a downward trend in cohorts with more recent birth years compared to those with earlier birth years. Specifically, the breakthrough varicella incidence rate was 53.16 per 100,000 individuals (range, 29.66 - 80.07 per 100,000) in 2009, 48.52 per 100,000 (range, 27.62 - 71.80 per 100,000) in 2010, 30.64 per 100,000 (range, 14.63 - 48.76 per 100,000) in 2011, 23.60 per 100,000 (range, 9.31 - 39.56 per 100,000) in 2012, and 17.53 per 100,000 (range, 5.06 - 32.91 per 100,000) in 2013.

XSL•FO
RenderX

Figure 4. The incidence and number of varicella breakthrough cases in Ningbo. (A) The annual incidence of breakthrough cases by birth cohorts. (B) The annual incidence of breakthrough cases by 1-dose and 2-dose varicella vaccination.



The average annual incidence of breakthrough cases following a single dose of the varicella vaccine was 83.40 per 100,000 (range, 51.21 - 119.50 per 100,000), significantly higher than the incidence after two doses, which was 24.80 per 100,000 (range, 17.67 - 32.90 per 100,000). Notably, after the implementation of the 2-dose immunization protocol, the incidence rate of breakthrough cases following the first dose showed a downward trend (Figure 4B).

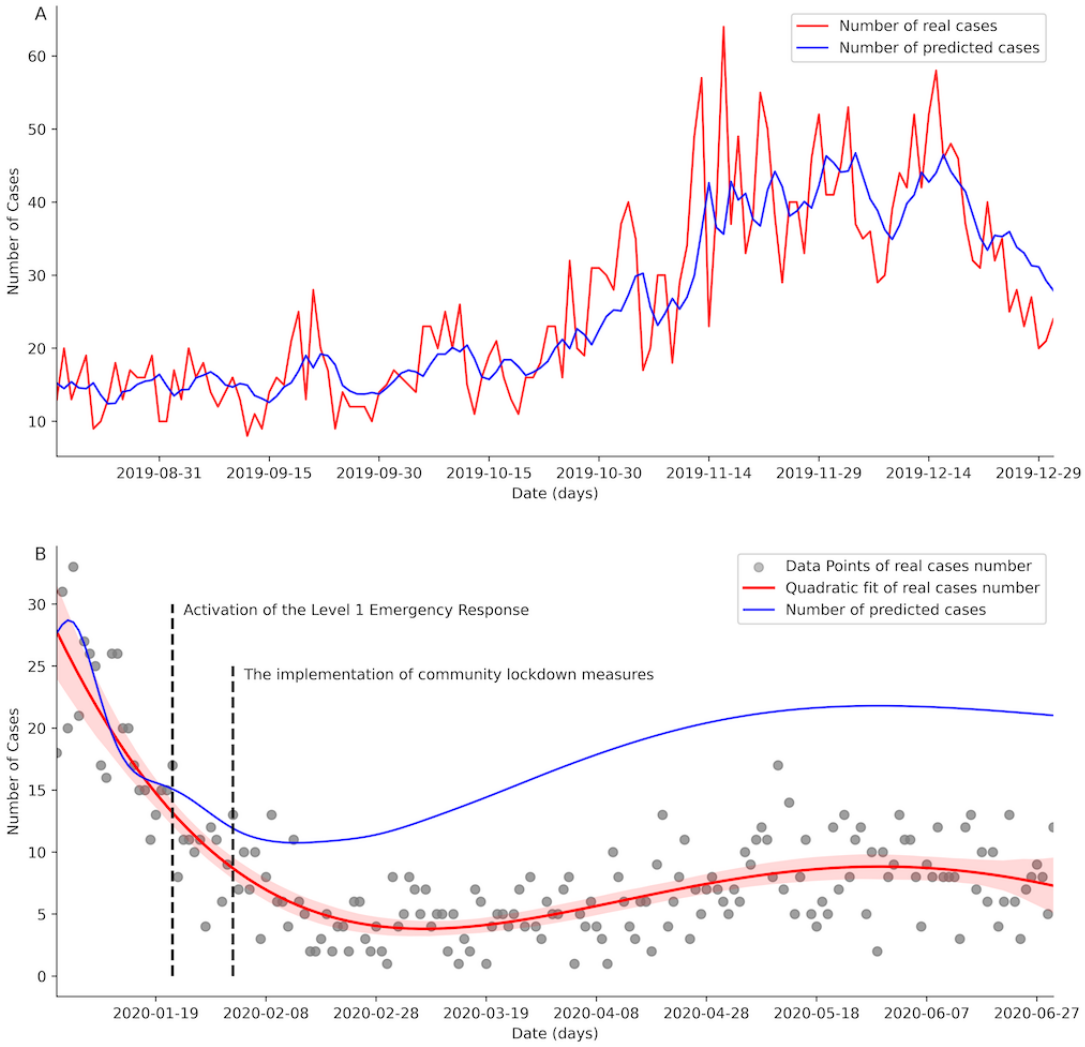
The median time from vaccination to the occurrence of breakthrough cases was 27 (IQR 17.50 - 48) months. For individuals who received a single dose, the median interval was 23 (IQR 17.25 - 31.50) months, whereas for those who received two doses, it was 46 (IQR 18 - 66) months.

Impact of NPIs

To further analyze the impact of NPIs on the incidence of varicella, we constructed a recurrent neural network model for deep learning. This model features a three-layer neural network architecture with a total of 469,633 trainable parameters. Figure 5A illustrates the comparison between the actual and predicted case numbers in the validation dataset, demonstrating the model's effectiveness and accuracy in identifying varicella incidence trends. The model achieved a mean squared error of 49.96, a root mean squared error of 7.07, and a mean absolute error of 5.36. Figure 5B presents a comparison of the predicted and actual varicella case numbers from January 1 to June 30, 2020. The results indicate that after the implementation of the Level 1 Emergency Response on January 22, the predicted case numbers gradually exceeded the actual case numbers. Following the implementation of community lockdown measures on February 2, the discrepancy between predicted and actual case numbers significantly increased.

XSL•FO
RenderX

Figure 5. Using recurrent neural network to analyze the impact of nonpharmacological interventions on varicella incidence. (A) The comparison between the actual and predicted case numbers in the validation dataset. (B) The comparison of the predicted and actual varicella case numbers from January 1 to June 30, 2020.



Discussion

Principal Findings

In China, varicella is not a legally regulated infectious disease, and the varicella vaccine, as the most effective control measure, is administered voluntarily and at personal expense. Varicella monitoring sites were established to conduct proactive monitoring in Ningbo City. Subsequently, health authorities required web-based direct reporting according to the requirements for reporting Class C infectious diseases. Compared to other studies based largely on voluntary case reporting [14-16], this method can more accurately reflect epidemiological characteristics. This study examined the epidemiological characteristics of varicella from 2010 to 2023, revealing significant changes. The administration of two doses of the varicella vaccine and NPIs had a substantial impact on the epidemiological characteristics.

Previous studies have consistently demonstrated a decline in varicella incidence following the introduction of the varicella vaccine [17-19]. However, our research presents a contrasting trend, revealing an overall increase in varicella incidence from 2010 to 2019. This observation, mirrored by the rising incidence reported in Anhui province from 2012 to 2019 [15], suggests that vaccine impact can be complex and influenced by local epidemiological factors. While our data show a 15% increase in varicella incidence following the implementation of a 2-dose immunization strategy, this overall trend masks important age-specific effects. Notably, we observed a significant decrease in varicella incidence among children younger than 9 years of age, with the most substantial decline (59.54%) in the 1 to 4

XSL•FO
RenderX

age group. This finding aligns with reports from Shi et al [17], who documented a similar reduction in incidence among children aged 0 - 4 years in Minhang District, Shanghai, and supports the efficacy of the vaccine in this young age group. However, the contrasting trend of increasing incidence among individuals aged 10 years and older warrants further investigation. This divergence from the findings of Marin et al [18], who reported declines across all groups, highlights the potential for waning immunity or incomplete vaccine coverage in older populations. Shi et al's [17] report of a 54.09% reduction in incidence among children aged 5 - 19 years further underscores the age-specific nature of vaccine effectiveness. Several factors may explain these discrepancies. The timing of 2-dose vaccination programs differs considerably between regions. The United States implemented a routine 2-dose program in 2007, with the second dose administered relatively early at ages 4 - 6 years, while this study area implemented the program in 2014. Shanghai's program, beginning in 2017 and integrated into the local immunization plan by August 2018 with over 80% coverage for the second dose [17], demonstrates the importance of both program timing and achieving high vaccination rates. Our continued observation of increasing incidence among those aged 10 - 19 years suggests that lower overall varicella vaccination rates in this cohort may be a key driver.

The optimal timing of the second dose also remains a point of discussion. While the United States and Shanghai recommend it at age 4, Ningbo City administers it at age 3. Sero-epidemiological studies offer conflicting insights: Sun et al [20] suggest a 3-year interval between doses maximizes antibody concentration, while Zhang et al [21] indicate higher seropositivity rates with earlier administration (ages 3 - 6). These findings emphasize the need for further research to refine immunization strategies and tailor them to local epidemiological contexts. Ultimately, understanding these nuanced interactions between vaccination programs, age-specific immunity, and population-level factors is crucial for optimizing varicella control and preventing future outbreaks.

During the COVID-19 pandemic, NPIs have demonstrated significant efficacy in controlling varicella outbreaks [22,23]. Originally intended to contain COVID-19, these interventions inadvertently reduced the incidence of other respiratory infectious diseases, such as varicella, due to similar transmission pathways [23]. Measures such as social distancing, mask-wearing, and enhanced hygiene practices led to a notable decline in varicella cases, underscoring their effectiveness in managing respiratory infections. In Xi'an, China, the incidence of varicella decreased by 43.18% in 2020 compared to 2019, attributed to the implementation of COVID-19 NPIs [14]. A European study also suggests that these infection control measures during the pandemic may have curtailed the spread of varicella [24]. Research by Wang et al [25] indicated that the actual number of varicella cases was 93.37% lower than expected, as predicted by the autoregressive integrated moving average model, highlighting a significant reduction in varicella activity due to COVID-19 NPIs. This study revealed that the incidence of chickenpox decreased by an average of 32.64% during the COVID-19 pandemic. Among individuals aged 15 years and older, the rate of increase in chickenpox cases continues to rise; however, this rate is significantly lower compared to the trends observed between 2014 and 2019. Additionally, predictions generated using the long short-term memory (LSTM) model demonstrated that prior to the implementation of COVID-19 prevention and control measures, the predicted number of chickenpox cases closely aligned with the actual reported cases. Following the adoption of these measures, the predicted case numbers consistently exceeded the actual reported cases. Notably, after the introduction of community closure management, the discrepancy between predicted and actual case numbers widened further. These findings suggest that NPIs implemented to curb the spread of COVID-19 also contributed to the prevention and control of chickenpox. Additionally, between 2020 and 2023, the low incidence of varicella from August to September was not significant. This finding aligns with the results of Sabale et al [24], indicating that infection control measures during the pandemic may have altered the seasonal transmission patterns of varicella and reduced its incidence rate.

This study reveals a notable shift in the age distribution of varicella incidence. While consistent with previous reports demonstrating a significant decline in cases among young children [19,26,27], we observed a gradual increase in varicella among individuals aged 10 - 14 years between 2010 and 2019, followed by a subsequent rise in cases among those aged 15 - 19 years between 2020 and 2023. This contrasts with findings from studies conducted in largely vaccinated populations [28], which have not reported a similar shift toward older age groups. The 95.5% decrease in incidence among children younger than 10 years of age reported by Moek and Siedler [19] highlights the effectiveness of childhood vaccination programs but also underscores the potential for a displacement of incidence to older, unvaccinated or incompletely vaccinated populations. Our finding of increasing incidence among adults aged 20 years and older further supports this hypothesis. The observed epidemiological pattern, with students representing the primary incidence group and a decline in cases among children in childcare, suggests that the implementation of the 2-dose vaccination strategy for children aged 3 years and older is successfully shifting the primary locus of transmission away from early childhood settings like kindergartens. This pattern is consistent with the concept of "herd immunity" waning in older, unvaccinated cohorts, leading to increased susceptibility and transmission within these groups. Further research is needed to investigate the long-term implications of this age-related shift and to inform targeted vaccination strategies for adolescents and adults.

Our initial observations, supported by extensive literature, indicate that the 2-dose varicella vaccine provides substantial epidemiological protection [8,10,29-31] and generally achieves a high vaccination rate. However, breakthrough cases of varicella continue to occur. Liu et al [32] reported that among 6961 varicella cases in Qingyang, China, 28.4% were breakthrough cases. Similarly, Zhu et al [33] documented an increase in the proportion of breakthrough cases during outbreaks, rising from 21.5% in 2008 to 86.1% in 2014, indicating an upward trend. Our findings also demonstrate an

XSL•FO
RenderX

increasing incidence of breakthrough cases over time following vaccination. Notably, the incidence of breakthrough cases with the 2-dose vaccine is lower than with a single dose. Furthermore, the implementation of a 2-dose immunization strategy significantly reduced the incidence of breakthrough cases associated with a single dose. The median interval between breakthrough cases was 23 months after a single dose and 46 months after two doses. This suggests that while breakthrough cases still occur after 2 doses, the 2-dose regimen is more effective than a single dose. It not only extends the duration of immune protection but also plays a crucial role in mass immunization efforts.

While the 2-dose varicella vaccine demonstrably provides substantial epidemiological protection and achieves high vaccination rates [8,10,29-31], recent data highlight the ongoing occurrence of breakthrough cases, a phenomenon increasingly recognized in the postvaccination era. Studies from China such as those by Liu et al [32] (28.4% breakthrough rate among 6961 cases) and Zhu et al [33] (increasing from 21.5% in 2008 to 86.1% in 2014) corroborate a growing trend of breakthrough infections, potentially linked to factors such as waning immunity or the emergence of vaccine-escape viral strains. Our findings align with this trend, demonstrating an increasing incidence of breakthrough cases over time following vaccination. Importantly, our data confirm the superior protection offered by the 2-dose regimen compared to a single dose, a finding consistent with established immunological principles regarding vaccine-induced immunity and memory cell development [34]. The significant reduction in breakthrough cases associated with the 2-dose strategy underscores its critical role in sustaining population immunity. The observed median interval of 23 months postsingle dose versus 46 months posttwo doses suggests that the second dose not only enhances initial protection but also extends the duration of immunity, a key consideration for long-term public health strategies. Further research is needed to investigate the factors contributing to breakthrough infections, including viral genetics, host immune responses, and the potential need for booster doses to maintain robust protection.

Limitations

This study has certain limitations. First, the identification of chickenpox cases in Ningbo primarily relies on clinical assessments. Cases with atypical presentations require evaluation in conjunction with epidemiological histories. However, when cases lack both characteristic clinical features and relevant epidemiological context, accurate identification becomes challenging. As vaccination coverage increases, breakthrough chickenpox cases often present with mild symptoms. Without comprehensive epidemiological investigations, such cases are likely to be overlooked. Furthermore, since 2012, chickenpox surveillance in Ningbo has predominantly relied on passive monitoring. This approach depends on health care providers voluntarily reporting diseases, symptoms, or health events, and it involves compiling data from existing medical records and self-reports from the public. Passive surveillance may lead to diagnostic inaccuracies, such as misclassifying nonchickenpox cases as chickenpox [35]. Additionally, it may result in underreporting, thereby underestimating the true incidence rate. The shift from active to passive monitoring introduces potential biases, particularly in terms of the completeness and representativeness of reports, which can obscure temporal trends and geographic distribution patterns in the data. Active monitoring, while more comprehensive, requires significant labor and time resources, making its consistent implementation across the city challenging. As a result, passive monitoring is widely adopted as a cost-effective alternative for tracking the epidemiological characteristics of infectious diseases over extended periods. To address the limitations of passive monitoring, implementing quality improvement measures is essential. For example, since 2012, Ningbo has integrated chickenpox case reporting into the management framework for Class C statutory infectious diseases. Disease control agencies at various levels are required to conduct regular audits to identify underreporting and provide feedback to improve surveillance accuracy. Based on monitoring outcomes, the system has not indicated significant underreporting of chickenpox cases. Third, LSTMs have demonstrated clear advantages in predicting long-term trends, but they remain unable to accurately model the overall spread of chickenpox during the COVID-19 pandemic. Consequently, the LSTM method is insufficient for evaluating the long-term impact of NPIs on chickenpox incidence when vaccine effects are excluded.

Conclusions

Through extensive longitudinal surveillance of varicella, we have identified significant changes in its epidemiological characteristics. These include the elimination of the seasonal incidence trough, a shift in the age of onset, and an increasing trend in adult morbidity. These findings highlight the critical need for ongoing epidemiological monitoring and underscore the importance of adapting vaccination strategies to address evolving epidemiological patterns, particularly in the context of postpandemic public health measures. Our investigation further revealed that the incidence of breakthrough cases was markedly lower among individuals who received the 2-dose varicella vaccination regimen compared to those who received only 1 dose. This emphasizes the importance of completing the full vaccination series to achieve higher levels of individual protection and herd immunity. Additionally, the application of machine learning algorithms has demonstrated significant potential in predicting chickenpox incidence. These algorithms enable more effective analysis of the impact of NPIs on disease incidence, providing evidence that NPIs have substantially reduced the burden of chickenpox. The growing use of machine learning in disease surveillance allows researchers to identify previously undetectable patterns and trends, improve predictive accuracy, and facilitate timely interventions. This, in turn, supports the development of more effective disease control measures. In summary, continuous epidemiological surveillance remains essential. We recommend enhanced monitoring of varicella in adults, increased vaccine coverage among individuals aged 15 years and older, and the implementation of 2-dose vaccination programs in high-risk settings such as high schools and universities. For individuals who have received only 1 dose, it is strongly advised to complete the second dose promptly to ensure optimal protection.

XSL•FO
RenderX

Acknowledgments

This research was supported by the Ningbo Science and Technology Program (2023S042), Zhejiang Province Medical and Health Science and Technology Program (2023KY1130), Ningbo Top Medical and Health Research Program (No.2023020713), and Science and Technology Program for Disease Prevention and Control in Zhejiang Province (No.2025JK277).

Authors' Contributions

XP and DZ designed the primary components of the study. XP, YZ, and XZ were responsible for data collection and manuscript preparation.

Conflicts of Interest

None declared.

References

1. Seward JF, Zhang JX, Maupin TJ, Mascola L, Jumaan AO. Contagiousness of varicella in vaccinated cases: a household contact study. JAMA 2004 Aug 11;292(6):704-708. [doi: 10.1001/jama.292.6.704] [Medline: 15304467]
2. Fairley CK, Miller E. Varicella-zoster virus epidemiology—a changing scene? J Infect Dis 1996 Nov;174 Suppl 3(Supplement 3):S314-S319. [doi: 10.1093/infdis/174.supplement_3.s314] [Medline: 8896538]
3. Varicella and herpes zoster vaccines: WHO position paper, June 2014. Wkly Epidemiol Rec 2014 Jun 20;89(25):265-287. [Medline: 24983077]
4. MM W, YH H, MY Y, et al. Epidemiological characteristics of public health emergency events of varicella in China, 2006-2021. Chin J Epidemiol 2023;12:1893-1898. [doi: 10.3760/cma.j.cn112338-20230411-00225]
5. Piazza MF, Amicizia D, Paganino C, et al. Has clinical and epidemiological varicella burden changed over time in children? overview on hospitalizations, comorbidities and costs from 2010 to 2017 in Italy. Vaccines (Basel) 2021 Dec 15;9(12):1485. [doi: 10.3390/vaccines9121485] [Medline: 34960231]
6. Wolfson LJ, Daniels VJ, Altland A, Black W, Huang W, Ou W. The impact of varicella vaccination on the incidence of varicella and herpes zoster in the United States: updated evidence from observational databases, 1991–2016. Clin Infect Dis 2020 Mar 3;70(6):995-1002. [doi: 10.1093/cid/ciz305] [Medline: 31147680]
7. Hu Y, Chen Y, Zhang B, Li Q. An evaluation of voluntary varicella vaccination coverage in Zhejiang Province, East China. Int J Environ Res Public Health 2016 Jun 3;13(6):560. [doi: 10.3390/ijerph13060560] [Medline: 27271649]
8. Shu M, Zhang D, Ma R, Yang T, Pan X. Long-term vaccine efficacy of a 2-dose varicella vaccine in China from 2011 to 2021: a retrospective observational study. Front Public Health 2022;10:1039537. [doi: 10.3389/fpubh.2022.1039537] [Medline: 36424959]
9. Shapiro ED, Marin M. The effectiveness of varicella vaccine: 25 years of postlicensure experience in the United States. J Infect Dis 2022 Oct 21;226(Suppl 4):S425-S430. [doi: 10.1093/infdis/jiac299] [Medline: 36265844]
10. Pan X, Shu M, Ma R, et al. Varicella breakthrough infection and effectiveness of 2-dose varicella vaccine in China. Vaccine (Auckl) 2018 Sep 5;36(37):5665-5670. [doi: 10.1016/j.vaccine.2018.05.025] [Medline: 30104113]
11. Dollard S, Chen MH, Lindstrom S, Marin M, Rota PA. Diagnostic and immunologic testing for varicella in the era of high-impact varicella vaccination: an evolving problem. J Infect Dis 2022 Oct 21;226(Suppl 4):S450-S455. [doi: 10.1093/infdis/jiac363] [Medline: 36265850]
12. Leung J, Broder KR, Marin M. Severe varicella in persons vaccinated with varicella vaccine (breakthrough varicella): a systematic literature review. Expert Rev Vaccines 2017 Apr;16(4):391-400. [doi: 10.1080/14760584.2017.1294069] [Medline: 28276305]
13. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020 Mar;17(3):261-272. [doi: 10.1038/s41592-019-0686-2] [Medline: 32015543]
14. Bai Y, Shen L, Sun M, et al. The short and long-term impact of nonpharmaceutical interventions on the prevalence of varicella in Xi'an during the COVID-19 pandemic. J Med Virol 2023 Aug;95(8):e29020. [doi: 10.1002/jmv.29020] [Medline: 37548166]
15. Xuan K, Zhang N, Li T, et al. Epidemiological characteristics of varicella in Anhui Province, China, 2012-2021: surveillance study. JMIR Public Health Surveill 2024 Apr 5;10:e50673. [doi: 10.2196/50673] [Medline: 38579276]
16. Wang M, Li X, You M, et al. Epidemiological characteristics of varicella outbreaks—China, 2006-2022. China CDC Wkly 2023 Dec 29;5(52):1161-1166. [doi: 10.46234/ccdcw2023.218] [Medline: 38164468]
17. Shi L, Lu J, Sun X, et al. Impact of varicella immunization and public health and social measures on varicella incidence: insights from surveillance data in Shanghai, 2013–2022. Vaccines (Basel) 2023 Nov 1;11(11):1674. [doi: 10.3390/vaccines11111674] [Medline: 38006006]

XSL•FO
RenderX

18. Marin M, Leung J, Anderson TC, Lopez AS. Monitoring varicella vaccine impact on varicella incidence in the United States: surveillance challenges and changing epidemiology, 1995-2019. J Infect Dis 2022 Oct 21;226(Suppl 4):S392-S399. [doi: 10.1093/infdis/jiac221] [Medline: 36265855]
19. Moek F, Siedler A. Trends in age-specific varicella incidences following the introduction of the general recommendation for varicella immunization in Germany, 2006-2022. BMC Public Health 2023 Nov 8;23(1):2191. [doi: 10.1186/s12889-023-17098-1] [Medline: 37936109]
20. Sun X, Wang Y, Zhang L, et al. Explore the optimal timing for administering the second dose of the varicella vaccine in China. J Med Virol 2023 Sep;95(9):e29119 [FREE Full text] [doi: 10.1002/jmv.29119] [Medline: 37737678]
21. Zhang Z, Zhang Y, Yu J, et al. Seroprevalence rates in children aged 3-6 years after implementing a two-dose varicella vaccination: a observational study. Hum Vaccin Immunother 2023 Dec 31;19(1):2211465. [doi: 10.1080/21645515.2023.2211465] [Medline: 37203320]
22. Bardsley M, Loveridge P, Bednarska NG, et al. The epidemiology of chickenpox in England, 2016-2022: an observational study using general practitioner consultations. Viruses 2023 Oct 27;15(11):2163. [doi: 10.3390/v15112163] [Medline: 38005841]
23. Xiao J, Dai J, Hu J, et al. Co-benefits of nonpharmaceutical intervention against COVID-19 on infectious diseases in China: a large population-based observational study. Lancet Reg Health West Pac 2021 Dec;17:100282. [doi: 10.1016/j.lanwpc.2021.100282] [Medline: 34611630]
24. Sabale U, Jarmale L, Murtagh J, Pawaskar M, Bencina G. Impact assessment of immunization and the COVID-19 pandemic on varicella across Europe using digital epidemiology methods: a descriptive study. PLOS ONE 2023;18(4):e0283465. [doi: 10.1371/journal.pone.0283465] [Medline: 37043422]
25. Wang M, Hu Y, Li X, Yin D, You M, Wang T. An ARIMA model to quantify the impact of COVID-19 nonpharmaceutical interventions on the epidemiology of varicella in china, 2005–2020. Research Square. Preprint posted Online on December 8, 2022. . [doi: 10.21203/rs.3.rs-2247295/v1]
26. Wen T, Fang Q, Fu C, et al. The impact of the implementation of the two-dose varicella vaccine immunization strategy in Quzhou: a retrospective birth cohort study. Hum Vaccin Immunother 2024 Dec 31;20(1):2408847. [doi: 10.1080/21645515.2024.2408847] [Medline: 39344170]
27. Chan DYW, Edmunds WJ, Chan HL, et al. The changing epidemiology of varicella and herpes zoster in Hong Kong before universal varicella vaccination in 2014. Epidemiol Infect 2018 Apr;146(6):723-734. [doi: 10.1017/S0950268818000444] [Medline: 29526171]
28. Baxter R, Tran TN, Ray P, et al. Impact of vaccination on the epidemiology of varicella: 1995-2009. Pediatrics 2014 Jul;134(1):24-30. [doi: 10.1542/peds.2013-4251] [Medline: 24913796]
29. Li Y, Xu F, Liu M, Teng S, Liang F, Wang F. Effectiveness of two-dose vs. one-dose varicella vaccine in children in Shanghai, China: a prospective cohort study. Front Public Health 2024;12:1320407. [doi: 10.3389/fpubh.2024.1320407] [Medline: 38894987]
30. Perella D, Wang C, Civen R, et al. Varicella vaccine effectiveness in preventing community transmission in the 2-dose era. Pediatrics 2016 Apr;137(4):e20152802. [doi: 10.1542/peds.2015-2802] [Medline: 26977081]
31. Lee YH, Choe YJ, Hong K, Yoon Y, Kim YK. The protective effectiveness of 2-dose varicella vaccination in children in Korea: a case-control study. Pediatr Infect Dis J 2023 Aug 1;42(8):719-722. [doi: 10.1097/INF.0000000000003958] [Medline: 37171976]
32. Liu B, Li X, Yuan L, et al. Analysis on the epidemiological characteristics of varicella and breakthrough case from 2014 to 2022 in Qingyang City. Hum Vaccin Immunother 2023 Aug 1;19(2):2224075. [doi: 10.1080/21645515.2023.2224075] [Medline: 37345683]
33. Zhu YF, Li YF, DU Y, Zeng M. Epidemiological characteristics of breakthrough varicella infection during varicella outbreaks in Shanghai, 2008-2014. Epidemiol Infect 2017 Jul;145(10):2129-2136. [doi: 10.1017/S0950268817000772] [Medline: 28446261]
34. Corrado M, Pearce EL. Targeting memory T cell metabolism to improve immunity. J Clin Invest 2022 Jan 4;132(1):e148546. [doi: 10.1172/JCI148546] [Medline: 34981777]
35. Viner K, Perella D, Lopez A, et al. Comparing active and passive varicella surveillance in Philadelphia, 2005-2010: recommendations for the transition to nationwide passive varicella disease surveillance. Public Health Rep 2014;129(1):47-54. [doi: 10.1177/003335491412900108] [Medline: 24381359]

Abbreviations

CDC: Center for Disease Prevention and Control
CISDCP: China Information System for Disease Control and Prevention
LSTM: long short-term memory
NPI: nonpharmaceutical intervention

XSL•FO
RenderX

Edited by A Mavragani; submitted 24.01.25; peer-reviewed by J Xiao, Z Li; revised version received 13.04.25; accepted 15.04.25; published 18.06.25.

Please cite as:
Pan X, Zhang Y, Zhao X, Zhang D
Changes in Epidemiological Characteristics of Varicella and Breakthrough Cases in Ningbo, China, From 2010 to 2023: Surveillance Study
JMIR Public Health Surveill 2025;11:e71691
URL: https://publichealth.jmir.org/2025/1/e71691
doi:10.2196/71691

© Xingqiang Pan, Yan Zhang, Xuefei Zhao, Dandan Zhang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Epidemic Trends and Spatial Distribution Characteristics of Hepatitis B in China: Surveillance Study

Xiaoxue Li[1,2], MPH; Jie Liang[3], BS; Lanqing Ma[1], MPH; Shijie Shen[1], MPH; Xueying Han[1], MPH; Tao Sun[3*], BS; Heng Guo[1,2*], MPH, DrPH

[1]Department of Public Health, Shihezi University School of Medicine, Xinjiang, No. 59, North 2nd Rd, Hong-Shan District, Shihezi, China

[2]Key Laboratory for Prevention and Control of Emerging Infectious Diseases and Public Health Security, The Xinjiang Production and Construction Corps, Shihezi, China

[3]Key Laboratory for Major Infectious Disease Monitoring in the Northern Frontier Region of the Corps, Shihezi, China

[*]these authors contributed equally

Corresponding Author:
Heng Guo, MPH, DrPH
Department of Public Health, Shihezi University School of Medicine, Xinjiang, No. 59, North 2nd Rd, Hong-Shan District, Shihezi, China

Abstract

Background: Hepatitis B is an important public health challenge facing China. Understanding the long-term epidemiological trends and evolving spatial distribution patterns is critical for optimizing prevention strategies and achieving the World Health Organization's 2030 hepatitis elimination targets.

Objective: This study aimed to explore the epidemic trends and spatial distribution characteristics of hepatitis B in China from 2004 to 2020.

Methods: This study used data on hepatitis B incidence from 2004 to 2020 from the China Public Health Science Data Center to analyze the time trend of hepatitis B incidence by joinpoint regression. The age-period-cohort model was used to analyze the age, period, and cohort effects of hepatitis B onset. Spatial autocorrelation analysis was used to explore the spatial distribution characteristics of hepatitis B in China.

Results: From 2004 to 2020, China reported a total of 17,449,842 cases of hepatitis B, with an average annual incidence rate of 76.30/100,000. The incidence of hepatitis B in China showed an increasing trend from 2004 to 2007, with an average annual percentage change (AAPC) of 9.49 (95% CI 2.12-17.39), and a decreasing trend from 2007 to 2014, with an AAPC of −3.77 (95% CI −5.93 to −1.55). The incidence of hepatitis B in China tended to be stable from 2014 to 2020, with an AAPC of −0.46 (95% CI −2.86 to 2.01). Age, period, and cohort effect significantly affect the incidence of hepatitis B. The age effect showed that the incidence rate of hepatitis B reached its peak at the age of 22 years, with an average incidence rate of 176.173/100,000. The period effect showed that the highest level during the study period occurred during 2004 - 2006. The cohort effect showed that the risk of hepatitis B increased first and then decreased, with the turning point of 1924 - 1974. The incidence of hepatitis B varies significantly among regions. The incidence in the northeast and northwest regions has decreased, that in the south and southwest regions has increased, and that in other regions has remained stable.

Conclusions: China has achieved remarkable results in the prevention and control of hepatitis B, but there are still differences in the incidence rate among different age groups and regions. These results suggest the need to further strengthen the identification and screening of high-risk populations and promote supplementary adult hepatitis B vaccination. Future intervention strategies should fully consider regional differences, implement precise intervention measures based on the epidemic trends and spatial distribution characteristics of each region, optimize resource allocation, and enhance the overall effectiveness of hepatitis B prevention and control.

(JMIR Public Health Surveill 2025;11:e70888) doi:10.2196/70888

KEYWORDS

hepatitis B; joinpoint regression analysis; APC model; spatial epidemiology; public health; vaccination

Introduction

Hepatitis B is a chronic infectious disease caused by hepatitis B virus (HBV) infection. Its onset is hidden and it is difficult to cure. Many infected people have no obvious symptoms in the early stage, and delayed treatment may lead to serious complications [1]. Hepatitis B is still an important public health problem in the world. According to the World Health

XSL•FO
RenderX

Organization, by 2022, a total of 254 million people worldwide were infected with chronic hepatitis B, with 1.2 million new infections and 1.1 million deaths; the main causes of death were cirrhosis and hepatocellular carcinoma (HCC) [2]. HCC is the sixth most common cancer and the third leading cause of cancer death worldwide [3]. HCC is also the fourth most common cancer and the third leading cause of cancer death in China [4]. In China, about 85% of patients with HCC have a history of HBV infection. HBV can induce HCC in various ways, and it is the most common and important factor in induced HCC [5]. HBV infection and its complications have brought a heavy economic burden to patients' families and society [6]. Therefore, understanding the epidemic characteristics of hepatitis B is helpful to provide a key basis for the optimization and precise implementation of hepatitis B prevention and control strategies.

Since China introduced the hepatitis B vaccine into the neonatal immunization program in 1992, the spread of hepatitis B has been controlled to some extent. The positive rate of hepatitis B surface antigen (HBsAg) in children under 5 years old has decreased from 9.9% in 1992 to 0.3% in 2014 [7]. However, due to China's large population base, HBV infection is still serious, and China's achievement of the global goal of eliminating hepatitis by 2030 is still hampered [8,9]. It is estimated that 90 million people are enduring chronic HBV infection in China, accounting for about 7% of the national population, and 330,000 people die of HBV-related cancer every year [10]. Although China has made some progress in hepatitis B control, it is still the country with the largest burden of HBV infection in the world. A series of health problems caused by HBV still pose a serious challenge to public health, and it is still necessary to continue to pay attention to high-risk regions and high-risk groups [11]. Therefore, gaining a comprehensive understanding of the epidemic trends and spatial distribution characteristics of hepatitis B incidence in China holds utmost significance for devising scientific, precise, and effective intervention strategies. It is pivotal for further optimizing China's hepatitis B prevention and control system, as well as driving deeper advancements in prevention and control efforts.

Current research on hepatitis B transmission, both domestically and internationally, predominantly concentrates on maternal-fetal prevention strategies and immunization programs. However, there remains a paucity of studies investigating the age-period-cohort effects on hepatitis B risk stratification. Furthermore, existing epidemiological investigations are largely confined to localized regions, specific demographic cohorts, or short-term observational periods, lacking systematic assessment of spatiotemporal dynamics and nationwide epidemiological patterns of hepatitis B in China. However, the spread of infectious diseases may vary greatly due to different spatial distributions, which makes the spatial distribution characteristics important factors to consider when formulating the response measures for infectious disease outbreaks [12]. For this reason, researchers use geographic information systems to examine the spatial distribution of various infectious diseases. Spatial autocorrelation analysis is a common method to analyze the spatial distribution characteristics of diseases and is widely used in the field of public health, especially the spatial cluster analysis of infectious diseases. This approach enables the identification of disease clusters, visualization of incidence and clustering trends over time and space, and clarification of geographical distribution patterns and risk levels across different regions [13].

In this study, the joinpoint regression model was used to analyze the long-term trend of the incidence of hepatitis B in Chinese residents from 2004 to 2020. The age-period-cohort model was used to estimate the effect of age, period, and birth cohort on the risk of hepatitis B. The spatial autocorrelation analysis was applied to visualize and spatially aggregate the incidence of hepatitis B in China from 2004 to 2020, and analyze the distribution characteristics of hepatitis B in different regions of China. This study aims to understand the spatiotemporal trends of hepatitis B in different regions of China and the impact of age, period, and birth cohort effect on the risk of hepatitis B, to provide a basis for clarifying the risk of hepatitis B in China and implementing accurate and effective prevention and control strategies. The results of this study will provide key evidence for the optimization of the hepatitis B prevention and control system in China, provide a scientific reference for the rational allocation of medical and health resources, and also contribute methodology and practical experience for the global realization of the World Health Organization's 2030 goal.

Methods

Data Source

Hepatitis B–related data in this study were sourced from the public database of the Public Health Science Data Center, which is hosted by the Chinese Center for Disease Control and Prevention [14]. The database is China's official authoritative public health data platform. Relevant data are collected based on China's notifiable infectious diseases reporting system, which requires medical institutions at all levels to confirm and report hepatitis B cases by the diagnostic criteria and management principles stipulated by the National Health Commission of the People's Republic of China. The dataset contains the hepatitis B data of 31 provinces, autonomous regions, and municipalities directly under the central government from 2004 to 2020 (excluding Hong Kong, Macao, and Taiwan). The data include the annual incidence and number of cases of hepatitis B in different regions and age groups in China, ensuring the comprehensiveness and representativeness of the study. The demographic data of China and its provinces in the same period (from 2004 to 2020) were obtained from the China Statistical Yearbook [15]. China's provincial vector maps are sourced from the National Basic Geographic Information System of China.

Joinpoint Regression Analysis

The Joinpoint regression software (version 5.0.2), developed by the National Cancer Institute, was used to analyze the temporal trend of hepatitis B incidence in China from 2004 to 2020. Joinpoint regression software uses the default grid search method to construct the model and uses the Monte Carlo permutation test method for model optimization. The Bayesian information criterion served as the criterion for evaluating the quality of model fit. Joinpoint regression analysis segments the study period into various intervals by identifying multiple connection points, fits and optimizes trends within each interval,

XSL•FO
RenderX

and thereby assesses the characteristics of hepatitis B incidence changes across different intervals. The number, location, and *P* value of the connection points were determined through a Monte Carlo permutation test (α=.05, 2-sided test). The average annual percentage changes (AAPCs) of the incidence rate of hepatitis B from 2004 to 2020, along with 95% CIs, were calculated. When AAPC>0 (*P*<.05), it indicates an increasing trend in incidence; when AAPC<0 (*P*<.05), it suggests a decreasing trend; and when *P*≥.05, it suggests a stable trend. The AAPC calculation formula is as follows:

(1)n(y|x)=b0+b1x

(2)AAPC={exp(wibiwi)−1}×100%

x represents the year, y denotes the incidence rate, b_0 signifies the intercept, b_i represents the regression coefficient, which is the slope coefficient of each line segment, and w_i indicates the number of years encompassed in each time segment.

Based on China's geographical characteristics and administrative divisions, this study divided China into 7 regions: the northern region (including Beijing, Tianjin, Hebei, Shanxi, and Inner Mongolia), the northeast region (including Heilongjiang, Jilin, and Liaoning), the northwest region (including Shaanxi, Gansu, Qinghai, Ningxia, and the Xinjiang), the eastern region (including Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Anhui, and Jiangxi), the central region (including Henan, Hubei, and Hunan), the southern region (including Guangdong, Hainan, and Guangxi), and the southwest region (including Chongqing, Sichuan, Guizhou, Yunnan, and the Xizang). It also analyzed the trend of hepatitis B incidence rate across these 7 regions.

Age-Period-Cohort Model

The age-period-cohort-model (APC model) is widely used to analyze incidence trends in epidemiology. Among them, the age effect reflects the difference in the risk of disease in different age groups. The period effect reflects the overall change in the incidence risk of all age groups at a specific time point. The cohort effect reflects the change in the risk of morbidity in people with the same birth year. This study used the online network analysis tool developed by Rosenberg, Check, and others to construct the APC model [16].

Due to the scarcity of cases among individuals aged 84 years and older, as well as the challenges in determining birth cohorts, this study adopted a 5-year age interval, dividing the age range of 0 - 84 years into 17 age groups. The data from 2004 to 2020 were divided into 5 groups: 2004 to 2008, 2009 to 2013, 2014 to 2018, and 2019 to 2020, to construct the APC model.

The model uses the Poisson log-linear model and uses the Akaike Information Criterion for a comprehensive evaluation of the model's goodness of fit. The formula for this model is expressed as:

(3)log(Eij)=log(Pij)+μ+αi+βi+γk

In this formula, E_{ij} and P_{ij} represent the total number of individuals who developed the disease in the ith age group during the jth period group, respectively. μ denotes the intercept term of the incidence rate. α_i signifies the regression coefficient for the ith age group, while β_i indicates the regression coefficient for the jth period group. γ_k stands for the regression coefficient of the kth birth cohort group.

Spatial Autocorrelation Analysis

In this study, ArcGIS 10.8 (Esri) software was used to describe the spatial distribution of hepatitis B in China from 2004 to 2020 through global spatial autocorrelation analysis. In addition, Moran's *I*, a global spatial autocorrelation coefficient, was used to assess the spatial distribution pattern of hepatitis B across various regions in China. Moran's *I* values range from −1 to 1. A positive value indicates spatial clustering, with a larger value indicating stronger spatial clustering. Negative values indicate spatial dispersion, with a smaller value indicating stronger spatial dispersion. A value of 0 indicates no spatial autocorrelation. The formula for calculating the Moran's *I* is as follows:

(4)I=n∑i=1n∑j=1nwij(xi−x^)(xj−x^)(∑i=1n∑j=1nwij)∑i=1n(xi−x^)2

In this formula, the incidence rate is represented by λ, the average incidence rate is represented by X, and the matrix of spatial weights is denoted as W_{ij}, representing the proximity relationship between region i and region j.

Local Moran's *I* was used to conduct a local analysis of the distribution of hepatitis B cases in each region, and the local autocorrelation type was determined by visualizing the local autocorrelation map. The aggregation types can be categorized into the following four types: "high-high," "high-low," "low-high," and "low-low."

Ethical Considerations

The data of this study were obtained from the public database authorized by the public health science data center of the Chinese Center for Disease Control and Prevention. According to the relevant provisions of the international code of Biomedical Research Ethics, this study meets the exemption conditions for ethical review, mainly for the following 3 reasons: (1) the research data are completely anonymized public information and do not contain any personally identifiable information; (2) this study did not involve any form of intervention, nor did it collect or use human biological samples; and (3) the processing and publication of the research results will not have any impact on the clinical diagnosis and treatment plan of the original data provider.

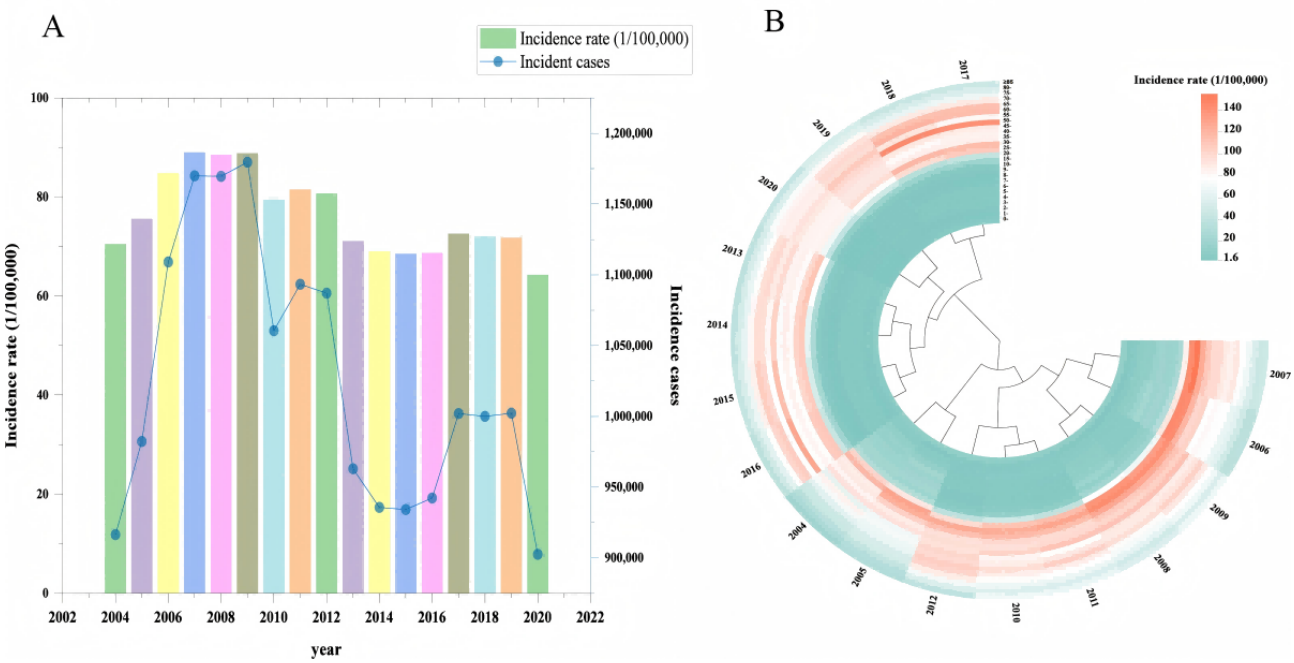
Results

Overall Description of Hepatitis B in China From 2004 to 2020

From 2004 to 2020, China reported a total of 17,449,842 cases of hepatitis B, with an average annual incidence rate of 76.3/100,000. The incidence rate fluctuated but exhibited an overall decreasing trend, decreasing from 70.5/100,000 in 2004 to 64.29/100,000 in 2020 (Figure 1A). From 2004 to 2020, the majority of hepatitis B cases occurred in the adult population, with 70.77% of cases falling within the age range of 20 - 50 years (Figure 1B).

XSL•FO
RenderX

Figure 1. Incidence of hepatitis B in China from 2004 to 2020. (A) Several cases and incidence rate of hepatitis B in China from 2004 to 2020. (B) Heat map of age-specific incidence rates of hepatitis B in China from 2004 to 2020.

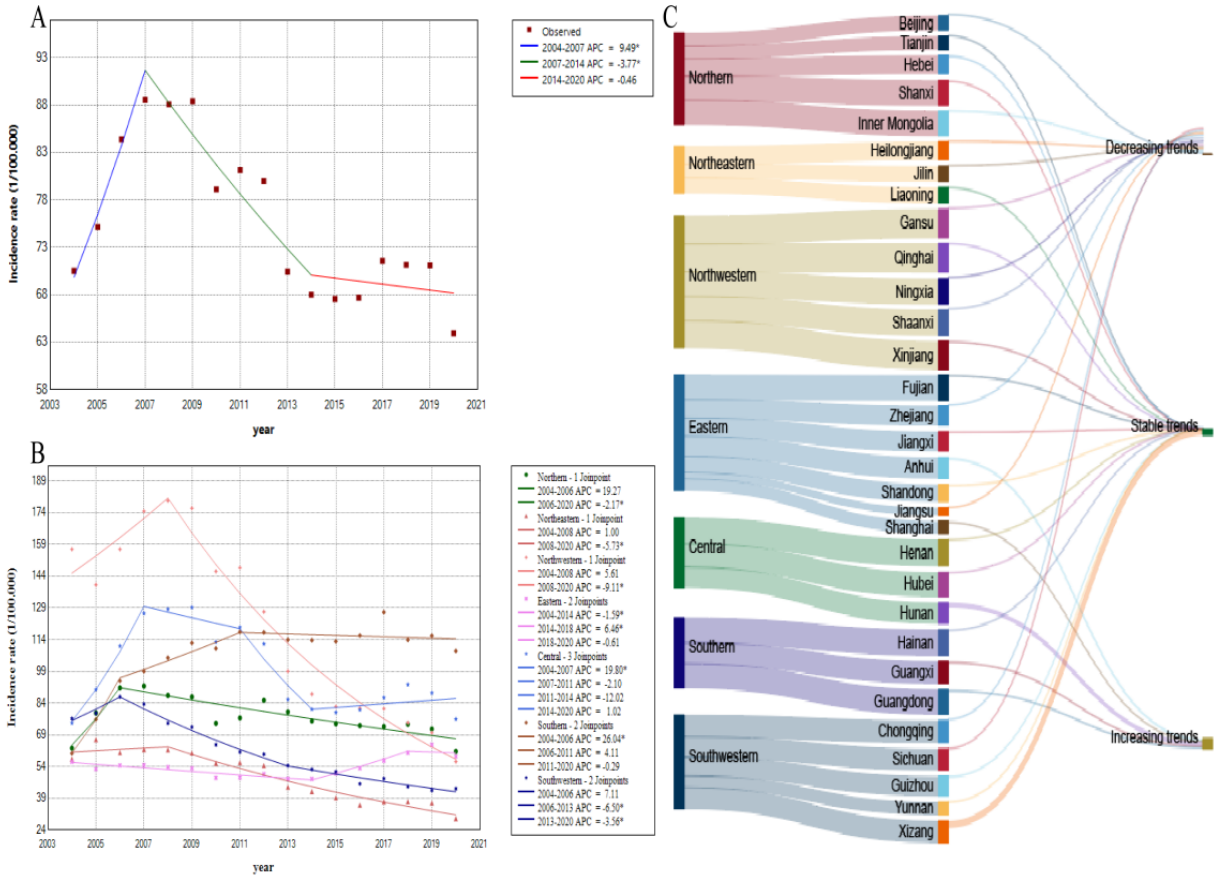


Temporal Trends of Hepatitis B Incidence Rate in China From 2004 to 2020

The incidence rate of hepatitis B in China increased from 70.5/100,000 in 2004 to 89/100,000 in 2007 (AAPC=9.49, 95% CI 2.12 to 17.39). Subsequently, it decreased from 2007 to 69.05/100,000 in 2014 (AAPC=–3.77, 95% CI −5.93 to −1.55). From 2014 to 2020, the incidence rate of hepatitis B in China exhibited a stabilizing trend (AAPC=−0.46, 95% CI −2.86 to 2.01) (Figure 2A).

XSL•FO
RenderX

Figure 2. Temporal trends of hepatitis B incidence rate in China from 2004 to 2020: (A) temporal trends in the incidence rate of hepatitis B in China, (B) temporal trends in the incidence rate of hepatitis B across 7 regions in China, and (C) temporal trends in the incidence rate of hepatitis B across 31 regions in China. APC: annual percentage change.



From 2004 to 2020, the incidence rate of hepatitis B decreased in Northeast China (AAPC=−4.09, 95% CI –5.60 to −2.56) and Northwest China (AAPC=−5.64, 95% CI −7.48 to −3.75). Conversely, the incidence rate of hepatitis B increased in Southern China (AAPC=4.07, 95% CI 1.41-6.80) and Southwest China (AAPC=3.60, 95% CI −4.86 to −2.33). However, the incidence of hepatitis B remained stable in the northern, eastern, and central regions (both *P*>.05; Figure 2B).

From 2004 to 2020, among all provinces, Xizang (AAPC=16.93, 95%CI 14.31-19.60) and Hunan (AAPC=10.52, 95% CI 7.91-13.20) experienced the most significant increases in hepatitis B incidence rates. Shaanxi (AAPC=−9.23, 95% CI −10.96 to −7.48), Chongqing (AAPC=−9.23, 95% CI −11.15 to −7.28), Beijing (AAPC=−9.05, 95% CI −15.36 to −2.26), Ningxia (AAPC=−9.84, 95% CI −11.37 to −8.28), Gansu (AAP=−8.31, 95% CI −12.63 to −3.77), and Zhejiang (AAPC=−8.43, 95% CI −11.25 to −5.52) saw a rapid decrease in hepatitis B incidence rates. The incidence rates of hepatitis B in other regions exhibited a stable trend (Figure 2C and Table 1).

XSL•FO
RenderX

Table . Joinpoint regression analysis of the incidence rate of hepatitis B in China from 2004 to 2020.

Region	AAPC[a] (95% CI)	*t* test	*P* value
Northern region	0.29 (−1.85 to 2.47)	0.26	.794
Beijing	−9.05 (−15.36 to −2.26)	2.58	.009
Tianjin	−5.08 (−11.02 to 1.27)	−1.58	.114
Hebei	0.98 (−1.38 to 3.39)	0.81	.419
Shanxi	3.21 (−0.41 to 6.97)	1.73	.082
Inner Mongolia	−2.26 (−3.92 to −0.56)	−2.60	.009
Northeast region	−4.09 (−5.60 to −2.56)	−5.17	<.001
Heilongjiang	−5.82 (−8.09 to −3.48)	−4.80	<.001
Jilin	−5.21 (−6.29 to −4.12)	−9.17	<.001
Liaoning	−1.06 (−3.40 to 1.33)	−0.87	.382
Northwest region	−5.64 (−7.48 to −3.75)	−5.75	<.001
Shaanxi	−9.23 (−10.96 to −7.48)	−9.91	<.001
Gansu	−8.31 (−12.63 to −3.77)	−3.52	<.001
Qinghai	−0.45 (−2.73 to 1.88)	−0.38	.701
Ningxia	−9.84 (−11.37 to −8.28)	−12.89	<.001
Xinjiang	0.36 (−3.43 to 4.30)	0.18	.854
Eastern region	0.49 (−0.82 to 1.81)	0.73	.466
Shanghai	3.81 (1.12 to 6.56)	3.04	.008
Jiangsu	−3.61 (−5.11 to −2.06)	4.57	<.001
Zhejiang	−8.43 (−11.25 to −5.52)	5.51	<.001
Fujian	0.18 (−1.0 to 1.41)	0.29	.771
Shandong	3.19 (−0.45 to 6.96)	1.71	.087
Anhui	5.46 (4.35 to 6.59)	9.81	<.001
Jiangxi	0.93 (−0.97 to 2.87)	0.96	.339
Central region	0.84 (−3.79 to 5.69)	0.35	.727
Henan	−3.00 (−8.32 to 2.64)	−1.06	.291
Hubei	-0.38 (−2.69 to 1.98)	−0.32	.750
Hunan	10.52 (7.91 to 13.20)	8.19	<.001
Southern region	4.07 (1.41 to 6.80)	3.02	.002
Guangdong	4.40 (2.01 to 6.83)	3.66	<.001
Hainan	1.38 (−2.62 to 5.55)	0.67	.505
Guangxi	3.22 (0.83 to 5.67)	2.65	.008
Southwest region	3.60 (−4.86 to −2.33)	−5.48	<.001
Chongqing	−9.23 (−11.15 to −7.28)	−8.9	<.001
Sichuan	−4.47 (−6.47 to −2.43)	−4.24	<.001
Guizhou	0.30 (−3.41 to 4.14)	0.15	.877
Yunnan	−2.52 (−5.40 to 0.45)	−1.67	.094
Xizang	16.93 (14.31 to 19.60)	14.73	<.001
Whole country	−0.15 (−1.77 to 1.49)	−0.18	.855

[a]AAPC: average annual percentage change.

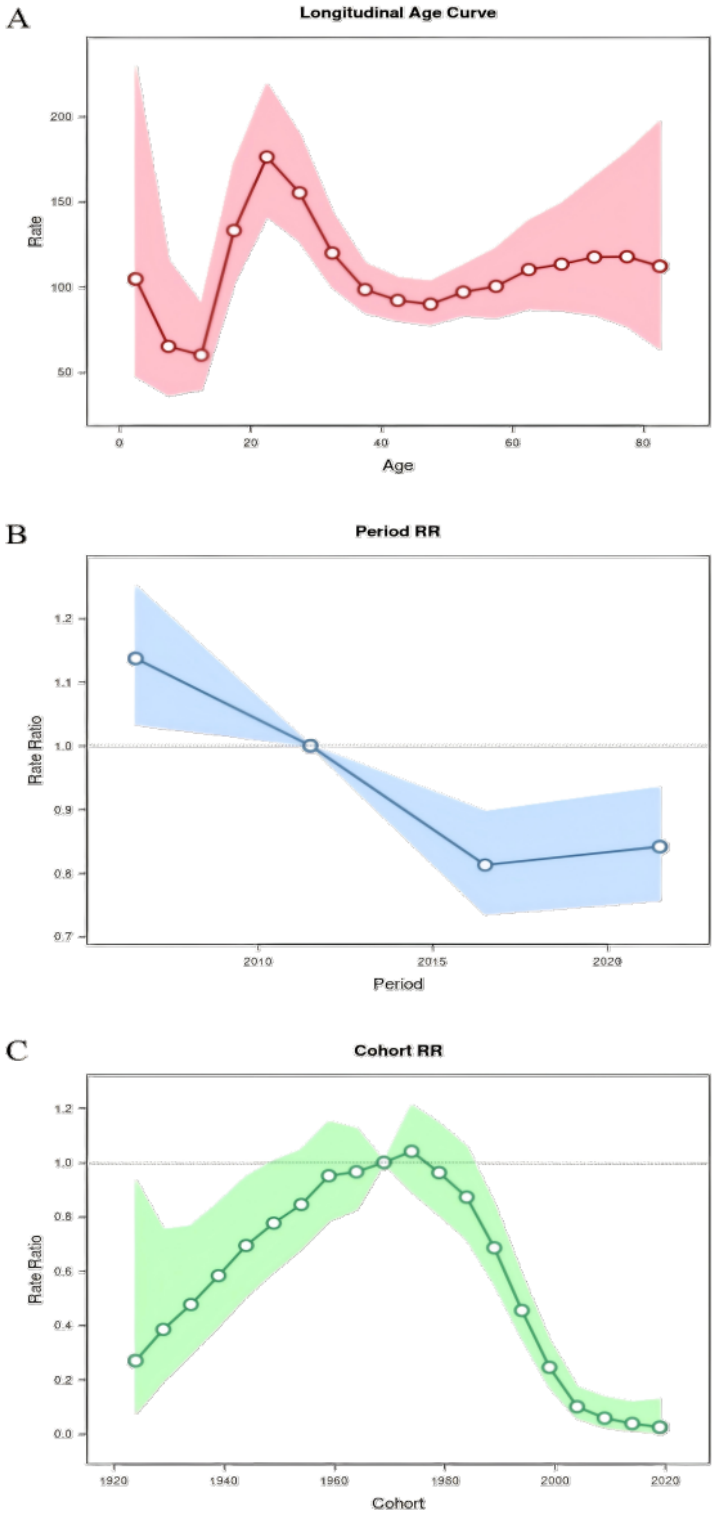
XSL•FO
RenderX

Age-Period-Cohort Model

Age Effect

APC model analysis indicates that the majority of hepatitis B cases occur in adults, with the incidence rate exhibiting fluctuations across different age groups. The incidence rate increased from childhood to adolescence. At the age of 22 years, the incidence rate of hepatitis B reached its peak, with an average incidence rate of 176.173/100,000. Subsequently, it began to decline, remained relatively stable in middle age, and slightly increased in old age (Figure 3A).

Figure 3. Average annual percentage change model of hepatitis B incidence rate in China from 2004 to 2020: (A) longitudinal age curve; (B) period rate ratio; (C) cohort rate ratio. RR: relative risk.



XSL•FO
RenderX

Period Effect

After adjusting for age and cohort effects, the relative risk (RR) value for the period effect of hepatitis B incidence rate from 2004 to 2020 exhibited a trend of initially decreasing and then increasing. Using 2011 as the reference period (RR=1), the peak incidence rate of hepatitis B was observed between 2004 and 2006, with an RR value of 1.14 (95% CI 1.03-1.25). Subsequently, the RR value for the period effect on the incidence rate of hepatitis B exhibited a decreasing trend, with the RR value for the incidence rate of hepatitis B during the period from 2011 to 2016 being 0.81 (95% CI 0.74-0.9). The RR value of hepatitis B incidence from 2016 to 2020 was 0.84 (95% CI 0.76-0.93), indicating a slight increase compared with the previous period (Figure 3B).

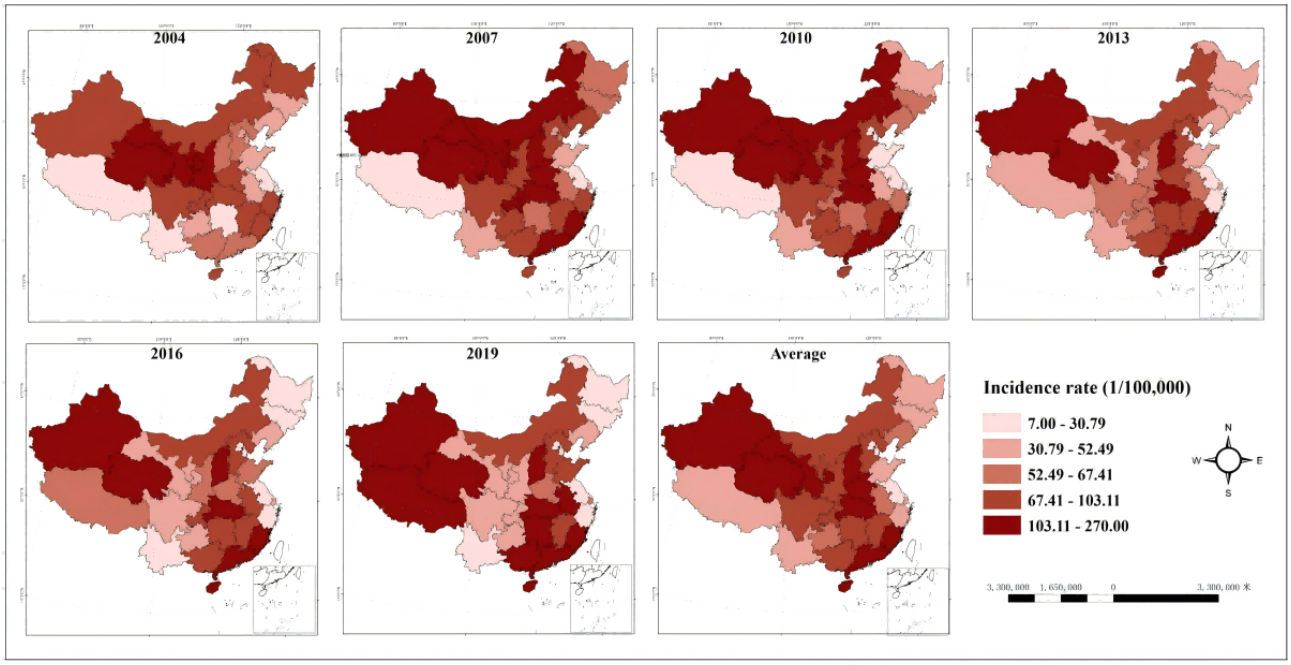
Cohort Effect

After adjusting for age and period effects, the cohort born in 1969 was set as the reference cohort (RR=1). From 1924 to 1974, the RR value of the cohort effect for hepatitis B in China exhibited an increasing trend, reaching its peak in 1974 with an RR value of 1.04 (95% CI 0.9-1.21). The RR value of the cohort effect for hepatitis B in China decreased after 1974, reaching its nadir in the 2019 cohort, with an RR value of 0.03 (95% CI 0.01-0.13; Figure 3C).

Spatial Distribution Characteristics of Hepatitis B in China From 2004 to 2020

Between 2004 and 2020, China witnessed significant regional disparities in the incidence rate of hepatitis B. The high-incidence regions of hepatitis B are primarily located in the northwest and central regions, while the incidence in the southwest and some parts of the eastern regions is gradually increasing over time. Among them, Qinghai (216.86/100,000), Xinjiang (170.21/100,000), and Gansu (130.23/100,000) exhibited the highest average incidence rate, whereas Jiangsu (18.56/100,000), Beijing (20.68/100,000), and Tianjin (23.54/100,000) had lower average incidence rate (Figure 4).

Figure 4. Spatial distribution characteristics of hepatitis B in China from 2004 to 2020.



Spatial Autocorrelation Analysis of Hepatitis B in China From 2004 to 2020

The global spatial autocorrelation analysis of hepatitis B incidence in 31 regions of China spanning from 2004 to 2020 revealed that the overall Moran's *I* values for the periods of 2004 to 2011 and 2020 were greater than 0, ranging from 0.1089 to 0.2084 (all $P<.05$), indicating a positive spatial autocorrelation. No significant spatial autocorrelation was observed from 2012 to 2019 (all $P>.05$; Table 2).

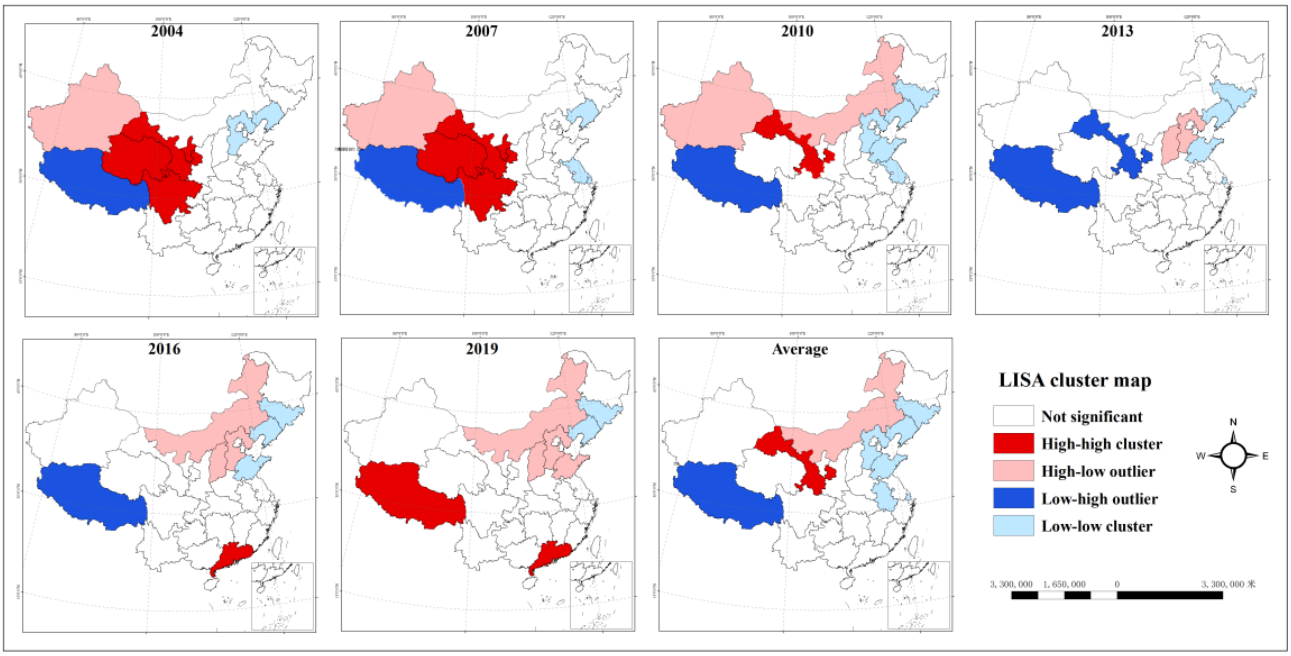
XSL•FO
RenderX

Table . Global spatial autocorrelation analysis of hepatitis B in China from 2004 to 2020.

Year	Moran's *I*	Expected index	Variance	Z values	*P* value
2004	0.2084	−0.0333	0.0057	3.2149	.001
2005	0.2024	−0.0333	0.0056	3.1603	.002
2006	0.1499	−0.0333	0.0056	2.4583	.014
2007	0.1304	−0.0333	0.0055	2.2073	.027
2008	0.1319	−0.0333	0.0053	2.2678	.023
2009	0.1089	−0.0333	0.0047	2.0844	.037
2010	0.1158	−0.0333	0.0054	2.0309	.042
2011	0.1227	−0.0333	0.0056	2.0895	.037
2012	0.0816	−0.0333	0.0055	1.5496	.121
2013	0.0581	−0.0333	0.0054	1.2471	.212
2014	0.1019	−0.0333	0.0057	1.7838	.074
2015	0.1109	−0.0333	0.0058	1.8991	.058
2016	0.0759	−0.0333	0.0057	1.4872	.149
2017	0.1082	−0.0333	0.0057	1.8735	.061
2018	0.1009	−0.0333	0.0058	1.7582	.079
2019	0.0999	−0.0333	0.0058	1.7432	.081
2020	0.1308	−0.0333	0.0058	2.1482	.032

The results of local spatial autocorrelation analysis indicate that low-low cluster regions are primarily located in Jilin and Liaoning in Northeast China, as well as Shandong in the eastern region. The high-low cluster regions have shifted from Xinjiang in the northwest to northern regions such as Inner Mongolia and Hebei. From 2004 to 2007, high-high cluster regions were primarily located in the northwest and southwest regions, specifically in Qinghai, Gansu, Ningxia, Sichuan, and Chongqing. After 2007, the high-high cluster regions phenomenon gradually dissipated in the northwestern region of China, except Ningxia. In 2013, Ningxia transformed from a high-high cluster region to a high-low cluster region, ultimately evolving into a region with no significant spatial clustering. Since 2015, Guangdong Province has transformed from a region with no significant spatial clustering to a high-high cluster region, whereas Xizang has shifted from a high-low cluster region to a high-high cluster region (Figure 5).

Figure 5. Local spatial autocorrelation analysis of hepatitis B in China from 2004 to 2020.



XSL•FO
RenderX

Discussion

Principal Findings

Using joinpoint regression analysis, the APC model, and spatial autocorrelation analysis, this study delves into the epidemic characteristics and spatial distribution patterns of hepatitis B in China from 2004 to 2020. The findings offer valuable data support and theoretical insights for comprehending the evolving epidemic trends of hepatitis B among Chinese residents and formulating targeted prevention and control strategies accordingly.

The study findings revealed that the overall incidence rate of hepatitis B in China exhibited a decreasing trend between 2004 and 2020, with a decrease from 70.5/100,000 in 2004 to 64.29/100,000 in 2020. This indicated that the prevention and control measures for hepatitis B in China from 2004 to 2020 had yielded certain positive outcomes, which align with the findings of other studies [17]. According to the results of joinpoint regression analysis, the incidence rate of hepatitis B in China exhibited an increasing trend between 2004 and 2007, subsequently experienced a decreasing trend, and ultimately stabilized after 2014. Since 2007, the incidence rate of hepatitis B decreased, primarily attributed to the neonatal hepatitis B vaccination program that was launched in China in 1992 [18,19]. Furthermore, China has been offering free hepatitis B vaccinations to newborns since 2002, significantly reducing the risk of hepatitis B infection among newborns. A study revealed that the hepatitis B vaccination coverage among individuals aged 1 to 59 in China increased significantly, escalating from 54.1% to 78.9% between 2006 and 2024 [20]. According to a study on the global burden of disease, the positive rate of HBsAg in China has undergone a significant decrease, decreasing from 11.9% in 1990 to 8.4% in 2015, and further decreasing to 7.8% in 2019 [7]. These data provide further evidence of China's significant advancements in the prevention and control of hepatitis B.

The APC model-based analysis of age effects revealed that the risk of hepatitis B is relatively low in childhood, rises during adolescence, peaks at age 22, and exhibits a decreasing trend in older adulthood. The increased incidence rate among young adults may be attributed to high-risk sexual behaviors (such as unprotected sex and multiple sexual partners), unhealthy lifestyles (including excessive stress, staying up late, alcohol abuse, and smoking), and specific occupational exposures [21,22]. The increased risk of hepatitis B in the elderly is significantly influenced by the low coverage of hepatitis B vaccination in this population. A study exploring the factors influencing hepatitis B vaccination coverage highlighted a negative correlation between vaccination coverage and age, which is likely linked to the national hepatitis B immunization program for newborns and infants instituted by the Chinese government in 1992 [23]. In addition, the decrease in immune function among the elderly further heightens their susceptibility to hepatitis B infection [24]. In terms of period effects, the highest hepatitis B incidence rate in China during 2004 - 2020 was observed between 2004 and 2006. Cohort analysis indicated that individuals born between 1924 and 1974 exhibited an increased trend in hepatitis B incidence risk, whereas a decreased trend was observed among those born after 1974. These shifts in period and cohort effects can be attributed to factors such as the widespread adoption of hepatitis B vaccines, advancements in medical technology, increased public awareness of hepatitis B prevention, and changes in lifestyle habits [25]. Therefore, an in-depth discussion of the age, period, and birth cohort effect of hepatitis B can not only help to locate the high-risk population of hepatitis B, but also provide important references for the formulation of public health policies in the future, to more effectively control the spread and risk of hepatitis B.

The overall incidence rate of hepatitis B in China exhibited a decreasing trend from 2004 to 2020, albeit with regional variations in its trajectory. Specifically, the northeastern and northwestern regions observed a decrease in hepatitis B incidence. Notably, the northwestern region experienced a steeper decrease in hepatitis B incidence rate compared with the northeastern region. However, the northwestern region consistently maintained a relatively high incidence rate, particularly in Qinghai, Xinjiang, and Gansu, where the average incidence rate from 2004 to 2020 ranked among the highest nationwide. This high incidence rate can be attributed to various factors such as geographical conditions, sluggish economic development, scarcity of medical resources, and varying levels of health literacy among residents [12]. Contrastingly, the southern and southwestern regions witnessed an increased trend in hepatitis B incidence. This increasing trend in the southern region is likely associated with factors like frequent population movements, uneven distribution of health care resources, and geographical considerations. Given the dense and mobile population in the southern region, disparities in health literacy, access to health care resources, and vaccination coverage among the mobile population may contribute to the widespread transmission of the HBV in this region. In addition, the humid and hot climate prevalent in the southern region may facilitate the easier spread of hepatitis B [26]. In addition, Liang et al [27] showed that some alleles are closely related to HBV infection. There are obvious differences in genotype distribution between southern and Northern Chinese populations. The difference in genetic background may be a potential factor for Southern Chinese populations to be more susceptible to hepatitis B. Simultaneously, the increasing incidence rate in Southern China in recent years can also be partially attributed to the enhanced quality of hepatitis B case reporting, which has significantly reduced previously overlooked cases [12,28]. The swift increase trend observed in Southwest China is primarily attributed to the rapid incidence rate in Xizang. Located on the southwestern border of China, the Xizang, owing to its vast territory, sparse population distribution, and inconvenient transportation, lags behind other provinces in immunization planning. Consequently, Xizang has consistently been identified as a high-prevalence area for hepatitis B [29].

Through global spatial autocorrelation analysis, this study revealed the spatial distribution characteristics and evolving trends of hepatitis B incidence in 31 regions of China from 2004 to 2020. The findings of the global spatial autocorrelation analysis indicated that there was a significant clustering of

XSL•FO
RenderX

hepatitis B incidence rate in China during the periods of 2004 - 2011 and 2020. These cluster regions could potentially be attributed to socioeconomic disparities, the distribution of medical resources, varying health conditions, and the implementation of preventive measures across different regions [30]. Based on the data from 2012 to 2019, no significant spatial autocorrelation was evident, indicating a more randomized geographical distribution of hepatitis B during this period and a narrowing of incidence rate disparities among regions. These changes may be attributed to the full implementation of "China's viral hepatitis prevention and control plan (2012 - 2020)" for gradually reducing the differences in hepatitis B prevalence between regions, as indicated by the high overall hepatitis B vaccination rate for newborns, which has been maintained at >90%; the popularization of hepatitis B prevention and control knowledge among the population; and the effective construction of the population immune barrier.[31].

Local spatial autocorrelation analysis has further illuminated the clustering patterns of hepatitis B incidence rates across diverse regions. Specifically, certain regions in the eastern part of the northeast exhibit low-low cluster regions, whereas regions in the northwest demonstrate high-low cluster regions. According to the findings of Ren et al [32] on the allocation of medical and health resources in China's provinces from 2010 to 2021, there has been an increasing trend in medical and health resources, and the allocation has essentially achieved parity in terms of quantity. However, high-quality medical and health resources in China are heavily concentrated in the eastern region, which may further explain the disparities in the spatial pattern of hepatitis B incidence among regions. Between 2004 and 2007, northwest and southwest regions, including Qinghai, Gansu, Ningxia, Sichuan, and Chongqing, exhibited high-high cluster regions. Nevertheless, this spatial aggregation gradually dissipated in 2007. This shift could be attributed to significant progress made in hepatitis B prevention and control efforts in these regions, leading to changes in the spatial clustering pattern of the disease. Since 2015, Guangdong and Xizang have transformed from areas with nonsignificant spatial cluster regions to high-high cluster regions. This transformation is likely linked to population movements, economic development, and improvements in health conditions in these regions. As an economically developed region, Guangdong Province has frequent population mobility, which is an important reason for increasing the risk of hepatitis B transmission [33]. However, due to the improvement of health conditions, the increase of medical resources, and the improvement of diagnostic levels in Xizang, the detection and reporting of hepatitis B is more accurate, thus showing a high-high clustering phenomenon [29].

This study offers fresh perspectives on the epidemic trends and spatial clustering patterns of hepatitis B in China, serving as a scientific foundation for developing targeted prevention and control measures. Future research should delve deeper into the factors influencing the variations in hepatitis B incidence rates across different regions, as well as explore more effective strategies for implementing preventive measures, particularly in regions with high incidence rates or evident increase trends. Concurrently, it is imperative to enhance vaccination and health education efforts targeting adults to mitigate the occurrence of hepatitis B.

This study has several limitations in analyzing hepatitis B surveillance data in China, which are worthy of further discussion. First of all, this study was limited to the analysis of the incidence of hepatitis B in China for 17 years (2004 to 2020) and failed to explore the influencing factors of the incidence of hepatitis B. In the future, a longitudinal research design should be considered to systematically assess the impact of various potential factors on the incidence rate of hepatitis B. Second, due to the limitations of data and methods, this study failed to fully analyze the differences between genders, differences between occupations, differences between urban and rural areas, and other factors that influence the incidence rate of hepatitis B in China. Therefore, future research should further improve the data and conduct a hierarchical analysis of these factors, to identify high-risk groups and develop targeted interventions. In addition, although economic factors, population density, medical service level, and medical resource allocation have an impact on the spatial clustering of hepatitis B, the impact of these factors on the regional incidence level still needs further research.

Conclusions

From 2004 to 2020, China witnessed a decrease in the overall incidence rate of hepatitis B, indicating remarkable progress in its hepatitis B prevention and control efforts. Nevertheless, challenges persist in specific populations and regions. Therefore, it is imperative to maintain vigilance in hepatitis B prevention and treatment, in addition to enhancing health education initiatives. Further, given the disparities in age and geographical distribution, efforts must be intensified to identify and screen high-risk groups and promote supplementary vaccination programs for adults. These comprehensive measures will alleviate the health burden of hepatitis B on the population, enhance the overall standard of healthy living, and establish a robust scientific foundation for the ongoing prevention and control of hepatitis B in China.

Acknowledgments

The authors would like to thank the Chinese Center for Disease Control and Prevention for sharing the valuable data, the fifth National Comprehensive Prevention and Control Demonstration Zone for AIDS (NO. State AIDS Office Letter [(2023) No. 4]), the Key Laboratory for Major Infectious Disease Monitoring in the Northern Xinjiang Area of the Corps (N0.2024PT01), and the Corps Guiding Science and Technology Program Project (N0.2024ZD026).

XSL•FO
RenderX

Data Availability

The data are available from the China Public Health Science Data Center [14].

Authors' Contributions

XL designed the study. JL and LM collected the data, and SS and XH performed the analysis and interpreted the results. TS and HG coordinated and drafted the manuscript. HG and TS contributed equally as corresponding authors.

Conflicts of Interest

None declared.

References

1. Rybicka M, Bielawski KP. Recent advances in understanding, diagnosing, and treating hepatitis B virus infection. Microorganisms 2020 Sep 15;8(9):1416. [doi: 10.3390/microorganisms8091416] [Medline: 32942584]
2. Hepatitis B. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/hepatitis-b [accessed 2024-08-09]
3. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
4. Liu F, Gai X, Wu Y, et al. Oncogenic β-catenin stimulation of AKT2-CAD-mediated pyrimidine synthesis is targetable vulnerability in liver cancer. Proc Natl Acad Sci U S A 2022 Sep 27;119(39):e2202157119. [doi: 10.1073/pnas.2202157119] [Medline: 36122209]
5. Shao Y, Su L, Hao R, Wang Q, Naranmandura H. Advances on molecular mechanism of hepatitis B virus-induced hepatocellular carcinoma. Zhejiang Da Xue Xue Bao Yi Xue Ban 2021 Feb 25;50(1):113-122. [doi: 10.3724/zdxbyxb-2021-0032] [Medline: 34117855]
6. Guo N, Wang J, Nicholas S, Maitland E, Zhu D. Behavioral differences in the preference for hepatitis B virus vaccination: a discrete choice experiment. Vaccines (Basel) 2020 Sep 14;8(3):527. [doi: 10.3390/vaccines8030527] [Medline: 32937824]
7. Zhang C, Liu Y, Zhao H, Wang G. Global, regional, and national burdens of cirrhosis in children and adolescents aged under 19 years from 1990 to 2019. Hepatol Int 2024 Feb;18(1):238-253. [doi: 10.1007/s12072-023-10531-y] [Medline: 37101103]
8. Liu J, Liang W, Jing W, Liu M. Countdown to 2030: eliminating hepatitis B disease, China. Bull World Health Organ 2019 Mar 1;97(3):230-238. [doi: 10.2471/BLT.18.219469] [Medline: 30992636]
9. Li T, Su S, Zhao Y, et al. Barriers to the prevention and control of hepatitis B and hepatitis C in the community of Southwestern China: a qualitative research. Int J Environ Res Public Health 2019 Jan 15;16(2):231. [doi: 10.3390/ijerph16020231] [Medline: 30650581]
10. Wang Y, Du Z, Lawrence WR, Huang Y, Deng Y, Hao Y. Predicting hepatitis B virus infection based on health examination data of community population. Int J Environ Res Public Health 2019 Dec 2;16(23):4842. [doi: 10.3390/ijerph16234842] [Medline: 31810204]
11. Cao G, Liu J, Liu M. Trends in mortality of liver disease due to hepatitis B in China from 1990 to 2019: findings from the Global Burden of Disease Study. Chin Med J (Engl) 2022 Sep 5;135(17):2049-2055. [doi: 10.1097/CM9.0000000000002331] [Medline: 36228164]
12. Jiao L, Shen T, Han Y, et al. The spatial-temporal distribution of hepatitis B virus infection in China,2006-2018. BMC Infect Dis 2024 Aug 11;24(1):811. [doi: 10.1186/s12879-024-09716-z] [Medline: 39129008]
13. Pavinati G, de Lima LV, Palmieri IGS, Magnabosco GT. Distribution and spatial autocorrelation of viral hepatitis B and C in Paraná, Brazil: an ecological study, 2011-2019. Epidemiol Serv Saude 2023;32(2):e2022888. [doi: 10.1590/S2237-96222023000200015] [Medline: 37556709]
14. Public Health Science Data Center. URL: https://www.phsciencedata.cn [accessed 2025-05-06]
15. China Statistical Yearbook. National Bureau of Statistics. URL: https://www.stats.gov.cn/sj/ndsj/ [accessed 2024-12-18]
16. Rosenberg PS, Check DP, Anderson WF. A web tool for age-period-cohort analysis of cancer incidence and mortality rates. Cancer Epidemiol Biomarkers Prev 2014 Nov;23(11):2296-2302. [doi: 10.1158/1055-9965.EPI-14-0300] [Medline: 25146089]
17. Liu Z, Li M, Hutton DW, et al. Impact of the national hepatitis B immunization program in China: a modeling study. Infect Dis Poverty 2022 Oct 11;11(1):106. [doi: 10.1186/s40249-022-01032-5] [Medline: 36221140]
18. Sun N, He F, Sun J, Zhu G. Viral hepatitis in China during 2002-2021: epidemiology and influence factors through a country-level modeling study. BMC Public Health 2024 Jul 8;24(1):1820. [doi: 10.1186/s12889-024-19318-8] [Medline: 38978017]
19. Wang X, Liu J, Wang Q, et al. Economic-related inequalities in hepatitis B virus infection among 115.8 million pregnant women in China from 2013 to 2020. EClinicalMedicine 2022 Jul;49(101465):101465. [doi: 10.1016/j.eclinm.2022.101465] [Medline: 35747197]

XSL•FO
RenderX

20. Gong X, Zheng C, Fang Q, Xu W, Yin Z. Survey of hepatitis B vaccination coverage and surface antibody-positive rates in people aged 1-59 years in 2006 and 2024. Open Forum Infect Dis 2024 Oct;11(10):ofae589. [doi: 10.1093/ofid/ofae589] [Medline: 39431151]
21. Ji W, Xie N, He D, Wang W, Li H, Wang K. Age-period-cohort analysis on the time trend of hepatitis B incidence in four prefectures of Southern Xinjiang, China from 2005 to 2017. Int J Environ Res Public Health 2019 Oct 14;16(20):3886. [doi: 10.3390/ijerph16203886] [Medline: 31615013]
22. Pinto CS, Costa GB, Allaman IB, Gadelha SR. Clinical, epidemiological aspects, and trends of hepatitis B in Brazil from 2007 to 2018. Sci Rep 2021 Jul 7;11(1):13986. [doi: 10.1038/s41598-021-93434-y] [Medline: 34234226]
23. Liang Y, Bai X, Liu X, et al. Hepatitis B vaccination coverage rates and associated factors: a community-based, cross-sectional study conducted in Beijing, 2019-2020. Vaccines (Basel) 2021 Sep 24;9(10):1070. [doi: 10.3390/vaccines9101070] [Medline: 34696178]
24. Fang K, Shi Y, Zhao Z, et al. Trends in disease burden of hepatitis B infection in Jiangsu Province, China, 1990-2021. Infect Dis Model 2023 Sep;8(3):832-841. [doi: 10.1016/j.idm.2023.07.007] [Medline: 37520113]
25. Zheng Z, Lin X, Huang Y, Zhang C, Zhang Z. Trends and age-period-cohort effect on incidence of hepatitis B from 2008 to 2022 in Guangzhou, China. Sci Rep 2024 Jun 11;14(1):13370. [doi: 10.1038/s41598-024-63796-0] [Medline: 38862511]
26. Su XJ, Yu Q. Epidemiological characteristics and spatiotemporal clustering analysis of hepatitis B virus infection in China from 2011 to 2021. Modern Preventive Medicine 2024;51(8):1384-1389. [doi: 10.20043/j.cnki.MPM.202312304]
27. Liang Z, Qiu J, Xiang Q, Yi J, Zhu J, Zhao Q. Epidemiology of hepatitis B virus infection among preconception couples in South China: a cross-sectional study. BMJ Open 2023 Jun 9;13(6):e061165. [doi: 10.1136/bmjopen-2022-061165] [Medline: 37295827]
28. Liu ZX, Wang YS, Wang WB. Study on changes in hepatitis B incidence in China, 1990-2017. Zhonghua Liu Xing Bing Xue Za Zhi 2021 Apr 10;42(4):613-619. [doi: 10.3760/cma.j.cn112338-20201026-01281] [Medline: 34814439]
29. Tian T, Hu YH, Jin M, Ciren D. Analysis of the trend of hepatitis B virus infection among residents in Tibet autonomous region from 2018 to 2022. Chinese Journal of Public Health 2024;40(7):782-785. [doi: 10.11847/zgggws1143128]
30. Qian J, Yue M, Huang P, et al. Spatiotemporal heterogeneity and impact factors of hepatitis B and C in China from 2010 to 2018: Bayesian space-time hierarchy model. Front Cell Infect Microbiol 2023;13(1115087):1115087. [doi: 10.3389/fcimb.2023.1115087] [Medline: 36923590]
31. Li J, Gao Z, Bai H, et al. Global, regional, and national total burden related to hepatitis B in children and adolescents from 1990 to 2021. BMC Public Health 2024 Oct 23;24(1):2936. [doi: 10.1186/s12889-024-20462-4] [Medline: 39443929]
32. Ren H, Li C, Huang Y. Spatial and temporal analysis of China's healthcare resource allocation measurements based on provincial data: 2010-2021. Front Public Health 2023;11(1269886):1269886. [doi: 10.3389/fpubh.2023.1269886] [Medline: 38074731]
33. Bai S, Dang W, Hong W, Liao W, Smith RD. The prevalence of hepatitis B in Chinese general population from 2018 to 2022: a systematic review and meta-analysis. BMC Infect Dis 2024 Feb 16;24(1):211. [doi: 10.1186/s12879-024-09103-8] [Medline: 38365596]

Abbreviations

AAPC: average annual percentage change
APC model: age-period-cohort-model
HBsAg: hepatitis B surface antigen
HBV: hepatitis B virus
HCC: hepatocellular carcinoma
RR: relative risk

Edited by A Mavragani; submitted 05.01.25; peer-reviewed by J Shao, Y Qiu; revised version received 10.03.25; accepted 11.03.25; published 12.05.25.

Please cite as:
Li X, Liang J, Ma L, Shen S, Han X, Sun T, Guo H
Epidemic Trends and Spatial Distribution Characteristics of Hepatitis B in China: Surveillance Study
JMIR Public Health Surveill 2025;11:e70888
URL: https://publichealth.jmir.org/2025/1/e70888
doi:10.2196/70888

© Xiaoxue Li, Jie Liang, Lanqing Ma, Shijie Shen, Xueying Han, Tao Sun, Heng Guo. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.5.2025. This is an open-access article distributed under the terms of

XSL•FO
RenderX

the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Spatiotemporal Trends in Deaths From External Causes in Brazil: 23-Year Ecological and Population-Based Study

Lucas Almeida Andrade[1*], MSc; Wandklebson Silva da Paz[2*], PhD; Luís Ricardo Santos de Melo[3], BSc; Débora dos Santos Tavares[4], PhD; Alvaro Francisco Lopes de Sousa[5,6*], PhD; Emerson Lucas Silva Camargo[7*], MSc; Carla Aparecida Arena Ventura[7*], PhD; Regina Claudia Silva Souza[5], PhD; Karina Conceição Gomes Machado de Araújo[3*], PhD; Carlos Dornels Freire de Souza[8*], PhD; Allan Dantas dos Santos[3,9*], PhD; Fagner Alfredo Ardisson Cirino Campos[7], PhD; Márcio Bezerra Santos[1,10], PhD

[1]Graduate Program in Health Sciences, Universidade Federal de Sergipe, Rosa Elze, São Cristóvão, Aracaju, Brazil

[10]Complex of Medical and Nursing Sciences, Federal University of Alagoas, Arapiraca, Alagoas, Brazil

[2]Graduate Program in Tropical Medicine, Center for Medical Sciences, Universidade Federal de Pernambuco, Recife, Brazil

[3]Graduate Program in Nursing, Universidade Federal de Sergipe, São Cristovão, Brazil

[4]Department of Health Education, Federal University of Sergipe, Lagarto, Sergipe, Brazil

[5]Institute of Teaching and Research, Hospital Sírio-Libanês, São Paulo, Brazil

[6]Public Health Research Centre, Comprehensive Health Research Center (CHRC), NOVA National School of Public Health, Universidade Nova de Lisboa, Lisbon, Portugal

[7]Ribeirão Preto College of Nursing, Universidade de São Paulo, Ribeirão Preto, Brazil

[8]College of Medicine, Universidade Federal do Vale do São Francisco, Petrolina, Brazil

[9]Department of Nursing, Federal University of Sergipe, Lagarto, Sergipe, Brazil

[*]these authors contributed equally

Corresponding Author:
Lucas Almeida Andrade, MSc
Graduate Program in Health Sciences, Universidade Federal de Sergipe, Rosa Elze, São Cristóvão, Aracaju, Brazil

Abstract

Background: Mortality from external causes is a major public health issue globally, with significant impacts on both lives and economies. In Brazil, external cause mortality has shown spatiotemporal variations across regions, influenced by social, economic, and demographic factors.

Objective: This study aimed to examine the spatiotemporal dynamics of mortality from external causes in Brazil for 23 years (2000 - 2022), identifying patterns across regions and demographic groups and assessing the major contributing causes of death.

Methods: This ecological study used data from the Brazilian Mortality Information System (SIM) and used joinpoint regression to analyze temporal trends, Moran I for spatial analysis, and Poisson scanning statistics for spatiotemporal patterns. A total of 3,240,023 deaths were analyzed, with specific attention given to regional and demographic disparities.

Results: The study found that mortality from external causes remained significant, with men and young adults (20 - 39 years) having the highest rates of death. The main causes of death were assaults (36.61%), transport accidents (26.55%), falls (7.83%), and self-harm (7.43%). Despite an overall decrease in mortality, increases were observed in the North and Northeast regions and among the older adults. High-risk areas were predominantly located in the North, Northeast, and Central-West regions. The mortality trends varied by region, with significant differences in risk across the country.

Conclusions: Although there was a general reduction in mortality from external causes in Brazil, this trend was not uniform across all regions. The North, Northeast, and Central-West regions showed the highest mortality risks, with men and young adults being the most affected demographic groups. These findings emphasize the need for targeted public health interventions that address the regional and demographic disparities in mortality from external causes.

(JMIR Public Health Surveill 2025;11:e68002) doi:10.2196/68002

KEYWORDS

mortality; accidents; violence; Brazilian; deaths; young adults; mortality information system; SIM; joinpoint regression; Moran's I; Poisson scanning statistics; population-based study; spatial analysis; public health; spatiotemporal analysis; historical series; self harm; assaults; temporal trends; external causes

XSL•FO
RenderX

Introduction

Mortality from external causes represents one of the main global public health problems. This phenomenon encompasses both unintentional events, such as traffic accidents, drownings, and falls, as well as intentional events, resulting from injuries caused by different forms of violence, such as suicide, homicide, femicide, and others [1]. Notably, these causes are impactful not only because of the number of lives lost but also due to the high socioeconomic burden they impose on countries and the negative consequences for affected families and health systems [1,2].

According to data from the World Health Organization (WHO), more than 4 million people die annually from external causes worldwide (about 8% of all global deaths) [1,3]. The WHO estimates that 29% of these fatal outcomes result from traffic accidents, 16% from suicides and falls, and 11% from homicides. Furthermore, teenagers and young adults aged 15-29 years are the most affected [3].

In Brazil, external causes constitute a significant issue in national mortality rates and represent substantial challenges to the national health system, being one of the major causes of disability and premature mortality [4,5]. In 2021, deaths from violence and accidents accounted for approximately 8% of all deaths across the country, ranking as the fourth leading cause of mortality, with particular emphasis on homicides and transport accidents, which together were responsible for more than 50% of deaths related to external factors in Brazilian territory [5].

Despite the significant socioeconomic impacts of accidents and violence worldwide, certain population groups and regions are more vulnerable [6,7]. Mortality patterns from external causes can be influenced by social, economic, and demographic determinants, such as income, age structure, the level of socioeconomic inequality, and the presence and efficiency of public policies in the areas of education, health, and safety [1,8]. In this context, deaths from external causes are not evenly distributed across Brazil and may present marked spatial heterogeneity across regions of the country [7,9,10]. This is due to Brazil's complex reality, which is manifested through distinct geographic, cultural, political, and socioeconomic characteristics, creating different mortality patterns throughout the country [11].

Importantly, in 2011, the Brazilian Ministry of Health launched the Strategic Action Plan to Combat Chronic Noncommunicable Diseases (NCDs) (2011 - 2022), with emphasis on effective, integrated, and sustainable strategies for the prevention and control of noncommunicable diseases and their risk factors. Nevertheless, as the plan's end year approached (in 2019), the Brazilian Ministry of Health began preparing a new document that reaffirmed and expanded the established proposals. The DANT Plan (Strategic Action Plan to Combat Chronic Diseases and Non-Communicable Diseases in Brazil, 2021 - 2030) also included external causes of death. This collective effort is part of the health agenda for the next 10 years, in line with the 2020 - 2030 Agenda on Sustainable Development Goals. The goals aim to reduce the mortality rate from traffic injuries by 50%, reduce the homicide mortality rate by one-third, and halt the growth of mortality from suicides and fall-related deaths among older adults in Brazil by 2030 [2].

In this regard, combating deaths from external causes requires a range of approaches that strengthen prevention and control strategies in Brazil. Therefore, spatial analysis techniques using geographic information systems are fundamental tools [12], as they provide a better understanding of the spatial dynamics of violent and accidental incidents, identifying high-mortality areas that require priority investment for effective interventions and public policies that consider the health realities of each region. This study aimed to analyze the spatiotemporal dynamics of mortality from external causes in Brazil and its regions for 23 years.

Methods

Type and Study Design

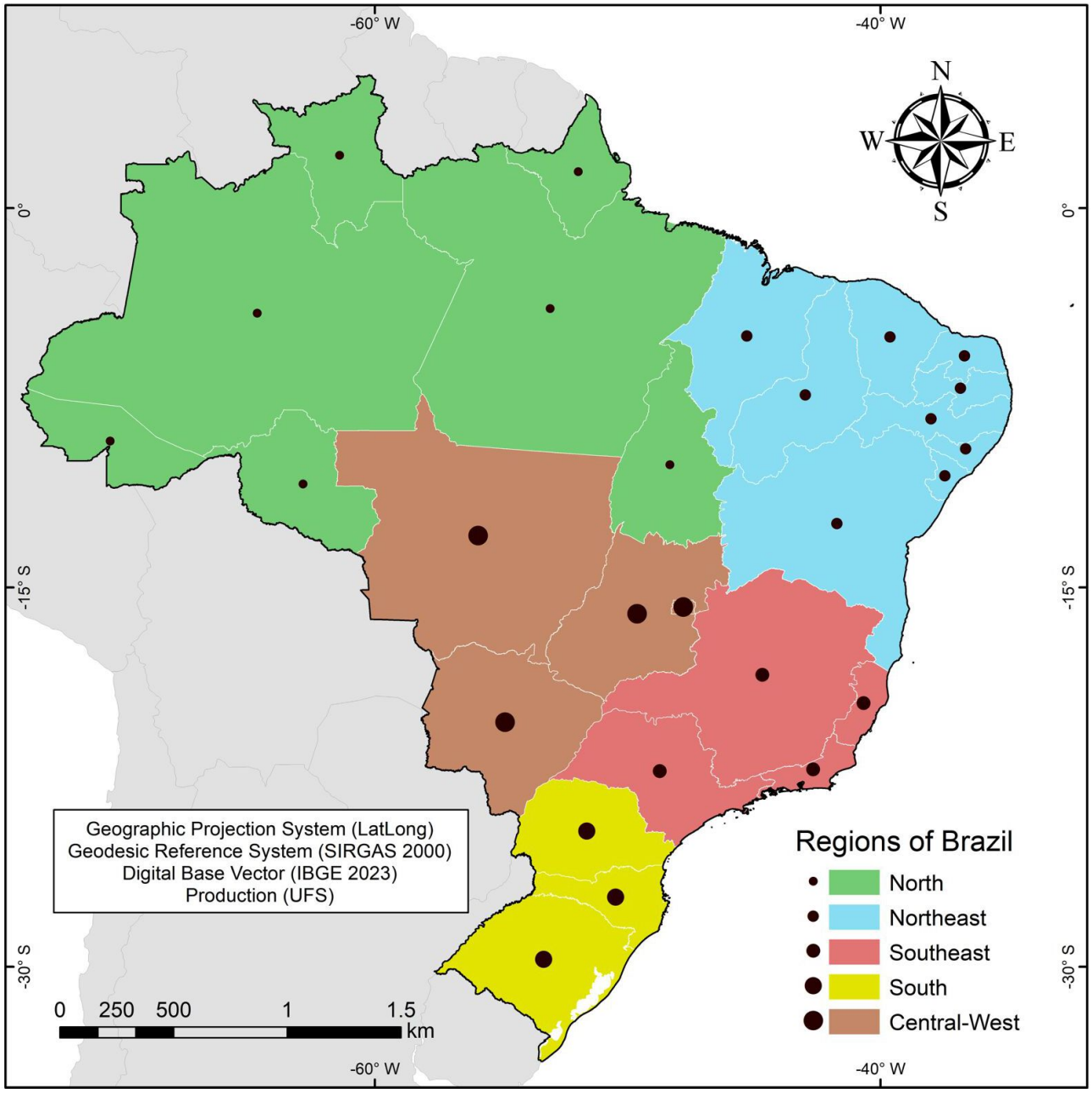
An ecological and population-based study was conducted using spatiotemporal analysis techniques. The study encompassed all deaths related to external causes in Brazil from 2000 to 2022. This time frame was established because, starting in 1999, Brazil adopted a new version of the Mortality Information System (SIM), which introduced an updated death certificate (DC). This update improved data recording, with more detailed completion of the DC [13]. When the data were collected for this study, the numbers of deaths from external causes were only available up to 2022. The 5 regions of the country and their 5570 municipalities were considered for all analyses.

Study Area

Brazil is located in South America, being the largest country on the continent, with a territorial area of 8,515,767.049 km². The Brazilian population is approximately 203 million inhabitants. The country is politically and administratively divided into 27 federative units (26 states and 1 federal district), with the capital city being Brasília. For political and operational purposes, the federative units are grouped into five regions (North, Northeast, Southeast, South, and Central-West) with distinct geographic, socioeconomic, and cultural characteristics [11] (Figure 1).

XSL•FO
RenderX

Figure 1. Study area: a map of Brazil divided into its 5 regions, 26 states, and 1 federal district. Coordinates based on latitude and longitude. Map elaborated by Universidade Federal de Sergipe. IBGE: Instituto Brasileiro de Geografia e Estatística; LatLong: latitude and longitude; SIRGAS: Sistema de Referência Geocêntrico para as Américas; UFS: Universidade Federal de Sergipe.



Data Source

Data regarding deaths from external causes were collected from the Brazilian SIM. The SIM plays a fundamental role in the process of collecting, storing, and managing death records in the country, using the DC as a standard document, which is a form filled out by medical professionals for all deaths occurring in Brazil. It is important to highlight that SIM data are in the public domain and available for access on the website of the Department of Informatics of the Unified Health System. For data acquisition, codes V01 to Y98 from the *ICD-10* (*International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*) were used.

In addition, population data were obtained from the Brazilian Institute of Geography and Statistics, based on information from the Brazilian population censuses conducted in 2000, 2010, and 2022, as well as intercensal population estimates for the years 2001 - 2009 and 2011 - 2019 [12]. The digital cartographic mesh of Brazil (divided by states and regions), based on the Geographic Projection System in shapefile format (Geodesic Reference System, SIRGAS/2000), was extracted from the Brazilian Institute of Geography and Statistics website for the preparation of spatial analysis maps.

Variables and Measures

The variables assessed were (1) number of deaths from external causes registered in the 5570 Brazilian municipalities; (2) age-standardized external cause mortality rates (ASMRs) based on the 1966 world standard population, as defined by Doll et al [14]. This standardization allows for data comparison at national and international levels, accounting for differences in

XSL•FO
RenderX

the age pyramids of diverse geographic areas and over time. ASMR rates were calculated in age-specific ranges (0 - 4, 5 - 9, ..., 75 - 79, and ≥80 years). The result was multiplied by 100,000 inhabitants in each municipality, state, and region. Rates were also calculated according to sex and age group.

Exploratory Data Analysis

The epidemiological variables used in the descriptive analysis were: sex (male and female), ethnicity/skin color ("White," "Black," "Yellow," "Brown," and "Indigenous"; in the Brazilian context, the terms White [*branco*], Black [*preto*], Brown [*pardo*], Yellow [*amarelo*], and Indigenous [*indígena*] are the official racial/ethnic categories established by the Brazilian Institute of Geography and Statistics and are systematically used in national censuses and health information systems, including the SIM), age group (0-9, 10-19, 20-39, 40-59, and ≥60 years), years of education (<8 and ≥8 years), and cause of death (according to the *ICD-10* code). These categorized variables were described for Brazil and its regions using absolute and relative frequencies.

Time Trend Analysis

Temporal trends were assessed through the joinpoint linear regression model (segmented linear regression), using standardized mortality rates from external causes by region, gender, and age group. This method allowed for the verification of changes in data trends over time by adjusting the data to a time series with the smallest possible number of joinpoints (zero, which indicates a line without inflection points) and testing whether the inclusion of more joinpoints was statistically significant [15].

The statistical significance test for choosing the best model was based on the Monte Carlo permutation method, considering $P<.05$ and 95% CI. To describe and quantify the trends, the annual percentage changes (APCs) and their respective 95% CIs were calculated. The average annual percentage change (AAPC) was calculated for the total period when more than one APC was obtained for segmented periods. The AAPC value was derived from the weighted geometric mean of the APCs, with weights corresponding to the length of each segment's time interval [16]. Trends were considered significant when APCs presented $P<.05$ and their 95% CI did not include a zero value. A positive and significant APC value indicated an increasing trend, while a negative and significant value indicated a decreasing trend. Nonsignificant trends were described as stable, regardless of APC values [16].

Spatial Analysis and Spatiotemporal Scanning

For spatial analyses, choropleth maps of Brazil were created to represent mortality from external causes across all municipalities. Initially, standardized mortality rates were used, but to minimize the instability caused by random fluctuations in cases, the local empirical Bayesian estimator was applied. This model smoothed the standardized rates through weighted averages, resulting in a new corrected coefficient that more accurately represented the epidemiological scenario [17]. This method also reduces data fluctuations in smaller areas. One of the advantages of Bayesian rates is the attribution of greater influence to neighboring municipalities, making the results more coherent at a regional level. To facilitate data visualization, the standardized and smoothed rates were presented in thematic maps, stratified into five categories of natural breaks (Jenks): (1) 0 - 45, (2) 45 - 61, (3) 61 - 77, (4) 77 - 97, and (5) 97 - 192 deaths per 100,000 inhabitants [17].

Subsequently, the Global Moran Index was calculated to investigate the existence of spatial autocorrelation in mortality from external causes and to determine if spatial patterns were present in the distribution of this variable. The Global Moran Index estimates the correlation between the values of a variable in different locations, ranging from −1 to +1. Values close to 0 indicate the absence of spatial autocorrelation, while positive values (0 to +1) indicate the presence of positive spatial autocorrelation. That is, areas with high values tend to be close to other areas with high values, and areas with low values tend to be close to other areas with low values. On the other hand, negative values (−1 to 0) indicate negative spatial autocorrelation, where areas with high values are close to areas with low values [17,18].

Once autocorrelation was identified, the occurrence of local autocorrelation was assessed by calculating the Local Moran Index (local indicators of spatial association), which determined the existence of patterns of spatial dependence. A scatter plot was also created with the following spatial quadrants: Q1 (high and high) and Q2 (low and low), indicating municipalities with similar values to their neighbors and with positive spatial association; Q3 (high and low) and Q4 (low and high), indicating municipalities with different values from their neighbors without spatial association. The results were considered statistically significant when $P<.05$ and represented in Moran maps [17,18].

Finally, spatiotemporal scanning analysis was applied to identify and evaluate spatiotemporal clusters of high risk for deaths from external causes. The identification of clusters occurred through scanning statistics (SaTScan; Kulldorff, 1997), using the retrospective space-time analysis type, through the Poisson probability distribution model, which met the following parameters: aggregation time of one year, no geographic or temporal overlap of clusters, circular clusters, maximum spatial cluster size of 50% of the population at risk, and a maximum temporal cluster size equal to 50% of the studied period [19]. Spatiotemporal clusters were detected using the log likelihood ratio test. Furthermore, the relative risks of mortality were calculated for each cluster in relation to its neighbors. Results with a $P<.05$ were considered significant, using 999 Monte Carlo simulations, and were represented in the form of maps and tables [19].

Software

Microsoft Excel 2017 was used for data tabulation and descriptive analysis, Joinpoint Regression Program 5.0.2 (Statistical Methodology and Applications Branch, Surveillance Research Program, National Cancer Institute) was used to assess time trends, TerraView 4.2.2 (Instituto Nacional de Pesquisas Espaciais – INPE, Brazil) was used to perform spatial analysis, QGIS 3.28.7 (QGIS Development Team, Open Source Geospatial Foundation) was used to create choropleth maps, and SaTScan 9.6 (developed by Martin Kulldorff and colleagues,

XSL•FO
RenderX

Harvard Medical School/Brigham and Women's Hospital, USA) was used to perform spatiotemporal scanning.

Ethical Considerations

This study used secondary data from the public domain, which did not contain any personal identifiers, and followed national and international ethical recommendations, including the rules of the Declaration of Helsinki and Resolution 466/2012 of the National Health Council.

Results

A total of 3,240,023 deaths related to external causes were registered in Brazil from 2000 to 2022. The regions with the highest percentages of deaths were the Southeast (n=1289,073, 39.79%) and the Northeast (n=945,850, 29.19%), together representing 68.98% of the total deaths due to external factors in the country. When assessing the sociodemographic characteristics of these deaths, the majority was among men (n=2,668,434, 82.36%), individuals identified as non-White (n=1,795,050, 55.11%), those aged between 20 and 39 years (1,432,969, 44.23%), and individuals with fewer than 8 years of education (n=1,554,174, 47.97%). Likewise, all Brazilian regions exhibited this profile described above, except for ethnicity in the Southeast and South regions, where deaths from external causes were higher among White individuals (n=656,634, 50.94% and n=385,393, 83%, respectively; Table 1).

Table . Sociodemographic characteristics of deaths from external causes in Brazil, 2000-2022.

Variable	North, n (%)	Northeast, n (%)	Southeast, n (%)	South, n (%)	Central-West, n (%)	Brazil, n (%)
Total	267,757 (8.26)	945,850 (29.19)	1,289,073 (39.79)	464,354 (14.33)	272,989 (8.43)	3,240,023 (100)
Sex						
Male	229,908 (85.86)	805,311 (85.14)	1,036,003 (80.37)	373,000 (80.33)	224,212 (82.13)	2,668,434 (82.36)
Female	37,463 (13.99)	139,786 (14.78)	251,498 (19.51)	91,026 (19.60)	48,501 (17.77)	568,274 (17.54)
Missing data	386 (0.15)	753 (0.08)	1572 (0.12)	328 (0.07)	276 (0.10)	3315 (0.10)
Ethnicity/skin color[a]						
White	36,861 (13.77)	118,525 (12.53)	656,634 (50.94)	385,393 (83.00)	89,309 (32.72)	1,286,722 (39.71)
Black	12,864 (4.80)	56,634 (5.99)	116,338 (9.02)	19,221 (4.14)	15,127 (5.54)	220,184 (6.80)
Yellow	578 (0.22)	1423 (0.15)	6604 (0.51)	1332 (0.29)	780 (0.29)	10,717 (0.03)
Brown	205,475 (76.74)	685,338 (72.46)	457,305 (35.48)	48,223 (10.38)	156,288 (57.25)	1,552,629 (47.92)
Indigenous	4976 (1.86)	1484 (0.16)	777 (0.06)	1132 (0.24)	3151 (1.15)	11,520 (0.36)
Missing data	7003 (2.61)	82,446 (8.71)	51,415 (3.99)	9053 (1.95)	8334 (3.05)	158,251 (5.18)
Age group (years)						
0 - 9	11,711 (4.37)	24,826 (2.62)	31,530 (2.45)	13,324 (2.87)	8856 (3.24)	90,247 (2.79)
10 - 19	36,250 (13.54)	123,432 (13.05)	134,550 (10.44)	47,736 (10.28)	31,168 (11.42)	373,136 (11.52)
20 - 39	134,872 (50.37)	459,366 (48.57)	528,599 (41.01)	187,859 (40.46)	122,273 (44.79)	1,432,969 (44.23)
40 - 59	54,991 (20.54)	196,959 (20.82)	294,582 (22.85)	117,050 (25.21)	63,516 (23.27)	727,098 (22.44)
≥60	26,529 (9.91)	134,032 (14.17)	273,152 (21.19)	95,183 (20.50)	43,933 (16.09)	572,829 (17.68)
Missing data	3404 (1.27)	7235 (0.77)	26,660 (2.06)	3202 (0.68)	3243 (1.19)	43,744 (1.34)
Years of study						
<8 years	150,569 (56.23)	496,858 (52.53)	561,106 (43.53)	219,420 (47.25)	126,221 (46.24)	1,554,174 (47.97)
≥8 years	63,800 (23.83)	145,772 (15.41)	301,412 (23.38)	118,270 (25.47)	67,569 (24.75)	696,823 (21.51)
Missing data	53,388 (19.94)	303,220 (32.06)	426,555 (33.09)	126,664 (27.28)	79,199 (29.01)	989,026 (30.52)

[a]In the Brazilian context, the terms White (*branco*), Black (*preto*), Brown (*pardo*), Yellow (*amarelo*), and Indigenous (*indígena*) are the official racial/ethnic categories established by the Brazilian Institute of Geography and Statistics and are systematically used in national censuses and health information systems, including the Mortality Information System.

Table 2 shows that assaults (n=1,186,257, 36.61%) and transport accidents (n=860,386, 26.55%) were the main causes of death due to external factors in Brazil and its regions. Other significant causes were tumbles/falls (n=253,705, 7.83%), intentional self-inflicted injuries (n=240,843, 7.43%), and drowning (n=127,547, 3.94%). The North and Northeast regions recorded the highest rates of deaths from assaults, at 46% (n=123,167) and 43.8% (n=414,239), respectively. In contrast, transport accidents were the main cause in the South (n=151,712, 32.67%), and intentional self-inflicted injuries were the third leading cause in this region, exhibiting a higher percentage than other regions (n=57,138, 12.3%). In addition, 267,419 (8.25%)

XSL•FO
RenderX

deaths were recorded as being of undetermined intent nationwide.

Regarding the time trend analyses, there was a slight decrease in standardized mortality rates in Brazil over the past 23 years, with an AAPC of −0.3 (95% CI −0.6 to −0.1; Table 3). Similarly, decreasing trends were observed in the Southeast (AAPC −1.8; 95% CI −1.9 to −1.6) and Central-West (AAPC −0.5; 95% CI −0.8 to −0.1) regions. In contrast, there was an increase in mortality rates in the Northeast (AAPC 1.2; 95% CI 0.9 to 1.6) and North (AAPC 0.8; 95% CI 0.5 to 1.2), while the South region showed stability (AAPC −0.2; 95% CI −0.4 to 0.1).

XSL•FO
RenderX

Table . Main causes of death from external causes in Brazil and its regions, 2000-2022.

Main causes of death from external causes	ICD-10[a] codes	Frequency, n (%)
Brazil		
Assaults	X85-Y09	1,186,257 (36.61)
Transport accidents	V01-V99	860,386 (26.55)
Events with undetermined intent	Y10-Y34	267,419 (8.25)
Tumbles/falls	W00-W19	253,705 (7.83)
Self-inflicted intentional injuries	X60-X84	240,843 (7.43)
Accidental drowning and submersion	W65-W74	127,547 (3.94)
Other accidental risks to breathing	W75-W84	67,476 (2.08)
Accidental exposure to other specified and unspecified factors	X58-X59	58,475 (1.80)
Exposure to electric current, radiation, and extreme temperatures and pressures	W85-W99	33,144 (1.02)
Complications of medical and surgical care	Y40-Y84	32,011 (0.99)
Other causes	Other codes	112,760 (3.48)
North		
Assaults	X85-Y09	123,167 (46.00)
Transport accidents	V01-V99	70,789 (26.44)
Self-inflicted intentional injuries	X60-X84	16,566 (6.19)
Accidental drowning and submersion	W65-W74	16,210 (6.05)
Tumbles/falls	W00-W19	12,597 (4.70)
Events with undetermined intent	Y10-Y34	7,344 (2.74)
Exposure to inanimate mechanical forces	W20-W49	5,276 (1.97)
Exposure to electric current, radiation, and extreme temperatures and pressures	W85-W99	3,944 (1.47)
Other accidental risks to breathing	W75-W84	2,248 (0.84)
Accidental exposure to other specified and unspecified factors	X58-X59	1,915 (0.72)
Other causes	Other codes	7,701 (2.88)
Northeast		
Assaults	X85-Y09	414,239 (43.80)
Transport accidents	V01-V99	234,451 (24.79)
Events with undetermined intent	Y10-Y34	76,489 (8.09)
Self-inflicted intentional injuries	X60-X84	54,545 (5.77)
Tumbles/falls	W00-W19	50,182 (5.31)
Accidental drowning and submersion	W65-W74	40,984 (4.33)
Other accidental risks to breathing	W75-W84	15,218 (1.61)
Exposure to electric current, radiation, and extreme temperatures and pressures	W85-W99	13,322 (1.41)
Accidental exposure to other specified and unspecified factors	X58-X59	11,940 (1.26)
Complications of medical and surgical care	Y40-Y84	9,242 (0.98)
Other causes	Other codes	25,238 (2.67)
Southeast		
Assaults	X85-Y09	417,643 (32.40)

XSL•FO
RenderX

Main causes of death from external causes	ICD-10[a] codes	Frequency, n (%)
Transport accidents	V01-V99	314,813 (24.42)
Events with undetermined intent	Y10-Y34	155,007 (12.02)
Tumbles/falls	W00-W19	122,794 (9.53)
Self-inflicted intentional injuries	X60-X84	90,545 (7.02)
Accidental drowning and submersion	W65-W74	42,164 (3.27)
Accidental exposure to other specified and unspecified factors	X58-X59	37,820 (2.93)
Other accidental risks to breathing	W75-W84	36,401 (2.82)
Complications of medical and surgical care	Y40-Y84	17,661 (1.37)
Legal interventions and operations of war	Y35-Y36	11,555 (0.90)
Other causes	Other codes	42,670 (3.31)
South		
Transport accidents	V01-V99	151,712 (32.67)
Assaults	X85-Y09	131,577 (28.34)
Self-inflicted intentional injuries	X60-X84	57,138 (12.30)
Tumbles/falls	W00-W19	46,077 (9.92)
Events with undetermined intent	Y10-Y34	19,896 (4.28)
Accidental drowning and submersion	W65-W74	17,955 (3.87)
Other accidental risks to breathing	W75-W84	9,582 (2.06)
Exposure to inanimate mechanical forces	W20-W49	5,315 (1.14)
Exposure to smoke, fire, and flames	X00-X09	4,759 (1.02)
Exposure to electric current, radiation, and extreme temperatures and pressures	W85-W99	4,586 (0.99)
Other causes	Other codes	15,757 (3.39)
Central-West		
Assaults	X85-Y09	99,631 (36.50)
Transport accidents	V01-V99	88,621 (32.46)
Tumbles/falls	W00-W19	22,055 (8.08)
Self-inflicted intentional injuries	X60-X84	22,049 (8.08)
Accidental drowning and submersion	W65-W74	10,234 (3.75)
Events with undetermined intent	Y10-Y34	8,683 (3.18)
Other accidental risks to breathing	W75-W84	4,027 (1.48)
Exposure to electric current, radiation, and extreme temperatures and pressures	W85-W99	3,497 (1.28)
Exposure to inanimate mechanical forces	W20-W49	3,328 (1.22)
Accidental exposure to other specified and unspecified factors	X58-X59	2,411 (0.88)
Other causes	Other codes	8,453 (3.10)

[a]*ICD-10*: *International Classification of Diseases, 10th Revision*.

XSL•FO
RenderX

Table . Temporal trends of mortality rates from external causes by region, sex, and age group in Brazil, 2000-2022.

Variable and period	Segmented period, APC[b] (95% CI)	Trend	Entire period, AAPC[c] (95% CI)	Trend
Brazil			−0.3[a] (−0.6 to −0.1)	Decreasing
2000 - 2017	0.0 (−0.3 to 0.3)	Stable		
2017 - 2020	−5.0 (−7.0 to 0.0)	Stable		
2020 - 2022	5.0 (−1.5 to 8.9)	Stable		
Region				
North			0.8[a] (0.5 to 1.2)	Increasing
2000 - 2016	2.2[a] (1.8 to 2.9)	Increasing		
2016 - 2022	−2.8[a] (−4.9 to −1.3)	Decreasing		
Northeast			1.2[a] (0.9 to 1.6)	Increasing
2000 - 2014	2.9[a] (2.3 to 3.7)	Increasing		
2014 - 2022	–1.6[a] (–3.3 to –0.4)	Decreasing		
Southeast			–1.8[a] (–1.9 to –1.6)	Decreasing
2000-2003	–0.3 (–1.7 to 1.9)	Stable		
2003 - 2007	–4.9[a] (–6.5 to –3.7)	Decreasing		
2007 - 2014	–0.9 (–1.4 to 0.7)	Stable		
2014 - 2020	–3.3[a] (–4.6 to –2.8)	Decreasing		
2020 - 2022	4.6[a] (1.8 to 6.5)	Increasing		
South			–0.2 (–0.4 to 0.1)	Stable
2000 - 2004	2.4[a] (0.8 to 6.0)	Increasing		
2004 - 2017	–0.8[a] (–1.2 to –0.4)	Decreasing		
2017 - 2020	–4.6[a] (–6.0 to –2.7)	Decreasing		
2020 - 2022	6.0[a] (2.4 to 9.2)	Increasing		
Central-West			–0.5[a] (–0.8 to –0.1)	Decreasing
2000 - 2004	2.0[a] (0.4 to 6.7)	Increasing		
2004 - 2007	–3.7[a] (–5.5 to –0.9)	Decreasing		
2007 - 2013	2.4[a] (1.1 to 5.7)	Increasing		
2013 - 2019	–4.1[a] (–7.0 to –3.0)	Decreasing		
2019 - 2022	0.9 (–1.8 to 4.9)	Stable		
Sex				
Males			–0.4[a] (–0.7 to –0.1)	Decreasing
2000 - 2017	0.0 (–0.3 to 0.3)	Stable		
2017 - 2020	–5.2[a] (–7.1 to -0.3)	Decreasing		
2020 - 2022	4.6 (–1.6 to 8.3)	Stable		
Females			0.2 (0.0 to 0.5)	Stable
2000 - 2007	–0.3 (–3.7 to 0.6)	Stable		
2007 - 2012	1.8[a] (0.8 to 3.3)	Increasing		
2012 - 2020	–2.1[a] (–2.8 to –1.6)	Decreasing		

XSL•FO
RenderX

Variable and period	Segmented period, APC[b] (95% CI)	Trend	Entire period, AAPC[c] (95% CI)	Trend
2020 - 2022	7.7[a] (4.7 to 9.7)	Increasing		
Age group (years)				
0 - 9			–1.5[a] (–1.8 to –1.3)	Decreasing
2000 - 2006	-4.3[a] (-5.5 to -3.5)	Decreasing		
2006-2011	1.3[a] (0.0 to 4.0)	Increasing		
2011 - 2020	-3.5[a] (–4.4 to –3.0)	Decreasing		
2020 - 2022	9.9[a] (5.3 to 12.9)	Increasing		
10-19			–1.1[a] (–1.6 to –0.6)	Decreasing
2000 - 2005	-2.0 (-8.2 to 0.7)	Stable		
2005 - 2016	2.7[a] (1.8 to 6.3)	Increasing		
2016 - 2022	–6.9[a] (–9.4 to –4.8)	Decreasing		
20-39			–0.4[a] (–0.9 to –0.1)	Decreasing
2000 - 2016	–0.1 (–2.3 to 0.6)	Stable		
2016 - 2019	–5.2 (–7.6 to 2.6)	Stable		
2019 - 2022	2.9 (–2.0 to 9.0)	Stable		
40-59			–0.5[a] (–0.7 to –0.3)	Decreasing
2000 - 2005	1.4[a] (0.6 to 3.0)	Increasing		
2005 - 2008	–4.7[a] (–6.0 to –2.6)	Decreasing		
2008 - 2012	1.6[a] (0.4 to 3.7)	Increasing		
2012 - 2019	–2.5[a] (–4.1 to –1.9)	Decreasing		
2019 - 2022	2.4[a] (0.3 to 5.6)	Increasing		
≥60			1.2[a] (0.8 to 1.8)	Increasing
2000 - 2004	4.5[a] (1.3 to 13.2)	Increasing		
2004 - 2022	0.5 (–0.1 to 0.8)	Stable		

[a]*P*<.05 (statistically significant).

[b]APC: annual percentage change.

[c]AAPC: average annual percentage change.

In addition, a reduction in mortality rates among men (AAPC −0.4) and stability among women (AAPC 0.2) was verified. Most importantly, from 2020 onward, a significant increase in death rates among women was observed (AAPC 7.7). A decreasing trend in mortality rates across all age groups, except for the older adults (aged ≥60 years; AAPC 1.2), was also evident.

Figure 2A–B shows maps of the spatial distribution of standardized and smoothed mortality rates due to external causes in Brazilian municipalities. Considering standardized rates, areas with high mortality (>97/100,000 inhabitants) were detected across all regions of the country. Nevertheless, when assessing the smoothed rates, municipalities with high mortality were mainly concentrated in the North and Central-West regions. A positive and significant spatial autocorrelation of mortality from external causes in Brazil was found (Moran I 0.53; *P*=.001). These results confirm the existence of spatial dependence in mortality from accidents and violence among Brazilian municipalities. In addition, 893 municipalities with high mortality rates (high and high–in red), forming clusters with a high risk of death due to violent and accidental incidents—especially in the North, Central-West, and Northeast regions—were identified (Figure 2C).

Finally, the spatiotemporal scanning analysis identified the formation of 3 statistically significant clusters of high-risk mortality from external causes. Notably, the primary cluster presented a relative risk of 1.22 and included a large portion of the national territory, encompassing 2922 municipalities and 21 out of 27 federative units, mainly located in the North, Northeast, and Central-West regions (Figure 2D and Table 4).

XSL•FO
RenderX

Figure 2. Spatial analysis of mortality rates from external causes in Brazil, 2000-2022. (A) Age-standardized mortality rate, (B) smoothed mortality rate, (C) Moran map (LISA cluster), and (D) spatiotemporal scanning analysis. LISA: Local Indicators of Spatial Association.

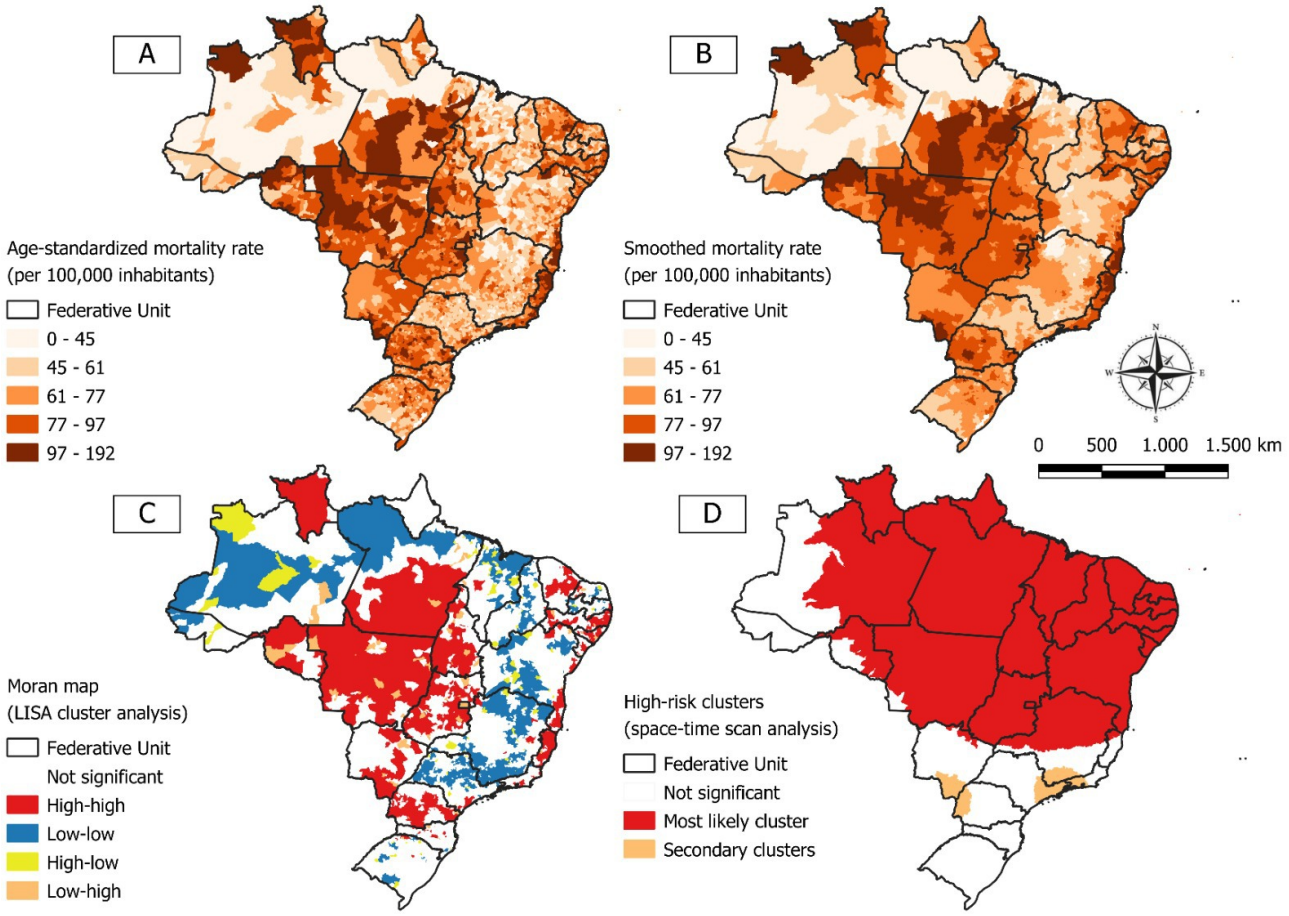


Table . Spatiotemporal scan analysis of mortality rates from external causes in Brazil, 2000-2022.

Cluster	Time period	Municipalities, n	States[a]	Observed	Expected	RR[b]	LLR[c]	*P* value
1	2012 - 2022	2922	AM, AP, RR, RO, PA, TO, MA, PI, CE, RN, PB, PE, AL, SE, BA, MG, ES, GO, DF, MT, MS	844,818	725,563	1.22	12,233.58	.001
2	2000 - 2003	299	MG, RJ, SP	147,643	113,725	1.31	4806.37	.001
3	2003 - 2013	81	MS, PR	21,915	15,341	1.43	1248.35	.001

[a]AM: Amazonas; AP: Amapá; RR: Roraima; RO: Rondônia; PA: Pará; TO: Tocantins; MA: Maranhão; PI: Piauí; CE: Ceará; RN: Rio Grande do Norte; PB: Paraíba; PE: Pernambuco; AL: Alagoas; SE: Sergipe; BA: Bahia; MG: Minas Gerais; ES: Espírito Santo; GO: Goiás; DF: Distrito Federal; MT: Mato Grosso; MS: Mato Grosso do Sul; RJ: Rio de Janeiro; SP: São Paulo; PR: Paraná.

[b]RR: relative risk.

[c]LLR: log likelihood ratio.

Discussion

Principal Findings

To the best of our knowledge, this is the first study to assess the dynamics of mortality from external causes across all regions of Brazil, applying different statistical and spatiotemporal techniques over an extensive period of 23 years. The findings indicate a slight decrease in the mortality rate from violent and accidental incidents in Brazil during this period. Furthermore, a reduction or stability was identified in almost all regions and population groups, except for the North and Northeast regions and the older adults, where increasing trends in these deaths were observed. Taken together, these results highlight the concerning scenario of deaths from external causes in Brazil

XSL•FO
RenderX

and the challenges the country will face in achieving the goals of the DANT Plan, agreed upon with the WHO by 2030 [2].

Notably, accidents and violence are considered serious public health problems and are among the leading causes of physical disability and mortality worldwide [1]. Faced with the dramatic magnitude of this phenomenon, several countries have implemented measures to reduce the number of incidents and deaths from these causes, such as the development of public safety policies and programs, the enactment of stricter laws [20,21], increased traffic inspection and education, and improvements in road infrastructure [22].

In this context, Brazil has also implemented strategies to mitigate these problems and reduce the number of deaths from external causes. In addition to the Dant Plan, other initiatives have been adopted, such as the National Policy for Reducing Morbidity and Mortality from Accidents and Violence [23], the Disarmament Statute [24], the implementation of qualified public safety actions and programs by some municipalities and states [25], and the National Policy for the Prevention of Self-Mutilation and Suicide [26]. Safety awareness campaigns have also been implemented, and more severe penalties have been applied for traffic violations across the country [27-29].

As also shown in previous studies [25,30,31], we strongly believe that these measures may have contributed to the reduction of mortality rates from violence and accidents in Brazil, as identified in this study. Nonetheless, it is important to highlight that significant and lasting changes in these rates occur slowly and gradually, and the general positive effects are intrinsically related to the consistency and effectiveness of these interventions. These effects can only be perceived in the long term [8].

Remarkably, reductions in death rates from external causes do not occur equally across different regions of the world. The problem involving external causes is complex and can be influenced by a variety of socioeconomic, demographic, and political factors [1]. Similarly, in Brazil, regions with greater socioeconomic vulnerability, such as the North and Northeast, showed an increase in mortality due to accidents and violence. Brazil has continental dimensions and is one of the most unequal countries in the world [11]. As a result, these disparities can influence rates and patterns of deaths from external causes across regions [1]. Unfortunately, certain population groups and regions are more affected, and the highest death rates are observed especially in low-income areas [32] and in those where prevention policies and strategies are absent or ineffective [33].

Regarding sociodemographic factors, a significant percentage of deaths occurred among men and young adults (aged 20-39 years). These findings are consistent with previous studies that identified men and young adults as the groups most affected by violent and accidental incidents [34,35]. Men and young adults are more vulnerable to deaths from external causes, largely due to their lifestyles, which often expose them to individual risk factors for accidents and violence, such as greater involvement with alcohol, drugs, firearms, participation in organized crime [9,36], and disregard for traffic laws and regulations [35]. This scenario has a significant impact on society, leading to the loss of economic and intellectual potential, which can hinder economic growth and delay the development of a region or country [2].

On the other hand, mortality from external causes among women showed a significant increase from 2020 onward. This increase is likely related to the evolution of women's roles in Brazilian society, leading them to occupy the same spaces and adopt lifestyles similar to those of men, which can increase risk behaviors for accidents and violence [37]. Furthermore, some authors believe that the reduction in the federal public budget for policies to combat violence against women, along with the encouragement of radical and conservative politicians, helped strengthen patriarchy and made access to firearms more flexible from 2019 to 2022, which likely contributed to the intensification of gender-based violence against women and femicide in recent years [25,38].

There was also an increasing trend in mortality rates from external causes among older adults, which can be attributed, in part, to the increase in life expectancy and the growth of this population in Brazil during recent decades. Older adults, due to the processes of senescence and senility, may experience a decrease in physical and cognitive capabilities, making them more susceptible to incidents such as falls [39], drownings [40], and traffic accidents [41]. Furthermore, aging can increase the occurrence of health conditions and stressors that raise the risk of suicide, such as a higher prevalence of mental disorders, marital problems, family losses, and social isolation [42].

As expected, assaults and transport accidents accounted for 63.16% of all deaths from external causes in Brazil and its regions. Undeniably, Brazil is widely recognized as one of the most violent countries in South America, being part of the group of nations with the highest risk of homicides [43]. This alarming scenario of violence is related to a series of factors, with social inequities emerging as a key element. Homicide rates are closely linked to structural inequalities, and violence is particularly intense in poorer areas, neglected urban spaces, or peripheral regions, such as favelas, where the population faces greater exposure to factors such as involvement with drugs, weapons, trafficking, and organized crime [9,44].

As for traffic accidents, several factors contribute to the high death rates in Brazil. In recent years, the country has seen exponential growth in the fleet of motor vehicles and motorcycles [45]. However, this increase has not been accompanied by improvements in road infrastructure, and many highways remain poorly maintained, with inadequate signage and a lack of safety devices, which considerably increases the risk of accidents [46]. Even more concerning, risky behaviors are quite common among drivers in the country, including driving under the influence of alcohol and drugs, speeding, and neglecting the proper use of safety equipment such as helmets and seat belts [47].

Spatial analysis and spatiotemporal scanning techniques used in this study enabled the identification of areas with a high risk of death related to external causes, concentrated mainly in the North, Northeast, and Central-West regions of Brazil. Previous studies have already classified these areas as having a high concentration of deaths from external causes [7,48].

XSL•FO
RenderX

In fact, the North, Northeast, and Central-West regions exhibited high-risk areas for mortality from external causes, and several factors are intrinsically linked to this concerning scenario. These Brazilian regions have experienced a process of rural exodus and intense, disorganized urbanization, which has driven the emergence of deprived areas lacking adequate public policies and infrastructure. As a result, there has been an increase in social inequities and structural determinants related to interpersonal and self-inflicted violence [49,50]. In addition, the migration of criminal factions from the Southeast region to states in the North and Northeast, along with the presence of border areas in the Central-West dominated by drug traffickers, has contributed to the expansion of the drug trafficking market, areas of armed conflict, and, ultimately, increased the risk of violent deaths [51]. Some states in the North and Central-West also face high rates of rural conflicts and territorial disputes, especially between Indigenous communities and large cattle ranchers and farmers [52,53]. Unfortunately, these conflicts result in an alarming number of violent deaths each year, especially among Indigenous people and rural producers, and significantly contribute to the increase in deaths from external causes in these regions of Brazil.

Despite advances in recent years, mortality from external causes remains a major public health challenge in Brazil. Accidents and violence are preventable, and it is the responsibility of policymakers and government officials to take effective measures to reduce them [8]. In this context, it is essential to improve current legislation and public security policies to combat drug trafficking and organized crime, as well as to enforce stricter regulation of civilian gun ownership [54,55]. It is equally important to strengthen violence prevention programs, including those targeting vulnerable populations such as women, Black people, and Indigenous communities [25], alongside expanding traffic enforcement and improving road infrastructure [22]. Other essential strategies may include improving urban mobility to reduce the risk of falls among older adults [56], ensuring adequate access to mental health services for individuals in psychological distress, and strengthening preventive actions against suicide [26].

Implications for Public Health and Surveillance

This study provides critical insights into the persistent public health challenge posed by mortality from external causes across Brazil. The findings emphasize the importance of targeted, region-specific public health interventions, particularly in high-risk areas such as the North, Northeast, and Central-West regions, where socioeconomic disparities and insufficient infrastructure significantly contribute to mortality rates.

The spatiotemporal patterns observed suggest the need for continuous and enhanced public health surveillance systems capable of detecting changes in mortality trends and risk factors specific to different demographic groups, such as young adults, men, and, increasingly, older adults. Surveillance systems that incorporate real-time data, predictive analytics, and spatial analysis will be essential in enabling timely, evidence-based public health responses to prevent violence and accidents.

Public policy makers and health program developers should consider these findings when planning interventions that address the intersection of multiple social determinants of health influencing mortality patterns, by adopting strategies that integrate cartographic approaches and analyses. In addition, efforts to prevent external causes of death should prioritize stricter enforcement of traffic laws, the development of violence prevention programs, and the implementation of policies targeted at high-risk populations and geographic areas. Enhanced public health strategies focused on injury and violence prevention, coupled with robust surveillance systems, are fundamental for Brazil to achieve its health targets under the WHO's DANT plan by 2030, improving overall population health and reducing preventable mortality across the country.

Limitations

This study has some limitations that should be acknowledged. Because secondary data were used, their quality cannot be guaranteed, and the values used in the analyses could be either underestimated or overestimated. As this is an ecological study, the results observed at the group level may not precisely reflect what occurs at the individual level. However, despite these limitations, the results described herein still provide valuable insights into the spatiotemporal panorama of mortality from external causes in Brazil and can support the development of public policies and interventions aimed at reducing mortality, particularly in high-risk areas.

Conclusion

Taken together, the results showed a slight reduction in mortality from external causes in Brazil. However, there was an increase in death rates in the North and Northeast regions of the country and among older adults, alongside high mortality rates among men and young adults. In addition, assaults and transport accidents remained the leading causes of death across all regions. Spatial analysis and spatiotemporal scanning revealed that the distribution of deaths from these causes is heterogeneous across Brazilian territory, with the primary risk areas concentrated in the North, Northeast, and Central-West regions. Given this scenario, we emphasize the need for intersectoral public policies that cover the entire national territory. These policies must involve greater allocation of resources and strategic targeting of prevention and control measures, focusing efforts on the most critical areas. Only in this way can the country achieve the goals of the Dant Plan, agreed with the WHO by 2030.

Acknowledgments

This research was funded by Hospital Sírio-Libanês and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brazil, with Finance Code 001.

XSL•FO
RenderX

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Injuries and violence. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/injuries-and-violence [accessed 2025-08-13]
2. Plano de Ações Estratégicas para o Enfrentamento das Doenças Crônicas e Agravos não Transmissíveis no Brasil 2021-2030 [Report in Portugese]. : Ministério da Saúde; 2021 URL: https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/svsa/doencas-cronicas-nao-transmissiveis-dcnt/09-plano-de-dant-2022_2030.pdf/view [accessed 2024-07-14]
3. World health statistics 2023: monitoring health SDGs, sustainable development goals. : World Health Organization; 2023 URL: https://www.who.int/publications/i/item/9789240074323 [accessed 2024-07-14]
4. França EB, Passos VMDA, Malta DC, et al. Cause-specific mortality for 249 causes in Brazil and states during 1990-2015: a systematic analysis for the global burden of disease study 2015. Popul Health Metr 2017 Nov 22;15(1):39. [doi: 10.1186/s12963-017-0156-y] [Medline: 29166948]
5. Mortes por causas externas: qualificação dos registros inespecíficos [Report in Portugese]. : Ministério da Saúde; 2024 URL: https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/svsa/vigilancia/mortes-por-causas-externas-qualificacao-dos-registros-inespecificos/view [accessed 2024-07-14]
6. Lawrence WR, Freedman ND, McGee-Avila JK, et al. Trends in mortality from poisonings, firearms, and all other injuries by intent in the US, 1999-2020. JAMA Intern Med 2023 Aug 1;183(8):849-856. [doi: 10.1001/jamainternmed.2023.2509] [Medline: 37399025]
7. Malta DC, Aquino ÉD, Veloso GA, et al. Mortality by road transport injury in Brazilian municipalities between 2000 and 2018. Public Health 2023 Jul;220:120-126. [doi: 10.1016/j.puhe.2023.04.013] [Medline: 37300976]
8. Nadanovsky P, Santos APP. Mortes por causas externas no Brasil: previsões para as próximas duas décadas [Report in Portuguese]. : Saude Amanha; 2021 URL: https://tinyurl.com/4dcnuftb [accessed 2024-07-14]
9. Silva CD, Souza KD, Paz WD, et al. Spatial modeling of homicide mortality in the Northeast region of Brazil. Rev Bras Enferm 2023;76(2):e20220182. [doi: 10.1590/0034-7167-2022-0182] [Medline: 36753255]
10. Maranhão TA, Sousa GJB, Silva TL, Dos Santos Araujo GA, et al. Temporal and spatial dynamics and factors related to suicide mortality among the elderly. J Bras Psiquiatr 2022;71:108-116. [doi: 10.1590/0047-2085000000367]
11. Censo demográfico 2022 [Web page in Portugese]. Instituto Brasileiro de Geografia e Estatística (IBGE). 2023. URL: https://www.ibge.gov.br/estatisticas/sociais/saude/22827-censo-demografico-2022.html [accessed 2024-07-14]
12. População. Instituto Brasileiro de Geografia e Estatística. 2021. URL: https://www.ibge.gov.br/estatisticas/sociais/populacao.html [accessed 2024-07-14]
13. Manual de procedimentos do sistema de informações sobre mortalidade [Report in Portugese]. : Ministèrio da Saúde. Fundaçao Nacional de Saúde; 2001 URL: https://bvsms.saude.gov.br/bvs/publicacoes/sis_mortalidade.pdf [accessed 2024-07-20]
14. Doll R, Payne P, Waterhouse J. Cancer Incidence in Five Continents: A Technical Report: Springer-Verlag (for UICC); 1966. URL: https://link.springer.com/book/9783540034759 [accessed 2024-05-13]
15. Joinpoint trend analysis software. : National Cancer Institute URL: https://surveillance.cancer.gov/joinpoint/ [accessed 2024-06-13]
16. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Statist Med 2000 Feb 15;19(3):335-351. [doi: 10.1002/(SICI)1097-0258(20000215)19:3<335::AID-SIM336>3.3.CO;2-Q]
17. Bailey TC, Gatrell AC. Interactive Spatial Data Analysis, 1st edition: Longman Scientific & Technical; 1995.
18. Anselin L. Exploring Spatial Data with GeoDa TM: A Workbook Center for Spatially Integrated Social Science 2005. URL: https://www.geos.ed.ac.uk/~gisteac/fspat/geodaworkbook.pdf [accessed 2024-05-16] [doi: 10.1201/9781003274919]
19. SaTScan. 2005. URL: https://www.satscan.org/ [accessed 2024-05-05]
20. Lee LK, Fleegler EW, Farrell C, et al. Firearm laws and firearm homicides: a systematic review. JAMA Intern Med 2017 Jan 1;177(1):106-119. [doi: 10.1001/jamainternmed.2016.7051] [Medline: 27842178]
21. Scope of the national firearms regulations. United Nations Office on Drugs and Crime. 2019. URL: https://www.unodc.org/e4j/en/firearms/module-6/key-issues/scope-of-the-national-firearms-regulations.html [accessed 2024-07-20]
22. Global status report on road safety 2018. : World Health Organization; 2018 URL: https://www.who.int/publications-detail-redirect/9789241565684 [accessed 2024-07-20]
23. Política nacional de redução da morbimortalidade por acidentes e violências: portaria MS/GM n.o 737 de 16/5/01 [Report in Portugese]. : Ministério da Saúde; 2001 URL: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2001/prt0737_16_05_2001.html [accessed 2024-07-20]

XSL•FO
RenderX

24. Dispõe sobre registro, posse e comercialização de armas de fogo e munição, sobre o sistema nacional de armas – sinarm, define crimes e dá outras providências [Report in Portugese]. : Presidência da República; 2003 URL: https://www.planalto.gov.br/ccivil_03/leis/2003/l10.826.html [accessed 2024-07-20]
25. Atlas da violência 2023 [Report in Portugese]. : Instituto de Pesquisa Econômica Aplicada (Ipea); 2023 URL: https://repositorio.ipea.gov.br/handle/11058/12614 [accessed 2024-07-16]
26. Institui a política nacional de prevenção da automutilação e do suicídio, a ser implementada pela união, pelos estados, pelo distrito federal e pelos municípios; e altera a lei n [Web page in Portugese]. Presidência da República. 2019. URL: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/lei/l13819.htm [accessed 2024-07-16]
27. Altera o Código de Trânsito Brasileiro (Lei nº 9.503/1997) e dispõe sobre limites de álcool no sangue e penalidades para condutores [Web page in Portugese]. Presidência da República. 2008. URL: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/l11705.htm [accessed 2024-07-16]
28. Altera dispositivos do Código de Trânsito Brasileiro (Lei nº 9.503/1997) e dá outras providências [Web page in Portugese]. Presidência da República. URL: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12971.htm [accessed 2024-07-16]
29. Estabelece a mensagem, os temas e o cronograma das campanhas educativas de trânsito a serem realizadas de janeiro a dezembro de 2023 [Report in Portugese]. : Conselho Nacional de Trânsito (CONTRAN); 2022 URL: https://www.gov.br/transportes/pt-br/assuntos/transito/conteudo-contran/resolucoes/Resolucao9802022.pdf 2022 [accessed 2024-07-17]
30. Abreu DRDOM, Souza EMD, Mathias TADF. Impact of the Brazilian Traffic Code and the Law Against Drinking and Driving on mortality from motor vehicle accidents. Cad Saude Publica 2018 Aug 20;34(8):e00122117. [doi: 10.1590/0102-311X00122117] [Medline: 30133659]
31. Atlas da violência 2020 [Report in Portugese]. : Instituto de Pesquisa Econômica Aplicada (Ipea); 2020 URL: https://repositorio.ipea.gov.br/handle/11058/10214 [accessed 2024-07-17]
32. He JY, Xiao WX, Schwebel DC, et al. Road traffic injury mortality and morbidity by country development status, 2011-2017. Chin J Traumatol 2021 Mar;24(2):88-93. [doi: 10.1016/j.cjtee.2021.01.007] [Medline: 33526264]
33. Matzopoulos R, Prinsloo M, Bradshaw D, Abrahams N. Reducing homicide through policy interventions: The case of gun control. S Afr Med J 2019 Dec 5;109(11b):63-68. [doi: 10.7196/SAMJ.2019.v109i11b.14256] [Medline: 32252871]
34. Haviland MJ, Rowhani-Rahbar A, Rivara FP. Age, period and cohort effects in firearm homicide and suicide in the USA, 1983–2017. Inj Prev 2021 Aug;27(4):344-348. [doi: 10.1136/injuryprev-2020-043714]
35. Tranchitella FB, Santos RSD, El Bacha JJSH, Sobrado JV, Santos MBSD, Colombo Souza P. Mortality due to transport accidents in the city of São Paulo: 2005-2015. Acta Ortop Bras 2021;29(4):193-196. [doi: 10.1590/1413-785220212904240552] [Medline: 34566477]
36. Malta DC, Minayo MCDS, Cardoso LSDM, et al. Mortality among Brazilian adolescents and young adults between 1990 to 2019: an analysis of the Global Burden of Disease study. Cien Saude Colet 2021 Sep;26(9):4069-4086. [doi: 10.1590/1413-81232021269.12122021] [Medline: 34586261]
37. de Souza AMG, Costa KDS, de Morais TNB, de Andrade FB. Overview of the mortality from external causes of reproductive-age women in Brazil. Medicine (Abingdon) 2022;101(1):e28508. [doi: 10.1097/MD.0000000000028508]
38. Tavares MS, Alves A. Gender violence in Bolsonaro's (un)government: license to kill. Rev Bras Hist 2023;43:41-61. [doi: 10.1590/1806-93472023v43n94-04]
39. Ang GC, Low SL, How CH. Approach to falls among the elderly in the community. Singapore Med J 2020 Mar;61(3):116-121. [doi: 10.11622/smedj.2020029] [Medline: 32488276]
40. Koon W, Stewart O, Brander R, Quan L, Peden AE. Burden of fatal drowning in California, 2005-2019. Inj Prev 2023 Oct;29(5):371-377. [doi: 10.1136/ip-2023-044862] [Medline: 37208006]
41. Furtado BMASM, Lima ACBD, Ferreira RCG. Road traffic accidents involving elderly people: an integrative review. Rev bras geriatr gerontol 2019;22(3). [doi: 10.1590/1981-22562019022.190053]
42. Beghi M, Butera E, Cerri CG, et al. Suicidal behaviour in older age: A systematic review of risk factors associated to suicide attempts and completed suicides. Neurosci Biobehav Rev 2021 Aug;127:193-211. [doi: 10.1016/j.neubiorev.2021.04.011] [Medline: 33878336]
43. Global study on homicide 2023. United Nations Office ond Drugs and Crime. 2023. URL: https://WwwUnodcOrg/Unodc/Data-and-Analysis/Global-Study-on-HomicideHtml [accessed 2024-07-17]
44. Bittencourt MB, Teixeira AN. Social structure and dynamics of violence: social determinants of intentional homicides in Brazilian micro-regions. Rev Bras Estud Popul 2023. [doi: 10.20947/S0102-3098a0240]
45. Mortes por acidentes de transporte terrestre no brasil: análise dos sistemas de informação do ministério da saúde [Report in Portugese]. : Instituto de Pesquisa Econômica Aplicada (Ipea) URL: https://repositorio.ipea.gov.br/bitstream/11058/6869/1/TD_2212.pdf [accessed 2024-07-17]
46. Pesquisa CNT de Rodovias 2023. 2023. URL: https://pesquisarodovias.cnt.org.br/ [accessed 2024-07-18]
47. Avaliação das políticas públicas de transporte e segurança nas rodovias federais [Web page in Portugese]. Ministério dos Transportes, Portos e Aviação Civil. 2018. URL: https://www.gov.br/transportes/pt-br/pt-br/centrais-de-conteudo/apt-seguranca-rodovias-federais-pdf [accessed 2024-07-18]

XSL•FO
RenderX

48. Filho AMS, Merchan-Hamann E, Vasconcelos CH. Expansion, displacement and interiorization of homicides in Brazil, between 2000 and 2015: A spatial analysis. Ciencia e Saude Coletiva 2020;25:3097-3105. [doi: 10.1590/1413-81232020258.32612018]
49. Wanzinack C, Signorelli MC, Reis C. Homicides and socio-environmental determinants of health in Brazil: a systematic literature review. Cad Saude Publica 2018 Nov 29;34(12):e00012818. [doi: 10.1590/0102-311X00012818] [Medline: 30517311]
50. Filho AMS, Duarte EC, Merchan-Hamann E. Trend and distribution of the homicide mortality rate in accordance with the size of the population of Brazilian municipalities-2000 and 2015. Ciencia e Saude Coletiva 2020;25:1147-1156. [doi: 10.1590/1413-81232020253.19872018]
51. Cerqueira D, Bueno S. Atlas da violência 2019 [Report in Portugese]. : Instituto de Pesquisa Econômica Aplicada (Ipea); 2019 URL: https://repositorio.ipea.gov.br/handle/11058/9489 [accessed 2024-07-18]
52. Wanzinack C, Signorelli MC, Shimakura S, Pereira PPG, Polidoro M, de OL, et al. Indigenous homicide in Brazil: geospatial mapping and secondary data analysis (2010 to 2014). Cien Saude Colet 2010;24:2637-2648. [doi: 10.1590/1413-81232018247.23442017]
53. Bragato FF, Neto PB. Conflitos territoriais indígenas no Brasil: entre risco e prevenção / Indigenous land conflicts in Brazil: between risk and prevention. Rev Direito Práx 2017;8(1). [doi: 10.12957/dep.2017.21350]
54. Plano Nacional de Políticas Sobre Drogas (PLANAD) 2022–2027 [Report in Portugese]. : Conselho Nacional de Políticas sobre Drogas; 2022 URL: https://www.gov.br/mj/pt-br/assuntos/sua-protecao/politicas-sobre-drogas/arquivo-manual-de-avaliacao-e-alienacao-de-bens/planad_set_2022.pdf [accessed 2024-07-18]
55. Governo Federal lança operação integrada de combate ao crime organizado [Web page in Portugese]. Governo Federal. 2023. URL: https://agenciagov.ebc.com.br/noticias/202311/governo-federal-lanca-operacao-integrada-de-combate-ao-crime-organizado [accessed 2024-07-18]
56. Abreu DDO, Novaes ES, Oliveira RD, Mathias TDF, Marcon SS. Fall-related admission and mortality in older adults in Brazil: trend analysis. Cien Saude Colet 2018 Apr;23(4):1131-1141. [doi: 10.1590/1413-81232018234.09962016] [Medline: 29694573]

Abbreviations

AAPC: average annual percentage change
APC: annual percentage change
DC: death certificate
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*
SIM: Mortality Information System
WHO: World Health Organization

Edited by A Mavragani; submitted 25.10.24; peer-reviewed by AA Villa-Rueda, Y Sarikhani; revised version received 18.12.24; accepted 19.12.24; published 15.09.25.

Please cite as:
Andrade LA, Silva da Paz W, Santos de Melo LR, dos Santos Tavares D, Lopes de Sousa AF, Silva Camargo EL, Arena Ventura CA, Silva Souza RC, Gomes Machado de Araújo KC, Freire de Souza CD, dos Santos AD, Campos FAAC, Bezerra Santos M
Spatiotemporal Trends in Deaths From External Causes in Brazil: 23-Year Ecological and Population-Based Study
JMIR Public Health Surveill 2025;11:e68002
URL: https://publichealth.jmir.org/2025/1/e68002
doi:10.2196/68002

© Lucas Almeida Andrade, Wandklebson Silva da Paz, Luís Ricardo Santos de Melo, Débora dos Santos Tavares, Alvaro Francisco Lopes de Sousa, Emerson Lucas Silva Camargo, Carla Aparecida Arena Ventura, Regina Claudia Silva Souza, Karina Conceição Gomes Machado de Araújo, Carlos Dornels Freire de Souza, Allan Dantas dos Santos, Fagner Alfredo Ardisson Cirino Campos, Márcio Bezerra Santos. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 15.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Epidemiology and Treatment Outcomes of Pulmonary Tuberculosis in Dazu District, Chongqing, China, 2005-2024: Surveillance Study

Yu Yu[1*], BA; Xiu-juan Hu[2*], BA; Xiao-man Fang[1*], BA; Jing Wu[1*], MD

[1]Infectious Department, The People's Hospital of Dazu, No.1073, Erhuan South Road, Tangxiang Subdistrict, Dazu District, Chongqing, China

[2]Research Department, The People's Hospital of Dazu, Chongqing, China

[*]all authors contributed equally

Corresponding Author:
Jing Wu, MD
Infectious Department, The People's Hospital of Dazu, No.1073, Erhuan South Road, Tangxiang Subdistrict, Dazu District, Chongqing, China

Abstract

Background: As a high tuberculosis (TB) burden area in China, Dazu District of Chongqing Municipality contains a large rural population and exhibits typical features of TB endemicity.

Objective: This study aimed to analyze the epidemiological characteristics and treatment outcomes of pulmonary tuberculosis (PTB) in Dazu District from 2005 to 2024, with the aim of supporting the optimization of regional TB control strategies.

Methods: Data on PTB cases in Dazu District from 2005 to 2024 were collected from the China Disease Control and Prevention Information System. Descriptive epidemiological methods were employed to analyze the temporal, demographic, and geographical distributions, along with trends in treatment outcomes. Global and local spatial autocorrelation analyses were performed using Moran *I* and Getis-Ord Gi* statistics, respectively.

Results: A total of 10,236 cases were reported, for an average annual notification incidence of 65.2 per 100,000 population. The annual average notification incidence decline rate was 7.7%. Joinpoint regression analysis revealed a statistically significant decline in annual incidence rates (average annual percent change=−6.81, 95% CI −7.25 to −6.30, *P*<.0001). The bacteriological positivity rate initially decreased before rising, reaching 81.6% in 2024. Reported case counts peaked in March, while relatively lower numbers were observed during October, November, and December. Cases were predominantly among male patients, with a male-to-female ratio of 3.57:1. The case composition ratio in the ≥65 years age group has gradually increased, from 13.8% in 2006 to 19.9% in 2015 and to 38.5% in 2024. Occupational distribution was primarily among farmers (77.6%, 7948/10,236), homemakers or unemployed individuals (5.6%, 570/10,236), and students (3%, 303/10,236). Cases were concentrated in Longshui Town (13.1%, 1339/10,236), Tangxiang Subdistrict (9.1%, 933/10,236), and Longgang Subdistrict (7.9%, 811/10,236)—areas with large population bases. Among these, Guoliang Town exhibited the highest average annual notification incidence (314.4/100,000). Treatment success rate reached 91.3%. Multivariate binary logistic regression revealed that age 25 - 44 years (OR 1.755; 95% CI 1.320 - 2.332; *P*<.0001), undergoing initial treatment (OR 3.786; 95% CI 2.524 - 5.680; *P*<.0001), absence of HIV coinfection (OR 2.499; 95% CI 1.714 - 3.643; *P*<.0001), negative bacteriologic test results (OR 2.841; 95% CI 2.214 - 3.646; *P*<.0001), and the receipt of full-course supervised treatment (OR 7.705; 95% CI 4.520 - 13.137; *P*<.0001) were significantly associated with treatment success.

Conclusions: The notification incidence of PTB in Dazu District, Chongqing, has gradually declined. Particular focus is required on the treatment of young children, elderly individuals, patients with HIV coinfection, those under intensive phase supervision, bacteriologically positive cases, and retreatment cases. These measures may reduce the incidence of PTB and improve treatment success rates in our district.

(JMIR Public Health Surveill 2025;11:e78564) doi:10.2196/78564

KEYWORDS

China; Chongqing; Dazu District; epidemiology; pulmonary tuberculosis; treatment outcome; tuberculosis

Introduction

Pulmonary tuberculosis (PTB), caused by *Mycobacterium tuberculosis* (*MTB*), is a chronic infection disease primarily affecting the lungs and represents the most common form of TB. As an airborne pathogen, *MTB* poses a serious global public health threat [1]. Globally, TB remains the leading cause of death from a single infectious agent [2]. According to the World Health Organization (WHO) Global Tuberculosis Report 2024 [3], there were approximately 10.8 million new TB cases and 1.25 million deaths in 2023. In response, the WHO and the

XSL•FO
RenderX

United Nations have set a goal to reduce TB deaths by 95% and new cases by 90% by 2035. China bears one of the highest TB burdens globally, reporting 741,000 new cases in 2023—an estimated incidence of 52 per 100,000 population. It ranks third among the 30 high-TB burden countries, accounting for 6.8% of the global cases. An estimated 25,000 TB-related deaths occurred, corresponding to a mortality rate of 2 per 100,000. The treatment success rate for drug-sensitive TB is approximately 88%. These figures indicate that China faces significant challenges in meeting the WHO's goal to end the TB epidemic by 2035.

The prevalence of TB exhibits notable spatial heterogeneity and demographic disparities. In China, regions with higher TB burdens are often associated with factors such as economic development, health care accessibility, and population mobility. Previous studies have indicated higher incidence rates in western regions and rural areas, with farmers, older adults, and male individuals being disproportionately affected [4-7]. As a municipality in western China with both extensive urban and rural populations, Chongqing plays a critical role in regional TB control efforts. However, compared with national or provincial macro-level data, long-term dynamic monitoring studies at the district and county level remain relatively scarce. Although several studies at the urban and county scale in Chongqing have been published [8-11], they are often limited by short time spans, small sample sizes, and insufficient depth.

District- and county-level dynamic surveillance of TB is critical for elucidating local transmission dynamics and evaluating the effectiveness of disease control interventions. In this study, we leverage 2 decades (2005 - 2024) of surveillance data from Dazu District, Chongqing, to track temporal trends with Joinpoint regression, map spatial hotspots through geospatial techniques, and identify key epidemiological and treatment-related factors. This study aims to provide a scientific basis for optimizing local TB control strategies and advancing the goal of TB elimination by elucidating the epidemiological characteristics and temporal dynamics of PTB at the district and county levels.

Methods

Data Sources

Patient data of PTB cases from Dazu District, Chongqing, with confirmed diagnosis dates between January 1, 2005, and December 31, 2024 were collected from the Tuberculosis Management Information System (TBIMS) subsystem of the China Information System for Disease Control and Prevention (retrieved 2025-03-28). Data fields included current address, gender, age, registration date, bacteriological results, population category, key populations, symptom onset date, initial consultation date, complications, diagnosis confirmation date, treatment initiation date, and reasons for treatment termination. The exclusion criteria comprised duplicate reports, errors in basic infectious disease report card information, missing essential data, diagnostic modifications, cases initially managed outside Dazu District, extrapulmonary TB, untreated cases, and nontuberculous mycobacterial infections.

Case management data were retrieved from the TBIMS, with initial registration management units being the Dazu District Tuberculosis Prevention and Control Center or Dazu District People's Hospital of Chongqing (designated medical institution). The TBIMS was launched in 2005 using paper-based reporting and manual entry and was upgraded to a web-based real-time electronic system in 2010, greatly reducing latency. It covers all TB control institutions across 31 provinces, over 300 prefectures, and 3000 counties in mainland China, mandating reporting for all bacteriologically confirmed or clinically diagnosed TB cases. Now the world's largest TB registration system, the TBIMS, is recommended by the WHO as a model for global TB surveillance.

This study analyzed incidence characteristics (temporal, geographical, and demographic distributions) and investigated influencing factors of treatment outcomes. Annual notification rates used end-of-year permanent-resident denominators from the *Dazu District Statistical Yearbooks* published by Dazu District, Chongqing, during 2005-2024 (retrieved 2025-03-31). Age-standardized incidence rates were calculated using the 2020 Seventh National Population Census data (National Bureau of Statistics) as the standard population. The names of the 27 townships or streets involved in this study are all standardized administrative division units, which do not reflect the differences in special geographical or social attributes. Its division is based on the "Chongqing Urban and County Administrative Division Code" (GB/T 2260), which only reflects the administrative management level.

Research Methods

Diagnosis and Treatment of PTB

Patients were diagnosed according to the 《*WS 288 - 2008 Pulmonary Tuberculosis Diagnostic Criteria*》, 《*WS 288 - 2017 Pulmonary Tuberculosis Diagnosis*》 [12,13], and 《*WS 196 - 2017 Tuberculosis Classification*》 [14] standards. All TB treatments were implemented in accordance with the 《*Technical Guidelines for Tuberculosis Prevention and Control in China (2021 Edition)*》 [15]. To maintain data comparability across the study period, TB cases in this research included tuberculous pleuritis cases throughout all calendar years.

Definitions

Definitions are as follows: (1) case detection methods: referral, active screening, direct consultation or referral, contact tracing, health examination, and others (cases with undetermined or unclassifiable sources). (2) Discovery delay: the total time interval between a patient's initial clinical symptom and confirmed TB diagnosis, encompassing both care-seeking and diagnostic confirmation delays. This study employed a 28-day threshold as the criterion for determining discovery delay. (3) Bacteriological results: from 2005 to 2018, smear-positive or culture-positive cases were classified as bacteriologically positive, while cases with negative results in both smear microscopy and culture were classified as bacteriologically negative. From 2019 to 2024, cases with any positive result in smear microscopy, culture, or molecular biological testing were classified as bacteriologically positive, whereas cases with negative results in all 3 methods were classified as

XSL•FO
RenderX

bacteriologically negative. Patients with tuberculous pleurisy were uniformly considered bacteriologically negative. (4) Number of patients with PTB: it refers to cases registered and reported between 2005 and 2024 with diagnostic confirmation of TB (including clinical diagnoses) and including patients with tuberculous pleurisy. (5) Treatment outcome determination: according to standards established in the 《*Technical Specifications for Tuberculosis Prevention and Control in China (2020 Edition*)》 [16], treatment outcomes are categorized into either treatment success (including cure and treatment completion) or unfavorable outcome (including mortality, loss to follow-up, treatment discontinuation, and revision of therapy to a multidrug-resistant TB regimen).

Observation Indicators

Observation indicators are as follows: (1) notification incidence rate: it refers to the proportion of patients with registered and reported PTB to the total population within a specified period. (2) Age-standardized incidence rate: it is an incidence rate calculated by weighting age-specific rates of a study population to a standard population structure, enabling comparisons unaffected by differences in age distribution. (3) Calculation formula for average annual decline rate of PTB notification incidence rate: average annual decline rate=(1 - Pn/P0n)×100%, where P_n represents the registered incidence rate in the nth year, P denotes the registered incidence rate in the base year, and n indicates the number of interval years. (4) Successful treatment rate = (Cured+Treatment completion)/(Number of pulmonary tuberculosis patients–Transferred to rifampicin-resistant treatment–Diagnosis changed)×100%.

Ethical Considerations

This study was submitted to the Ethics Commission of the People's Hospital of Dazu District. This study does not involve human participants or animals. Ethics approval was not required for this study because we did not include any identifiable data of patients' personal information, including name, identity information, address, and telephone number. This study reviewed only secondary aggregated data at the population level; therefore, the need for written informed consent was waived. The Ethics Committee of the People's Hospital of Dazu District approved this secondary analysis of deidentified TBIMS data (Approval 2025LLSC158; 2025-03-12) and granted a waiver of informed consent.

Statistical Analysis

Descriptive Analysis

Data were organized, and a database was established using Microsoft Excel 2021. Statistical analysis was performed using SPSS (version 27.0; IBM SPSS Statistics,) and Joinpoint (version 5.2.0; National Cancer Institute of America) software. Descriptive epidemiologic methods were employed. Count data were described using "number of cases, incidence rate (/100,000), and composition ratio (%)." Intergroup comparisons were analyzed using the χ^2 test.

Joinpoint Regression

The temporal trend change points in age-standardized incidence rates were identified using the Joinpoint Regression Program (version 5.40; National Cancer Institute of America). This method fits the natural logarithm of the incidence rate as a piecewise linear function. The optimal number of joinpoints was determined and adopted using a Monte Carlo permutation test (10,000 iterations, α=.05). The independent variable was the calendar year, the interval type was set to annual, and the data point offset was 0. The dependent variables were the annual notification incidence rate and the age-standardized notification incidence rate, with data points where the incidence rate was 0 being excluded. The annual percent change (APC) along with its 95% CI was calculated for each trend segment, and the average annual percent change (AAPC) for the entire study period was also computed. The overall significance level for the permutation test was set at 0.0500. The maximum number of joinpoints was constrained to 3 based on the 19-year observation period, and model selection was optimized using the Bayesian information criterion.

Spatial Analysis

Global and local spatial autocorrelation analyses were performed using Moran *I* and Getis-Ord Gi* statistics, respectively. The Moran *I* index ranges from −1 to 1, where values approaching +1 indicate positive spatial clustering (similar values cluster together), values near 0 suggest a random distribution, and values approaching −1 indicate dispersion. Statistically significant Giclusters were classified as hotspots (high-value clusters; P<.05, $|Z|$>1.96) or coldspots (low-value clusters; P<.05, $|Z|$<1.96).

Logistic Regression

A binary multivariate logistic regression analysis model was used to analyze the influencing factors of treatment outcomes; a P<.05 was considered statistically significant.

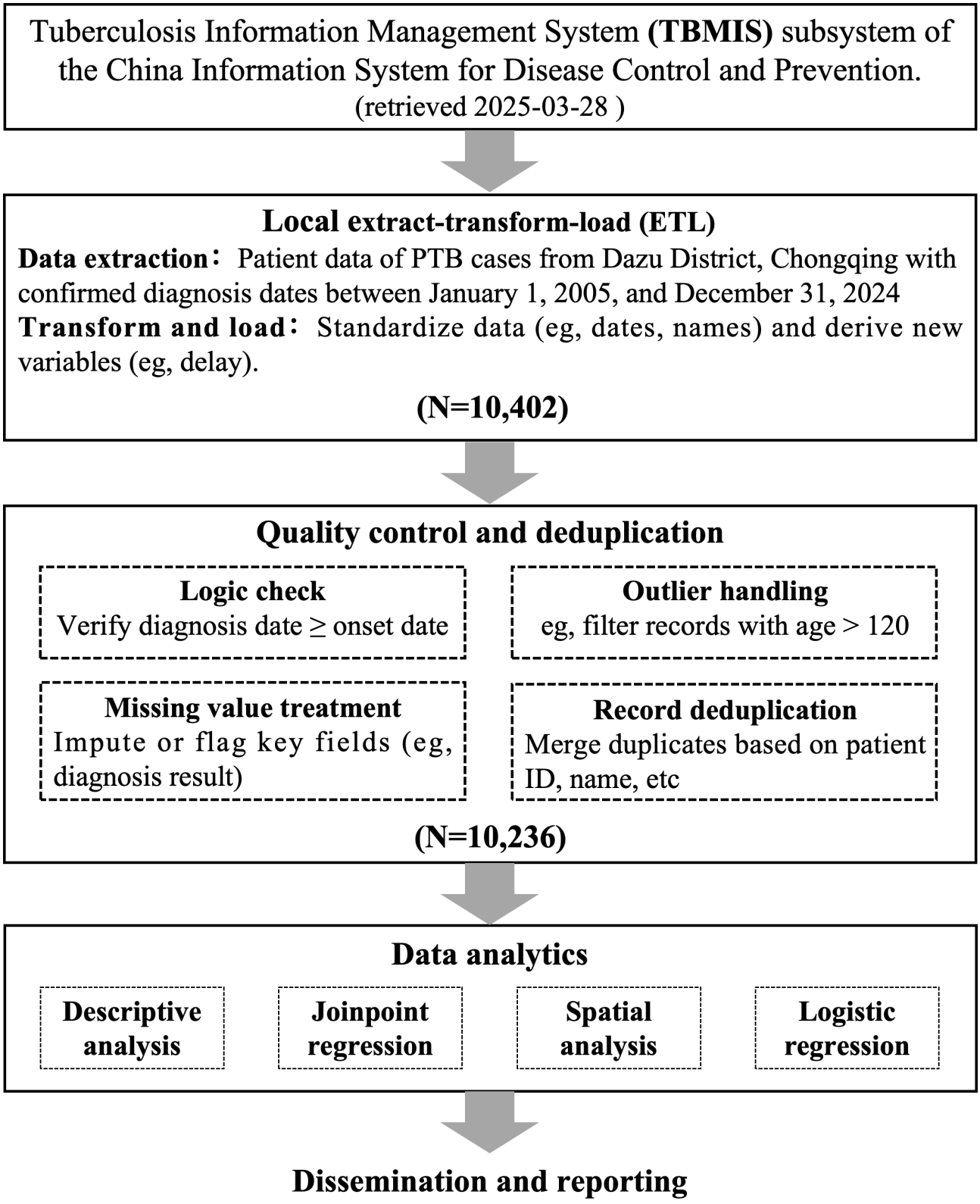
Results

Reported Incident Case Count, Notification Incidence Rates, and Bacteriological Positivity Rates

All regions share the same case reporting system and diagnostic criteria (Figure 1 and Checklist 1). A total of 10,236 cases were reported. The highest reported case count was in 2005 with 1000 cases, followed by a gradual downward trend, reaching a nadir of 234 cases in 2024. The notification incidence similarly decreased from 131.2/100,000 in 2005 to 28.4/100,000 in 2024, with an average annual notification incidence of 65.2/100,000. The average annual decline rate of reported TB incidence was 7.7%. The bacteriological positivity rate gradually declined from 46.9% in 2005 to 31.6% in 2017 and then subsequently increased to 81.6% by 2024. Monthly analysis revealed that March had the highest cumulative reported case count (1098 cases), while October, November, and December showed relatively lower counts, with 720, 750, and 748 cases, respectively (Tables 1 and 2).

XSL•FO
RenderX

Figure 1. Data flow diagram. PTB: pulmonary tuberculosis.



XSL•FO
RenderX

Table . Reported incident case count, annual and age-standardized incidence rates, and bacteriological positivity rates of pulmonary tuberculosis in Dazu District, Chongqing, China, 2005 - 2024.

Year	Permanent resident population (100,000 persons)	Reported cases (n)		Bacteriological positivity (%)	Annual notification incidence rate (per 100,000 population)	Age-standardized incidence rate (per 100,000 population)		
		Reported case count	Bacteriologically positive case count			<15	15 - 64	≥65
2005	76.2	1000	469	46.9	131.2	2.6	15.8	13.4
2006	76.0	701	294	41.9	92.2	1.0	11.1	10.8
2007	76.0	810	373	46.0	106.6	0.9	12.3	15.3
2008	76.7	747	365	48.9	97.4	0.6	11.5	12.8
2009	77.0	663	356	53.7	86.1	0.4	10.1	11.8
2010	67.1	630	346	54.9	93.9	0.0	11.1	13.2
2011	73.6	508	156	30.7	69	0.3	8.2	9.3
2012	74.9	594	73	12.3	79.3	0.4	9.1	12.0
2013	76.1	587	134	22.9	77.1	0.2	9.0	11.0
2014	77.2	554	103	18.6	71.8	0.1	8.6	9.3
2015	78.4	534	118	22.1	68.1	0.1	7.9	10.0
2016	80.0	508	124	24.4	63.5	0.2	7.3	9.6
2017	80.5	351	111	31.6	43.6	0.1	5.1	6.3
2018	81.9	357	151	42.3	43.6	0.1	4.8	7.7
2019	82.7	347	175	57.8	42	0.2	4.6	7.4
2020	83.6	285	149	52.3	34.1	0.1	3.7	6.5
2021	83.6	305	177	58.0	36.5	0.1	3.6	8.7
2022	83.4	255	168	65.9	30.6	0.1	2.8	8.4
2023	82.1	266	177	66.5	32.4	0.1	2.9	8.9
2024	82.3	234	191	81.6	28.4	0.0	2.6	8.1

XSL•FO
RenderX

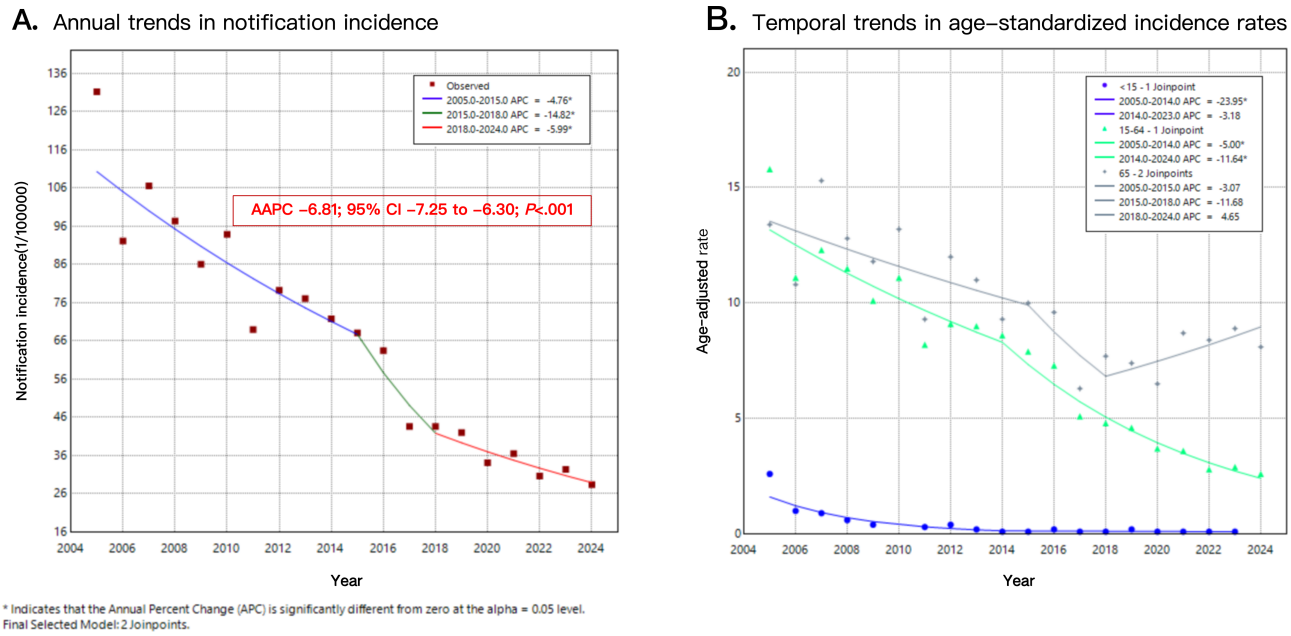
Table . Monthly cumulative reported tuberculosis (TB) cases in Dazu District, Chongqing, China, 2005-2024.

Year	Month (n)												Annual total cases (n)
	January	February	March	April	May	June	July	August	September	October	November	December	
2005	70	65	141	111	73	91	63	84	80	54	84	84	1000
2006	84	93	85	75	42	51	42	51	46	42	46	44	701
2007	61	50	81	76	77	81	74	75	59	64	61	51	810
2008	48	68	87	59	67	56	46	80	74	56	42	64	747
2009	30	66	80	64	68	74	60	36	47	40	43	55	663
2010	61	41	69	48	44	67	67	45	55	42	50	41	630
2011	38	39	40	43	44	44	49	44	54	42	41	30	508
2012	40	63	76	54	56	59	48	48	42	38	45	25	594
2013	43	41	74	58	49	45	50	60	57	31	46	33	587
2014	42	53	47	55	49	54	49	41	63	40	30	31	554
2015	35	39	45	39	37	52	48	50	30	53	49	57	534
2016	30	43	63	46	43	52	37	35	42	46	36	35	508
2017	34	55	42	24	28	24	19	25	20	30	24	26	351
2018	31	26	26	31	42	28	35	30	33	27	10	38	357
2019	27	30	38	27	26	32	33	7	61	8	32	26	347
2020	5	0	6	68	39	19	19	26	30	21	25	27	285
2021	26	14	31	22	28	28	24	27	34	22	14	35	305
2022	29	14	22	22	26	18	22	33	18	19	22	10	255
2023	19	26	19	24	24	19	27	31	20	24	21	12	266
2024	14	14	26	11	16	20	15	18	26	21	29	24	234
Monthly total cases (n)	767	840	1098	957	878	914	827	846	891	720	750	748	10,236

The Joinpoint regression analysis revealed that the annual notification incidence of PTB in Dazu District exhibited a significant overall downward trend (AAPC=−6.81; 95% CI −7.25 to −6.30; *P*<.001). This overall trend was characterized by a fluctuating decline across 3 distinct phases: a slight decrease from 2005 to 2015 (APC=−4.76; 95% CI −5.58 to −3.63; *P*=.002), an accelerated decline from 2015 to 2018 (APC=−14.82; 95% CI −16.96 to −10.04; *P*=.002), followed by a moderated rate of decline from 2018 to 2024 (APC=−5.99; 95% CI −7.83 to −0.38; *P*=.04; Figure 2A).

XSL•FO
RenderX

Figure 2. Trend chart of annual notification pulmonary tuberculosis (PTB) incidence in Dazu District, Chongqing, China, 2005-2024: Joinpoint regression.



The trends in age-standardized notification incidence of PTB in Dazu District varied across different age groups. The <15 years age group experienced a significant initial decline (2005 - 2014: APC=−23.95; 95% CI −37.04 to −18.47; *P*<.001), which then plateaued (2014 - 2023: APC=−3.18; 95% CI −9.89 to 13.39; *P*=.48). The 15 - 64 years age group showed the largest overall decrease, with an accelerating downward trend (2005 - 2014: APC=−5.00; 95% CI −6.51 to −2.73; *P*=.004; 2014 - 2024: APC=−11.64; 95% CI −13.47 to −10.36; *P*<.001). In contrast, the ≥65 years age group, after a period of fluctuating decline (2005 - 2015: APC=−3.07; 95% CI −5.32 to −10.80; *P*=.14; 2015 - 2018: APC=−11.68; 95% CI −17.20 to 0.15; *P*=.05), transitioned to a slow increase (2018 - 2024: APC=4.65; 95% CI −1.78 to 18.40; *P*=.08; Figure 2B).

Demographic Distribution

Reported patients comprised 7998 cases among male patients and 2238 cases among female patients, yielding a male-to-female ratio of 3.57:1. When categorized in 5-year intervals, male and female cases predominantly occurred in the 15 - 64 age group, with relatively uniform case distribution across subintervals within this range (Figure 3). Throughout the observation period and across all populations, children <15 years accounted for 102 cases (0.9%), young adults aged 15 - 44 years comprised 4028 cases (39.4%), middle-aged adults 45 - 64 years constituted 4000 cases (39.1%), and older patients ≥65 years represented 2106 cases (20.6%). The proportion of patients ≥65 years showed a gradually increasing trend, while the incidence proportion among young adults 15 - 44 years progressively declined. The proportion among middle-aged adults 45 - 64 years remained stable (Table 3, Figure 3).

XSL•FO
RenderX

Figure 3. Reported case counts and proportions of pulmonary tuberculosis (PTB) across age groups by year in Dazu District, Chongqing, China, 2005-2024.

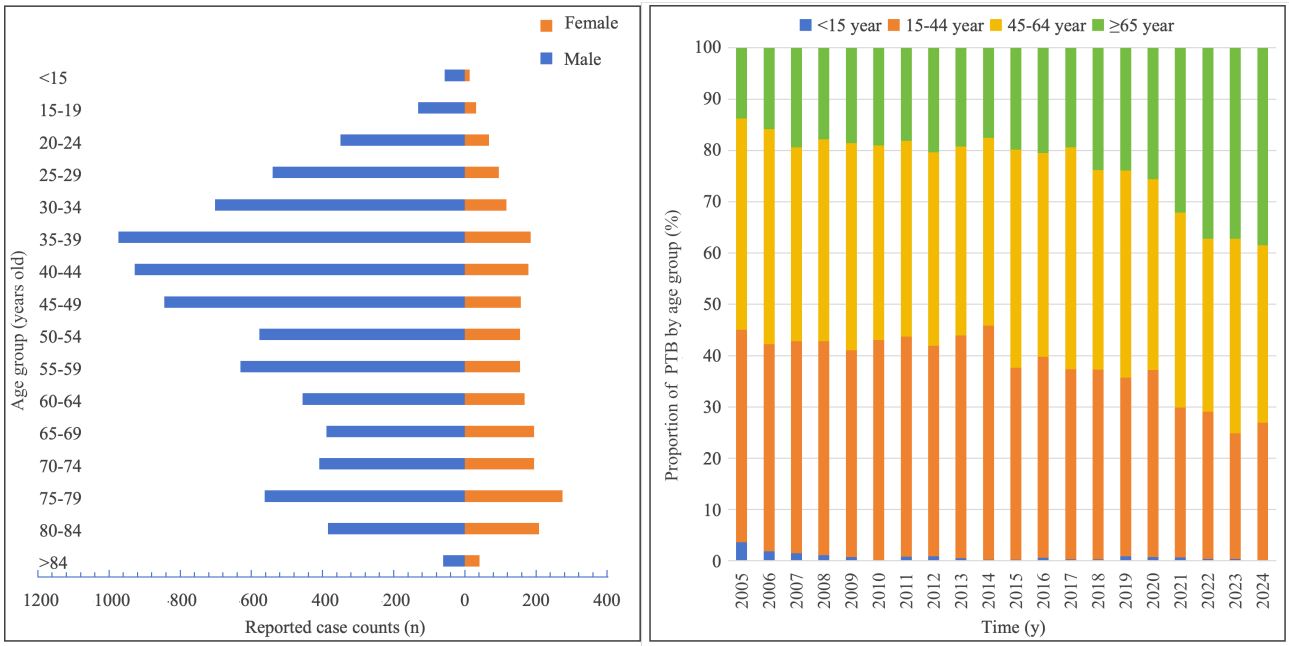


Table . Reported case counts and percentages of pulmonary tuberculosis (PTB) across age groups in Dazu District, Chongqing, 2005 to 2024.

Year	Age group (y), n (%)			
	<15	15 - 44	45 - 64	≥65
2005	36 (3.6)	414 (41.4)	412 (41.2)	138 (13.8)
2006	13 (1.8)	283 (40.4)	294 (41.9)	111 (15.8)
2007	12 (1.5)	335 (41.4)	306 (37.8)	157 (19.4)
2008	8 (1.1)	312 (41.8)	294 (39.4)	133 (17.8)
2009	5 (0.8)	267 (40.3)	268 (40.4)	123 (18.6)
2010	0 (0)	271 (43)	239 (37.9)	120 (19.1)
2011	4 (0.8)	218 (42.9)	194 (38.2)	92 (18.1)
2012	5 (0.8)	244 (41.1)	224 (37.7)	121 (20.4)
2013	3 (0.5)	255 (43.4)	216 (36.8)	113 (19.3)
2014	1 (0.2)	253 (45.7)	203 (36.6)	97 (17.5)
2015	1 (0.2)	200 (37.5)	227 (42.5)	106 (19.9)
2016	3 (0.6)	199 (39.2)	202 (39.8)	104 (20.5)
2017	1 (0.3)	130 (37)	152 (43.3)	68 (19.4)
2018	1 (0.3)	132 (37)	139 (38.9)	85 (23.8)
2019	3 (0.9)	121 (34.9)	140 (40.3)	83 (23.9)
2020	2 (0.7)	104 (36.5)	106 (37.2)	73 (25.6)
2021	2 (0.7)	89 (29.2)	116 (38.0)	98 (32.1)
2022	1 (0.4)	73 (28.6)	86 (33.7)	95 (37.3)
2023	1 (0.4)	65 (24.4)	101 (38)	99 (37.2)
2024	0 (0)	63 (26.9)	81 (34.6)	90 (38.5)

The top 3 occupational groups were farmers (77.6%, 7948/10,236), homemakers or unemployed individuals (5.6%, 570/10,236), and students (5.2%, 532/10,236), accounting for 88.4% (9050/10,236) of the total reported cases. These were followed by workers and migrant workers (3.0%, 303/10,236), commercial service personnel (1.1%, 116/10,236), and government employees (0.9%, 93/10,236). The remaining

XSL•FO
RenderX

occupations accounted for less than 7% collectively (Multimedia Appendix 1).

Geographical Distribution

All towns or subdistricts in Dazu District except for Longtanzi Subdistrict reported cases annually. The highest reported case count was in Longshui Town (13.1%, 1339/10,236), followed by Longgang Subdistrict (9.1%, 933/10,236) and Tangxiang Subdistrict (7.9%, 811/10,236). The lowest case count occurred in Longtanzi Subdistrict (0.7%, 72/10,236).

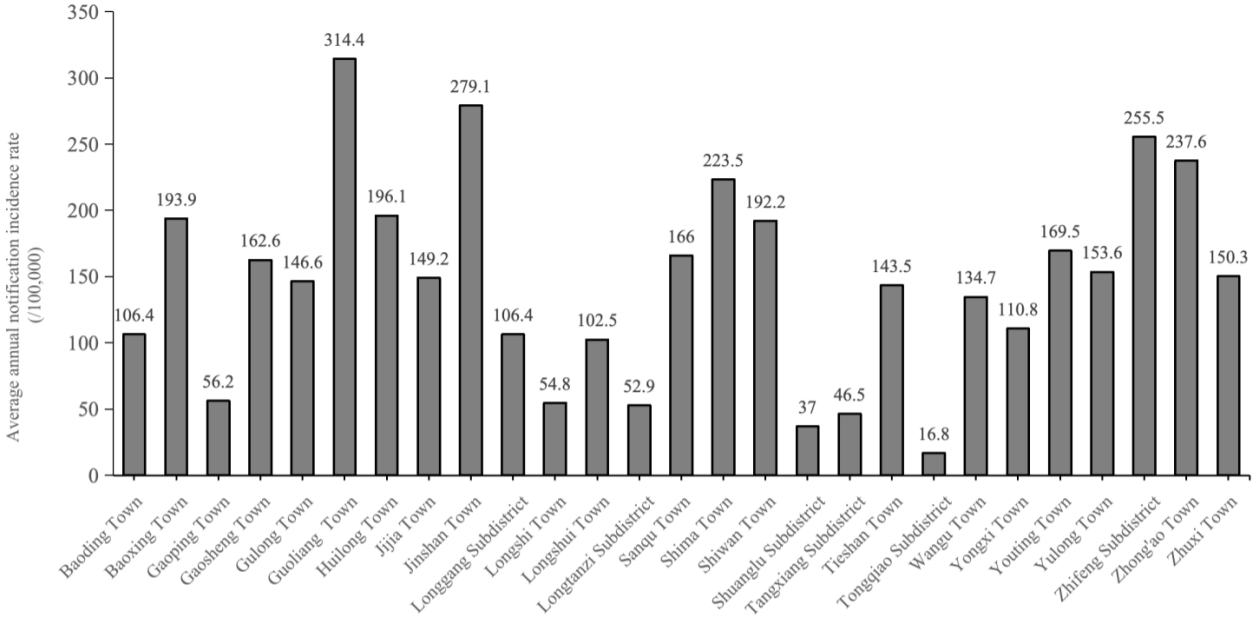
Due to the difficulty in obtaining accurate annual resident population data for each town and street, the proportion of reported cases in each town relative to the district's total annual cases was used as a substitute for the annual incidence rate (Table 4 and Figure 4). Longshui Town had the highest proportion of the reported cases in most years, exceeding 10% (47/234) in all years except 2011, although its case numbers showed a clear downward trend; it was followed by Longgang Street. In Tangxiang Street, the annual case count declined from 2005 to 2017 and then gradually fluctuated upward. By 2024, it had the highest number of reported cases in the district, accounting for 20.1% (Table 5, Figures 5 and 6).

Table . Annual pulmonary tuberculosis (PTB) average annual notification incidence rates by town or subdistrict in Dazu District.

Region	Average annual notification incidence rate (/100,000)	Resident population	Total case count (n)	Average annual case count (/100,000)
Baoding Town	106.4	15,037	274	13.7
Baoxing Town	193.9	17,534	375	18.7
Gaoping Town	56.2	12,451	140	7.0
Gaosheng Town	162.6	13,534	171	8.55
Gulong Town	146.6	4774	107	5.35
Guoliang Town	314.4	10,178	209	10.4
Huilong Town	196.1	13,260	243	12.1
Jijia Town	149.2	11,392	169	8.4
Jinshan Town	279.1	12,900	268	13.4
Longgang Subdistrict	106.4	88,330	933	46.6
Longshi Town	54.8	10,952	130	6.5
Longshui Town	102.5	1,12,170	1339	66.9
Longtanzi Subdistrict	52.9	17,023	72	3.6
Sanqu Town	166.0	30,722	495	24.7
Shima Town	223.5	17,452	370	18.5
Shiwan Town	192.2	14,565	290	14.5
Shuanglu Subdistrict	37.0	40,575	250	12.5
Tangxiang Subdistrict	46.5	1,50,488	811	40.5
Tieshan Town	143.5	16,721	227	11.3
Tongqiao Subdistrict	16.8	29,758	113	5.6
Wangu Town	134.7	37,108	516	25.8
Yongxi Town	110.8	17,154	270	13.5
Youting Town	169.5	32,443	543	27.15
Yulong Town	153.6	14,978	251	12.5
Zhifeng Subdistrict	255.5	27,786	462	23.1
Zhong'ao Town	237.6	35,360	635	31.7
Zhuxi Town	150.3	29,947	573	28.6

XSL•FO
RenderX

Figure 4. Average annual notification incidence rate of pulmonary tuberculosis (PTB), Dazu District, Chongqing, China, 2005-2024.



XSL•FO
RenderX

Table . Annual pulmonary tuberculosis (PTB) incident case counts by town or subdistrict in Dazu District, Chongqing, China, 2005 - 2024.

	Year, n (%)																			
Region	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Baoding Town	16 (1.6)	16 (2.3)	18 (2.2)	23 (3.1)	25 (3.8)	19 (3.0)	24 (4.7)	22 (3.7)	21 (3.6)	11 (2.0)	13 (2.4)	14 (2.8)	4 (1.1)	9 (2.5)	9 (2.6)	7 (2.5)	5 (1.6)	5 (2.0)	5 (1.9)	8 (3.4)
Baoxing Town	34 (3.4)	34 (4.9)	40 (4.9)	22 (2.9)	34 (5.1)	21 (3.3)	16 (3.1)	18 (3.0)	20 (3.4)	20 (3.6)	19 (3.6)	9 (1.8)	13 (3.7)	8 (2.2)	17 (4.9)	11 (3.9)	14 (4.6)	13 (5.1)	10 (3.8)	2 (0.9)
Gaoping Town	7 (0.7)	5 (0.7)	6 (0.7)	6 (0.8)	7 (1.1)	8 (1.3)	10 (2.0)	6 (1.0)	8 (1.4)	14 (2.5)	8 (1.5)	11 (2.2)	8 (2.3)	9 (2.5)	6 (1.7)	3 (1.1)	3 (1.0)	6 (2.4)	7 (2.6)	2 (0.9)
Gaosheng Town	22 (2.2)	18 (2.6)	12 (1.5)	8 (1.1)	14 (2.1)	11 (1.7)	11 (2.2)	17 (2 9)	11 (1.9)	10 (1.8)	5 (0.9)	6 (1.6)	8 (2.3)	2 (0.6)	4 (1.2)	4 (1.4)	3 (1.0)	1 (0.4)	3 (1.1)	1 (0.4)
Gulong Town	7 (0.7)	8 (1.1)	18 (2.2)	11 (1.5)	8 (1.2)	11 (1.7)	4 (0.8)	9 (1.5)	6 (1.0)	4 (0.7)	2 (0.4)	2 (0.4)	3 (0.9)	2 (0.6)	3 (0.9)	1 (0.4)	1 (0.3)	1 (0.4)	4 (1.5)	2 (0.9)
Guoliang Town	32 (3.2)	12 (1.7)	18 (2.2)	11 (1.5)	15 (2.3)	15 (2.4)	9 (1.8)	17 (2.9)	13 (2.2)	7 (1.3)	17 (3.2)	12 (2.4)	6 (1.7)	3 (0.8)	3 (0.9)	3 (1.1)	5 (1.6)	5 (2.0)	3 (1.1)	3 (1.3)
Huilong Town	26 (2.6)	17 (2.4)	28 (3.5)	10 (1.3)	15 (2.3)	14 (2.2)	15 (3.0)	12 (2.0)	14 (2.4)	11 (2.0)	10 (1.9)	18 (3.5)	11 (3.1)	8 (2.2)	9 (2.6)	3 (1.1)	10 (3.3)	5 (2.0)	5 (1.9)	2 (0.9)
Jijia Town	17 (1.7)	13 (1.9)	15 (1.9)	11 (1.5)	12 (1.8)	11 (1.7)	9 (1.8)	13 (2.2)	5 (0.9)	5 (0.9)	9 (1.7)	8 (1.6)	5 (1.4)	5 (1.4)	7 (2.0)	5 (1.8)	9 (3.0)	1 (0.4)	3 (1.1)	6 (2.6)
Jinshan Town	36 (3.6)	19 (2.7)	18 (2.2)	23 (3.1)	20 (3.0)	13 (2.1)	14 (2.8)	14 (2.4)	14 (2.4)	20 (3.6)	17 (3.2)	14 (2.8)	6 (1.7)	7 (2.0)	10 (2.9)	7 (2.5)	8 (2.6)	2 (0.8)	5 (1.9)	1 (0.4)
Longgang Subdistrict	94 (9.4)	55 (7.8)	64 (7.9)	69 (9.2)	60 (9.0)	59 (9.4)	59 (11.6)	51 (8.6)	52 (8.9)	46 (8.3)	36 (6.7)	37 (7.3)	29 (8.3)	50 (14.0)	42 (12.1)	27 (9.5)	35 (11.5)	28 (11.0)	23 (8.6)	17 (7.3)
Longshi Town	6 (0.6)	11 (1.6)	8(1.0)	6 (0.8)	8 (1.2)	12 (1.9)	4 (0.8)	5 (0.8)	11 (1.9)	12 (2.2)	6 (1.1)	12 (2.4)	6 (1.7)	5 (1.4)	3 (0.9)	3 (1.1)	6 (2.0)	2 (0.8)	2 (0.8)	2 (0.9)
Longshui Town	115 (11.5)	83 (11.8)	112 (13.8)	88 (11.8)	89 (13.4)	95 (15.1)	47 (9.3)	75 (12.6)	85 (14.5)	73 (13.2)	69 (12.9)	69 (13.6)	61 (17.4)	50 (14.0)	36 (10.4)	32 (11.2)	40 (13.1)	46 (18.0)	39 (14.7)	35 (15.0)
Longtanzi Subdistrict	9 (0.9)	5 (0.7)	8 (1.0)	8 (1.1)	0	3 (0.5)	3 (0.6)	3 (0.5)	0	4 (0.7)	4 (0.7)	4 (0.8)	1 (0.3)	4 (1.1)	3 (0.9)	2 (0.7)	2 (0.7)	4 (1.6)	1 (0.4)	4 (1.7)
Sanqu Town	51 (5.1)	38 (5.4)	43 (5.3)	43 (5.8)	31 (4.7)	28 (4.4)	18 (3.5)	28 (4.7)	28 (4.8)	34 (6.1)	33 (6.2)	21 (4.1)	20 (5.7)	14 (3.9)	11 (3.2)	15 (5.3)	11 (3.6)	10 (3.9)	8 (3.0)	10 (4.3)
Shima Town	39 (3.9)	28 (4.0)	35 (4.3)	24 (3.2)	21 (3.2)	20 (3.2)	19 (3.7)	30 (5.1)	26 (4.4)	23 (4.2)	20 (3.7)	14 (2.8)	10 (2.8)	13 (3.6)	12 (3.5)	5 (1.8)	9 (3.0)	5 (2.0)	13 (4.9)	4 (1.7)
Shiwan Town	28 (2.8)	22 (3.1)	26 (3.2)	23 (3.1)	20 (3.0)	29 (4.6)	16 (3.1)	14 (2.4)	15 (2.6)	19 (3.4)	20 (3.7)	12 (2.4)	7 (2.0)	4 (1.1)	12 (3.5)	7 (2.5)	6 (2.0)	4 (1.6)	3 (1.1)	3 (1.3)

XSL•FO
RenderX

	Year, n (%)																			
Region	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
Shuanglu Subdistrict	15 (1.5)	8 (1.1)	11 (1.4)	10 (1.3)	10 (1.5)	12 (1.9)	7 (1.4)	10 (1.7)	17 (2.9)	25 (4.5)	19 (3.6)	11 (2.2)	6 (1.7)	13 (3.6)	13 (3.7)	18 (6.3)	17 (5.6)	7 (2.7)	8 (3.0)	13 (5.6)
Tangxiang Subdistrict	70 (7.0)	21 (3.0)	37 (4.6)	48 (6.4)	37 (5.6)	30 (4.8)	36 (7.1)	26 (4.4)	41 (7.0)	45 (8.1)	39 (7.3)	34 (6.7)	32 (9.1)	44 (12.3)	47 (13.5)	49 (17.2)	46 (15.1)	38 (14.9)	44 (16.5)	47 (20.1)
Tashan Town	24 (2.4)	20 (2.9)	12 (1.5)	16 (2.1)	14 (2.1)	10 (1.6)	14 (2.8)	19 (3.2)	16 (2.7)	13 (2.3)	9 (1.7)	17 (3.3)	6 (1.7)	5 (1.4)	3 (0.9)	8 (2.8)	4 (1.3)	5 (2.0)	6 (2.3)	6 (2.6)
Tongjia Subdistrict	5 (0.5)	4 (0.6)	7 (0.9)	6 (0.8)	6 (0.9)	6 (1.0)	5 (1.0)	13 (2.2)	4 (0.7)	5 (0.9)	3 (0.6)	5 (1.0)	4 (1.1)	8 (2.2)	6 (1.7)	8 (2.8)	7 (2.3)	4 (1.6)	2 (0.8)	5 (2.1)
Wangu Town	50 (5.0)	46 (6.6)	42 (5.2)	50 (6.7)	30 (4.5)	32 (5.1)	30 (5.9)	34 (5.7)	27 (4.6)	24 (4.3)	26 (4.9)	17 (3.3)	15 (4.3)	13 (3.6)	20 (5.8)	11 (3.9)	10 (3.3)	15 (5.9)	17 (6.4)	7 (3.0)
Yongxi Town	19 (1.9)	21 (3.0)	24 (3.0)	21 (2.8)	17 (2.6)	17 (2.7)	15 (3.0)	27 (4.5)	13 (2.2)	11 (2.0)	13 (2.4)	15 (3.0)	12 (3.4)	10 (2.8)	6 (1.7)	3 (1.1)	3 (1.0)	8 (3.1)	10 (3.8)	5 (2.1)
Youting Town	55 (5.5)	38 (5.4)	44 (5.4)	50 (6.7)	33 (5.0)	30 (4.8)	28 (5.5)	27 (4.5)	40 (6.8)	29 (5.2)	38 (7.1)	37 (7.3)	20 (5.7)	13 (3.6)	10 (2.9)	12 (4.2)	13 (4.3)	8 (3.1)	7 (2.6)	11 (4.7)
Yulong Town	23 (2.3)	14 (2.0)	15 (1.9)	21 (2.8)	24 (3.6)	16 (2.5)	18 (3.5)	17 (2.9)	14 (2.4)	12 (2.2)	16 (3.0)	18 (3.5)	4 (1.1)	8 (2.2)	8 (2.3)	7 (2.5)	3 (1.0)	2 (0.8)	6 (2.3)	5 (2.1)
Zhifeng Subdistrict	71 (7.1)	37 (5.3)	33 (4.1)	30 (4.0)	36 (5.4)	26 (4.1)	20 (3.9)	23 (3.9)	28 (4.8)	27 (4.9)	25 (4.7)	23 (4.5)	13 (3.7)	10 (2.8)	14 (4.0)	10 (3.5)	9 (3.0)	12 (4.7)	10 (3.8)	5 (2.1)
Zongbo Town	84 (8.4)	62 (8.8)	63 (7.8)	63 (8.4)	38 (5.7)	45 (7.1)	24 (4.7)	35 (5.9)	27 (4.6)	22 (4.0)	27 (5.1)	34 (6.7)	19 (5.4)	13 (3.6)	16 (4.6)	13 (4.6)	13 (4.3)	11 (4.3)	12 (4.5)	14 (6.0)
Zhuxi Town	45 (4.5)	46 (6.6)	53 (6.5)	46 (6.2)	39 (5.9)	37 (5.9)	33 (6.5)	29 (4.9)	31 (5.3)	28 (5.1)	31 (5.8)	34 (6.7)	22 (6.3)	27 (7.6)	17 (4.9)	11 (3.9)	13 (4.3)	7 (2.7)	10 (3.8)	14 (6.0)

XSL•FO
RenderX

Figure 5. Trends in reported tuberculosis (TB) case counts by town or subdistrict and year in Dazu District, Chongqing, China.

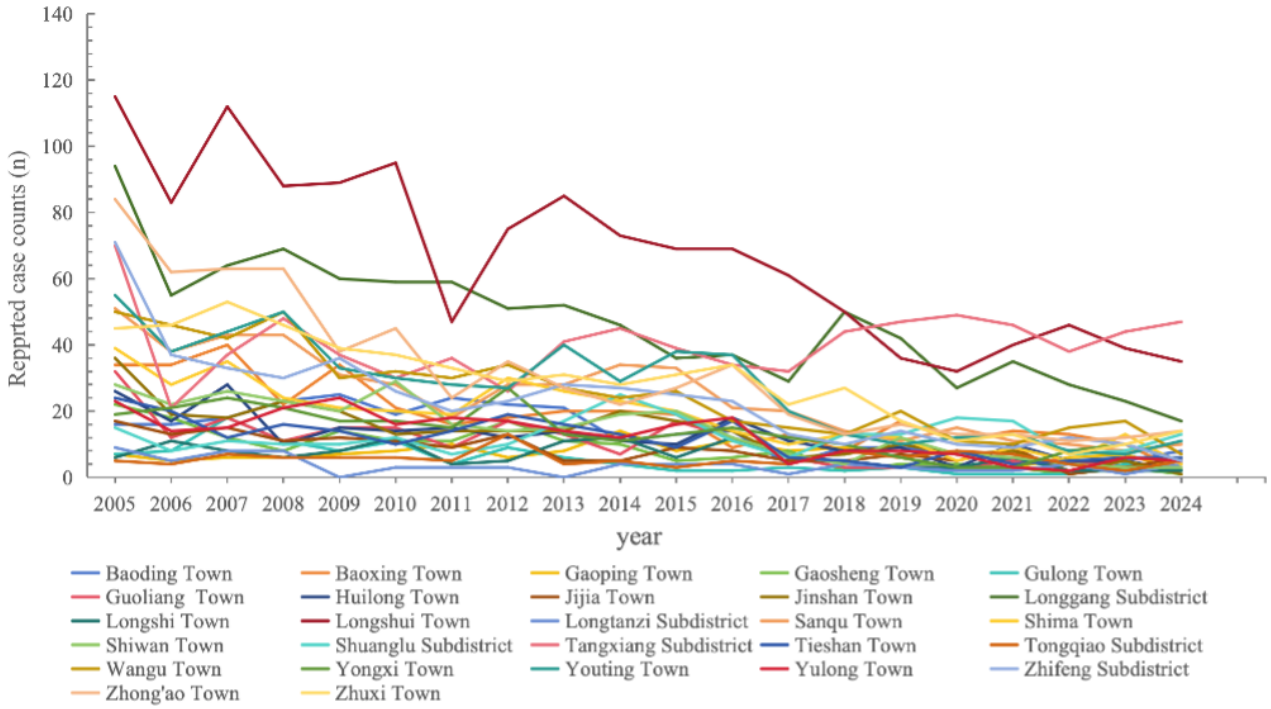
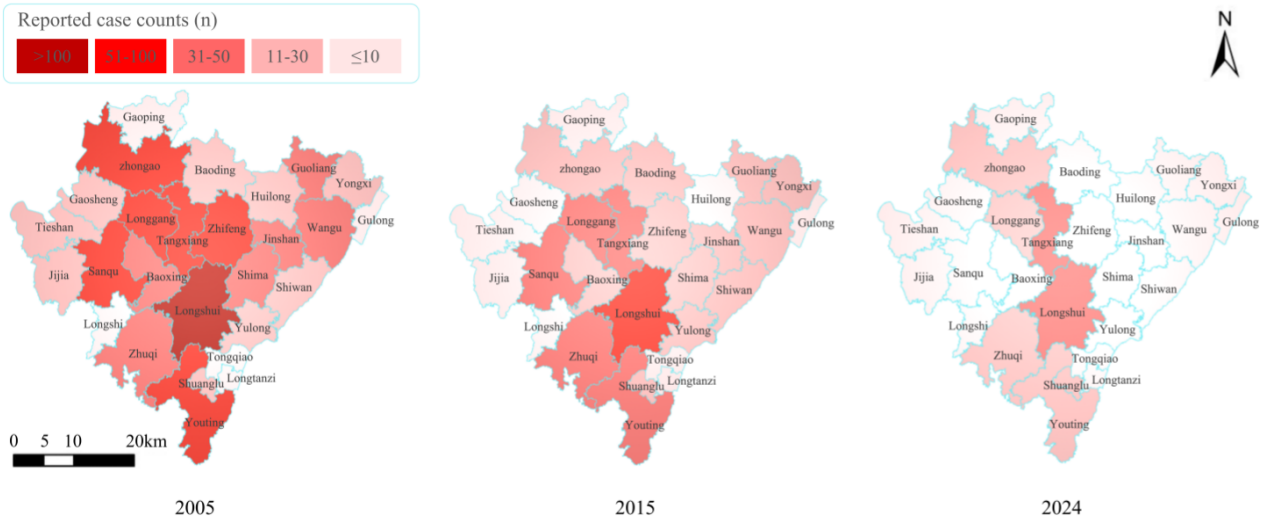


Figure 6. Spatiotemporal evolution of reported tuberculosis cases: Geographic heatmap across town/subdistrict in Dazu District, 2005 - 2024.



We selected annual reported cases from 2005, 2015, and the most recent year (2024) to create geographical heatmaps (Figure 6). As shown, in 2005, Dazu District exhibited widespread high-intensity clustering of cases (>50 cases), with 7 towns and streets along the central axis (including Longshui, Longgang, and Tangxiang) each reporting over 50 cases. By 2015, the TB burden had decreased significantly across the region. Only Longshui Town remained a high-intensity cluster, while the other 6 areas had declined to medium (31 - 50 cases) or medium-low intensity (11 - 30 cases). By 2024, except for Tangxiang Street and Longshui Town, which were classified as medium-intensity areas, all other towns and streets fell into medium-low or low-intensity categories, of which 20 are low-intensity areas (<10 cases), with a coverage rate of more than 74% (Figure 6).

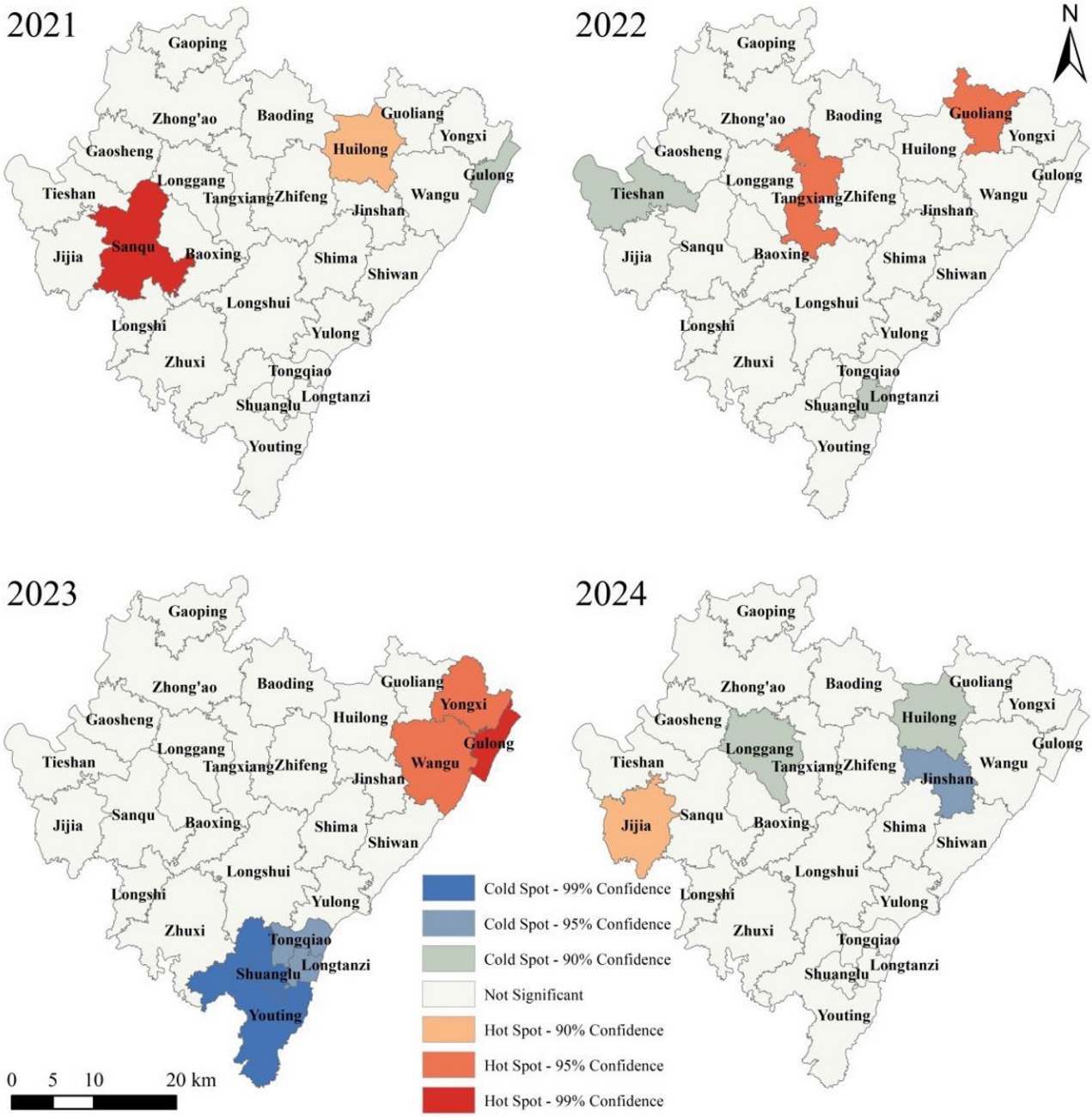
We calculated the notification incidence rate for each township and subdistrict in Dazu District from 2021 to 2024, using the permanent resident population data from the 2020 Seventh National Population Census of China as the denominator, and subsequently performed a spatial autocorrelation analysis. The global spatial autocorrelation analysis (Table 6) indicated that the spatial clustering of township-level PTB incidence was not statistically significant overall during the 2021 - 2024 period, with the exception of the year 2023, which exhibited significant spatial clustering rather than a random distribution (P<.001). Local spatial autocorrelation analysis (Figure 7) revealed indistinct spatial segregation of cold-spot or hot-spot areas for PTB incidence among the townships and subdistricts of Dazu District, with the identified cold-spot and hot-spot areas varying across different years.

XSL•FO
RenderX

Table . Global spatial autocorrelation analysis of pulmonary tuberculosis (PTB) incidence in Dazu District, Chongqing 2021 - 2024.

Year	Moran *I*	*Z* value	*P* value
2021	0.232	1.44	.15
2022	0.197	1.28	.20
2023	0.669	3.831	<.001
2024	0.017	0.292	.77

Figure 7. Local spatial autocorrelation analysis of pulmonary tuberculosis incidence in Dazu District, Chongqing, 2021 - 2024.



Univariate Analysis of Treatment Outcomes

Among 10,236 reported cases, excluding 159 cases still under treatment, a total of 10,077 cases were included in the treatment outcome analysis. There were 9203 successfully treated cases, yielding a treatment success rate of 91.3% (9203/10,077). Univariate analysis revealed that age (χ^2_3=78.265, *P*<.001), occupation (χ^2_3=10.228, *P*=.02), patient origin (χ^2_5=11.59, *P*=.04), treatment classification (initial treatment vs retreatment, χ^2_1=10.205, *P*=.001), diagnostic results (bacteriologically negative vs positive, χ^2_1=52.723, *P*<.001), HIV coinfection status (χ^2_1=14.724, *P*<.001), and treatment management modality (intensive phase supervision vs full-course supervision,

XSL•FO
RenderX

χ^2_2=430.013, *P*<.001) were risk factors influencing treatment outcomes (Table 7).

XSL•FO
RenderX

Table . Univariate analysis of treatment outcomes among patients with pulmonary tuberculosis (PTB) in Dazu District, Chongqing, China, 2005 - 2024.

Variable	Outcome		Chi-square (*df*)	*P* value
	Treatment success (n)	Unfavorable outcome (n)		
Gender			0.098 (1)	.75
Male	7186	687		
Female	2017	187		
Age (y)			78.265 (3)	<.001
<25	1414	110		
25 - 44	2375	183		
45 - 64	3642	303		
≥65	1772	278		
Occupation			10.228 (3)	.02
Farmers	7135	701		
Homemakers and unemployed	526	31		
Students	493	36		
Others	1049	106		
Case source			11.59 (5)	.04
Direct hospital visits	3291	301		
Referral	3408	346		
Tracking	1417	108		
Recommendation	937	97		
Health checkup	112	15		
Active screening	38	7		
Diagnostic delay			1.199 (2)	.55
≤28 days	2582	254		
29 days to 3 months	5570	513		
≥3 months	1051	107		
Treatment classification			10.205 (1)	<.001
Initial treatment	8676	800		
Retreatment	527	74		
Diagnostic results			52.723 (1)	<.001
Bacteriologically positive	3831	253		
Bacteriologically negative	5372	621		
Primary-level management			3.498 (1)	.06
Yes	5808	580		
No	3395	294		
HIV coinfection			14.724 (1)	<.001
Yes	218	40		
No	8985	834		
Comorbid diabetes mellitus			0.83 (1)	.36
Yes	254	19		
No	8949	855		
Treatment management modality			430.013 (2)	<.001

XSL•FO
RenderX

Variable	Outcome		Chi-square (*df*)	*P* value
	Treatment success (n)	Unfavorable outcome (n)		
Intensive phase supervision	4054	668		
Full-course supervision	5149	206		
Complicated by extrapulmonary tuberculosis			0.243 (1)	.62
Yes	440	38		
No	8763	836		

Multivariate Analysis of Treatment Outcomes

Collinearity analysis was performed for all variables. All variables showed tolerance >0.1, variance inflation factor <10, eigenvalues substantially greater than 0, and condition indices <30, indicating no significant multicollinearity issues. Using treatment outcome as the dependent variable (1=Treatment success, 0=Unfavorable treatment outcome), variables showing significance in univariate analysis were included as independent variables in multivariate binary logistic regression with the enter method. Multivariate logistic regression revealed that age, treatment classification, HIV coinfection status, bacteriologic test results, and treatment management modality were significant determinants of treatment success (Table 8).

Table . Binary multivariate logistic regression analysis of treatment outcomes.

Variable	*B*	SE	Wald chi-square (*df*)	*P* value	Exp (*B*)	95% CI of Exp (*B*)	
						Lower limit	Upper limit
Age (y)							
<25	__[a]	—	56.590 (3)	<.001	—	—	—
25 - 44	0.562	0.145	15.012 (1)	<.001	1.755	1.320	2.332
45 - 64	0.680	0.106	41.055 (1)	<.001	1.974	1.603	2.431
≥65	0.598	0.092	41.860 (1)	<.001	0.819	1.518	2.181
Treatment classification (retreatment, initial)	1.331	0.207	41.375 (1)	<.001	3.786	2.524	5.680
Bacteriologic test results (positive, negative)	1.044	0.127	67.293 (1)	<.001	2.841	2.214	3.646
HIV test results (positive, negative)	0.916	0.192	22.655 (1)	<.001	2.499	1.714	3.643
Treatment management modality (intensive phase supervision)	—	—	343.845 (2)	<.001	—	—	—
Full-course supervision	2.042	0.272	56.275 (1)	<.001	7.705	4.520	13.137
Constant	−2.795	0.594	22.140 (1)	<.001	0.061	—	—

[a]not applicable.

Discussion

Our analysis of 2 decades of surveillance data yielded 3 main findings that align with the study objectives. First, Joinpoint regression confirmed a significant decline in the incidence of PTB (AAPC=–6.81; 95% CI –7.25 to –6.30%; *P*<.0001), with bacteriological positivity rates following a U-shaped trend over the study period. Second, spatial analysis revealed persistent high-burden areas, with Longshui Town and Tangxiang Subdistrict exhibiting the highest disease burden. Third, treatment success was significantly associated with younger age, initial treatment, HIV-negative status, bacteriologically negative results, and full-course supervised management.

Accurate data on TB incidence are difficult to obtain; consequently, reported incident case numbers serve as crucial indicators of disease burden and epidemiologic trends. From 2005 to 2024, a total of 10,236 patients with PTB were reported in our district. The annual notification incidence rate decreased from an initial rate of 131.2/100,000 population to 28.4/100,000 population in 2024, with an average annual rate of 65.2/100,000

XSL•FO
RenderX

population. This rate is slightly higher than the national average (60.77/100,000 population) [4] and the average annual notification incidence rate of Chongqing (54.6/100,000 population) [17] but lower than that of the Western China region (87.35/100,000 population) [4]. The notification incidence demonstrated an overall downward trend with an average annual decline rate of 7.7%, comparable to the annual decline rate observed in Chengdu City, China (7.75%) [18]. From an annual perspective, the TB case count in our district was particularly high in 2005, reaching 1000 cases, or 9.8% of the total cases reported during the entire study period. This historic peak is attributed to the full implementation of the Directly Observed Therapy-Short-course strategy [19] and improvements of the infectious disease reporting system that year. However, with the refinement of PTB prevention and control measures, annual case counts have gradually decreased, dropping to only 234 cases in 2024. The bacteriological positivity rate in our district initially decreased from 46.9% to 24.4% in 2016 and then gradually increased to 81.6% by 2024. This trend aligns with national reports on positivity rate changes [4], reflecting that our district's epidemic prevention and control strategies have kept pace with national efforts. The increase is also associated with the widespread adoption of molecular biology detection technologies in recent years, particularly second-generation gene sequencing of bronchoscopic alveolar lavage fluid, which has significantly enhanced the bacteriological positivity rate of suspected PTB cases [20].

The incidence of PTB exhibits seasonal patterns [21], with the highest numbers of reported cases in our district occurring in March, predominantly concentrated during spring and summer, while winter case counts remained relatively lower. This phenomenon aligns with the findings of an international study [22]. Potential contributing factors include (1) cold winter weather increases indoor congregation with poor ventilation, facilitating *MTB* transmission; (2) reduced sunlight exposure during the winter decreases vitamin D synthesis; low vitamin D levels may increase susceptibility to TB [23]; (3) January and February coincide with China's Spring Festival period. As TB symptoms often present atypically during early disease, many patients delay medical consultations until post-festival periods, leading to a surge in reported cases by March.

Our finding that cases were predominantly among male patients (male-to-female ratio of 3.57:1) was consistent with the findings of domestic and international studies [24,25]. This phenomenon may be attributed to gender-specific differences in hormone levels and genetic polymorphisms that influence susceptibility to *MTB* [26]. Many studies in humans and experimental animals have established clear links between sex-specific factors and the differential susceptibility of male and female individuals to a number of infectious diseases. Sex-specific determinants of Semin Immunopathol immunity include effects of sex steroid hormones as well as sex chromosome–encoded genes and microRNA [27]. Additionally, male patients often face systemic disadvantages in seeking or accessing TB care across various social contexts [28]. These findings highlight the necessity of incorporating gender-sensitive approaches into PTB prevention and control measures, particularly by enhancing diagnostic accessibility and screening services for male populations, to ensure equitable health care delivery across genders. This study revealed that reported PTB cases, in both male and female patients, were predominantly concentrated in the 15‐ to 64-year-old age group, consistent with previous research [29]. Notably, the proportion of cases among children less than 15 years of age is gradually decreasing in our district, while reported cases among those over the age of 65 years are progressively increasing. This trend suggests that our district must place greater emphasis on TB prevention and control in the older population. The observed pattern may be attributed to accelerated population aging, declining physiological function, and higher comorbidity rates among older adults, increasing susceptibility to TB. Conversely, the gradual reduction in cases among children under 15 years of age likely results from protection conferred by Bacillus Calmette Guerin vaccination and enhanced TB control measures targeting student populations. These findings highlight the necessity to increase the frequency of PTB screening among the older adults, particularly among male seniors, to reduce incidence rates in this vulnerable demographic.

Our finding that farmers constituted the majority (77.6%, 7948/10,236) of reported cases in our district is consistent with the findings of a previous study [30]. Potential explanations for this pattern include the following: (1) China's status as an agricultural nation with a large rural population and high demographic proportion; (2) rural areas may experience relatively underdeveloped economic conditions, limited access to medical services, and inadequate prevention and control measures; (3) a generally lower educational status and limited awareness of PTB-related knowledge among farmers. TB-related knowledge, educational levels, and awareness of transmission risk factors correlate with disease incidence [31,32]. Therefore, enhanced education regarding PTB prevention and control measures is essential in rural communities.

Our observation of the second-highest case count (5.6%, 570/10,236) among homemakers and unemployed may reflect an association between relatively low income levels and unwillingness to seek medical care when symptoms emerge, which may facilitate TB transmission [33]. The third-highest case count (5.2%, 532/10,236) among students may be related to school population densities, intense academic pressure, and sleep deprivation [34] that contribute to both *MTB* transmission and increased susceptibility to post-infection disease onset and progression. These findings collectively highlight the necessity for targeted prevention and control measures focused on these high-risk populations in our district.

Significant geographical disparities of incidence were observed across towns and subdistricts. The highest reported case count (13.1%, 1339/10,236) observed in Longshui Town, followed by Longgang Subdistrict and Tangxiang Subdistrict, may be related to the large population bases of these 3 towns and subdistricts. Guoliang Town exhibited the highest average annual incidence rate (314.4/100,000), while Tongqiao Subdistrict showed the lowest (16.8/100,000). A total of 21 towns or subdistricts had incidence rates exceeding 100/100,000, highlighting the persistently severe epidemic in our district. This necessitates intensified prevention and control measures in these areas, with particular attention required to mitigate

XSL•FO
RenderX

future epidemic trends in Guoliang Town. Global spatial autocorrelation analysis of township-level TB incidence in Dazu District (2021 - 2024) showed significant clustering only in 2023 (P<.001). Local spatial autocorrelation revealed indistinct hot-spot or cold-spot segregation. We speculate that spatial autocorrelation may have limited sensitivity at this fine scale due to low case counts leading to unstable rates and high geographical proximity with shared risk factors among adjacent townships. Creating geographic heatmaps using annual case counts may better illustrate spatial variations in the epidemic.

Our treatment success rate of 91.3% was slightly lower than reported in previous studies [35,36]. Multivariate analysis revealed that age, diagnostic criteria, management approaches, HIV coinfection status, and bacteriological test results were risk factors that influenced treatment success. Variations in economic status, external environmental exposures, and stress levels across different age groups, combined with age-related differences in immune function, contribute to the heterogeneous manifestations of PTB pathogenesis. Compromised immune function and chronic comorbidities among people over 65 years of age increase susceptibility to TB [37].

Positive bacteriologic test results indicate a higher bacterial load that may be more pathogenic and pose greater therapeutic challenges. Retreatment cases exhibited lower treatment success rates due to higher drug resistance rates, poor adherence, prolonged treatment duration, and more extensive lung pathology. Patients enrolled solely in intensive phase supervision programs may initially benefit from health care worker engagement that may facilitate more regular medication intake and heightened awareness of their condition. However, once the intensive phase concludes without continued supervision, patients may mistakenly believe they are cured of TB, resulting in reduced adherence, irregular medication intake, and ultimately unfavorable treatment outcomes. Individuals with HIV coinfection experience varying degrees of immune suppression, resulting in higher susceptibility to *MTB* infection, lower treatment success rates, and higher mortality rates [38,39].

This study has certain limitations. First, when analyzing annual incidence rates across towns and subdistricts, we were unable to obtain specific permanent resident population figures for corresponding years, and the proportion of reported cases in each town relative to the district's total annual cases was used as a substitute for the annual incidence rate. Second, potentially relevant variables such as nutritional status, silicosis comorbidity, and specific extrapulmonary TB manifestations were not included in the analysis of treatment outcomes. Third, underreporting inherent to passive surveillance systems, particularly among asymptomatic cases, may lead to an underestimation of true incidence. Fourth, the evolution of diagnostic techniques (including the introduction of molecular assays like Xpert *MTB* or rifampicin) over the 20-year period may influence the temporal comparability of bacteriological results. Fifth, univariate analyses involved multiple comparisons without statistical adjustment (eg, Bonferroni), increasing the risk of type I errors; thus, multivariate results should be prioritized. Finally, the findings, derived from Dazu District—an urban-rural transition zone in western Chongqing with a high farmer proportion (77.6%) and aging trend—may be most applicable to agricultural regions in central and western China and should be generalized to metropolitan core areas with caution.

In conclusion, the incidence of PTB in our district gradually decreased. Special focus should be directed toward monitoring epidemic trends among farmers and male individuals. Concurrently, treatment protocols should be enhanced for patients with PTB across different age groups, those with HIV coinfection, individuals receiving only intensive phase supervision, patients undergoing retreatment, and bacteriologically confirmed cases. These measures will facilitate the reduction of PTB incidence and the improvement of treatment success rates for patients in our district. Future research could focus on high-burden towns, explore and construct an active screening model for TB suitable for the grassroots level, and deeply study the impact mechanism and intervention strategies of labor mobility on the cross-regional transmission of TB, so as to provide assistance for the formation of collaborative prevention and control strategies at the district and county scale.

Acknowledgments

We would like to express our gratitude to the doctors at the Tuberculosis Clinic in our hospital, as well as the health care workers throughout the district, for their contributions in collecting the follow-up data of patients with tuberculosis. We thank the Medjaden Academy & Research Foundation for Young Scientists for its assistance in the preparation of this manuscript.

Funding

The study was supported by the first batch of key Disciplines on Public Health in Chongqing (YWBF2022072).

Data Availability

The raw data collected and used in this study can be obtained from the corresponding author upon reasonable request.

Authors' Contributions

YY and XJH jointly designed the study. JW conducted the data collection, while XMF performed data analysis. YY and XJH wrote the manuscript, which was reviewed and revised by JW. All authors reviewed and approved the final version for publication.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Occupational distribution and case count percentages of reported pulmonary tuberculosis (PTB) cases in Dazu District, Chongqing, China, 2005-2024.
[PNG File, 434 KB - publichealth_v11i1e78564_app1.png]

Checklist 1
STROBE checklist.
[DOCX File, 31 KB - publichealth_v11i1e78564_app2.docx]

References

1. Trajman A, Campbell JR, Kunor T, et al. Tuberculosis. The Lancet 2025 Mar;405(10481):850-866. [doi: 10.1016/S0140-6736(24)02479-6]
2. Anderer S. New global tuberculosis cases hit record high. JAMA 2025 Jan 7;333(1):12. [doi: 10.1001/jama.2024.24261]
3. Global tuberculosis report 2024. : World Health Organization; 2024 URL: https://www.who.int/teams/global-tuberculosis-programme/tb-reports/global-tuberculosis-report-2024 [accessed 2025-10-01]
4. Song YY, Li T, Xia H, et al. Analysis of the epidemiological characteristics of national reported pulmonary tuberculosis incidence, 1997–2023 [Article in Chinese]. Chin J Antituberc 2024;46(10):1198-1208. [doi: 10.19982/j.issn.1000-6621.20240382]
5. Chen J, Qiu Y, Wu W, et al. Trends and projection of the incidence of active pulmonary tuberculosis in southwestern China: age-period-cohort analysis. JMIR Public Health Surveill 2023 Dec 29;9:e48015. [doi: 10.2196/48015] [Medline: 38157236]
6. Jiang H, Zhang G, Yin J, et al. Assessment of strategies and epidemiological characteristics of tuberculosis in Henan province, China: observational study. JMIR Public Health Surveill 2021 Jan 22;7(1):e24830. [doi: 10.2196/24830] [Medline: 33480857]
7. Hu M, Feng Y, Li T, et al. Unbalanced risk of pulmonary tuberculosis in China at the subnational scale: spatiotemporal analysis. JMIR Public Health Surveill 2022 Jul 1;8(7):e36242. [doi: 10.2196/36242] [Medline: 35776442]
8. Liao Y, Pang Y, Zhao J, et al. Reporting status and characteristics of diagnosis delay among pulmonary tuberculosis patients in Liangping District, Chongqing Municipality from 2018 to 2023 [Article in Chinese]. J Tuberc Lung Dis 2025(1). [doi: 10.19983/j.issn.2096-8493.2024137]
9. He J, Zhang Y. Epidemiological features of pulmonary tuberculosis in Dianjiang County, Chongqing Municipality from 2016 to 2023 [Article in Chinese]. Paras Dis Infect Dis 2025(1):31-36. [doi: 10.19972/j.cnki.1672-2116.240056]
10. Xie JW, Xiong HL, Jiang H, et al. Spatial epidemiological characteristics of pulmonary tuberculosis in Rongchang District, Chongqing Municipality, 2011-2021 [Article in Chinese]. Pract Prev Med (4):424-428. [doi: 10.3969/j.issn.1006-3110.2023.04.009]
11. Zheng JY, Tang XQ, Xu R, et al. Epidemiological characteristics of pulmonary tuberculosis in Nan'an District, Chongqing Municipality, 2017-2021 [Article in Chinese]. J Appl Prev Med (3):177-179. [doi: 10.3969/j.issn.1673-758X.2023.03.013]
12. Zou JQ. Diagnostic criteria for pulmonary tuberculosis. Tuberc Health Educ 2008(1):3 [FREE Full text]
13. National Health and Family Planning Commission of the People's Republic of China. WS 288-2017 Diagnosis of Pulmonary Tuberculosis. J Tuberc Lung Dis 2024;5(4):376-378 [FREE Full text] [doi: 10.19983/j.issn.2096-8493.2024022]
14. National Health and Family Planning Commission of the People's Republic of China. Classification of Tuberculosis (WS196-2017). Emerg Infect Dis 2018;3(3) [FREE Full text]
15. Zhao YL, Chen MT. Technical Guidelines for Tuberculosis Prevention and Control in China: People's Medical Publishing House; 2021.
16. Zhou L, Liu EY. Early diagnosis and standardized treatment of pulmonary tuberculosis patients embodies the essence of implementing the latest national industry standards and technical guidelines. Chin J Antituberc 2020;42(9):2. [doi: 10.3969/j.issn.1000-6621.2020.09.002]
17. Lv J, Chen ZL, Wang QY, et al. Epidemiological characteristics of pulmonary tuberculosis among healthcare workers in Chongqing Municipality from 2017 to 2021. Disease Surveillance 2024(2):39 [FREE Full text]
18. Duan XF, Wu GH, Yang J, et al. Epidemiological characteristics and treatment outcomes of pulmonary tuberculosis in Chengdu City from 2012 to 2022 [Article in Chinese]. Emerg Infect Dis 2024;9(5):19-24. [doi: 10.19871/j.cnki.xfcrbzz.2024.05.004]
19. Jiang SW, Zhang H, Wang DM, et al. Contribution of the Global Fund tuberculosis program to tuberculosis control in China [Article in Chinese]. Chin J Antituberc 2014;36(5):333-339. [doi: 10.3969/j.issn.1000-6621.2014.05.001]
20. Dai XW, Wang NH, Chen SS, et al. Preliminary evaluation of next-generation sequencing for detecting Mycobacterium tuberculosis in clinical sputum specimens [Article in Chinese]. Chin J Antituberc 2022(7). [doi: 10.19982/j.issn.1000-6621.20220097]

XSL•FO
RenderX

21. Tedijanto C, Hermans S, Cobelens F, Wood R, Andrews JR. Drivers of seasonal variation in tuberculosis incidence: insights from a systematic review and mathematical model. Epidemiology (Sunnyvale) 2018 Nov;29(6):857-866. [doi: 10.1097/EDE.0000000000000877] [Medline: 29870427]
22. Taylan M, Dogru S, Sezgi C, Yilmaz S. Epidemiological trends and seasonal dynamics of tuberculosis in Southeastern Turkey. Niger J Clin Pract 2023 Jul;26(7):928-933. [doi: 10.4103/njcp.njcp_629_22] [Medline: 37635576]
23. Martineau AR, Nhamoyebonde S, Oni T, et al. Reciprocal seasonal variation in vitamin D status and tuberculosis notifications in Cape Town, South Africa. Proc Natl Acad Sci USA 2011 Nov 22;108(47):19013-19017. [doi: 10.1073/pnas.1111825108]
24. Korzeniewska-Koseła M. Tuberculosis in Poland in 2017. Przegl Epidemiol 2019;73(2):211-226. [doi: 10.32394/pe.73.20]
25. Fan YZ, Huang Y, Yang YR, et al. Epidemiological characteristics of pulmonary tuberculosis in Nanshan district of Shenzhen, 2016−2020. Disease Surveillance 2022;37(12):1583-1587 [FREE Full text]
26. Nhamoyebonde S, Leslie A. Biological differences between the sexes and susceptibility to tuberculosis. J Infect Dis 2014 Jul 15;209 Suppl 3:S100-S106. [doi: 10.1093/infdis/jiu147] [Medline: 24966189]
27. Hertz D, Schneider B. Sex differences in tuberculosis. Semin Immunopathol 2019 Mar;41(2):225-237. [doi: 10.1007/s00281-018-0725-6] [Medline: 30361803]
28. Horton KC, MacPherson P, Houben RMGJ, White RG, Corbett EL. Sex differences in tuberculosis burden and notifications in low- and middle-income countries: a systematic review and meta-analysis. PLoS Med 2016 Sep 6;13(9):e1002119. [doi: 10.1371/journal.pmed.1002119]
29. Wang L, Wang W. Temporal trends in notification and mortality of tuberculosis in China, 2004-2019: a joinpoint and age-period-cohort analysis. Int J Environ Res Public Health 2021 May 24;18(11):5607. [doi: 10.3390/ijerph18115607] [Medline: 34073943]
30. Zhang GY, Liu J, Shi SQ, et al. Trends in pulmonary tuberculosis epidemiology and analysis of epidemiological characteristics in Anhui Province, 2011–2020. Disease Surveillance 2021;36(10):1025-1030 [FREE Full text]
31. Bäckdahl T, Sharma M. Knowledge and transmission risk awareness of tuberculosis among the pilgrims attending a religious mass gathering in India: a cross-sectional study. BMC Public Health 2021 Nov 22;21(1):2141. [doi: 10.1186/s12889-021-12192-8] [Medline: 34809593]
32. Zhang X, Zhu X, Sun HY, et al. Analysis of the incidence and influencing factors of active tuberculosis among 199,182 residents in Mianyang, Sichuan. Chin J Evid Based Med 2021;21(6):642-647 [FREE Full text]
33. Zhang X. Research progress on epidemiological features and influencing factors of pulmonary tuberculosis. Disease Surveillance 2024;39(2):207-214 [FREE Full text]
34. Ibarra-Coronado EG, Pantaleón-Martínez AM, Velazquéz-Moctezuma J, et al. The bidirectional relationship between sleep and immunity against infections. J Immunol Res 2015;2015(4):678164. [doi: 10.1155/2015/678164] [Medline: 26417606]
35. Tian XM, Ma QY, Lei J, et al. Evaluation of epidemiological characteristics and treatment outcomes of pulmonary tuberculosis patients with diabetes mellitus in Ningxia, 2019-2022 [Article in Chinese]. Ningxia Med J 2024;46(5):438-441. [doi: 10.13621/j.1001-5949.2024.05.0438]
36. Yang JY, Xiao XC, Du YH, et al. Analysis of epidemiological characteristics and treatment outcomes of pulmonary tuberculosis patients complicated with diabetes in Guangzhou, 2014-2023 [Article in Chinese]. Chin J Antituberc 2024;46(12):1504-1510. [doi: 10.19982/j.issn.1000-6621.20240232]
37. Yu Y, Chen J, Xu B, et al. Analysis of treatment outcomes and influencing factors in Rifampicin-sensitive pulmonary tuberculosis patients in Chongqing Municipality [Article in Chinese]. Chin J Antituberc 2024;46(12):1511-1518. [doi: 10.19982/j.issn.1000-6621.20240273]
38. Mo SL, Li MJ, Huang XH, et al. Advances in diagnosis and treatment of HIV/AIDS complicated with pulmonary tuberculosis [Article in Chinese]. Hebei Med J 2021;43(16):2526-2530. [doi: 10.3969/j.issn.1002-7386.2021.16.030]
39. Liu LD, Zhang LY. Epidemiological features and treatment outcomes of Mycobacterium tuberculosis/human immunodeficiency virus co-infection in Jingzhou, Hubei Province. Infect Dis Inf 2024;37(5):405-407 [FREE Full text] [doi: 10.3969/j.issn.1007-8134.2024.05.004]

Edited by A Mavragani; submitted 04.06.25; peer-reviewed by A Alabi, A Jamali, S Su, Y Zhang; revised version received 07.10.25; accepted 27.10.25; published 28.11.25.

Please cite as:
Yu Y, Hu XJ, Fang XM, Wu J
Epidemiology and Treatment Outcomes of Pulmonary Tuberculosis in Dazu District, Chongqing, China, 2005-2024: Surveillance Study
JMIR Public Health Surveill 2025;11:e78564
URL: https://publichealth.jmir.org/2025/1/e78564
doi:10.2196/78564

XSL•FO
RenderX

© Yu Yu, Xiu-juan Hu, Xiao-man Fang, Jing Wu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Primary Health Care Data Quality on Infectious Disease Surveillance in Brazil: Case Study

Pilar Tavares Veras Florentino[1], PhD; Juracy Bertoldo Junior[1], MSc; George Caique Gouveia Barbosa[1], MSc; Thiago Cerqueira-Silva[2], PhD; Vinicius de Araújo Oliveira[1,3], MD; Marcio Henrique de Oliveira Garcia[4], MSc; Gerson Oliveira Penna[5,6], PhD; Viviane Boaventura[7], PhD; Pablo Ivan Pereira Ramos[1], PhD; Manoel Barral-Netto[1,7], PhD; Izabel Marcilio[1,8], PhD

[1]Centro de Integração de Dados e Conhecimento em Saúde (CIDACS), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, R. Mundo, 121 - sala 315 - Trobogy, Salvador, Brazil

[2]Department of Medical Statistics, London School of Hygiene & Tropical Medicine, London, United Kingdom

[3]Secretaria de Atenção Primária, Ministério da Saúde, Brasília, Brazil

[4]Secretaria de Vigilância em Saúde e Ambiente, Ministério da Saúde, Brasília, Brazil

[5]Escola Fiocruz de Governo, Fundação Oswaldo Cruz (Fiocruz), Rio de Janeiro, Brazil

[6]Núcleo de Medicina Tropical, Universidade de Brasília, Brasília, Brazil

[7]Laboratório de Medicina e Saúde Pública de Precisão, Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil

[8]Departamento de Epidemiologia, Escola Bahiana de Medicina e Saúde Pública, Salvador, Brazil

Corresponding Author:
Pilar Tavares Veras Florentino, PhD
Centro de Integração de Dados e Conhecimento em Saúde (CIDACS), Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, R. Mundo, 121 - sala 315 - Trobogy, Salvador, Brazil

Abstract

Background: The increase in emerging and re-emerging infectious disease outbreaks underscores the need for robust early warning systems (EWSs) to guide mitigation and response measures. Administrative health care databases provide valuable epidemiological insights without imposing additional burdens on health services. However, these datasets are primarily collected for operational use, making data quality assessment essential to ensure an accurate interpretation of epidemiological analysis. This study focuses on the development and implementation of a data quality index (DQI) for surveillance integrated into an EWS for influenza-like illness (ILI) outbreaks using Brazil's a nationwide Primary Health Care (PHC) dataset.

Objective: We aimed to evaluate the impact of data completeness and timeliness on the performance of an EWS for ILI outbreaks and establish optimal thresholds for a suitable DQI, thereby improving the accuracy of outbreak detection and supporting public health surveillance.

Methods: A composite DQI was established to measure the completeness and timeliness of PHC data from the Brazilian National Information System on Primary Health Care. Completeness was defined as the proportion of weeks within an 8-week rolling window with any register of encounters. Timeliness was calculated as the interval between the date of encounter and its corresponding registry in the information system. The backfilled PHC dataset served as the gold standard to evaluate the impact of varying data quality levels from the weekly updated real-time PHC dataset on the EWS for ILI outbreaks across 5570 Brazilian municipalities from October 10, 2023, to March 10, 2024.

Results: During the study period, the backfilled dataset recorded 198,335,762 ILI-related encounters, averaging 8,623,294 encounters per week. The EWS detected a median of 4 (IQR 2 - 5) ILI outbreak warnings per municipality using the backfilled dataset. Using the real-time dataset, 12,538 (65%) warnings were concordant with the backfilled dataset. Our analysis revealed that 100% completeness yielded 76.7% concordant warnings, while 80% timeliness resulted in at least 50% concordant warnings. These thresholds were considered optimal for a suitable DQI. Restricting the analysis to municipalities with a suitable DQI increased concordant warnings to 80.4%. A median of 71% (IQR 54%-71.9%) of municipalities met the suitable DQI threshold weekly. Municipalities with ≥60% of weeks achieving a suitable DQI demonstrated the highest concordance between backfilled and real-time datasets, with those achieving ≥80% of weeks showing 82.3% concordance.

Conclusions: Our findings highlight the critical role of data quality in improving the EWS' performance based on PHC data for detecting ILI outbreaks. The proposed framework for real-time DQI monitoring is a practical approach and can be adapted to other surveillance systems, providing insights for similar implementations. We demonstrate that optimal completeness and timeliness of data significantly impact the EWS' ability to detect ILI outbreaks. Continuous monitoring and improvement of data quality should remain a priority to strengthen the reliability and effectiveness of surveillance systems.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e67050) doi:10.2196/67050

KEYWORDS

primary health care; data quality; infectious disease surveillance; Brazil; early warning system

Introduction

In recent decades, the world has witnessed an unprecedented surge of emerging and re-emerging infectious disease outbreaks, underscoring the need for stronger early warning systems (EWSs) [1,2]. The widespread and growing use of electronic health records (EHRs) has heightened the demand for automated processes in disease surveillance [3,4].

Systematic monitoring of administrative health care databases provides valuable epidemiological insights [5,6]. Importantly, using administrative data for health surveillance avoids the overburden of surveillance teams while ensuring timeliness, as no duplication of registry is required [3]. This cost-efficient approach enhances the ability to detect outbreaks, particularly in low-resource settings, thus contributing to global security [3].

However, an effective automated EWS based on administrative datasets requires that a real-time data quality assessment algorithm is set within the EWS pipeline. Since administrative data are primarily collected for operational purposes, assessing their quality is crucial to accurate interpretation of epidemiological analysis [7,8]. A systematic review on EHR data quality assessment studies found 14 articles describing dedicated data quality programs deployed in real-world settings, while only 4 produced results generally applicable in diverse settings. Ozonze et al [9] suggest there is an absence of comprehensive tools for facilitating reliable and consistent data quality assessments.

Moreover, despite existing methods for evaluating the quality of administrative health data, including EHR data quality assessment [8] and indicators for specific programs such as the Data Quality Audit and the Data Quality Self-Assessment for immunization data [10], there remains a gap in applying similar methods to data used for health surveillance. Although metrics for assessing the quality of surveillance systems are well established [3], to the best of our knowledge, these have not been applied to evaluate administrative data when used for epidemiological surveillance purposes.

This paper describes the development and implementation of a data quality index (DQI) to assess the quality of administrative data used in epidemiological surveillance systems. We focus on applying the DQI to nationwide Brazilian primary health care (PHC) administrative data integrated into an EWS for influenza-like illness (ILI) outbreaks. The study compares the EWS performance across different DQI levels, addressing a critical gap in current research by establishing metrics that ensure accurate and timely outbreak detection while leveraging the cost-efficiency of administrative health databases.

Methods

Study Design

We developed and implemented a data quality assessment algorithm within ÆSOP (Alert-Early System of Outbreaks with Pandemic Potential), a previously validated EWS [11]. This EWS applies aberration detection algorithms, such as the Early Aberration Reporting System (C2) [12], to a time series consisting of weekly counts of ILI-related PHC encounters per municipality, aiming at the early detection of outbreaks. To assess the data quality of the PHC data stream, we established the composite indicator DQI to measure the completeness and timeliness of the data. Using the backfilled PHC dataset as a gold standard, we evaluated the impact of data quality in the EWS' performance using different levels of data quality of the weekly updated real-time PHC dataset across all 5570 Brazilian municipalities from October 10, 2023, to March 10, 2024.

Data Source

Brazil is an upper middle-income country with approximately 212.6 million people living in 5570 municipalities [13], and we included all ILI-related PHC encounters occurring during the study period in our analysis. We analyzed data from the Brazilian Unified Health System (SUS), which stands as one of the largest public health systems globally, providing comprehensive and universal health care to the entire population. The effective management of SUS relies on diverse information systems, among which the Brazilian National Information System on Primary Health Care (SISAB [Sistema de Informação em Saúde para a Atenção Básica]) plays a crucial role. SISAB is a hierarchical, decentralized information system maintained and managed by the Ministry of Health (MoH), and harbors data on all publicly funded PHC encounters in the country. Data registration is mandatory for the allocation of financial resources from the federal to the municipal level. All encounters are coded by the *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) or the International Classification of Primary Care (ICPC-2).

According to the MoH's guidelines, municipalities are requested to update the system at least on a monthly basis, with a window of 4 months for amendments following each monthly submission. This operational guideline aligns with the SISAB's purpose of informing decision-making for the management of the PHC system in the country. However, the EWS uses weekly updates of the SISAB database to detect ILI outbreaks. Therefore, this real-time, weekly updated dataset may present incompleteness and a temporal lag between the dates of encounter and data registration into the system (Figure S1 in Multimedia Appendix 1).

The DQI

We defined the dimensions of completeness and timeliness to develop quantitative indicators for monitoring data quality in

XSL•FO
RenderX

the EWS. Completeness is one of the most commonly used dimensions in data quality assessment and may be defined as the proportion of data filled with values for each attribute or entity in the database, while timeliness can be defined as the availability of data for decision-making, measured by the time interval between the occurrence of the measured event and its capture in an information system [14].

In our study, completeness refers to the proportion of weeks in each 8-week rolling window with any register of a PHC encounter. The indicator is measured as a fraction, with the numerator ranging from 0 to 8, and the denominator is 8 weeks, which is expressed as a percentage. Timeliness refers to the time interval, in number of weeks, between the date of the PHC encounter and its registry in the database. The indicator is represented by the proportion of registries occurring in 2 weeks or less from the PHC encounter in the same 8-week rolling window.

As it is recommended that the diverse quality dimensions should be collectively analyzed for a more comprehensive evaluation of data quality [15], we combined the 2 selected indicators in a composite measure, named DQI. The DQI is assessed weekly, for each municipality, once the PHC data are updated into the EWS pipeline.

Impact of DQI on the EWS' Performance

To decide on the minimum required threshold of completeness and timeliness to derive trustworthy results with the EWS, we applied the EWS algorithm to the retrospectively gathered, backfilled PHC dataset. We compared the results to those obtained when applying the EWS to the weekly updated, herein named real-time PHC dataset (Figure S1 in Multimedia Appendix 1). Using the backfilled dataset as a reference, we calculated the proportion of concordant warnings detected in the real-time dataset. Accordingly, the DQI is expressed as either "suitable" or "unsuitable" when the minimum threshold of both completeness and timeliness is reached, indicating that the data quality may not be adequate for reliable EWS outputs.

Analyses were performed using Python (version 3.9) and R (version 4.3.1) software. The database's description and the scripts are available on GitHub [16].

Ethical Considerations

The study protocol and procedures were reviewed and approved by the Ethical Review Board of Oswaldo Cruz Foundation – Fiocruz Bahia (protocol CAAE 61444122.0.0000.0040).

Data on publicly funded PHC encounters were collected and compiled by the MoH for funding reasons. No consent was needed for data collection at this administrative level. For this study, we accessed an aggregated database consisting of the number of encounters per epidemiological week, per municipality, and per diagnostic code. The accessed database has no information at the individual level, and given that this study involves secondary analysis of existing deidentified data and does not involve direct interaction with human participants, it is classified as exempt from the requirement for informed consent under applicable ethical guidelines.

Results

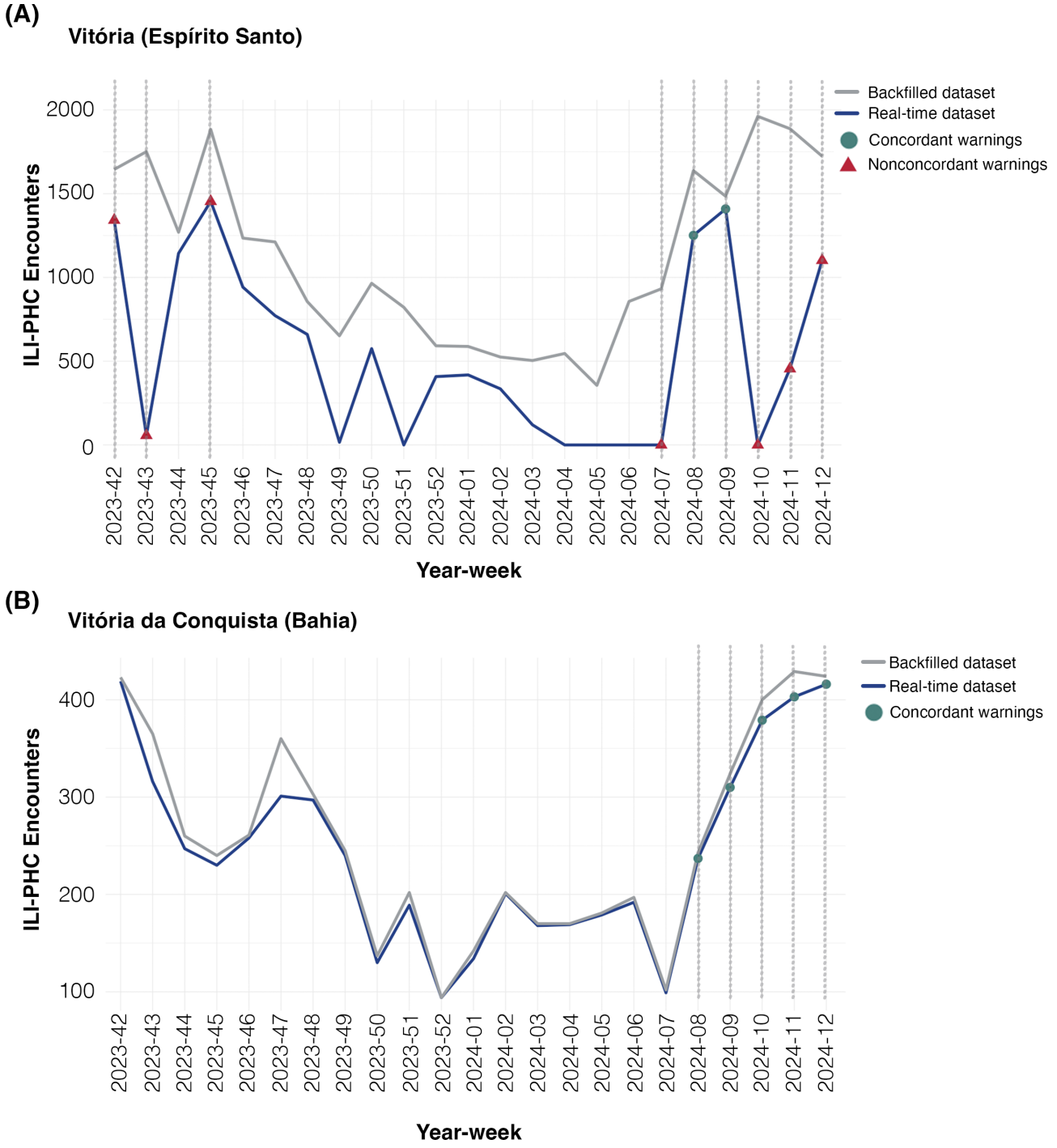
There were 198,335,762 recorded ILI-related encounters in the backfilled PHC dataset, which corresponds to an average of 8,623,294 encounters per week between October 10, 2023, and March 10, 2024. Using the backfilled dataset, the EWS detected a median of 4 (IQR 2 - 5) warnings of ILI outbreaks per municipality in the study period.

Figure 1 illustrates the impact of the DQI on the ability of the EWS to correctly identify potential ILI outbreaks. Using the real-time dataset, the EWS detected 12,538 (65%) warnings of ILI-outbreaks that were concordant with warnings detected in the backfilled dataset (Table S1 in Multimedia Appendix 1). The proportion of concordant warnings detected in the real-time dataset, based on different levels of completeness (Figure 2A and Table S1 in Multimedia Appendix 1) and timeliness (Figure 2B and Table S1 in Multimedia Appendix 1), indicated that 100% completeness and a minimum of 80% timeliness yielded the highest percentage of concordant warnings. Therefore, these values were established as the thresholds for grading the DQI as suitable or unsuitable for the EWS. Restricting the EWS analysis to municipalities with a suitable DQI, the proportion of warnings for ILI outbreaks concordant to the backfilled dataset increased to 80.4% (Table S2 in Multimedia Appendix 1). We found a median of 71% (IQR 54% - 71.9%) of Brazilian municipalities with a suitable DQI per week in the study period (Figure 3A and Table S2 in Multimedia Appendix 1).

Additionally, we analyzed concordant warnings by grouping municipalities based on the proportion of weeks in which they exhibited a suitable DQI (≤20%, 20% - 40%, 40% - 60%, 60% - 80%, and ≥80%). Our findings revealed that municipalities with over 60% of weeks featuring a suitable DQI had the highest proportion of concordant warnings between the backfilled and real-time datasets (Figure 3B, Table S2 in Multimedia Appendix 1).

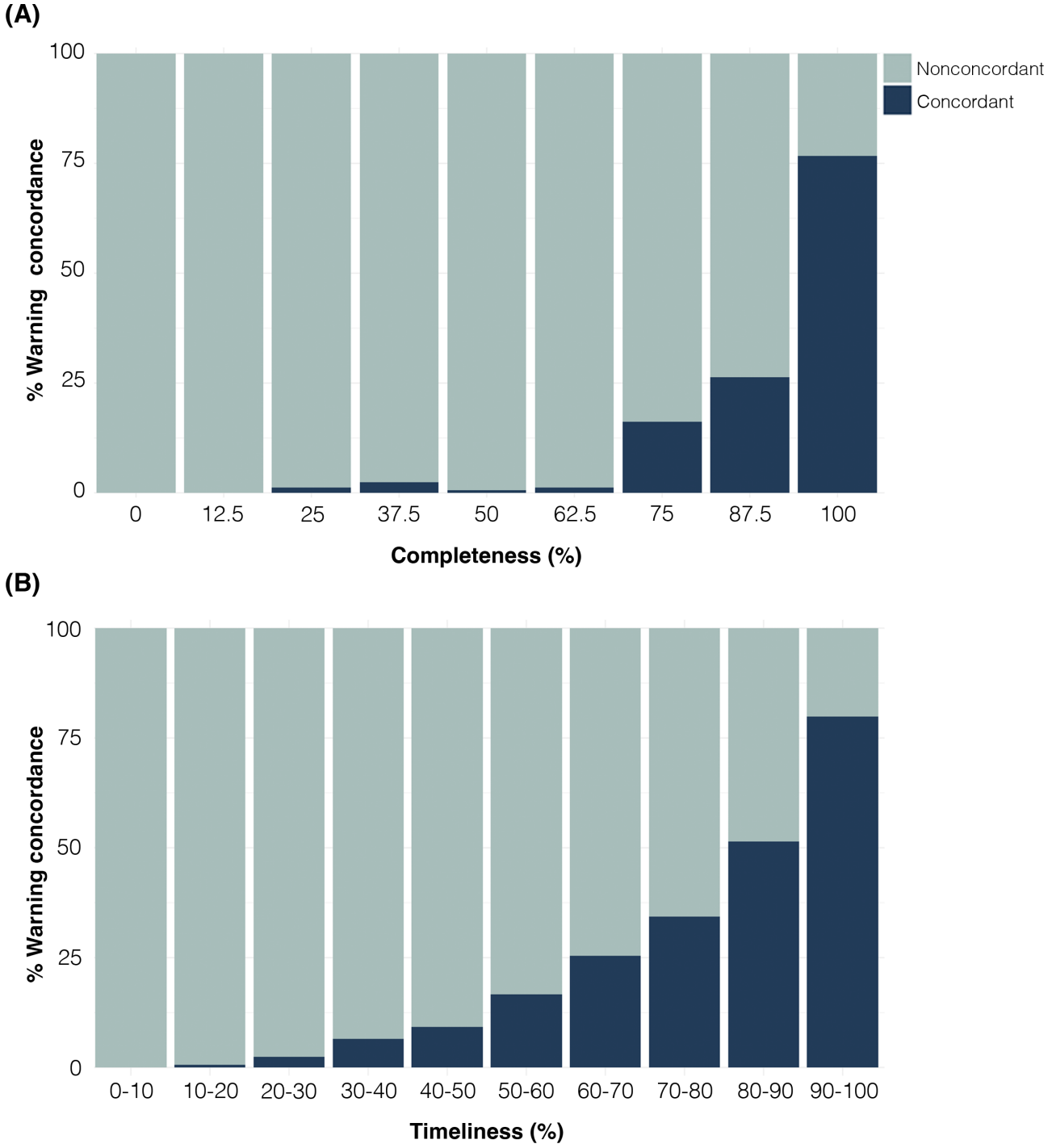
XSL•FO
RenderX

Figure 1. Primary Health Care encounters due to influenza-like illness per week in (A) Sao Paulo (municipality with a suitable data quality index in less than 60% of the 23 weeks in study period) and (B) Vitoria da Conquista (municipality with a suitable data quality index for over 80% of the study period). Plots show backfilled (gray line) and real-time (green line) PHC datasets for influenza-like illness (ILI) encounters. Vertical dashed lines show all detected warnings with the Early Aberration Reporting System (EARS), red triangles show nonconcordant warnings between backfield and real-time datasets, and blue circles show concordant warnings between them. ILI: influenza-like illness; PHC: Primary Health Care.



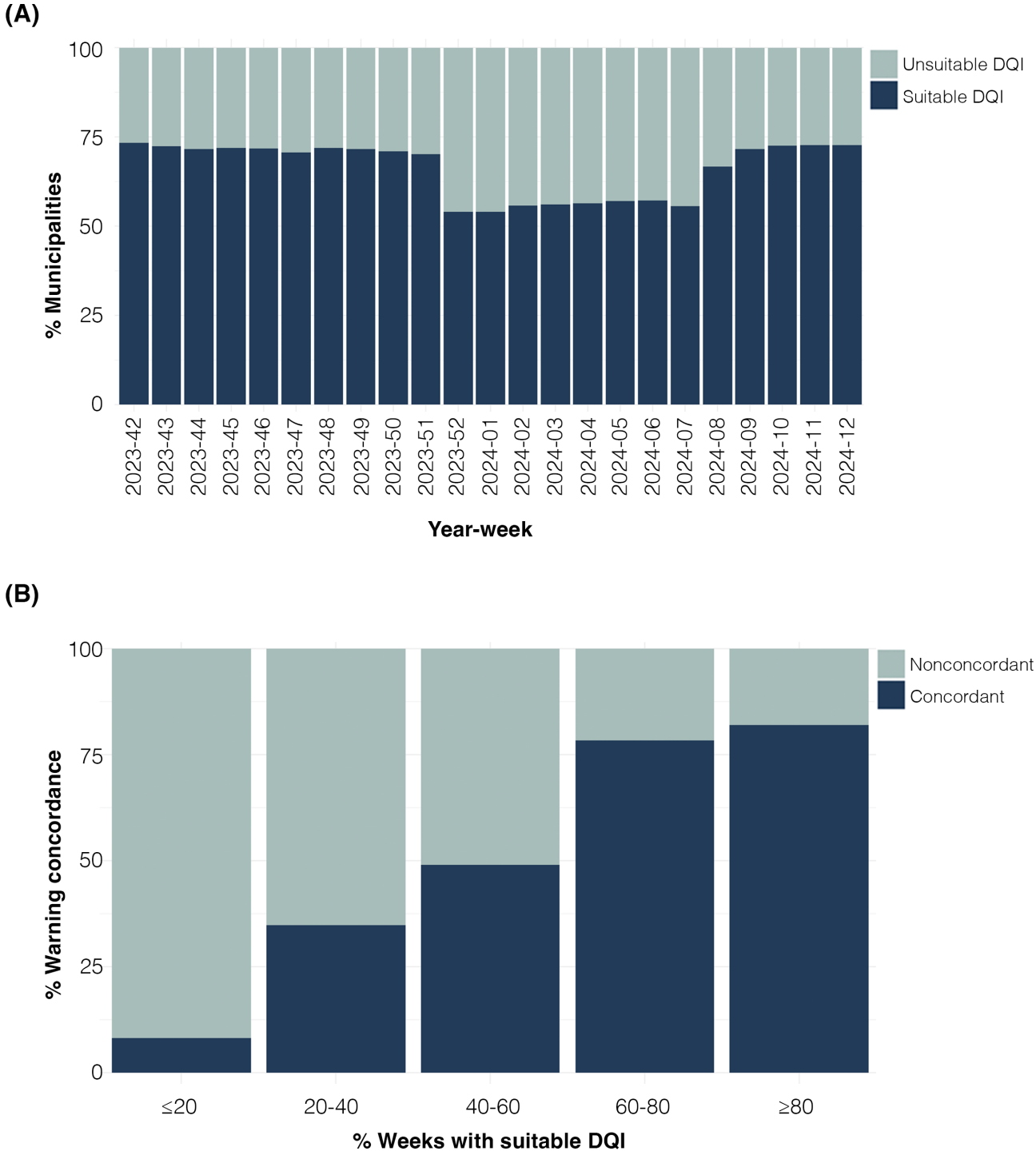
XSL•FO
RenderX

Figure 2. Proportion of concordant outbreak warnings detected in backfilled and real-time Brazilian Primary Health Care datasets. Outbreak warnings generated by the Early Aberration Reporting System (EARS) were identified in both backfilled and real-time datasets. Concordant warnings, detected in both datasets within the same week, are represented in dark blue, while nonconcordant warnings, identified only in the backfilled dataset, are shown in light blue. The analysis considers the proportion of concordant and nonconcordant warnings based on real-time dataset. (**A**) Completeness: the percentage of records from the real-time dataset in each 8-week rolling window (ranging from 0% to 100%) and (**B**) timeliness: the proportion of records registered within 2 weeks or less of the PHC encounter, measured within the same 8-week rolling window (ranging from 0% to 100%).



XSL•FO
RenderX

Figure 3. Proportion of Brazilian municipalities with suitable data quality index (DQI) and concordant warnings over time. (**A**) Weekly analysis of the proportion of municipalities with a suitable DQI from epidemiological week 42 of 2023 to week 12 of 2024. (**B**) Proportion of concordant warnings (dark blue), identified in both datasets within the same week, and nonconcordant warnings (light blue), detected only in the backfilled dataset. The analysis is based on the proportion of weeks with a suitable DQI in the real-time dataset from Brazilian municipalities.



Discussion

Principal Findings

Our study highlights the critical role of data quality in the performance of the EWS for infectious disease surveillance using PHC data. In addition, we provide a practical approach for monitoring data quality in real time, which can be adapted to other settings and data types. Our findings revealed that municipalities with over 60% of weeks featuring a suitable DQI had the highest proportion of concordant warnings between the backfilled and real-time datasets. Introducing the DQI as an algorithm integrated into the EWS can guide data management practices and inform decision-making processes.

Similar to our findings, a recent systematic review of the effectiveness of EWS found that the improvement of data is pivotal for emergency department–based surveillance [17]. However, efforts for automatization of data quality assessment are typically scattered [9], and the literature on the operationalization of data quality assessment remains scarce.

XSL•FO
RenderX

A study on data quality assessment for public health information systems found a lack of systematic procedures for quality assessment. While quality assessment of quantitative data generally used descriptive surveys, the authors argued about the importance of systematic scientific data quality assessment [18]. To the best of our knowledge, this is the first publication to assess the importance of integrating data quality monitoring into an EWS.

Fulcher et al [19] demonstrated how administrative health data were successfully used to implement a syndromic surveillance system during the COVID-19 pandemic. However, the process of cleaning data and handling missed data was carried out by a dedicated analyst once the updated database became available [19]. We anticipate that the framework for a data quality assessment integrated to the EWS pipeline presented here can be adapted to other surveillance systems and can provide insights for similar implementations.

Using a retrospectively gathered, backfilled PHC dataset, we evaluated the EWS based on optimal data quality conditions. However, administrative data usually exhibit incompleteness and delays, and the EWS should be capable of detecting outbreaks using the available dataset in real time. Our analysis revealed that high levels of completeness (100%) and timeliness (at least 80%) are necessary to achieve the highest proportion of concordant warnings between backfilled and real-time datasets. Additionally, our results indicate that even incremental data quality improvements substantially enhance the EWS' performance. Achieving such high standards may pose challenges, particularly in low-resource settings that potentially face limitations due to infrastructure such as unreliable internet connectivity and insufficient computer power. Despite these challenges, we found a weekly median of 71% of Brazilian municipalities achieving the threshold for a suitable DQI for the EWS. This result suggests that a significant proportion of municipalities met the minimum threshold for data quality even in constrained settings.

In this study, we used the SUS database, which covers approximately 75% of the Brazilian population, with great granularity, reaching underserved rural and remote regions [20]. This approach allowed us to assess the performance of the EWS across different regions and health service contexts. However, these findings may not be directly applicable to other countries. It is likely that the use of the EWS in different health system structures and data management practices will need adjustments and may require distinct data quality requirements [8].

Another limitation of this study is that we could not access other dimensions of data quality. Specifically, we could not access the accuracy of registers in the PHC dataset. Accuracy represents the extent to which the data are free of error and reliable [8,9]. We worked with aggregated, secondary data, and did not have access to the complete EHRs, which precluded us from verifying whether the diagnostic codes in the database accurately reflected patients' main clinical problems. It is our perspective that evaluating the accuracy of the *ICD-10* and ICPC-2 is of great importance. However, given the large numbers of PHC encounters registered weekly, misclassifications of the reason of encounter are likely to be nondifferential. Additionally, syndromic surveillance systems are designed to operate effectively even with some level of imprecision, as their primary purpose is to detect patterns and trends rather than to provide definitive diagnoses.

Conclusion

Our findings demonstrate that implementing a robust and integrated DQI analysis can significantly enhance the EWS' ability to detect ILI outbreaks, contributing to better public health outcomes and ultimately to global health security. Beyond contributing to the existing literature on EWS, this study highlights the importance of systematic data quality assessment. Continuous monitoring and improvement of data quality should be prioritized to ensure the reliability and effectiveness of surveillance systems. Additionally, our study suggests that similar frameworks can be adapted to different contexts. As health systems increasingly use digital health data for decision-making, our approach represents a model for integrating data quality monitoring into surveillance systems, ultimately enhancing the capacity to detect and respond to infectious disease outbreaks effectively.

Acknowledgments

This study was funded by the Rockefeller Foundation's Health Initiative (grant 2023-PPI-007 awarded to MBN). MBN, PIPR, and VB are research fellows from the National Council for Scientific and Technological Development (CNPq, Brazil). TCS acknowledges funding from the Royal Society (NIF\R1\231435). The funders did not interfere in the analysis, interpretation, or decision to submit the manuscript for publication. This study is part of the Alert-Early System of Outbreaks with Pandemic Potential (ÆSOP), an initiative under development by Brazil's Fundação Oswaldo Cruz (Fiocruz) and the Federal University of Rio de Janeiro, and financially supported by Rockefeller Foundation's Health Initiative. Figure S1 in Multimedia Appendix 1 was created with Biorender.com and was published with permission. Generative artificial intelligence was not used for ideation or any part of the study design. Additionally, it was not used for reference searches. Its use was limited to grammatical revisions in the manuscript, with the prompt, "Please check for readability and possible grammatical mistakes" (a transcript of the conversation with the chatbot is provided in Multimedia Appendix 2).

Data Availability

Our agreement with the Brazilian Ministry of Health (MoH) for accessing the referenced databases patently denies authorization of access to any third parties. All requests to access these databases must be addressed to the Brazilian MoH.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: PTVF, IM, and MBN
Data acquisition: VdAO
Data curation and processing: VdAO, JBJ, and GCGB
Formal analysis: PTVF
Script verification: JBJ, GCGB, and TCS
Study design: PTVF, JBJ, GCGB, TCS, VdAO, MHdOG, GOP, VB, PIPR, MBN, and IM
Writing—original draft: PTVF and IM
Writing—review and editing: GOP, TCS, VB, MHdOG, VdAO, and PIPR

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional material.
[DOCX File, 141 KB - publichealth_v11i1e67050_app1.docx]

Multimedia Appendix 2
Conversation with the chatbot for grammatical revision.
[PDF File, 562 KB - publichealth_v11i1e67050_app2.pdf]

References

1. Baker RE, Mahmud AS, Miller IF, et al. Infectious disease in an era of global change. Nat Rev Microbiol 2022 Apr;20(4):193-205. [doi: 10.1038/s41579-021-00639-z] [Medline: 34646006]
2. Morgan OW, Aguilera X, Ammon A, et al. Disease surveillance for the COVID-19 era: time for bold changes. The Lancet 2021 Jun 19;397(10292):2317-2319. [doi: 10.1016/S0140-6736(21)01096-5] [Medline: 34000258]
3. Data quality monitoring and surveillance system evaluation - a handbook of methods and applications. European Centre for Disease Prevention and Control. 2014. URL: https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/Data-quality-monitoring-surveillance-system-evaluation-Sept-2014.pdf [accessed 2025-02-06]
4. Ramos PIP, Marcilio I, Bento AI, et al. Combining digital and molecular approaches using health and alternate data sources in a next-generation surveillance system for anticipating outbreaks of pandemic potential. JMIR Public Health Surveill 2024 Jan 9;10:e47673. [doi: 10.2196/47673] [Medline: 38194263]
5. Ulrich EH, So G, Zappitelli M, Chanchlani R. A review on the application and limitations of administrative health care data for the study of acute kidney injury epidemiology and outcomes in children. Front Pediatr 2021;9:742888. [doi: 10.3389/fped.2021.742888] [Medline: 34778133]
6. Tassi MF, le Meur N, Stéfic K, Grammatico-Guillon L. Performance of French medico-administrative databases in epidemiology of infectious diseases: a scoping review. Front Public Health 2023;11:1161550. [doi: 10.3389/fpubh.2023.1161550] [Medline: 37250067]
7. Shaw RJ, Harron KL, Pescarini JM, et al. Biases arising from linked administrative data for epidemiological research: a conceptual framework from registration to analyses. Eur J Epidemiol 2022 Dec;37(12):1215-1224. [doi: 10.1007/s10654-022-00934-w] [Medline: 36333542]
8. Lewis AE, Weiskopf N, Abrams ZB, et al. Electronic health record data quality assessment and tools: a systematic review. J Am Med Inform Assoc 2023 Sep 25;30(10):1730-1740. [doi: 10.1093/jamia/ocad120] [Medline: 37390812]
9. Ozonze O, Scott PJ, Hopgood AA. Automating electronic health record data quality assessment. J Med Syst 2023 Feb 13;47(1):23. [doi: 10.1007/s10916-022-01892-2] [Medline: 36781551]
10. Bloland P, MacNeil A. Defining & assessing the quality, usability, and utilization of immunization data. BMC Public Health 2019 Apr 4;19(1):380. [doi: 10.1186/s12889-019-6709-1] [Medline: 30947703]
11. Cerqueira-Silva T, Oliveira JF, de Araújo Oliveira V, et al. Early warning system using primary health care data in the post-COVID-19 pandemic era: Brazil nationwide case-study. Cad Saude Publica 2024;40(11):e00010024. [doi: 10.1590/0102-311XEN010024] [Medline: 39775767]
12. Zhu Y, Wang W, Atrubin D, Wu Y. Initial evaluation of the early aberration reporting system--Florida. MMWR Suppl 2005 Aug 26;54:123-130. [Medline: 16177703]
13. 2022 population census: main results. Brazilian Institute of Geography and Statistics. 2022. URL: https://www.ibge.gov.br/en/statistics/social/health/22836-2022-census-3.html [accessed 2024-12-18]
14. Fox C, Levitin A, Redman T. The notion of data and its quality dimensions. Inf Process Manag 1994 Jan;30(1):9-19. [doi: 10.1016/0306-4573(94)90020-5]
15. Hassenstein MJ, Vanella P. Data quality—concepts and problems. Encyclopedia 2022 Feb;2(1):498-510. [doi: 10.3390/encyclopedia2010032]

XSL•FO
RenderX

16. AESOP-data-documentation. GitHub. URL: https://github.com/cidacslab/AESOP-Data-Documentation/tree/main/DataPipeline [accessed 2025-02-06]
17. Meckawy R, Stuckler D, Mehta A, Al-Ahdal T, Doebbeling BN. Effectiveness of early warning systems in the detection of infectious diseases outbreaks: a systematic review. BMC Public Health 2022 Nov 29;22(1):2216. [doi: 10.1186/s12889-022-14625-4] [Medline: 36447171]
18. Chen H, Hailey D, Wang N, Yu P. A review of data quality assessment methods for public health information systems. Int J Environ Res Public Health 2014 May 14;11(5):5170-5207. [doi: 10.3390/ijerph110505170] [Medline: 24830450]
19. Fulcher IR, Boley EJ, Gopaluni A, et al. Syndromic surveillance using monthly aggregate health systems information data: methods with application to COVID-19 in Liberia. Int J Epidemiol 2021 Aug 30;50(4):1091-1102. [doi: 10.1093/ije/dyab094]
20. Macinko J, Harris MJ, Rocha MG. Brazil's National Program for Improving Primary Care Access and Quality (PMAQ): fulfilling the potential of the world's largest payment for performance system in primary care. J Ambul Care Manage 2017;40 Suppl 2 Supplement(2 Suppl):S4-S11. [doi: 10.1097/JAC.0000000000000189] [Medline: 28252498]

Abbreviations

DQI: data quality index
EHR: electronic health record
EWS: early warning system
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
ICPC-2: International Classification of Primary Care
ILI: influenza-like illness
MoH: Ministry of Health
PHC: Primary Health Care
SISAB: Sistema de Informação em Saúde para a Atenção Básica
SUS: Brazilian Unified Health System
***Æ*SOP:** Alert-Early System of Outbreaks with Pandemic Potential

Edited by A Mavragani; submitted 30.09.24; peer-reviewed by C Haridas, O Olasunkanmi; revised version received 19.12.24; accepted 20.12.24; published 21.02.25.

Please cite as:
Florentino PTV, Bertoldo Junior J, Barbosa GCG, Cerqueira-Silva T, Oliveira VDA, Garcia MHDO, Penna GO, Boaventura V, Ramos PIP, Barral-Netto M, Marcilio I
Impact of Primary Health Care Data Quality on Infectious Disease Surveillance in Brazil: Case Study
JMIR Public Health Surveill 2025;11:e67050
URL: https://publichealth.jmir.org/2025/1/e67050
doi:10.2196/67050

© Pilar Tavares Veras Florentino, Juracy Bertoldo Junior, George Caique Gouveia Barbosa, Thiago Cerqueira-Silva, Vinicius de Araújo Oliveira, Marcio Henrique de Oliveira Garcia, Gerson Oliveira Penna, Viviane Boaventura, Pablo Ivan Pereira Ramos, Manoel Barral-Netto, Izabel Marcilio. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Frequent Point-of-Care Molecular COVID-19 Surveillance in a Rural Workplace: Nonrandomized Controlled Clinical Trial Among Miners

Akshay Sood[1], MPH, MD; William "Cotton" Jarrell[2], CMSP; Xin W Shore[3], MSc; Nestor Sosa[3], MD; Alisha Parada[3], MD; Nicholas Edwardson[4], PhD; Alexandra V Yingling[3], PhD; Teah Amirkabirian[3], BSc; Qiuying Cheng[3], PhD; Ivy Hurwitz[3], PhD; Linda S Cook[5], PhD; Shuguang Leng[3], PhD; Orrin B Myers[3], PhD; Douglas J Perkins[6], PhD

[1]Department of Internal Medicine, School of Medicine, University of New Mexico Health Sciences Center, 1 University of New Mexico MSC 10 5550, Albuquerque, NM, United States

[2]Miners Colfax Medical Center, Raton, NM, United States

[3]University of New Mexico Health Sciences Center, Albuquerque, NM, United States

[4]School of Public Administration, University of New Mexico, Albuquerque, NM, United States

[5]Department of Epidemiology, University of Colorado School of Public Health, Aurora, CO, United States

[6]Center for Global Health, University of New Mexico Health Sciences Center, Albuquerque, NM, United States

Corresponding Author:
Akshay Sood, MPH, MD
Department of Internal Medicine, School of Medicine, University of New Mexico Health Sciences Center, 1 University of New Mexico MSC 10 5550, Albuquerque, NM, United States

Abstract

Background: Numerous studies have assessed the risk of SARS-CoV-2 exposure and infection among health care workers during the pandemic. However, far fewer studies have investigated the impact of SARS-CoV-2 on essential workers in other sectors. Moreover, guidance for maintaining a safely operating workplace in sectors outside of health care remains limited. Workplace surveillance has been recommended by the Centers for Disease Control and Prevention, but few studies have examined the feasibility or effectiveness of this approach.

Objective: The objective of this study was to investigate the feasibility and effectiveness of using frequent point-of-care molecular workplace surveillance as an intervention strategy to prevent the spread of SARS-CoV-2 at essential rural workplaces (mining sites) where physical distancing, remote work, and flexible schedules are not possible.

Methods: In this nonrandomized controlled clinical trial conducted from February 2021, to March 2022, 169 miners in New Mexico (intervention cohort) and 61 miners in Wyoming (control cohort) were enrolled. Investigators performed point-of-care rapid antigen testing on midnasal swabs (NSs) self-collected by intervention miners. Our first outcome was the intervention acceptance rate in the intervention cohort. Our second outcome was the rate of cumulative postbaseline seropositivity to SARS-CoV-2 nucleocapsid protein, which was analyzed in the intervention cohort and compared to the control cohort between baseline and 12 months. The diagnostic accuracy of detecting SARS-CoV-2 using rapid antigen testing on NSs was compared to laboratory-based reverse transcriptase polymerase chain reaction (RT-PCR) on nasopharyngeal swabs (NPSs) in a subset of 68 samples.

Results: Our intervention had a mean acceptance rate of 96.4% (11,413/11,842). The intervention miners exhibited a lower cumulative postbaseline incident seropositivity at 12 months compared to control miners (14/97, 14% vs 17/45, 38%; *P*=.002). Analysis of SARS-CoV-2 antigen detection in self-administered NSs revealed 100% sensitivity and specificity compared to laboratory-based RT-PCR testing on NPSs.

Conclusions: Our findings establish frequent point-of-care molecular workplace COVID-19 surveillance as a feasible option for keeping essential rural workplaces open and preventing SARS-CoV-2 spread. These findings extend beyond this study, providing valuable insights for designing interventions to maintain employees' safety at other essential workplaces during an infectious disease outbreak.

Trial Registration: ClinicalTrials.gov NCT04977050; https://clinicaltrials.gov/study/NCT04977050

(JMIR Public Health Surveill 2025;11:e59845) doi:10.2196/59845

XSL•FO
RenderX

KEYWORDS

point-of-care; seroprevalence; SARS-CoV-2; coronavirus; COVID-19; surveillance; rural workplace; miners; infectious disease; pandemic; antigen testing; midnasal swabs; public health

Introduction

Thus far, an estimated 55 million US workers have provided essential services during the COVID-19 pandemic [1]. While the significant occupational risk of SARS-CoV-2 infection faced by essential health care workers has been well-documented [2-5], far fewer studies have investigated SARS-CoV-2 infection rates in other essential industries [6-10]. Moreover, guidance for maintaining a safely operating workplace in sectors outside of health care remains limited. Workplace surveillance has been recommended by the Centers for Disease Control and Prevention (CDC) [11], but few studies have examined the feasibility or effectiveness of this approach [12-14]. The lack of adequate data in this regard constitutes a critical gap in the literature that needs to be addressed to help devise pandemic intervention strategies for workplaces outside of health care.

Laboratory-based nucleic acid amplification tests, such as reverse transcriptase polymerase chain reaction (RT-PCR) assays performed on nasopharyngeal swabs (NPS) collected by trained professionals, are considered the gold standard for SARS-CoV-2 diagnostic testing based on accuracy, sensitivity, and specificity. However, essential workers outside the health care sector, particularly in rural and remote minority communities, may encounter issues with access, cost, and wait time for results associated with this gold standard testing. Consequently, inconsistent or absent gold standard testing has placed rural essential workers outside of health care at an increased risk for SARS-CoV-2 infection and associated complications [15].

One strategy to overcome challenges with gold standard testing is to use self-collected nasal swabs (NS) and point-of-care qualitative SARS-CoV-2 rapid antigen testing (QAT). QAT offers several advantages, such as low cost, enhanced testing accessibility, and rapid turnaround time, making it well-suited for resource-limited settings. Moreover, the use of self-collected NSs is comparable to clinician-collected NPSs for detecting SARS-CoV-2 in controlled settings [16,17]. In this study, we evaluate the diagnostic accuracy, reliability, feasibility, and effectiveness of frequent point-of-care molecular workplace COVID-19 surveillance in a diverse population of coal miners in a remote community in New Mexico, compared to a cohort of coal miners from a rural mining site in Campbell County, Wyoming. The novelty of the research question and the involvement of a diverse, understudied, and underserved population contribute to the innovation of the study. The long-term goal of the study was to inform strategies for future pandemic mitigation efforts in mining settings and other essential rural workplaces.

Methods

Study Design

To assess the effectiveness of implementing point-of-care QAT for SARS-CoV-2 on self-collected NS in rural workplaces, we conducted a nonrandomized controlled clinical trial involving two cohorts of coal miners from a mining site in McKinley County, New Mexico (intervention site) and from a mining site in Campbell County, Wyoming (control site). The primary recruitment period was between February 2021 and March 2022 with additional miners entering the study after this time. Inclusion criteria were as follows: male or female miners currently employed at the above-mentioned mines, at least 18 years of age, willing and able to consent to study participation, and willing and able to comply with study procedures. Exclusion criteria were as follows: unable or unwilling to consent and less than 18 years of age.

Both intervention and control sites were surface coal mines in rural and remote mountainous locations operated by the same company with similar working conditions and policies. They shared similar engineering and administrative controls and had similar protective measures on screening for SARS-CoV-2 symptoms and fever and for using face masks on-site. Workplace policy at both sites precluded miners with COVID-19 symptoms from reporting for duty, and symptom screening was conducted by mine safety personnel at the beginning of each work shift before miners could enter the mines.

Trained miner investigators collected information using phone- and computer-based web applications developed by Ingenuity Software Labs (Albuquerque, NM), paper-based questionnaires, and body temperature measurements using a no-touch forehead infrared thermometer. The information collected included a 24-hour history of SARS-CoV-2 symptoms, any provider diagnosis of infection, potential close contact with infected individuals, vaccination status, and use of cloth face covering in public settings outside the mine. Data were transferred manually by the miner investigators to the university-based investigators, who entered the information into a secure REDCap database. All information entered was double-checked for accuracy.

Miner Testing

Qualitative SARS-CoV-2 Rapid Antigen Testing

For rapid detection of the SARS-CoV-2 antigen, eligible intervention miners were provided with an NS at the mine entrance at the beginning of each alternate work shift, in a schedule which consisted of 4 days on and 4 days off, resulting in 2 NSs being collected in every 8-day cycle. NSs were self-collected by the miner subjects, with trained miner investigators supervising the process. Subsequently, the 10-minute Quidel Quickvue SARS-CoV-2 QAT (ie, intervention test; Quidel Corporation, San Diego, CA) was conducted on the collected swabs by trained miner investigators at a temporary

XSL•FO
RenderX

laboratory created at the intervention mine entrance. Test results were available in about 10 minutes, roughly the time required to drive from the mine entrance to the parking lot. Based on the CDC guidelines, those who tested positive were requested to return home, isolate themselves, and contact their primary care providers [11].

Seropositivity Testing

Venipuncture blood samples of 8 mL were obtained from intervention and control miner subjects at the time of enrollment and at 3, 6, and 12 months. Whole blood was centrifuged at the mine site, and plasma was aspirated, aliquoted, and frozen by trained nurse technicians. Plasma samples were tested to determine the seroprevalence of COVID-19 using the immunoglobulin G antibody to the nucleocapsid protein of SARS-CoV-2, and tests were performed by a qualitative chemiluminescent immunoassay (ARUP Laboratory, Salt Lake City, UT), under the Emergency Use Authorization. Unlike the antibodies to spike protein, antibodies to nucleocapsid protein develop in response to natural infection and not due to COVID-19 vaccination.

Diagnostic Accuracy of NSs Versus NPSs

To determine whether self-collected NSs were comparable to NPSs collected by health care providers for detecting SARS-CoV-2, a subset of miners were requested to perform a repeat self-administered NS on both nostrils, while 2 NPSs were administered by a skilled health care provider from separate nostrils, all in one setting. All swabs were placed separately into appropriately labeled tubes with viral transport media for transportation to the laboratories. The first NPS sample was used for running the confirmatory qualitative SARS-CoV-2 RT-PCR test at TriCore Diagnostic Laboratory (Albuquerque, NM)—the gold standard—performed under Emergency Use Authorization from the Food and Drug Administration. Sensitivity, specificity, and predictive values were calculated for the intervention test compared to the gold standard test.

The NS and the second NPS were used for quantitative RT-PCR (RT-qPCR) tests. Aliquots of viral transport media were inactivated with DNA/RNA Shield (Zymo Research, CA) and stored at –80°C until bulk processing using the Quick-RNA Viral kit (Zymo Research). RT-qPCR assays were performed using the N1 and RP primer/probe sets from the CDC 2019-nCoV-01 diagnostic panel. No target controls were included in all amplification reactions. The N1 primer/probe set is designed to specifically detect and amplify a region within the nucleocapsid gene of SARS-CoV-2, while the RP primer/probe set targets a portion of the human RNAse P (RNP) gene. The RP primer/probe set is used in all clinical samples to assess specimen quality. Samples with a cycle threshold (Ct) values ≥40 were considered negative. For viral load quantification, standard curves were generated from the Ct values of N1 and RP using known concentrations of SARS-CoV-2 RNA and plasmids harboring the RNP gene fragment, respectively. Viral load per sample was determined as described by Perkins et al [18]. The Cohen κ coefficient was used to determine the intrarater agreement between the NS and NPS sample collection methodologies.

Statistical Analysis

Miner Characteristics

Characteristics of miners participating at the intervention and control mines were summarized as frequencies and percentages and means and SDs. Frequencies of categorical variables for the two study arms were compared using χ^2 tests and Fisher exact tests. Continuous age was compared using a Wilcoxon rank sum test.

Intervention Acceptance

Intervention test acceptance by miners was measured for each eligible work shift, and the outcome variable and cumulative acceptance rate were calculated over the study time frame (frequency of tests accepted/number of tests offered). Predictors influencing the outcome were examined in the intervention cohort, including self-reported educational status, racial and ethnic minority status, prior SARS-CoV-2 infection or vaccination, and evidence of previous infection (from a prior positive test). Nonparametric Wilcoxon and Kruskal-Wallis tests were used to assess the association of miner characteristics with cumulative acceptance rates, since acceptance rates had a skewed distribution. We also computed the total number of tests accepted per miner for each week of the study (total number of tests accepted in each week/number of consented miners). Weekly tests per miner were summarized graphically, and a test for trend was made assuming residual errors had a first-order autocorrelation.

Seropositivity Analysis

Seroprevalence patterns of miners enrolled and tested between February 23 and March 6, 2021 were compared between intervention and control miners over time intervals using χ^2 tests and time-to-first seropositivity by Kaplan-Meier analysis and log-rank tests. Kaplan-Meier analyses accounted for dropouts during follow-up after baseline. We used logistic regression to compare the following binary outcome variables: baseline (prevalence) seropositivity, postbaseline incident seropositivity, and combined cumulative seropositivity (as positive at any time during the study). Combined cumulative seropositivity analyses also included miners that were enrolled after March 6, 2021. Vaccination status and select host susceptibility and vulnerability factors were added to multivariable logistic regression models. Odds ratios (ORs) and 95% CIs were computed to describe association strength. SAS version 9.4 (SAS Institute) and SPSS version 28.0.0.0 (IBM Corp) were used for statistical analyses, and a *P* value of less than .05 was considered statistically significant.

Power Calculation

The study was planned to estimate nasal swab acceptance rates with precision of ±4.6% from a sample of up to 250 intervention group miners, which also would have at least 80% power for subgroup analyses. A comparison group with up to n=350 was estimated to have at least 80% power to detect group differences in baseline seropositivity and incident seropositivity. However, sample sizes recruited were lower than planned.

XSL•FO
RenderX

Ethical Considerations

The University of New Mexico Health Sciences Center's Human Research Protections Office and Institutional Review Board (HRPO 20 - 680) approved the study. All participants provided informed consent with an ability to opt out. The data were deidentified for analysis. Participants were compensated with a $10 electronic merchandise card for each specimen and survey collected.

Results

Study Design: Cohort Demographics

The study included 115 intervention miners and 60 miners recruited during February and March 2021 and another 54 intervention miners and 1 control miner recruited during the rest of the study (n=169 intervention and n=61 control miners). Baseline characteristics of subjects in the two cohorts are shown in Table 1. The two cohorts were similar in age distribution, with a mean age of 44.4 (SD 11.2) and 45.2 (SD 11.9) years in the intervention and control miner groups, respectively. When comparing the distribution of comorbidities, including asthma, chronic obstructive pulmonary disease, chronic lung diseases, hypertension, diabetes, and depression, 43.8% (74/169) of intervention miners reported at least 1 comorbidity, similar to 46% (28/61) of control miners.

XSL•FO
RenderX

Table . Comparison of baseline characteristics between New Mexico–based (intervention cohort) and Wyoming-based (control cohort) participants.

Characteristic	All (n=230)	New Mexico (n=169)	Wyoming (n=61)	*P* value
Age (years), mean (SD)	44.6 (11.3)	44.4 (11.2)	45.2 (12)	.57
Age group (years), n (%)				.34
<40	93 (40.4)	73 (43.2)	20 (33)	
40 - 50	57 (24.8)	41 (24.3)	16 (26)	
>50	80 (34.8)	55 (32.5)	25 (41)	
Education >12th grade, n (%)	121 (52.6)	69 (40.9)	52 (85)	<.001
Annual household income ≥$80,000, n (%)	111 (48.2)	70 (41.4)	41 (67)	.005
Race, n (%)				<.001
White	152 (66.1)	96 (56.8)	56 (92)	
American Indian or Alaska Native	29 (12.6)	27 (16)	2 (3)	
Asian	1 (0.4)	1 (0.6)	0 (0)	
Other race	34 (14.8)	31 18.3)	3 (5)	
Prefer not to answer or missing	14 (6.1)	14 (8.3)	0 (0)	
Ethnicity, n (%)				<.001
Not Hispanic, Latino, or Spanish origin	117 (50.9)	63 (37.3)	54 (88)	
Hispanic, Latino, or Spanish origin	101 (43.9)	97 (57.4)	4 (7)	
Prefer not to answer or missing	12 (5.2)	9 (5.3)	3 (5)	
Biological sex at birth, n (%)				.02
Male	202 (87.8)	154 (91.1)	48 (79)	
Female	25 (10.9)	14 (8.3)	11 (18)	
Intersex, prefer not to answer or missing	3 (1.2)	1 (0.6)	2 (3)	
COVID-19-related behavior and vaccination characteristics, n (%)				
Close contact with confirmed or suspected COVID-19 since the last shift	19 (8.3)	13 (7.7)	6 (10)	.76
Frequency of ride-sharing to and from mine site often or very often	81 (35.2)	75 (44.4)	6 (10)	<.001
Frequency of face covering outside the mine often or very often	162 (70.4)	142 (84.1)	20 (33)	<.001
COVID-19 vaccination (yes)	52 (22.6)	43 (25.4)	9 (15)	.07
Comorbidities, n (%)				
Previously tested positive for COVID-19	42 (18.3)	24 (14.2)	18 (30)	.08
Asthma, COPD[a], and other chronic lung diseases	30 (13.1)	23 (13.6)	7 (12)	.67
Hypertension	47 (20.4)	34 (20.1)	13 (21)	.87

XSL•FO
RenderX

Characteristic	All (n=230)	New Mexico (n=169)	Wyoming (n=61)	*P* value
Diabetes	21 (9.1)	17 (10.1)	4 (7)	.41
Depression	11 (4.8)	7 (4.1)	4 (7)	.46
Household wood smoke exposure (ever), n (%)	62 (26.9)	51 (30.3)	11 (18)	.08

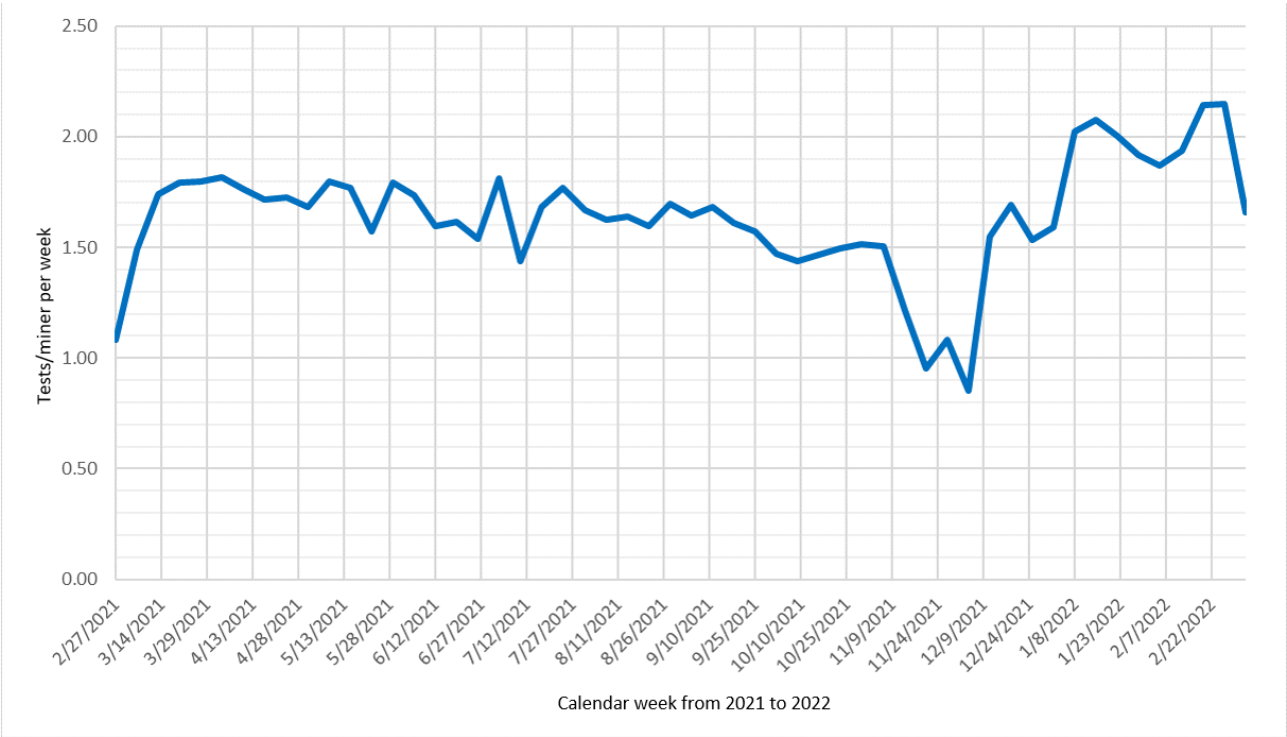
[a]COPD: chronic obstructive pulmonary disease.

A significant difference was observed between the cohorts regarding racial and ethnic composition, with 72.8% (123/169) of intervention miners being from underrepresented racial and ethnic minorities (including American Indians and Hispanics), compared to 10% (6/61) of control miners (*P*<.001). The intervention cohort consisted of more male participants than the control cohort (154/169, 91% vs 48/61, 79%; *P*=.02). Furthermore, intervention miners had a lower annual income, with 58.6% (99/169) of intervention miners earning less than $80,000 annually versus 33% (20/61) of control miners (*P*=.005) and a lower education level, with 40.9% (69/169) of intervention miners having postsecondary education versus 85% (52/61) of control miners (*P*<.001). In terms of COVID-19–related behavior and vaccination characteristics, intervention miners were more likely to ride-share to and from the mining site (75/169, 44.4% vs 6/61, 10%; *P*<.001) but were also more likely to wear face coverings outside of the mine often or very often (142/169, 84.1% vs 30/61, 33%; *P*<.001). Moreover, intervention miners had a 25.4% (43/169) baseline vaccination rate against COVID-19 versus 15% (9/61) for control miners (*P*=.07). The intervention miners may have been more likely to have been exposed to household wood smoke than control miners (51/169, 30% vs 11/61, 18%; *P*=.08).

Intervention Acceptance

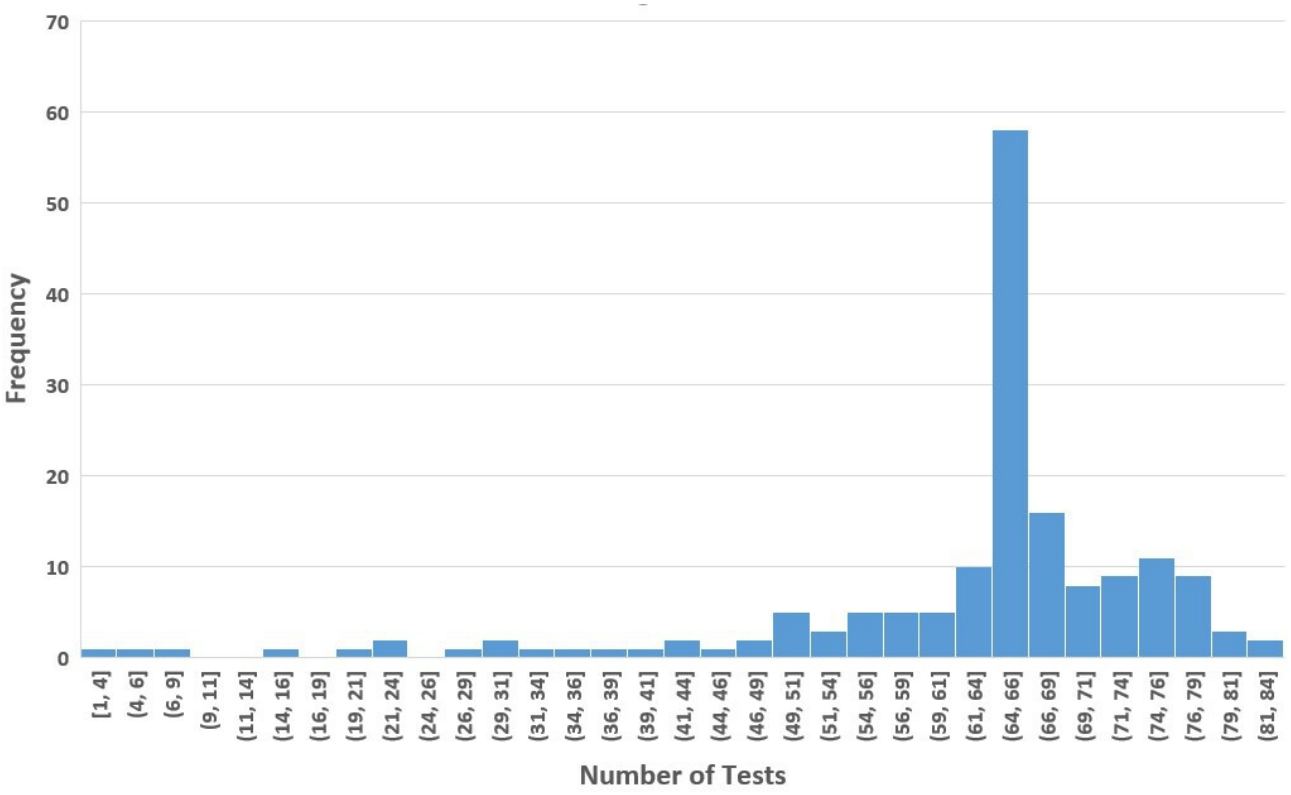
The mean intervention test acceptance rate was 96.4% (11,413/11,842). The high test acceptance was unaffected by demographic or medical history characteristics or baseline COVID-19 vaccine uptake (*P*>.05 for all, data not shown). The number of tests per miner per week (mean 1.64, SD 0.26) did not show a trend over the study (Figure 1; slope=0.004; SE=0.005; *P*=.46; temporal autocorrelation=–0.73). The total number of testing opportunities varied by the individual, with an average of 84 opportunities per miner over 267 study days. Participants voluntarily performed an average of 62.0 (SD 14.4) tests, ranging from 1 to 84 tests (Figure 2).

Figure 1. Number of SARS-CoV-2 antigen tests performed per miner per week. Intervention tests were performed at each alternate work shift (schedule 4 days on and 4 days off) during the 2021 calendar weeks 8-52 and 2022 calendar weeks 1-9 (ie, from the end of February 2021, to the end of February 2022; n=169 miners at the intervention site).



XSL•FO
RenderX

Figure 2. Distribution of the number of SARS-CoV-2 antigen tests per miner. There were approximately 84 opportunities over 267 study days. Participants performed an average of 62.0 (SD 14.4) tests and a median of 65 tests (IQR 68.25-61.75), ranging from 1 to 84 tests (n=169 miners at the intervention site).



Diagnostic Accuracy of the Intervention Test Compared to RT-PCR Gold Standard Testing

Sensitivity, specificity, and predictive values of the intervention (QAT on self-collected NSs at the mine site analyzed by trained miner investigators in the temporary field laboratory) were established through comparison with the gold standard test (ie, qualitative RT-PCR conducted on an NPS collected on the same day by trained health personnel and analyzed in a laboratory). Compared to gold standard testing on a sample of 68 pairs of tests, the intervention test in the field setting showed 100% sensitivity and 100% specificity (Tables 2 and 3).

Table . Comparison of test characteristics between the qualitative rapid antigen test on midnasal swabs in the field and the gold standard test in a laboratory, in a sample of 68 pairs of tests.

Test result	Gold standard test (qualitative RT-PCR[a] on NPSs[b])		
	Disease (positive)	Nondisease (negative)	Total
Antigen test on NS[c]			
Positive	38	0	38
Negative	0	30	30
Total	38	30	68

[a]RT-PCR:reverse transcriptase polymerase chain reaction.

[b]NPS: nasopharyngeal swab.

[c]NS: nasal swab.

XSL•FO
RenderX

Table . Sensitivity, specificity, and predictive values for the qualitative rapid antigen test on midnasal swabs in the field, compared to the gold standard test in a laboratory, in a sample of 68 pairs of tests.

Parameters	Equations	Calculated sensitivity (95% CI)
Sensitivity (%)	True positive/(true positive + false negative)	100 (91-100)
Specificity (%)	True negative/(true negative + false positive)	100 (88-100)
Positive predictive value (%)	True positive/(true positive + false positive)	100[a]
Negative predictive value (%)	true negative/(true negative + false negative)	100[a]

[a]95% CI was not calculated for these values.

Diagnostic Accuracy of Self-Collected NS Versus NPS Collected by Health Personnel Using RT-qPCR

For this analysis, 46 paired samples (NS and NPS) were collected. Among these swabs, the human RNP gene was not detected in 4 samples: 3 from self-collected samples and 1 NPS collected from a health care professional following RT-qPCR conducted in the Center for Global Health Perkins laboratories. These samples were excluded from further analyses. The RT-qPCR results for the 42 remaining samples are shown in Table 4.

Table . Diagnostic performance of nasal swab versus nasopharyngeal swab using RT-qPCR[a] for SARS-CoV-2 (42 matched samples). Samples with a cycle threshold value of ≥40 were considered negative.

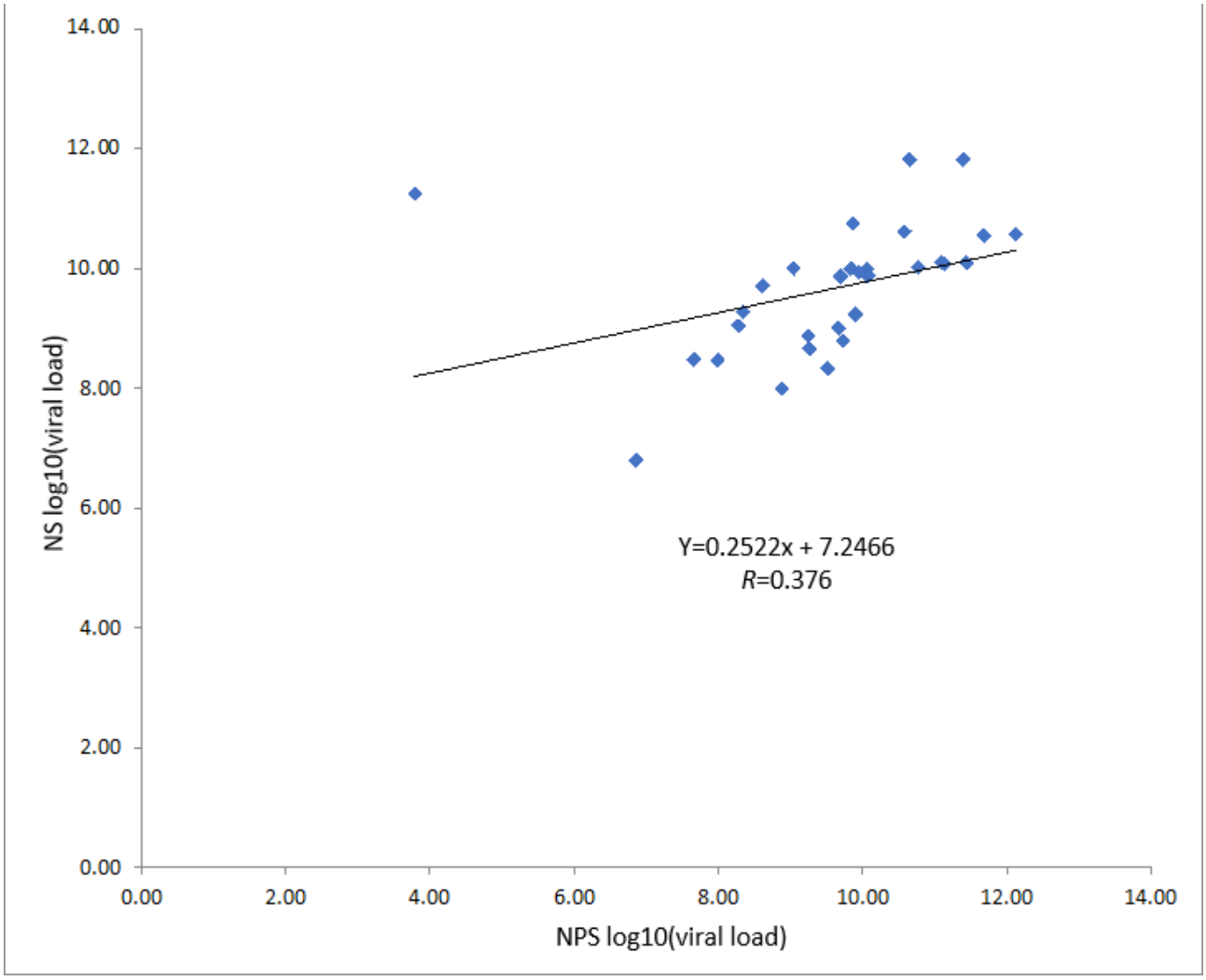
Test result	Nasopharyngeal swab		Concordance rate (%)	Kappa coefficient (95% CI)
	Negative	Positive		
Nasal swab			80.9	0.39 (0.04 - 0.73)
Negative	4	5		
Positive	3	30		

[a]RT-qPCR: quantitative reverse transcriptase polymerase chain reaction.

The concordance rate was 80.9% between the two sampling methods with a κ coefficient of 0.39 (95% CI 0.04 - 0.73; *P*=.01). Simple linear regression was performed to further investigate the relationship between viral loads and the two sample collection methodologies (Figure 3). A modest correlation (*R*=0.38; *P*=.04) for SARS-CoV-2 viral load was detected between 30 paired NP and NS samples. The Ct values and computed viral loads for the discordant samples are shown in Table 5. The NPS virus nucleoprotein gene N1 Ct values for the discordant samples were ≥35, which, after normalizing to the respective RNP sample [18], corresponded to viral loads between log 3 to log 6.5 copies of viral RNA per 1000 cells. This magnitude of viral load was not detectable in self-collected NS samples, suggesting that an NPS was better than an NS for detecting lower viral loads.

XSL•FO
RenderX

Figure 3. Modest correlation between SARS-CoV-2 viral load from nasopharyngeal swabs and nasal swabs. A linear regression model was performed from nasopharyngeal swab and nasal swab paired samples SARS-CoV-2 viral load (log10) to examine the relationship between viral loads and the two sample collection methodologies (n=30; y=0.2522x+7.2466; *R*=0.376). NPS: nasopharyngeal swab; NS: nasal swab.



XSL•FO
RenderX

Table . Cycle threshold values and corresponding calculated viral loads of discordant self-collected nasal swab and nasopharyngeal samples collected by trained health care providers using RT-qPCR[a].

Sample ID	NS[b]		NPS[c]		Corresponding viral load (log_{10})	
	N1 Ct[d]	RP[e] Ct	N1[f] Ct	RP Ct	NS	NP
221	>40	37.17	36.82	36.59	0[g]	7.31
263	>40	38.27	37.14	33.96	0	6.17
398	>40	30.85	38.43	25.07	0	3.09
420	>40	34.15	37.07	37.47	0	6.26
738	>40	35.64	37.14	33.01	0	5.78
749	38.52	31.72	>40	33.27	5.74	0
759	35.95	26.96	>40	24.58	4.42	0
967	28.59	31.39	>40	32.15	8.43	0

[a]RT-PCR: reverse transcriptase polymerase chain reaction.

[b]NS: nasal swab.

[c]NPS: nasopharyngeal swab.

[d]Ct: cycle threshold.

[e]RP: human RNase P gene assay control.

[f]N1: virus nucleoprotein gene N1 assay.

[g]Cycle threshold values ≥40 were considered negative.

Seropositivity Testing

Seropositivity patterns of miners enrolled and tested during the initial recruitment period (baseline, February to March 2021) were compared between intervention and control miners using χ^2 tests for prevalent seropositivity and cumulative postbaseline incidence during the study. At baseline, prevalent seropositivity was seen among 18 out of 115 (15.7%) intervention miners and 15 out of 60 (25%) control miners (OR 0.56, 95% CI 0.26-1.19; *P*=.13), and cumulative postbaseline incidence was found in 14 out of 97 (14%) intervention miners compared to 17 out of 45 (38%) control miners (OR 0.28, 95% CI 0.12-0.63; *P*=.002; Table 6). We also used Kaplan–Meier analysis and log-rank tests, which accounted for dropouts, to visualize and analyze whether seropositivity patterns were different for intervention and control miners. Figure 4 shows that seropositivity was greater in control miners than in intervention miners at all time periods (log-rank test *P*=.02). The 12-month cumulative incidence was 48.2% (95% CI 35.4-62.9) among intervention miners and 65.3% (95% CI 51.1-79.2) among control miners when estimated by Kaplan-Meier analysis. When only postbaseline incidence was analyzed, the log-rank test had a lower power and was not significant (*P*=.09) with a 12-month seropositivity of the intervention group of 38.6% (95% CI 24.9-56.4) and 53.8% (95% CI 37.2-72.1) for the control group.

Table . Prevalent and incident seropositivity patterns between the intervention and control miners enrolled and tested during the initial recruitment period (baseline, February to March 2021).

Study timepoint	Intervention (New Mexico–based) mine		Control (Wyoming-based) mine (n=35)	
	New seropositivity, n	Cumulative seropositivity, n/N (%)	New seropositivity, n	Cumulative seropositivity, n/N (%)
Prevalent seropositivity (baseline)	18	18/115 (15.7)	15	15/60 (25)
Postbaseline incident seropositivity[a]				
3 mo	0	0/97 (0)	0	0/45 (0)
6 mo	1	1/97 (1)	6	6/45 (13)
12 mo	13	14/97 (14)	11	17/45 (38)

[a]Postbaseline incident seropositivity was calculated using the total sample size minus the baseline sample size.

XSL•FO
RenderX

Figure 4. Cumulative incidence of SARS-CoV-2 infection in miners in intervention and control mine sites. Data shown are the cumulative incidences of seropositivity (1 – survival estimated by Kaplan-Meier analysis) for New Mexico–based miners from the intervention mine (solid blue line) and Wyoming-based miners from the control mine (dashed green line). Plus symbols indicate miners that were not followed in subsequent testing periods or were negative at the end of follow-up. Numbers at the bottom of the graph are the number of miners in the analysis at each time. NM: New Mexico; S: survival; WY: Wyoming.

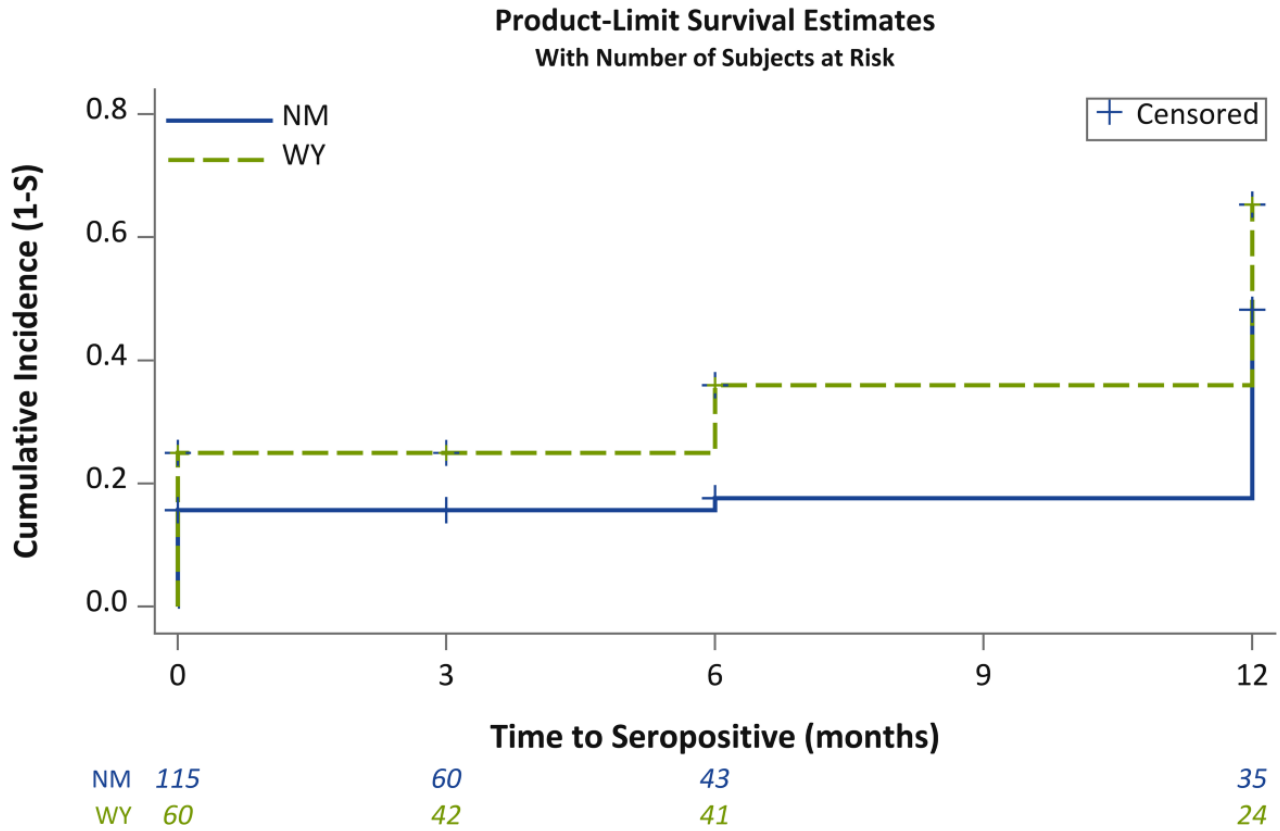


Table S1 in Multimedia Appendix 1 shows unadjusted logistic regression analyses for prevalent and postbaseline incidence, as well as the combined cumulative seropositivity. Combined cumulative seropositivity analyses included miners enrolled at any time during the study. When testing differences in the incidence of seropositivity after baseline, intervention miners (14/97, 14%) had a lower incidence compared to control miners (17/45, 38%; OR 0.28, 95% CI 0.12-0.64; *P*=.002). Incident seropositivity was lower in those who self-reported using masks very often (6/67, 9%) or often (11/39, 28%) compared to those using masks less often (13/35, 37%; *P*=.004; Table S1 in Multimedia Appendix 1). A previous positive COVID-19 test was associated with higher odds of baseline seropositivity (OR 12.72, 95% CI 5.00-32.31; *P*<.001) but not for incident seropositivity (OR 1.34, 95% CI 0.45-4.15; *P*=.59). Age, race and ethnicity, sex, ride-sharing, and COVID-19 vaccination status were not associated with baseline or incident seropositivity. Seropositivity at any time during the study (combined cumulative seropositivity) was lower among intervention miners (OR 0.23, 95% CI 0.12-0.43; *P*<.001), lower among miners who self-reported wearing masks very often (OR 0.19, 95% CI 0.09-0.38; *P*<.001), and lower among miners vaccinated at baseline (OR 0.41, 95% CI 0.19-0.89; *P*=.03). Combined cumulative seropositivity was lower among miners with any nasal testing, which was confounded by the intervention and baseline seropositivity, and odds of cumulative seropositivity did not show a dose-response relationship, with more tests having a lower OR than fewer tests (*P*>.05, data not shown). We also used multivariable logistic models with covariates for age, race and ethnicity, sex, masking frequency, ride-sharing, and COVID-19 vaccination status to assess factors associated with seropositivity (Table S2 in Multimedia Appendix 1). More frequent masking was significantly associated with lower odds of seropositivity for incident and combined cumulative seropositivity measures. Further, more frequent masking was more common in the intervention group, had a strong protective effect on seroprevalence, and confounded the protective effect of the intervention on incident seropositivity and combined cumulative seropositivity.

Discussion

Principal Results

Workers in essential sectors have experienced higher COVID-19 infection and mortality rates throughout the pandemic [19]. Much attention has been given to the significant occupational risk of infection among health care workers, but essential work extends beyond health care. Understanding the feasibility and effectiveness of implementing point-of-care testing in essential non–health care professions is crucial for preventing or slowing the transmission of infectious diseases.

Miners, deemed essential workers, face unique challenges that make them more vulnerable and susceptible to COVID-19 [20]. Mining environments often have decreased ventilation and limited availability for physical distancing, remote work, or

XSL•FO
RenderX

flexible schedules. Miners are also exposed to higher levels of air pollutant particulates, which may increase their susceptibility to respiratory infections or worsen health outcomes from respiratory infections [21]. Moreover, mine sites are situated in rural areas with limited access to health care and testing facilities. Despite these disproportionate challenges, limited guidelines are available for maintaining mining and other non–health care operations effectively and safely during an infectious disease outbreak. Thus, the objective of this study was to investigate the feasibility and effectiveness of using frequent point-of-care molecular workplace surveillance as an intervention strategy to prevent the spread of SARS-CoV-2 at rural mining sites.

The intervention had a high mean acceptance rate of 96.4% (11,413/11,842), reflecting the involvement of mine safety personnel in the design and execution of the study. Additionally, frequent point-of-care workplace surveillance was associated with a lower cumulative SARS-CoV-2 seropositivity rate in intervention miners compared to control miners (P=.002). This study further examined the reliability of self-collected NS versus NPS samples collected by trained health care personnel to detect SARS-CoV-2 using RT-qPCR. The Cohen κ coefficient of 0.39 indicates fair agreement between the two collection methods. However, we observed that NPS samples that tested positive for SARS-CoV-2 by RT-qPCR but had lower viral loads were frequently associated with negative results when tested using NSs instead. While less reliable when compared to NPSs for detecting low levels of SARS-CoV-2, self-collected NS can provide diagnostic performance comparable to NPSs in participants with higher viral loads. As such, NS is a suitable and practical option for intervention test use, especially in resource-limited settings where point-of-care QAT is more accessible, cost-effective, and faster than the gold standard test.

Limitations

A limitation of this study is that protective behaviors influence COVID-19 health outcomes, including confounding seropositivity rates [22]. While the CDC has encouraged the public use of face coverings during high spikes in community infection rates and vaccination against COVID-19, the use of these protective measures can vary between mine sites based on individual and community factors [23]. We observed that intervention miners were more likely to wear face coverings in public and more willing to receive a vaccine against COVID-19. This difference may reflect partisan variation in official messaging related to COVID-19 protective behaviors to the public between Democratic-governed (New Mexico) and Republican-governed (Wyoming) states [24].

It is also important to note that our study did not involve randomization, so the unmatched distribution of covariates between the two cohorts may result in confounding bias which may have influenced our findings. However, several differences related to demographics and social determinants of health between the mine sites placed the intervention cohort at a higher risk for SARS-CoV-2 infection than the control miners. For example, intervention miners were more likely to be male, which is associated with an increased risk for SARS-CoV-2 infection [25]. Additionally, intervention miners were more likely to be racial and ethnic minorities. Given the disparities observed in minority populations during the pandemic [26,27], this selection bias may have placed the intervention miners at an increased inherent risk for infection and associated complications, which may have biased the protective effect of the intervention toward the null value. Intervention miners were more likely to ride-share and report previous exposure to woodsmoke, a risk factor for increased respiratory disease susceptibility [28]. Moreover, data previously published on our New Mexico cohort indicate that minority miners are at greater risk for developing lung disease than non-Hispanic White miners [29].

Intervention miners also had a lower mean income and educational status than control miners. A study analyzing the joint effects of socioeconomic position, race and ethnicity, and sex on mortality in the first year of the COVID-19 pandemic revealed that COVID-19 mortality was 5 times higher in individuals in low versus high socioeconomic positions [30]. The above-mentioned biases may impact the study's conclusions, making it difficult to determine whether the observed effects are genuinely due to the intervention or other factors. Although randomization is a key method to mitigate these biases and enhance the credibility of clinical research, community investigators did not consider this study design feasible during the pandemic. Challenges encountered by the study included the lack of easily available intervention tests and community research personnel in the field during the early stage of the pandemic, which limited study recruitment and expansion to other mine sites. The recruitment of a sample size that was lower than planned limited the power of the study. Additionally, the remote location of the study sites increased study costs, and the lack of cellphone and internet connectivity required adaptive strategies by the investigators. Although our intervention can be replicated in other geographic areas and workers in other public and private essential service settings, scaling the intervention would require careful consideration of these unique challenges and a study of its cost-effectiveness.

Comparison With Prior Work

Since the onset of the pandemic, several papers have been published supporting routine testing to limit the spread of SARS-CoV-2 in health care settings [31-33]. However, studies evaluating frequent point-of-care surveillance, particularly in sectors outside of health care, are still limited. One study, which included a variety of essential workers, including health care workers, found that regular testing of all key workers was associated with reduced transmission of approximately 67 individuals per 1000 tests, with high accuracy of testing (87.1% - 99.9%) [34]. In this study, researchers used RT-PCR rather than QAT, likely due to the limited availability of these tests at the study's onset. However, a later study evaluating the sensitivity and specificity of the Boson Rapid SARS-CoV-2 rapid antigen detection test revealed an overall sensitivity of 63.04% for anterior NSs and 73.33% for NPSs, with 100% test specificity [35]. A similar study, also showing 100% specificity, compared the Medomics SARS-CoV-2 antigen test device to gold standard testing and reported that rapid antigen detection tests using self-collected anterior NSs proved to be as sensitive as and more tolerable than professionally collected NPSs for Ct values up to 30, determined by RT-PCR [36]. Supporting

XSL•FO
RenderX

these results and our own, additional studies have emerged revealing that NS self-sampling yields comparable results to NPS sampling using both RT-PCR [37,38] and QAT for analysis [39]. Studies have established the effectiveness of other field interventions, such as PCR-based wastewater surveillance of SARS-CoV-2 [40]. However, such samples are difficult to collect and analyze in rural areas and provide community-level and not individual-level surveillance data, limiting their impact.

Conclusion

Our data firmly establish SARS-CoV-2 QAT on self-collected NSs as a feasible alternative to laboratory-based RT-PCR on NPSs for preventing the spread of COVID-19 at essential workplaces within vulnerable communities at an increased risk of being adversely affected by the COVID-19 pandemic because of their occupation, demographics, and rural and remote locations. Additionally, our findings demonstrate the excellent performance of the intervention test in a real-world setting compared to the gold standard. The study findings support developing and implementing policy measures for workplace surveillance and other protective interventions against the spread of SARS-CoV-2 among rural non–health care essential workers. The implications of our findings extend beyond this study, providing valuable insights for designing interventions to maintain employees' safety at other essential workplaces during an infectious disease outbreak of respiratory pathogens.

Acknowledgments

The research reported in this publication was supported in part by the National Institutes of Health (NIH) under Award Number 3U01GM132175. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Authors' Contributions

AS, WCJ, NS, AP, NE, IH, LSC, SL, OBM, and DJP made substantial contributions to the conception or design of the work. AS, WCJ, AVY, TA, QC, IH, and DJP contributed to data acquisition, and AS, XWS, OBM, QC, IH, and DJP contributed to data analysis. AS, WCJ, XWS, NS, AP, NE, IH, LSC, SL, OBM, and DJP contributed to the interpretation of the data. All authors contributed to drafting the work or reviewing it critically for important intellectual content and gave final approval of the version to be published. All authors agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. We would also like to thank Hesch Consulting LLC for their support during the preparation of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Univariate and multivariable analysis of the association between prevalence seropositivity, postbaseline incident seropositivity, combined cumulative seropositivity, and miner characteristics at baseline.
[DOCX File, 52 KB - publichealth_v11i1e59845_app1.docx]

References

1. Current population survey extracts. Economic Policy Institute. 2020. URL: https://microdata.epi.org [accessed 2023-08-06]
2. Dzinamarira T, Murewanhema G, Mhango M, et al. COVID-19 prevalence among healthcare workers. A systematic review and meta-analysis. Int J Environ Res Public Health 2021 Dec 23;19(1):146. [doi: 10.3390/ijerph19010146] [Medline: 35010412]
3. Gómez-Ochoa SA, Franco OH, Rojas LZ, et al. COVID-19 in health-care workers: a living systematic review and meta-analysis of prevalence, risk factors, clinical characteristics, and outcomes. Am J Epidemiol 2021 Jan 4;190(1):161-175. [doi: 10.1093/aje/kwaa191] [Medline: 32870978]
4. Manai MV, Shaholli D, La Torre G. Contact tracing as an essential prevention tool for the spreading of COVID-19 among healthcare workers. Clin Ter 2022;173(5):396-397. [doi: 10.7417/CT.2022.2452] [Medline: 36155722]
5. Al-Oraibi A, Naidu JS, Chaka A, et al. Prevalence of long COVID-19 among healthcare workers: a systematic review and meta-analysis protocol. BMJ Open 2022 Dec 6;12(12):e065234. [doi: 10.1136/bmjopen-2022-065234] [Medline: 36600349]
6. Song H, McKenna R, Chen AT, David G, Smith-McLallen A. The impact of the non-essential business closure policy on Covid-19 infection rates. Int J Health Econ Manag 2021 Dec;21(4):387-426. [doi: 10.1007/s10754-021-09302-9] [Medline: 33792808]
7. Naleway AL, Grant L, Caban-Martinez AJ, et al. Incidence of SARS-CoV-2 infection among COVID-19 vaccinated and unvaccinated healthcare personnel, first responders, and other essential and frontline workers: eight US locations, January-September 2021. Influenza Other Respir Viruses 2022 May;16(3):585-593. [doi: 10.1111/irv.12956] [Medline: 35023288]

XSL•FO
RenderX

8. Liu Y, Yang B, Liu L, Jilili M, Yang A. Occupational characteristics in the outbreak of the COVID-19 delta variant in Nanjing, China: rethinking the occupational health and safety vulnerability of essential workers. Int J Environ Res Public Health 2021 Oct 13;18(20):10734. [doi: 10.3390/ijerph182010734] [Medline: 34682477]
9. Lutrick K, Ellingson KD, Baccam Z, et al. COVID-19 infection, reinfection, and vaccine effectiveness in a prospective cohort of Arizona frontline/essential workers: the AZ HEROES research protocol. JMIR Res Protoc 2021 May 26;10(6):e28925. [doi: 10.2196/28925] [Medline: 34057904]
10. Alsharef A, Banerjee S, Uddin SMJ, Albert A, Jaselskis E. Early impacts of the COVID-19 pandemic on the United States construction industry. Int J Environ Res Public Health 2021 Feb 6;18(4):1559. [doi: 10.3390/ijerph18041559] [Medline: 33562127]
11. Guidance for SARS-Cov-2 rapid testing performed in point-of-care settings. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/covid/php/lab/point-of-care-testing.html [accessed 2024-01-17]
12. Larremore DB, Wilder B, Lester E, et al. Test sensitivity is secondary to frequency and turnaround time for COVID-19 surveillance. medRxiv 2020 Sep 8:2020.06.22.20136309. [doi: 10.1101/2020.06.22.20136309] [Medline: 32607516]
13. Paltiel AD, Zheng A, Walensky RP. Assessment of SARS-CoV-2 screening strategies to permit the safe reopening of college campuses in the United States. JAMA Netw Open 2020 Jul 1;3(7):e2016818. [doi: 10.1001/jamanetworkopen.2020.16818] [Medline: 32735339]
14. Connor BA, Rogova M, Garcia J, et al. Comparative effectiveness of single vs repeated rapid SARS-CoV-2 antigen testing among asymptomatic individuals in a workplace setting. JAMA Netw Open 2022 Mar 1;5(3):e223073. [doi: 10.1001/jamanetworkopen.2022.3073] [Medline: 35302635]
15. The Lancet. The plight of essential workers during the COVID-19 pandemic. Lancet 2020 May 23;395(10237):1587. [doi: 10.1016/S0140-6736(20)31200-9] [Medline: 32446399]
16. McCulloch DJ, Kim AE, Wilcox NC, et al. Comparison of unsupervised home self-collected midnasal swabs with clinician-collected nasopharyngeal swabs for detection of SARS-CoV-2 infection. JAMA Netw Open 2020 Jul 1;3(7):e2016382. [doi: 10.1001/jamanetworkopen.2020.16382] [Medline: 32697321]
17. Lim HJ, Baek YH, Park MY, et al. Performance analysis of self-collected nasal and oral swabs for detection of SARS-CoV-2. Diagnostics (Basel) 2022 Sep 21;12(10):2279. [doi: 10.3390/diagnostics12102279] [Medline: 36291968]
18. Perkins DJ, Yingling AV, Cheng Q, et al. Elevated SARS-CoV-2 in peripheral blood and increased COVID-19 severity in American Indians/Alaska Natives. Exp Biol Med (Maywood) 2022 Jul;247(14):1253-1263. [doi: 10.1177/15353702221091180] [Medline: 35491994]
19. Chen YH, Riley AR, Duchowny KA, et al. COVID-19 mortality and excess mortality among working-age residents in California, USA, by occupational sector: a longitudinal cohort analysis of mortality surveillance data. Lancet Public Health 2022 Sep;7(9):e744-e753. [doi: 10.1016/S2468-2667(22)00191-8] [Medline: 36057273]
20. Evans K, Lerch S, Boyce TW, et al. An innovative approach to enhancing access to medical screening for miners using a mobile clinic with telemedicine capability. J Health Care Poor Underserved 2016;27(4A):62-72. [doi: 10.1353/hpu.2016.0182] [Medline: 27818414]
21. Laney AS, Weissman DN. Respiratory diseases caused by coal mine dust. J Occup Environ Med 2014 Oct;56 Suppl 10(10):S18-S22. [doi: 10.1097/JOM.0000000000000260] [Medline: 25285970]
22. Green H, Fernandez R, MacPhail C. The social determinants of health and health outcomes among adults during the COVID-19 pandemic: a systematic review. Public Health Nurs 2021 Nov;38(6):942-952. [doi: 10.1111/phn.12959] [Medline: 34403525]
23. Fisher KA, Barile JP, Guerin RJ, et al. Factors associated with cloth face covering use among adults during the COVID-19 pandemic - United States, April and May 2020. MMWR Morb Mortal Wkly Rep 2020 Jul 17;69(28):933-937. [doi: 10.15585/mmwr.mm6928e3] [Medline: 32673303]
24. Golos AM, Hopkins DJ, Bhanot SP, Buttenheim AM. Partisanship, messaging, and the COVID-19 vaccine: evidence from survey experiments. Am J Health Promot 2022 May;36(4):602-611. [doi: 10.1177/08901171211049241] [Medline: 35232232]
25. Bwire GM. Coronavirus: why men are more vulnerable to Covid-19 than women? SN Compr Clin Med 2020;2(7):874-876. [doi: 10.1007/s42399-020-00341-w] [Medline: 32838138]
26. Feldman JM, Bassett MT. Variation in COVID-19 mortality in the US by race and ehnicity and educational attainment. JAMA Netw Open 2021 Nov 1;4(11):e2135967. [doi: 10.1001/jamanetworkopen.2021.35967] [Medline: 34812846]
27. Kakol M, Upson D, Sood A. Susceptibility of southwestern American Indian tribes to coronavirus disease 2019 (COVID-19). J Rural Health 2021 Jan;37(1):197-199. [doi: 10.1111/jrh.12451] [Medline: 32304251]
28. Assad NA, Kapoor V, Sood A. Biomass smoke exposure and chronic lung disease. Curr Opin Pulm Med 2016 Mar;22(2):150-157. [doi: 10.1097/MCP.0000000000000246] [Medline: 26814722]
29. Kocher E, Rendon KJ, Kesler D, et al. Uranium workers demonstrate lower lobe predominant irregular pneumoconiotic opacities on chest radiographs. J Health Care Poor Underserved 2016;27(4A):116-127. [doi: 10.1353/hpu.2016.0193] [Medline: 27818418]
30. Pathak EB, Menard JM, Garcia RB, Salemi JL. Joint effects of socioeconomic position, race/ethnicity, and gender on COVID-19 mortality among working-age adults in the United States. Int J Environ Res Public Health 2022 Apr 30;19(9):5479. [doi: 10.3390/ijerph19095479] [Medline: 35564872]

XSL•FO
RenderX

31. Maya S, Padda G, Close V, et al. Optimal strategies to screen health care workers for COVID-19 in the US: a cost-effectiveness analysis. Cost Eff Resour Alloc 2022 Jan 15;20(1):2. [doi: 10.1186/s12962-021-00336-x] [Medline: 35033100]
32. Sanchez-Taltavull D, Castelo-Szekely V, Murugan S, et al. Regular testing of asymptomatic healthcare workers identifies cost-efficient SARS-CoV-2 preventive measures. PLoS ONE 2021;16(11):e0258700. [doi: 10.1371/journal.pone.0258700] [Medline: 34739484]
33. Larese Filon F, Rui F, Ronchese F, De Michieli P, Negro C. Incidence of COVID-19 infection in hospital workers from March 1, 2020 to May 31, 2021 routinely tested, before and after vaccination with BNT162B2. Sci Rep 2022 Feb 15;12(1):2533. [doi: 10.1038/s41598-021-04665-y] [Medline: 35169127]
34. Sandmann FG, White PJ, Ramsay M, Jit M. Optimizing benefits of testing key workers for infection with SARS-CoV-2: a mathematical modeling analysis. Clin Infect Dis 2020 Dec 15;71(12):3196-3203. [doi: 10.1093/cid/ciaa901] [Medline: 32634823]
35. Leventopoulos M, Michou V, Papadimitropoulos M, et al. Evaluation of the Boson rapid Ag test vs RT-PCR for use as a self-testing platform. Diagn Microbiol Infect Dis 2022 Nov;104(3):115786. [doi: 10.1016/j.diagmicrobio.2022.115786] [Medline: 35998553]
36. Wölfl-Duchek M, Bergmann F, Jorda A, et al. Sensitivity and specificity of SARS-CoV-2 rapid antigen detection tests using oral, anterior nasal, and nasopharyngeal swabs: a diagnostic accuracy study. Microbiol Spectr 2022 Feb 23;10(1):e0202921. [doi: 10.1128/spectrum.02029-21] [Medline: 35107327]
37. Labhardt ND, González Fernández L, Katende B, et al. Head-to-head comparison of nasal and nasopharyngeal sampling using SARS-CoV-2 rapid antigen testing in Lesotho. PLoS ONE 2023;18(3):e0278653. [doi: 10.1371/journal.pone.0278653] [Medline: 36862684]
38. Nahar K, Begum MN, Tony SR, et al. Nasal swab as an alternative specimen for the detection of severe acute respiratory syndrome coronavirus 2. Health Sci Rep 2023 Apr;6(4):e1213. [doi: 10.1002/hsr2.1213] [Medline: 37077182]
39. Klein JAF, Krüger LJ, Tobian F, et al. Head-to-head performance comparison of self-collected nasal versus professional-collected nasopharyngeal swab for a WHO-listed SARS-CoV-2 antigen-detecting rapid diagnostic test. Med Microbiol Immunol 2021 Aug;210(4):181-186. [doi: 10.1007/s00430-021-00710-9] [Medline: 34028625]
40. Toledo DM, Robbins AA, Gallagher TL, et al. Wastewater-based SARS-CoV-2 surveillance in northern New England. Microbiol Spectr 2022 Apr 27;10(2):e0220721. [doi: 10.1128/spectrum.02207-21] [Medline: 35412387]

Abbreviations

CDC: Centers for Disease Control and Prevention
Ct: cycle threshold
NPS: nasopharyngeal swab
NS: nasal swab
OR: odds ratio
QAT: qualitative SARS-CoV-2 rapid antigen testing
RNP: RNAse P
RT-PCR: reverse transcriptase polymerase chain reaction
RT-qPCR: quantitative reverse transcriptase polymerase chain reaction

Edited by A Mavragani; submitted 23.04.24; peer-reviewed by D Laborde, X Sun; revised version received 19.11.24; accepted 20.11.24; published 27.01.25.

Please cite as:
Sood A, "Cotton" Jarrell W, Shore XW, Sosa N, Parada A, Edwardson N, Yingling AV, Amirkabirian T, Cheng Q, Hurwitz I, Cook LS, Leng S, Myers OB, Perkins DJ
Effectiveness of Frequent Point-of-Care Molecular COVID-19 Surveillance in a Rural Workplace: Nonrandomized Controlled Clinical Trial Among Miners
JMIR Public Health Surveill 2025;11:e59845
URL: https://publichealth.jmir.org/2025/1/e59845
doi:10.2196/59845

© Akshay Sood, William "Cotton" Jarrell, Xin W Shore, Nestor Sosa, Alisha Parada, Nicholas Edwardson, Alexandra V Yingling, Teah Amirkabirian, Qiuying Cheng, Ivy Hurwitz, Linda S Cook, Shuguang Leng, Orrin B Myers, Douglas J Perkins. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which

XSL•FO
RenderX

permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Updated Surveillance Metrics and History of the COVID-19 Pandemic (2020-2023) in East Asia and the Pacific Region: Longitudinal Trend Analysis

Alexander L Lundberg[1,2], PhD, MSc; Alan G Soetikno[3], BS; Scott A Wu[3], BA; Egon Ozer[4,5], MD, PhD; Sarah B Welch[1], MPH; Yingxuan Liu[1], MA; Claudia Hawkins[4,6], MD, MPH; Maryann Mason[1,2], PhD; Robert Murphy[4,7], MD; Robert J Havey[7,8], MD; Charles B Moss[9], PhD; Chad J Achenbach[4,7], MD, MPH; Lori Ann Post[1,2], PhD

[1]Buehler Center for Health Policy and Economics, Robert J Havey, MD Institute for Global Health, Northwestern University, Chicago, IL, United States

[2]Department of Emergency Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[3]Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[4]Department of Medicine, Division of Infectious Diseases, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[5]Center for Pathogen Genomics and Microbial Evolution, Robert J. Havey, MD Institute for Global Health, Northwestern University, Chicago, IL, United States

[6]Center for Global Communicable and Emerging Infectious Diseases, Robert J Havey, MD Institute for Global Health, Northwestern University, Chicago, IL, United States

[7]Robert J Havey, MD Institute for Global Health, Northwestern University, Chicago, IL, United States

[8]Department of Medicine, General Internal Medicine and Geriatrics, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States

[9]Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, United States

Corresponding Author:
Lori Ann Post, PhD
Buehler Center for Health Policy and Economics
Robert J Havey, MD Institute for Global Health
Northwestern University
420 E Superior St
Chicago, IL, 60611
United States
Phone: 1 3125031706
Email: lori.post@northwestern.edu

Abstract

Background: This study updates the COVID-19 pandemic surveillance in East Asia and the Pacific region that we first conducted in 2020 with 2 additional years of data for the region.

Objective: First, we aimed to measure whether there was an expansion or contraction of the pandemic in East Asia and the Pacific region when the World Health Organization (WHO) declared the end of the COVID-19 *public health emergency of international concern* on May 5, 2023. Second, we used dynamic and genomic surveillance methods to describe the dynamic history of the pandemic in the region and situate the window of the WHO declaration within the broader history. Finally, we aimed to provide historical context for the course of the pandemic in East Asia and the Pacific region.

Methods: In addition to updates of traditional surveillance data and dynamic panel estimates from the original study, this study used data on sequenced SARS-CoV-2 variants from the Global Initiative on Sharing All Influenza Data to identify the appearance and duration of variants of concern. We used Nextclade nomenclature to collect clade designations from sequences and Pangolin nomenclature for lineage designations of SARS-CoV-2. Finally, we conducted a 1-sided *t* test to determine whether the regional weekly speed was greater than an outbreak threshold of 10. We ran the test iteratively with 6 months of data across the sample period.

Results: Several countries in East Asia and the Pacific region had COVID-19 transmission rates above an outbreak threshold at the point of the WHO declaration (Brunei, New Zealand, Australia, and South Korea). However, the regional transmission rate had remained below the outbreak threshold for 4 months. In the rolling 6-month window *t* test for regional outbreak status, the final *P* value ≤.10 implies a rejection of the null hypothesis (at the α=.10 level) that the region as a whole was not in an outbreak

XSL•FO
RenderX

for the period from November 5, 2022, to May 5, 2023. From January 2022 onward, nearly every sequenced SARS-CoV-2 specimen in the region was identified as the Omicron variant.

Conclusions: While COVID-19 continued to circulate in East Asia and the Pacific region, transmission rates had fallen below outbreak status by the time of the WHO declaration. Compared to other global regions, East Asia and the Pacific region had the latest outbreaks driven by the Omicron variant. COVID-19 appears to be endemic in the region, no longer reaching the threshold for a pandemic definition. However, the late outbreaks raise uncertainty about whether the pandemic was truly over in the region at the time of the WHO declaration.

(JMIR Public Health Surveill 2025;11:e53214) doi:10.2196/53214

KEYWORDS

SARS-CoV-2; COVID-19; East Asia; Pacific; American Samoa; Australia; Brunei Darussalam; Cambodia; China; Fiji; French Polynesia; Guam; Hong Kong; Indonesia; Japan; Kiribati; People's Democratic Republic of Korea; Republic of Korea; Lao People's Democratic Republic; Macao; Malaysia; Marshall Islands; Federated States of Micronesia; Mongolia; Myanmar; Nauru, New Caledonia; New Zealand; Northern Mariana Islands; Palau; Papua New Guinea; Philippines; Samoa; Singapore; Solomon Islands; Thailand; Timor-Leste; Tonga; Tuvalu; Vanuatu; Vietnam; pandemic; surveillance; public health; speed; acceleration; deceleration; jerk; dynamic panel; generalized method of moments; Arellano-Bond; 7-day lag

Introduction

Background

COVID-19, the disease caused by SARS-CoV-2, was first detected in Wuhan, China, in the fall of 2019 [1-5]. Our research team conducted an analysis of the pandemic in East Asia and the Pacific region one year into the pandemic [6]. This study provides 2 additional years of updated surveillance and analysis for the region.

We adopted the World Bank's definition of East Asia and the Pacific region, which is based on economic development and geographical proximity, encompassing American Samoa, Australia, Brunei Darussalam, Cambodia, China, Fiji, French Polynesia, Guam, Hong Kong, Indonesia, Japan, Kiribati, People's Democratic Republic of Korea, Republic of Korea, Lao People's Democratic Republic, Macao, Malaysia, Marshall Islands, Federated States of Micronesia, Mongolia, Myanmar, Nauru, New Caledonia, New Zealand, Northern Mariana Islands, Palau, Papua New Guinea, the Philippines, Samoa, Singapore, Solomon Islands, Thailand, Timor-Leste, Tonga, Tuvalu, Vanuatu, and Vietnam [7]. North Korea neither collects nor reports COVID-19 caseloads and deaths.

The World Health Organization (WHO) and Director-General Ghebreyesus declared the end of *COVID-19 as a public health emergency of international concern* on May 5, 2023 [8-10], based on the recommendation of the COVID-19 Emergency Committee [10]. We compared the progression of the pandemic before and after this declaration.

Empirical Definition of Pandemic Versus Epidemic Versus Outbreak Versus Endemic

Epidemiological terms, such as pandemic, epidemic, outbreak, and endemic, are used to describe the occurrence and spread of disease [11,12]. The distinctions between these terms lie in their scope, geographic extent, and severity. An epidemic refers to a sudden increase in the number of disease cases in a specific population or region. If the epidemic spreads across several countries or continents, it becomes a pandemic. An outbreak, on the other hand, describes a sudden increase in a concentrated setting, usually involving a more limited geographic area than an epidemic. Endemic refers to the constant presence of a disease in a particular geographic region or population, with no sudden increases in case volume [13,14]. Field epidemiology defines these terms based on transmission metrics and geographical distribution and uses transmission metrics and geographical distribution to differentiate the nuances of these terms. Although public health surveillance data may not capture all cases, they are the best proxy for monitoring and tracking disease spread within a population in near real time. Surveillance is crucial for the timely response to health threats [15].

Traditional Surveillance Versus Enhanced Surveillance

Public health surveillance is the "ongoing, systematic collection, analysis, and interpretation of health-related data essential to planning and evaluation of public health practice" [16]. Not only does surveillance explain the burden of death and disease transmission but also it generates research questions and guides investigators on topics that require further investigation [17-31]. Surveillance allows us to compare the burden of disease between geographical regions and to understand which regions are most impacted. The impact can be measured through per capita rates of how many people contract a disease, how many die, and affiliated costs.

However, traditional surveillance carries several limitations that this study has addressed. Traditional surveillance provides a snapshot of past events [17-31], meaning it is static and only reflects historical data. In the middle of a burgeoning pandemic, policy makers and public health practitioners also need to understand what is about to happen. Is an outbreak increasing? Will growth switch from linear to exponential? Are more people dying from that particular condition in one place than another? To inform health policy and practice, knowledge of what is about to happen is often more valuable than knowledge of what did happen. To that end, we have developed enhanced surveillance metrics that reflect the dynamics of a pandemic and inform imminent growth—most importantly, where along the epidemiological outbreak curve a particular region is situated. In addition, we included dynamic metrics about the speed of the pandemic at the national, regional, and global level. We measured how acceleration of speed one week compared

XSL•FO
RenderX

to the prior week, as well as how novel infections in a prior week predicted new cases the following week. We can think of the latter measure as the echoing forward of cases. These metrics were tested and validated in prior research [6,32-42].

For the purpose of this study, standard surveillance metrics explain what has already happened in East Asia and the Pacific region, while enhanced surveillance metrics provide insights into future trends or the current position of a country along an epidemiological curve. We used both types of metrics to analyze the possible end to the pandemic in East Asia and the Pacific region.

Objectives

This study has 3 objectives. First, we aimed to measure whether there was an expansion or contraction in the pandemic in East Asia and the Pacific region at the time the WHO declared the end of the *COVID-19 pandemic as a public health emergency of international concern* on May 5, 2023 [43-51]. At both the region and country level, we used advanced surveillance and analytical techniques to describe the status of the pandemic in a 2-week window around the WHO declaration. From a public health perspective, we needed to know whether the rate of new COVID-19 cases was increasing, decreasing, or stable from week to week, and if any changes in the transmission rate indicated an acceleration or deceleration of the pandemic. Statistical insignificance is significant—it can signal the epidemiological *end* to the pandemic if the rate of new cases is 0 (or very low) and stable, meaning the number of new cases is neither accelerating nor decelerating.

Second, we used dynamic and genomic surveillance methods to describe the history of the pandemic in the region and situate the time window around the WHO declaration within the broader history. We included the ratio of COVID-19 deaths to the number of transmissions as a proxy for the mortality risk from infection at the population level. We also included a historical record of genomic surveillance from sequenced viral specimens to identify the appearance and spread of variants of concern in the region.

Third, we aimed to provide historical context for the course of the pandemic in East Asia and the Pacific region. We have addressed several questions. How did countries respond to the pandemic? How did the region fare in terms of disease burden? And what social, economic, and political factors shaped the course of COVID-19 in the region? This context can provide important lessons for disease prevention and mitigation in future pandemics.

Methods

Overview

We conducted trend analyses with longitudinal COVID-19 data from Our World in Data (OWID) [52]. This study provides updates of traditional surveillance data and dynamic panel estimates from the original work by Post et al [6,40,41,53-55]. For East Asia and the Pacific region, the data comprised an unbalanced panel of 36 countries and territories, running from August 14, 2020, to May 12, 2023. Because several countries around the world switched from daily to weekly reports at various points in 2023, we used a cubic spline to interpolate daily new cases and deaths if any country had 4 consecutive periods of nonzero new cases interspersed by 6 days of 0 new cases [56-58].

To identify the appearance and duration of variants of concern, we also used data on sequenced SARS-CoV-2 variants from the Global Initiative on Sharing All Influenza Data (GISAID), which is an effective and trusted web-based resource for sharing genetic, clinical, and epidemiological COVID-19 data [59-62]. We used Nextclade nomenclature [63] to collect clade designations from sequences and Pangolin nomenclature for lineage designations of SARS-CoV-2 [64,65]. Metadata for the study period were collected on June 22, 2023. To avoid low frequency or potentially erroneous samples, the dataset was further filtered to exclude months with fewer than 100 available samples, variant groups with fewer than 5 samples in a month, and variant groups representing <0.5% of the total samples in a month. The final dataset consisted of 184,386 total samples available on GISAID [59-62].

Traditional surveillance metrics include the *speed* of transmission or the rate of new COVID-19 cases per 100,000 population. Novel metrics include *acceleration*, *jerk*, and *1- and 7-day persistence* estimates. *Acceleration* is the change in speed from one time unit to the next. Acceleration can measure whether transmission rates are rising, falling, or stable. *Jerk* is the change in acceleration from one time unit to the next. The term *jerk* is adopted from physics nomenclature, as a large jerk can signal explosive growth in transmissions. The *1- and 7-day persistence* measures provide the impact of the 1- and 7-day lag of speed on current speed. For another interpretation, these estimates capture how COVID-19 cases echo forward to cases either 1 or 7 days later. They are coefficient estimates from an Arellano-Bond dynamic panel data model [66]:

in which the dependent variable is speed, the independent variables include weekend and recent week indicators, α_i is a country fixed effect, and u_{it} is the idiosyncratic error term. Advantages of the Arellano-Bond estimates include a correction for time-invariant omitted variables and the ability to assess both the predictive ability of the model and validity of its specification [41].

We analyzed the potential *statistical end* to the pandemic with a 1-sided *t* test for whether the mean of speed was equal to or greater than the outbreak threshold of 10. We ran the test on a rolling 6-month window over weekly speed for the region, and we plotted the *P* values from the test over time. The test can assess whether the region as a whole was experiencing an outbreak over a defined time window, during which the speed may sometimes exceed the outbreak threshold and at other times remain below it. All statistical analyses were conducted in R (version 4.2.1; R Foundation for Statistical Computing) with the *plm* package (version 2.6-2) [53,54].

Ethical Considerations

All data used in this study are publicly available and contain no identifiable, private information. Therefore, the study is

XSL•FO
RenderX

deemed exempt research with human subjects as defined by the US Government Code of Federal Regulations 45CFR46 and so was not submitted the Northwestern University Institutional Review Board. However, the authors note that anonymized COVID-19 data surveillance systems can generate local and global ethical questions beyond the scope of this study [67].

Results

Tables 1-6 present traditional (ie, static) and novel (ie, dynamic) COVID-19 surveillance metrics with OWID transmission and death data for East Asia and the Pacific region. These results focus on the 2-week period around the time when the WHO declared COVID-19 was no longer a public health emergency of international concern. The week before the declaration was April 28, 2023, and the week after was May 5, 2023.

Table 1. Arellano-Bond dynamic panel data estimates of the number of daily COVID-19 infections reported by country in East Asia and the Pacific region from April 28 to May 12, 2023a,b.

Variables	Values	*P* values[c]
1-day lag coefficient	0.171	.46
7-day lag coefficient	0.877	<.001
Shift parameter, week of April 28	−0.488	<.001
Shift parameter, week of May 5	0.491	<.001
Weekend effect	−1.215	.28

[a]Wald test: χ^2_6=730,275, P≤.001.

[b]Sargan: χ^2_{540}=24, P=.99.

[c]Contains estimates from the model in equation (1).

Figure 1 plots the novel surveillance metrics over time, from the start of the pandemic up to the WHO declaration in May 2023. Figure 2 uses GISAID data on sequenced SARS-CoV-2 variants to present the proportion of variants of concern returned among specimens over a similar period. Figure 3 returns to regional transmission data to present the results of a rolling 6-month window *t* test for whether the region was in a state of outbreak over each period. Finally, Figure 4 presents a qualitative timeline of key events in the COVID-19 pandemic in East Asia and the Pacific region.

Table 1 presents the dynamic panel estimates for the most recent time window. The Wald test for the regression was significant (P<.001), and the Sargan test failed to reject the validity of the overidentification restrictions (P=.99). The 1-day lag coefficient was not statistically significant (P=.46), but the 7-day lag coefficient was (P<.001), suggesting a cluster effect in which cases on a given day impact cases 7 days later. The shift parameter for the week of April 28 was negative (−0.488), while the more recent shift parameter for the week of May 5 was positive (0.491). The roughly equal magnitudes suggest the echo-forward effect of cases remained roughly stable from the start to the end of the 2-week period.

Standard surveillance metrics for the weeks of April 28 and May 5, 2023, are provided in Tables 2 and 3. Although most countries in the region had a small number of new COVID-19 cases, several countries had high rates of transmission. For the week of April 28, Brunei had by far the highest rate of new COVID-19 cases at 1431 per 100,000 population. The next highest rate was 204 in New Zealand, followed by 115 in Australia and 36 in South Korea. These rates exceeded the outbreak threshold according to the Centers for Disease Control and Prevention [32-42,68]. Specifically, *low* transmission is defined as no more than 10 cases per 100,000 people per week, *moderate* transmission ranges from 10 to 50 cases per 100,000 people per week, and *substantial* transmission is between 50 to 100 cases per 100,000 people per week [68,69]. Thus, Brunei, New Zealand, and Australia were in greater-than-substantial outbreaks, while South Korea was in a moderate outbreak.

Outbreak status for countries in the region looked roughly the same for the week of May 5. Brunei had somewhat lower transmission rate than in the prior week, while Australia had a higher rate. The other exception is Guam, who entered an outbreak with a speed of 61 (Table 3). The speed in Guam was 17 for the week of April 21. The dip to zero in the intermediate week (Table 2) most likely represents a disruption in data reports, but it is also important to note that transmission speed can be more variable in small island territories. For the region as a whole, COVID-19 may have still met the pandemic definition at the time of the WHO declaration, with several countries experiencing sizeable outbreaks. However, transmission rates were low for the rest of the region, suggesting a transition from pandemic to endemic.

XSL•FO
RenderX

Table 2. Static COVID-19 surveillance metrics for East Asia and Pacific countries in the week of April 28, 2023.

Country	New COVID-19 cases, n	Cumulative COVID-19 cases, n	7-day moving average of new cases	Infection rate per 100,000 individuals	New weekly deaths, n	Cumulative deaths, n	7-day moving average of deaths	Death rate per 100,000 individuals	Conditional death rate
Australia	4301	11,272,355	4273.57	115.01	0	20,613	18	0.52	0
Brunei	918	292,644	721.29	1431.20	0	158	0.14	0.30	0
Cambodia	0	138,733	0.14	0	0	3056	0	0	0.02
China	462	99,250,200	511.14	0.03	7	120,984	5.86	0	0
Fiji	0	68,921	0	0	0	883	0	0	0.01
French Polynesia	6	78,518	9	0	0	649	0	0	0.01
Guam	0	51,240	8.71	0	0	413	0	0	0.01
Indonesia	2122	6,784,170	1871.14	0.77	20	161,404	20.86	0.01	0.02
Japan	7343	33,766,957	11,059.29	5.92	18	74,614	21	0.01	0
Laos	4	218,085	1.29	0.05	0	671	0	0	0
Malaysia	0	5,071,840	709	0	0	37,020	1.29	0	0.01
Mongolia	19	1,008,038	9.43	0	0	2136	0	0	0
Myanmar	208	635,660	140.57	0.38	0	19,492	0.14	0	0.03
New Caledonia	0	80,058	0	0	0	314	0	0	0
New Zealand	1509	2,261,126	1423.57	203.73	0	2762	3.71	0.51	0
Northern Mariana Islands	3	13,849	1.86	0	0	41	0	0	0
Palau	0	6000	0	5.53	0	9	0	0	0
Papua New Guinea	1	46,850	1.14	0.01	0	670	0	0	0.01
Philippines	1190	4,097,525	932.71	1.03	0	66,444	0	0	0.02
Singapore	3201	2,391,770	3316.71	0	0	1722	0	0	0
South Korea	18,752	31,251,203	15,477.57	36.19	6	34,518	6.71	0.01	0
Thailand	242	4,732,301	252.29	2.36	0	33,957	1.43	0.01	0.01
Timor	2	23,431	0.43	0.15	0	138	0	0	0.01
Vietnam	2233	11,567,728	1836.14	2.27	0	43,195	1	0	0

XSL•FO
RenderX

Table 3. Static COVID-19 surveillance metrics for East Asia and Pacific countries in the week of May 5, 2023.

Country	New COVID-19 cases, n	Cumulative COVID-19 cases, n	7-day moving average of new cases	Infection rate per 100,000 individuals	New weekly deaths, n	Cumulative deaths, n	7-day moving average of deaths	Death rate per 100,000 individuals	Conditional death rate
Australia	5076	11,303,671	4643.29	135.74	0	20,751	19.71	0.43	0
Brunei	816	299,505	927.86	1271.72	0	160	0.29	0.54	0
Cambodia	0	138,736	0.43	0	0	3056	0	0	0.02
China	572	99,254,488	612.57	0.04	12	121,048	9.14	0	0
Fiji	0	68,921	0	0	0	883	0	0	0.01
French Polynesia	3	78,545	3.57	0	0	649	0	0	0.01
Guam	105	51,345	15	61.12	0	413	0	0	0.01
Indonesia	1471	6,795,221	1578.71	0.53	22	161,574	24.29	0.01	0.02
Japan	833	33,803,572	5230.71	0.67	25	74,694	11.43	0.02	0
Laos	4	218,096	1.57	0.05	0	671	0	0	0
Malaysia	0	5,079,436	1085.14	0	0	37,028	1.14	0	0.01
Mongolia	49	1,008,265	36.43	0	0	2136	0	0	0
Myanmar	157	636,967	186.71	0.29	0	19,493	0.14	0	0.03
New Caledonia	0	80,058	0	0	0	314	0	0	0
New Zealand	1535	2,272,229	1564.43	207.28	0	2792	4.29	0.89	0
Northern Mariana Islands	4	13,872	3.57	0	0	41	0	0	0
Palau	1	6000	0.14	0	0	9	0	0	0
Papua New Guinea	0	46,864	2	0	0	670	0	0	0.01
Philippines	1940	4,108,914	1627	1.68	0	66,453	1.29	0	0.02
Singapore	3386	2,414,394	3277	0	0	1722	0	0	0
South Korea	19,989	31,371,347	17,163.43	38.58	8	34,591	10.43	0.02	0
Thailand	320	4,734,000	275.43	3.12	0	33,967	1.43	0.03	0.01
Timor	1	23,435	0.57	0.07	0	138	0	0	0.01
Vietnam	2823	11,585,390	2523.14	2.88	0	43,200	0.71	0	0

A comparison of Tables 2 and 3 demonstrates little to no change before and after the WHO declared an end to COVID-19 as a public health emergency. China, Japan, and South Korea had the most cumulative cases of COVID-19 transmissions, but these ranks are a function of population size. Thus, a better measure is the number of COVID-19 cases and deaths per 100,000 population. Moreover, death is often a better proxy for the state of an outbreak than transmissions because deaths are less likely to be undercounted [70]. Undercounting may be due to poor public health infrastructure, increased use of home antigen testing, or a dearth in polymerase chain reaction testing or other resources. China, Japan, and South Korea reported <.01 deaths per population. When we control for risk of death given the number of COVID-19 transmissions, we find that Myanmar had the highest conditional death rate, at 0.03 deaths per 100,000 population, while Cambodia, Indonesia, and the Philippines each reported a rate of 0.02.

XSL•FO
RenderX

Table 4. Novel COVID-19 surveillance metrics for East Asia and Pacific countries for the week of April 28, 2023.

Country	Speed	Acceleration	Jerk	7-day persistence effect on speed
Australia	114.28	0.31	−0.02	39.51
Brunei	1124.51	111.65	−4.97	134.57
Cambodia	0	0	0	0
China	0.04	0	0	0.01
Fiji	0	−0.05	−0.05	0.02
French Polynesia	3.26	0	0	0.51
Guam	5.07	−2.49	0.75	2.52
Indonesia	0.68	0	−0.03	0.15
Japan	8.92	−0.51	−0.90	2.88
Laos	0.02	0	−0.01	0.01
Malaysia	2.09	0	0	0.71
Mongolia	0.16	0	0	0.04
Myanmar	0.26	0.03	0	0.04
New Caledonia	0	0	0	0.02
New Zealand	192.18	2.01	1.31	72.76
Northern Mariana Islands	2.88	0	0	2.54
Palau	0.79	0.79	0.79	0
Papua New Guinea	0.01	0	0	0
Philippines	0.81	0.05	0.01	0.16
Singapore	59.34	0	0	24.83
South Korea	29.87	1.37	−0.17	8.68
Thailand	2.46	−0.01	−0.02	0.71
Timor	0.03	0.02	0.02	0.01
Vietnam	1.87	−0.11	0.12	0.85

Tables 4 and 5 present enhanced dynamic surveillance metrics for the weeks before and after May 5. For most countries, speed was low and stable, which implied values of acceleration and jerk close to 0. Four countries were in a state of outbreak: Australia, Brunei, New Zealand, and South Korea.

From the first to the second week, acceleration was positive for Australia but negative for Brunei, which means the outbreak was worsening in Australia but improving in Brunei. The near-zero value of jerk for Australia suggests an inflection point of the outbreak, which may have been on the verge of reduction. Because acceleration was negative for Brunei, the negative value of jerk suggests a continued, but slowed, downward trend in the transmission rate. However, the 7-day persistence effect remained high for both countries.

The outbreaks in New Zealand and South Korea remained stable, with near-zero acceleration and jerk, although the 7-day persistence effect showed a slight increase. Because 4 out of the 36 countries and territories were experiencing outbreaks, the distinction between pandemic and endemic for the region remains unclear. On the basis of trends in other global regions, these outbreaks are likely to be the last outbreaks of this magnitude for these 4 countries; however, the scale of these outbreaks means that transmissions could spread across borders, potentially leading to new outbreaks in neighboring countries.

Finally, we note that the figures in Tables 4 and 5 are not calculated as day-over-day averages across the week, as in Tables 2 and 3. Thus, the magnitudes of speed may not exactly match those presented in Tables 2 and 3.

XSL•FO
RenderX

Table 5. Novel COVID-19 surveillance metrics for East Asia and Pacific countries for the week of May 5, 2023.

Country	Speed	Acceleration	Jerk	7-day persistence effect on speed
Australia	124.16	2.96	0.61	44.39
Brunei	1446.57	−22.78	−25.05	436.81
Cambodia	0	0	0	0
China	0.04	0	0	0.01
Fiji	0	0	0	0
French Polynesia	1.26	0	0	1.27
Guam	8.73	8.73	8.73	1.97
Indonesia	0.57	−0.03	0.05	0.26
Japan	4.22	−0.75	1.17	3.47
Laos	0.02	0	0.01	0.01
Malaysia	3.20	0	0	0.81
Mongolia	0.95	0	0	0.06
Myanmar	0.34	−0.01	−0.01	0.10
New Caledonia	0	0	0	0
New Zealand	211.21	0.51	−1.12	74.65
Northern Mariana Islands	6.63	0	0	1.12
Palau	0	−0.79	−0.79	0.31
Papua New Guinea	0.02	0	0	0
Philippines	1.41	0.09	0.02	0.31
Singapore	57.34	0	0	23.05
South Korea	33.12	0.34	0.14	11.60
Thailand	2.69	0.11	0.03	0.96
Timor	0.04	−0.01	−0.01	0.01
Vietnam	2.57	0.09	−0.10	0.73

Table 6. East Asia and Pacific countries with the highest 7-day persistence estimate in the week of May 5, 2023.

Country	7-day persistence (May 5, 2023)
Brunei	436.81
New Zealand	74.65
Australia	44.39
Singapore	23.05
South Korea	11.60

Table 6 compares the 7-day persistence effect on speed for the top 5 countries for the week of May 5. Unsurprisingly, persistence was highest for Brunei, which had the largest outbreak at the time. Interestingly, Singapore appears on the list despite not having reached outbreak status. The persistence suggests that an outbreak could have been imminent for Singapore. Alternatively, the persistence could signify a strong, continued presence of COVID-19 as endemic in the area.

Figure 1 plots regional speed, acceleration, jerk, and 7-day persistence metrics from August 14, 2020, to May 12, 2023. The dashed gray line denotes the informal Centers for Disease Control and Prevention outbreak threshold of speed equal to 10. The region saw 4 outbreaks over the period, but the first and third were small, barely eclipsing a speed of 10. The first outbreak happened in August 2021. The second happened in February 2022, ultimately reaching a peak speed of 38. A third minor outbreak occurred in July 2022. The final outbreak began in December 2022, reaching a peak speed of 243 at the end of the year. This outbreak was one of the largest in the world over the course of the pandemic. However, since the end of January 2023, the region has not been in a state of outbreak.

XSL•FO
RenderX

Figure 1. Novel surveillance metrics (speed, acceleration, jerk, and 7-day persistence) for COVID-19 infections in East Asia and the Pacific region from August 2020 to May 2023.

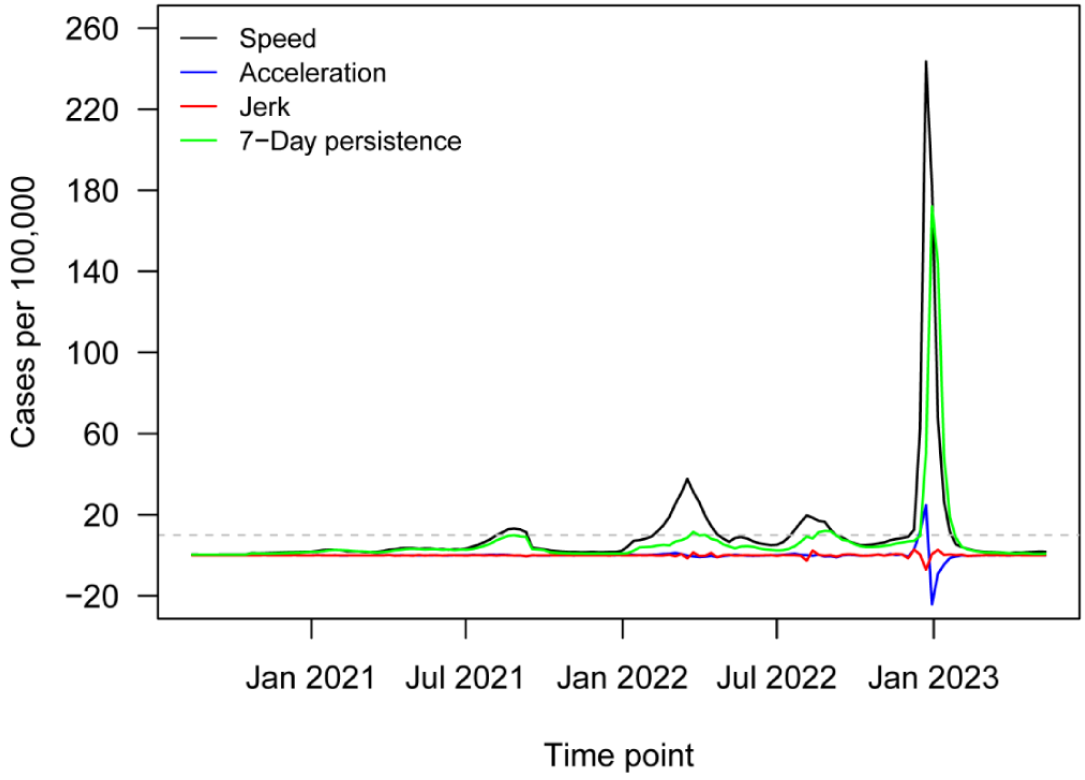


Figure 2 plots variant groups as a proportion of all viral specimens collected and sequenced in the region (and made available through GISAID) each month. The first outbreak occurred shortly after the appearance of the Delta variant. Each subsequent outbreak occurred after the dominant Omicron variant arrived. East Asia and the Pacific region, like much of the rest of the world, saw a surge in cases amid the heightened transmissibility of Omicron [71]. In contrast, the large outbreaks driven by Omicron did not occur until the end of 2022. This delay was achieved by the effective zero–COVID-19 policy in China, which was finally abandoned at that point. The outbreak hit in force shortly afterward.

Another potential indication of the end to the pandemic is the continued dominance of the Omicron variant. While the region saw a mixture of the ancestral, Alpha, Beta, and Delta variants before the arrival of Omicron in November of 2021, viral sequences have almost exclusively been identified as Omicron and its subvariants ever since.

Figure 3 plots *P* values from a series of 1-sided *t* tests of whether speed for the region was equal to or greater than the threshold outbreak of 10. These tests were conducted on a rolling 6-month window of weekly regional speed. The test can assess whether the region was in an outbreak over a period in which speed may sometimes eclipse an outbreak threshold and other times remain below the threshold. Statistical insignificance indicates no evidence against the null hypothesis of below-outbreak regional speed over a period. For example, at the α=.10 level, the final *P* value ≤.10 in the figure implies a rejection of the null hypothesis that the region as a whole was not in an outbreak for the period from November 5, 2022, to May 5, 2023.

In Figure 3, the dashed gray line denotes the least restrictive conventional significance level threshold of α=.10. The test first rejected the null in favor of the alternative for the 6-month period ending in early May of 2021. The test lost significance again in mid-September, before regaining significance at the very end of 2022. This second period of rejection was driven by the large Omicron outbreak around the time. More recently, the test was on the verge of insignificance at the time of WHO declaration. Although the test failed to reject the null at the α=.10 level, it did reject at the α=.05 level. Compared to other global regions, East Asia and the Pacific region are the latest to potentially transition from the pandemic to endemic COVID-19 status.

Within the historical context of enhanced surveillance metrics, the region may be at the end stage of the pandemic. Speed has not been this low for this long since the start of the pandemic. However, we do note that several countries in the region remained in substantial outbreaks at the time of the WHO declaration. This reality brings some uncertainty to the ostensible end to the pandemic in the region.

XSL•FO
RenderX

Figure 2. Variants of concern (VOCs) as a proportion of all sequenced SARS-CoV-2 specimens from March 2020 to May 2023 in East Asia and the Pacific region (n=1,234,493).

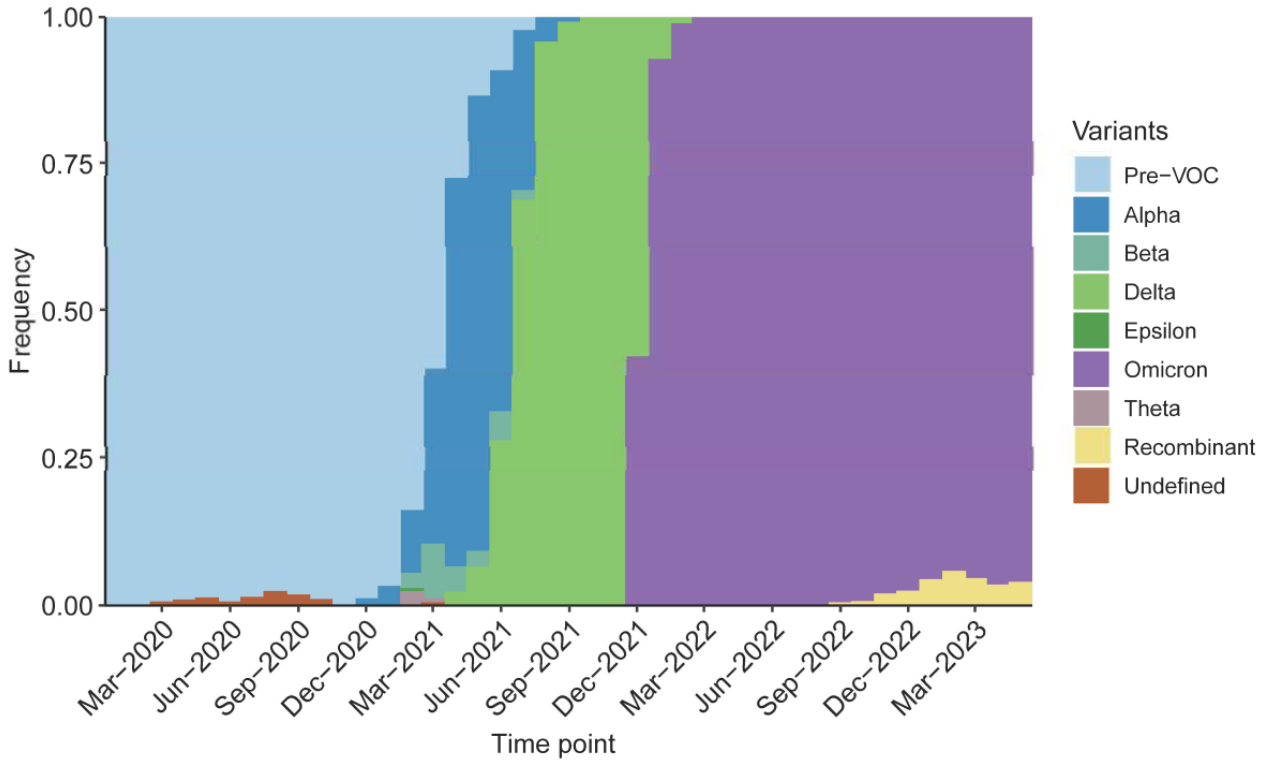


Figure 3. *P* values from t tests of weekly COVID-19 transmissions per 100,000 population equal to 10 over a rolling 6-month window in East Asia and the Pacific region.

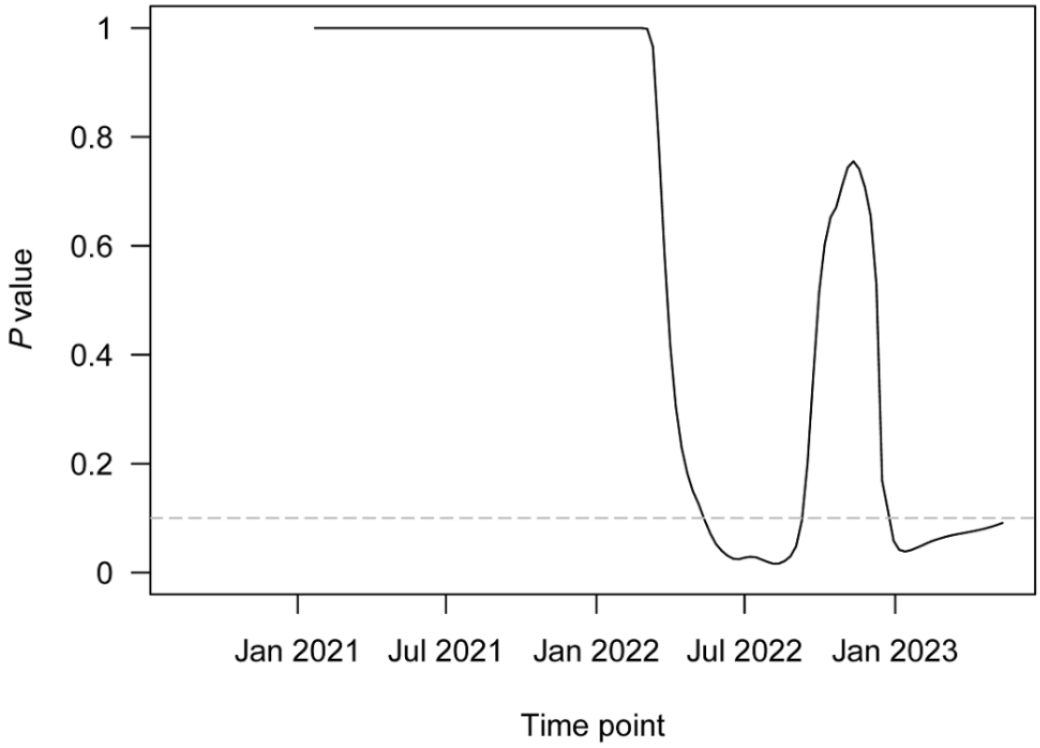
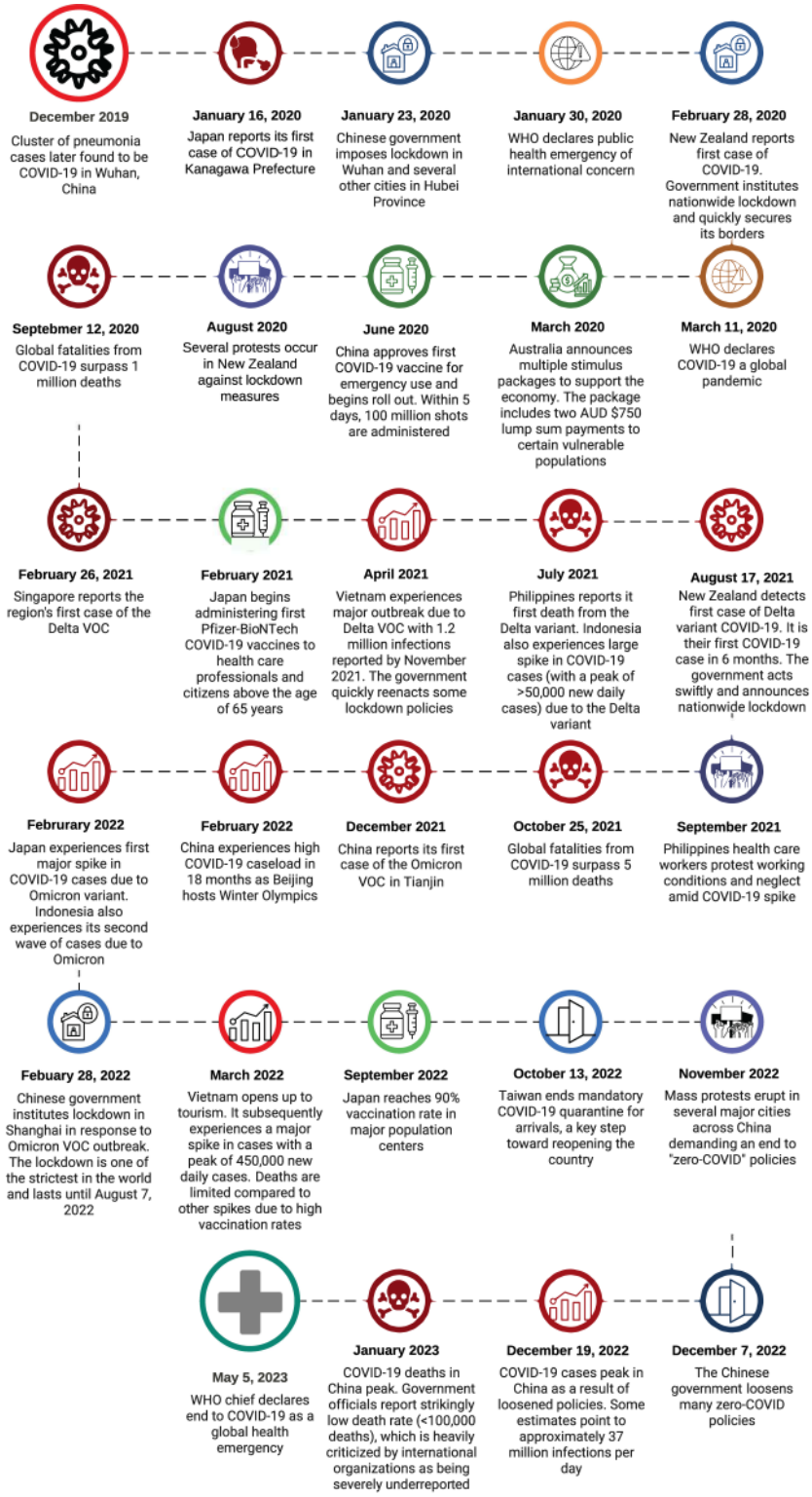


Figure 4 provides a timeline of the onset of COVID-19 in East Asia and the Pacific region as well as vaccination programs and major events that shaped the progression of the pandemic in the region. Most notably, the zero–COVID-19 policy in China effectively contained the disease in the country and therefore to an extent in the region. Upon the abandonment of the policy, a large wave of cases driven by Omicron finally hit the country, much later than in other parts of the world.

XSL•FO
RenderX

Figure 4. Timeline of the COVID-19 pandemic in East Asia and the Pacific region. VOC: variant of concern; WHO: World Health Organization.



Discussion

Principal Findings

COVID-19 has affected all countries in East Asia and the Pacific region. As of March 16, 2023, the WHO had reported a total of 213,883,614 COVID-19 cases and 622,018 deaths in East Asia and the Pacific region [72]. At the same point, the WHO reported 760,360,956 confirmed cases of COVID-19, including 6,873,477 deaths worldwide [72].

The results indicate a potential COVID-19 transition from pandemic to endemic at the time the WHO removed the designation of COVID-19 as a public health emergency of international concern on May 5, 2023. Regional speed had remained below the outbreak threshold of 10 for several months ahead of the WHO declaration. The rolling *t* test of speed, equal to the outbreak threshold, suggests that the region was in an

XSL•FO
RenderX

outbreak over the six-month period starting at the end of 2022 when the Omicron variant drove a large increase in cases. However, the result was only significant at the α=.10 level, not the 0.05 level. However, several countries in the region had transmission rates consistent with a substantial outbreak in the 2 weeks around the WHO declaration. These outbreaks make an apparent end to the pandemic in the region unclear at the time.

As another indicator, nearly all sequenced SARS-CoV-2 specimens in the region were identified as the Omicron variant from early 2022 onward. This transition is consistent with a more clearly defined end to the pandemic in other global regions [73-75]. The relative delay for East Asia and the Pacific region may be attributable to a late Omicron outbreak in the region when China relaxed its stringent and effective zero–COVID-19 policy [76].

Policies Implemented to Control and Mitigate the Transmission of COVID-19

Many countries in the region experienced multiple waves of COVID-19 infections, with some waves more severe than others. During the initial outbreaks, countries in East Asia and the Pacific region were quick to respond with policies that were highly effective at curbing transmission rates and limiting mortality. Compared to other regions of the world, Asian countries experienced significantly lower mortality early in the pandemic (2.7 deaths per million in Asian countries vs 197 deaths per million in European countries) [77]. This disparity has been attributed to several factors.

First, countries in the region were able to quickly formulate effective strategies that were guided by their prior experiences with other respiratory pandemics, such as the 2003 outbreak of SARS-CoV-1 and the 2015 outbreak of Middle East respiratory syndrome [78-80].

Second, cultural differences between Eastern and Western societies may have allowed countries to more easily adopt face masking and social distancing policies [77]. Where Western countries such as the United States pride themselves on individual freedoms, many Asian countries have cultures that value collective societal well-being [81].

Third, the region as a whole tended to implement strict containment policies, which typically included mask mandates, social distancing, mass testing requirements, contact tracing, quarantine protocols, border lockdowns, and eventually mass vaccination programs to control the virus.

China initially implemented a strict zero–COVID-19 policy with the aim to keep cases as close to zero as possible. At times, entire cities, such as Shanghai, Chengdu, and Wuhan, entered citywide lockdown to prevent the spread of the virus [82-84]. Amid these lockdowns, economic activities were halted, public transportation was suspended, and residents were restricted to their homes. Meanwhile, individuals who tested positive for COVID-19 were taken to large isolation facilities to undergo quarantine [82,85].

New Zealand and Taiwan also initially implemented strict COVID-19 policies including mask mandates, tight border control, contact tracing, and tightly targeted quarantines for local outbreaks. These strategies were highly effective in containing the virus during major outbreaks for both countries [86,87]. However, both eventually shifted away from these strategies with the emergence of more contagious variants such as Delta and Omicron. Instead, they focused more on achieving herd immunity through widespread vaccination campaigns and protecting populations [86,88].

Vietnam, having learned lessons from the SARS outbreak in 2003, began preparing for COVID-19 weeks before the first cases were reported in their country. Knowing that their health care resources were limited, Vietnam's strategic goal was to keep COVID-19 cases as low as possible. They began stockpiling personal protective equipment and created local medical response teams to address outbreaks. When Vietnam detected its first cases in the community, it acted decisively and swiftly by locking down areas and performing meticulous contact tracing [89,90]. Vietnam never relied on widespread testing to curb infections [89].

The Philippines and Indonesia were COVID-19 hotspots throughout the pandemic due to limited public health responses by their governments. Both countries were late to institute COVID-19 containment measures, such as contact tracing, testing, and isolation, leading to high rates of infections and deaths relative to other countries in the region [91-94]. In the Philippines, the country's limited health care infrastructure left it ill-prepared for the pandemic, leading to an exponential growth in cases during the first wave [94,95].

Many Pacific Island countries were quick to implement strategies to prevent COVID-19 infections. Given their geographical isolation, these countries went almost completely free of COVID-19 cases during the initial wave of infections through the use of strict border closures and restrictions on international travel [96-98]. For example, the Solomon Islands reported its first COVID-19 case in October 2020 in a student who was returning from the Philippines [99]. However, it was not until 2022 that the Solomon Islands faced its first widespread community outbreak [100]. These strict border closures, while effective, came with a natural economic cost as many island nations rely heavily on tourism, which virtually evaporated during the pandemic [97].

Subsequent waves of infection were primarily driven by local easing of COVID-19 precautions and the emergence of new and more contagious variants such as Delta and Omicron. For example, Japan has experienced several waves of COVID-19 infections, with the largest wave occurring in mid-2022 due to the Omicron variant [101]. China, on the other hand, has reported fewer waves of COVID-19 infections due to its strict containment measures, although it is not clear how accurate their data are given their lack of consistent reporting. However, when China eased their zero–COVID-19 policy in December 2022, it subsequently experienced an intense surge in cases and related deaths [76]. Other countries in Asia, such as Vietnam and Taiwan, have been relatively successful in controlling the virus and have experienced fewer waves of infection.

While countries enjoyed relative success early on in controlling disease, they had to grapple with the economic and social

XSL•FO
RenderX

consequences of their restrictive containment policies. These outbreaks and their associated governmental responses led to substantial sociopolitical turmoil throughout the region. Protests were widespread and focused on the government's mishandling of economic recovery and overly restrictive COVID-19 policies. Thailand and Malaysia saw protests break out in 2021 due to the government's perceived lackluster pandemic response [102,103]. The Philippines saw protesters calling for an end to quarantine rules, which led to President Duterte to issue a "shoot to kill" order against the protesters [104,105]. In China, antigovernment protests broke out throughout the country in late 2022 calling for an end to China's zero–COVID-19 policy, which used strict and draconian city-wide lockdowns to curb the spread of the virus. In response, China eased their policy in December 2022 and subsequently experienced an intense surge in cases and deaths. The WHO criticized China for underrepresenting the true number of deaths, with one senior official stating that China was using too narrow of a definition to count deaths [106,107]. New Zealand saw protesters converge on Parliament to oppose COVID-19 vaccines mandates [108].

In terms of the economy, East Asia and the Pacific region experienced a 0.2% decline in gross domestic product (GDP) in 2020 driven by containment measures that limited economic activity. However, as vaccines became more prevalent and countries began easing restrictions, economies rebounded with a GDP growth of 5.8% in 2021 [109].

To combat the economic ramifications of their containment strategies, governments in the region focused on reopening their economies and using large fiscal packages to stimulate economic recovery. These fiscal packages typically included direct cash support for affected workers, tax relief for individuals and corporations, support for their strained health care systems, and economic investment in specific economic sectors and public works. In 2020, the Chinese government launched several fiscal measures estimated to be worth approximately 4.9 trillion RMB (4.7% of GDP; US $672.3 billion). Key aspects of these measures included tax relief, investment in epidemic prevention and control, and medical equipment [110]. In Vietnam, the government introduced a fiscal support package worth 291.7 trillion VND (3.6% of GDP; US $11.6 billion) with measures that included deferment of value-added, personal income, and corporate income taxes [110]. The Australian government approved a substantial fiscal stimulus package worth AUS $312 billion (15.75% of GDP; US $196 billion) to be used through the financial year 2025, which included an AUS $20 billion (US $12.5 billion) health response package to secure vaccines and strengthen the existing health care system [110]. Singapore announced an SGD $92 billion (US $67.9 billion) fiscal package that included a cash payout to all Singaporeans [110]. In 2020, Japan approved multiple stimulus packages, including the Emergency Economic Package against COVID-19 (worth ¥117.1 trillion; US $750.5 billion) and the Comprehensive Economic Measures to Secure People's Lives and Livelihoods toward Relief Hope package (worth ¥73.6 trillion; US $471.7 billion). These measures were intended to protect businesses and support wages while also focusing on building a more resilient future economy [110].

Vaccination Efforts

Efforts to create a vaccine started immediately. SinoVac, a Chinese pharmaceutical company, began development of an inactivated COVID-19 vaccine called CoronaVac as early as January 2020. By fall 2020, SinoVac launched phase 3 trials in several countries, including Brazil, Indonesia, the Philippines, and Turkey [111]. The large phase 3 trial in Brazil, which used a 2-dose series administered 14 days apart, demonstrated an efficacy of 51% against symptomatic COVID-19 infections and 100% against severe infections [112]. By early 2021, many countries in East Asia and the Pacific region, as well as in South America and Eastern Europe, began mass vaccination programs using CoronaVac [113-118]. By April 2021, SinoVac had announced that it had production capacity to produce 2 billion doses per year [111]. On June 1, 2021, the WHO validated the vaccine for emergency use. Other countries, such as Taiwan, Japan, and Korea, purchased millions of vaccines from western countries, including the AstraZeneca, Pfizer-BioNTech, and Moderna vaccines in early 2021 [119-122].

With many countries now having reached herd immunity through a combination of several available vaccines, there has been widespread easing of border restrictions, as seen in Japan, China, Taiwan, Singapore, Australia, and New Zealand [123-127].

Conclusions

Concerns about potential resurgences of the virus remain valid. As long as COVID-19 continues to spread and mutate, the possibility of new variants emerging remains. These variants could potentially be more transmissible, be resistant to vaccines, or cause more severe illness. This underlines the importance of continued vigilance, vaccination efforts, and global cooperation to control the spread of the virus [39].

Limitations

COVID-19 data became less frequently reported worldwide by the time the WHO declared an end to the pandemic as a public health emergency of international concern [128]. In addition, more people began using at-home tests as the pandemic evolved [129], and experts believe the Chinese government has underreported COVID-19 deaths [130]. Because the enhanced surveillance metrics of speed, acceleration, jerk, and 7-day persistence are based on rates and not total counts, statistical bias caused by countries dropping in or out of the sample is mitigated. However, if a nonincluded country is unrepresentative of the region in terms of disease burden, its omission can still influence historical data comparisons. Viral specimen tests for variants of concern in GISAID depend on testing and sequencing capacity, which varied by country across the region.

XSL•FO
RenderX

Acknowledgments

Research reported in this publication was generously supported by the Robert J Havey, MD, Institute for Global Health's Innovation Challenge Award at Northwestern University Feinberg School of Medicine.

The authors gratefully acknowledge all data contributors, including the authors and their originating laboratories responsible for obtaining the specimens and their submitting laboratories for generating the genetic sequence and metadata and sharing them via the Global Initiative on Sharing All Influenza Data, on which this research is based.

Conflicts of Interest

None declared.

References

1. Muralidar S, Ambi SV, Sekaran S, Krishnan UM. The emergence of COVID-19 as a global pandemic: understanding the epidemiology, immune response and potential therapeutic targets of SARS-CoV-2. Biochimie 2020 Dec;179:85-100 [FREE Full text] [doi: 10.1016/j.biochi.2020.09.018] [Medline: 32971147]
2. Sharma A, Ahmad Farouk I, Lal SK. COVID-19: a review on the novel coronavirus disease evolution, transmission, detection, control and prevention. Viruses 2021 Jan 29;13(2):202 [FREE Full text] [doi: 10.3390/v13020202] [Medline: 33572857]
3. Chilamakuri R, Agarwal S. COVID-19: characteristics and therapeutics. Cells 2021 Jan 21;10(2):206 [FREE Full text] [doi: 10.3390/cells10020206] [Medline: 33494237]
4. Hu B, Guo H, Zhou P, Shi Z. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 2021 Mar 06;19(3):141-154 [FREE Full text] [doi: 10.1038/s41579-020-00459-7] [Medline: 33024307]
5. Seyed Hosseini E, Riahi Kashani N, Nikzad H, Azadbakht J, Hassani Bafrani H, Haddad Kashani H. The novel coronavirus disease-2019 (COVID-19): mechanism of action, detection and recent therapeutic strategies. Virology 2020 Dec;551:1-9 [FREE Full text] [doi: 10.1016/j.virol.2020.08.011] [Medline: 33010669]
6. Post LA, Lin JS, Moss CB, Murphy RL, Ison MG, Achenbach CJ, et al. SARS-CoV-2 wave two surveillance in east Asia and the Pacific: longitudinal trend analysis. J Med Internet Res 2021 Feb 01;23(2):e25454 [FREE Full text] [doi: 10.2196/25454] [Medline: 33464207]
7. East Asia and Pacific. The World Bank Group. URL: https://data.worldbank.org/region/east-asia-and-pacific [accessed 2024-04-29]
8. Smith-Schoenwalder C. When Does the COVID-19 pandemic end?: neither WHO nor any other organization decides when a pandemic begins or ends. U.S. News & World Report. 2023. URL: https://www.usnews.com/news/health-news/articles/2023-05-12/when-does-the-covid-19-pandemic-end [accessed 2024-04-29]
9. Burki T. WHO ends the COVID-19 public health emergency. Lancet Respir Med 2023 Jul;11(7):588 [FREE Full text] [doi: 10.1016/S2213-2600(23)00217-5] [Medline: 37247628]
10. WHO chief declares end to COVID-19 as a global health emergency. United Nations. URL: https://news.un.org/en/story/2023/05/1136367 [accessed 2024-04-29]
11. Lesson 1: introduction to epidemiology. Centers for Disease Control and Prevention. URL: https://archive.cdc.gov/#/details?url=https://www.cdc.gov/csels/dsepd/ss1978/lesson1/section11.html [accessed 2024-04-29]
12. Epidemic, endemic, pandemic: what are the differences? Columbia University. 2021 Feb 19. URL: https://www.publichealth.columbia.edu/public-health-now/news/epidemic-endemic-pandemic-what-are-differences [accessed 2024-04-29]
13. The Lancet Infectious Diseases. Transitioning to endemicity with COVID-19 research. Lancet Infect Dis 2022 Mar;22(3):297 [FREE Full text] [doi: 10.1016/S1473-3099(22)00070-6] [Medline: 35151377]
14. Britto D, Delorme J, Miller B, Sleat D, Summers H, Wain R. Pandemic to endemic: the race against time. Tony Blair Institute for Global Change. 2021 Sep 19. URL: https://institute.global/policy/pandemic-endemic-race-against-time [accessed 2024-04-29]
15. Gilbert R, Cliffe SJ. Public health surveillance. In: Regmi K, Gee I, editors. Public Health Intelligence. Cham, Switzerland: Springer; Apr 27, 2016:91-110.
16. Hong R, Walker R, Hovan G, Henry L, Pescatore R. The power of public health surveillance. Dela J Public Health 2020 Jul;6(2):60-63 [FREE Full text] [doi: 10.32481/djph.2020.07.016] [Medline: 34467112]
17. Teutsch SM, Churchill RE. Principles and Practice of Public Health Surveillance 2nd Edition. Oxford, UK: Oxford University Press; 2000.
18. Teutsch SM. Considerations in planning a surveillance system. In: Lee LM, Teutsch SM, Thacker SB, St. Louis ME, editors. Principles & Practice of Public Health Surveillance. 3rd edition. Oxford, UK: Oxford Academic Press; 2010:18-28.
19. Teutsch SM, Thacker SB. Planning a public health surveillance system. Epidemiol Bull 1995 Mar;16(1):1-6. [Medline: 7794696]
20. Thacker SB, Qualters JR, Lee LM, Centers for Disease ControlPrevention. Public health surveillance in the United States: evolution and challenges. MMWR Suppl 2012 Jul 27;61(3):3-9. [Medline: 22832990]

XSL•FO
RenderX

21. Centers for Disease Control (CDC). Guidelines for evaluating surveillance systems. MMWR Suppl 1988 May 06;37(5):1-18. [Medline: 3131659]
22. Lee LM, Thacker SB. Public health surveillance and knowing about health in the context of growing sources of health data. Am J Prev Med 2011 Dec;41(6):636-640. [doi: 10.1016/j.amepre.2011.08.015] [Medline: 22099242]
23. Davis AM, Dunet DO, Keaton R, Snider DE, Sosin DM, Stroup DF, et al. CDC guidelines: improving the quality. Centers for Disease Control and Prevention. 1996. URL: https://stacks.cdc.gov/view/cdc/24407#moretextPAmods.subject_name [accessed 2024-04-29]
24. Thacker SB, Stroup DF. Future directions for comprehensive public health surveillance and health information systems in the United States. Am J Epidemiol 1994 Sep 01;140(5):383-397. [doi: 10.1093/oxfordjournals.aje.a117261] [Medline: 8067331]
25. Nsubuga P, White M, Thacker S, Anderson M, Blount S, Broome C, et al. Public health surveillance: a tool for targeting and monitoring interventions. In: Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, et al, editors. Disease Control Priorities in Developing Countries. 2nd edition. New York, NY: Oxford University Press; 2006.
26. Lee LM, Thacker SB, St. Louis ME, Teutsch SM. Principles and Practice of Public Health Surveillance 3rd Edition. New York, NY: Oxford University Press; 2010.
27. Thacker SB, Stroup DF, Rothenberg RB. Public health surveillance for chronic conditions: a scientific basis for decisions. Stat Med 1995 Feb 28;14(5-7):629-641. [doi: 10.1002/sim.4780140520] [Medline: 7792453]
28. Perry HN, McDonnell SM, Alemu W, Nsubuga P, Chungong S, Otten MW, et al. Planning an integrated disease surveillance and response system: a matrix of skills and activities. BMC Med 2007 Aug 15;5(1):24 [FREE Full text] [doi: 10.1186/1741-7015-5-24] [Medline: 17697387]
29. Koo D, Thacker SB. In snow's footsteps: commentary on shoe-leather and applied epidemiology. Am J Epidemiol 2010 Sep 15;172(6):737-739. [doi: 10.1093/aje/kwq252] [Medline: 20720100]
30. Romaguera RA, German RR, Klaucke DN. Evaluating public health surveillance. In: Teutsch SM, Churchill RE, editors. Principles and Practice of Public Health Surveillance. Oxford, UK: Oxford University Press; 2000:176-193.
31. Pappaioanou M, Malison M, Wilkins K, Otto B, Goodman RA, Churchill R, et al. Strengthening capacity in developing countries for evidence-based public health: the data for decision-making project. Soc Sci Med 2003 Nov;57(10):1925-1937. [doi: 10.1016/s0277-9536(03)00058-3] [Medline: 14499516]
32. Post L, Boctor MJ, Issa TZ, Moss CB, Murphy RL, Achenbach CJ, et al. SARS-CoV-2 surveillance system in Canada: longitudinal trend analysis. JMIR Public Health Surveill 2021 May 10;7(5):e25753 [FREE Full text] [doi: 10.2196/25753] [Medline: 33852410]
33. Post L, Culler K, Moss CB, Murphy RL, Achenbach CJ, Ison MG, et al. Surveillance of the second wave of COVID-19 in Europe: longitudinal trend analyses. JMIR Public Health Surveill 2021 Apr 28;7(4):e25695 [FREE Full text] [doi: 10.2196/25695] [Medline: 33818391]
34. Post L, Marogi E, Moss CB, Murphy RL, Ison MG, Achenbach CJ, et al. SARS-CoV-2 surveillance in the Middle East and North Africa: longitudinal trend analysis. J Med Internet Res 2021 Jan 15;23(1):e25830 [FREE Full text] [doi: 10.2196/25830] [Medline: 33302252]
35. Post L, Ohiomoba RO, Maras A, Watts SJ, Moss CB, Murphy RL, et al. Latin America and the Caribbean SARS-CoV-2 surveillance: longitudinal trend analysis. JMIR Public Health Surveill 2021 Apr 27;7(4):e25728 [FREE Full text] [doi: 10.2196/25728] [Medline: 33852413]
36. Post LA, Argaw ST, Jones C, Moss CB, Resnick D, Singh LN, et al. A SARS-CoV-2 surveillance system in Sub-Saharan Africa: modeling study for persistence and transmission to inform policy. J Med Internet Res 2020 Nov 19;22(11):e24248 [FREE Full text] [doi: 10.2196/24248] [Medline: 33211026]
37. Post LA, Benishay ET, Moss CB, Murphy RL, Achenbach CJ, Ison MG, et al. Surveillance metrics of SARS-CoV-2 transmission in central Asia: longitudinal trend analysis. J Med Internet Res 2021 Feb 03;23(2):e25799 [FREE Full text] [doi: 10.2196/25799] [Medline: 33475513]
38. Post LA, Issa TZ, Boctor MJ, Moss CB, Murphy RL, Ison MG, et al. Dynamic public health surveillance to track and mitigate the US COVID-19 epidemic: longitudinal trend analysis study. J Med Internet Res 2020 Dec 03;22(12):e24286 [FREE Full text] [doi: 10.2196/24286] [Medline: 33216726]
39. Post LA, Lorenzo-Redondo R. Omicron: fewer adverse outcomes come with new dangers. Lancet 2022 Apr 02;399(10332):1280-1281 [FREE Full text] [doi: 10.1016/S0140-6736(22)00514-1] [Medline: 35305298]
40. Oehmke JF, Moss CB, Singh LN, Oehmke TB, Post LA. Dynamic panel surveillance of COVID-19 transmission in the United States to inform health policy: observational statistical study. J Med Internet Res 2020 Oct 05;22(10):e21955 [FREE Full text] [doi: 10.2196/21955] [Medline: 32924962]
41. Oehmke JF, Oehmke TB, Singh LN, Post LA. Dynamic panel estimate-based health surveillance of SARS-CoV-2 infection rates to inform public health policy: model development and validation. J Med Internet Res 2020 Sep 22;22(9):e20924 [FREE Full text] [doi: 10.2196/20924] [Medline: 32915762]
42. Oehmke TB, Post LA, Moss CB, Issa TZ, Boctor MJ, Welch SB, et al. Dynamic panel data modeling and surveillance of COVID-19 in metropolitan areas in the United States: longitudinal trend analysis. J Med Internet Res 2021 Feb 09;23(2):e26081 [FREE Full text] [doi: 10.2196/26081] [Medline: 33481757]

XSL•FO
RenderX

43. Zarocostas J. With the COVID-19 PHEIC over, what next? Lancet 2023 May 20;401(10389):1642-1643 [FREE Full text] [doi: 10.1016/S0140-6736(23)01003-6] [Medline: 37211013]
44. Sarker R, Roknuzzaman AS, Hossain MJ, Bhuiyan MA, Islam MR. The WHO declares COVID-19 is no longer a public health emergency of international concern: benefits, challenges, and necessary precautions to come back to normal life. Int J Surg 2023 Sep 01;109(9):2851-2852 [FREE Full text] [doi: 10.1097/JS9.0000000000000513] [Medline: 37222700]
45. Fan VY, Cash R, Bertozzi S, Pate M. The when is less important than the what: an epidemic scale as an alternative to the WHO's Public Health Emergency of International Concern. Lancet Glob Health 2023 Oct;11(10):e1499-e1500 [FREE Full text] [doi: 10.1016/S2214-109X(23)00314-5] [Medline: 37598696]
46. Torner N. The end of COVID-19 public health emergency of international concern (PHEIC): and now what? Vacunas 2023 Jul;24(3):164-165. [doi: 10.1016/j.vacune.2023.05.001]
47. Wise J. COVID-19: WHO declares end of global health emergency. BMJ 2023 May 09;381:1041. [doi: 10.1136/bmj.p1041] [Medline: 37160309]
48. Harris E. WHO declares end of COVID-19 global health emergency. JAMA 2023 Jun 06;329(21):1817. [doi: 10.1001/jama.2023.8656] [Medline: 37195734]
49. Okesanya OJ, Manirambona E, Buban J, Olabode ON, Lucero-Prisno DE. Coronavirus disease 2019 emergency is over but the pandemic is not: implications for a new global order. Int J Surg Glob Health 2023;6(4):e0207. [doi: 10.1097/gh9.0000000000000207]
50. Lenharo M. WHO declares end to COVID-19's emergency phase. Nature 2023 May 05:882. [doi: 10.1038/d41586-023-01559-z] [Medline: 37147368]
51. Sarker R, Roknuzzaman AS, Nazmunnahar, Hossain MJ, Islam MR. Benefits and probable ill effects of WHO's declaration of end of COVID-19 pandemic: a way back to pandemic-free normal life. Ann Med Surg (Lond) 2023 Jun;85(6):3199-3201 [FREE Full text] [doi: 10.1097/MS9.0000000000000848] [Medline: 37363491]
52. Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Gavrilov D, Giattino C, et al. Coronavirus pandemic (COVID-19). Our World in Data. URL: https://ourworldindata.org/coronavirus [accessed 2024-04-29]
53. Croissant Y, Millo G. Panel data econometrics in R: the plm package. J Stat Soft 2008;27(2):1-43. [doi: 10.18637/jss.v027.i02]
54. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2023. URL: http://www.R-project.org/ [accessed 2024-04-14]
55. Hansen LP. Large sample properties of generalized method of moments estimators. Econometrica 1982 Jul;50(4):1029. [doi: 10.2307/1912775]
56. Eales O, Ainslie KE, Walters CE, Wang H, Atchison C, Ashby D, et al. Appropriately smoothing prevalence data to inform estimates of growth rate and reproduction number. Epidemics 2022 Sep;40:100604 [FREE Full text] [doi: 10.1016/j.epidem.2022.100604] [Medline: 35780515]
57. Elliott P, Eales O, Steyn N, Tang D, Bodinier B, Wang H, et al. Twin peaks: the Omicron SARS-CoV-2 BA.1 and BA.2 epidemics in England. Science 2022 Jun 24;376(6600):eabq4411 [FREE Full text] [doi: 10.1126/science.abq4411] [Medline: 35608440]
58. Shaddick G, Zidek JV, Schmidt AM. Spatio–Temporal Methods in Environmental Epidemiology with R. New York, NY: CRC Press; 2023.
59. Global influenza surveillance and response system (GISRS). Global Initiative on Sharing All Influenza Data (GISAID). URL: https://www.gisaid.org/ [accessed 2024-04-29]
60. Khare S, Gurry C, Freitas L, Schultz MB, Bach G, Diallo A, et al. GISAID's role in pandemic response. China CDC Wkly 2021 Dec 03;3(49):1049-1051 [FREE Full text] [doi: 10.46234/ccdcw2021.255] [Medline: 34934514]
61. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data - from vision to reality. Euro Surveill 2017 Mar 30;22(13):30494 [FREE Full text] [doi: 10.2807/1560-7917.ES.2017.22.13.30494] [Medline: 28382917]
62. Nasereddin A, Golan Berman H, Wolf DG, Oiknine-Djian E, Adar S. Identification of SARS-CoV-2 variants of concern using amplicon next-generation sequencing. Microbiol Spectr 2022 Aug 31;10(4):e0073622 [FREE Full text] [doi: 10.1128/spectrum.00736-22] [Medline: 35758686]
63. Huddleston J, Hadfield J, Sibley T, Lee J, Fay K, Ilcisin M, et al. Augur: a bioinformatics toolkit for phylogenetic analyses of human pathogens. J Open Source Softw 2021 Jan;6(57):2906 [FREE Full text] [doi: 10.21105/joss.02906] [Medline: 34189396]
64. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. Nat Microbiol 2020 Nov 15;5(11):1403-1407 [FREE Full text] [doi: 10.1038/s41564-020-0770-5] [Medline: 32669681]
65. O'Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. Virus Evol 2021;7(2):veab064 [FREE Full text] [doi: 10.1093/ve/veab064] [Medline: 34527285]
66. Arellano M, Bond S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Econ Stud 1991 Apr;58(2):277-297. [doi: 10.2307/2297968]
67. Hu M. Pandemic Surveillance: Privacy, Security, and Data Ethics Elgar Law, Technology and Society Series. New York, NY: Monograph Book; 2022.

XSL•FO
RenderX

68. Levenson E, Firger J. What the CDC's 'substantial' and 'high' levels of COVID-19 transmission actually mean. CNN Health. 2021. URL: https://edition.cnn.com/2021/07/28/health/substantial-or-high-covid-19-transmission-wellness/index.html [accessed 2024-04-29]
69. Christie A, Brooks JT, Hicks LA, Sauber-Schatz EK, Yoder JS, Honein MA, et al. Guidance for implementing COVID-19 prevention strategies in the context of varying community transmission levels and vaccination coverage. MMWR Morb Mortal Wkly Rep 2021 Jul 27;70(30):1044-1047 [FREE Full text] [doi: 10.15585/mmwr.mm7030e2] [Medline: 34324480]
70. Stoto MA. Public health assessment in the 1990s. Annu Rev Public Health 1992;13:59-78. [doi: 10.1146/annurev.pu.13.050192.000423] [Medline: 1599602]
71. Lundberg AL, Lorenzo-Redondo R, Ozer EA, Hawkins CA, Hultquist JF, Welch SB, et al. Has Omicron changed the evolution of the pandemic? JMIR Public Health Surveill 2022 Jan 31;8(1):e35763 [FREE Full text] [doi: 10.2196/35763] [Medline: 35072638]
72. Coronavirus disease (COVID-19) dashboard. World Health Organization. URL: https://covid19.who.int/table [accessed 2024-04-29]
73. Post L, Wu SA, Soetikno AG, Ozer EA, Liu Y, Welch SB, et al. Updated surveillance metrics and history of the COVID-19 pandemic (2020-2023) in Latin America and the Caribbean: longitudinal trend analysis. JMIR Public Health Surveill 2024 May 17;10:e44398 [FREE Full text] [doi: 10.2196/44398] [Medline: 38568194]
74. Lundberg AL, Wu SA, Soetikno AG, Hawkins C, Murphy RL, Havey RJ, et al. Updated surveillance metrics and history of the COVID-19 pandemic (2020-2023) in Europe: longitudinal trend analysis. JMIR Public Health Surveill 2024 Jun 21;10:e53551 [FREE Full text] [doi: 10.2196/53551] [Medline: 38568186]
75. Soetikno AG, Lundberg AL, Ozer EA, Wu SA, Welch SB, Mason M, et al. Updated surveillance metrics and history of the COVID-19 pandemic (2020-2023) in the Middle East and North Africa: longitudinal trend analysis. JMIR Public Health Surveill 2024 Jun 12;10:e53219 [FREE Full text] [doi: 10.2196/53219] [Medline: 38568184]
76. Woodruff J, Warsi Z. COVID rapidly spreads in China as government eases strict quarantine rules. Public Broadcasting Service. URL: https://tinyurl.com/yc3xb293 [accessed 2024-04-29]
77. Landoni G, Maimeri N, Fedrizzi M, Fresilli S, Kuzovlev A, Likhvantsev V, et al. Why are Asian countries outperforming the western world in controlling COVID-19 pandemic? Pathog Glob Health 2021 Feb 26;115(1):70-72 [FREE Full text] [doi: 10.1080/20477724.2020.1850982] [Medline: 33241776]
78. Acter T, Uddin N, Das J, Akhter A, Choudhury TR, Kim S. Evolution of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as coronavirus disease 2019 (COVID-19) pandemic: a global health emergency. Sci Total Environ 2020 Aug 15;730:138996 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.138996] [Medline: 32371230]
79. To KK, Sridhar S, Chiu KH, Hung DL, Li X, Hung IF, et al. Lessons learned 1 year after SARS-CoV-2 emergence leading to COVID-19 pandemic. Emerg Microbes Infect 2021 Dec;10(1):507-535 [FREE Full text] [doi: 10.1080/22221751.2021.1898291] [Medline: 33666147]
80. Khan M, Adil SF, Alkhathlan HZ, Tahir MN, Saif S, Khan M, et al. COVID-19: a global challenge with old history, epidemiology and progress so far. Molecules 2020 Dec 23;26(1):39 [FREE Full text] [doi: 10.3390/molecules26010039] [Medline: 33374759]
81. Jennings R. How cultural differences help Asian countries beat COVID-19, while US struggles. Voice of America (VOA). URL: https://www.insidevoa.com/p/5831.html [accessed 2024-04-29]
82. Brant R. Shanghai Covid: China announces largest city-wide lockdown. BBC News. 2022 Mar 27. URL: https://www.bbc.com/news/world-asia-china-60893070 [accessed 2024-04-29]
83. Woo R. China's Chengdu exits full citywide Covid lockdown on Thursday. Reuters. 2022 Sep 14. URL: https://www.reuters.com/world/china/chinas-chengdu-lift-citywide-covid-lockdown-thursday-2022-09-14/ [accessed 2022-09-14]
84. China ends 76-day lockdown of Wuhan city, where the virus outbreak started. CNBC. 2020 Apr 07. URL: https://www.cnbc.com/2020/04/08/coronavirus-china-ends-lockdown-of-wuhan-city-where-outbreak-started.html [accessed 2024-04-29]
85. What is China's 'zero covid' policy and why did it trigger protests? The Washington Post. 2022 Nov 09. URL: https://www.washingtonpost.com/world/2022/11/29/china-zero-covid-policy-protests-explained/ [accessed 2024-04-29]
86. Chen SC. Taiwan's experience in fighting COVID-19. Nat Immunol 2021 Apr 26;22(4):393-394. [doi: 10.1038/s41590-021-00908-2] [Medline: 33772224]
87. Qin A, Chien AC. Covid? What Covid? Taiwan thrives as a bubble of normality. The Washington Post. 2021 Aug 23. URL: https://www.nytimes.com/2021/03/13/world/asia/taiwan-covid.html [accessed 2024-04-29]
88. Frost N. Battling delta, New Zealand abandons its zero-covid ambitions. The NewYork Times. 2021 Oct 04. URL: https://www.nytimes.com/2021/10/04/world/australia/new-zealand-covid-zero.html [accessed 2024-04-29]
89. Gan N. How Vietnam managed to keep its coronavirus death toll at zero. CNN World. 2020 May 30. URL: https://edition.cnn.com/2020/05/29/asia/coronavirus-vietnam-intl-hnk/index.html [accessed 2024-04-29]
90. Willoughby E. An ideal public health model? Vietnam's state-led, preventative, low-cost response to COVID-19. Brookings. 2021 Jun 29. URL: https://www.brookings.edu/blog/order-from-chaos/2021/06/29/an-ideal-public-health-model-vietnams-state-led-preventative-low-cost-response-to-covid-19/ [accessed 2024-04-29]
91. Indonesia: weak public health response to COVID-19. Human Rights Watch. 2021 Jan 13. URL: https://www.hrw.org/news/2021/01/13/indonesia-weak-public-health-response-covid-19 [accessed 2024-04-29]

XSL•FO
RenderX

92. Sullivan M. The Philippines becomes coronavirus hot spot in Southeast Asia. National Public Radio. 2020 Oct 08. URL: https://www.npr.org/2020/08/10/900279722/the-philippines-becomes-coronavirus-hot-spot-in-southeast-asia [accessed 2024-04-29]
93. Philippines now has worst COVID-19 outbreak in Southeast Asia. The Japan Times. 2020 Aug 07. URL: https://www.japantimes.co.jp/news/2020/08/07/asia-pacific/philippines-coronavirus-southeast-asia/ [accessed 2024-04-29]
94. Amit AM, Pepito VC, Dayrit MM. Early response to COVID-19 in the Philippines. Western Pac Surveill Response J 2021 Mar 31;12(1):56-60 [FREE Full text] [doi: 10.5365/wpsar.2020.11.1.014] [Medline: 34094626]
95. -. COVID-19: an ongoing public health crisis in the Philippines. Lancet Reg Health West Pac 2021 Apr;9:100160 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100160] [Medline: 33937888]
96. Ryall J. Many Pacific Island nations have kept Covid at bay. Deutsche Welle. 2020 May 19. URL: https://www.dw.com/en/zero-cases-how-pacific-islands-kept-coronavirus-at-bay/a-53495263 [accessed 2024-04-29]
97. Addressing COVID-19's impacts and structural crisis in the Pacific. The Food and Agriculture Organization. 2020 May 06. URL: https://www.fao.org/newsroom/detail/Addressing-COVID-19-s-impacts-and-structural-crisis-in-the-Pacific/en [accessed 2024-04-29]
98. McClure H. Surviving the pandemic: COVID-19 in the Pacific and health security. Development Bulletin. URL: https://pacificsecurity.net/wp-content/uploads/2021/10/DB82_Part17.pdf [accessed 2024-04-29]
99. The Solomon Islands reports its first coronavirus case, a student returning from the Philippines. SBS News. 2020 Oct 04. URL: https://www.sbs.com.au/news/article/the-solomon-islands-reports-its-first-coronavirus-case-a-student-returning-from-the-philippines/x20eytmn1 [accessed 2024-04-30]
100. Rising D. Solomon Islands' 1st virus outbreak causes growing concern. The Associated Press. 2022 Feb 17. URL: https://apnews.com/article/coronavirus-pandemic-health-solomon-islands-honiara-30d3df8cb0f809c9e3e5a0cc5c6f886f [accessed 2024-04-29]
101. McCurry J. Japan urges ‘highest level of vigilance’ as Omicron subvariant drives record Covid surge. The Gurdian. 2022 Jul 22. URL: https://www.theguardian.com/world/2022/jul/22/japan-urges-highest-level-of-vigilance-as-omicron-subvariant-drives-record-covid-surge [accessed 2024-04-29]
102. Suhartono M, Beech H. Thai protesters are back, and angrier, as government fumbles on Covid. The NewYork Times. 2021 Sep 05. URL: https://www.nytimes.com/2021/09/05/world/asia/thailand-protesters-covid.html [accessed 2024-04-29]
103. Malaysia: hundreds take to the streets in anti-government protest. Al Jazeera Media Network. 2021 Jul 21. URL: https://www.aljazeera.com/news/2021/7/31/hundreds-take-to-streets-in-anti-government-protest-in-malaysia [accessed 2024-04-29]
104. Philippines: president Duterte gives “shoot to kill” order amid pandemic response. Amnesty International. 2020 Apr 02. URL: https://www.amnesty.org/en/latest/news/2020/04/philippines-president-duterte-shoot-to-kill-order-pandemic/ [accessed 2024-04-29]
105. Lo B, Ott H. The Philippines has the worst coronavirus outbreak in Southeast Asia, and critics warn the government could use it to crack down on dissent. CBS News. 2020 Aug 07. URL: https://www.cbsnews.com/news/covid-19-philippines-worst-coronavirus-outbreak-in-southeast-asia-crackdown-on-dissent/ [accessed 2024-04-29]
106. Branswell H. Senior WHO official faults China for undercounting Covid deaths. STAT News. 2023 Jun 04. URL: https://www.statnews.com/2023/01/04/who-official-faults-china-undercounting-covid-deaths/ [accessed 2024-04-29]
107. China might be underreporting its COVID-related death toll. Euronews. 2022 Dec 21. URL: https://www.euronews.com/2022/12/21/china-might-be-underreporting-its-covid-related-death-toll-experts [accessed 2024-04-29]
108. Anti-mandate protesters converge on New Zealand parliament. VOA News. 2022 Aug 23. URL: https://www.voanews.com/a/anti-mandate-protesters-converge-on-new-zealand-parliament-/6712836.html [accessed 2024-04-29]
109. GDP growth (annual %) - East Asia and Pacific. World Bank Group. URL: https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?locations=Z4 [accessed 2024-04-29]
110. Policy responses to COVID-19. International Monetary Fund. URL: https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19#C [accessed 2024-04-29]
111. Corum J, Zimmer C. How the Sinovac vaccine works. The NewYork Times. 2021 May 07. URL: https://www.nytimes.com/interactive/2020/health/sinovac-covid-19-vaccine.html [accessed 2024-04-29]
112. The Sinovac-CoronaVac COVID-19 vaccine: what you need to know. World Health Organization. 2022 Jun 10. URL: https://www.who.int/news-room/feature-stories/detail/the-sinovac-covid-19-vaccine-what-you-need-to-know#:~:text=The%20vaccine%20is%20safe%20and,COVID%2D19%20in%20the%20past [accessed 2024-04-29]
113. Lyudmila K. Vaccination with CoronaVac launched in Ukraine on April 13 – Health minister. UNIAN. 2021 Apr 13. URL: https://www.unian.info/society/covid-19-vaccination-with-coronavac-launched-in-ukraine-on-april-13-health-minister-11386060.html [accessed 2024-04-29]
114. Miranda NA. Chile receives 2 million-dose first delivery of Sinovac COVID-19 vaccine. Reuters. 2021 Jan 28. URL: https://www.reuters.com/business/healthcare-pharmaceuticals/chile-receives-2-million-dose-first-delivery-sinovac-covid-19-vaccine-2021-01-28/ [accessed 2024-04-29]

XSL•FO
RenderX

115. Mano A, Simões E. Chinese vaccine draws demand across Latin America, say Brazilian officials. Reuters. 2020 Dec 10. URL: https://www.reuters.com/article/business/healthcare-pharmaceuticals/chinese-vaccine-draws-demand-across-latin-america-say-brazilian-officials-idUSKBN28K1IJ/ [accessed 2024-04-29]
116. Nguyen A. Thailand finally kicks off COVID vaccinations: 5 things to know. NIKKEI Asia. 2021 Mar 01. URL: https://asia.nikkei.com/Spotlight/Coronavirus/COVID-vaccines/Thailand-finally-kicks-off-COVID-vaccinations-5-things-to-know [accessed 2024-04-29]
117. Indonesia starts mass COVID vaccinations over vast territory. The Associated Press. 2021 Jan 13. URL: https://apnews.com/article/asia-pacific-indonesia-coronavirus-pandemic-joko-widodo-3ff30c85dc6943e2a06db470a9d04931 [accessed 2024-04-29]
118. Brazil health ministry says plans to order 30 million more Coronavac doses. Reuters. 2021 Feb 19. URL: https://www.reuters.com/article/business/healthcare-pharmaceuticals/brazil-health-ministry-says-plans-to-order-30-million-more-coronavac-doses-idUSKBN2AJ1D3/ [accessed 2024-04-29]
119. Takenaka K. Japan begins COVID-19 vaccination in 'first major step' to halt pandemic. Reuters. 2021 Feb 16. URL: https://www.reuters.com/article/world/japan-begins-covid-19-vaccination-in-first-major-step-to-halt-pandemic-idUSKBN2AH00O/ [accessed 2024-04-29]
120. Cha S. S.Korea to secure more mRNA vaccines for COVID-19 booster shot. Reuters. 2021 Jun 29. URL: https://www.reuters.com/world/asia-pacific/skorea-secure-more-mrna-vaccines-covid-19-booster-shot-2021-06-29/ [accessed 2024-04-29]
121. Lee IC. Taiwan begins COVID-19 vaccinations. Taipei Times. 2021 Mar 23. URL: https://www.taipeitimes.com/News/front/archives/2021/03/23/2003754315 [accessed 2024-04-24]
122. Taiwan clears AstraZeneca vaccine, shots might start on Monday. Al Jazeera. 2021 Mar 18. URL: https://www.aljazeera.com/news/2021/3/18/taiwan-clears-astrazeneca-vaccine-shots-might-start-on-monday [accessed 2024-04-29]
123. Travelers rush to enter China as it eases pandemic border restrictions. PBS News. 2023 Jan 08. URL: https://www.pbs.org/newshour/world/travelers-rush-to-enter-china-as-it-eases-pandemic-border-restrictions [accessed 2024-04-29]
124. Coronavirus (COVID-19) measures by the government. Japan National Tourism Organization. 2023 May 09. URL: https://www.japan.travel/en/practical-coronavirus-information/government-measures/ [accessed 2024-04-29]
125. COVID-19 and the border. Government of Australia. 2023 Mar 11. URL: https://www.homeaffairs.gov.au/portfolio-functions/travel-and-crossing-the-border/covid-19-and-the-border [accessed 2024-04-29]
126. You want to travel to New Zealand. New Zealand Customs. URL: https://covid19.govt.nz/international-travel/#:~:text=Travellers%20and%20crew%20do%20not,still%20require%20proof%20of%20vaccination [accessed 2024-04-29]
127. COVID-19 travel information. Ministry of Foreign Affairs Singapore. URL: https://www.mfa.gov.sg/Consular-Services/Singapore-Citizens/COVID-19-Information [accessed 2024-04-29]
128. Stein R. As the pandemic ebbs, an influential COVID tracker shuts down. KERA News. 2023 Feb 10. URL: https://www.keranews.org/2023-02-10/as-the-pandemic-ebbs-an-influential-covid-tracker-shuts-down [accessed 2024-04-29]
129. Ritchey MD, Rosenblum HG, Del Guercio K, Humbard M, Santos S, Hall J, et al. COVID-19 self-test data: challenges and opportunities - United States, October 31, 2021-June 11, 2022. MMWR Morb Mortal Wkly Rep 2022 Aug 12;71(32):1005-1010 [FREE Full text] [doi: 10.15585/mmwr.mm7132a1] [Medline: 35951486]
130. Brown RB. Biases in COVID-19 case and death definitions: potential causes and consequences. Disaster Med Public Health Prep 2022 Dec 12;17:e313 [FREE Full text] [doi: 10.1017/dmp.2022.281] [Medline: 36503702]

Abbreviations

GDP: gross domestic product
GISAID: Global Initiative on Sharing All Influenza Data
OWID: Our World in Data
VOC: variant of concern
WHO: World Health Organization

Edited by A Mavragani; submitted 03.10.23; peer-reviewed by S Thawillarp, A Coristine; comments to author 09.10.24; revised version received 06.01.25; accepted 12.01.25; published 21.02.25.

Please cite as:
Lundberg AL, Soetikno AG, Wu SA, Ozer E, Welch SB, Liu Y, Hawkins C, Mason M, Murphy R, Havey RJ, Moss CB, Achenbach CJ, Post LA
Updated Surveillance Metrics and History of the COVID-19 Pandemic (2020-2023) in East Asia and the Pacific Region: Longitudinal Trend Analysis
JMIR Public Health Surveill 2025;11:e53214
URL: https://publichealth.jmir.org/2025/1/e53214
doi:10.2196/53214
PMID:39804185

XSL•FO
RenderX

©Alexander L Lundberg, Alan G Soetikno, Scott A Wu, Egon Ozer, Sarah B Welch, Yingxuan Liu, Claudia Hawkins, Maryann Mason, Robert Murphy, Robert J Havey, Charles B Moss, Chad J Achenbach, Lori Ann Post. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Availability and Key Characteristics of National Early Warning Systems for Emerging Profiles of Antimicrobial Resistance in High-Income Countries: Systematic Review

Jessica Iera[1,2], MD; Claudia Isonne[1,3], MD; Chiara Seghieri[2], PhD; Lara Tavoschi[4], PhD; Mariateresa Ceparano[1], BSc; Antonio Sciurti[1], MD; Alessia D'Alisera[5], MD; Monica Sane Schepisi[5], MSc, MD; Giuseppe Migliara[6], MD, PhD; Carolina Marzuillo[1], BSc; Paolo Villari[1], MPH, MD; Fortunato D'Ancona[3], MD; Valentina Baccolini[1], MPH, MD

[1]Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy

[2]Department EMbeDS, Sant'Anna School of Advanced Studies, Management and Health Laboratory, Institute of Management, Pisa, Italy

[3]Department of Infectious Diseases, Istituto Superiore di Sanità, Rome, Italy

[4]Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy

[5]Italian Ministry of Health, General Directorate for Health Prevention, Rome, Italy

[6]Department of Life Sciences, Health and Health Professions, Link Campus University, Rome, Italy

Corresponding Author:
Jessica Iera, MD
Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy

Abstract

Background: The World Health Organization (WHO) recently advocated an urgent need for implementing national surveillance systems for the timely detection and reporting of emerging antimicrobial resistance (AMR). However, public information on the existing national early warning systems (EWSs) is often incomplete, and a comprehensive overview on this topic is currently lacking.

Objective: This review aimed to map the availability of EWSs for emerging AMR in high-income countries and describe their main characteristics.

Methods: A systematic review was performed on bibliographic databases, and a targeted search was conducted on national websites. Any article, report, or web page describing national EWSs in high-income countries was eligible for inclusion. EWSs were identified considering the emerging AMR-reporting WHO framework.

Results: We identified 7 national EWSs from 72 high-income countries: 2 in the East Asia and Pacific Region (Australia and Japan), 3 in Europe and Central Asia (France, Sweden, and the United Kingdom), and 2 in North America (the United States and Canada). The systems were established quite recently; in most cases, they covered both community and hospital settings, but their main characteristics varied widely across countries in terms of the organization and microorganisms under surveillance, with also different definitions of emerging AMR and alert functioning. A formal system assessment was available only in Australia.

Conclusions: A broader implementation and investment of national surveillance systems for the early detection of emerging AMR are still needed to establish EWSs in countries and regions lacking such capabilities. More standardized data collection and reporting are also advisable to improve cooperation on a global scale. Further research is required to provide an in-depth analysis of EWSs, as this study is limited to publicly available data in high-income countries.

(JMIR Public Health Surveill 2025;11:e57457) doi:10.2196/57457

KEYWORDS

early warning system; surveillance; emerging AMR; high-income countries; antimicrobial resistance

Introduction

Antimicrobial resistance (AMR) is a global public health emergency [1,2] that increases the number of deaths, length of hospital stays, and health care costs expenditure [3-8]. Worldwide, a total of 4.95 million deaths per year are associated with bacterial AMR, 1.27 million of which are directly attributable to it [9]. Gram-negative bacteria resistant to last-resort antibiotics such as carbapenem-resistant *Acinetobacter baumannii* and carbapenem-resistant and third-generation cephalosporin-resistant *Enterobacterales* have been listed as of critical priority for public health measures, due to their significant global burden and the ability to transfer resistance genes [10]. As the threat of AMR continues to grow, more efforts are needed to strengthen infection prevention and

XSL•FO
RenderX

control strategies as well as to enhance surveillance systems [11].

AMR surveillance plays a pivotal role in strengthening health system resilience and preparedness, thereby supporting effective antibiotic stewardship and leading to optimized patient health outcomes [12]. Enhanced investments are considered essential to improve the comparability, quantity, and quality of AMR data [13]. Current trends, such as the rising prevalence of carbapenem-resistant *Acinetobacter spp*. isolates, underline the need to intensify efforts for the early detection of drug resistance [13]. Several international initiatives aimed at improving epidemiological and microbiological discussion for coordinated actions are ongoing [14,15]. Within this context, the World Health Organization (WHO) launched the Global Antimicrobial Resistance and Use Surveillance System (GLASS) in 2015 [16,17], included the development of national surveillance systems capable of timely detecting and reporting emerging resistance among the goals outlined in the Global Plan of Action on AMR published in 2016 [18], and more recently, developed a focused surveillance with a specific reporting component for novel (emerging) AMR (ie, GLASS-EAR), in order to support early data sharing and coordinated actions among member states [19].

However, while well-established routine AMR surveillance systems are available in high-income countries (HIC), implementing AMR surveillance has been particularly challenging in low- and middle-income countries, where the number of surveillance sites contributing to national surveillance is often not representative [20,21], and a limited number of referral hospitals report AMR data to GLASS [22]. Therefore, studies have mainly focused on countries where the laboratory capabilities required to strengthen national surveillance systems are more likely to be in place (ie, HIC) and publicly available information on the implementation of tools for the early detection and reporting of emerging AMR profiles is often fragmented [23]. Even when findings are available, understanding these systems is challenging due to different approaches used for data collection, reporting, and interpretation of definitions [24,25]. Given the lack of a clear and comprehensive overview on the topic, the aim of our study was to map existing national early warning systems (EWSs) for emerging AMR in HIC as well as describe and compare their main characteristics. Findings from our study could provide relevant information to stakeholders engaged in AMR surveillance, informing the development and/or optimization of EWSs in several countries, including low- and middle-income countries.

Methods

Ethical Considerations

This study was performed according to the Cochrane Handbook for Systematic Reviews and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement (Checklist 1) [26,27]. The protocol was registered on the Open Science Framework via the Center for Open Science (210 Ridge McIntire Road, Suite 500, Charlottesville, VA 22903 - 5083; identifier: pr6a8). As the study did not involve the collection of primary data, it did not require informed consent or the submission for institutional review board approval.

Search Strategy and Study Selection

Three bibliographic databases (ie, PubMed, Web of Science, and Scopus) were searched. Search strings were adapted to fit the search criteria of each database (Table S1 in Multimedia Appendix 1). Taking into account the topic under investigation and the expected challenges in retrieving information, we intentionally developed broad search strings, prioritizing sensitivity over specificity. The search was conducted among records published from database inception to July 3, 2024, without restrictions such as language or date. After the removal of duplicate articles, language restriction was applied in the context of the title and abstract of all the retrieved records during screening, according to the eligibility criteria described below. The full texts of potentially relevant articles were examined by three researchers, who resolved disagreements through discussion and recorded reasons for exclusion. In addition, relevant national websites (ie, the Ministries of Health and/or National Health Institutes websites) were also explored, focusing on thematic areas involving AMR surveillance (Table S2 in Multimedia Appendix 1).

Eligibility Criteria

Considering the Emerging Antimicrobial Resistance Reporting framework within the GLASS developed by the WHO [19], EWSs were defined, for the purposes of this study, as any system able to timely detect, provide verification, and report emerging AMR events. This broad definition was adopted as we expected to encounter systems with diverse characteristics, including variations in design, data flow requirements, implementation, and functionality. Emerging AMR was defined as unusual AMR findings in bacteria and fungi causing infections in humans with the potential impact on public health, such as new types of phenotypic resistance (not previously reported or very rare) and new genetic determinants of AMR that may have high potential for spread. As a specific resistant phenotype/genotype may be emerging in one area but may already be endemic in another one, a list of AMR events was not developed by considering both newly detected AMR and known resistance profiles detected for the first time in new geographical areas.

We included articles, web pages, and/or technical reports with the following characteristics: (1) reported in English, French, Spanish, or Italian, aiming at broadening our investigation by taking into account the language capabilities of the co-authors; (2) describing national EWSs for emerging AMR currently implemented or under development; and (3) referring to HIC, as identified by the World Bank classification of countries by income [28]. We excluded articles, web pages, and/or technical reports that (1) described only standard national AMR surveillance systems, for example routine surveillance systems that retrospectively analyzed data on an annual basis and/or that lacked active alert components; (2) focused on EWSs limited to subnational levels; or (3) reported data in languages other than English, French, Spanish, or Italian.

XSL•FO
RenderX

Data Collection and Synthesis

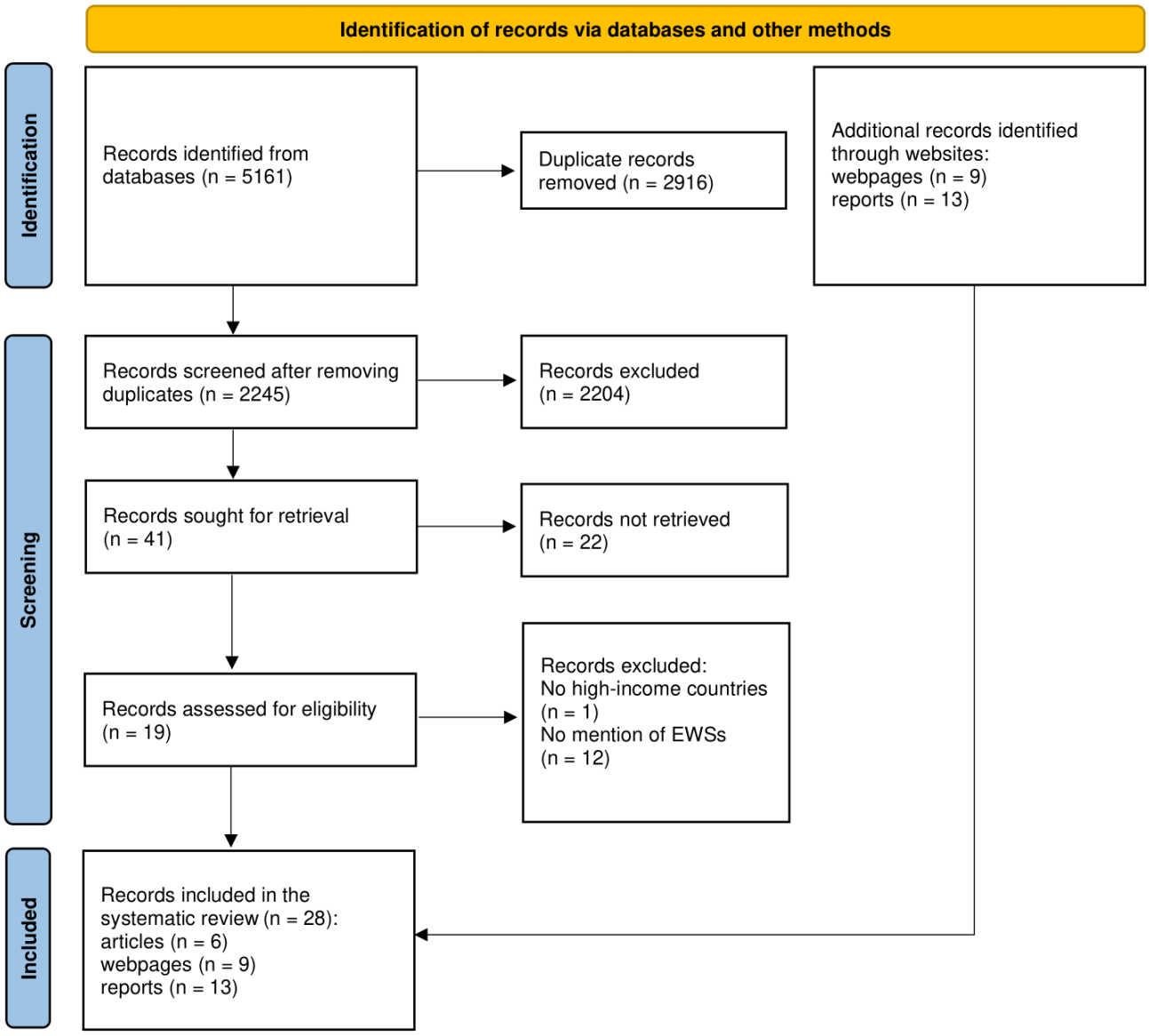
For each record, two reviewers independently retrieved the following information using a standardized data abstraction form: identification of the report/study, if applicable (title, first author, year of publication, and DOI); country; key characteristics of the EWS (eg, coordinating institution, year of establishment, sector involved—human/animal/food/environment, setting where specimens are collected, specimen types, and availability of genomic data); key characteristics of the microorganisms under surveillance (eg, resistance profiles); alert timing (eg, real-time, daily, or weekly); and conduction of performance evaluation of the EWS. Countries were grouped by region, according to the World Bank classification. Data were descriptively synthesized through tables and graphical representations using R software (version 4.3.1; R Foundation for Statistical Computing).

Results

Records Identification and Screening

Overall, 5161 records were identified via database searching (Figure 1). After duplicate removal and screening by the title and abstract, 41 articles were eligible for full-text analysis. Of these, 35 were excluded with reasons, providing a total of 6 articles ultimately included in the systematic review [24,29-33]. A targeted search on relevant national websites allowed for the identification of thematic areas involving AMR from 72 out of 86 HIC (15 countries were excluded due to language restrictions or the inability to retrieve websites), with 22 records finally included in the systematic review (ie, 9 web pages [34-42] and 13 reports [25,43-54]).

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the review process. EWSs: early warning systems.



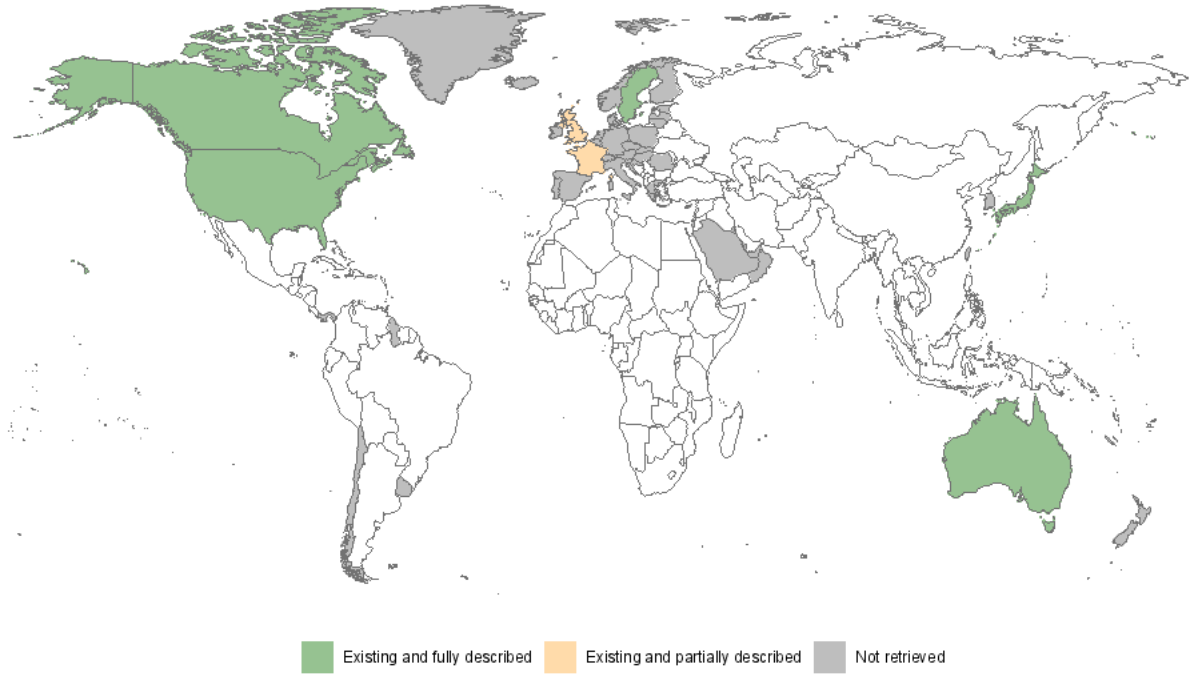
XSL•FO
RenderX

Availability of National EWSs for Emerging AMR in HIC

Overview

Nine web pages and 19 publications described 7 national EWSs in 72 (9.7%) HIC (Figure 2). Of these, 2 were in the East Asia and Pacific Region (ie, Australia [29,34,44,45,53,54] and Japan [24,30,31,33]), 3 were in Europe and Central Asia (ie, France [36-38,48,49], Sweden [24,39,51,52], and the United Kingdom [32,40,46]), and 2 were in North America (ie, United States [25,41,42,50] and Canada [35,43,47]). A partial description of the systems was available in 2 cases (ie, the United Kingdom and France).

Figure 2. Availability of early warning systems for emerging antimicrobial resistance at the national level in high-income countries.



East Asia and Pacific Region

Australia: National Alert System for Critical Antimicrobial Resistances

The National Alert System for Critical Antimicrobial Resistances (CAR-Alert) was established by the Australian Commission on Safety and Quality in Health Care in March 2016 as part of the Antimicrobial Use and Resistance in Australia (ie, AURA) Surveillance System and reviewed in 2018 and 2022 (Table 1) [54]. By collecting data on nationally agreed priority microorganisms with critical resistance to last-line antimicrobial agents, it is a coordinated national system that allows both the collection and the communication of information on confirmed critical antimicrobial resistance (CAR) and potential outbreaks, as close as possible to the time of confirmation [29,34]. At the national level, 28 confirming laboratories provided data to the system in 2021 and 2022, with at least 1 confirming laboratory in each state and territory [53]. In 2023, 26 confirming laboratories participated in the CAR-Alert [54].

XSL•FO
RenderX

Table . National EWSs[a] for emerging AMR[b] identified in high-income countries.

Country	EWS	Institution	Year of establishment	Brief description
East Asia and Pacific Region				
Australia	CAR-Alert National Alert System for Critical Antimicrobial Resistances	Australian Commission on Safety and Quality in Health Care	2016 Revised in 2018 and 2022	CAR-Alert allows for the timely surveillance of critical nationally agreed antimicrobial resistances, integrating the AURA[c] surveillance system. The system involves laboratories conducting confirmatory susceptibility testing on human isolates.
Japan	JANIS-CL Japan nosocomial infection surveillance-Clinical Laboratory division	National Institute of Infectious Diseases	2000 Revised in 2007	The system involves several laboratories across the country. Alert emails are automatically sent to contact personnel of facilities that report unusual combinations of bacterial species and antimicrobial susceptibility, defined as “unusual AMR.”
North America				
Canada	AMRnet Canadian Antimicrobial Resistance Network	Public Health Agency of Canada	NA[d]	Laboratory-based surveillance system capturing information on AMR susceptibility testing from laboratory information systems in both public and private clinical and veterinary laboratories, including reference laboratories. The system aims at collecting and disseminating real time data to allow stakeholders response to emerging AMR issues.
United States	AR Lab Network Antimicrobial Resistance Laboratory Networks	Center for Disease Control and Prevention	2016	Provides nationwide lab capacity to rapidly detect antimicrobial resistance. Laboratories cooperate to rapidly identify AMR profiles causing hard-to-treat or potentially untreatable infections. Protocols to immediately notify health department, health care provider, and infection control staff of unusual resistance are provided.
Europe and Central Asia				

XSL•FO
RenderX

Country	EWS	Institution	Year of establishment	Brief description
France	e-SIN Signalement Externe des Infections Nosocomiales	Ministry of Health	2011 Revised in 2017	Application that allows the monitoring of infectious events in health care facilities, providing data on isolated cases/outbreaks of bacteria with particular AMR profiles. A specific reporting form for emerging highly resistant bacteria is available since 2017. The National Reference Center for Antibiotic Resistance is integrated into this alert system, in particular for events involving emerging mechanisms of resistance. Reports may result in national or international health alerts.
Sweden	Svebar	Public Health Agency of Sweden	2015	Svebar consists of an IT system for early alerts and continuous resistance monitoring. All culture findings from the country's laboratories are automatically transferred on a daily basis, allowing an early warning on findings of serious antibiotics resistance.
United Kingdom	Antimicrobial resistance alerts (ARAs)	Public Health England Antimicrobial resistance and healthcare associated infections reference unit (AMRHAI)	NA	AMRHAI detects new and emerging AMR via interpretive analysis of MIC[e] profiles/molecular investigation. Unusual isolates are sent for testing from diagnostic laboratories to AMRHAI, where an appropriate reference unit confirms relevant isolates resistance. If deemed of public health impact, ARAs inform microbiologists of emerging AMR pathogen strains that could spread in the United Kingdom health service.

[a]EWS: early warning system.

[b]AMR: antimicrobial resistance.

[c]AURA: Antimicrobial Use and Resistance in Australia.

[d]NA: not available.

[e]MIC: minimum inhibitory concentration.

In order to identify and confirm potential CAR, pathology laboratories perform a routine three-step process based on (1) isolate collection from the patient and routine testing; (2) confirmation by a laboratory with the capacity to identify CAR, if a critical AMR profile is suspected by the originating laboratory; and (3) data submission to the CAR-Alert web portal by the confirming laboratory and feedback to the originating laboratory, allowing the implementation of proper infection prevention and control by the health service caring for the patient [44,45].

Japan: Japan Nosocomial Infections Surveillance - Clinical Laboratory Division

The Ministry of Health, Labour, and Welfare of Japan provided two national surveillance systems to monitor AMR in bacteria, including the Japan Nosocomial Infections Surveillance (ie, JANIS) (Table 1) [31]. JANIS was launched in 2000 as a voluntary surveillance system focusing on infections in health care settings and includes several divisions [30,33]. Among them, the JANIS clinical laboratory division (JANIS-CL) specifically focuses on AMR bacteria, and it collects comprehensive specimen-based data from participating hospitals on a monthly basis through a member-restricted website.

XSL•FO
RenderX

Submitted files from participating hospitals are automatically processed to check data structure, interpret antimicrobial susceptibility, remove duplicates, and calculate the prevalence of AMR [30]. The system has the capacity to detect unusual combinations of bacterial species and antimicrobial susceptibility, defined as unusual AMR, and generates alert emails that are automatically sent to the contact personnel of the facilities [31].

North America

Canada: Antimicrobial Resistance Network

The Antimicrobial Resistance Network (AMRNet) is a collaboration between the Public Health Agency of Canada, provincial and territorial public health organizations, and clinical and veterinary laboratories across the country (Table 1). The system has been designed to detect the emergence and spread of AMR and to timely disseminate data to stakeholders, in order to address emerging AMR issues in Canada [35,47].

Originally started as a series of pilot projects, the AMRNet transitioned to collecting routine data (antimicrobial susceptibility testing results from bacterial and fungal pathogens, together with a defined set of patient or animal characteristics) from a select group of provinces in 2022. Currently, data collection involves approximately 1.5 million bacterial and fungal isolates per year from Ontario, Saskatchewan, and Prince Edward Island [43].

United States: The Antimicrobial Resistance Laboratory Network

The Antimicrobial Resistance Laboratory Network (AR Lab Network) was established and is coordinated by the Centers for Disease Control and Prevention (CDC) to rapidly detect existing and emerging resistance nationwide (Table 1). The network includes 7 regional laboratories, the National Tuberculosis Molecular Surveillance Center, and laboratories across 50 states, 5 cities, and Puerto Rico [25,50]. Laboratories cooperate, from the local to the national level, to provide the rapid identification and response to antimicrobial-resistant genes and microorganisms that cause hard-to-treat or potentially untreatable infections. When unusual resistance is detected, the Healthcare and Clinical Laboratories notify the Health Department. Public health laboratories then confirm the bacterial species identity, perform additional testing to characterize the isolates received, provide notification of unusual resistance to CDC and, when appropriate, send the isolates to regional laboratories for additional testing [42]. Every regional laboratory performs core testing, including the detection of new and emerging threats. If needed, regional laboratories send unusual isolates to the CDC to confirm testing and conduct additional tests. When threats are reported, the CDC provides infection prevention and control recommendations and supports outbreak responses [41].

Europe and Central Asia

France: E-SIN (Signalement Externe Des Infections Nosocomiales - External Reporting on Healthcare-Associated Infections)

Since 2011, the reporting of health care–associated infections, initially introduced in 2001, has been based on a digital tool called e-SIN (Signalement Externe Des Infections Nosocomiales - External Reporting on Healthcare-Associated Infections; Table 1) [48]. The e-SIN application monitors infectious cases reported by health care facilities and allows for the identification of national or even international health alerts, thereby providing data on isolated cases or outbreaks of bacteria with exceptional AMR profiles such as emerging highly resistant bacteria (defined as microorganisms with susceptibility to only one or two classes of antibiotics) [37,38,49]. The reporting process, moreover, enables a simultaneous reporting to both the Regional Coordination Centers for Healthcare-Associated Infections and the Regional Health Agency and includes an internal alert system that allows the notification of the members of the Epidemiology and Hospital Hygiene Team [36].

Sweden: Svebar

The Public Health Agency of Sweden is responsible for national AMR monitoring and analysis (Table 1) [39]. The IT system Svebar was developed in 2015 to expand and strengthen existing surveillance systems both at the national and local levels [24,52]. This system relies on the voluntary participation of clinical microbiology laboratories. All data on clinical isolates from humans are transferred to Svebar by connected laboratories on a daily basis. Every night, the participating laboratories automatically send to the system a file, saved in a short-term storage, with the culture findings from the past 14 days. Data are reported according to standard definitions and saved in a short-term storage for 13 days before being transferred to a long-term storage, which retains the data from the previous night's report, thereby receiving more processed data [51]. The system allows the generation of early alerts and the continuous monitoring of serious antibiotic resistance or other undesired changes in resistance conditions, through pre-set alert functions [51]. Currently, 22 laboratories provide data to Svebar [52].

United Kingdom: Antimicrobial Resistance Alerts

The UK antimicrobial resistance alerts (ARAs) provide microbiologists with information of emerging antimicrobial-resistant pathogen strains that could spread in the UK health service (Table 1) [32]. When a suspicious resistant strain is identified, isolates are sent to Public Health England's AMRHAI reference unit for further testing. The AMRHAI reference unit includes the national reference laboratory for AMR investigation and detection of new and emerging resistances: the unit investigate isolates found to have unusual resistances by diagnostic laboratories, aiming to identify emerging resistance of public health importance, underlying resistance mechanisms and clonal spread of resistant strains [46].

Proposals for ARAs can be submitted by any interested party and should be directed to the co-chairs of a standing committee, which review the proposals and determine its public health

XSL•FO
RenderX

implications, in order to assess the necessity of issuing a national alert [40].

Characteristics of National EWSs for Emerging AMR in HIC

Health Sectors Involved, Specimen Type, and Setting

The human sector was involved in all the identified EWSs, and, in the case of Canada, the system also included the animal sector [47]. All the systems reported data from clinical isolates, while 2 also included screening samples (ie, Australia and the United States; Table 2) [45,50].

XSL•FO
RenderX

Table . Main characteristics of national early warning systems for emerging AMR[a] in high-income countries.

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
East Asia and Pacific Region							

XSL•FO
RenderX

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
Australia (CAR-Alert)	Human	Hospital-based Community-based	Clinical isolate screening		Yes	Weekly summary alert emails	Yes

XSL•FO
RenderX

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
				• Carbapenemase-producing *Acinetobacter baumannii* complex • *Candida auris* • Carbapenemase-producing *Enterobacterales* and/or ribosomal methyltransferase-producing *Enterobacterales*, with transmissible colistin-resistance • Linezolid-resistant *Enterococcus* species • Multidrug-resistant *Mycobacterium tuberculosis* – resistant to at least rifampicin and isoniazid • Ceftriaxone- and/or azithromycin-nonsusceptible *Neisseria gonorrhoeae*; gentamicin-resistant *Neisseria gonorrhoeae* • Ciprofloxacin-nonsusceptible *Neisseria meningitidis* • Carbapenemase-producing *Pseudomonas aeruginosa* • Ceftriaxone-nonsusceptible *Salmonella* species			

XSL•FO
RenderX

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
				• Multidrug-resistant *Shigella* species • Vancomycin- or linezolid-nonsusceptible *Staphylococcus aureus* complex (*S argenteus* and *S schweitzeri*) • Penicillin-reduced susceptibility *Streptococcus pyogenes*			
Japan (Janis-CL)	Human	Hospital-based Community-based	Clinical isolates	• AMR never officially reported (eg, vancomycin-resistant *Staphylococcus aureus*) • AMR reported but rare (eg, multidrug-resistant *Acinetobacter spp*. and vancomycin-resistant enterococci)	Yes	Alert emails within 1 - 2 h	NA[b]
North America							

XSL•FO
RenderX

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
Canada (AMRnet)	Human Animal	Hospital-based Community-based	Clinical isolates	• All bacterial and fungal pathogens tested for AMR, including priority pathogens[c], as classified by the World Health Organization	NA	NA	NA
United States (AR Lab Network)	Human	Hospital-based Community-based	Clinical isolate screening	• Carbapenemase-producing organism (CPO), including carbapenem-resistant *Enterobacterales* (CRE) • *Candida* species, to identify emerging resistance • Emerging and concerning threats, such as mcr-1 and carbapenem-resistant *Acinetobacter baumannii*, and changes in known threats, including MRSA • Pandrug-resistant (PDR) bacteria to new antibiotics	Yes	NA	NA
Europe and Central Asia							

XSL•FO
RenderX

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
France (e-SIN)	Human	Hospital-based	Clinical isolates	• Carbapenemase-producing *Enterobacterales* • Glycopeptide-resistant *Enterococcus faecium*	NA	NA	NA
Sweden (Svebar)	Human	Hospital-based Community-based	Clinical isolates	• Bacterial species resistant to one or more antibiotics (ie, *E coli* resistant to carbapenems) • A trend (ie, resistance to ampicillin in more than 30 percent of cases of *E coli*).	Yes	Daily	NA
United Kingdom (ARAs)	Human	Hospital-based	Clinical isolates		Yes	NA	NA

XSL•FO
RenderX

Country	Sector	Setting	Specimen type	AMR profiles of microorganisms	Genomic data	Alert timing	Performance evaluation
				The AMRHAI[d] advises the referral of isolates with exceptional resistance phenotypes. A list of unusual combinations of resistance and organisms is available for diagnostic laboratories[e], including the following microorganisms: • *Acinetobacter spp*. • Coagulase-negative *staphylococci* • *Enterobacterales* • *Enterococci* • *Haemophilus influenzae* • *Moraxella catarrhalis* • Organisms or antibiotics for which there are no EUCAST[f] clinical breakpoints (invasive sites) • *Pseudomonas aeruginosa* • *Staphylococcus aureus* • *Streptococci* (groups A, B, C, and G, β-haemolytic) • *Streptococcus pneumoniae*			

[a]AMR: antimicrobial resistance.

[b]NA: not available.

[c]*Acinetobacter* spp; *Escherichia coli*; *Klebsiella pneumoniae*; *Neisseria gonorrhoeae*; *Salmonella* spp; *Shigella* spp; *Staphylococcus aureus*; *Streptococcus pneumoniae*.

[d]AMRHAI: antimicrobial resistance and healthcare-associated infections reference unit.

[e]United Kingdom Health Security Agency Reference Laboratories Colindale. Bacteriology Reference Department user manual. Appendix 1 (Antimicrobial resistance and mechanisms service). Version 15, October 10, 2022.

[f]EUCAST: European Committee on Antimicrobial Susceptibility Testing.

XSL•FO
RenderX

The setting where specimens are collected was hospital based and community based in all the identified systems except for e-SIN (France) and ARAs (United Kingdom), where the collection was mainly hospital based [36,46]. More specifically, CARAlert (Australia) was the most comprehensive alert system, including public and private hospitals, general practices, aged care homes, community health services, and hospital nonadmitted care services [54]; JANIS-CL (Japan) included outpatients and inpatients data from large (usually tertiary care) and small hospitals (usually privately owned hospitals with less than 200 beds, providing both acute and long-term care) [31]; and AMRnet (Canada) included AMR data from clinical laboratories in both public and private facilities [43]. Detailed published information was not available referring to the remaining EWSs.

AMR Events Reported and Availability of Genomic Data

Reported AMR events varied widely among the identified EWSs (Table 2). In the majority of instances, emerging AMR was classified in terms of broader categories of unusual events (eg, AMR never previously officially reported, AMR reported but rarely, or a trend). Notably, among the 7 systems examined, 3 systems (ie, Australia, France, and the UK) provided a list of unusual combinations of microorganisms and resistance, including in most cases, carbapenemase-producing *Acinetobacter baumannii* and carbapenemase-producing *Enterobacterales* [38,44,46]. Two countries (ie, Australia and the United States) included emerging fungal pathogens such as *Candida auris* in their systems [25,44].

Genomic data were provided by 5 out of 7 EWSs, while this information was not retrievable for 2 countries, specifically Canada and France [25,33,44,46,52].

Alert Functions

Alert timing and functioning were clearly defined in 3 out of 7 EWSs (ie, Australia, Japan, and Sweden), with several differences among these systems. Concerning CAR-Alert (Australia), data are submitted within 7 days of the isolate being confirmed as a CAR (Table 2). The system generates a weekly summary email alert to report information on confirmed CARs to state and territory health authorities, the Australian Government Department of Health, and confirming laboratories, supporting timely responses to CARs by hospitals and state and territory health departments [45].

In the case of JANIS-CL (Japan), a data format was developed as a unified standard for collecting electronic data from microbiology laboratories. Microbiological alerts can be checked on the JANIS member-restricted website within a few hours after data submission. Furthermore, alert emails reporting unusual profiles of AMR are automatically sent within 1 or 2 hours to facilities contact personnel while monthly feedback emails are sent within 48 hours [30].

Regarding Svebar (Sweden), preset alerting algorithms search through the short-term storage and generate an alert in specific conditions, for example when the system discovers a trend. National alerts are sent via e-mail to the contact person at the local laboratory and to administrators at the Public Health Agency of Sweden, and if needed, they can discuss the alert, while local alerts only reach the contact person in the local laboratory. By systematically tracking the progression of resistance, the detection of early-stage changes might justify modifications to the alert system's settings [51].

Performance Evaluation

Based on publicly available information, the formal assessment of national EWSs has been conducted in only 1 country, specifically Australia, where the Department of Health and Aged Care maintains a routine process of evaluating national surveillance systems to ensure they align with their intended purpose and objectives (Table 2). The evaluation of CAR-Alert took place in 2022 - 2023, with the aim of assessing the system's operational efficiency in fulfilling its objectives, evaluating their appropriateness, and identifying potential enhancements to improve the system's ability to achieve its goals [54].

To assess the system's overall effectiveness and performance, the evaluation process was based on the "Updated guidelines for evaluating public health surveillance systems" from the CDC [55].

Discussion

Principal Findings

Our systematic review allowed the identification of national EWSs for emerging profiles of AMR in a limited number of HIC, approximately 10%. In addition, only half of the 6 World Bank regions that we investigated had functional EWSs in place, revealing a disparity in the availability of these systems at the global level. Given that several regions, including Latin America and the Caribbean, are currently facing alarming developments with national authorities raising concerns about the emergence of carbapenemase-producing *Enterobacterales* not previously described, or about the increasing number of isolates that coexpress two or more of these enzymes [56], the lack of available information on national EWSs in these countries is concerning. Even in the European region, where countries are mostly high-income and have advanced surveillance and health care systems, national EWSs were not retrievable in those areas where AMR rates have been found to be extremely high by the European Center for Disease Prevention and Control (ECDC) [13]. Such an absence of EWSs in high AMR burden countries reveals a critical gap in the region's preparedness and response to the growing challenge of AMR [54], making it difficult for these nations to not only track the evolution of resistance patterns over time, but also early detect and mitigate the spread of unusual resistant bacteria, implement targeted interventions in high-risk settings, and contribute to shared data on a global scale [19]. In 2021, the ECDC launched an online portal (EpiPulse) for European public health authorities, aiming at enhancing the early detection and assessment of threats due to infectious diseases. However, its effectiveness relies on information from member states concerning microbiological alerts, and without national EWSs in place, the management of threats at the international level becomes challenging [15].

XSL•FO
RenderX

Comparison With Previous Studies

Notably, while identified EWSs were established quite recently and shared the overarching goal of monitoring and responding to emerging AMR threats, our study showed that their main characteristics varied significantly across countries. This heterogeneity might be attributable to several factors, including the unique health care infrastructure, different public health policies and funding, and the specific challenges and priorities faced by each country [57-59]. The heterogeneity in identified EWSs also reflects the specific AMR trends and epidemiological situation within each state, considering that AMR profiles can vary significantly, being unusual in one country and endemic in another [60]. Nevertheless, some common ground in monitoring, understanding, and responding to AMR emerged in a few aspects, especially if we consider the inclusion of the community setting and the availability of genomic data in most of the identified EWSs. Several reasons can explain this approach: community-based surveillance is crucial because resistant microorganisms can spread within health care and community settings, with early detection allowing the proper prevention and control of outbreaks [61], while genomic data play a pivotal role in understanding AMR trends and developing targeted prevention and control interventions [62]. As for the availability of performance evaluation of EWSs, which is crucial to assess their effectiveness, we found that Australia was the only country known to conduct a system evaluation in accordance with the CDC guidelines, even if related data are not yet available [54]. Considering that regular evaluations can lead to improvements in the system design, data collection, and response mechanisms, and ultimately enhance the ability to respond effectively to emerging AMR profiles [55], more efforts should be made to institutionalize regular assessments as an integral part of any strategy aimed to manage AMR.

Lastly, although the existence of EWSs for emerging AMR was not clearly reported in most cases, this review found that many countries indicated, in their National Action Plans against AMR, the goal of strengthening national surveillance through the implementation of these systems in the next years. This is comforting, as this underlines their commitment in addressing this issue, in line with the WHO recommendations [18,63-65]. Furthermore, as the landscape of infectious disease surveillance is rapidly evolving owing to the advancements in new technologies, there might be a significant potential for progress in bolstering the efforts against emerging AMR [66]. Artificial intelligence, for example, could enhance AMR surveillance systems given its ability to integrate and analyze data from various sources, such as clinical data and microbiology reports, recognize AMR patterns, develop predictive models that signal the early emergence of resistant strains or the spread of resistance in specific regions or health care facilities, or provide real-time insights into AMR trends [67]. The incorporation of machine learning algorithms to predict trends in resistance development based on historical data could also improve AMR surveillance systems capacity in early warning [68]. Therefore, as artificial intelligence technologies continue to improve, they are definitely an option to be considered as a supporting tool in the early detection of emerging AMR in the near future, enabling a more timely and effective response also by facilitating the sharing and analysis of data on a global scale [69].

Limitations

To the best of our knowledge, this is the first synthesis review of publicly available information on national EWSs for emerging AMR in HIC. The limitations of our study mostly rely on the inclusion criteria we adopted. First, by focusing on human surveillance, our study did not include other relevant sectors in the context of the “One Health” approach [70,71]. Second, by limiting evidence collection to HIC, we did not address challenges faced by countries, such as low- and middle-income countries, where AMR has a disproportionate impact due to higher burden of infections, reduced laboratory capabilities, and limited regulations involving antimicrobial use [72] Third, the restriction to the national level did not account for regional or local-level EWSs, which could be available in some countries [23]. However, the decision to focus on national rather than subnational EWSs relies on the aim of mapping systems capable of centralizing AMR-related data through a comprehensive and coordinated approach, facilitating global data sharing and response strategies, in accordance with the WHO recommendations [18]. Moreover, we are aware that, as this study relies on publicly available information, and not all EWSs might be documented or publicly disclosed [73], our review may not fully represent the extent of EWSs across different countries and across various levels of implementation, such as national and subnational levels. Similarly, the partial availability of information on a few systems limited our ability to fully describe the different strategies adopted in the investigated areas. Nevertheless, our interest was focused on mapping the efforts made by countries at the national level, and we simultaneously searched multiple data sources as an attempt to retrieve all available information. Findings from our study could support countries currently lacking EWSs in the process of strengthening national surveillance systems and contribute to global effort, through improved detection, reporting, and data sharing, in tackling AMR. Additional information is certainly required to provide a complete overview on available EWSs. To this end, our review could be integrated with a country consultation and/or international survey aimed at retrieving additional findings from relevant stakeholders, such as the Ministries of Health and National Health Institutes representatives. Further investigations could also provide an in-depth analysis on the availability of EWSs at subnational (local or regional) levels and in low- and middle-income countries.

Conclusions

These findings highlight the urgent need for a broader implementation of surveillance systems that allow for the early detection of emerging AMR, with increased investments and collaborative efforts to establish EWSs in countries and regions lacking such capabilities to date. Platforms such as EpiPulse within the European region could enhance the collection and analysis of AMR surveillance data, taking into account that each country should implement tools capable of detecting and managing alert signals. In addition, given the heterogeneity in national health care systems and their different epidemiological

XSL•FO
RenderX

contexts, tailored approaches that enable the collection of standardized and comparable AMR data are strongly encouraged to help promote the global preparedness for AMR.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies used in the systematic review and high-income countries' institutional websites explored.
[PDF File, 197 KB - publichealth_v11i1e57457_app1.pdf]

Checklist 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File, 81 KB - publichealth_v11i1e57457_app2.pdf]

References

1. Antimicrobial resistance. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance [accessed 2024-09-30]
2. About antimicrobial resistance. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/antimicrobial-resistance/about/index.html [accessed 2024-09-30]
3. O'Neill J. Tackling drug-resistant infections globally: final report and recommendations. : Review on Antimicrobial Resistance; 2016 URL: https://amr-review.org/sites/default/files/160525_Final%20paper_with%20cover.pdf [accessed 2024-12-24]
4. O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations the review on antimicrobial resistance. : Review on Antimicrobial Resistance; 2014 URL: https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf [accessed 2024-12-24]
5. Antibiotic resistance threats in the United States, 2019. : Centers for Disease Control and Prevention (CDC); 2019 URL: https://stacks.cdc.gov/view/cdc/82532 [accessed 2024-09-30]
6. Naylor NR, Atun R, Zhu N, et al. Estimating the burden of antimicrobial resistance: a systematic literature review. Antimicrob Resist Infect Control 2018;7:58. [doi: 10.1186/s13756-018-0336-y] [Medline: 29713465]
7. Cassini A, Högberg LD, Plachouras D, et al. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. Lancet Infect Dis 2019 Jan;19(1):56-66. [doi: 10.1016/S1473-3099(18)30605-4] [Medline: 30409683]
8. Migliara G, Di Paolo C, Barbato D, et al. Multimodal surveillance of healthcare associated infections in an intensive care unit of a large teaching hospital. Ann Ig 2019;31(5):399-413. [doi: 10.7416/ai.2019.2302] [Medline: 31304521]
9. Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. The Lancet 2022 Feb;399(10325):629-655. [doi: 10.1016/S0140-6736(21)02724-0]
10. WHO Bacterial Priority Pathogens List, 2024: bacterial pathogens of public health importance to guide research, development and strategies to prevent and control antimicrobial resistance. : World Health Organization; 2024.
11. Hay SI, Rao PC, Dolecek C, et al. Measuring and mapping the global burden of antimicrobial resistance. BMC Med 2018 Jun 4;16(1):78. [doi: 10.1186/s12916-018-1073-z] [Medline: 29860943]
12. Tacconelli E, Sifakis F, Harbarth S, et al. Surveillance for control of antimicrobial resistance. Lancet Infect Dis 2018 Mar;18(3):e99-e106. [doi: 10.1016/S1473-3099(17)30485-1] [Medline: 29102325]
13. European Centre for Disease Prevention and Control. Antimicrobial resistance in the EU/EEA (EARS-net) - annual epidemiological report 2022. Stockholm. : ECDC; 2023.
14. Emerging antimicrobial resistance reporting: guide for emerging AMR event sharing. : World Health Organization; 2018 URL: https://www.who.int/publications/i/item/9789241514583 [accessed 2024-12-24]
15. EpiPulse - the European surveillance portal for infectious diseases. European Centre for Disease Prevention and Control. 2021. URL: https://www.ecdc.europa.eu/en/publications-data/epipulse-european-surveillance-portal-infectious-diseases [accessed 2024-09-30]
16. Global antimicrobial resistance and use surveillance system (GLASS) report 2022. : World Health Organization; 2022.
17. GLASS manual for antimicrobial resistance surveillance in common bacteria causing human infection. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240076600 [accessed 2024-09-30]
18. Global action plan on antimicrobial resistance. World Health Organization. URL: https://www.who.int/publications/i/item/9789241509763 [accessed 2024-09-30]
19. GLASS emerging antimicrobial resistance reporting framework (GLASS-EAR). World Health Organization. 2018. URL: https://www.who.int/publications/i/item/9789241514590 [accessed 2024-09-30]

XSL•FO
RenderX

20. Sulis G, Sayood S, Gandra S. Antimicrobial resistance in low- and middle-income countries: current status and future directions. Expert Rev Anti Infect Ther 2022 Feb;20(2):147-160. [doi: 10.1080/14787210.2021.1951705] [Medline: 34225545]
21. Gandra S, Alvarez-Uria G, Turner P, Joshi J, Limmathurotsakul D, van Doorn HR. Antimicrobial resistance surveillance in low- and middle-income countries: progress and challenges in eight South Asian and Southeast Asian countries. Clin Microbiol Rev 2020 Sep 16;33(3):e00048-19. [doi: 10.1128/CMR.00048-19] [Medline: 32522747]
22. Implementing the global action plan on antimicrobial resistance: first quadripartite biennial report. : World Health Organization, Food and Agriculture Organization of the United Nations, United Nations Environment Programme and World Organisation for Animal Health; 2023 URL: https://www.who.int/publications/i/item/9789240074668 [accessed 2024-12-23]
23. Iera J, Seghieri C, Tavoschi L, et al. Early warning systems for emerging profiles of antimicrobial resistance in Italy: a national survey. Int J Environ Res Public Health 2023 Apr 24;20(9):5623. [doi: 10.3390/ijerph20095623] [Medline: 37174143]
24. Vong S, Anciaux A, Hulth A, et al. Using information technology to improve surveillance of antimicrobial resistance in South East Asia. BMJ 2017 Sep 5;358:j3781. [doi: 10.1136/bmj.j3781] [Medline: 28874344]
25. About CDC's AR lab network. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/drugresistance/solutions-initiative/ar-lab-networks.html [accessed 2024-09-30]
26. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021). : Cochrane; 2021.
27. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
28. Fantom NJ, Serajuddin U. The World Bank's classification of countries by income. Policy research working paper, no. WPS 7528 Washington, D.C. World Bank Group. URL: http://documents.worldbank.org/curated/en/408581467988942234/The-World-Banks-classification-of-countries-by-income [accessed 2024-09-30]
29. Turnidge JD, Meleady KT. Antimicrobial Use and Resistance in Australia (AURA) surveillance system: coordinating national data on antimicrobial use and resistance for Australia. Aust Health Rev 2018;42(3):272. [doi: 10.1071/AH16238]
30. Tsutsui A, Suzuki S. Japan nosocomial infections surveillance (JANIS): A model of sustainable national antimicrobial resistance surveillance based on hospital diagnostic microbiology laboratories. BMC Health Serv Res 2018 Oct 20;18(1):799. [doi: 10.1186/s12913-018-3604-x] [Medline: 30342499]
31. Suzuki S. A view on 20 years of antimicrobial resistance in Japan by two national surveillance systems: the national epidemiological surveillance of infectious diseases and Japan nosocomial infections surveillance. Antibiotics (Basel) 2021 Sep 30;10(10):1189. [doi: 10.3390/antibiotics10101189] [Medline: 34680770]
32. Johnson AP. Surveillance of antibiotic resistance. Philos Trans R Soc Lond B Biol Sci 2015 Jun 5;370(1670):20140080. [doi: 10.1098/rstb.2014.0080] [Medline: 25918439]
33. Kajihara T, Yahara K, Hirabayashi A, Shibayama K, Sugai M. Japan Nosocomial Infections Surveillance (JANIS): current status, international collaboration, and future directions for a comprehensive antimicrobial resistance surveillance system. Jpn J Infect Dis 2021 Mar 24;74(2):87-96. [doi: 10.7883/yoken.JJID.2020.499] [Medline: 32863357]
34. National alert system for critical antimicrobial resistances (caralert). Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/our-work/antimicrobial-resistance/antimicrobial-use-and-resistance-australia-surveillance-system/national-alert-system-critical-antimicrobial-resistances-caralert [accessed 2024-09-30]
35. One-health surveillance of antimicrobial resistance in Canada: spotlight on AMRNet. Government of Canada. URL: https://health-infobase.canada.ca/amrnet/?utm_source=stakeholder-weekly-newsletter&utm_medium=link&utm_content=en-02-07-23&utm_campaign=st-stakeholder-weekly-newsletter&wbdisable=true [accessed 2024-09-30]
36. E-sin: signalement externe des infections nosocomiales [Article in French]. Santé publique France. URL: https://www.santepubliquefrance.fr/maladies-et-traumatismes/infections-associees-aux-soins-et-resistance-aux-antibiotiques/infections-associees-aux-soins/articles/e-sin-signalement-externe-des-infections-nosocomiales [accessed 2024-09-30]
37. Un système d'alerte: le signalement externe des infections associées aux soins [Article in French]. Santé publique France. URL: https://www.santepubliquefrance.fr/maladies-et-traumatismes/infections-associees-aux-soins-et-resistance-aux-antibiotiques/resistance-aux-antibiotiques/articles/un-systeme-d-alerte-le-signalement-externe-des-infections-associees-aux-soins [accessed 2024-09-30]
38. Les bhre, bactéries hautement Résistantes émergentes [Article in French]. Centre d'appui pour la prévention des infections associées aux soins. URL: https://www.cpias-ile-de-france.fr/appui/bhre.php [accessed 2024-09-30]
39. Multisector surveillance of antibiotic sales and resistance. The Public Health Agency of Sweden. URL: https://www.folkhalsomyndigheten.se/the-public-health-agency-of-sweden/communicable-disease-control/antibiotics-and-antimicrobial-resistance/overview-of-swedens-one-health-response-to-antibiotic-resistance/swedish-work-against-antibiotic-resistance--a-one-health-approach/multisector-surveillance-of-antibiotic-sales-and-resistance/ [accessed 2024-09-30]

XSL•FO
RenderX

40. UK public health antimicrobial resistance alerts. Public Health England. URL: https://www.gov.uk/government/publications/uk-public-health-antimicrobial-resistance-alerts [accessed 2024-09-30]
41. Antibiotic resistance laboratory network. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/drugresistance/ar-lab-networks/domestic.html [accessed 2024-09-30]
42. How labs work together. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/drugresistance/ar-lab-networks/domestic/how-labs-work-together.html [accessed 2024-09-30]
43. Rudnick W, Mukhi SN, Reid-Smith RJ, et al. Overview of Canada's Antimicrobial Resistance Network (AMRNet): A data-driven One Health approach to antimicrobial resistance surveillance. Can Commun Dis Rep 2022 Nov 3;48(11-12):522-528. [doi: 10.14745/ccdr.v48i1112a05] [Medline: 38173468]
44. National alert system for critical antimicrobial resistances (CARAlert) laboratory handbook. Australian Commission on Safety and Quality in Health Care. 2022. URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/caralert-laboratory-handbook-0 [accessed 2024-09-30]
45. CARAlert standard operating procedures. : Australian Commission on Safety and Quality in Health Care; 2022 URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/caralert-standard-operating-procedures [accessed 2024-12-24]
46. Doshi N. Bacteriology reference department user manual. GOV-13422. : UK Health Security Agency; 2022.
47. Canadian antimicrobial resistance surveillance system (CARSS) report 2022. Government of Canada. URL: https://www.canada.ca/en/public-health/services/publications/drugs-health-products/canadian-antimicrobial-resistance-surveillance-system-report-2022.html [accessed 2024-09-30]
48. Berger-Carbonne A, et al. La lettre du signalement. Infections associées aux soins. : Santé publique France; 2024.
49. Actualisation des recommandations relatives à la maîtrise de la diffusion des bactéries hautement résistantes aux antibiotiques émergentes (BHRe) [Article in French]. Conseil de la santé publique. URL: https://www.hcsp.fr/explore.cgi/avisrapportsdomaine?clefr=758 [accessed 2024-09-30]
50. National action plan for combating antibiotic-resistant bacteria. Year 5 report. : Task Force for Combating Antibiotic-Resistant Bacteria; 2022.
51. Swedish work on containment of antibiotic resistance. Tools, methods and experiences. : Public Health Agency of Sweden; 2014.
52. Aspevall O, Dobric S, Jagdmann J, Nilsson O, Pringle M. A report on Swedish antibiotic sales and resistance in human medicine (Swedres) and Swedish veterinary antibiotic resistance monitoring (Svarm) editors. Swedish Veterinary Agency. URL: https://www.sva.se/en/what-we-do/antibiotics/svarm-resistance-monitoring/swedres-svarm-reports/ [accessed 2024-09-30]
53. AURA 2023: fifth Australian report on antimicrobial use and resistance in human health. Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/aura-2023-fifth-australian-report-antimicrobial-use-and-resistance-human-health-report [accessed 2024-09-30]
54. CARAlert annual report: 2023. Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/caralert-annual-report-2023 [accessed 2024-09-30]
55. Updated guidelines for evaluating public health surveillance systems recommendations from the guidelines working group. Centers for Disease Control and Prevention. 2001. URL: https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm [accessed 2024-09-30]
56. Emergence and increase of new combinations of carbapenemases in enterobacterales in Latin America and the Caribbean. Pan American Health Organization. 2021. URL: https://www.paho.org/sites/default/files/2021-10/2021-October_PHE-Epi%20Alert_carbapenemases_EN.pdf [accessed 2024-12-24]
57. Lessa FC, Sievert DM. Antibiotic resistance: a global problem and the need to do more. Clin Infect Dis 2023 Jul 5;77(Suppl 1):S1-S3. [doi: 10.1093/cid/ciad226] [Medline: 37406051]
58. Saltman RB, Bankauskaite V, Vrangbaek K, editors. Decentralization in Health Care: Strategies and Outcomes: European Observatory on Health Systems and Policies; 2007:1-293 URL: https://eurohealthobservatory.who.int/publications/m/decentralization-in-health-care-strategies-and-outcomes [accessed 2024-12-24]
59. James C, Beazley I, Penn C, Philips L, Dougherty S. Decentralisation in the health sector and responsibilities across levels of government. OECD J Budget 2019 Dec 13;19(3) [FREE Full text] [doi: 10.1787/c2c2058c-en]
60. Bonomo RA, Burd EM, Conly J, et al. Carbapenemase-producing organisms: a global scourge. Clin Infect Dis 2018 Apr 3;66(8):1290-1297. [doi: 10.1093/cid/cix893] [Medline: 29165604]
61. Embracing a one health framework to fight antimicrobial resistance. Organisation for Economic Co-operation and Development (OECD). 2023. URL: https://www.oecd-ilibrary.org/social-issues-migration-health/embracing-a-one-health-framework-to-fight-antimicrobial-resistance_ce44c755-en [accessed 2024-09-30]
62. ECDC public health microbiology strategy 2018 – 2022. European Centre for Disease Prevention and Control (ECDC). URL: https://www.ecdc.europa.eu/sites/default/files/documents/ECDC-public-health-microbiology-strategy-2018-2022.pdf [accessed 2024-09-30]
63. Piano nazionale di contrasto all'Antimicrobico-resistenza (PNCAR) 2022-2025 [Article in Italian]. : Ministero della Salute; 2022 URL: https://www.salute.gov.it/imgs/C_17_pubblicazioni_3294_allegato.pdf [accessed 2024-09-30]

XSL•FO
RenderX

64. New Zealand antimicrobial resistance action plan. Ministry of Health. URL: https://www.health.govt.nz/strategies-initiatives/programmes-and-initiatives/antimicrobial-resistance/nz-antimicrobial-resistance-action-plan [accessed 2024-09-30]
65. Ireland's national action plan on antimicrobial resistance 2017 - 2020. : Department of Health; 2017 URL: https://assets.gov.ie/9519/afcba9bce7c54bf9bcbe9a74f49fdaf2.pdf [accessed 2024-09-30]
66. Digital technologies for infectious disease surveillance, prevention and control - a scoping review of the research literature 2015−2019. : European Centre for Disease Prevention and Control URL: https://www.ecdc.europa.eu/en/publications-data/digital-technologies-surveillance-prevention-and-control-infectious-diseases [accessed 2024-09-30]
67. Brownstein JS, Rader B, Astley CM, Tian H. Advances in artificial intelligence for infectious-disease surveillance. N Engl J Med 2023 Apr 27;388(17):1597-1607. [doi: 10.1056/NEJMra2119215] [Medline: 37099342]
68. Sakagianni A, Koufopoulou C, Feretzakis G, et al. Using machine learning to predict antimicrobial resistance-a literature review. Antibiotics (Basel) 2023 Feb 24;12(3):452. [doi: 10.3390/antibiotics12030452] [Medline: 36978319]
69. MacIntyre CR, Chen X, Kunasekaran M, et al. Artificial intelligence in public health: the potential of epidemic early warning systems. J Int Med Res 2023 Mar;51(3):3000605231159335. [doi: 10.1177/03000605231159335] [Medline: 36967669]
70. A One Health Priority Research Agenda for Antimicrobial Resistance: World Health Organization, Food and Agriculture Organization of the United Nations, United Nations Environment Programme, World Organisation for Animal Health; 2022. URL: https://www.who.int/publications/i/item/9789240075924 [accessed 2024-09-30]
71. One Health High-Level Expert Panel (OHHLEP), Adisasmito WB, Almuhairi S, et al. One Health: a new definition for a sustainable and healthy future. PLoS Pathog 2022 Jun;18(6):e1010537. [doi: 10.1371/journal.ppat.1010537] [Medline: 35737670]
72. Loftus M, Stewardson A, Naidu R, et al. Antimicrobial resistance in the Pacific Island countries and territories. BMJ Glob Health 2020 Apr;5(4):e002418. [doi: 10.1136/bmjgh-2020-002418] [Medline: 32349993]
73. Núñez-Núñez M, Navarro MD, Palomo V, et al. The methodology of surveillance for antimicrobial resistance and healthcare-associated infections in Europe (SUSPIRE): a systematic review of publicly available information. Clin Microbiol Infect 2018 Feb;24(2):105-109. [doi: 10.1016/j.cmi.2017.07.014] [Medline: 28750921]

Abbreviations

AMR: antimicrobial resistance
AMRNet: Antimicrobial Resistance Network
AR Lab Network: Antimicrobial Resistance Laboratory Network
ARAs: antimicrobial resistance alerts
CAR: critical antimicrobial resistances
CAR-Alert: National Alert System for Critical Antimicrobial Resistances
CDC: Centers for Disease Control and Prevention
ECDC: European Center for Disease Prevention and Control
EWSs: early warning systems
GLASS: Global Antimicrobial Resistance and Use Surveillance System
HIC: high-income countries
JANIS-CL: Japan Nosocomial Infections Surveillance - Clinical Laboratory Division
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
WHO: World Health Organization

Edited by AS Bhagavathula; submitted 17.02.24; peer-reviewed by A Rosso, B Tornimbene, FC Dezza; revised version received 30.09.24; accepted 04.10.24; published 15.01.25.

Please cite as:
Iera J, Isonne C, Seghieri C, Tavoschi L, Ceparano M, Sciurti A, D'Alisera A, Sane Schepisi M, Migliara G, Marzuillo C, Villari P, D'Ancona F, Baccolini V
Availability and Key Characteristics of National Early Warning Systems for Emerging Profiles of Antimicrobial Resistance in High-Income Countries: Systematic Review
JMIR Public Health Surveill 2025;11:e57457
URL: https://publichealth.jmir.org/2025/1/e57457
doi:10.2196/57457

© Jessica Iera, Claudia Isonne, Chiara Seghieri, Lara Tavoschi, Mariateresa Ceparano, Antonio Sciurti, Alessia D'Alisera, Monica Sane Schepisi, Giuseppe Migliara, Carolina Marzuillo, Paolo Villari, Fortunato D'Ancona, Valentina Baccolini. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 15.1.2025. This is an open-access article

XSL•FO
RenderX

distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Wastewater Monitoring During the COVID-19 Pandemic in the Veneto Region, Italy: Longitudinal Observational Study

Honoria Ocagli[1*], PhD; Marco Zambito[1*], MSc; Filippo Da Re[2], MSc; Vanessa Groppi[2], BSc; Marco Zampini[3], MSc; Alessia Terrini[3], BSc; Franco Rigoli[3], MSc; Irene Amoruso[4], PhD; Tatjana Baldovin[4], PhD; Vincenzo Baldo[4], MD; Francesca Russo[2*], MD; Dario Gregori[1*], PhD

[1]Unit of Biostatistics, Epidemiology and Public Health, Department of Cardio-Thoraco-Vascular Sciences and Public Health, University of Padova, Via Loredan 18, Padova, Italy

[2]Directorate of Prevention, Food Safety and Veterinary Public Health-Veneto Region, Venice, Italy

[3]ARPAV - Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto, U.O. Biologia, Padova, Italy

[4]Unit of Hygiene and Public Health, Department of Cardiac, Thoracic, Vascular Sciences, and Public Health, University of Padova, Padova, Italy

[*]these authors contributed equally

Corresponding Author:
Dario Gregori, PhD
Unit of Biostatistics, Epidemiology and Public Health, Department of Cardio-Thoraco-Vascular Sciences and Public Health, University of Padova, Via Loredan 18, Padova, Italy

Abstract

Background: As the COVID-19 pandemic has affected populations around the world, there has been substantial interest in wastewater-based epidemiology (WBE) as a tool to monitor the spread of SARS-CoV-2. This study investigates the use of WBE to anticipate COVID-19 trends by analyzing the correlation between viral RNA concentrations in wastewater and reported COVID-19 cases in the Veneto region of Italy.

Objective: We aimed to evaluate the effectiveness of the cumulative sum (CUSUM) control chart method in detecting changes in SARS-CoV-2 concentrations in wastewater and its potential as an early warning system for COVID-19 outbreaks. Additionally, we aimed to validate these findings over different time periods to ensure robustness.

Methods: This study analyzed the temporal correlation between SARS-CoV-2 RNA concentrations in wastewater and COVID-19 clinical outcomes, including confirmed cases, hospitalizations, and intensive care unit (ICU) admissions, from October 2021 to August 2022 in the Veneto region, Italy. Wastewater samples were collected weekly from 10 wastewater treatment plants and analyzed using a reverse transcription–quantitative polymerase chain reaction. The CUSUM method was used to detect significant shifts in the data, with an initial analysis conducted from October 2021 to February 2022, followed by validation in a second period from February 2022 to August 2022.

Results: The study found that peaks in SARS-CoV-2 RNA concentrations in wastewater consistently preceded peaks in reported COVID-19 cases by 5.2 days. Hospitalizations followed with a delay of 4.25 days, while ICU admissions exhibited a lead time of approximately 6 days. Notably, certain health care districts exhibited stronger correlations, with notable values in wastewater anticipating ICU admissions by an average of 13.5 and 9.5 days in 2 specific districts. The CUSUM charts effectively identified early changes in viral load, indicating potential outbreaks before clinical cases increased. Validation during the second period confirmed the consistency of these findings, reinforcing the robustness of the CUSUM method in this context.

Conclusions: WBE, combined with the CUSUM method, offers valuable insight into the level of COVID-19 outbreaks in a community, including asymptomatic cases, thus acting as a precious early warning tool for infectious disease outbreaks with pandemic potential.

(JMIR Public Health Surveill 2025;11:e58862) doi:10.2196/58862

KEYWORDS

wastewater-based epidemiology; SARS-CoV-2; COVID-19; CUSUM; WBE; cumulative sum chart

Introduction

The COVID-19 pandemic has had a profound impact on public health, with more than 200 million confirmed cases and more than 4 million deaths worldwide as of August 2021. As the virus spread, concerns about possible transmission through wastewater became a significant public health issue [1]. Wastewater-based epidemiology (WBE) plays a pivotal role in various health-related studies. Research highlights several applications. The estimation of global nicotine consumption through WBE

XSL•FO
RenderX

was reviewed in the work of Asadi et al [2], highlighting its diverse applications, shedding light on trends in nicotine consumption worldwide and the associated risks, and stressing the urgency of global action. Antibiotic-resistant bacteria are another topic of interest in WBE, as shown in the work of Tiwari et al [3], whose systematic review underscores the need for active monitoring of antibiotic-resistant bacteria, offering insights into their clinical utility, sensitivity, and uniformity. In addition, antimicrobial concentration was investigated. The systematic review by Chau et al [4] examines the concordance between estimates of the prevalence of antimicrobial resistance in wastewater and humans, while Holton et al [5] explore community-wide antimicrobial use via WBE, evaluating 16 antimicrobials and their metabolites, providing insight for improving precision in assessing drug intake. Wastewater-based antimicrobial resistance surveillance, as demonstrated in the most recent literature, appears to be an effective and promising approach to confronting such phenomena in support of public health actions. In addition, wastewater has long been useful for the surveillance of different pathogens.

Recent studies have shown that SARS-CoV-2, a COVID-19 viral strain, can be detected in wastewater samples from infected individuals, even when they are asymptomatic or have mild symptoms [1,6]. Wastewater surveillance is based on the assumption that viral RNA can be detected in the fecal matter of asymptomatic people, as the shedding of viral particles can occur before the onset of symptoms [7] and since diarrhea is a symptom reported by a significant proportion of patients with SARS-CoV-2 [8]. Furthermore, wastewater can also contain fragments of viral RNA secreted by oral or nasal routes that can enter wastewater when hands are washed. The virus can persist in wastewater for several days, although higher temperatures and lower pH levels can reduce its survival [9].

In the literature, various studies offer information on multiple aspects of COVID-19 and wastewater, including the detection of SARS-CoV-2 in wastewater [10], the persistence of the virus in wastewater [11], and the potential use of WBE as a tool to monitor the prevalence of the virus in communities [12]. The study by Ciannella et al [13] focuses on analytical procedures and epidemiological modeling in WBE for COVID-19. They highlight the lack of standardization in wastewater analytical methods, with reverse transcription–quantitative polymerase chain reactions (PCRs) being the most widely used technique for the detection and quantification of viral RNA. They suggest exploring the solid portion of wastewater due to its higher viral load and advocate the development of cost-effective portable detection devices [13]. The study by Peccia et al [14] investigates the potential of wastewater surveillance as a tool for the early detection of COVID-19 outbreaks in a community. The results show that WBE could provide early warnings of outbreaks before clinical cases are detected.

In Italy, the study by La Rosa et al [10] suggests that SARS-CoV-2 was present in northern Italy even in December 2019. The researchers analyzed wastewater samples from October 2019 to February 2020 and found positive results for the virus in multiple cities. This challenged the prior understanding that the geo-warning system primarily served to detect the virus' presence and to monitor outbreaks before they were reported to the health care system.

In WBE, a variety of statistical methods have traditionally been employed to analyze the relationship between viral RNA concentrations in wastewater and epidemic incidence, along with other public health outcomes. Common techniques include correlation analysis, regression models [15], and principal component analysis [16], which are often used to establish associations between wastewater parameters and COVID-19 case numbers, hospitalizations, or deaths. For example, studies have frequently utilized methods such as ANOVA [17], Gaussian distribution models [18], and autoregressive integrated moving average (ARIMA) time series models to understand the temporal dynamics and predict trends in viral loads [19]. Advanced techniques include the Efficient and Practical Virus Identification System with Enhanced Sensitivity for Membrane (EPISENS-M) method, which combines highly sensitive RNA detection with mathematical modeling to predict cases of COVID-19 [20]. Additionally, the framework utilized by Dai et al [21] uses functional data analysis to detect true trends in viral concentration, incorporating known covariates such as sample storage temperature and influent volume, and employs Markov chain Monte Carlo methods to forecast future viral concentrations.

However, while these approaches are effective for retrospective analysis and prediction, they often focus on broad trend identification rather than real-time detection. Our study introduces the cumulative sum (CUSUM) chart method, a statistical process control tool commonly used in quality control [22], which offers unique advantages in WBE by detecting small shifts in viral load trends. Unlike traditional methods, CUSUM charts are particularly useful for identifying early deviations from established baselines, thus allowing earlier detection of potential outbreaks.

This study aims to investigate whether the time series of viral RNA concentrations in wastewater, expressed as genome copies (gc) per microliter, can serve as an early indicator of COVID-19–positive cases, including total positives, hospitalized cases, and those admitted to intensive care units (ICUs), in the Veneto region using CUSUM charts.

Methods

Overview

This study was conducted in the Veneto region of Italy, focusing on a specific timeframe during the COVID-19 pandemic. While the number of data points is limited, the study aims to provide a detailed analysis of the correlation between the viral RNA load of wastewater and the COVID-19 clinical outcomes, offering valuable information for regional public health strategies.

Our research evaluated this hypothesis by comparing the temporal trends of these case categories to SARS-CoV-2 concentrations in wastewater samples.

XSL•FO
RenderX

Ethical Considerations

This study was conducted using publicly available, anonymized, and deidentified data, ensuring that no personal information of individuals was accessed or processed. Wastewater data were derived from environmental samples (ie, untreated wastewater) and COVID-19 epidemiological data were aggregated at the population level, preventing any direct identification of individuals. Additionally, the analysis adhered to ethical guidelines for research using secondary data, ensuring that all data were used responsibly and in accordance with the applicable privacy laws. No personal or sensitive information was collected and the study design did not involve any interventions or direct interaction with human subjects, thus exempting it from requiring formal ethical approval.

Data Sources

This study used data from multiple sources to analyze the correlation between SARS-CoV-2 RNA concentrations in urban wastewater and COVID-19 clinical outcomes. The primary data sources included analytical data from raw wastewater samples and clinical data, including COVID-19 case data, hospitalizations, and ICU admissions, from the Presidency of the Council of Ministers' Civil Protection Department. These datasets were analyzed to assess temporal correlations and to identify potential early warning signals for COVID-19 surges based on wastewater monitoring.

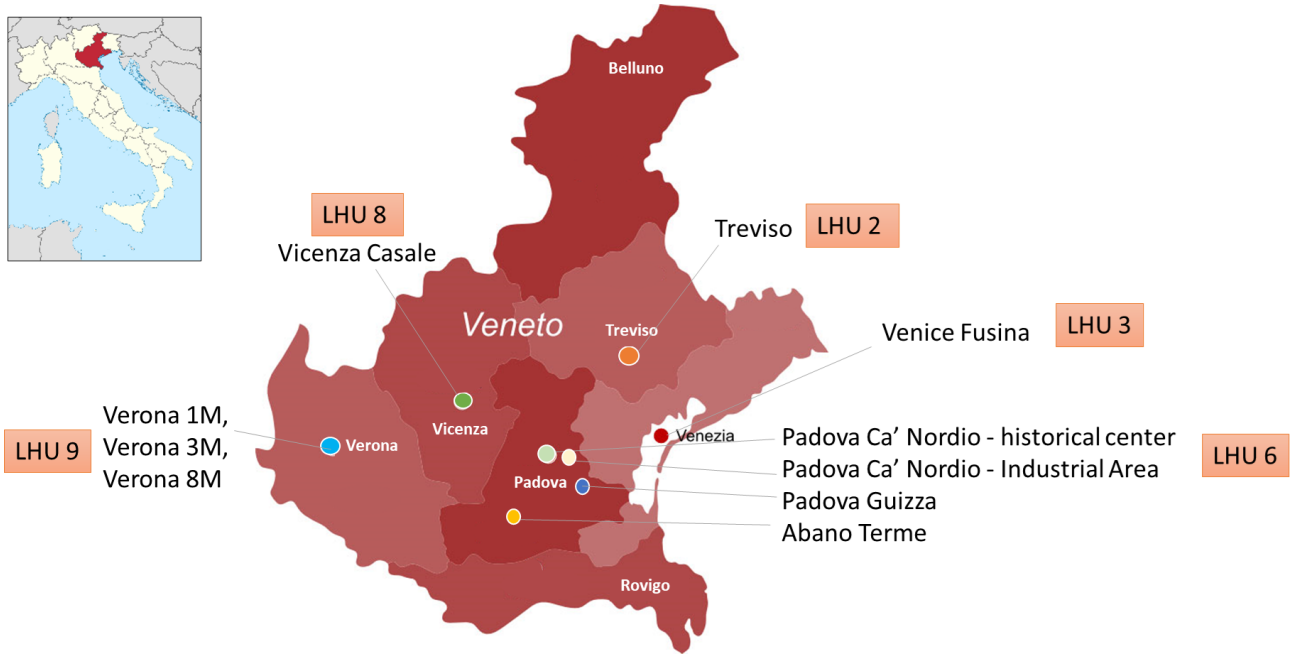
Wastewater Data and Study Design

The study focused on the Veneto region, selecting 10 wastewater treatment plants (WTPs) in 5 provinces, collectively serving more than 4 million inhabitants. Weekly samples were collected from October 5, 2021, to August 2, 2022.

Sampling Equipment and Sites

The samples were collected using an automated sampler (compact portable sampler 6712C, ISCO) to obtain 100 mL of a 24-hour composite raw wastewater sample [10], ensuring representative capture of daily variations in wastewater composition. Each sample consisted of 20 mL subsamples of raw wastewater collected every 30 minutes for the 24-hour period. After collection, the samples were immediately refrigerated at 4 °C to preserve viral RNA integrity until analysis. Sample concentration, extraction, and viral RNA detection protocols are described in the study by Baldovin et al [23]. The selected WTPs included: Abano Terme, Padova Ca' Nordio - historical center, Padova Ca' Nordio - Industrial Area, Padova Guizza, Treviso, Venice Fusina, Verona 1M, Verona 3M, Verona 8M, and Vicenza Casale (Figure 1).

Figure 1. Geographic distribution of wastewater treatment plants and correspondence with LHUs in the Veneto region, Italy. LHU: local health unit.



RNA Detection and Laboratory Methods

Following collection, wastewater samples were concentrated using the Epidemiological Surveillance for SARS-CoV-2 in Urban Sewage in Italy, third revision (SARI_rev3) method [24], which includes a sample concentration step involving polyethylene glycol precipitation. This process yielded a concentrated pellet from a 45 mL sample volume, which was then subjected to RNA extraction using 2 different commercial kits, following the manufacturer's instructions. For the University of Padova samples, RNA extraction was performed using the QIAamp viral RNA mini kit (Qiagen). The samples were collected at the following locations: Abano Terme, Padova Ca' Nordio - historical center, Padova Ca' Nordio - Industrial Area, and Padova Guizza. For the Regional Agency for Environmental Protection and Prevention of Veneto samples, RNA extraction was conducted using the eGENE-UP system (Biomerieux), a well-established method for SARS-CoV-2 RNA recovery. These samples were collected from Treviso, Venice Fusina, Verona 1M, Verona 3M, Verona 8M, and Vicenza Casale. For all samples, RNA was treated with the OneStep PCR Inhibitor Removal Kit (Zymo Research) after extraction to remove PCR inhibitors. The final volume of the RNA elution was standardized to 100 µL.

XSL•FO
RenderX

Reverse transcription–quantitative PCR assays targeting the SARS-CoV-2 open reading frames 1ab (ORF1ab) (nonstructural protein 14) gene were used for quantification. A limit of detection was established at 2 gc/µL. For measurements below this threshold, the values were recorded as half of the limit of detection to account for the presence of the low level. PCR inhibition was routinely assessed using an internal Mengovirus control and recovery efficiency was confirmed for each sample using this control as well.

Replicates and Quality Control

Each wastewater sample was processed in duplicate to ensure the reproducibility and accuracy of viral load quantification. Additionally, a spiking experiment was performed periodically using known concentrations of synthetic SARS-CoV-2 RNA to verify the precision of the assay on different matrices. This approach enabled the detection of any potential assay failures or matrix effects that could impact quantification. To account for the variability in influent flow rates and population size, viral load data were normalized per liter of wastewater and per capita. This normalization facilitated comparison across different WTPs and over time. Detailed records of each WTP's daily influent volume, as well as specific operational characteristics, were maintained and integrated into the data analysis to ensure that any observed variations in viral concentrations could be accurately interpreted.

Positive COVID-19 Case Data

The Presidency of the Council of Ministers' Civil Protection Department diligently managed, processed, and provided data on active and cumulative confirmed COVID-19 cases and daily COVID-19-related deaths for public use under the Creative Commons Attribution 4.0. The population of confirmed COVID-19 cases in the area served by each WTP was identified based on local health units (LHUs) as follows: LHU 2 (Treviso collector); LHU 3 (Venice Fusina collector); LHU 6 (Abano Terme, Padova Ca' Nordio - historical center, Padova Ca' Nordio - Industrial Area, and Padova Guizza collectors); LHU 8 (Vicenza Casale collector); and LHU 9 (Verona 1M, 3M, and 8M collectors) (Figure 1).

Statistical Analysis

Continuous variables were reported as median and first and third quartiles, and categorical variables as percentages and absolute numbers. The trend of wastewater was reported as an absolute value for each measurement, and the trend in the number of positive cases was shown as positives per province population multiplied by 100,000.

CUSUM Charts

To compare COVID-19 cases and trends in wastewater, we used CUSUM charts. Originally developed by Page in 1954 [25], CUSUM charts have since been widely adopted across various fields, including health care, for quality control purposes. The key advantage of a CUSUM chart lies in its cumulative nature, where each plotted point reflects not only the current value but also the cumulative effect of all previous data points [22]. This makes CUSUM charts especially sensitive to small deviations that could indicate a gradual trend change, positive or negative, within a dataset [26].

This graphical tool detects slight deviations from a target value (in this case, the mean) by considering current and previous values in the series. Specifically, the CUSUM chart is based on a statistic derived from the standard deviations of the process, assuming that the empirical distribution follows a normal distribution with a mean of zero. Statistic values that exceed the upper and lower limits (typically set at 4σ, where σ represents the standard deviation of the process) are marked as "out of control" points. These points can be interpreted as "structural breakpoints" or "turning points," indicating significant shifts in the series' trend, either positive or negative. The "differential" signifies the difference between the time ("t") and the value at the values of time ("t-1"). The CUSUM charts for positive case differentials employ upper and lower limits at 4σ, while those for gc/µL concentrations use a threshold at ± 2.782 [27] due to lower data variability (varpositive=16,048,128; varRNA=3070.525 gc/µL).

The analyses were performed using R [28], with the *qcc* package used for the CUSUM chart [29]; *ggplot* [30] and *ggplot2* [31] were used to generate the graphs.

Results

Overview

Table 1 provides an overview of the WTPs considered during the sampling period. It offers information on their key characteristics, including flow rates (m^3/24 h) and total suspended solids concentrations (mg/L). These parameters exhibit variations at different locations, indicating the diverse compositions of the wastewater. The table also highlights the volume of collected samples (mL) and the corresponding distribution across different sample sizes. The concentrations of total suspended solids showed diversity, ranging from 94.25 mg/L in Treviso to 207.00 mg/L in Verona 8M.

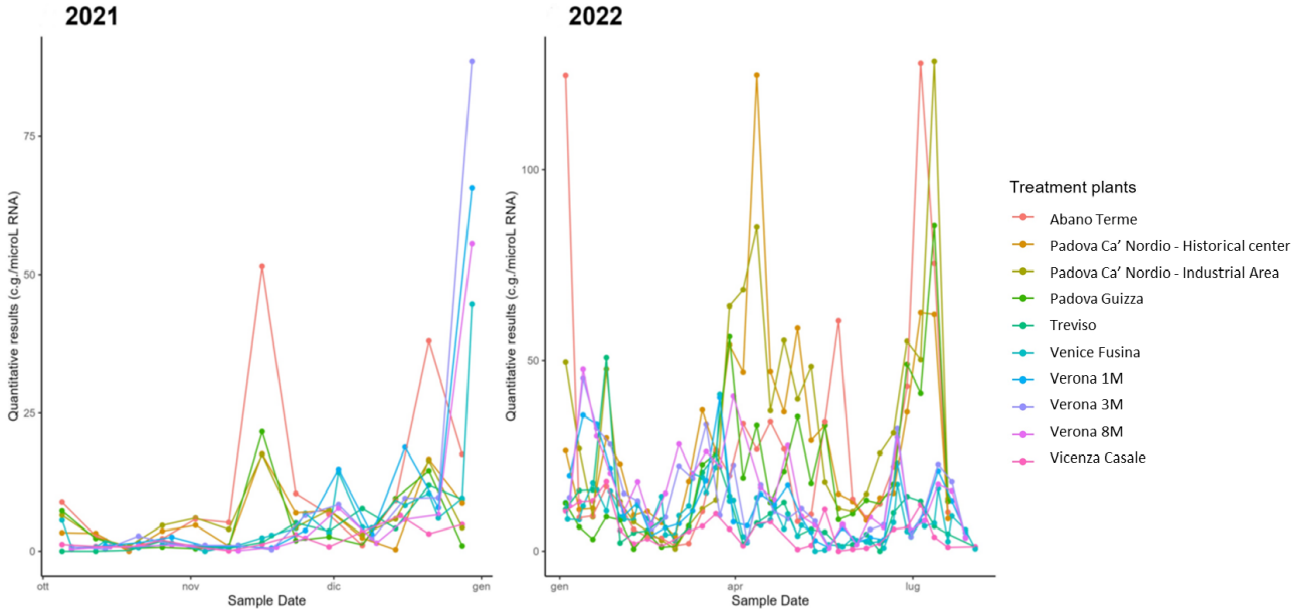
Figure 2 shows the trend of wastewater for 2021 and 2022, divided by the different collectors. It should be noted that there are different peaks in wastewater load during specific periods, including November 2021, late December 2021, and January 2022, followed by recurring peaks in April and July 2022.

XSL•FO
RenderX

Table . Descriptive characteristics of wastewater treatment plants in the Veneto region, Italy, considered in the study. Continuous variables are reported as I, II, and III quartiles, whereas categorical variables are reported as percentages and absolute numbers.

	Flow rates (m^3/24 h), I/II/III quartile	Total suspended solids (mg/L), I/II/III quartile	Sample volume (125 mL), n (%)	Sample volume (250 mL), n (%)	Sample volume (500 mL), n (%)
Abano Terme (n=44)	4927.00/5428.50/6059.50	170.00/232.00/276.00	0 (0)	0 (0)	44 (100)
Padova Ca' Nordio - historical center (n=44)	22,074.25/24,677.00/27,760.75	112.00/172.00/234.00	0 (0)	0 (0)	44 (100)
Padova Ca' Nordio - Industrial Area (n=44)	21,626.75/24,677.00/27,760.75	112.00/170.00/235.00	0 (0)	0 (0)	44 (100)
Padova Guizza (n=44)	3087.25/3275.50/3406.25	16.00/47.00/74.50	0 (0)	0 (0)	44 (100)
Treviso (n=44)	13,477.50/14,365.00/14,830.00	94.25/123.50/169.75	0 (0)	0 (0)	44 (100)
Venice Fusina (n=87)	72,906.50/87,562.00/100,155.25	68.50/116.00/193.00	34 (39)	51 (59)	2 (2)
Verona 1M (n=43)	12,657.50/12,874.00/13,089.50	117.00/181.00/230.00	0 (0)	0 (0)	43 (100)
Verona 3M (n=43)	11,232.00/14,688.00/17,237.00	132.50/183.00/203.50	0 (0)	0 (0)	43 (100)
Verona 8M (n=43)	33,739.50/34,474.00/39,571.00	207.00/231.00/265.00	0 (0)	0 (0)	43 (100)
Vicenza Casale (n=44)	27,052.75/30,018.50/32,466.75	85.75/101.00/119.25	0 (0)	0 (0)	44 (100)

Figure 2. Temporal trends of SARS-CoV-2 RNA concentrations in wastewater throughout the Veneto region, Italy, during 2021 and 2022.



CUSUM Chart Results in the First Period

Daily total COVID-19–positive cases, hospitalized cases, and admissions to the ICU were compared with daily viral concentrations. The analyses considered recovery efficiency (percentage recovery value), which was used as a weighting factor. The results obtained, particularly with respect to the peak concentrations, remained consistent even after adjustment for recovery values.

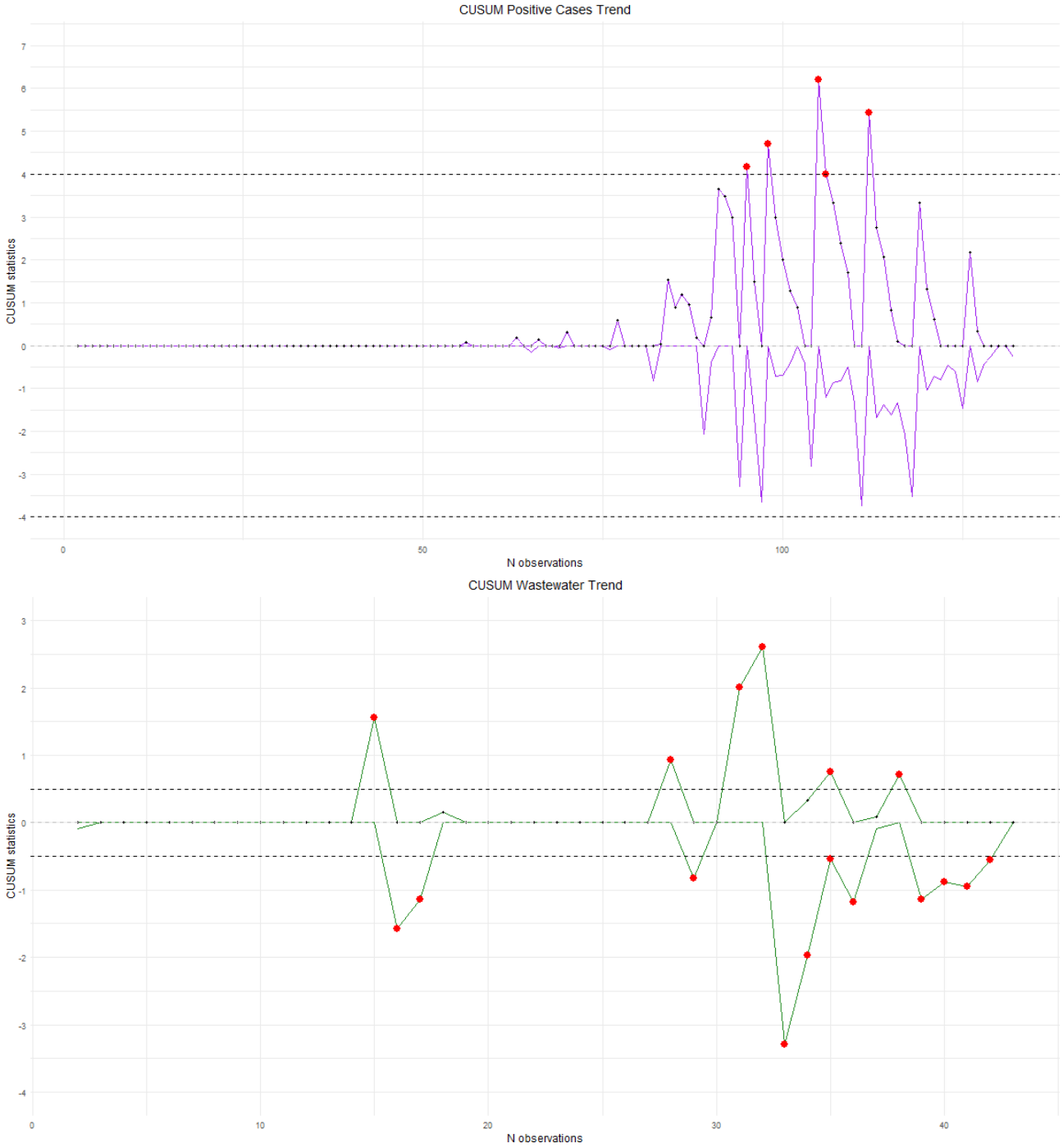
Figure 3 illustrates the CUSUM charts that show the interaction between positive cases and wastewater trends, plotting progressive observations on the x-axis, while the y-axis shows the cumulative deviation from the target, highlighting shifts and trends over time. Notable spikes indicate significant increases, suggesting potential outbreaks or changes in viral concentration.

In Figure 3, we observe a significant positive differential value beyond the upper limit at the 106th data point in the CUSUM Positive Cases Trend chart, which corresponds to the observation made on January 19, 2022. This indicates a substantial increase in positive cases. When comparing this with the CUSUM Wastewater Trend chart, several positive and negative peaks appear. In particular, there is a peak in wastewater effluent between December 30, 2021, and January 25, 2022, and each of these effluent peaks aligns with a subsequent peak in positive case differentials between January 3 and 31, 2022.

Table 2 presents data on the correspondence between significant peaks in wastewater SARS-CoV-2 RNA concentrations and subsequent peaks in COVID-19 cases, distinguishing between total cases, hospitalized cases, and admissions to the ICU.

XSL•FO
RenderX

Figure 3. CUSUM analysis of COVID-19–positive cases (top) and SARS-CoV-2 RNA detected in wastewater (bottom) in the Veneto region, covering the period from October 5, 2021, to February 14, 2022. CUSUM: cumulative sum.



XSL•FO
RenderX

Table . Temporal correspondence between peaks in SARS-CoV-2 RNA concentrations in wastewater and COVID-19 cases in the Veneto region (October 5, 2021, to February 14, 2022).

Dates of wastewater peaks	SARS-CoV-2 RNA concentration, gc/µL	Dates of COVID-19 peak cases	COVID-19 cases, n	Difference in days, n
Total positive COVID-19 cases				
12/21/2021	100.75	12/27/2021	8883	6
12/30/2021	188.73	1/03/2022	16,848	4
1/04/2022	247.79	1/10/2022	22,024	6
1/13/2022	140.62	1/17/2022	25,283	4
1/25/2022	172.71	1/31/2022	15,372	6
Hospitalized COVID-19 cases				
12/21/2021	100.75	12/27/2021	284	6
12/30/2021	188.73	1/03/2022	265	4
1/04/2022	247.79	1/07/2022	253	3
1/13/2022	140.62	1/17/2022	254	4
Hospitalized in intensive care units COVID-19 cases				
12/21/2021	100.75	12/27/2021	18	6

For example, for total positive cases, it should be noted that effluent peaks on December 30, 2021, and January 4, 2022, were followed by significant increases in positive cases, with a delay of 4 to 6 days. This suggests a correlation between wastewater changes and surges in total cases of COVID-19.

Analysis of the CUSUM charts reveals that effluent concentration peaks consistently preceded positive case peaks by an average of 5.2 days for total positive cases. In the case of hospitalized patients, the average time lag between the detection of a significant effluent value and the subsequent detection of a significant value in hospitalizations was 4.25 days.

Similarly, for positive cases that required hospitalization, effluent peaks on December 30, 2021, and January 4, 2022, were followed by case peaks, although with a slightly shorter delay of 3 to 4 days. This close temporal relationship between effluent peaks and hospitalizations is a significant finding.

Furthermore, data indicate that effluent peaks on December 21 and 27, 2021, coincided with the increase in ICU admissions 6 days later, underlining a consistent trend in which wastewater changes anticipated changes in the need for critical care.

CUSUM Chart Results in the Second Period

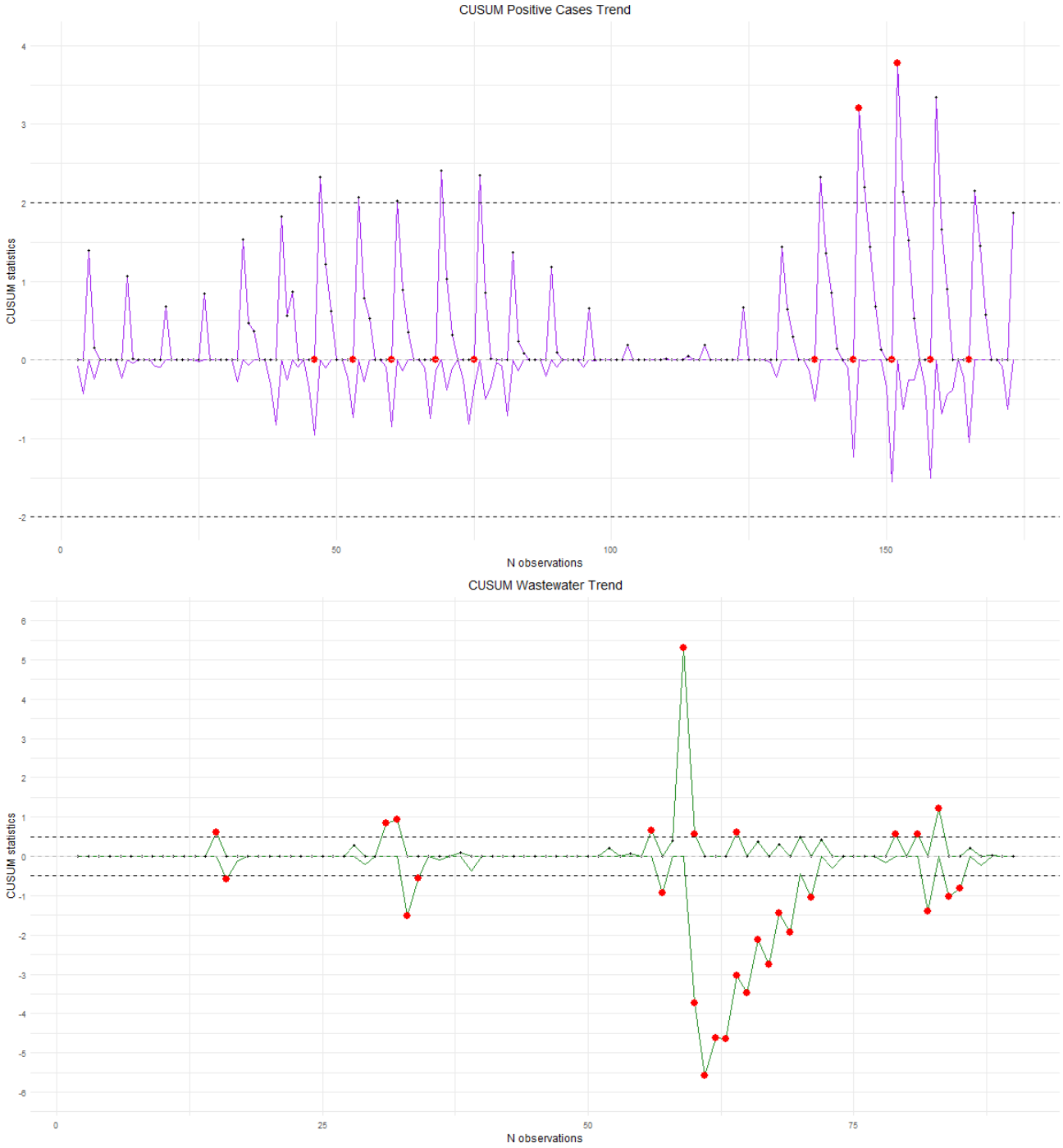
To validate the hypotheses and results, the same analysis was performed on the series of positive cases within the time frame from January 14 to August 2, 2022, (Figure 4). Furthermore, Figure 4 shows the complete series of the differential of viral load in wastewater effluent.

As shown in Table 3, seven turning points in the effluent series corresponded to turning points in the positive case series. The average number of days between a wastewater peak and a clinically confirmed case peak was 6. In particular, a peak in the effluent dated April 7, 2022, was associated with an elevated positive case count, although it did not differ significantly from other outbreak values.

The results of the comparison of the CUSUM charts for total positive cases and effluents until August 2, 2022, align with the findings observed in the analysis up to February 13, 2022. The average number of days between an effluent peak and a positive case peak was 6.

XSL•FO
RenderX

Figure 4. CUSUM analysis of positive cases for COVID-19 (top) and SARS-CoV-2 RNA trends in wastewater (bottom) in the Veneto region, Italy, February 14, 2022, to August 2, 2022. CUSUM: cumulative sum.



XSL•FO
RenderX

Table . Temporal correspondence between peaks in SARS-CoV-2 RNA concentrations in wastewater and COVID-19 cases in the Veneto region (February 14, 2022, to August 2, 2022).

Dates of wastewater peaks	SARS-CoV-2 RNA concentration, gc/µL	Dates of COVID-19 peak cases	Total positive COVID-19 cases, n	Difference in days, n
3/29/2022	227.96	4/05/2022	8205	7
4/07/2022	715.86	4/12/2022	7854	5
4/12/2022	290.86	4/20/2022	8922	8
4/26/2022	158.38	4/27/2022	8814	1
6/21/2022	109.60	6/28/2022	8753	7
6/28/2022	209.81	7/05/2022	10,884	7
7/05/2022	189.48	7/12/2022	12,571	7

Discussion

Principal findings

This study provides valuable information on the use of WBE as an early warning system for COVID-19 outbreaks in the Veneto region of Italy. Using CUSUM charts, we were able to detect significant changes in viral RNA concentrations in wastewater and assess their correlation with COVID-19 clinical outcomes, such as positive cases, hospitalizations, and admissions to the ICU.

Our findings confirm the potential of WBE to anticipate COVID-19 trends, with viral peaks in wastewater consistently preceding peaks in reported cases by an average of 5.2 days. Similarly, wastewater data provided early indications of increases in hospitalizations and admissions to the ICU, with a lag time of 4.25 and 6 days, respectively. These lead times are crucial for public health preparedness, allowing authorities to mobilize resources and implement containment measures before the number of clinical cases surges.

Our results align with previous WBE studies for the monitoring of SARS-CoV-2. Henriques et al [32], in their 2023 study, underscore the importance of WBE in assessing the presence and prevalence of SARS-CoV-2, particularly in regions with limited clinical data. WBE is a crucial complementary tool for the early detection and surveillance of COVID-19 outbreaks. It helps to estimate clinical cases and evaluate the effectiveness of vaccination programs [32].

The systematic review and meta-analysis by Li et al [33] emphasize the varying correlation between SARS-CoV-2 RNA concentrations in wastewater and clinically confirmed cases of COVID-19. Environmental factors, WBE sampling designs, and epidemiological conditions influence the strength of this correlation. Their study underscores the need to understand the dynamics of viral shedding, in-sewer decay, and sampling strategies to accurately estimate COVID-19 cases through WBE [33]. Furthermore, a study by Sodhi and Singh in 2022 [34] explores SARS-CoV-2 in wastewater, highlighting the presence of the virus even before symptoms. They advocate revised guidelines and advanced viral remediation techniques.

The temporal correlations between wastewater trends and clinical outcomes underscore the potential usefulness of WBE as a supplementary tool for public health surveillance. However, our analysis revealed a lead time of only 5 days when considering COVID-19–positive cases. Although a 5-day warning can help prepare health care resources, it comes with certain limitations. Relatively short lead times can lead to a faster but less reliable response, especially given the complexities of reorganizing health care resources and managing intensive care. Therefore, while we acknowledge the value of short-term alerts, it is crucial to strike a balance between timeliness and reliability when implementing preventive and resource management measures. To better focus on the increase in health care resources, it is essential to obtain more data on hospitalized cases and admissions to the ICU for COVID-19. However, during the considered period, these data are not robust enough to suggest a mean number of days to prepare for an increase in the number of beds for admitted patients.

The analysis in Table 2 shows a 13-day gap between the peak in the SARS-CoV-2 RNA concentration in wastewater on January 4, 2022, and the peak in reported cases of COVID-19 on January 17, 2022. This longer-than-expected delay highlights the variability inherent in WBE. Several factors could explain this discrepancy, including complex disease dynamics, inconsistent lag times due to varying public health responses, and potential delays in testing and reporting during periods of high case numbers. These factors underscore the need for careful interpretation of the correlations between wastewater data and clinical case trends, as the relationship may not always be linear or immediate.

The comparison between the fall of 2021 and the spring of 2022 reveals a potential gap in clinical surveillance, identified through the analysis of divergent trends between wastewater concentrations and reported cases of COVID-19. Specifically, while wastewater data indicated fluctuations in viral RNA concentrations, the number of reported COVID-19 clinical cases began to decline from April 2022, suggesting either a decrease in testing or reporting accuracy or a genuine decrease in cases not fully captured by the wastewater analysis. This divergence highlights the importance of integrating multiple data sources, including wastewater and clinical data, to provide a comprehensive overview of disease trends.

Despite current challenges, future research aims to improve sampling and analysis protocols, understand viral dynamics in sewer systems, and exploit artificial intelligence and big data to better monitor wastewater. The study by Nayak et al [18]

XSL•FO
RenderX

highlights the potential of machine learning, particularly in diverse sewer sheds, using time series models such as long short-term memory and various data parameters for accurate trend forecasting. In parallel, the work by Jiang et al [35] underscores the crucial role of WBE in tracking COVID-19 transmission.

In the field of WBE, several methodologies have been developed to predict future COVID-19 cases, including our method, the EPISENS-M technique of Ando et al, and the statistical framework proposed by Dai et al. Although all 3 methods share the objective of using wastewater data to monitor and forecast the evolution of the pandemic, they differ significantly in their approach, robustness, and applicability.

The EPISENS-M method, developed by Ando et al [20], represents a significant advancement in the sensitive detection of SARS-CoV-2 RNA in wastewater. This method utilizes membrane adsorption and direct RNA extraction to enhance detection sensitivity, allowing for accurate prediction of clinical cases even in low-prevalence settings. However, a key limitation of this approach is its reliance on recent clinical data for calibration, which may reduce its effectiveness in situations where such data are not readily available.

On the other hand, Dai et al [21] developed a sophisticated statistical framework based on Bayesian models and functional principal component analysis to address the challenges posed by noisy and sparsely sampled data. Although their approach is extremely useful for interpreting viral concentration data in wastewater and mitigating fluctuations due to technical errors, it is not designed to directly predict clinical cases of COVID-19. This limits its applicability as a direct epidemiological forecasting tool.

The use of the CUSUM method in this study marks a novel approach to identifying subtle shifts in SARS-CoV-2 trends, offering a more sensitive detection of deviations from established baselines. Unlike EPISENS-M, our method is less dependent on recent clinical data, making it potentially more robust in contexts with limited clinical monitoring. Furthermore, compared to the method of Dai et al, our approach focuses specifically on epidemiological prediction, providing a direct advantage for public health planning.

Limitations

Despite its potential, our study has limitations. The reliance on a single region and a specific time frame restricts the generalizability of our findings.

Furthermore, the significant heterogeneity observed in flow rates, suspended solid concentrations, and volume distribution provides information on the differences in the sites considered. The sites reflect the data on specific cities; so, while we have the number of positive cases per province, we need to exercise caution when generalizing these findings to the whole region. Additionally, the study's reliance on SARS-CoV-2 RNA concentrations as an early indicator may overlook other potential contributors to wastewater dynamics and require further validation through complementary approaches.

This study focused on the correlation between SARS-CoV-2 RNA concentrations in wastewater and clinical indicators of COVID-19, without accounting for the dynamics of population mobility or additional biomarkers in wastewater. While these factors can influence wastewater analysis and provide deeper insights, they were beyond the scope of the current study. Future research could integrate mobility data and additional biomarkers to improve the predictive capacity of WBE.

Validation analyses were performed only for positive cases due to a low number of peaks for hospitalized and ICU cases. Therefore, the study acknowledges the need for further validation and standardization of analytical methods in wastewater analysis. Additionally, while the study highlights correlations, the cause cannot be definitively established. Future research could explore the generalizability of these findings in different regions and populations.

Conclusions

This study confirms that WBE, even when applied at a regional level with limited data points, can provide early warnings for COVID-19 outbreaks. The use of CUSUM charts enabled the sensitive and timely detection of shifts in viral load trends, offering crucial lead times for public health interventions. Our findings highlight the value of integrating WBE into public health surveillance systems, especially in areas where clinical data may not be immediately available.

The results confirm the reliability of wastewater analysis in anticipating COVID-19 trends, particularly at the LHU level, which is vital for informed public health planning and proactive responses. The potential of WBE to monitor the spread of SARS-CoV-2, including emerging variants, underscores the importance of continued efforts to improve wastewater monitoring and treatment.

More research is essential to refine the detection and quantification methods for pathogens and viruses in wastewater, as well as to evaluate their feasibility in different contexts. Ultimately, this approach offers a robust and valuable tool for epidemiological surveillance, contributing significantly to the early detection and management of infectious disease outbreaks with pandemic potential.

Acknowledgments

This work was supported by the scope of the Centro Nazionale per la Prevenzione e il Controllo delle Malattie of the Italian Ministry of Health 2020 Project "Modelli organizzativi per l'efficientamento delle attività di contact tracing" (Organizational models to improve the efficiency of contact tracing activities) funded by the Italian Ministry of Health. We appreciate the assistance of the staff of the Consorzio per la Ricerca Sanitaria, who helped us with the administrative aspects of the project.

XSL•FO
RenderX

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

HO and DG contributed to the conception and design of the study. IA, TB, FR, MZ, and VG collected the data. IA, TB, FR, and MZ conducted the laboratory analysis. HO and MZ conducted the statistical analysis. HO, MZ, and all other authors interpreted the results. HO prepared the draft manuscript. FDR and VB provided resources. All authors reviewed the results and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Ahmad J, Ahmad M, Usman ARA, Al-Wabel MI. Prevalence of human pathogenic viruses in wastewater: a potential transmission risk as well as an effective tool for early outbreak detection for COVID-19. J Environ Manage 2021 Nov 15;298:113486. [doi: 10.1016/j.jenvman.2021.113486] [Medline: 34391102]
2. Asadi A, Fakhri Y, Salimi Y, Daglioglu N, Tahmasebifard M, Aghajarinezhad M. Nicotine consumption rate through wastewater-based epidemiology: a systematic review, meta-analysis and probabilistic risk assessment. Environ Sci Pollut Res Int 2023 May;30(23):63416-63426. [doi: 10.1007/s11356-023-27017-x] [Medline: 37084052]
3. Tiwari A, Kurittu P, Al-Mustapha AI, et al. Wastewater surveillance of antibiotic-resistant bacterial pathogens: a systematic review. Front Microbiol 2022 Dec 15;13:977106. [doi: 10.3389/fmicb.2022.977106] [Medline: 36590429]
4. Chau KK, Barker L, Budgell EP, et al. Systematic review of wastewater surveillance of antimicrobial resistance in human populations. Environ Int 2022 Apr;162:107171. [doi: 10.1016/j.envint.2022.107171] [Medline: 35290866]
5. Holton E, Sims N, Jagadeesan K, Standerwick R, Kasprzyk-Hordern B. Quantifying community-wide antimicrobials usage via wastewater-based epidemiology. J Hazard Mater 2022 Aug 15;436:129001. [doi: 10.1016/j.jhazmat.2022.129001] [Medline: 35594673]
6. La Rosa G, Iaconelli M, Mancini P, et al. First detection of SARS-CoV-2 in untreated wastewaters in Italy. Sci Total Environ 2020 Sep 20;736:139652. [doi: 10.1016/j.scitotenv.2020.139652] [Medline: 32464333]
7. Giacobbo A, Rodrigues MAS, Zoppas Ferreira J, Bernardes AM, de Pinho MN. A critical review on SARS-CoV-2 infectivity in water and wastewater. What do we know? Sci Total Environ 2021 Jun 20;774:145721. [doi: 10.1016/j.scitotenv.2021.145721] [Medline: 33610994]
8. Han C, Duan C, Zhang S, et al. Digestive symptoms in COVID-19 patients with mild disease severity: clinical presentation, stool viral RNA testing, and outcomes. Am J Gastroenterol 2020 Jun;115(6):916-923. [doi: 10.14309/ajg.0000000000000664] [Medline: 32301761]
9. Rimoldi SG, Stefani F, Gigantiello A, et al. Presence and infectivity of SARS-CoV-2 virus in wastewaters and rivers. Sci Total Environ 2020 Nov 20;744:140911. [doi: 10.1016/j.scitotenv.2020.140911] [Medline: 32693284]
10. La Rosa G, Mancini P, Bonanno Ferraro G, et al. SARS-CoV-2 has been circulating in northern Italy since December 2019: evidence from environmental monitoring. Sci Total Environ 2021 Jan 1;750:141711. [doi: 10.1016/j.scitotenv.2020.141711] [Medline: 32835962]
11. Randazzo W, Truchado P, Cuevas-Ferrando E, Simón P, Allende A, Sánchez G. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. Water Res 2020 Aug 15;181:115942. [doi: 10.1016/j.watres.2020.115942] [Medline: 32425251]
12. Wu F, Xiao A, Zhang J, et al. Wastewater surveillance of SARS-CoV-2 across 40 U.S. states from February to June 2020. Water Res 2021 Sep 1;202:117400. [doi: 10.1016/j.watres.2021.117400] [Medline: 34274898]
13. Ciannella S, González-Fernández C, Gomez-Pastora J. Recent progress on wastewater-based epidemiology for COVID-19 surveillance: a systematic review of analytical procedures and epidemiological modeling. Sci Total Environ 2023 Jun 20;878:162953. [doi: 10.1016/j.scitotenv.2023.162953] [Medline: 36948304]
14. Peccia J, Zulli A, Brackney DE, et al. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. Nat Biotechnol 2020 Oct;38(10):1164-1167. [doi: 10.1038/s41587-020-0684-z] [Medline: 32948856]
15. Vallejo JA, Trigo-Tasende N, Rumbo-Feal S, et al. Modeling the number of people infected with SARS-COV-2 from wastewater viral load in Northwest Spain. Sci Total Environ 2022 Mar 10;811:152334. [doi: 10.1016/j.scitotenv.2021.152334] [Medline: 34921882]
16. Chekkala A, Atasoy M, Williams C, Cetecioglu Z. Statistical analysis of SARS-CoV-2 using wastewater-based data of Stockholm, Sweden. Int J Environ Res Public Health 2023 Feb 26;20(5):4181. [doi: 10.3390/ijerph20054181] [Medline: 36901194]
17. Howard I, Espigares E, Lardelli P, Martín JL, Espigares M. Evaluation of microbiological and physicochemical indicators for wastewater treatment. Environ Toxicol 2004 Jun;19(3):241-249. [doi: 10.1002/tox.20016] [Medline: 15101039]

XSL•FO
RenderX

18. Nayak SR, Arora V, Sinha U, Poonia RC. A statistical analysis of COVID-19 using Gaussian and probabilistic model. J Interdiscip Math 2020 Nov 25;24(1):19-32. [doi: 10.1080/09720502.2020.1833442]
19. Karthikeyan S, Ronquillo N, Belda-Ferre P, et al. High-throughput wastewater SARS-CoV-2 detection enables forecasting of community infection dynamics in San Diego County. mSystems 2021 Mar 2;6(2):e00045-21. [doi: 10.1128/mSystems.00045-21] [Medline: 33653938]
20. Ando H, Murakami M, Ahmed W, Iwamoto R, Okabe S, Kitajima M. Wastewater-based prediction of COVID-19 cases using a highly sensitive SARS-CoV-2 RNA detection method combined with mathematical modeling. Environ Int 2023 Mar;173:107743. [doi: 10.1016/j.envint.2023.107743] [Medline: 36867995]
21. Dai X, Champredon D, Fazil A, et al. Statistical framework to support the epidemiological interpretation of SARS-CoV-2 concentration in municipal wastewater. Sci Rep 2022 Aug 5;12(1):13490. [doi: 10.1038/s41598-022-17543-y] [Medline: 35931713]
22. Noyez L. Control charts, Cusum techniques and funnel plots. A review of methods for monitoring performance in healthcare. Interact Cardiovasc Thorac Surg 2009 Sep;9(3):494-499. [doi: 10.1510/icvts.2009.204768] [Medline: 19509097]
23. Baldovin T, Amoruso I, Fonzo M, et al. Trends in SARS-CoV-2 clinically confirmed cases and viral load in wastewater: a critical alignment for Padua city (NE Italy). Heliyon 2023 Oct;9(10):e20571. [doi: 10.1016/j.heliyon.2023.e20571] [Medline: 37822618]
24. La Rosa G, Bonadonna L, Suffredini E. Protocollo della sorveglianza di SARS-cov-2 in reflui urbani (SARI) - rev. 3 [Article in Italian]. Zenodo. 2021. URL: https://zenodo.org/records/5758725 [accessed 2024-12-09]
25. Page ES. Continuous Inspection Schemes. Biometrika 1954 Jun;41(1-2):100-115. [doi: 10.1093/biomet/41.1-2.100]
26. Novoa NM, Varela G. Monitoring surgical quality: the cumulative sum (CUSUM) approach. Mediastinum 2020 Mar 25;4:4. [doi: 10.21037/med.2019.10.01] [Medline: 35118272]
27. Govindaraju K, Lai CD. Run length variability and three sigma control limits. Stoch Qual Control 2004;19(2):175-184. [doi: 10.1515/EQC.2004.175]
28. The R Project for Statistical Computing. URL: https://www.R-project.org/ [accessed 2024-11-19]
29. Scrucca L. Qcc: an R package for quality control charting and statistical process control. R News 2004 Jun;4(1):11-17 [FREE Full text]
30. Wickham H. Data analysis. In: ggplot2: Elegant Graphics for Data Analysis, 2nd edition: Springer; 2016:189-201. [doi: 10.1007/978-3-319-24277-4_9]
31. Wickham H. ggplot2: Elegant Graphics for Data Analysis, 2nd edition: Springer; 2016. [doi: 10.1007/978-3-319-24277-4]
32. Henriques TB, Cassini ST, de Pinho Keller R. Contribution of wastewater-based epidemiology to SARS-CoV-2 screening in Brazil and the United States. J Water Health 2023 Mar;21(3):343-353. [doi: 10.2166/wh.2023.260] [Medline: 37338314]
33. Li X, Zhang S, Sherchan S, et al. Correlation between SARS-CoV-2 RNA concentration in wastewater and COVID-19 cases in community: a systematic review and meta-analysis. J Hazard Mater 2023 Jan 5;441:129848. [doi: 10.1016/j.jhazmat.2022.129848] [Medline: 36067562]
34. Sodhi KK, Singh CK. A systematic review on the occurrence, fate, and remediation of SARS-CoV-2 in wastewater. Int J Environ Sci Technol (Tehran) 2023;20(7):8073-8086. [doi: 10.1007/s13762-022-04326-1] [Medline: 35755183]
35. Jiang G, Liu Y, Tang S, et al. Moving forward with COVID-19: future research prospects of wastewater-based epidemiology methodologies and applications. Curr Opin Environ Sci Health 2023 Jun;33:100458. [doi: 10.1016/j.coesh.2023.100458] [Medline: 37034453]

Abbreviations

ARIMA: autoregressive integrated moving average
CUSUM: cumulative sum
EPISENS-M: Efficient and Practical Virus Identification System with Enhanced Sensitivity for Membrane
gc: genome copies
ICU: intensive care unit
LHU: local health unit
ORF1ab: open reading frames 1ab
PCR: polymerase chain reaction
WBE: wastewater-based epidemiology
WTP: wastewater treatment plant

XSL•FO
RenderX

Edited by A Mavragani; submitted 27.06.24; peer-reviewed by A Berzins, RK Mishra, RDF Bueno; revised version received 31.08.24; accepted 02.09.24; published 14.01.25.

Please cite as:
Ocagli H, Zambito M, Da Re F, Groppi V, Zampini M, Terrini A, Rigoli F, Amoruso I, Baldovin T, Baldo V, Russo F, Gregori D
Wastewater Monitoring During the COVID-19 Pandemic in the Veneto Region, Italy: Longitudinal Observational Study
JMIR Public Health Surveill 2025;11:e58862
URL: https://publichealth.jmir.org/2025/1/e58862
doi:10.2196/58862

© Honoria Ocagli, Marco Zambito, Filippo Da Re, Vanessa Groppi, Marco Zampini, Alessia Terrini, Franco Rigoli, Irene Amoruso, Tatjana Baldovin, Vincenzo Baldo, Francesca Russo, Dario Gregori. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Monitoring Public Health Through a Comprehensive Primary Care Database in the Netherlands: Overview of the Nivel Syndromic Surveillance System

Christos Baliatsas[1], PhD; Jojanneke van Summeren[1], PhD; Sander van Beusekom[1]; Amy Matser[1,2,3,4], PhD; Mariette Hooiveld[1], PhD

[1]Nivel – Netherlands Institute for Health Services Research, Otterstraat 118, Utrecht, The Netherlands

[2]Infectious Diseases, Amsterdam UMC, location University of Amsterdam, Amsterdam, The Netherlands

[3]Infectious Diseases, Amsterdam Institute for Infection and Immunity, Amsterdam, The Netherlands

[4]Amsterdam Public Health, Amsterdam, The Netherlands

Corresponding Author:
Christos Baliatsas, PhD
Nivel – Netherlands Institute for Health Services Research, Otterstraat 118, Utrecht, The Netherlands

Abstract

Background: Syndromic surveillance systems are crucial for the monitoring of population health and the early detection of emerging health problems. Internationally, there are numerous established systems reporting on different types of data. In the Netherlands, the Nivel syndromic surveillance system provides real-time monitoring on all diseases and symptoms presented in general practice.

Objective: The present article introduces the national syndromic surveillance system in primary care, emphasizing its role in providing real-time information on infectious diseases and various health problems at the population level, in the Netherlands. In addition, we report on the central role of the participating general practices in data provision, and discuss the applicability of the syndromic surveillance data in different contexts of public health research.

Methods: The Nivel syndromic surveillance system is part of the Nivel Primary Care Database (Nivel-PCD) that collects routinely recorded data from electronic health records of about 10% of the Dutch population, on the basis of approximately 500 practices. This translates to approximately 1.9 million citizens. Since 2010, the surveillance system relies on representative, pseudonymized data collected on a weekly basis from a subset of about 400 practices in the Nivel-PCD, for the entire practice population. Health problems are registered according to the International Classification of Primary Care, applied in all general practices in the Netherlands. Prevalence rates are recalculated and reported every week in the form of figures, also stratified by age, sex, and region. Weekly rates are defined as the number of people that consulted the general practitioner in a certain week for a specific health problem, divided by the total number of registered individuals in the practice.

Results: While utilizing data from general practitioners' electronic health records, the system allows for the timely monitoring and identification of symptom and disease patterns and trends, not only among individuals who seek primary health care, but the entire registered population. Besides their use in disease monitoring, syndromic surveillance data are useful in various public health research contexts, such as environmental health and disaster research.

Conclusions: The Nivel syndromic surveillance system serves as a valuable tool for health monitoring and research, offering valuable insights into public health.

(JMIR Public Health Surveill 2025;11:e58767) doi:10.2196/58767

KEYWORDS

surveillance; monitoring; general practice; public health

Introduction

Syndromic surveillance comprises the systematic, automated collection, analysis, and evaluation of real-time health data for the early monitoring and detection of changes in health indicators [1-3]. Data can be routinely collected from different sources, such as primary (general practitioners; GPs) or secondary care (emergency department visits) [4-6], depending on the health system structure. Other information sources can also be considered, such as helpline calls, medication sales, and internet searches, although they are less commonly used [4,6,7]. The broad scope of syndromic surveillance and the rapid data collection procedures can provide great flexibility for timely assessment of different health outcomes on a large population

XSL•FO
RenderX

scale [6-8]. Many countries, also in Europe, have successfully incorporated syndromic surveillance into national surveillance systems [5,9-12], the role of which was crucial during the COVID-19 pandemic, despite inherent limitations [7,8,13,14].

There are well-established national networks that make use of primary care data under strict privacy regulations, for health (services) research and monitoring, for instance in the United Kingdom and Belgium [5,15,16], offering important insights into population health. In the United States, the integration of routine anonymized primary care data is more limited, primarily due to regulatory restrictions, with greater reliance on secondary health care data [17]. In the Netherlands, the Nivel (Netherlands Institute for Health Services Research) Primary Care Database (Nivel-PCD; formerly known as the Netherlands Information Network of General Practice) [17,18] was developed in the early 1990s.

Within this infrastructure, the Nivel syndromic surveillance system provides real-time information on all diseases and symptoms presented in general practice, utilizing data from a large network of GPs. Despite its important role nationwide, there is currently relatively limited international visibility of this infrastructure as well as its applicability in different research contexts. The present paper has the following main objectives: (1) to provide an overview of the Nivel syndromic surveillance system, detailing its methodology; (2) to discuss the important role of the participating general practices in providing data; and (3) to outline the diverse applications of syndromic surveillance data in public health research, highlighting its utility in identifying emerging health threats and supporting research on the health impacts of various exposures.

Methods

Standardized Registration in Dutch General Practices

In the Netherlands, GPs play a vital role as the first point of contact for health care services, acting as gatekeepers for specialized secondary care. Almost every individual is registered at a general practice located in the broader vicinity of their residence. GPs keep detailed electronic health records (EHRs) of consultations, diagnoses, prescribed medications, and referrals. Dutch GPs adhere to the NHG (Dutch College of General Practitioners) guidelines, specifically the ADEPD (adequate record keeping in electronic patient records) guideline, which ensures consistency, reproducibility, and standardization in medical registries, on the basis of a well-structured classification system [19,20]. More specifically, health problems and symptoms documented in general practice adhere to the International Classification of Primary Care (ICPC) [21,22], while medication prescriptions are classified according to the anatomical therapeutic chemical (ATC) classification system [23]. This EHR registration system ensures that information on both consulting and nonconsulting populations is available, which is of high importance for epidemiological research and monitoring. The data recorded by general practices, therefore, provide a comprehensive view of the health characteristics of not only primary care patients but also the broader population.

Nivel-PCD: A Large, Nationally Representative Primary Care Database on Routinely Recorded Data

The Nivel is an independent, nonprofit, research institute that is part of the Dutch Ministry of Health's knowledge infrastructure. The Nivel-PCD [18] is a large research infrastructure based on real world data extracted from routinely kept EHRs of GPs; primary care out-of-hours services; and paramedical care physicians such as physiotherapists, exercise therapists, and dieticians. There are approximately 500 general practices across the Netherlands that participate yearly in the Nivel-PCD [18], while the total number of general practices in the country is nearly 5000, comprising about 12,000 GPs [24].

Data Collection and Processing

The Nivel gathers data from approximately 10% of the Dutch population (~1.9 million citizens) registered in general practices and about two-thirds of the population considering the use of out-of-hours services. Additionally, data are collected from sentinel general practices (Nivel Peilstations), where a subset of patients with influenza-like illness or other acute respiratory infections provide samples for laboratory testing. These samples are analyzed by the National Institute for Public Health and the Environment (RIVM) to identify the presence of influenza and other respiratory viruses.

For the Nivel syndromic surveillance system, EHR data are collected, extracted, and delivered by the EHR software providers used by the general practices. The system relies on pseudonymized data, available weekly since 2010, on patients consulting for certain health problems and conditions recorded in approximately 400 general practices from all over the country that provide data on a weekly basis. Pseudonymization refers to assigning nonidentifiable identification numbers to the registered practice population. The focus lies on observing and documenting trends in symptoms and diseases, presenting a timely and realistic view of health problems presented in primary care, including influenza-like illness, COVID-19, and other infectious and noninfectious diseases in the Dutch population.

The Nivel-PCD adheres to Dutch data protection regulations and laws for health data use in epidemiological research (Dutch Civil Law, Article 7:458). Medical and personal information are separated with Trusted Third Party support, preventing access to identifiable patient details. Participating health care providers can withdraw from the Nivel-PCD anytime. An opt-out system is available for patients who do not want their data to be used. The Nivel-PCD is ISO27001 certified.

Ethical Considerations

The use of EHRs for research purposes is allowed under certain conditions. When these conditions are fulfilled, neither obtaining informed consent from patients nor approval by a medical ethics committee is obligatory for this type of observational studies containing no directly identifiable data (art. 24 GDPR Implementation Act jo art. 9.2 sub j GDPR).

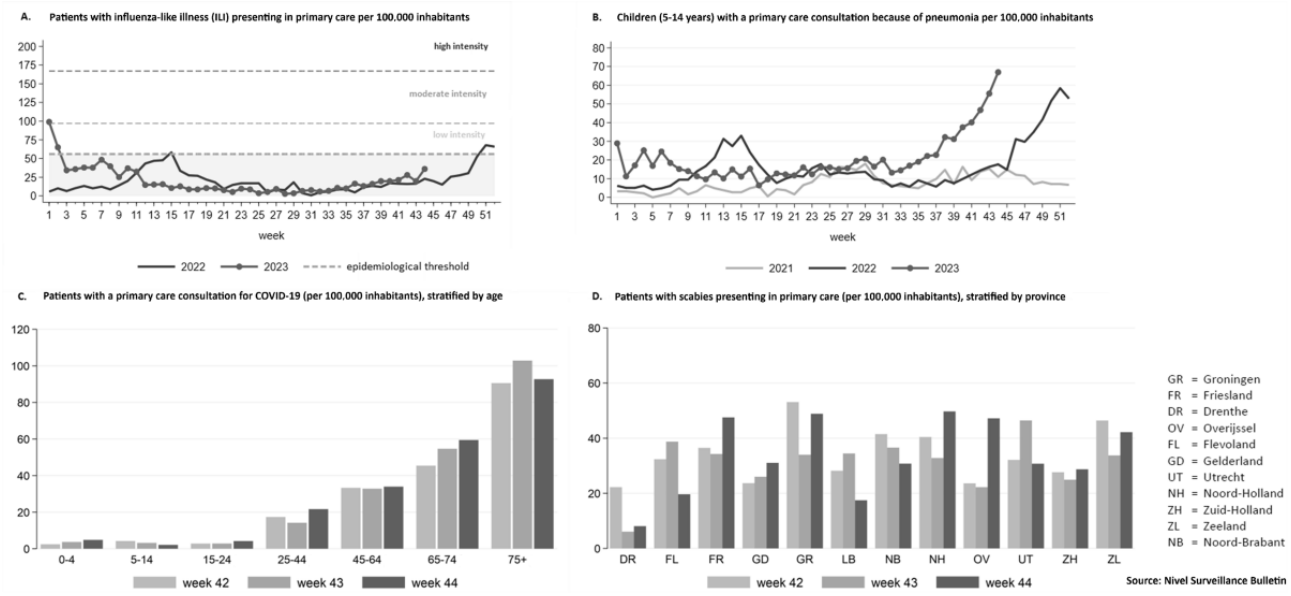
XSL•FO
RenderX

Results

Analysis and Reporting

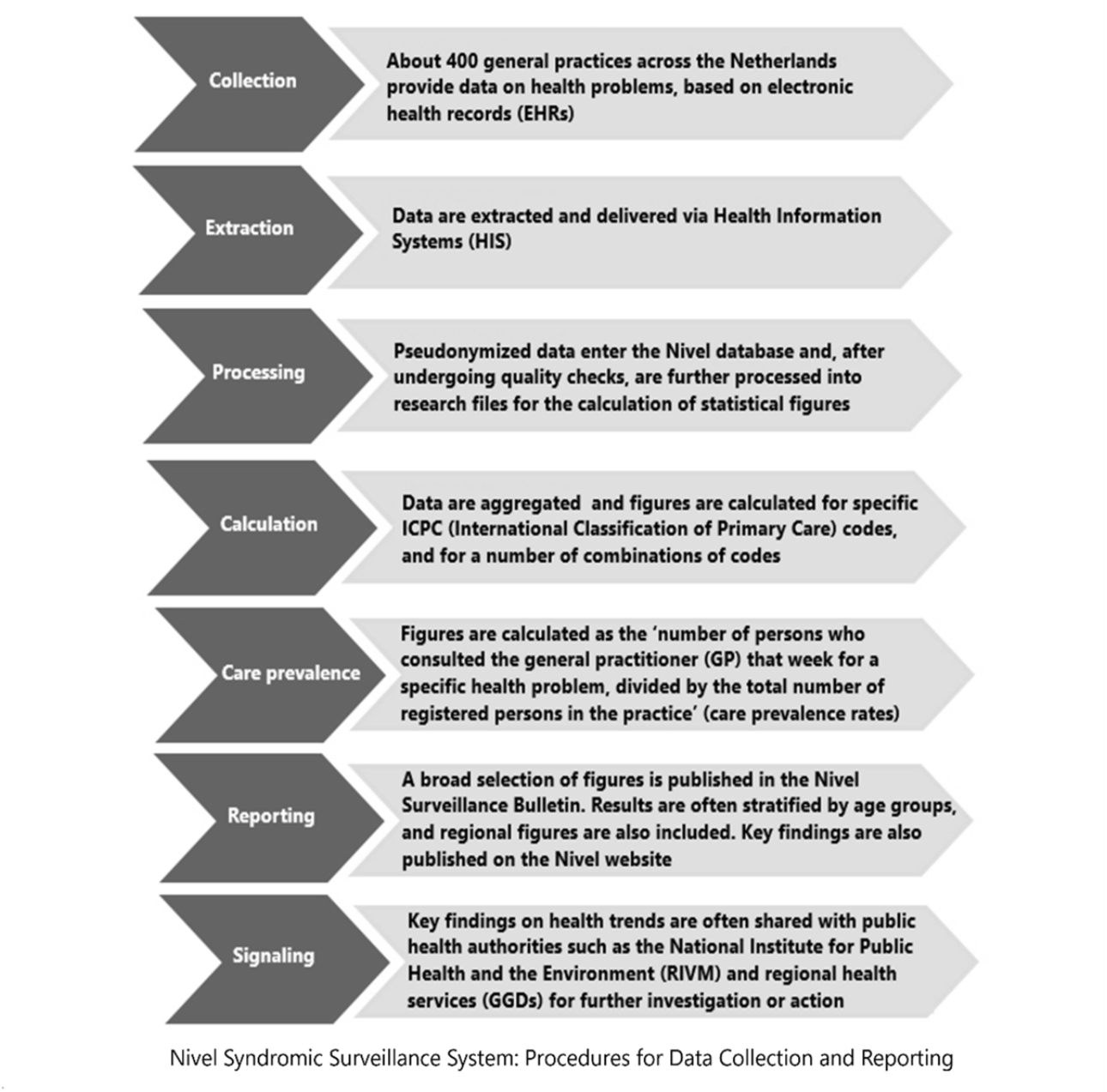
Rates are generated for specific diagnostic codes and for a number of ICPC code combinations (eg, R74, R75, R77, R78, R80, and R83.03 for the cluster "acute respiratory infections excluding pneumonia"); the rates are defined as the number of people who have consulted the GP in that week for a certain health problem, divided by the total number of registered people in the practice. The weekly care prevalence rates are also stratified by age and sex, and by province or public health regions where the local Public Health Services operate, to monitor regional differences. The figures obtained in the previous week are recalculated weekly with any additional data provided, in order to be as up-to-date and complete as possible. The key findings are summarized weekly and published in the Nivel Surveillance Bulletin, accessible on the Nivel website [25] (see example in Figure 1). The report includes selected figures, often categorized by age groups and regions, with additional figures available upon request. An example is the notable increase in the number of children with pneumonia, reported by the Nivel syndromic surveillance system (Figure 1). This early warning initiated further investigation by the Nivel and National Institute for Public Health and the Environment (RIVM) to determine the cause of this increase. Since the start of the COVID-19 pandemic in early 2020, the system has been monitoring cases in primary care. First, the system used the free text describing the disease episode and manually categorized patients clinically diagnosed with COVID-19 or COVID-19-like illness. After the introduction of the ICPC subcode R83.03 in November 2020, the monitoring continued using the registration of this diagnosis code. Figure 2 summarizes the main steps in data collection, analysis and reporting of the Nivel syndromic surveillance system.

Figure 1. Example of four data representations from the Nivel Surveillance Bulletin as of week 44, 2023: **A.** Number of patients with influenza-like illness presenting in primary care per week; **B.** Weekly number of patients in the age group 5 - 14 years with a primary care consultation because of pneumonia; **C.** Weekly number of patients with a primary care consultation because of COVID-19, by age group; **D.** Number of patients with scabies presenting in primary care per week, by province. All numbers are expressed per 100,000 inhabitants.



XSL•FO
RenderX

Figure 2. Overview of the main steps and tasks involved in data collection and report of the Nivel syndromic surveillance system.



Diverse Applications of Syndromic Surveillance in Different Population Research Contexts

Syndromic surveillance also extends beyond monitoring health problems reported in the Surveillance bulletin. Its infrastructure has been used for studies on different topics and contexts, for instance on the physical and psychosocial impact of the COVID-19 pandemic [26,27] as well as to monitor the presentation of symptoms in primary care in relation to environmental exposures and heat waves [28,29]. The comprehensive patient registry in the Nivel-PCD enables continuous tracking of patients over time, linked with additional sources such as socioeconomic data from Statistics Netherlands (CBS), geographic information, and environmental exposure data. Integration occurs across aggregated levels using pseudonyms to ensure patient privacy.

Discussion

Study Findings and Comparisons With Previous Works

The Nivel syndromic surveillance system provides large scale monitoring of diseases and health symptoms, based on representative population health data routinely registered in general practice in the Netherlands, on the basis of unified registration guidelines and strict privacy regulations. The system's infrastructure can also be used in different contexts of public health research, including studies on the physical and psychosocial impact of pandemics as well as environmental exposures. It provides information to the stakeholders through the openly accessible Surveillance Bulletin and signals early warnings to the national and local public health authorities.

XSL•FO
RenderX

Nivel syndromic surveillance allows for the observation of patterns and trends not only among individuals who seek care, but the entire registered population. This is an important strength compared to syndromic surveillance systems that solely rely on emergency department visits [10], which provide a rapid picture of health trends but are limited to the severe cases. The monitoring of broad categories of health problems such as respiratory infections enables syndromic surveillance to serve as an early warning mechanism for novel threats, providing information earlier than laboratory confirmation [6]. The data of the Nivel syndromic surveillance system can also be used in different public health research contexts, providing insight into the extent of public health impacts, identify vulnerable populations, and guide appropriate mitigation strategies.

Limitations

It is important to acknowledge the limitations and inherent complexities associated with syndromic surveillance. One of them is the reliance of syndromic surveillance on reporting clinical diagnoses or symptoms rather than diagnoses confirmed by a laboratory. This could introduce some degree of uncertainty, given that symptoms may not always be associated with a specific disease. In addition, it could increase the possibility of a delayed diagnosis registration compared to surveillance infrastructures that use laboratory-confirmed cases or emergency visits. Identifying rare or emerging diseases without pre-established diagnostic codes or historical data poses another challenge; a recent example is the miscoding and misclassification of COVID-19 cases in primary care, during the pandemic [30]. Establishing a reliable baseline for certain health problems can also be difficult due to different factors such as seasonality and changes in health care–seeking behavior. Technical aspects such as updates in guidelines and coding practices in primary care may also disrupt trends. Moreover, there is currently a lack of quantitative data from other surveillance systems to support a structured comparative evaluation [31] of the system's effectiveness regarding indicators such as early detection.

Future Directions

As the field of syndromic surveillance evolves, there are several challenges to be addressed that could substantially strengthen the utility of the existing system in public health monitoring. Despite the satisfactory national coverage rate of general practices, participation of practices in the system is voluntary and can vary per region [32]. The expansion of the system's geographical coverage with the inclusion of more general practices across the country would strengthen the sample power as well as representativeness, allowing for a more accurate detection of disease trends and variation at the regional and population group levels. In light of climate change and emerging crises, linking syndromic surveillance with publicly available environmental exposure data could allow for the monitoring of the physical and mental impact on the population's health more accurately, especially during seasonal events such as heatwaves or acute incidents of hazardous emissions. Another aspect would be the integration of additional data sources in the syndromic surveillance system such as those on emergency visits, similar to other established surveillance systems [9], in order to provide a more comprehensive monitoring of the population's health. Strengthening and expanding international collaborations and networks with other syndromic surveillance systems could also help identify existing caveats and develop best practices to address them and would improve early warning of potential health threats across countries [33]. Finally, innovation should be an indispensable part of syndromic surveillance.

In particular, machine learning approaches as well as artificial intelligence are promising areas that could contribute to more accurate and efficient health monitoring, by facilitating analysis processes, improving detection algorithms, and even the development of predictive models [34].

Conclusions

The Nivel syndromic surveillance system is an important infrastructure in providing monitoring of real world data on the occurrence of all symptoms and diseases presented in general practice, at the population level. Further implementation and refinement of syndromic surveillance methodologies towards data completeness, quality, and representativeness will enhance the system's effectiveness in contributing to public health research and response strategies.

Acknowledgments

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Data Availability

Data are available upon request (zorgregistraties@nivel.nl).

Conflicts of Interest

None declared.

References

1. Triple S Project. Assessment of syndromic surveillance in Europe. Lancet 2011 Nov 26;378(9806):1833-1834. [doi: 10.1016/S0140-6736(11)60834-9] [Medline: 22118433]

XSL•FO
RenderX

2. Ziemann A, Fouillet A, Brand H, Krafft T. Success factors of European syndromic surveillance systems: a worked example of applying qualitative comparative analysis. PLoS ONE 2016;11(5):e0155535. [doi: 10.1371/journal.pone.0155535] [Medline: 27182731]
3. Elliot AJ, Hughes HE, Harcourt SE, Morbey RA, Smith S, Smith GE. Real-time syndromic surveillance. In: WHO Guidance on Research Methods for Health Emergency and Disaster Risk Management: World Health Organization; 2021. URL: https://extranet.who.int/kobe_centre/sites/default/files/chapters/Chapter%204.9%20Slideshow.pdf
4. Duijster JW, Doreleijers SDA, Pilot E, et al. Utility of emergency call centre, dispatch and ambulance data for syndromic surveillance of infectious diseases: a scoping review. Eur J Public Health 2020 Aug 1;30(4):639-647. [doi: 10.1093/eurpub/ckz177] [Medline: 31605491]
5. Bednarska NG, Smith S, Bardsley M, et al. Trends in general practitioner consultations for hand foot and mouth disease in England between 2017 and 2022. Epidemiol Infect 2025 Jan 13;153:e22. [doi: 10.1017/S095026882400181X] [Medline: 39801026]
6. Adedire O, Love NK, Hughes HE, Buchan I, Vivancos R, Elliot AJ. Early detection and monitoring of gastrointestinal infections using syndromic surveillance: a systematic review. Int J Environ Res Public Health 2024 Apr 17;21(4):489. [doi: 10.3390/ijerph21040489] [Medline: 38673400]
7. Corrao G, Bonaugurio AS, Bagarella G, et al. Does syndromic surveillance assist public health practice in early detecting respiratory epidemics? Evidence from a wide Italian retrospective experience. J Infect Public Health 2025 Feb;18(2):102621. [doi: 10.1016/j.jiph.2024.102621] [Medline: 39729672]
8. Vos B, Debouverie L, Doggen K, et al. Monitoring COVID-19 in Belgian general practice: a tool for syndromic surveillance based on electronic health records. Eur J Gen Pract 2024 Dec;30(1):2293699. [doi: 10.1080/13814788.2023.2293699] [Medline: 38186340]
9. Smith GE, Elliot AJ, Lake I, et al. Syndromic surveillance: two decades experience of sustainable systems - its people not just data!. Epidemiol Infect 2019 Jan;147:e101. [doi: 10.1017/S0950268819000074] [Medline: 30869042]
10. Hughes HE, Edeghere O, O'Brien SJ, Vivancos R, Elliot AJ. Emergency department syndromic surveillance systems: a systematic review. BMC Public Health 2020 Dec 9;20(1):1891. [doi: 10.1186/s12889-020-09949-y] [Medline: 33298000]
11. Boender TS, Cai W, Schranz M, et al. Using routine emergency department data for syndromic surveillance of acute respiratory illness, Germany, week 10 2017 until week 10 2021. Euro Surveill 2022 Jul;27(27):2100865. [doi: 10.2807/1560-7917.ES.2022.27.27.2100865] [Medline: 35801521]
12. de Fougerolles TR, Damm O, Ansaldi F, et al. National influenza surveillance systems in five European countries: a qualitative comparative framework based on WHO guidance. BMC Public Health 2022 Jun 9;22(1):1151. [doi: 10.1186/s12889-022-13433-0] [Medline: 35681199]
13. Thiam MM, Pontais I, Forgeot C, et al. Syndromic surveillance: a key component of population health monitoring during the first wave of the COVID-19 outbreak in France, February-June 2020. PLoS ONE 2022;17(2):e0260150. [doi: 10.1371/journal.pone.0260150] [Medline: 35143501]
14. Brainard J, Lake IR, Morbey RA, Jones NR, Elliot AJ, Hunter PR. Comparison of surveillance systems for monitoring COVID-19 in England: a retrospective observational study. Lancet Public Health 2023 Nov;8(11):e850-e858. [doi: 10.1016/S2468-2667(23)00219-0] [Medline: 37832574]
15. Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: clinical practice research datalink (CPRD). Int J Epidemiol 2015 Jun;44(3):827-836. [doi: 10.1093/ije/dyv098] [Medline: 26050254]
16. Bartholomeeusen S, Kim CY, Mertens R, Faes C, Buntinx F. The denominator in general practice, a new approach from the Intego database. Fam Pract 2005 Aug;22(4):442-447. [doi: 10.1093/fampra/cmi054] [Medline: 15964863]
17. Verheij RA, Curcin V, Delaney BC, McGilchrist MM. Possible sources of bias in primary care electronic health record data use and reuse. J Med Internet Res 2018 May 29;20(5):e185. [doi: 10.2196/jmir.9134] [Medline: 29844010]
18. Vanhommerig JW, Verheij RA, Hek K, Ramerman L, Hooiveld M, Veldhuijzen NJ, et al. Data Resource Profile: Nivel Primary Care Database (Nivel-PCD). Int J Epidemiol 2024. [doi: 10.1093/ije/dyaf017]
19. Duineveld B, Kole HM, Werven H, NHG-richtlijn. Adequate dossiervorming met het elektronisch patiëntendossier (ADEPD). Volledig gereviseerde versie 2019. [Adequate record keeping in electronic patient records (In Dutch)]. : Nederlands Huisartsen Genootschap; 2019 URL: https://www.nhg.org/wp-content/uploads/2022/11/NHG-Richtlijn-ADEPD-2019.pdf
20. Smeets HM, Kortekaas MF, Rutten FH, et al. Routine primary care data for scientific research, quality of care programs and educational purposes: the Julius General Practitioners' Network (JGPN). BMC Health Serv Res 2018 Sep 25;18(1):735. [doi: 10.1186/s12913-018-3528-5] [Medline: 30253760]
21. Sean E. The anatomical therapeutic chemical classification system. J Pharmacol Ther 2021;1(5):1 [FREE Full text]
22. Soler JK, Okkes I, Wood M, Lamberts H. The coming of age of ICPC: celebrating the 21st birthday of the International Classification of Primary Care. Fam Pract 2008 Aug;25(4):312-317 [FREE Full text] [doi: 10.1093/fampra/cmn028] [Medline: 18562335]
23. World Health Organization. WHO collaborating centre for drug statistics methodology. guidelines for ATC classification and DDD assignment. : Norwegian Institute of Public Health; 2024. URL: https://atcddd.fhi.no/filearchive/publications/1_2024_guidelines__final_web.pdf [accessed 2025-02-20]

XSL•FO
RenderX

24. Batenburg RS, Flinterman LE, Vis EB, et al. Cijfers uit de nivel- registratie van huisartsen en huisartsenpraktijken: een actualisering voor de periode 2020-2022. [figures from the nivel registration of general practitioners and general practices: an update for the period 2020-2022 (in dutch)]. : Nivel; 2022 URL: https://www.nivel.nl/sites/default/files/bestanden/1004340.pdf [accessed 2025-02-20]
25. Hooiveld M, Baliatsas C, Summeren J, van J, Stelma FF. Nivel-cijfers Ziekten per week in Nederland - actueel: wat heerst er? [Nivel Figures Weekly Diseases in the Netherlands (In Dutch)]. URL: https://www.nivel.nl/nl/zorg-en-ziekte-in-cijfers/actuele-cijfers-ziekten-per-week [accessed 2024-03-10]
26. van Duinkerken A, Bosmans M, Baliatsas C, et al. The integrated health monitor COVID-19: a protocol for a comprehensive assessment of the short- and long-term health impact of the pandemic in the Netherlands. Methods Protoc 2023 Dec 2;6(6):117. [doi: 10.3390/mps6060117] [Medline: 38133137]
27. Rahmon I, Bosmans M, Baliatsas C, Hooiveld M, Marra E, Dückers M. COVID-19 health impact: a use case for syndromic surveillance system monitoring based on primary care patient registries in the Netherlands. JMIR Public Health Surveill 2024 Sep 26;10:e53368. [doi: 10.2196/53368] [Medline: 39588660]
28. Baliatsas C, Hooiveld M, Kenens R, et al. Gezondheidsproblemen in gebieden met hoge plaagdruk van de eikenprocessierups: de grote overlast in 2019 [Health issues in areas heavily infested with the oak processionary: the significant nuisance in 2019] (In Dutch).: Nivel; 2022. URL: https://www.nivel.nl/nl/publicatie/gezondheidsproblemen-gebieden-met-hoge-plaagdruk-van-de-eikenprocessierups-de-grote [accessed 2025-02-20]
29. Hondema L, Hoek G, Kraaij-Dirkzwager M, Fischer P, Hooiveld M. Heat wave effects on morbidity in the Netherlands using general practice data. Environmental Epidemiology 2019 Oct;3(Supplement 1):161-162. [doi: 10.1097/01.EE9.0000607552.35945.d9]
30. de Lusignan S, Williams J. To monitor the COVID-19 pandemic we need better quality primary care data. BJGP Open 2020;4(2):bjgpopen20X101070. [doi: 10.3399/bjgpopen20X101070] [Medline: 32295793]
31. European Centre for Disease Prevention and Control (ECDC). Data quality monitoring and surveillance system evaluation: A handbook of methods and applications. : Stockholm: ECDC; 2014. URL: https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/Data-quality-monitoring-surveillance-system-evaluation-Sept-2014.pdf [accessed 2025-02-20] [doi: 10.2900/35329]
32. Smith S, Morbey R, de Lusignan S, Pebody RG, Smith GE, Elliot AJ. Investigating regional variation of respiratory infections in a general practice syndromic surveillance system. J Public Health (Oxf) 2021 Jun 7;43(2):e153-e160. [doi: 10.1093/pubmed/fdaa014] [Medline: 32009178]
33. Zucs P, Beauté J, Palm D, Spiteri G. Focus, vigilance, resilience: towards stronger infectious disease surveillance, threat detection and response in the EU/EEA. Euro Surveill 2024 Aug;29(34):2400066. [doi: 10.2807/1560-7917.ES.2024.29.34.2400066] [Medline: 39176987]
34. Chowdhury AT, Newaz M, Saha P, Pedersen S, Khan MS, Chowdhury MEH. Use of artificial intelligence in the surveillance of seasonal respiratory infections. In: Chowdhury MEH, Kiranyaz S, editors. Surveillance, Prevention, and Control of Infectious Diseases: Springer; 2024. [doi: 10.1007/978-3-031-59967-5_10]

Abbreviations

ADEPD: adequate record keeping in electronic patient records
ATC: anatomical therapeutic chemical
CBS: Statistics Netherlands
EHR: electronic health record
GP: general practitioner
ICPC: International Classification of Primary Care
NHG: Dutch College of General Practitioners
Nivel: Netherlands Institute for Health Services Research
Nivel-PCD: Netherlands Institute for Health Services Research-primary care database
RIVM: National Institute for Public Health and the Environment

Edited by A Mavragani; submitted 24.03.24; peer-reviewed by S Mukhida, S Jonathan; revised version received 29.01.25; accepted 29.01.25; published 12.03.25.

Please cite as:
Baliatsas C, van Summeren J, van Beusekom S, Matser A, Hooiveld M
Monitoring Public Health Through a Comprehensive Primary Care Database in the Netherlands: Overview of the Nivel Syndromic Surveillance System
JMIR Public Health Surveill 2025;11:e58767
URL: https://publichealth.jmir.org/2025/1/e58767
doi:10.2196/58767

XSL•FO
RenderX

© Christos Baliatsas, Jojanneke van Summeren, Sander van Beusekom, Amy Matser, Mariette Hooiveld. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Experiences and Perceptions About Death Reporting and Notification Among Rural Communities on the Islands of Lake Victoria, Uganda: Qualitative Study

Steven Ndugwa Kabwama[1,2,3], MSc; Douglas Bulafu[3,4], MSc; Rawlance Ndejjo[3,4,5], PhD; Rose Nampeera[6], MPH; Christine Kihembo[6], MPH; Chukwuma David Umeokonkwo[6], MPH; Simon Antara[6], MPH; Sheba Nakacubo Gitta[6,7], MPH; Caroline Kyozira[8], MPH; Allan Muruta[8], MPH; Jean-Edgard Nguessan[9], MPH; Rhoda K Wanyenze[3,4], PhD

[1]Department of Global Public Health, Karolinska Institutet, Tomtebodavägen 18A, Solna, Stockholm, Sweden

[2]Department of Community Health and Behavioral Sciences, School of Public Health, Makerere University, Kampala, Uganda

[3]Cluster of Research Excellence for Preparedness and Response to Pandemics and Shocks, Kampala, Uganda

[4]Department of Disease Control and Environmental Health, School of Public Health, Makerere University, Kamplala, Uganda

[5]Department of Preventive Medicine, College of Medicine, Korea University, Seoul, Republic of Korea

[6]Department of Programs, African Field Epidemiology Network, Kampala, Uganda

[7]Department of Epidemiology and Biostatistics, School of Public Health, Makerere University, Kampala, Uganda

[8]Department of Integrated Epidemiology, Surveillance and Public Health Emergencies, Ministry of Health, Kampala, Uganda

[9]Department of Noncommunicable Diseases, CDC Foundation, Atlanta, United States

Corresponding Author:
Steven Ndugwa Kabwama, MSc
Department of Global Public Health, Karolinska Institutet, Tomtebodavägen 18A, Solna, Stockholm, Sweden

Abstract

Background: Mortality data are critical for planning and prioritization of public health interventions and are generated through civil registration and vital statistics systems like mortality surveillance systems. However, frameworks for strengthening mortality surveillance systems do not acknowledge the cultural relativism surrounding death and how it influences strategies to improve mortality surveillance systems.

Objective: This paper aims to describe the experiences and perceptions about death reporting and notification among rural dwellers on the islands of Lake Victoria in Central Uganda.

Methods: The study was conducted in Buvuma and Kalangala Districts on Lake Victoria using a phenomenological qualitative research design. We conducted 12 in-depth interviews with community members who were purposively identified by village leaders and had experienced the death of a next of kin and reported and notified, and 8 in-depth interviews with those who had experienced the loss of a next of kin but did not notify and report the death. Key informant interviews were also conducted with 2 police officers and 2 cultural leaders. A total of 4 focus group discussions were conducted among village leaders. Interviews were abductively analyzed to generate grand narratives.

Results: The findings revealed 6 grand narratives of the perceptions and experiences of the process of death reporting and notification among the rural dwellers. These include (1) death reporting and notification are preceded by a tragic event that affects how, when, and if it is conducted; (2) a long and cumbersome process; (3) a process that involves multiple stakeholders with official and unofficial roles and responsibilities; (4) a process with little perceived individual or societal value; (5) a process with several mandatory but unofficial costs; and (6) a process preceded by events with deep cultural undertones.

Conclusions: Death reporting and notification are perceived to be tedious and cumbersome, which discourages community members from conducting them. There is a need to evaluate the process to remove any perceived or actual barriers through strategies such as decentralization of the process to lower levels of political administration. Death reporting and notification are also part of a broader social context that includes cultural beliefs, norms, and traditions. Efforts to strengthen mortality surveillance systems would profit from acknowledging the broader sociocultural issues around death and grieving and the role that cultural and religious institutions can contribute to addressing misconceptions and articulating the benefit of the process to society.

(JMIR Public Health Surveill 2025;11:e77135) doi:10.2196/77135

KEYWORDS

mortality; surveillance; death notification; death reporting; Uganda

XSL•FO
RenderX

Introduction

Mortality data are critical for planning, monitoring, and evaluating the implementation and impacts of national health policies and programs [1]. The data can be analyzed to identify changes in disease burden due to specific public health policies or interventions that can inform reprioritization and reallocation of resources. Mortality data are generated through national civil registration and vital statistics (CRVS) systems that provide information on death by sex, age, and cause of death. Ecological analyses have shown that good CRVS systems are associated with better population level health outcomes such as lower risk of child mortality, lower maternal mortality ratio, and better healthy life expectancy [2]. However, many low- and middle-income countries, particularly in Sub-Saharan Africa, lack comprehensive CRVS systems [3]. According to the 2022 United Nations Populations and Vital Statistics Report, many low- and middle-income countries, including Uganda, did not report any vital statistics such as live births and deaths in the 15 years of the reporting period between 2006 and 2020 [4]. Although some countries have made progress in improving maternal and perinatal death surveillance [5,6], challenges remain regarding the establishment of mortality surveillance systems that comprehensively record all deaths that occur outside of health facility settings. The 2016 Uganda Demographic Health Survey noted that only 24.2% of all deaths that occur in the country are reported, notified, and registered [7].

In 2023, the Africa Centres for Disease Control and Prevention (Africa CDC) published a continental framework for strengthening mortality surveillance systems in countries in Africa [8]. The framework contextualizes the mortality surveillance process by emphasizing the systematic collection of death-related data through leveraging community structures to implement an active case search approach and linking the information to policy and decision makers for public health action. It specifies the core functions of mortality surveillance systems at various levels of the health system, such as the detection of events and notification to national civil registration systems at both the health facility and community level. However, aside from the policy and infrastructural investments required to improve mortality surveillance systems, the framework does not acknowledge the cultural relativism [9], surrounding death or the wide societal and cultural variations in the interpretation, articulation, and meaning of death as a concept, and the bereavement process [10], and how these relate with, and influence strategies to improve mortality surveillance systems on the continent. From an anthropological standpoint, mortality is associated with grief, mourning, and bereavement, which are complex social phenomena shaped by deeply entrenched cultural norms, attitudes, and beliefs about death and the dead [11]. These norms, attitudes, and beliefs can influence how the process of death reporting and notification is perceived, interpreted, experienced, and conducted. Island communities often have unique cultural identities [12] and logistical challenges to implementing public health interventions [13] like surveillance. We conducted a study to describe what death reporting and notification mean for rural dwellers on the islands of Lake Victoria in Central Uganda to inform strategies for aligning and contextualizing interventions to improve mortality surveillance systems in Africa.

Methods

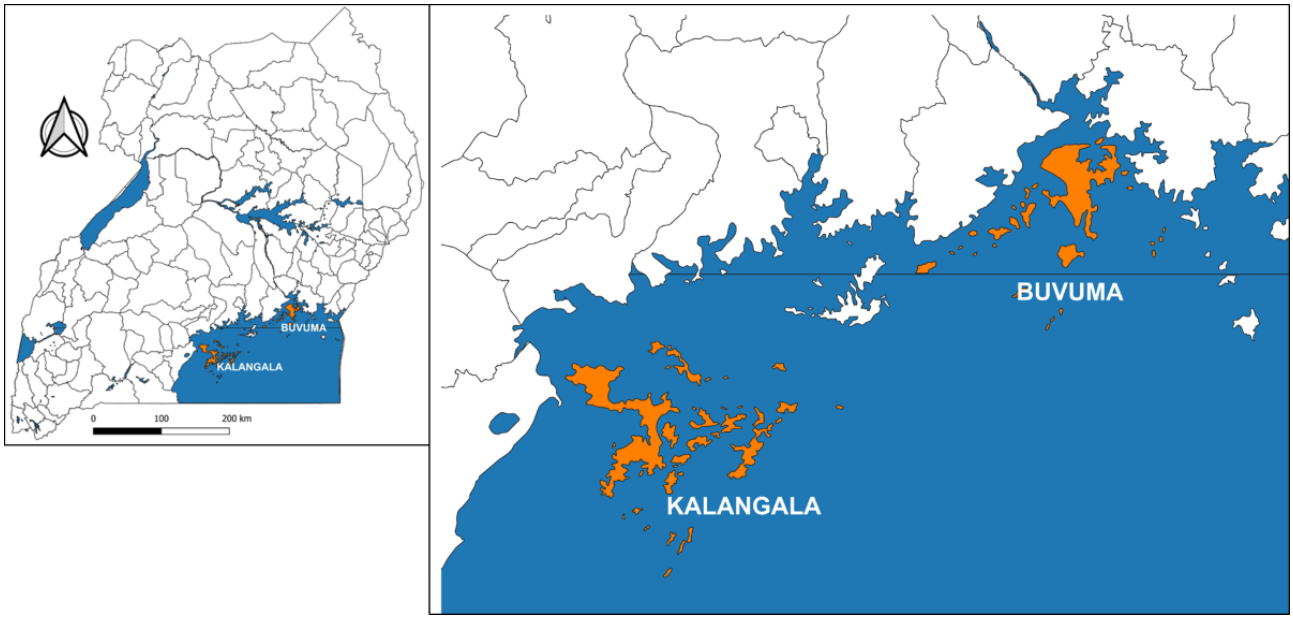
Study Area

The study was conducted in Buvuma and Kalangala Districts on Lake Victoria (see Figure 1). Buvuma District comprises 9 subcounties on 52 islands with a population of about 120,000 people, while Kalangala District consists of 7 subcounties on 84 islands and a population of about 70,000 people [14]. Buvuma and Kalangala Districts are among the least populated districts in Uganda and have been described as hard-to-reach [13] because of their geographical location on Lake Victoria, which presents unique challenges in terms of service delivery, including access to health services.

The islands are part of the Buganda Kingdom in Central Uganda. The kingdom is a constitutional monarchy composed of 52 clans and headed by a king (Kabaka), who is supported by a prime minister (Katikkiro), a legislative assembly, and leaders at the county, subcounty, and parish levels. The subjects of the kingdom are called Baganda and have cultural norms and traditions related to language, family, identity and genealogy, fashion and food, spirituality and beliefs, and ceremonies such as marriage, birth, and death.

XSL•FO
RenderX

Figure 1. Map showing Buvuma and Kalangala Districts on Lake Victoria, Central Uganda.



Study Design

The study used a phenomenological qualitative research design [15,16], which aimed to distill the perceptions and lived experiences of individuals in a rural community regarding the process of death reporting and notification. The choice of this study design was based on social constructivism as a knowledge claim [15], acknowledging the cultural relativism [9], and the various contexts where people live and work that determine and shape the experiences of the process of death reporting and notification.

Study Population

The study population was the inhabitants of Buvuma and Kalangala Districts. The study participants were community members who had experienced the death of a next of kin and were purposively selected by village leaders. The village leaders are part of the local administrative structure that oversees public health, education, and social welfare [17], and participate in all burial ceremonies in their respective villages.

Sampling and Sample Size

In-depth interviews were conducted with 20 community members who had experienced the death of a next of kin in the 12 months that preceded the date of the interview. These included 12 community members who had experienced the death of a next of kin and reported and notified the death as well as 8 community members who had lost a next of kin but did not notify and report the death. Community members who were sick, ages 18 years and younger, or could not provide consent for the interview were excluded from the study. Table 1 shows the characteristics of participants included in the in-depth interviews.

To enhance the trustworthiness of the study [18,19], focus group discussions (FGDs), and key informant interviews (KIIs) were also conducted. A total of 4 FGDs were conducted among local administrative leaders with each group comprising 8 - 10 participants. Key informant interviews were also held with 2 police officers and 2 cultural leaders at the subcounty level. The KIIs were equally distributed across the 2 districts, with at least one in each of the selected subcounties. The selection of the police officers and cultural leaders at the subcounty level was purposive as stakeholders involved in funeral ceremonies [20].

XSL•FO
RenderX

Table . Characteristics of the in-depth interview participants.

Characteristic	Reported and notified (n=12), n (%)	Did not report (n=8), n (%)
Sex		
Male	1 (8)	3 (38)
Female	11 (92)	5 (62)
Age categories (years)		
20 - 29	0 (0)	1 (13)
30 - 39	3 (25)	2 (25)
40 - 49	4 (33)	2 (25)
50 - 59	4 (33)	2 (25)
60+	1 (8)	1 (13)
Education level		
No formal education	4 (33)	3 (38)
Primary education	6 (50)	4 (50)
Secondary education	2 (17)	1 (12)
Tertiary education	0 (0)	0 (0)
Employment		
Peasant or farmer	12 (100)	8 (100)

Death Reporting and Notification in Uganda

The Uganda National Identification and Registration Authority (NIRA) published a death registration handbook [20], which outlines the steps of reporting, notifying, and registering a death. Briefly, when a death occurs in the community, the deceased's next of kin obtains a recommendation letter from the administrative leadership at the village level. The next of kin then takes this letter to the administrative leadership at the subcounty level for validation and issuance of a notification record. Thereafter, they take the notification record to the NIRA district office where the death is registered and subsequently a death certificate is issued.

If the death occurs in a health facility, the health worker fills out the death notification record and then a medical officer fills the medical certification of cause of death. A records officer enters this information into the national health information system which is linked to the national information and registration authority database. The next of kin then goes to the NIRA district office to obtain a death certificate.

For both processes, the death certificate issued by NIRA is free if it is requested within a year of the death of the next of kin. Beyond this, the certificate is issued after payment of 5000 UGX, equivalent to US $1.40.

Data Collection

The in-depth interview guides were developed following the constructs of the socioecological model [21]. The model acknowledges that actions such as engaging in notifying and reporting a death are influenced by factors that occur at the individual level (eg, knowledge about the process), at the interpersonal level (eg, a family member who knows the process), at the community level (eg, how cultures and traditions align with the requirement to notify and report death), and at the policy level (eg, the legal and policy environment for reporting and notification). In addition, reference was made to the Uganda NIRA Form 12 for community death notification and the HMIS Form 100 for health facility death notification to obtain information regarding the various steps and stakeholders involved in the reporting and notification process.

The guides for the key informant interviews, FGDs, and in-depth interviews were pretested through continuous role plays among the research assistants. This strategy has been used in previous studies and served to assess the clarity of the translations, skip patterns, and comprehensiveness of the guide to address all research objectives. The in-depth interviews, key informant interviews, and FGDs were conducted within the community at a place convenient for all interviewees. The key informant interviews and in-depth interviews lasted an average of about 40 minutes, while the FGDs lasted an average of about 60 minutes. The KIIs and FGDs were conducted by 6 female research assistants and 5 male research assistants who had at least Bachelor's Degree training and were fluent in Luganda—the locally spoken language in the region. In addition, interviewers received additional training on the principles of research ethics as well as qualitative data collection techniques such as probing and taking important field notes to supplement the analysis. SNK, DB, and Ndejjo R held regular meetings with the research assistants to discuss and address any challenges regarding participant recruitment and the emerging themes as well as the rate at which new information was being generated. The data collection process was halted after SNK, DB, and Ndejjo R discussed with the research assistants and concluded that saturation was reached.

XSL•FO
RenderX

Data Management and Analysis

All interviews and FGDs were audio recorded and transcribed verbatim. All transcripts were then entered into ATLAS.ti (version 8.0; Lumivero) and analyzed thematically. We used an abductive approach to data analysis, which involved both deductive and inductive processes. The first and second author gave an overview of the analytical approach to the team that conducted the analysis. A total of 8 people who were involved in the data collection supported the generation of codes and themes from the data. This allowed for researcher triangulation where the codes and themes were compared across the analysts to generate a narrative. Subsequent meetings were held to discuss the application of the framework and any improvements to improve the synthesis and generation of themes. Deductively, we generated themes following the Socioecological Model [21] as a guide, but with the flexibility to include themes and codes beyond the framework. We then merged these themes into several grand narratives [22] to construct the meaning of death reporting and notification based on the experiences and perceptions of the study participants. Grand narratives are broad descriptions of a phenomenon. In our case, we generated broad descriptions of what emerged as experiences and perceptions as told by our participants.

To enhance the completeness with which the methodological approach has been described, we have filled the Consolidated Criteria for Reporting Qualitative Studies checklist (see Checklist 1).

Ethical Considerations

The study received ethical approval from Makerere University School of Public Health Research and Ethics Committee (Approval SPH-2024 - 553). The study was also registered by the Uganda National Council for Science and Technology (HS3966ES). Permission to conduct the study in the selected districts was also obtained from Buvuma and Kalangala District Local Government authorities. Previous written informed consent was obtained from all study participants and participation was voluntary after establishing rapport and explaining the rationale for conducting the study. No individually identifiable information was collected, and each participant in the study was assigned a number to protect their identity. Each participant in the interviews was reimbursed with US $14, while those in the FGDs were reimbursed with US $6, to compensate their time for participating in the study. Only research assistants with a certificate in human subjects' protection or good clinical practice participated in the study.

Results

Participants and Analysis

The study involved 56 participants. These included 12 participants in the in-depth interviews, 4 participants in the key informant interviews, and 40 participants in the FGDs.

The analysis of results from the interviews and FGDs revealed 6 grand narratives of the perceptions and experiences of the process of death reporting and notification among the rural dwellers in our study.

Death Reporting and Notification Is Antedated by a Tragic Event That Affects How, When, and if it is Conducted

In-depth interview participants expressed that their experiences associated with death, loss, and the grief by individuals and families determine whether they conduct the process of death reporting and notification. Relatedly, the manner in which the death occurs also determines how it is reported and notified. For example, participants expressed feelings of shock and inertia when the death occurred suddenly. Initiating the process of reporting and notification was impeded in a context of sudden shock, grief, and mourning. Others noted that in circumstances where the death event is sudden and the cause is unclear, participants felt that initiating the process of death reporting and notification might imply guilt and culpability on the part of the notifier. In fact, there was a general lack of knowledge of who, how, and when to initiate the process of death reporting and notification.

> *We experienced the death of my son who was involved in an accident with his friends while riding a motorcycle… I didn't notify my son's death because I couldn't think straight at the scene of the accident. I simply informed the local leaders of the area. I also didn't know what was required to register a death* [In-depth interview, IDI 02, Did not report and notify]

Similar sentiments were expressed by participants who had experienced the death of a next of kin from a health facility:

> *[my sister] needed to return to the health facility to acquire a specific form required by NIRA for the certificate. However, due to this grief, the process has been slow. As a result, we have not yet finalized obtaining the death certificate…* [IDI 03, Reported but did not notify]

Experience: A Long and Cumbersome Process

Participants noted that they found the process of death reporting and notification to be very long and cumbersome. In addition to the logistical challenges of movement in the context of the area being an island, participants noted that there were many offices one needed to go to and demanded too much in terms of time and financial resources. For example, the process involves organizing and attending family meetings, getting letters from the administrative leadership at village level, the notification form from the health facility, which are then taken to the administrative offices at the subcounty level.

> *It's a long process, full of bureaucratic systems, at all levels, the hospital, the community, police letters, clan meetings, and NIRA also takes time. So, the process is too long... the officers are often too busy to attend to death registrations. They require lots of confirmations from family members which would be difficult to collect from all the family members at a single time.* [KII cultural leader 01]

Relatedly, participants noted that the death reporting and notification process required several documents and letters whose content and format were not clear. For example, participants noted that the process starts with a family meeting

XSL•FO
RenderX

whose minutes are required for the administrative leadership to provide a recommendation letter. However, little guidance is provided for who should be in the family meeting and what the minutes should describe. Furthermore, the village leader (lowest administrative level leadership) provides a recommendation letter but little guidance is provided in terms of the format or the content of the letter.

> *When I lost my brother last year, I had not heard anything about the death notification process. I didn't even know where to start. I believe it's important to complete the notification process because I've lost six family members, yet no one outside our immediate circle is aware of it.* [IDI 04, Did not report and notify]

A Process Involving Multiple Stakeholders With Official and Unofficial Roles and Responsibilities

Participants in the interviews described the death reporting and notification process as involving several stakeholders such as the family, relatives, friends, and clan members of the deceased, clan leaders and elders, administrative leadership at village and subcounty level, community health workers, and the police. The family and clan leaders have the responsibility of ensuring that cultural norms and values are followed in the funeral processes, the administrative leadership at the village level provides a letter of support that accompanies the death notification form, while community health workers provide technical support to community members going through this process. However, the role of the police and the administrative leadership at the subcounty level was unclear as noted by a respondent in the quote below:

> *The problem with police is that it deals with those who are able to give them money in exchange for a service! If you don't have money, they will ignore you!! So if this police we have were the conventional type where every person is listened to equally and fairly, collaborating with them on matters of death registration in the community would be a very good thing. Otherwise, as of now, we fear to interact with the police.* [KII cultural leader 02]

A Process With Little Perceived Individual or Societal Value

Generally, participants expressed little appreciation of the value of notifying and reporting a death. The general sentiment was that this was a process that was required if the deceased left an estate to be administered by a surviving next of kin. In circumstances of the death of a minor or where the deceased did not leave any estate, there was no motivation from the next of kin to report and notify the death. Participants felt that in addition to being cumbersome, there was no real direct benefit at both individual and societal levels.

> *I saw no value to register that death; my person died… there was nothing to gain whatsoever. Now, my person was dead and the sub-county is far away! Why should I incur costs on transport to register this death yet I have children to provide for sugar and school fees?.* [IDI 03, Did not report and notify]

Other participants perceived that death reporting and notification were only important in circumstances where the deceased had property because of the need to have an administrator who will take care of that property.

An FGD participant mentioned that:

> *People who register are the people who have money and hope to get other property but for those without property, after burial; it is over and even forget about the name. …not all who lose people come to register their people but those who have need or when the deceased had some property. In most cases, if they have not called us to join their meetings, that person comes with a family letter confirming that he/she was selected to work on the deceased's property.* [FGD among local leaders]

A Process With Several Mandatory but Unofficial Costs

Interview participants noted that the process of death reporting and notification involved several mandatory steps that required unofficial costs. For example, the process begins with a letter of support from the administrative leadership at village level. This letter is seldom given free of charge. The administrative leaders argue that the payments made are to support the procurement of stationery for issuing the letters. Similarly, the involvement of the police, particularly for deaths that involve accidents like drowning, will attract a cost on the part of the person trying to notify and report that death.

> *We lost our dear child who had been suffering from an illness. We needed a support letter from the chairman confirming that the child had died at home… The chairman provided the letter without any issues as we had already informed him about the death. We also obtained a support letter from the police detailing the cause of death, the date, and the place of residence. At the chairman's place, we paid US $1.4 for the letter. However, at the police station, they initially asked for US $5.5, which I couldn't afford. I told them I could only manage US $2.8, and they accepted that amount.* [IDI 06, Reported but did not notify]

Worse still, the communities in the 2 districts were generally poor, and any official or unofficial costs associated with death reporting and notification will only discourage anyone from following through with it. Death reporting and notification involves meeting different stakeholders including the village local council chairperson, the health facility, and NIRA officials, among others. Some of the relatives of the deceased were too poor to afford these costs associated with death reporting and notification.

Death Reporting and Notification Are Preceded by Events With Deep Cultural Undertones

Participants noted that death and burial of the deceased were cultural events with several rites and rituals involving members of the family, community, and clan as well as cultural leaders. They noted the importance of adhering to prescribed cultural

XSL•FO
RenderX

norms and practices when grieving and mourning the death of a loved one. Others were concerned that death reporting and notification was contrary to traditionally accepted cultural practices and might disturb the peace of the deceased, which discouraged some next of kin from reporting and notifying deaths.

> *As per the culture here, they will say; Leave the person to rest, leave the person to rest. Why do you disturb him?*. [KII Police Officer]

These sentiments were expressed by clan leaders and representatives of cultural leaders.

> *"Because people don't know as the Baganda say "Omuffu tebamubala" meaning "the dead cannot be counted". So, they take registration of deaths as "counting the dead" people will be like why should you register a dead person? ...why do you make a death notification…* [KII cultural leader 01]

Discussion

Principal Findings and Comparison With Previous Works

This study aimed to describe the meaning of death reporting and notification among rural dwellers on the islands of Lake Victoria in Central Uganda to inform strategies for aligning and contextualizing interventions to improve mortality surveillance systems in Africa. Our results revealed that death reporting and notification is preceded by a tragic event that affects how, when, and if it is conducted, a long and cumbersome process, which involves multiple stakeholders with official and unofficial roles and responsibilities, a process with little perceived value, involving several costs, and yet it is preceded by events with deep cultural undertones. The fact that participants felt that the death reporting and notification was difficult to initiate and conduct emerged as an experience among some groups and a perception among others. The challenge of complex systems and processes for reporting and notifying deaths has been reported in other countries like Mozambique [23] and Gambia [24] and has been exacerbated by poor logistical infrastructure coupled with little investment in technology, equipment, and human resources. The fact that this theme emerged from informants who had attempted to conduct the process, as well as those who only perceived it to be difficult, calls for a systematic approach to identify gaps and recommend adjustments for improvement. First, we propose a review of the entire process to clarify the contribution and role of each step, identify bottlenecks and opportunities for improved efficiency, and the subsequent removal of steps with unclear roles. In addition to addressing the practical or logistical difficulties of conducting the process, there is a need to improve how the process is appreciated and interpreted by the public. There are several theories in behavioral science and psychology that incorporate constructs of perception and motivation as indispensable precursors to action. For example, self-efficacy theory [25] posits that the belief that a task is difficult disincentivizes attempts to act. Perceptions and beliefs about the difficulty of a task are also critical in the psychological theory of cognitive dissonance [26], where the human mind engages in a struggle to align perceptions and beliefs with behavior. Psychological discomfort (dissonance) occurs where attempts are made to engage in an action that the mind perceives as difficult, resulting in inaction as the behavior to alleviate this discomfort. Relatedly, knowledge has been described as an indispensable precursor to perceptions [27] and beliefs. Interventions to improve awareness about and appreciation of the process of death reporting and notification in these communities should entail a clear articulation of its value to individuals, societies, and the health system. Such strategies could involve provision of specific examples of how mortality data are used to inform public health actions like identification of disease hot spots, planning, monitoring, and evaluation of interventions, policies, and health programs [28]. People are more likely to engage in tasks that they perceive to have high use value [29] for themselves, the community, or society.

Our study found that the process of death reporting and notification was described as one that involved multiple officials and stakeholders that were both official and unofficial. In addition, the process attracted several charges, most of which were unofficial. These 2 results corroborate our earlier findings that the process was perceived to be cumbersome. Aside from a review of the roles and relevance of the various steps, we propose further decentralization of the death reporting and notification process to the lowest levels of the political administrative structure. Here, a next of kin would report and notify a death at the village administrative officer or the parish administrative officer instead of the subcounty administrative officer. Although complex to implement [30], decentralization would alleviate the logistical challenges of movement that are particular to the communities in this context by reducing the administrative approvals one had to have before completing the process of death reporting and notification.

Participants also noted that death reporting and notification are preceded by a series of other norms and rituals with profound cultural significance. Efforts to improve mortality surveillance systems should avoid compartmentalizing the death reporting and notification process by ignoring the cultural significance of the events that precede it. Many cultures in Africa have norms and rituals that relate to recognizing the change of the body from physical to spirit form and ensuring that a good relationship remains between the spirit of the dead and the living [31]. If these processes are not properly conducted and followed, there are beliefs that the dead can have powers over and negatively impact the lives of the living [31-33]. In fact, the reservations about conducting the death reporting and notification process expressed by the participants in our study stemmed from concerns about offending the dead. Participants expressed sentiments about the cultural inappropriateness of talking about the dead or counting them, which discouraged some next of kin from conducting the process. Interventions and strategies to improve mortality surveillance systems on the continent should recognize the role of traditions and cultures and how these influence the understanding and appreciation of death reporting and notification. This can be operationalized through partnering with traditional and cultural institutions and leaders to promote death reporting and notification and its relevance to demystify the process and remove cultural barriers as one of the strategies

XSL•FO
RenderX

for strengthening mortality surveillance. Acknowledging cultural norms and values to advance public health goals has been successfully used to improve access to maternal health services in Indonesia [34], and scale up male circumcision in Lesotho and Swaziland [35]. Policies to improve mortality surveillance should be sensitive to cultural nuance by recognizing the culture, people, food, language, and other contextual factors within which the policy is developed [36].

The findings from this work should be interpreted while considering some limitations. First, we included participants who described their perceptions and experiences with death reporting and notification within the year that preceded the interview. This could introduce some information bias regarding the comprehensiveness with which events could be recalled. However, we recruited research assistants with the relevant experience in conducting qualitative interviews and used techniques such as probing to enhance the completeness with which events were recalled. Also, the results presented in this paper are synthesized from multiple sources which improves the credibility of the research. In addition, the authors from the Ministry of Health could have introduced reflexivity bias when describing the challenges with the process of reporting and notifying deaths. However, these authors were not involved in the data collection and analysis process. The other limitation from this study could be the overrepresentation of females among the in-depth interview participants. This could imply that the results regarding the experience of death reporting and notification lacked generalizability to both males and females. However, the rationale for selecting participants was solely on the experience of having lost a next of kin within the year that preceded the interview and the gender imbalance in the participants was not intentional from the researcher's perspective. Additional research would be required to assess whether there is nuance that was missed due to the gender imbalance in the in-depth interview participants.

Conclusions

Death reporting and notification is an initial step in efforts to collate health-related data that is used for planning and evaluating health programs. Our study in this rural setting on the islands of Lake Victoria has established that the process is perceived to be tedious and cumbersome, which discourages community members from reporting and notifying deaths. There is a need to reevaluate the process to remove any perceived or actual barriers through strategies such as decentralization of the entire process to lower administrative levels. Death reporting and notification are also seen as being part of a broader context that includes cultural beliefs, norms, and traditions. Efforts to strengthen reporting and notification should acknowledge that death is a sociocultural issue and actively involve cultural and traditional institutions. This will be particularly relevant in efforts to dispel myths and misconceptions about death reporting and notification as well as articulating the benefit of the process to society.

Acknowledgments

We are grateful to all individuals who participated in this study. In addition, we are grateful to the research assistants: Divine Katiiti, Deo Tomusange, Patience Oputan, Kenneth Sebukeera, Aminah Nambuusi, Alex Mulyowa, Nankya Hindu, Rogers Ssenyondo, Max Nalwoga, Junior Mike Wejuli, and Kisakye Shilla who supported the data collection and analysis. We also thank Mr. Steven Kasumba from the National Identification Authority and Dr. Carol Nanziri from the Uganda National Institute of Public Health who supported the implementation of the research project.

This work was supported by the Data for Health Initiative, a joint project of the CDC Foundation and Bloomberg Philanthropies through the African Field Epidemiology Network. The funders had no role in the design, implementation, or decision to publish the work. The views and opinions in this work are those of the authors and not official views of AFENET, CDC Foundation, Bloomberg Philanthropies, or the Uganda Ministry of Health.

Data Availability

All data generated or analyzed during this study are included in this published article and Multimedia Appendix 1.

Authors' Contributions

SNK, DB, Ndejjo R, and RKW contributed to funding acquisition, conceptualization of the idea, and design of research methodology. AM, J-EN, and RKW provided project supervision and technical oversight. SNK, Nampeera R, DB, and Kyozira C contributed to project management, planning, and execution. SNK wrote the original draft. Kihembo C, CDU, SA, SNG, and RKW were involved in technical review and editing, as well as validation and ensuring reproducibility of results. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Analyzed data.
[PDF File, 372 KB - publichealth_v11i1e77135_app1.pdf]

XSL•FO
RenderX

Checklist 1
COREQ (Consolidated Criteria for Reporting Qualitative Studies) checklist.
[PDF File, 174 KB - publichealth_v11i1e77135_app2.pdf]

References

1. Nyakeriga E, Waruiru W, Opollo V, et al. Mortuary and hospital-based HIV mortality surveillance among decedents in a low-resource setting: lessons from Western Kenya. BMC Public Health 2022 Mar 29;22(1):609. [doi: 10.1186/s12889-022-12909-3] [Medline: 35351057]
2. Phillips DE, AbouZahr C, Lopez AD, et al. Are well functioning civil registration and vital statistics systems associated with better health outcomes? The Lancet 2015 Oct;386(10001):1386-1394. [doi: 10.1016/S0140-6736(15)60172-6]
3. Ye Y, Wamukoya M, Ezeh A, Emina JBO, Sankoh O. Health and demographic surveillance systems: a step towards full civil registration and vital statistics system in sub-Sahara Africa? BMC Public Health 2012 Sep 5;12:1-11. [doi: 10.1186/1471-2458-12-741] [Medline: 22950896]
4. Population and vital statistics report. : United Nations Department of Economic and Social Affairs; 2022 URL: https://unstats.un.org/unsd/demographic-social/products/vitstats/sets/Series_A_2022.pdf [accessed 2025-07-23]
5. Whiting-Collins L, Serbanescu F, Moller AB, et al. Maternal death surveillance and response system reports from 32 low-middle income countries, 2011-2020: What can we learn from the reports? PLOS Glob Public Health 2024;4(3):e0002153. [doi: 10.1371/journal.pgph.0002153] [Medline: 38442110]
6. Zemen E, Yimer YS, Deyessa N, Abebe Y. Evaluation of surveillance system and response for maternal and perinatal mortalities in Meda Welabu District, Ethiopia, 2023. Pan Afr Med J 2024;47:133. [doi: 10.11604/pamj.2024.47.133.42963] [Medline: 38881771]
7. Uganda Bureau of Statatistics. The uganda demographic and health survey 2016. : ICF International Inc; 2017 URL: https://dhsprogram.com/pubs/pdf/FR333/FR333.pdf [accessed 2025-07-29]
8. Continental framework strengthening mortality surveillance systems in africa. : Africa Centers for Disease Control and Prevention; 2023 URL: https://africacdc.org/download/continental-framework-on-strengthening-mortality-surveillance-systems-in-africa/ [accessed 2025-07-23]
9. Tilley JJ. Cultural Relativism. hrq 2000 May;22(2):501-547. [doi: 10.1353/hrq.2000.0027]
10. Baloyi L, Makobe-Rabothata M. The african conception of death: a cultural implication. 2014 Presented at: International Association of Cross Cultural Psychology Congress; Jul 15-19, 2014; Reims, France URL: https://scholarworks.gvsu.edu/iaccp_proceedings/1 [doi: 10.4087/FRDW2511]
11. Silverman GS, Baroiller A, Hemer SR. Culture and grief: Ethnographic perspectives on ritual, relationships and remembering. Death Stud 2021;45(1):1-8. [doi: 10.1080/07481187.2020.1851885] [Medline: 33272138]
12. Grydehøj A, Nadarajah Y, Markussen U. Islands of indigeneity: Cultural distinction, indigenous territory and island spatiality. Area 2020 Mar;52(1):14-22. [doi: 10.1111/area.12520]
13. Ssetaala A, Ssempiira J, Wambuzi M, et al. Improving access to maternal health services among rural hard-to-reach fishing communities in Uganda, the role of community health workers. Womens Health (Lond) 2022;18:17455057221103993. [doi: 10.1177/17455057221103993] [Medline: 35748585]
14. The Uganda national population and housing census 2024 final report. : Uganda Bureau of Statistics; 2024 URL: https://www.ubos.org/wp-content/uploads/2024/12/National-Population-and-Housing-Census-2024-Final-Report-Volume-1-Main.pdf [accessed 2025-07-29]
15. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing among Five Approaches: Sage publications; 2016.
16. Greening N. Phenomenological research methodology. Scientific Research Journal 2019;VII(V):88-92. [doi: 10.31364/SCIRJ/v7.i5.2019.P0519656]
17. Johnson ML. The role of decentralization in local government empowerment and effective service delivery: a case study in Uganda. IJFMR 2023;5(6). [doi: 10.36948/ijfmr.2023.v05i06.8863]
18. Elo S, Kääriäinen M, Kanste O, Pölkki T, Utriainen K, Kyngäs H. Qualitative content analysis: a focus on trustworthiness. SAGE Open 2014;4(1):2158244014522633. [doi: 10.1177/2158244014522633]
19. Ahmed SK. The pillars of trustworthiness in qualitative research. Journal of Medicine, Surgery, and Public Health 2024 Apr;2:100051. [doi: 10.1016/j.glmedi.2024.100051]
20. Death registration handbook. Uganda National Identification and Registration Authority. 2022. URL: https://www.nira.go.ug/media/2022/11/DEATH-REGISTRATION-HANDBOOK-final.pdf [accessed 2025-07-23]
21. Kilanowski JF. Breadth of the socio-ecological model. J Agromedicine 2017;22(4):295-297. [doi: 10.1080/1059924X.2017.1358971] [Medline: 28742433]
22. Bernstein JM. Grand narratives. In: On Paul Ricoeur: Routledge; 2002:116-137 URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9780203416815-10/grand-narratives-bernstein [accessed 2025-08-07]
23. Gudo ES. Mozambique sample vital statistics system: filling the gaps for mortality data. Am J Trop Med Hyg 2023 May 2;108(5_Suppl):3-4. [doi: 10.4269/ajtmh.23-0094] [Medline: 37037427]

XSL•FO
RenderX

24. Wutor BM, Osei I, Sarwar G, et al. Under-5 mortality surveillance in low-income and middle-income countries: insights from two health and demographic surveillance systems in rural Gambia. BMJ Glob Health 2024 Apr 2;9(4):e014937. [doi: 10.1136/bmjgh-2023-014937] [Medline: 38569661]
25. Maddux JE. Self-Efficacy Theory: An Introduction Self-Efficacy, Adaptation, and Adjustment: Theory, Research: Springer; 1995:3-33. [doi: 10.1007/978-1-4419-6868-5]
26. Aronson E. The Theory of Cognitive Dissonance: A Current Perspective Advances in Experimental Social Psychology: Elsevier; 1969:1-34 URL: https://www.sciencedirect.com/science/article/abs/pii/S0065260108600751 [accessed 2025-08-07]
27. Ndugwa Kabwama S, Berg-Beckhoff G. The association between HIV/AIDS-related knowledge and perception of risk for infection: a systematic review. Perspect Public Health 2015 Nov;135(6):299-308. [doi: 10.1177/1757913915595831]
28. Groenewald P, Azevedo V, Daniels J, et al. The importance of identified cause-of-death information being available for public health surveillance, actions and research. S Afr Med J 2015 Sep 21;105(7):528-530. [doi: 10.7196/SAMJnew.8019] [Medline: 26428743]
29. Wigfield A, Tonks S, Klauda SL. Expectancy-value theory. In: Handbook of Motivation at School: Routledge; 2009:69-90 URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9781315773384-5/expectancy-value-theory-allan-wigfield-stephen-tonks-susan-lutz-klauda [accessed 2025-08-07]
30. Liwanag HJ, Wyss K. What conditions enable decentralization to improve the health system? Qualitative analysis of perspectives on decision space after 25 years of devolution in the Philippines. PLoS ONE 2018;13(11):e0206809. [doi: 10.1371/journal.pone.0206809] [Medline: 30395625]
31. Ekore RI, Lanre-Abass B. African cultural concept of death and the idea of advance care directives. Indian J Palliat Care 2016;22(4):369-372. [doi: 10.4103/0973-1075.191741] [Medline: 27803556]
32. Isidienu IC, Onyekelu AC. Ancestral cults in African Traditional Religion: their relevance in the contemporary African society. JASSD 2021;4(4) [FREE Full text]
33. Sekagya YHK, Muchunguzi C, Unnikrishnan P, Mulogo EM. A grounded theory study of beliefs underlying use of ancestral spirits for healing among baganda traditional spiritual healers in central uganda. Primary Care Research. p. 03. [doi: 10.1101/2024.04.03.24304941]
34. Titaley CR, Hunter CL, Dibley MJ, Heywood P. Why do some women still prefer traditional birth attendants and home delivery?: a qualitative study on delivery care services in West Java Province, Indonesia. BMC Pregnancy Childbirth 2010 Aug 11;10:1-14. [doi: 10.1186/1471-2393-10-43] [Medline: 20701762]
35. Bulled N, Green EC. Making voluntary medical male circumcision a viable HIV prevention strategy in high prevalence countries by engaging the traditional sector. Crit Public Health 2016 May 1;26(3):258-268. [doi: 10.1080/09581596.2015.1055319] [Medline: 27110065]
36. Resnicow K, Baranowski T, Ahluwalia JS, Braithwaite RL. Cultural sensitivity in public health: defined and demystified. Ethn Dis 1999;9(1):10-21. [Medline: 10355471]

Abbreviations

Africa CDC: Africa Centers for Disease Control and Prevention
CRVS: civil registration and vital statistics
FGD: focus group discussion
KII: key informant interview
NIRA: Uganda National Identification and Registration Authority

Edited by A Mavragani, T Sanchez; submitted 08.05.25; peer-reviewed by C Obianyo, S Sahai; revised version received 04.07.25; accepted 06.07.25; published 13.08.25.

Please cite as:
Kabwama SN, Bulafu D, Ndejjo R, Nampeera R, Kihembo C, Umeokonkwo CD, Antara S, Gitta SN, Kyozira C, Muruta A, Nguessan JE, Wanyenze RK
Experiences and Perceptions About Death Reporting and Notification Among Rural Communities on the Islands of Lake Victoria, Uganda: Qualitative Study
JMIR Public Health Surveill 2025;11:e77135
URL: https://publichealth.jmir.org/2025/1/e77135
doi:10.2196/77135

© Steven Ndugwa Kabwama, Douglas Bulafu, Rawlance Ndejjo, Rose Nampeera, Christine Kihembo, Chukwuma David Umeokonkwo, Simon Antara, Sheba Nakacubo Gitta, Caroline Kyozira, Allan Muruta, Jean-Edgard Nguessan, Rhoda K Wanyenze. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.8.2025. This is an open-access

XSL•FO
RenderX

article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of Virus Watch COVID-19 Positivity, Incidence, and Hospitalization Rates With Other Surveillance Systems: Surveillance Study

Wing Lam Erica Fong[1], MSc; Vincent Grigori Nguyen[1,2,3], PhD; Sarah Beale[1], PhD; Thomas E Byrne[1], PhD; Cyril Geismar[4], PhD; Ellen Fragaszy[1,5], PhD; Jana Kovar[1], PhD; Alexei Yavlinsky[1], PhD; Ibrahim Abubakar[6], PhD; Andrew C Hayward[2,7], MD; Robert W Aldridge[1,8], PhD

[1]Institute of Health Informatics, University College London, 222 Euston Road, London, United Kingdom

[2]Institute of Epidemiology and Health Care, University College London, London, United Kingdom

[3]Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, United Kingdom

[4]NIHR Health Protection Research Unit in Modelling and Health Economics, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom

[5]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, United Kingdom

[6]Faculty of Population Health Sciences, University College London, London, United Kingdom

[7]NIHR Health Protection Research Unit in Health Analytics and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom

[8]Institute for Health Metrics and Evaluation, University of Washington, Seattle, United States

Corresponding Author:
Wing Lam Erica Fong, MSc
Institute of Health Informatics, University College London, 222 Euston Road, London, United Kingdom

Related Article:

This is a corrected version. See correction statement: https://publichealth.jmir.org/2025/1/e85512

Abstract

Background: Effective disease surveillance is essential for understanding pathogens' epidemiology, detecting outbreaks, and enabling timely public health responses. In the United Kingdom, large-scale studies, such as the Office for National Statistics COVID-19 Infection Survey (CIS), have monitored SARS-CoV-2 transmission but required significant resources, making them challenging to sustain when pandemic-specific funding ends and also in resource-limited settings. In contrast, the Virus Watch study, at lower cost, relied on self-reported and linked national testing data as well as symptomatic testing, while Severe Acute Respiratory Infections Watch (SARI) leveraged hospital data for cost-effective surveillance.

Objective: This study aimed to evaluate the effectiveness of Virus Watch as a surveillance system in monitoring COVID-19 positivity, incidence, and hospitalization rates in England and Wales, using data from the CIS and SARI as benchmarks for comparison, while considering the key differences in the study designs, including recruitment strategies, incentives, and testing criteria.

Methods: We used the Virus Watch prospective community cohort study to estimate COVID-19 positivity, incidence, and hospitalization rates in England and Wales from June 2020 to March 2023. Rate estimates were compared with CIS modeled positivity and incidence rates, and with SARI COVID-19 hospitalization rates. Global synchrony between datasets was measured using overall Spearman and local synchrony using 9-week rolling Spearman . For England, comparisons with CIS estimates used Virus Watch rates calculated with and without linked national testing data. Positivity rates were also assessed overall and separately before and after the end of free national testing.

Results: A total of 58,628 participants were recruited into the Virus Watch study, of whom 52,526 (89.6%) were resident in England and 1532 (2.6%) in Wales; region was missing for the remainder. Virus Watch–estimated COVID-19 positivity and incidence rates in England, calculated with and without linked testing data, showed strong global synchrony with CIS estimates (positivity : 0.91 and 0.90; both $P<.001$ and incidence : 0.92 and 0.90; both $P<.001$) and strong local synchrony (positivity : median 0.75, IQR 0.53 - 0.85 and median 0.67, IQR 0.47 - 0.83, and incidence : median 0.76, IQR 0.49 - 0.88 and median 0.66, IQR 0.45 - 0.82), despite having lower absolute values. Global and local synchrony of positivity rates were similar for periods before and after the end of free national testing, although the difference between Virus Watch and CIS estimates was greater

XSL•FO
RenderX

post–free testing. COVID-19 hospitalization rates were also lower and less synchronized with SARI estimates. In Wales, Virus Watch estimates exhibited greater variability (positivity ρ: 0.75, *P*<.001; incidence rate ρ: 0.85, *P*<.001) and lower local synchrony (positivity ρ: median 0.61, IQR 0.34 - 0.74, and incidence ρ: median 0.52, IQR 0.38 - 0.71) compared to England.

Conclusions: Our results highlight the effectiveness of the Virus Watch approach in providing accurate estimates of COVID-19 positivity and incidence rates, even in the absence of national surveillance systems. This low-cost method can be adapted to various settings, particularly low-resource ones, to strengthen public health surveillance and inform timely interventions.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2020-048042

(JMIR Public Health Surveill 2025;11:e69655) doi:10.2196/69655

KEYWORDS

surveillance; COVID-19; SARS-CoV-2; public health monitoring; study comparison

Introduction

Surveillance systems are essential for epidemic and pandemic response, as they are required for detecting outbreaks and establishing fundamental epidemiological parameters to inform and evaluate effective and efficient public health interventions [1]. Comparing and learning from different surveillance strategies used during the COVID-19 pandemic provides important insights to prepare for future pandemics. In the United Kingdom, large-scale epidemiological studies, such as the Office for National Statistics (ONS) COVID-19 infection survey (CIS) and the Real-Time Assessment of Community Transmission-1 (REACT-1) study, along with sentinel surveillance systems such as Severe Acute Respiratory Infections Watch (SARI), played a crucial role in monitoring SARS-CoV-2 transmission [2,3].

The ONS CIS, considered a "best standard" surveillance study and used as a benchmark in this analysis, involved a substantial investment of £988.5 million (US ~$1.3 billion) since study inception from April 2020 to March 2023 [4]. The study enrolled a representative sample of children and adults (aged 2 years and older) in private residential households (excluding care homes and other communal establishment settings) [2]. They performed weekly nose and throat swabs to test for SARS-CoV-2, irrespective of symptoms. Initially, approximately 150,000 polymerase chain reaction (PCR) tests were conducted monthly in September 2020. By May 2021, the study expanded its testing capacity to analyze an average of 390,300 PCR tests monthly, continuing through March 2022 [2]. To encourage the return of swabs, the ONS CIS offered vouchers as a financial incentive. Up to November 2021, nearly 7 million e-vouchers, which can be spent at a range of retailers online and in-store, worth £211.5 million (US ~$285 million USD) were issued [5]. Despite its valuable insights, ONS CIS is resource-intensive, costly, and requires sufficient testing coverage and capacity, which may be challenging to sustain in many settings, especially where health care resources are limited [6,7].

The Virus Watch study, a community household study conducted in England and Wales, was less resource-intensive, with a total expenditure of £4.89 million (US ~$6.6 million USD). Its design was based on an earlier study tracking influenza transmission, Flu Watch, and it relied on self-reported data and linkage to national COVID-19 testing and hospital admission databases [8,9]. SARS-CoV-2 test results primarily came from free PCR or lateral flow tests (LFTs) made widely available through the national "Test Trace and Isolate Program" launched in May 2020. The program recommended testing for individuals experiencing a high temperature, a new, continuous cough, or a loss or change in their sense of smell or taste or if they have had close contact with someone who has COVID-19, meaning that most testing was symptomatic [10]. No incentives were provided for nasal swab testing. However, reliance on self-reported data and symptomatic testing might limit the accuracy and completeness of the data [9].

In contrast to ONS CIS and Virus Watch, the SARI surveillance system was established in 2020 to monitor hospitalizations and critical care admissions for COVID-19 and other respiratory illnesses (eg, influenza and respiratory syncytial virus [RSV]) across all acute National Health Service (NHS) trusts (hospital trusts that provide secondary health care services) in England [11]. Unlike survey-based systems, SARI is inexpensive and sustainable as it uses existing publicly funded health care infrastructure to estimate the COVID-19 disease burden without extensive community testing [1]. While it only captures data on severe cases of respiratory infections, the data collected is likely to be comprehensive and consistent, offering an overview of epidemiological trends, as around 75% of all acute trusts report hospital and critical care admissions for COVID-19 and other respiratory diseases [12]. However, evaluating only the most severe component of the "clinical iceberg'" of infection may introduce well-recognized bias into key epidemiological parameters, such as potentially overestimating the case fatality rate and underestimating the infection fatality rate, which are critical for informing public health responses [13].

To prepare for future pandemics, it is essential to explore efficient, scalable, and cost-effective methods for collecting and analyzing epidemiological data for acute respiratory disease surveillance that can be implemented in diverse settings. In this analysis, we aimed to evaluate the effectiveness of Virus Watch as a surveillance system in monitoring COVID-19 positivity, incidence, and hospitalization rates in England and Wales. We used data from the CIS and SARI as a benchmark for comparison, while also considering the key differences in the study designs, including recruitment strategies, incentives, and testing criteria.

XSL•FO
RenderX

Methods

Study Design and Participants

Virus Watch was a prospective cohort study that was launched on June 24, 2020, and ran until April 01, 2025. The study recruited 28,527 whole households and 58,628 participants aged 0 to 98 years (with a mean age of 48) up to March 2022 [8]. To participate, households were required to have access to a mobile phone, tablet, or a computer with an internet connection, a valid email address, and at least one household member who could read and respond in English to complete regular surveys. Whole households were required to provide informed consent or, where applicable, assent to participate. Eligible household sizes ranged from one to a maximum of 6 members due to limitations of the REDCap (Research Electronic Data Capture) survey platform (Vanderbilt University) used for data collection. Individuals living in institutional settings, such as care homes and university halls of residence, were not eligible to participate. The Virus Watch cohort was also linked to national COVID-19 testing and vaccination records, and hospital admission and mortality records covering all causes. A detailed description of the study design and methodology can be found in previous publications, with key elements relevant to this analysis outlined here [8,14].

Identifying SARS-CoV-2 Infections and COVID-19-Related Hospitalizations

Multiple sources were used to identify SARS-CoV-2 infections among Virus Watch participants. These included:

- Linked national testing data: National testing data is obtained from the COVID-19 Second Generation Surveillance system, which contains information on swab tests for SARS-CoV-2, from Pillar 1 (hospital testing) and Pillar 2 (community testing). NHS England linked SARS-CoV-2 test results from both Pillar 1 and 2 to participant data in March 2021. Hospital testing data encompassed results from March 2020 to July 2023, and community testing data from June 2020 to August 2023. The linked data was only available for participants with English postcodes.
- Self-reported test results: Participants reported positive PCR or LFTs in weekly illness surveys. Tests were obtained outside of the study, such as through the UK Test-Trace-Isolate program or privately.
- Virus Watch PCR swab testing: A total of 2 subsets of participants were provided nasopharyngeal swabs for PCR assays for SARS-CoV-2. They self-administered PCR swabs if they experienced symptoms including fever, cough, or loss of taste or smell between October 2020 and May 2021 (n=12,877) and between January and June 2023 (n=2851). All swabs were tested for SARS-CoV-2 RNA via reverse transcriptase–polymerase chain reaction (RT-PCR).

To avoid double-counting tests from the same infection episode, we defined each episode as beginning on the date of the first positive result from any of the data sources. Each episode was assumed to last 90 days from the start date; any subsequent positive tests within the 90-day window were considered part of the same episode. A new episode was only recorded if a positive test occurred more than 90 days after the start of the previous episode. This approach aligns with the UK Health Security Agency definition [15].

We used linked Hospital Episode Statistics Admitted Patient Care data to identify any hospitalizations for or with COVID-19 between March 2020 and March 2024. Hospitalization records provide detailed information on all admissions to NHS hospitals in England, as well as admissions to private or charitable hospitals funded by the NHS. For this analysis, only completed consultations were included. COVID-19-related hospital admissions were identified using the *ICD-10* code U07.1, which indicates confirmed COVID-19. Any admission with U07.1 recorded as a diagnostic code, regardless of whether it was listed as the primary or a secondary diagnosis, was considered a hospitalization for or with COVID-19. Consecutive admissions with a discharge and readmission occurring within a single day were treated as a single admission.

ONS CIS Modelled Positivity and Incidence Rates

CIS COVID-19 positivity estimates for England were available from April 27, 2020, to January 31, 2023, and incidence rate estimates were available for June 8, 2020, to June 15, 2022. For Wales, positivity estimates were reported from July 27, 2020, to January 31, 2023, and incidence estimates from October 25, 2020 to June 15, 2022 [16].

SARI Surveillance System Hospitalization Rates

The SARI surveillance system from the UK Health Security Agency provided the overall hospital admission rate of patients with confirmed COVID-19 in England [17]. Hospitalization rates data were available from August 9, 2020, to March 19, 2023.

Statistical Analysis

Virus Watch COVID-19 Positivity, Incidence, and Hospitalization Rates

We calculated 2 weekly positivity and incidence rates: one including linked hospital and community testing data and another based only on self-reported data. Positivity and incidence rates are rates of participants aged 2 years and above, which is in line with ONS CIS study design.

Weekly COVID-19 positivity rates incorporating linked national testing data were calculated as the proportion of participants who tested positive for SARS-CoV-2, identified through either Virus Watch self-reported data or linked national testing data, in a given week, divided by the total number of participants who either completed the weekly Virus Watch survey or had any test result (positive or negative) recorded in linked testing data during that same week. Positivity rates based only on self-reported data were calculated as the proportion of participants who self-reported testing positive in a given week, divided by the total number of participants who responded to the weekly survey during the same week.

The 7-day average incidence rate per 10,000 people, based on self-reported and linked testing data, was calculated by identifying the weekly number of participants who tested positive for SARS-CoV-2 for the first time in ≥90 days. This

XSL•FO
RenderX

weekly total, the numerator, was then divided by the weekly number of participants who had either reported no illness in the weekly survey or had only negative SARS-CoV-2 test results recorded in linked testing data in the preceding 90 days. Similarly, the incidence rate based solely on self-reported data were calculated using positive tests reported through weekly surveys. The numerator included participants who self-reported a positive test, having not reported a previous positive in the previous 90 days. The denominator consisted of participants who completed the weekly survey and had either not reported any illness or had not self-reported a positive test in the preceding 90 days. Both incidence rates were divided by 7 to obtain a 7-day average and scaled per 10,000 individuals.

Weekly COVID-19 hospitalization rates per 100,000 people were calculated using the subset of participants with any linked data (hospital records, COVID-19 testing, vaccination, and mortality data). Participants without linked data were excluded from this hospital rate estimation, as their hospitalization status during follow-up could not be determined; their inclusion would risk underestimating rates by inflating the denominator. Hospitalization rates were also restricted to participants in England, as hospital data were not available for Welsh residents. Rates were calculated by identifying the weekly total number of participants hospitalized for or with COVID-19. The denominator included all eligible participants who were alive and under follow-up during each week of the analysis period. Participants entered the denominator from the date of study registration and remained each week until the end of the analysis period or death.

ONS CIS Positivity and Incidence Rates

National positivity rates were generated using Bayesian multilevel regression poststratification (MRP) modeling, which were adjusted for potential biases from differences in participation across age, sex, and region. This approach improves the accuracy of positivity estimates by providing a more precise estimate of the population at-risk used as the denominator [18,19]. Nonlinear functions were used to account for time and regional interactions. To estimate COVID-19 incidence rate (per 100,000 people per week), the study first estimated clearance time estimates (the time period someone remains positive). The posterior samples from the MRP models were then combined with the estimated distribution of clearance times, generating incidence estimates of COVID-19 incidence (more details described here [18]).

SARI Hospitalization Rates

SARI weekly hospitalization rates are based on the trust catchment population of NHS trusts that actively reported new COVID-19, influenza, or RSV cases for that week [17]. Each week, counts of new COVID-19 admissions to hospitals for all levels of care were summed and converted to rates per 100,000 people based on catchment populations for the trusts for those participating in each reporting week [17].

Assessing Global and Local Synchrony Between Datasets

We assessed the agreement in trends between Virus Watch estimated and national surveillance data (CIS and SARI estimates) by calculating both global and local synchrony using Spearman correlation separately for England and Wales. Spearman correlation was used due to the nonparametric nature of the rate estimates. Global synchrony was defined as the overall Spearman across the full comparison period, and local synchrony as the median (IQR) of 9-week (roughly 60 d) rolling correlations [20]. Lagged correlations were not tested in this analysis, as the date of SARS-CoV-2 testing was available from both self-reported and linked national testing data. Using the test date as the proxy for infection onset in all cases minimized the potential for temporal misalignment between datasets.

For England, we compared Virus Watch positivity estimates calculated with linked national testing data to CIS positivity estimates, and we also compared Virus Watch positivity estimates calculated from only self-reported testing data to CIS positivity estimates. These comparisons with CIS estimates were repeated separately for the periods before the end of free national testing (June 22, 2020 - March 31, 2022) and after the end of free national testing (April 1, 2022 - January 31, 2023), for both linked and unlinked Virus Watch positivity estimates. In addition, we compared Virus Watch incidence rates calculated with linked national testing data to CIS incidence estimates, and Virus Watch incidence rates based on self-reported testing data only to CIS incidence estimates. We did not conduct comparisons with CIS estimates pre- and post-testing withdrawal for incidence rates, as the post-testing period was too short (April 1, 2022 - June 20, 2022) to calculate meaningful rolling correlations. Finally, we compared Virus Watch–estimated COVID-19 hospitalization rates with SARI estimates between August 9, 2020, and March 19, 2023.

For Wales, only self-reported testing data were available; therefore, we compared self-reported positivity and incidence rates with CIS estimates. We repeated positivity comparisons with CIS estimates before (June 22, 2020 - March 31, 2022) and after the end of free national testing (April 1, 2022 - January 31, 2023). No pre- and post-testing withdrawal comparisons with CIS estimates were conducted. Furthermore, linked hospitalization data were not available for Wales; therefore, no comparisons were made with SARI.

Weekly survey response data were extracted from REDCap, linked, and analyzed in R (version 4.4.1; R Foundation for Statistical Computing) and RStudio (version 2025.05.1+513; Posit).

Ethical Considerations

Virus Watch was approved by the Hampstead NHS Health Research Authority Ethics Committee (20/HRA/2320) and conformed to the ethical standards set out in the Declaration of Helsinki. Participants provided informed consent and, where relevant, for children that they were responsible for, for all aspects of the study. The survey dataset was deidentified by removing participants' names, dates of birth, and residential addresses. These were stored in another dataset with matched study IDs to preserve anonymity. All datasets were stored in a secure remote server and only accessed by designated research staff and investigators.

XSL•FO
RenderX

Results

Overview

The Virus Watch study recruited 58,628 participants between June 2020 and February 2023, primarily residents of England (52,529/58,628, 89.6%) with a smaller proportion (1521/58,628, 2.6%) from Wales. Compared to the general population, the study cohort was older, with a higher representation of individuals aged 45‐64 years and 65 years and older. While the cohort was generally balanced in terms of sex, there were notable disparities in regional and ethnic representation. Certain regions, such as the North West, South West, West Midlands, and Yorkshire and the Humber, were underrepresented, while others, like the East of England and South East, were overrepresented. Ethnic minority groups, including Black and other Asian populations, were also underrepresented in the study (see Table 1).

XSL•FO
RenderX

Table . Demographic characteristics of Virus Watch study participants at baseline recruitment compared to the Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) and 2021 census.

Characteristic	Virus Watch participants (n=58,628)	ONS CIS (n=535,121) [2]	Percentage of England population (ONS census), % [21,22]
Age group (y), n (%)			
2 - 11	4617 (7.9)	33,507 (6.3)	13
12 - 16	2623 (4.5)	29,147 (5.4)	6
17 - 24	3000 (5.1)	35,403 (6.6)	9
25 - 34	4973 (8.5)	52,359 (9.8)	14
35 - 49	10,406 (17.7)	103,068 (19.3)	20
50 - 69	22,789 (38.9)	176,814 (33.0)	25
70+	9589 (16.4)	104,823 (19.6)	14
Sex (including derived)[a], n (%)			
Male	26,274 (44.8)	252,047 (47.1)	49
Female	31,533 (53.8)	283,074 (52.9)	51
Other or Missing or Prefer not to say	821 (1.4)	—[b]	—
Ethnicity, n (%)			
White	43,968 (75.0)	493,026 (92.1)	86
Asian/ Asian British	3155 (5.4)	22,128 (4.1)	7
Black, African, Caribbean, or Black British	493 (0.8)	5476 (1.0)	3
Mixed or multiple ethnic groups	998 (1.7)	9598 (1.8)	2
Other ethnic groups	288 (0.5)	4804 (0.9)	2
Missing	9726 (16.6)	89 (0.01)	—
Region, n (%)			
North East	2528 (4.3)	—	5
North West	5572 (9.5)	—	12
Yorkshire and The Humber	3035 (5.2)	—	9
East Midlands	4945 (8.4)	—	8
West Midlands	3020 (5.2)	—	10
East of England	10,545 (18.0)	—	11
London	9083 (15.5)	—	15
South East	9845 (16.8)	—	15
South West	3956 (6.7)	—	10
Wales	1532 (2.6)	—	5
Missing	4567 (7.8)	—	—

[a]Sex at birth was self-reported. If missing, sex was supplemented via data linkage or derived via name-sex matching based on US names from 1930 - 2015 [23].

[b]Not available.

The study identified a total of 30,031 COVID-19 cases in England and 570 in Wales between June 2020 and February 2023. These cases were identified through self-reported positive PCR or LFT results (n=28,463), positive results from Virus Watch swabs (n=92), and data linkage (n=29,609). It is important to note that these sources were not mutually exclusive, meaning some cases may have been identified through multiple sources. Among the 35,551 participants with linked data, 642 COVID-19-related hospitalizations were identified.

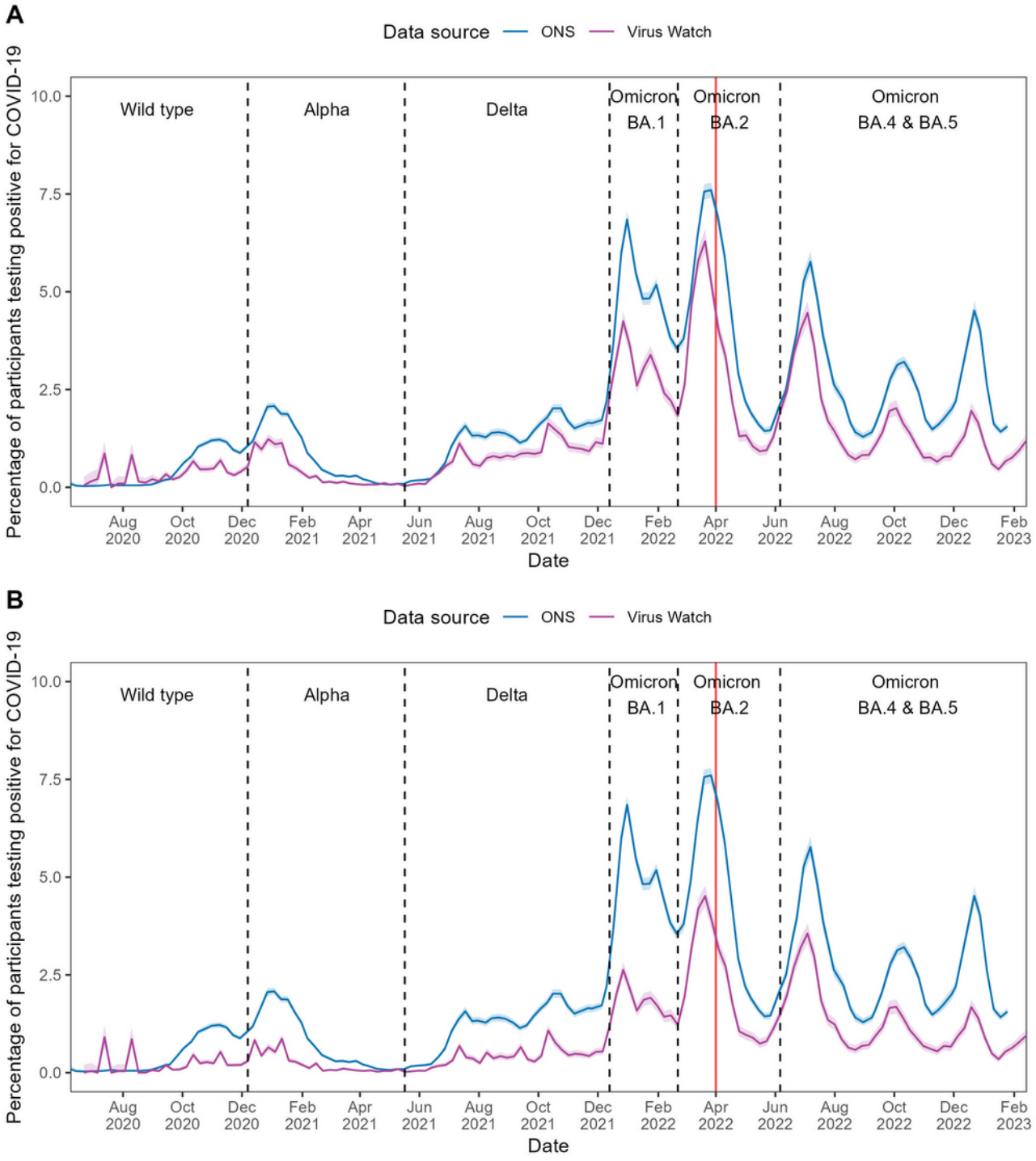
Assessing Global and Local Synchrony Between Datasets

CIS and Virus Watch–estimated positivity rates calculated with linked national testing data exhibited similar trends across all

XSL•FO
RenderX

waves of infection, although Virus Watch consistently reported lower peaks, particularly during the Omicron BA.1 wave (see Figure 1A). Overall, Spearman correlation demonstrated high global synchrony between Virus Watch and CIS estimates, with =0.92 (*P*<.001). The 9-week rolling correlations also demonstrated strong local synchrony, with a median of 0.75 (IQR 0.53 - 0.85; see Table 2 and Figure S1 in Multimedia Appendix 1). Assessing synchrony separately before and after the end of free national testing, before the end of free national testing, strong global synchrony (=0.89; *P*<.001) was observed, alongside a 9-week rolling correlation median of 0.72 (IQR 0.47 - 0.85; see Table 2 and Figure S1 in Multimedia Appendix 1). Following the end of free testing, the global synchrony decreased (=0.74; *P*<.001), while local synchrony remained similar to before the end of free national testing (median =0.76, IQR 0.58 - 0.88).

Figure 1. COVID-19 positivity rates in England estimated using Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) and Virus Watch data from June 2020 to February 2023, with the dominant variant of concern for each period indicated. Panel A presents Virus Watch estimates calculated with linked national testing data, while Panel B presents Virus Watch estimates calculated without linked national testing data. Shaded areas represent 95% CIs for the positivity rate estimates, and the red line denotes the end of free national COVID-19 testing.



XSL•FO
RenderX

Table . Overall and 9-week rolling correlation between Virus Watch and Office for National Statistics (ONS), COVID-19 Infection Survey (CIS), and Severe Acute Respiratory Infections Watch (SARI) estimates. The median (IQR) of correlation coefficients was reported for the 9-week rolling correlation. The time period indicates when each correlation was measured.

Region and rate estimates	Time period	Overall		9-week rolling
			P value	median (IQR)
England				
Positivity rates with linked data				
Overall	June 22, 2020-January 31, 2023	0.91	<.001	0.75 (0.53 - 0.85)
Before end of free national testing	June 22, 2020-March 31, 2022	0.89	<.001	0.72 (0.47 - 0.85)
After end of free national testing	April 01, 2022-January 31, 2023	0.74	<.001	0.76 (0.58 - 0.88)
Positivity rates without linked data				
Overall	June 22, 2020-January 31, 2022	0.90	<.001	0.67 (0.47 - 0.83)
Before end of free national testing	June 22, 2020-March 31, 2022	0.87	<.001	0.57 (0.40 - 0.76)
After end of free national testing	April 01, 2022-January 31, 2023	0.72	<.001	0.72 (0.53 - 0.83)
Incidence rates with linked data	June 22, 2020-June 20, 2022	0.91	<.001	0.77 (0.53 - 0.85)
Incidence rates without linked data	June 22, 2020-June 20, 2022	0.90	<.001	0.66 (0.45 - 0.82)
Hospitalization rates	August 09, 2020-March 19, 2023	0.74	<.001	0.49 (0.15 - 0.70)
Wales				
Positivity rates				
Overall	July 27, 2020-March 31, 2022	0.77	<.001	0.63 (0.40 - 0.76)
Before end of free national testing	July 27, 2020-March 31, 2022	0.67	<.001	0.55 (0.26 - 0.70)
After end of free national testing	April 01, 2022-January 31, 2022	0.63	<.001	0.68 (0.51 - 0.82)
Incidence rates	October 26, 2020-June 20, 2022	0.86	<.001	0.52 (0.38 - 0.71)

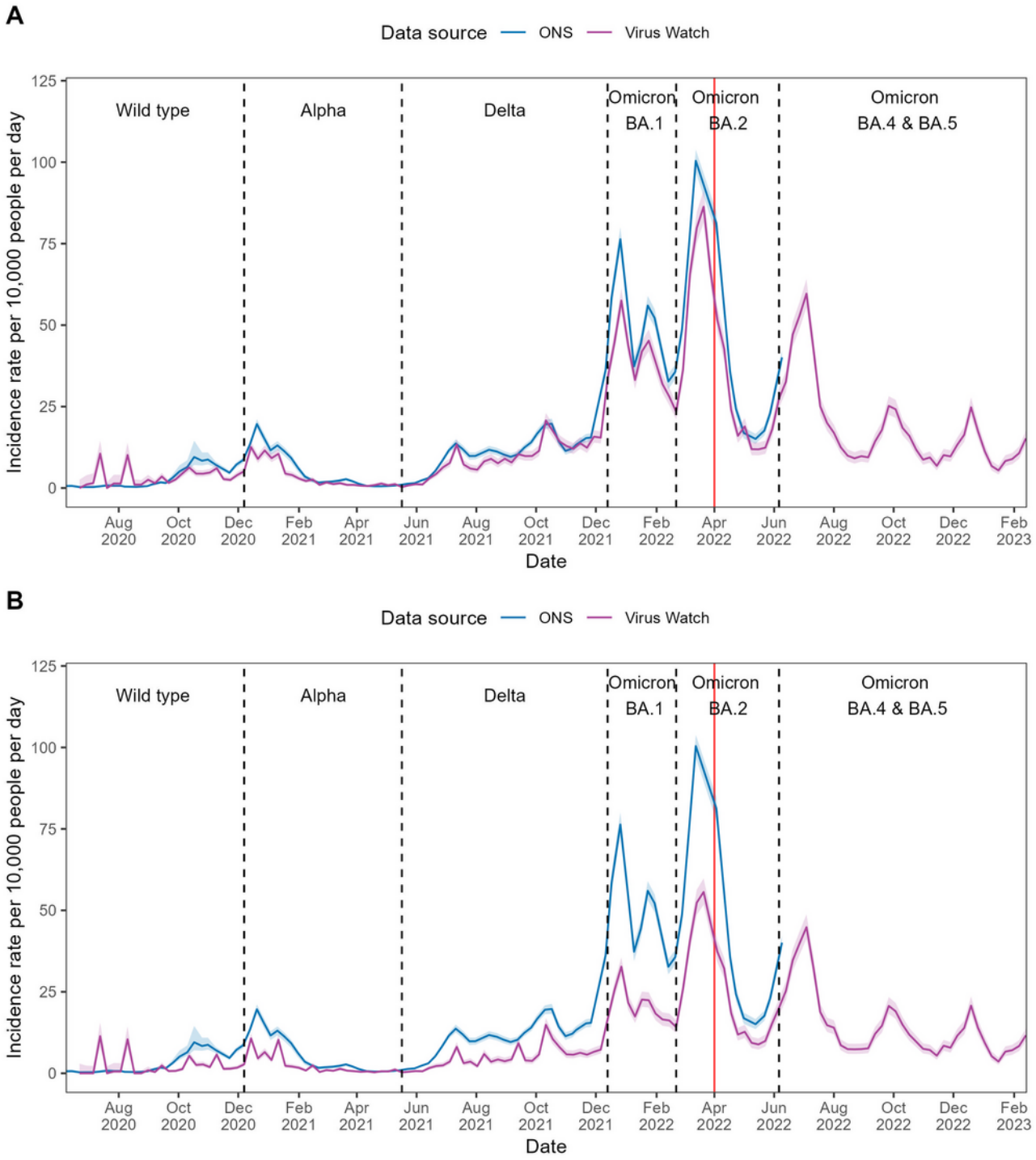
Without linked testing data, Virus Watch–estimated positivity rates were lower than estimates calculated with linked data. Compared to CIS estimates, overall trends remained broadly similar but showed less alignment during the Wild type, Alpha, and Delta waves. However, overall, Virus Watch positivity estimates were consistently lower than those from CIS (see Figure 1B). Across the full study period, Virus Watch estimates without linked testing data maintained strong global synchrony with CIS estimates (positivity =0.90; *P*<.001). The 9-week rolling correlation was comparatively weaker (median =0.67, IQR 0.47 - 0.83) than for comparisons using rates calculated with linked data; yet, trends over time remained similar (see Figure 1B, Table 2, and Figure S1 in Multimedia Appendix 1). Global synchrony was stronger before the end of free national testing than after, while local synchrony was higher after free national testing than before (see Table 2).

In England, Virus Watch–estimated incidence rates calculated with linked national testing data demonstrated strong concordance with CIS estimates across all waves of infection (see Figure 2A). While the overall magnitude of both rates was generally comparable, Virus Watch–estimated peaks during the Omicron BA.1 and BA.2 waves were slightly lower than in the CIS estimates. Spearman correlation demonstrated high global synchrony between Virus Watch and CIS estimates (=0.91; *P*<.001). The 9-week rolling correlation also indicated strong local synchrony, as seen by a high median (median 0.76, IQR 0.49 - 0.88; see Table 2 and Figure S1 in Multimedia Appendix 1). Without linked national testing data, the Virus Watch-estimated incidence rates were consistently lower than those of CIS (see Figure 2B). However, high global synchrony (=0.90; *P*<.001) was maintained, demonstrating similar temporal trends. The strength of local synchrony was reduced compared with estimates including linked data, but still reflected

XSL•FO
RenderX

moderate agreement (median =0.66, IQR 0.45 - 0.82) (see Table 2 and Figure S1 in Multimedia Appendix 1).

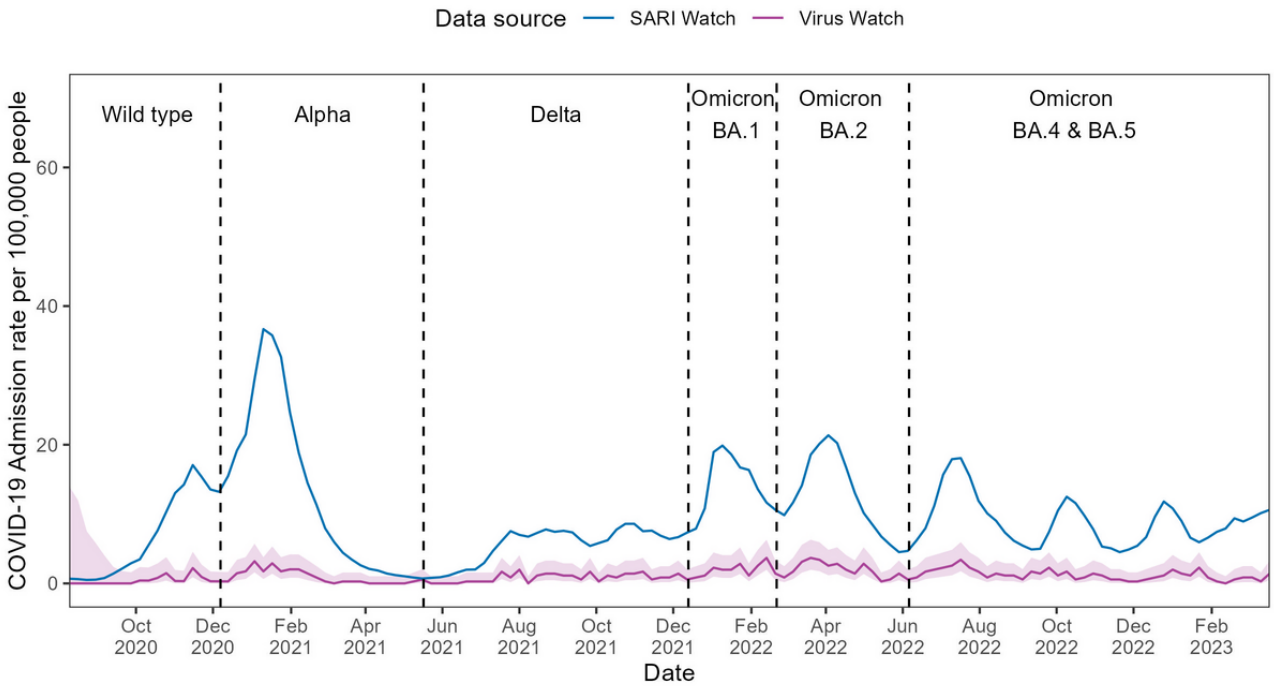
Figure 2. COVID-19 incidence rates in England estimated using Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) and Virus Watch data from June 2020 to February 2023, with the dominant variant of concern for each period indicated. Panel A presents Virus Watch estimates calculated with linked national testing data, while Panel B presents Virus Watch estimates calculated without linked national testing data. Shaded areas represent 95% CIs for the incidence rate estimates, and the red line denotes the end of free national COVID-19 testing.



The Virus Watch–estimated hospitalization rates with COVID-19 were significantly lower than SARI estimates and did not follow the same trend, failing to capture any peaks of admission observed in the SARI data across all waves of infection (see Figure 3). While Spearman correlation showed an overall positive correlation with SARI estimates (=0.74; *P*<.001), the 9-week rolling correlation showed weaker correlation and greater variability (median =0.49, IQR 0.15 - 0.70; see Table 2 and Figure S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 3. COVID-19 hospitalization rates in England estimated using Severe Acute Respiratory Infections Watch (SARI) and Virus Watch data from June 2020 to March 2023, with the dominant variant of concern for each period indicated. Shaded areas represent 95% CIs for the hospitalization rate estimates.



In Wales, Virus Watch-estimated positivity (see Figure 4) and incidence (see Figure 5) rates exhibited greater variability and wider confidence intervals compared to CIS estimates. Both estimates demonstrated broadly similar temporal trends to CIS estimates, with peaks corresponding to major waves of infection. However, Virus Watch estimates were consistently lower than those from CIS. For positivity rates, the difference between the two estimates appeared largest during the Omicron BA.1 wave, which narrowed during Omicron BA.2 and later waves. The overall trend of Virus Watch positivity rates was strongly correlated with CIS estimates (=0.77; *P*<.001), while a median of 0.63 (IQR 0.40 - 0.76) suggested moderate local synchrony. Global synchrony was stronger in the period before the end of free national testing, whereas local synchrony was higher after free testing ended (see Table 2). Furthermore, the overall Spearman for incidence rate comparison indicated high global synchrony between the 2 estimates (= 0.86; *P*<.001). Local synchrony (median =0.52, IQR 0.38 - 0.71) indicated a moderate correlation, lower than that observed for incidence rate comparisons restricted to England (see Table 2).

XSL•FO
RenderX

Figure 4. COVID-19 positivity rates in Wales estimated using Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) and Virus Watch data from June 2020 to February 2023, with the dominant variant of concern for each period indicated. Shaded areas represent 95% CIs for the positivity rate estimates, and the red line denotes the end of free national COVID-19 testing.

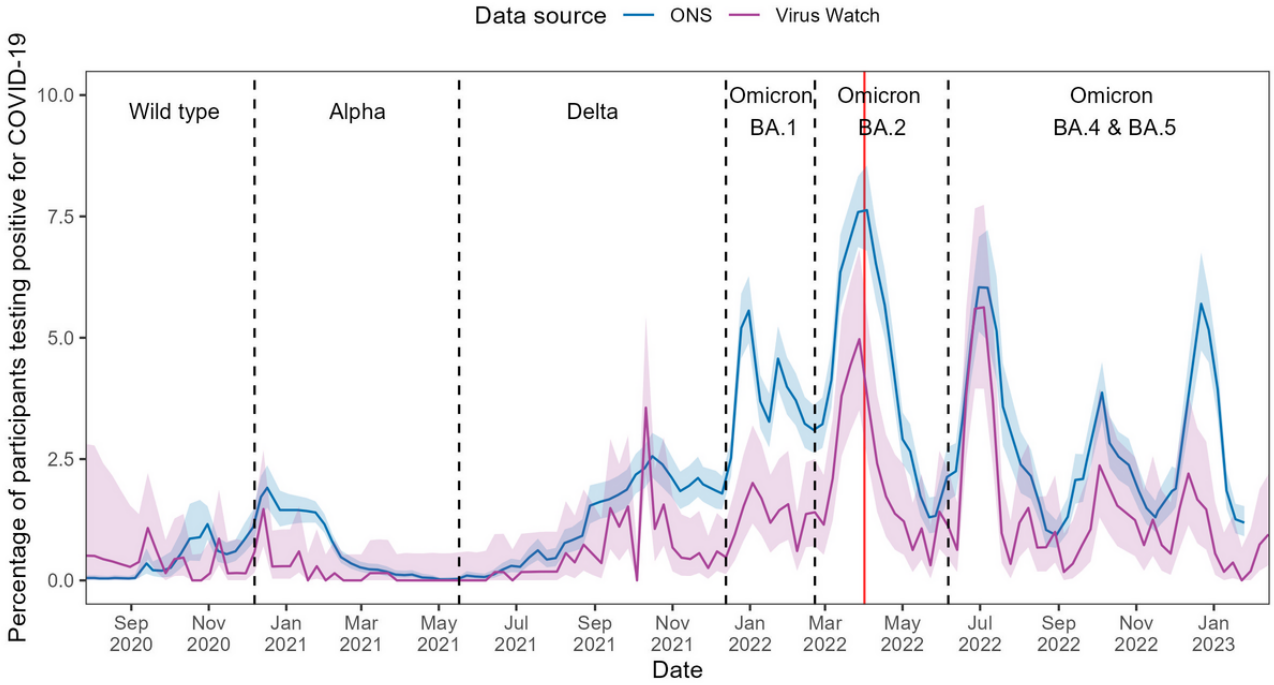
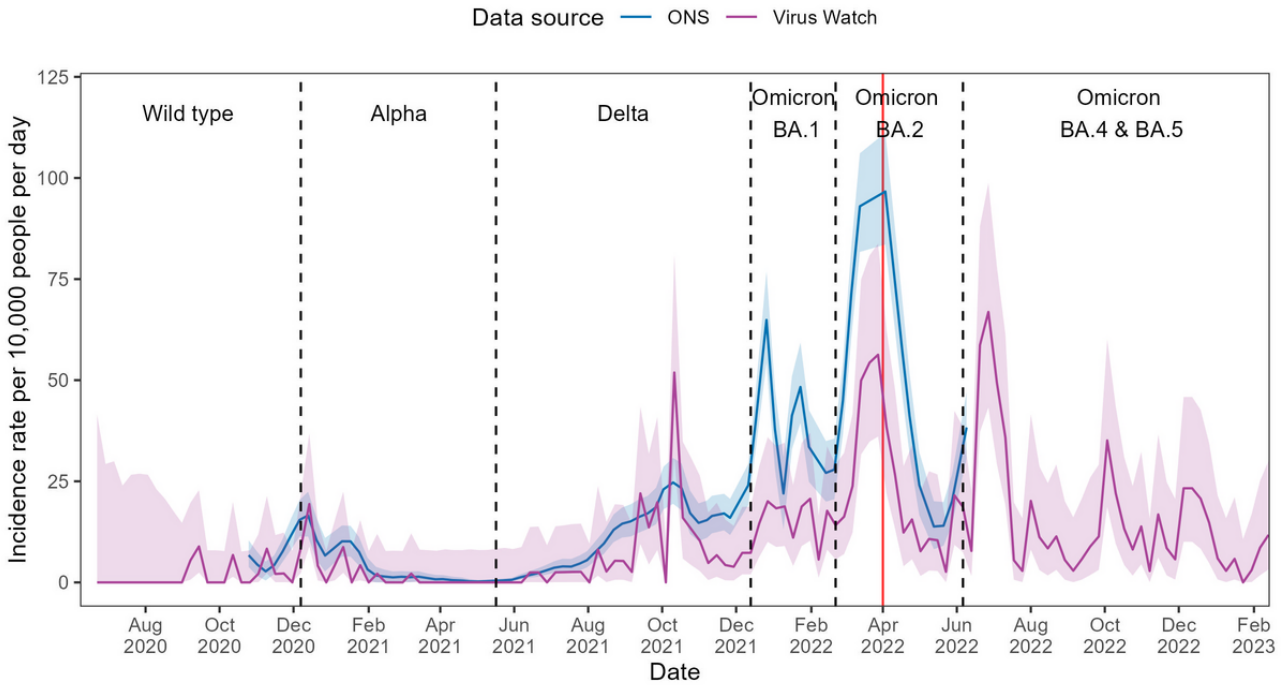


Figure 5. COVID-19 incidence rates in Wales estimated using Office for National Statistics (ONS) COVID-19 Infection Survey (CIS) and Virus Watch data from June 2020 to February 2023, with the dominant variant of concern for each period indicated. Shaded areas represent 95% CIs for the incidence rate estimates, and the red line denotes the end of free national COVID-19 testing.



Discussion

Principal Findings

Using Virus Watch data, we estimated COVID-19 positivity, incidence, and hospitalization rates in England and Wales, comparing our findings with CIS and SARI estimates. In England, we observed a strong overall and 9-week rolling correlation between Virus Watch and CIS positivity and incidence rates, particularly when using Virus Watch estimates with linked national testing data. Correlations were weaker for Wales, likely due to fluctuations in the Virus Watch rates driven by the small, nonrepresentative Welsh cohort and the lack of linked national testing data for this region. For both countries,

XSL•FO
RenderX

global synchrony of positivity rates was stronger before the free testing ended, while local synchrony was stronger after, possibly reflecting greater disagreement in early pandemic trends when testing access and behaviors were less established. Correlations between Virus Watch and SARI hospitalization rates were also positive but weaker than those observed with CIS positivity and incidence estimates.

The feasibility and effectiveness of internet-based, population-based surveillance studies were also demonstrated early in the pandemic through the COVID-19 Symptom Study, which recruited over 2.8 million users and received more than 120 million daily reports of symptoms or swab test results on their mobile app [20]. Although the study did not link to national testing data, it was still able to estimate COVID-19 prevalence, incidence, and effective reproduction number that were in good agreement with estimates from CIS. These findings reinforce the potential of participatory surveillance to track population-level infection patterns.

Key Differences in Study Designs and Their Implications

ONS CIS, SARI, and Virus Watch differ significantly in study designs, resulting in Virus Watch reporting lower COVID-19 positivity, incidence, and hospitalization rates than CIS and SARI. CIS, with its large, population-based sample and regular and incentivized PCR testing, captured both symptomatic and asymptomatic cases [2]. Virus Watch, by contrast, primarily captured symptomatic cases through self-testing, with incentives provided only during a specific recruitment phase [8]. Without continuous incentives for engagement and following the end of free national testing, Virus Watch yielded lower estimates of COVID-19 positivity and incidence. For SARI, data reporting is mandatory from NHS trusts; however, between 2023 and 2024, about 75% of trusts contributed data [12]. Although trust coverage is incomplete, SARI is still able to provide indicative estimates of hospitalization rates in England [12]. Virus Watch hospitalization data was obtained through linkage to national hospital surveillance data, which reduces the risk of underreporting compared to self-report. However, since the cohort is not fully representative of the general population, particularly underrepresenting individuals who are very old or clinically extremely vulnerable, Virus Watch–estimated hospitalization rates are lower than those reported by SARI.

Surveillance systems are essential for timely outbreak detection, identifying at-risk populations, and evaluating population-level public health interventions. In addition, factors such as timeliness and cost-effectiveness should be considered for their implementation [20]. While the CIS can be considered the “best standard” surveillance system, its limitations, including timeliness and cost, can restrict its implementation in any setting once pandemic-specific funding ends and may pose even greater challenges in resource-limited settings [20]. Passive surveillance systems, such as SARI, are relatively inexpensive and easy to implement as they use real-time data from existing publicly funded hospital-based systems [1]. However, SARI estimates, similar to Virus Watch estimates, are contingent on national testing guidelines and practices, which may impact its sensitivity [12]. Furthermore, while SARI remains a reliable surveillance system for monitoring hospitalizations due to respiratory infections, its usability for tracking disease spread within the community using hospitalization rates may be affected, particularly as the clinical pyramid of infection evolves. For example, vaccination rollouts and the emergence of less virulent variants can reduce disease severity, resulting in fewer admissions despite ongoing community transmission [20]. In addition, the determinants of exposure can differ substantially from those of severe disease requiring hospitalization. Using only hospitalization data to monitor disease spread without understanding patterns of community transmission creates an information bottleneck in epidemic and pandemic contexts. It can also overestimate virulence, distorting key epidemiological metrics and public health strategies [13]. Alternatively, surveillance using national testing data has been shown to be a reliable indicator of epidemic progress and offers timely and extensive coverage. However, it may lack the information necessary to identify at-risk populations and risk factors for infection [24]. The reliance on a comprehensive national testing infrastructure also limits the viability of this approach.

Given the limitations of individual surveillance systems, a complementary approach, combining multiple methods, is preferable. Virus Watch used a multifaceted approach that integrated community-based testing, internet-based weekly surveys, and linkage to national datasets to estimate COVID-19 positivity, incidence, and hospitalization rates. Data linkage reduced the cost of the study as well as the risk of underreporting associated with self-testing, enhancing the study results’ reliability. However, while Virus Watch hospitalization estimates are less suitable for tracking absolute rates, they can provide some indication of trends over time and remain valuable for identifying relative differences across demographic groups.

The collection of sociodemographic and clinical data of participants also allowed for the analysis of clinical risk factors and symptom profiles across different SARS-CoV-2 variants [25-31]. The Virus Watch approach can act as a low-cost supplementary surveillance method during periods of national testing. Without a national testing program, Virus Watch’s symptomatic self-swabbing strategy provides reliable estimates for community positivity and incidence rates at significantly lower costs compared to larger studies that involve swabbing regardless of symptoms. Our analysis also suggests that, even after the end of free national testing in England, the reduction in global synchrony with CIS estimates was minimal and did not undermine the reliability of overall trends. These findings support the validity of the Virus Watch approach as a reliable source of proxy estimates of community-level incidence and prevalence. Furthermore, the Flu Watch study, on which our study methodology was based, demonstrated the effectiveness of this approach for monitoring influenza transmission during and after the H1N1 pandemic in England without extensive testing systems. Flu Watch–estimated PCR-confirmed influenza rates were, on average, 22 times higher (95% CI 17 - 28) than rates reported through primary-care–based surveillance during the winter seasons from 2009 to 2011 [9]. The adaptable and cost-efficient design of Virus Watch ensures its sustainability beyond peak outbreaks, enabling the monitoring and guiding of target responses to emerging variants and future acute

XSL•FO
RenderX

respiratory diseases. While our findings suggest that infection trends can be robustly monitored even in the absence of a national testing program, further investigation is needed to determine whether the Virus Watch approach can be effectively adapted for use in low-resource settings.

Limitations

Several limitations should be considered. First, due to differences in study design, we were unable to directly compare COVID-19 positivity and incidence rates with ONS CIS, as we could not use the same methodology for estimating these rates. This is because we were unable to model the clearance time, as ONS CIS did, since participants do not test regularly for COVID-19. As a result, we lack data on the date they stop testing positive. While the difference in methodology limits the comparability of our estimates, the similarity observed between the 2 studies remains encouraging. In addition, Virus Watch likely underestimated positivity and incidence rates, as it mainly captured symptomatic cases and did not incentivize swab returns or weekly reporting, in contrast to ONS CIS. The accuracy of our estimated rates may also be impacted by the self-reporting nature of our study, since reports serve as the first recorded sign of infection, which may not align with the actual onset of infection. Furthermore, the smaller sample size in our study also affects the generalizability of our findings, particularly in the lower age groups.

The self-selected nature of the Virus Watch cohort, while broadly representative of the general English population, introduces potential biases. Our cohort is likely biased toward individuals with a greater interest in COVID-19 and health research, potentially skewing toward a more cautious and health-conscious demographic. In addition, the underrepresentation of younger and the oldest age groups and ethnic minority groups, particularly Black and other Asian groups, and the exclusion of households with more than 6 members and individuals residing in institutional settings further restricts the generalizability of our COVID-19 estimates. Furthermore, the significant attrition observed over the 3-year study period, particularly among younger, ethnically diverse, and London-based participants, underscores the challenges of long-term cohort studies [8]. While this pattern is consistent with other longitudinal studies, this limitation may impact the representativeness of the cohort over time at faster rates compared to ONS CIS due to their continual incentivization [32].

Conclusion

In this study, we used Virus Watch data to estimate COVID-19 positivity, incidence, and hospitalization rates in England and Wales, comparing our findings with estimates from ONS CIS and SARI. Despite differences in study design, Virus Watch demonstrated strong correlations with ONS CIS infection rates, highlighting its usability as a complementary surveillance method. Each surveillance system has unique strengths and limitations, and our findings underscore the need for complementary surveillance approaches to address gaps in individual systems. Virus Watch's sustainable and low-cost design is suited for settings without extensive national testing infrastructure. Beyond COVID-19, this surveillance approach may also effectively track other acute respiratory infections, guiding public health responses in both high- and low-resource settings.

Acknowledgments

This work was supported by the Medical Research Council (grant MC_PC 19070) awarded to University College London on March 30, 2020, and the Medical Research Council (grant MR/V028375/1) awarded on August 17, 2020. The study also received US $15,000 of advertising credit from Facebook to support a pilot social media recruitment campaign on August 18, 2020. This study was supported by the Wellcome Trust through a Wellcome Clinical Research Career Development Fellowship to RWA (206602). From May 1, 2022, Virus Watch received funding from the European Union (project 101046314). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them.

Data Availability

Data from the Virus Watch cohort are available (excluding any data or variables originating from linkage via NHS Digital) on the Office for National Statistics Secure Research Service (SRS) [33]. The data are available under restricted access and can be obtained by submitting a request directly to the SRS. The data citation should include a persistent identifier (eg, web URL or DOI) cited per journal style in the reference list.

Authors' Contributions

Conceptualization: EF, VGN, RWA
Methodology: ACH, EF, JK, VGN, SB, TEB, WLEF, CG, RWA
Project administration: ACH, EF, JK, VGN, SB, TEB, WLEF, CG, RWA
Data curation: VGN, WLEF, SB, CG, TEB, RWA
Writing – original draft: WLEF
Writing – review & editing: all authors

XSL•FO
RenderX

Conflicts of Interest

ACH serves on the UK New and Emerging Respiratory Virus Threats Advisory Group. None declared by the other authors.

Multimedia Appendix 1
Supplementary figures showing the 9-week rolling Spearman correlation coefficients over time of COVID-19 positivity, incidence, and hospitalization rates in England and Wales.
[DOCX File, 1947 KB - publichealth_v11i1e69655_app1.docx]

References

1. Nsubuga P, White ME, Thacker SB, et al. Public health surveillance: a tool for targeting and monitoring interventions. In: Disease Control Priorities in Developing Countries: Oxford University Press; 2006. URL: https://www.ncbi.nlm.nih.gov/books/NBK11770 [accessed 2024-11-19]
2. Coronavirus (COVID-19) infection survey: quality and methodology information (QMI). Office for National Statistics. 2021. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/methodologies/coronaviruscovid19infectionsurveyqmi [accessed 2024-11-18]
3. Elliott P, Whitaker M, Tang D, et al. Design and Implementation of a National SARS-CoV-2 Monitoring Program in England: REACT-1 Study. Am J Public Health 2023 May;113(5):545-554. [doi: 10.2105/AJPH.2023.307230] [Medline: 36893367]
4. Total cost of the COVID-19 infection survey. Office for National Statistics. 2023. URL: https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/totalcostofthecovid19infectionsurvey [accessed 2024-11-18]
5. COVID-19 infection survey vouchers and retailers. Office for National Statistics. 2022. URL: https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/covid19infectionsurveyvouchersandretailers [accessed 2024-11-18]
6. Ibrahim NK. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. J Infect Public Health 2020 Nov;13(11):1630-1638. [doi: 10.1016/j.jiph.2020.07.019] [Medline: 32855090]
7. Ricoca Peixoto V, Nunes C, Abrantes A. Epidemic surveillance of COVID-19: considering uncertainty and under-ascertainment [Article in Portuguese]. Port J Public Health 2020 Apr 9:1-7. [doi: 10.1159/000507587] [Medline: 40504207]
8. Byrne T, Kovar J, Beale S, et al. Cohort Profile: Virus Watch-understanding community incidence, symptom profiles and transmission of COVID-19 in relation to population movement and behaviour. Int J Epidemiol 2023 Oct 5;52(5):e263-e272. [doi: 10.1093/ije/dyad087] [Medline: 37349899]
9. Hayward AC, Fragaszy EB, Bermingham A, et al. Comparative community burden and severity of seasonal and pandemic influenza: results of the Flu Watch cohort study. Lancet Respir Med 2014 Jun;2(6):445-454. [doi: 10.1016/S2213-2600(14)70034-7] [Medline: 24717637]
10. NHS test and trace: what to do if you are contacted. GOV.UK. 2020. URL: https://www.gov.uk/guidance/nhs-test-and-trace-how-it-works [accessed 2024-11-22]
11. National influenza and COVID-19 surveillance report: week 27 report (up to week 26 2024 data). UK Health Security Agency. 2024. URL: https://assets.publishing.service.gov.uk/media/6686a730d9d35187868f44ad/weekly-flu-and-COVID-19-report-week-27.pdf [accessed 2024-11-18]
12. Data quality report: national flu and COVID-19 surveillance report. GOV.UK. 2024. URL: https://tinyurl.com/3xnwz6ke [accessed 2024-11-18]
13. Stoto MA, Kraemer JD, Piltch-Loeb R. Better metrics to guide public health policy: lessons learned from COVID-19 for data systems improvement. Harvard Data Science Review 2023;5(1). [doi: 10.1162/99608f92.3e516c04]
14. Hayward A, Fragaszy E, Kovar J, et al. Risk factors, symptom reporting, healthcare-seeking behaviour and adherence to public health guidance: protocol for Virus Watch, a prospective community cohort study. BMJ Open 2021 Jun 23;11(6):e048042. [doi: 10.1136/bmjopen-2020-048042] [Medline: 34162651]
15. Changing the COVID-19 case definition. UK Health Security Agency. 2022 Feb 4. URL: https://ukhsa.blog.gov.uk/2022/02/04/changing-the-covid-19-case-definition/ [accessed 2025-08-12]
16. Coronavirus (COVID-19) infection survey: England. Office for National Statistics. 2023 Mar. URL: https://tinyurl.com/yuhec5fe [accessed 2024-11-18]
17. Coronavirus (COVID-19) latest insights. Office for National Statistics. 2023. URL: https://tinyurl.com/9j2vjaey [accessed 2024-11-18]
18. Coronavirus (COVID-19) infection survey: methods and further information. Office for National Statistics. 2023. URL: https://tinyurl.com/23p2h4wx [accessed 2024-11-18]
19. Pouwels KB, House T, Pritchard E, et al. Community prevalence of SARS-CoV-2 in England from April to November, 2020: results from the ONS Coronavirus Infection Survey. Lancet Public Health 2021 Jan;6(1):e30-e38. [doi: 10.1016/S2468-2667(20)30282-6] [Medline: 33308423]

XSL•FO
RenderX

20. Brainard J, Lake IR, Morbey RA, Jones NR, Elliot AJ, Hunter PR. Comparison of surveillance systems for monitoring COVID-19 in England: a retrospective observational study. Lancet Public Health 2023 Nov;8(11):e850-e858. [doi: 10.1016/S2468-2667(23)00219-0] [Medline: 37832574]
21. Population of England and Wales - ethnicity facts and figures. GOV.UK. 2022. URL: https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/population-of-england-and-wales/latest/ [accessed 2024-11-22]
22. Population and household estimates, England and Wales: census 2021. Office for National Statistics. 2022. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationandhouseholdestimatesenglandandwales/census2021 [accessed 2024-11-22]
23. Howard D. Gender by name. data.world. 2016. URL: https://data.world/howarder/gender-by-name [accessed 2024-11-22]
24. Clare T, Twohig KA, O'Connell AM, Dabrera G. Timeliness and completeness of laboratory-based surveillance of COVID-19 cases in England. Public Health 2021 May;194:163-166. [doi: 10.1016/j.puhe.2021.03.012] [Medline: 33945929]
25. Navaratnam AMD, Beale S, Boukari Y, et al. Assessing the causal effect of air pollution on risk of SARS-cov-2 infection. Epidemiology: medRxiv Preprint posted online on Oct 26, 2023. [doi: 10.1101/2023.10.26.23297598]
26. Beale S, Hoskins S, Byrne T, et al. Workplace contact patterns in England during the COVID-19 pandemic: Analysis of the Virus Watch prospective cohort study. Lancet Reg Health Eur 2022 May;16:100352. [doi: 10.1016/j.lanepe.2022.100352] [Medline: 35475035]
27. Hoskins S, Beale S, Nguyen V, et al. Relative contribution of essential and non-essential activities to SARS-CoV-2 transmission following the lifting of public health restrictions in England and Wales. Epidemiol Infect 2023;151:e3. [doi: 10.1017/S0950268822001832]
28. Fong WLE, Nguyen VG, Burns R, et al. The incidence of COVID-19-related hospitalisation in migrants in the UK: Findings from the Virus Watch prospective community cohort study. J Migr Health 2024;9:100218. [doi: 10.1016/j.jmh.2024.100218] [Medline: 38559897]
29. Fragaszy E, Shrotri M, Geismar C, et al. Symptom profiles and accuracy of clinical case definitions for COVID-19 in a community cohort: results from the Virus Watch study. Wellcome Open Res 2022;7:84. [doi: 10.12688/wellcomeopenres.17479.1] [Medline: 37745779]
30. Aldridge RW, Pineo H, Fragaszy E, et al. Household overcrowding and risk of SARS-CoV-2: analysis of the Virus Watch prospective community cohort study in England and Wales. Wellcome Open Res 2021;6(347). [doi: 10.12688/wellcomeopenres.17308.1] [Medline: 38807847]
31. Hoskins S, Beale S, Nguyen V, et al. Settings for non-household transmission of SARS-CoV-2 during the second lockdown in England and Wales - analysis of the Virus Watch household community cohort study. Wellcome Open Res 2022;7:199. [doi: 10.12688/wellcomeopenres.17981.1] [Medline: 36874571]
32. Teague S, Youssef GJ, Macdonald JA, et al. Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. BMC Med Res Methodol 2018 Nov 26;18(1):151. [doi: 10.1186/s12874-018-0586-7] [Medline: 30477443]
33. Virus Watch: England and Wales. Office of National Statistics. 2022 May 19. URL: https://ons.metadata.works/browser/dataset?id=89201 [accessed 2025-09-23]

Abbreviations

CIS: Community Infection Survey
HES: Hospital Episode Statistics
LFT: lateral flow test
NHS: National Health Service
ONS: Office for National Statistics
PCR: polymerase chain reaction
REACT: Real-Time Assessment of Community Transmission
SARI: Severe Acute Respiratory Infections Watch
SGSS: Second Generation Surveillance System

Edited by C Argyropoulos; submitted 05.12.24; peer-reviewed by A Carnahan, S Farley; revised version received 27.08.25; accepted 01.09.25; published 29.09.25.

Please cite as:
Fong WLE, Nguyen VG, Beale S, Byrne TE, Geismar C, Fragaszy E, Kovar J, Yavlinsky A, Abubakar I, Hayward AC, Aldridge RW Comparison of Virus Watch COVID-19 Positivity, Incidence, and Hospitalization Rates With Other Surveillance Systems: Surveillance Study
JMIR Public Health Surveill 2025;11:e69655
URL: https://publichealth.jmir.org/2025/1/e69655
doi:10.2196/69655

XSL•FO
RenderX

© Wing Lam Erica Fong, Vincent Grigori Nguyen, Sarah Beale, Thomas E Byrne, Cyril Geismar, Ellen Fragaszy, Jana Kovar, Alexei Yavlinsky, Ibrahim Abubakar, Andrew C Hayward, Robert W Aldridge. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Analyzing Satellite Imagery to Target Tuberculosis Control Interventions in Densely Urbanized Areas of Kigali, Rwanda: Cross-Sectional Pilot Study

Mauro Faccin[1,2], PhD; Caspar Geenen[3], MD; Michiel Happaerts[3,4], MD; Sien Ombelet[3,5*], MD, PhD; Patrick Migambi[6*], MD; Emmanuel André[3,5*], MD, PhD

[1]Department of Medical and Surgical Sciences, University of Bologna, Via Massarenti 9, Bologna, Italy

[2]Istituto Nazionale di Fisica Nucleare, Sezione di Bologna, Bologna, Italy

[3]Department of Microbiology, Immunology and Transplantation, Laboratory of Clinical Microbiology, KU Leuven, Leuven, Belgium

[4]Department of General Internal Medicine, University Hospital Leuven, Leuven, Belgium

[5]Department of Laboratory Medicine, University Hospital Leuven, Leuven, Belgium

[6]Tuberculosis and Other Respiratory Disease Division, Department of HIV/AIDS Diseases Prevention and Control, Rwanda Biomedical Center, Kigali, Rwanda

[*]these authors contributed equally

Corresponding Author:
Mauro Faccin, PhD
Department of Medical and Surgical Sciences, University of Bologna, Via Massarenti 9, Bologna, Italy

Abstract

Background: Early diagnosis and treatment initiation for tuberculosis (TB) not only improve individual patient outcomes but also reduce circulation within communities. Active case-finding (ACF), a cornerstone of TB control programs, aims to achieve this by targeting symptom screening and laboratory testing for individuals at high risk of infection. However, its efficiency is dependent on the ability to accurately identify such high-risk individuals and communities. The socioeconomic determinants of TB include difficulties in accessing health care and high within-household contact rates. These two determinants are common in the poorest neighborhoods of many sub-Saharan cities, where household crowding and lack of health-care access often coincide with malnutrition and HIV infection, further contributing to the TB burden.

Objective: In this study, we propose a new approach to enhance the efficacy of ACF with focused interventions that target subpopulations at high risk. In particular, we focus on densely inhabited urban areas, where the proximity of individuals represents a proxy for poorer neighborhoods with enhanced contact rates.

Methods: To this end, we used satellite imagery of the city of Kigali, Rwanda, and computer-vision algorithms to identify areas with a high density of small residential buildings. We subsequently screened 10,423 people living in these areas for TB exposure and symptoms and referred patients with a higher risk score for polymerase chain reaction testing.

Results: We found autocorrelation in questionnaire scores for adjacent areas up to 782 meters. We removed the effects of this autocorrelation by aggregating the results based on H3 hexagons with a long diagonal of 1062 meters. Out of 324 people with high questionnaire scores, 202 underwent polymerase chain reaction testing, and 9 people had positive test results. We observed a weak but statistically significant correlation (r=0.28; P=.04) between the mean questionnaire score and the mean urban density of each hexagonal area.

Conclusions: Nine previously undiagnosed individuals had positive test results through this screening program. This limited number may be due to low TB incidence in Kigali, Rwanda, during the study period. However, our results suggest that analyzing satellite imagery may allow the identification of urban areas where inhabitants are at higher risk of TB. These findings could be used to efficiently guide targeted ACF interventions.

(JMIR Public Health Surveill 2025;11:e68355) doi:10.2196/68355

KEYWORDS

tuberculosis; Rwanda; satellite image; TB; PCR testing; PCR; questionnaire; satellite; active case-finding; diagnostic; urban; Africa; TB screening; ACF; polymerase chain reaction

XSL•FO
RenderX

Introduction

Tuberculosis (TB) is one of the major global health challenges that could be managed through prompt diagnosis and treatment initiation. An estimated 10.6 million people became ill, and 1.6 million people died from TB in 2021 [1]. Despite international efforts, the burden of disease remains uncontrolled in many of the most fragile communities around the world. As poverty-affected communities experience major difficulties in accessing quality medical care, screening for TB in these populations cannot rely solely on health care–centered diagnosis. Considering that disease notification by routine diagnostic activities imperfectly reflects the burden of disease in such populations, alternative case detection approaches are needed.

To this end, community-based active case-finding (ACF) can be an effective approach if targeted at high-risk subgroups or implemented with high coverage and intensity [1-5]. However, the yield of ACF is difficult to predict, and cost-efficiency can be questioned when a large number of people need to be screened to identify a few infected individuals [6]. In practice, large-scale screening at the community level is usually done using a simple triage tool, followed by selective laboratory testing. These triage tools typically include AI-assisted chest X-ray radiography or symptom screening [7,8]. However, even these triage tests can be difficult to scale up and are inefficient if deployed in low-risk communities. In a previous study [8], we showed that the burden of TB could be accurately predicted in rural environments of the Democratic Republic of the Congo by integrating historical disease notification, distance of villages to the nearest health care center, and proximity to mining activities.

Africa accounted for a quarter of all new TB cases worldwide in 2022 and has been at the center of many efforts to eradicate TB [1]. Since 1990, the continent has seen rapid urbanization, with 0.5 billion more people now living in urban areas [9]. While population density is generally associated with easier geographical access to health services, rapid urbanization is also intrinsically associated with socioeconomic disparities, including the development of crowded townships [10]. Due to these disparities and the absence of specific interventions targeting the specific health needs of poverty-affected neighborhoods, TB control in urban areas may be delayed as pockets of TB are maintained by the cycle of disease and poverty.

Rwanda has moved into a pre-eradication phase regarding TB, with an estimated incidence of 56 patients per 100,000 population in 2021, a decline of more than 40% over the past two decades [1]. Nevertheless, the city of Kigali has a higher incidence of reported TB (108 per 100,000 inhabitants yearly) compared to the rest of the country.

In previous studies, metrics derived from satellite images have correlated with the prevalence of both noncommunicable and communicable diseases, particularly when affected by climate [11,12].

The aim of this study was to assess the feasibility of guiding targeted ACF interventions in an urban environment (ie, Kigali, Rwanda) using satellite images [13]. We hypothesized that those urban areas with a high density of small residential buildings—as a proxy for poverty and crowded households—are associated with a high risk of undiagnosed TB [14]. We constructed an urban density index (UDI) to highlight such areas, dispersed throughout the city, and investigate the correlation with TB screening results.

Methods

Study Design

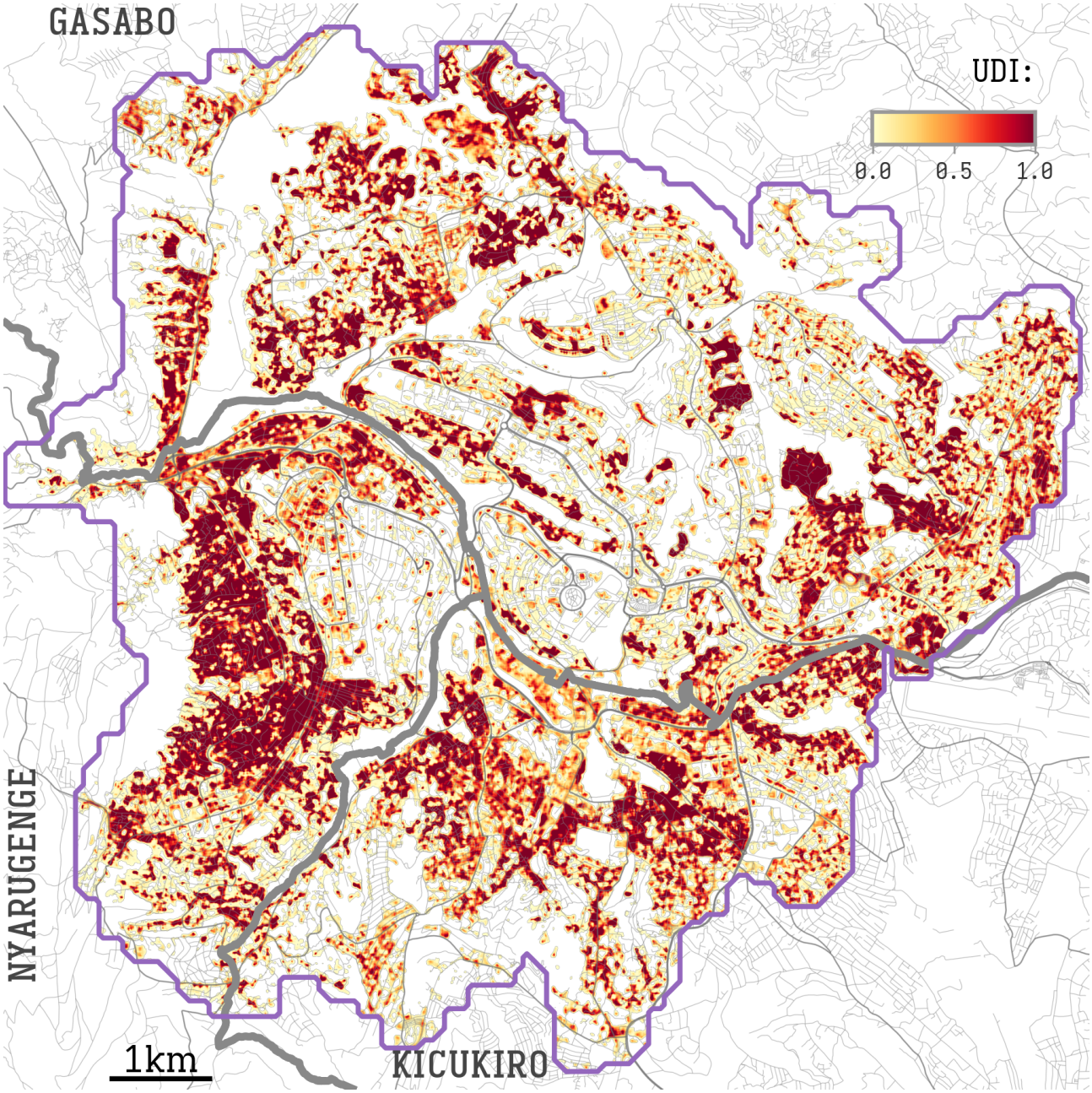
In this cross-sectional pilot study, we combined analysis of satellite images with results of ACF for TB in July 2021 in Kigali, Rwanda.

Urban Density Index Based on Satellite Images

The urban area to be analyzed was delineated based on the population density provided by WorldPop [15], selecting the cells that are estimated to have more than 8000 inhabitants per square kilometer (see the purple boundary in Figure 1). This area was extended to surrounding cells by filling holes and dilating to neighboring cells within 5 steps (approximately 500m).

XSL•FO
RenderX

Figure 1. Urban density index (UDI) overlaid on a map of Kigali. A white-to-red color scale indicates the UDI. High UDI values correspond to a high density of small buildings and a low density of green spaces. A purple line indicates the edge of the city, as defined by an estimated population density over 6000 per square kilometer.



We used a two-step computer-vision approach applied to the satellite imagery to calculate a UDI. Satellite images of the urban area of interest were retrieved from ArcGIS (source: Esri, Maxar, Earthstar Geographics, and the GIS User Community) at zoom level 17 (approximately 1 meter per pixel). All tiles intersecting the area of interest were joined in a large image.

The first step detects object contours from the prepared image with the aim to highlight zones with a high density of small houses or shelters. To perform such detection, we used the contour-detection Canny algorithm. The latter is a multistage algorithm that detects edges of an image by computation of the intensity gradient. Cells with high values of intensity gradient were selected to form an edge image. After blurring (blurring radius of approximately 50m), areas with a higher density of edges had higher values.

In the second step, we selected and discarded all green areas in the satellite images. As a result, in areas such as open fields and gardens, as well as within wealthier neighborhoods—where one can assume larger houses surrounded by gardens or public green—the index was set to zero. The green areas were detected applying a hue, saturation and value color filter to the satellite images, selecting all pixels with hue in the range of 15 - 120 and value below 120.

The result was a spatial UDI map with a resolution of approximately 1m per pixel, with the UDI ranging from 0 to 1.

Active Case Finding

Based on the selection of high-risk areas by the UDI, seven sectors (ie, Nyakabanda, Kimisagara, Rwezamenyo, Nyaregenge, Kimironko, Remera, and Niboye) were visited by

XSL•FO
RenderX

10 community health workers (CHW) during 21 days in July 2021 to perform door-to-door screening surveys. All persons residing at the visited residential locations were eligible to participate. The number of participants was determined by how many persons the CHWs could visit during this time frame.

In total, 10,423 persons were screened using MediScout (Savics, Belgium), a digitally supported questionnaire adapted from a previous study [8]. The screening and selection process was conducted by the local National TB Program after they were given access to the UDI map.

The questionnaire (see Table S1 in Multimedia Appendix 1 for the list of questions) was composed of three sets of questions evaluating the symptoms, environment, and health history of the individual. This questionnaire was composed of a set of seven questions formulated in Kinyarwanda and was administered by trained CHWs.

The weighted responses assigned to each screened individual provided a personal total score between 0 and 18 (Qscore). When the Qscore ≥4, the CHW suggested that the individual provide a sputum sample, which was analyzed using an Xpert MTB/Rif Ultra polymerase chain reaction (PCR) assay (Cepheid) [16]. The samples collected were transferred to the laboratory, and tests were performed within 24 hours after collection.

The answers to the questionnaire, together with the exact location where the survey took place, were collected using the MediScout app (Savics, version 1.0).

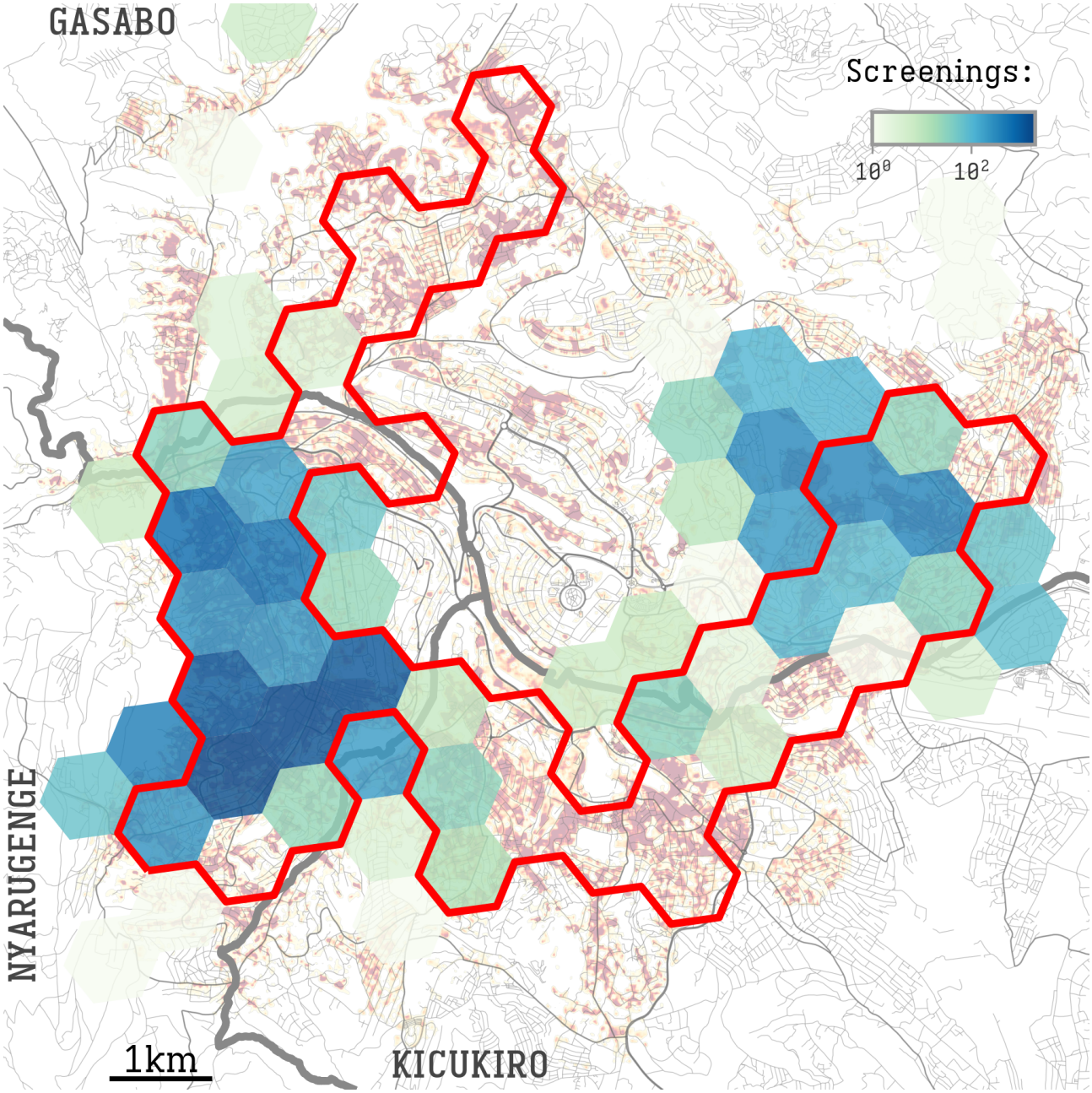
Disaggregation Length-Scale

Analysis at the level of the three city districts did not reveal major disparities in average Qscores. We then investigated whether UDI correlated with tuberculosis risk within a more restricted spatial range. To determine the optimal aggregation level, we investigated the spatial autocorrelation of risk scores. To this end, we used the global Moran's I [17,18], a statistical measure of autocorrelation of spatial data. We confirmed their results through a variogram analysis [19] (see Figure S1 in Multimedia Appendix 1). A variogram is a tool from geostatistical modeling that quantitatively assesses semivariance over a variable lag distance. This tool assesses the spatial autocorrelation of the data and estimates the distance at which those correlations vanish. We found that the estimated maximum distance at which nearby risk scores were correlated was approximately 782 meters.

Therefore, we performed a segmentation of the city area based on H3 hexagons [20] with a long diagonal of 1062 meters, the smallest segmentation level that removed the observed autocorrelation. We assumed that the risk inside each hexagon was homogeneous, despite the heterogeneity of the enclosed neighborhoods. These hexagons had a mean estimated population of 8385 (SD 4688) inhabitants. A hexagon was labeled as "high UDI" if it had a mean UDI >0.353, the 75th percentile of the city surface UDI shown in Figure 1. Figure 2 overlays these high UDI hexagons with the total number of screenings performed in each hexagon. See Figure S2 in Multimedia Appendix 1 for deeper analysis. The resulting hexagons with the aggregated screenings and UDI are included in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 2. Distribution of screenings in Kigali neighborhoods. A color scale shows the number of screening interviews performed per hexagonal area, overlaid on the urban density index (UDI) map of Figure 1. Screenings were broadly targeted to high-UDI neighborhoods, after visual inspection of Figure 1, and then aggregated in H3 hexagons. The area delineated in red shows hexagons with a high UDI (>0.353).



Statistical Analysis

Proportions were reported using binomial test confidence intervals and compared with the Fisher exact test. The aggregation of the dataset on the hexagonal grid allowed us to treat the data points as statistically independent. In regression analysis, we used the t statistic to compute the *P* values of the slope.

Persons who completed the questionnaire but did not complete PCR testing were considered lost to follow-up, and were not included in further analysis.

Ethical Considerations

The study protocol was approved by the Rwanda National Ethical Committee (Institutional Review Board 00001497 of IORG0001100). Informed consent was electronically obtained from all screened participants through the MediScout app. No personal identifiers were collected except for the location of the screening, sex, and age; these data were securely stored in a password-protected database. No compensation was provided to the participants.

Results

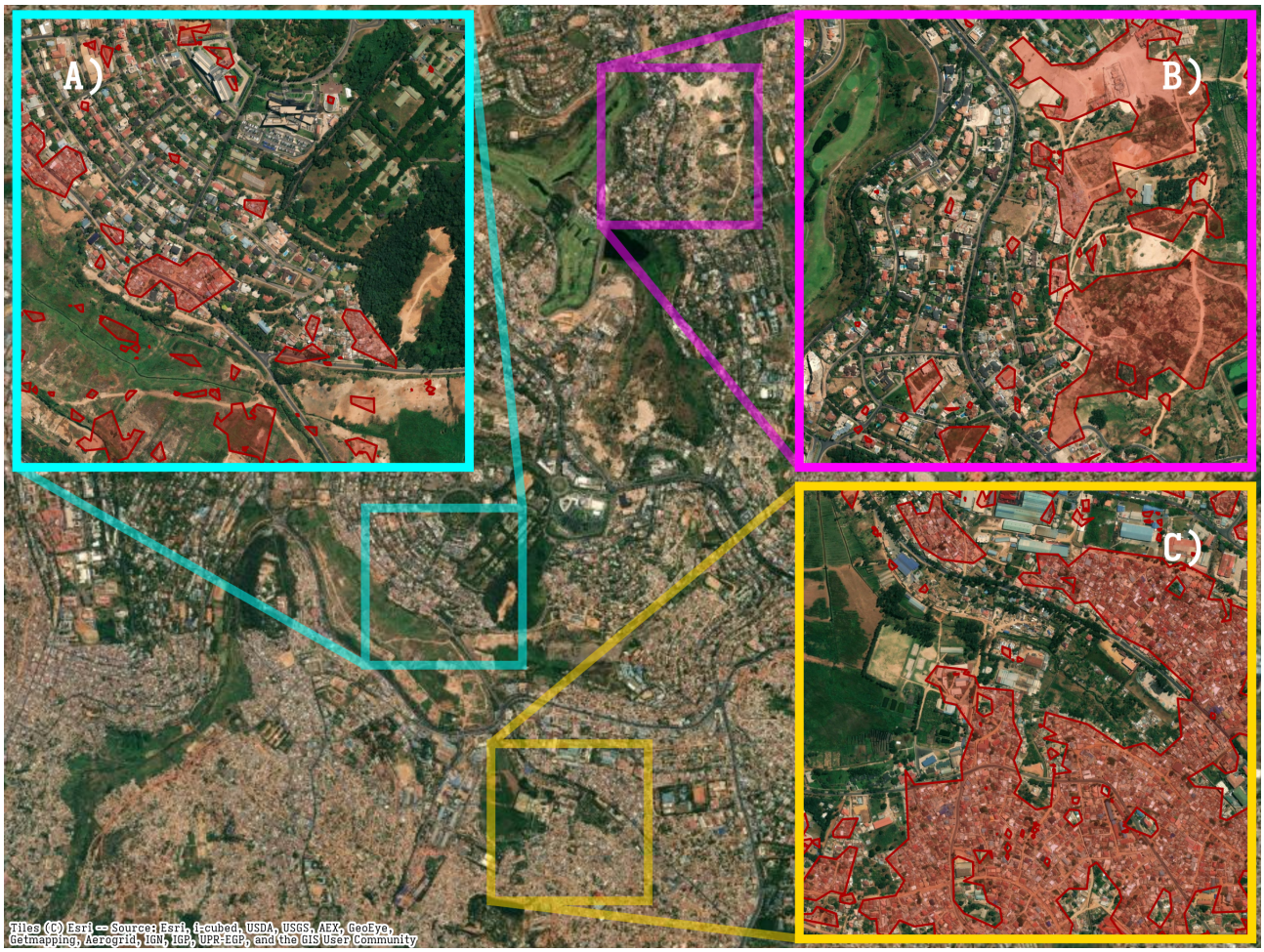
Visual Validation of the Urban Density Index

First, we analyzed satellite images to produce a UDI map of Kigali for each square meter (Figure 1). A visual inspection of selected areas (Figure 3) confirmed a clear overlap between high UDI and the density of small buildings and scarcity of green spaces. This figure highlights three areas of Kigali with

XSL•FO
RenderX

different building density distributions: small but dense clusters of buildings within healthier neighborhoods (Figure 3A); adjacent neighborhoods with differing density of buildings (Figure 3B); and dense neighborhoods surrounding green areas (Figure 3C). In all cases, the UDI identified areas with a high density of buildings.

Figure 3. Overlay of the UDI (red highlights indicate UDI >0.5) onto satellite images of Kigali, providing a visual indication of the algorithm's ability to automatically identify areas with a high density of small residential buildings and few green spaces. Inlays highlight (A) areas with small clusters of dense shelters within a wealthy neighborhood; (B) a wealthy and poor neighborhood side by side; (C) dense neighborhoods surrounding green areas.



Locations with UDI >0.5 accounted for 29% of the total surface of the city. Table 1 provides an overview of notified TB cases and UDI in the different sectors of Kigali. For each sector, the table shows the population estimate reported by the Rwandan National Institute of Statistics [21], the total number of confirmed cases reported by the Ministry of Health [22], and the resulting estimate of incidence (per year per 100,000 inhabitants). The surface area with UDI greater than 0.5 is also reported as a summary statistic.

Table . Urban density index and incidence of confirmed cases of tuberculosis by sector of Kigali.

Sector	Population	TB[a] cases (2017)	Incidence rate	Surface UDI[b] >0.5 (%)
Gasabo	879,505	570	65	27
Kicukiro	491,731	533	108	34
Nyarugenge	374,319	790	211	39
Kigali (aggregated)	1,745,555	1893	108	31

[a]TB: tuberculosis.

[b]UDI: urban density index.

Targeting of Active Case Finding Interventions Based on Urban Density Index

In total, the CHWs screened 10,423 individuals (Table 2, Figure S3 and S4 in Multimedia Appendix 1). In Nyarugenge district, which had the highest TB notification rate, 3.3% (n=270) of screened individuals had a Qscore ≥4, and this proportion was similar in the Gasabo district, which had the lowest notified TB incidence: 2.6% (n=51; *P*=.13). Mean Qscores in Nyarugenge and Gasabo among screened individuals were also similar (0.64

XSL•FO
RenderX

and 0.61, respectively; *P*=.34). Age was positively associated with Qscore (*P*<.001).

Table . Tuberculosis screening results aggregated by administrative district, age, and sex.

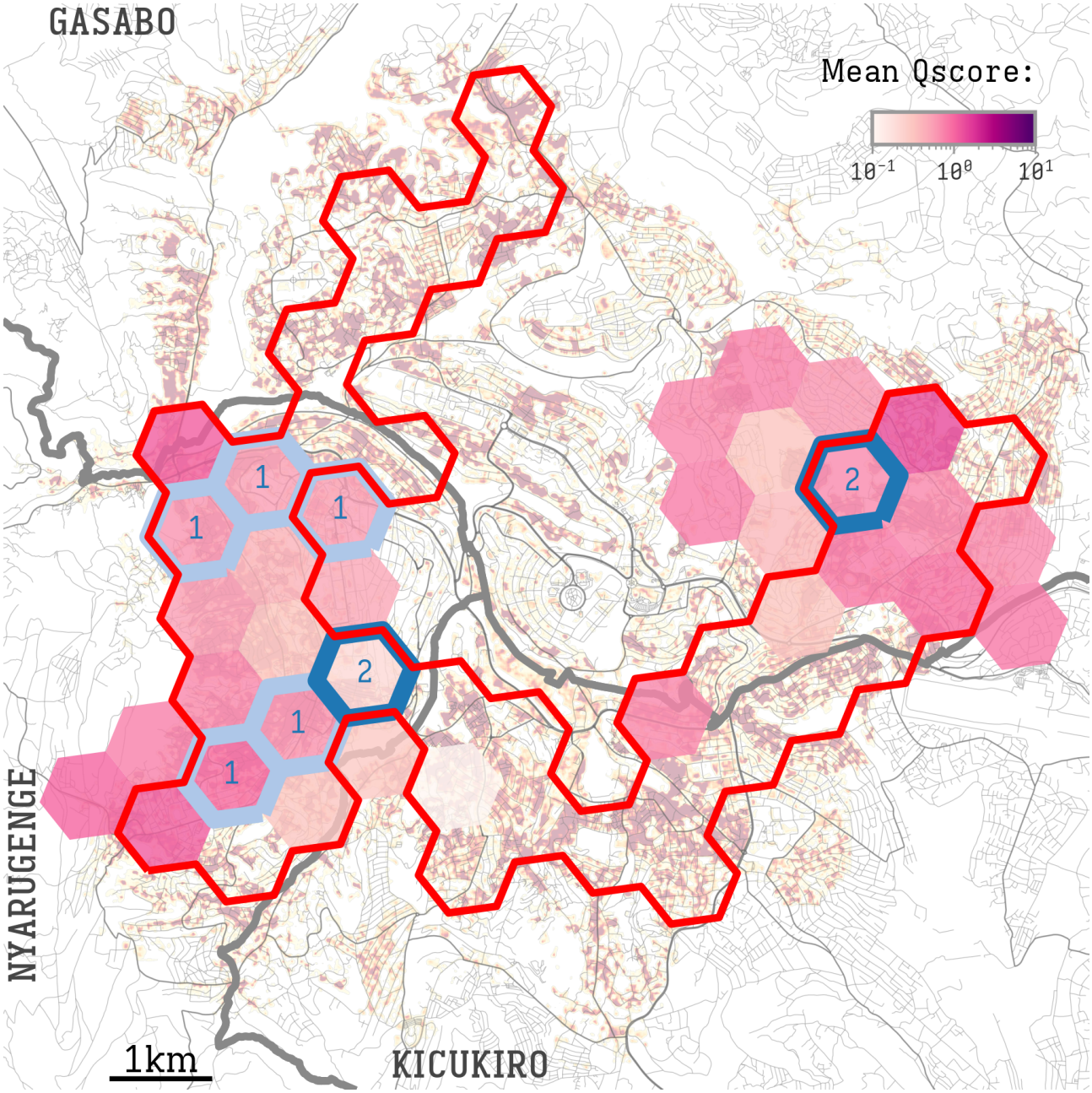
Variables	Screenings (N=10,423)	Screenings Q score ≥4	Molecular tests (n=202)	Positive tests (n=9)	Mean Qscore
Sex					
Female	5222	138	77	2	0.60
Male	5197	186	125	7	0.67
Missing	4	0	0	0	0.50
Age (years)					
1 - 10	435	7	4	0	0.34
11 - 20	1539	32	17	0	0.43
21 - 30	3098	62	40	2	0.54
31 - 40	2619	91	57	3	0.69
41 - 50	1483	75	47	4	0.82
51 - 60	672	31	18	0	0.84
61 - 70	336	16	11	0	0.94
71+	165	9	8	0	1.04
Missing	76	1	0	0	0.76
Districts					
Gasabo	1988	51	30	2	0.61
Kicukiro	129	3	3	0	0.53
Nyarugenge	8306	270	169	7	0.64
Kigali (total)	10,423	324	202	9	0.63

Correlation Between Tuberculosis Risk and Urban Density Index

Of all screenings (N=10,423), 8585 (82.4%) were performed in high UDI hexagon cells, and 1838 (17.6%) in low UDI cells (Figure 2). Among the screened individuals, 324 (3.1%) were considered at risk for TB based on the questionnaire-derived Qscore ≥4: 254 (3.0% of screenings) in high UDI cells and 70 (3.8% of screenings) in low UDI cells (*P*=.06) (Figure 4). Of the 324 at-risk individuals, 202 (62.3%) provided a sputum sample: 166 (82.2%) in high UDI cells and 36 (17.8%) in low UDI cells.

XSL•FO
RenderX

Figure 4. Mean Qscore results. A color scale shows the mean Qscore in each hexagonal cell, overlaid on the UDI map in Figure 1. Only cells with at least 10 screenings are displayed. Tuberculosis cases were found in the cells highlighted in blue, labeled with the number of cases detected. The area delineated in red shows hexagons with a high urban density index (>0.353).

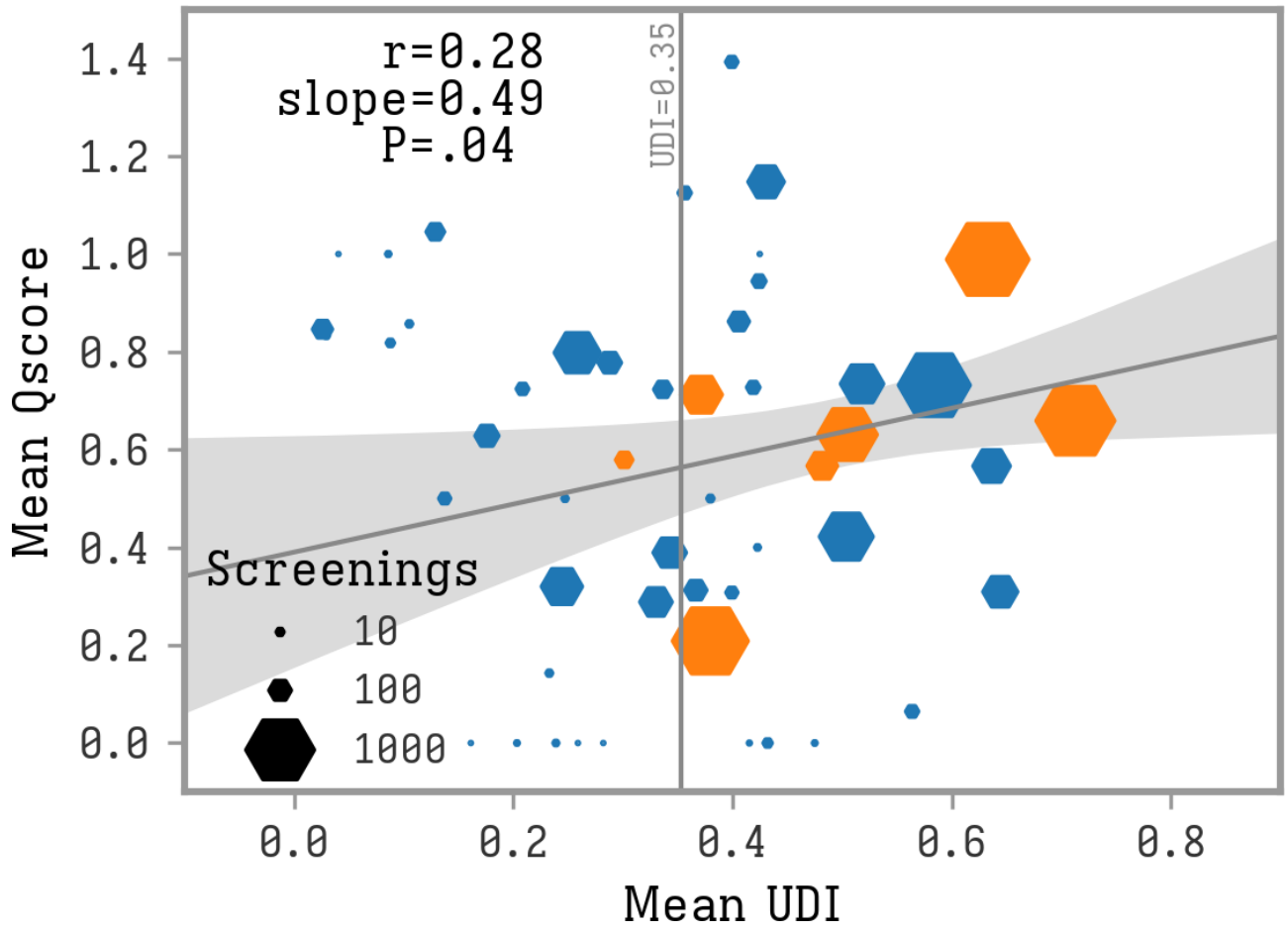


Mycobacterium tuberculosis complex was detected in 9 of 202 (4.5%) individuals using the GeneXpert assay on sputum samples. Eight of the newly detected TB cases lived in high UDI areas, corresponding to a prevalence of 93 in 100,000 (95% 47 - 175). In low UDI cells, one individual had positive results, indicating a prevalence of newly detected TB of 54 in 100,000 (95% CI 0 - 270).

We observed a weak (r=0.28) but statistically significant correlation between the mean Qscore and mean UDI of each hexagon (P=.04) (Figure 5). This suggests that locations with high UDI values, indicating dense areas within the city, tend to have higher Qscores. Given that the Qscore serves as a proxy for TB risk, this correlation implies that areas with higher UDI might experience a greater number of possibly undiagnosed cases.

XSL•FO
RenderX

Figure 5. Correlation between mean UDI and mean Qscore. Each hexagon represents a hexagonal geographical area, orange if at least one positive case was found, in blue otherwise. The size of the hexagons indicates the number of screenings performed. A black line and gray confidence band indicate the result of a linear regression analysis weighted by the number of screenings in each area. The threshold to consider an area as high UDI (mean UDI >0.353) is indicated with a vertical line. UDI: urban density index.



Discussion

In this study, we investigated the value of satellite imagery to guide targeted TB control interventions in Kigali, Rwanda. The central finding is that satellite-derived UDIs, which highlight densely populated areas with limited green spaces, could be used to identify communities that share a higher risk of TB.

We assessed this risk using a questionnaire administered to more than 10,000 participants across the city. Individual TB risk was estimated using a personalized score based on a weighted-answer questionnaire previously validated in other contexts [8]. We observed a weak but significant correlation between UDI and TB risk scores. Participants with an elevated risk for TB were offered a PCR test. Of 202 individuals referred for testing, nine cases of TB were confirmed. This corresponds to a prevalence of undetected TB of 86 per 100,000 inhabitants in the general population. Although we observed a correlation between UDI and TB risk scores, this study did not clearly demonstrate the efficient detection of TB cases due to the low number of actual cases detected. However, these findings suggest that satellite imagery could play a valuable role in identifying areas where targeted TB interventions are needed, particularly in resource-constrained urban settings. Our results highlighted that individuals living near each other had a risk for TB, which auto correlated within a radius of around 0.8 km. This information may be useful in planning and delineating targeted screening interventions in areas where new TB cases are identified. In general, only households and close contacts are currently included in ACF interventions. Our results suggest that the perimeter of ACF interventions could be extended to the neighborhood level.

The main limitation of this study was the low number of new TB cases detected, probably due to the relatively low overall prevalence of TB in Kigali during the study. Therefore, the observed correlations of TB risk with UDI were based on questionnaire-based risk scores rather than microbiological confirmation. This study also did not establish a causal link between high-density residential building and prevalence of undetected TB. However, for the purpose of efficiently targeting screening to specific urban areas, a causal link does not need to be shown. It should also be noted that this study took place in a single city, and the results may not be directly generalizable to other settings. Finally, as areas for screening were selected by health care workers based on the UDI map, high UDI areas were overrepresented in the study, resulting in selection bias and reduced statistical power.

We expect that the inclusion of more low UDI areas could have resulted in a stronger observed correlation between UDI and TB risk.

XSL•FO
RenderX

Overall, our study suggests that, with or without satellite image analysis, targeting ACF for TB to small geographic entities (eg, smaller than 1 km) could help leverage geographic disparities to increase the yield of community-based TB control. While local health care systems typically aggregate cases by larger administrative regions, TB infections might exhibit higher incidence within certain smaller areas. Nevertheless, this approach should not substitute individualized contact tracing efforts, considering the population mobility and the contact networks outside the living areas, which are also involved in disease transmission.

Acknowledgments

We thank the participants, the community health workers, National Tuberculosis Program staff, and health facility staff who supervised the data collection. We also thank Savics (Belgium) for their technological support. Part of this work was funded by Innoviris (Belgium).

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files. The datasets generated or analyzed during this study are available in the GitHub repository [23].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional information: detection of local aggregation; hexagonal grid choice; questionnaire replies frequencies; questionnaire original questions.
[PDF File, 461 KB - publichealth_v11i1e68355_app1.pdf]

Multimedia Appendix 2
Aggregated data: urban density index and aggregated screening results.
[ZIP File, 23 KB - publichealth_v11i1e68355_app2.zip]

Checklist 1
STROBE checklist.
[PDF File, 113 KB - publichealth_v11i1e68355_app3.pdf]

References

1. Global Tuberculosis Report: World Health Organization; 2022. URL: https://www.who.int/publications/i/item/9789240061729 [accessed 2025-04-15]
2. Ayabina DV, Gomes MGM, Nguyen NV, et al. The impact of active case finding on transmission dynamics of tuberculosis: a modelling study. PLOS ONE 2021;16(11):e0257242. [doi: 10.1371/journal.pone.0257242] [Medline: 34797864]
3. Bohlbro AS, Hvingelby VS, Rudolf F, Wejse C, Patsche CB. Active case-finding of tuberculosis in general populations and at-risk groups: a systematic review and meta-analysis. Eur Respir J 2021 Oct;58(4):2100090. [doi: 10.1183/13993003.00090-2021] [Medline: 33766950]
4. Burke RM, Nliwasa M, Feasey HRA, et al. Community-based active case-finding interventions for tuberculosis: a systematic review. Lancet Public Health 2021 May;6(5):e283-e299. [doi: 10.1016/S2468-2667(21)00033-5] [Medline: 33765456]
5. MacPherson P, Shanaube K, Phiri MD, et al. Community-based active-case finding for tuberculosis: navigating a complex minefield. BMC Glob Public Health 2024 Feb 8;2(1):9. [doi: 10.1186/s44263-024-00042-9] [Medline: 39681899]
6. Ogbudebe C, Jeong D, Odume B, et al. Identifying hot spots of tuberculosis in Nigeria using an early warning outbreak recognition system: retrospective analysis of implications for active case finding interventions. JMIR Public Health Surveill 2023 Feb 8;9:e40311. [doi: 10.2196/40311] [Medline: 36753328]
7. Lee Y, Raviglione MC, Flahault A. Use of digital technology to enhance tuberculosis control: scoping review. J Med Internet Res 2020 Feb 13;22(2):e15727. [doi: 10.2196/15727] [Medline: 32053111]
8. Faccin M, Rusumba O, Ushindi A, et al. Data-driven identification of communities with high levels of tuberculosis infection in the Democratic Republic of Congo. Sci Rep 2022 Mar 10;12(1):3912. [doi: 10.1038/s41598-022-07633-2] [Medline: 35273212]
9. OECD/UN and ECA/AfDB. Africa's Urbanisation Dynamics 2022: The Economic Power of Africa's Cities: OECD; 2022. [doi: 10.1787/3834ed5b-en] [Medline: 38483583]

XSL•FO
RenderX

10. Prasad A, Ross A, Rosenberg P, Dye C. A world of cities and the end of TB. Trans R Soc Trop Med Hyg 2016 Mar;110(3):151-152. [doi: 10.1093/trstmh/trw004] [Medline: 26884491]
11. Folkmann EJ, Hughes MC, Khan UA, Vaezi M. Examining noncommunicable diseases using satellite imagery: a systematic literature review. BMC Public Health 2024 Oct 10;24(1):2774. [doi: 10.1186/s12889-024-20316-z] [Medline: 39390457]
12. Ford TE, Colwell RR, Rose JB, Morse SS, Rogers DJ, Yates TL. Using satellite images of environmental changes to predict infectious disease outbreaks. Emerg Infect Dis 2009 Sep;15(9):1341-1346. [doi: 10.3201/eid1509.081334] [Medline: 19788799]
13. Deonarine A, Lyons G, Lakhani C, De Brouwer W. Identifying communities at risk for COVID-19-related burden across 500 US Cities and within New York City: unsupervised learning of the coprevalence of health indicators. JMIR Public Health Surveill 2021 Aug 26;7(8):e26604. [doi: 10.2196/26604] [Medline: 34280122]
14. Nordholm AC, Joergensen A, Holm LH, et al. The impact of living conditions and health interventions on tuberculosis, Denmark, 1876 to 2022. Euro Surveill 2024 Jun;29(24):2300652. [doi: 10.2807/1560-7917.ES.2024.29.24.2300652] [Medline: 38873798]
15. Stevens FR, Gaughan AE, Linard C, Tatem AJ. Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data. PLOS ONE 2015;10(2):e0107042. [doi: 10.1371/journal.pone.0107042] [Medline: 25689585]
16. Chakravorty S, Simmons AM, Rowneki M, et al. The new Xpert MTB/RIF Ultra: improving detection of Mycobacterium tuberculosis and resistance to rifampin in an assay suitable for point-of-care testing. MBio 2017 Aug 29;8(4):e00812-17. [doi: 10.1128/mBio.00812-17] [Medline: 28851844]
17. MORAN PAP. Notes on continuous stochastic phenomena. Biometrika 1950 Jun;37(1-2):17-23. [doi: 10.1093/biomet/37.1-2.17] [Medline: 15420245]
18. Lai P, Cai W, Qu L, et al. Pulmonary tuberculosis notification rate within Shenzhen, China, 2010-2019: spatial-temporal analysis. JMIR Public Health Surveill 2024 Jun 14;10:e57209. [doi: 10.2196/57209] [Medline: 38875687]
19. Matheron G. Principles of geostatistics. Economic Geology 1963 Dec 1;58(8):1246-1266. [doi: 10.2113/gsecongeo.58.8.1246]
20. Uher V, Gajdoš P, Snášel V, Lai YC, Radecký M. Hierarchical hexagonal clustering and indexing. Symmetry (Basel) 2019;11(6):731. [doi: 10.3390/sym11060731]
21. MAIN INDICATORS: 5th Rwanda Population and Housing Census (PHC).: National Institute of Statistics of Rwanda; 2022 URL: https://statistics.gov.rw/publication/main_indicators_2022 [accessed 2023-10-06]
22. Rwanda integrated health management information system. Ministry of Health. 2021. URL: https://hmis.moh.gov.rw/ [accessed 2023-10-06]
23. CartoTB satellite - python module to find densely inhabited areas from satellite imagery. GitHub. 2024. URL: https://github.com/maurofaccin/cartotb_satellite [accessed 2024-01-01]

Abbreviations

ACF: active case-finding
CHW: community health workers
HSV: hue saturation and value
PCR: polymerase chain reaction
TB: tuberculosis
UDI: urban density index

Edited by A Mavragani; submitted 04.11.24; peer-reviewed by T Sagna, T Awofala; revised version received 18.02.25; accepted 19.02.25; published 24.04.25.

Please cite as:
Faccin M, Geenen C, Happaerts M, Ombelet S, Migambi P, André E
Analyzing Satellite Imagery to Target Tuberculosis Control Interventions in Densely Urbanized Areas of Kigali, Rwanda: Cross-Sectional Pilot Study
JMIR Public Health Surveill 2025;11:e68355
URL: https://publichealth.jmir.org/2025/1/e68355
doi:10.2196/68355

© Mauro Faccin, Caspar Geenen, Michiel Happaerts, Sien Ombelet, Patrick Migambi, Emmanuel André. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and

XSL•FO
RenderX

Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Balancing Human Mobility and Health Care Coverage in Sentinel Surveillance of Brazilian Indigenous Areas: Mathematical Optimization Approach

Juliane Fonseca Oliveira[1], PhD; Adriano O Vasconcelos[2], PhD; Andrêza L Alencar[3], PhD; Maria Célia S L Cunha[2], PhD; Izabel Marcilio[1,4], MD, PhD; Manoel Barral-Netto[1,5*], MD, PhD; Pablo Ivan P Ramos[1*], PhD

[1]Center for Data and Knowledge Integration for Health, Gonçalo Moniz Institute, Fundação Oswaldo Cruz, Parque Tecnológico da Edf. Tecnocentro, R. Mundo, 121 - sala 315 - Trobogy, Salvador, Brazil

[2]Luiz Coimbra Institute of Graduate and Engineering Research, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

[3]Department of Computer Science, Federal Rural University of Pernambuco, Recife, Brazil

[4]Bahiana School of Medicine and Public Health (EBMSP), Salvador, Brazil

[5]Medicine and Precision Public Health Laboratory, Gonçalo Moniz Institute, Fundação Oswaldo Cruz, Salvador, Brazil

[*]these authors contributed equally

Corresponding Author:
Juliane Fonseca Oliveira, PhD
Center for Data and Knowledge Integration for Health, Gonçalo Moniz Institute, Fundação Oswaldo Cruz, Parque Tecnológico da Edf. Tecnocentro, R. Mundo, 121 - sala 315 - Trobogy, Salvador, Brazil

Abstract

Background: Optimizing sentinel surveillance site allocation for early pathogen detection remains a challenge, particularly in ensuring coverage of vulnerable and underserved populations.

Objective: This study evaluates the current respiratory pathogen surveillance network in Brazil and proposes an optimized sentinel site distribution that balances Indigenous population coverage and national human mobility patterns.

Methods: We compiled Indigenous Special Health District (Portuguese: Distrito Sanitário Especial Indígena [DSEI]) locations from the Brazilian Ministry of Health and estimated national mobility routes by using the Ford-Fulkerson algorithm, incorporating air, road, and water transportation data. To optimize sentinel site selection, we implemented a linear optimization algorithm that maximizes (1) Indigenous region representation and (2) human mobility coverage. We validated our approach by comparing results with Brazil's current influenza sentinel network and analyzing the health attraction index from the Brazilian Institute of Geography and Statistics to assess the feasibility and potential benefits of our optimized surveillance network.

Results: The current Brazilian network includes 199 municipalities, representing 3.6% (199/5570) of the country's cities. The optimized sentinel site design, while keeping the same number of municipalities, ensures 100% coverage of all 34 DSEI regions while rearranging 108 (54.3%) of the 199 cities from the existing flu sentinel system. This would result in a more representative sentinel network, addressing gaps in 9 of 34 previously uncovered DSEI regions, which span 750,515 km² and have a population of 1.11 million. Mobility coverage would improve by 16.8 percentage points, from 52.4% (4,598,416 paths out of 8,780,046 total paths) to 69.2% (6,078,747 paths out of 8,780,046 total paths). Additionally, all newly selected cities serve as hubs for medium- or high-complexity health care, ensuring feasibility for pathogen surveillance.

Conclusions: The proposed framework optimizes sentinel site allocation to enhance disease surveillance and early detection. By maximizing DSEI coverage and integrating human mobility patterns, this approach provides a more effective and equitable surveillance network, which would particularly benefit underserved Indigenous regions.

(JMIR Public Health Surveill 2025;11:e69048) doi:10.2196/69048

KEYWORDS

representative sentinel surveillance; early pathogen detection; indigenous health; human mobility; surveillance network optimization; infectious disease surveillance; public health strategy; Brazil

Introduction

Sentinel surveillance, which involves systematic and regular clinical sample collection for monitoring the emergence of infectious diseases across a network of sites, is a key component of public health strategies [1]. Ideally, a sentinel surveillance network should be representative of the general population while

XSL•FO
RenderX

also prioritizing high-risk subpopulations, such as Indigenous groups [2-5].

Major global sentinel networks include the Global Influenza Surveillance and Response System established by the World Health Organization, which operates across 127 countries [6]; the European Influenza Surveillance Network [7], which relies on reports from general practitioners to monitor flu activity; and SARInet plus (Severe Acute Respiratory Infections Network) in the Americas, which is coordinated by the Pan American Health Organization [8]. These networks are essential for monitoring flu cases, detecting emerging respiratory viruses, guiding public health interventions and vaccine development, and standardizing global sentinel surveillance practices.

The effectiveness of and capability for early pathogen detection depend on local routine surveillance, which varies widely across regions and faces significant challenges [6,9]. In fact, despite the falling costs of genomic sequencing and advancements in metagenomics allowing for the monitoring of multiple pathogens and the identification of novel agents, low- and middle-income countries still struggle to scale up sentinel surveillance due to financial constraints, as well as a lack of political support for and investment in technological and scientific developments [10-13]. As a result, sentinel sites are often chosen based on convenience, leaving populations at high risk and important target regions, such as Indigenous communities, underserved. In this context, the use of alternative data streams to understand pathogen emergence and the likely pathways of disease spread may maximize the available information for improving the data-driven allocation of surveillance resources [11,13-16]. In particular, this approach enables the incorporation of epidemiological intelligence into planning the target population and geographical location for sentinel clinical sample collection. Motivated by this, we integrated human mobility data with Indigenous population coverage to refine sentinel site selection in Brazil.

The boundaries between wild landscapes and human settlements have been increasingly blurred by climate change, agricultural and urban expansion, deforestation, and landscape fragmentation. These compounding factors heighten Indigenous communities' vulnerability to infectious disease spillover, which can rapidly spread to urban areas [17]. In addition, Brazilian Indigenous populations face significant barriers to health care access, limiting their awareness of circulating diseases and disproportionately increasing their disease burden [2-4,18]. During the COVID-19 pandemic, the incidence of and mortality resulting from the disease among Indigenous and traditional people were higher when compared to those among the general population, highlighting the need for tailored public health strategies to address health disparities in these communities [19-22].

We recently showed how human mobility patterns can inform the redesign of existing sentinel networks to improve early pathogen detection [13]. However, the high risk of pathogen emergence in Indigenous communities may remain undetected when relying solely on mobility data, potentially delaying the identification of health threats in these populations. To address this gap, we aimed to build and expand on our mobility-based model by explicitly incorporating the geographical distribution of Indigenous communities in Brazil. In particular, we aimed to provide an optimized list of cities that could serve as strategic candidate sites for early pathogen detection in Brazil, while also adding an equity component into sentinel planning.

Methods

Data Sources

For our analyses, we collected data on cities within Brazilian Indigenous Special Health Districts (Portuguese: Distrito Sanitário Especial Indígena [DSEI]), which were obtained from the Brazilian Ministry of Health, along with mobility coverage data derived from the Ford-Fulkerson algorithm [13,23].

To ensure the practical applicability of our results, we also collected information on the composition of Brazil's current influenza sentinel surveillance network, which was obtained through direct communication with the Ministry of Health. We also incorporated a health attraction index, which quantifies a city's potential to attract individuals seeking health care services. This index, as estimated by the Brazilian Institute of Geography and Statistics [24], categorizes cities based on their capacity to provide low-, medium-, and high-complexity health services, and its incorporation further refined our sentinel site selection process. The complete dataset is available in a GitHub directory [25].

Optimization Problem

To identify the most suitable sentinel sites for maximizing both human mobility coverage and the representation of DSEIs, we applied an optimization approach.

Brazil's 34 DSEIs are administrative regions designed to manage health activities for Indigenous territories in an ethnically and culturally sensitive manner. DSEIs encompass 1470 of Brazil's 5570 municipalities, with partial overlap in 46 others. Overlaps occur because Indigenous territories often span multiple municipalities or states, and DSEIs are designed to align with Indigenous geography rather than political borders. In our work, we considered a DSEI to be covered in the sentinel network if at least 1 municipality completely located within a DSEI was included in the design of the network.

To define mobility coverage, we constructed an intercity mobility network, which was represented as graph G=(V, E), where the set of nodes (V) corresponds to cities, and the set of links (E) represents movement frequencies between them. Using this network, we applied the Ford-Fulkerson algorithm [13] to model potential disease spread pathways, identifying the most probable routes for pathogen transmission from a source city. This approach allowed us to determine key locations for early detection by analyzing how frequently cities appeared at different stages in these transmission paths. Further methodological details are available in our previously published paper [13]. Based on the output of the Ford-Fulkerson algorithm, we defined a city's mobility coverage over another city, as follows: city A covers city B if n% of the most likely transmission paths originating from city B include city A as their first step; n is a value between 0 and 100, and when n=100, it is considered that city A fully covers city B.

XSL•FO
RenderX

To maximize mobility coverage while ensuring that all DSEI regions are represented, we formulated this as an optimization problem. Mathematically, we let x_i be a binary variable, where x_i=1 if city i is selected for the sentinel network and x_i=0 otherwise. Further, we let ω_i represent the mobility coverage weight of city i. Additionally, we let R_j denote the set of cities within DSEI region j and let N be the total number of cities to be selected for the sentinel network. The optimization function was then defined as follows:

$$\max(\sum_i \omega_i x_i)$$

subject to

1. $\sum_i x_i = N$, that is, select N cities to compose the sentinel.

2. $\sum_{j \in R_j} x_j \geq 1$, for all $j \in 1,2,...,34$, that is, for each DSEI region j, at least 1 city in R_j is selected.

The optimization problem was implemented in Python and solved by using the Python Universal Linear Programming library [25].

Practical Applicability

Brazil's flu sentinel network primarily requires sites in capital cities, with 1 site per 500,000 inhabitants, except in the south region, where 1 site is implemented for every 300,000 inhabitants, regardless of metropolitan status. Additionally, sample collection occurs in intensive care unit services that cover approximately 10% of the available intensive care unit beds in each municipality, ensuring coverage across age groups [26].

To assess the practical applicability of our proposed sentinel network, we compared it with Brazil's existing flu sentinel network to evaluate improvements in mobility and Indigenous region coverage. We also assessed whether the newly identified cities serve as key access points for health care services across different levels of complexity. This evaluation reinforced that the proposed sentinel locations align with urban centers that attract individuals seeking health care, making them strategic candidates for genomic sample collection.

Ethical Considerations

This study used publicly available secondary data sources. No personal data were collected or analyzed, ensuring the complete anonymity of all data. Ethical approval was not required, as the dataset consisted of anonymized, aggregated information, following resolutions 466/2012 and 510/2016 (article 1, sections III and V) from the National Health Council (Conselho Nacional de Saúde), Brazil. All analyses were carried out in accordance with relevant ethical guidelines for data protection and research integrity.

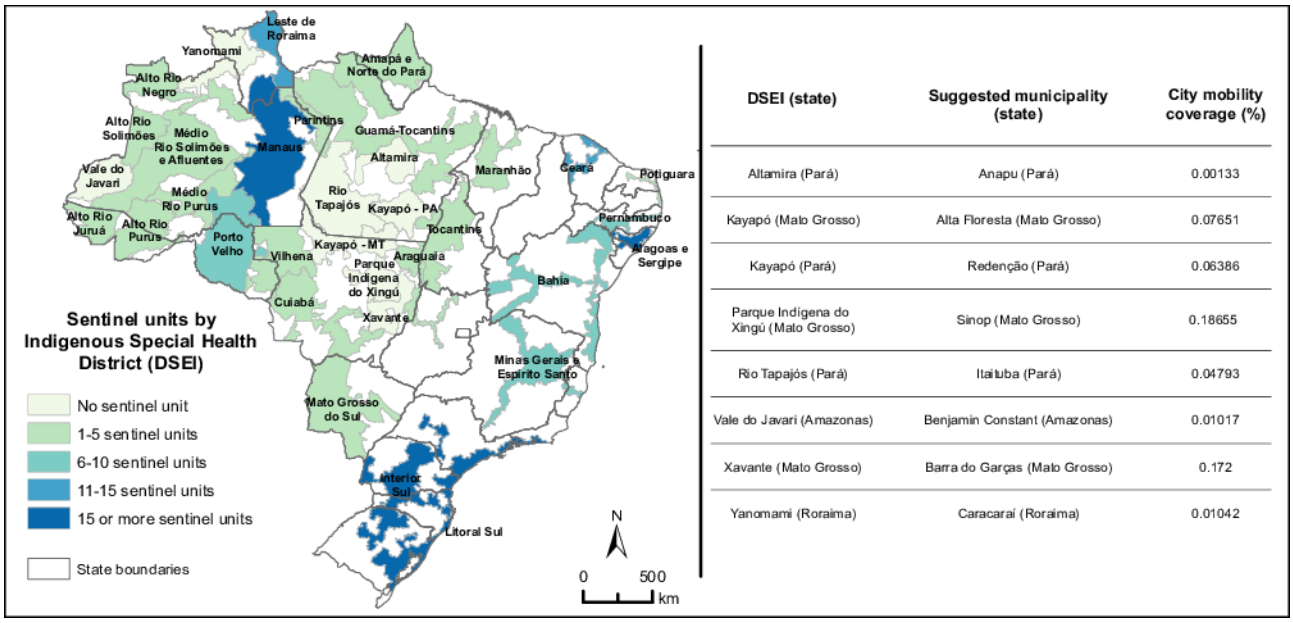
Results

The current network includes 310 sentinel sites in 199 municipalities, representing only 3.6% (199/5570) of the country's cities. To compose a more representative sentinel network for the country, we optimized the selection of 199 cities—the same number considered in Brazil's current flu sentinel network. In the optimized network, we achieved 100% DSEI coverage by selecting the cities (87/199, 43.7%) that are fully inside of DSEI regions. In addition, 69.2% of the country's mobility pattern is covered by these cities, translating to 6,078,747 paths (of 8,780,046 total paths) originating from each city in the country.

In the current network, 53.8% (107/199) of the municipalities overlap 25 DSEIs (Figure 1). However, the following nine DSEIs lack any sentinel coverage: Altamira (Pará), Kayapó (Mato Grosso), Kayapó (Pará), Médio Rio Purus (Amazonas), Parque Indígena do Xingu (Mato Grosso), Rio Tapajós (Pará), Vale do Javari (Amazonas), Xavante (Mato Grosso), and Yanomami (Roraima). These regions span more than 750,515 km² and are home to more than 1.11 million people. In addition, Brazil's current flu sentinel network covers only 52.4% (4,598,416 of 8,780,046 paths) of mobility patterns (considering air, road, and waterway transportation modes).

XSL•FO
RenderX

Figure 1. Left: distribution of active surveillance sites within DSEIs in Brazil (2025). This map illustrates the current coverage of DSEIs by the national influenza surveillance network, highlighting the geographical distribution of sentinel sites across Indigenous regions. Colored areas represent DSEI regions according to the number of sentinel units. Right: proposed optimized list of municipalities within underserved Indigenous regions not currently covered by the Brazilian flu sentinel network, along with their corresponding mobility coverage. This list, which was generated from the optimization process, aims to improve DSEI coverage while incorporating Brazil's national human mobility patterns. The proposed cities guide resource allocation to enhance surveillance effectiveness in the country's underserved Indigenous areas. DSEI: Distrito Sanitário Especial Indígena.



DSEI (state)	Suggested municipality (state)	City mobility coverage (%)
Altamira (Pará)	Anapu (Pará)	0.00133
Kayapó (Mato Grosso)	Alta Floresta (Mato Grosso)	0.07651
Kayapó (Pará)	Redenção (Pará)	0.06386
Parque Indígena do Xingú (Mato Grosso)	Sinop (Mato Grosso)	0.18655
Rio Tapajós (Pará)	Itaituba (Pará)	0.04793
Vale do Javari (Amazonas)	Benjamin Constant (Amazonas)	0.01017
Xavante (Mato Grosso)	Barra do Garças (Mato Grosso)	0.172
Yanomami (Roraima)	Caracaraí (Roraima)	0.01042

By rearranging 108 cities in Brazil's flu sentinel network—keeping 91 cities that are already included in the current network—we achieved 100% DSEI coverage and increased national mobility coverage by 16.8 percentage points (from 52.4% to 69.2%). Figure 1 shows the list of municipalities that would ensure coverage for DSEIs without sentinel sites in Brazil's flu sentinel network.

Finally, we analyzed the movement of people seeking health care in the cities selected for the optimized sentinel network. None of the identified cities had a low health care attraction index. The majority (192/199, 96.5%) of cities serve as hubs for high-complexity health care services, while 3.5% (7/199) are classified as medium-complexity hubs. Among cities within DSEIs, 33% (29/87) have a high health care attraction index and are not currently covered by Brazil's flu sentinel network. Additionally, 67% (58/87) have a medium attraction index, with 4 lacking sentinel coverage.

Discussion

In this study, we present an optimized list of 199 cities that would enhance early pathogen detection by strategically balancing Indigenous area coverage and human mobility patterns. Our selection ensures 100% coverage of Indigenous regions through 87 key cities while also capturing 69.2% of the country's human mobility flows. Additionally, the feasibility of implementing sentinel surveillance in these locations is supported by their high health care attraction index, which indicates their role as hubs for complex health care services.

Our findings introduce a novel approach that integrates human mobility and Indigenous population coverage to optimally rearrange Brazil's current flu sentinel network locations for clinical sample collection. By focusing on the historically excluded Indigenous communities, the proposed approach aligns with expert recommendations to enhance surveillance in underserved areas [2-4,10], thereby ultimately protecting vulnerable communities while improving Brazil's ability to detect and respond to emerging health threats.

By ensuring that all DSEIs are represented by at least 1 sentinel site, we addressed surveillance gaps in 9 previously uncovered regions that, collectively, are home to over 1 million people. This outcome is particularly significant, given the increased disease risks faced by Indigenous populations residing at the wildlife-urban interface. Our mobility-based optimization, which was combined with a restriction to 199 sentinel cities, reduced the number of cities within DSEIs when compared to Brazil's current flu sentinel network. This highlights the following two key considerations: (1) some DSEI regions covered in the existing network may be overserved, and (2) the total number of cities selected for sample collection may be insufficient to account for Brazil's geographic and epidemiological diversity, potentially limiting early pathogen detection. Nevertheless, maintaining or expanding routine sentinel surveillance is paramount to establishing a baseline of case incidence and tracking circulating viruses [1], which enable the early detection of outbreaks and the introduction of emergent pathogens. For instance, the absence of baseline testing likely contributed to the delay in the recognition of the Zika virus outbreak in the Americas, which peaked in 2015 to 2016 but probably spread undetected since late 2013 [27,28].

The optimized sentinel network achieved 69.2% mobility coverage—an improvement over Brazil's current flu sentinel network. This shows that incorporating Indigenous population criteria resulted in the optimized network's mobility coverage being only slightly lower than the 70% we reported [13] for a network based solely on mobility patterns. This underscores

XSL•FO
RenderX

the need to examine the key epidemiological priorities (which may include socioeconomic deprivation and regions with a high risk of climate-related disasters) that should be taken into account when designing a representative sentinel network for early pathogen detection, as these factors may impact coverage metrics. Consequently, identifying an optimal balance between the minimum number of cities required for sentinel surveillance and the diversity of epidemiological priorities remains a critical area for future research.

Establishing a sentinel surveillance network, including a pipeline from sample collection to laboratory analysis, demands substantial resources, especially in regions with limited infrastructure. To address this challenge, we strategically centered our focus around DSEIs—an already in-place health administrative structure in Brazil that provides primary care to Indigenous populations and articulates with other networks in the Unified Health System to guarantee access to medium- and high-complexity services. Additionally, by evaluating the health attraction index and taking into account the outdated requirement for sentinel network implementation that was established by the Brazilian Ministry of Health [26], we found that the newly selected cities already offer medium- to high-complexity health care services—an indication that may support genomic sample collection. Leveraging this existing infrastructure not only facilitates routine surveillance but also strengthens overall health care capacity.

This work contributes novel insights to the literature by addressing the need for more effective selection of epidemiologically relevant criteria for early pathogen detection, while promoting a more equitable distribution of sentinel surveillance sites. To our knowledge, no previous study has explicitly optimized both underserved community coverage and human mobility patterns within a sentinel network framework. Existing research may only indicate standard criteria, such as gender, ethnicity, age, population coverage, and incidence of a specific disease [1,5,10], or rely on a single data source with anticipatory potential [29], without assessing whether these criteria are sufficient to truly maximize the chances of early threat detection. Our approach bridges this gap by offering a data-driven framework to enhance public health preparedness, and our findings may encourage the scientific community to discuss more about the importance of setting up and optimizing sentinel surveillance.

Our work has some limitations. Despite the significance of our findings, the data used in this study primarily identify the geographical regions encompassing Indigenous populations, without capturing other factors, such as climate-related risks or socioeconomic disparities, that may worsen these communities' vulnerability to emerging new or unknown pathogens. Further research is needed to refine surveillance strategies tailored to these populations. Additionally, as we previously highlighted [13], access to up-to-date mobility data remains a key challenge. Finally, we emphasize that this study represents a step toward more data-driven approaches in designing sentinel surveillance networks, demonstrating how different epidemiological criteria (exemplified by the two considered herein) can be integrated into the optimization process to enhance disease monitoring.

To conclude, our results indicate that our proposed framework can be used to effectively guide the optimization of sentinel site allocation, and the framework can be further validated for use during new emergencies or in the monitoring of flu syndrome in Brazil. Our results contribute to increasing the likelihood of early pathogen detection without requiring an expansion in the number of sentinel site locations that have already been put into place by the Brazilian Ministry of Health, thereby addressing the critical challenge of limited funding—a well-documented barrier to improving global epidemiological surveillance, especially in low- and middle-income countries.

Acknowledgments

The findings reported are part of the Alert-Early System of Outbreaks with Pandemic Potential (ÆSOP) [30], an initiative under development by Brazil's Fundação Oswaldo Cruz (Fiocruz) and the Federal University of Rio de Janeiro, with financial support from the Rockefeller Foundation's Health Initiative (grant 2023-PPI-007 awarded to MBN).

Data Availability

All datasets collected and generated during this study, along with the corresponding code to ensure reproducibility of our results, are available in a GitHub repository [25].

Authors' Contributions

Conceptualization: JFO, MBN, PIPR
Formal analysis: ALA, AOV, JFO, MCSLC, PIPR
Project administration: IM, MBN, PIPR
Resources: MBN
Supervision: JFO, MBN, PIPR
Writing – original draft: IM, JFO, MBN, PIPR
Writing – review & editing: IM, JFO, MBN, PIPR

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Global epidemiological surveillance standards for influenza. : World Health Organization; 2013 URL: https://apps.who.int/iris/bitstream/handle/10665/311268/9789241506601-eng.pdf [accessed 2025-03-20]
2. Cardoso AM, Resende PC, Paixao ES, et al. Investigation of an outbreak of acute respiratory disease in an indigenous village in Brazil: contribution of Influenza A(H1N1)pdm09 and human respiratory syncytial viruses. PLoS One 2019 Jul 8;14(7):e0218925. [doi: 10.1371/journal.pone.0218925] [Medline: 31283762]
3. La Ruche G, Tarantola A, Barboza P, et al. The 2009 pandemic H1N1 influenza and indigenous populations of the Americas and the Pacific. Euro Surveill 2009 Oct 22;14(42):19366. [doi: 10.2807/ese.14.42.19366-en] [Medline: 19883543]
4. Groom AV, Jim C, Laroque M, et al. Pandemic influenza preparedness and vulnerable populations in tribal communities. Am J Public Health 2009 Oct;99 Suppl 2(Suppl 2):S271-S278. [doi: 10.2105/AJPH.2008.157453] [Medline: 19461107]
5. Leston M, Elson WH, Watson C, et al. Representativeness, vaccination uptake, and COVID-19 clinical outcomes 2020-2021 in the UK Oxford-Royal College of General Practitioners Research and Surveillance network: cohort profile summary. JMIR Public Health Surveill 2022 Dec 19;8(12):e39141. [doi: 10.2196/39141] [Medline: 36534462]
6. Ziegler T, Moen A, Zhang W, Cox NJ. Global Influenza Surveillance and Response System: 70 years of responding to the expected and preparing for the unexpected. Lancet 2022 Sep 24;400(10357):981-982. [doi: 10.1016/S0140-6736(22)01741-X] [Medline: 36154679]
7. de Fougerolles TR, Damm O, Ansaldi F, et al. National influenza surveillance systems in five European countries: a qualitative comparative framework based on WHO guidance. BMC Public Health 2022 Jun 9;22(1):1151. [doi: 10.1186/s12889-022-13433-0] [Medline: 35681199]
8. Rondy M, Azziz-Baumgartner E, Kondor R, Palekar RS, Zhang W, Vicari AS. 10 years of Severe Acute Respiratory Infections Network (SARInet plus): accomplishments and way forward. J Infect Dis 2025 Mar 10;231(Supplement_2):S87-S89. [doi: 10.1093/infdis/jiaf039] [Medline: 39891533]
9. Baltrusaitis K, Vespignani A, Rosenfeld R, Gray J, Raymond D, Santillana M. Differences in regional patterns of influenza activity across surveillance systems in the United States: comparative evaluation. JMIR Public Health Surveill 2019 Sep 14;5(4):e13403. [doi: 10.2196/13403] [Medline: 31579019]
10. Pei S, Teng X, Lewis P, Shaman J. Optimizing respiratory virus surveillance networks using uncertainty propagation. Nat Commun 2021 Jan 11;12(1):222. [doi: 10.1038/s41467-020-20399-3] [Medline: 33431854]
11. Tsui JLH, Zhang M, Sambaturu P, et al. Toward optimal disease surveillance with graph-based active learning. Proc Natl Acad Sci U S A 2024 Dec 24;121(52):e2412424121. [doi: 10.1073/pnas.2412424121] [Medline: 39700149]
12. Ramos PIP, Marcilio I, Bento AI, et al. Combining digital and molecular approaches using health and alternate data sources in a next-generation surveillance system for anticipating outbreaks of pandemic potential. JMIR Public Health Surveill 2024 Jan 9;10(1):e47673. [doi: 10.2196/47673] [Medline: 38194263]
13. Oliveira JF, Alencar AL, Cunha MCLS, et al. Human mobility patterns in Brazil to inform sampling sites for early pathogen detection and routes of spread: a network modelling and validation study. Lancet Digit Health 2024 Aug;6(8):e570-e579. [doi: 10.1016/S2589-7500(24)00099-2] [Medline: 39059889]
14. El-Khatib Z, Richter L, Reich A, Benka B, Assadian O. Implementation of a surveillance system for severe acute respiratory infections at a tertiary care hospital in Austria: protocol for a retrospective longitudinal feasibility study. JMIR Res Protoc 2023 Aug 3;12:e47547. [doi: 10.2196/47547] [Medline: 37535414]
15. Gu X, Watson C, Agrawal U, et al. Postpandemic sentinel surveillance of respiratory diseases in the context of the World Health Organization mosaic framework: protocol for a development and evaluation study involving the English primary care network 2023-2024. JMIR Public Health Surveill 2024 Apr 3;10:e52047. [doi: 10.2196/52047] [Medline: 38569175]
16. de Lusignan S, Lopez Bernal J, Byford R, et al. Influenza and respiratory virus surveillance, vaccine uptake, and effectiveness at a time of cocirculating COVID-19: protocol for the English primary care sentinel system for 2020-2021. JMIR Public Health Surveill 2021 Feb 19;7(2):e24341. [doi: 10.2196/24341] [Medline: 33605892]
17. Gibb R, Redding DW, Chin KQ, et al. Zoonotic host diversity increases in human-dominated ecosystems. Nature 2020 Aug;584(7821):398-402. [doi: 10.1038/s41586-020-2562-8] [Medline: 32759999]
18. Browne AJ, Smye VL, Rodney P, Tang SY, Mussell B, O'Neil J. Access to primary care from the perspective of Aboriginal patients at an urban emergency department. Qual Health Res 2011 Mar;21(3):333-348. [doi: 10.1177/1049732310385824] [Medline: 21075979]
19. Vittor AY, Laporta GZ, Sallum MAM, Walker RT. The COVID-19 crisis and Amazonia's indigenous people: implications for conservation and global health. World Dev 2021 Sep;145:105533. [doi: 10.1016/j.worlddev.2021.105533] [Medline: 36570383]
20. Salvalaggio G, Hyshka E, Brown C, et al. A comparison of the COVID-19 response for urban underserved patients experiencing healthcare transitions in three Canadian cities. Can J Public Health 2022 Dec;113(6):846-866. [doi: 10.17269/s41997-022-00651-7] [Medline: 35771364]
21. Fellows M, Paye V, Alencar A, et al. Under-reporting of COVID-19 cases among indigenous peoples in Brazil: a new expression of old inequalities. Front Psychiatry 2021 Apr 12;12:638359. [doi: 10.3389/fpsyt.2021.638359] [Medline: 33912084]

XSL•FO
RenderX

22. Pereira FAC, Filho FMHS, de Azevedo AR, et al. Profile of COVID-19 in Brazil-risk factors and socioeconomic vulnerability associated with disease outcome: retrospective analysis of population-based registers. BMJ Glob Health 2022 Dec;7(12):e009489. [doi: 10.1136/bmjgh-2022-009489] [Medline: 36517111]
23. Distrito Sanitário Especial Indígena (DSEI) [Article in Portuguese]. Ministério da Saúde. URL: https://www.gov.br/saude/pt-br/composicao/sesai/estrutura/dsei [accessed 2025-03-20]
24. REGIC - Regiões de Influência das Cidades [Article in Portuguese]. Instituto Brasileiro de Geografia e Estatística. URL: https://www.ibge.gov.br/geociencias/cartas-e-mapas/redes-geograficas/15798-regioes-de-influencia-das-cidades.html?=&t=o-que-e [accessed 2025-03-20]
25. Oliveira JF, Vasconcelos AO, Alencar AL, et al. Surveillance_Network_Indigenous_Mobility. GitHub. URL: https://github.com/Julian-sun/Surveillance_Network_Indigenous_Mobility.git [accessed 2025-03-20]
26. Portaria no 183, de 30 de Janeiro de 2014 [Article in Portuguese]. Ministério da Saúde. 2014. URL: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2014/prt0183_30_01_2014.html [accessed 2025-03-20]
27. Campos GS, Bandeira AC, Sardi SI. Zika virus outbreak, Bahia, Brazil. Emerg Infect Dis 2015 Oct;21(10):1885-1886. [doi: 10.3201/eid2110.150847] [Medline: 26401719]
28. Faria NR, Quick J, Claro IM, et al. Establishment and cryptic transmission of Zika virus in Brazil and the Americas. Nature 2017 Jun 15;546(7658):406-410. [doi: 10.1038/nature22401] [Medline: 28538727]
29. Cerqueira-Silva T, Oliveira JF, de Araújo Oliveira V, et al. Early warning system using primary health care data in the post-COVID-19 pandemic era: Brazil nationwide case-study. Cad Saude Publica 2024 Dec 20;40(11):e00010024. [doi: 10.1590/0102-311XEN010024] [Medline: 39775767]
30. ÆSOP | Alert-Early System of Outbreaks with Pandemic Potential [Article in Portuguese]. Centro de Integração de Dados e Conhecimentos para Saúde. URL: https://cidacs.bahia.fiocruz.br/aesop/ [accessed 2025-03-26]

Abbreviations

DSEI: Distrito Sanitário Especial Indígena
SARInet plus: Severe Acute Respiratory Infections Network

Edited by A Mavragani; submitted 20.11.24; peer-reviewed by H Kim, R Saadane; revised version received 18.02.25; accepted 19.02.25; published 01.04.25.

Please cite as:
Oliveira JF, Vasconcelos AO, Alencar AL, Cunha MCSL, Marcilio I, Barral-Netto M, P Ramos PI
Balancing Human Mobility and Health Care Coverage in Sentinel Surveillance of Brazilian Indigenous Areas: Mathematical Optimization Approach
JMIR Public Health Surveill 2025;11:e69048
URL: https://publichealth.jmir.org/2025/1/e69048
doi:10.2196/69048

© Juliane Fonseca Oliveira, Adriano O Vasconcelos, Andrêza L Alencar, Maria Célia S L Cunha, Izabel Marcilio, Manoel Barral-Netto, Pablo Ivan P Ramos. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 1.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Statistical Relationship Between Wastewater Data and Case Notifications for COVID-19 Surveillance in the United States From 2020 to 2023: Bayesian Hierarchical Modeling Approach

Masahiko Haraguchi[1], PhD; Fayette Klaassen[1], PhD; Ted Cohen[2], MD, DPH; Joshua A Salomon[3], PhD; Nicolas A Menzies[1], PhD

[1]Department of Global Health and Population, Harvard T H Chan School of Public Health, 677 Huntington Ave, Boston, MA, United States
[2]Department of Epidemiology of Microbial Diseases and Public Health Modeling Unit, Yale School of Public Health, New Haven, CT, United States
[3]Department of Health Policy, Stanford University, Stanford, CA, United States

Corresponding Author:
Masahiko Haraguchi, PhD
Department of Global Health and Population, Harvard T H Chan School of Public Health, 677 Huntington Ave, Boston, MA, United States

Abstract

Background: During the COVID-19 pandemic, several US jurisdictions began to regularly report levels of SARS-CoV-2 in wastewater as a proxy for SARS-CoV-2 incidence. Despite the promise of this approach for improving COVID-19 situational awareness, the degree to which wastewater surveillance data agree with other data has varied, and better evidence is needed to understand the situations in which wastewater surveillance data track closely with traditional surveillance data.

Objective: In this study, we quantified the statistical relationship between wastewater data and traditional case-based surveillance data for multiple jurisdictions.

Methods: We collated data on wastewater SARS-CoV-2 RNA levels and COVID-19 case reports from July 2020 to March 2023 for 107 counties representing a range in terms of geographic location, population size, and urbanicity. For these counties, we used Bayesian hierarchical regression modeling to estimate the statistical relationship between wastewater data and reported cases, allowing for variation in this relationship across counties. We compared different model structural approaches and assessed how the strength of the estimated relationships varied between settings and over time.

Results: Our analyses revealed a strong positive relationship between wastewater data and COVID-19 cases for the majority of locations, with a median correlation coefficient between observed and predicted cases of 0.904 (IQR 0.823 - 0.943). In total, 23/107 counties (21.5%) had correlation coefficients below 0.8, and 3/107 (2.8%) had values below 0.6. Across locations, the COVID-19 case rate associated with a given level of wastewater SARS-CoV-2 RNA concentration declined over the study period. Counties with greater population size (P<.001) and higher levels of urbanicity (P<.001) had stronger concordance between wastewater data and COVID-19 cases. Measures of model fit, and relationships with urbanicity and population size, were robust to sensitivity analyses in which we varied the time period of analysis and the sample of counties used for model fitting.

Conclusions: In a sample of 107 US counties, wastewater surveillance had a close relationship with COVID-19 cases reported for the majority of locations, with these relationships found to be stronger in counties with greater population size and urbanicity. In situations where routine COVID-19 surveillance data are less reliable, wastewater surveillance may be used to track local SARS-CoV-2 incidence trends.

(JMIR Public Health Surveill 2025;11:e68213) doi:10.2196/68213

KEYWORDS

wastewater epidemiology; COVID-19 surveillance; environmental surveillance; SARS-CoV-2; Bayesian modeling; pandemic; wastewater; surveillance; USA; United States of America; Bayesian; hierarchical model; regression model; coronavirus; respiratory; infectious; pulmonary; COVID-19; environment; sewer; emission; sewerage

Introduction

Background

The practice of using wastewater data to track pathogens has gained significant interest as an innovative method of infectious disease surveillance, with nearly 80% of the US population connected to public wastewater systems [1,2]. During the COVID-19 crisis, local health agencies started to gather and report data on COVID-19 concentrations in wastewater, using

XSL•FO
RenderX

the trends in these data as an indirect indicator of SARS-CoV-2 transmission patterns.

Tracking the presence and concentration of pathogens in wastewater, a passive method of environmental surveillance, has been used for several decades to track infectious diseases, such as polio [3], gastroenteritis [4], hepatitis E [5], and acute diarrhea [6], among others. However, the COVID-19 pandemic accelerated interest in this approach worldwide [7], as it offers multiple benefits as a complement to traditional surveillance systems. First, it can serve as an early warning system for SARS-CoV-2 transmission, sensitive to asymptomatic and presymptomatic cases 4-10 days prior to clinical testing signals [8]. Second, it can monitor community-level transmission when implemented at downstream locations such as wastewater treatment plants [9]. The fact that it does not require individual testing circumvents the challenges created by variable supply of and demand for COVID-19 diagnostic testing (as has been observed over the pandemic), and the decline in reporting of test results [10]. Third, wastewater surveillance programs with specific methods can track virus variants in the early stage of their evolution, allowing for early identification of emerging variants [9,11,12]. Finally, it provides an early indicator of epidemiological changes (as compared to hospitalization and death data), so that mitigation and other response measures can be deployed more rapidly. With these advantages, tracking wastewater COVID-19 data is a potentially powerful tool for COVID-19 surveillance.

Despite the potential advantages of SARS-CoV-2 wastewater surveillance, significant challenges remain. Existing studies have generally considered the relationship between reported cases and wastewater metrics at a limited number of locations, such as university dorms [13], nursing homes [14], university campuses [15], and municipalities (eg, Oklahoma City, Oklahoma [16] and Louisville, Kentucky [17]). While these studies provide valuable insights, they often do not fully account for heterogeneities across different geographical locations. Furthermore, the time periods covered by existing studies are limited. For example, Xiao, Wu [18] associated clinical case data with wastewater data within 3 Massachusetts counties from March 2020 through May 2021. Similarly, Weidhaas and Aanderud [19] analyzed 9 weeks of wastewater and COVID-19 case data related to 10 wastewater treatment facilities in Utah. Notably, reported cases themselves are not a perfect indicator for true infection situations, given that they depend on various factors such as testing availability and access as well as the number of asymptomatic infections. As such, while quantifying the relationship between reported cases and wastewater metrics provides valuable insights, it should be noted as the next-best alternative. Studies that compare the performance of COVID-19 wastewater surveillance data across sewersheds over extended time frames remain scarce [20]. Such research is valuable for establishing the statistical basis for real-time trend analysis and describing the conditions under which wastewater surveillance performs well.

Objectives

This study explored the quantitative relationship between SARS-CoV-2 wastewater surveillance data and reported COVID-19 diagnoses across multiple sewer-sheds over the initial years of the COVID-19 pandemic. Using weekly aggregated wastewater and case report data for 107 US counties over 2020 - 2023, we used Bayesian hierarchical modeling to establish the statistical relationship between these 2 data sources, describe changes in these relationships over time and across locations, and identify how the strength of these relationships varied systematically by county characteristics. A visual summary of the study design and findings is provided in Multimedia Appendix 1.

Methods

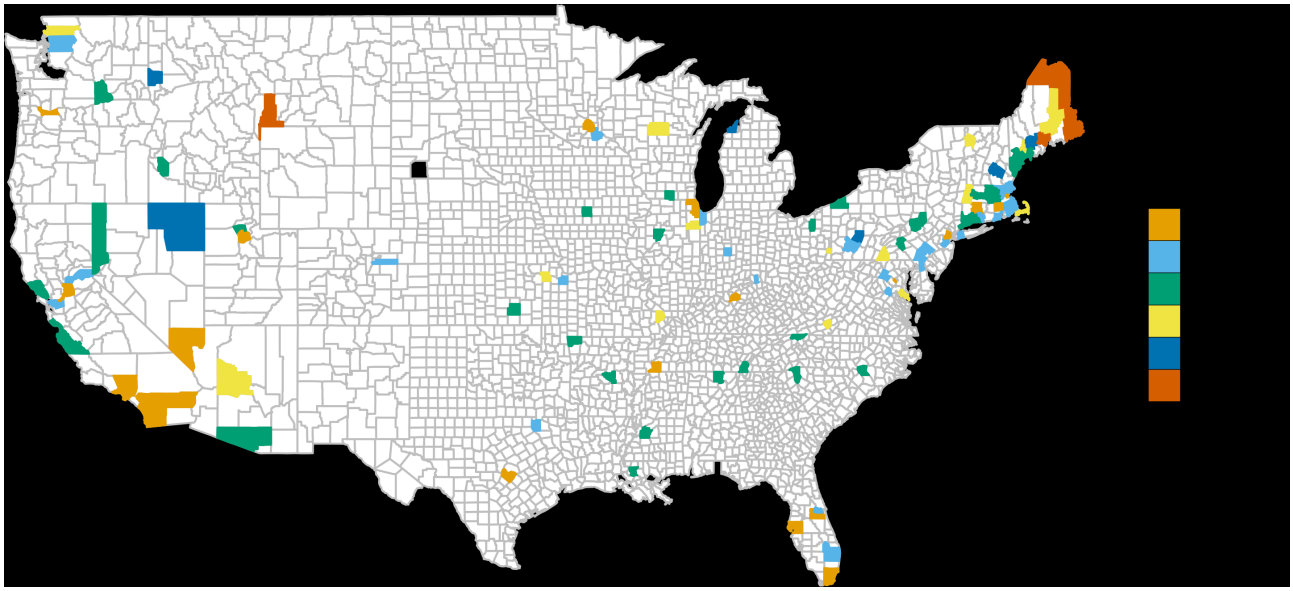
Data Sources

For the study period July 1, 2020, to March 1, 2023, we collated county-level SARS-CoV-2 wastewater surveillance data reported by Biobot Analytics, including 254 counties covering approximately 30% of the US population. These data represent RNA copies per milliliter, normalized by the concentration of pepper mild mottle virus to correct for variability in fecal content, which is influenced by environmental factors such as stormwater [21]. Surveillance data on county-level weekly COVID-19 reported case totals were extracted from the COVID-19 data repository in the Center for Systems Science and Engineering at Johns Hopkins University. This repository aggregates data from a large number of national and subnational US sources, including the CDC, state and country health departments, and nongovernment COVID-19 surveillance projects [22]. Aggregated and harmonized data represent all COVID-19 cases diagnosed and reported to these sources for each reporting period, and therefore will reflect variation in the coverage of COVID-19 testing and completeness of reporting over time and between locations. Both the Biobot Analytics data and the COVID-19 data repository at Johns Hopkins University consist of aggregated and deidentified data. No individual-level or personally identifiable information was used in the analysis.

Due to varied implementation of wastewater surveillance operations, wastewater data were not available for all county-weeks. We restricted the analysis to counties with a minimum of 50 weeks of available wastewater data during the study period, which resulted in 107 counties being included, covering a range of geographic areas within the United States, and with periods of data incompleteness for the majority of counties. We assessed the correlation between wastewater metrics and case totals with different time lags (0 wk, 1 wk, 2 wk, etc), and found that a 0 week time lag has the highest correlation. Therefore, we adopted this for the analysis. We grouped counties into 6 ordinal urbanicity categories as defined by the National Center for Health Statistics (NCHS). Categories ranged from NCHS category 1 (large central metro) as the most urban to NCHS category 6 (noncore) as the most rural. Figure 1 shows the geographic distribution of the 107 counties included in the analysis, coded by NCHS urbanicity category.

XSL•FO
RenderX

Figure 1. This shows geographic location and National Center for Health Statistics urbanicity category for the 107 counties included in the analysis. NCHS: National Center for Health Statistics.



Hierarchical Regression Models

Using the wastewater and COVID-19 case data, we constructed hierarchical Bayesian regression models to capture key features of each data source, allowing for differences in the estimated relationship between these data over time and between modeled counties. Accounting for temporal variations is vital as the dynamics of the pandemic changed over time due to factors such as the emergence of new variants, public health interventions, vaccination rollouts, and population immunity. Also, allowing for spatial variations accounts for local differences, such as vaccination coverage, socioeconomic and demographic characteristics, and variation in COVID-19 testing and reporting practices. As all of them influence the relationship between wastewater metrics and reported cases, a model that can account for the complex and evolving nature of the pandemic is critical. Namely, we modelled the weekly COVID-19 case reports for each county using a negative binomial likelihood, allowing for over-dispersion in these data:

$$(1) Yit \sim NegBin(\mu it, \phi)$$

$$(2) \mu it = \exp(\alpha t + \beta 0i + \beta 1i * Xit) * ni$$

In Equation 1, Yit represents reported COVID-19 cases for county i and week t. We used an alternative parameterization of the negative binomial in which μit parameterizes the mean of the likelihood and ϕ parameterizes the extra-Poisson variation. The mean was specified as a function of αt, a time-varying coefficient given a random-walk prior; β0i, a county-specific intercept; and β1i,a county-specific slope term applied to Xit, the demeaned wastewater COVID-19 concentration value for each county and week (Equation 2). Both β0i and β1i were specified as random effects to pool information across counties. ni represented the population of each county, as reported by the US Census Bureau. We fit this model to the COVID-19 case and wastewater data using the RStan package in R (Stan Development Team) [23]. The prior distributions used in this model, including weakly informative priors for fixed effects and random effects coefficients, were selected based on their widespread use in Bayesian modeling. These priors serve to regularize parameter estimates without excessively constraining the model [24]. Sensitivity checks confirmed that the priors do not strongly influence the posterior estimates. Additional details on the specification of this regression model, including prior distributions, are provided in the Multimedia Appendix 2.

Analysis of Fitted Models

We used several approaches to assess model fit. First, we visually compared the COVID-19 case time series for each county to the fitted values from the regression model. Second, we calculated a quantitative measure of model fit: the median absolute deviation (MAD). We estimated this value for each county and used them to describe the overall level of model fit and how this varied across counties. We used the MAD to investigate whether the strength of estimated relationships differed systematically as a function of county characteristics (urbanicity and population size), examining these relationships visually and via univariable and multivariable regression models. Finally, we calculated the correlation coefficient between modeled values and raw case totals as a simple summary measure of model fit.

Coefficient Estimates

We used the fitted values of αt (the temporal trend in the regression model) to understand how the relationship between wastewater and COVID-19 case totals varied over the study period (from July 2020 to March 2023). We used the fitted values of β0i and β1i to understand how the relationship between wastewater concentration and COVID-19 case totals varied within each county.

Sensitivity Analyses

We conducted 2 sets of sensitivity analyses to evaluate the robustness of the model results. First, we divided the dataset into 3 equal 11-month time periods (early, mid, and late pandemic). We refit the model for each of these time periods, and compared the goodness-of-fit metrics (MAD) and analyses of county characteristics (urbanicity and population size) from

XSL•FO
RenderX

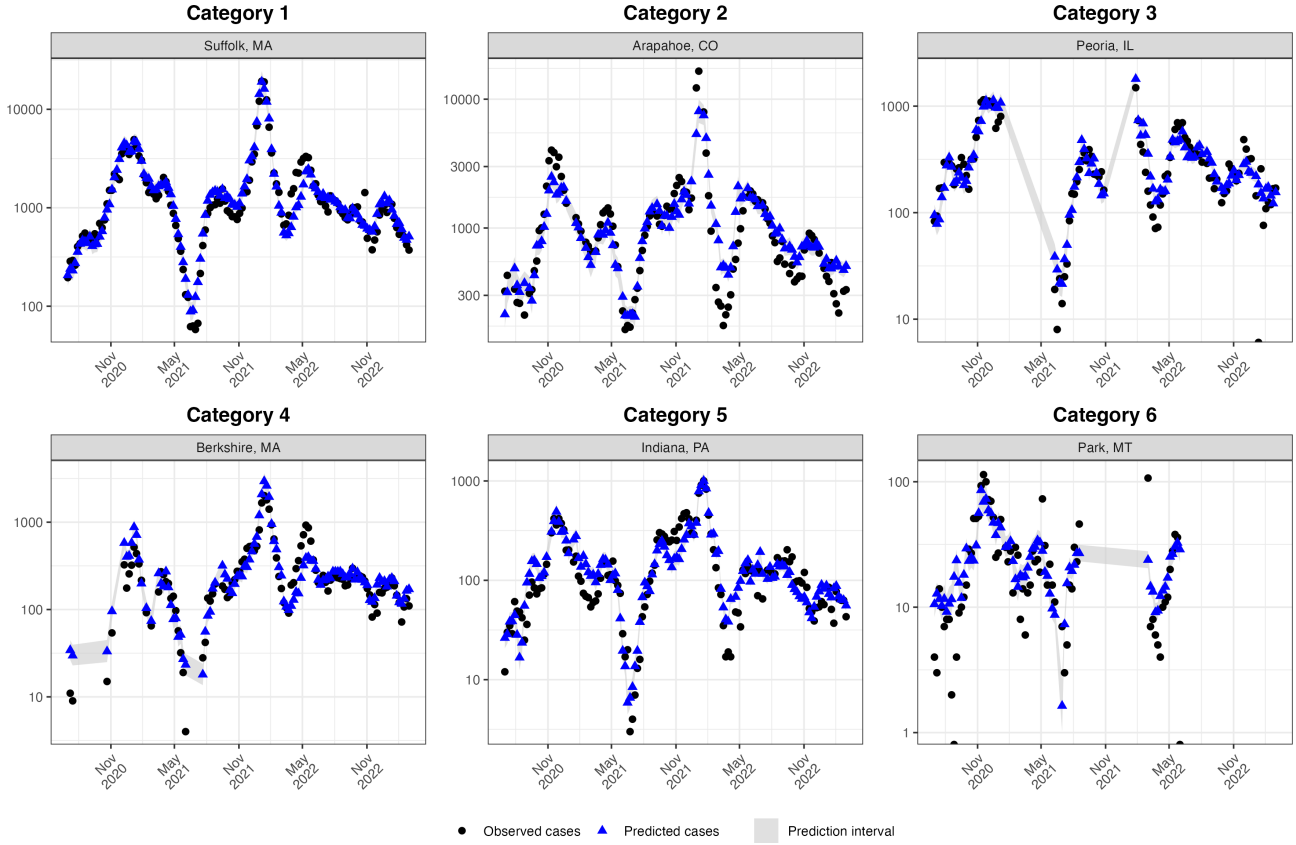
these 3 models to the results of the main analysis. Second, we identified counties with high data completeness (at least 100 wk of data out of a total of 140 wk in the study period) and refit the model using only these counties, then compared these results to those of the main analysis.

Results

Fitted Relationship Between COVID-19 Cases and Wastewater Concentration

For the majority of modeled locations, we estimated a relationship between wastewater concentration and weekly COVID-19 cases, indicating that wastewater concentration serves as a useful predictor of case trends. Figure 2 shows the temporal trend in reported COVID-19 cases and fitted model estimates for each of the 6 example countries, representing a range in terms of urbanicity and population size, which are often correlated but not identical. For each of these example countries the fitted model values (blue symbols) follow the empirical case data (black symbols) closely, with occasional deviations (eg, late 2021 - early 2022 estimates for Arapahoe County, CO). Figures for other counties are available in Figure S1 in Multimedia Appendix 2. We also calculated correlation coefficients comparing observed and predicted values of COVID-19 case counts. Across counties, the median of these correlation coefficients was 0.904 (IQR 0.823-0.943). In total, 23/107 counties (21.5% of the sample) had correlation coefficients below 0.8, and 3 had values below 0.6.

Figure 2. Comparison of observed and predicted COVID-19 case counts for a select group of counties within each urban-rural category as defined by the National Center for Health Statistics (NCHS). Category 1 represents the most urban counties, while Category 6 denotes the most rural ones.



Systematic Differences in Model Fit Across Counties

To further evaluate how the model performed in each county, we calculated the MAD for which smaller values indicate better model fit. For MAD, the median value was 0.259 (IQR 0.201-0.301). When we compared MAD to country population size (Figure 3A), we found that model fit was better for counties with higher population numbers (P<.001).

Furthermore, we found that MAD was associated with urbanicity (Figure 3B), with more urban counties (Categories 1, 2, and 3) having lower MAD (P<.001) and therefore better model fit. More rural counties had poorer model fits (higher values of MAD), with the exception of Chittenden VT.

When we fit a multivariable regression model including both logged population and urbanicity categories as predictors, we found both coefficients to have the same sign as in the univariate analyses but were no longer significant (population size: P=.09; urbanicity: P=.13).

XSL•FO
RenderX

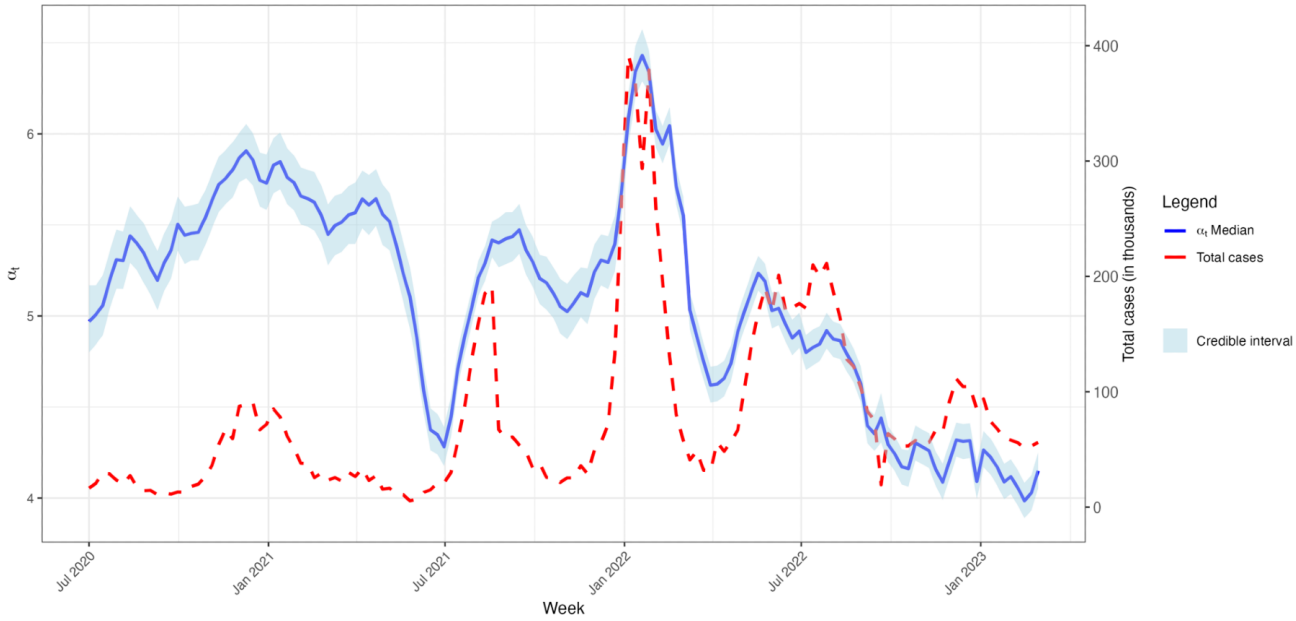
Figure 3. (A) How the quality of model fit (MAD) varies with county population size. (B) How the quality of model fit (MAD) varies with country urbanicity category. Category 1 represents the most urban counties, while Category 6 denotes the most rural ones.

XSL•FO
RenderX

Time Trends

Figure 4 shows changes in the estimated relationship between wastewater concentration and COVID-19 case reports over the study period, quantified as the level of logged COVID-19 case totals consistent with a given wastewater concentration, shown in blue. Notably, the fluctuations in this relationship are closely associated with the significant US waves at the end of 2020 to the beginning of 2021 (Alpha wave), the summer of 2021 (Delta wave), and the beginning of 2022 (Omicron wave). During these periods, the ratio of COVID-19 cases to wastewater concentration is relatively high as compared to the months before and afterwards. After early 2022, the estimated trend shows a progressive decline.

Figure 4. Time series of the time-varying coefficient (αt) alongside the aggregated case counts across counties from July 2020 to February 2023.

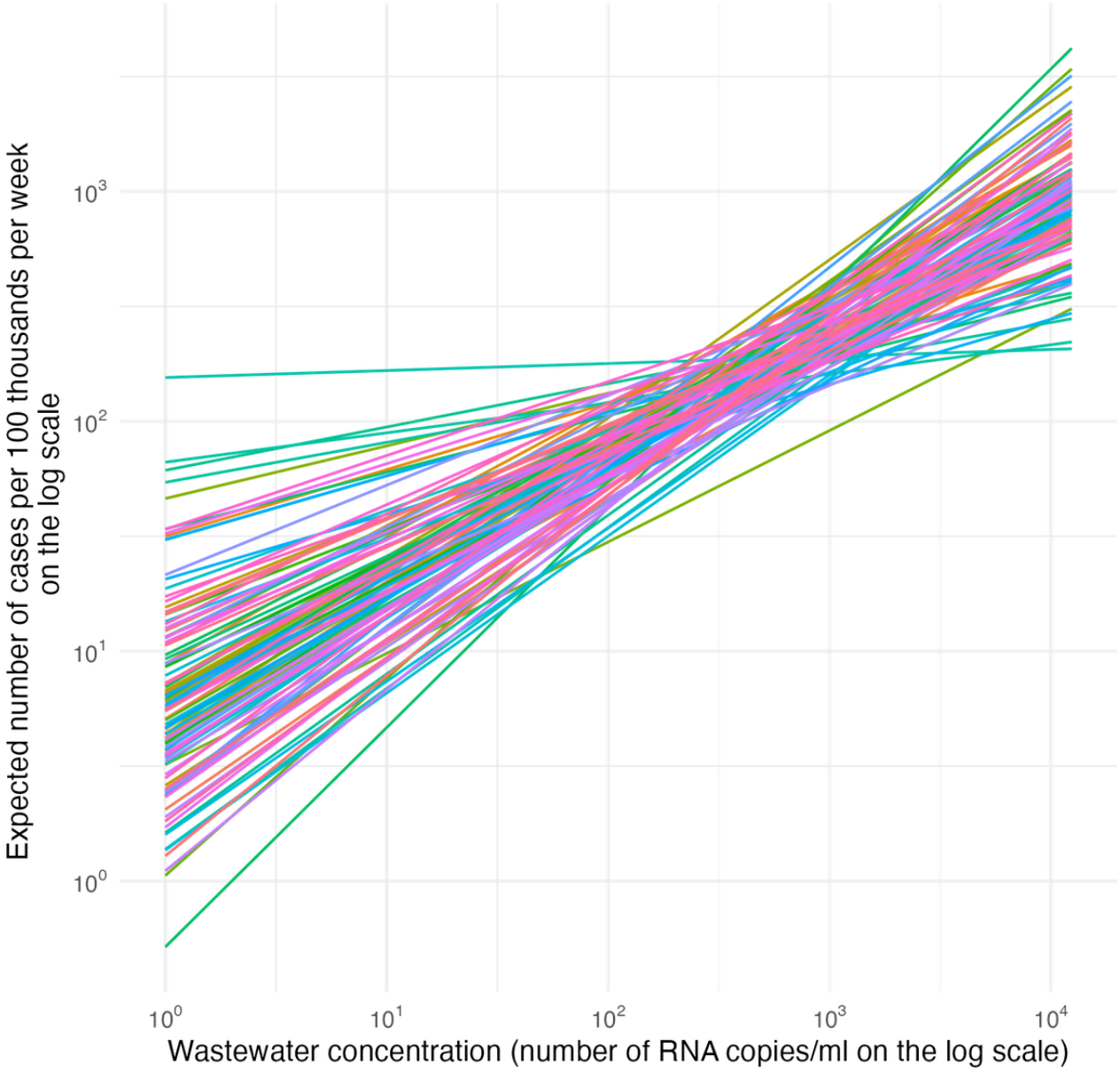


Differences Across Counties

Figure 5 shows fitted estimates for how changes in wastewater concentration values are associated with changes in COVID-19 cases for each county. As expected, the slopes shown in Figure 5 are positive (indicating that an increase in wastewater concentration was associated with an increase in reported cases), and relatively consistent across counties. In all cases, the slope of these lines was estimated to be less than 1 (median 0.551, IQR 0.447-0.632), indicating that the fitted relationship between wastewater concentration and COVID-19 case totals is less than proportional (eg, a 50% increase in wastewater concentration is associated with a <50% increase in case totals).

XSL•FO
RenderX

Figure 5. Estimated relationship between wastewater concentration and mean weekly cases per 100 thousand people for each county.



Results of Sensitivity Analysis

In the first sensitivity analysis, we re-estimated key results after fitting the study model to each of three equal 11-month time periods subdividing the overall study period into early, mid, and late pandemic. Estimated median MAD values from these sensitivity analyses were 0.19 (IQR 0.15‐0.26) for the early-pandemic period (July 2020-May 2021), 0.29 (IQR 0.22‐0.38) for the mid-pandemic period (June 2020-April 2022), and 0.2 (IQR 0.14‐0.26) for the late-pandemic period April 2022-March 2023. Overall, their results are similar to the MAD estimates from the main analysis (median 0.26, IQR 0.2-0.3), and show the general quality of model fit was generally consistent over the study period. Similarly, estimated trends in MAD as a function of population size and urbanicity matched those of the main analysis (MAD lower with larger population size and greater urbanicity), with results shown in Figures S2-S7 in Multimedia Appendix 2.

In the second sensitivity analysis, we re-estimated results after fitting the study model to countries with high data completeness (19 counties included). From this sensitivity analysis, we found a median MAD value of 0.22 (IQR 0.2‐0.26) comparable to the value of 0.26 (IQR 0.2-0.3) from the main analysis. Estimated trends in MAD as a function of population size and urbanicity matched those of the main analysis (Figures S8 and S9 in Multimedia Appendix 2).

Discussion

Principal Findings

Wastewater data have been used extensively during the SARS-CoV-2 pandemic to monitor disease trends and provide

XSL•FO
RenderX

early evidence of rising community transmission. However, there is limited information on the statistical relationship between wastewater metrics and reported COVID-19 cases, and how this relationship varies over time and across jurisdictions [20]. This knowledge is valuable for making decisions about how best to use wastewater data, and to understand the settings in which these data provide accurate information about COVID-19 case trends. In this study, we modeled the relationship between wastewater metrics and clinical cases at the county level in the United States from July 2020 to March 2023.

The results of our analysis show that models fit to wastewater data are better able to predict case counts in urban counties (based on NCHS categorization) as compared to more rural counties. This may be due to rural areas having lower levels of connection to centralized sewage systems, the source of wastewater surveillance data [10,25]. We also noted a reduction in model performance among counties with smaller population sizes. This is consistent with other studies that have reported wastewater surveillance to have limited sensitivity as an early warning indicator in smaller geospatial scales [17,26]. We also estimated differences in the quality of model fit that were not explained by urbanicity and population size—these may relate to local differences in the coverage of wastewater surveillance, the processing of wastewater samples, or the quality of COVID-19 case reporting.

In addition to intercounty differences, the results also revealed fluctuations in the relationship between wastewater concentration and COVID-19 case totals over the course of the pandemic. Several factors could account for these temporal trends. First, viral shedding patterns among individuals developing COVID-19 in recent years likely differ from those developing the disease in the early stage of the pandemic, as immunity in the population through previous infections and vaccination increased significantly throughout the pandemic [27]. Second, the transition between different dominant variants may have also influenced the dynamics of discharged RNA copies in human waste. For example, the viral and antibody dynamics are distinct between omicron and delta variants [28]. Also, as mutations may affect the quantification of SARS-CoV-2 concentration in wastewater, such viral changes may need to be accounted for in estimating the relationship between wastewater levels and case notifications [29] . Finally, case reporting systems have changed over time, influencing the ratio of reported and unreported cases [30,31]. In particular, the declining trend over the final year of the time series (ie, a declining number of reported COVID-19 cases for a given wastewater level) likely relates to changes in COVID-19 testing and reporting practices, with a progressively smaller fraction of COVID-19 cases diagnosed and reported to public health authorities.

Future Directions

Designing effective wastewater surveillance systems requires trade-offs among cost-effectiveness, speed, and local feasibility [32]. Most current sequencing is implemented with hundreds to thousands of samples in parallel with expensive machinery and intensive investment in human resources [32]. This implies that counties with fewer resources may find it difficult to finance and support the required laboratory infrastructure and human resources.

While the quality of model fit was generally good, our analyses revealed substantial variation in the utility of wastewater surveillance across counties. It is also important to note that reported case counts, which we used as a proxy for infection trends, are an imperfect measure of true incidence [33], as they will reflect variation in COVID-19 diagnosis and reporting practices over time and across locations. Ideally, a population-based survey, such as the United Kingdom's Office for National Statistics' COVID-19 Infection Survey, would provide more accurate information for assessing the predictive performance of wastewater surveillance. However, without such data in the United States, we rely on case reports as the best available data.

The pepper mild mottle virus normalization method applied to the Biobot wastewater data corrects for variability in fecal content due to environmental factors such as stormwater. However, it may not fully account for seasonal changes or weather-related fluctuations that could influence the observed RNA concentrations. While this method is widely accepted in the field, future research could explore advanced normalization techniques to further refine nationwide wastewater surveillance models.

Further investigation, validation, and standardized data collection frameworks are required to better understand the relationship between wastewater and epidemiological data. The incorporation of next-generation sequencing and automation of wastewater data collection processes could enhance the effectiveness of wastewater surveillance. [34,35].

Conclusions

The SARS-CoV-2 pandemic made it clear that traditional event-based surveillance systems have critical deficiencies for providing prompt and valid information about the local epidemiological situation. Wastewater surveillance may provide health agencies with another early detection and effective surveillance tool, unaffected by several of the deficiencies of traditional surveillance data. As of March 2024, more than 1300 locations in the United States and over 72 countries globally conducted wastewater surveillance [36,37]. When implemented effectively, these data can provide a comprehensive picture of SARS-CoV-2 transmission, capturing asymptomatic and nontested infections. Our study demonstrates that analyzing wastewater metrics across multiple jurisdictions can establish the relationship between wastewater and potential cases, and how these differ across locations and over time. However, the missing data in wastewater and uncertainty in case data require future efforts to make the relationship between them more established. Efforts to collect wastewater data in a more standardized manner should be enhanced further to fully realize their potential.

XSL•FO
RenderX

Acknowledgments

This project is supported by cooperative agreement NU38OT000297 from the Centers for Disease Control and Prevention (CDC) and the Council of State and Territorial Epidemiologists (CSTE), and SHEPheRD contract 200-2016-91779 from the CDC. This work does not necessarily represent the views of the CDC or CSTE. The authors would like to express their gratitude for the valuable discussions with Dr Noriko Endo during the early stages of this work. They also acknowledge the insightful contributions through seminars from the Center for Communicable Disease Dynamics at Harvard University.

Data Availability

The COVID-19 datasets analyzed in this study are publicly available in the repository of the Center for Systems Science and Engineering at Johns Hopkins University [38]. The wastewater datasets used in this study were obtained from Biobot Analytics and are not publicly available. Access to these data may be requested directly from Biobot Analytics.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Visual abstract.
[PNG File, 911 KB - publichealth_v11i1e68213_app1.png]

Multimedia Appendix 2
Supplementary tables and figures on the prior distributions used in the Bayesian model, comparisons of observed and predicted COVID-19 cases across all studied counties, and sensitivity analyses of how MAD varies with population size and urbanicity. MAD: median absolute deviation.
[DOCX File, 5322 KB - publichealth_v11i1e68213_app2.docx]

References

1. Science & tech spotlight: wastewater surveillance. : US Government Accountability Office; 2022.
2. Yu Q, Olesen SW, Duvallet C, et al. Assessment of sewer connectivity in the United States and its implications for equity in wastewater-based epidemiology. PLOS Glob Public Health 2024;4(4):e0003039. [doi: 10.1371/journal.pgph.0003039] [Medline: 38630670]
3. Matrajt G, Naughton B, Bandyopadhyay AS, et al. A review of the most commonly used methods for sample collection in environmental surveillance of poliovirus. Clin Infect Dis 2018 Oct 30;67(suppl_1):S90-S97. [doi: 10.1093/cid/ciy638] [Medline: 30376094]
4. Kazama S, Miura T, Masago Y, et al. Environmental surveillance of norovirus genogroups I and II for sensitive detection of epidemic variants. Appl Environ Microbiol 2017 May 1;83(9):e03406-16. [doi: 10.1128/AEM.03406-16] [Medline: 28213546]
5. Alfonsi V, Romanò L, Ciccaglione AR, et al. Hepatitis E in Italy: 5 years of national epidemiological, virological and environmental surveillance, 2012 to 2016. Euro Surveill 2018 Oct;23(41):1700517. [doi: 10.2807/1560-7917.ES.2018.23.41.1700517] [Medline: 30326991]
6. Prevost B, Lucas FS, Goncalves A, et al. Large scale survey of enteric viruses in river and waste water underlines the health status of the local population. Environ Int 2015 Jun;79:42-50. [doi: 10.1016/j.envint.2015.03.004] [Medline: 25795193]
7. Naughton CC, Roman FA Jr, Alvarado AGF, et al. Show us the data: global COVID-19 wastewater monitoring efforts, equity, and gaps. FEMS Microbes 2023;4:xtad003. [doi: 10.1093/femsmc/xtad003] [Medline: 37333436]
8. Wu F, Xiao A, Zhang J, et al. SARS-CoV-2 RNA concentrations in wastewater foreshadow dynamics and clinical presentation of new COVID-19 cases. Sci Total Environ 2022 Jan 20;805:150121. [doi: 10.1016/j.scitotenv.2021.150121] [Medline: 34534872]
9. McClary-Gutierrez JS, Mattioli MC, Marcenac P, et al. SARS-CoV-2 Wastewater Surveillance for Public Health Action. Emerg Infect Dis 2021 Sep;27(9):1-8. [doi: 10.3201/eid2709.210753]
10. National Academies of Sciences, Engineering, and Medicine. Wastewater-based disease surveillance for public health action. : The National Academies Press; 2023.
11. Jahn K, Dreifuss D, Topolsky I, et al. Early detection and surveillance of SARS-CoV-2 genomic variants in wastewater using COJAC. Nat Microbiol 2022 Aug;7(8):1151-1160. [doi: 10.1038/s41564-022-01185-x] [Medline: 35851854]
12. Karthikeyan S, Levy JI, De Hoff P, et al. Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. Nature 2022 Sep;609(7925):101-108. [doi: 10.1038/s41586-022-05049-6] [Medline: 35798029]
13. Kotay SM, Tanabe KO, Colosi LM, et al. Building-level wastewater surveillance for SARS-CoV-2 in occupied university dormitories as an outbreak forecasting tool: one year case study. ACS EST Water 2022 Nov 11;2(11):2094-2104. [doi: 10.1021/acsestwater.2c00057]

XSL•FO
RenderX

14. Davó L, Seguí R, Botija P, et al. Early detection of SARS-CoV-2 infection cases or outbreaks at nursing homes by targeted wastewater tracking. Clin Microbiol Infect 2021 Jul;27(7):1061-1063. [doi: 10.1016/j.cmi.2021.02.003] [Medline: 33601008]
15. Karthikeyan S, Nguyen A, McDonald D, et al. Rapid, large-scale wastewater surveillance and automated reporting system enable early detection of nearly 85% of COVID-19 cases on a university campus. mSystems 2021 Aug 31;6(4):e0079321. [doi: 10.1128/mSystems.00793-21] [Medline: 34374562]
16. Kuhn KG, Jarshaw J, Jeffries E, et al. Predicting COVID-19 cases in diverse population groups using SARS-CoV-2 wastewater monitoring across Oklahoma City. Sci Total Environ 2022 Mar 15;812:151431. [doi: 10.1016/j.scitotenv.2021.151431] [Medline: 34748841]
17. Klaassen F, Holm RH, Smith T, et al. Predictive power of wastewater for nowcasting infectious disease transmission: a retrospective case study of five sewershed areas in Louisville, Kentucky. Environ Res 2024 Jan 1;240(Pt 2):117395. [doi: 10.1016/j.envres.2023.117395] [Medline: 37838198]
18. Xiao A, Wu F, Bushman M, et al. Metrics to relate COVID-19 wastewater data to clinical testing dynamics. Water Res 2022 Apr 1;212:118070. [doi: 10.1016/j.watres.2022.118070] [Medline: 35101695]
19. Weidhaas J, Aanderud ZT, Roper DK, et al. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. Sci Total Environ 2021 Jun 25;775:145790. [doi: 10.1016/j.scitotenv.2021.145790] [Medline: 33618308]
20. Dai X, Acosta N, Lu X, et al. A Bayesian framework for modeling COVID-19 case numbers through longitudinal monitoring of SARS-CoV-2 RNA in wastewater. Stat Med 2024 Mar 15;43(6):1153-1169. [doi: 10.1002/sim.10009] [Medline: 38221776]
21. Duvallet C, Wu F, McElroy KA, et al. Nationwide trends in COVID-19 cases and SARS-CoV-2 RNA wastewater concentrations in the United States. ACS EST Water 2022 Nov 11;2(11):1899-1909. [doi: 10.1021/acsestwater.1c00434] [Medline: 36380771]
22. Dong E, Ratcliff J, Goyea TD, et al. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard: data collection process, challenges faced, and lessons learned. Lancet Infect Dis 2022 Dec;22(12):e370-e376. [doi: 10.1016/S1473-3099(22)00434-0] [Medline: 36057267]
23. Stan Development Team. RStan: the R interface to Stan. CRAN. URL: https://cran.r-project.org/web/packages/rstan/vignettes/rstan.html [accessed 2025-04-14]
24. Lemoine NP. Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. Oikos 2019 Jul;128(7):912-928. [doi: 10.1111/oik.05985]
25. Varkila M, Montez-Rath M, Salomon J, et al. Use of wastewater metrics to track COVID-19 in the U.S.: a national time-series analysis over the first three quarters of 2022. medRxiv. 2023 Feb 8. [doi: 10.1101/2023.02.06.23285542] [Medline: 36798337]
26. Gamage SD, Jinadatha C, Rizzo V Jr, et al. Nursing home wastewater surveillance for early warning of SARS-CoV-2-positive occupants-Insights from a pilot project at 8 facilities. Am J Infect Control 2024 Jun;52(6):701-706. [doi: 10.1016/j.ajic.2023.12.016] [Medline: 38181902]
27. Puhach O, Meyer B, Eckerle I. SARS-CoV-2 viral load and shedding kinetics. Nat Rev Microbiol 2023 Mar;21(3):147-161. [doi: 10.1038/s41579-022-00822-w] [Medline: 36460930]
28. Yang Y, Guo L, Yuan J, et al. Viral and antibody dynamics of acute infection with SARS-CoV-2 omicron variant (B.1.1.529): a prospective cohort study from Shenzhen, China. Lancet Microbe 2023 Aug;4(8):e632-e641. [doi: 10.1016/S2666-5247(23)00139-8] [Medline: 37459867]
29. Endo N, Nihei Y, Fujita T, et al. Explaining the impact of mutations on quantification of SARS-CoV-2 in wastewater. Sci Rep 2024 May 30;14(1):12482. [doi: 10.1038/s41598-024-62659-y] [Medline: 38816525]
30. Alvarez E, Bielska IA, Hopkins S, et al. Limitations of COVID-19 testing and case data for evidence-informed health policy and practice. Health Res Policy Syst 2023 Jan 25;21(1):11. [doi: 10.1186/s12961-023-00963-1] [Medline: 36698202]
31. Silk BJ, Scobie HM, Duck WM, et al. COVID-19 surveillance after expiration of the public health emergency declaration - United States, May 11, 2023. MMWR Morb Mortal Wkly Rep 2023 May 12;72(19):523-528. [doi: 10.15585/mmwr.mm7219e1] [Medline: 37167154]
32. Levy JI, Andersen KG, Knight R, et al. Wastewater surveillance for public health. Science 2023 Jan 6;379(6627):26-27. [doi: 10.1126/science.ade2503] [Medline: 36603089]
33. Varkila MRJ, Montez-Rath ME, Salomon JA, et al. Use of wastewater metrics to track COVID-19 in the US. JAMA Netw Open 2023 Jul 3;6(7):e2325591. [doi: 10.1001/jamanetworkopen.2023.25591] [Medline: 37494040]
34. Iwamoto R, Yamaguchi K, Katayama K, et al. Identification of SARS-CoV-2 variants in wastewater using targeted amplicon sequencing during a low COVID-19 prevalence period in Japan. Sci Total Environ 2023 Aug 20;887:163706. [doi: 10.1016/j.scitotenv.2023.163706] [Medline: 37105480]
35. Singer AC, Thompson JR, Filho CRM, et al. A world of wastewater-based epidemiology. Nat Water 2023;1(5):408-415. [doi: 10.1038/s44221-023-00083-8]
36. COVID data tracker. Centers for Disease Control and Prevention. 2025 Apr 14. URL: https://covid.cdc.gov/covid-data-tracker/#wastewater-surveillance [accessed 2025-05-09]

XSL•FO
RenderX

37. COVIDPoops19. ArcGIS. 2025 Apr 14. URL: https://www.arcgis.com/apps/dashboards/c778145ea5bb4daeb58d31afee389082 [accessed 2025-05-09]
38. CSSEGISandData/COVID-19. GitHub. URL: https://github.com/CSSEGISandData/COVID-19 [accessed 2025-05-09]

Abbreviations

CDC: Centers for Disease Control and Prevention
MAD: median absolute deviation
NCHS: National Center for Health Statistics

Edited by A Mavragani; submitted 01.11.24; peer-reviewed by D Han, X Dai; revised version received 17.02.25; accepted 18.02.25; published 22.05.25.

Please cite as:
Haraguchi M, Klaassen F, Cohen T, Salomon JA, Menzies NA
Statistical Relationship Between Wastewater Data and Case Notifications for COVID-19 Surveillance in the United States From 2020 to 2023: Bayesian Hierarchical Modeling Approach
JMIR Public Health Surveill 2025;11:e68213
URL: https://publichealth.jmir.org/2025/1/e68213
doi:10.2196/68213

© Masahiko Haraguchi, Fayette Klaassen, Ted Cohen, Joshua A Salomon, Nicolas A Menzies. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 22.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

SARS-CoV-2 Detection in International Travelers Through Wastewater-Based Epidemiology at the Kigali International Airport: Genomic Surveillance Study

Rwagasore Edson[1], MSc, MD; Robert Rutayisire[1], PhD; Ziad El-Khatib[2], MSc, EMBA, PhD; Olivier Nsekuye[1], MSc; Hugues Valois Mucunguzi[1], MSc, MD; Raissa Muvunyi[1], MSc; Esperance Umumararungu[1], MSc; Eric Remera[1], PhD

[1]Rwanda Biomedical Center, KG 644 St, Kimihurura, Kigali, Rwanda

[2]Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Rwagasore Edson, MSc, MD
Rwanda Biomedical Center, KG 644 St, Kimihurura, Kigali, Rwanda

Abstract

Background: Traditional infectious disease surveillance systems face significant limitations, including delayed detection, underreporting of asymptomatic cases, and inequitable health care access. Wastewater-based epidemiology (WBE), enhanced with genomic analysis, offers a noninvasive and cost-effective alternative for early pathogen detection and variant characterization, particularly valuable for monitoring international disease transmission.

Objective: This study aimed to implement and evaluate a genomics-enhanced WBE surveillance system for detecting and characterizing SARS-CoV-2 variants among international travelers at the Kigali International Airport, Rwanda, and to assess its potential as an early warning system for pandemic preparedness.

Methods: Between May and December 2023, we collected wastewater samples from international flights arriving at the Kigali International Airport under Rwanda's National One Health strategy. Molecular detection was performed using polymerase chain reaction (PCR) assays, followed by whole-genome sequencing of positive samples. Bioinformatics analysis included quality assessment with Nanoplot (version 1.41.6), genome mapping using minimap2 (version 2.26), and lineage identification using the Freyja tool (version 1.4.5). Spatial and temporal analyses were used to identify transmission patterns and variant origins.

Results: Of 630 wastewater samples collected from flights originating from 9 countries, 603 were successfully processed, with 21% (132/617) testing positive for SARS-CoV-2. Whole-genome sequencing was conducted on 33 samples, yielding an average viral sequence depth of 1250 reads with 92% genome coverage (range 78% - 97%). Genomic analysis identified 7 SARS-CoV-2 variants, including Omicron subvariants XBB.1.5, XBB.1.16.6 (eg, 5.1), GE.1, and FE.1.1.1. Notably, 70% (23/33) of sequenced samples could not be assigned to existing lineages, suggesting potential novel variants. Most samples came from Qatar (21.4%, 135/1630), the United Arab Emirates (19.5%, 123/1630), and the United Kingdom (19.4%, 122/1630). Positive samples were detected from 11 countries, with variants frequently found in flights from the United Kingdom, France, Belgium, Kenya, Tanzania, and South Africa. Sample collection capacity increased from 6 in week 1 to 33 by week 27. SARS-CoV-2 positivity rates showed seasonal variation, with a marked decline in June-July 2023.

Conclusions: Genomics-enhanced WBE demonstrated a high sensitivity for the early detection of SARS-CoV-2 variants among international travelers, including potential novel variants undetectable through traditional surveillance. Its noninvasive and cost-effective nature, combined with the ability to generate population-level epidemiological insights, makes it particularly suitable for resource-limited settings. This approach supports Rwanda's National One Health strategy and offers a scalable model for advancing global health security in Sub-Saharan Africa through innovative surveillance tools.

(JMIR Public Health Surveill 2025;11:e71104) doi:10.2196/71104

KEYWORDS

COVID-19 pandemic; SARS-CoV-2; community-based surveillance; building health systems resilience; low-income countries; pandemic preparedness

XSL•FO
RenderX

Introduction

The emergence of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which led to the global COVID-19 pandemic, represents one of the most significant global health crises of the 21st century. The pandemic challenged health systems worldwide in their preparedness, readiness, and capacity to respond effectively to a crisis of such magnitude [1]. More importantly, the rapid spread of the pandemic exposed delays and limitations in existing surveillance systems, underscoring the urgent need to restructure and incorporate advanced tools and novel strategies to improve public health surveillance and response systems [2-4].

The conventional approach to disease surveillance typically relies on clinical reporting and individual testing, which faces significant limitations including delayed detection, underreporting due to asymptomatic cases, and inequitable access to health care services. These limitations became particularly evident during the COVID-19 pandemic, highlighting the need for complementary surveillance methods that can overcome these challenges.

As part of global efforts to enhance epidemiological surveillance, the approach of wastewater-based epidemiology (WBE) has emerged. Therefore, WBE is the current science of analyzing wastewater to harvest the biological and chemical indicators of community health and risks. By detecting and quantifying biomarkers such as chemicals, pharmaceuticals, pathogens, and antimicrobial resistance genes in wastewater, WBE offers a powerful surveillance tool [5,6].

Recent advances in WBE have demonstrated its effectiveness in monitoring various pathogens beyond SARS-CoV-2, including influenza viruses, norovirus, and antimicrobial resistance genes. WBE can detect pathogen signals 5 - 7 days before clinical cases are reported, providing crucial early warning for public health interventions. Furthermore, the integration of WBE with digital disease surveillance and artificial intelligence has shown promise in enhancing the sensitivity and specificity of outbreak detection systems. Additionally, WBE can be expanded to include the surveillance of disease vectors with aquatic life stages, such as mosquitoes and snails, through the detection and characterization of environmental DNA [7-9].

The central premise of WBE is that wastewater provides a noninvasive method to collect pooled community samples containing epidemiological data, which is critical for monitoring public health and associated risks within a given area. Furthermore, combining WBE with advanced diagnostic tools, such as genomics and molecular diagnostics, enhances surveillance systems by enabling the early detection of locally circulating pathogens. This integrated approach provides an effective early warning system for disease outbreaks [5]. Molecular and genomic sequencing tools have proven sensitive enough to detect SARS-CoV-2 in sewage and wastewater, validating the utility of WBE in pandemic response efforts [10,11].

In response to the COVID-19 pandemic, Rwanda invested in strengthening its surveillance systems following the World Health Organization's (WHO) declaration of COVID-19 as a public health emergency of international concern in early 2020. Rwanda monitored the spread of COVID-19 through molecular detection and contact tracing, deploying locally adapted measures to control transmission. According to the WHO, Rwanda reported 133,194 confirmed cases and 1468 deaths attributable to COVID-19 [12].

Like many other low-income countries, Rwanda faces significant public health challenges, including recurrent infectious disease outbreaks, antimicrobial resistance, and the impacts of climate change. The early detection of health threats is critical to enabling rapid responses that limit further spread and mitigate negative outcomes. However, traditional surveillance approaches are often limited by the data they can collect and provide, resulting in reduced effectiveness in informing policymakers and guiding public health interventions. This highlights the urgent need to adopt innovative approaches and robust tools such as WBE, which can generate real-time, community-based epidemiological data.

To address these challenges, Rwanda is investing in implementing a genomics-enhanced WBE surveillance system as part of its National One Health strategy for early preparedness and response [13,14]. This strategy integrates genomic surveillance with public health response, universal health coverage, and the use of advanced technologies, such as drones, to improve animal, human, and environmental health services. In this study, we presented the integration of WBE and genomic surveillance in Kigali, Rwanda, to monitor the dynamics of infectious diseases, with a focus on COVID-19 among international travelers.

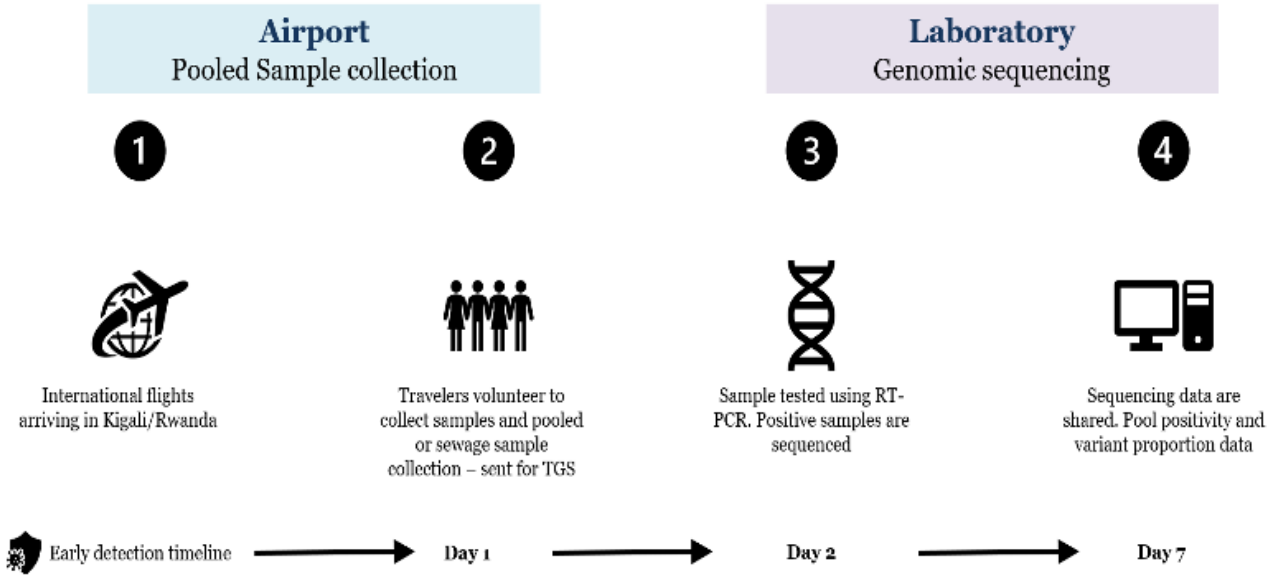
Methods

Overview

The Rwanda Pathogen Monitoring program, based on WBE, was launched in May 2023. This program implemented WBE at the Kigali International Airport, the main international airport in Rwanda, to detect and monitor variants of severe acute respiratory SARS-CoV-2 among international travelers. Wastewater monitoring for COVID-19 on flights involved the collection and preservation of sewage samples from aircraft wastewater systems. These samples were analyzed using molecular and genomics tools to detect SARS-CoV-2. The program was complemented by collecting individual biological samples from international travelers who voluntarily participated in the study. Genomic analysis was performed on these individual samples, and the results were compared to those obtained from wastewater samples to assess the sensitivity and capacity of WBE for monitoring cross-border transmission of SARS-CoV-2. Figure 1 illustrates the structure of the genomics-enhanced WBE surveillance system implemented in the Kigali International Airport for monitoring the dynamics of SARS-CoV-2 among international passengers entering Rwanda.

XSL•FO
RenderX

Figure 1. Implementation of the genomics-enhanced wastewater-based epidemiology (WBE) surveillance system for SARS-CoV-2 detection at the Kigali International Airport, Rwanda (May-December 2023).



Aircraft Wastewater Sampling

The WBE approach was implemented following a well-established protocol developed by the Rwanda Biomedical Center (RBC). Briefly, aircraft wastewater sampling involved the collection and analysis of wastewater from grounded aircraft to monitor the presence of pathogens. The process began by accessing the aircraft's wastewater system and connecting it to the waste valve of the aircraft lavatory system. Samples were collected using aseptic techniques to prevent contamination, using sterilized containers and tools to ensure the representativeness of the wastewater in the system. Following collection, the samples were transported to the Rwanda National Reference Laboratory under strict conditions to maintain appropriate storage temperatures and preserve sample integrity. At the laboratory, the samples underwent molecular and genomic analysis to detect circulating variants of interest. Results were interpreted in alignment with established protocols to ensure consistency and reliability.

Detection of SARS-CoV-2

Molecular analysis using polymerase chain reaction (PCR) assays was used to detect and quantify SARS-CoV-2 viral loads. All samples testing positive for SARS-CoV-2 via PCR were further subsequently subjected to genomic sequencing to identify viral lineages. A genomic surveillance pipeline was developed and implemented to detect the virus in travelers and wastewater samples and to characterize circulating variants. This pipeline maintained established links between passengers and flights to allow for validation of findings.

Data Analysis of Raw Reads

Bioinformatics tools were used for data analysis. Amplicon reads were assessed for quality using Nanoplot (version 1.41.6). Reads were then mapped to the SARS-CoV-2 reference genome (GenBank accession no NC_045512) using minimap2 (version 2.26). Subsequently, genome indexing was performed to produce sorted and indexed BAM files using SAMtools (version 1.17). The Freyja tool (version 1.4.5) was used to capture the abundance of variants lineages in wastewater samples.

Ethical Considerations

This study was conducted within the public health surveillance mandate of the RBC and was deemed exempt from ethical review, as it involved only the analysis of wastewater samples collected from aircrafts, with no human subject involvement whatsoever. No personal data were collected, processed, or stored at any point during this research. The wastewater sampling was performed as part of routine public health surveillance activities, in accordance with international regulations for disease monitoring at points of entry. All collected samples were completely anonymized, with only flight origin data retained for epidemiological analysis. The study adhered to established biosafety protocols for handling wastewater samples potentially containing infectious materials.

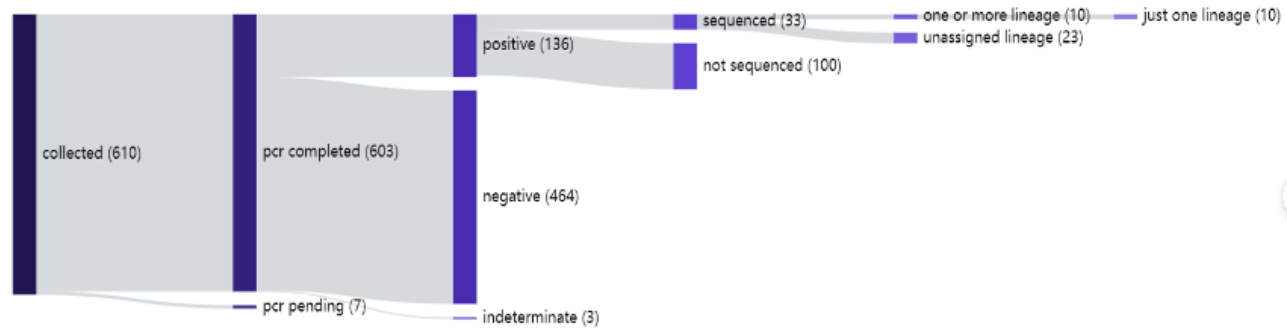
Results

Overview

A total of 630 wastewater samples were collected, with 603 being successfully tested. Among these, 21% (132/617) tested positive for SARS-CoV-2. Whole-genome sequencing was performed on 33 of the positive samples. Figure 2 outlines the surveillance algorithm for genomics-enhanced WBE used to monitor SARS-CoV-2 among international travelers arriving in Kigali.

XSL•FO
RenderX

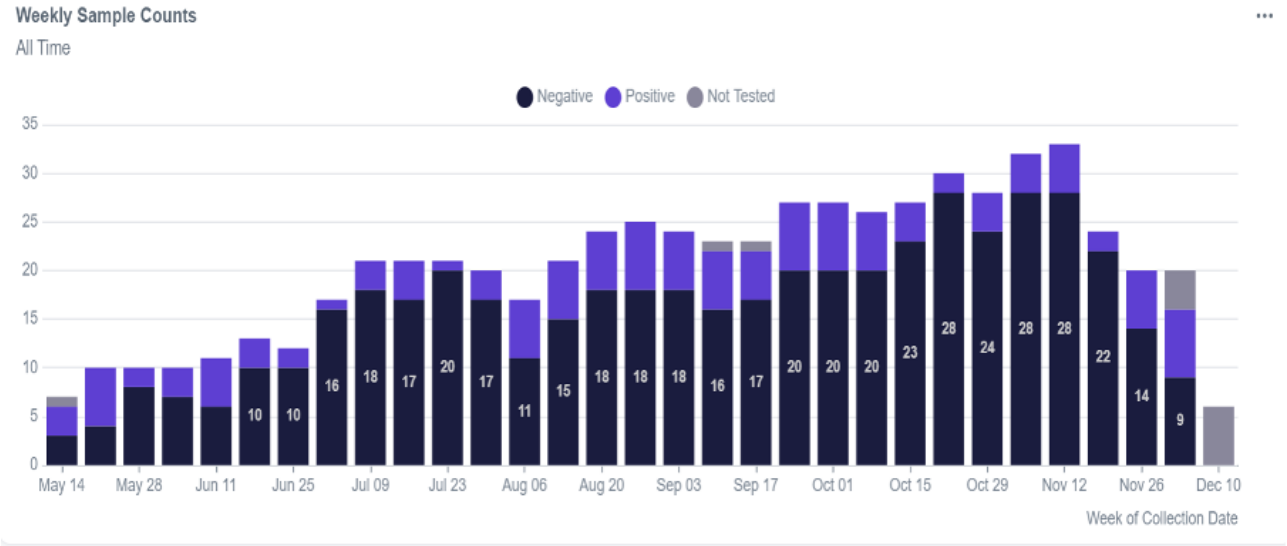
Figure 2. Surveillance algorithm for COVID-19 detection in wastewater samples collected from international flights arriving at the Kigali International Airport, Rwanda (May-December 2023).



Temporal Analysis of the Genomics-Enhanced WBE Surveillance for COVID-19 Among International Flights Arriving in Rwanda

The capacity of the surveillance system improved over time, as indicated by a significant increase in the weekly number of wastewater samples collected. Sample collection increased from 6 samples in the first week (May 14 - 19, 2023) to 33 samples in week 27 (November 12 - 18, 2023; Figure 3). However, a major decline in the SARS-CoV-2 positivity rate was observed between June and July, with only a single positive sample recorded during the period of 26 June to 23 July. Figure 3 illustrates biweekly changes in wastewater sample collection and seasonal variations in SARS-CoV-2 positivity rates among commercial flights arriving at the Kigali International Airport.

Figure 3. Temporal trends of SARS-COV-2 testing and positivity rate in wastewater samples from international flights arriving at the Kigali International Airport between May 14 and December 10, 2023.

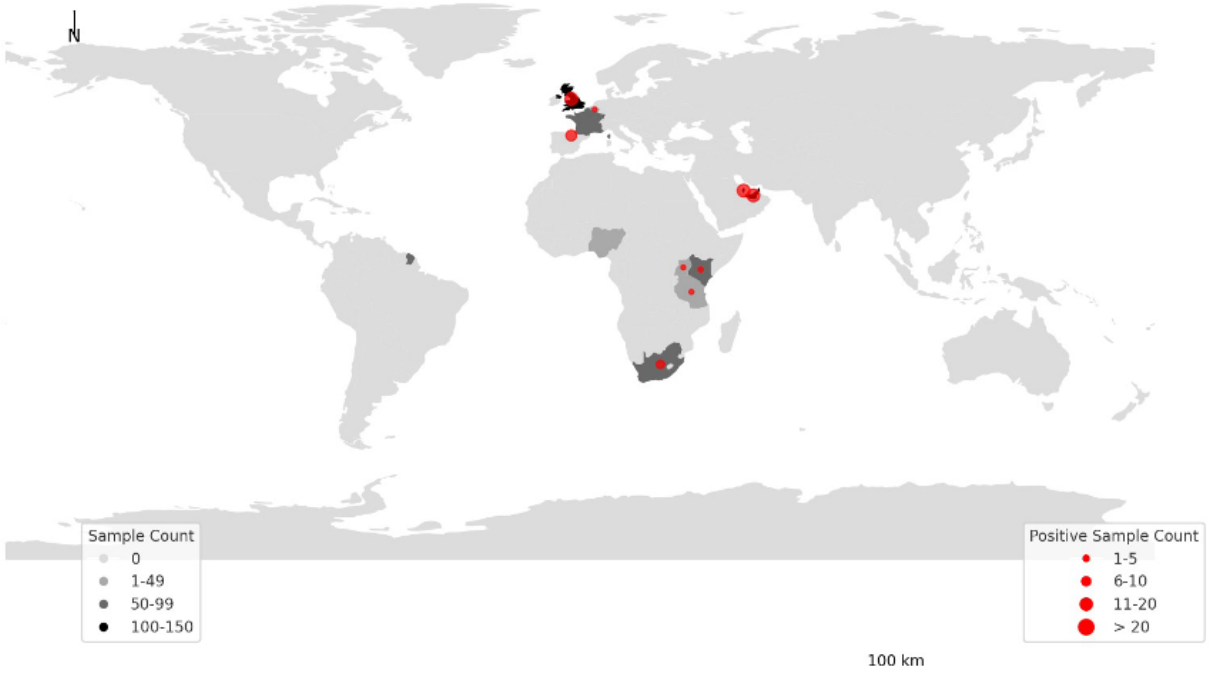


Spatial Analysis of the Genomics-Enhanced WBE Surveillance for COVID-19 Among International Flights Arriving in Rwanda

The 630 wastewater samples were collected from flights arriving from 9 countries. The majority of samples obtained from flights arriving from Qatar (21.4%, 135/630) followed by the United Arab Emirates (19.5%, 123/630) and the United Kingdom (19.4%, 122/630). Positive samples were identified from flights arriving from 11 countries, including the United Kingdom, France, Belgium, Kenya, Tanzania, the United Arab Emirates, Qatar, Tanzania, South Africa, and Uganda (Figure 4).

XSL•FO
RenderX

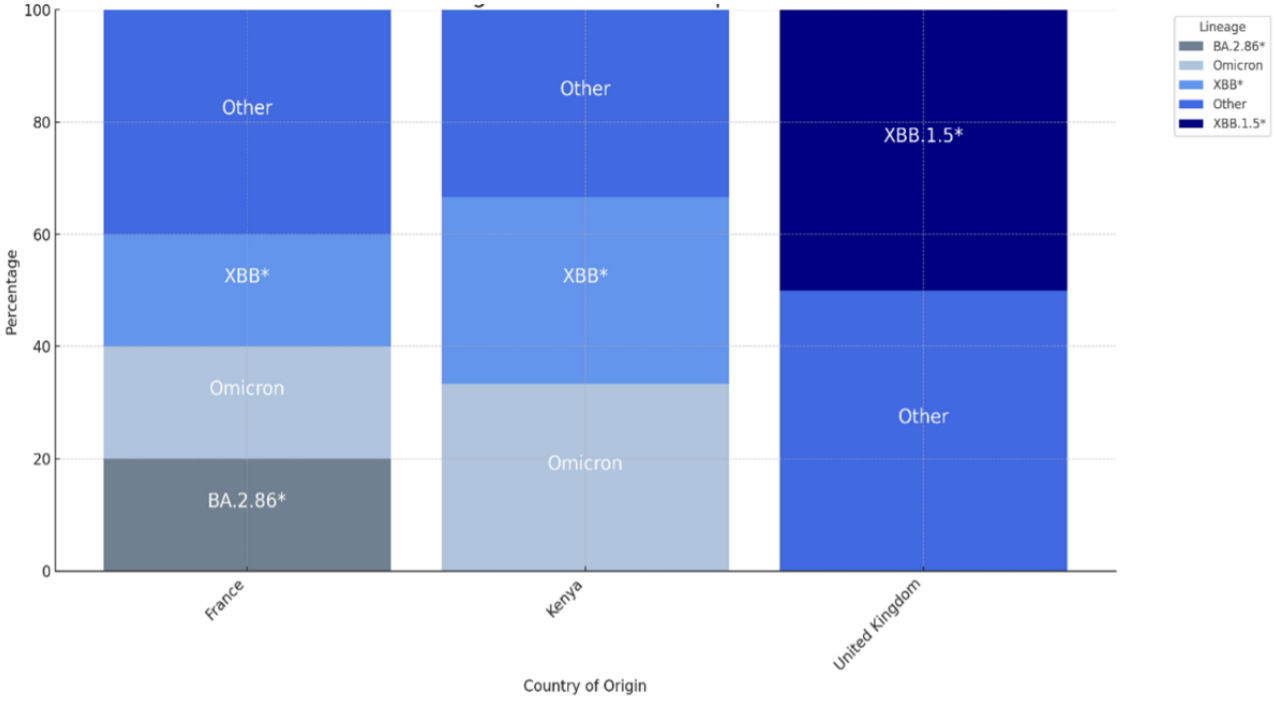
Figure 4. World map highlights the distribution of the collected and COVID-19 positive wastewater samples by the country of origin that was detected by the genomics-enhanced wastewater-based epidemiology surveillance system in Rwanda in 2023.



Lineage Description

Of the 33 COVID-19 positive wastewater samples sequenced, 23 could not be assigned to any known lineages. Genomic analysis of the remaining 10 samples revealed a high diversity of currently circulating internationally variants of SARS-CoV-2. Seven distinct Omicron subvariants were detected, including XBB.1.5, XBB.1.16.6 (eg, 5.1), GE.1, and FE.1.1.1. This included both recombinant variants such as XBB.1.5 and XBB.1.16.6, as well as subvariants from established Omicron lineages, including BA.5 (eg, 5.1), BA.2 (eg, GE.1) and BA.1 (eg, FE.1.1.1). Notably, XBB, XBB.1.5, BA.2.86, and other Omicron subvariants were the most prevalent and were primarily detected in wastewater from flights arriving from France, Kenya, and the United Kingdom (Figure 5).

Figure 5. The distribution of SARS-CoV-2 lineages distributed per the country of origin.



The high proportion (70%, 23/33) of unassigned sequences suggests the potential emergence of novel variants. However, further investigation is required to understand the evolution and dynamics of these variants.

XSL•FO
RenderX

Discussion

Principal Findings and Comparison With Previous Works

In this study, we presented a proof-of-concept study that emphasized the potential of genomics-enhanced WBE surveillance in monitoring the dynamics of the COVID-19 pandemic across countries via international flights and in detecting the spread of SARS-CoV-2 variants. The findings demonstrate the feasibility and effectiveness of WBE in the early detention of SARS-CoV-2 and circulating variants, highlighting its value in contributing to global heath security and pandemic preparedness. The diversity of SARS-CoV-2 origins detected underscores the utility of genomics-enhanced WBE for infectious disease surveillance.

Genomics-enhanced WBE offers an opportunity to track and monitor foodborne and waterborne diseases at the community level. This innovative approach aligns with Rwanda's National One Health strategy, which aims to improve human, animal, and environmental health through an integrated, transdisciplinary approach and multisectoral collaboration [15-17]. Additionally, it holds potential for advancing the control of neglected tropical diseases and reducing their socioeconomic burden [15]. Findings of this analysis demonstrate the feasibility and effectiveness of the use of WBE for the early detection of SARS-CoV-2 and the circulating variants. Furthermore, it highlights the diversity of origins in the source of SARS-CoV-2, contributing to global health security and pandemic preparedness.

The high proportion (70%, 23/33) of unassigned sequences suggests the potential emergence of novel variants. However, further investigation is required to understand the evolution and dynamics of these variants. This finding underscores the significant potential of WBE as an early warning system for detecting novel pathogen variants before they are identified through clinical surveillance. The ability to detect unassigned SARS-CoV-2 sequences through airport wastewater monitoring demonstrates how this approach could serve as a critical first line of defense in identifying emerging pandemic threats at their earliest stages, potentially providing weeks of additional preparation time for public health responses before widespread community transmission occurs.

This project has built the national capacity in Rwanda, a low-income country, to implement an innovative surveillance system that successfully detected new SARS-CoV-2 variants in Rwanda for the first time. The robustness and cost-effectiveness of this system highlight its potential for scaling. The integration of this surveillance system with immediate response measures—such as community engagement, public awareness campaigns on respiratory disease prevention, and COVID-19 vaccination efforts—has strengthened Rwanda's pandemic preparedness. These measures not only protect Rwandans but also safeguard international travelers by reducing the risk of disease spread or introduction of new variants into previously unaffected regions.

The implementation of genomics-enhanced WBE in Rwanda serves as a model for enhancing global health security in Sub-Saharan Africa. By building local capacity and adopting innovative approaches, Rwanda demonstrates how to effectively implement the One Health Strategy for improved health outcomes [18]. Expanding this cost-effective strategy both locally and internationally could help establish an early warning and response system to save lives and resources [4,19]. Such systems could be instrumental in densely populated areas with limited access to health care, particularly for monitoring infectious diseases with a high epidemic potential [5].

The emergence of WBE has transformed public health surveillance by providing real-time data on community health through the analysis of biomarkers in wastewater. This noninvasive approach reduces health care access challenges and socioeconomic barriers; public health systems often face challenges such as underreporting due to stigmatization or limited health care access, particularly in low-income settings. In contrast, WBE bypasses these bottlenecks by offering an unbiased snapshot of community health. Moreover, genomics-enhanced WBE aligns with the International Health Regulations [20] by facilitating cross-border data sharing and collaboration, improving global preparedness for future pandemics [21-23].

In this study, WBE not only identified new SARS-CoV-2 variants but also pinpointed their countries of origin (Figure 5). This highlights the importance of regionally coordinated data-sharing strategies to strengthen cross-border cooperation and enhance pandemic preparedness [21-23]. Integrating WBE with Rwanda's existing digital disease surveillance infrastructure would further maximize its impact. Linking wastewater findings to individual diagnostic syndromic surveillance data could validate signals and reveal broader transmission patterns. Additionally, a digitized WBE system incorporating big data analysis and artificial intelligence can automate sample processing and interpretation, rapidly uncovering health threats and trends.

Beyond infectious diseases, WBE holds promise for monitoring antimicrobial resistance, substance use trends, disease vector composition, and the unregulated use of pesticides and fertilizers. This could inform strategies to combat antibiotics, insecticides and pesticides resistance, as well as addiction-related harms while guiding vector control efforts. When applied ethically, WBE generates actionable and equitable health insights unattainable through individual testing alone.

Focusing on the Kigali International Airport underscores Rwanda's strategic approach to infectious disease surveillance. As a primary point of entry, the airport allows for the swift detection and mitigation of potential threats. The detection of 7 different Omicron subvariants in wastewater samples from international flights demonstrates the technique's sensitivity and ability to identify diverse viral strains. This not only supports domestic public health response strategies but also contributes to global knowledge on viral evolution and transmission.

Using WBE for SARS-CoV-2 detection at airports has numerous advantages. It enables early detection, even before travelers develop symptoms or widespread community transmission occurs. Given that a significant proportion of asymptomatic

XSL•FO
RenderX

travelers may carry SARS-CoV-2 [24,25]. WBE provides a cost-effective and noninvasive alternative to individual testing [26]. Unlike health declarations and temperature checks, which rely on passenger honesty and symptomatic presentations, WBE offers an objective measure of viral prevalence. Additionally, its nonintrusive nature respects passenger privacy (because data are collected in an anonymous fashion), making it an ethical surveillance tool. As new variants of SARS-CoV-2 emerge, WBE can detect shifts in viral strains within wastewater, enabling timely responses to potential outbreaks [27].

Limitations

A key limitation of this study is the relatively small number of samples (n=33) that underwent full genomic sequencing, which may not capture the complete diversity of SARS-CoV-2 variants entering Rwanda through air travel. Additionally, while we collected complementary individual samples from travelers, we were unable to perform a direct comparison between individual and wastewater results for all flights due to logistical constraints and varying participation rates.

Furthermore, our sequencing data varied in quality across samples, with genome coverage ranging from 78% to 97% (median 92%, IQR 87%–95%) and depth of coverage varying from 850 to 1650 reads (mean [SD] 1250 [200] reads). The sequenced samples were not evenly distributed across the study period, with more samples sequenced during July to September and November, which may have affected our ability to detect temporal trends in variant circulation. Future studies should aim for a more consistent temporal distribution of sequencing to better track variant evolution over time.

However, the expansion of genomics-enhanced WBE in Rwanda faces funding challenges that could hinder scaling efforts. Sustainable financing models, such as donor-implementer partnerships and public-private collaborations, are essential for institutionalizing the system. Capacity building through knowledge training and knowledge exchange will foster self-sufficiency. Addressing these challenges through locally tailored strategies will ensure the long-term success of WBE. We call upon all health stakeholders to mobilize resources and strengthen collaboration to support pandemic preparedness through investments in genomics-enhanced WBE surveillance, harnessing wastewater surveillance to improve global health.

Conclusions

The implementation of WBE for SARS-CoV-2 surveillance at the Kigali International Airport showed a strong potential of wastewater surveillance for public health in Rwanda. The system's sensitivity in detecting COVID-19 prevalence and emerging variants among international travelers underscores WBE's capability for real-time monitoring of public health threats. Expanding wastewater monitoring to communities across Rwanda could provide invaluable insights into population health, particularly among marginalized and underserved groups.

To fully realize the potential of WBE nationwide, further research and multi-stakeholder collaboration is essential. Optimizing decentralized sampling strategies and integrating WBE findings with clinical diagnostics and digital disease surveillance systems would significantly enhance its impact. Sustainable financing models and capacity-building initiatives are critical to ensuring that WBE becomes a cornerstone of Rwanda's public health infrastructure.

Pilot WBE programs at other points of entry, densely populated urban areas, public institutions, and isolated rural communities could provide important insights for scaling up the approach. Initially, with additional support, Rwanda could prioritize the use of genomics-enhanced WBE to detect high-risk pathogens of global concern, such as polio virus, cholera, and typhoid. Over time, this system could be expanded to include broader disease surveillance, antimicrobial resistance monitoring, and the tracking of biochemical markers.

Acknowledgments

We would like to thank Ginkgo Bioworks for their training in wastewater sampling protocols and bioinformatics analysis.

Funding

No external funding was received to conduct this study.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to being used for operational reasons at the Rwanda Biomedical Center, but are available from the corresponding author on reasonable request. Authors agree to share data via email requests to the corresponding author.

Authors' Contributions

ER, YB, SN and CM took part of the study design and implementation; ER, ON and ZEK took part of the study analysis; All authors took part of the writing and revision of the text.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Myers LC, Liu VX. The COVID-19 pandemic strikes again and again and again. JAMA Netw Open 2022 Mar 1;5(3):e221760. [doi: 10.1001/jamanetworkopen.2022.1760] [Medline: 35262720]
2. Adebisi YA, Rabe A, Lucero-Prisno Iii DE. COVID-19 surveillance systems in African countries. Health Promot Perspect 2021;11(4):382-392. [doi: 10.34172/hpp.2021.49] [Medline: 35079582]
3. Ibrahim NK. Epidemiologic surveillance for controlling Covid-19 pandemic: types, challenges and implications. J Infect Public Health 2020 Nov;13(11):1630-1638. [doi: 10.1016/j.jiph.2020.07.019] [Medline: 32855090]
4. Ali Y, Ahmed A, Siddig EE, Mohamed NS. The role of integrated programs in the prevention of COVID-19 in a humanitarian setting. Trans R Soc Trop Med Hyg 2022 Mar 2;116(3):193-196. [doi: 10.1093/trstmh/trab119] [Medline: 34358320]
5. Sims N, Kasprzyk-Hordern B. Future perspectives of wastewater-based epidemiology: monitoring infectious disease spread and resistance to the community level. Environ Int 2020 Jun;139:105689. [doi: 10.1016/j.envint.2020.105689] [Medline: 32283358]
6. Choi PM, Tscharke BJ, Donner E, et al. Wastewater-based epidemiology biomarkers: past, present and future. TrAC Trends in Analytical Chemistry 2018 Aug;105:453-469. [doi: 10.1016/j.trac.2018.06.004]
7. Schneider J, Valentini A, Dejean T, et al. Detection of invasive mosquito vectors using environmental DNA (eDNA) from water samples. PLoS ONE 2016;11(9):e0162493. [doi: 10.1371/journal.pone.0162493] [Medline: 27626642]
8. Sengupta ME, Hellström M, Kariuki HC, et al. Environmental DNA for improved detection and environmental surveillance of schistosomiasis. Proc Natl Acad Sci U S A 2019 Apr 30;116(18):8931-8940. [doi: 10.1073/pnas.1815046116] [Medline: 30975758]
9. Amarasiri M, Furukawa T, Nakajima F, Sei K. Pathogens and disease vectors/hosts monitoring in aquatic environments: potential of using eDNA/eRNA based approach. Sci Total Environ 2021 Nov 20;796:148810. [doi: 10.1016/j.scitotenv.2021.148810] [Medline: 34265610]
10. Tran HN, Le GT, Nguyen DT, et al. SARS-CoV-2 coronavirus in water and wastewater: a critical review about presence and concern. Environ Res 2021 Feb;193:110265. [doi: 10.1016/j.envres.2020.110265] [Medline: 33011225]
11. Randazzo W, Truchado P, Cuevas-Ferrando E, Simón P, Allende A, Sánchez G. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. Water Res 2020 Aug 15;181:115942. [doi: 10.1016/j.watres.2020.115942] [Medline: 32425251]
12. WHO COVID19 dashboard. World Health Organization. 2023 Jun 4. URL: https://covid19.who.int/region/afro/country/rw [accessed 2023-10-03]
13. Djordjevic SP, Jarocki VM, Seemann T, et al. Genomic surveillance for antimicrobial resistance - a One Health perspective. Nat Rev Genet 2024 Feb;25(2):142-157. [doi: 10.1038/s41576-023-00649-y] [Medline: 37749210]
14. Gardy JL, Loman NJ. Towards a genomics-informed, real-time, global pathogen surveillance system. Nat Rev Genet 2018 Jan;19(1):9-20. [doi: 10.1038/nrg.2017.88] [Medline: 29129921]
15. Global report on neglected tropical diseases 2023. : World Health Organization; 2023 URL: https://www.who.int/teams/control-of-neglected-tropical-diseases/global-report-on-neglected-tropical-diseases-2023 [accessed 2025-11-18]
16. Bordier M, Delavenne C, Nguyen DTT, Goutard FL, Hendrikx P. One health surveillance: a matrix to evaluate multisectoral collaboration. Front Vet Sci 2019;6:109. [doi: 10.3389/fvets.2019.00109] [Medline: 31106210]
17. World Health Organization. One Health Joint Plan of Action (2022 2026): Working Together for The Health of Humans, Animals, Plants and the Environment 2022. URL: https://www.who.int/publications/i/item/9789240059139 [accessed 2025-11-18]
18. Giovanetti M, Cella E, Moretti S, et al. Monitoring monkeypox: safeguarding global health through rapid response and global surveillance. Pathogens 2023 Sep 11;12(9):1153. [doi: 10.3390/pathogens12091153] [Medline: 37764961]
19. EWARS- improving early detection and prompt response to acute public health events. World Health Organization. 2019 Nov 1. URL: https://www.afro.who.int/news/ewars-improving-early-detection-and-prompt-response-acute-public-health-events [accessed 2025-11-18]
20. International Health Regulations (2005) – Third Edition: World Health Organization; 2005. URL: https://www.who.int/publications/i/item/9789241580496 [accessed 2025-11-20]
21. Ahmed A. Urgent call for a global enforcement of the public sharing of health emergencies data: lesson learned from serious arboviral disease epidemics in Sudan. Int Health 2020 Jul 1;12(4):238-240. [doi: 10.1093/inthealth/ihz122] [Medline: 32142110]
22. The next pandemic: rift valley fever? GAVI TGAfVaI. 2021 May 27. URL: https://www.gavi.org/vaccineswork/next-pandemic/rift-valley-fever [accessed 2025-11-18]
23. Preventing the next pandemic - zoonotic diseases and how to break the chain of transmission. UNEP. 2020. URL: https://www.unep.org/resources/report/preventing-future-zoonotic-disease-outbreaks-protecting-environment-animals-and [accessed 2025-11-18]
24. Al-Qahtani M, AlAli S, AbdulRahman A, Salman Alsayyad A, Otoom S, Atkin SL. The prevalence of asymptomatic and symptomatic COVID-19 in a cohort of quarantined subjects. Int J Infect Dis 2021 Jan;102:285-288. [doi: 10.1016/j.ijid.2020.10.091] [Medline: 33157290]
25. Bai Y, Yao L, Wei T, et al. Presumed asymptomatic carrier transmission of COVID-19. JAMA 2020 Apr 14;323(14):1406-1407. [doi: 10.1001/jama.2020.2565] [Medline: 32083643]

XSL•FO
RenderX

26. Dollard P, Griffin I, Berro A, et al. Risk assessment and management of COVID-19 among travelers arriving at designated U.S. airports, January 17-September 13, 2020. MMWR Morb Mortal Wkly Rep 2020 Nov 13;69(45):1681-1685. [doi: 10.15585/mmwr.mm6945a4] [Medline: 33180758]
27. Karthikeyan S, Levy JI, De Hoff P, et al. Wastewater sequencing reveals early cryptic SARS-CoV-2 variant transmission. Nature New Biol 2022 Sep;609(7925):101-108. [doi: 10.1038/s41586-022-05049-6] [Medline: 35798029]

Edited by A Mavragani, T Sanchez; submitted 10.01.25; peer-reviewed by CGG Bamford, L Henszel; revised version received 08.07.25; accepted 09.07.25; published 27.11.25.

Please cite as:
Edson R, Rutayisire R, El-Khatib Z, Nsekuye O, Mucunguzi HV, Muvunyi R, Umumararungu E, Remera E
SARS-CoV-2 Detection in International Travelers Through Wastewater-Based Epidemiology at the Kigali International Airport: Genomic Surveillance Study
JMIR Public Health Surveill 2025;11:e71104
URL: https://publichealth.jmir.org/2025/1/e71104
doi:10.2196/71104

© Rwagasore Edson, Robert Rutayisire, Ziad El-Khatib, Olivier Nsekuye, Hugues Valois Mucunguzi, Raissa Muvunyi, Esperance Umumararungu, Eric Remera. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Gaps and Challenges in Acute Hepatitis B Surveillance in the Country of Georgia: Comprehensive Surveillance System Evaluation

Lika Karichashvili[1,2], MPH; Ketevan Galdavadze[1], MD; Khatuna Zakhashvili[1], MD; Maia Tsereteli[1], MD, PhD; Ekaterine Ruadze[1,2], PhD; Sophia Surguladze[3], MMedSc; Pawel Stefanoff[2], MD, PhD, MS

[1]National Center for Disease Control and Public Health, Kakheti Highway N99, Tbilisi, Georgia

[2]Mediterranean and Black Sea Programme in Intervention Epidemiology Training (MediPIET), European Centre for Disease Prevention and Control (ECDC), Stockholm, Sweden

[3]Integral Global Health, Tbilisi, Georgia

Corresponding Author:
Lika Karichashvili, MPH
National Center for Disease Control and Public Health, Kakheti Highway N99, Tbilisi, Georgia

Abstract

Background: In 2012, the country of Georgia established an electronic integrated disease surveillance system (EIDSS) for acute hepatitis B virus (HBV) infection. All medical facilities must report suspected and confirmed acute HBV cases to the regional public health centers within 24 hours, which are subsequently registered in the EIDSS.

Objective: This study aims to evaluate the acute hepatitis B surveillance system in Georgia in order to identify areas for improvement and develop recommendations that enhance its capacity to inform prevention and response efforts, supporting the elimination of viral hepatitis.

Methods: For the evaluation of the acute HBV surveillance system from 2015 to 2020, we used the US Centers for Disease Control and Prevention updated guidelines. We assessed data quality by calculating the percentage of missing values for key variables. We assessed simplicity, acceptability, and flexibility by describing surveillance processes and by surveying public health center epidemiologists. We evaluated representativeness by comparing cases registered in EIDSS with cases registered in hospital discharges. We assessed timeliness by calculating the number of days from the date of diagnosis to the date of notification in EIDSS. We calculated the positive predictive value as the proportion of cases notified during 2018‐2020 having documentation of confirmatory tests in their medical records, meeting the confirmed case definition.

Results: During 2015‐2020, 270 cases of acute viral hepatitis B were reported to the EIDSS. All notified cases were hepatitis B surface antigen positive. However, only 10 of the 19 (53%) key variables were complete. Hepatitis B test results were missing in most reported cases, despite 82% (223/270) being classified as “confirmed.” Simplicity and acceptability of the system were affected by 30% (31/104) of the respondents experiencing challenges with the EIDSS reporting form. The system had limited flexibility due to cumbersome procedures to implement any changes. Representativeness was limited, as only 41% (270/657) of confirmed cases recorded in the hospital discharge database were reported to the EIDSS. The average notification delay was 72 hours. Among 104 cases notified in 2018‐2020, 66 met the case definition, leading to a positive predictive value of 63%.

Conclusions: The surveillance system for acute HBV infection was timely, although only 51% (139/270) of the cases were reported within the 24-hour notification target. The system was not representative and did not correctly ascertain cases. We recommend reconsidering the statutory notification time of 24 hours, revising notification forms and providing clear guidelines for data entry, and reporting all test results needed for adequate case classification to enhance data completeness and reliability of case classification.

(JMIR Public Health Surveill 2025;11:e72888) doi:10.2196/72888

KEYWORDS

acute hepatitis B; country of Georgia; evaluation; surveillance system; HBV

XSL•FO
RenderX

Introduction

Hepatitis B as a Global Health Challenge

Hepatitis B is a potentially life-threatening liver infection caused by the hepatitis B virus (HBV). It can cause acute and chronic disease and increases the risk of death from cirrhosis and liver cancer [1]. Hepatitis B is transmitted by exposure to contaminated blood (injection drug use, unsafe medical injections, unscreened blood and blood products) and body fluids (perinatal, sexual transmission). The incubation period for HBV ranges from 2 to 6 months [2]. Laboratory tests including IgM to hepatitis B core antibodies (IgM anti-HBc) and hepatitis B surface antigen (HBsAg) are needed to confirm acute HBV infection and differentiate acute from chronic infections. Diagnosis of acute hepatitis B is more common in adults than in children, who remain mostly asymptomatic following exposure and have a high risk of progressing to chronic HBV infection [2]. Hepatitis B is preventable by vaccination, and the World Health Organization recommends 3 doses of hepatitis B vaccine for all children, health care workers, and persons at high risk of HBV infection [3].

The Role of Viral Hepatitis Surveillance

Viral hepatitis surveillance is essential to detect outbreaks, monitor disease trends, identify risk factors for recent infections, estimate the burden of viral hepatitis and its sequelae, and assess the impact of vaccination and other preventive services. Surveillance for acute viral hepatitis through reporting and investigation of patients with symptoms of acute hepatitis presenting to health care facilities can help detect outbreaks and identify leading risk factors for infection, which informs the development of prevention and response activities. The prevalence of chronic HBV infection is mainly assessed through nationally representative serological surveys, which can monitor disease trends over time [2].

Hepatitis B in Georgia

Georgia implemented both acute hepatitis B surveillance and nationally representative hepatitis B serosurveys [4]. In 2015 and 2021, the Georgian National Centre for Disease Control and Public Health (NCDC) and the US Centers for Disease Control and Prevention (CDC) conducted 2 serosurveys throughout the country to determine the burden of viral hepatitis. The results showed that in 2015 the prevalence of HBsAg was 2.8% in adults. By 2021, the prevalence of HBsAg in adults and children was 2.7% and 0.03%, respectively, indicating the need to scale up the prevention, screening, and treatment of hepatitis B in adults [4].

Acute Hepatitis B Notification and Reporting System in Georgia

In 2005, Georgia included acute hepatitis B in the list of notifiable diseases, and health care providers were mandated to report cases to the national surveillance system [5]. The acute hepatitis B surveillance system is comprehensive and is designed to cover the entire country, including all hospitals nationwide. Since 2012, reporting of acute hepatitis B cases is done using electronic case reporting forms in the electronic integrated disease surveillance system (EIDSS). EIDSS is directly linked to the Georgian civil registry, meaning that the national unique identification number (ID) and date of birth in the notification section are automatically identified. This guarantees the accuracy of the personal ID number belonging to the notified case. If the case subject does not have a personal ID number (immigrant or refugee), they cannot be linked to the Georgian civil registry. In such situations, the epidemiologist is required to register the case in EIDSS by noting the passport number in the annotation. In addition, reporting of confirmed acute cases of hepatitis B is done using an electronic module of registration of discharged inpatients (IV-066). This is a separate register that collects the aggregated number of medical records from all patients discharged from hospitals across the country by the end of each month.

Evaluating the Acute Hepatitis B Surveillance System

The World Health Organization recommends assessing viral hepatitis surveillance to identify potential barriers that might impact data quality and influence the use of data for action [2]. Surveillance of acute hepatitis B in Georgia has never been evaluated. We conducted an evaluation of the acute hepatitis B surveillance system in Georgia to identify areas for improvement and provide recommendations to ensure surveillance can inform prevention and response efforts and advance the elimination of viral hepatitis.

Methods

Description of the Surveillance System

Suspected cases of acute hepatitis B are usually identified in outpatient and inpatient facilities or private laboratories. The physician and laboratory personnel requesting the diagnostic tests are responsible for providing the patient with feedback regarding their test results. The physician is required to report each suspected or confirmed acute hepatitis B case within 24 hours to the municipal or regional primary health care center (PHC). Subsequently, within 72 hours of receiving the notification, the PHC epidemiologist starts a case investigation to collect relevant epidemiological data and enters the collected information in the EIDSS.

The EIDSS case reporting form is part of the public health surveillance system and contains 3 sections: (1) notification of data, (2) case investigation, and (3) laboratory results (Table 1). The notification section includes general demographic information (personal ID number, age, date of birth, sex, current residence, permanent residence, telephone number), symptom history, and the primary diagnosis. The case investigation section includes details of patient symptoms, exposures, hospitalization, vaccination history, information about contact, and final case classification. The laboratory investigation section includes test results.

XSL•FO
RenderX

Table . Description of variables in the electronic integrated disease surveillance system (EIDSS) case reporting form, Georgia, 2022.

Section	Subsection	Description of contents
Notification	Demographic information	Personal ID number, age, date of birth, sex, current residence, permanent residence, telephone number (If not residing in Georgia, records the country of residence)
Notification	General information	Date of completion of the paper form, reporting institution and receiving institution information, as well as investigator, investigator organization, initial date of investigation
Notification	Clinical information	Date of symptoms onset, status of the patient at the time of notification (alive or dead), initial diagnosis (if known), date of diagnosis and current location of the patient (hospital or house)
Notification	Note	Epidemiologists used this space for any additional comments. Additionally, if the case was not a resident of Georgia, epidemiologists entered the passport ID here
Case investigation	Clinical information	Initial case classification, date of exposure, date of symptoms onset, location of exposure, facility where patient first sought care, date patient first sought care, nonreported diagnosis from facility where patient first sought care, hospitalization, date of hospitalization, and antibiotic or antiviral therapy administered before specimen collection
Case investigation	Samples collection	Specimen type, date of reception or registration
Case investigation	Contact list	Demographic information of the patient's close contacts
Case investigation	Case clinical classification	Clinical symptoms selected from a general list composed of all the symptoms characteristic of 74 communicable diseases
Case investigation	Epidemiological links and risk factors	Hepatitis B immunization history, risk factors and source of infection, information on medical manipulations and safety practices for potential sources of infection
Case investigation	Final case classification and outcome	Final case classification, final diagnosis, link to an outbreak, outcome, and additional comments
Laboratory results	Tests	Results of diagnostic tests, selected from a list

Data Sources

All epidemiologists working at municipal, regional, and central levels have access to the EIDSS. There is a web version of EIDSS accessible from any device connected to the internet. Each epidemiologist has an individual username and password. The web version and analytical module are available in 3 languages (Georgian, English, Russian). Furthermore, epidemiologists at the NCDC have access to the EIDSS analytical module, allowing further analysis and preparation of annual surveillance reports.

Another source of data on confirmed acute hepatitis B cases is the hospital discharge form IV-066, which collects information from 265 hospitals across all 11 regions of Georgia. Each hospital is required to submit the list of discharge records to the NCDC's statistics department no later than the 10th day of the month following the discharge of the patient. The IV-066 form consists of patient demographics, facility information, medical chart number, and confirmed diagnosis. The IV-066 database is linked to and validated against the Ministry of Health's database, which is used to determine financial reimbursements to health care facilities. The database is strictly monitored by the Ministry of Health to ensure that diagnoses are correctly reported.

Case Definitions

During 2015 - 2020, the following case classifications for acute hepatitis B were used in Georgia:

- Suspected case: A person who has the following signs and symptoms: fever, headache, weakness, thirst, nausea, vomiting, diarrhea, abdominal pain, and one of the following: jaundice or a serum alanine aminotransferase level >100 IU/L—or in the absence of clinical signs, if there is a documented negative HBsAg test (HbsAg, "e" antigen—HbeAg or HBV nucleic acid testing) within 6

XSL•FO
RenderX

months of receiving a positive test result, the case was considered a suspected acute viral hepatitis B.
- Confirmed case: A suspected case having both IgM anti-HBc and HBsAg positive.

Analysis of Surveillance Attributes

We restricted the evaluation of the acute hepatitis B surveillance system to the period from January 1, 2015, to December 31, 2020, as the case definition for acute hepatitis B was revised in January 2015.

To assess the acute HBV surveillance system, we used the CDC updated guideline "Updated Guidelines for Evaluating Public Health Surveillance Systems" [6]. We considered the following system attributes: flexibility, simplicity, acceptability, data quality, representativeness, timeliness, and positive predictive value (PPV).

Flexibility

We assessed the flexibility of the EIDSS by describing the procedures required to adapt the system to new requirements, including the process for modifying questionnaires.

Simplicity and Acceptability

To assess the simplicity and acceptability of EIDSS, during March-August 2022, we distributed an online questionnaire (Multimedia Appendix 1) to epidemiologists working across 59 PHCs in 10 regions and 2 out of 9 regional branches of the NCDC. We did not estimate a sample size due to the qualitative nature of the survey and attempted to include all PHC epidemiologists. We allowed the respondents 2 months to complete the questionnaire. During this period, we called and reminded them once every 2 weeks and sent a reminder email once a month.

Data Quality and Representativeness

We assessed data quality in EIDSS by calculating the percentage of complete information for selected variables (demographic, laboratory, and case classification).

We evaluated representativeness by comparing the number of hospitalized cases recorded in the IV-066 database and those notified to EIDSS by year and region, and calculated reporting ratios by dividing cases registered in EIDSS by cases registered in IV-066. We also calculated the incidence of cases recorded in the EIDSS and IV-066 databases. Even though the IV-066 database has never been evaluated, for this calculation, we assumed that cases registered in the IV-066 database reflect HBV infections correctly diagnosed among hospitalized patients, as the database is linked to and validated against the Ministry of Health's reimbursement system and is strictly monitored to ensure that diagnoses are correctly reported.

Timeliness

To assess timeliness, we calculated delays between diagnosis and notification (target: maximum 24 h) and between notification and the start of the epidemiological investigation (target: maximum 72 h). The information on the start time of the epidemiological investigation is recorded by the PHC epidemiologists directly in the EIDSS during the case investigation process.

Positive Predictive Value

To assess the PPV, we reviewed medical histories of confirmed acute hepatitis B cases registered in EIDSS during 2018 - 2020. This period was selected because recent patient documentation was available in the hospital archives. We compiled a list of confirmed acute hepatitis B cases registered in the EIDSS during 2018 - 2020 based on the location of hospitalization, and if this information was missing, based on the region of residence. We distributed the list to hospitals or regional PHCs according to the patient's region of residence and hospital information and asked them to complete the following information based on review of medical records: date of birth, age, place of residence, date of notification, date when the case was registered in EIDSS, date of onset of symptoms, date of the patient's first visit to the first care facility, name of the medical facility, whether or not there was hospitalization, and if so the date of hospitalization and date of discharge, date of final diagnosis, presence of clinical signs, IgM anti-HBc and HBsAg laboratory test results, outcome (recovered, died, or unknown), and final case classification. Then, we calculated PPV as the proportion of cases reported to EIDSS that met the criteria of a confirmed acute hepatitis B case.

Ethical Considerations

The survey of regional epidemiologists who reported on their experience with EIDSS and notification was determined by the US CDC to represent a nonresearch public health activity and, therefore, did not require the CDC's institutional review board (IRB) review. The protocol was approved by the NCDC IRB (registration number for the IRB decision letter: IRB #2022 - 027). The study used routinely collected secondary surveillance data, and no personal or identifiable information was accessed. The questionnaire administered to regional epidemiologists was anonymous. The survey link was distributed to contact persons in each regional PHC together with a description of the survey's purpose, intended data use, and assurance of anonymity. Participation required confirming consent on the survey landing page before proceeding. No IP addresses or other identifiers were collected. The NCDC had no influence over how the link was shared within regional departments, and it was assumed that only individuals who consented chose to participate. No incentives or compensation were provided.

Results

During 2015 - 2020, 270 suspected or confirmed cases of acute hepatitis B were reported to EIDSS, while 657 confirmed cases were reported in the hospital discharge registries (IV-066). A total of 104 regional epidemiologists and PHC employees responded to the survey. Most respondents (85%, 88/104) were female, and 79% (82/104) were epidemiologists.

Analysis of Surveillance Attributes

Flexibility

To incorporate changes to information collected in EIDSS, NCDC's epidemiologists need to submit an Excel form with all proposed changes. Their proposed changes need to be approved by the head of the communicable diseases department.

XSL•FO
RenderX

After approval, the list of approved changes is sent by mail to the developer of the EIDSS database. Any changes to the form usually take 2 - 3 months. Following each major change, PHC epidemiologists need to be trained on the new changes.

Simplicity and Acceptability

Of 104 survey respondents, 30% (31/104) found that the EIDSS case reporting form was difficult to complete. Most respondents (65/104, 62%) believed that information on "safe practices" in health facilities was obsolete as municipal public health services do not have the mandate to evaluate infection prevention and control measures in medical facilities. Only 45% (47/104) of the respondents reported sufficient knowledge of an acute hepatitis B investigation, including details needed to obtain from the patient's medical documentation and during case interviews. Of the 33 respondents who were not able to thoroughly investigate the notified cases, 16 (48%) could not access relevant laboratory results. The predefined list of tests included in the EIDSS case investigation form does not include the IgM anti-HBc for acute hepatitis B, required for the confirmed case ascertainment. If the PHC epidemiologists obtain the laboratory test results, they can enter this information in the free-text "Note" field. However, this is not a standardized data entry.

Data Quality

While demographic information such as sex, age, and place of residence was recorded with completeness rates of 96%, 99%, and 100%, respectively, critical variables, such as hepatitis B vaccination date and laboratory test results, were entirely missing. Reasons for missing hepatitis B vaccinations were noted in only 21% (58/270) of the cases. While 82% (223/270) of the cases were classified as "confirmed," this was often based only on HBsAg positivity without results of IgM anti-HBc, undermining the accuracy of this classification (Table 2).

Table . Completeness of key variables, in the acute hepatitis B case reporting form submitted to the electronic integrated disease surveillance system, Georgia, 2015 - 2020 (N=270).

Variable	Values, n (%)
Sex	260 (96)
Age	267 (99)
Place of residence	270 (100)
Hospitalized	231 (86)
Outpatients	17 (6)
Information about last vaccination date	0 (0)
Reason why hepatitis B vaccination was not given	58 (21)
Laboratory test (HBsAg[a], IgM anti-HBc[b], HBeAg[c])	0 (0)
Final classification "probable"	5 (2)
Final classification "possible/suspicious" "lost from supervision" " rejected"	8 (3)
Final classification "confirmed"	223 (82)

[a]HBsAg: hepatitis B surface antigen.

[b]IgM anti-HBc: immunoglobulin M antibody to hepatitis B core antigen.

[c]HBeAg: hepatitis B e antigen.

Representativeness

Based on the review of hospital discharge forms (IV-066), during 2015 - 2020, 657 confirmed cases of acute hepatitis B were diagnosed in hospitals, of which 270 (41%) were registered in EIDSS. The incidence of cases recorded in hospital discharges (IV-066) varied between 2 and 4 cases per 100,000 inhabitants. In 2015, the incidence was 2 per 100,000, reaching its peak at 4 per 100,000 in 2017. Subsequently, the incidence slightly declined to 3 per 100,000 in 2018 and 2019, and further dropped to 2 per 100,000 in 2020. As for the EIDSS data, incidence remained constant at 1 case per 100,000 each. This comparison suggests that EIDSS reporting did not capture the trend in incidence of suspected acute hepatitis B that was observed in hospital discharge forms.

Timeliness

The average delay between diagnosis and notification to the PHCs was 72 hours, while the national guideline required notification within 24 hours. Almost half (139/270, 51%) of the reported acute HBV cases were notified within 24 hours. Furthermore, on average, there was a 48-hour time lapse between case notification and the beginning of the epidemiological investigation, which falls within the national guideline that mandates the investigation to begin within 72 hours of receiving notification.

Positive Predictive Value

Medical chart reviews were used to validate 104 patients who were registered as confirmed cases of acute hepatitis B during 2018 - 2020. Of these, 39 (37.5%) lacked sufficient information on clinical or laboratory data to determine if the patient was a confirmed case of acute hepatitis B.

XSL•FO
RenderX

Thus, we estimated the PPV at 63%, since only 66 cases met the criteria for a confirmed case. All confirmed acute HBV cases presented with jaundice and dark urine. Their average age was 36(SD 14.3) years, and 44 (66%) were men. Most (53/67, 79%) resided in Tbilisi (Table 3). All true positive cases were hospitalized. The average alanine aminotransferase level for these 66 true positive cases was 2313 IU/L (SD 1094). HBV DNA testing was performed in 46 out of 66 cases (Table 3).

Table . Demographic and clinical characteristics of confirmed (true positive) acute hepatitis B cases in Georgia, 2018 - 2020 (N=67).

Variable	Values
Gender, n (%)	
Men	44 (66)
Women	22 (33)
Region of residency, n (%)	
Guria	1 (2)
Ajara	12 (18)
Tbilisi	53 (79)
Symptoms, n (%)	
Fever	27 (41)
Stomach ache	37 (56)
Nausea	53 (80)
Lack of appetite	63 (95)
Weakness	64 (97)
Fatigue	65 (98)
Jaundice	66 (100)
Dark urine	66 (100)
Mean age (y) (SD)	36 (14.3)
Mean ALT[a] level (IU/L) (SD)	2313 (1094)

[a]ALT: alanine aminotransferase.

Discussion

Principal Findings

The comprehensive evaluation of the acute hepatitis B surveillance system in Georgia revealed areas of improvement. Georgia's acute hepatitis B surveillance reporting in EIDSS was timely, as 51% (139/270) of the acute hepatitis B cases were reported within 24 hours, as required by national guidelines. However, the system was not sufficiently representative and did not correctly ascertain cases.

While Georgia's system adapted to the new requirements, implementing changes is complicated and takes time (2 - 3 mo). Accelerating adaptation timelines would enhance the Georgian system's flexibility.

The simplicity of Georgia's EIDSS case reporting system was another area of concern. Although the reporting form for acute hepatitis B aimed to collect comprehensive data on transmission modes, risk factors, and contacts, public health specialists lacked access to key required information, which resulted in incomplete forms; this limited the utility of the epidemiological investigation section at the national level. Additionally, local epidemiologists reported technical difficulties collecting and entering key variables like IgM anti-HBc results into the EIDSS. Without access to laboratory data, PHC epidemiologists often submitted incomplete notifications, complicating case classification at the national level, which led to a decreased PPV of 63%. This highlighted the importance of reporting all test results to enable accurate case classification and public health action. Additionally, variables such as the date of last hepatitis B vaccination are crucial for assessing vaccine effectiveness and identifying potential cases of vaccine failure, which are essential for guiding immunization policy and response strategies.

Our survey of regional epidemiologists found that only 45% (47/104) had sufficient knowledge of acute hepatitis B investigation, leading to potential misclassification of cases and reduced validity of notifications. These findings emphasize the need for further training at the regional and health facility levels to improve the understanding of case definitions, investigation methods, and data reporting to EIDSS.

Incomplete data was an issue, with important fields such as vaccination status and the source of infection often left blank. Improving data completeness in Georgia is essential for ensuring the accuracy and reliability of the surveillance system. Underreporting undermined the representativeness of the system. The overall reporting ratio was 41%, with significant variation across regions. Discrepancies between the number of suspected cases reported to EIDSS and those recorded in the hospital

XSL•FO
RenderX

discharge database highlighted the need for cross-referencing data from multiple systems to ensure the representativeness and completeness of case reporting.

The survey of regional epidemiologists may not reflect the experience of all surveillance users. We selected a convenience sample of employees of PHCs and collected mostly qualitative feedback. To address this potential shortcoming, we tried to contact all PHCs and send several reminders to include as many eligible respondents as possible. In addition, a substudy validating the classification of acute hepatitis B cases utilized medical histories from reported cases between 2018 and 2020. This period did not cover the full period of our investigation. However, as we obtained complete information on these cases, we believe our results reflect true case classification practices. Access to diagnostic testing is another aspect of surveillance that can affect the reliability of surveillance notifications. Although we did not investigate access to diagnosis, it is generally recommended to be performed in all regional hospitals. The completeness of data relies heavily on the accuracy of entries by epidemiologists. If there are inconsistencies or gaps in data entry, the assessment may underestimate or overestimate the actual completeness. The assumption that all hospitalizations in IV-066 are correctly classified as acute HBV infections might introduce errors. Misclassifications or omissions in either database (IV-066 or EIDSS) may affect the accuracy of the representativeness analysis, leading to an under- or overestimation of reporting ratios. Additionally, since IV-066 includes only patients who have been admitted to hospitals, our comparison was limited to hospitalized cases in both the IV-066 and EIDSS databases to ensure consistency and accuracy in the representativeness analysis.

Comparison With Similar Studies

Similar evaluations were performed in China [7], Armenia [8], Germany [9], and South Korea [10]. Similar to Georgia, China's hepatitis B surveillance system, particularly in Fujian, Hainan, and Gansu provinces, relied on case reporting through the National Notifiable Disease Reporting System [7]. The system depended on clinicians to report hepatitis B cases and classified them as acute or chronic, according to national guidelines, as it does in Georgia. Both Georgia and China had electronically integrated surveillance data, allowing public health authorities to gather, monitor, and analyze data efficiently, providing near real-time data access. Both China and Georgia faced challenges with misclassification and data quality in hepatitis B surveillance. While China's system demonstrated problems with underreporting and inconsistent testing practices, Georgia also had data quality issues, especially with incomplete laboratory confirmations and misclassification of cases based on HBsAg without IgM anti-HBc.

Both Georgia's and South Korea's hepatitis B surveillance systems focused on acute cases. South Korea, which transitioned to mandatory reporting in 2011, demonstrated a high positive predictive value (94.4%) and good data quality [10]. However, it faced challenges with chronic cases being misreported as acute, a misclassification also identified in Georgia. Unlike South Korea, Georgia faced more significant challenges in data completeness and misclassification.

Similar to Georgia, Armenia's surveillance systems reflected the legacy challenge from the Soviet public health framework. Armenia's system, modeled on the Soviet-era framework, was extensive and involved reporting at various administrative levels [8]. It collected both aggregated and case-based data for numerous infectious diseases. It lacked complex procedures and standardized case definitions. Georgia, on the other hand, struggled with incomplete reporting of key variables and regional underreporting.

Both Germany and Georgia faced challenges with data completeness, though the focus of these issues differed: Germany struggled with recording transmission routes and vaccination status, while Georgia faced gaps in laboratory results and vaccination information. Germany operated a mandatory surveillance system under the Protection Against Infection Act, requiring physicians and laboratories to report acute hepatitis B cases [9]. The system was well established with a clear data flow through local, state, and national public health authorities. Although Germany's system performed well in terms of timeliness, it has experienced a decline in data quality over the 10-year evaluation period, especially in capturing transmission routes and vaccination history.

Conclusions

Acute hepatitis B surveillance in Georgia suffers from unnecessarily complex investigation procedures. The complicated case reporting forms are time-consuming and cumbersome to complete. The users are not able to complete the information necessary for classifying cases. In particular, the information on confirmatory testing is not recorded reliably, leading to the misclassification of a large proportion of notified cases. In addition, a large proportion of regional epidemiologists are not aware of the notification procedures and case definition criteria. Despite these challenges, the system demonstrated positive results in terms of timeliness. This investigation provides essential insights into the strengths and weaknesses of the Georgian hepatitis B surveillance system, offering targeted recommendations to enhance data accuracy, improve case classification, and ultimately strengthen public health response capabilities.

Recommendations

We recommend revising the EIDSS case reporting form to simplify data entry and ensure that only essential information for case classification is required, improving data completeness, simplicity of EIDSS, and acceptability. Additionally, to improve representativeness and PPV, we recommend developing clear guidelines on proper case notification and investigation procedures, particularly focusing on the use of confirmatory tests, disseminating to regional epidemiologists and reporting physicians, and providing easy-to-access training in notification procedures. We recommend conducting targeting training sessions for public health specialists to improve timeliness.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank Rania Tohme, US Centers for Disease Control and Prevention.

LK is a fellow of the MediPIET program, and PS is a Scientific Coordinator of the MediPIET program, supported financially by the European Centre for Disease Prevention and Control. The views and opinions expressed herein do not state or reflect those of the European Centre for Disease Prevention and Control. The European Centre for Disease Prevention and Control is not responsible for the data and information collation and analysis, and cannot be held liable for conclusions or opinions drawn.

Data Availability

The data that support the findings of this study were collected from the electronic integrated disease surveillance system managed by the Georgian National Centre for Disease Control and Public Health, medical chart reviews, and questionnaires. Due to Georgian legislation on data confidentiality and restrictions imposed by the Georgian Government, these data are not publicly available. However, data may be available upon reasonable request to the corresponding author, subject to approval from the Ministry of Internally Displaced Persons from the Occupied Territories, Labour, Health and Social Affairs of Georgia, and the National Centre for Disease Control and Public Health.

Authors' Contributions

Conceptualization: LK, KG, KZ, PS, SS
Final approval of the manuscript: ER, KZ, LK, MT, PS, SS
Formal analysis (data analysis and interpretation): ER, LK, KG, PS
Investigation (data collection/implementation): LK, KG
Methodology: KG, LK, PS
Supervision: KZ, SS, ER
Writing original draft: LK, PS
Writing review & editing: ER, KZ, LK, PS, SS

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study questionnaire.
[DOCX File, 89 KB - publichealth_v11i1e72888_app1.docx]

References

1. WHO overview on hepatitis B. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/hepatitis-b [accessed 2024-03-30]
2. Technical considerations and case definitions to improve surveillance for viral hepatitis. World Health Organization. 2016. URL: https://www.who.int/publications/i/item/technical-considerations-and-case-definitions-to-improve-surveillance-for-viral-hepatitis-report [accessed 2025-11-18]
3. Hepatitis B vaccines: WHO position paper. World Health Organization. 2017 Jul. URL: https://www.who.int/publications/i/item/WER9227 [accessed 2025-11-12]
4. Khetsuriani N, Gamkrelidze A, Shadaker S, et al. Toward reaching hepatitis B goals: hepatitis B epidemiology and the impact of two decades of vaccination, Georgia, 2021. Euro Surveill 2023 Jul;28(30):2200837. [doi: 10.2807/1560-7917.ES.2023.28.30.2200837] [Medline: 37498531]
5. The Ministry of Internally Displaced Persons from the Occupied Territories HL and SA of G. Order of the ministry of internally displaced persons from the occupied territories, health, labour and social affairs of Georgia N101/n April 5, 2005 st. : Tbilisi on the Manner of Production and Delivery of Medical Statistical Information (GEO); 2012.
6. Armstrong G, Birkhead GS, Horan JM, Herrera G, Lee LM, Milstein RL, et al. Updated guidelines for evaluating public health surveillance systems recommendations from the guidelines working group. Centers for Disease Control and Prevention. 2001. URL: https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm
7. Zheng H, Millman AJ, Rainey JJ, et al. Using a hepatitis B surveillance system evaluation in Fujian, Hainan, and Gansu provinces to improve data quality and assess program effectiveness, China, 2015. BMC Infect Dis 2020 Jul 25;20(1):547. [doi: 10.1186/s12879-020-05265-3] [Medline: 32711465]
8. Wuhib T, Chorba TL, Davidiants V, Mac Kenzie WR, McNabb SJN. Assessment of the infectious diseases surveillance system of the Republic of Armenia: an example of surveillance in the Republics of the former Soviet Union. BMC Public Health 2002;2:3. [doi: 10.1186/1471-2458-2-3] [Medline: 11914147]

XSL•FO
RenderX

9. Boes L, Houareau C, Altmann D, et al. Evaluation of the German surveillance system for hepatitis B regarding timeliness, data quality, and simplicity, from 2005 to 2014. Public Health (Fairfax) 2020 Mar;180:141-148. [doi: 10.1016/j.puhe.2019.11.012] [Medline: 31918048]
10. Chung J, Yu J, Cheon M, Tak S. Evaluation of the acute hepatitis B surveillance system in the Republic of Korea following the transition to mandatory surveillance. Osong Public Health Res Perspect 2024 Aug;15(4):353-363. [doi: 10.24171/j.phrp.2024.0083] [Medline: 39091164]

Abbreviations

CDC: US Centers for Disease Control and Prevention
EIDSS: electronic integrated disease surveillance system
HBsAg: hepatitis B surface antigen
HBV: hepatitis B virus
IgM anti-HBc: immunoglobulin M antibody to hepatitis B core antigen
IRB: institutional review board
NCDC: National Centre for Disease Control and Public Health
PHC: public health center
PPV: positive predictive value

Edited by DI Yaras; submitted 22.02.25; peer-reviewed by A Litzroth, KM Angelo; revised version received 31.07.25; accepted 26.09.25; published 01.12.25.

Please cite as:
Karichashvili L, Galdavadze K, Zakhashvili K, Tsereteli M, Ruadze E, Surguladze S, Stefanoff P
Identifying Gaps and Challenges in Acute Hepatitis B Surveillance in the Country of Georgia: Comprehensive Surveillance System Evaluation
JMIR Public Health Surveill 2025;11:e72888
URL: https://publichealth.jmir.org/2025/1/e72888
doi:10.2196/72888

© Lika Karichashvili, Ketevan Galdavadze, Khatuna Zakhashvili, Maia Tsereteli, Ekaterine Ruadze, Sophia Surguladze, Pawel Stefanoff. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 1.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Equivalence of Type 2 Diabetes Prevalence Estimates: Comparative Study of Similar Phenotyping Algorithms Using Electronic Health Record Data

Muchiri E Wandai[1], MSc, PhD; Katie S Allen[1], PhD; Ashley Wiensch[1], MPH; John Price[1], MA; Brian E Dixon[1,2], MPA, PhD

[1]Regenstrief Institute, 1101 West 10th Street, Indianapolis, IN, United States

[2]Richard M. Fairbanks School of Public Health (FSPH), Indiana University, Indianapolis, IN, United States

Corresponding Author:
Muchiri E Wandai, MSc, PhD
Regenstrief Institute, 1101 West 10th Street, Indianapolis, IN, United States

Abstract

Background: Timely surveillance of diabetes mellitus remains a challenge for public health agencies. In this study, researchers compared type 2 diabetes (T2D) prevalence estimates using electronic health record (EHR) data and computable phenotypes (CPs) as defined and applied by 2 independent networks. One network, Diabetes in Children, Adolescents, and Young Adults, was a research consortium, and the other, the Multi-State EHR-Based Network for Disease Surveillance, is a practice-based public health surveillance network.

Objective: This study sought to determine the equivalence of T2D prevalence estimates generated by 2 distinct, yet conceptually related, CPs using EHR data.

Methods: Each network used diagnostic, laboratory, and medication data for young adults (aged 18-44 years) extracted from the Indiana Network for Patient Care (INPC) to independently calculate prevalence of T2D using distinct CPs for the year 2022. The INPC is a statewide health information exchange that receives EHR data from multiple health care systems and supports public health use cases such as surveillance. The two one-sided tests method for independence with a predefined margin of –2.5 to +2.5 percentage points was used to compare the estimated prevalence as previously derived from the Multi-State EHR-Based Network for Disease Surveillance and Diabetes in Children, Adolescents, and Young Adults networks. The two one-sided tests for equivalence show that any observed difference between 2 estimates is small and practically insignificant. Results at the overall level, and stratified by sex, age, and race or ethnicity, were examined.

Results: Overall prevalence estimates for 2022 were 4.1% for CP 1 and 2.4% for CP 2. Although prevalence estimates for CP 1 were consistently higher than those for CP 2, absolute differences were generally less than 2.5 percentage points, which did not result in a statistically significant (P<.001) difference between estimates. The only exception was for Hispanic individuals, where prevalence was significantly different (P=0.2) for CP 1 (5.4%) versus CP 2 (3.0%), yielding a margin of 2.4 (95% CI 2.2-2.6) percentage points. Other groups that had relatively higher but statistically nonsignificant prevalence included male individuals (4.6% for CP 1 vs 2.3% for CP 2), individuals aged 35-44 years (6.9% for CP 1 vs 4.9% for CP 2), and African American individuals (5.5% for CP 1 vs 3.7% for CP 2). Therefore, we concluded that the 2 CPs largely produced equivalent estimates of T2D prevalence.

Conclusions: The 2 independent CPs demonstrated equivalent T2D prevalence estimates, except in Hispanic individuals. Although the CPs can be considered statistically equivalent, the data driving each CP may impact accuracy and completeness. CP 1 was broader, incorporating clinical diagnoses, laboratory data, and medication, whereas CP 2 used clinical diagnostic codes alone. These results have implications for improving harmonization of CPs for public health surveillance.

(JMIR Public Health Surveill 2025;11:e79653) doi:10.2196/79653

KEYWORDS

type 2 diabetes; computable phenotype; prevalence; equivalence; two one-sided tests; TOST; electronic health record; EHR; public health surveillance; public health research; public health practice; health information exchange

XSL•FO
RenderX

Introduction

Background

Diabetes is a major public health concern as it poses numerous health challenges that increase the risk of complications such as heart disease, kidney failure, and vision loss [1]. This chronic disease also has a significant impact on individuals' quality of life [2,3], health care systems, and economies, particularly when left untreated or poorly managed [4]. Timely interventions to prevent or delay the onset of diabetes are a common, critical function performed by public health organizations.

In the United States, there has been a rising prevalence of diabetes over the past few decades [5]. Between August 2021 and August 2023, the prevalence of diabetes among the adult population (aged ≥20 years) was 15.8%, which included both diagnosed (11.3%) and undiagnosed (4.5%) cases [6], with type 2 diabetes (T2D) accounting for approximately 90% to 95% of these instances [7]. The prevalence of diabetes varies by sex, age, race, and ethnicity, with higher rates observed among certain groups, including African American, Hispanic, and American Indian or Alaska Native individuals [8]. Risk factors contributing to diabetes include obesity, physical inactivity, and a family history of the disease [9].

In the US state of Indiana, the prevalence of diabetes mirrors national trends. According to recent data (2023) from the Behavioral Risk Factor Surveillance System, 13.2% of Indiana adults (aged ≥18 years) have been diagnosed with diabetes, with T2D being the most common [10]. The state exhibits notable disparities in diabetes prevalence based on factors such as age, race, and socioeconomic status. Higher rates are often observed among African American and Hispanic populations, as well as in individuals with lower incomes and educational levels [11]. Estimating prevalence at the state and substate levels is challenging due to the small sample sizes in existing surveillance systems. The National Health and Nutrition Examination Survey and the National Health Interview Survey that monitor the prevalence and incidence of diabetes in the United States have been shown to fall short of large enough sample sizes to accurately assess diabetes prevalence and incidence in youth at the state or national levels, as well as in race and ethnicity subgroups [12]. In addition, differentiating diagnosis by diabetes type in children, adolescents, and young adults has also been challenging due to the co-occurrence between T2D and obesity [13]. Therefore, minor variations in the way in which a case is defined for these age groups could yield huge differences in prevalence, unlike that in the middle-aged and older adults.

Given the limitations of existing surveillance systems, there is an increasing interest in using electronic health record (EHR) data to estimate prevalence at these granular levels. It is posited that EHR systems may capture more cases with distinguishing characteristics relative to a population-based survey as data are obtained from all individuals interacting with health care services [14]. Accurate estimates of the burden of diseases are critical in policy advising, planning of prevention and control programs, and optimal allocation of scarce resources [15-17]. Specifically, accurately estimating the burden of diabetes in unique geographic areas is essential for improving overall health outcomes and reducing health care costs [11].

One aspect of accurately estimating the burden is ensuring precision in estimating the prevalence of individuals with diabetes. This is particularly true when leveraging EHR data, which may be derived from various sources such as hospital records and administrative data. Identifying accurate case definitions is a critical component in estimating prevalence using EHR data [18]. It has been shown that the prevalence of diabetes can vary based on differences in computable phenotype (CP) definitions, which are sets of clinical features used to identify a patient with a particular condition [19]. The variation is dependent on component criteria and timing of observations and measurements, and as such, phenotype definition requires careful consideration if EHR data are to be useful for individual and population health improvement [20].

Objectives

Understanding case definitions for a particular health condition is important, as this would determine the degree to which disease surveillance measures could be under-, over-, or accurately estimated. EHRs contain a variety of data, including diagnoses, laboratory test results, and medications. The methods through which these data are combined in a CP could affect the accuracy of measures such as incidence and prevalence. The purpose of this study was to compare 2 CP methods for estimating the prevalence of diabetes using EHR data. These methods have been proposed for use by researchers as well as local and state public health agencies. The findings could be useful for public health practitioners in diabetes prevention and control programs to identify the method that best captures their populations and to understand the strengths and limitations of each approach.

Methods

Study Overview

The primary objective of this study was to compare T2D prevalence rates calculated from EHR data during the same period (2022) for young adults living in the Indianapolis metropolitan area using 2 previously defined CPs. The rates were calculated using Indiana-based data included in 2 national networks that seek to develop generalizable methods for using EHR data to enhance surveillance of chronic disease in the United States.

The first CP (CP 1) is based on the Multi-State EHR-Based Network for Disease Surveillance (MENDS) led by the National Association of Chronic Disease Directors, with 5 distinct geographic sites [21]. The second CP (CP 2) was developed by the Diabetes in Children, Adolescents, and Young Adults (DiCAYA) network (coordinated by New York University Langone), which has 8 distinct geographic sites [22].

Both projects are funded by the US Centers for Disease Control and Prevention and have been previously described in the literature. Indiana is a common site in both the networks. Through secondary analysis, we sought to understand whether the 2 sets of estimates converged or diverged from one another and examine the potential reasons for convergence or divergence. The two one-sided tests (TOST) method was used

XSL•FO
RenderX

to find a minimum, non-nil effect size between the prevalences produced by the 2 algorithms. Expert knowledge based on the available literature was used to determine whether the minimum effect size obtained can be used to infer practically equivalent prevalences.

Data Sources

The data used in each source study were extracted from the Indiana Network for Patient Care (INPC), the EHR repository at the center of the Indiana Health Information Exchange. The Indiana Health Information Exchange is a network of >20 years, including >100 hospitals, clinical laboratories, outpatient clinics, and other health care organizations that exchange EHR data for clinical and population health purposes [23,24]. The INPC is routinely used to support research [25]. The INPC hosts a variety of EHR data, both structured and unstructured, including diagnosis codes; laboratory tests; clinical observation reports in text form; and vital measurements such as blood pressure, body weight, and height. Data on diabetes diagnosis, treatment, and laboratory measurements previously extracted by the 2 research consortiums were used in this study.

Data from the INPC were extracted independently by each network study to calculate their respective CPs. These data were subsequently available to the researchers to use in this comparative study. The time windows for observation and prevalence measurement overlapped in the year 2022. Therefore, we used January 1, 2022, to December 31, 2022, as the prevalence window of observation.

Populations

CP 1 used a subset of the INPC data, focusing on individuals who sought care from 2 health systems that collectively deliver care to at least two-thirds of the population in central Indiana. One system, the largest in the state, has 17 hospitals, 15 emergency rooms, 11 urgent care centers, nearly 3000 hospital beds, and hundreds of primary care and specialty clinics. The second system, based in the city of Indianapolis, has over 300 beds, an emergency room, outpatient and inpatient care, specialty clinics, and 12 community health centers. Combined, the 2 health systems delivered care to 2.5 million individuals during the observation window. CP 1 included patients with at least one clinical encounter during a 2-year period, including the year of observation for diabetes prevalence, to capture a representative number of clinical encounters since individual health care use may not take place annually, especially for individuals who consider themselves to be healthy.

CP 2 used a subset of the INPC data that included individuals aged 18 to 44 years living in the metropolitan statistical area (MSA) of Indianapolis. The project included data from all health care providers during the observation window to patients living in the 11 counties that constitute the MSA. The total population in the MSA is 2.7 million. Included patients were required to have had at least one clinical encounter in the previous 3 years, which included the year of prevalence observation and the previous 2 years.

Because the DiCAYA study that developed CP 2 limited its scope to patients aged 18 to 44 years, we compared prevalence measures for only that age range. This allowed the study to be conducted without reidentifying patients or pulling additional data from the INPC. Furthermore, ,the distribution of the INPC encounter data (used as the denominator) by sex, race or ethnicity, and age was compared with that of the American Community Survey. This helped in understanding whether any group was nonrepresentative of the underlying population.

CP Description

Textbox 1 provides a summary of each CP algorithm, allowing for a comparison of the distinct approaches used in each surveillance project. We kept the geographic and CP distinctions intact.

XSL•FO
RenderX

Textbox 1. Comparison of the type 2 diabetes (T2D) computable phenotype (CP) algorithms examined in this study.

CP 1 from Multi-State Electronic Health Record–Based Network for Disease Surveillance [26]

Starts by diabetes defined as: Any one of the following:

1. Laboratory:
   - Hemoglobin A_{1C} ≥6.5
   - Fasting glucose ≥126
   - Random glucoses ≥200 on 2 or more occasions within a 2-year period
2. Diagnostic codes: *International Classification of Diseases* (*ICD*) codes on 2 or more occasions within a 2-year period
   - *ICD-9* code 250.x or
   - *ICD-10* code (E10.x, E11.x, or E14.x)
3. Prescriptions:Prescription for INSULIN outside of pregnancyOther prescriptions: albiglutide, alogliptin, chlorpropamide, dulaglutide, ertugliflozin, exenatide, glimepiride, glipizide, glyburide, liraglutide, lixisenatide, pramlintide, rosiglitazone, semaglutide, sitagliptin, tolazamide, tolbutamide, acarbose, canagliflozin, dapagliflozin, empagliflozin, gliclazide, linagliptin, miglitol, nateglinide, pioglitazone, repaglinide, saxagliptin, tirzepatide

Type 1

Among patients meeting the definition for diabetes above, define as type 1 diabetes (T1D) if any of the following apply:

1. C-peptide test <0.8
2. Diabetes auto-antibodies positive
3. Prescription for URINE ACETONE TEST STRIPS
4. Ratio of T1D: T2D *ICD-9* or *ICD-10* codes >50% and prescription for GLUCAGON
5. Ratio of T1D: T2D *ICD-9* or *ICD-10* codes >50% and never prescribed oral hyperglycemic medications

Type 2: if not T1D

CP 2 from Diabetes in Children, Adolescents, and Young Adults [27]

Step 1: Starts by a wide-net definition of potential patients with diabetes mellitus (DM) meeting any of the following:

1. ≥1 Hemoglobin A1c ≥6.5% (≥42 mmol/mol)
2. ≥1 Fasting glucose ≥126 mg/dL (≥7.0 mmol/L)
3. ≥1 Random plasma glucose ≥200 mg/dL (≥11.1 mmol/L)
4. ≥1 Diabetes-related diagnosis code from an inpatient or outpatient encounter
5. ≥1 Prescribed, administered (ie, provided during a hospitalization), or dispensed diabetes-related medication

Medications: ≥1 prescribed, dispensed, or administered: exenatide, gliclazide, glimepiride, glipizide, glyburide, insulin, meglitinide, metformin, nateglinide, pioglitazone, pramlintide, repaglinide, rosiglitazone, sitagliptin, sulfonylurea, thiazolidinediones.

Step 2: Uses the pool of patients in Step 1 to identify DM cases using diagnostic codes only.*ICD-9* codes: 250.x0 and 250.x2*ICD-10* codes: E08.x, E10.x, E11.x, E13.x

Prevalent T2D:

1. # type 1 diagnosis codes# DM diagnosis codes≤0.5; OR
2. Evidence of an anti-diabetes medication besides insulin or metformin; OR
3. No evidence of glucagon medication

CP 1 distinguishes diabetes by type (type 1 diabetes [T1D] or T2D) based on laboratory measurements (glycated hemoglobin [HbA_{1c}]≥6.5%, fasting plasma glucose ≥126 mg/dL, and random plasma glucose ≥200 mg/dL) on 2 or more occasions within a 2-year period; C-peptide and diabetes auto-antibody tests; *International Classification of Diseases, Tenth Revision* (*ICD-10*), code (E10.x, E11.x, or E14.x) on 2 or more occasions within a 2-year period; and prescriptions of insulin outside of pregnancy, oral hyperglycemic medications, or urine acetone test strips. Patients are classified as having T1D if any of the following criteria are met: C-peptide test of <0.8 ng/mL, positive diabetes auto-antibody test, prescription for urine acetone test strips, ratio of T1D to T2D *ICD-10* codes of >50%, and prescription for glucagon and never having been prescribed oral hypoglycemic medications. A patient who does not meet any

XSL•FO
RenderX

of the aforementioned criteria for T1D is classified as having T2D.

CP 2 uses a 2-step process to determine whether a patient has T2D. First, data on diagnoses (*ICD-10, Clinical Modification*, codes E08.x, E09.x, E10.x, E11.x, and E13.x); laboratory test results (HbA_{1c}>6.5% and fasting glucose≥126 mg/dL); or medications (eg, metformin) for patients with at least 2 clinical encounters are extracted. Next, the CP classifies individuals by diabetes type based only on the *ICD-10* diagnosis codes. A patient is classified as having T2D if the ratio of the number of T1D codes to the total number of diabetes codes per patient is less than half (<0.50) [22].

Textbox 1 shows that the main difference between the CPs is that CP 2 does not include medication and laboratory data, whereas CP 1 excludes the E08, E09, and E13 *ICD-10* diagnostic codes. Diagnostic codes E10.x and E11.x are included in both CPs.

Statistical Analysis

The TOST method of equivalence for independent samples was used to test whether diabetes prevalence differed by a prespecified margin of equivalence based on the definition of T2D by the different CPs from MENDS and DiCAYA. The TOST procedure is used to statistically reject the presence of effects that are large enough to be considered worthwhile [28]. This method shows whether the difference between groups is smaller than a tolerably small amount [29]. T2D prevalence was stratified by sex, age group, and race. The choice of the equivalence interval for statistical and practical significance is critical [29,30]. In this study, we used a margin of 2.5 percentage points as recommended by Tatem et al [31] for estimates that are below 10%. A significance level (α) of .05 was used, yielding 90% CIs for equivalence testing because the formula for the CI in the TOST method is $100(1 - 2\alpha)\%$ [32]. We hypothesized that prevalence estimates based on CP 1 and CP 2 would differ by more than 2.5 percentage points (as recommended by Tatem et al [31]); that is, the difference in the prevalences would lie outside the −2.5 to +2.5 bounds.

The comparison involved the difference in prevalence calculated as a result of the independent CP methods used in the distinct studies. Therefore, a TOST of independence was used given that each network study pulled data and calculated prevalence independently. Although the population in the MENDS study accounts for 80% of the geographic region covered by the DiCAYA study, which implies that all MENDS patients would likely be included in the DiCAYA study, the researchers could not link individual patient records used in the 2 network studies. Therefore, we treated the datasets from each network study as independent.

Ethical Considerations

Ethics Review and Approval

This study was conducted under an approved research protocol by the Human Research Protection Program at Indiana University (protocol 2008160551), which received exemption under Category 4(iii) of the revised Common Rule.

Privacy and Informed Consent

This study represents a secondary analysis of deidentified, population-level EHR data obtained from the INPC, which is governed by HIPAA (Health Insurance Portability and Accountability Act) protections as it is a health information exchange (HIE) that uses and discloses protected health information. The data used in this research contained no personally identifiable information as all data were anonymized, and the researchers did not have access to any patient identifiers. Because no human participants were directly involved in the research, the institutional review board waived the need for researchers to obtain informed consent.

Results

The overall prevalence of T2D among young adults in the year 2022 was 4.1% for CP 1 and 2.4% for CP 2 (Table 1). A breakdown by the demographic variables (age, sex, and race) also showed that CP 1–based prevalences were higher than those estimated using CP 2. However, despite numeric differences in the prevalence estimates, the variation was not more than 2 percentage points, with 3 notable exceptions: individuals aged 35 to 44 years, Hispanic individuals, and male individuals. Moreover, the only group for which the null hypothesis was not rejected was Hispanic individuals (difference of 2.4%, 95% CI 2.2%-2.6%). Both CPs produced narrow CIs under the assumption of normally distributed prevalences. Therefore, despite the differences, both CPs on average showed statistically equivalent estimates of diabetes prevalence for the Indianapolis MSA. However, it is important to note that the difference in prevalence between the 2 CPs increased with age. Therefore, there is a likelihood that the CPs could show no equivalent estimates for older ages.

XSL•FO
RenderX

Table . Estimated prevalence of type 2 diabetes derived from 2 different computable phenotypes (CPs) for the Indianapolis metropolitan statistical area population by age, sex, and race (2022).

Demographic characteristic	CP 1, n/N (%)	CP 2, n/N (%)	Percentage point difference (90% CI)	Reject null hypothesis?
Age group (y)				
18-24	1506/88,952 (1.7)	412/196,113 (0.2)	1.5 (1.4 to 1.6)	Yes
25-34	4105/128,848 (3.2)	4568/302,385 (1.5)	1.7 (1.6 to 1.8)	Yes
35-44	8394/121,714 (6.9)	13,589/276,832 (4.9)	2 (1.9 to 2.1)	Yes
Race				
African American	3401/62,064 (5.5)	4393/118,634 (3.7)	1.8 (1.6 to 2)	Yes
Asian	353/8900 (4.0)	696/28,771 (2.4)	1.6 (1.2 to 2)	Yes
Hispanic	2063/38,432 (5.4)	2253/74,144 (3.0)	2.4 (2.2 to 2.6)	No
White	7853/210,135 (3.7)	9706/457,682 (2.1)	1.6 (1.5 to 1.7)	Yes
Other	335/19,983 (1.7)	1521/96,099 (1.6)	0.1 (–0.1 to 0.3)	Yes
Sex				
Female	8009/207,979 (3.9)	10,488/428,534 (2.4)	1.5 (1.4 to 1.6)	Yes
Male	5996/131,535 (4.6)	8081/346,796 (2.3)	2.3 (2.2 to 2.4)	Yes
Overall	14,005/339,514 (4.1)	18,569/775,330 (2.4)	1.7 (1.6 to 1.8)	Yes

Discussion

Principal Findings

This study compared prevalence estimates of T2D for a single observation period using EHR data among young adults from 2 networks (MENDS and DiCAYA) that use distinct CP algorithms. The 2 CPs were developed independently to achieve generalizable methods for surveillance of diabetes mellitus (DM) by type using EHR data. CP 1 was developed in concert with local health departments and was designed to be used at the local and state levels. This CP incorporates diagnostic codes, laboratory measures, and medications to classify patients with T2D. CP 2 was developed by diabetes and epidemiological researchers to be used for obtaining state and national estimates of prevalence. This CP uses only diagnostic codes in its determination of T2D classification. On the basis of the results, the CPs produced equivalent estimates of the prevalence of T2D (except among Hispanic individuals) despite differences in their individual point estimates and in their approaches.

Our methodological approach to equivalence testing and interpretation of the findings is supported by prior work using the TOST method. Using a difference of –2.5 to +2.5 percentage points between prevalence estimates to indicate equivalence is reasonable and in line with prior work that tested various margins. For example, a study comparing equivalence of prevalence estimates from population survey data found a difference of –5 to +5 percentage points useful in establishing equivalence but a lower margin of –2.5 to +2.5 percentage points in cases in which prevalence was below 10% [31]. A margin of –5 to +5 percentage points was also used in a study investigating disparity in pediatric immunization coverage by race and ethnicity [33]. Given the variation in ranges, we based our decision on the recommendations of Tatem et al [31] for estimates that are below 10%.

The findings of this study have implications for EHR-based CP development and refinement, as well as for their (CPs) use in research. These findings also have implications for public health practice, which seeks to use CPs off the shelf for the surveillance of diabetes and other chronic diseases.

Considerations for CP Algorithms

Although the CPs were equivalent, they are fundamentally different, and these differences have implications for the accuracy of the prevalence estimates. On average, CP 1 produced higher estimates of T2D prevalence than CP 2. There are several reasons for these differences. First, although the CP 2 algorithm initially considers a range of clinical variables (eg, fasting glucose, HbA_{1c}, and medications) in determining the population of patients likely to have diabetes, at the classification stage of the algorithm, it only uses *ICD-10* codes to distinguish between T1D and T2D. On the other hand, CP 1 includes individuals in the numerator based on the diagnostic codes, presence of 1 or more abnormal laboratory measurements, and medications. Second, the CP 1 and CP 2 DM diagnostic codes differ slightly. The CP 1 algorithm does not use the E08.x, E09.x, and E13.x codes but does include E14.x, all of which represent other or unidentified DM types. On the basis of CP 1, it is likely that patients with E08.x, E09.x, or E13.x who have an abnormal laboratory test result will be classified as having T2D and not other DM types. This could be a factor contributing to Hispanic individuals showing a nonequivalent prevalence between the CPs, especially if they (Hispanics) were more sensitive to the discrepant items defining prevalence between the CPs (they could be having a slightly higher number of patients that are flagged as having diabetes based on laboratory tests and on *ICD-10* diagnostic codes E08.x, E13.x, or E14.x). This difference might not be important except when considering accuracy, as E09.x might be temporary due to the influence of other drugs rather than an indication of chronic illness.

XSL•FO
RenderX

As work continues to improve the accuracy, sensitivity, and specificity for diabetes CPs and other chronic diseases, these differences should be noted and further analyzed. Our results demonstrate not just equivalence but the need for refinement and harmonization. Using a broader set of diagnostic criteria, such as the laboratory data used by CP 1, likely improves accuracy and sensitivity as prior work has found that using diagnostic codes alone in EHR data can generate inaccurate estimates of prevalence [34,35]. Moreover, the value of medication data should be reconsidered in light of prior work as some diabetes medications are prescribed for nondiabetes conditions such as obesity and heart failure [36].

Considerations for Public Health Practice

Public health indicators, such as the incidence and prevalence of chronic diseases, provide insights into the health status of a population, allowing public health officials to monitor trends, identify areas of concern, evaluate the effectiveness of interventions, and inform policy decisions by highlighting key health issues and areas requiring further action [37-39]. The ultimate objective in calculating health indicators is to improve the health of populations and reduce preventable inequalities [40]. These indicators are estimates of a given health dimension in a target population and are therefore prone to some degree of imprecision [41]. Thus, it is essential to consider a test of equivalence of these indicators when measuring from different data sources or when using different methodological algorithms.

This study found that diabetes prevalence estimated from 2 slightly divergent phenotype algorithms can be considered equivalent, except when measuring prevalence in Hispanic populations. This implies that some jurisdictions, which might only have *ICD-10* codes available or data from a limited number of health systems in their geographic area, could estimate prevalence in a reasonably equivalent manner to those jurisdictions with access to near complete coverage or with full EHR data available. This could be an option for some public health organizations as prevalence computation may be quicker. Jurisdictions may now be able to implement routine estimates for chronic illness with limited data, improving timeliness while they continue efforts to implement data modernization plans. In addition, leaner approaches to population estimates may be necessary if funding for chronic disease or surveillance programs is reduced in the following months or years.

We note that although the CPs are implementable, they are complex. The sites involved in both the MENDS and DiCAYA studies are sophisticated health systems, academic medical centers, or HIE networks. Public health does not yet have scalable, easy-to-implement applications that analyze *ICD-10* data for chronic disease surveillance. Moreover, no library of CPs for chronic disease surveillance exists that could be easily downloaded and implemented within jurisdictions' computing environments. More work is needed to make CP algorithms available to public health jurisdictions for implementation, along with tools that support knowledge management and integration into the existing surveillance infrastructure. Future work should focus on improving not only the accuracy of CP algorithms but also the distribution, implementation, and maintenance of such algorithms for use by thousands of public health agencies. Although platforms such as the National Coalition of Certification Centers (NC3) knowledge base and the eMERGE network exist for clinical research studies, these efforts may not meet the needs of public health agencies. Therefore, public health jurisdictions may need to create a separate platform for sharing CP algorithms or partner with efforts such as NC3 and eMERGE to distinguish clinical CPs from those designed for public health use.

Limitations

This study has several limitations of note. First, the denominators for each CP were not the same. CP 2 leveraged the entire population of the Indianapolis MSA who sought care during the observation window, whereas CP 1 used only patients from 2 health systems. This limitation is mitigated by the fact that the 2 health systems, data from which were used by both CPs, represent more than two-thirds of the metropolitan population. However, it remains a potential source of bias for the CP 1–based estimates, which could be responsible for the significant difference in the Hispanic population between the 2 CPs (the numerators were comparable, but the CP 2 denominator was twice that of CP 1).

Moreover, it was not possible to link individual patient records between the MENDS and DiCAYA study data to independently verify whether the same patients with T2D were correctly identified by each CP algorithm as this was a secondary analysis using limited datasets generated for each independent network study. Applying the CP algorithms on the same set of patients and examining reasons for differences in classifications would strengthen reliability in the TOST findings, and this approach is recommended for future studies that compare CPs.

Additionally, medication data from all sources were limited. Medication data, including orders and retail pharmacy dispensing data, are not routinely available in the HIE network that provided the EHR data for both CPs. More complete medication data might improve the accuracy of the CPs, although prior work has found that few patients only have documentation of medication use without a chronic illness diagnosis [34]. However, relying on medication data to identify T2D cases may introduce misclassification bias as several antidiabetic medications such as metformin, glucagon-like peptide–1 receptor agonists, and sodium/glucose cotransporter 2 inhibitors are also prescribed for nondiabetic conditions, including polycystic ovary syndrome, obesity, and heart failure. This overlap can result in the inclusion of individuals without T2D, potentially inflating prevalence estimates and affecting phenotype specificity.

Finally, it was not possible to conduct extrinsic validation of the results using external datasets such as those from the Behavioral Risk Factor Surveillance System, claims data, or the US Diabetes Surveillance System (DSS). The 2022 data from the DSS showed that type-insensitive DM prevalence for the entire state of Indiana was 3.4% (95% CI 2.7%-4.3%) for individuals aged 18 to 44 years. However, it was not possible to obtain data from DSS stratified by sex, race or ethnicity, and granular age groups to compare with data from the MENDS or DiCAYA network studies. Comparisons with other public health

XSL•FO
RenderX

data, such as in the study by Allen et al [42], are recommended in future research to assess external validity.

Conclusions

As work continues to expand the use of EHR data for routine public health surveillance, such as the active work by jurisdictions within the Centers for Disease Control and Prevention data modernization initiative, refining CP algorithms for chronic diseases will be necessary to achieve the right balance between sensitivity and specificity. Efforts by both MENDS and DiCAYA demonstrate useful approaches that, through harmonization and refinement, can create practical and effective metrics that enable the tracking of diabetes burden over time. Future work is needed to standardize CP processes to enable scaling across the public health enterprise. This will enable not only surveillance but also action to reduce burden and improve health in communities.

Acknowledgments

The authors thank the Regenstrief Data Services team and the following individuals specifically who supported data acquisition and management: Anna Roberts; Jeff Warvel; and Acatia Greenwell, Master of Public Health. The authors further thank Kate Hohman from the National Association of Chronic Disease Directors for her review and comments on the early draft of this manuscript. This research was supported in part by Lilly Endowment Inc through its support for the Indiana University Pervasive Technology Institute. This work also received funding support from the Centers for Disease Control and Prevention (5U18DP006500) and the National Association of Chronic Disease Directors (via the Centers for Disease Control and Prevention award NU38PW000036). None of these bodies had any role in the study design, data collection, data analysis, or decision to publish. The authors confirm that this manuscript, or any similar version, has not been published elsewhere and is not under review by any other journal.

Data Availability

The clinical data used in this study are not publicly available, and restrictions apply to their use; however, the deidentified dataset may be available from the Regenstrief Institute, subject to individual institution and network-wide ethics approvals. Interested researchers should contact the corresponding author.

Authors' Contributions

Conceptualization: KSA and BED originally conceptualized the idea and shared it with MEW
Data curation: JP and MEW
Formal analysis: MEW
Funding acquisition: BED
Methodology: MEW and BED
Supervision: BED
Project administration: AW
Software: MEW
Writing – original draft: MEW
Writing – review & editing: MEW, KSA, AW, AS, JP, BED

Conflicts of Interest

None declared.

References

1. Papatheodorou K, Banach M, Bekiari E, Rizzo M, Edmonds M. Complications of diabetes 2017. J Diabetes Res 2018;2018:3086167. [doi: 10.1155/2018/3086167] [Medline: 29713648]
2. Al-Khaledi M, Al-Dousari H, Al-Dhufairi S, et al. Diabetes self-management: a key to better health-related quality of life in patients with diabetes. Med Princ Pract 2018;27(4):323-331. [doi: 10.1159/000489310] [Medline: 29669330]
3. Hurtado MD, Vella A. What is type 2 diabetes? Medicine (Abingdon) 2019 Jan;47(1):10-15. [doi: 10.1016/j.mpmed.2018.10.010]
4. O'Connell JM, Manson SM. Understanding the economic costs of diabetes and prediabetes and what we may learn about reducing the health and economic burden of these conditions. Diabetes Care 2019 Sep;42(9):1609-1611. [doi: 10.2337/dci19-0017] [Medline: 31431494]
5. Neupane S, Florkowski WJ, Dhakal C. Trends and disparities in diabetes prevalence in the United States from 2012 to 2022. Am J Prev Med 2024 Aug;67(2):299-302. [doi: 10.1016/j.amepre.2024.04.010] [Medline: 38648908]
6. Gwira JA, Fryar CD, Gu Q. Prevalence of total, diagnosed, and undiagnosed diabetes in adults: United States, August 2021-August 2023. NCHS Data Brief 2024 Nov(516):516. [doi: 10.15620/cdc/165794] [Medline: 40085919]

XSL•FO
RenderX

7. A report card: diabetes in the United States infographic. Center for Disease Control (CDC). 2021. URL: https://www.cdc.gov/diabetes/communication-resources/diabetes-statistics.html [accessed 2025-10-13]
8. Menke A, Casagrande S, Geiss L, Cowie CC. Prevalence of and trends in diabetes among adults in the United States, 1988-2012. JAMA 2015 Sep 8;314(10):1021-1029. [doi: 10.1001/jama.2015.10029] [Medline: 26348752]
9. Ismail L, Materwala H, Al Kaabi J. Association of risk factors with type 2 diabetes: a systematic review. Comput Struct Biotechnol J 2021;19:1759-1785. [doi: 10.1016/j.csbj.2021.03.003] [Medline: 33897980]
10. BRFSS prevalence & trends data. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/brfss/brfssprevalence [accessed 2025-05-19]
11. Disparities in diabetes care: an analysis of the Indiana Behavioral Risk Factor Surveillance System diabetes module. Indiana State Government. 2023. URL: https://www.in.gov/health/oda/files/BRFSS_Diabetes_Module_Analysis.pdf [accessed 2025-05-29]
12. Bjornstad P, Chao LC, Cree-Green M, et al. Youth-onset type 2 diabetes mellitus: an urgent challenge. Nat Rev Nephrol 2023 Mar;19(3):168-184. [doi: 10.1038/s41581-022-00645-1] [Medline: 36316388]
13. Unnikrishnan R, Shah VN, Mohan V. Challenges in diagnosis and management of diabetes in the young. Clin Diabetes Endocrinol 2016;2:18. [doi: 10.1186/s40842-016-0036-6] [Medline: 28702252]
14. Sauer CM, Chen LC, Hyland SL, Girbes A, Elbers P, Celi LA. Leveraging electronic health records for data science: common pitfalls and how to avoid them. Lancet Digit Health 2022 Dec;4(12):e893-e898. [doi: 10.1016/S2589-7500(22)00154-6] [Medline: 36154811]
15. Meadows AJ, Oppenheim B, Guerrero J, et al. Infectious disease underreporting is predicted by country-level preparedness, politics, and pathogen severity. Health Secur 2022;20(4):331-338. [doi: 10.1089/hs.2021.0197] [Medline: 35925788]
16. Anjana RM, Ali MK, Pradeepa R, et al. The need for obtaining accurate nationwide estimates of diabetes prevalence in India - rationale for a national study on diabetes. Indian J Med Res 2011 Apr;133(4):369-380. [Medline: 21537089]
17. Mardon R, Marker D, Nooney J, et al. Novel methods and data sources for surveillance of state-level diabetes and prediabetes prevalence. Prev Chronic Dis 2017 Nov 2;14:E106. [doi: 10.5888/pcd14.160572] [Medline: 29101768]
18. Ali MK, Siegel KR, Laxy M, Gregg EW. Advancing measurement of diabetes at the population level. Curr Diab Rep 2018 Sep 19;18(11):108. [doi: 10.1007/s11892-018-1088-z] [Medline: 30232630]
19. Richesson RL, Rusincovitch SA, Wixted D, et al. A comparison of phenotype definitions for diabetes mellitus. J Am Med Inform Assoc 2013 Dec;20(e2):e319-e326. [doi: 10.1136/amiajnl-2013-001952] [Medline: 24026307]
20. Spratt SE, Pereira K, Granger BB, et al. Assessing electronic health record phenotypes against gold-standard diagnostic criteria for diabetes mellitus. J Am Med Inform Assoc 2017 Apr 1;24(e1):e121-e128. [doi: 10.1093/jamia/ocw123] [Medline: 27616701]
21. Hohman KH, Martinez AK, Klompas M, et al. Leveraging electronic health record data for timely chronic disease surveillance: the multi-state EHR-based network for disease surveillance. J Public Health Manag Pract 2023;29(2):162-173. [doi: 10.1097/PHH.0000000000001693] [Medline: 36715594]
22. Hirsch AG, Conderino S, Crume TL, et al. Using electronic health records to enhance surveillance of diabetes in children, adolescents and young adults: a study protocol for the DiCAYA network. BMJ Open 2024 Jan 17;14(1):e073791. [doi: 10.1136/bmjopen-2023-073791] [Medline: 38233060]
23. Overhage JM, Kansky JP. The Indiana health information exchange. In: Dixon BE, editor. Health Information Exchange: Elsevier; 2023:471-487. [doi: 10.1016/B978-0-323-90802-3.00022-8]
24. Williams KS, Rahurkar S, Grannis SJ, Schleyer TK, Dixon BE. Evolution of clinical health information exchanges to population health resources: a case study of the Indiana network for patient care. BMC Med Inform Decis Mak 2025 Feb 24;25(1):97. [doi: 10.1186/s12911-025-02933-9] [Medline: 39994604]
25. Dixon BE, Whipple EC, Lajiness JM, Murray MD. Utilizing an integrated infrastructure for outcomes research: a systematic review. Health Info Libraries J 2016 Mar;33(1):7-32. [doi: 10.1111/hir.12127]
26. Klompas M, Eggleston E, McVetta J, Lazarus R, Li L, Platt R. Automated detection and classification of type 1 versus type 2 diabetes using electronic health record data. Diabetes Care 2013 Apr;36(4):914-921. [doi: 10.2337/dc12-0964] [Medline: 23193215]
27. Shao H, Thorpe LE, Islam S, et al. Developing a computable phenotype for identifying children, adolescents, and young adults with diabetes using electronic health records in the DiCAYA network. Diabetes Care 2025 Jun 1;48(6):914-921. [doi: 10.2337/dc24-1972] [Medline: 40163581]
28. Lakens D. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. Soc Psychol Personal Sci 2017 May;8(4):355-362. [doi: 10.1177/1948550617697177] [Medline: 28736600]
29. Barker LE, Luman ET, McCauley MM, Chu SY. Assessing equivalence: an alternative to the use of difference tests for measuring disparities in vaccination coverage. Am J Epidemiol 2002 Dec 1;156(11):1056-1061. [doi: 10.1093/aje/kwf149] [Medline: 12446263]
30. Limentani GB, Ringo MC, Ye F, Berquist ML, McSorley EO. Beyond the t-test: statistical equivalence testing. Anal Chem 2005 Jun 1;77(11):221A-226A. [doi: 10.1021/ac053390m] [Medline: 15957231]

XSL•FO
RenderX

31. Tatem KS, Romo ML, McVeigh KH, et al. Comparing prevalence estimates from population-based surveys to inform surveillance using electronic health records. Prev Chronic Dis 2017 Jun 8;14:E44. [doi: 10.5888/pcd14.160516] [Medline: 28595032]
32. Wable Grandner G, Dickin K, Kanbur R, Menon P, Rasmussen KM, Hoddinott J. Assessing statistical similarity in dietary intakes of women of reproductive age in Bangladesh. Matern Child Nutr 2021 Apr;17(2):e13086. [doi: 10.1111/mcn.13086]
33. Walker E, Nowacki AS. Understanding equivalence and noninferiority testing. J Gen Intern Med 2011 Feb;26(2):192-196. [doi: 10.1007/s11606-010-1513-8] [Medline: 20857339]
34. Hohman KH, Zambarano B, Klompas M, et al. Development of a hypertension electronic phenotype for chronic disease surveillance in electronic health records: key analytic decisions and their effects. Prev Chronic Dis 2023 Sep 14;20:E80. [doi: 10.5888/pcd20.230026] [Medline: 37708339]
35. Lenoir KM, Wagenknecht LE, Divers J, et al. Determining diagnosis date of diabetes using structured electronic health record (EHR) data: the SEARCH for diabetes in youth study. BMC Med Res Methodol 2021 Oct 10;21(1):210. [doi: 10.1186/s12874-021-01394-8] [Medline: 34629073]
36. Tan H, Lim R, Ang P, et al. Phenotyping diabetes mellitus on aggregated electronic health records from disparate health systems. Pharmacoepidemiology 2023;2(3):223-235. [doi: 10.3390/pharma2030019]
37. World Health Organization (WHO). Health in sustainable development planning: the role of indicators. Policy Commons. 2002. URL: https://policycommons.net/artifacts/589829/health-in-sustainable-development-planning/1568949/ [accessed 2025-10-13]
38. Tulchinsky TH, Varavikova EA. Measuring, monitoring, and evaluating the health of a population. New Public Health 2014 Oct 10:91-147. [doi: 10.1016/B978-0-12-415766-8.00003-3]
39. Broyles D, et al. Shared longitudinal health records for clinical and population health. In: Dixon BE, editor. Health Information Exchange, 2nd edition: Academic Press; 2023:257-273. [doi: 10.1016/B978-0-323-90802-3.00025-3]
40. Etches V, Frank J, Di Ruggiero E, Manuel D. Measuring population health: a review of indicators. Annu Rev Public Health 2006;27:29-55. [doi: 10.1146/annurev.publhealth.27.021405.102141] [Medline: 16533108]
41. Health indicators: conceptual and operational considerations. Pan American Health Organization. 2018. URL: https://www.paho.org/en/documents/health-indicators-conceptual-and-operational-considerations [accessed 2025-10-13]
42. Allen KS, Valvi N, Gibson PJ, McFarlane T, Dixon BE. Electronic health records for population health management: comparison of electronic health record-derived hypertension prevalence measures against established survey data. Online J Public Health Inform 2024 Mar 13;16:e48300. [doi: 10.2196/48300] [Medline: 38478904]

Abbreviations

CP: computable phenotype
DiCAYA: Diabetes in Children, Adolescents, and Young Adults
DM: diabetes mellitus
DSS: Diabetes Surveillance System
EHR: electronic health record
HbA_{1c}: glycated hemoglobin
HIE: health information exchange
HIPAA: Health Insurance Portability and Accountability Act
***ICD-10*:** *International Classification of Diseases, Tenth Revision*
INPC: Indiana Network for Patient Care
MENDS: Multi-State Electronic Health Record–Based Network for Disease Surveillance
MSA: metropolitan statistical area
T1D: type 1 diabetes
T2D: type 2 diabetes
TOST: two one-sided tests

Edited by A Mavragani; submitted 30.06.25; peer-reviewed by K Adegoke, S Goyal; revised version received 22.09.25; accepted 24.09.25; published 27.10.25.

Please cite as:
Wandai ME, Allen KS, Wiensch A, Price J, Dixon BE
Equivalence of Type 2 Diabetes Prevalence Estimates: Comparative Study of Similar Phenotyping Algorithms Using Electronic Health Record Data
JMIR Public Health Surveill 2025;11:e79653
URL: https://publichealth.jmir.org/2025/1/e79653
doi:10.2196/79653

XSL•FO
RenderX

© Muchiri E Wandai, Katie S Allen, Ashley Wiensch, John Price, Brian E Dixon. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing Usefulness of the Dashboard Instrument to Review Equity (DIRE) Checklist to Evaluate Equity in Public Health Dashboards: Reliability Study

Paulina Sosa[1], DrPH, MPH; Emir A Syailendra[2], MS, MD; Harold P Lehmann[2], MD, PhD; Hadi Kharrazi[1,2], MD, PhD

[1]Department of Health Policy and Management, Bloomberg School of Public Health, Johns Hopkins University, 615 N Wolfe St, Baltimore, MD, United States

[2]Section of Biomedical Informatics and Data Science, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Paulina Sosa, DrPH, MPH
Department of Health Policy and Management, Bloomberg School of Public Health, Johns Hopkins University, 615 N Wolfe St, Baltimore, MD, United States

Abstract

Background: The COVID-19 pandemic was a critical time for public health, and though dashboards remained a source of critical health information for decision-makers, key gaps in equity-based decision support were revealed. The DIRE (Dashboard Instrument to Review Equity) Framework and Checklist tool was developed to be a practical tool for public health departments to use in evaluating equity-based decision support mechanisms in their dashboards.

Objective: The objective of this agreement and reliability study was to validate the DIRE Checklist tool as a practical and reliable instrument for data practitioners to use in evaluating dashboards.

Methods: This study was divided into 5 steps to conduct the necessary analysis for agreement and reliability. Step 1 completed the development of the DIRE Checklist tool in Qualtrics (Qualtrics International Inc). Step 2 focused on the parameters required for the selection of the 26 US state–based dashboards. Step 3 was the user testing and assessment process during which each reviewer applied the DIRE tool to each dashboard. Step 4 involved conducting different assessment methods to specifically calculate the comparative analysis, interrater agreement, intraclass correlation coefficients, and the cosine similarity for the Qualtrics, reviewer, and categorical scores. Finally, Step 5 involved conducting any qualitative assessment required on the notes.

Results: A total of 26 dashboards were evaluated using the DIRE Checklist tool by 2 reviewers. The overall percentage comparison for the Qualtrics Score was 31.7% (28.24/89) for Reviewer 1 and 41.8% (37.16/89) for Reviewer 2, resulting in a relative percent agreement of 72.7%. Additionally, the categorical scores showed substantial to high agreement across most categories based on percent agreement within each category. The intraclass correlation coefficient scores indicated varying levels of agreement across different categories, with good agreement observed for the Qualtrics score.

Conclusions: The reliability and agreement result of the study confirmed strong performance of the DIRE checklist tool. The scores calculated were evaluated consistently and reliably by both raters—demonstrating the DIRE Checklist tool's ability to robustly evaluate different dashboards across a number of different categories and parameters.

(JMIR Public Health Surveill 2025;11:e71094) doi:10.2196/71094

KEYWORDS

public health; dashboards; equity; COVID-19; informatics; public health informatics; public health data; decision support

Introduction

The COVID-19 pandemic provided a pivotal juncture to assess the reliability and effectiveness of traditional public health practices [1]. Between 2020 and 2023, public health departments, decision-makers, and data practitioners were faced with time-sensitive information requiring urgent decisions to be made. Public health dashboards played a key role in the decisions made by public health officials [2]; however, shortcomings in the design and implementation of these dashboards created challenges in decision-making, thus increasing the possibility of unintentional and undesired outcomes [3,4]. Specifically, while dashboards were the focal source for many of these critical decisions [5], several gaps in data and information limited their effectiveness as a tool for decision support [6].

Several studies have confirmed that minority and vulnerable populations in the United States faced higher rates of hospitalization and death during the COVID-19 pandemic, but

XSL•FO
RenderX

these populations (eg, Black communities, Hispanic communities, and Indigenous communities) continued to receive less funding, vaccination opportunities, and health care access [7-10]. The evidence of public health inequities during the COVID-19 pandemic has further supported the efforts to integrate and communicate such health disparity information in public health dashboards.

Incorporating basic health equity information (eg, stratification of outcomes by demographics and geography) in public health dashboards has been shown to be key in communicating potential disparities across population groups and communities [11]. Additionally, to more cohesively address equity gaps, data sources representing social determinants of health (SDOH), also referred to as social and structural drivers of health, could provide additional equity dimensions (eg, housing, education, income, and employment) for public health decision-makers to address the needs of communities affected by health inequities [12]. Previous studies have confirmed the need for data practitioners to integrate data representing these gaps into public health dashboards [12].

To address the integration of health equity data and information in public health dashboards, in a previous study, the Dashboard Instrument to Review Equity (DIRE) framework was developed and refined [11]. The refinement of the DIRE framework was accomplished by collecting quantitative and qualitative data from public health data practitioners and decision-makers, which was [11] used to develop a user-friendly checklist (the DIRE checklist) that can be used by data practitioners to apply the DIRE framework principles to their public health dashboards.

The DIRE checklist gives data practitioners and developers a streamlined and useful tool to ensure that equity-based data and decision support guidelines are included and added to public health dashboards. Despite these promising potential outcomes, previous studies have not assessed the usability and reliability [13] of the DIRE checklist when applied by different users to real-world public health dashboards (eg, COVID-19 dashboards). Once reliability and usability are proven, a tool like DIRE could transform the preparedness of dashboards, visualizations, and potentially decision support for future emergencies.

This study aims to assess the reliability of the DIRE checklist as a practical instrument for data practitioners to assess the integration of health equity information in their own dashboards [13]. To achieve this, we applied the DIRE checklist to available US COVID-19 and respiratory dashboards and evaluated its reliability through 2 independent raters while also identifying general areas for improvement. In choosing dashboards for evaluation, we constructed a representative set of state-based dashboards, enabling us to provide a qualitative assessment of the current state of dashboards in incorporating health equity information. The findings of this study could be used to unveil common gaps in COVID-19 dashboards in representing health equity information for decision-making purposes. The DIRE checklist provides data practitioners with an approach to implement equity-based data guidelines into their dashboards to help decision-makers identify timely trade-offs to prevent unintended inequities during similar future emergencies.

Methods

The DIRE Checklist

In our previous studies, the checklist was adapted to directly reflect the DIRE framework components, resulting in 17 questions—recording the presence of six high-level DIRE framework categories: (1) data availability and sources, (2) visualizations and analyses, (3) human-computer interface (HCI), (4) decision support visualizations, (5) equity-based decision support, and (6) community interventions (Multimedia Appendix 1). Upon adaptation from the framework, the DIRE checklist underwent a series of internal reviews and edits by several users with varying backgrounds in public health and dashboard development [11].

Upon completion of the DIRE checklist design and development, the checklist was implemented in Qualtrics (Qualtrics International Inc [14]), a web-based survey administration software, for ease of use and scoring. The DIRE checklist is specifically built to be used by data practitioners, regardless of their public health background. The checklist gives data practitioners an easy-to-use list of questions to assess if their dashboards meet the DIRE criteria [11] (Multimedia Appendix 2). The study team adhered to the GRRAS (Guidelines for Reporting Reliability and Agreement Studies) in conducting the study and reporting the results [13].

US State-Based Dashboard Selection

Several state-based COVID-19 and respiratory disease dashboards were selected using the purposive sampling technique from across the United States in 2024 [15]. This approach allowed the deliberate selection of dashboards that reflected diversity in key characteristics relevant to the assessment process and analysis thereafter. Dashboards were selected based on several diversity parameters to ensure that the testing of the checklist was applied across different demographics of US states. Dashboards were selected based on diversity and representation of the following parameters: (1) population size, based on the US Census Bureau [16] (ie, large, medium, or small); (2) demographic diversity, based on Census ACS (American Community Survey) 2019 estimates [17] (eg, ethnicity, race, and age); (3) median gross domestic product (GDP) and income of the state [18]; and (4) geographic diversity as defined by the Bureau of Economic Analysis (BEA) Regions (ie, Far West, Great Lakes, Mideast, New England, Plains, Rocky Mountain, Southeast, Southwest, and Territories). By using these criteria, we ensured that our sample represented a wide range of state demographics for more comprehensive testing of the DIRE checklist's health equity measures. Selection of the state-based dashboards was not planned as a factorial design, but as one of overlapping DIRE categories; thus, US-related diversity attributes were represented.

Upon categorization of states by these parameters, we chose a selection of at least 1 - 3 dashboards per region and territory, based on the inclusion of all categories and upon the dashboard being (1) available for public viewing and (2) fitting the definition of a dashboard, "a visual display of the most important information needed to achieve one or more objectives, consolidated and arranged on a single screen so the information

XSL•FO
RenderX

can be monitored at a glance" [5]. Upon completion of this process, the top 26 dashboards that satisfied these parameters were selected for user testing and application of the DIRE checklist (Multimedia Appendix 3).

User Testing and Assessment Process

The study consisted of 2 test users: Reviewer 1 (PS) and Reviewer 2 (EAS). Both reviewers applied the DIRE checklist to the 26 state dashboards to showcase the average "score" of dashboards against the categories of the DIRE framework and to test the checklist's reliability. Each reviewer was provided with the same set of instructions, a list of selected dashboards, and a Qualtrics link for the DIRE checklist to conduct the assessment. This process aimed to (1) evaluate the ease of use of the checklist by assessing the time required to complete it, (2) examine the checklist's reliability by measuring interrater agreement when applied by 2 different users on the same dashboard, (3) evaluate the validity of the tool in measuring what it was built to measure, and (4) assess the general state of currently available US COVID-19 dashboards and whether they address DIRE framework elements.

The primary objective of applying the checklist to the dashboards was to assess the level of agreement between the 2 reviewers. This process not only evaluated the consistency of their evaluations but also confirmed the checklist's clarity and comprehensibility between 2 reviewers with different backgrounds and expertise. By comparing their assessments, we aimed to ensure that both reviewers interpreted the checklist in the same way, thereby confirming its effectiveness as a reliable evaluation tool. Inherently, the validity of the tool was also assessed in this process by confirming whether it measured the components of the DIRE framework and health equity that it was meant to measure.

Dashboard identifiers (eg, states and target population) were replaced in the analysis via a code to avoid possible reviewer bias in the analysis. Each reviewer followed a standard operating procedure (Multimedia Appendix 4) to ensure consistency of methods, conditions, and parameters and to mitigate potential deviations or biases in use. Each reviewer used a shared spreadsheet to track dashboard assessment times, scores, and any relevant observation notes. Each reviewer tracked time, at the start and close of each assessment, to calculate the average time needed to use the checklist. Each reviewer took notes to assess the usability, efficiency, validity, and reliability of the checklist. Additionally, each reviewer used the Qualtrics-based DIRE checklist to mark and submit the present parameters observed in each dashboard, which would automatically calculate the total DIRE checklist score. Finally, each reviewer provided an overall grade for the dashboard using a subjective approach. This process resulted in both quantitative and qualitative data.

Assessment Analysis

Data Collection and Organization

Completed checklists of both users were consolidated and organized into one spreadsheet for comparison and analysis. Raw data collected in Qualtrics from each of the completed checklists were downloaded into a unified spreadsheet. The data were then organized in the spreadsheet for analysis in Tableau (version 2024.2.2; Tableau Software LLC, Salesforce [19]) and R (version 4.2.1; R Foundation for Statistical Computing [20]).

Scoring and Comparative Analysis

Multiple scores were generated using the DIRE checklist for comparative analysis. First, each of the DIRE framework categories and subcategories was mapped to DIRE checklist questions to evaluate relevant aspects of the dashboard. Each category was then analyzed as a whole or as a subcategory by the reviewer (ie, the categorical and subcategorical scores). Second, 2 overall scores were calculated after applying the DIRE checklist. One score was done quantitatively by Qualtrics as the sum of scores captured by all checklist questions (ie, the DIRE checklist or Qualtrics score). The other score was generated qualitatively using the reviewer's subjective assessment of the availability of DIRE attributes and user preferences in navigating the dashboard (ie, the Reviewer score or grade).

Interrater Agreement Measurement

The primary goal of the interrater calculation was to determine the degree to which both reviewers provided consistent ratings when applying the checklist to the same dashboard. High agreement levels would indicate whether the checklist was interpreted similarly by both reviewers across dashboards, which would reinforce the checklist's reliability. Interrater agreement was calculated using several methods to assess the consistency of the DIRE checklist scores, categorical scores, and textual notes between the 2 reviewers. The percentage agreement was also calculated for the reviewer score; however, the subjective nature of this score limited its interpretation. The methods used to measure interrater agreement were (1) percent agreement, (2) intraclass correlation coefficients (ICCs), and (3) cosine similarity.

The comparison of percentage agreement was conducted to quantify and assess the level of agreement between the 2 reviewers for the DIRE checklist and categorical scores. For the first comparison, the average DIRE checklist score per dashboard was divided by the total number of dashboards to yield the percentage of DIRE checklist score by reviewer. For the second comparison, the average categorical score per dashboard (ie, the sum of all subcategories divided by the total possible sum) was divided by the total number of subcategories to yield the categorical score percentage by reviewer. Each of these percentages was then compared between the 2 reviewers to evaluate how closely their assessments aligned. This approach provided a general measure of interrater agreement in the DIRE checklist and categorical scores.

To assess the consistency of the average percentages for each score, the percentage agreement was calculated by comparing the difference between their score percentages to the average of those score percentages. A validated formula [21] was applied to the DIRE checklist and categorical scores across all dashboards to quantify the level of agreement between reviewers. The thresholds that were used to interpret results were based on the threshold recommendations provided by McHugh [21].

XSL•FO
RenderX

Agreement between the 2 reviewers on continuous outcomes was analyzed using the ICC3 and ICC3k approaches [22]. The ICC measure is ideal for scenarios with 2 independent raters evaluating the same participants. ICC accounts for both consistency and absolute agreement in ratings, making it a suitable methodology for evaluating how similarly the 2 reviewers assessed the outcomes. The ICC calculation was performed using R version 4.2.1 [20]. Thresholds used to interpret results were based on the threshold recommendations provided by Bartko [22].

The semantic similarity of the reviewers' textual notes was assessed using the cosine similarity measurement of their vector embeddings. To obtain meaningful representations, we used the all-mpnet-base-v2 Transformer-based model [23], a state-of-the-art sentence embedding model designed to capture deep semantic meanings within text. The model produces dense vector embeddings that effectively capture the nuances of language, allowing the analysis of subtle semantic differences and assessment of the textual agreement between the reviewers. The cosine similarity values were then converted into angles to provide a more intuitive understanding of alignment [24]. Semantic similarity analysis was accomplished using Python 3.12 (Python Software Foundation [25]).

Descriptive Analyses

After calculating the interrater agreement, the reviewers conducted further descriptive analysis to gauge the performance of the dashboards using the checklist. The following descriptive statistics were calculated to summarize the data: average, median, SD, and 2 potential baseline thresholds (ie, an ideal baseline based on the total possible score and a realistic baseline based on the median). Fixed base tertiles were also calculated to categorize the average total percentage between both reviewers across 26 dashboards. Dashboards were organized into 3 tertiles: low (tertile 1: 0% - 33.3%), medium (tertile 2: 33.3% - 67%), and high (tertile 3: 67% - 100%). Deciles were also calculated to visualize the spread of dashboard scores on a map. Descriptive analyses were conducted using the DIRE checklist, categorical, and reviewer scores.

Ethical Considerations

This study was reviewed and deemed exempt by the Johns Hopkins School of Public Health's Institutional Review Board (IRB00018696). This study analyzed publicly available dashboards and did not involve human participants or private health information. No individual-level protected health information was collected in this study.

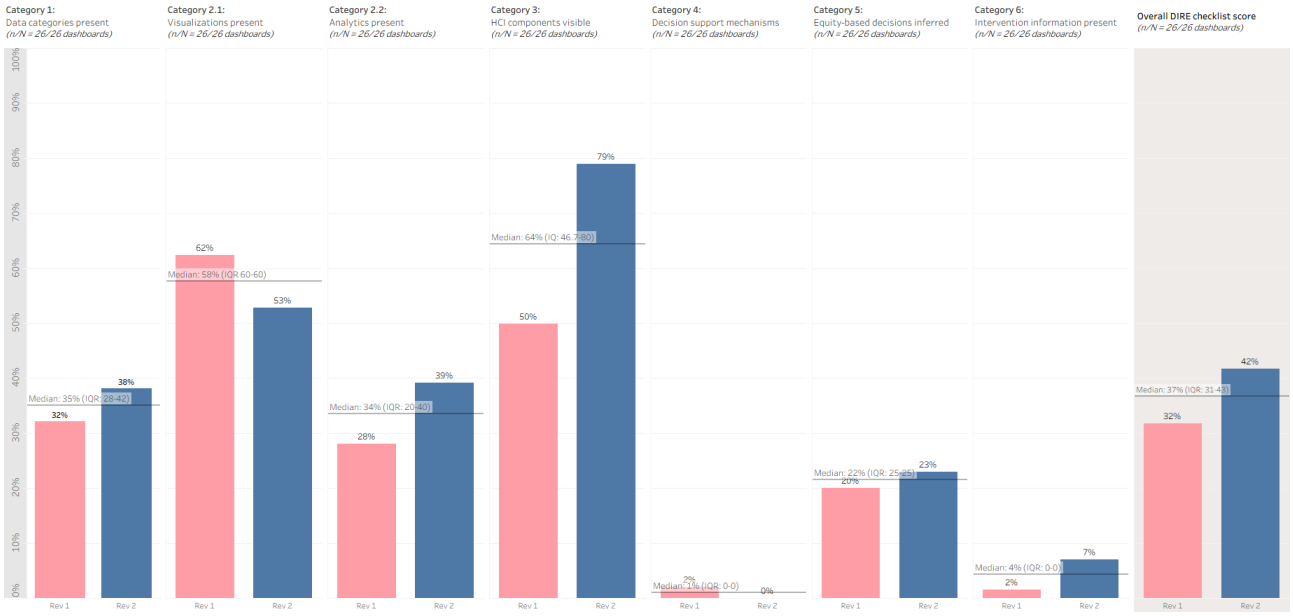
Results

Interrater Agreement

The overall percentage of the DIRE checklist score was 31.7% (28.24/89) for PS and 41.8% (37.16/89) for ES. To assess the consistency of scores between the 2 reviewers, percent agreement was calculated between their 2 average percentages, resulting in a percent agreement of 72.7% (1-(|41.8-31.7|/(31.7+41.8/2)); Multimedia Appendix 5) [21].

The percentage comparison of the categorical score of each reviewer was applied by calculating the percentage comparison for each of the 6 DIRE framework's categories. These comparisons yielded varying levels of agreement across the 6 categories. For example, the percentage comparison for category 1 was 32.1% (96.4/3) for PS and 38.1% (114.4/3) for ES, while category 2.1 showed scores of 62.4% (78/125) and 52.8% (66/125), respectively (Figure 1).

Figure 1. Overall comparison of category and DIRE (Dashboard Instrument to Review Equity) checklist scores across reviewers (N=26 dashboards). HCI: human-computer interface; Rev: Reviewer.



DIRE checklist categories 1 (data representation), 2.1 (visualizations), 2.2 (analyses), 3 (HCI components), and 5 (equity-based decisions) showed moderate to substantial percent agreement. Categories 4 (decision support) and 6 (interventions) were both graded too low to compute an accurate percent agreement; however, both reviewers rated these categories as extremely low (Table 1).

XSL•FO
RenderX

Table . Overall percentage comparison, percentage agreements, and agreement levels.

DIRE[a] category	Category name	Reviewer 1, n/N (%)	Reviewer 2, n/N (%)	Percent agreement (%)[b]	Agreement level [21]
Categories 1 - 6	DIRE checklist score	28.24/89 (31.7)	37.16/89 (41.8)	72.72	Substantial
Category 1	Data representation	96.4/3 (32.1)	114.4/3 (38.1)	82.9	High
Category 2.1	Visualizations	78/125 (62.4)	66/125 (52.8)	83.3	High
Category 2.2	Analyses	35/125 (28)	49/125 (39.2)	66.7	Substantial
Category 3	HCI[c] components	187/375 (49.9)	296/375 (78.9)	54.9	Moderate
Category 4	Decision support	1/50 (2)	0/50 (0)	–100	High
Category 5	Equity-based decisions	20/100 (20)	23/100 (23)	86	High
Category 6	Interventions	4/253 (1.6)	18/253 (7.1)	–27.3	High

[a]DIRE: Dashboard Instrument to Review Equity.

[b]Formula: 1-(|PS%- ES%|/(ES%+PS%/2.

[c]HCI: human-computer interface.

ICC results showed varying levels of agreement across different categories, with moderate to good agreement in the DIRE Qualtrics score (ICC3=0.47; ICC3k=0.64) and category 1 (data representation; ICC3=0.59; ICC3k=0.74), indicating consistency between the 2 reviewers. In contrast, categories 2 (visualizations), 3 (HCI components), and 4 (decision support) showed slight agreement (ICC3=0; ICC3k=0) due to low variation in scores between the 2 reviewers, making it difficult for the ICC to capture any meaningful agreement. Notably, categories 5 (equity-based decisions) and 6 (interventions) showed fair to moderate agreements, with ICC values ranging between 0.34 and 0.57 (Table 2).

XSL•FO
RenderX

Table . Intraclass correlation cooefficient calculations for the Dashboard Instrument to Review Equity checklist and category scores.

DIRE[a] category and type	ICC[b] (95% CI)	*P* value	Agreement
DIRE checklist score (overall)			
ICC3	0.47 (0.10 to 0.73)	.007	Moderate
ICC3k	0.64 (0.19 to 0.84)	.007	Moderate
1 - Data representation			
ICC3	0.59 (0.26 to 0.80)	<.001	Moderate
ICC3k	0.74 (0.41 to 0.89)	<.001	Good
2.1 - Visualization			
ICC3	0 (–0.39 to 0.39)	.50	Slight
ICC3k	0 (–1.27 to 0.56)	.50	Slight
2.2 - Analyses			
ICC3	0.25 (–0.15 to 0.58)	.11	Fair
ICC3k	0.40 (–0.36 to 0.74)	.11	Fair
3 - HCI representation			
ICC3	0.00 (–0.39 to 0.39)	.50	Slight
ICC3k	0.00 (–1.27 to 0.56)	.50	Slight
4 - Decision support			
ICC3	0.00 (–0.39 to 0.30)	.50	Slight
ICC3k	0.00 (–1.27 to 0.56)	.50	Slight
5 - Equity-based decisions			
ICC3	0.34 (–0.06 to 0.64)	.05	Fair
ICC3k	0.50 (–0.13 to 0.78)	.05	Moderate
6 - Interventions			
ICC3	0.40 (0.02 to 0.69)	.02	Fair
ICC3k	0.57 (0.04 to 0.81)	.02	Moderate

[a]DIRE: Dashboard Instrument to Review Equity.

[b]ICC: intraclass correlation coefficient.

The observed pattern of the cosine results highlighted varying degrees of semantic agreement between the notes. The highest cosine similarity was observed for dashboard 9 (0.96), indicating a strong alignment between the 2 reviewers' notes. Dashboard 2 (0.69), 13 (0.67), and 26 (0.66) also exhibited moderate similarity levels, suggesting that the notes share some common features, but with more variation. Other dashboards showed lower similarities between the reviewers' notes (Table 3).

XSL•FO
RenderX

Table . Cosine similarity of reviewer notes across dashboards (N=26).

Rank	Dashboard number	Cosine similarity
1	9	0.96
2	2	0.69
3	13	0.68
4	26	0.67
5	10	0.64
6	16	0.62
7	3	0.61
8	4	0.6
9	11	0.59
10	1	0.58
11	23	0.57
12	6	0.56
13	22	0.53
14	15	0.52
15	20	0.49
16	5	0.49
17	8	0.48
18	14	0.45
19	25	0.44
20	21	0.42
21	18	0.42
22	17	0.42
23	24	0.39
24	7	0.38
25	12	0.34
26	19	0.31

Descriptive Statistics

DIRE categories 1 (data), 2 (visualization and analytics), and 3 (HCI components) were rated comparably high by the 2 reviewers. For category 1, reviewers agreed that clinical data are more frequently present than social or environmental data (Figure S1 in Multimedia Appendix 5). This category is where data sources reflecting different types of SDOH would typically appear, which is an area directly aligned with the DIRE checklist. For category 2, both reviewers agreed that variations of visualization techniques (ie, maps, graphs, and charts) were more prominent than analytic methods (eg, descriptive and predictive). Indeed, most dashboards only displayed descriptive analyses (Figure S2 in Multimedia Appendix 5). For category 3, despite some disagreements, both reviewers agreed that dashboards were at most moderately showcasing HCI components. Reviewers agreed that certain HCI components were more visible than others, including feedback and visual cues, system flexibility, consistency, and simplicity (Figure S3 in Multimedia Appendix 5).

In contrast to categories 1-3, DIRE categories 4 (decision support), 5 (equity decisions), and 6 (interventions) were rated low across the dashboards. For category 4, reviewers agreed that all dashboards were lacking any form of decision support (Figure S4 in Multimedia Appendix 5). For category 5, both reviewers agreed that very few dashboards showed any form of equity-based decision-making. However, based on the demographic data presented, Reviewer 1 denoted 20 dashboards, and Reviewer 2 denoted 19 dashboards presenting visualizations where equity-based decisions could potentially be inferred (Figure S5 in Multimedia Appendix 5). For category 6, both reviewers agreed that all dashboards were lacking detailed information on interventions (Figure S6 in Multimedia Appendix 5).

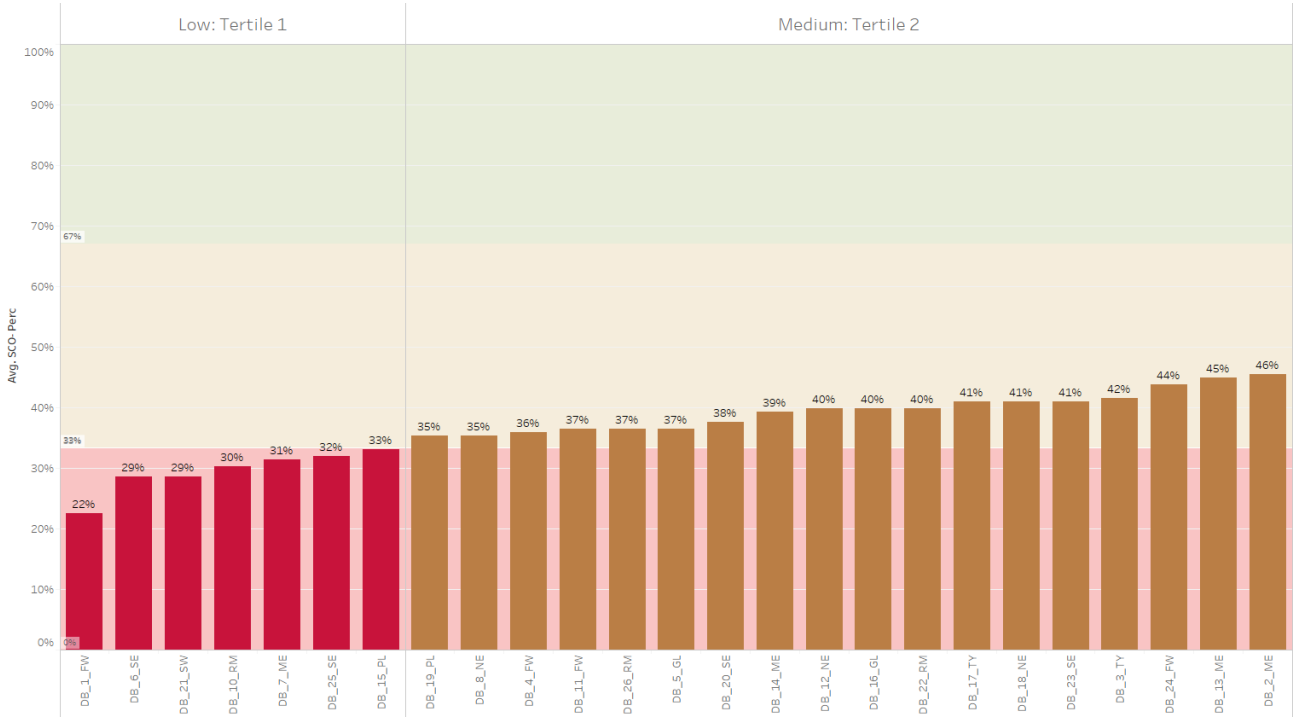
The average DIRE checklist score was calculated as 37% (ie, 32.7 out of a total possible score of 89) when measured across all dashboards. The average reviewer score was 9.24 out of the highest grade of 13. More specifically, the median reviewer score for PS was 9 (IQR 7-10) and for ES was 11 (IQR 8-12), as illustrated in Figure S7 in Multimedia Appendix 5.

XSL•FO
RenderX

Baseline thresholds (ie, both ideal and realistic) were calculated for each score (ie, DIRE checklist, category, and reviewer). The ideal baseline is defined at 80% [21,26]; however, none of the scores reached this threshold. Thus, baselines were developed based on the median scores to provide a more realistic baseline of current dashboard expectations. The median-derived realistic baseline for category 1 was calculated at 35.1% (IQR 32.1%-38.1%), category 2.1 at 57.6% (IQR 52.8%-62.4%), category 2.2 at 33.6% (IQR 28%-39.2%), and category 3 at 64.4% (IQR 49.9%-78.9%). Categories 4-6 were scored too low to develop a relevant baseline (Figure 1). The overall DIRE checklist score's realistic baseline was 37% (Figure 1).

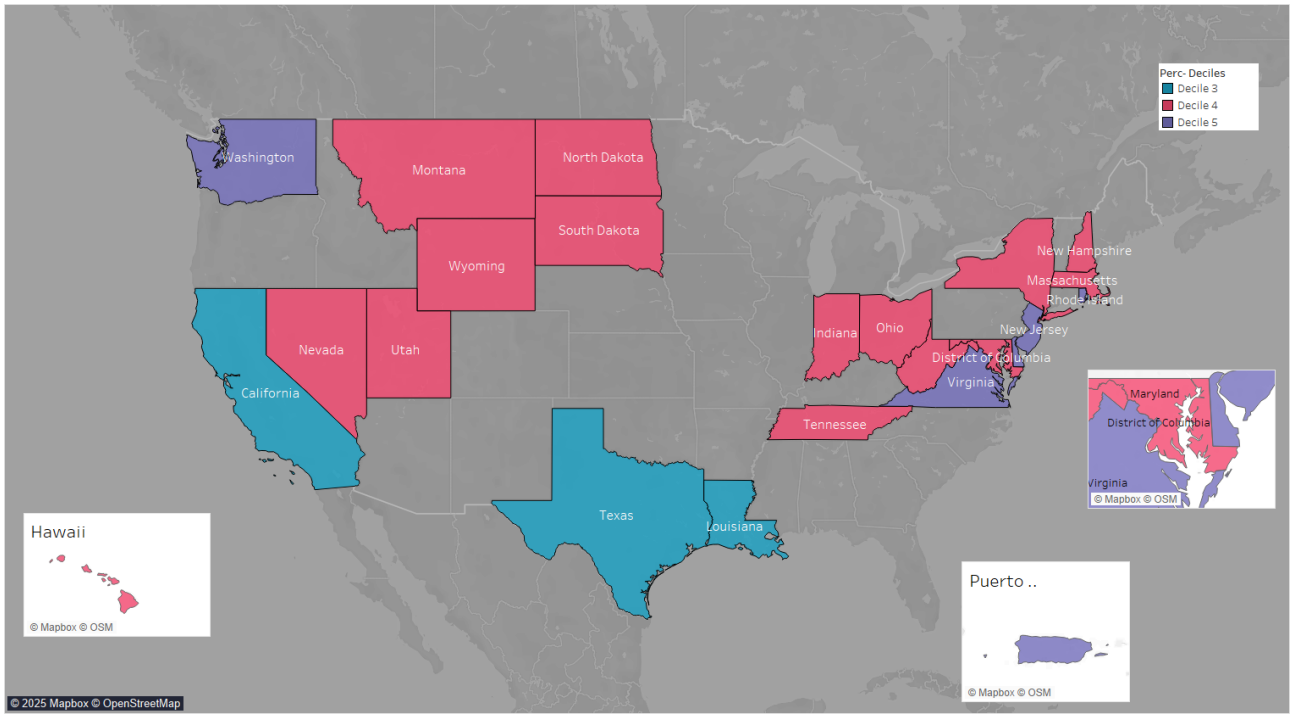
Seven state-based dashboards were scored in the low tertile (tertile 1: 0% - 33.3%) while 18 dashboards were placed in the medium tertile (tertile 2: 33.3% - 67%). None of the dashboards scored high enough to reach the high tertile (tertile 3: 67% - 100%), so hence the highest tertile is not shown in Figure 2. Dashboards were visualized on the US map to depict the spread by decile. Three dashboards ranged in decile 3 (20% - 30%), 15 dashboards fell in decile 4 (31% - 40%), and 7 dashboards fell in decile 5 (41% - 50%; Figure 3). No state-based dashboards ranged above decile 5, with the highest score at 45.5%. Multimedia Appendix 6 includes the full list of dashboards, states, and the total average percent scores.

Figure 2. Average total scores by dashboard and tertile. FW: Far West; GL: Great Lakes; ME: Mideast; NE: New England; PL: Plains; RM: Rocky Mountain; SE: Southeast; SW: Southwest; TY: Territory.



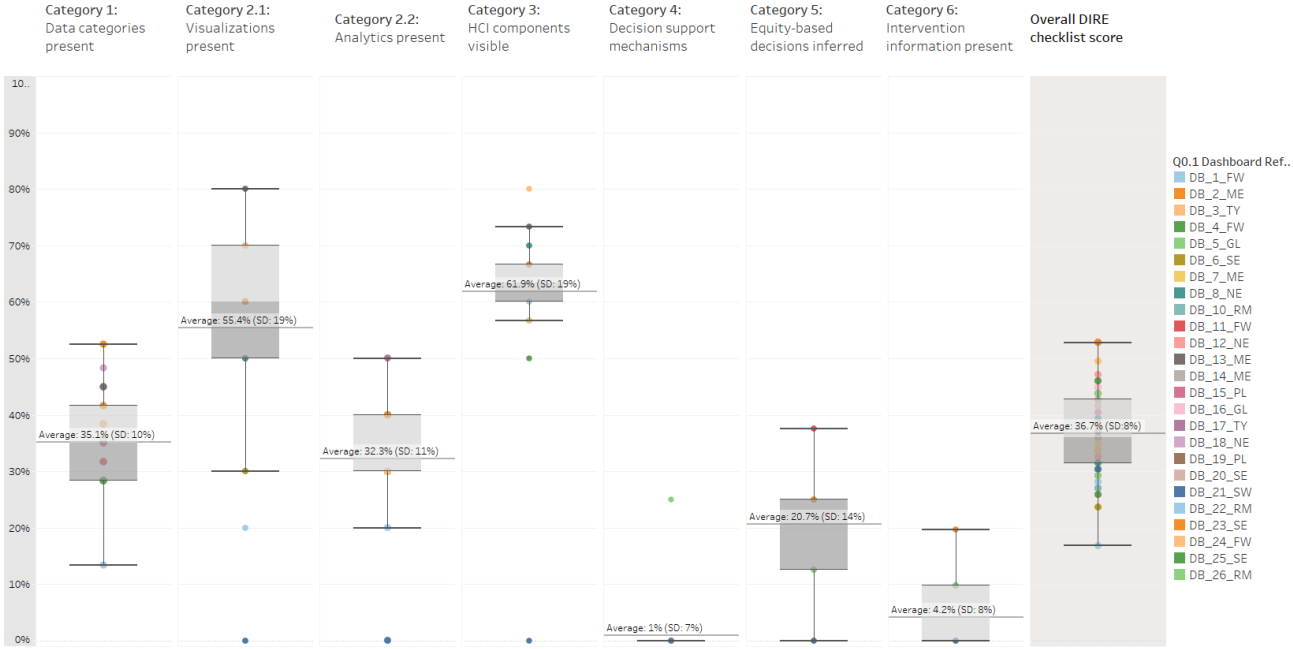
XSL•FO
RenderX

Figure 3. Average total scores by dashboard, state, and decile.



Category and DIRE checklist scores were visualized using boxplots (Figure 4). The boxplots displayed the range of scores, outliers, and the average for each score. Certain categories scored higher (eg, visualizations and HCI components), while other categories scored very low (eg, decision support mechanisms and intervention).

Figure 4. Boxplot of DIRE (Dashboard Instrument to Review Equity) checklist and category scores (N=26 dashboards). FW: Far West; GL: Great Lakes; ME: Mideast; NE: New England; PL: Plains; RM: Rocky Mountain; SE: Southeast; SW: Southwest; TY: Territory.



Comments entered by the reviewers were explored for common themes; however, due to the limited textual data, a full thematic analysis was not feasible (Multimedia Appendix 7).

Discussion

The DIRE Checklist

The goal of this study was to refine and launch the DIRE checklist, a user-friendly decision-support tool designed to help data practitioners and dashboard developers integrate health equity into public health dashboards. While the primary goal was to assess the reliability and validity of the DIRE checklist, this study also highlighted common trends in the design and functionality of state-based COVID-19 dashboards.

XSL•FO
RenderX

DIRE Checklist Performance and Review

DIRE Checklist Reliability and Validity

The interrater agreement results of the study confirmed the reliability of the DIRE checklist. The dashboards were evaluated consistently by both raters, demonstrating the checklist's robustness. The percent agreement of 72.7% indicated a substantial level of agreement within the DIRE categories of data representation, visualizations, and equity-based decisions [21]. The results, however, also identified challenges in other categories where score variability was minimal. The ICC values showed "Moderate" agreement for the DIRE checklist score, indicating the overall consistency of the DIRE checklist. The variability found in certain categories for the ICC calculation reflects differences in reviewer perspectives but also suggests areas where the DIRE checklist may require further refinement. Additionally, minimal score variation in some categories was simply too difficult to calculate because no attributes were present for that category, yet the checklist effectively identified the presence or absence of these attributes. While this study focused on 26 out of more than 50 US states and territories, the purposive sampling that was used provides a diverse sampling of dashboards to establish reliability for the DIRE checklist. These findings underscore the importance of observing reviewer variability as part of the evaluation and application process and confirm the DIRE checklist's capacity to support equity-based dashboard evaluation, while also considering whether certain refinements to the checklist could improve comparative results.

In this study, the validity of the DIRE checklist was confirmed via the application of the checklist. The process of application demonstrated whether intended parameters of health equity data representation in dashboards were present according to the categories of the DIRE framework. The checklist aligned with the DIRE framework and detected variability between dashboard content and context, thus confirming construct validity. The checklist also adequately covered key components of health equity and the DIRE framework in its assessment of dashboards, thus confirming content validity. Given this demonstrated validity of the checklist to determine whether DIRE categories were sufficiently addressed in the dashboard, the checklist can be used to strengthen and enhance the readiness of public health dashboards to support decision-makers with health equity data and decision support in public health emergencies and routine surveillance.

Implications for Public Health Departments

One of the key roles of dashboards is to "employ a visual medium for communication information… relevant for a concrete problem." [5] Indeed, while dashboards are a critical part of communication, dashboards are also one of the most effective ways to communicate data for decision-making, compared to data reports or briefs. The reason for this is due to a dashboard's ability to present relevant data that can visually provide information on inequities and to integrate real-time aspects of predictive and diagnostic tools to aid in decision support [27]. While several stakeholders are involved, decision support assists the decision-maker to make critical decisions that can support community efforts and interventions, leading to the anticipated result of a healthier community.

The reliability of the DIRE checklist as an equity-based evaluation tool supports its broader implications for public health dashboards. The study findings underscore the need for standardized equity assessment tools, like the DIRE checklist, in dashboard development to ensure a comprehensive evaluation of health disparities and inequities that disproportionately affect underserved communities [27]. As shown in Figures 2-4, most dashboards had an average total score of 31% - 60% (ie, medium tertile), with no dashboards scoring higher than an average of 45%. These results highlight the need for a sustainable and standardized tool to address health equity data and decision support in public health dashboards. Integrating the DIRE-based checklist in dashboard design and development could enable public health departments with equity-based decision support during public health emergencies.

A key realization from the findings is the need to improve access to multiple types of data sources that capture dimensions of SDOH to more effectively implement equity-based parameters at all levels of intervention. Collaboration among public health stakeholders is key to integrating DIRE-based metrics into public health dashboards and to better addressing the different types of inequities that affect communities [28]. Collaboration between health care and nontraditional stakeholders is critical to improve access to multiple types of data sources (eg, education and housing), which remains a significant challenge for data practitioners. Access to SDOH-related data sources remains limited, which constrains the comprehensive equity-based insights that the DIRE checklist encourages. By fostering a strong sense of partnership, stakeholders can revolutionize how different data are accessed, integrated, and used for empowering equity-based visualizations and decision support. The study findings confirm the need for strengthened collaboration between decision-makers and data practitioners to ensure meaningful progress in incorporating equity data in dashboards.

To promote equity considerations in public health resource allocations and interventions, policymakers could encourage equity-based evaluations for health care IT systems, including dashboards. This effort can enhance transparency and accountability in addressing health disparities [29,30]. Public health decision-makers might also consider encouraging equity-based evaluations using checklists, such as DIRE for public health systems and dashboards. Indeed, through the integration of the DIRE checklist, equity recommendations could more holistically be integrated into public health dashboards.

The findings of this study also revealed that the DIRE checklist is a practical tool to assess equity-based aspects of public health dashboards across different states. Figure 3 provides a snapshot of the DIRE checklist's assessment of state-based dashboards across the United States and underscores the need for widespread improvement. While the DIRE checklist is not meant to rank dashboards, it is meant to help data practitioners and related stakeholders keep track of the data and information to include in their dashboards. In fact, the DIRE checklist should be used as a guide when building and integrating new health equity components during dashboard design and development. The DIRE checklist's primary use case is to assist dashboard

XSL•FO
RenderX

developers and data teams in public health departments and health departments to identify current and missing data available for dashboard integration. These stakeholders will also be able to identify potential barriers to certain missing data or information. Additionally, the checklist provides guidelines to ensure that dashboard design is considerate of other important equity components, including the incorporation of accessibility, diverse representation of demographic data, decision support visualizations, data transparency of sources and analyses, and integration of continuous user feedback to address disparities and evolving community needs [31,32].

Evaluation and Insights of State-Based Dashboards

Positive Practice 1: Clinical, Demographic, and Geographic Data

This study showed that all 26 dashboards integrated clinical data sources. The most prevalent clinical data sources were electronic health records, hospital discharges, and data provided by other sources, such as laboratories. Examples of clinical variables included test results, hospital bed numbers, and percentage positivity for the disease being monitored. Findings also noted that over 80% (22/26) of dashboards included basic demographic data (ie, age, sex, and ethnicity). Over 88% (23/26) of dashboards included geographic data and visualization features (eg, maps) to highlight hotspots for required support. Although the availability of demographic and geographic data in dashboards is promising for health equity decision support, not all dashboards provided the option to stratify the clinical data, such as patient outcomes or health trends, using such data.

Public health dashboards tend to have simple descriptive visualizations and basic analyses to describe the current situation [33-35]. Although data representations are expanding in dashboards, gaps still exist in data systems used by the dashboards. With the DIRE checklist, we were able to disaggregate the types of data used in the dashboard. While it is a positive trend to see an increase in use of demographic and geographic data, this discovery underscores the need to identify solutions to remove barriers to data accessibility for other equity-relevant data sources, such as SDOH data (eg, housing, transportation, income, education, and employment) that would amplify equity in public health dashboards.

Positive Practice 2: Range of Dashboard Visualization Techniques

Results showed that at least 20 (out of 26) dashboards integrated maps, graphs, charts, and tables showcasing a wide range of visualization techniques. Different visualizations show different aspects of the data, thus providing an interactive approach to support decision-making regarding health equity challenges. Given the use of commercial visualization tools by many of the dashboards, most dashboards provided the same basic visualization features for each of the techniques (eg, most maps included a zoom option). Despite the increasing use of artificial intelligence in enhancing data visualization techniques [36,37], none of the reviewed dashboards incorporated such features to provide additional information on health equity challenges.

Positive Practice 3: HCI Components (User Interface Navigation)

To enhance the user experience and improve efficiency, developers have been motivated to integrate interactivity across pages and visualizations within their public health dashboards [38,39]. The majority of the reviewed state dashboards presented the following user interface components: (1) feedback cues and icons (ie, a progress bar as you scroll through the dashboard), (2) affordance or visual cues (ie, a clickable hyperlink underlined in blue), (3) flexibility of system (ie, interactive maps or graphs vs static images of data), and (4) visibility of components (ie, a “click here to filter” button that explains how it customizes the results or maps). These are critical attributes for user interface and underscore the considerable improvements across dashboard interface and interactivity in general. However, reviewers could not identify any dashboard that presented all possible HCI components (ie, 14 features) as listed in the DIRE checklist. Additionally, 10 of these HCI components were minimally, or not at all, integrated into the dashboards, which included features, such as error prevention, system accessibility, and system documentation of user actions. These additional components might require additional information, training, and direct support for developers, presenting a unique opportunity to strengthen future dashboards as a tool to support decision-making regarding public health equity.

Areas of Growth 1: Widespread Absence of Decision Support and Equity-Based Decisions

The study findings revealed that only 21.5% (5/26) of state dashboards incorporated equity-related information. Moreover, none of the dashboards incorporated equity-based decision support mechanisms. Indeed, none of the reviewed dashboards had any form of decision support built in, and only 4.4% (1/26) of the dashboards included general information on community-based interventions or related information. Results also revealed a general lack of comprehensive equity factors incorporated in existing public health dashboards. While some dashboards incorporated basic demographic stratification, they still failed to capture the multifaceted aspects of health equity adequately [27]. Previous studies have emphasized the importance of data disaggregation and visual representation in identifying disparities [27]. However, the study findings underscored the need for explicit equity-focused metrics and interpretability features to facilitate targeted interventions and policy decisions [27]. Overall, despite an increase in equity-based data and metrics in some of the reviewed dashboards, particularly concerning demographic data [40], more work is needed by state-based public health departments to implement the required adjustments to best address health equity needs of underserved communities in state and local dashboards.

The lack of equity-related decision support features in public health dashboards is a widespread issue [12] and raises 2 critical points of consideration. First, the widespread absence of decision support, equity-based decisions, and community intervention recommendations in all reviewed state-based dashboards raises concerns on whether these gaps reflect oversights by data practitioners or broader systemic issues within public health

XSL•FO
RenderX

department operations. One possibility is whether public health departments inherently isolate decision support to a selective group of decision-makers, creating an internalized boundary or siloed approach, which excludes data practitioners. Alternatively, this issue may instead indicate a need within public health departments to collaborate more effectively with data practitioners to ensure they have the information required to integrate equity-based decision support components. Second, data practitioners may not have access to the data needed to develop these mechanisms. Data source fragmentation makes it difficult, if not impossible, to collect and use additional equity-based data required for decision-support visualizations and community interventions.

Areas of Growth 2: Innovatively Improving Data Accessibility (Including Social and Environmental)

The findings showed that many dashboards presented high-level social (eg, demographics) and environmental data (eg, geographical boundaries), but only 9 dashboards included census, survey, or economic data. Only one dashboard presented SDOH data, such as neighborhood or housing data, the social vulnerability index (SVI) [41], or assisted living residence data. These unique types of data sources are critical to elevating equity-based visualizations in dashboards, beyond current expectations but do also highlight the data access barrier faced by data practitioners. While data fragmentation is a well-known challenge, the DIRE checklist lists these variables to strengthen the type of equity-based decision support developers can integrate as innovatively as possible. Data fragmentation remains a significant challenge in public health data systems, requiring a national collaboration and solution to bridge the gap between data systems and increase communications between systems.

Areas of Growth 3: Expansion to Different Types of Analytics

Another finding highlighted the absence of certain analytics in dashboards, particularly the use of diagnostic (average of 12/26 dashboards), predictive (average of 2/26 dashboards), and prescriptive (0/26 dashboards) analytics. While the analytic methods used for reporting public health data have grown considerably, our results revealed gaps in providing such analytics in public health dashboards when addressing health equity data. First, despite the expansion of prescriptive and predictive analytics, especially in the wake of AI innovations [36,37,42], these analytics were not integrated into the dashboards. Second, despite the general acceptance of using traditional analytics for public health dashboards and decision support [43], none of the dashboards had leveraged the traditional analytical methods (eg, regression methods) to provide more insightful results regarding health equity. Given the additional training, resources, and support for data teams needed to integrate such analytical methods [44], it is important for data practitioners and public health departments to start integrating the use of these analytics and the supporting data sources, now rather than during a public health emergency. It is important to note that several barriers can play a role in a health department's ability to provide certain data, analyses, or graphics, including capacity, budgets, and data accessibility.

Limitations and Future Research

Despite the study results supporting the ease of use and reliability of the DIRE checklist in assessing the utility of public health dashboards for health equity decision-making, this study has limitations that should be considered when interpreting its results. First, the assessment of the DIRE checklist's validity was limited to content and construct validities. Although the checklist is not considered a survey and should instead be treated as a guideline, future studies should further evaluate other dimensions of the checklist's validity that may enhance its usability (eg, making the checklist shorter without losing any of the equity categories as defined by the DIRE framework). Second, the DIRE checklist's reliability was assessed by calculating the interrater reliability measures between 2 reviewers. Additional prior knowledge and exposure to the tool may have affected the results of the study, even though this was mitigated through the operating procedure. Additional reviewers, especially with diverse backgrounds, technical expertise, and operational barriers, can increase the strength of the DIRE checklist's reliability assessment. Third, qualitative results of the study (eg, textual notes, reviewer score, and grade) might have been affected by the inherent bias of the reviewers. Although this study mitigated bias as much as possible by clearly identifying the evaluation steps and instructions for notating observations, an inevitable level of bias is still built into this approach. This potential bias may have also affected quantitative results, specifically certain category scores (eg, HCI category showed high variability among the reviewers). This observation highlights the need to clarify the definition and requirements of each checklist item to mitigate any potential reviewer bias and individual interpretations. Fourth, this study contained a limited sample size of US-based COVID-19 dashboards (ie, 26 state dashboards, out of a total of 50 states and 16 territories), potentially impacting the generalizability of the study findings regarding the status of health equity data integration in other countries or non–COVID-19 dashboards.

To advance the DIRE checklist and strengthen the utility of public health dashboards for improving health equity during an emergency response, future studies could (1) apply and evaluate the checklist against a larger sample of dashboards representing more diverse populations and geographical boundaries (eg, city, county, and international dashboards), (2) conduct longitudinal studies to evaluate the long-term impact of the checklist on dashboard improvements in reducing health disparities and promoting equitable interventions and decisions, and (3) explore the development and validation of other technologies (which could include solutions, such as NLP technologies) to streamline and facilitate an easier use of the application of the DIRE checklist for dashboard developers.

Conclusion

This study verified the reliability of the DIRE checklist by having 2 reviewers apply it to 26 US statewide COVID-19 dashboards. The results demonstrated the checklist's strong reliability and its potential to assist data practitioners in integrating health equity and decision support into their future dashboards. The study also underscored key strengths and opportunities to strengthen the DIRE checklist for broader public

XSL•FO
RenderX

use, while highlighting valuable insights on the current state of public health dashboards in health equity, decision support, and public health preparedness. The findings of this study suggest that most dashboards lack adequate integration of equity considerations, decision support, and community intervention recommendations, underscoring the need to continue strengthening these areas. The DIRE checklist offers a practical tool that can address these gaps, thus empowering public health departments to strengthen their dashboards for routine public health surveillance as well as address health equity challenges in future public health emergencies.

Acknowledgments

We acknowledge the support of public health officials for their guidance and advice on the development of this study and the DIRE (Dashboard Instrument to Review Equity) tool. Their time, support, and expertise were critical. No generative AI was used to generate, analyze, or write the study's content for this manuscript.

Data Availability

Data are presented in the main manuscript. Deidentified rating sheets, the DIRE checklist, and analysis code are available for public use and review in the manuscript tables and Multimedia Appendices.

Authors' Contributions

PS contributed to conceptualization, methodology, investigation, data analysis, and writing–original draft, formatting, and writing–editing. EAS was involved in research, data analysis, and writing–original draft. HL was responsible for conceptualization, project supervision, and writing–review and editing. HK provided supervision and contributed to writing–review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
DIRE (Dashboard Instrument to Review Equity) framework.
[PDF File, 730 KB - publichealth_v11i1e71094_app1.pdf]

Multimedia Appendix 2
DIRE (Dashboard Instrument to Review Equity) checklist tool.
[PDF File, 173 KB - publichealth_v11i1e71094_app2.pdf]

Multimedia Appendix 3
Table of state dashboards used.
[XLSX File, 19 KB - publichealth_v11i1e71094_app3.xlsx]

Multimedia Appendix 4
Dashboard review: standard operating procedure (SOP).
[DOCX File, 8 KB - publichealth_v11i1e71094_app4.docx]

Multimedia Appendix 5
Analysis of results by category.
[PDF File, 1138 KB - publichealth_v11i1e71094_app5.pdf]

Multimedia Appendix 6
Average scores by dashboard and by quartile.
[PNG File, 52 KB - publichealth_v11i1e71094_app6.png]

Multimedia Appendix 7
Thematic analysis.
[DOCX File, 15 KB - publichealth_v11i1e71094_app7.docx]

References

XSL•FO
RenderX

1. Gostin LO, Cohen IG, Koplan JP. Universal Masking in the United States: The Role of Mandates, Health Education, and the CDC. JAMA 2020;324(9):837-838. [doi: 10.1001/jama.2020.15271]
2. Dong E, Ratcliff J, Goyea TD, et al. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 Dashboard: data collection process, challenges faced, and lessons learned. Lancet Infect Dis 2022 Dec;22(12):e370-e376. [doi: 10.1016/S1473-3099(22)00434-0] [Medline: 36057267]
3. Meesters K, Wijnen J, Visser M, Boersma K. The challenge and value of dashboard development during the COVID-19 pandemic. 2022 Presented at: 2022 IEEE Global Humanitarian Technology Conference (GHTC); Sep 8-11, 2022; Santa Clara, CA p. 9-16. [doi: 10.1109/GHTC55712.2022.9910610]
4. Badker R, Miller K, Pardee C, et al. Challenges in reported COVID-19 data: best practices and recommendations for future epidemics. BMJ Glob Health 2021 May;6(5):e005542. [doi: 10.1136/bmjgh-2021-005542] [Medline: 33958393]
5. Few S. Information Dashboard Design: The Effective Visual Communication of Data: O'Reilly Media; 2006.
6. Improving COVID-19 epidemiology through a health equity lens. COVID-19 Real-Time Learning Network. 2024 May 10. URL: https://web.archive.org/web/20240523184104/https://www.idsociety.org/covid-19-real-time-learning-network/COVID-Health-Equity-Resources/improving-covid-19-epidemiology-through-a-health-equity-lens/ [accessed 2024-06-05]
7. Schoch-Spana M. Equity in vaccination: a plan to work with communities of color toward COVID-19 recovery and beyond. Johns Hopkins Bloomberg School of Public Health. URL: https://publichealth.jhu.edu/2021/equity-in-vaccination-a-plan-to-work-with-communities-of-color-toward-covid-19-recovery-and-beyond [accessed 2024-06-05]
8. Nguyen A, Buhr RG, Fonarow GC, Hsu JJ, Brown AF, Ziaeian B. Racial and ethnic disparities and the national burden of COVID-19 on inpatient hospitalizations: a retrospective study in the United States in the year 2020. J Racial Ethn Health Disparities 2025 Aug;12(4):2509-2520. [doi: 10.1007/s40615-024-02069-y] [Medline: 39316343]
9. Addressing disparities in COVID-19 impact on minorities. : US Department of Health and Human Services; 2020 URL: https://www.hhs.gov/sites/default/files/hhs-fact-sheet-addressing-disparities-in-covid-19-impact-on-minorities.pdf [accessed 2024-06-05]
10. Bauchner H, Fontanarosa PB, Flanagin A. Advancing racial equity in U.S. health care: the Commonwealth Fund 2024 state health disparities report. : The Commonwealth Fund; 2024 URL: https://www.commonwealthfund.org/publications/fund-reports/2024/apr/advancing-racial-equity-us-health-care
11. Sosa P, Kharrazi H, Lehmann H. A framework to integrate equity in public health emergency response dashboards: dashboard instrument to review equity (DIRE). Public Health (Fairfax) 2025 Mar;240:182-194. [doi: 10.1016/j.puhe.2024.12.053] [Medline: 39919521]
12. Bell K, White S, Diaz A, et al. Can evidence drive health equity in the COVID-19 pandemic and beyond? J Public Health Pol 2024 Mar;45(1):137-151. [doi: 10.1057/s41271-023-00452-3]
13. Kottner J, Audigé L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011 Jan;64(1):96-106. [doi: 10.1016/j.jclinepi.2010.03.002] [Medline: 21130355]
14. Qualtrics XM. Qualtrics. 2024. URL: https://www.qualtrics.com [accessed 2024-05-05]
15. Etikan I, Musa SA, Alkassim RS. Comparison of convenience sampling and purposive sampling. AJTAS 2016;5(1):1. [doi: 10.11648/j.ajtas.20160501.11]
16. DP05 ACS demographic and housing estimates. United States Census Bureau. URL: https://data.census.gov/table/ACSDP1Y2019.DP05?g=010XX00US&tid=ACSDP1Y2019.DP05 [accessed 2024-01-05]
17. DP05 ACS demographic and housing estimates database, Hispanic or Latino filter. URL: https://data.census.gov/cedsci/table?q=hispanic%20or%20latino%20by%20state&t=Hispanic%20or%20Latino&g=0100000US.04000.001_0200000US1,2,3,4&tid=ACSDP1Y2019.DP05&hidePreview=true [accessed 2024-01-05]
18. GDP by state. Bureau of Economic Analysis. URL: https://www.bea.gov/data/gdp/gdp-state [accessed 2024-01-05]
19. Tableau Software. 2024. URL: https://www.tableau.com [accessed 2024-05-05]
20. R Core Team. R: a language and environment for statistical computing (version 4.2.1). R Foundation for Statistical Computing. 2022. URL: https://www.r-project.org/
21. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282. [doi: 10.11613/BM.2012.031] [Medline: 23092060]
22. Bartko JJ. The intraclass correlation coefficient as a measure of reliability. Psychol Rep 1966 Aug;19(1):3-11. [doi: 10.2466/pr0.1966.19.1.3] [Medline: 5942109]
23. Wolf T, Debut L, Sanh V, et al. Transformers: state-of-the-art natural language processing. 2020 Presented at: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing; Nov 16-20, 2020; online p. 38-45. [doi: 10.18653/v1/2020.emnlp-demos.6]
24. Xu S, Wu Z, Zhao H, et al. Reasoning before comparison: LLM-enhanced semantic similarity metrics for domain specialized text analysis. arXiv. Preprint posted online on Feb 20, 2024 URL: https://arxiv.org/abs/2402.11398 [accessed 2025-11-24]
25. The Python language reference. Python Software Foundation. 2023. URL: http://www.python.org [accessed 2024-06-05]
26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-174. [doi: 10.2307/2529310] [Medline: 843571]

XSL•FO
RenderX

27. Gallifant J, Kistler EA, Nakayama LF, et al. Disparity dashboards: an evaluation of the literature and framework for health equity improvement. Lancet Digit Health 2023 Nov;5(11):e831-e839. [doi: 10.1016/S2589-7500(23)00150-4] [Medline: 37890905]
28. Yi S, Burke C, Reilly A, Straube S, Graterol J, Peabody CR. Designing and developing a digital equity dashboard for the emergency department. J Am Coll Emerg Physicians Open 2023 Aug;4(4):e12997. [doi: 10.1002/emp2.12997] [Medline: 37397184]
29. Martino SC, Mathews M, Beckett MK, et al. Development of a Medicare plan dashboard to promote health equity. Am J Manag Care 2023 Mar 1;29(3):e91-e95. [doi: 10.37765/ajmc.2023.89335] [Medline: 36947022]
30. Agniel D, Cabreros I, Damberg CL, Elliott MN, Rogers R. A formal framework for incorporating equity into health care quality measurement: study proposes a framework to incorporate equity into quality measurement. Health Aff (Millwood) 2023 Oct;42(10):1383-1391. [doi: 10.1377/hlthaff.2022.01483] [Medline: 37782880]
31. Gallifant J, Zhang J, Whebell S, et al. A new tool for evaluating health equity in academic journals; the Diversity Factor. PLOS Glob Public Health 2023;3(8):e0002252. [doi: 10.1371/journal.pgph.0002252] [Medline: 37578942]
32. Migita D, Cooper A, Barry D, Bettinger B, Tieder A, Sharek PJ. Equity dashboards: data visualizations for assessing inequities in a hospital setting. Pediatrics 2023 Mar 1;151(3):e2022058848. [doi: 10.1542/peds.2022-058848] [Medline: 36751915]
33. Rabiei R, Bastani P, Ahmadi H, Dehghan S, Almasi S. Developing public health surveillance dashboards: a scoping review on the design principles. BMC Public Health 2024 Feb 6;24(1):392. [doi: 10.1186/s12889-024-17841-2] [Medline: 38321469]
34. Dixon BE, Dearth S, Duszynski TJ, Grannis SJ. Dashboards are trendy, visible components of data management in public health: sustaining their use after the pandemic requires a broader view. Am J Public Health 2022 Jun;112(6):900-903. [doi: 10.2105/AJPH.2022.306849] [Medline: 35446601]
35. Elgendi M, Howard N, Hussain A, Menon C, Ward R. From ancient times to modern: realizing the power of data visualization in healthcare and medicine. Big Data Anal 2020 Dec;5(1):4. [doi: 10.1186/s41044-020-00047-z]
36. Xu P, Zhang Y, Chen W, et al. AI4VIS: survey on artificial intelligence approaches for data visualization. . Preprint posted online on Jul 19, 2021. [doi: 10.48550/arXiv.2102.01330]
37. Liu Z, Ren D, Wu Y, et al. HAIChart: human and AI paired visualization system. Arxiv. Preprint posted online on Sep 7, 2024. [doi: 10.48550/arXiv.2406.11033]
38. Weissgerber TL, Garovic VD, Savic M, Winham SJ, Milic NM. From static to interactive: transforming data visualization to improve transparency. PLoS Biol 2016 Jun;14(6):e1002484. [doi: 10.1371/journal.pbio.1002484] [Medline: 27332507]
39. Rathod Y. Understanding the evolution of data visualization techniques: from static to dynamic visualizations. IJARCCE 2024 Jan 30;13(2):99-103 [FREE Full text] [doi: 10.17148/IJARCCE.2024.13218]
40. Thorpe LE, Gourevitch MN. Data dashboards for advancing health and equity: proving their promise? Am J Public Health 2022 Jun;112(6):889-892. [doi: 10.2105/AJPH.2022.306847] [Medline: 35446603]
41. United States. Agency for Toxic Substances and Disease Registry. Geospatial Research, Analysis & Services Program. The Social Vulnerability Index (SVI). Centers for Disease Control and Prevention. 2014. URL: https://stacks.cdc.gov/view/cdc/27762 [accessed 2024-05-15]
42. Corpuz JCG. Artificial intelligence (AI) and public health. J Public Health (Oxf) 2023 Nov 29;45(4):e783-e784. [doi: 10.1093/pubmed/fdad074]
43. Fisher S, Rosella LC. Priorities for successful use of artificial intelligence by public health organizations: a literature review. BMC Public Health 2022;22(1):2146. [doi: 10.1186/s12889-022-14422-z]
44. Ramezani M, Takian A, Bakhtiari A, Rabiee HR, Ghazanfari S, Mostafavi H. The application of artificial intelligence in health policy: a scoping review. BMC Health Serv Res 2023 Dec 15;23(1):1416. [doi: 10.1186/s12913-023-10462-2] [Medline: 38102620]

Abbreviations

ACS: American Community Survey
BEA: Bureau of Economic Analysis
DIRE: Dashboard Instrument to Review Equity
GDP: gross domestic product
GRRAS: Guidelines for Reporting Reliability and Agreement Studies
HCI: human-computer interface
ICC: intraclass correlation coefficient
SDOH: social determinants of health
SVI: Social Vulnerability Index

XSL•FO
RenderX

Edited by A Mavragani; submitted 20.08.25; peer-reviewed by MS Mekbib, T Okumura; revised version received 30.10.25; accepted 31.10.25; published 04.12.25.

Please cite as:
Sosa P, Syailendra EA, Lehmann HP, Kharrazi H
Assessing Usefulness of the Dashboard Instrument to Review Equity (DIRE) Checklist to Evaluate Equity in Public Health Dashboards: Reliability Study
JMIR Public Health Surveill 2025;11:e71094
URL: https://publichealth.jmir.org/2025/1/e71094
doi:10.2196/71094

© Paulina Sosa, Emir A Syailendra, Harold P Lehmann, Hadi Kharrazi. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 4.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Nonrandom Missingness in Child Race and Ethnicity Records and the US Federal Data Standards: Pooled Analysis of Community-Based Child Health Studies

Danielle M Krobath[1,2], PhD; Norbert L W Wilson[3], PhD; Adolfo G Cuevas[4,5], PhD; Elena N Naumova[2,6], PhD; Jennifer M Sacheck[7], PhD; Sydney A Tyler[8], BS; Christina D Economos[2], PhD

[1]Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, 915 Greene Street, Columbia, SC, United States

[2]Friedman School of Nutrition Science and Policy, Tufts University, Boston, MA, United States

[3]Duke Divinity School and Sanford School of Public Policy, Duke University, Durham, NC, United States

[4]Department of Social and Behavioral Sciences, School of Global Public Health, New York University, New York, NY, United States

[5]Center for Anti-racism, Social Justice, and Public Health, New York University, New York, NY, United States

[6]Department of Community Health, Tufts University, Medford, MA, United States

[7]Department of Social and Behavioral Sciences, School of Public Health, Brown University, Providence, RI, United States

[8]Department of Public Health and Community Medicine, School of Medicine, Tufts University, Boston, MA, United States

Corresponding Author:
Danielle M Krobath, PhD
Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, 915 Greene Street, Columbia, SC, United States

Abstract

Background: Racism perpetuates the unequal distribution of power, resources, and privilege within and between societies to the detriment of marginalized groups. Racialization involves categorizing people based on traits to which socially constructed meaning and value have been ascribed. In public health, this process can manifest when tracking racial health disparities in children, which requires aggregating parent-reported race and ethnicity data into federally recognized categories. The demographic surveys used to characterize children's identity in the United States mirror those administered in adults and typically follow federal race and ethnicity data standards, which include ambiguous response options (eg, other race), "select all that apply" directives, and open-ended fields followed by a request specification, with limited guidance for coding and interpretation. These methodological challenges could contribute to nonrandom data missingness and misclassification bias and must be resolved to better harmonize historic data, especially given recent revisions to the country's federal race and ethnicity data standards.

Objective: We aimed to explore the prevalence of systematic bias within past, current, and recently revised federal race and ethnicity data standards in the United States and develop a standardized method for improving the reporting of child race and ethnicity in public health research, policy, and practice.

Methods: We developed a replicable decision-making process to uncover racial heterogeneity obscured by key components of US federal race and ethnicity data standards (open-ended and ambiguous response fields). We applied it to a pooled sample of 8 community-based child health studies with 8087 participants and examined changes in the dataset's racial and ethnic diversity.

Results: Overall, 93.11% (7530/8087) of parents provided child race and ethnicity data, with 3.73% (281/7530) identified as other race and 9.72% (732/7530) identified as multiracial. In total, 101 distinct open-ended written responses (eg, "Haitian") were provided. The replicable decision-making process resulted in 4.02% (303/7530) of children being reallocated from their parent-reported race or ethnicity category, of whom 38.6% (117/303) were moved into the Black category based on written responses. Within the multiracial group, we identified 22 unique combinations, including White-Hispanic (269/732, 36.7%) and White-Black (169/732, 23.08%).

Conclusions: These findings demonstrate how the current paradigm of assessing race and ethnicity in the United States may contribute to the erasure and further marginalization of individuals disproportionately enduring the effects of racism. While updated federal race and ethnicity data standards may soon take effect, persistent gaps in demographic and health surveillance will remain. Our data reallocation decision-making process offers a novel and practical framework for harmonizing race and ethnicity data across time, populations, and datasets, emphasizing the relevance and longevity of preexisting datasets and tools. Efforts to build equitable public health surveillance and data systems should expand the survey response options, avoid aggregating diverse populations, and develop new statistical techniques for data analysis.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e65660) doi:10.2196/65660

KEYWORDS

race and ethnicity; health disparities; population health; health surveillance; child health; racism; survey research methods; demography; data harmonization; data crosswalk

Introduction

Conceptualizing Race, Ethnicity, and Health Inequities

Race and ethnicity are social constructs associated with health inequities through the immediate and intergenerational effects of racism, discrimination, and other social hierarchies. Multiple forms of racism permeate society and impede the health of racially minoritized populations, including through interpersonal discrimination, segregated schools and neighborhoods, and discriminatory health care systems [1-6]. These long-standing injustices received renewed global attention in response to the COVID-19 pandemic and ongoing police killings of Black individuals in the United States. Many major institutions (governments, funders, journals, and universities) have also pledged to enact change [7,8]. For instance, the United Nations recently called upon countries to use a "whole-of-government" approach to dismantle structural racism, beginning with the population-level collection, surveillance, and public reporting of racial and ethnic identity data [9].

Global Approaches to Race and Ethnicity Data Collection and Implications for Health Equity

According to Morning [10], who analyzed unpublished data compiled by the United Nations, only 63% of the 141 countries that fielded national census surveys asked about ethnicity, race, or a related identity construct. However, these terms were differently conceptualized and highly conflated. Most national censuses used the term "ethnicity," but "race" and "nationality" were often either excluded or measured with identical response categories. Limited quality or nonexistent racial and ethnic data infrastructure allows systemic racism to manifest, driving racial health inequities in the United States and elsewhere around the world [11,12].

Racialization is the process by which individuals are ascribed to a racial or ethnic group based on both self-identified and socially perceived (ie, the categorization of an individual by members of a society) categorizations. Race and ethnicity are complex and dynamic constructs that change by time, place, and historical context [4,13]. For example, Morning [10] found that 13 nations included a census question assessing "race" alone, 11 of which were New World former slaveholding nations (ie, the United States) or their territories [14]. Conversely, many governments—predominantly in Western Europe and West Africa—prohibit official racial or ethnic categorization by the government due to violent histories built upon racial caste systems. The tangible consequences of racialization on health disparities are numerous. Some of the social and economic ramifications in the United States include unequal resource allocation (eg, health care providers may be less likely to accept Medicaid in low-income, racially segregated neighborhoods where the majority of residents are non-White) [15] as well as disproportionately heightened exposure to stress among Black individuals with darker skin tones compared to their White and lighter-skinned counterparts, which can increase allostatic load (ie, the cumulative "wear and tear" on the body resulting from prolonged stress), [16].

Categorizing people by race is crucial for cultivating representative study samples, developing inclusive and effective policies and programs, and alleviating racial health disparities [17-19]. Williams and Jackson [20] have long substantiated that making progress toward racial equity requires collecting and reporting on racial identity. Given that funders either require or encourage collecting data disaggregated by race and ethnicity and medical journals have growing attention on the subject of racism [4,21], researchers and policy makers across the globe may begin or continue reporting race and ethnicity in the context of health outcomes and disease prevalence [21,22].

Race and Ethnicity Data Systems in the United States

The US government has a well-established demographic surveillance system that simultaneously guides national race and ethnicity data standards. The entities and individuals worldwide that begin collecting race and ethnicity data may turn to investigators experienced in demographic enumeration for methodological insights. While this decision should be made cautiously, the federal data systems in the United States may serve as a starting point for surveilling racial and ethnic health outcomes.

In the United States, the Office of Management and Budget (OMB) Directive 15, *Race and Ethnic Standards for Federal Statistics and Administrative Reporting*, defines the data standards and minimum categories for reporting race and ethnicity for all federal agencies and federally funded health researchers. In 2024, the OMB announced several updates aimed at improving question format, wording, and inclusivity of federal race and ethnicity data standards which are expected to take effect with the 2030 census [23]. Changes will include the creation of a new racial group specifically for Middle Eastern and North African (MENA) individuals and assessing race and Hispanic or Latino ethnicity with a single combined question. However, enumeration of race and ethnicity will still rely on nondescript collective terminology, such as "other race," and instructing respondents to "select all that apply" to assess and subsequently aggregate diverse individuals into a single residual multiracial category. The multiracial population in the US census increased by 276% over the last decade [24]. The National Institutes of Health acknowledges that individuals of two or more races have unique health risks in the United States, yet standard guidance in data reporting is nonexistent beyond aggregation into a finite category [25]. The OMB standards will continue to allow for open-ended written responses to capture additional information; however, standard guidance for handling subsequent information is nonexistent.

XSL•FO
RenderX

Limitations of Current Race and Ethnicity Data Collection and Reporting Procedures

While federally-funded health researchers in the United States are required to collect demographic data from study participants, it appears that most peer-reviewed studies published in major epidemiology and public health journals in recent years have not actually used these variables as they fail to report the racial and ethnic composition of study samples or clarify how these constructs were defined and operationalized in practice [26,27]. In pediatric health research, the racial identities of children are usually reported by parents in adherence with the standard OMB Directive 15 guidelines. Rees et al [28] found that most youth identified by their parents as other race were either excluded entirely from analyses or aggregated inconsistently across pediatric health studies. In the 2020 census, the other race was the fastest-growing racial group in the United States: 49.9 million Americans identified as other race either alone or in combination with another racial category [29,30]. Among those classified as other race, 90.8% of individuals were Hispanic or Latino [30]. Hispanic and Latino individuals experience significant health disparities, including the onset of hypertension, obesity, anxiety, and posttraumatic stress disorders at rates that outpace their non-Hispanic White counterparts. The risk for developing these poor health outcomes is even more pronounced among undocumented populations [31]. Even less is known about the implications of aggregating multiracial individuals into a single combined category because very few previous studies in either the pediatric or general health literature have attempted to systematically examine the heterogeneity hidden within this population. As a result, the interrelationship between survey response formats, nonrandom race and ethnicity data missingness, and health outcomes across and within racialized groups remains poorly understood, hindering the potential for public health systems to equitably protect population health.

Overall, the absence of, opposition to, and significant differences between racial and ethnic data collection systems across the globe leave racialized health inequities unrecognized and overlooked. Greater attention by the public health community to race and ethnicity data standards could contribute toward eliminating preventable racial health inequities by accurately identifying and tracking health outcomes to inform inclusive and equitable programs, policies, and interventions [17,32,32,33]. Therefore, this study developed a standardized and replicable method for improving the use and specificity of child race and ethnicity data collected from parent-reported demographic surveys. We demonstrate the implications of investigating the hidden diversity within the open-ended written response fields, other race group, and aggregate multiracial category on demographic surveys. We hypothesized that these oft-applied survey procedures systematically fail to identify specific marginalized populations that are disproportionately exposed to racism. To test our hypotheses, we retrospectively pooled the demographic data of children we enrolled in health and nutrition studies across the United States and examined whether common survey procedures were associated with the misrepresentation of minoritized racial and ethnic groups.

Methods

Study Design and Data Source

This cross-sectional study used data from 8087 parents or guardians who reported child demographics at or before baseline in 8 pediatric health studies led by members of our study team. The studies used a community-based design in 6 states across the United States (California, Kentucky, Massachusetts, Mississippi, South Carolina, and Rhode Island) from 1999 - 2018. The goal of each study was to improve children's nutritional health outcomes or health behaviors. Additional details on the individual studies in this pooled analysis are in Table 1.

XSL•FO
RenderX

Table . Descriptions of the 8 studies included in the pooled analysis of parent-reported race and ethnicity of child participants (N=8087).

Study[a]	Years	State(s)[b]	Setting	Target population	Consented participants, n (%)	Study design	Study outcome(s)
A	2016 - 2018	MA	Urban and suburban	Children in grades 3 - 4	983 (12.1)	Randomized control trial	Impact of 2 physical activity programs on school time and total daily moderate-to-vigorous physical activity
B	2013	MA	Urban	Children in grades 3 - 6	448 (5.5)	Observational cross-section study	Associations between physical activity and academic achievement
C	2011 - 2012	MA	Urban	Children in grades 4 - 8	693 (8.6)	Randomized control trial	Impact of vitamin D supplementation on serum 25(OH)D and cardiovascular disease risk factors
D	2009 - 2010	MA	Urban	Children in grades 4 - 8	560 (6.9)	Observational cross-section study	Differences in physical fitness levels and health status
E	2008 - 2009	CA, KY, MS, and SC	Rural	Children in grades 1 - 6	1301 (16.1)	Randomized control trial	Community environmental change intervention on undesirable weight gain
F	2007 - 2008	CA, KY, MS, and SC	Rural	Children in grades 1 - 6	959 (11.8)	Prepost program evaluation	Changes in nutrition behaviors following an after-school health program implementation
G	2003 - 2004	MA	Urban and suburban	Children in grades 1 - 3	1726 (21.3)	Community-level intervention with matched control communities	Influence of environmental change intervention on preventing excess weight gain
H	1999 - 2002	MA and RI	Urban and suburban	Children aged 6-9 years	1417 (17.5)	Group randomized trial	Impact of a nutrition and physical activity intervention on bone quality and muscular strength

[a]Study names have been deidentified to protect participant privacy and unpublished data.

[b]MA: Massachusetts; CA: California; KY: Kentucky; MS: Mississippi; SC: South Carolina; RI: Rhode Island.

Ethical Considerations

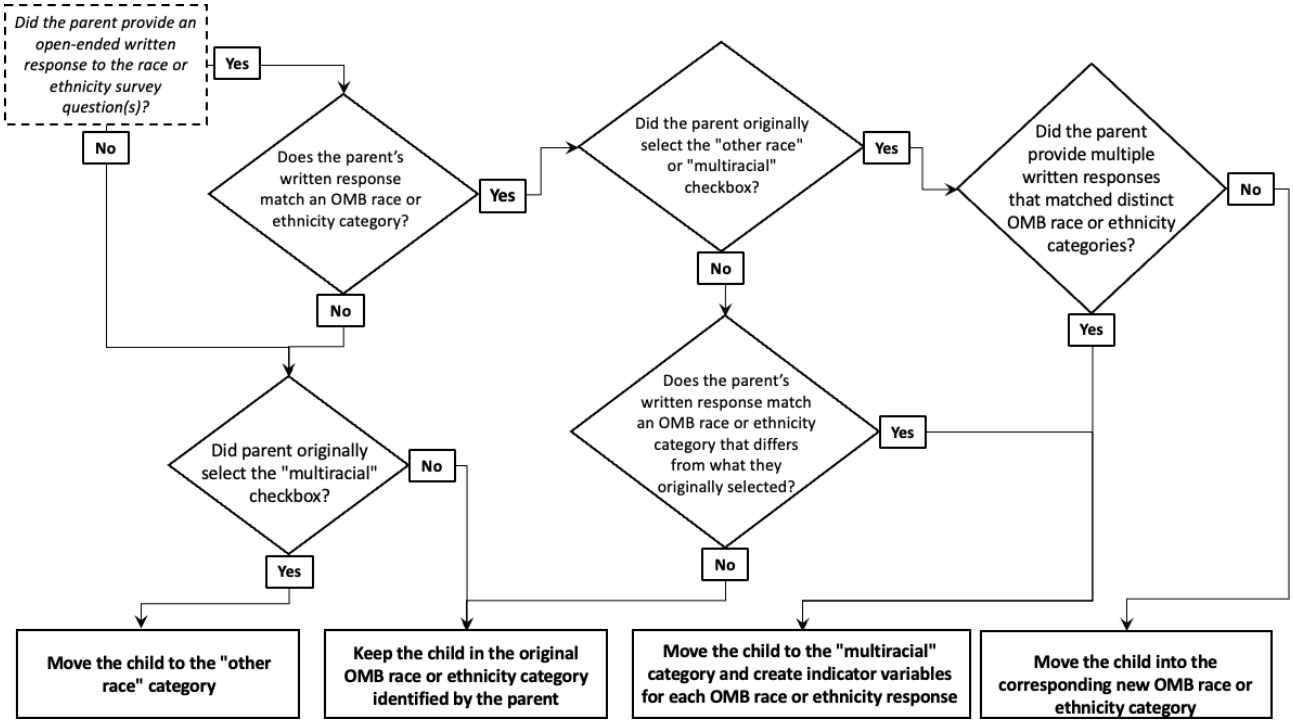
The institutional review board of Tufts University exempted this study from review because it was deemed not human participant research. Study data were deidentified before being shared for pooled analysis, and no direct identifiers (names and contact information) were included. For the 8 pediatric research studies described herein, we obtained informed consent from the parents or caregivers and verbal assent from the youth participants. Across each of the 8 studies, participants received stipends or compensation that varied slightly by protocol (cash, gift cards, or equivalent stipends). All compensation procedures were approved by the Tufts University institutional review board. We followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines (Checklist 1).

XSL•FO
RenderX

Harmonizing Data Across Diverse Race and Ethnicity Survey Response Formats

Inevitably, surveys that ask diverse individuals with unique life experiences to self-report their racial and ethnic identities in limited detail to facilitate group comparisons flatten the complex identities of respondents [10,34,35]. However, our overall objective is to illustrate how the current paradigm of collecting, reporting, and operationalizing the social constructs of race and ethnicity in the United States can bias demographic estimates of populations and may produce inaccurate research findings, which ultimately obscures understanding of the prevalence of racial health disparities. To harmonize children's race and ethnicity data, we pooled and assessed parent-reported survey responses in a double-blinded process (researchers were unaware of the categorical responses when allocating written responses). Figure 1 depicts our replicable decision-making process.

Figure 1. Decision-making process used to allocate all parent-provided written and categorical responses to child race and ethnicity to the Office of Management and Budget (OMB) Directive 15 standards. This flowchart illustrates the standardized procedure we created to ensure consistent categorization of race and ethnicity data across 8 community-based studies with varying race and ethnicity data collection instruments. The decision-making process represents a replicable coding procedure for mapping parent or caregiver-reported child race and ethnicity responses to the federal OMB Directive 15 standards. The process accounts for single- and multiple-race responses, open and close-ended responses, use of the other race and multiracial checkboxes, and written responses that match or diverge from existing OMB categories.



The individual race and ethnicity survey response formats replicated or exceeded the federal minimum reporting categories in OMB Directive 15 [26]. Details on survey response options are in Table 2. As observed in surveys globally [10], all demographic surveys offered a set of closed-ended race and ethnicity categories and the other race checkbox option [36]. The closed-ended race and ethnicity categories included White (White or Caucasian), Black or African American, Asian or Pacific Islander (Asian, Asian American, Asian Indian, or Native Hawaiian or Other Pacific Islander [NHPI]), American Indian or Alaska Native (Native American, American Indian, or Alaska Native), and Hispanic or Latino (Hispanic, Latino, Spanish, or Mexican). All surveys accepted open-ended written responses designated with "please specify." We allocated individuals to the multiracial category if parents or guardians (1) selected the multiracial checkbox (as was the status quo for studies conducted in early years), (2) selected multiple categories, or (3) selected a single category and wrote a distinct race or ethnicity in the "please specify" field (Table 2).

XSL•FO
RenderX

Table . Examples of open-ended responses in "please specify" fields and how they were allocated according to federal US race and ethnicity standards.[a,b]

Example written responses	Official race or ethnicity category definition in the federal register
Any federally recognized tribe (ie, "Cherokee")	American Indian or Alaska Native: A person having origins in any of the original peoples of North and South America (including Central America) and who maintains tribal affiliation or community attachment.
Indian, Japanese, or Vietnamese	Asian: A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent, including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.
Haitian, Malian, or West African	Black or African American: A person having origins in any of the Black racial groups of Africa. Terms such as "Haitian" or "Negro" can be used in addition to "Black or African American."
Central American, El Salvadorean, or Puerto Rican	Hispanic or Latino: A person of Cuban, Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race. The term, ''Spanish origin,'' can be used in addition to ''Hispanic or Latino.''
Polynesian, Hawaiian, or Tahitian	Native Hawaiian or Other Pacific Islander: A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.
Austrian, Dutch, or Lebanese	White: A person having origins in any of the original peoples of Europe, the Middle East, or North Africa.

[a]The table provides example written responses that parents provided regarding their child's race and ethnicity, along with the corresponding federal race or ethnicity category into which the answer was categorized.

[b]Example written responses that do not map to federal race or ethnicity categories: American, Brazilian, Jewish, Muslim, the United States, Massachusetts, and mixed.

In 6 of the 8 studies, demographic surveys assessed race and ethnicity (Hispanic or Latino, not Hispanic or Latino) using a single combined question, reflecting best practices during the respective year(s) of implementation. In studies A and B (1431/7530; 17.68% of the pooled dataset), which were the only ones that assessed race and ethnicity separately, 51.40% (735/1431) of children were identified as Hispanic or Latino. However, racial identity data were missing for 66.8% (491/735) of those identified as Hispanic or Latino in these studies A (the question was left blank or the other race box was selected). To ensure consistency across studies, we grouped ethnicity and race responses such that multiracial includes children identified as two or more federally recognized races, including Hispanic or Latino.

Using Open-Ended Responses to Improve the Accuracy of Race and Ethnicity Data

Next, we individually examined and harmonized written and categorical race or ethnicity responses in a similarly double-blinded fashion to understand the groups misrepresented by response formats. We attempted to allocate the open-ended written responses according to federal race and ethnicity categories, modeled after similar approaches by the US Census Bureau [37-39]. The research team discussed discrepancies and resolved them through consensus, guided by the OMB Directive 15 standards [40]. Our use of the term "[re]allocation" in the context of this study is not meant to discount the parent or guardian's original response. Rather, we recognize that the open-ended responses from respondents reflect essential aspects of children's identity that warrant consideration by researchers.

Disaggregating Individual Race or Ethnicity Responses Within the Aggregate Multiracial Group

To demonstrate the breadth and depth of information lost by grouping diverse individuals in a single multiracial category, we disaggregated and plotted the individual responses within the multiracial group using an UpSet plot. This approach visually highlights both the diversity and the structure of the multiracial category.

Statistical Analysis

Individuals were the primary unit of analysis. We examined changes in racial and ethnic group frequencies (as defined by the OMB 15 Standards) before and after applying our standardized allocation process described in Figure 1. Chi-square tests determined whether there was a significant association between the original parent-reported race or ethnicity category and the reallocated category. Analyses were conducted from June 1, 2022, to July 26, 2023, using Stata/SE software (version 17; StataCorp LLC).

Results

Analytic Sample Characteristics

In total, 93.11% (7530/8087) of parents provided data on the child's race or ethnicity, and 6.88% (557/8087) did not (fields were left blank or demographic surveys were not returned). Altogether, the open-ended written responses to race or ethnicity spanned 101 distinct groups, including responses such as "Haitian," "Mexican," and others (Table 2). Racial and ethnic

XSL•FO
RenderX

characteristics of the sample before and after the harmonized reallocation process are in Table 3. Implementing this process (Figure 1) revealed that 95.97% (7277/7530) of children remained in their original parent-reported race or ethnicity category; the chi-square test indicated a strong and significant association between the original and reallocated racial categories (χ^2_{36}=37,000; P<.001). Figure 2 shows that of the 4.02% (303/7530) of participants who were reallocated, 89.4% (271/303) were originally identified as multiracial or other race by their parents. Figure 2 also illustrates the movement across racial groups among the subsample of participants who were reallocated out of the original parent-reported category.

Table . Cross-tabulation of original parent-reported child race or ethnicity with child race or ethnicity after the replicable reallocation process[a].

	Reallocated racial or ethnic group, n							
	White	Black	Hispanic	Asian or NH-PI[b]	AIAN[c]	Multiracial	Other race	Total
Original parent-provided categorical response for the race or ethnicity of the child, n								
White	2791	0	0	0	0	1	0	2792
Black	0	1477	0	0	0	5	0	1482
Hispanic	0	0	1871	0	0	23	0	1894
Asian or NHPI	0	0	0	314	0	0	0	314
AIAN	0	0	0	0	35	0	0	35
Multiracial	9	29	15	1	2	621	55	732
Other race	33	88	18	9	0	15	118	281
Total	2833	1594	1904	324	37	665	173	7530

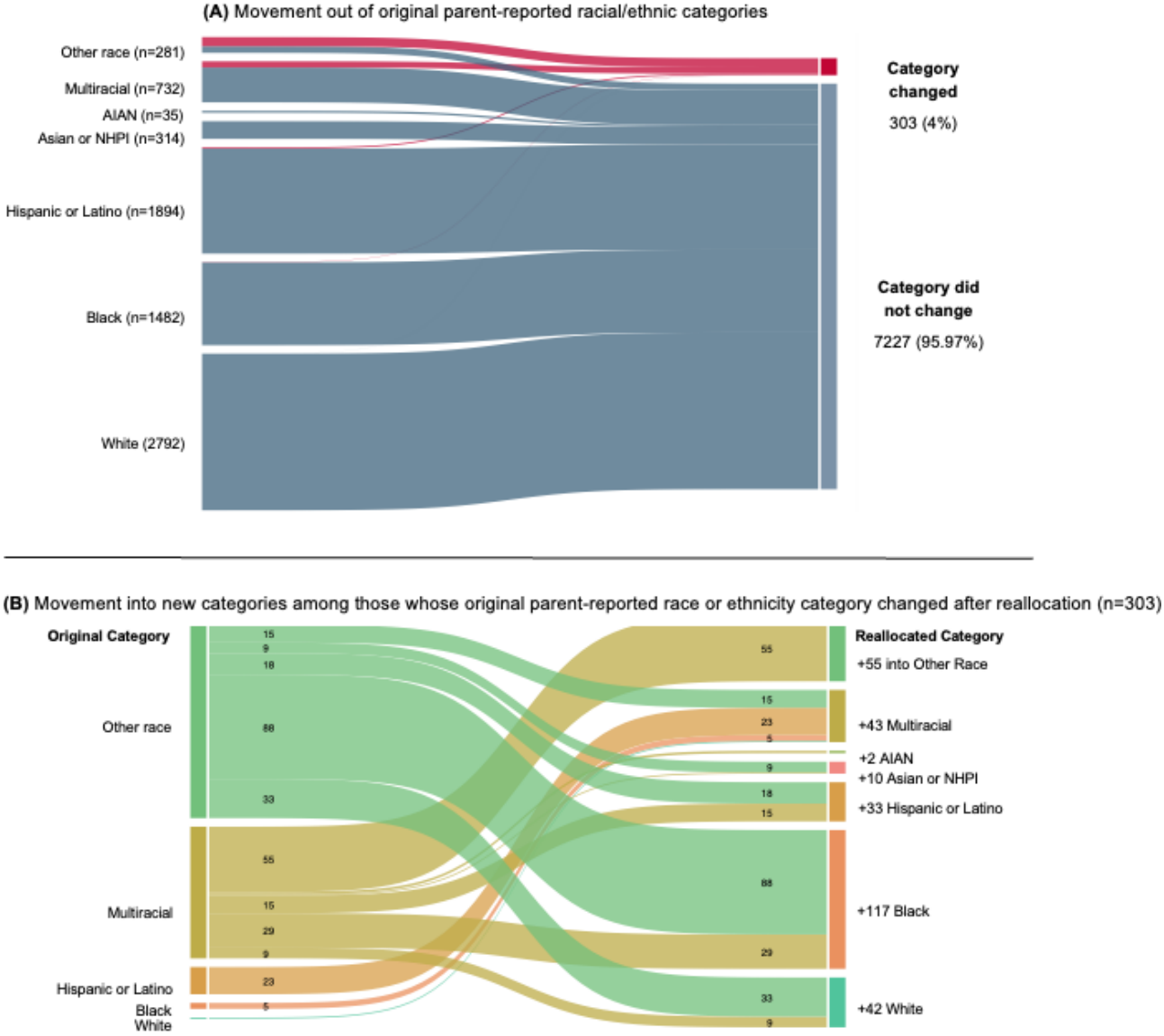
[a]χ^2_{36}=37,000; P<.001.

[b]NHPI: Native Hawaiian or Other Pacific Islander.

[c]AIAN: American Indian or Alaska Native.

XSL•FO
RenderX

Figure 2. Results of reallocation process on frequency of and movement between racial and ethnic groups. (A) Movement out of original parent-reported racial or ethnic categories. (B) Movement into new categories among those whose original parent-reported race or ethnicity category changed after reallocation (n=303). AIAN: American Indian or Alaska Native; NHPI: Native Hawaiian or Other Pacific Islander.



Reallocation Results From the Harmonized Reallocation Process

Among youth originally categorized as other race, only 41.9% (118/281) remained in this category after harmonization, while 31.3% (88/281) were moved into Black, 11.7% (33/281) into White, 6.4% (18/281) into Hispanic or Latino, 5.3% (15/281) into multiracial, and 3.2% (9/281) into Asian or NHPI. Moreover, 84.8% (621/732) of children who were originally identified as multiracial remained in the same group after the harmonization process, while 3.9% (29/732) were reallocated to Black, 2% (15/732) to Hispanic or Latino, and 7.5% (55/732) to other race.

Following the reallocation process, the Black category experienced the largest increase, with 117 additional children reclassified into this group, increasing from 1482 to 1594 (a +7.42% change). This change reflected the movement of 88 children originally categorized as other race and 29 as multiracial into the Black category, while 5 children originally identified as Black were reallocated to the multiracial group. Alternatively, the other race category experienced the largest decrease, going from 281 to 173 (a –38.43% change). This change reflected that 55 children originally identified as multiracial were moved into the other race category because their parents selected the multiracial checkbox without providing additional written details to create indicator variables. However, 163 children were moved out of the other race group into a different OMB category (15 into multiracial, 9 into Asian or NHPI, 18 into Hispanic or Latino, 88 into Black, and 33 into White). Moreover, the multiracial category declined from 732 to 665 children (a –9.5% change) following the reallocation process for several reasons. As previously described, 109 children left the category (9 moved into White, 29 into Black, 15 into Hispanic or Latino, 1 into Asian or NHPI, and 55 into other race). In addition, 15 children moved into the multiracial category from other race because the parent checked the other race box but provided two or more written responses that were mapped to distinct OMB racial groups, such as French (White) and Japanese (Asian). Another 23 children were reallocated into

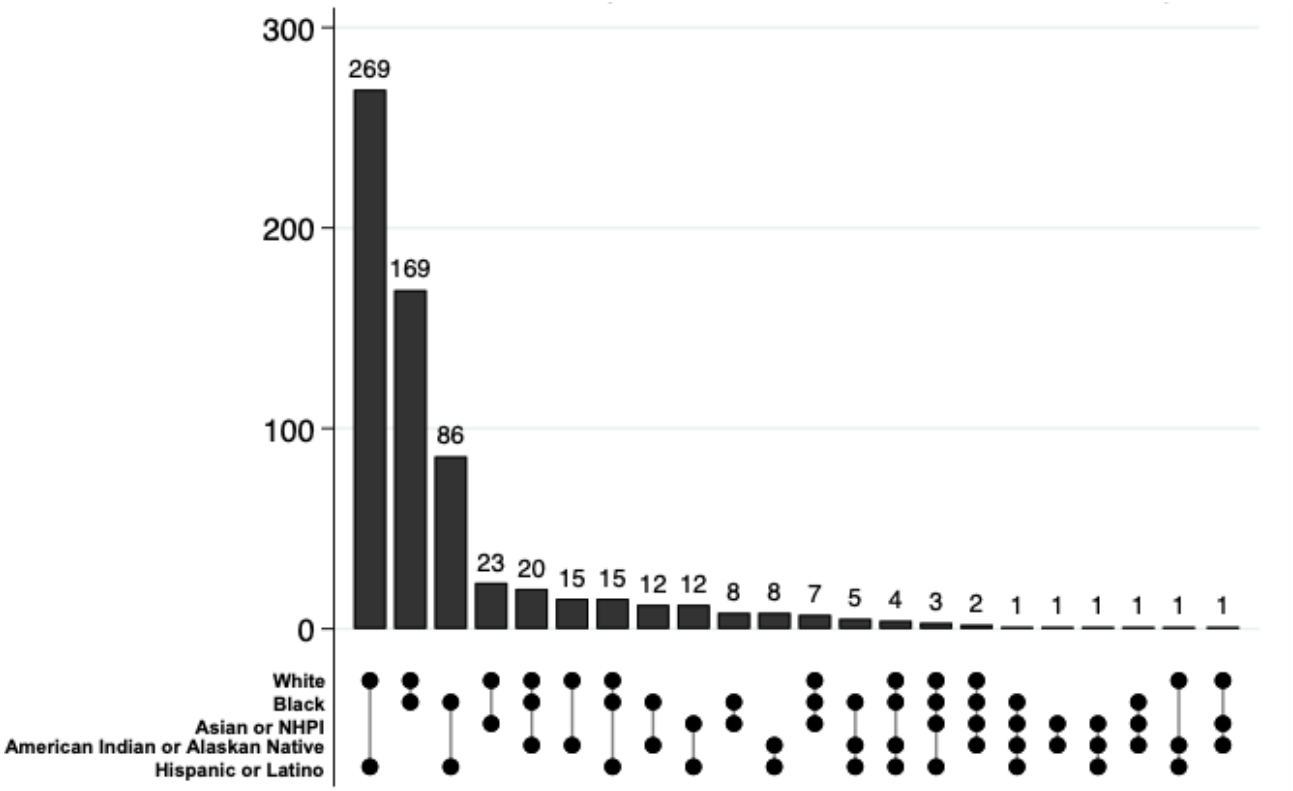
XSL•FO
RenderX

the multiracial category from Hispanic or Latino because the parent provided an open-ended written response that mapped to a distinct OMB category and either checked the Hispanic box or answered affirmatively to the question of whether the child was Hispanic or Latino (a child with a check in the Hispanic box with "White and Black" written would be allocated to multiracial with indicators for Hispanic or Latino, White, and Black). In addition, 5 children moved from Black to multiracial because the parent checked Black and provided a written response that would be defined as a distinct category per the OMB standards. The White category increased by 41 children, from 2792 to 2833 (a +1.47% change), due to 33 children joining this category who were originally identified as other race (parent could have checked the other race box and wrote "German") and 9 as multiracial (in studies prior to 2000, the parent checked multiracial and could have written "German"), while 1 was moved from White into multiracial. The Asian or NHPI category increased by 10 children, rising from 314 to 324 (a 3.2% increase), largely due to the 9 children who were reallocated to this group from other race.

Disaggregating the Racial or Ethnic Combinations Within the Multiracial Category

Figure 3 presents an UpSet plot illustrating the overlap among the 22 distinct racial and ethnic identity combinations we uncovered within the 8.83% of participants (665/7530) categorized as multiracial in the reallocated dataset (N=8087). For example, the bar with markers for White, Black, and Hispanic or Latino represents multiracial participants who were reported with all 3 identities. Among the multiracial participants, most were White and Hispanic or Latino (269/665, 40.5%). The next most common combination was White and Black (169/665, 25.4%) individuals, followed by Black and Hispanic or Latino (86/665, 12.9%). The American Indian or Alaska Native identity contributed to 13 of the 22 combination categories, but only 12.2% (81/665) of multiracial children held this identity.

Figure 3. Prevalence and overlap of racial or ethnic identity combinations within the multiracial group (n=665). The vertical bars represent the number of participants identified with each specific combination of racial and ethnic identities, with bar height indicating the number of participants within that combination. The horizontal matrix below the bars shows which identities are included in each of the 22 distinct combinations. NHPI: Native Hawaiian or Other Pacific Islander.



Discussion

Principal Findings

Retrospectively harmonizing survey data across 8 community-based child health studies revealed new insights into how past, current, and new federal data standards can influence the representation of marginalized racial and ethnic populations in public health research, policy, and practice. We observed significant changes in the racial and ethnic composition of our sample after reallocating parent-provided written and categorical responses on child race or ethnicity surveys into federally defined categories. Our findings add to the existing literature underscoring how specific components of survey research design disproportionately threaten the quality, completeness, and internal validity of social determinants of health data for minoritized racial populations [41]. The inconsistencies in survey design and methodology we identified may reinforce racial health disparities at the population level through racial group misclassification, nonresponse bias, and the obscuring of diversity through the aggregate multiracial

XSL•FO
RenderX

category, thereby warranting further hypothesis testing [42]. Continuing to overlook these survey research elements—open-ended responses, the other race category, and multiracial populations—may hinder racial equity by contributing to an evidence base that lacks validity and reliability for racial and ethnic groups that are exposed to multiple forms of discrimination. The trends we observed between written and categorical responses underscore that race and ethnicity are context-specific, fluid, and dynamic social constructs, which researchers and policymakers should carefully consider during data collection, analysis, and dissemination [5].

Currently, public health researchers and surveillance methods may not adequately account for the confusion that can result from inconsistent or conflated terminology, government standards, and societal norms around race and ethnicity. In addition, internalized racism may lead many immigrants and members of marginalized racial groups to reject or not identify with federal racial or ethnic categories. For example, all surveys in our study enumerated the single combined Black or African American categorical response option; however, many children were reallocated into this category because their parents checked other race and provided written data, such as "Haitian," "Cape Verdean," or "Black." Documented differences in group consciousness between foreign and US-born Black individuals could partially explain this discrepancy [43]. Intraracial relations also play a prominent role in how Caribbean immigrants, African immigrants, and US-born Black people view themselves and each other, underscoring that heterogeneity within racial, ethnic, and nativity groups shapes individual identity formation [44-47]. Cultural demands and expectations profoundly affect identity formation, particularly in how an individual chooses to identify with a larger demographic group [46,48-50]. The findings of this study may also inform broader efforts to assess data related to descendants of formerly enslaved individuals in the United States, which was a major point of contention during the OMB's process of updating the race and ethnicity data standards announced in 2024. While this was identified as a priority area for further research, the final decision was not to require the collection of such data given concerns that doing so could make accurately counting the Black community, and particularly African immigrants, even more difficult [23]. Overall, our results raise important directions for new research examining how racism, and specifically anti-Blackness and settler colonialism, impact demographic surveillance systems and racial health research in the United States and elsewhere, across generations and age groups [43].

Our findings also underscore the need to improve the measurement and reporting of child race and ethnicity. Following the current best practice in pediatric health research, all surveys asked parents or caregivers to report the racial or ethnic identity of their child. However, racial identity formation is influenced by societal norms and intrafamilial characteristics, such as maternal race and acculturation levels. Therefore, many children, especially those from immigrant, Black, and Hispanic families, might have self-identified their race differently than did their parents if they were administered similar demographic surveys. There is a major gap in knowledge regarding the prevalence and implications of discrepancies between how parents and children identify themselves and each other. Therefore, the understanding of demographic trends among youth and the ability to address racial health disparities could be improved by collecting the self-identified and parent-identified race and ethnicity of children, and recollecting this data over time to account for changes during critical periods of social identity development [51-53].

Our results also reinforce the broader movement by MENA individuals in the United States for official federal recognition with their own ethnicity category [54]. We allocated 33 participants from other race to White due to written responses indicating MENA identity (eg, "Lebanese"). People of MENA descent have diverse backgrounds, cultures, and experiences [54]. MENA individuals in the United States and other Western countries do not identify as White, are often not socially perceived as White, and endure pervasive yet specific forms of discrimination and xenophobia, and the federal government's categorization of them as "White" is inappropriate [54]. Aggregation of heterogeneous individuals into broad racial or ethnic groups mutes the diversity of populations. Some states have enacted laws to prevent the harms of aggregation for some but not all minoritized racial groups. For example, the *California Government Code* Section 8310.5 requires all state agencies in California to collect data for each major Asian and Pacific Islander group (ie, including, but not limited to, Asian Indian, Cambodian, Chinese, Filipino, Guamanian, Hawaiian, Hmong, Japanese, Korean, Laotian, Samoan, Tahitian, and Vietnamese) [55]. Detailed subgroup reporting should be mandated nationally, extending beyond Asian subpopulations to include MENA, people racialized as Black, diverse individuals of Hispanic and Latino heritage, and others. By examining written responses, we substantiated the high probability of widespread nonrandom missingness of race or ethnicity data associated with close-ended survey response options. Essentially, these survey techniques erase the existence and experiences of members of minoritized racial groups, representing a ubiquitous form of racialized social hierarchies.

The geographic areas within which our studies worked included large populations of Brazilians. Many individuals checked other race and wrote in "Brazilian" but could not be allocated to a specific racial or ethnic group because the current OMB race and ethnicity standards do not recognize Brazilians. Some Brazilians may self-identify as Latino due to Brazil's geographic location in South America; however, the OMB definition of Hispanic or Latino, which "refers to a person of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race," specifically excludes this group [56]. Brazil was colonized by the Portuguese, who the OMB currently defines as White, or "a person having origins in any of the original peoples of Europe, the Middle East, or North Africa," which overlooks and fails to acknowledge the effects of colonization over history. The complexities of racialization among Brazilians in the United States are far-reaching. For instance, the US Census Bureau explicitly excludes Brazilians from its definition of "Hispanic or Latino" and even back codes their data (ie, recategorizes Brazilians as non-Hispanic or Latino even if they self-identified as Hispanic or Latino) in federal reports [57]. However, the

XSL•FO
RenderX

National Institutes of Health style guide explicitly mentions Brazilians when describing the Latino category [58]. Nonetheless, federally recognized racial and ethnic groups are not inclusive toward Brazilians, who are recategorized as other race, non-Hispanic during US Census Bureau data cleaning and processing, effectively rendering them invisible. Further research with Brazilian communities is needed to guide the creation of more inclusive race and ethnicity data standards and to avoid additional obscurement of the unique health risks and protective factors within this population, which rose by nearly 50% in the United States between 2010 and 2019 and is projected to continue growing [59].

Marginalized populations deeply affected by racism are obscured by the aggregate multiracial category. Aggregation involves the loss of information and decreased ability of statistical and epidemiological tools to analyze and identify patterns in the outcome of interest [32]. Frequently, a simplistic aggregation approach of multiracial individuals is used to simplify data collection or, presumably, to protect individual privacy. However, these aggregation decisions accomplish neither in practice nor in theory [60]. The disaggregation approach reported in our study could be replicated at a national scale using existing cohorts to improve our understanding of the experiences of populations historically excluded from health research. Because very little attention has been paid to the heterogeneity within biracial and multiracial populations, the scope and specific implications of the aggregate category's interference with representation and demographic surveillance remain unknown. Critically, more research is needed to develop clear analytical guidelines to improve the representation of diverse individuals.

Our findings have direct application to the procedural transitions that institutions will be faced with implementing as the OMB's 2024 revisions to race and ethnicity data standards take effect in 2030. Data crosswalks, such as the translational harmonization process we developed, are essential for accurately comparing temporal changes in (1) population prevalence and demographic shifts and (2) health risks, behaviors, and outcomes across racial or ethnic groups. In our pooled analysis, child race and ethnicity data were measured using methods consistent with how child health and demographic estimates have been reported in the United States for more than 3 decades. The approach presented in our study provides a feasible framework to support large-scale harmonization of child race and ethnicity data across time, populations, and datasets. Adopting this process could help mitigate systematic racial bias in federal data standards and safeguard the validity of child health data across past and future generations. Overall, our study highlights and offers a solution for persistent gaps in race and ethnicity data practices that remain despite forthcoming OMB revisions.

Generating Future Research Directions

This study pooled and harmonized parent-reported data on children's racial and ethnic identity. Additional research is needed to determine whether distinct patterns emerge when this process is applied to adult populations or to self-reported data from children. Nonetheless, OMB race and ethnicity standards do not differentiate racial or ethnic group definitions or data handling techniques (collection, cleaning, or reporting) based on the age of the population. Therefore, we posit that standard survey formats—a set of close-ended race or ethnicity categories with other race, a "select all that apply" instruction, and an open-ended "please specify" field—incur significant limitations for accurately enumerating race and ethnicity in public health research and surveillance. As a result of identifying specific mechanisms in data collection and reporting, we present 3 hypotheses for future investigation into the overlooked drivers of structural racism in health research.

First, we hypothesize that when survey questions and response formats change to reduce the stigmatizing, incomplete, inconsistent, or unclear language (including language barriers), respondents are less likely to withhold information or identify as a collective category such as other race. [61] Thus, researchers can reduce bias in the data collection stage. With targeted investments by funders, researchers could improve the inclusion of underrepresented groups at the design and input stages. Toward this hypothesis, investigators could systematically test for interlinkages between close- and open-ended responses using our decision-making process (Figure 1) to uncover how aggregate and constrained race and ethnicity categories hide or exclude specific groups. Input from diverse community members must be prioritized during this investigation and when designing race and ethnicity survey questions to prevent additional stigma and bias and build trust with communities [18,62,63].

Second, we hypothesize that when crude aggregation of individuals of multiple races or ethnicities into residual categories is avoided, the bias associated with aggregation at the data analysis and presentation stages could be substantially reduced. Researchers could substantially reduce race-related bias by making relevant decisions that better capture the diversity of multiracial participants and examine the specificity of treatment effect differences. Reliance on aggregation as the practical method of planning and implementing public health research ignores minoritized groups and blocks paths for further investigation. Testing these hypotheses by the global research community may uncover local and universal mechanisms and pathways of how identity obfuscation and invalid assumptions contribute to structural racism and decrease equitable representation in health surveillance, research, policies, and programs.

Third, we hypothesize that the significant health differences that exist across racial and ethnic groups could be more accurately understood and addressed through the large-scale application of this data harmonization and disaggregation method. This approach would allow researchers and health agencies to reanalyze population cohorts, providing deeper insight into how our understanding of health status may be shaped or obscured by racial aggregation, particularly for individuals identified as other race and multiracial.

Limitations

Our exploratory process aimed to provide a method to inform improved race and ethnicity data systems to reflect more accurately the identities and backgrounds of participants, but this work presents some limitations. First, we used a convenience sample of data by pooling children's demographic

XSL•FO
RenderX

surveys from studies conducted by investigators at Tufts University in Boston, Massachusetts. Thus, while these studies did recruit children living in 6 US states, our results were not meant to be generalizable to the broader pediatric population in the United States or elsewhere. In addition, our harmonized replication process required changing some parents' initial categorical race and ethnicity responses to different categories. However, the overall implications of this work could allow researchers and policy makers to address health inequities with a consistent and inclusive approach to data collection and reporting. Specifically, retrospectively applying this method to population datasets will reveal groups that require new unique categories instead of being forced to omit data or identify with the long-standing yet ambiguous other race option. Moreover, children whose parents checked the multiracial checkbox but provided a single written response were reallocated into the corresponding single-race category. A drawback of this decision was not preserving these children's multiracial identity. We acknowledge that grouping multiracial individuals with their monoracial counterparts is not the appropriate solution since histories, experiences, and self-identities within groups differ [34,64]. Going forward, researchers should avoid using a single checkbox to assess the identities of multiracial individuals and work to develop equitable and reflective measurement and analysis practices. Our goal was to contribute to local, national, and global conversations that address the stigmatizing and exclusionary nature of current race and ethnicity data standards while providing a framework to increase data visibility for investigators and governments who may implement new and existing practices.

Conclusions

Rectifying the past and present harms of racism begins with ensuring the race and ethnicity data that inform research and policy are reliable, replicable, valid, and inclusive. The current paradigm of assessing these social constructs on surveys in the United States significantly impedes the ability to identify, track, and improve health outcomes across and within diverse racial and ethnic groups. The results of our study could inform future changes to the OMB's formal process for collecting and reporting race or ethnicity in the United States. Our study underscores that Black subpopulations in the United States are also at a disproportionately high risk of misclassification. The implications of federal racial data standards for people racialized as non-Hispanic Black have been largely excluded from policy discussions and recent OMB data standard revisions. Creating new categories that reflect the overwhelming heterogeneity within the Black population (eg, Afro-Caribbeans) must be central within any forthcoming changes to federal race and ethnicity data standards. Simultaneously, researchers across disciplines must develop new methods to account for growing multiracial populations worldwide.

To address growing racial health inequities, we advocate for the large-scale application of our standardized method to improve understandings of disease disparities and inform more inclusive public health research and policy. Such methodological decisions are essential for understanding the causes and consequences of racism and forming evidence-based guidelines to address inequities across the globe. Minoritized racial groups have long endured mistreatment in science and society, hence the ongoing global reckoning with racism [3,63,65,66]. Targeted investments and collective efforts by entire research teams, funders, journals, institutions, and governments are necessary to find, test, and implement solutions that uncover the unique contributions of multiple forms of racism to health outcomes, which are currently unknown due to nonrandom structural missingness of racial identity data. Race, ethnicity, and other forms of identity data inform health policy, research, and programs, and unless corrected, the problems we have identified will exacerbate long-standing health disparities and leave experiences of racial discrimination undocumented and overlooked.

Data Availability

The datasets analyzed for this study are not publicly available due to restrictions in the original study protocols, which did not include consent for data sharing.

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[DOCX File, 34 KB - publichealth_v11i1e65660_app1.docx]

References

1. Williams DR, Mohammed SA. Racism and health I: Pathways and scientific evidence. Am Behav Sci 2013 Aug 1;57(8):1152-1173. [doi: 10.1177/0002764213487340] [Medline: 24347666]
2. Hardeman RR, Homan PA, Chantarat T, Davis BA, Brown TH. Improving the measurement of structural racism to achieve antiracist health policy. Health Aff (Millwood) 2022 Feb;41(2):179-186. [doi: 10.1377/hlthaff.2021.01489] [Medline: 35130062]
3. Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. The Lancet 2017 Apr;389(10077):1453-1463. [doi: 10.1016/S0140-6736(17)30569-X]

XSL•FO
RenderX

4. Flanagin A, Frey T, Christiansen SL, AMA Manual of Style Committee. Updated guidance on the reporting of race and ethnicity in medical and science journals. JAMA 2021 Aug 17;326(7):621-627. [doi: 10.1001/jama.2021.13304] [Medline: 34402850]
5. Cuevas AG, Dawson BA, Williams DR. Race and skin color in Latino health: an analytic review. Am J Public Health 2016 Dec;106(12):2131-2136. [doi: 10.2105/AJPH.2016.303452] [Medline: 27736206]
6. Jones CP, Truman BI, Elam-Evans LD, et al. Using "socially assigned race" to probe White advantages in health status. Ethn Dis 2008;18(4):496-504. [doi: 10.13016/36IU-SC8C] [Medline: 19157256]
7. Krieger N. ENOUGH: COVID-19, structural racism, police brutality, plutocracy, climate change—and time for health justice, democratic governance, and an equitable, sustainable future. Am J Public Health 2020 Nov;110(11):1620-1623. [doi: 10.2105/AJPH.2020.305886] [Medline: 32816556]
8. McAvey K. Investing in the data systems we need to create the health system we deserve. Health Aff Forefront 2021. [doi: 10.1377/forefront.20210528.555489]
9. United Nations General Assembly. A global call for concrete action for the elimination of racism, racial discrimination, xenophobia and related intolerance and the comprehensive implementation of and follow-up to the Durban Declaration and Programme of Action: Resolution A/RES/75/237. United Nations Official Document System. 2020 Dec 31. URL: https://undocs.org/pdf?symbol=en/A/HRC/47/53 [accessed 2022-01-04]
10. Morning A. Ethnic classification in global perspective: a cross-national survey of the 2000 Census Round. Popul Res Policy Rev 2008 Apr;27(2):239-272. [doi: 10.1007/s11113-007-9062-5]
11. Yi SS, Kwon SC, Suss R, et al. The mutually reinforcing cycle of poor data quality and racialized stereotypes that shapes Asian American health. Health Aff (Millwood) 2022 Feb;41(2):296-303. [doi: 10.1377/hlthaff.2021.01417] [Medline: 35130076]
12. Geneviève LD, Martani A, Wangmo T, Elger BS. Precision public health and structural racism in the United States: promoting health equity in the COVID-19 pandemic response. JMIR Public Health Surveill 2022 Mar 4;8(3):e33277. [doi: 10.2196/33277] [Medline: 35089868]
13. O'Connell HA, Bratter JL, Casarez RS. One drop on the move: historical legal context, racial classification, and migration. Ethn Racial Stud 2022 Apr 4;45(5):809-828. [doi: 10.1080/01419870.2020.1761554]
14. Simon P, Piché V, Gagnon AA. Social Statistics and Ethnic Diversity: Cross-National Perspectives in Classifications and Identity Politics: Springer International Publishing; 2015. [doi: 10.1007/978-3-319-20095-8]
15. Greene J, Blustein J, Weitzman BC. Race, segregation, and physicians' participation in Medicaid. Milbank Q 2006;84(2):239-272. [doi: 10.1111/j.1468-0009.2006.00447.x] [Medline: 16771818]
16. Cobb RJ, Thomas CS, Laster Pirtle WN, Darity WA. Self-identified race, socially assigned skin tone, and adult physiological dysregulation: assessing multiple dimensions of "race" in health disparities research. SSM Popul Health 2016 Dec;2:595-602. [doi: 10.1016/j.ssmph.2016.06.007] [Medline: 29349174]
17. Neighbors HW, Mattingly DT, Johnson J, Morse K. The contribution of research to racial health equity? Blame and responsibility in navigating the status quo of anti-black systemic racism. Soc Sci Med 2023 Jan;316:115209. [doi: 10.1016/j.socscimed.2022.115209]
18. Kader F, Đoàn LN, Lee M, Chin MK, Kwon SC, Yi SS. Disaggregating race/ethnicity data categories: criticisms, dangers, and opposing viewpoints. Health Aff Forefront 2022. [doi: 10.1377/forefront.20220323.555023]
19. Krobath DM, Naumova EN, Cuevas AG, Sacheck JM, Wilson NLW, Economos CD. Use of Bland-Altman analysis to examine the racial and ethnic representativeness of study populations in community-based pediatric health research. JAMA Netw Open 2023 May 1;6(5):e2312920. [doi: 10.1001/jamanetworkopen.2023.12920] [Medline: 37166796]
20. Williams DR, Jackson JS. Race/ethnicity and the 2000 Census: Recommendations for African American and other Black populations in the United States. Am J Public Health 2000 Nov;90(11):1728-1730. [doi: 10.2105/ajph.90.11.1728] [Medline: 11076240]
21. Powell RA, Njoku C, Elangovan R, et al. Tackling racism in UK health research. BMJ 2022 Jan 18;376:e065574. [doi: 10.1136/bmj-2021-065574] [Medline: 35042720]
22. Proposed standards for race-based and indigenous identity data collection and health reporting in Canada. : Canadian Institute for Health Information; 2020 URL: https://tinyurl.com/bddp2wap [accessed 2022-01-31]
23. Revisions to OMB's Statistical Policy Directive No.15: standards for maintaining, collecting, and presenting federal data on race and ethnicity. Federal Register. 2024. URL: https://tinyurl.com/5bu347dh [accessed 2025-05-01]
24. Thompson JH. Preparing for the 2020 census: measuring race and ethnicity in America. United States Census Bureau. URL: https://www.census.gov/newsroom/blogs/director/2014/10/preparing-for-the-2020-census-measuring-race-and-ethnicity-in-america.html [accessed 2021-04-27]
25. Glossary of NIH terms. National Institutes of Health. URL: https://grants.nih.gov/grants/glossary.htm#E [accessed 2022-01-30]
26. Martinez RAM, Andrabi N, Goodwin AN, Wilbur RE, Smith NR, Zivich PN. Conceptualization, operationalization, and utilization of race and ethnicity in major epidemiology journals, 1995-2018: a systematic review. Am J Epidemiol 2023 Feb 24;192(3):483-496. [doi: 10.1093/aje/kwac146] [Medline: 35938872]

XSL•FO
RenderX

27. Swilley-Martinez ME, Coles SA, Miller VE, et al. “We adjusted for race”: Now what? A systematic review of utilization and reporting of race in American Journal of Epidemiology and Epidemiology, 2020-2021. Epidemiol Rev 2023 Dec 20;45(1):15-31. [doi: 10.1093/epirev/mxad010] [Medline: 37789703]
28. Rees CA, Stewart AM, Mehta S, et al. Reporting of participant race and ethnicity in published US pediatric clinical trials from 2011 to 2020. JAMA Pediatr 2022 May 1;176(5):e220142. [doi: 10.1001/jamapediatrics.2022.0142] [Medline: 35311946]
29. Jones N, Marks R, Ramirez R, Rios-Vargas M. 2020 census illuminates racial and ethnic composition of the country. United States Census Bureau. 2021 Aug 12. URL: https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html [accessed 2022-10-03]
30. Peña JE, Lowe RHJ, Sánchez-Rivera AI. Over half a million people self-identified as Brazilian in 2020 census. United States Census Bureau. URL: https://www.census.gov/library/stories/2023/10/2020-census-dhc-a-some-other-race-population.html [accessed 2025-05-01]
31. Cabral J, Cuevas AG. Health inequities among Latinos/Hispanics: documentation status as a determinant of health. J Racial Ethn Health Disparities 2020 Oct;7(5):874-879. [doi: 10.1007/s40615-020-00710-0] [Medline: 32026286]
32. Wang K, Grossetta Nardini H, Post L, Edwards T, Nunez-Smith M, Brandt C. Information loss in harmonizing granular race and ethnicity data: descriptive study of standards. J Med Internet Res 2020 Jul 20;22(7):e14591. [doi: 10.2196/14591] [Medline: 32706693]
33. Cook L, Espinoza J, Weiskopf NG, et al. Issues with variability in electronic health record data about race and ethnicity: descriptive analysis of the National COVID Cohort Collaborative Data Enclave. JMIR Med Inform 2022 Sep 6;10(9):e39235. [doi: 10.2196/39235] [Medline: 35917481]
34. Charmaraman L, Woo M, Quach A, Erkut S. How have researchers studied multiracial populations? A content and methodological review of 20 years of research. Cultur Divers Ethnic Minor Psychol 2014 Jul;20(3):336-352. [doi: 10.1037/a0035437] [Medline: 25045946]
35. Sanchez GE. The politics of Cape Verdean American identity. Transform Anthropol 1997 Jan 1;6(1-2):54-71. [doi: 10.1525/tran.1997.6.1-2.54]
36. Revisions to the standards for the classification of federal data on race and ethnicity. Office of Management and Budget. URL: https://obamawhitehouse.archives.gov/node/15626 [accessed 2021-08-19]
37. Buescher P, Gizlice Z, Self-Reported JVK. Self-Reported Versus Published Data on Racial Classification in North Carolina Birth Records: The State Center for Health Statistics; 2002, Vol. 139.
38. Rodriguez-Lainz A, McDonald M, Fonseca-Ford M, et al. Collection of data on race, ethnicity, language, and nativity by US public health surveillance and monitoring systems: gaps and opportunities. Public Health Rep 2018;133(1):45-54. [doi: 10.1177/0033354917745503] [Medline: 29262290]
39. Marrow H. To be or not to be (Hispanic or Latino): Brazilian racial and ethnic identity in the United States. Ethnicities 2003;3(4):427-464. [doi: 10.1177/1468796803003004001]
40. Wallman KK, Evinger S, Schechter S. Measuring our nation’s diversity: developing a common language for data on race/ethnicity. Am J Public Health 2000 Nov;90(11):1704-1708. [doi: 10.2105/ajph.90.11.1704] [Medline: 11076235]
41. Cook LA, Sachs J, Weiskopf NG. The quality of social determinants data in the electronic health record: a systematic review. J Am Med Inform Assoc 2021 Dec 28;29(1):187-196. [doi: 10.1093/jamia/ocab199]
42. Naumova EN. Public health inequalities, structural missingness, and digital revolution: time to question assumptions. J Public Health Pol 2021 Dec;42(4):531-535. [doi: 10.1057/s41271-021-00312-y]
43. Smith CW. Ethnicity and the role of group consciousness: a comparison between African Americans and Black immigrants. Polit Groups Identities 2013 Jun;1(2):199-220. [doi: 10.1080/21565503.2013.786650]
44. Asante G, Sekimoto S, Brown C. Becoming “Black”: Exploring the racialized experiences of African immigrants in the United States. Howard J Commun 2016 Oct;27(4):367-384. [doi: 10.1080/10646175.2016.1206047]
45. Okonofua BA. I Am Blacker Than You”: Theorizing conflict between African immigrants and African Americans in the United States. SAGE Open 2013;3(3):2158244013499162. [doi: 10.1177/2158244013499162]
46. Clark MK. Identity among first and second generation African immigrants in the United States. African Identities 2008 May;6(2):169-181. [doi: 10.1080/14725840801933999]
47. Nunnally SC. Linking blackness or ethnic othering?: African Americans’ diasporic linked fate with West Indian and African peoples in the United States. Du Bois Rev 2010;7(2):335-355. [doi: 10.1017/S1742058X10000305]
48. Rong XL, Brown F. Socialization, culture, and identities of Black immigrant children: What educators need to know and do. Educ Urban Soc 2002;34(2):247-273. [doi: 10.1177/0013124502342008]
49. Habecker S. Becoming African Americans: African immigrant youth in the United States and hybrid assimilation. J Pan Afr Stud 2017;10(1) [FREE Full text]
50. Borrell LN. Racial identity among Hispanics: implications for health and well-being. Am J Public Health 2005 Mar;95(3):379-381. [doi: 10.2105/AJPH.2004.058172] [Medline: 15727961]
51. Sanchez DT, Gaither SE, Albuja AF, Eddy Z. How policies can address multiracial stigma. Policy Insights Behav Brain Sci 2020 Oct;7(2):115-122. [doi: 10.1177/2372732220943906]

XSL•FO
RenderX

52. Gaither SE. “Mixed” results: Multiracial research and identity explorations. Curr Dir Psychol Sci 2015;24(2):114-119. [doi: 10.1177/0963721414558115]
53. Pauker K, Meyers C, Sanchez DT, Gaither SE, Young DM. A review of multiracial malleability: identity, categorization, and shifting racial attitudes. Soc Personal Psychol Compass 2018 Jun;12(6):e12392. [doi: 10.1111/spc3.12392]
54. Maghbouleh N, Schachter A, Flores RD. Middle Eastern and North African Americans may not be perceived, nor perceive themselves, to be White. Proc Natl Acad Sci USA 2022 Feb 15;119(7):e2117940119. [doi: 10.1073/pnas.2117940119]
55. FAQs - race and ethnicity collection and reporting - longitudinal education data systems (CA Dept of Education). California Department of Education. URL: https://www.cde.ca.gov/ds/dc/es/refaq.asp [accessed 2021-11-01]
56. Lopez MH, Krogstad JM, Passel JS. Who is Hispanic? Pew Research Center. 2024 Sep 12. URL: https://www.pewresearch.org/short-reads/2024/09/12/who-is-hispanic/ [accessed 2025-04-30]
57. Passel JS, Krogstad JM. How a coding error provided a rare glimpse into Latino identity among Brazilians in the US. Pew Research Center. 2023 Apr 19. URL: https://www.pewresearch.org/short-reads/2023/04/19/how-a-coding-error-provided-a-rare-glimpse-into-latino-identity-among-brazilians-in-the-u-s/ [accessed 2025-04-30]
58. Race and National Origin: National Institutes of Health; 2025. URL: https://www.nih.gov/nih-style-guide/race-national-origin [accessed 2025-04-30]
59. Simon SS, Brucki SMD, Fonseca LM, et al. The (in)visible Brazilians: a perspective review on the need for brain health and dementia research with Brazilian immigrants in the United States. Alzheimers Dement (N Y) 2023 Sep 23;9(3):e12425. [doi: 10.1002/trc2.12425] [Medline: 37744309]
60. Fefferman NH, O’Neil EA, Naumova EN. Confidentiality and confidence: Is data aggregation a means to achieve both? J Public Health Pol 2005 Dec;26(4):430-449. [doi: 10.1057/palgrave.jphp.3200029]
61. Link MW, Mokdad AH, Stackhouse HF, Flowers NT. Race, ethnicity, and linguistic isolation as determinants of participation in public health surveillance surveys. Prev Chronic Dis 2006 Jan;3(1):A09. [Medline: 16356362]
62. Diaz V. Encouraging participation of minorities in research studies. Ann Fam Med 2012;10(4):372-373. [doi: 10.1370/afm.1426] [Medline: 22778129]
63. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. Am J Public Health 2014 Feb;104(2):e16-e31. [doi: 10.2105/AJPH.2013.301706] [Medline: 24328648]
64. Brunsma DL. Public categories, private identities: exploring regional differences in the biracial experience. Soc Sci Res 2006 Sep;35(3):555-576. [doi: 10.1016/j.ssresearch.2004.10.002]
65. Kressin NR, Chang BH, Hendricks A, Kazis LE. Agreement between administrative data and patients’ self-reports of race/ethnicity. Am J Public Health 2003 Oct;93(10):1734-1739. [doi: 10.2105/ajph.93.10.1734] [Medline: 14534230]
66. Krieger N, Van Wye G, Huynh M, et al. Structural racism, historical redlining, and risk of preterm birth in New York City, 2013-2017. Am J Public Health 2020 Jul;110(7):1046-1053. [doi: 10.2105/AJPH.2020.305656] [Medline: 32437270]

Abbreviations

MENA: Middle Eastern and North African
NHPI: Native Hawaiian or Other Pacific Islander
OMB: Office of Management and Budget
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani; submitted 21.08.24; peer-reviewed by CH Tejera, K Wang; revised version received 15.09.25; accepted 17.09.25; published 13.11.25.

Please cite as:
Krobath DM, Wilson NLW, Cuevas AG, Naumova EN, Sacheck JM, Tyler SA, Economos CD
Nonrandom Missingness in Child Race and Ethnicity Records and the US Federal Data Standards: Pooled Analysis of Community-Based Child Health Studies
JMIR Public Health Surveill 2025;11:e65660
URL: https://publichealth.jmir.org/2025/1/e65660
doi:10.2196/65660

© Danielle M Krobath, Norbert L W Wilson, Adolfo G Cuevas, Elena N Naumova, Jennifer M Sacheck, Sydney A Tyler, Christina D Economos. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Predictors of Cervical Cancer Screening Uptake in Sub-Saharan Africa Using Machine Learning: Cross-Sectional Study

Nebebe Demis Baykemagn[1], MPH; Mekuriaw Nibret Aweke[2], MPH; Amare Mesfin[2], MPH; Lemlem Daniel Baffa[2], MPH; Muluken Chanie Agimas[3], MPH; Habtamu Wagnew Abuhay[3], MPH; Dagnew Getnet Adugna[4], MSc; Tewodros Getaneh Alemu[5], MSc; Alemu Teshale Bicha[6], MD, MPH; Gebrie Getu Alemu[3], MPH

[1]Department of Health Informatics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[2]Department of Human Nutrition, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[3]Department of Epidemiology and Biostatistics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[4]Department of Human Anatomy, School of Medicine, College of Medicine and Health Science, University of Gondar, Gondar, Ethiopia

[5]Department of Pediatrics and Child Health Nursing, School of Nursing, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

[6]Department of Obstetrics and Gynecology, School of Medicine, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

Corresponding Author:
Nebebe Demis Baykemagn, MPH
Department of Health Informatics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia

Abstract

Background: Cervical cancer has been ranked as the fourth most common cancer affecting women, contributing to approximately 660,000 new diagnoses and 350,000 fatalities worldwide. Effective early screening has been shown to reduce cervical cancer incidence by up to 80% and prevent more than 40% of new cases.

Objective: This study aims to assess a machine learning–based prediction model and identify the key predictors influencing cervical cancer screening uptake among women aged 30 - 49 years in sub-Saharan Africa.

Methods: For this study, a weighted dataset of 33,952 individuals from the 2022 Demographic and Health Survey in Ghana, Kenya, Mozambique, and Tanzania was used. STATA version 17 (StataCorp) and Python 3.10 (Python Software Foundation) were used for data preprocessing and analysis. MinMax and standard scaler were applied for feature scaling, and recursive feature elimination was used for feature selection. An 80:20 ratio was applied for data splitting. Tomek links with random oversampling were used for handling class imbalance. A total of 7 models were selected and trained using both balanced and unbalanced datasets. Model evaluation was performed using area under the receiver operating characteristic curve, accuracy, and a confusion matrix.

Results: The proportion of cervical cancer screening in sub-Saharan Africa was 13%, which is lower than reported in previous studies. Random forest was the best-performing model, achieving an accuracy of 78%, an area under the curve of 86%, an F_1-score of 79%, a recall of 81%, and a precision of 77%. The waterfall plot's Shapley Additive Explanations analysis showed that wealth status, awareness of sexually transmitted infections, HIV testing exposure, age at first sexual intercourse, educational level, residency, smartphone ownership, having a single sexual partner, and previous health status were predictors of cervical cancer screening.

Conclusions: Improving education and awareness, expanding access to screening (especially in rural areas), leveraging both digital health and community-based outreach, integrating screening with other health services, and addressing socioeconomic barriers are recommended strategies to increase cervical cancer screening rates in sub-Saharan Africa.

(JMIR Public Health Surveill 2025;11:e71677) doi:10.2196/71677

KEYWORDS

cervical cancer; digital health; machine learning; woman; artificial intelligence

XSL•FO
RenderX

Introduction

Cervical cancer has been ranked as the fourth most common cancer affecting women, contributing to approximately 660,000 new diagnoses and 350,000 fatalities globally [1]. About 75% - 85% of these cases occur in low-income countries, with 23% specifically in Africa (as reported by the World Health Organization [WHO]). The incidence and mortality rates of cervical cancer are 2-3 times higher in low-income countries compared to high-income countries [2].

Cervical cancer ranks as the second leading cause of cancer-related death globally [3]. In 2025, studies from Africa showed that cervical cancer screening prevalence differs significantly between high-income countries (84%) and low- and middle-income countries (15%) [4]. Also, in Africa, more than 53% of cervical cancer cases are identified at an advanced stage (stage III-IV) [5].

Cervical cancer is "a malignant tumor that forms when the cells in the tissue surrounding the cervix grow and multiply uncontrollably by passing the normal process of cell division" [6]. Effective early screening has been shown to reduce cervical cancer incidence by up to 80% and prevent more than 40% of new cases [2,7].

With an average cure rate of 80%, radiation therapy effectively treats cervical cancer in its early stages [8]. However, the practice of early screening, before the disease reaches an advanced stage, continues to be a major problem in many parts of Africa [5].

Despite the implementation of screening initiatives and the introduction of the human papillomavirus (HPV) vaccine, cervical cancer remains a significant public health concern in Africa [9].

Early detection and timely treatment of cervical cancer are crucial [10]. The HPV has more than 100 serotypes, but the HPV vaccine targets only 9 of them [11]. Also, early screening is a key approach to achieving 1 of the 3 WHO 90-70-90 pillars: "90% of girls fully vaccinated against HPV, 70% of women screened by age 35 and 45 with a high-performance test, and 90% of women with precancerous lesions receiving treatment" [12,12].

Therefore, early screening is essential to enhance the probability of early diagnosis and treatment. Early screening is guaranteed to identify precancerous lesions and cancer before symptoms appear, which is the goal of screening, and enhances the chances of successful treatment and reduces maternal mortality [12-14].

However, studies show that screening in Africa is conducted at less than 70% of the WHO target [15]. This is due to a perceived lack of expertise, limited knowledge about cervical cancer, inadequate screening tools and supplies, and insufficient funding [5,16,17]. In the recent study, the prevalence of precancerous cervical lesions and cancer is relatively high, at 18% and 14%, respectively [18].

According to previous studies conducted in Africa, age, education level, employment status, wealth index, place of residence, and access to health care facilities were common predictors of cervical cancer screening among women of reproductive age [19-22].

Although studies have been conducted in African countries, they were limited to small sample sizes and single countries, lacking representativeness. This study, however, is representative, with a large dataset covering multiple countries, providing reliable evidence for policymakers and researchers to inform interventions.

The health care system produces vast amounts of data that must be analyzed to support effective decision-making. As the volume and complexity of this data grow, relying on traditional analytical methods without the aid of machine learning is becoming increasingly impractical.

A key strength of machine learning is its ability to handle big datasets while delivering reliable performance metrics. Although this study focuses on sub-Saharan Africa, the application of machine learning in this context allows for the extraction of meaningful patterns from complex health data.

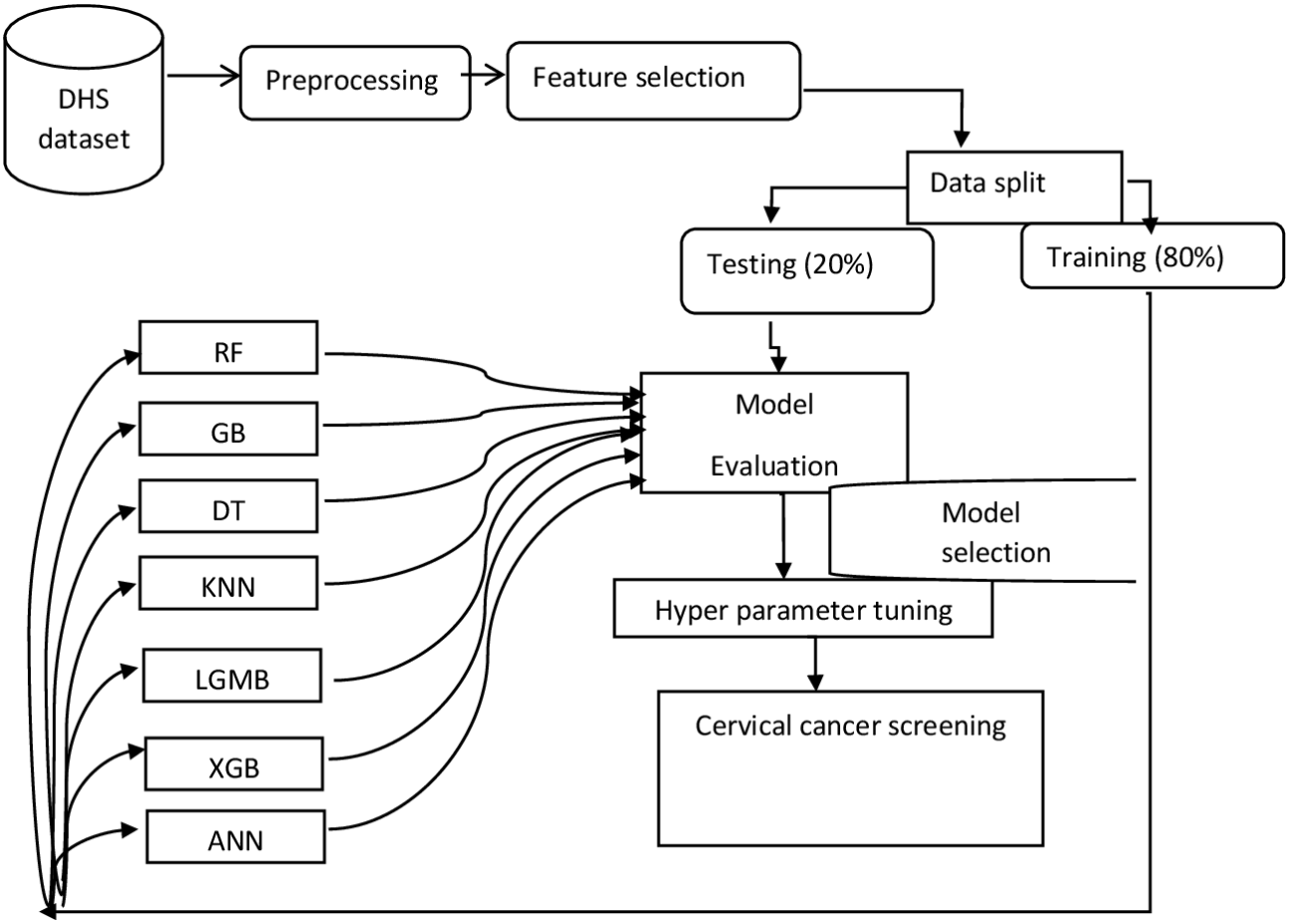
Methods

Data Processing and Management

After deciding on the title by considering the current public health issue, data were extracted from the individual record dataset. Then, the relationships between all variables and the outcome were reviewed based on previous study evidence and expert opinions. The similarity of the selected variables across 4 countries (Ghana, Kenya, Mozambique, and Tanzania) that had recent (2022 - 2023) Demographic and Health Surveys (DHS) data available was checked. STATA version 17 (StataCorp) and Python 3.10 (Python Software Foundation) were used for data processing and management. To handle missing data, mode imputation and K-nearest neighbors imputation were used to maintain the data distribution, ensure statistical consistency, and preserve the integrity of the dataset (Figure 1).

XSL•FO
RenderX

Figure 1. Overview flowchart of data preparation and analysis plan applied. ANN: artificial neural network; DHS: Demographic and Health Surveys; DT: decision tree; GB: gradient boosting; KNN: K-nearest neighbors; LGMB: light gradient boosting machine; RF: random forest; XGB: extreme gradient boosting.



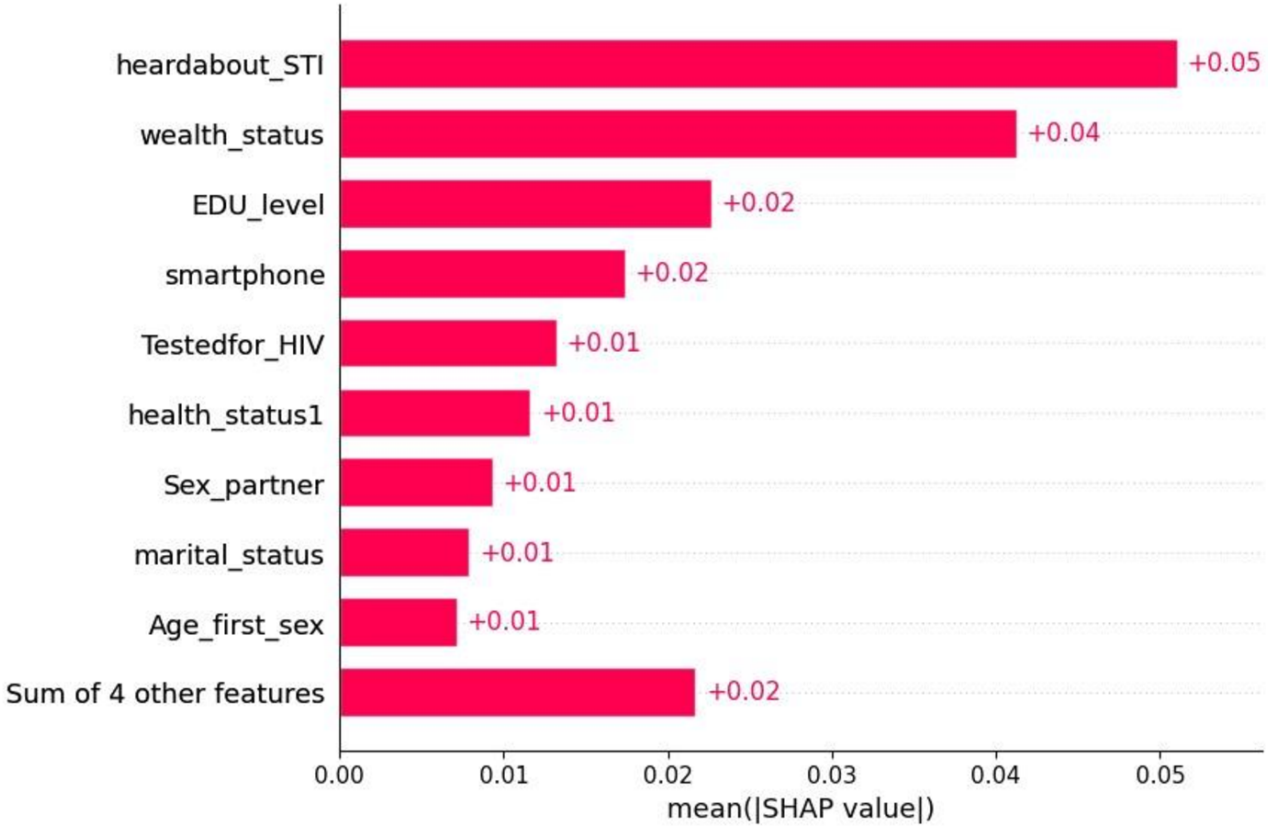
Many machine learning algorithms are sensitive to the scale of features and need normalization before modeling [23]. MinMaxScaler and StandardScaler were applied to normalize or standardize our data, ensuring a consistent scale and improving the performance of modeling tasks. The goal is to minimize performance bias in models that are sensitive to feature scaling, as the absence of standardization can negatively impact model performance.

Feature Selection

Recent studies indicate that irrelevant variables weaken the model's capacity for generalization, raise its overall complexity, and possibly lower a classifier's overall accuracy in a machine learning study [24]. Recursive feature elimination was used to methodically remove irrelevant features, decreasing dimensionality while retaining the most important predictive variables, thereby enhancing the model's efficiency and ability to generalize (Figure 2).

XSL•FO
RenderX

Figure 2. The top 10 features of cervical cancer screening. SHAP: Shapley Additive Explanations.



Splitting the Data

Previous studies indicate that data splitting offers a trustworthy estimation of the model's performance on unseen data and aids in the detection of overfitting [25]. For data splitting, an 80:20 ratio was used for training and testing, respectively, to assess the model's performance and prevent overfitting in machine learning algorithms.

Handling Imbalanced Data

In machine learning, class imbalance is frequently encountered due to unbalanced data, and this can have a big influence on model accuracy [26].

To improve model generalization and minimize noise and class overlap, we applied the synthetic minority oversampling technique (SMOTE) in combination with Tomek links undersampling. This process resulted in a balanced and cleaner dataset, where the original distribution of 3878 "yes" cases and 30,074 "no" cases was adjusted to 30,074 instances in each class.

Model Selection

After a detailed review of machine learning studies conducted on reproductive health, 7 models were selected based on simplicity, accuracy, robustness, and computational efficiency, including decision tree, random forest (RF), K-nearest neighbors, extreme gradient boosting, light gradient boosting machine, adaptive boosting, and gradient boosting (GB).

Model Training

Both balanced and unbalanced datasets were used to train the selected classifiers after the model was selected. They used 10-fold cross-validation to assess their performance. In order to generate final predictions on test data that had not yet been observed, the best successful predictive model was selected after the comparison and trained using balanced training data.

Hyperparameter Tuning

A crucial step in maximizing the potential of machine learning models and ensuring their accuracy, efficiency, and robustness is hyperparameter tuning. Hyperparameter tuning by RF is listed in Table 1.

XSL•FO
RenderX

Table . Hyperparameters and values for all models.

Classifier	Hyperparameter and value
DT[a]	max_depth=10, criterion=gini, min_samples_split=10, min_sample_leaf=1,
RF[b]	max_depth=20,min_samples_split=10, n_estimators=300
KNN[c]	n_neighbors=5, weights='uniform', metric='minkowski'
ANN[d]	hidden_layer_sizes=(100,), activation='relu', solver='adam', learning_rate='adaptive', max_iter=300
XGB[e]	colsample_bytree=0.8, learning_rate=0.05, max_depth=3, n_estimators=100, subsample=1.0
LGBM[f]	learning_rate=0.1, n_estimators=20, num_leaves=31
ADA[g]	n_estimators=300, learning_rate=1.0
GB[h]	n_estimators=300, learning_rate=0.1, max_depth=3, min_samples_split=10

[a]DT: decision tree.
[b]RF: random forest.
[c]KNN: K-nearest neighbors.
[d]ANN: artificial neural network.
[e]XGB: extreme gradient boosting.
[f]LGBM: light gradient boosting machine.
[g]ADA: adaptive boosting.
[h]GB: gradient boosting.

Model Evaluation

Model evaluation was performed using confusion matrix, F_1-score, area under the receiver operating characteristic curve, accuracy, precision, and recall in this study. Ultimately, the RF model outperformed others due to its effectiveness in handling nonlinear relationships, evaluating feature importance, and minimizing overfitting through the ensemble of multiple decision trees. For results, see Figure 3 and Tables 2 and 3.

XSL•FO
RenderX

Figure 3. ROC curve for model comparison to predict cervical cancer screening. ADA: adaptive boosting; AUC: area under the curve; DT: decision tree; GB: gradient boosting; KNN: K-nearest neighbors; LGMB: light gradient boosting machine; RF: random forest; ROC: receiver operating characteristic; XGB: extreme gradient boosting.

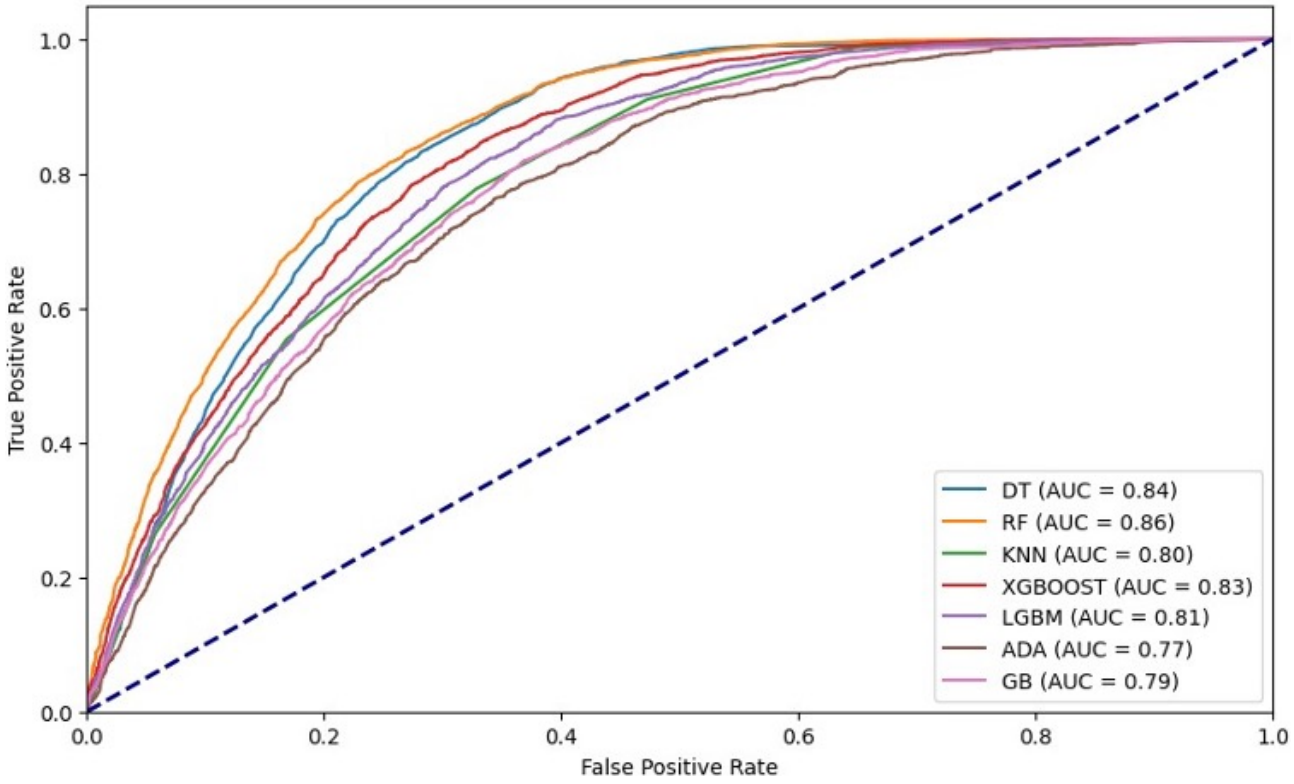


Table . Model evaluation after synthetic minority oversampling technique.

Model	Accuracy (%)	AUC[a] (%)	F_1-score (%)	Recall (%)	Precision (%)
DT[b]	78	85	78	80	76
RF[c]	78	86	79	81	77
KNN[d]	72	80	74	75	73
XGBoost[e]	76	83	76	78	75
LGBM[f]	74	81	75	76	74
ADA[g]	72	77	71	73	69
GB[h]	72	79	72	74	70

[a]AUC: area under the curve.
[b]DT: decision tree.
[c]RF: random forest.
[d]KNN: K-nearest neighbors.
[e]XGBoost: extreme gradient boosting.
[f]LGBM: light gradient boosting machine.
[g]ADA: adaptive boosting.
[h]GB: gradient boosting.

Table . Confusion matrix for cervical cancer prediction by random forest accuracy (2802+2494+837+658=6791=78%).

	Predicted positive	Predicted negative
Actual positive	2802 (true positive)	658 (false negative)
Actual negative	837 (false positive)	2494 (true negative)

XSL•FO
RenderX

Ethical Considerations

This study did not require institutional review board approval because it involved secondary analysis of publicly available and deidentified data from the DHS Program. As a result, a consent letter for data access was obtained from a major health and demographic survey program via a web-based request submitted at the DHS Program website [27]. All data used in this study consisted of nonidentifiable information.

Results

Cervical Cancer Screening Distribution by Country

The proportion of women screening rates in Kenya (14%), Mozambique (13%), Tanzania (12%), and Ghana (7%) remains low (Table 4).

Table . Cervical cancer screening distribution by country.

Country	Categories	
	Yes, n (%)	No, n (%)
Ghana	530.8 (7)	6767.2 (92)
Kenya	2053 (14)	12,389 (85)
Mozambique	669 (13)	4575.8 (88)
Tanzania	842.6 (12)	6068.1 (88)

Individual Characteristics of Cervical Cancer Screening

The majority of women living in rural (17,987.6/19,690.1, 91%) and urban areas (11,812.7/14,205.8, 83%) have not undergone cervical cancer screening. Among women who own a smartphone, 2170.8 of 11,628.9 (19%) have been screened for cervical cancer, while only 1924.8 of 22,267.0 (9%) of those without a smartphone have been screened.

Women who self-report good health status have a cervical cancer screening rate of 15% (156.9/1,056.6) compared to 11% (2762.9/24,302.6) among those who report poor health status. Women with high wealth status have a cervical cancer screening rate of 18% (2884.8/15,923.9), compared to 5% (604.5/11,398.8) among those in the poorest wealth category.

Only 286.1 of 1875.2 (15%) women who use family planning methods have been screened for cervical cancer. Those who do not consider distance a big problem had a slightly higher screening rate (3226.1/26,048.9, 12%). Women who initiated sexual activity after the age of 30 years had a screening rate of 14% (1932.9/13,527.5), higher than the 11% (2162.7/20,368.2) among those who began before the age of 30 years.

Women reporting multiple sexual partners had a significantly higher screening rate (793.2/4218, 19%) than those who did not (3302.5/29,677.9, 11%). Women who have been tested for HIV have a cervical cancer screening rate of 13% (3958.3/29,777), compared to only 3% (137.3/4118.9) among those who have not been tested. Divorced and widowed women show slightly higher screening rates at 15% (676.3/4638.9) and 14% (258.5/1879.7), respectively.

Women with tertiary education have a screening rate of 22% (838.7/3738.8), compared to only 4% (238.6/6572.2) among those with no education. Women who have heard about sexually transmitted infections (STIs) have a screening rate of 17% (3900.8/23,313.8), compared to only 1.9% (194.8/10,582.1) among those who have not (Table 5).

XSL•FO
RenderX

Table . Individual characteristics of cervical cancer screening.

Variables and categories	Yes, n (%)	No, n (%)
Residence		
Urban	2393.1 (17)	11,812.7 (83)
Rural	1702.5 (9)	17,987.6 (91)
Frequently social media usage		
No	2873.5 (12)	20,629.5 (88)
Yes	1222.2 (12)	9170.8 (88)
Availability of smartphone		
No	1924.8 (9)	20,342.2 (91)
Yes	2170.8 (19)	9458.1 (81)
Self-report of health status		
Good	156.9 (15)	899.7 (85)
Moderate	1175.9 (14)	7360.8 (86)
Poor	2762.9 (11)	21,539.7 (88)
Wealth		
Lowest	604.5 (5)	10,794.3 (94)
Middle	606.3 (9)	5966.9 (91)
Highest	2884.8 (18)	13,039.1 (82)
Family planning method use		
No	3809.5 (12)	28,211.2 (88)
Yes	286.1 (15)	1589.1 (84)
Distance to health facility		
Big problem	869.6 (11)	6977.4 (88)
Not a big problem	3226.1 (12)	22,822.8 (87)
Age at first sexual intercourse (years)		
>30	1932.9 (14)	11,594.6 (85)
<30	2162.7 (11)	18,205.5 (89)
Have multiple sexual partners		
No	3302.5 (11)	26,375.4 (89)
Yes	793.2 (19)	3424.8 (81)
Tested for HIV/AIDS		
No	137.3 (3)	3981.6 (96)
Yes	3958.3 (13)	25,818.7 (86)
Marital status		
Single	254 (13)	1751.7 (87)
Married	2906.7 (12)	22,464.6 (89)
Widowed	258.5 (14)	1621.2 (86)
Divorced	676.3 (15)	3962.6 (85)
Educational status		
No education	238.6 (4)	6333.6 (96)
Primary	1668.7 (12)	12,746.3 (88)
Secondary	1349.6 (15)	7820.2 (85)
Tertiary	838.7 (22)	2900.1 (78)

XSL•FO
RenderX

Variables and categories	Yes, n (%)	No, n (%)
Heard about STI[a]		
No	194.8 (1.9)	10,387.3 (98)
Yes	3900.8 (17)	19,413 (83)

[a]STI: sexually transmitted infection.

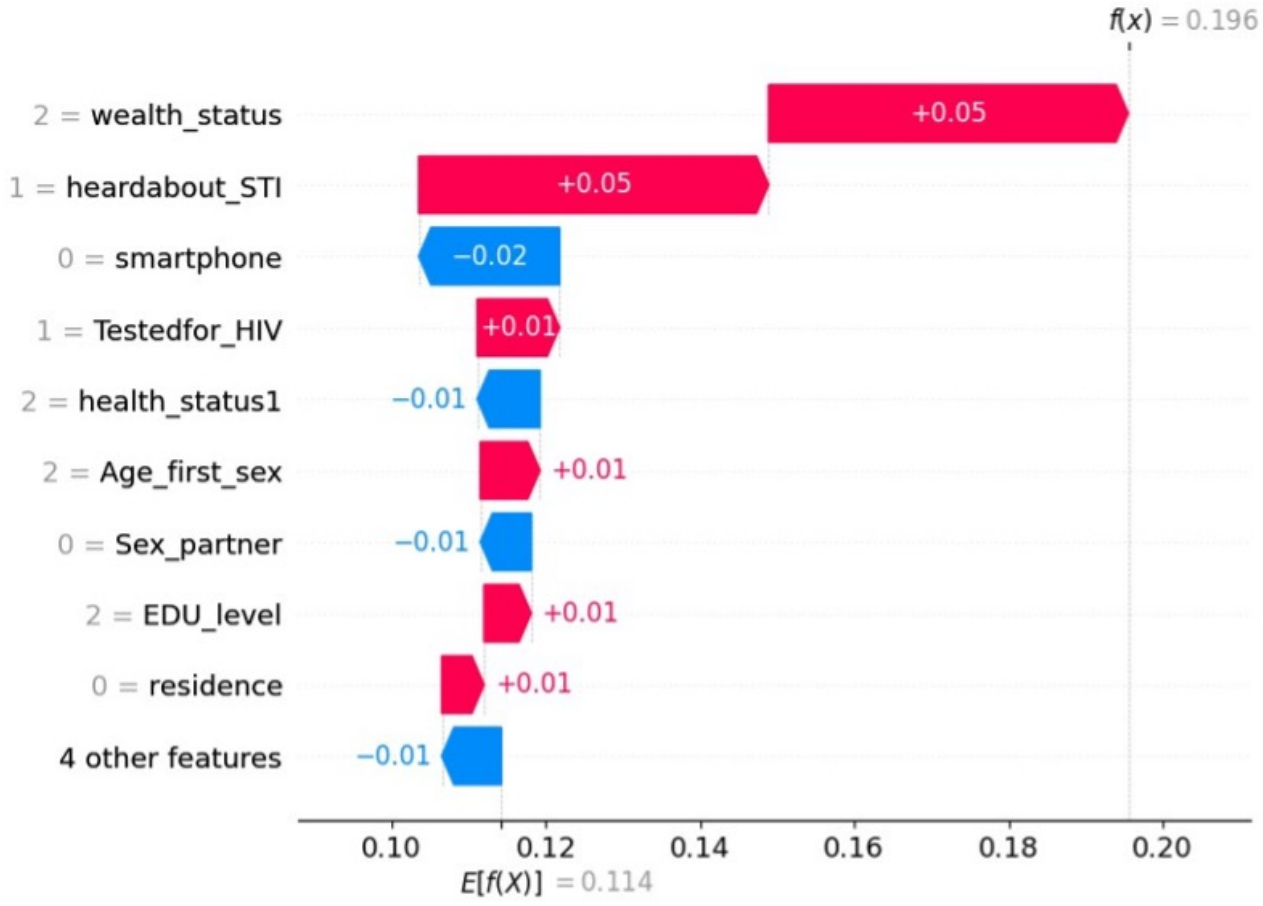
Determinants of Cervical Cancer Screening

In machine learning, Shapley Additive Explanations (SHAP) values indicate how each feature impacts the output variable. This study examines the impact of the top 10 features on women's cervical cancer screening status. These insights are crucial for specific interventions and policy formulation aimed at enhancing cervical cancer screening practices, as well as reducing morbidity and mortality among women.

In Figure 4, a feature contributing positively to the predicted screening practice is represented by a positive SHAP value (red), while a feature contributing negatively to the predicted screening outcome is represented by a negative SHAP value (blue).

The best predictive model (RF) showed that wealth status, awareness of STIs, HIV testing, age at first sex, primary education and above, and living in urban areas are significant factors associated with increased cervical cancer screening. However, factors such as not owning a smartphone, having a single sexual partner, and unknown health status are associated with a decrease in cervical cancer screening (Figure 4).

Figure 4. Determinants of cervical cancer screening visualized by a waterfall plot.



Discussion

Principal Findings

In sub-Saharan Africa, cervical cancer is the second most common cause of cancer-related death on the continent and the fourth leading cause of cancer-related death globally [28,29]. Low- and middle-income countries account for nearly 85% of cervical cancer cases and related deaths [28]. However, according to this study, a small percentage of women were screened for cervical cancer in Kenya (14.22%), Mozambique (12.76%), Tanzania (12.20%), and Ghana (7.27%). This low screening rate contributes to delays in early cancer management and highlights the urgent need for targeted interventions to improve screening behaviors among women.

The predictive evidence from this machine is used to support the realization of the 2030 plan [30]: 90% of cervical diseases

XSL•FO
RenderX

identified, 90% of precancerous cases treated, and 90% of women with invasive cancer managed.

The RF is the best-performing model, achieving an accuracy of 78%, an area under the curve of 86%, an F_1-score of 79%, a recall of 81%, and a precision of 77%. The waterfall plot's SHAP analysis indicates that variables, such as middle and higher wealth status, awareness of STIs, HIV testing, age at first sex, primary education and above, and living in urban areas are significant factors associated with increased cervical cancer screening.

However, not having a smartphone, having a single sexual partner, and being unaware of health status are associated with a decrease in cervical cancer screening.

In this study, the proportion of cervical cancer screening is 13%, which is higher than findings from studies conducted in Cameroon (4%) [2], South India (7.1%) [31], and lower than those reported in Ethiopia (14.79%) [32], and sub-Saharan Africa (19%) [33]. Possible justifications for this difference may include public awareness levels, accessibility of screening services, outreach programs, community health education, socioeconomic status, and geographic barriers. Differences in study populations, sampling methods, and time frames may also contribute to these variations [34].

According to this finding, women with middle and higher wealth status have increased cervical cancer screening rates, which is supported by previous research conducted in the United States, Ethiopia, Uganda, and Cameroon [2,35-37]. One explanation could be that women with middle-class and upper-class incomes have easier access to health care because they can afford the direct and indirect costs of medical care, including transportation and other costs.

This finding shows that women who have information about STIs are more likely to undergo cervical cancer screening. This is supported by previous studies conducted in Southern Ethiopia, Eastern and Southern Africa, Haiti, and Morocco [38-41]. This is because women who have information about STIs are also exposed to information about cervical cancer, including the benefits of early screening, and they tend to have better health-seeking behavior.

According to this SHAP finding, women who do not have smartphones show decreased cervical cancer screening rates. This is supported by previous studies conducted in Japan and England [42,43]. Those who have smartphones can easily access health information, including preventive health care such as screenings, through websites, mobile apps, and social media. Those who do not have smartphones may miss this information and, as a result, may not pay attention to screening [44].

The SHAP finding showed that women who have previously been tested for HIV are more likely to undergo cervical cancer screening. This is consistent with previous studies conducted in Uganda, China, Kenya, and Ethiopia [45-48].

Also, this study indicated that women who did not have information about their health status had decreased cervical cancer screening practices. This is supported by previous studies conducted in Nigeria, Ghana, Australia, India, and Kenya [49-53]. Women who do not have information about their health status are less likely to undergo cervical cancer screening because a lack of awareness reduces their understanding of risk, engagement with health care, and motivation to seek preventive services.

This study showed that women who had their first sexual intercourse before the age of 30 years had an increased cervical screening status. This is in line with studies conducted in East Gojjam Zone, Ethiopia, Africa, and Nigeria [54-56]. A possible explanation for this is that early sexual activity is associated with a higher likelihood of STIs, which in turn leads to more frequent visits to health facilities due to increased awareness of cervical cancer risk.

In addition, women in this age group are more educated and more likely to understand messages from social media and health professionals, including information about cervical cancer risk and the benefits of early screening and treatment [56].

SHAP findings showed that women who have a single sexual partner have decreased cervical cancer screening rates. This finding was supported by studies conducted in Korea, Jordan, Ethiopia, and Uganda [57-60].

Women with a single partner have high confidence that they are free from cervical cancer risk factors because, during health facility visits and through social media information, they understand that women with single sexual partners are at a lower risk of contracting HPV and other STIs, which are the main causes of cervical cancer [61].

This finding also showed that women with primary and higher education levels have increased cervical cancer screening status, which is in line with previous studies conducted in Romania, the Southwest United States, Mayuge, Ghana, and Korea [62-66].

A possible explanation for this could be that educated women have greater health literacy, allowing them to easily understand basic health information and services, including information about cervical cancer. Educated women also tend to have better access to resources such as health care providers, smartphones, and official websites like those of the WHO.

In addition, they may be less influenced by cultural misconceptions or stigma surrounding cervical cancer screening [67].

SHAP findings also showed that women residing in urban areas had increased cervical cancer screening rates. This is in line with previous studies conducted in Nepal, Nigeria, China, and Ghana [55,68-70]. Women residing in urban areas have improved access to health information and services through social media, public health campaigns, and proximity to health care facilities and professionals. Consequently, they tend to have a better understanding of cervical cancer compared to their rural counterparts [71].

Strengths and Limitations

One of the study's strengths is that it makes use of a large and varied dataset, which improves the findings' applicability to a larger geographic area. Furthermore, complex patterns and

XSL•FO
RenderX

relationships that traditional statistics can miss might be found through the use of machine learning models, and the limitations of this include issues that can impact prediction accuracy, such as self-reported information, recall bias, and missing data. Furthermore, it is challenging to determine causal links due to the cross-sectional nature of data, and the findings based on machine learning outputs should be interpreted as associations rather than causations.

Conclusions

Wealth status, awareness of STIs, HIV testing exposure, age at first sexual intercourse, educational level, residency, smartphone ownership, having a single sexual partner, and previous health status are predictors of cervical cancer screening.

Improving education and awareness, expanding access to screening, especially in rural areas, leveraging both digital health and community-based outreach, integrating screening with other health services, and addressing socioeconomic barriers are recommended strategies to increase cervical cancer screening rates in sub-Saharan Africa.

Acknowledgments

The authors would like to express their gratitude for the valuable time and expertise provided by everyone involved in this study. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Data Availability

The datasets analyzed in this study are available in the public domain through the Measure DHS website [72].

Authors' Contributions

NDB, GGA, and MNA developed the concept for the study. AM, LDB, MCA, HWA, DGA, TGA, and ATB reviewed the literature. NDB, GGA, MNA, and HWA conducted the data analysis. NDB, GGA, MNA, HWA, and DGA discussed the findings.

Conflicts of Interest

None declared.

References

1. Cervical cancer. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/cervical-cancer [accessed 2025-08-18]
2. Okyere J, Duodu PA, Aduse-Poku L, Agbadi P, Nutor JJ. Cervical cancer screening prevalence and its correlates in Cameroon: secondary data analysis of the 2018 demographic and health surveys. BMC Public Health 2021 Jun 5;21(1):1071. [doi: 10.1186/s12889-021-11024-z] [Medline: 34090372]
3. Amorha KC, Ozota GO, Ndunwere MGO, Anyaji UL, Egbo OF, Ogugofor OA. Knowledge, attitudes, and practices of adolescent girls regarding cervical cancer: a cross-sectional study in Enugu State, Nigeria. Pan Afr Med J 2024;47:17. [doi: 10.11604/pamj.2024.47.17.41087] [Medline: 38524107]
4. Olivieri DJ, Eastment MC, Mugisha N, Menon MP. Correlates of cervical cancer awareness among women aged 30-49 in five sub-Saharan African nations: evidence from the Demographic and Health Survey (DHS)-2017-2023. PLOS Glob Public Health 2025;5(5):e0003344. [doi: 10.1371/journal.pgph.0003344] [Medline: 40333666]
5. Mantula F, Toefy Y, Sewram V. Barriers to cervical cancer screening in Africa: a systematic review. BMC Public Health 2024 Feb 20;24(1):525. [doi: 10.1186/s12889-024-17842-1] [Medline: 38378542]
6. Uddin KMM, Al Mamun A, Chakrabarti A, Mostafiz R, Dey SK. An ensemble machine learning-based approach to predict cervical cancer using hybrid feature selection. Neuroscience Informatics 2024 Sep;4(3):100169. [doi: 10.1016/j.neuri.2024.100169]
7. Saaka SA, Hambali MG. Factors associated with cervical cancer screening among women of reproductive age in Ghana. BMC Womens Health 2024 Sep 18;24(1):519. [doi: 10.1186/s12905-024-03367-7] [Medline: 39289681]
8. Shakil R, Islam S, Akter B. A precise machine learning model: detecting cervical cancer using feature selection and explainable AI. J Pathol Inform 2024 Dec;15:100398. [doi: 10.1016/j.jpi.2024.100398] [Medline: 39494050]
9. Ong SK, Abe SK, Thilagaratnam S, et al. Towards elimination of cervical cancer - human papillomavirus (HPV) vaccination and cervical cancer screening in Asian National Cancer Centers Alliance (ANCCA) member countries. Lancet Reg Health West Pac 2023 Oct;39:100860. [doi: 10.1016/j.lanwpc.2023.100860] [Medline: 37576906]
10. Advancing the cervical cancer elimination agenda in the African region. : WHO Regional Office for Africa; 2024 Nov URL: https://www.afro.who.int/media-centre/statements-commentaries/advancing-cervical-cancer-elimination-agenda-african-region [accessed 2025-09-10]

XSL•FO
RenderX

11. Human Papillomavirus (HPV) Infection: International Agency for Research on Cancer; 2007. URL: https://www.ncbi.nlm.nih.gov/books/NBK321770/ [accessed 2025-09-10]
12. Beyene T, Akibu M, Bekele H, Seyoum W. Risk factors for precancerous cervical lesion among women screened for cervical cancer in south Ethiopia: unmatched case-control study. PLoS One 2021;16(7):e0254663. [doi: 10.1371/journal.pone.0254663] [Medline: 34265004]
13. Holcakova J, Bartosik M, Anton M, et al. New trends in the detection of gynecological precancerous lesions and early-stage cancers. Cancers (Basel) 2021 Dec 17;13(24):6339. [doi: 10.3390/cancers13246339] [Medline: 34944963]
14. Yan T, Wong PK, Qin YY. Deep learning for diagnosis of precancerous lesions in upper gastrointestinal endoscopy: a review. World J Gastroenterol 2021 May 28;27(20):2531-2544. [doi: 10.3748/wjg.v27.i20.2531] [Medline: 34092974]
15. Yang L, Boily MC, Rönn MM, et al. Regional and country-level trends in cervical cancer screening coverage in sub-Saharan Africa: a systematic analysis of population-based surveys (2000-2020). PLoS Med 2023 Jan;20(1):e1004143. [doi: 10.1371/journal.pmed.1004143] [Medline: 36634119]
16. Adewumi K, Nishimura H, Oketch SY, Adsul P, Huchko M. Barriers and facilitators to cervical cancer screening in Western Kenya: a qualitative study. J Cancer Educ 2022 Aug;37(4):1122-1128. [doi: 10.1007/s13187-020-01928-6] [Medline: 33411253]
17. Tshabalala G, Blanchard C, Mmoledi K, et al. A qualitative study to explore healthcare providers' perspectives on barriers and enablers to early detection of breast and cervical cancers among women attending primary healthcare clinics in Johannesburg, South Africa. PLOS Glob Public Health 2023;3(5):e0001826. [doi: 10.1371/journal.pgph.0001826] [Medline: 37159437]
18. Tenkir L, Mamuye A, Jemebere W, Yeheyis T. The magnitude of precancerous cervical lesions and its associated factors among women screened for cervical cancer at a referral center in Southern Ethiopia, 2021: a cross-sectional study. Front Glob Womens Health 2023;4:1187916. [doi: 10.3389/fgwh.2023.1187916] [Medline: 37664421]
19. Ampofo AG, Boyes AW, Asibey SO, Oldmeadow C, Mackenzie LJ. Prevalence and correlates of modifiable risk factors for cervical cancer and HPV infection among senior high school students in Ghana: a latent class analysis. BMC Public Health 2023 Feb 15;23(1):340. [doi: 10.1186/s12889-022-14908-w] [Medline: 36793003]
20. El-Zein M, Ramanakumar AV, Naud P, et al. Determinants of acquisition and clearance of human papillomavirus infection in previously unexposed young women. Sex Transm Dis 2019 Oct;46(10):663-669. [doi: 10.1097/OLQ.0000000000001053] [Medline: 31464859]
21. Guo C, Zhan B, Li MY, Yue L, Zhang C. Association between oral contraceptives and cervical cancer: a retrospective case-control study based on the National Health and Nutrition Examination Survey. Front Pharmacol 2024;15:1400667. [doi: 10.3389/fphar.2024.1400667] [Medline: 39086392]
22. Torres-Poveda K, Ruiz-Fraga I, Madrid-Marina V, Chavez M, Richardson V. High risk HPV infection prevalence and associated cofactors: a population-based study in female ISSSTE beneficiaries attending the HPV screening and early detection of cervical cancer program. BMC Cancer 2019 Dec 10;19(1):1205. [doi: 10.1186/s12885-019-6388-4] [Medline: 31823749]
23. Feature scaling in machine learning: standardization, MinMaxScaling and more. Train In Data. URL: https://www.blog.trainindata.com/feature-scaling-in-machine-learning/#:~:text=Many%20machine%20learning%20algorithms%20are%20sensitive [accessed 2025-08-18]
24. Vyas A, Raman S, Sen S, et al. Machine learning-based diagnosis and ranking of risk factors for diabetic retinopathy in population-based studies from South India. Diagnostics (Basel) 2023 Jun 16;13(12):2084. [doi: 10.3390/diagnostics13122084] [Medline: 37370980]
25. Muraina I. Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts. Presented at: 7th International Mardin Artuklu scientific research conference; Dec 10-12, 2021; Mardin, Turkey URL: https://tinyurl.com/2j7j9ezm [accessed 2025-09-04]
26. 10 techniques to handle imbalanced classes in machine learning. Analytics Vidhya. URL: https://www.analyticsvidhya.com/articles/class-imbalance-in-machine-learning [accessed 2025-08-18]
27. The DHS Program. URL: https://dhsprogram.com/ [accessed 2025-08-18]
28. Cervical cancer. World Health Organization African region. URL: https://www.afro.who.int/health-topics/cervical-cancer [accessed 2025-08-18]
29. Shocking statistics: cervical cancer claims 70% of cases in Africa, experts call for action. Top Africa news. URL: https://www.topafricanews.com/2025/04/24 [accessed 2025-08-18]
30. Continental consultative meeting report: accelerating the plan to eliminate cervical cancer in Africa by 2030. : AfricaCDC URL: https://africacdc.org/download/continental-consultative-meeting-report-accelerating-the-plan-to-eliminate-cervical-cancer-in-africa-by-2030/ [accessed 2025-09-04]
31. Reichheld A, Mukherjee PK, Rahman SM, David KV, Pricilla RA. Prevalence of cervical cancer screening and awareness among women in an urban community in South India-a cross sectional study. Ann Glob Health 2020 Mar 16;86(1):30. [doi: 10.5334/aogh.2735] [Medline: 32211300]

XSL•FO
RenderX

32. Desta M, Getaneh T, Yeserah B, et al. Cervical cancer screening utilization and predictors among eligible women in Ethiopia: a systematic review and meta-analysis. PLoS One 2021;16(11):e0259339. [doi: 10.1371/journal.pone.0259339] [Medline: 34735507]
33. Ba DM, Ssentongo P, Musa J, et al. Prevalence and determinants of cervical cancer screening in five sub-Saharan African countries: a population-based study. Cancer Epidemiol 2021 Jun;72:101930. [doi: 10.1016/j.canep.2021.101930] [Medline: 33756434]
34. Asgedom YS, Kassie GA, Habte A, Ketema DB, Aragaw FM. Socioeconomic inequality in cervical cancer screening uptake among women in sub-Saharan Africa: a decomposition analysis of Demographic and Health Survey data. BMJ Open 2024 Dec 10;14(12):e088753. [doi: 10.1136/bmjopen-2024-088753] [Medline: 39658286]
35. Isabirye A, Mbonye MK, Kwagala B. Predictors of cervical cancer screening uptake in two districts of Central Uganda. PLoS ONE 2020;15(12):e0243281. [doi: 10.1371/journal.pone.0243281] [Medline: 33270792]
36. Miles-Richardson S, Allen S, Claridy MD, Booker EA, Gerbi G. Factors associated with self-reported cervical cancer screening among women aged 18 years and older in the United States. J Community Health 2017 Feb;42(1):72-77. [doi: 10.1007/s10900-016-0231-5] [Medline: 27496176]
37. Misgun T, Demissie DB. Knowledge, practice of cervical cancer screening and associated factors among women police members of Addis Ababa police commission Ethiopia. BMC Cancer 2023 Oct 10;23(1):961. [doi: 10.1186/s12885-023-11478-x] [Medline: 37817082]
38. Amado G, Weldegebreal F, Birhanu S, Dessie Y. Cervical cancer screening practices and its associated factors among females of reproductive age in Durame town, Southern Ethiopia. PLoS ONE 2022;17(12):e0279870. [doi: 10.1371/journal.pone.0279870] [Medline: 36584208]
39. Bendahhou K, Serhier Z, Diouny S, et al. Women's knowledge and attitudes towards cervical cancer screening in Morocco. Cureus 2023 Apr;15(4):e37989. [doi: 10.7759/cureus.37989] [Medline: 37223139]
40. Macleod CI, Reynolds JH. Human papilloma virus infection and cervical cancer among women who sell sex in Eastern and Southern Africa: a scoping review. Womens Health (Lond) 2021;17:17455065211058349. [doi: 10.1177/17455065211058349] [Medline: 34775848]
41. McCarthy SH, Walmer KA, Boggan JC, et al. Awareness of cervical cancer causes and predeterminants of likelihood to screen among women in Haiti. J Low Genit Tract Dis 2017 Jan;21(1):37-41. [doi: 10.1097/LGT.0000000000000281] [Medline: 27906806]
42. Groves S, Brooks J. What do young women below national screening age in England think about cervical cancer and cervical screening? A qualitative study. J Clin Nurs 2022 Jun;31(11-12):1588-1597. [doi: 10.1111/jocn.16012] [Medline: 34418192]
43. Irino S, Ose H, Takata N, Kamoshida S, Ohsaki H. Barriers to undergoing cervical cancer screening among health sciences university students in Japan: a cross-sectional study. Nurs Health Sci 2023 Sep;25(3):466-473. [doi: 10.1111/nhs.13043] [Medline: 37581364]
44. Ghahramani F, Wang J. Impact of smartphones on quality of life: a health information behavior perspective. Inf Syst Front 2020 Dec;22(6):1275-1290. [doi: 10.1007/s10796-019-09931-z]
45. Choi Y, Ibrahim S, Park LP, Cohen CR, Bukusi EA, Huchko MJ. Uptake and correlates of cervical cancer screening among women attending a community-based multi-disease health campaign in Kenya. BMC Womens Health 2022 Apr 18;22(1):122. [doi: 10.1186/s12905-022-01702-4] [Medline: 35436908]
46. Emru K, Abebaw TA, Abera A. Role of awareness on cervical cancer screening uptake among HIV positive women in Addis Ababa, Ethiopia: a cross-sectional study. Womens Health (Lond) 2021;17:17455065211017041. [doi: 10.1177/17455065211017041] [Medline: 34096400]
47. Lin S, Chen WT, Gu C, Cheng HL, Wang H, Tang S. Knowledge, perception of HIV symptom severity and cervical cancer screening behaviour among women living with HIV in China. Eur J Cancer Care (Engl) 2022 Mar;31(2):e13542. [doi: 10.1111/ecc.13542] [Medline: 34961999]
48. Ninsiima M, Nyabigambo A, Kagaayi J. Acceptability of integration of cervical cancer screening into routine HIV care, associated factors and perceptions among HIV-infected women: a mixed methods study at Mbarara Regional Referral Hospital, Uganda. BMC Health Serv Res 2023 Apr 3;23(1):333. [doi: 10.1186/s12913-023-09326-6] [Medline: 37013535]
49. Brown RF, Muller TR, Olsen A. Australian women's cervical cancer screening attendance as a function of screening barriers and facilitators. Soc Sci Med 2019 Jan;220:396-402. [doi: 10.1016/j.socscimed.2018.11.038] [Medline: 30529797]
50. Eastment MC, Wanje G, Richardson BA, et al. A cross-sectional study of the prevalence, barriers, and facilitators of cervical cancer screening in family planning clinics in Mombasa County, Kenya. BMC Health Serv Res 2022 Dec 23;22(1):1577. [doi: 10.1186/s12913-022-08984-2] [Medline: 36564740]
51. Ghosh S, Mallya SD, Shetty RS, et al. Knowledge, attitude and practices towards cervical cancer and its screening among women from tribal population: a community-based study from Southern India. J Racial Ethn Health Disparities 2021 Feb;8(1):88-93. [doi: 10.1007/s40615-020-00760-4] [Medline: 32333376]
52. Rimande-Joel R, Ekenedo GO. Knowledge, belief and practice of cervical cancer screening and prevention among women of Taraba, North-East Nigeria. Asian Pac J Cancer Prev 2019 Nov 1;20(11):3291-3298. [doi: 10.31557/APJCP.2019.20.11.3291] [Medline: 31759351]

XSL•FO
RenderX

53. Stuart A, Obiri-Yeboah D, Adu-Sarkodie Y, Hayfron-Benjamin A, Akorsu AD, Mayaud P. Knowledge and experience of a cohort of HIV-positive and HIV-negative Ghanaian women after undergoing human papillomavirus and cervical cancer screening. BMC Womens Health 2019 Oct 23;19(1):123. [doi: 10.1186/s12905-019-0818-y] [Medline: 31647013]
54. Mekonnen AG, Mittiku YM. Early-onset of sexual activity as a potential risk of cervical cancer in Africa: a review of literature. PLOS Glob Public Health 2023;3(3):e0000941. [doi: 10.1371/journal.pgph.0000941] [Medline: 36962975]
55. Olubodun T, Ogundele OO, Salisu ZA, Odusolu YO, Caleb-Ugwuowo UU. Cervical cancer awareness and risk factors among women residing in an urban slum in Lagos, Southwest Nigeria. Afr Health Sci 2023 Sep;23(3):269-279. [doi: 10.4314/ahs.v23i3.33] [Medline: 38357134]
56. Tesfaw K, Kindie W, Mulatu K, Bogale EK. Utilisation of cervical cancer screening and factors associated with screening utilisation among women aged 30-49 years in Mertule Mariam Town, East Gojjam Zone, Ethiopia, in 2021: a cross-sectional survey. BMJ Open 2022 Nov 22;12(11):e067229. [doi: 10.1136/bmjopen-2022-067229] [Medline: 36414295]
57. Isabirye A. Individual and intimate-partner factors associated with cervical cancer screening in Central Uganda. PLoS One 2022;17(9):e0274602. [doi: 10.1371/journal.pone.0274602] [Medline: 36108074]
58. Kim S, Lee SY, Choi-Kwon S. Cervical cancer screening and human papillomavirus vaccination among Korean sexual minority women by sex of their sexual partners. Int J Environ Res Public Health 2020 Nov 30;17(23):8924. [doi: 10.3390/ijerph17238924] [Medline: 33266281]
59. Teame H, Gebremariam L, Kahsay T, Berhe K, Gebreheat G, Gebremariam G. Factors affecting utilization of cervical cancer screening services among women attending public hospitals in Tigray region, Ethiopia, 2018; case control study. PLoS One 2019;14(3):e0213546. [doi: 10.1371/journal.pone.0213546] [Medline: 30870497]
60. Urquhart G, Maclennan SJ, Guntupalli AM. Is there an association between intimate partner violence and the prevalence of cervical cancer screening in Jordan? PLoS One 2023;18(8):e0290678. [doi: 10.1371/journal.pone.0290678] [Medline: 37651440]
61. CervixCheck information for healthcare providers. CancerCare Manitoba. URL: https://www.cancercare.mb.ca/screening/hcp/cervix [accessed 2025-09-04]
62. Cartwright K, Kosich M, Gonya M, et al. Cervical cancer knowledge and screening patterns in Zuni Pueblo Women in the Southwest United States. J Cancer Educ 2023 Oct;38(5):1531-1538. [doi: 10.1007/s13187-023-02295-8] [Medline: 37046142]
63. Covaliu BF, Forray AI, Tomic M, et al. Understanding cervical cancer screening attendance: barriers and facilitators in a representative population survey. Cancers (Basel) 2025 Feb 19;17(4):706. [doi: 10.3390/cancers17040706] [Medline: 40002299]
64. Fang CY, Ma GX, Handorf EA, et al. Addressing multilevel barriers to cervical cancer screening in Korean American women: a randomized trial of a community-based intervention. Cancer 2017 May 15;123(6):1018-1026. [doi: 10.1002/cncr.30391] [Medline: 27869293]
65. Nakisige C, Trawin J, Mitchell-Foster S, et al. Integrated cervical cancer screening in Mayuge District Uganda (ASPIRE Mayuge): a pragmatic sequential cluster randomized trial protocol. BMC Public Health 2020 Jan 31;20(1):142. [doi: 10.1186/s12889-020-8216-9] [Medline: 32005202]
66. Williams MS, Kenu E, Adanu A, et al. Awareness and beliefs about cervical cancer, the HPV vaccine, and cervical cancer screening among Ghanaian women with diverse education levels. J Cancer Educ 2019 Oct;34(5):897-903. [doi: 10.1007/s13187-018-1392-y] [Medline: 29974412]
67. Rosyda R, Santoso B, Yunitasari E. Is an educational level affect women's participation on cervical cancer screening? - a systematic review. Presented at: The 9th International Nursing Conference; Apr 7-8, 2018; Surabaya East Java, Indonesia. [doi: 10.5220/0008327204820488]
68. Adzigbli LA, Aboagye RG, Adeleye K, Osborne A, Ahinkorah BO. Cervical cancer screening uptake and its predictors among women aged 30-49 in Ghana: providing evidence to support the World Health Organization's cervical cancer elimination initiative. BMC Infect Dis 2025 Feb 21;25(1):246. [doi: 10.1186/s12879-025-10485-6] [Medline: 39984848]
69. Paneru B, Karmacharya A, Bharati A, et al. Association between cancer stigma and cervical cancer screening uptake among women of Dhulikhel and Banepa, Nepal. PLoS ONE 2023;18(5):e0285771. [doi: 10.1371/journal.pone.0285771] [Medline: 37200234]
70. Zhang M, Zhong Y, Zhao Z, et al. Cervical cancer screening rates among Chinese women - China, 2015. China CDC Wkly 2020 Jun 26;2(26):481-486. [doi: 10.46234/ccdcw2020.128] [Medline: 34594684]
71. Zhetpisbayeva I, Kassymbekova F, Sarmuldayeva S, Semenova Y, Glushkova N. Cervical cancer prevention in rural areas. Ann Glob Health 2023;89(1):75. [doi: 10.5334/aogh.4133] [Medline: 37928103]
72. Measuredhs. URL: https://www.measuredhs.com/ [accessed 2025-08-19]

Abbreviations

DHS: Demographic and Health Survey
LR: logistic regression
RF: random forest
SHAP: Shapley Additive Explanations

XSL•FO
RenderX

STI: sexually transmitted infection
WHO: World Health Organization

Edited by JT Lee; submitted 23.01.25; peer-reviewed by A Gebeyew, L Raymond Guo; revised version received 28.05.25; accepted 06.06.25; published 17.09.25.

Please cite as:
Baykemagn ND, Aweke MN, Mesfin A, Baffa LD, Agimas MC, Abuhay HW, Adugna DG, Alemu TG, Bicha AT, Getu Alemu G Identifying Predictors of Cervical Cancer Screening Uptake in Sub-Saharan Africa Using Machine Learning: Cross-Sectional Study JMIR Public Health Surveill 2025;11:e71677
URL: https://publichealth.jmir.org/2025/1/e71677
doi:10.2196/71677

© Nebebe Demis Baykemagn, Mekuriaw Nibret Aweke, Amare Mesfin, Lemlem Daniel Baffa, Muluken Chanie Agimas, Habtamu Wagnew Abuhay, Dagnew Getnet Adugna, Tewodros Getaneh Alemu, Alemu Teshale Bicha, Gebrie Getu Alemu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Dietary Risk Factors and Associated Disease Burden Among Chinese Adults Aged 25 Years and Older: Systematic Analysis of the Global Burden of Disease Study 2021

Yuze Xin[1*], BMed; Dong Shui[1*], BMed; Guangcan Yan[1], BMed; Wei Tian[1,2], PhD; Ning Tang[1], MPH; Jinyu Liang[1], BMed; Junyi Peng[1], BMed; Hongru Sun[1], PhD; Anqi Ge[1], PhD; Xinyan Liu[1], PhD; Katrina Kissock[3,4], PhD; Kathy Trieu[3,4], PhD; Jing Zhang[1], PhD; Xinyi Zhang[1], PhD; Pengpeng Ye[5], PhD; Maoyi Tian[1,3,6], PhD

[1]School of Public Health, Harbin Medical University, 157 Baojian Road, Nangang District, Harbin, China

[2]School of Basic Medicine, Harbin Medical University, Harbin, China

[3]The George Institute for Global Health, Faculty of Medicine and Health, University of New South Wales, New South Wales, Australia

[4]School of Population Health, Faculty of Medicine and Health, University of New South Wales, New South Wales, Australia

[5]National Center for Chronic and Non-communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

[6]Department of General Practice, The Second Affiliated Hospital of Harbin Medical University, Harbin, China

[*]these authors contributed equally

Corresponding Author:
Maoyi Tian, PhD
School of Public Health, Harbin Medical University, 157 Baojian Road, Nangang District, Harbin, China

Abstract

Background: With rapid economic development and lifestyle changes, diet-related diseases have become a major public health concern globally. China is experiencing significant dietary transitions. From 2001 to 2021, the intake of staple foods declined, while the consumption of animal-based foods and ultraprocessed foods increased significantly. But comprehensive assessments of major dietary risk factors and the long-term health impacts of shifting dietary patterns in China remain unclear.

Objective: This study aims to assess the disease burden attributable to dietary risk factors in China using data from the Global Burden of Diseases, Injuries, and Risk Factors Study 2021 (GBD 2021), and to examine long-term trends over the past 3 decades. In addition, it provides an in-depth analysis of the 3 major diet-related disease categories in China: cardiovascular diseases, neoplasms, and diabetes and kidney diseases.

Methods: We extracted data from GBD 2021, focusing on diet-related health outcomes in China across 33 provinces and regions. Measures included deaths, years of life lost (YLLs), years lived with disability (YLDs), and disability-adjusted life years (DALYs), stratified by age, sex, and region. Age-standardized rates (ASRs) were calculated, and temporal trends were analyzed using estimated annual percentage change (EAPC).

Results: In 2021, dietary risk factors accounted for 1.70 million deaths and 38.39 million DALYs among Chinese adults aged 25 years and older. The leading contributors were high sodium intake, low fruit consumption, and low whole grain intake. Cardiovascular diseases were the largest contributors to diet-related DALYs. The burden was more pronounced in males than in females and highest among older adults aged 80 years and older. Substantial regional variation was observed, with the Northeastern and Western regions showing higher burden. From 1990 to 2021, overall disease burden due to dietary risks declined steadily, as reflected by decreasing ASR-DALYs (EAPC= -1.76), YLLs, and death rates. In contrast, YLDs showed a slight upward trend (EAPC=0.75), indicating a shift toward increased years lived with disability. In addition, the relative contributions of specific dietary risk factors changed considerably. Low vegetable intake, once ranked the third in 1990, dropped to the 12th by 2021, while high red meat consumption rose from the 15th to 7th place. Although the ranking of high-sugar beverage consumption did not change, the ASR-DALYs rate increased significantly, with a percentage change of 689.14% between 1990 and 2021.

Conclusions: In China, the burden of diet-related diseases remains substantial. While the overall age-standardized disease burden has declined, marked regional and demographic disparities persist. Certain dietary risks, such as high red meat and sugar-sweetened beverage consumption, are rising, and high sodium intake remains a serious concern. These trends highlight the urgent need for comprehensive, adaptable, and evidence-based nutrition policies to be implemented to address the evolving diet-related disease burden across diverse populations in China.

(JMIR Public Health Surveill 2025;11:e72978) doi:10.2196/72978

XSL•FO
RenderX

KEYWORDS

global burden of disease; dietary risk factors; diet-related diseases; ultraprocessed food; China

Introduction

Noncommunicable diseases (NCDs) are the leading causes of death globally, accounting for approximately 74% of all deaths worldwide [1]. Among these, cardiovascular diseases, neoplasms, diabetes, and kidney diseases represent the most prominent contributors to both mortality and disability [2,3]. The role of dietary risks in the development of NCDs has been extensively documented through long-term prospective studies and randomized controlled trials [4]. Unhealthy dietary patterns, characterized by high intake of sodium, trans fats, sugar-sweetened beverages, and red or processed meats, as well as insufficient consumption of fruits, vegetables, fiber, and whole grains, have emerged as the most critical and modifiable risk factors for NCDs and premature death [5-7]. As dietary risk factors have imposed an increasing health and economic burden worldwide, improving diet quality has consequently become a central focus of global health efforts to reduce the burden of NCDs [8].

Socioeconomic transformations have driven substantial changes in the national dietary landscape over the past few decades in China [9,10]. For instance, between 2000 and 2021, the consumption of traditional staple foods declined by 36.9%, while the intake of animal-based products, such as meat, eggs, milk, and fish and seafood, increased by 44.7%, 39.9%, 406.4%, and 67.6%, respectively [11]. Concurrently, the expansion of the food industry has contributed to the widespread availability and growing consumption of ultraprocessed foods, which are often high in salt, sugar, and unhealthy fats [12,13]. Moreover, the transition has not occurred uniformly across the country. Significant regional disparities in dietary patterns persist due to variations in economic development, food availability, cultural preferences, and local agricultural practices [14]. These differences pose substantial challenges for developing and implementing unified national dietary guidelines and nutrition interventions that are both culturally appropriate and context-specific.

In response to growing concerns over population-level nutrition and diet-related health risks, China has implemented a series of comprehensive measures aimed at improving dietary quality nationwide [15]. One of the cornerstones of these efforts is the development and periodic revision of the Chinese Dietary Guidelines, which were first published in 1989 and subsequently updated in 1997, 2007, 2016 [16]. The most recent version (the 5th edition) was released in 2022 [17]. These guidelines provide recommendations for nutrient intake for the general population and specific population groups, such as pregnant women or breastfeeding mothers, infants, children, and adolescents [18,19]. Complementing these technical guidelines, the Chinese government launched the Healthy China Action Plan (2019 - 2030), which identifies “healthy diet promotion” as one of its key strategic initiatives [20]. This plan underscores a multisectoral approach that calls for joint efforts from government bodies, health professionals, schools, and the food industry to improve dietary behaviors and create supportive environments for nutrition across all life stages [20-22]. However, despite these policy advancements and public health campaigns, significant challenges remain [23].

Analyzing the disease burden associated with dietary risk factors in China is important to inform the development of targeted and context-specific nutrition policies [24,25]. Although previous studies have explored the relationship between individual dietary risk factors and specific diseases, or between overall dietary patterns and single disease categories, such as high salt intake and cardiovascular disease, or the association between dietary patterns and kidney disease [26,27]. A comprehensive assessment of the cumulative burden of all major dietary risks and their long-term health consequences remains limited. This has resulted in a gap in understanding the broader and long-term impact of dietary risks on population health in China. To address this evidence gap, we used data from the Global Burden of Diseases, Injuries, and Risk Factors Study 2021 (GBD 2021) to conduct a systematic assessment of the disease burden attributable to dietary risk factors in China. This study also analyzed temporal trends over the past 3 decades and focused on the 3 leading disease categories contributing to diet-related disease burden in the country: cardiovascular diseases, neoplasms, and diabetes and kidney diseases [25,28,29].

Methods

Overview

Using data from the GBD 2021, the study provided estimates of disease burden related to dietary risks in China, by sex and age, between 1990 and 2021. A detailed description of the methods for estimating the disease burden attributable to risk factors has been published elsewhere [4]. We summarize the methods for estimating disease burden related to dietary risks below.

Briefly, this study analyzed data on dietary risk factors associated with all diseases among adults aged 25 years and older across 33 provinces in China from 1990 to 2021. These include all 31 mainland provinces, autonomous regions, municipalities, and two special administrative regions, Hong Kong and Macao (Taiwan Province was excluded due to data unavailability). Data were segmented by age, year, sex, geographical region, and different dietary risk factors [30]. Data were segmented by age, year, sex, geographical region, and different dietary risk factors.

Data Source and Definitions

GBD 2021 consisted of the most recent epidemiological data and had improved standardized methodologies compared to previous GBD data [2]. In this study, data classification conformed to the GBD database classification standards, encompassing age, region, and disease categories. This study not only analyzed the all-cause disease burden caused by dietary risk factors but also separately analyzed the 3 major disease categories: cardiovascular diseases, neoplasms, and diabetes

XSL•FO
RenderX

and kidney diseases. All 3 disease categories are classified as Level 2 causes within the GBD classification system [2].

Dietary risk factors were behavioral factors that were correlated with the increased or decreased risk of developing diseases [2]. The selection of dietary risk factors includes 15 specific dietary risks that meet GBD criteria for risk factor selection. These criteria considered the significance of the risk factors in contributing to disease burden, as well as the availability of adequate data to estimate exposure to these risks. The selection process was further guided by the strength and consistency of epidemiological evidence supporting a causal relationship between dietary factors and disease. The method used to evaluate the strength of this evidence for causality is detailed elsewhere [4] and summarized in Table S3 in Multimedia Appendix 1.

Measurement of Disease Burden

GBD 2021 quantified the proportion of disease burden attributable to each dietary factor that could be prevented if exposure levels were maintained at the minimum risk level, defined as the theoretical minimum risk exposure level (TMREL) [31]. Assuming exposure levels for other risk factors remain constant, the population attributable fraction (PAF) for a given risk factor was estimated by comparing the TMREL with the exposure level in a specific population [32]. GBD 2021 applied a comparative risk assessment framework to calculate the disease burden attributable to major dietary risk factors.

Statistical Analysis

This study used various epidemiological indicators, including the number of deaths, mortality rate, years of life lost (YLLs), years lived with disability (YLDs), and disability-adjusted life years (DALYs) to describe the disease burden attributable to dietary factors. When comparing populations across different regions or age groups over time, we used age-standardized rates (ASRs), calculated as a weighted average of age-specific rates. All measures were reported as unstandardized and ASRs (rate per 100,000 persons), including their 95% uncertainty intervals (UIs), representing the range within which the true value is expected to lie with 95% probability. To explore the trends of ASRs from 1990 to 2021, we calculated the estimated annual percentage change (EAPC) of ASRs based on the formula $EAPC=100\times(\exp(\beta)-1)$, where $Y=(\alpha+\beta X+\varepsilon)=\ln(ASR)$, X=calendar y, and ε=the error term. This model is based on the assumption that a natural logarithmic scale will show a linear trend of ASR over a specified time. The ASR is considered to exhibit an increasing trend when EAPC and the lower bound of the UI are positive, whereas the ASR is considered to exhibit a decreasing trend when EAPC and the upper bound of the UI are negative. All data were analyzed using RStudio (Version 2024.04.0).

Ethical Considerations

This study used data from the GBD study, which was approved by the institutional review board of the University of Washington School of Medicine. The original data collection obtained informed consent from study participants or was granted exemptions by the institutional review board. As this is a secondary analysis of existing data, no additional human participant research ethics review or informed consent was required. Study data were anonymized and deidentified to protect the privacy and confidentiality of study participants. Our study adhered to the GBD protocol and followed the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) [33]. The GATHER checklist is provided in Table S1 in Multimedia Appendix 1.

Results

Overall Disease Burden Attributable to Dietary Risk Factors in 2021

In 2021, 1.70million deaths (95% UI 0.69-2.68) from all diseases were attributed to dietary risk factors among people aged 25 years and older in China. The YLLs due to dietary factors amounted to 34.18million person-years (95% UI 14.57-52.84). The YLDs were 4.21million person-years (95% UI 1.57-6.73), and the DALYs totaled 38.39million person-years (95% UI 16.21-58.61). YLLs accounted for 89% of DALYs. The ASR-YLDs were 204.17 per 100,000 persons (95% UI 75.67-326.07), and the ASR-DALYs were 1892.69 per 100,000 persons (95% UI 775.77-2904.23) (Table 1).

XSL•FO
RenderX

Table . Overall and cause-specific disease burden attributable to dietary risk factors in China, 1990 - 2021: deaths, years of life lost, years lived with disability, disability-adjusted life years (95% uncertainty intervals), and estimated annual percentage change.

Burden of disease indicators	Absolute number, thousands		Age-standardized rate, per 100,000 people		EAPC[a] of ASR[bc]
	1990	2021	1990	2021	1990 - 2021
Deaths (95% UI)[d]					
All causes	1011.13 (518.27-1474.48)	1704.02 (687.92-2684.78)	152.13 (76.40-222.36)	90.37 (35.35-143.15)	−1.68 (-2.49 to −1.42)
Cardiovascular diseases	829.87 (471.40-1149.29)	1449.88 (594.79-2244.04)	128.62 (69.70-181.22)	77.76 (30.45-121.22)	−1.62 (−2.67 to −1.30)
Diabetes and kidney diseases	31.94 (16.00-48.16)	76.77 (32.01-121.77)	4.77 (2.41-7.06)	3.91 (1.65-6.18)	−0.64 (−1.22 to −0.43)
Neoplasms	147.02 (29.72-306.47)	176.53 (60.80-367.46)	18.44 (3.89-38.01)	8.66 (2.98-17.89)	−2.43 (−2.41 to −0.86)
YLLs[e] (95% UI)					
All causes	25261.26 (12317.56-36672.75)	34186.08 (14575.66-52839.91)	3103.51 (1545.93-4508.14)	1688.52 (689.96-2601.5)	−1.96 (−2.60 to −1.77)
Cardiovascular diseases	20213.91 (11357.25-27664.76)	28353.16 (12271.3-42425.37)	2536.16 (1418.19-3489.77)	1409.55 (584.61-2114.00)	−1.89 (−2.86 to −1.62)
Diabetes and kidney diseases	793.65 (397.55-1209.75)	1581.38 (654.40-2528.97)	97.41 (49.20-146.73)	75.81 (31.53-120.59)	−0.81 (−1.44 to −0.63)
Neoplasms	4189.84 (855.34-8742.78)	4229.48 (1429.36-8797.48)	462.79 (95.46-964.05)	202.08 (68.33-417.31)	−2.67 (−2.70 to −1.08)
YLDs[f] (95% UI)					
All causes	1440.14 (648.54-2200.15)	4207.73 (1567.38-6728.75)	161.77 (72.93-250.42)	204.17 (75.67-326.07)	0.75 (0.12-0.85)
Cardiovascular diseases	824.48 (425.83-1258.08)	1913 (786.69-3080.93)	94.68 (47.86-145.38)	90.03 (36.97-145.22)	−0.16 (−0.83 to 0.01)
Diabetes and kidney diseases	550.76 (125.63-981.24)	2109.35 (403.13-3857.85)	59.72 (14.66-105.88)	105.35 (20.53-193.47)	1.83 (1.09-1.94)
Neoplasms	58.94 (15.08-114.05)	177.77 (47.72-340.07)	6.70 (1.77-12.80)	8.41 (2.25-16.11)	0.74 (0.73-0.77)
DALYs[g] (95% UI)					
All causes	26701.40 (13025.80-38381.93)	38393.81 (16213.07-58607.05)	3265.28 (1625.52-4693.91)	1892.69 (775.77-2904.23)	−1.76 (−2.39 to −1.55)
Cardiovascular diseases	21038.39 (11853.04-28797.03)	30266.17 (13228.42-45225.09)	2630.84 (1471.93-3614.80)	1499.58 (632.89-2247.22)	−1.81 (−2.72 to −1.53)
Diabetes and kidney diseases	1344.40 (546.84-2147.75)	3690.73 (1021.61-6162.48)	157.13 (66.38-247.18)	181.17 (50.05-301.66)	0.46 (−0.91 to 0.64)
Neoplasms	4248.78 (868.87-8847.49)	4407.25 (1479.11-9106.09)	469.49 (97.07-975.92)	210.49 (70.5-431.87)	−2.59 (−2.63 to −1.03)

[a]EAPC: estimated annual percentage change.

[b]ASR: age-standardized rate.

[c]EAPC of ASR: The EAPC of ASR to explore the trends of ASRs from 1990 to 2021.

[d]UI: uncertainty interval.

[e]YLL: year of life lost.

[f]YLD: year lived with disability.

[g]DALY: disability-adjusted life year.

XSL•FO
RenderX

Regional Distribution of Disease Burden Attributable to Dietary Risk Factors

There are significant regional differences in the disease burden attributed to dietary factors in China. Fujian, Shanghai, Macao, and Hong Kong experience a lower diet-related disease burden, while Hebei, Heilongjiang, Jilin, Inner Mongolia, and Qinghai have a higher diet-related disease burden. Figure 1 displays the distribution of age-standardized death rate (ASDR) and ASR-DALY attributed to dietary factors for all diseases. The contribution of individual dietary factors to the overall diet-related disease burden varies across regions. Figure 2 ranks the different dietary factors contributing to the disease burden in various regions of China, and the high-sodium diet emerges as the top risk factor for disease burden in all regions. A diet high in sodium and low in whole grains and fruit is the leading dietary risk factor in almost every region. In Shanghai and Beijing, high red meat consumption ranks third in contributing to the burden of diet-related diseases, while low fruit intake ranks fourth. In Macao and Hong Kong, low dietary fiber intake ranks third, with low whole-grain intake and low fruit intake ranking fourth, respectively.

Figure 1. Regional distribution of age-standardized death rates and age-standardized disability-adjusted life year rates attributable to dietary risk factors in China, 2021. ASDR: age-standardized death rates; ASR-DALY: age-standardized disability-adjusted life year rates;

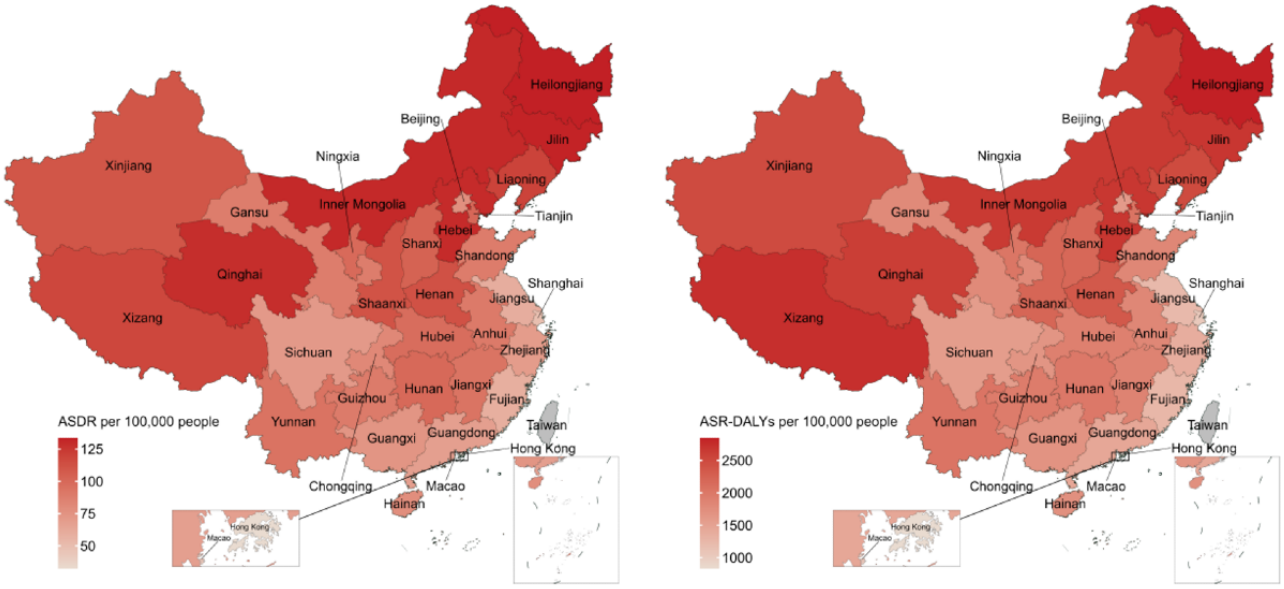


Figure 2. Provincial ranking of age-standardized disability-adjusted life year rates attributable to dietary risk factors in China, 2021.

	China	Anhui	Beijing	Chongqing	Fujian	Gansu	Guangdong	Guangxi	Guizhou	Hainan	Hebei	Heilongjiang	Henan	Hong Kong	Hubei	Hunan	Inner Mongolia	Jiangsu	Jiangxi	Jilin	Liaoning	Macao	Ningxia	Qinghai	Shaanxi	Shandong	Shanghai	Shanxi	Sichuan	Tianjin	Xinjiang	Xizang	Yunnan	Zhejiang
High sodium	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Low whole grains	2	2	2	2	2	3	2	2	3	2	3	2	2	2	3	3	2	2	3	2	2	4	2	3	2	2	2	2	3	2	3	3	3	2
Low fruits	3	3	4	3	3	2	3	3	2	3	2	3	3	4	2	2	3	3	2	3	3	2	3	2	3	3	4	3	2	4	2	2	2	3
Low polyunsaturated fat	4	4	5	4	5	4	4	4	5	4	4	4	4	10	4	4	4	5	4	4	4	6	4	4	4	4	6	4	4	3	4	5	4	6
Low nuts and seeds	5	5	6	5	7	5	6	5	6	6	5	5	5	13	5	5	5	8	5	5	5	9	5	5	5	5	8	5	6	6	5	6	6	8
Low fiber	6	6	9	6	8	6	8	6	4	7	6	6	7	3	6	6	6	6	6	7	7	3	7	6	7	7	9	7	5	8	6	4	5	5
High red meat	7	8	3	9	4	14	5	7	15	5	10	8	9	5	10	7	7	4	11	8	6	5	8	15	9	6	3	8	12	5	7	15	13	4
Low seafood omega-3 fat	8	7	10	8	12	7	11	8	7	9	7	7	6	14	7	8	9	10	7	6	9	12	6	7	6	9	13	6	7	12	8	8	7	13
Low legumes	9	11	11	11	10	9	10	10	9	10	9	9	8	8	9	9	8	11	10	9	8	10	9	8	8	8	12	9	10	7	9	9	9	11
Low milk	10	10	7	7	6	10	7	9	8	8	11	10	11	7	8	10	10	7	9	10	10	7	10	9	10	10	5	11	8	10	12	10	10	7
High processed meat	11	12	8	10	9	11	9	12	11	11	12	11	12	6	11	12	11	9	12	11	11	11	11	11	12	11	7	12	11	9	11	11	11	9
Low vegetables	12	9	14	12	13	8	14	11	10	12	8	12	10	9	12	11	12	14	8	12	13	8	12	10	11	12	14	10	9	15	10	7	8	12
Low calcium	13	13	13	13	11	12	12	13	12	13	13	13	13	12	13	13	13	13	13	13	12	13	13	12	13	14	10	13	13	13	13	12	12	10
High sugar beverages	14	14	12	14	14	13	13	14	13	14	14	14	14	11	14	14	14	12	14	14	14	14	14	13	14	13	11	14	14	11	14	13	14	14
High trans fat	15	15	15	15	15	15	15	15	14	15	15	15	15	15	15	15	15	15	15	15	15	15	15	14	15	15	15	15	15	14	15	14	15	15

Dietary Risk Factors Associated With Individual Disease Burden in 2021

Cardiovascular diseases attributable to dietary risk factors were the leading cause of death and DALYs amongst disease groups, with 1.45million deaths (95% UI 0.59-2.24) and 30.26million person-years of DALYs (95% UI 13.23-45.23). The ASR-DALYs rate was 1499.58 per 100,000 persons (95% UI 632.89-2247.22). Meanwhile, diabetes and kidney diseases are attributed to dietary risk factors, with 0.08million deaths (95%

XSL•FO
RenderX

UI 0.03-0.12) and 3.69million person-years of DALYs (95% UI 1.02-6.16). The ASR-DALYs rate was 181.17 per 100,000 persons (95% UI 50.05-301.66). Neoplasms attributed to dietary risk factors lead to 0.18million deaths (95% UI 0.06-0.37) and 4.41million person-years of DALYs (95% UI 1.47-9.11). The ASR-DALYs rate was 210.49 per 100,000 persons (95% UI 70.50,431.87).

Sex and Age Distribution of Disease Burden Attributable to Dietary Risk Factors

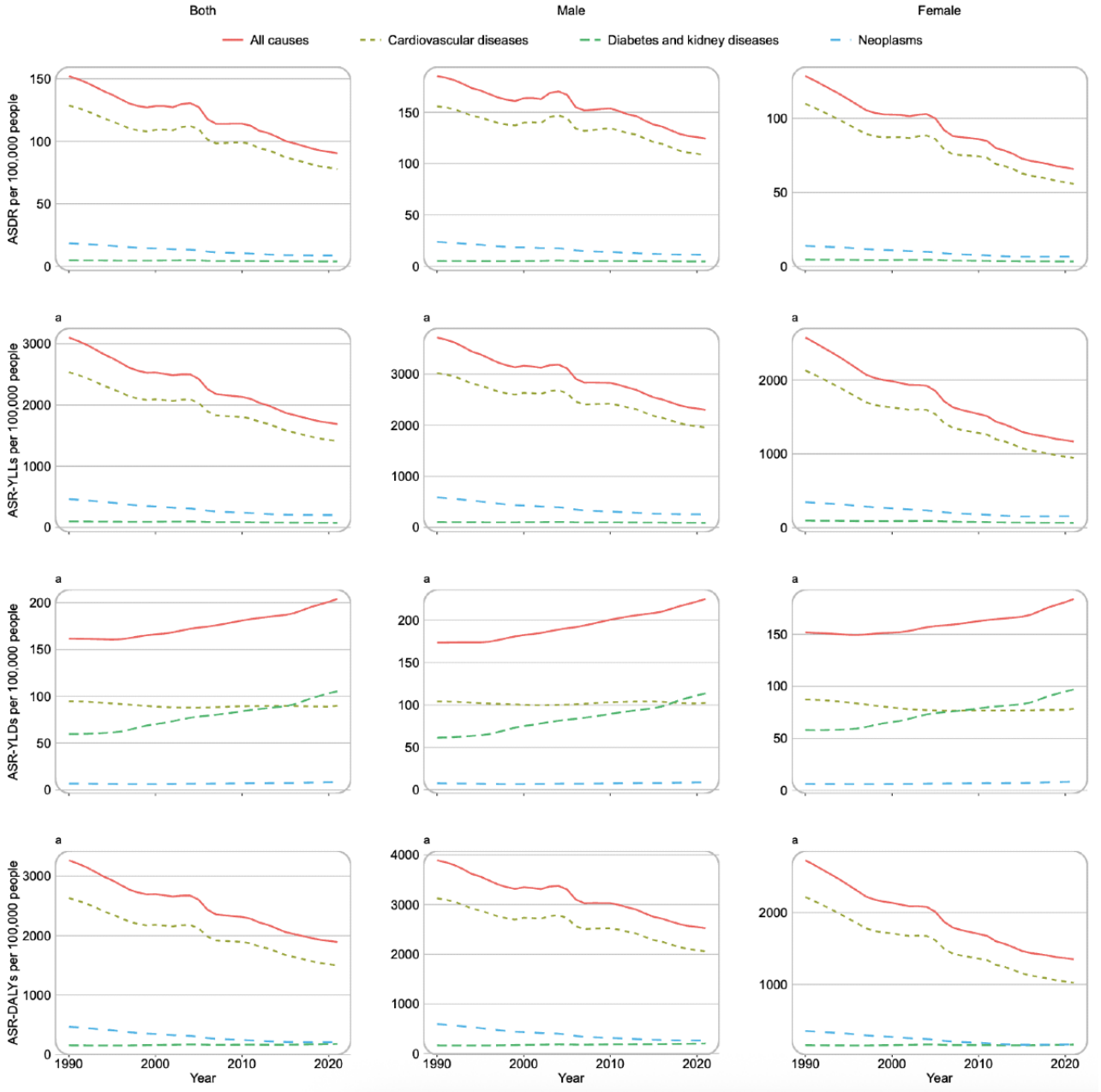
The death rate for males due to dietary risk factors was 124.40 per 100,000 persons (95% UI 49.56-199.28), 1.89 times higher than females (65.73 per 100,000 persons, 95% UI 23.36-110.28). The YLDs rate for males was 225.21 per 100,000 persons (95% UI 86.13-355.83), and the DALYs rate was 2524.01 per 100,000 persons (95% UI 1034.57-3878.49). Whereas for females, the YLDs rate and DALYs rate were 183.97 per 100,000 persons (95% UI 64.49-299.26) and 1350.20 per 100,000 persons (95% UI 509.30-2179.77), respectively. The proportion of DALYs due to YLLs from dietary risk factors was 91.1% for males and 86.4% for females, with males having a slightly higher proportion compared to females. The disease burden attributable to dietary risk factors increases with age, particularly among individuals aged 60 years and older. Death rate, YLL rate, and DALY rate are the highest in the 80 years and older age group, while YLD rate was highest in the 75 - 79 age group (1038.88 per 100,000 persons for the 75 - 79 years age group compared with 964.02 per 100,000 persons for the 80 years and older age group). The proportion of YLLs to DALYs gradually increased with age, reaching 95.4% in the 80 years and older age group. Detailed data are presented in Tables S5-6 in Multimedia Appendix 1. The figures illustrating the trends are presented in Figures S1–5 in Multimedia Appendix 2.

Trends of Disease Burden Attributed to Dietary Risk Factors From 1990 to 2021

Figure 3 illustrates the temporal changes in the disease burden caused by dietary risk factors amongst Chinese adults aged 25 years and older from 1990 to 2021. The ASDR, ASR-YLLs, and ASR-DALYs for all diseases show a significant downward trend over time, with an EAPC of −1.76 (95% UI −2.39 to −1.55) for ASR-DALY. In contrast, ASR-YLDs exhibit an upward trend, with an EAPC of 0.75 (95% UI 0.12-0.85). Figure 4 illustrates the changes in the proportion of disease burden caused by different dietary risk factors. Over the past 30 years, there have been notable changes in dietary risk factors contributing to the diet-related disease burden. In 1990, low vegetable intake ranked third in its contribution but dropped to 12th place by 2021. In contrast, high red meat consumption rose from the lowest rank (15th) to seventh place. Although the ranking of sugar-sweetened beverages (SSBs) consumption as a dietary risk remained unchanged, the age-standardized DALY rate increased substantially, with a percentage change of 689.14% (95% UI 296.03%-512.16%) from 1990 to 2021.

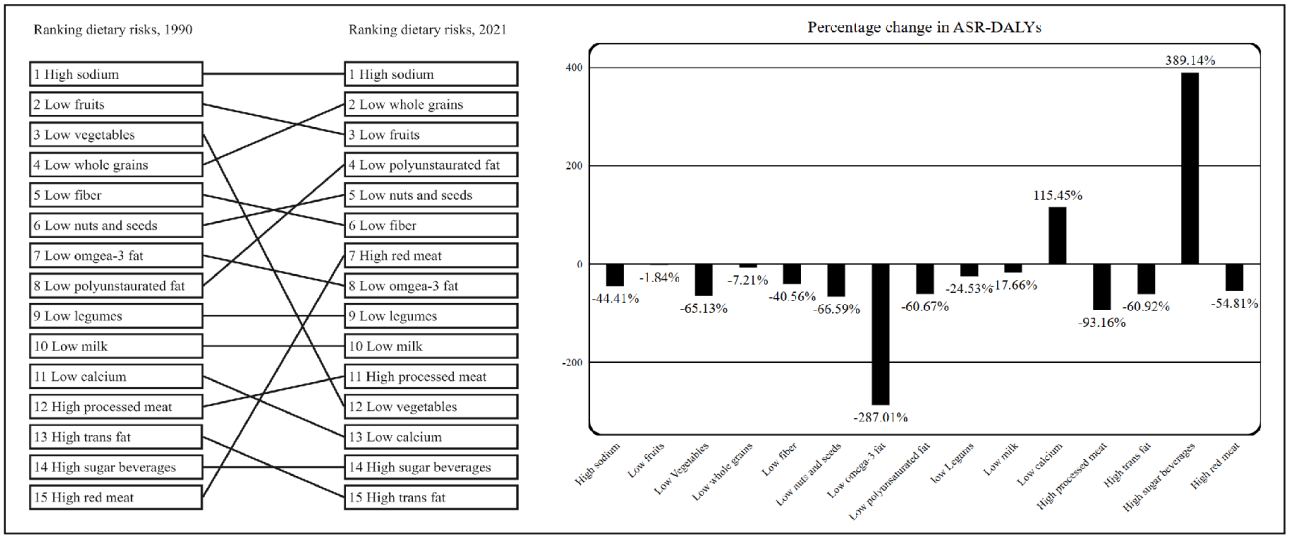
XSL•FO
RenderX

Figure 3. Temporal changes in disease burden (deaths, years of life lost, years lived with disability, and disability-adjusted life years) attributable to dietary risk factors in China, 1990 - 2021. ASR-DALY: age-standardized rate-disability-adjusted life year; ASR-YLD: age-standardized rate-year lived with disability; ASR-YLL: age-standardized rate-year of life lost.



XSL•FO
RenderX

Figure 4. Changes in the ranking of dietary risk factors contributing to the proportion of disease burden (Age-standardized rates -disability-adjusted life years) in China, 1990 - 2021.



Discussion

Principal Findings

In 2021, the dietary factors contributing to disease burden in China were more severe for males and older age groups. Cardiovascular diseases remain the primary contributors to the burden attributed to dietary risk factors. Key dietary contributors to disease burden include high sodium intake, low fruit consumption, and low whole-grain intake. The burden of disease, as measured by ASR-DALYs, is remarkably higher in certain areas of Northern and Western China compared with other regions.

The Chinese government has been striving to enhance the dietary health of its population [34], and since 1989, the Chinese Residents' Dietary Guidelines have been issued and revised multiple times [20]. These initiatives have demonstrated effectiveness, as evidenced by improvements in disease burden over time. This progress is closely linked to China's rapid economic expansion in recent years, resulting in greater dietary variety among the population and a relatively heightened consumption of specific food items, both healthful and unhealthful in nature [35,36]. One significant transformation is observed in vegetable consumption, which has transitioned in disease burden rankings from third place in 1990 to twelfth place. The discrepancy in the increasing trajectory of ASR-YLDs attributed to dietary factors, in contrast to other metrics, may be explained by advancements in medical treatment that prolong patient survival [37]. Furthermore, the rise in the prevalence of chronic diseases stemming from overnutrition could also play a role in this situation [38].

Although the overall disease burden from dietary factors has declined over time, the burden attributable to certain specific dietary risks, such as diets high in red meat and low in whole grains, has increased. In most provinces in China, especially in inland cities, red meat is still the main meat consumed [39,40], and economic development has also stimulated an increase in food variety, especially in meat consumption [41]. According to data from the China Statistical Yearbook, from 1990 to 2021, China's red meat consumption showed a continuous upward trend except for a slight decline in 2019 [42,43]. In addition, the increased proportion of disease burden attributable to a diet low in whole grains may also be associated with economic development, which has led to a shift towards more refined grains in the diet. In economically underdeveloped periods, whole grains were more readily available than refined grains. As a result, higher whole-grain intake was often associated with increased poverty among the Chinese population [44]. This consciousness often results in reduced consumption of coarse grains and other whole-grain foods, exacerbating health risks [45].

Another change is the rise in the disease burden caused by a diet high in SSBs. The consumption of SSBs is rapidly increasing worldwide, especially among adolescents [46]. The rise in SSB intake is associated with various health issues, such as obesity, insulin resistance, and dental caries [47]. This has become a significant public health concern. In China, the intake of SSBs is also particularly concerning among young children, highlighting the need for early intervention [48]. To minimize the consumption of SSBs, the World Health Organization (WHO) released a manual in 2022 recommending taxation as an effective strategy to reduce SSB intake [49]. China may also adopt such measures in the future to curb SSB consumption [50]. The above findings suggest that as socioeconomic status improves, the nature of diet-related disease burden shifts. Instead of being driven by the lack of specific nutrients, it increasingly stems from imbalances in the overall dietary structure [51,52]. However, achieving this shift in dietary patterns is challenging for China, where economic growth is just beginning [53].

It is also noteworthy that from 1990 to 2021, high sodium intake and low fruit consumption consistently ranked among the leading dietary risk factors contributing to disease burden. Studies indicate that the average fruit intake among the Chinese population is around 64.30g per day [54], below both the recommendations of the Chinese Dietary Guidelines (daily intake for adults aged 18 - 64 years is 200 - 350 g of fruits) and those of the WHO (recommends a minimum combined daily intake of 400 g of fruits and vegetables) [17,55]. This can

XSL•FO
RenderX

be attributed to factors such as low economic status, poor dietary literacy, and a lack of fruit consumption habits among residents [56]. Addressing these issues requires further efforts and plans at the national level, along with more frequent and in-depth health education initiatives. Educating individuals about healthy habits from a personal lifestyle perspective is relatively challenging and demands greater resourcing [57]. On the other hand, current sodium intakes of the Chinese population are high at around 4g per day [58], far exceeding the WHO maximum recommended level of <2g per day. Sodium intake in China typically comes from home cooking practices [59], which are deeply rooted in family cooking traditions and regional dietary cultures [60]. A more effective approach to reduce sodium is to promote the substitution of regular salt with potassium-enriched salt [61]. Addressing these factors also requires long-term health education efforts [60].

An analysis of population subgroups revealed variation in the disease burden attributable to dietary risk factors across different demographic characteristics. Our findings revealed notable differences in disease burden attributable to dietary risk factors by sex and age. Specifically, males experienced a substantially higher burden compared to females. This disparity may be partially explained by sex-specific differences in dietary behaviors, with men in China generally consuming higher amounts of salt, red and processed meats, and lower amounts of fruits and vegetables [62,63]. Thus, it potentially leads to a higher prevalence of diet-related NCDs among men [14]. In addition, we observed that the disease burden associated with dietary risks increases with age, particularly among individuals aged 60 years and older. This trend may reflect the cumulative effect of long-term exposure to unhealthy dietary habits, as well as increased physiological vulnerability and comorbidities in older adults [64].

The distribution of disease burden caused by dietary risk factors in China varies significantly between regions. According to the National Bureau of Statistics of China, the country is divided into Eastern, Central, Western, and Northeastern regions [65]. The death rate and DALYs are disproportionately higher in the Western and Northeastern regions, while the Eastern and Central regions exhibit comparatively lower disease burden, particularly the eastern coastal areas, which have the lowest overall disease burden. This can be attributed to both socioeconomic status [66] and dietary habits, such that diets are relatively healthful in Eastern regions such as Shanghai, Guangdong, Jiangsu, and Zhejiang, where coastal cities consume more seafood [67]. In contrast, the Northeastern region has heavier flavor profiles in its diet, which means a higher sodium intake [68]. The varying levels of economic development and dietary habits across regions can lead to an emphasis on differing dietary factors that require specific interventions. For instance, in Beijing and Shanghai, where economic levels are high, the disease burden caused by high red meat consumption is higher than in other areas [69]. These factors can serve as focal points for dietary interventions in different regions.

This study also has certain limitations. First, it lacks data on food intake levels across China, which hinders comprehensive analysis. In addition, although stratified analyses by urban-rural classification and socioeconomic status could provide more granular insights, such analyses were not feasible due to the lack of data in the GBD 2021 dataset. Future studies incorporating primary data collection or additional contextual datasets may help address these gaps. However, this study provides an analysis of the burden of diet-related diseases amongst Chinese adults aged 25 years and older, with a particular focus on regional variations and the burden attributed to specific dietary risk factors. This evidence can better inform the development of targeted dietary interventions.

Conclusion

In conclusion, in China, the burden of diseases related to diet remains significant. With rapid economic growth and shifting dietary patterns, the disease burden caused by dietary risk factors should receive greater attention. In response, tailored and impactful nutrition policies and strategies that address diet-related disease burdens in China need to be developed and implemented.

Acknowledgments

We appreciate the work done by the 2021 Global Burden of Diseases study collaborations.

Data Availability

All data generated or analyzed during this study are included in this published article.

Authors' Contributions

MT and PY were responsible for the conceptualization and methodology of the study. DS conducted the formal analysis. YX drafted the original manuscript. MT, PY, NT, JL, JP, GY, WT, HS, AG, XL, KK, KT, JZ, and XZ contributed to the review and editing of the manuscript. All authors reviewed and approved the final version of the manuscript.

PY is the co-corresponding author and can be reached at The National Center for Chronic and Noncommunicable Disease Control and Prevention; The Chinese Center for Disease Control and Prevention; 27 Nanwei Road; Beijing, 100050; China; Email: smztsmzt@163.com; Phone: 86 18701355688.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Guidelines for Accurate and Transparent Health Estimates Reporting checklist, disease classifications & levels, dietary risk–disease references, and supplementary tables.
[DOCX File, 231 KB - publichealth_v11i1e72978_app1.docx]

Multimedia Appendix 2
Deaths number and ASR-deaths.
[DOCX File, 2265 KB - publichealth_v11i1e72978_app2.docx]

References

1. The top 10 causes of death. World Health Organization. 2020. URL: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death [accessed 2025-07-24]
2. Murray CJL, Aravkin AY, Zheng P, et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 2020 Oct;396(10258):1223-1249. [doi: 10.1016/S0140-6736(20)30752-2]
3. Dunkler D, Kohl M, Teo KK, et al. Dietary risk factors for incidence or progression of chronic kidney disease in individuals with type 2 diabetes in the European Union. Nephrol Dial Transplant 2015 Aug;30 Suppl 4:iv76-iv85. [doi: 10.1093/ndt/gfv086] [Medline: 26209742]
4. Afshin A, Sur PJ, Fay KA, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 2019 May;393(10184):1958-1972. [doi: 10.1016/S0140-6736(19)30041-8]
5. Sun X, Yon DK, Nguyen TT, et al. Dietary and other lifestyle factors and their influence on non-communicable diseases in the Western Pacific region. Lancet Reg Health West Pac 2024 Feb;43:100842. [doi: 10.1016/j.lanwpc.2023.100842] [Medline: 38456094]
6. Kimokoti RW, Newby PK. Dietary patterns, smoking, and cardiovascular diseases: a complex association. Curr Nutr Rep 2013 Jun;2(2):113-125. [doi: 10.1007/s13668-013-0043-5]
7. Babalola OO, Akinnusi E, Ottu PO, et al. The impact of ultra-processed foods on cardiovascular diseases and cancer: Epidemiological and mechanistic insights. Aspects of Molecular Medicine 2025 Jun;5:100072. [doi: 10.1016/j.amolm.2025.100072]
8. Micha R, Peñalvo JL, Cudhea F, Imamura F, Rehm CD, Mozaffarian D. Association between dietary factors and mortality from heart disease, stroke, and type 2 diabetes in the United States. JAMA 2017 Mar 7;317(9):912-924. [doi: 10.1001/jama.2017.0947] [Medline: 28267855]
9. Zhang J, Wang Z, Du W, et al. Twenty-five-year trends in dietary patterns among Chinese adults from 1991 to 2015. Nutrients 2021;13(4):1327. [doi: 10.3390/nu13041327]
10. Zhai FY, Du SF, Wang ZH, Zhang JG, Du WW, Popkin BM. Dynamics of the Chinese diet and the role of urbanicity, 1991-2011. Obes Rev 2014 Jan;15 Suppl 1(1):16-26. [doi: 10.1111/obr.12124] [Medline: 24341755]
11. Wang R, Xiong F, Cheng G, Wang H, Liu G. Province-specific sustainable diets in China considering nutrition, environment, affordability, and acceptability. J Clean Prod 2024 Aug;468:143141. [doi: 10.1016/j.jclepro.2024.143141]
12. Li S, Ma J, Wen J, et al. Trends in consumption of ultra-processed foods among adults in Southern China: analysis of serial cross-sectional health survey data 2002-2022. Nutrients 2024 Nov 23;16(23):4008. [doi: 10.3390/nu16234008] [Medline: 39683401]
13. Yi B, Xu H. Research and development status of prepared foods in China: a Review. Appl Sci (Basel) 2023;13(14):7998. [doi: 10.3390/app13147998]
14. Hu XF, Zhang R, Chan HM. Identification of Chinese dietary patterns and their relationships with health outcomes: a systematic review and meta-analysis. Public Health Nutr 2024;27(1). [doi: 10.1017/S1368980024001927]
15. Jiang K, Zhang Z, Fullington LA, et al. Dietary patterns and obesity in Chinese adults: a systematic review and meta-analysis. Nutrients 2022 Nov 20;14(22):4911. [doi: 10.3390/nu14224911] [Medline: 36432596]
16. Yang YX, Wang XL, Leong PM, et al. New Chinese dietary guidelines: healthy eating patterns and food-based dietary recommendations. Asia Pac J Clin Nutr 2018;27(4):908-913. [doi: 10.6133/apjcn.072018.03] [Medline: 30045438]
17. Food-based dietary guidelines - China. Food and Agriculture Organization of the United Nations. 2022. URL: https://www.fao.org/nutrition/education/food-dietary-guidelines/regions/countries/china/en/#:~:text=The%20dietary%20guidelines%20for%20Chinese,edition)%20was%20released%20in%202022 [accessed 2025-07-24]
18. Ding Y, Xu F, Zhong C, et al. Association between Chinese dietary guidelines compliance index for pregnant women and risks of pregnancy complications in the Tongji maternal and child health cohort. Nutrients 2021 Mar 3;13(3):829. [doi: 10.3390/nu13030829] [Medline: 33802324]
19. Yang Y, Yang X, Zhai F, Cheng Y. Dietary guidelines for Chinese (2016). J Acad Nutr Diet 2016 Sep;116(9):A37. [doi: 10.1016/j.jand.2016.06.127]

XSL•FO
RenderX

20. Wang L, Wang H, Wang Z, et al. Interpretation of healthy diet campaign in healthy China initiative 2019-2030. China CDC Wkly 2021 Apr 16;3(16):346-349. [doi: 10.46234/ccdcw2021.092] [Medline: 34594881]
21. Jiang Z, Jiang W. Health Education in the Healthy China Initiative 2019-2030. China CDC Wkly 2021 Jan 22(4):78-80. [doi: 10.46234/ccdcw2021.018] [Medline: 34595007]
22. Gao C, Xu J, Liu Y, Yang Y. Nutrition policy and healthy China 2030 building. Eur J Clin Nutr 2021 Feb;75(2):238-246. [doi: 10.1038/s41430-020-00765-6] [Medline: 33219269]
23. Ye Y, Leeming J. By the numbers: China's changing diet. Nature New Biol 2023 Jun 22. [doi: 10.1038/d41586-023-02060-3] [Medline: 37349552]
24. Chen Y, Perez-Cueto FJA, Giboreau A, Mavridis I, Hartwell H. The promotion of eating behaviour change through digital interventions. IJERPH ;17(20):7488. [doi: 10.3390/ijerph17207488]
25. Zhong F, Chen X, Li J. The burden of type 2 diabetes attributable to dietary risks in China: insights from the global burden of disease study 2021. Public Health (Fairfax) 2024 Dec;237:122-129. [doi: 10.1016/j.puhe.2024.09.026]
26. Xu SS, Hua J, Huang YQ, Shu L. Association between dietary patterns and chronic kidney disease in a middle-aged Chinese population. Public Health Nutr 2020 Apr;23(6):1058-1066. [doi: 10.1017/S1368980019002805] [Medline: 31576799]
27. Wang YJ, Yeh TL, Shih MC, Tu YK, Chien KL. Dietary sodium intake and risk of cardiovascular disease: a systematic review and dose-response meta-analysis. Nutrients 2020 Sep 25;12(10):2934. [doi: 10.3390/nu12102934] [Medline: 32992705]
28. Fang Y, Xia J, Lian Y, et al. The burden of cardiovascular disease attributable to dietary risk factors in the provinces of China, 2002–2018: a nationwide population-based study. The Lancet Regional Health - Western Pacific 2023 Aug;37:100784. [doi: 10.1016/j.lanwpc.2023.100784]
29. Xiang H, Tao X, Guan X, et al. Contemporary Chinese dietary pattern: Where are the hidden risks? Front Nutr 2022;9:997773. [doi: 10.3389/fnut.2022.997773] [Medline: 36211490]
30. Tian W, Yan G, Xiong S, et al. Burden of depressive and anxiety disorders in China and its provinces, 1990-2021: findings from the Global Burden of Disease Study 2021. Br J Psychiatry 2025 Jan 22:1-11. [doi: 10.1192/bjp.2024.267B1] [Medline: 39838265]
31. Mansournia MA, Altman DG. Population attributable fraction. BMJ 2018 Feb 22;360:k757. [doi: 10.1136/bmj.k757] [Medline: 29472187]
32. Plass D, Hilderink H, Lehtomäki H, et al. Estimating risk factor attributable burden - challenges and potential solutions when using the comparative risk assessment methodology. Arch Public Health 2022 May 27;80(1):148. [doi: 10.1186/s13690-022-00900-8] [Medline: 35624479]
33. Stevens GA, Alkema L, Black RE, et al. Guidelines for accurate and transparent health estimates reporting: the GATHER statement. The Lancet 2016 Dec;388(10062):e19-e23. [doi: 10.1016/S0140-6736(16)30388-9]
34. Jiang K, Wen Y, Li S, et al. Differences in awareness of Chinese dietary guidelines among urban and rural residents: a cross-sectional survey in Southwest China. Int J Public Health 2023;68:1605344. [doi: 10.3389/ijph.2023.1605344] [Medline: 36712819]
35. Yang G, Wang Y, Zeng Y, et al. Rapid health transition in China, 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet 2013 Jun 8;381(9882):1987-2015. [doi: 10.1016/S0140-6736(13)61097-1] [Medline: 23746901]
36. Zhai F, Wang H, Du S, et al. Lifespan nutrition and changing socio-economic conditions in China. Asia Pac J Clin Nutr 2007;16 Suppl 1:374-382. [Medline: 17392135]
37. Johnson CO, Nguyen M, Roth GA, et al. Global, regional, and national burden of stroke, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2019 May;18(5):439-458. [doi: 10.1016/S1474-4422(19)30034-1]
38. Su B, Guo S, Zheng X. Transitions in chronic disease mortality in China: evidence and implications. China CDC Wkly 2023 Dec 15;5(50):1131-1134. [doi: 10.46234/ccdcw2023.211] [Medline: 38124884]
39. Kantono K, Hamid N, Ma Q, Chadha D, Oey I. Consumers' perception and purchase behaviour of meat in China. Meat Sci 2021 Sep;179:108548. [doi: 10.1016/j.meatsci.2021.108548] [Medline: 33993070]
40. Popkin BM. Will China's nutrition transition overwhelm its health care system and slow economic growth? Health Aff (Millwood) 2008;27(4):1064-1076. [doi: 10.1377/hlthaff.27.4.1064] [Medline: 18607042]
41. Wang Q, Liu S, Wang H, Su C, Liu A, Jiang L. Consumption of aquatic products and meats in Chinese residents: A nationwide survey. Front Nutr 2022;9:927417. [doi: 10.3389/fnut.2022.927417] [Medline: 35938120]
42. (China) NBoS. NBoS China Statistical Yearbook 2022: China Statistics Press; 2022. URL: https://www.stats.gov.cn/sj/ndsj/2022/indexch.htm [accessed 2025-08-20]
43. Liu J, Liu B, Shan B, et al. Prevalence of African Swine fever in China, 2018-2019. J Med Virol 2020 Aug;92(8):1023-1034. [doi: 10.1002/jmv.25638] [Medline: 31769521]
44. Zhang X, Wang J, Tan F, Gao H, Fan S. The potential impact of increased whole grain consumption among Chinese adults on reducing healthcare costs and carbon footprint. Journal of Integrative Agriculture 2024 Aug;23(8):2842-2852. [doi: 10.1016/j.jia.2023.12.006]
45. Yu D, Zhao L, Zhao W. Status and trends in consumption of grains and dietary fiber among Chinese adults (1982-2015). Nutr Rev 2020 Aug 1;78(Suppl 1):43-53. [doi: 10.1093/nutrit/nuz075] [Medline: 32728754]

XSL•FO
RenderX

46. Lara-Castor L, Micha R, Cudhea F, et al. Intake of sugar sweetened beverages among children and adolescents in 185 countries between 1990 and 2018: population based study. BMJ 2024 Aug 7;386:e079234. [doi: 10.1136/bmj-2024-079234] [Medline: 39111807]
47. Bleich SN, Vercammen KA. The negative impact of sugar-sweetened beverages on children's health: an update of the literature. BMC Obes 2018;5(1):6. [doi: 10.1186/s40608-017-0178-9] [Medline: 29484192]
48. Yu P, Chen Y, Zhao A, et al. Consumption of sugar-sweetened beverages and its association with overweight among young children from China. Public Health Nutr 2016 Sep;19(13):2336-2346. [doi: 10.1017/S1368980016001373]
49. WHO Manual on Sugar-Sweetened Beverage Taxation Policies to Promote Healthy Diets: World Health Organization; 2022. URL: https://iris.who.int/handle/10665/365285 [accessed 2025-08-20]
50. Zhai T, Wang J, Li L, Si W. Taxing sugar-sweetened beverages in China: by volume or sugar content? A consumer welfare approach. Journal of Integrative Agriculture 2024 Dec;23(12):4237-4249. [doi: 10.1016/j.jia.2024.10.006]
51. Han Z, Zheng X, Hou L, Xiao N, Deng X. Changes in China's food security driven by nutrition security and resource constraints. Environ Dev Sustain 2023 Feb 25:1-19. [doi: 10.1007/s10668-023-03042-1] [Medline: 37363032]
52. Zhou Y, Du S, Su C, Zhang B, Wang H, Popkin BM. The food retail revolution in China and its association with diet and health. Food Policy 2015 Aug;55:92-100. [doi: 10.1016/j.foodpol.2015.07.001]
53. Yin J, Zhang X, Huang W, et al. The potential benefits of dietary shift in China: synergies among acceptability, health, and environmental sustainability. Sci Total Environ 2021 Jul 20;779:146497. [doi: 10.1016/j.scitotenv.2021.146497] [Medline: 33752013]
54. Guo Q, Fang H, Zhao L, Ju L, Xu X, Yu D. Level of fruit and vegetable intake and its relationship with risk for malnutrition in China's adult labor force: China nutrition and health surveillance, 2015-2017. Nutrients 2023 Mar 16;15(6):1431. [doi: 10.3390/nu15061431] [Medline: 36986161]
55. Global Action Plan for the Prevention and Control of Non Communicable Diseases 2013–2020: World Health Organization; 2023.
56. Luo L, Jiang J, Yu C, et al. Stroke mortality attributable to low fruit intake in China: a joinpoint and age-period-cohort analysis. Front Neurosci 2020;14:552113. [doi: 10.3389/fnins.2020.552113] [Medline: 33335466]
57. Kickbusch IS. Health literacy: addressing the health and education divide. Health Promot Int 2001 Sep 1;16(3):289-297. [doi: 10.1093/heapro/16.3.289]
58. Fang K, He Y, Fang Y, Lian Y. Dietary sodium intake and food sources among Chinese adults: data from the CNNHS 2010-2012. Nutrients 2020 Feb 11;12(2):453. [doi: 10.3390/nu12020453] [Medline: 32054013]
59. Zhang J, Wang H, Wang Z, et al. Trends in adult cooking salt intake - China, 1991-2018. China CDC Wkly 2020 Feb 14;2(7):104-108. [Medline: 34594834]
60. Zhang P, Sun J, Li Y, et al. An mHealth-based school health education system designed to scale up salt reduction in China (EduSaltS): A development and preliminary implementation study. Front Nutr 2023;10:1161282. [doi: 10.3389/fnut.2023.1161282] [Medline: 37139455]
61. Neal B, Wu Y, Feng X, et al. Effect of salt substitution on cardiovascular events and death. N Engl J Med 2021 Sep 16;385(12):1067-1077. [doi: 10.1056/NEJMoa2105675] [Medline: 34459569]
62. Guo L, Huang F, Liu M, et al. Generational differences in food consumption among Chinese adults of different ages. Nutrients 2023 Oct 20;15(20):4451. [doi: 10.3390/nu15204451] [Medline: 37892526]
63. Zhang G, Kang J, Jing S, et al. Meat, vegetable, and fruit consumption among urban and rural elders aged 60+ years in regional China: a population-level nutritional study. J Nutr Sci 2023;12:e120. [doi: 10.1017/jns.2023.104] [Medline: 38155810]
64. Wang M, Ma H, Qin C, et al. Burden of diet-related chronic diseases in Chinese and Japanese adults attributable to dietary risk factors from 1990 to 2021: a systematic analysis of the Global Burden of Disease Study 2021. Front Nutr 2024;11:1472451. [doi: 10.3389/fnut.2024.1472451]
65. Press. CS. China Statistical Yearbook 2023: China Statistics Press; 2023. URL: https://www.stats.gov.cn/sj/ndsj/2023/indexch.htm [accessed 2025-08-20]
66. Liu W, Liu Z, Wang L, Liu H, Wang Y. Regional social development gap and regional coordinated development based on mixed-methods research: evidence from China. Front Psychol ;13. [doi: 10.3389/fpsyg.2022.927011]
67. Shi Y, Kan J, Li X, et al. Eastern diet - a healthful dietary pattern from Eastern China: its characteristics and relation to adiposity, cardiometabolic diseases, mortality, and gut microbiota. Nutrition. . [doi: 10.1101/2024.07.10.24310112]
68. Bai L, Tang H, Wang M. Dietary behaviors of rural residents in northeastern China: implications for designing intervention information and targeting high-risk population. Front Public Health ;12. [doi: 10.3389/fpubh.2024.1239449]
69. Wang HH. The perspective of meat and meat-alternative consumption in China. Meat Sci 2022 Dec;194:108982. [doi: 10.1016/j.meatsci.2022.108982]

Abbreviations

ASR: age-standardized rate
DALY: disability-adjusted life year

XSL•FO
RenderX

EAPC: estimating annual percentage change
GATHER: Guidelines for Accurate and Transparent Health Estimates Reporting
GBD: The Global Burden of Disease
NCD: noncommunicable diseases
PAF: population attributable fraction
SSB: sugar-sweetened beverage
STROBE: STrengthening the Reporting of OBservational studies in Epidemiology
TMREL: theoretical minimum risk exposure level
UI: uncertainty interval
WHO: World Health Organization
YLD: year lived with disability
YLL: year of life lost

Edited by A Mavragani, T Sanchez; submitted 22.02.25; peer-reviewed by N Xiao, S Huang; revised version received 05.07.25; accepted 06.07.25; published 25.08.25.

Please cite as:
Xin Y, Shui D, Yan G, Tian W, Tang N, Liang J, Peng J, Sun H, Ge A, Liu X, Kissock K, Trieu K, Zhang J, Zhang X, Ye P, Tian M
Dietary Risk Factors and Associated Disease Burden Among Chinese Adults Aged 25 Years and Older: Systematic Analysis of the Global Burden of Disease Study 2021
JMIR Public Health Surveill 2025;11:e72978
URL: https://publichealth.jmir.org/2025/1/e72978
doi:10.2196/72978

© Yuze Xin, Dong Shui, Guangcan Yan, Wei Tian, Ning Tang, Jinyu Liang, Junyi Peng, Hongru Sun, Anqi Ge, Xinyan Liu, Katrina Kissock, Kathy Trieu, Jing Zhang, Xinyi Zhang, Pengpeng Ye, Maoyi Tian. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 25.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Characterizing US Spatial Connectivity and Implications for Geographical Disease Dynamics and Metapopulation Modeling: Longitudinal Observational Study

Giulia Pullano[1], PhD; Lucila Gisele Alvarez-Zuzek[2], PhD; Vittoria Colizza[3], PhD; Shweta Bansal[1], PhD

[1]Department of Biology, Georgetown University, 37th and O Streets NW, Washington, DC, United States

[2]Fondazione Bruno Kessler, Trento, Italy

[3]Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique (IPLESP), Paris, France

Corresponding Author:
Shweta Bansal, PhD
Department of Biology, Georgetown University, 37th and O Streets NW, Washington, DC, United States

Abstract

Background: Human mobility is expected to be a critical factor in the geographic diffusion of infectious diseases, and this assumption led to the implementation of social distancing policies during the early fight against the COVID-19 emergency in the United States. Yet, because of substantial data gaps in the past, what still eludes our understanding are the following questions: (1) How does mobility contribute to the spread of infection within the United States at local, regional, and national scales? (2) How do seasonality and shifts in behavior affect mobility over time? (3) At what geographic level is mobility homogeneous across the United States?

Objective: This study aimed to address the questions that are critical for developing accurate transmission models, predicting the spatial propagation of disease across scales, and understanding the optimal geographical and temporal scale for the implementation of control policies.

Methods: We analyzed high-resolution mobility data from mobile app usage from SafeGraph Inc, mapping daily connectivity between the US counties to grasp spatial clustering and temporal stability. Integrating this into a spatially explicit transmission model, we replicated SARS-CoV-2's first wave invasion, assessing mobility's spatiotemporal impact on disease predictions.

Results: Analysis from 2019 to 2021 showed that mobility patterns remained stable, except for a decline in April 2020 due to lockdowns, which reduced daily movements from 45 million to approximately 25 million nationwide. Despite this reduction, intercounty connectivity remained seasonally stable, largely unaffected during the early COVID-19 phase, with a median Spearman coefficient of 0.62 (SD 0.01) between daily connectivity and gravity networks. We identified 104 geographic clusters of US counties with strong internal mobility connectivity and weaker links to counties outside these clusters. These clusters were stable over time, largely overlapping state boundaries (normalized mutual information=0.82) and demonstrating high temporal stability (normalized mutual information=0.95). Our findings suggest that intercounty connectivity is relatively static and homogeneous at the substate level. Furthermore, while county-level, daily mobility data best captures disease invasion, static mobility data aggregated to the cluster level also effectively models spatial diffusion.

Conclusions: Our work demonstrates that intercounty mobility was negligibly affected outside the lockdown period in April 2020, explaining the broad spatial distribution of COVID-19 outbreaks in the United States during the early phase of the pandemic. Such geographically dispersed outbreaks place a significant strain on national public health resources and necessitate complex metapopulation modeling approaches for predicting disease dynamics and control design. We thus inform the design of such metapopulation models to balance high disease predictability with low data requirements.

(JMIR Public Health Surveill 2025;11:e64914) doi:10.2196/64914

KEYWORDS

geographical disease dynamics; spatial connectivity; mobility data; metapopulation modeling; COVID-19; human mobility; infectious diseases; social distancing; epidemic; mobile apps; SafeGraph; SARS-CoV-2; coronavirus; pandemic; spatio-temporal; US; public health; mobile health; mHealth; digital health; health informatics

XSL•FO
RenderX

Introduction

Human mobility plays a crucial role in the spread of respiratory diseases [1]. The combination of regional travel and local commuting represents the spatial connectivity between locations, serving as the main driver in the geographic diffusion of infectious diseases. Characterizing the spatial dynamics of pathogen transmission is, therefore, intricately tied to unraveling human mobility patterns. Such a task has proven to be challenging due to the inherent complexity and privacy-related limitations on collecting mobility data [2]. Over the past few decades, researchers have extensively relied on mobility data obtained from census records, surveys, transportation statistics, commuting data, and international air traffic data. Such datasets have widely contributed to a better understanding of human mobility patterns and their impact on the epidemic spread [3-6], but can be limited in their resolution or scale. More recently, this gap has been filled by the use of mobile phone data [7,8], primarily based on phone records, but no such data have been available in the United States.

The global health crisis triggered by COVID-19 has underscored the critical need for swift access to mobility to help mitigate the spread of the virus. The urgency of the situation prompted an unprecedented sharing of data by private companies worldwide, through legally and ethically compliant agreements. These data were based on mobile location–based app usage and thus provided incomparable access to high-resolution, large-scale, and near-real-time mobility data that have expanded human mobility science [9] and computational epidemiology [10,11]. The availability of these data has especially represented a shift in the US public health and it has been used to inform epidemic models and reveal the impact of mitigation strategies on behavior [12-16]. Although the association between mobility patterns and COVID-19 transmission in the United States has been extensively explored, for example, in studies by Pei et al [14], Badr et al [16], Xiong et al [17], Tokey [18], Zheng et al [19], and Gao et al [20], no studies have been devoted to assessing when the underlying mobility network needs to be embedded into models to characterize epidemic spread.

Furthermore, the effects of control measures such as social distancing and travel restrictions as well as the most suitable geographical and temporal granularity for implementing these measures still lack clarity. This gap in understanding the characteristic spatio-temporal scale of mobility not only limits target control policies but also our ability to model transmission dynamics effectively. To date, mobility data have been integrated into epidemic models without due consideration for the optimal geographical (eg, municipalities, regions, and states) and temporal resolution (eg, day, week, and month) required to accurately capture epidemic spread. The level of granularity used in these models has consistently been dictated by a priori assessments from data providers [21,22].

To address these gaps, this study made 3 key contributions. First, for characterizing spatiotemporal scales, we systematically analyzed human mobility patterns across the United States using high-resolution mobile app–based location data to characterize intercounty connectivity at different spatial and temporal scales before and after the COVID-19 pandemic. Second, for quantifying mobility's impact on the geographic spread of disease, we integrated these mobility data into spatially explicit transmission models to assess its role in the geographic diffusion of SARS-CoV-2 during the first wave. Finally, for evaluating the predictive power across scales, we analyzed how the predictive ability of epidemic models changes with varying resolutions of mobility data, identifying the spatial and temporal scales of intercounty connectivity that significantly influence disease dynamics. This analysis allowed us also to explore the trade-offs between using fine-grained mobility data and aggregated data at coarser scales.

While our approach yields valuable insights into human behavior and disease dynamics and makes a theoretical contribution to the field, we acknowledge its limitations—most notably the simplifications in the disease model, such as assuming homogeneous mixing within counties. However, our focus was on understanding heterogeneity at scales larger than counties, as most public health data are typically available at the county level. This approach represents a significant step toward understanding the role of mobility in geographical disease diffusion and optimizing the integration of mobility data into epidemic models to better inform public health policies.

Methods

Study Design and Population

This study investigated the role of human mobility at various temporal and spatial scales in the spread of COVID-19 across the US counties, using mobility and disease incidence data. First, we examined human displacements and the resulting intercounty connectivity patterns in the United States based on the daily number of visits between census blocks. We also analyzed disease incidence during the early phases of the COVID-19 pandemic, from January 2019 to July 2020, and used an inference framework to estimate underreporting based on the number of new deaths. In the second step, we developed a metapopulation model to simulate disease transmission, incorporating intercounty mobility patterns and epidemiological data, adjusted for underreporting. The study focused on US counties with populations greater than 10,000 to minimize biases in mobility and incidence data.

Data Collection

Characterizing Intercounty Connectivity With Mobility Data

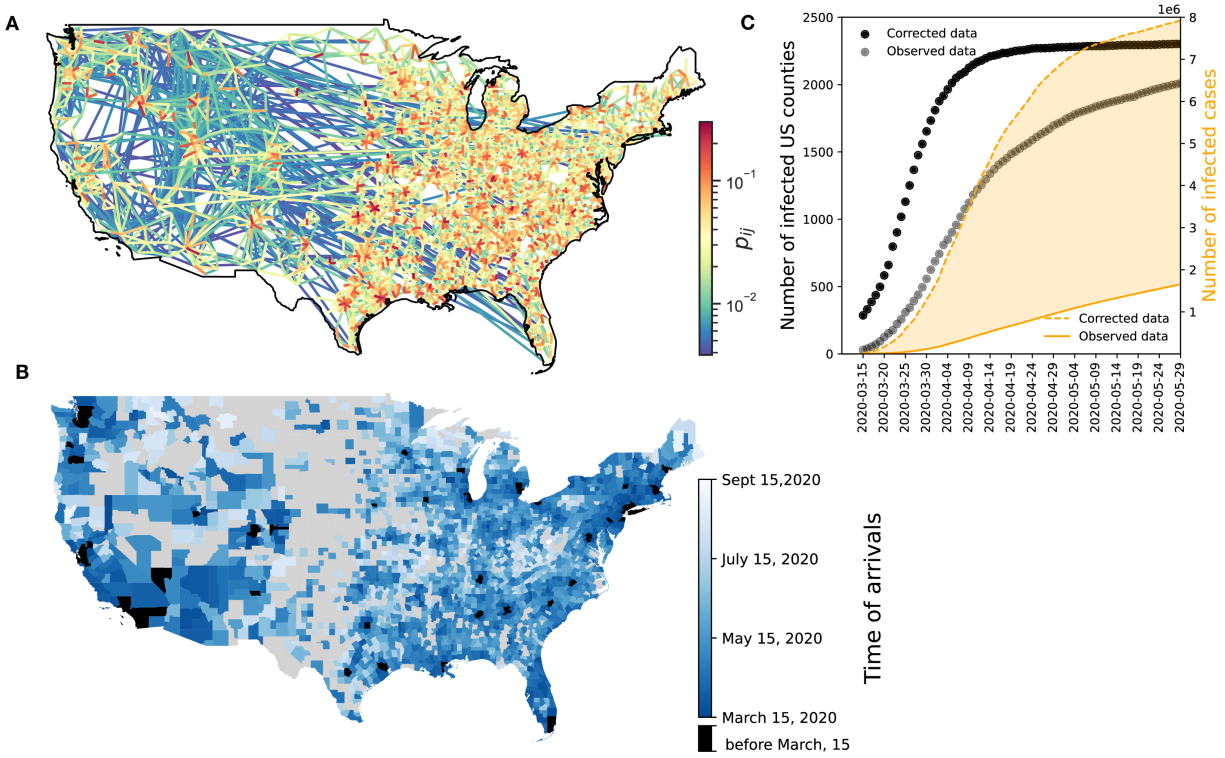
We used data from SafeGraph [23] (now Advan Patterns [Advan Inc]), a platform gathering mobility data from location-based mobile app usage. Specifically, we used the daily Social Distancing dataset provided by SafeGraph (refer to Multimedia Appendix 1 for dataset details). This dataset includes information on the number of mobile devices with a home from a census block group visiting another census block group or staying in the originating one for at least a minute. The data covered the period from January 2019 to April 2021 on a daily basis. These data were aggregated to the US county level to ensure consistency in geographic scale for disease surveillance

XSL•FO
RenderX

and public health decision-making. A correction factor addresses spatial and temporal variations (Multimedia Appendix 1).

We quantified monthly intercounty connectivity by normalizing visits from an origin county to destination counties and calculated average daily visits for all county pairs from January 2019 to March 2021. We then normalized the results so that the sum of probabilities for each county equals 1. The resulting time-evolving network reflected daily coupling probabilities. Figure 1A shows the connectivity network for March 2020. Figure S2 in Multimedia Appendix 1 compares monthly and daily datasets.

Figure 1. Data sources and epidemic context. (A) Mobility data. The figure shows the spatial connectivity network between US counties in March 2020. The map shows only links with the top 1% of coupling probability (p_{ij}>.0038). (B-C) Public health data. (B) The map shows the time of arrival by county, defined as the time when the total number of cases (observed+unreported) in counties that reached at least 10 cases. Black-colored counties are counties that have been infected before March 15, 2020. (C) Black dots show the corrected number of infected counties over time nationally. Gray dots show the observed number of infected counties over time nationally. The orange solid line shows the daily number of new confirmed cases nationally, while the orange dotted line shows the real number of cases nationally accounting for underreporting.



Characterizing Early Phase of COVID-19 in the United States With Public Health Data

The first confirmed case of COVID-19 in the United States emerged in Washington state on January 21, 2020, quickly leading to local transmission. Guidelines promoting social distancing and discouraging gatherings were issued on March 16, 2020. While many European countries enacted nationwide restrictions, the United States implemented localized measures at varying times. Lockdowns peaked in the United States in April 2020, with over 40 states enforcing stay-at-home or shelter in-place orders [24]. Despite these efforts, COVID-19 continued to spread, with most US counties reporting cases by June 2020. As cases surged again in October 2020, recommendations for social distancing were made to keep epidemic activity in check.

The COVID-19 disease incidence data were derived from the Centers for Disease Control and Prevention (CDC) data [24]. We used daily reports of new cases and the timing of their arrival in each US county, defined as the day when at least 10 cases were recorded in that area. Daily new-reported cases and time of arrivals were adjusted for potential underreporting at the county level, estimated using global data on COVID-19 cases and fatalities, as outlined in a study by Russell et al [25] (Figures 1B and 1C).

Statistical Analysis

Describing Temporal and Spatial Variability in the Mobility Network

We examined the monthly network structure to evaluate the temporal mobility patterns, quantifying each county's degree (number of connections to other counties). We also defined link persistence as the probability links with nonzero mobility in a month of 2019 that remained in the same month of 2020 and 2021. We also fit a gravity model to the intercounty connectivity network for each month. Community detection analysis using the stochastic InfoMap algorithm developed by Rosvall et al identifies regions with more frequent internal movements [26]. We used a bootstrap resampling method to account for stochasticity (refer to Multimedia Appendix 1 for details). Urban and rural classifications are based on the National Center for Health Statistics Urban-Rural Classification Scheme. All network analyses were performed using Python's NetworkX

XSL•FO
RenderX

library [27], which is widely used for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. Gravity model fitting was conducted using the scikit-mobility library, a Python package designed for mobility data analysis [28].

Incorporating Human Mobility Into Infectious Disease Models

We used a stochastic non-Markovian transmission model with a metapopulation structure at the US county level [21]. The model accounts for disease transmission proportional to (1) infected residents not moving, (2) infected visitors coming from other counties, and (3) returning residents previously infected in other counties. The resulting force of infection in the county i is defined as follows:

$$\lambda i=\lambda ii+\sum i\neq j\lambda jiv+\sum i\neq j\lambda ijr$$

$$\lambda ii=\beta pii2IiNi\wedge$$

$$\lambda jiv=\beta piipjiIjN\wedge i;\lambda ijr=\beta pijI\wedge jN\wedge j$$

where p_{ij} is the coupling probability extracted from the intercounty connectivity network. The effective population, and effective number of infections are, respectively, defined as follows:

$$Ni\wedge=piiNi+\sum i\neq jpjiNj$$

$$Ii\wedge=piiIi+\sum i\neq jpjiIj$$

We considered the susceptible-exposed-infectious-recovered epidemic dynamics specific to COVID-19. The epidemics parameters were described by Pullano et al [29].

The detailed mathematical framework, model calibration, and implementation details can be found in Multimedia Appendix 1. The model and inference framework were developed in C++ to optimize performance and efficiency, particularly when working with large-scale data.

Inference Framework and Goodness of Fit

To calibrate the epidemic pathway, we used the cumulative number of infected US counties (Figure 1C). Calibration covered March 14 to July 15, 2020, when all counties reported infections. Parameter estimates β_{pre-LD} (March 15 - 31) and $\beta_{post-LD}$ (March 31-May 15) were derived using maximum likelihood, assuming a Poisson distribution for reported cases.

We assessed model performance by comparing the modeled invasion probability $p_{i,inv}(t)$ with the observed early phase COVID-19 spatial invasion. $p_{i,inv}(t)$ denotes the probability for a county i reporting at least 10 infected cases on day t in the simulation. The goodness of fit is defined as follows:

$$G(t)=\sum i(Iilog\ pi,inv(t)+(1-Ii)log\ (1-pi,inv(t)))$$

I_i=1 if the county i reported at least 10 infected cases at the day t; I_i=0 otherwise.

Comparing Models Across Geographical Scales

In designing the metapopulation structure at different spatial scales, we aimed to better understand the role of mobility heterogeneity that matters in disease diffusion. Beyond this primary goal, this approach also highlighted the data requirements necessary for accurate modeling at different spatial scales. Specifically, it explored the implications of using mobility data at a coarser regional scale (eg, cluster or state), assuming higher-resolution flow data are unavailable.

For a given spatial scale R (eg, cluster and state), we randomized the mobility links among counties within a region R, preserving the number of links. Coupling probabilities for connected counties within each region R are equal to the average coupling probability of the links within that region, as follows:

$$pijR=\sum l,k\in RplkN$$

with i,j, $\in$, and R, while coupling probabilities across regions are not changed.

Ethical Considerations

This study used publicly available, deidentified, and aggregated data from SafeGraph (now Advan Patterns) and the CDC. SafeGraph ensures user privacy by implementing differential privacy techniques, which involve adding Laplacian noise to anonymize data at the Census Block level. In addition, SafeGraph excluded connections associated with census block groups containing data from fewer than 2 devices. Ethical review for this study was sought from the Institutional Review Board at Georgetown University and the study was approved on October 14, 2020 (STUDY00003041).

Results

Analyzing the US county connectivity via mobility data revealed temporal and geographical stability, identifying the key scale driving COVID-19's initial spread. Our findings addressed public health needs and determined the optimal scale for metapopulation model design.

Temporal Stability of the Intercounty Connectivity Network

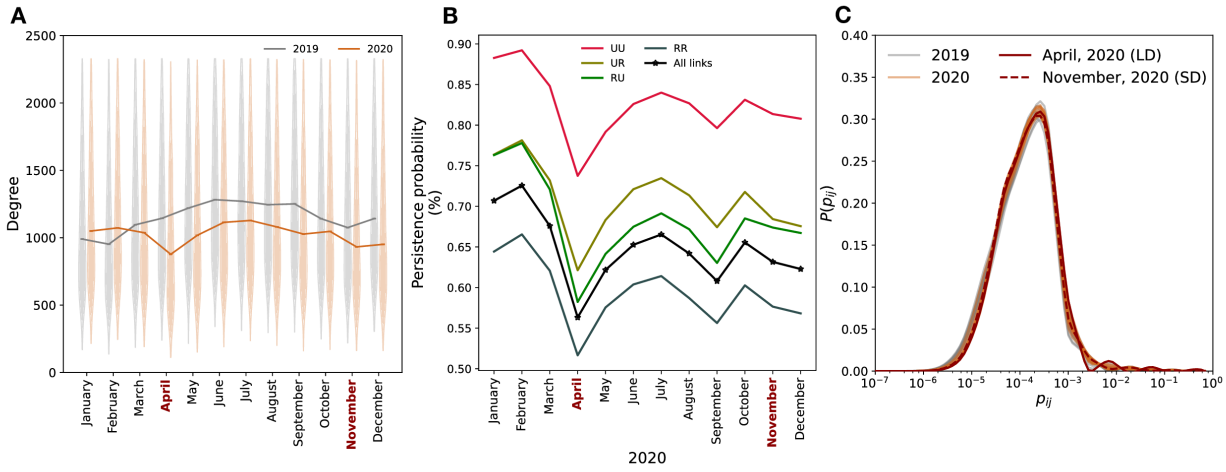
From January 2019 to March 2021, mobility remained relatively stable, except for a significant drop in April 2020 coinciding with lockdown measures, reducing daily visits from 45 million to around 25 million nationwide (Figure 2A). The mobility shock extended throughout the month, encompassing a transitional period (Figures 2A and 2B). Analysis of the intercounty connectivity network's temporal evolution revealed a consistent seasonal pattern in degree distribution and link persistence. Only April 2020 showed local variations, with a 23% reduction in degree (from 1144 to 877) compared with 2019 and a 20% reduction in link persistence (from 0.70 to 0.56) compared to previous months in 2020. Surprisingly, no variation occurred in November 2020, despite strong social distancing recommendations preceding the winter surge of SARS-CoV-2. The reduction in rural-urban connections was particularly pronounced, with a 22% decrease (from 0.75 to 0.58). The decrease stabilized in May and beyond. Notably, urban-urban connections exhibited greater resilience over time when compared with connections involving rural areas. We hypothesized that rural-rural connections were systematically less persistent due to the inherently less stable nature of these links. Factors such as lower population density, reduced economic activity, and less frequent interconnectivity in rural areas contributed to this instability. Furthermore, while coupling

XSL•FO
RenderX

probabilities stayed consistent over the study period (Figure 1C), the probability of staying in the home location exhibited larger variability (Multimedia Appendix 1).

Despite occasional local fluctuations, the intercounty connectivity network demonstrated temporal stability and high predictability through a gravity fit model (Figures S7-S9 in Multimedia Appendix 1). The Spearman coefficient between the original and modeled intercounty connectivity network remained constant over time, with a median of 0.62 (SD 0.02).

Figure 2. Temporal stability in intercounty mobility. (A) The monthly degree distribution of the intercounty network is shown as a violin plot with 95% CIs. It demonstrates significant variability while remaining consistent across months. (B) The persistence probability of links is illustrated, denoting the likelihood that a connection existing in 2019 remains present in 2020 and 2021. The plot provides a breakdown for different link types: urban-urban (UU), urban-rural (UR), rural-urban (RU), and rural-rural (RR) links. (C) Distribution of coupling probabilities in the connectivity network by month. We highlighted in dark red: April, 2020 (LD) that represents the peak time of number of US states in lockdown, and November, 2020 (SD) that represents the period when social distancing recommendations were in place. LD: lockdown; RR: rural-rural; RU: rural-urban; SD: social distancing; UR: urban-rural; UU: urban-urban.

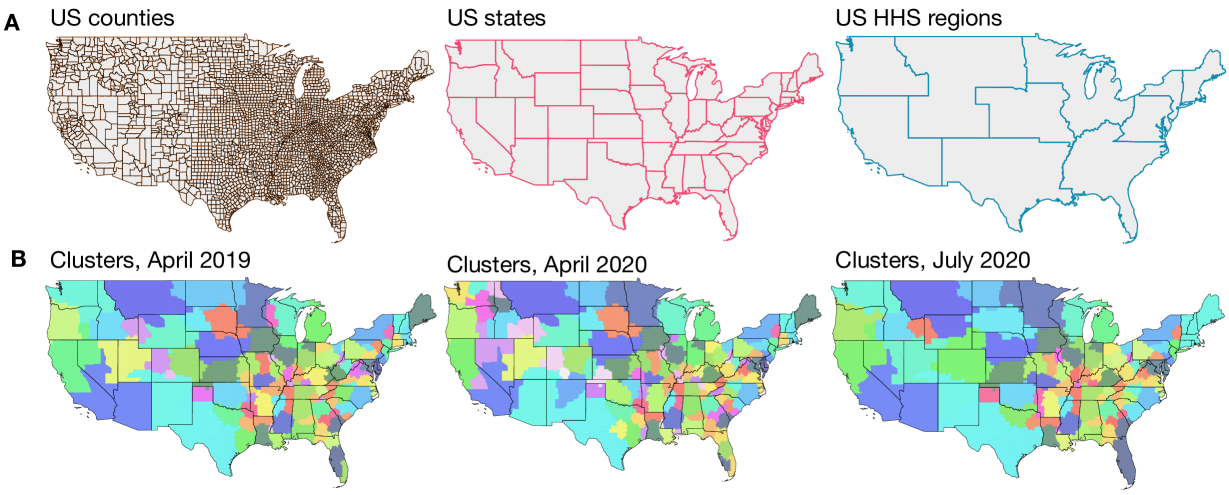


Spatial Stability of the Intercounty Connectivity Network

To identify the geographic scale at which mobility is highly connected, we detected clusters of counties that were more connected via mobility within the cluster than outside the clusters, we used a network community detection algorithm. Our hypothesis was that this partitioning of the United States would be at a geographic scale larger than 3143 US counties but smaller than 50 US states or 10 Health and Human Services (HHS) regions. Indeed, we found that based on human mobility, the United States can be partitioned into around 100 regions that split most US states into multiple clusters (Figure 3). We also found that these clusters were highly and spatially contiguous and respected state boundaries (with a similarity measured by normalized mutual information as 0.82). Furthermore, these regions demonstrated stability over time (normalized mutual information=0.95) despite the perturbations of the early phases of the COVID-19 pandemic (Figure 3B and in Figures S9-S10 in Multimedia Appendix 1). Thus, we identified a persistent geographic partitioning of the United States in which clusters were more connected within than between, and hypothesized that the relevance of mobility to the spatial diffusion of infectious diseases occured at a mesoscale.

XSL•FO
RenderX

Figure 3. Analysis of spatial stability. (A) Geographical subdivisions at the county and state levels within the United States, as well as the division into Health and Human Services regions, used for health administration purposes. (B) We partition the intercounty mobility network so that each cluster of counties is more strongly connected to each other via mobility than to counties in other clusters. We find highly consistent portioning based on mobility networks from April 2019, April 2020 (during the mandated lockdown period of the early COVID-19 pandemic), and July 2020 (after the lockdown period of the early COVID-19 pandemic). Clusters are colored to delineate cluster boundaries and do not represent any other information. Counties colored in gray have populations of fewer than 11,000 inhabitants and are excluded from the analysis. HHS: Health and Human Services.



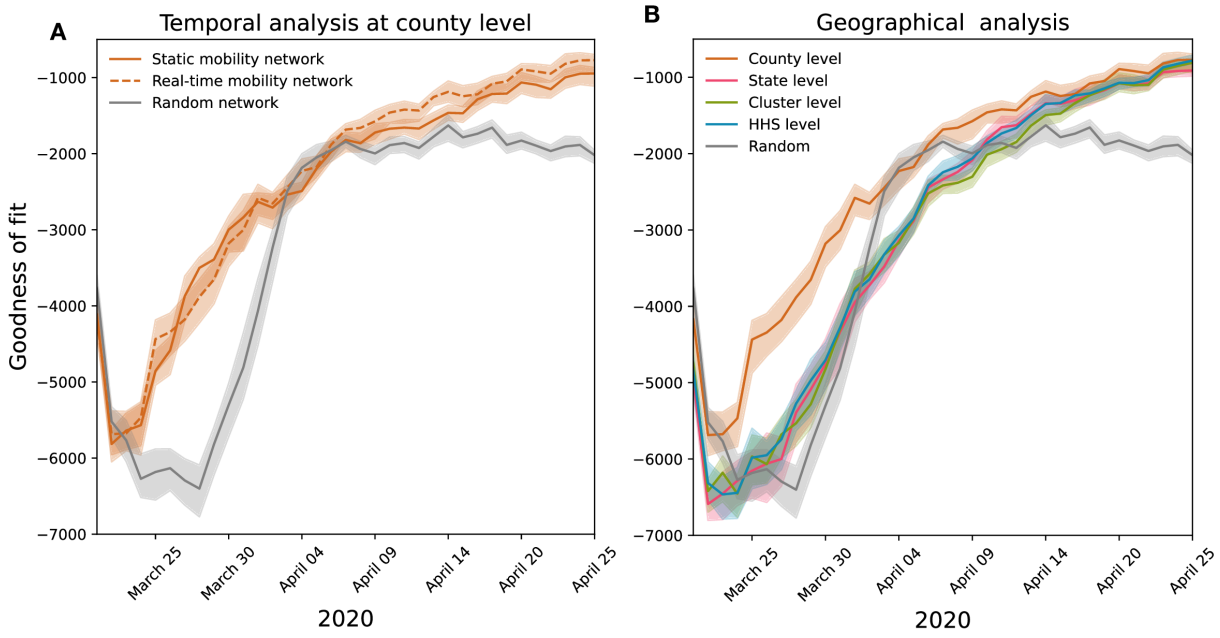
Implications for Metapopulation Disease Models

After examining the stability of mobility patterns over time and space, we evaluated how spatiotemporal scales of mobility affect disease modeling in a metapopulation framework by integrating connectivity networks into a model simulating the initial SARS-CoV-2 spread in the United States. We investigated the influence of geographic scale by homogenizing networks at different spatial levels, maintaining the county resolution. In addition, to gauge the impact of temporal scale, we provided the model with either a time-evolving or static connectivity network representing mobility patterns since March 2020, considering the network's temporal stability. Goodness of fit was assessed by comparing predicted and observed disease arrival times in counties.

Figure 4A shows that a county-level metapopulation model using empirical connectivity networks accurately predicts early COVID-19 diffusion, outperforming a Erdős–Rényi random network. Indeed, the empirical mobility network had a stronger goodness of fit throughout the early phase of the pandemic. This emphasized the crucial role of human mobility in the spatial spread of the initial SARS-CoV-2 wave and underscored the importance of mobility data for modeling SARS-CoV-2 spread. In addition, predictions based on static and time-varying mobility networks were comparable, suggesting that static data were sufficient to accurately capture epidemic spatial heterogeneity. Following the initial invasion phase in March, transmission became more widespread, reducing the importance of spatial connections. At this stage, random networks and county-level networks yielded similar performance, indicating that local epidemic growth outweighed the influence of importations and exportations. Figure 4B shows that county-level mobility data predicted spatial diffusion better than data at larger scales, such as US HHS regions, states, and clusters (as defined in Figure 3). All larger-scale data performed similarly to a random network, indicating they lack the granularity needed to capture diverse mobility patterns effectively. Sensitivity analyses on the definition of the time of arrivals and its impact on the goodness of fit are reported in Multimedia Appendix 1, more precisely in Figure S14B.

XSL•FO
RenderX

Figure 4. Implications of temporal and spatial scale of mobility data for the prediction with metapopulation disease models. (A) The goodness of fit (median and 95% CI) for the time of arrivals for metapopulation models at a county level, informed by a time-evolving intercounty connectivity network, a static intercounty connectivity network, and a random intercounty network. (B) Solid lines show the goodness of the fit (median and 95% CI) for metapopulation models informed with state-, cluster-, and Health and Human Services–level static mobility network. HHS: Health and Human Services.



Discussion

Principal Results

Our findings revealed significant insights into the dynamics of human mobility and their implications for infectious disease modeling. In contrast to findings from other countries (eg, France [29], India [30], Germany [31], and Spain [32]), we observed that despite the implementation of local social distancing measures and lockdowns, intercounty connectivity remained largely unperturbed, leading to rapid geographic diffusion of SARS-CoV-2. Mobility patterns experienced only marginal changes before and after the early-stage COVID-19 pandemic. The most notable disruption occurred during the first lockdown period in April 2020, when mobility sharply declined. Although a temporary reduction in mobility was observed, it proved to be short-lived as mobility patterns quickly returned to prerestriction levels. Importantly, this decline in mobility did not alter the underlying intercounty connectivity structure, potentially diminishing the overall effectiveness of the implemented travel restrictions. Even when mobility is limited—resulting in low-strength edges between locations—these connections can still facilitate the spread of disease, introducing pathogens into new areas [33]. This underscores the challenge of relying solely on travel restrictions as a mitigation strategy, as they may fail to disrupt the pathways that enable disease transmission, particularly when the connectivity network remains intact. Such insights emphasize the need for more nuanced approaches that address the persistence of network structures during public health interventions. Notably, even during periods of social distancing recommendations, the mobility network remains relatively stable. Assuming the lockdown represents the most extreme form of mobility disruption, the temporal stability findings suggest that global human mobility demonstrates resilience against short-term changes.

We also assessed the spatial stability of the intercounty connectivity network by detecting spatial communities based on mobility patterns. Our results indicate that mobility-driven clusters align closely with state boundaries, reflecting the influence of administrative and geographical factors on human movement, accordingly with Steinegger et al [31]. These clusters exhibited remarkable stability over time, reinforcing the idea that spatial mobility patterns are deeply ingrained and relatively resistant to abrupt changes. The fact that mobility patterns are highly correlated with state boundaries suggests that state-level structures could be effective for designing target public health interventions based on travel reductions. Our findings underscore the importance of considering mobility patterns when designing interventions, resource allocation, and disease control strategies.

As shown in the context of COVID-19 pandemic in France [34], we also demonstrated that incorporating high-resolution human mobility data are crucial for accurately capturing the spatial spread of infectious diseases. Our findings indicate that county-level, daily mobility data offer the most accurate representation of the spatial spread of disease in the United States. Notably, static county-level mobility data achieve similar model performance to real-time data, suggesting that an undisturbed representation of reality is adequate for reproducing spatial spread. More interestingly, our exploration of various spatial scales for metapopulation models underscores the significance of aligning the model's structure with the inherent spatial scale of human mobility. While county-level mobility data yield the most accurate depiction, mobility databased

XSL•FO
RenderX

clusters, US states, and HHS regions do not capture the heterogeneity of the COVID-19 geographical diffusion.

Limitations

While our study provides valuable insights, it is not without limitations. Our work focused on the early phase of the pandemic, during which response measures (eg, social distancing, closures, and lack of masking) were largely homogeneous in the United States, and pharmaceutical measures (eg, vaccination and antivirals) were not available; thus, these findings are not generalizable to later stages. Furthermore, we assume homogeneity within the US counties. In addition, Safegraph mobility data, like all mobile app–based location data, exhibit sampling biases. Ongoing efforts to comprehend these biases are crucial for developing better correction methods [35]. An independent analysis by Safegraph revealed the underrepresentation of older and non-White individuals in point of interest-specific analyses, though the panel is representative of race, education, and income [35,36].

Comparison With Previous Work

Since the onset of the COVID-19 pandemic, mobile phone data have played a crucial role in addressing the public health crisis [10,11,13-15,22,29,30,34,37]. During this period, numerous network operators and private enterprises have made considerable efforts to swiftly share their data within the confines of legal regulations. Consequently, researchers worldwide have embarked on working with this data, monitoring human behavior caused by containment measures and adaptive responses to the epidemic, and using it to enhance epidemic models in order to increase their reliability [14,15,22,29,34].

While static mobility data have predominantly been analyzed and integrated into models over the past decades [3,4], the current accessibility to real-time human behavior data prompts an essential investigation into the optimal scenarios for using this dynamic information versus relying solely on static representations of reality [38]. Equally important is the exploration of the characteristic mobility scale to comprehensively capture the intricate coupling between different locations, a consideration with potential implications for target control policies to reduce epidemic activity, and for improving epidemic model forecasting. Furthermore, numerous researchers have emphasized the pressing necessity to implement standardized strategies that facilitate rapid data access while upholding stringent data privacy measures [2]. To address this gap in the literature, in this study, we investigated the spatial connectivity of the US counties during the early phase of the COVID-19 pandemic using high-resolution real-time human mobility data obtained from mobile phone usage.

Conclusions

While characterizing the key role of mobility in the spatial invasion of the COVID-19 pandemic in the United States, our study sheds light on the global stability of human mobility patterns, and the relevant information needed to design a reliable predictive model. This result may be specific to countries, such as the United States, in which mobility restrictions were not stringent, specified for intercounty mobility, nor enforced. Metapopulation models that incorporate accurate mobility data can provide valuable insights into disease dynamics and enhance our ability to predict and control the spread of future infectious disease outbreaks. Furthermore, standardized data extraction and sharing that we introduced might help facilitate the timelines associated with legal agreements for data sharing, which do not always align with the rapid spread of epidemics, thus diminishing the feasibility of timely responses to such outbreaks.

Acknowledgments

The authors thank the teams at Safegraph (now Advan Patterns) for sharing mobility data. The research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under awards R01GM123007 and R35GM153478.

Authors' Contributions

GP and SB conceived of and designed the study. GP and LGA-Z analyzed the data and did the analysis. GP, VC, and SB interpreted the results. GP drafted the manuscript. All authors contributed to the writing of the final version of the manuscript.

Conflicts of Interest

SB has received funding from the Merck Pharmaceuticals Investigator Studies Program.

Multimedia Appendix 1
Supplementary material.
[DOCX File, 8862 KB - publichealth_v11i1e64914_app1.docx]

References

1. Findlater A, Bogoch II. Human mobility and the global spread of infectious diseases: a focus on air travel. Trends Parasitol 2018 Sep;34(9):772-783. [doi: 10.1016/j.pt.2018.07.004] [Medline: 30049602]
2. de Montjoye YA, Gambs S, Blondel V, et al. On the privacy-conscientious use of mobile phone data. Sci Data 2018 Dec 11;5(1):180286. [doi: 10.1038/sdata.2018.286] [Medline: 30532052]

XSL•FO
RenderX

3. Colizza V, Barrat A, Barthélemy M, Vespignani A. Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study. BMC Med 2007 Nov 21;5(1):34-7015. [doi: 10.1186/1741-7015-5-34] [Medline: 18031574]
4. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. Proc Natl Acad Sci U S A 2009 Dec 22;106(51):21484-21489. [doi: 10.1073/pnas.0906910106] [Medline: 20018697]
5. Barbosa H, Barthelemy M, Ghoshal G, et al. Human mobility: models and applications. Phys Rep 2018 Mar;734:1-74. [doi: 10.1016/j.physrep.2018.01.001]
6. Viboud C, Bjørnstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science 2006 Apr 21;312(5772):447-451. [doi: 10.1126/science.1125237] [Medline: 16574822]
7. Blondel VD, Decuyper A, Krings G. A survey of results on mobile phone datasets analysis. EPJ Data Sci 2015 Dec;4(1):10-1127. [doi: 10.1140/epjds/s13688-015-0046-0]
8. Maxmen A. Can tracking people through phone-call data improve lives? Nature New Biol 2019 May 30;569(7758):614-617. [doi: 10.1038/d41586-019-01679-5]
9. Pappalardo L, Manley E, Sekara V, Alessandretti L. Future directions in human mobility science. Nat Comput Sci 2023 Jul;3(7):588-600. [doi: 10.1038/s43588-023-00469-4] [Medline: 38177737]
10. Oliver N, Lepri B, Sterly H, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. Sci Adv 2020 Jun;6(23):eabc0764. [doi: 10.1126/sciadv.abc0764] [Medline: 32548274]
11. Grantz KH, Meredith HR, Cummings DAT, et al. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. Nat Commun 2020;11(1):4961-1723. [doi: 10.1038/s41467-020-18190-5]
12. Garnier R, Benetka JR, Kraemer J, Bansal S. Socioeconomic disparities in social distancing during the COVID-19 pandemic in the United States: observational study. J Med Internet Res 2021 Jan 22;23(1):e24591. [doi: 10.2196/24591] [Medline: 33351774]
13. Susswein Z, Rest EC, Bansal S. Disentangling the rhythms of human activity in the built environment for airborne transmission risk: An analysis of large-scale mobility data. elife 2023 Apr 4;12:e80466. [doi: 10.7554/eLife.80466] [Medline: 37014055]
14. Pei S, Kandula S, Shaman J. Differential effects of intervention timing on COVID-19 spread in the United States. Sci Adv 2020 Dec 4;6(49):2375-2548. [doi: 10.1126/sciadv.abd6370]
15. Chang S, Pierson E, Koh PW, et al. Mobility network models of COVID-19 explain inequities and inform reopening. Nature New Biol 2021 Jan;589(7840):82-87. [doi: 10.1038/s41586-020-2923-3] [Medline: 33171481]
16. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis 2020 Nov;20(11):1247-1254. [doi: 10.1016/S1473-3099(20)30553-3] [Medline: 32621869]
17. Xiong C, Hu S, Yang M, Luo W, Zhang L. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. Proc Natl Acad Sci USA 2020 Nov 3;117(44):27087-27089. [doi: 10.1073/pnas.2010836117]
18. Tokey AI. Spatial association of mobility and COVID-19 infection rate in the USA: a county-level study using mobile phone location data. J Transp Health 2021 Sep;22:101135. [doi: 10.1016/j.jth.2021.101135] [Medline: 34277349]
19. Zheng Z, Xie Z, Qin Y, Wang K, Yu Y, Fu P. Exploring the influence of human mobility factors and spread prediction on early COVID-19 in the USA. BMC Public Health 2021 Dec;21(1):1-13. [doi: 10.1186/s12889-021-10682-3]
20. Gao S, Rao J, Kang Y, et al. Association of mobile phone location data indications of travel and stay-at-home mandates with COVID-19 infection rates in the US. JAMA Netw Open 2020 Sep 1;3(9):e2020485. [doi: 10.1001/jamanetworkopen.2020.20485] [Medline: 32897373]
21. Pullano G. Human mobility and epidemics. PhD thesis.: Sorbonne Université; 2021 URL: https://theses.hal.science/tel-04099742
22. Zhang M, Wang S, Hu T, et al. Human mobility and COVID-19 transmission: a systematic review and future directions. Ann GIS 2022 Oct 2;28(4):501-514. [doi: 10.1080/19475683.2022.2041725]
23. Safegraph. URL: https://www.safegraph.com [accessed 2025-02-10]
24. CDC. COVID Data Tracker. 2020. URL: https://covid.cdc.gov/covid-data-tracker [accessed 2025-02-10]
25. Russell TW, Golding N, Hellewell J, et al. Reconstructing the early global dynamics of under-ascertained COVID-19 cases and infections. BMC Med 2020 Oct 22;18(1):332. [doi: 10.1186/s12916-020-01790-9] [Medline: 33087179]
26. Rosvall M, Axelsson D, Bergstrom CT. The map equation. Eur Phys J Spec Top 2009 Nov;178(1):13-23. [doi: 10.1140/epjst/e2010-01179-1]
27. NetworkX Python Library. URL: https://networkx.org/ [accessed 2025-02-10]
28. Scikit-Mobility Python Library. URL: https://scikit-mobility.github.io/scikit-mobility/ [accessed 2025-02-10]
29. Pullano G, Valdano E, Scarpa N, Rubrichi S, Colizza V. Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the COVID-19 epidemic in France under lockdown: a population-based study. Lancet Digit Health 2020 Dec;2(12):e638-e649. [doi: 10.1016/S2589-7500(20)30243-0] [Medline: 33163951]

XSL•FO
RenderX

30. Saha J, Barman B, Chouhan P. Lockdown for COVID-19 and its impact on community mobility in India: an analysis of the COVID-19 Community Mobility Reports, 2020. Child Youth Serv Rev 2020 Sep;116:105160. [doi: 10.1016/j.childyouth.2020.105160] [Medline: 32834269]
31. Sobolevsky S, Szell M, Campari R, Couronné T, Smoreda Z, Ratti C. Delineating geographical regions with networks of human interactions in an extensive set of countries. PLoS One 2013;8(12):e81707. [doi: 10.1371/journal.pone.0081707] [Medline: 24367490]
32. Steinegger B, Granell C, Rapisardi G, et al. Joint analysis of the epidemic evolution and human mobility during the first wave of COVID-19 in Spain: retrospective study. JMIR Public Health Surveill 2023 May 22;9:e40514. [doi: 10.2196/40514] [Medline: 37213190]
33. Colizza V, Vespignani A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: theory and simulations. J Theor Biol 2008 Apr 7;251(3):450-467. [doi: 10.1016/j.jtbi.2007.11.028] [Medline: 18222487]
34. Pullano G, Di Domenico L, Sabbatini CE, et al. Underdetection of cases of COVID-19 in France threatens epidemic control. Nature New Biol 2021 Feb 4;590(7844):134-139. [doi: 10.1038/s41586-020-03095-6]
35. Coston A, Guha N, Ouyang D, Lu L, Chouldechova A, Ho DE. Leveraging administrative data for bias audits: assessing disparate coverage with mobility data for COVID-19 policy. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency FAccT '21: Association for Computing Machinery; 2021:173-184. [doi: 10.1145/3442188.3445881]
36. Fox R. "What about bias in your dataset?": Quantifying sampling bias in SafeGraph Patterns. URL: https://medium.com/safegraph/what-about-bias-in-the-safegraph-dataset-b95b7208893c [accessed 2025-02-10]
37. Xia J, Yin K, Yue Y, et al. Impact of human mobility on COVID-19 transmission according to mobility distance, location, and demographic factors in the Greater Bay Area of China: population-based study. JMIR Public Health Surveill 2023 Apr 26;9:e39588. [doi: 10.2196/39588] [Medline: 36848228]
38. Baker RE, Mahmud AS, Miller IF, et al. Infectious disease in an era of global change. Nat Rev Microbiol 2022 Apr;20(4):193-205. [doi: 10.1038/s41579-021-00639-z] [Medline: 34646006]

Abbreviations

CDC: Centers for Disease Control and Prevention
HHS: Health and Human Services

Edited by A Mavragani; submitted 30.07.24; peer-reviewed by C Argyropoulos, M Mazzoli; revised version received 09.12.24; accepted 25.12.24; published 18.02.25.

Please cite as:
Pullano G, Alvarez-Zuzek LG, Colizza V, Bansal S
Characterizing US Spatial Connectivity and Implications for Geographical Disease Dynamics and Metapopulation Modeling: Longitudinal Observational Study
JMIR Public Health Surveill 2025;11:e64914
URL: https://publichealth.jmir.org/2025/1/e64914
doi:10.2196/64914

© Giulia Pullano, Lucila Gisele Alvarez-Zuzek, Vittoria Colizza, Shweta Bansal. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Behavioral and Demographic Profiles of HIV Transmission and Exposure Networks in Florida: Network Analysis of HIV Contact Tracing Data

Yiyang Liu[1*], MPH, PhD; Christina Parisi[1*], MPH, PhD; Chaoyue Sun[2], MSc; Rebecca Fisk-Hoffman[1], MPH, PhD; Marco Salemi[3,4], PhD; Diego Viteri[1], MSc; Brandi Danforth[5], MPH; Mattia Prosperi[1,4], MEng, PhD; Simone Marini[1,4], MSc, PhD

[1]Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, 2004 Mowry Road, Gainesville, FL, United States

[2]Department of Electrical and Computer Engineering, College of Engineering, University of Florida, Gainesville, FL, United States

[3]Department of Pathology, Immunology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, FL, United States

[4]Emerging Pathogens Institute, University of Florida, Gainesville, FL, United States

[5]STD Section, Florida Department of Health, Tallahassee, FL, United States

[*]these authors contributed equally

Corresponding Author:
Simone Marini, MSc, PhD
Department of Epidemiology, College of Public Health and Health Professions and College of Medicine, University of Florida, 2004 Mowry Road, Gainesville, FL, United States

Abstract

Background: To complete the Ending the HIV Epidemic initiative in areas with high HIV incidence, there needs to be a greater understanding of the demographic, behavioral, and geographic factors that influence the rate of new HIV diagnoses. This information will aid the creation of targeted prevention and intervention efforts.

Objective: This study aims to identify the geographic distribution of risk groups and their role within potential transmission networks in Florida.

Methods: Public data from the Florida Department of Health and behavioral data from the Surveillance Tools and Reporting System between 2012 and 2022 were used in these analyses. We analyzed records as a combination of variables of interest (gender, age, race or ethnicity, and HIV risk group) to create demographic-behavioral profiles (DBPs) that represent the profiles of people newly diagnosed with HIV. We then used the resulting DBPs to characterize Florida counties and HIV coordination areas and calculated the county-to-county and area-to-area rank (Spearman) correlation. We then drew a dendrogram based on the correlation matrix and identified clusters of similar counties and areas. Lastly, network analysis used HIV contact tracing data from the Surveillance Tools and Reporting System to identify HIV transmission and exposure contact networks and characterized large networks by DBPs and geolocation.

Results: We identified 37 DBPs. The largest DBPs were Hispanic and non-Hispanic Black males aged 25 - 49 reporting male-to-male sexual contact (n=7539 and n=4329, respectively), non-Hispanic White males aged 25 - 49 reporting male-to-male sexual contact (n=4221), and non-Hispanic Black females aged 25 - 49 reporting heterosexual contact (n=3371). The state could be broken up generally into 2 transmission and exposure clusters by region: Northwestern or Northern and Central or Southern. We identified several counties with similar DBPs that were not in the same HIV coordination area. A total of 3097 contact networks were identified among 7944 people with HIV contact tracing data. Most (n=2508, 81%) networks involve only 2 people, 11% (n=349) involve 3 people, 7% (n=224) involve 4 to 19 people, and 6 networks involve 20 or more people. As network size increases, the proportion of people within the network who identify as female, non-Hispanic Black, aged older than 50 years, and exposed to HIV via heterosexual contact decreases.

Conclusions: We identified distinct risk groups and clusters of transmission and exposure throughout Florida. These results can help regions identify health disparities and allocate their HIV prevention and intervention resources accordingly. The goal of this work was to highlight areas of need in a high-incidence setting, not to contribute to existing stigma against vulnerable groups, and it is important to consider the ethics and possible harm of advanced methodologies such as contact network analysis when addressing public health problems.

(JMIR Public Health Surveill 2025;11:e65573) doi:10.2196/65573

XSL•FO
RenderX

KEYWORDS

population health; demographic-behavioral profiles; HIV; HIV prevention; Florida; Surveillance Tools and Reporting System; STARS; United States; public health; health informatics; infectious disease

Introduction

The burden of HIV has long been felt disproportionately across gender, sexual, and racial or ethnic minority groups in the United States. Historically, men who have sex with men (MSM) have had the highest risk of acquiring HIV and made up a majority (69%) of new HIV diagnoses in 2019 [1]. Among MSM, non-Hispanic Black MSM account for a large proportion of new diagnoses. Furthermore, many non-Hispanic Black and Hispanic MSM are unaware of their HIV status, which can lead to further transmission [1]. While the number of new diagnoses among MSM has generally decreased, in 2019 the number of diagnoses among this group aged 55 years or older increased by 5% [1], so there is heterogeneity among these groups by demographic and behavioral factors. Women make up nearly 20% of new HIV diagnoses, and a majority of transmissions among women are due to heterosexual contact (hetero) [1,2]. Among women, nearly 60% of new diagnoses are among non-Hispanic Black women [1,3]. While most transmissions of HIV are due to sexual contact, injection drug use (IDU) accounts for about 10% of new HIV diagnoses [1]. There is great variability in the demographic composition and behavioral risk factors for HIV, and this variability is even more apparent when considering geographic differences.

The Southern United States experiences higher rates of HIV diagnoses and worse HIV-related outcomes compared with the rest of the nation [1,4,5]. Within the South, Florida has consistently been among the states with the highest HIV incidence for the past decade, and ranked third in new diagnoses in 2019 [1]. Florida also experiences high rates of poverty and low rates of health insurance coverage, and has a fluctuating population due to migration, seasonal residents, and tourism. As is the case nationwide, new HIV diagnoses in Florida disproportionately affect Hispanic and non-Hispanic Black MSM and non-Hispanic Black women [6-8]. In 2019, a total of 7 counties in Florida were selected as geographic priority areas in phase 1 of the Ending the HIV Epidemic (EHE) initiative [9,10]. These counties have a high incidence and prevalence of HIV and include some of the most populous in the state: Miami-Dade, Broward, Palm Beach, Orange, Duval, Pinellas, and Hillsborough. The counties highlighted in the EHE initiative are largely urban, however, rural areas in Florida and the rest of the Southern United States have relatively high HIV prevalence and are disproportionately impacted by HIV [11-13]. Therefore, intervention efforts need to extend beyond the 7 selected counties. Compared with urban areas, rural areas of Florida differ demographically, and past research has found that there are demographic differences between rural and urban areas in Florida in terms of those diagnosed with HIV/AIDS [12,14]. There are also different structural and social barriers that can prevent people from receiving HIV care in urban compared with rural areas. Additionally, stigma and discrimination due to HIV and racial or ethnic, gender, and sexual minority groups can vary by location, within groups, and across cultures, and be substantial barriers to knowledge about HIV, engagement in HIV prevention behaviors, and receipt of HIV health care [15].

There are many effective evidence-based strategies for reducing HIV incidence, such as pre-exposure prophylaxis (PrEP) for people in high-risk social networks [16,17]. Additionally, implementing routine screenings among high-risk groups can aid in identifying HIV infection at an earlier stage, which may lead to the timely initiation of treatment and lower the risk of HIV transmission [18-20]. However, the effectiveness of these prevention efforts is limited by resource availability. Further, these interventions must be targeted and require a thorough assessment of the at-risk population. Although the incidence of HIV has gradually declined within Florida and nationwide in the past decade, new and effectively targeted interventions are needed to achieve the EHE goal of reducing HIV diagnoses by 90% by 2030 [9]. The HIV epidemic impacts people from a wide variety of areas and backgrounds, so a one-size-fits-all approach is unlikely to be effective in meeting these goals. There is a need to identify and understand the diversity of factors that can influence the impact of HIV prevention efforts to create informed and impactful interventions.

Past research has identified clusters of HIV risk groups to reduce transmission and inform targeted interventions [21-23]. Predicting the trajectory of HIV transmission clusters is a pivotal step for both prevention and treatment interventions. In this study, we aimed to use a retrospective public health dashboard and surveillance data to characterize people newly diagnosed with HIV by their demographics, behavioral profile, geolocation clusters, and contact network. Additionally, this study serves as proof of concept for leveraging HIV contact tracing data from surveillance tools to enhance insights into HIV transmission dynamics.

Methods

Data Source

We downloaded a public dashboard dataset from the Florida Department of Health (FDOH) through publicly available FLHealthCHARTS. The dataset included all documented new HIV diagnoses between January 1, 2012, and June 30, 2022. Records including diagnoses among those aged 12 years or younger were excluded.

The Surveillance Tools and Reporting System (STARS), formerly known as PRISM (Patient Reporting Investigation Surveillance Manager), is managed by the FDOH. STARS contains relational data collected during contact tracing interviews, where people newly diagnosed with HIV identify recent partners, both through sexual contact and needle-sharing, who are at higher risk of HIV. Information on both people with HIV and their reported contacts is recorded in STARS. These contact tracing interviews are conducted by disease intervention specialists (DIS), who also obtain additional information on demographics and risk behaviors. Partners who are identified

XSL•FO
RenderX

are then contacted by DIS to encourage HIV testing, and, if positive, are interviewed again to identify contacts. In general, DIS successfully contacted 90% of people newly diagnosed with HIV and contacted 65% - 70% of the named partners in Florida. STARS data elements used in the current analyses include individual ID and ID from their linked partners, demographic information (age, gender, or race or ethnicity), county of residence, and HIV risk information (male-to-male sexual contact [MMSC], hetero, IDU, etc). To extract data from STARS for network analysis, we retained only records for individuals that could be attributed to a single, specific demographic-behavioral profile (DBP; ie, without missing DBP attributes), and had at least 1 connection with another individual that could also be attributed to a single, specific DBP. We filtered out the records reporting modes of exposure not consistent with the DBPs, such as blood transfusions or prenatal exposures. The remaining modes of exposure matched the modes of exposure in the FDOH public dataset. If an individual showed multiple records over time, we used the median age from these records. Unlike the FDOH public dataset, STARS reported combinations of race and ethnicity. To be able to apply the DBPs to STARS, all people who self-identified as Hispanic were classified as Hispanic, regardless of race. Those who identified as non-Hispanic were further classified based on the race reported in STARS.

Ethical Considerations

The authors abide by the Declaration of Helsinki. This study protocol was approved by the University of Florida's Institutional Review Board (IRB) and by FDOH's IRB (protocol #IRB201901041 and #2020 - 069, respectively) as exempt. We received data extracts from FDOH's STARS in a fully deidentified format according to the HIPAA (Health Insurance Portability and Accountability Act).

Data Availability

The dashboard data used in this study are publicly available at FLHealthCHART.gov. The STARS data are managed by FDOH. Researchers seeking access for replication purposes need to submit a data request to the FDOH (research@flhealth.gov) by state, federal regulations, and required ethical and privacy policies, including IRB approval by FDOH and execution of a data user agreement. Requests are independently reviewed by FDOH.

Data Analyses

We identified a total of 4 demographic and behavioral variables common in both FDOH public data and STARS and used their combinations to create DBPs. The four variables are (1) gender: male (M), female (F); (2) age: <25, 25 - 49, 50+ years; (3) race or ethnicity: Hispanic, non-Hispanic Black, non-Hispanic White, and non-Hispanic other; and (4) HIV risk group (ie, mode of exposure): MMSC, hetero, IDU, MMSC or IDU, and other. We explored the option of extracting DBPs via unsupervised clustering. Both elbow and silhouette analyses indicated 8 as the suggested number of clusters (data not shown). With the number of clusters (k=8), both agglomerative and divisive clustering approaches based on Gower distance split the data into groups perfectly and uniformly characterized by gender and race or ethnicity (ie, clusters where all the members belong to the same combination of gender and race or ethnicity). When using k>8, data were further partitioned by age (data not shown). However, none of the profile identities from the unsupervised clustering methods include the only behavioral factor (ie, HIV risk group), which we believe is important to include as it is a modifiable factor. Therefore, we decided to extract DBPs based on their frequency, as explained above, instead of using unsupervised clustering.

Descriptive analysis was used to demonstrate the frequencies for identified DBPs and our study sample characteristics. Multiple correspondence analysis (MCA) was implemented using the FactoMineR and factoextra libraries in R Studio (R Foundation). MCA is an extension of simple correspondence analysis and is used to summarize and visualize data tables with more than 2 categorical variables. This analysis was used to show a graphical representation of the similarity between identified DBPs.

Additionally, we used the identified DBPs to characterize Florida counties and HIV coordination areas. HIV coordination areas ("areas") are groups of counties that operate under the leadership of an HIV/AIDS program coordinator who oversees the region's HIV health care and prevention efforts within the FDOH [24]. Each county and area can be represented by an array of 37 slots, each representing the counts for a specific patient type. To compare counties and areas, we calculated the county-to-county and area-to-area rank (Spearman) correlation. We then drew the dendrogram based on the correlation matrix and found groups of similar counties and areas.

Lastly, we performed network analysis using STARS data based on the DBPs identified. The Leiden algorithm [25] is a network node partitioning method that focuses on identifying clusters characterized by maximized intracluster connections and minimized intercluster connections. In this study, the algorithm was implemented using the *leidenalg* library in Python (Python Software Foundation) [26], with modularity serving as the optimization function and setting the resolution parameter to 0.2 to control cluster counts.

Results

Sample Characteristics and Identified DBPs

Sample characteristics are shown in Table 1 stratified by data source. The Florida Chart public data and STARS data show similar sample characteristics: mostly male, aged 25 - 49 years, exposed to HIV via MMSC. Non-Hispanic Black is the most common race or ethnicity group in both data sources.

Using the Florida public data, we identified 84 different DBPs among 44,090 people newly diagnosed with HIV between 2012 and 2022. Of these 84 DBPs, 36 cover 97% of the HIV incident cases (42,734/44,090), and each included 100 or more individuals. The remaining DBPs were combined into an "other" group, making 37 DBPs total. These DBPs are reported in Table 2. The results from the MCA, showing a graphical representation of the similarity between these DBPs, are provided in Figure S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Sociodemographic comparison between the FDOH[a] public dataset and STARS[b]. Numerical description of the datasets.

	Florida chart public data (n=44,090), n (%)	STARS data (n=7944), n (%)
Gender		
Female	9526 (21.6)	1530 (19.3)
Male	34,564 (74.4)	6414 (80.7)
Race or ethnicity		
Hispanic	14,043 (31.9)	2039 (25.7)
Non-Hispanic Black	18,343 (41.6)	3839 (48.3)
Non-Hispanic White	10,709 (24.3)	1866 (23.5)
Non-Hispanic other	995 (2.3)	200 (2.5)
Age group (years)		
<25	7654 (17.4)	1695 (21.3)
25 - 49	26,925 (61.1)	5908 (74.4)
50+	9511 (21.6)	1151 (14.5)
HIV risk group		
Hetero[c]	14,918 (33.8)	2422 (30.5)
MMSC[d]	26,498 (60.1)	4919 (61.9)
MMSC or IDU[e]	898 (2)	291 (3.7)
IDU	1718 (3.9)	312 (3.9)
Other	58 (0.1)	0 (0)

[a]FDOH: Florida Department of Health.

[b]STARS: Surveillance Tools and Reporting System.

[c]Hetero: heterosexual contact.

[d]MMSC: male-to-male sexual contact.

[e]IDU: injection drug use.

XSL•FO
RenderX

Table . Thirty-six DBPs[a] were extracted from the FDOH[b] public dataset among people newly diagnosed with HIV between 2012 and 2023, covering 97% of cases. A 37th "other" DBP included the remaining cases.

DBP	Gender	Ethnicity or race	Age group (years)	Exposure	Cases, n
1	Male	Hispanic	25 - 49	MMSC[c]	7539
2	Male	Black Non-Hispanic	25 - 49	MMSC	4329
3	Male	White Non-Hispanic	25 - 49	MMSC	4221
4	Female	Black Non-Hispanic	25 - 49	Hetero[d]	3371
5	Male	Black Non-Hispanic	<25	MMSC	3024
6	Male	Black Non-Hispanic	25 - 49	Hetero	2204
7	Male	White Non-Hispanic	50+	MMSC	2066
8	Male	Hispanic	<25	MMSC	1744
9	Female	Black Non-Hispanic	50+	Hetero	1652
10	Male	Hispanic	50+	MMSC	1404
11	Male	Black Non-Hispanic	50+	Hetero	1400
12	Female	Hispanic	25 - 49	Hetero	939
13	Male	White Non-Hispanic	<25	MMSC	870
14	Male	Hispanic	25 - 49	Hetero	831
15	Female	Black Non-Hispanic	<25	Hetero	748
16	Female	White Non-Hispanic	25 - 49	Hetero	690
17	Male	Black Non-Hispanic	50+	MMSC	641
18	Male	Other Non-Hispanic	25 - 49	MMSC	444
19	Male	White Non-Hispanic	25 - 49	Hetero	444
20	Female	Hispanic	50+	Hetero	431
21	Male	Hispanic	50+	Hetero	418
22	Male	White Non-Hispanic	25 - 49	MMSC or IDU[e]	391
23	Female	White Non-Hispanic	25 - 49	IDU	380
24	Female	White Non-Hispanic	50+	Hetero	377
25	Male	White Non-Hispanic	50+	Hetero	370
26	Male	Black Non-Hispanic	<25	Hetero	333
27	Male	White Non-Hispanic	25 - 49	IDU	303
28	Female	Hispanic	<25	Hetero	165
29	Male	Other Non-Hispanic	<25	MMSC	155
30	Female	White Non-Hispanic	<25	Hetero	144
31	Male	Hispanic	25 - 49	MMSC or IDU	130
32	Male	Black Non-Hispanic	25 - 49	IDU	123
33	Male	Hispanic	25 - 49	IDU	122
34	Male	Black Non-Hispanic	50+	IDU	119
35	Male	Black Non-Hispanic	25 - 49	MMSC or IDU	111
36	Female	Other Non-Hispanic	25 - 49	Hetero	101

[a]DBP: demographic-behavioral profile.

[b]FDOH: Florida Department of Health.

[c]MMSC: male-to-male sexual contact.

[d]Hetero: heterosexual contact.

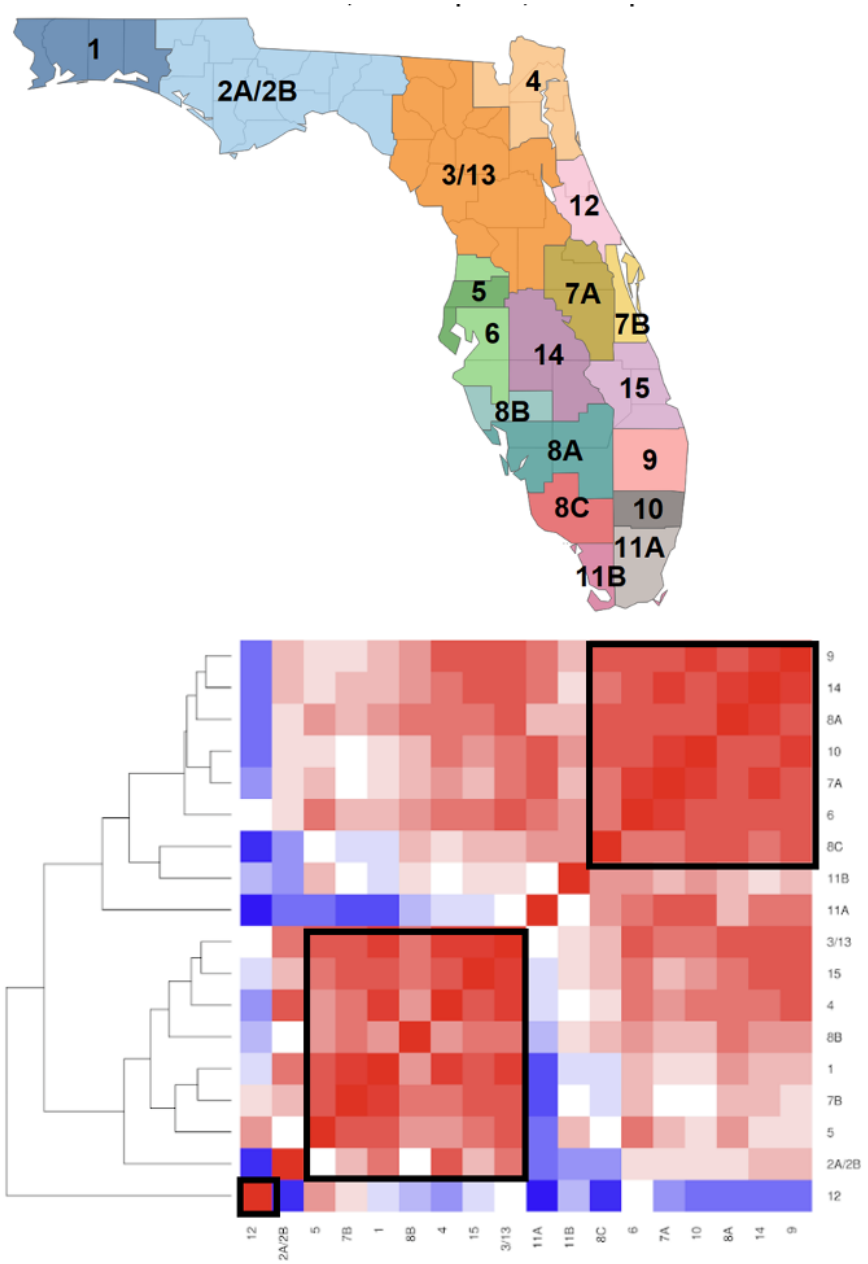
[e]IDU: injection drug use

XSL•FO
RenderX

Characterization of County and Area Similarity Based on the DBPs

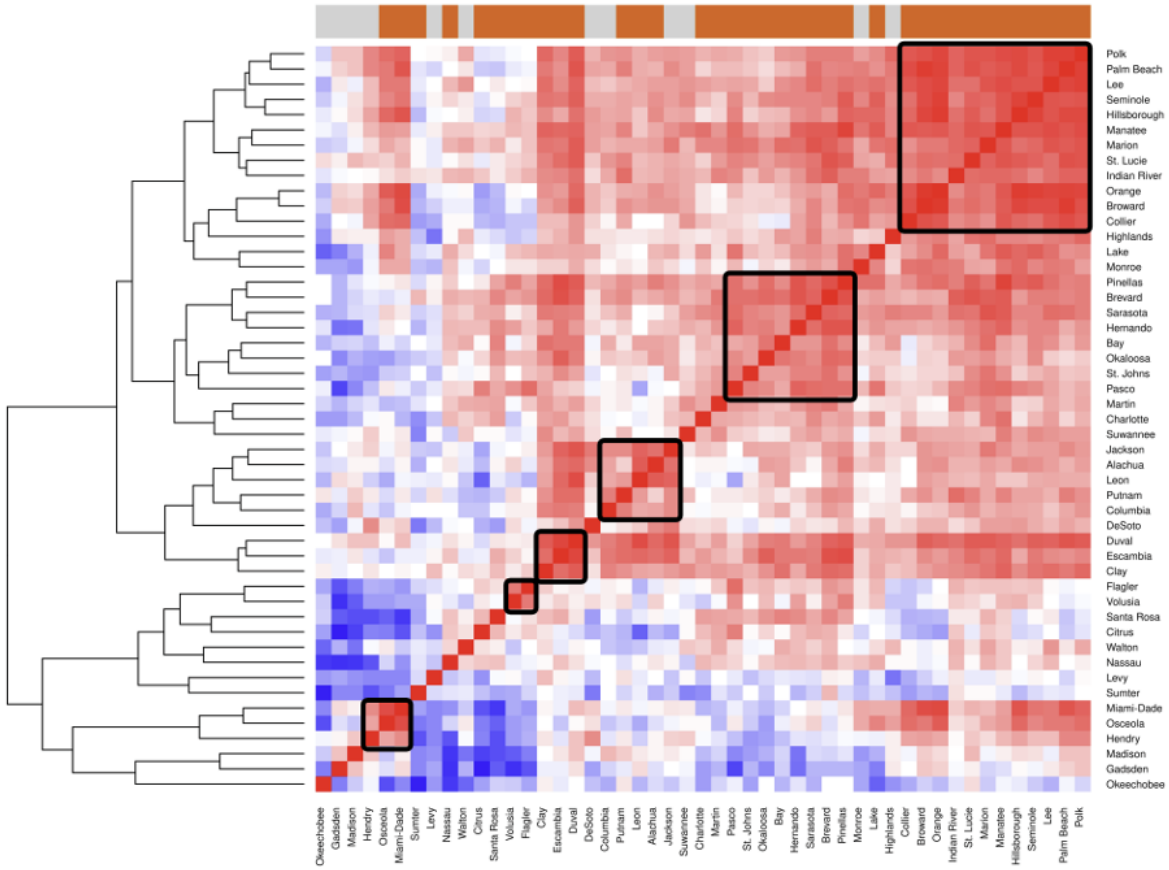
The dendrogram showing the correlation matrix between county and HIV coordination areas in Florida using the identified DBPs, is depicted in Figures 1 and 2. Many HIV areas share similar distributions of DBPs and predominant DBP types, and 2 main clusters were identified. The first area cluster included many counties in the western part of Central Florida and Southern Florida (areas 9, 14, 8A, 10, 7A, and 6) and the second represented counties in Northern Florida, the Panhandle (Northwestern Florida), and the east coast of Central Florida (areas 5, 7B, 1 8B, 4, 15, and 3/13). Area 12 (Flagler and Volusia) was dissimilar from most of the other areas. Area 11A (Miami-Dade) was similar to the Central or Southern Florida cluster but very dissimilar from the Northern Florida or Northwestern Florida cluster. Among the counties, rural counties tended to be more similar but there was no clear delineation between rural and urban counties. We identified 6 clusters of varying sizes. The largest included 12 counties: Polk, Palm Beach, Lee, Seminole, Hillsborough, Manatee, Marion, St. Lucie, Indian River, Orange, Broward, and Collier. The next largest cluster included 8 counties: Pinellas, Brevard, Sarasota, Hernando, Bay, Okaloosa, St. Johns, and Pasco. The remaining clusters included between 6 and 2 counties: Jackson, Alachua, Leon, Putnam, and Columbia; Duval, Escambia, and Clay; Flagler and Volusia; and Miami-Dade, Osceola, and Hendry. The Duval, Escambia, and Clay cluster were also similar to the 2 largest clusters, while the 1 containing Miami-Dade also shared many similarities with the counties in the largest cluster.

Figure 1. Area similarity based on the sociodemographic profiles. Top panel: map of HIV coordination areas in Florida. Bottom panel: heatmap of the area similarity. In the heatmap, each area is compared to the others per DBPs, that is, with higher similarity shifting toward red, and lower similarity shifting toward blue. Of note, areas are arranged via a data-driven process (hclust) so that similar ones will be closer together, as indicated by the dendrogram on the right. Visual inspection reveals 3 clusters, indicated by the black squares. DBP: demographic-behavioral profile.



XSL•FO
RenderX

Figure 2. Heatmap of county similarity based on the sociodemographic profiles and by rural or urban (bar on top of figure: orange-urban and gray-rural). Each county is compared to the others per DBPs, that is, with higher similarity shifting toward red, and lower similarity shifting toward blue. Of note, counties are arranged via a data-driven process (hclust) so that similar ones will be closer together, as indicated by the dendrogram on the right. Visual inspection reveals 6 clusters, indicated by the black squares. DBP: demographic-behavioral profile.



Analysis of HIV Contacts Reveals Mode-of-Exposure Characterizing Both Network Size and Composition

STARS represents contact tracing data to create networks of people with HIV and people exposed to HIV. We extracted 7944 individuals from STARS, sharing 4928 connections. The DBP counts in STARS correlate strongly with the counts of the FDOH public dataset (Spearman correlation=0.93). STARS individuals formed 3097 subnetworks (ie, networks that are not connected to one another). Of those, 2518 (~81%) are 2-node networks, 349 (~11%) are 3-node networks, 224 (~7%) networks have 4 - 19 nodes, while the remaining 6 networks have 20 - 246 nodes (shown in Figure 3). Notably, hetero decreases decisively with the network size, is found more commonly in smaller networks, and its representation decreases with network growth. Table 3 reports the characteristics of the STARS population based on the network size. In the 2 largest contact networks (Figures 3A and 3B), males of Black, Hispanic, and White ethnicities aged 25 - 49 years predominantly constitute both networks. Interestingly, when focusing on hub nodes (nodes with more than 2 links in the network), aside from the dominant groups, young Hispanic and Black males younger than the age of 25 years also comprise a significant portion.

XSL•FO
RenderX

Table . Characteristics of contact networks by number of nodes. Numerical description of extracted networks. For the first 2 rows, the row percentages were calculated. For the rest of the rows, column percentages were calculated.

	Values, n (%)	Values, n (%)	Values, n (%)	Values, n (%)
Number of nodes[a]	2	3	4 to 19	20 to 246
Total networks	2518 (81)	349 (11)	224 (7)	6 (0.2)
Total nodes (individual)	5036 (63)	1047 (13)	1350 (17)	511 (7)
Gender				
Female	1261 (25)	193 (18)	76 (6)	0 (0)
Male	3775 (75)	854 (82)	1274 (94)	511 (100)
Race or ethnicity				
Non-Hispanic Black	2501 (50)	511 (49)	644 (48)	183 (36)
Non-Hispanic White	1157 (23)	268 (26)	327 (24)	114 (22)
Hispanic	1258 (25)	242 (23)	336 (25)	203 (40)
Non-Hispanic other	120 (2)	26 (2)	43 (3)	11 (2)
Age group (years)				
<25	829 (16)	247 (24)	448 (33)	171 (33)
25 - 49	3282 (65)	689 (66)	815 (60)	312 (61)
50+	925 (18)	111 (11)	87 (6)	28 (5)
HIV risk group				
Hetero[b]	2022 (40)	276 (26)	119 (9)	5 (1)
IDU[c]	260 (5)	33 (3)	17 (1)	2 (0)
MMSC[d]	2602 (52)	698 (67)	1140 (84)	479 (94)
MMSC or IDU	152 (3)	40 (4)	74 (5)	25 (5)

[a]This line does not report any percentage and is, in fact, the label for each column, that is, the size in terms of total number of nodes we are describing.

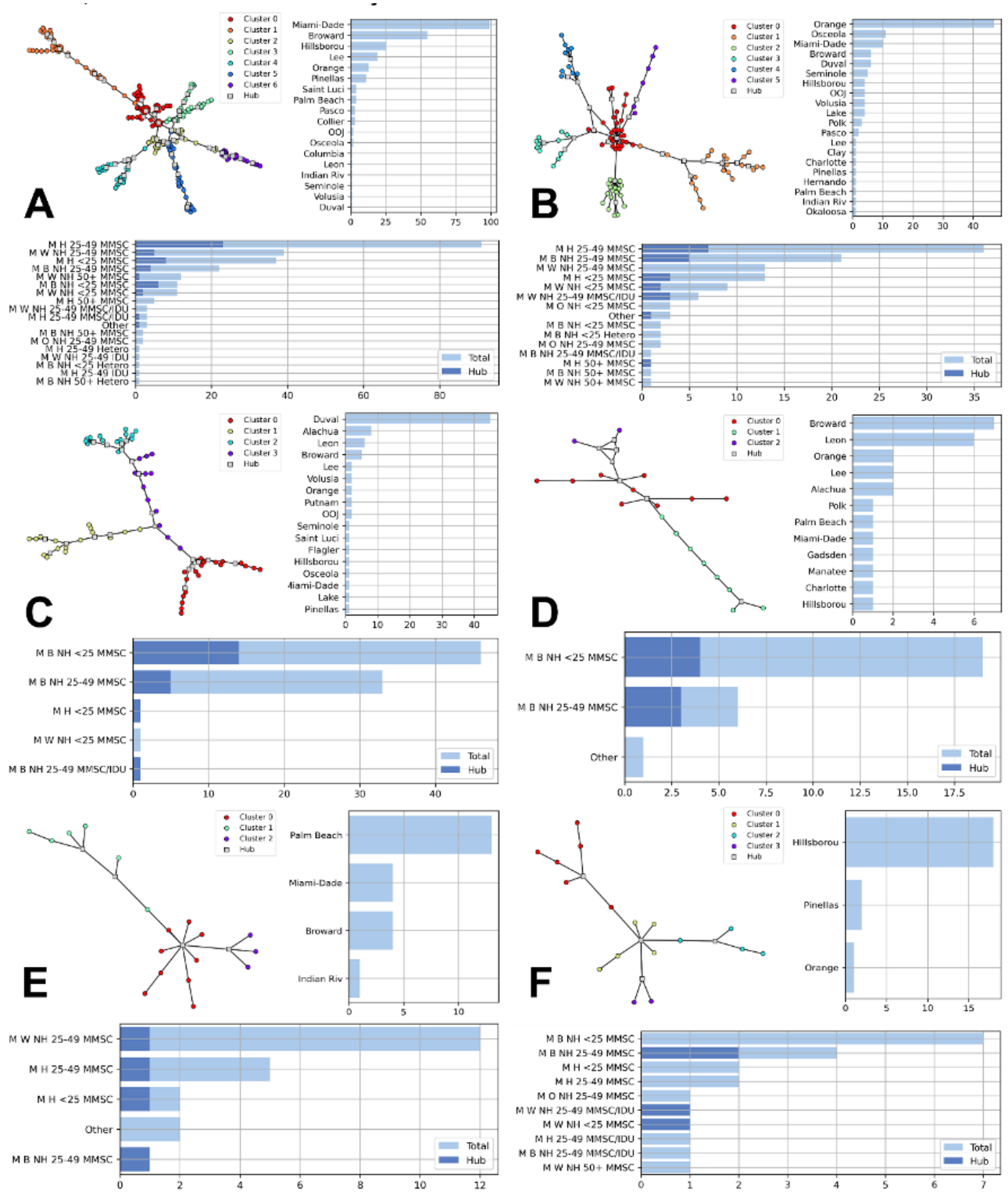
[b]Hetero: heterosexual contact.

[c]IDU: injection drug use.

[d]MMSC: male-to-male sexual contact.

XSL•FO
RenderX

Figure 3. The 6 largest contact networks (20+ nodes), characterized by hub nodes, Leiden Clusters, identified DBPs and county, are represented here ordered on size (largest to smallest), labelled A to F. Each panel represents a network of individuals in our dataset. Nodes (dots) represent individuals, and undirected edges represent connections between individuals based on contact tracing data. Individuals are grouped into clusters using the Leiden algorithm and are colored accordingly—nodes with the same color belong to the same cluster. Notably, this clustering is based solely on the network structure, not on individual characteristics. Some individuals are identified as hubs, defined as nodes with more than 2 connections; these represent individuals with active transmission-related behavior and warrant closer attention. We also provide a graphical description of the aggregated data for each network, shown via bar charts that display the individuals' counties as well as the identified DBPs within the network, including the fraction of each DBP covered by hub individuals. These results highlight potential differences in transmission patterns. B: Black; DBP: demographic-behavioral profile; H; Hispanic; IDU: injection drug use; MMSC: male-to-male sexual contact; NH: non-Hispanic; O: other; W: White.



XSL•FO
RenderX

Discussion

Overview

Using the FDOH public dataset, we characterized DBPs for all documented people newly diagnosed with HIV in Florida between 2012 and 2022. A total of 37 DBPs were identified. The attempt to extract HIV behavioral risk groups via unsupervised clustering stratified data based on demographics only, as the behavioral component (ie, the mode of exposure) was not a discriminating feature even when deliberately overclustering. Therefore, a list of all combinations of variable categories was used to characterize DBP. DBPs were further used to examine area and county similarity across Florida and characterize contact networks. The results from these analyses can be used to inform future HIV prevention and intervention efforts by guiding public health professionals in making decisions on resource allocation and policy for their strategies to end the HIV epidemic. These findings can also potentially help public health workers in the field, allowing them to better understand trends in the HIV epidemic in their areas to identify and help community members most in need. Moreover, these findings show that contact tracing surveillance data can be used to investigate the dynamics of HIV transmission and support further research and development in this area for the creation of more informative models. Integrating these models into routine surveillance would enable regular communication of results to program staff and field workers, ensuring the continued effectiveness of current efforts while also allowing for strategies to be adapted in response to evolving needs.

The 6 largest networks (30+ nodes and 511 individuals) are almost exclusively composed of individuals with MMSC (ie, MSM). While the female, non-Hispanic Black, aged 25 - 49 years, and hetero profile is the fourth largest DBP, none of the large networks involve females of any race or age. This may indicate that HIV transmission to women does not involve large contact networks with many potentially at-risk partners. If this is the case, then interventions that focus solely on increasing prevention efforts among those with multiple sexual partners may still leave many women vulnerable to HIV. The implementation of past prevention efforts has made similar missteps. For example, the original guidance for prescribing PrEP focused largely on and marketed toward MSM and included language that excluded at-risk women. As a consequence, PrEP awareness and uptake have remained low among at-risk women, although the guidelines have been updated to try and address this [27,28].

Those with MMSC comprise the dominant HIV risk group among the large networks identified. This matches national trends, with 67% of new HIV diagnoses in the United States in 2021 involving MMSC [29]. Additionally, in the United States, 79% of new HIV cases were among men, 40% were among Black individuals, 29% among Hispanic individuals, and 37% among those aged 25 - 34 years [29], which closely mirrors the largest transmission networks that we identified in Florida. People with IDU may appear in the network, but no major network is characterized solely by IDU. This might be because IDU is not a common behavior, with less than 2% of the US population reporting IDU in 2018 [30], and ~4% in the extracted STARS DBPs report IDU. However, IDU is still a prominent means of HIV acquisition and transmission, accounting for about 10% of new cases [1], and should still be considered in prevention and intervention efforts. Hetero exposure is also not common in the identified networks, again indicating that exposure via hetero may occur mainly in small networks.

All of the large HIV transmission networks we identified involve individuals from multiple counties. Most of these networks were localized by region or represented multiple urban centers in the state. While neighboring counties are often grouped together into HIV coordination areas to address HIV, not all of these networks are contained in 1 area. Moreover, Florida has a high population who are seasonal residents, migrants, and transient tourists from all over the world. Close communication between county health officials is essential to address the HIV epidemic as these transmission networks extend beyond immediate neighboring counties. One of the goals of the HIV coordination areas is to help facilitate between-county communication, even though some areas are only 1 county (eg, areas 7B and 11A).

We identified 2 distinct clusters of HIV areas that divide the state roughly in half (Northern or Northwestern and Central or Southern) and 6 county clusters. Another level of categorization into 2 larger HIV area groups might also help increase communication between the counties that we identified as being within similar clusters. For example, the largest county cluster, represented counties within areas 6, 7A, 8A, 9, and 15, which are all in the Central or Southern Florida cluster, plus a county in area 14. Improving both interarea and intercounty communication could help improve information sharing between counties with similar DBPs, and information shared could include successful interventions and lessons learned from interventions that were less successful to improve future decision-making and resource allocation. Increasing discussions within these clusters could also improve information sharing between neighboring counties that happen to be in different HIV areas.

There are existing evidence-based and impactful strategies for addressing HIV, such as PrEP [16,17], postexposure prophylaxis [31], condoms [32], and syringe services programs [33,34]. Our findings support increasing the use of these and other prevention strategies in areas and among groups we identified as having high transmission. For example, our findings identified those with MMSC in Miami-Dade as a high-transmission contact network, and past research has found that there is low uptake of PrEP among MSM in Miami-Dade [35]. Additionally, testing efforts among high-risk groups can be increased in geographic areas and among groups identified in this study. Our findings can enable focused and efficient resource allocation, prioritizing efforts and funding for groups at high HIV risk and areas with increased HIV transmission.

When characterizing DBPs in identified contact networks, we intentionally only include these DBPs at an aggregated level due to ethical concerns. Moreover, we would like to emphasize that many of the groups and behaviors that we have identified in these clusters are groups that already experience high levels of stigma and discrimination. The goal of this work is not to

XSL•FO
RenderX

contribute to this existing stigma, but to highlight areas of need and groups who might need targeted intervention to reduce their risk of HIV. There is a continuous need to conduct ethical discussions to avoid the possible harm of advanced methodologies to vulnerable groups, especially in the context of HIV stigmatization [22].

Our results should be interpreted in consideration of some limitations. We removed STARS records without linked partners or who had missing or inconsistent records, so our network analyses were conducted on a subset of STARS data. The extracted STARS DBPs have a high correlation with DBPs identified from the Florida public dataset [13], indicating good representativeness. The data are retrospective, which could limit their usefulness for public health workers in the field if they simply confirm what they already know about the HIV epidemic in their area rather than provide new information. However, these findings support further investigation and application of contact tracing surveillance data to provide a more up-to-date and accurate understanding of transmission dynamics, allowing more targeted and impactful prevention and intervention efforts. A final consideration is that extracting DBPs from the FDOH public dataset via unsupervised clustering leads to clusters solely characterized by nonbehavioral components, such as gender and race or ethnicity, therefore missing the mode of exposure, a key feature to explore the HIV contact networks. A primary strength of this study is our public health data, with high representativeness of people with HIV in the last 10 years in Florida.

Conclusions

In our contact network analysis, we identified several distinct risk groups (young MSM, non-Hispanic Black MSM, Hispanic MSM, Hispanic women, and people with IDU) and clusters of transmission in Florida (divided generally into Northern or Northwestern and Central or Southern, with some similarities identified at the county level). These results can be used to inform targeted interventions aimed at reducing HIV diagnoses in the state. These groups we identified are already recognized as having a disproportionately high risk of HIV in Florida and nationwide, and our findings have shown areas in the state where these disparities might be exacerbated. Inequitable distribution of the burden of HIV among these groups contributes to existing health disparities and shows that the HIV epidemic will not be ended by a one-size-fits-all approach. Network analysis can help state and county health departments not only identify disparities and areas or groups with high transmission but also make the best use of their HIV intervention and prevention resources to help their community members.

Acknowledgments

This work was supported in part by the National Institutes of Health (R01AI145552, R01AI172875, R34MH135768, 1F31AA030518-01, and 1F31AA030733-01). The findings and conclusions in this work are those of the authors and do not necessarily represent the views of the Florida Department of Health (FDOH). Generative artificial intelligence has been used to improve the readability of this paper. All contents have been thoroughly reviewed and verified by the authors before being incorporated into this paper.

Data Availability

The dashboard data used in this study are publicly accessible at FLHealthCHARTS [13]. The STARS data are managed by the Florida Department of Health (FDOH) and are not publicly available. Researchers interested in accessing STARS data for replication or further analysis need to submit a data request to FDOH. Access requires independent review by FDOH.

Authors' Contributions

Conceptualization – SM, MP, MS, YL
Data curation – YL, CS, CP, RFH, DV, SM, BD
Formal analysis, Methodology, Software, Visualization – CS, SM, YL, CP, RFH, DV
Funding acquisition - MP, MS
Project administration – SM, MP, MS, BD
Resources, Investigation – BD, YL, MP
Supervision – SM, MP, MS
Writing – original draft: YL, CS, SM CP, RFH, MP
Writing – review & editing: all authors

Conflicts of Interest

None declared.

Multimedia Appendix 1
Similarity between DBPs from an MCA. DBP: demographic-behavioral profile; MCA: multiple correspondence analysis.
[DOCX File, 274 KB - publichealth_v11i1e65573_app1.docx]

XSL•FO
RenderX

References

1. Diagnoses of HIV infection in the United States and dependent areas, 2019. : Centers for Disease Control and Prevention; 2021.
2. Martin EG, Ansari B, Hart-Malloy R, et al. Racial and ethnic disparities in HIV diagnoses among heterosexually active persons in the United States nationally and by state, 2018. PLOS ONE 2021;16(9):e0257583. [doi: 10.1371/journal.pone.0257583] [Medline: 34543322]
3. Ojikutu BO, Mayer K. HIV prevention among black women in the US-time for multimodal integrated strategies. JAMA Netw Open 2021 Apr 1;4(4):e215356. [doi: 10.1001/jamanetworkopen.2021.5356] [Medline: 33835180]
4. Gant Z, Dailey A, Wang S, et al. Trends in HIV care outcomes among adults and adolescents in the U.S. South, 2015-2019. Ann Epidemiol 2022 Jul;71:15-22. [doi: 10.1016/j.annepidem.2022.04.011] [Medline: 35525379]
5. Prejean J, Tang T, Hall HI. HIV diagnoses and prevalence in the southern region of the United States, 2007-2010. J Community Health 2013 Jun;38(3):414-426. [doi: 10.1007/s10900-012-9633-1] [Medline: 23179388]
6. Epidemiological profile: continuum of HIV care by county. : Florida Department of Health; 2017.
7. Lieb S, White S, Grigg BL, Thompson DR, Liberti TM, Fallon SJ. Estimated HIV incidence, prevalence, and mortality rates among racial/ethnic populations of men who have sex with men, Florida. J Acquir Immune Defic Syndr 2010 Aug;54(4):398-405. [doi: 10.1097/QAI.0b013e3181d0c165] [Medline: 20182358]
8. Wright IA, Reid R, Shahid N, et al. Neighborhood characteristics, intersectional discrimination, mental health, and HIV outcomes among black women living with HIV, southeastern United States, 2019 2020. Am J Public Health 2022 Jun;112(S4):S433-S443. [doi: 10.2105/AJPH.2021.306675] [Medline: 35763751]
9. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV Epidemic: a plan for the United States. JAMA 2019 Mar 5;321(9):844-845. [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
10. Florida's unified Ending the HIV Epidemic plan (Division of Disease Control and Health Protection Bureau of Communicable Diseases, HIV/AIDS section). : Florida Department of Health; 2020.
11. Watson M, Johnson SD, Zhang T, Oster AM. Characteristics of and trends in HIV diagnoses in the deep south region of the United States, 2012-2017. AIDS Behav 2019 Oct;23(Suppl 3):224-232. [doi: 10.1007/s10461-019-02659-6] [Medline: 31473847]
12. Trepka MJ, Fennie KP, Sheehan DM, Lutfi K, Maddox L, Lieb S. Late HIV diagnosis: differences by rural/urban residence, Florida, 2007-2011. AIDS Patient Care STDS 2014 Apr;28(4):188-197. [doi: 10.1089/apc.2013.0362] [Medline: 24660767]
13. Persons with HIV (PWH). FLHealthCHARTS. URL: https://www.flhealthcharts.gov/ChartsDashboards/rdPage.aspx?rdReport=HIVAIDS.DataViewer&cid=9866 [accessed 2025-04-19]
14. Trepka MJ, Niyonsenga T, Maddox LM, Lieb S. Rural AIDS diagnoses in Florida: changing demographics and factors associated with survival. J Rural Health 2013 Jun;29(3):266-280. [doi: 10.1111/j.1748-0361.2012.00449.x] [Medline: 23802929]
15. Tan RKJ, Tang W, Tucker JD. Public health services and intersectional stigma: a social sciences perspective with implications for HIV service design and delivery. Curr Opin HIV AIDS 2023 Jan 1;18(1):18-26. [doi: 10.1097/COH.0000000000000769] [Medline: 36444656]
16. Fonner VA, Dalglish SL, Kennedy CE, et al. Effectiveness and safety of oral HIV preexposure prophylaxis for all populations. AIDS 2016 Jul 31;30(12):1973-1983. [doi: 10.1097/QAD.0000000000001145] [Medline: 27149090]
17. Spinner CD, Boesecke C, Zink A, et al. HIV pre-exposure prophylaxis (PrEP): a review of current knowledge of oral systemic HIV PrEP in humans. Infection 2016 Apr;44(2):151-158. [doi: 10.1007/s15010-015-0850-2] [Medline: 26471511]
18. Valentine SS, Caldwell J, Tailor A. Effect of CDC 2006 revised HIV testing recommendations for adults, adolescents, pregnant women, and newborns on state laws, 2018. Public Health Rep 2020;135(1_suppl):189S-196S. [doi: 10.1177/0033354920930146] [Medline: 32735201]
19. The role of testing in HIV prevention. Lancet HIV 2017 May;4(5):e189. [doi: 10.1016/S2352-3018(17)30076-0] [Medline: 28460691]
20. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011 Aug 11;365(6):493-505. [doi: 10.1056/NEJMoa1105243] [Medline: 21767103]
21. Wilbourn B, Saafir-Callaway B, Jair K, et al. Characterization of HIV risk behaviors and clusters using HIV-transmission cluster engine among a cohort of persons living with HIV in Washington, DC. AIDS Res Hum Retroviruses 2021 Sep;37(9):706-715. [doi: 10.1089/AID.2021.0031] [Medline: 34157853]
22. Rich SN, Prosperi MCF, Dellicour S, et al. Molecular epidemiology of HIV-1 subtype b infection across Florida reveals few large superclusters with metropolitan origin. Microbiol Spectr 2022 Dec 21;10(6):e0188922. [doi: 10.1128/spectrum.01889-22] [Medline: 36222706]
23. Rich SN, Poschman K, Hu H, et al. Sociodemographic, ecological, and spatiotemporal factors associated with human immunodeficiency virus drug resistance in Florida: a retrospective analysis. J Infect Dis 2021 Mar 3;223(5):866-875. [doi: 10.1093/infdis/jiaa413] [Medline: 32644119]
24. State of Florida Integrated HIV Prevention and Care Plan, 2022-2026. : Florida Department of Health; 2023.
25. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep 2019 Mar 26;9(1):5233. [doi: 10.1038/s41598-019-41695-z] [Medline: 30914743]

XSL•FO
RenderX

26. Leidenalg 0.10.2. Python Package Index (PyPI). URL: https://pypi.org/project/leidenalg/ [accessed 2025-05-08]
27. Goparaju L, Praschan NC, Warren-Jeanpiere L, Experton LS, Young MA, Kassaye S. Stigma, partners, providers and costs: potential barriers to PrEP uptake among US women. J AIDS Clin Res 2017 Sep;8(9):730. [doi: 10.4172/2155-6113.1000730] [Medline: 29201531]
28. Auerbach JD, Kinsky S, Brown G, Charles V. Knowledge, attitudes, and likelihood of pre-exposure prophylaxis (PrEP) use among US women at risk of acquiring HIV. AIDS Patient Care STDS 2015 Feb;29(2):102-110. [doi: 10.1089/apc.2014.0142] [Medline: 25513954]
29. HIV surveillance report, 2021. : Centers for Disease Control and Prevention; 2023 URL: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2023-09-07]
30. Bradley H, Hall EW, Asher A, et al. Estimated number of people who inject drugs in the United States. Clin Infect Dis 2023 Jan 6;76(1):96-102. [doi: 10.1093/cid/ciac543] [Medline: 35791261]
31. DeHaan E, McGowan JP, Fine SM, et al. PEP to Prevent HIV Infection: Johns Hopkins University; 2022.
32. Pinkerton SD, Abramson PR. Effectiveness of condoms in preventing HIV transmission. Soc Sci Med 1997 May;44(9):1303-1312. [doi: 10.1016/S0277-9536(96)00258-4]
33. Vlahov D, Junge B. The role of needle exchange programs in HIV prevention. Public Health Rep 1998 Jun;113(Suppl 1):75-80. [Medline: 9722812]
34. Broz D, Carnes N, Chapin-Bardales J, et al. Syringe services programs' role in Ending the HIV Epidemic in the U.S.: why we cannot do it without them. Am J Prev Med 2021 Nov;61(5 Suppl 1):S118-S129. [doi: 10.1016/j.amepre.2021.05.044] [Medline: 34686281]
35. Gonzales-Zamora JA, Ponce-Rosas L, Martinez R. Determinants of public health and interventions to address HIV infection among men who have sex with men in Miami-Dade County, Florida, USA. Infez Med 2022;30(3):392-402. [doi: 10.53854/liim-3003-7] [Medline: 36148172]

Abbreviations

DBP: demographic-behavioral profile
DIS: disease intervention specialist
EHE: Ending the HIV Epidemic
FDOH: Florida Department of Health
hetero: heterosexual contact
HIPAA: Health Insurance Portability and Accountability Act
IDU: injection drug use
IRB: institutional review board
MCA: multiple correspondence analysis
MMSC: male-to-male sexual contact
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
PRISM: Patient Reporting Investigation Surveillance Manager
STARS: Surveillance Tools and Reporting System

Edited by A Mavragani; submitted 19.08.24; peer-reviewed by D Laborde, T Finlayson; revised version received 27.02.25; accepted 04.03.25; published 25.06.25.

Please cite as:
Liu Y, Parisi C, Sun C, Fisk-Hoffman R, Salemi M, Viteri D, Danforth B, Prosperi M, Marini S
Behavioral and Demographic Profiles of HIV Transmission and Exposure Networks in Florida: Network Analysis of HIV Contact Tracing Data
JMIR Public Health Surveill 2025;11:e65573
URL: https://publichealth.jmir.org/2025/1/e65573
doi:10.2196/65573

© Yiyang Liu, Christina Parisi, Chaoyue Sun, Rebecca Fisk-Hoffman, Marco Salemi, Diego Viteri, Brandi Danforth, Mattia Prosperi, Simone Marini. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 25.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

COVID-19–Relevant Insights Into the Elevated Risk of Accidental Injuries in Survivors of SARS and Their Relatives in Taiwan: Retrospective Cohort Study

Chieh Sung[1,2*], PhD; Chi-Hsiang Chung[3], PhD; Chien-An Sun[4,5], PhD; Chang-Huei Tsao[2,6], PhD; Daphne Yih Ng[7], MS; Tsu-Hsuan Weng[2,8], MS; Li-Yun Fann[9,10], PhD; Fu-Huang Lin[3], PhD; Wu-Chien Chien[2,3,8*], PhD

[1]Department of Senior Citizen Care and Welfare, Deh Yu College of Nursing and Health, Keelung City, Taiwan

[10]Department of Nursing, Tzu Chi University, Hualien, Taiwan

[2]Department of Medical Research, Tri-Service General Hospital, National Defense Medical Center, No. 325, Section 2, Chenggong Road, Taipei, Taiwan

[3]School of Public Health, National Defense Medical Center, Taipei, Taiwan

[4]Department of Public Health, College of Medicine, Fu-Jen Catholic University, New Taipei City, Taiwan

[5]Big Data Research Center, College of Medicine, Fu-Jen Catholic University, New Taipei City, Taiwan

[6]Department of Microbiology & Immunology, National Defense Medical Center, National Defense Medical Center, Taipei, Taiwan

[7]Department of Family Medicine, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan

[8]Graduate Institute of Medical Sciences, National Defense Medical Center, Taipei, Taiwan

[9]Department of Nursing, Taipei City Hospital, 145 Zhengzhou Road, Datong District, Taipei, Taiwan

[*]these authors contributed equally

Corresponding Author:
Li-Yun Fann, PhD
Department of Nursing, Taipei City Hospital, 145 Zhengzhou Road, Datong District, Taipei, Taiwan

Abstract

Background: The 2003 outbreak of severe acute respiratory syndrome (SARS), caused by a novel coronavirus, heavily impacted Taiwan's health care system, triggering clinical crises and lasting effects among affected individuals and families. The first case in Taiwan was identified on February 25, 2003, and the final case was reported on June 15, 2003. During the epidemic, 346 people were diagnosed with SARS, leading to 37 deaths. Outbreaks also occurred in China, Singapore, and Toronto (Canada), showing the vulnerability of global health systems to new zoonotic diseases. Clinically, SARS causes high fever and severe lung inflammation. Survivors often had long-term lung problems, including fibrosis, and bone issues like osteonecrosis, mostly due to high-dose steroid treatment. Although studies have looked at long-term outcomes—especially lung and bone issues—none followed patients beyond 7 years. The COVID-19 pandemic further revealed gaps in understanding how serious viral infections affect wider health areas, including unintentional and intentional injuries. Data on related hospitalizations also remain limited.

Objective: This study aimed to investigate the long-term risk of both unintentional and intentional injuries among survivors of SARS and their relatives, using a nationwide population-based cohort.

Methods: This retrospective cohort study used data from Taiwan's National Health Insurance Research Database, focused on 285 individuals diagnosed with SARS in 2003 and 699 of their relatives, matched in a 1:10 ratio with controls. Injury risks were assessed using Fine and Gray's competing risk models, adjusting for sociodemographic and clinical covariates, over a follow-up period of up to 15 years.

Results: Survivors of SARS exhibited a significant increase in the risk of accidental injuries, with an adjusted hazard ratio (AHR) of 1.631 (95% CI 1.184-2.011; *P*<.001), indicating persistent physiological vulnerabilities postinfection. Family members of survivors of SARS also had elevated injury risk (AHR 1.572, 95% CI 1.148-1.927; *P*<.001), possibly due to stress and caregiving burdens. Subgroup analysis showed increased risks for poisoning (AHR 2.701, 95% CI 1.956-4.084; *P*<.001) and falls (AHR 1.524, 95% CI 1.102-1.878; *P*=.003) among survivors. Relatives faced higher risks for traffic incidents (AHR 2.003, 95% CI 1.462-2.459), poisoning (AHR 1.531, 95% CI 1.120-1.886), medical incidents, falls (AHR 1.802, 95% CI 1.324-2.214), and crushing injuries (AHR 2.469, 95% CI 1.803-3.026; all *P*<.001). These findings highlight the need for targeted preventive measures to address long-term health risks in both survivors of SARS and their families.

Conclusions: Survivors of SARS and their relatives face increased injury risks, highlighting long-term physical and psychosocial vulnerabilities after severe infectious outbreaks. These findings suggest that health care systems should provide preventive and supportive measures to mitigate long-term impacts for those affected by pandemics.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e70608) doi:10.2196/70608

KEYWORDS

SARS; accidental injuries; COVID-19 pandemic; psychosocial impacts; intentional injuries; survivors of SARS; accidental falls; poisoning; psychological stress; caregivers; population surveillance; severe acute respiratory syndrome

Introduction

An outbreak of severe acute respiratory syndrome (SARS) caused by a novel coronavirus severely affected Taiwan in 2003 [1], with the first confirmed diagnosis of SARS on February 25, 2003, and the last diagnosis on June 15, 2003. Through this endemic outbreak, there were 346 patients diagnosed with SARS, and 37 died among these patients [2], and there were also SARS outbreaks in China, Singapore, and Toronto, Canada, at about the same time [2]. The COVID-19 pandemic has had a global impact on all facets of medicine, resulting in significant stress and challenges in the medical community.

The main clinical manifestations after infection were high fever and severe lung inflammation [3,4]. The patients who survived had residual pulmonary fibrosis, as well as osteonecrosis, resulting from treatment with large doses of steroid pulse therapy [5]. Several studies have investigated the long-term outcomes of recovered patients with SARS, especially with respect to lung and bone damage, but no follow-up periods of more than 7 years have been reported [6,7]. Furthermore, there is a scarcity of research regarding the effects of COVID-19 on unintentional injuries and intentional injuries, as well as admission rates [8]. Survivors of respiratory infections like SARS and COVID-19 frequently report persistent symptoms that severely impact their daily functioning and quality of life. These symptoms include chronic fatigue, muscle weakness, and neurological manifestations such as “brain fog” and cognitive deficits, which are well-documented consequences of both diseases [9-11]. Beyond the physical impairments, there is growing evidence highlighting the long-term mental health burden of these infections, with elevated rates of anxiety, depression, posttraumatic stress disorder, and sleep disturbances reported to persist for months or even years after recovery [12-14]. These multifaceted impairments may heighten the risk of accidental injuries in affected individuals. Specifically, cognitive dysfunction, such as reduced attention, slower processing speed, and impaired judgment, is known to increase vulnerability to various types of accidents, including falls and motor vehicle incidents [15,16]. Moreover, physical limitations like chronic fatigue and muscle weakness directly compromise balance, coordination, and reaction time, thus elevating the likelihood of falls, particularly in older adults [17,18]. The increased prevalence of mental health disorders postpandemic may also lead to greater reliance on psychotropic medications, which carry their own risks, including sedation and impaired gait, further contributing to the elevated risk of falls [19]. While the direct link between SARS and subsequent accidental injuries has been underexplored, the parallels in the long-term physical and psychological consequences observed in survivors of COVID-19 strongly suggest that survivors of SARS may similarly face increased vulnerability to injuries [20,21].

The implementation of social distancing and quarantine measures during COVID-19 likely led to a decrease in trauma surgery cases, as restrictions reduced high-risk activities [8]. We hypothesized that SARS is associated with an increased risk of these injuries and tested this hypothesis through a nationwide cohort study using the National Health Insurance Research Database (NHIRD), which includes claims data for Taiwan’s entire population [22].

Methods

Data Sources

Established in 1995, Taiwan’s National Health Insurance (NHI) Program is a mandatory and universal health insurance system that covers approximately 23 million beneficiaries—over 99% of the population [22]. The NHIRD contains comprehensive and detailed data regarding the total outpatients and inpatients [23-33].

We selected an inpatient dataset from 2000 to 2015 from the NHIRD, with individual diagnoses coded using the *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*). The details of the NHI program have been documented in several previous studies [34-39]. Several studies have also confirmed the precision and validity of these diagnoses [40-42].

Ethical Considerations

This study was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). The Institutional Review Board of the Tri-Service General Hospital approved this study (IRB: TSGHIRB No. E202416040) on November 4, 2024, and waived the need for individual consent since all identification data were encrypted in the NHIRD. The study was granted review exemption status, and no artificial intelligence tools were used in data analysis or study preparation. As this was a secondary data analysis of existing administrative data, no direct patient recruitment or contact was performed. The database provides a representative, real-world population sample, allowing for longitudinal follow-up and comprehensive capture of medical events without selection bias inherent in traditional recruitment. No generative artificial intelligence tools were used in any part of the study writing process.

Study Design and Sampled Participants

This retrospective matched-cohort study used inpatient data from January 1, 2000, to December 31, 2015. Each patient with SARS was required to have a diagnosis in an inpatient setting with *ICD-9-CM* codes 480.8 and 480.9. We selected 285 patients with SARS from 2003 and matched them with 2850 controls at a 1:10 ratio. Additionally, we identified 699 relatives of patients with SARS and matched them with 6990 controls, also at a 1:10 ratio. Propensity score matching was used using

XSL•FO
RenderX

variables including gender, age, insurance premium, comorbidities, location, level of care, and index date. Exclusion criteria for the cohorts included individuals with unknown gender and those diagnosed with psychiatric disorders or pneumonia and influenza (*ICD-9-CM* codes 480 - 488) before the index date. The index date was defined as the time when individuals were first diagnosed with SARS within the 1-year study period (Table S1 in Multimedia Appendix 1).

All SARS participants and their matched controls were followed from the index date until the occurrence of an injury, including unintentional injuries (traffic accidents, poisoning, medical-related incidents, falls, burns and fires, drowning, suffocation, crushing injuries, adverse drug reactions, and other unintentional injuries), intentional injuries (suicide, or homicide or abuse), or until death, withdrawal from the NHI program, or the end of 2015.

The covariates include sociodemographic characteristics and comorbidities. Sociodemographic characteristics included gender, age (18 - 44, 45 - 64, ≥65 y), education (<12 years; ≥12 y), monthly insured premiums, urbanization levels, regions of residence, and levels of medical care. The monthly insured premiums have been divided into three categories in New Taiwan Dollars: <18,000, 18,000 - 34,999, and ≥35,000. The urbanization level was defined by population and certain indicators of the city's level of development. Level 1 urbanization was defined as having a population greater than 1,250,000 people. Level 2 urbanization was defined as having a population between 500,000 and 1,250,000. Urbanization levels 3 and 4 were defined as having a population between 150,000 and 500,000 and less than 150,000, respectively [43]. The Charlson Comorbidity Index (CCI) is one of the most widely used comorbidity indices [44,45], which consists of 22 conditions [46]. We used the CCI to quantify the comorbidities since it could predict the in-hospital mortality or outcome in patients with severe adult respiratory infection and other infections [47-49].

Statistical Analysis

Statistical analyses were conducted using SPSS (version 22; IBM Corp). Categorical variables were examined with the chi-square test, while continuous variables, expressed as means (SDs), were compared using 2-sample *t* tests. To evaluate the risk of injuries among patients with and without exposure to patients with SARS or their relatives, competing risk analysis was performed using Fine and Gray's model to calculate adjusted hazard ratios (AHRs) and 95% CIs, adjusting for sociodemographic characteristics and comorbidities. Additionally, Cox proportional hazards regression and sensitivity analyses were used to assess injury risk through the end of 2015. The Kaplan-Meier method was used to compare the cumulative incidence of injuries between the exposed groups and control cohorts, with differences evaluated using the log-rank test. A *P* value of <.05 was considered statistically significant.

Results

Study Cohort Characteristics

Table 1 presents the demographic and clinical characteristics—including gender, age, education level, monthly insured premiums, urbanization level, region of residence, comorbidities, and levels of medical care—of patients with SARS and their relatives, both with and without SARS. Compared to the control group, there were no significant differences in these covariates between patients with SARS, their relatives, and the controls. The participant flowchart is presented in Figure 1.

XSL•FO
RenderX

Table . Baseline demographic and clinical characteristics of survivors of SARS, first-degree relatives, and matched controls in a nationwide retrospective cohort study (Taiwan, 2000 - 2015)[a].

	Patients with SARS				Relatives of patients with SARS			
	Total (n=3135)	With (n=285)	Without (n=2850)	*P* value[b]	Total (n=7689)	With (n=699)	Without (n=6990)	*P* value[b]
Gender, n (%)				≥.99				≥.99
Male	1166 (37.19)	106 (37.19)	1060 (37.19)		4422 (57.51)	402 (57.51)	4020 (57.51)	
Female	1969 (62.81)	179 (62.81)	1790 (62.81)		3267 (42.49)	297 (42.49)	2970 (42.49)	
Age (years), mean (SD)	49.67 (12.24)	49.12 (12.18)	49.73 (12.25)	.42	46.72 (12.20)	46.69 (12.15)	46.72 (12.21)	.95
Age groups (years), n (%)				≥.99				≥.99
18 - 44	1562 (49.82)	142 (49.82)	1420 (49.82)		4411 (57.37)	401 (57.37)	4010 (57.37)	
45 - 64	1012 (32.28)	92 (32.28)	920 (32.28)		2695 (35.05)	245 (35.05)	2450 (35.05)	
≥65	561 (17.89)	51 (17.89)	510 (17.89)		583 (7.58)	53 (7.58)	530 (7.58)	
Insured premium (New Taiwan $)[c], n (%)				.48				.98
<18,000	3068 (97.86)	281 (98.60)	2787 (97.79)		7573 (98.49)	689 (98.57)	6884 (98.48)	
18,000 - 34,999	48 (1.53)	2 (0.70)	46 (1.61)		59 (0.77)	5 (0.72)	54 (0.77)	
≥35,000	19 (0.61)	2 (0.70)	17 (0.60)		57 (0.74)	5 (0.72)	52 (0.74)	
Marital status, n (%)				.38				.92
Without	1465 (46.73)	126 (44.21)	1339 (46.98)		3556 (46.25)	322 (46.07)	3234 (46.27)	
With	1670 (53.27)	159 (55.79)	1511 (53.02)		4133 (53.75)	377 (53.93)	3756 (53.73)	
Education levels (years), n (%)				.95				≥.99
<12	1625 (51.83)	147 (51.58)	1478 (51.86)		3941 (51.26)	359 (51.36)	3582 (51.24)	
≧12	1510 (48.17)	138 (48.42)	1372 (48.14)		3748 (48.74)	340 (48.64)	3408 (48.76)	
CCI_R[d], n (%)				≥.99				≥.99
0	1738 (55.44)	158 (55.44)	1580 (55.44)		5288 (68.77)	482 (68.96)	4806 (68.76)	
1	638 (20.35)	58 (20.35)	580 (20.35)		1471 (19.13)	131 (18.74)	1340 (19.17)	
≥2	759 (24.21)	69 (24.21)	690 (24.21)		930 (12.10)	86 (12.30)	844 (12.07)	
Location, n (%)				.93				≥.99
Northern Taiwan	1437 (45.84)	132 (46.32)	305 (45.79)		3579 (46.55)	325 (46.49)	3254 (46.55)	
Middle Taiwan	597 (19.04)	54 (18.95)	543 (19.05)		1464 (19.04)	133 (19.03)	1331 (19.04)	
Southern Taiwan	961 (30.65)	87 (30.53)	874 (30.67)		2352 (30.59)	214 (30.62)	2138 (30.59)	
Eastern Taiwan	135 (4.31)	11 (3.86)	124 (4.35)		284 (3.69)	26 (3.72)	258 (3.69)	

XSL•FO
RenderX

	Patients with SARS				Relatives of patients with SARS			
	Total (n=3135)	With (n=285)	Without (n=2850)	*P* value[b]	Total (n=7689)	With (n=699)	Without (n=6990)	*P* value[b]
Outlying islands	5 (0.16)	1 (0.35)	4 (0.14)		10 (0.13)	1 (0.14)	9 (0.13)	
Urbanization level, n (%)				.95				≥.99
1 (the highest)	1651 (52.66)	154 (54.04)	1497 (52.53)		4154 (54.03)	378 (54.08)	3776 (54.02)	
2	779 (24.85)	69 (24.21)	710 (24.91)		1875 (24.39)	170 (24.32)	1705 (24.39)	
3	535 (17.07)	46 (16.14)	489 (17.16)		1245 (16.19)	113 (16.17)	1132 (16.19)	
4 (the lowest)	170 (5.42)	16 (5.61)	154 (5.40)		415 (5.40)	38 (5.44)	377 (5.39)	
Level of care, n (%)				.98				≥.99
Hospital center	1634 (52.12)	149 (52.28)	1485 (52.11)		4013 (52.19)	365 (52.22)	3648 (52.19)	
Regional hospital	1170 (37.32)	107 (37.54)	1063 (37.30)		2889 (37.57)	262 (37.48)	2627 (37.58)	
Local hospital	331 (10.56)	29 (10.18)	302 (10.60)		787 (10.24)	72 (10.30)	715 (10.23)	

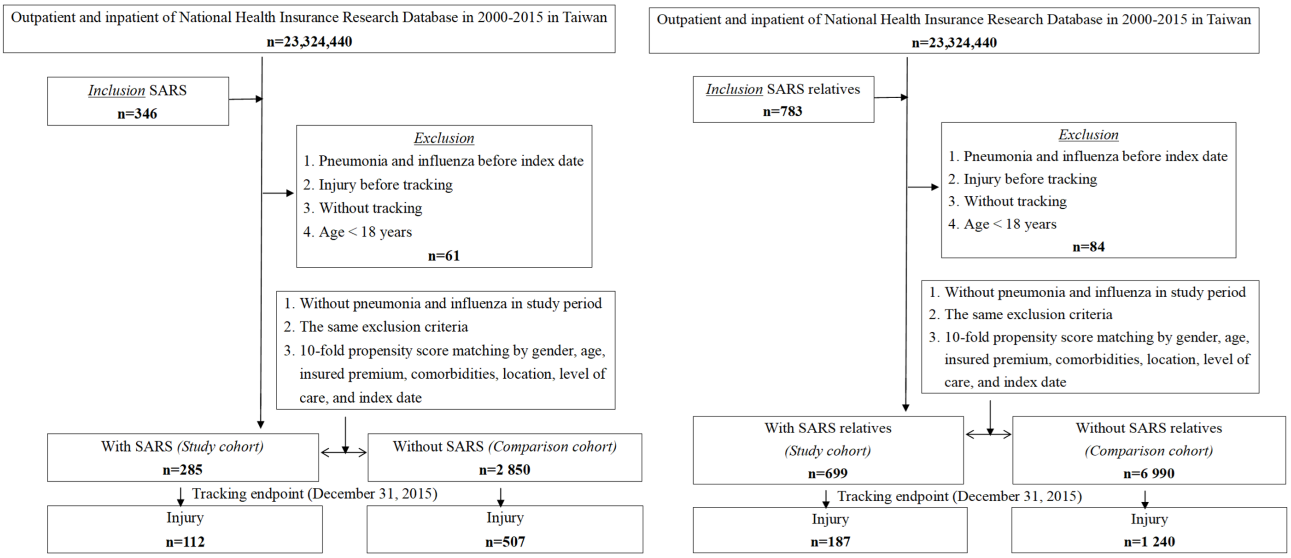
[a]This table presents the demographic and clinical baseline characteristics of survivors of SARS (n=285), their first-degree relatives (n=699), and matched non-SARS controls (n=2850 and n=6990, respectively), extracted from the National Health Insurance Research Database in Taiwan. The data spans from 2000 to 2015 and includes variables such as gender, age, marital status, education level, comorbidity index, insurance premiums, and health care access. The groups were matched 1:10 based on age, gender, comorbidities, and other demographic factors. Statistical significance was assessed using 2-tailed chi-square or Fisher exact tests for categorical variables and 2-tailed *t* tests for continuous variables.

[b]*P* values: chi-square and Fisher exact test on category variables and *t* test on continue variables.

[c]NT$1=US $0.035.

[d]CCI: Charlson Comorbidity Index, Revised.

Figure 1. Flowchart detailing the selection of survivors of severe acute respiratory syndrome (SARS) and their first-degree relatives and matched control cohorts from the National Health Insurance Research Database in a nationwide retrospective cohort study conducted in Taiwan (2000 - 2015).

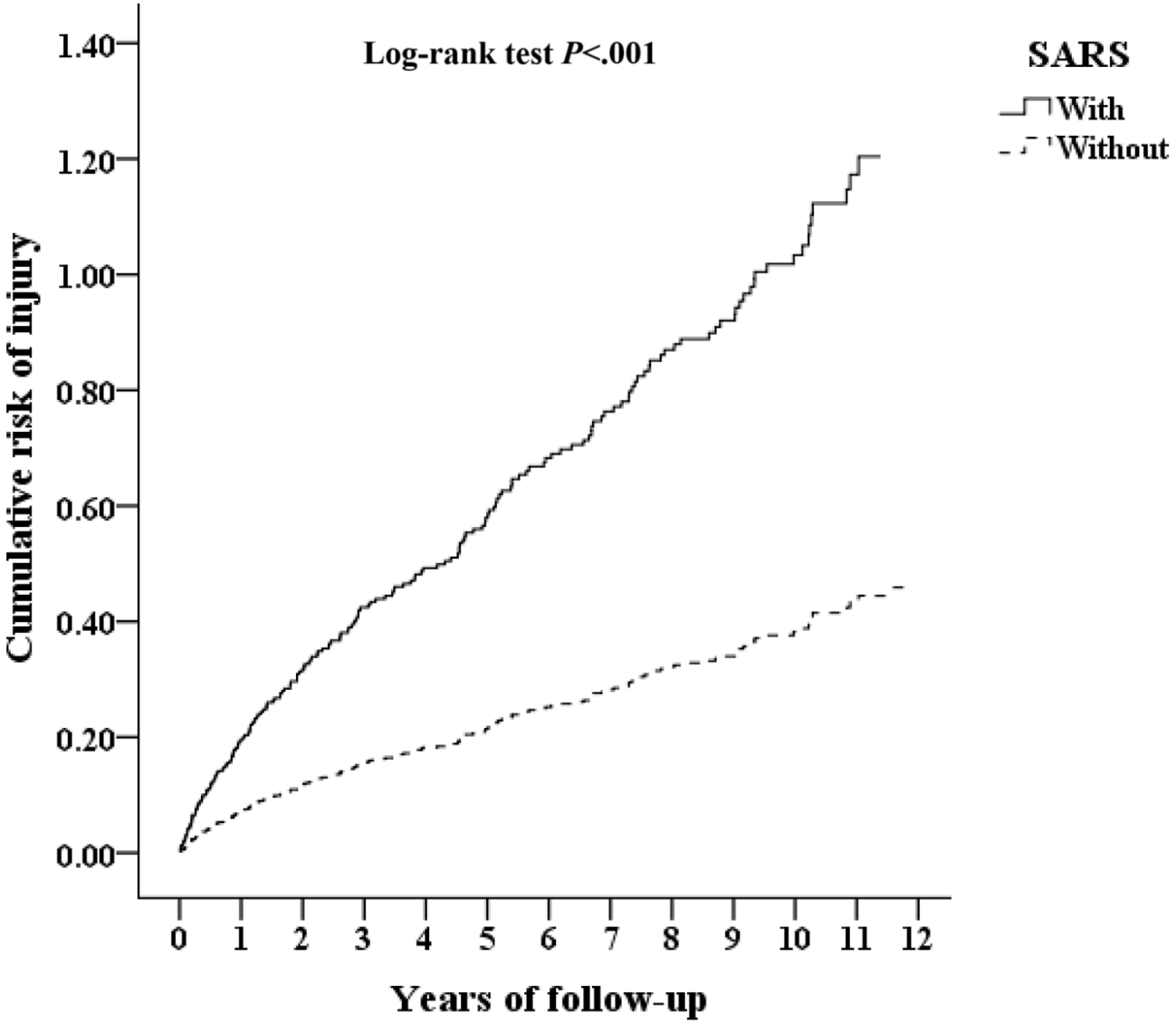


XSL•FO
RenderX

Patients With SARS Kaplan-Meier Curves for the Cumulative Incidence of Injury

In total, 285 patients were diagnosed with SARS during the study period. During the follow-up period, 112 in the SARS group (n=285) and 507 in the control group (n=2850) developed injury (5560.38 vs 2284.63 per 100,000 person-years). Figure 2 reveals that the difference between the two cohorts in the injury was significant (log-rank test; *P*<.001).

Figure 2. Kaplan-Meier survival curves illustrating the cumulative incidence of injury among the survivors of severe acute respiratory syndrome (SARS) and matched controls aged 18 years or older in a nationwide retrospective cohort study (Taiwan, 2000–2015). Statistical significance was determined using the log-rank test (*P*<.001).

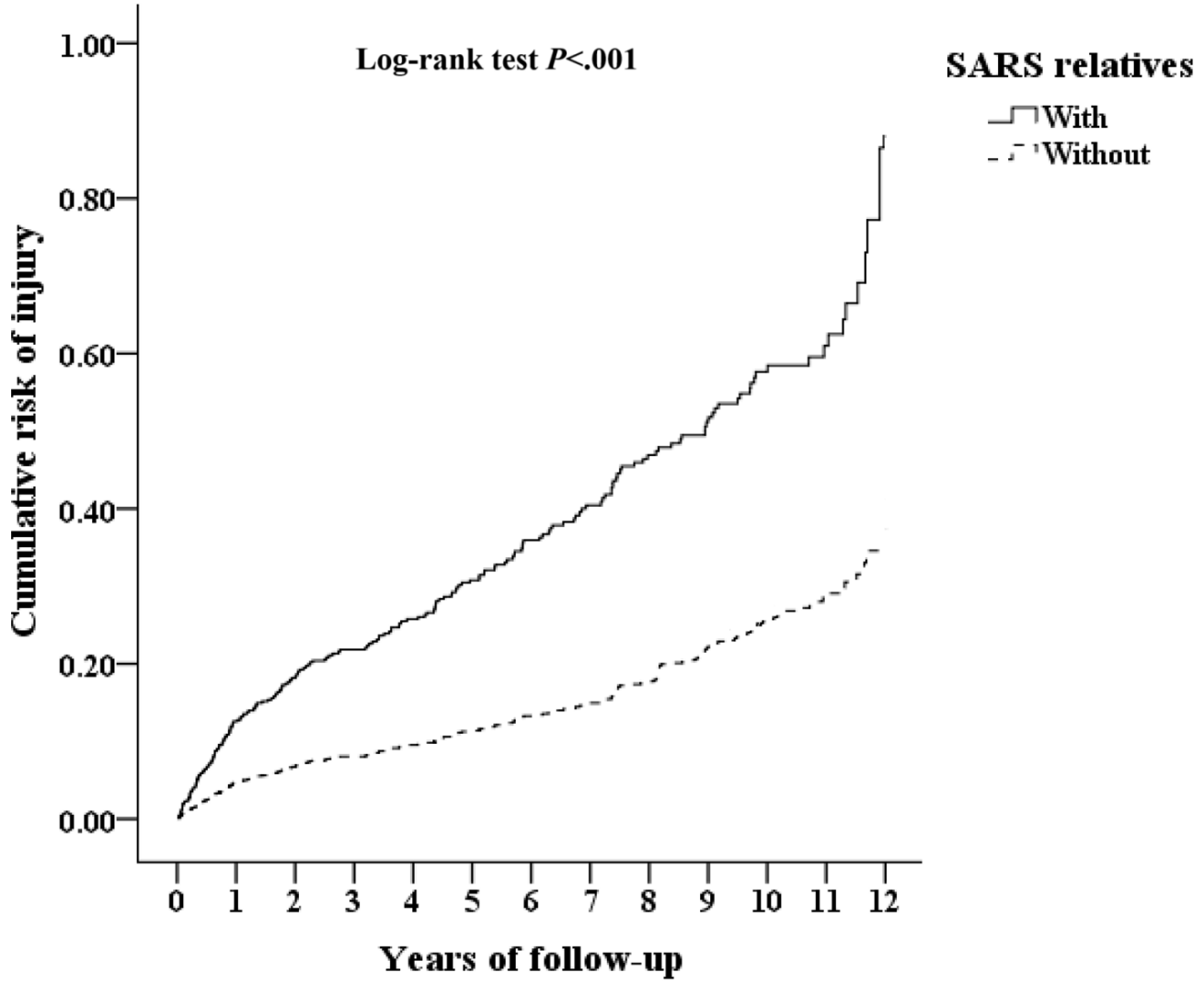


Relatives of Patients With SARS Kaplan-Meier Curves for the Cumulative Incidence of Injury

In total, 699 were relatives of patients with SARS during the study period. During the follow-up period, 187 in the relatives of patients with SARS group (n=699) and 1240 in the control group (n=6990) developed injury (3044.57 vs 2236.71 per 100,000 person-years). Figure 3 reveals that the difference between the two cohorts in the injury was significant (log-rank test; *P*<.001).

XSL•FO
RenderX

Figure 3. Kaplan-Meier survival curves demonstrating the cumulative incidence of injury among first-degree relatives of survivors of severe acute respiratory syndrome (SARS) compared to matched controls aged 18 years or older in a nationwide retrospective cohort study (Taiwan, 2000 - 2015). Statistical significance was evaluated by the log-rank test (*P*<.001).



Years From Patients With SARS and Their Relatives to Injuries

The mean time from the index date to the diagnosis of injuries after diagnosis was 3.60 (SD 3.45) years. The mean years to developed injuries in patients with SARS were 3.47 (SD 3.02) years, which was earlier than in patients without SARS (3.61, SD 3.49 years; Table S2 in Multimedia Appendix 1). The mean time from the index date to the diagnosis of injuries after diagnosis was 3.61 (SD 3.45) years. The mean years to the developed injuries in relatives of patients with SARS were 3.54 (SD 3.07) years, which was earlier than relatives of patients without SARS (3.62, SD 3.49 years; Table S2 in Multimedia Appendix 1).

Analysis of Injury in Patients With SARS and Their Relatives

Table 2 shows the factors of SARS and relatives of patients with SARS, using Fine and Gray's survival analysis, of the factors associated with the risk of injury. In patients with SARS, the crude AHR was 1.684 (95% CI 1.197 - 2.001; *P*<.001), and after adjusting for gender, age, education, monthly insured premiums, urbanization levels, regions of residence, comorbidities, and levels of medical care, the adjusted AHR was 1.631 (95% CI 1.184 - 2.011; *P*<.001) for injury. In relatives of patients with SARS, the crude AHR was 1.670 (95% CI 1.283 - 2.025; *P*<.001), and after adjusting for gender, age, education, monthly insured premiums, urbanization levels, regions of residence, comorbidities, and levels of medical care, the adjusted AHR was 1.572 (95% CI 1.148 - 1.927; *P*<.001) for injury. Factors associated with an increased risk of injury among patients with SARS and their relatives included being male, aged 45–64 years or older than 65 years, having a CCI score of ≥2, residing in urbanization levels 1 and 2, and receiving care from medical centers and regional hospitals.

XSL•FO
RenderX

Table . AHRs[a] for factors associated with injury risk among survivors of severe acute respiratory syndrome (SARS) and their first-degree relatives compared to matched controls using Fine and Gray's competing risk model in a nationwide retrospective cohort study (Taiwan, 2000 - 2015)[b].

	Patients with SARS						Relatives of patients with SARS					
	Competing risk in the model						Competing risk in the model					
	Crude HR[c]	95% CI	*P* value	Adjusted HR	95% CI	*P* value	Crude HR	95% CI	*P* value	Adjusted HR	95% CI	*P* value
SARS (reference: without SARS)	1.684	1.197-2.001	<.001	1.631	1.184-2.011	<.001	1.670	1.283-2.025	<.001	1.572	1.148-1.927	<.001
Male (reference: female)	1.454	1.396-1.572	<.001	1.386	1.267-1.498	<.001	1.428	1.371-1.555	<.001	1.352	1.211-1.409	<.001
Age (reference: 18-44 years)												
45-64 years	1.428	1.199-1.426	<.001	1.401	1.184-1.382	<.001	1.411	1.201-1.529	<.001	1.382	1.10-1.486	<.001
≥65 years	1.677	1.503-1.786	<.001	1.623	1.490-1.777	<.001	1.637	1.512-1.803	<.001	1.597	1.482-1.770	<.001
CCI[d] (reference: CCI=0)												
1	1.325	1.234-1.496	<.001	1.312	1.204-1.452	<.001	1.311	1.201-1.489	<.001	1.297	1.184-1.446	<.001
≥2	1.620	1.591-1.701	<.001	1.607	1.558-1.693	<.001	1.568	1.401-1.698	<.001	1.525	1.372-1.687	<.001
Urbanization level (reference level: 4)												
1 (the highest)	1.397	1.222-1.489	<.001	1.328	1.201-1.412	<.001	1.375	1.273-1.586	<.001	1.303	1.229-1.512	<.001
2	1.286	1.137-1.402	<.001	1.246	1.101-1.354	<.001	1.312	1.146-1.503	<.001	1.289	1.111-1.487	<.001
Level of care (reference: local hospital)												
Hospital center	1.979	1.733-2.079	<.001	1.928	1.701-2.016	<.001	1.889	1.684-2.011	<.001	1.813	1.607-1.978	<.001
Regional hospital	1.633	1.401-1.976	<.001	1.604	1.389-1.921	<.001	1.606	1.375-1.970	<.001	1.597	1.342-1.784	<.001

[a]AHR: adjusted hazard ratio.

[b]This table reports the results of multivariable Cox regression analysis using Fine and Gray's competing risk model to evaluate the factors influencing injury risk among survivors of SARS (n=285) and their first-degree relatives (n=699), compared to matched non-SARS controls. Hazard ratios and 95% CIs are presented for each factor, including SARS exposure, gender, age, comorbidity index, urbanization level, and hospital level of care. The analysis adjusts for baseline characteristics such as demographic variables, comorbidity burden, and access to health care. Data were collected from the National Health Insurance Research Database during the period 2000 - 2015, with statistical significance set at *P*<.05.

[c]HR: hazard ratio.

[d]CCI: Charlson Comorbidity Index.

Subgroup Analysis of Unintentional Injuries and Intentional Injuries in the SARS Cohort and Controls

Table S3 in Multimedia Appendix 1 shows that the SARS cohort had an overall AHR of 1.631 (95% CI 1.184 - 2.011; *P*<.001). The AHRs for these unintentional injuries were: overall unintentional injuries 1.711 (95% CI 1.204 - 2.112; *P*<.001), poisoning 2.701 (95% CI 1.956 - 4.084; *P*<.001), and falls 1.524 (95% CI 1.102 - 1.878; *P*=.003). For intentional injuries, the overall AHR was 2.232 (95% CI 1.695-2.879; *P*<.001), with suicide at 2.685 (95% CI 1.947-3.313; *P*<.001) and homicide or abuse at 1.846 (95% CI 1.341-2.277; *P*<.001).

XSL•FO
RenderX

Subgroup Analysis of Unintentional Injuries and Intentional Injuries in the Relatives of Patients With SARS Cohort and Controls

Table S4 in Multimedia Appendix 1 presents the AHR for the relatives of patients with SARS cohort, which exhibited a significantly increased overall injury risk with an AHR of 1.572 (95% CI 1.148-1.927; *P*<.001). In the category of unintentional injuries, the AHRs were notably elevated: overall unintentional injuries had an AHR of 1.742 (95% CI 1.270-2.135; *P*<.001); traffic-related injuries, 2.003 (95% CI 1.462-2.459; *P*<.001); poisoning, 1.531 (95% CI 1.120 - 1.886; *P*<.001); falls, 1.802 (95% CI 1.324 - 2.214; *P*<.001); and crushing injuries, 2.469 (95% CI 1.803 - 3.026; *P*<.001). Similarly, intentional injuries demonstrated significantly higher AHR: overall intentional injuries, 2.876 (95% CI 2.101 - 3.529; *P*<.001); suicide, 2.197 (95% CI 1.603 - 2.678; *P*<.001); and homicide or abuse, 4.163 (95% CI 3.032 - 5.010; *P*<.001).

Sensitivity Test for Analysis of the Risk of Injury in Patients With SARS and Their Relatives

Patients with SARS and their relatives were associated with an increased risk of injury compared to the control group. For patients with SARS, the AHR for these injuries was 1.631 (95% CI 1.184 - 2.011; *P*<.001). For relatives of patients with SARS, the AHR was 1.572 (95% CI 1.148 - 1.927; *P*<.001). Sensitivity analyses revealed that the associations between injury and both the patients with SARS and their relative cohorts remained significant after excluding individuals diagnosed with injury within the first year, as well as when those diagnosed within the first 5 years were excluded (Table 3).

Table . Sensitivity analysis of injury risk excluding early injury events in survivors of severe acute respiratory syndrome (SARS) and their first-degree relatives compared to matched controls using Fine and Gray's competing risk model in a nationwide retrospective cohort study (Taiwan, 2000 - 2015)[a].

	Patients with SARS									Relatives of patients with SARS								
	With			Without (reference)			Competing risk in the model			With			Without (reference)			Competing risk in the model		
Sensitivity test	Events	PYs[b]	Rate (per 10^5 PYs)	Events	PYs	Rate (per 10^5 PYs)	AHR[c]	95% CI	*P* value	Events	PYs	Rate (per 10^5 PYs)	Events	PYs	Rate (per 10^5 PYs)	AHR	95% CI	*P* value
Overall	112	20425	548.38	507	229177	221.28	1.631	1.184-2.011	<.001	187	61429	304.57	1240	549857	225.51	1.572	1.148-1.927	<.001
In the first year excluded	101	18280	552.49	483	197230	244.85	1.618	1.177-2.000	<.001	167	55281	302.09	1189	489425	243.04	1.529	1.124-1.906	<.001
In the first 5 years excluded	62	11325	547.47	303	124533	243.61	1.429	1.051-1.867	.001	108	36270	301.49	803	329218	243.86	1.498	1.095-1.873	<.001

[a]This table displays results of sensitivity analyses evaluating the robustness of injury risk estimates among survivors of SARS and their relatives by excluding injury events occurring within the first year and first 5 years of follow-up. Person-years, event rates per 100,000 person-years, and adjusted hazard ratios with 95% CIs are reported. Data were derived from the National Health Insurance Research Database in Taiwan and analyzed using Fine and Gray's model, adjusting for relevant baseline characteristics.

[b]PY: person-year.

[c]AHR: adjusted hazard ratio.

Discussion

Summary of Main Findings

Our study conducted long-term follow-ups of patients with SARS and their relatives. This longitudinal approach provides a comprehensive understanding of the enduring impacts of SARS on both patients and their families. A systematic review conducted in Singapore observed significant decreases in elective, traumatic, and outpatient orthopedic surgeries by 80%, 47%, and 63%, respectively, since the onset of the pandemic. This review found an overall proportional increase in domestic injuries and polytraumas, which was attributed to patients with minor injuries potentially avoiding treatment to minimize their risk of COVID-19 exposure [50]. Additionally, a study from Michigan reported an overall 45% reduction in orthopedic trauma encounters at a level II trauma center following the implementation of a shelter-in-place order [51]. Other research has similarly demonstrated a trend of decreased pediatric and adult trauma and orthopedic admissions since the beginning of the COVID-19 pandemic. However, none of these studies have

XSL•FO
RenderX

specifically examined the incidence of injuries or changes in trauma admissions in the northwest Ohio region [52-58].

In this study, we identified several noteworthy findings regarding injury risks within the SARS cohort and their relatives. Specifically, the SARS cohort exhibited a 1.631-fold increased risk of overall injuries compared to the control group. Delving into the causes of these injuries, unintentional injuries presented a 1.711-fold increase, poisoning incidents rose by 2.701-fold, and falls were associated with a 1.524-fold higher risk. Additionally, the cohort of relatives of patients with SARS demonstrated a 1.572-fold elevated risk of overall injuries. Breaking down the injury causes within this group, unintentional injuries showed a 1.742-fold increase, traffic-related injuries doubled with a 2.003-fold rise, poisoning incidents were up by 1.531-fold, medical-related injuries increased by 1.362-fold, falls were 1.802-fold more likely, and crushing injuries saw the highest increase at 2.469-fold. In this study, "relatives" are defined as first-degree family members (spouses, children, parents, or siblings) sharing the same NHI household registration [7]. Because the NHIRD does not capture cohabitation or caregiving roles, we cannot determine whether these relatives lived with or provided direct care to the patients with SARS. Subgroup analysis (Table S4 in Multimedia Appendix 1) shows that relatives of deceased patients with SARS had an AHR of 2.34 (95% CI 2.10‐2.61), compared to 1.78 (95% CI 1.56‐2.03) for relatives of surviving patients. The higher risk among bereaved relatives may reflect grief-related trauma and family disruption, though increased health care interactions or reporting bias cannot be excluded. Future research should incorporate data on household residency and caregiving responsibilities—such as via linkage to household registries or targeted surveys—to clarify the pathways by which SARS exposure affects relatives' injury risk.

Detailed Discussion and Interpretation

Interestingly, in other studies, we have found that falls encounters increased during the COVID-19 pandemic. The impact of COVID-19 on the older population included not only detrimental health effects but also societal effects, such as increased isolation leading to higher rates of depression, anxiety, and dementia [5]. We hypothesize that the incidence of falls increased among the older population due to the decreased availability of professional caregivers during COVID-19. With fewer family members and professional help available, older patients may have been more susceptible to falling. Additionally, in this study, the risk of falls remained elevated 15 years after tracking patients with SARS, which we attribute to the physical weakness resulting from the illness. In contrast, among relatives of patients with SARS, the risk of falls did not increase. These findings highlight the long-term physical vulnerabilities in patients with SARS and suggest that the broader societal disruptions caused by pandemics can have lasting impacts on injury risks within vulnerable populations. The literature indicates that fatigue was the most common symptom among survivors of SARS 18 years after discharge, with osteoporosis and femoral head necrosis being the primary sequelae. The respiratory and hip function scores of survivors of SARS were significantly lower compared to those of the control group. Physical and social functioning at 18 years postdischarge showed improvement compared to that at 12 years but still remained inferior to the controls. Emotional and mental health were fully recovered. Lung lesions detected by computed tomography scans remained consistent at 18 years, particularly in the right upper lobe and left lower lobe. Plasma multiomics analysis indicated abnormal metabolism of amino acids and lipids, an enhanced host immune defense response to bacteria and external stimuli, B-cell activation, and increased cytotoxicity of CD8+ T cells, while showing impaired antigen presentation capacity of CD4+ T cells [6,59-61]. These factors may contribute to the increased risk of falls among patients with SARS.

The persistence of suicide and homicide, or abuse risk morbidities among survivors of SARS and their relatives who participated in this study was alarming. This study identified a particularly noteworthy finding: among patients with SARS, the risk of homicide or abuse was elevated by a factor of 1.846. However, this risk was significantly higher among the relatives of patients with SARS, increasing to 4.163 times. Reviewing results from previous post-SARS cohorts, psychiatric morbidities measured by standardized questionnaires ranged from 10% to 35% in the acute phase of the infection (from the acute phase to 1 month after SARS) [62,63], increasing to 64% at the 1-year follow-up [64]. In the context of the COVID-19 pandemic, research has demonstrated a significant change in the number of nonaccidental trauma encounters, with more cases occurring during the pre-COVID-19 period than during COVID-19 among both adult (15 vs 11 cases) and pediatric (29 vs 9 cases) populations. There is substantial concern that individuals may have been subjected to increased neglect and abuse due to the heightened time spent at home during COVID-19 [8]. Regarding the observed decrease in nonaccidental trauma encounters during COVID-19, concerns have been raised about whether this reduction was due to fewer patients seeking treatment for these traumas rather than a true decrease in the incidence of such events. During the pandemic, time spent engaging in activities outside the home—where nonaccidental trauma could be identified and reported—considerably decreased. In addition to the psychosocial stressors associated with SARS and COVID-19, the long-term adverse health outcomes experienced by survivors of SARS may also serve as a risk factor for accidental injuries among their relatives [65,66]. The rise in suicide rates following the SARS outbreak was a critical public health concern. Previous studies have explored the potential increased risk of suicide among older populations [3,4] and emergency department visitors [67] during and after the SARS outbreak. To our knowledge, this study represents the first investigation into suicide specifically among survivors of SARS. As previously discussed, the association between SARS and suicide risk was found to be statistically significant, even after excluding suicides occurring within the first year postinfection. However, this association was not significant after excluding suicides that occurred within the first 5 years. These findings highlight the importance of vigilant assessment of suicide risk in the years following SARS infection.

To elucidate pathways linking SARS exposure to elevated injury risks, we propose a 3-fold framework. Biological: chronic fatigue, pulmonary fibrosis, and immune dysregulation in

XSL•FO
RenderX

survivors of SARS may impair physical resilience and coordination, contributing to a higher fall risk (AHR 1.524, 95% CI 1.102 - 1.878; Table S3 in Multimedia Appendix 1) and paralleling neurological sequela observed in survivors of COVID-19 [68]. Psychological: quarantine stress, fear of morbidity or mortality, and economic hardship exacerbate mental-health complications, elevating self-harm risk (suicide AHR 2.685, 95% CI 1.947 - 3.313; Table S3 in Multimedia Appendix 1) and accidental injuries. Social: reduced social support and disrupted routines further heighten vulnerability, particularly among isolated individuals. Preventive interventions: regular physical activity has demonstrated efficacy in preventing postinfection mental-health complications [69], suggesting a modifiable strategy to mitigate injury risk in both SARS and COVID-19 populations [70,71].

Factors contributing to this increased burden may include stress from quarantine and isolation, fear of uncertainty and death, impaired physical health following severe viral infection, the economic consequences of SARS, and even adverse emotional responses to SARS-related information [71-73]. These factors collectively underscore the necessity for careful monitoring and proactive intervention to mitigate psychiatric risks in affected populations.

Limitations

First, although we achieved an excellent balance on observed covariates—gender, age, insurance premium, CCI, geographic location, level of care, and index date (all $P>.05$; Table 1)—our retrospective matched-cohort design cannot definitively establish causality. Unmeasured confounders—such as occupational exposures (eg, health care workers at hospitals and airport or transit staffs) or underlying socioeconomic and environmental factors—may bias our estimates. Indeed, the AHR for homicide or abuse among relatives (AHR 4.163; 95% CI 3.032 - 5.010) substantially exceeds that for patients (AHR 1.846; 95% CI 1.341 - 2.277; Table S4 in Multimedia Appendix 1), highlighting potential residual confounding. Moreover, our analysis is based exclusively on Taiwan's 2003 SARS outbreak, which featured distinctive health care responses, social-distancing policies, and cultural norms; therefore, generalizing to other populations or to the COVID-19 pandemic—despite emerging parallels in postinfection neurological sequela—requires caution [68].

Second, the NHIRD does not capture validated injury severity measures (eg, standardized scales), nor does it include laboratory or imaging results or confirmatory diagnostic-test data, limiting our capacity to explore underlying injury mechanisms or to stratify outcomes by clinical severity [40].

Third, patients treated for SARS may have been more health-conscious, potentially leading to an overestimation of injury incidence. Additionally, clinicians may have exercised heightened vigilance in managing survivors of SARS, possibly increasing inpatient referrals. However, sensitivity analyses excluding injuries occurring within the first year and first 5 years after SARS still showed significantly elevated risks (AHR 1.618, 95% CI 1.177 - 2.000; AHR 1.429, 95% CI 1.051 - 1.867). This suggests that differential treatment decisions alone are unlikely to explain the higher hazard ratios observed.

Finally, our 15-year follow-up spanned periods of major shifts in health care delivery (eg, rapid adoption of telemedicine and reorganization of hospital services for infection control), social-support structures (eg, community quarantine measures and establishment of outreach networks), and injury-reporting workflows (eg, revised triage protocols and updated insurance-claim coding). Such changes likely affected both care-seeking behaviors and the administrative capture of injuries; for instance, outpatient visits decreased by approximately 45% during the SARS outbreak [74], indicating potential underascertainment of minor injuries. Similar patterns of health care avoidance have been observed during the COVID-19 pandemic, highlighting the importance of contextualizing our results within evolving health care and reporting environments.

Future research should include more detailed clinical data, address potential confounders, and consider expanding the cohort to include family members and health care workers for a more comprehensive understanding.

Conclusions

This study found that survivors of SARS and their relatives had a significantly increased risk of both unintentional and intentional injuries, persisting over 15 years. Relatives faced a heightened risk of homicide or abuse, while survivors showed increased risks of poisoning, falls, and suicide. These findings highlight the need for clinicians to closely monitor the psychological and injury risks in survivors of SARS and their relatives, especially during future outbreaks like COVID-19, to ensure adequate mental health and safety support.

Acknowledgments

The authors would like to express their gratitude and appreciation to the Health and Welfare Data Science Center, Ministry of Health and Welfare, Taiwan, for providing access to the National Health Insurance Research Database. This research was funded by the Tri-Service General Hospital (grants TSGH-B-114022 and TSGH-A-114010 special plan).

Data Availability

This study uses third-party data. Taiwan launched a single-payer National Health Insurance (NHI) program on March 1, 1995. The database of this program contains registration files and original claim data for reimbursement. Large computerized databases are derived from this system by the NHI Administration. Investigators interested may submit a formal proposal to the NHI Research Database. The authors confirm they did not have any special access privileges.

XSL•FO
RenderX

Authors' Contributions

CS wrote the Introduction, Discussion, and Conclusion sections, contributed to drafting the Abstract, and prepared Table 1 and Table 2. CHC implemented the software, verified data accuracy, and contributed to data validation. FHL participated in methodology development, conducted data validation, and supported the interpretation of statistical results. DYN cowrote the Results section and contributed to drafting the Abstract and Supplementary Tables S2 to S4. CAS designed the study methodology, curated the dataset, and constructed Table S1. CHT contributed to the study design and performed cohort matching and data curation. LYF conceptualized the study, supervised the research process, and critically reviewed the manuscript. WCC coconceptualized the study, coordinated the research project, secured funding, and reviewed and edited the final manuscript. All authors reviewed and approved the final version of the manuscript

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File, 27 KB - publichealth_v11i1e70608_app1.docx]

References

1. Ruuskanen O, Lahti E, Jennings LC, Murdoch DR. Viral pneumonia. Lancet 2011 Apr 9;377(9773):1264-1275. [doi: 10.1016/S0140-6736(10)61459-6] [Medline: 21435708]
2. Acute respiratory syndrome. China, Hong Kong Special Administrative Region of China, and Viet Nam. Wkly Epidemiol Rec 2003 Mar 14;78(11):73-74. [Medline: 12674023]
3. Chan SMS, Chiu FKH, Lam CWL, Leung PYV, Conwell Y. Elderly suicide and the 2003 SARS epidemic in Hong Kong. Int J Geriatr Psychiatry 2006 Feb;21(2):113-118. [doi: 10.1002/gps.1432] [Medline: 16416469]
4. Cheung YT, Chau PH, Yip PSF. A revisit on older adults suicides and severe acute respiratory syndrome (SARS) epidemic in Hong Kong. Int J Geriatr Psychiatry 2008 Dec;23(12):1231-1238. [doi: 10.1002/gps.2056] [Medline: 18500689]
5. Abbasi J. Social isolation—the other COVID-19 threat in nursing homes. JAMA 2020 Aug 18;324(7):619. [doi: 10.1001/jama.2020.13484]
6. Dhawan RT, Gopalan D, Howard L, et al. Beyond the clot: perfusion imaging of the pulmonary vasculature after COVID-19. Lancet Respir Med 2021 Jan;9(1):107-116. [doi: 10.1016/S2213-2600(20)30407-0] [Medline: 33217366]
7. Tzeng NS, Chung CH, Chang CC, et al. What could we learn from SARS when facing the mental health issues related to the COVID-19 outbreak? A nationwide cohort study in Taiwan. Transl Psychiatry 2020 Oct 6;10(1):339. [doi: 10.1038/s41398-020-01021-y] [Medline: 33024072]
8. Sanford EL, Zagory J, Blackwell JM, Szmuk P, Ryan M, Ambardekar A. Changes in pediatric trauma during COVID-19 stay-at-home epoch at a tertiary pediatric hospital. J Pediatr Surg 2021 May;56(5):918-922. [doi: 10.1016/j.jpedsurg.2021.01.020] [Medline: 33516579]
9. Panagea E, Messinis L, Petri MC, et al. Neurocognitive Impairment in long COVID: a systematic review. Arch Clin Neuropsychol 2025 Jan 21;40(1):125-149. [doi: 10.1093/arclin/acae042] [Medline: 38850628]
10. Haunhorst S, Bloch W, Wagner H, et al. Long COVID: a narrative review of the clinical aftermaths of COVID-19 with a focus on the putative pathophysiology and aspects of physical activity. Oxf Open Immunol 2022;3(1):iqac006. [doi: 10.1093/oxfimm/iqac006] [Medline: 36846561]
11. Subramaniam A, Shekar K, Afroz A, et al. Frailty and mortality associations in patients with COVID-19: a systematic review and meta-analysis. Intern Med J 2022 May;52(5):724-739. [doi: 10.1111/imj.15698] [Medline: 35066970]
12. Ceban F, Ling S, Lui LMW, et al. Fatigue and cognitive impairment in post-COVID-19 syndrome: a systematic review and meta-analysis. Brain Behav Immun 2022 Mar;101:93-135. [doi: 10.1016/j.bbi.2021.12.020] [Medline: 34973396]
13. Taquet M, Geddes JR, Husain M, Luciano S, Harrison PJ. 6-month neurological and psychiatric outcomes in 236 379 survivors of COVID-19: a retrospective cohort study using electronic health records. Lancet Psychiatry 2021 May;8(5):416-427. [doi: 10.1016/S2215-0366(21)00084-5] [Medline: 33836148]
14. AlRasheed MM, Fekih-Romdhane F, Jahrami H, et al. The prevalence and severity of insomnia symptoms during COVID-19: a global systematic review and individual participant data meta-analysis. Sleep Med 2022 Dec;100:7-23. [doi: 10.1016/j.sleep.2022.06.020] [Medline: 36030616]
15. Arenas DJ, Thomas A, Wang J, DeLisser HM. A systematic review and meta-analysis of depression, anxiety, and sleep disorders in US adults with food insecurity. J Gen Intern Med 2019 Dec;34(12):2874-2882. [doi: 10.1007/s11606-019-05202-4] [Medline: 31385212]
16. Paterson RW, Brown RL, Benjamin L, et al. The emerging spectrum of COVID-19 neurology: clinical, radiological and laboratory findings. Brain (Bacau) 2020 Oct 1;143(10):3104-3120. [doi: 10.1093/brain/awaa240]
17. Ambrose AF, Paul G, Hausdorff JM. Risk factors for falls among older adults: a review of the literature. Maturitas 2013 May;75(1):51-61. [doi: 10.1016/j.maturitas.2013.02.009] [Medline: 23523272]

XSL•FO
RenderX

18. Woodrow M, Carey C, Ziauddeen N, et al. Systematic review of the prevalence of long COVID. Open Forum Infect Dis 2023 Jul;10(7):ofad233. [doi: 10.1093/ofid/ofad233] [Medline: 37404951]
19. Seppala LJ, Wermelink A, de Vries M, et al. Fall-risk-increasing drugs: a systematic review and meta-analysis: ii. psychotropics. J Am Med Dir Assoc 2018 Apr;19(4):371. [doi: 10.1016/j.jamda.2017.12.098] [Medline: 29402652]
20. Keshta AS, Mallah SI, Al Zubaidi K, et al. COVID-19 versus SARS: a comparative review. J Infect Public Health 2021 Jul;14(7):967-977. [doi: 10.1016/j.jiph.2021.04.007] [Medline: 34130121]
21. Low ZY, Yip AJW, Sharma A, Lal SK. SARS coronavirus outbreaks past and present-a comparative analysis of SARS-CoV-2 and its predecessors. Virus Genes 2021 Aug;57(4):307-317. [doi: 10.1007/s11262-021-01846-9] [Medline: 34061288]
22. Chan WSH. Taiwan's healthcare report 2010. EPMA J 2010 Dec;1(4):563-585. [doi: 10.1007/s13167-010-0056-8] [Medline: 23199110]
23. Chien WC, Chung CH, Lin FH, Chang HA, Kao YC, Tzeng NS. Is weight control surgery associated with increased risk of newly onset psychiatric disorders? A population-based, matched cohort study in Taiwan. J Med Sci 2017;37(4):137. [doi: 10.4103/jmedsci.jmedsci_94_16]
24. Chen TY, Huang CH, Chung CH, et al. Sex and age differences in the association between anxiety disorders and narcolepsy: a nationwide population-based case control study. J Affect Disord 2020 Mar 1;264:130-137. [doi: 10.1016/j.jad.2019.12.010] [Medline: 32056742]
25. Lin CH, Chien WC, Chung CH, et al. Increased risk of dementia in patients with genital warts: a nationwide cohort study in Taiwan. J Dermatol 2020 May;47(5):503-511. [doi: 10.1111/1346-8138.15277] [Medline: 32189395]
26. Lin YC, Chen TY, Chien WC, et al. Stimulants associated with reduced risk of hospitalization for motor vehicle accident injury in patients with obstructive sleep apnea—a nationwide cohort study. BMC Pulm Med 2020 Feb 3;20(1):28. [doi: 10.1186/s12890-019-1041-1] [Medline: 32013932]
27. Liu YP, Chien WC, Chung CH, Chang HA, Kao YC, Tzeng NS. Are anticholinergic medications associated with increased risk of dementia and behavioral and psychological symptoms of dementia? a nationwide 15-year follow-up cohort study in Taiwan. Front Pharmacol 2020;11:30. [doi: 10.3389/fphar.2020.00030] [Medline: 32116707]
28. Wan FJ, Chien WC, Chung CH, Yang YJ, Tzeng NS. Association between traumatic spinal cord injury and affective and other psychiatric disorders—a nationwide cohort study and effects of rehabilitation therapies. J Affect Disord 2020 Mar 15;265:381-388. [doi: 10.1016/j.jad.2020.01.063] [Medline: 32090763]
29. Wang DS, Chung CH, Chang HA, et al. Association between child abuse exposure and the risk of psychiatric disorders: a nationwide cohort study in Taiwan. Child Abuse Negl 2020 Mar;101:104362. [doi: 10.1016/j.chiabu.2020.104362]
30. Yang CC, Chien WC, Chung CH, et al. No association between human immunodeficiency virus infections and dementia: a nationwide cohort study In Taiwan. Neuropsychiatr Dis Treat 2019;15:3155-3166. [doi: 10.2147/NDT.S225584] [Medline: 31814723]
31. Yeh TC, Chien WC, Chung CH, et al. Psychiatric disorders after traumatic brain injury: a nationwide population-based cohort study and the effects of rehabilitation therapies. Arch Phys Med Rehabil 2020 May;101(5):822-831. [doi: 10.1016/j.apmr.2019.12.005] [Medline: 31917196]
32. Kao LC, Chien WC, Chung CH, et al. The newly diagnosed amnestic disorders and dementia: a nationwide, cohort study in Taiwan. Taiwan J Psychiatry 2018;32(1):18-28 [FREE Full text]
33. Chiou YE, Chien WC, Chung CH, et al. New users of herbal medicine containing aristolochic acids and the risk of dementia in the elderly: a nationwide, population-based study in Taiwan. Neuropsychiatr Dis Treat 2020;16:1493-1504. [doi: 10.2147/NDT.S250659] [Medline: 32606702]
34. Chien WC, Chung CH, Lin FH, et al. The risk of injury in adults with attention-deficit hyperactivity disorder: a nationwide, matched-cohort, population-based study in Taiwan. Res Dev Disabil 2017 Jun;65:57-73. [doi: 10.1016/j.ridd.2017.04.011] [Medline: 28458048]
35. Chien WC, Chung CH, Lai CH, Chou CH. A retrospective population-based study of injury types among elderly in Taiwan. Int J Inj Contr Saf Promot 2014;21(1):3-8. [doi: 10.1080/17457300.2012.717084] [Medline: 22924733]
36. Chang SY, Chien WC, Chung CH, et al. Risk of dementia after charcoal-burning suicide attempts: a nationwide cohort study in Taiwan. J Investig Med 2018 Oct;66(7):1070-1082. [doi: 10.1136/jim-2018-000759] [Medline: 29884660]
37. Chu CW, Chien WC, Chung CH, et al. Electroconvulsive therapy and risk of dementia—a nationwide cohort study in Taiwan. Front Psychiatry 2018;9:397. [doi: 10.3389/fpsyt.2018.00397] [Medline: 30245639]
38. Tzeng NS, Chung CH, Lin FH, et al. Anti-herpetic medications and reduced risk of dementia in patients with herpes simplex virus infections—a nationwide, population-based cohort study in Taiwan. Neurotherapeutics 2018 Apr;15(2):417-429. [doi: 10.1007/s13311-018-0611-x] [Medline: 29488144]
39. Tzeng NS, Chung CH, Lin FH, et al. Risk of dementia in adults with ADHD: a nationwide, population-based cohort study in Taiwan. J Atten Disord 2019 Jul;23(9):995-1006. [doi: 10.1177/1087054717714057] [Medline: 28629260]
40. Hsieh CY, Su CC, Shao SC, et al. Taiwan's National Health Insurance Research Database: past and future. Clin Epidemiol 2019;11:349-358. [doi: 10.2147/CLEP.S196293] [Medline: 31118821]
41. Cheng CL, Kao YHY, Lin SJ, Lee CH, Lai ML. Validation of the National Health Insurance Research Database with ischemic stroke cases in Taiwan. Pharmacoepidemiol Drug Saf 2011 Mar;20(3):236-242. [doi: 10.1002/pds.2087] [Medline: 21351304]

XSL•FO
RenderX

42. Liang JA, Sun LM, Muo CH, Sung FC, Chang SN, Kao CH. The analysis of depression and subsequent cancer risk in Taiwan. Cancer Epidemiol Biomarkers Prev 2011 Mar 1;20(3):473-475. [doi: 10.1158/1055-9965.EPI-10-1280]
43. Chang CY, Chen WL, Liou YF, et al. Increased risk of major depression in the three years following a femoral neck fracture—a national population-based follow-up study. PLoS One 2014;9(3):e89867. [doi: 10.1371/journal.pone.0089867] [Medline: 24626193]
44. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
45. de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity. a critical review of available methods. J Clin Epidemiol 2003 Mar;56(3):221-229. [doi: 10.1016/s0895-4356(02)00585-1] [Medline: 12725876]
46. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. J Clin Epidemiol 1994 Nov;47(11):1245-1251. [doi: 10.1016/0895-4356(94)90129-5] [Medline: 7722560]
47. Ternavasio-de la Vega HG, Castaño-Romero F, Ragozzino S, et al. The updated Charlson Comorbidity Index is a useful predictor of mortality in patients with Staphylococcus aureus bacteraemia. Epidemiol Infect 2018 Dec;146(16):2122-2130. [doi: 10.1017/S0950268818002480]
48. Baiomi A, Abbas H, Pirzada U, Zaidi B, Saturno D, Nayudu SK. Charlson Comorbidity Index (CCI): an independent predictor of outcomes in Clostridium difficile infection (CDI). Am J Gastroenterol 2018 Oct;113(Supplement):S1525. [doi: 10.14309/00000434-201810001-02742]
49. Setter NW, Peres ML, de Almeida BMM, Petterle RR, Raboni SM. Charlson Comorbidity Index scores and in-hospital prognosis of patients with severe acute respiratory infections. Intern Med J 2020 Jun;50(6):691-697. [doi: 10.1111/imj.14398] [Medline: 31180163]
50. Ow ZGW, Cheong CK, Chin YH, Chin BZ. A look at the global impact of SARS CoV-2 on orthopedic services. J Clin Orthop Trauma 2021 Jan;12(1):33-39. [doi: 10.1016/j.jcot.2020.10.052] [Medline: 33191995]
51. Stoker S, McDaniel D, Crean T, et al. Effect of shelter-in-place orders and the COVID-19 pandemic on orthopaedic trauma at a community level ii trauma center. J Orthop Trauma 2020 Sep;34(9):e336-e342. [doi: 10.1097/BOT.0000000000001860] [Medline: 32815848]
52. Bram JT, Johnson MA, Magee LC, et al. Where have all the fractures gone? The epidemiology of pediatric fractures during the COVID-19 pandemic. J Pediatr Orthop 2020;40(8):373-379. [doi: 10.1097/BPO.0000000000001600]
53. Druel T, Andeol Q, Rongieras F, Bertani A, Bordes M, Alvernhe A. Evaluation of containment measures' effect on orthopaedic trauma surgery during the COVID-19 pandemic: a retrospective comparison between 2019 and 2020. Int Orthop 2020 Nov;44(11):2229-2234. [doi: 10.1007/s00264-020-04712-9] [Medline: 32696333]
54. Kale NN, Patel AH, Leddy MJ 3rd, Savoie FH 3rd, Sherman WF. The effect of COVID-19 on orthopedic practices and surgeons in Louisiana. Orthopedics 2020 Nov 1;43(6):351-355. [doi: 10.3928/01477447-20201023-01] [Medline: 33211903]
55. Fyllos A, Varitimidis S, Papageorgiou F, et al. Upper extremity emergencies during SARS-COV-2 pandemic: turnout trends. Injury 2021 Mar;52(3):387-394. [doi: 10.1016/j.injury.2020.11.045] [Medline: 33246643]
56. Gumina S, Proietti R, Polizzotti G, Carbone S, Candela V. The impact of COVID-19 on shoulder and elbow trauma: an Italian survey. J Shoulder Elbow Surg 2020 Sep;29(9):1737-1742. [doi: 10.1016/j.jse.2020.05.003] [Medline: 32713663]
57. Lubbe RJ, Miller J, Roehr CA, et al. Effect of statewide social distancing and stay-at-home directives on orthopaedic trauma at a southwestern level 1 trauma center during the COVID-19 pandemic. J Orthop Trauma 2020 Sep;34(9):e343-e348. [doi: 10.1097/BOT.0000000000001890] [Medline: 32815849]
58. Wallace CN, Kontoghiorghe C, Kayani B, Chang JS, Haddad FS. The impact of COVID-19 on trauma and orthopaedic surgery in the United Kingdom. Bone Jt Open 2020 Jul;1(7):420-423. [doi: 10.1302/2633-1462.17.BJO-2020-0102] [Medline: 33215133]
59. Li K, Wu Q, Li H, et al. Multiomic characterisation of the long-term sequelae of SARS survivors: a clinical observational study. EClinicalMedicine 2023 Apr;58:101884. [doi: 10.1016/j.eclinm.2023.101884] [Medline: 36873427]
60. Tansey CM, Louie M, Loeb M. One-year outcomes and health care utilization in survivors of severe acute respiratory syndrome. Arch Intern Med 2007 Jun 25;167(12):1312. [doi: 10.1001/archinte.167.12.1312] [Medline: 17592106]
61. Lam MHB, Wing YK, Yu MWM, et al. Mental morbidities and chronic fatigue in severe acute respiratory syndrome survivors: long-term follow-up. Arch Intern Med 2009 Dec 14;169(22):2142-2147. [doi: 10.1001/archinternmed.2009.384] [Medline: 20008700]
62. Wu KK, Chan SK, Ma TM. Posttraumatic stress after SARS. Emerg Infect Dis 2005 Aug;11(8):1297-1300. [doi: 10.3201/eid1108.041083] [Medline: 16102324]
63. Sheng B, Cheng SKW, Lau KK, Li HL, Chan ELY. The effects of disease severity, use of corticosteroids and social factors on neuropsychiatric complaints in severe acute respiratory syndrome (SARS) patients at acute and convalescent phases. Eur Psychiatry 2005 May;20(3):236-242. [doi: 10.1016/j.eurpsy.2004.06.023] [Medline: 15935422]
64. Lee AM, Wong J, McAlonan GM, et al. Stress and psychological distress among SARS survivors 1 year after the outbreak. Can J Psychiatry 2007 Apr;52(4):233-240. [doi: 10.1177/070674370705200405] [Medline: 17500304]
65. Lee DTS, Wing YK, Leung HCM, et al. Factors associated with psychosis among patients with severe acute respiratory syndrome: a case-control study. Clin Infect Dis 2004 Oct 15;39(8):1247-1249. [doi: 10.1086/424016] [Medline: 15486852]

XSL•FO
RenderX

66. Phua DH, Tang HK, Tham KY. Coping responses of emergency physicians and nurses to the 2003 severe acute respiratory syndrome outbreak. Acad Emerg Med 2005 Apr;12(4):322-328. [doi: 10.1197/j.aem.2004.11.015] [Medline: 15805323]
67. Huang CC, Yen DHT, Huang HH, et al. Impact of severe acute respiratory syndrome (SARS) outbreaks on the use of emergency department medical resources. J Chin Med Assoc 2005 Jun;68(6):254-259. [doi: 10.1016/S1726-4901(09)70146-7] [Medline: 15984818]
68. Park JM, Woo W, Lee SC, et al. Prevalence and mortality risk of neurological disorders during the COVID-19 pandemic: an umbrella review of the current evidence. Neuroepidemiology 2023;57(3):129-147. [doi: 10.1159/000530536] [Medline: 37044073]
69. Rahmati M, Lee S, Yon DK, et al. Physical activity and prevention of mental health complications: an umbrella review. Neurosci Biobehav Rev 2024 May;160:105641. [doi: 10.1016/j.neubiorev.2024.105641] [Medline: 38527637]
70. Chua SE, Cheung V, McAlonan GM, et al. Stress and psychological impact on SARS patients during the outbreak. Can J Psychiatry 2004 Jun;49(6):385-390. [doi: 10.1177/070674370404900607] [Medline: 15283533]
71. Tsang HWH, Scudds RJ, Chan EYL. Psychosocial impact of SARS. Emerg Infect Dis 2004 Jul;10(7):1326-1327. [doi: 10.3201/eid1007.040090] [Medline: 15338536]
72. Sim K, Chua HC. The psychological impact of SARS: a matter of heart and mind. CMAJ 2004 Mar 2;170(5):811-812. [doi: 10.1503/cmaj.1032003] [Medline: 14993176]
73. Sin SS, Huak CY. Psychological impact of the SARS outbreak on a Singaporean rehabilitation department. Int J Ther Rehabil 2004 Sep;11(9):417-424. [doi: 10.12968/ijtr.2004.11.9.19589]
74. Chu D, Chen RC, Ku CY, Chou P. The impact of SARS on hospital performance. BMC Health Serv Res 2008 Nov 6;8(1):228. [doi: 10.1186/1472-6963-8-228] [Medline: 18990210]

Abbreviations

AHR: adjusted hazard ratio
CCI: Charlson Comorbidity Index
***ICD-9-CM*:** *International Classification of Diseases, Ninth Revision, Clinical Modification*
NHI: National Health Insurance
NHIRD: National Health Insurance Research Database
SARS: severe acute respiratory syndrome

Edited by A Mavragani, T Sanchez; submitted 26.12.24; peer-reviewed by SW Lee, W Yang; revised version received 05.06.25; accepted 08.06.25; published 08.07.25.

Please cite as:
Sung C, Chung CH, Sun CA, Tsao CH, Ng DY, Weng TH, Fann LY, Lin FH, Chien WC
COVID-19–Relevant Insights Into the Elevated Risk of Accidental Injuries in Survivors of SARS and Their Relatives in Taiwan: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e70608
URL: https://publichealth.jmir.org/2025/1/e70608
doi:10.2196/70608

© Chieh Sung, Chi-Hsiang Chung, Chien-An Sun, Chang-Huei Tsao, Daphne Yih Ng, Tsu-Hsuan Weng, Li-Yun Fann, Fu-Huang Lin, Wu-Chien Chien. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An Explainable Artificial Intelligence Text Classifier for Suicidality Prediction in Youth Crisis Text Line Users: Development and Validation Study

Julia Thomas[1,2,3], MA; Antonia Lucht[3], MSci; Jacob Segler[4], MA; Richard Wundrack[3], PhD; Marcel Miché[1], PhD; Roselind Lieb[1], Prof Dr; Lars Kuchinke[5], Prof Dr; Gunther Meinlschmidt[2,6,7,8], Prof Dr

[1]Division of Clinical Psychology and Epidemiology, Faculty of Psychology, University of Basel, Basel, Switzerland

[2]Division of Clinical Psychology and Cognitive Behavioural Therapy, International Psychoanalytic University Berlin, Berlin, Germany

[3]Department of Research, Analytics and Development, krisenchat gGmbH, Berlin, Germany

[4]Division of Child and Adolescent Psychiatry/Psychotherapy, Universitätsklinikum Ulm, Ulm, Germany

[5]Division of Methods and Statistics, International Psychoanalytic University Berlin, Berlin, Germany

[6]Clinical Psychology and Psychotherapy, Methods and Approaches, Department of Psychology, Trier University, Trier, Germany

[7]Department of Digital and Blended Psychosomatics and Psychotherapy, Psychosomatic Medicine, University Hospital and University of Basel, Basel, Switzerland

[8]Department of Psychosomatic Medicine, University Hospital Basel, University of Basel, Basel, Switzerland

Corresponding Author:
Julia Thomas, MA
Division of Clinical Psychology and Epidemiology
Faculty of Psychology
University of Basel
Missionsstrasse 60/62
Basel, 4055
Switzerland
Phone: 49 30 57714627
Email: julia.thomas@krisenchat.de

Abstract

Background: Suicide represents a critical public health concern, and machine learning (ML) models offer the potential for identifying at-risk individuals. Recent studies using benchmark datasets and real-world social media data have demonstrated the capability of pretrained large language models in predicting suicidal ideation and behaviors (SIB) in speech and text.

Objective: This study aimed to (1) develop and implement ML methods for predicting SIBs in a real-world crisis helpline dataset, using transformer-based pretrained models as a foundation; (2) evaluate, cross-validate, and benchmark the model against traditional text classification approaches; and (3) train an explainable model to highlight relevant risk-associated features.

Methods: We analyzed chat protocols from adolescents and young adults (aged 14-25 years) seeking assistance from a German crisis helpline. An ML model was developed using a transformer-based language model architecture with pretrained weights and long short-term memory layers. The model predicted suicidal ideation (SI) and advanced suicidal engagement (ASE), as indicated by composite Columbia-Suicide Severity Rating Scale scores. We compared model performance against a classical word-vector-based ML model. We subsequently computed discrimination, calibration, clinical utility, and explainability information using a Shapley Additive Explanations value-based post hoc estimation model.

Results: The dataset comprised 1348 help-seeking encounters (1011 for training and 337 for testing). The transformer-based classifier achieved a macroaveraged area under the curve (AUC) receiver operating characteristic (ROC) of 0.89 (95% CI 0.81-0.91) and an overall accuracy of 0.79 (95% CI 0.73-0.99). This performance surpassed the word-vector-based baseline model (AUC-ROC=0.77, 95% CI 0.64-0.90; accuracy=0.61, 95% CI 0.61-0.80). The transformer model demonstrated excellent prediction for nonsuicidal sessions (AUC-ROC=0.96, 95% CI 0.96-0.99) and good prediction for SI and ASE, with AUC-ROCs of 0.85 (95% CI 0.97-0.86) and 0.87 (95% CI 0.81-0.88), respectively. The Brier Skill Score indicated a 44% improvement in classification performance over the baseline model. The Shapley Additive Explanations model identified language features predictive of SIBs, including self-reference, negation, expressions of low self-esteem, and absolutist language.

XSL•FO
RenderX

Conclusions: Neural networks using large language model–based transfer learning can accurately identify SI and ASE. The post hoc explainer model revealed language features associated with SI and ASE. Such models may potentially support clinical decision-making in suicide prevention services. Future research should explore multimodal input features and temporal aspects of suicide risk.

(JMIR Public Health Surveill 2025;11:e63809) doi:10.2196/63809

KEYWORDS

deep learning; explainable artificial intelligence (XAI); large language model (LLM); machine learning; neural network; prevention; risk monitoring; suicide; transformer model; suicidality; suicidal ideation; self-murder; self-harm; youth; adolescent; adolescents; public health; language model; language models; chat protocols; crisis helpline; help-seeking behaviors; German; Shapley; decision-making; mental health; health informatics; mobile phone

Introduction

Suicide, the third leading cause of premature mortality among German adolescents [1], encompasses a spectrum from suicidal ideation (SI) to preparatory actions [2-5]. Adolescents' vulnerability to suicidal phenomena stems from a complex interplay of biological, genetic, psychological, and social factors [6-9]. Alarmingly, less than 50% of adolescents who attempt suicide receive appropriate psychiatric intervention [10]. This gap between need and care emphasizes the critical importance of accurately identifying and timely identification of at-risk individuals, challenging mental health care providers and educational institutions.

The digital age, characterized by the frequent use of the internet and smartphones, has transformed youth help-seeking behaviors, with online text-based services becoming the preferred communication mode [11-13]. These platforms align with young people's digital environments [12,14], reducing help-seeking barriers such as limited service availability and stigma [15]. Their immediacy and anonymity are crucial for crisis help-seeking, establishing text-based helplines as critical public health measures for suicide-related issues [15,16]. While presenting challenges for clinicians in rapidly assessing risks and implementing interventions [17], these digitized services also offer opportunities. They enable traditional clinical risk monitoring models [18] to be complemented or potentially superseded by machine learning (ML) approaches.

ML has shown promise in identifying risk factors associated with suicide risk [19-21] and suicidal outcomes, including SI [22], behaviors [21,23], attempts [24], and completed suicides [25,26]. ML models use various input data types, from electronic health records [14,24,27] to textual data [19,28], collected in diverse settings [12,19,24,28,29]. These "theory-free" approaches [30] often outperform traditional detection methods [31].

Natural language processing advancements, particularly pretrained large language models based on transformer architectures, have significantly enhanced language classification tasks [32,33]. These "foundational models" [34], such as BERT, RoBERTa, LLaMA, and GPT, provide probabilistic language representations applicable to diverse tasks, opening new avenues for ML-based prediction in language-based psychological domains [34].

Language models provide rich text embeddings for transfer learning and fine-tuning [35,36], potentially enabling more accurate detection of SI and behaviors, even with limited domain-specific data [37]. Recent research has highlighted ML techniques, particularly foundational language models, in comprehensive suicidality prediction [38,39] and extraction of relevant textual indicators [40,41]. Transformer-based models outperform traditional ML approaches in suicide prediction using textual data, especially in social media datasets [38,42-45], though their real-world clinical applicability has remained unexplored.

The clinical implementation of ML models requires rigorous validation [46]. The "black box" nature of ML decisions necessitates explainable AI to emphasize interpretability for clinicians [30,31]. While traditional rule-based ML approaches such as decision trees offer straightforward explainability, the complexity of deep neural networks in foundational models presents interpretability challenges [47]. Model-agnostic post hoc explanation approaches, such as Shapley Additive Explanations (SHAPs), address this complexity [48-50], potentially elucidating key language features associated with suicide risk.

Deep neural network models have been applied to social media data [29,51-56], and explainable AI studies have been conducted on medical tabular data [57] and electronic health records [58]. However, studies using textual data from clinical populations, particularly helpline data, have not adequately addressed explainability.

Clinical implementations of ML as a diagnostic tool require rigor and transparency beyond interpretability. Ensuring model stability, tuning, and clinical value is crucial when considering the cost-benefit ratio of diagnostic decisions [59,60]. Metrics such as calibration and clinical utility [61-64] can emphasize the reliability and practical value of ML-based decisions in suicide risk prediction, though they remain underused in this context.

Our study aims to bridge the gap between state-of-the-art natural language processing techniques and real-world clinical applications in suicide prevention, addressing the lack of explainable transformer predictions for suicide risk assessment using clinical textual inputs validated for clinical utility. We develop and test a model for real-time crises within crisis text lines, integrating transformer-based models with SHAPs and incorporating clinical utility metrics.

XSL•FO
RenderX

We hypothesize that our transformer-based model will outperform the traditional baseline model in predicting SI and advanced suicidal engagement (ASE). Additionally, the SHAPs application is expected to highlight interpretable language features influencing the classification process, providing clinically relevant insights for decision-making in suicide prevention. This research aims to advance both the technical capabilities and practical applicability of ML in mental health assessment, potentially enhancing early intervention strategies in suicide prevention.

Methods

Overview

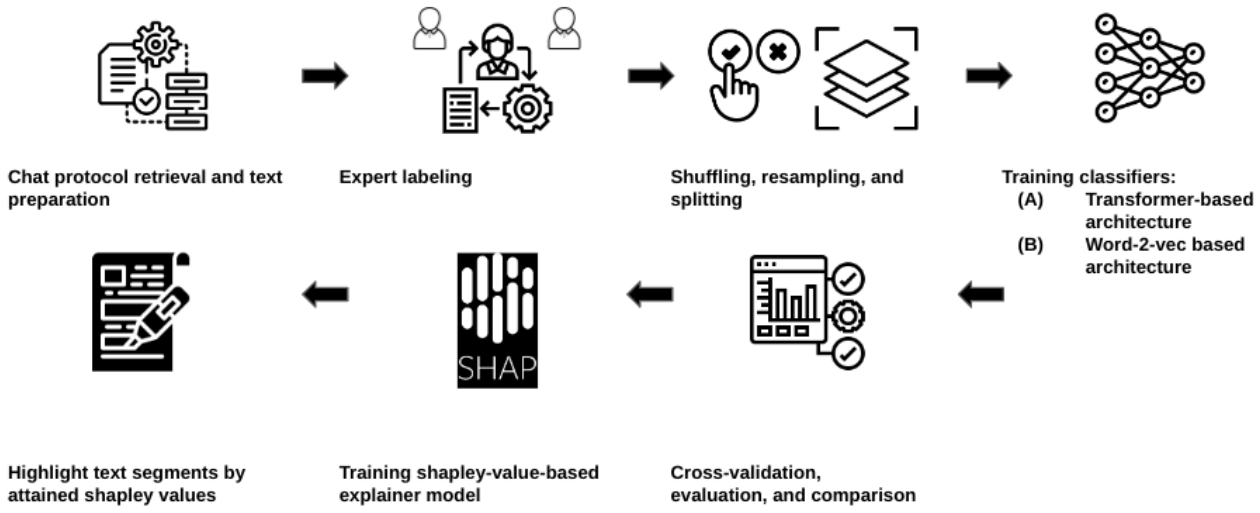
This study used a comprehensive approach to develop and evaluate ML models for identifying SI and behaviors in crisis helpline conversations. Our methodology encompassed several vital components, including data acquisition and preparation, model development and training, model evaluation and comparison, and explainability analysis. The statistical analysis comprised three main stages: preliminary sample analysis, development and training of two neural networks, and comprehensive evaluation of model performance. The following sections detail each component of our methodology, including data collection, preprocessing, model architecture, training procedures, and evaluation metrics.

Study Design

This study consists of several steps (Figure 1 [65-71]). We obtained chat protocols from krisenchat, a German crisis text line for youth. The dataset underwent cleaning, preprocessing, and expert labeling.

Post labeling, the data were randomized, partitioned, resampled, and encoded for model training. We then trained two text classification models and assessed model performance using various metrics. Lastly, we used a SHAPs explainer model to interpret the classification results.

Figure 1. Study design for suicidality prediction model development and evaluation using German youth crisis helpline session transcripts collected between November 11, 2021, and April 30, 2022: data flow, model development, and evaluation process for image sources.



Data Collection and Preprocessing

We obtained chat transcripts from krisenchat, a prominent German preclinical crisis intervention service for individuals up to 25 years of age [72]. The dataset comprised sessions conducted between November 30, 2021, and April 30, 2022. We excluded sessions from individuals younger than 14 years of age and those lacking age information due to informed consent considerations. The final dataset included transcripts from individuals experiencing suicidal thoughts and behaviors, as well as those seeking advice on other topics.

We defined a chat session as all messages from a single individual during a counseling session, typically at most 2 hours. We removed all personal identifiers and preprocessed the text data to enhance readability and reduce variability. This process eliminated extraneous elements such as links, HTML tags, and special characters beyond typical German punctuation.

From sessions marked with SI by krisenchat counselors, we identified relevant conversations addressing SI within this study's period. We excluded sessions too brief to assess suicidal ideation and behaviors (SIBs) or those terminating prematurely. Of 14,073 sessions, we identified 3193 control sessions, with 2886 meeting length and age criteria. We randomly selected 500 nonsuicidal sessions for manual review, replacing any indicating suicidal thoughts or severe self-harm with nonsuicidal cases. We used the imbalanced-learn library to address the class imbalance in our dataset by applying random oversampling of minority classes in the test set after splitting.

Measures and Labeling

We used the Columbia-Suicide Severity Rating Scale (C-SSRS), the gold standard for assessing SIB [73,74]. The C-SSRS categorizes ideation levels from a general desire to die to actual suicide attempts, quantifying severity and intensity. It defines suicidal behavior as any life-ending action, including nonharmful behaviors, interrupted attempts, and preparatory actions.

To address data imbalance and align with recent suicidology theories [75,76], we applied composite categories of SI and

XSL•FO
RenderX

ASE based on C-SSRS levels. This approach reflects the distinct processes of SI development and progression to suicidal behaviors [77]. The probability of acting on suicidal thoughts is influenced by an individual's capacity to endure pain and access to lethal means [78].

We defined SI using C-SSRS items "wish to be dead" and "nonspecific active suicidal thoughts," encompassing thoughts from the desire not to wake up to explicit life-ending statements. ASE comprised C-SSRS items involving specific methods, intent, or attempts, including contemplation of lethal methods, setting definite dates for action, and engaging in preparatory behaviors. The consideration of lethal means marks the transition to suicidal behaviors, defining the ASE category.

Three independent expert raters, extensively trained in C-SSRS, labeled the sessions. Their training involved mutual panel ratings and group discussions across 50 trial conversations not included in this study. Each rater subsequently assigned ratings independently. Raters conducted integrity discussions between rating sessions to maintain consistency and mitigate observer drift, exchanging thoughts and rating standards. Upon completing all session ratings, the raters convened to address discrepancies and reach a consensus on final ratings through comprehensive discussions.

Reporting Standards

This study adheres to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis for artificial intelligence (AI) abstract and reporting guidelines checklist for prediction model development and validation [79] (Tables S1 and S2 in Multimedia Appendix 1) and the CONSORT-AI (Consolidated Standards of Reporting Trials for Artificial Intelligence) checklist [80] (Table S3 in Multimedia Appendix 1).

Ethical Considerations

This study was conducted using ethical standards for human participant research. This study's protocol received approval from the Ethics Committee of the International Psychoanalytic University Berlin (2023_08). Informed consent was obtained through krisenchat's terms of service, which explicitly state that user data may be used for research purposes without direct identification of individuals. All personally identifiable information was removed from the chat transcripts during data preprocessing to protect privacy and confidentiality. Participants were not compensated directly, as this study used existing data from the crisis helpline. No images of participants were used in this study; the source text was solely used for training the classifier and was cleaned of all personally identifiable information. These measures ensure that no identification of individual participants is possible.

Software

All preprocessing, modeling, and evaluation tasks were performed using TensorFlow's [81] official Docker Image to construct a containerized GPU runtime on a Linux machine, using Python (version 3.8.1; Python Software Foundation) and CUDA (version 11.4; NVIDIA Corp). The machine had a 4 GB GPU and 16 GB of RAM. We used NumPy [82] and pandas [83] for preprocessing, while the Imblearn library was used for resampling. Model architectures and weights were sourced from the Hugging Face Transformers [35] library, and we used TensorFlow's Keras [84] module for modeling. Visualization and metrics were facilitated through matplotlib [85], seaborn [86], sci-kit-learn [87], and sklearn [88]. The final script is openly accessible on GitHub [89].

Model Architecture

Overview

One-way ANOVA to evaluate differences in word usage and age across 3 groups: nonsuicidal, SI, and ASE.

We developed two neural network models to classify SI and ASE in crisis interactions. Both models share a similar architecture, comprising three main blocks: an embedding block for the mathematical representation of messages, a long short-term memory (LSTM) block to model temporal dependencies, and a classification block.

Model 1, a transformer-multilayer perceptron (T-MLP), uses a pretrained transformer encoder (XLM-RoBERTa-base) [90], available on Hugging Face [35]. This multilingual model, trained on 2.5 TB of CommonCrawl data in 100 languages, tokenizes and encodes input text into a 768-dimensional embedding. Each message is embedded separately and then attached to an array of embedded messages. The embeddings feed into a time-distributed dense layer, followed by an LSTM encoder with dropout regularization, and finally, a multilayer perceptron for classification.

Model 2, a word2vector-multilayer perceptron (W2V-MLP), uses pretrained German word-vector embeddings (300-dimensional) processed through convolutional filters. These word-wise embeddings, trained using FastText on standard crawl and Wikipedia datasets, form an embedding matrix. This matrix is processed through a convolutional layer, max pooling, dropout, an LSTM encoder, and a multilayer perceptron layer for classification.

Both models output probabilities for three categories: not suicidal, SI, and ASE. Detailed model parameters are provided in Table S1 in Multimedia Appendix 2.

Model Training, Cross-Validation, and Evaluation

We implemented a three-phase evaluation and cross-validation protocol: (1) cross-validation of training, (2) post hoc estimation of overall evaluation metrics, and (3) assessment of calibration and clinical utility metrics.

To evaluate model stability, we performed 5 iterations of repeated cross-validation. We initialized data at random seeds, then shuffled and partitioned it (1011 for training and 337 for testing), maintaining consistent class ratios through stratified shuffling. We applied random oversampling to both sets individually to ensure balanced class ratios.

We limited training to a maximum of 100 epochs, incorporating an early stopping mechanism triggered by test accuracy, with a patience setting of 5 epochs. We selected the best model based on its area under the curve (AUC) performance, derived from receiver operating characteristics (ROCs).

XSL•FO
RenderX

We assessed model performance using standard class-wise metrics, including AUC, precision, recall, and F_1-scores, computing these metrics macroaveraged across classes. We determined class-wise metrics using a one-versus-all approach, binarizing predictions for each class. AUC measures the model's discriminative ability, with scores of 0.7-0.8 considered acceptable and 0.9 indicating high discrimination. We computed ROC curves using the one-versus-all method, quantifying the model's classification ability across all thresholds [80]. We applied bootstrapping with 1000 subsamples from the test dataset to establish 95% CIs [81]. We provided confusion matrices to visualize misclassifications [82].

To demonstrate clinical utility, we conducted decision curve analysis (DCA) for the SI and ASE classes [61,62]. DCA calculates a "net benefit" (NB) value based on the harm-to-benefit ratio of clinical decisions. An NB of 0.2 indicates the model detects 20 true positives without increasing false positives. We set threshold probabilities for SI between 0.3 and 0.5 and for ASE between 0.15 and 0.30 based on our sample's event rates and clinical considerations [63].

We assessed model calibration by calculating the overall Brier score (BS), which measures the mean square error of prediction probabilities against actual outcomes. Lower BS values indicate better calibration accuracy. We decomposed the BS into reliability, resolution, and uncertainty to offer insights beyond AUC-ROC [61]. We also calculated the Brier skill score to compare performance improvements of our transformer model over the word2vec model in calibration.

We visually assessed calibration using plots depicting the relationship between predicted probabilities and observed outcome frequencies [86]. A perfectly calibrated model aligns with the plot's diagonal; deviations indicate under- or overprediction.

Detailed DCA calculations, rationale, interpretations, and BS decomposition analysis are available in Table S3 in Multimedia Appendix 2.

Explainability Analysis

We used a SHAPs model-agnostic post hoc explainer to enhance our understanding of the models' prediction processes. Shapley values, derived from cooperative game theory [91], provide local explanations for each prediction [92]. These values quantify each feature's impact on an instance's prediction relative to its average impact across other feature combinations. The algorithm uses tokenized text vectors as inputs and considers all possible feature combinations to calculate these contributions.

In addition to examining individual results, we determined the overall importance of specific language features in predicting nonsuicidal, SI, or ASE categories. This approach provides a comprehensive understanding of the linguistic markers that influence the model's decision-making process.

Results

Overview

The final dataset comprised 1348 eligible sessions, partitioned into 1011 training and 337 testing cases. After resampling, the training set expanded to 1254 cases, and the test set to 420 cases. We provide dataset characteristics in Table S2 in Multimedia Appendix 2. The mean participant age was 17.9 (95% CI 17.7-18.1) years. We observed significant differences in word count across groups ($F_{2,97}$=48.34, P<.001).

Model Training, Evaluation, and Cross-Validation and Evaluation

The training curves demonstrated that the transformer model learned consistently and stably (Figure S1 in Multimedia Appendix 3). In contrast, the baseline word vector model rapidly reached its learning peak before exhibiting overfitting, halted only by the early stopping mechanism (Figure S1 in Multimedia Appendix 3). The Training Curves section in Multimedia Appendix 3 provides a more detailed interpretation of these learning curves.

The transformer model (T-MLP) achieved superior performance with an overall accuracy of 0.79 (95% CI 0.73-0.99) and a macroaveraged AUC-ROC of 0.89 (95% CI 0.81-0.91). It demonstrated excellent discrimination for nonsuicidal sessions (AUC-ROC=0.96, 95% CI 0.96-0.99) and good discrimination for SI and ASE with AUC-ROCs of 0.85 (95% CI 0.97-0.86) and 0.87 (95% CI 0.81-0.88), respectively. The baseline word2vec model (W2V-MLP) showed lower performance across all metrics (Table 1).

XSL•FO
RenderX

Table 1. Classification performance metrics of transformer-based (T-MLP[a]) and word vector-based (W2V-MLP[b]) models for predicting NS[c], SI[d], and ASE[e] in German youth crisis helpline users (N=337 evaluation set).

Model and class	Precision (positive predictive value)	Recall (sensitivity)	F_1-score (harmonic mean of precision and recall)	AUC[f]-ROC[g] (95% CI)
T-MLP				0.89 (0.81-0.91)
NS	0.97	1	0.98	0.96 (0.96-0.99)
SI	0.71	0.65	0.68	0.85 (0.97-0.86)
ASE	0.69	0.72	0.71	0.87 (0.81-0.88)
W2V-MLP				0.77 (0.64-0.9)
NS	0.61	0.86	0.71	0.89 (0.84-0.9)
SI	0.59	0.69	0.61	0.78 (0.68-0.78)
ASE	0.67	0.28	0.39	0.64 (0.62-0.72)

[a]T-MLP: transformer-multilayer perceptron.

[b]W2V-MLP: word2vector-multilayer perceptron.

[c]NS: not suicidal.

[d]SI: suicidal ideation.

[e]ASE: advanced suicidal engagement.

[f]AUC: area under the curve.

[g]ROC: receiver operating characteristic.

Evaluation and Cross-Validation

Class-wise metrics demonstrate the transformer model's consistent superiority over the baseline model across all classification metrics. Figure 2 illustrates the performance of the T-MLP model. The AUC-ROC analysis (Figure 2A) confirms the model's high discriminative ability across different classes and threshold probabilities. Precision-recall curves (Figure 2B) reveal some uncertainty at lower recall thresholds, suggesting potential for improved calibration. The confusion matrix (Figure 2C) demonstrates the model's effectiveness in distinguishing nonsuicidal cases while showing some inconsistency in classifying SI and ASE.

The T-MLP model exhibits a low false negative rate of 0.01, indicating high accuracy in identifying suicidal sessions. Conversely, the W2V-MLP model (Figure 3) demonstrates inferior performance. Its AUC-ROC analysis (Figure 3A) and precision-recall curves (Figure 3B) show lower discriminative ability. The confusion matrix (Figure 3C) reveals a tendency to over-predict ASE, with a higher false negative rate of 0.17 and a false positive rate of 25% for nonsuicidal outcomes.

These results underscore the T-MLP model's enhanced capability in accurately classifying SIBs compared to the baseline W2V-MLP model while also highlighting areas for potential improvement in model calibration.

DCA demonstrates the T-MLP model's superior clinical utility. For SI prediction, the T-MLP model yields higher NBs than treat-all and treat-none strategies within 0.3-0.5 threshold probabilities, identifying 10-18 true positives without increasing false positives (Figure 4A). The W2V-MLP model, while initially beneficial, exhibits potentially detrimental clinical decisions beyond 0.4 thresholds (Figure 5A).

For ASE, the T-MLP model outperforms alternative strategies within 0.15-0.3 thresholds, achieving an NB of 0.20-0.25 (Figure 4B). Conversely, the W2V-MLP model underperforms relative to the treat-all strategy, potentially leading to suboptimal clinical decisions (Figure 5B). These findings underscore the T-MLP model's enhanced clinical applicability in suicide risk assessment.

We assessed model calibration using Brier and Log loss scores. The T-MLP model achieved lower scores (Brier: 0.10, Log loss: 0.50) compared to the W2V-MLP model (Brier: 0.18, Log loss: 0.92), indicating superior calibration performance. However, both models exhibited resolution challenges in specific classes (Table 2, and Figures 5 and 6).

The Brier Skill Score revealed a 44.4% improvement in classification performance by the T-MLP model over the baseline. Both models demonstrated low variability in predicted probabilities across samples, with an overall uncertainty of 0.22 due to class rebalancing.

Calibration curve analysis (Figure 6) exposed distinct prediction patterns. The T-MLP model consistently overpredicted for the nonsuicidal class across all decision thresholds (Figure 6A) while underpredicting SI and ASE at probabilities above 0.5 and overpredicting at lower thresholds. The W2V-MLP model underpredicted for the nonsuicidal class up to a 0.6 threshold and overpredicted suicidal outcomes at lower thresholds (Figure 6B).

These calibration patterns suggest potential overestimation of clinical utility within predefined thresholds for both models, necessitating cautious interpretation of the DCA results.

XSL•FO
RenderX

Figure 2. Performance analysis of the transformer-based model (T-MLP) for predicting not suicidal, SI, and ASE based on session transcripts of German youth crisis helpline users between November 30, 2021, and April 30, 2022 (N=337 evaluation set). (A) Class-wise one-versus-all AUC-ROC analysis with 95% CIs derived from 1000 bootstrap samples. (B) Class-wise precision-recall analysis with 95% CIs derived from 1000 bootstrap samples. (C) Confusion matrix of the test set: light colors represent low numbers, dark colors represent high numbers; correct classifications are on the diagonal from top-left to bottom-right. ASE: advanced suicidal engagement; AUC: area under the curve; SI: suicidal ideation; ROC: receiver operating characteristic; T-MLP: transformer-multilayer perceptron.

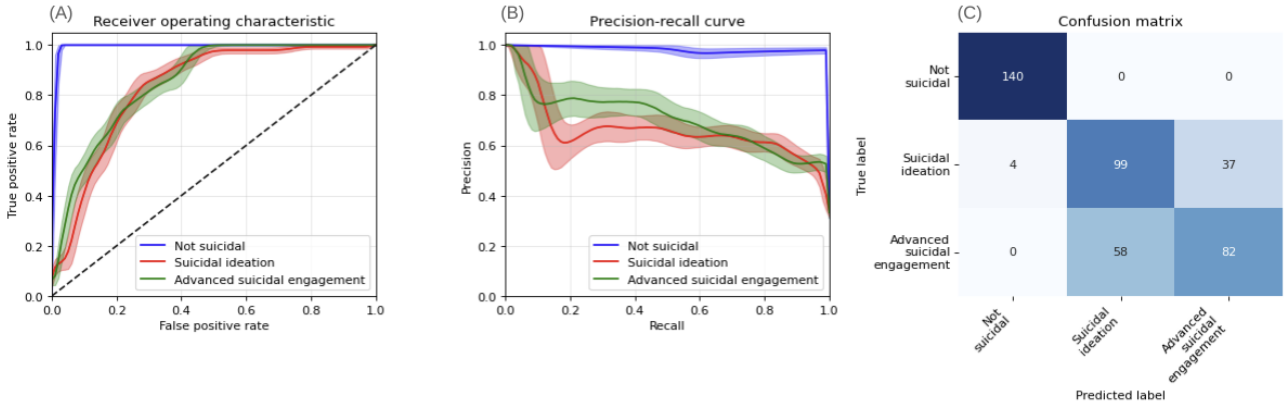


Figure 3. Performance analysis of word vector-based model (W2V-MLP) for predicting not suicidal, SI, and ASE in German youth crisis helpline users (N=337 evaluation set). (A) Class-wise one-versus-all AUC-ROC analysis with 95% CIs derived from 1000 bootstrap samples. (B) Class-wise sensitivity-specificity analysis with 95% CIs derived from 1000 bootstrap samples. (C) Confusion matrix of the test set: light colors represent low numbers, dark colors represent high numbers; correct classifications are on the diagonal from top-left to bottom-right. ASE: advanced suicidal engagement; AUC: area under the curve; SI: suicidal ideation; ROC: receiver operating characteristic; W2V-MLP: word2vector-multilayer perceptron.

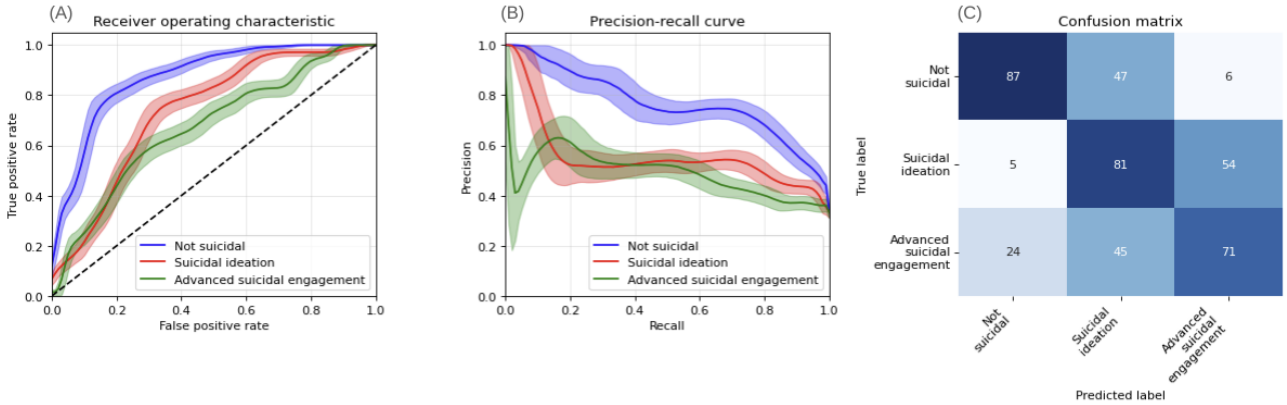
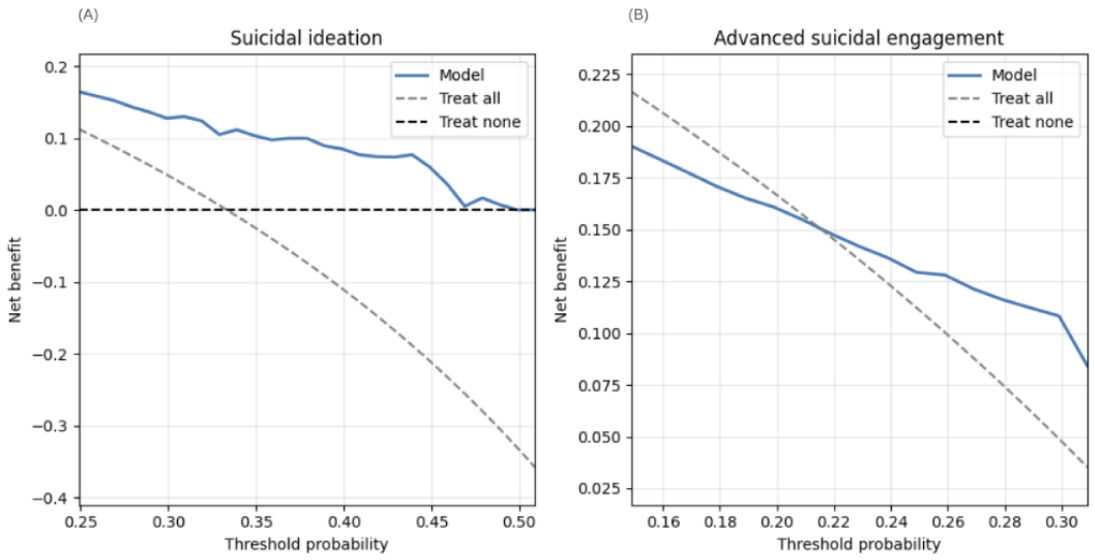


Figure 4. Decision curve analysis of the transformer-based model (T-MLP) for predicting suicidal ideation and advanced suicidal engagement in German youth crisis helpline users (N=337 evaluation set). (A) Class-wise decision curve analysis for suicidal ideation. (B) Class-wise decision curve analysis for advanced suicidal engagement. Color coding: treat-all strategy in gray, treat-none strategy in black, classifier (T-MLP) in blue, theoretical maximal benefit in red. Irrelevant threshold segments are grayed out. T-MLP: transformer-multilayer perceptron.



XSL•FO
RenderX

Figure 5. Decision curve analysis of word-vector model (T-MLP) for predicting suicidal ideation and advanced suicidal engagement in German youth crisis helpline users (N=337 evaluation set). (A) Class-wise decision curve analysis for advanced suicidal engagement. (B) Class-wise decision curve analysis for advanced suicidal engagement. Color coding: treat-all strategy in gray, treat-none strategy in black, classifier (T-MLP) in blue, theoretical maximal benefit in red. Irrelevant threshold segments are grayed out. T-MLP: transformer-multilayer perceptron.

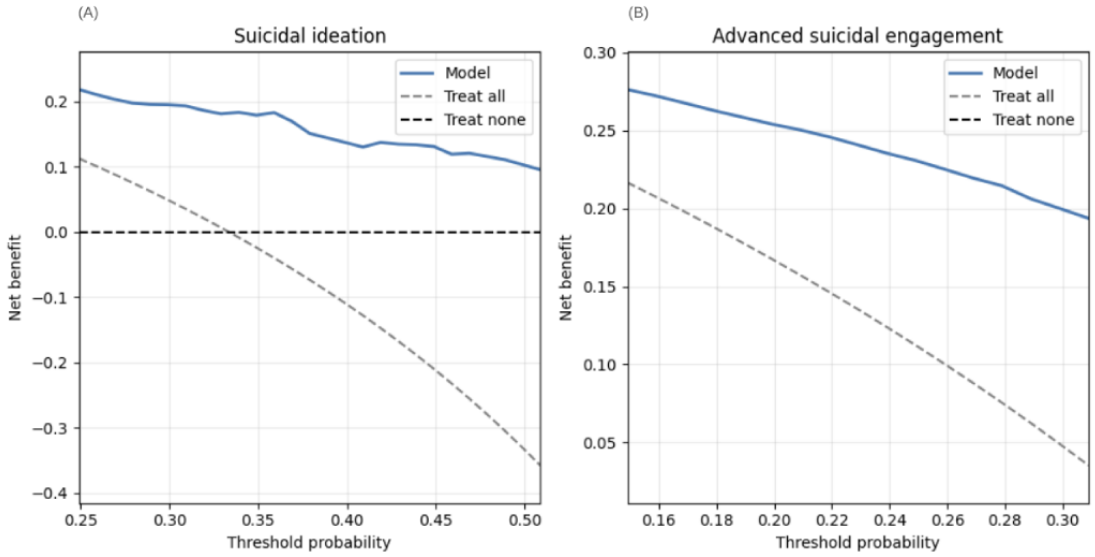
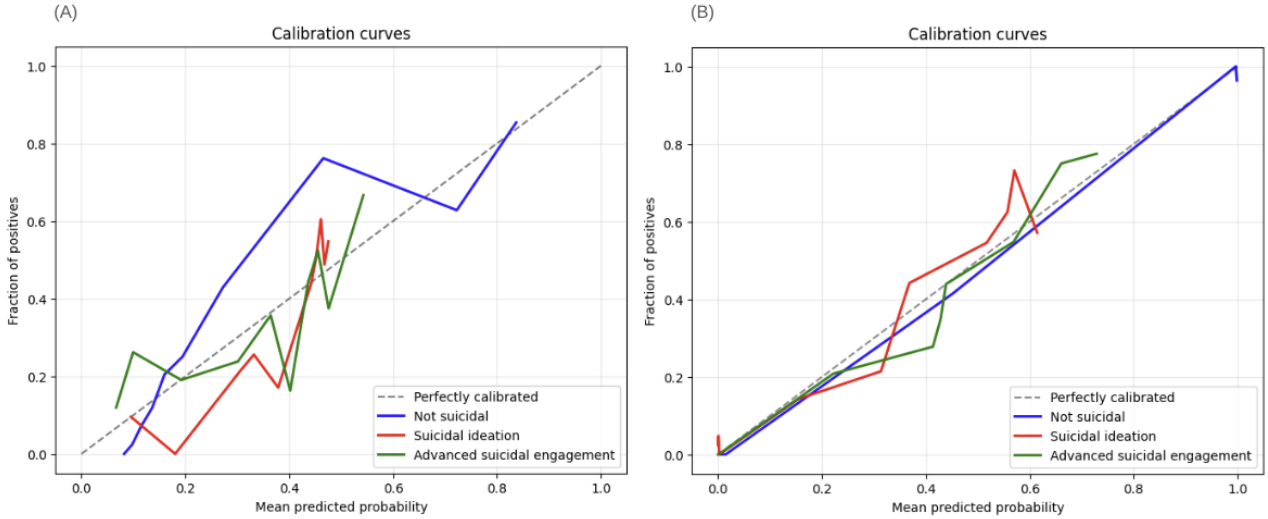


Figure 6. Calibration analysis of transformer-based (T-MLP) and word vector-based (W2V-MLP) models for predicting nonsuicidal, suicidal ideation, and advanced suicidal engagement based on session transcripts of German youth crisis helpline users between November 30, 2021, and April 30, 2022 (N=337 evaluation set). (A) Class-wise calibration curves for the T-MLP model. (B) Class-wise calibration curves for the W2V-MLP model. Color coding: not suicidal (blue), suicidal ideation (orange), and advanced suicidal engagement (green). T-MLP: transformer-multilayer perceptron; W2V-MLP: word2vector-multilayer perceptron.



XSL•FO
RenderX

Table 2. Brier score decomposition and model comparison for transformer-based (T-MLP[a]) and word vector-based (W2V-MLP[b]) models in predicting NS[c], SI[d], and ASE[e] based on session transcripts of German youth crisis helpline users between November 30, 2021, and April 30, 2022 (N=337 evaluation set).

	Transformer[f]			word2vec[g]		
Class	NS	SI	ASE	NS	SI	ASE
Reliability	0.002	0.007	0.008	0.001	0.010	0.011
Resolution	0.219	0.863	0.106	0.06	0.015	0.023
Uncertainty	0.222	0.222	0.222	0.222	0.222	0.222

[a]T-MLP: transformer-multilayer perceptron.

[b]W2V-MLP: word2vector-multilayer perceptron.

[c]NS: not suicidal

[d]SIS: suicidal ideation.

[e]ASE: advanced suicidal engagement.

[f]Average Brier loss=0.10; Brier skill score=44.4%.

[g]Average Brier loss=0.18; Brier skill score=44.4%.

Explainability Analysis

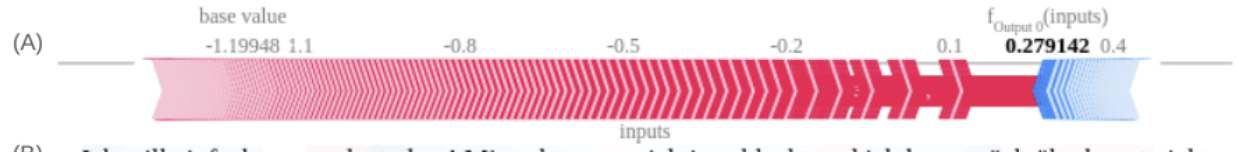
We implemented SHAPs to generate additive text explanations for the transformer model. Due to privacy constraints, we present a vignette illustrating the model's interpretation of feature importance rather than specific outputs.

Global feature importance analysis revealed linguistic traits associated with SI, including self-reference, negation, expressions of low self-esteem, problem emphasis, and articulations of isolation. ASE texts exhibited similar characteristics with additional absolutist language and frequent references to lethal methods and injuries.

Figures 7A and 7B depict a synthetic session vignette predicting the nonsuicidal group, highlighting de-escalating statements in the latter half of the conversation as influential for a positive prediction. Figures 8A, 8B, 9A, and 9B illustrate tokens associated with SI and ASE, respectively. These figures emphasize self-referential statements, expressions indicating a desire to end one's life, and words associated with negative sentiment.

It is noteworthy that in transformer architectures, word tokens do not have fixed impacts on the output due to their context-dependent evaluation. This contextual approach, while representing an advancement over other word-based methods, limits the generalizability of derived Shapley values and plots. Interestingly, the model did not assign high importance to the misspelled word "sschluss gemacht" (intended as "schluss gemacht" meaning "broke up"), demonstrating its robustness to typographical errors.

Figure 7. Shapley values-based text explanation plot for nonsuicidal outcome prediction in German youth crisis helpline users, plot based on synthetic case vignette. (A) Force plot of aggregated Shapley values indicating the predominance of not suicidal features over SI-associated risk markers. (B) The text highlights illustrating content associated with the target class (not suicidal) in red and suicidal content in blue. SI: suicidal ideation.



(B) Ich will einfach nur noch sterben! Mir geht es sor richtig schlecht und ich kann müch überhaupt nicht mehr einkriegen. Hab einfach die ganze Nacht geheult und so. Mein Freund hat vorgestern sschluss gemacht und ich weiß nicht wie ich ohne ihn leben soll! Hm. Ja, na mit mit meiner Mama kann üch berhaupt nicht reden, die sagt mir nur die ganze Zeit was ich falsch gemacht hab. Ne, ich hab mir noch nichts überlegt wie ich es mache. Ich bin mir auch nicht so sicher, ich glaube ich würde mich das auch nicht trauen, es würde meine Omi mega unglücklich machen und das will ich nicht. Danke fürs zuhören, es hat schon sehr geholfen dass du einfach da warst.

XSL•FO
RenderX

Figure 8. Shapley values-based text explanation plot for suicidal ideation outcome prediction in German youth crisis helpline users, plot based on synthetic case vignette. (A) Force plot of aggregated Shapley values indicating the balance between suicidal ideation-associated risk markers and nonsuicidal features. (B) Text highlights illustrating content associated with the target class (suicidal ideation) in red and nonsuicidal content in blue.

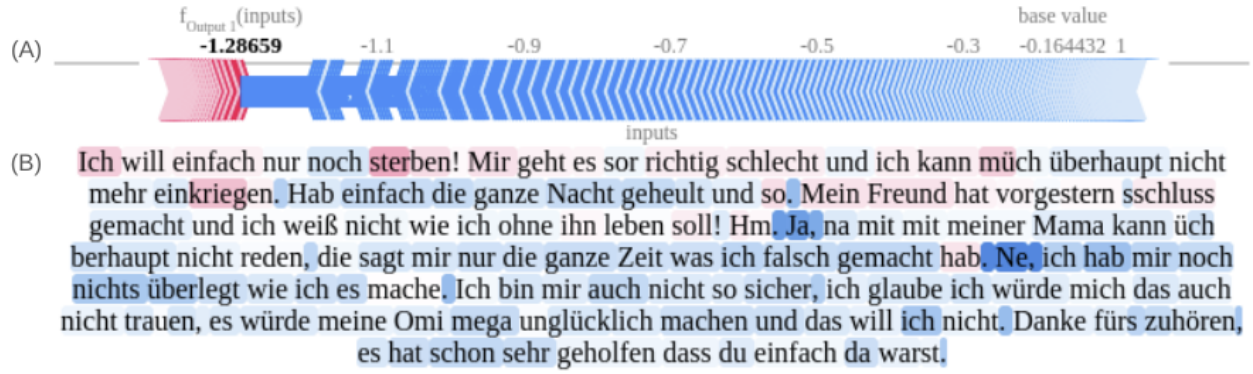
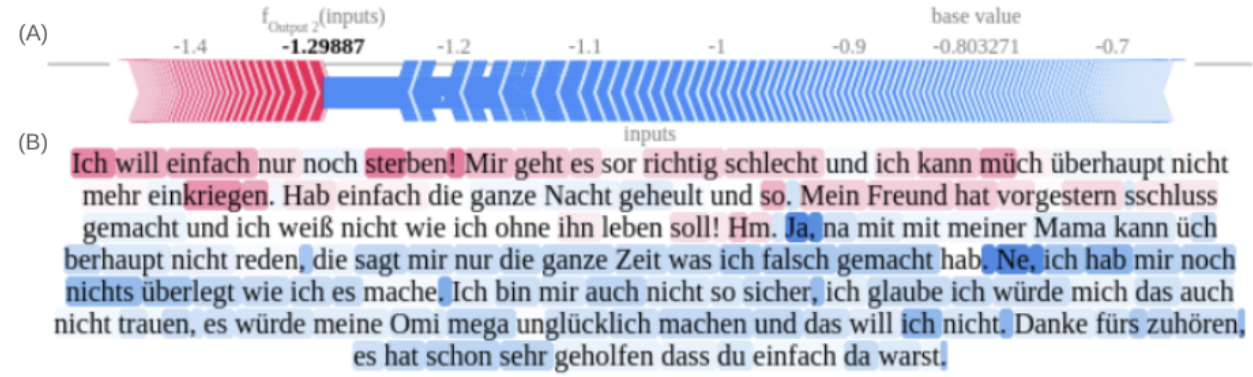


Figure 9. Shapley values-based text explanation plot for advanced suicidal engagement outcome prediction in German youth crisis helpline users, plot based on synthetic case vignette. (A) Force plot of aggregated Shapley values indicating the balance between advanced suicidal engagement–associated risk markers and nonsuicidal features. (B) Text highlights illustrate content associated with the target class (advanced suicidal engagement) in red and nonsuicidal content in blue. Translation: “I just want to die right now. I am feeling pretty bad, and I can’t calm myself down. I was crying the whole night and stuff. My boyfriend dumped me the day before yesterday, and I don't know how to live without him!” “Hm. Yeah, I can't talk to my mum about such stuff. She just points out my flaws all the time.” “No, I don't have a specific Plan of how to do it; I am also pretty unsure - I think I would never have the courage to do it, my Granny would be super sad and I don't want that. Thank you for listening to me. Thank you for being there for me. This helped a lot.”.



Discussion

Overview

We aimed to develop, evaluate, cross-validate, and benchmark a transformer-based prediction model for detecting SI and ASE among adolescents seeking emergency online counseling services in Germany.

The primary findings of this study demonstrate the proficiency of our transformer-based prediction model in identifying SI and ASE among adolescents using online counseling services. Our model outperformed the baseline word-vector approach, exhibiting excellent to very good performance in distinguishing between nonsuicidal sessions, SI, and ASE. Additionally, the Shapley analysis successfully identified word tokens associated with suicidal outcomes, providing valuable insights into language features indicative of SI. These findings highlight the potential of ML models in enhancing suicide prevention efforts.

Principal Results

Our key findings underscore the superior performance of our transformer-based prediction model in identifying SI and ASE using natural language data from adolescents. The model was trained on rigorously collected data evaluated by 3 expert raters using an established and validated scale. We adhered to a strict validation protocol, including repeated cross-validation and clinical utility metrics, to ensure our results are reliable, interpretable, and applicable in clinical settings.

The transformer model showed strong capabilities in differentiating between session transcripts containing nonsuicidal statements and those featuring discussions of SI or ASE, despite some calibration challenges. These findings suggest the model’s potential as a valuable screening tool for identifying individuals who may be at risk for SI and ASE, potentially supporting clinical decision-making and enabling timely interventions. The model excelled at recognizing transcripts with nonsuicidal content, though it faced some difficulties in precisely distinguishing more nuanced levels of

XSL•FO
RenderX

SIB. Notably, our model substantially surpassed the performance of the baseline model that used pretrained word vectors, demonstrating improved overall metrics and a reduced risk of misclassification. While the model achieved high sensitivity and specificity in identifying nonsuicidal cases, its specificity for SI and ASE was comparatively lower.

Analysis of calibration curves indicates a necessity for recalibration before clinical implementation, as current reliability metrics suggest some degree of instability. Although the transformer model exhibits evident clinical utility, the observed calibration issues may lead to an overestimation of its practical performance. Furthermore, while the application of SHAPs analysis effectively identified word tokens associated with suicidal outcomes, it demonstrated reduced effectiveness in interpreting out-of-context tokens. This limitation potentially constrains the model's capacity to comprehensively capture the dynamic nature of youth language patterns.

Overall, the transformer model outperformed the word-vector model, demonstrating superior performance and higher clinical utility, even when trained on a relatively small, domain-specific dataset. This advantage primarily stems from the transformer encoders' advanced ability to represent natural language, partly due to their contextual embedding strategy and training on larger datasets. Additionally, the architecture inherently excels at analyzing deeper levels of language structure.

Our findings offer strong support for the use of ML models, particularly those using transformers and transfer learning, in detecting and classifying levels of SI in clinical psychological and psychiatric research that relies heavily on text and speech data. The transformer model surpasses both the treat-all and treat-none strategies and the baseline model for detecting SI and ASE. Including an explainability component provides clinicians with insights into the specific language features associated with SI or advanced engagement [93].

The SHAPs algorithm generated local explanations for the model's predictions, which we aggregated to gain insight into the global importance of features in the language of suicidal adolescents. The explainer model effectively pinpointed specific language markers that influence the model's predictions. The importance of explainability is underscored by regulations such as the AI Act of the European Union, which emphasizes transparency in AI applications [94].

The language markers identified by the model are consistent with existing research on suicidal language markers, validating our results. Notably, our analysis confirmed that the usage of first-person singular pronouns and absolutist language is linked to SI [95-97].

This study demonstrates the potential of our model to accurately assess and identify suicidal thoughts in adolescents, with significant implications for prevention strategies addressing SIBs. The model's high applicability to the online crisis counseling context facilitates its potential adoption by other online counseling services, potentially aiding in risk mitigation and supporting clinical decision-making processes. The model's capacity to identify key risk markers may enable more precise risk assessments and inform targeted interventions for individuals at elevated risk of ASE. Furthermore, the model's ability to track and detect changes in language usage over time could contribute to earlier identification of individuals experiencing increased risk.

In summary, the developed model can significantly enhance the effectiveness and efficiency of prevention efforts at any level of SI.

Limitations

The primary limitation of this study is the small sample size, which may restrict the generalizability of the model's performance. As evidenced by the learning curves (Figures S1 and S2 in Multimedia Appendix 3), the transformer model converged with minimal overfitting in contrast to the baseline model. The limited sample size might have influenced class learnability, particularly for the word-vector approach, which requires more parameter estimation. Repeated cross-validation revealed training instability and sensitivity to weight initialization, potentially leading to class-specific biases. Calibration metrics indicated significant reliability issues for both models, suggesting the need for recalibration before clinical implementation. Due to these calibration concerns, the clinical utility of the models should be interpreted cautiously, as small sample sizes introduce more noise to neural networks [98]. These issues could be mitigated by using larger datasets, implementing robust weight initialization, or applying additional regularization techniques.

Furthermore, due to metric requirements, we applied resampling techniques to the rarer outcomes, which may lead to overfitting minority classes and lower generalization. Performance improvements over the baseline may be overstated due to less emphasis on feature engineering in the word-vector model. Incorporating n-grams, tf-idf (term frequency–inverse document frequency), or dictionary-based features could address the lack of sequential text structure in word embeddings, which this study did not fully address.

Potential age or gender biases in the expression of suicidality may exist, as the sample was not gender-balanced, potentially skewing results toward female expressions of SIBs. The class imbalance necessitated the use of oversampling techniques, which might introduce biases. Technical limitations led to session truncations, affecting data quality and representativeness.

This study's exclusive focus on textual data, while appropriate for the crisis text line setting, may limit the comprehensive assessment of suicidality. Suicidal behavior is multifaceted, and other data types such as audio or visual cues could provide additional valuable information. The absence of these modalities in our analysis is a limitation inherent to the text-based nature of the crisis line service studied.

Token-based post hoc explanations might not fully capture the sequential nature of speech, possibly missing complex effects detected by transformers and LSTM layers. Future research should explore specific techniques for transformer-based explainability and evaluate the linguistic features of highlighted tokens. Limiting analysis to text tokens may overlook the importance of other communication forms, such as emoticons,

XSL•FO
RenderX

which play a significant role in conveying meaning. Future models should consider these aspects.

Additionally, the vignette demonstrated that despite being trained on more extensive data, the transformer encoder still will not fully capture irregular or misspelled language. Adapting to new language variations, often developed among youth, heavily depends on collecting domain-specific datasets for embedding models.

Comparison to Prior Work

The discriminative abilities of our model are comparable, albeit slightly lower, to those reported by Zhang et al [99], who analyzed 659 suicide notes, 431 last statements, and 2000 neutral posts. Their model achieved 95%, 94.9%, and 94.9% in precision, recall, and F_1-score, respectively, although it was limited to binary targets and explicitly focused on suicide notes. Similarly, Broadbent et al [28] found that using sentence-wise embeddings of counseling sessions, a transformer embedding surpassed a baseline tf-idf model in reducing the false-negative rate.

In contrast, Aldhyani et al [100] demonstrated that a convolutional neural network and bidirectional LSTM model excelled over the XGBoost model, achieving 95% accuracy in detecting SI compared to the latter's 91.5%. However, when using validated dictionary features such as Linguistic Inquiry and Word Count, boosted tree algorithms performed better than a neural network architecture combining convolutional and LSTM layers (convolutional neural network and bidirectional LSTM) on binary targets. The key distinction of this study from Aldhyani et al [100] lies in the emphasis on feature engineering and the use of dictionaries, which were leading methods for building text classifiers before the emergence of transfer learning with large pretrained language models [62].

Unique to our study, we also include metrics for calibration and clinical utility, enhancing the practical value of our findings.

Generalizability

The generalizability of our findings varies by domain. Given the small and domain-specific sample, the results may primarily apply to German-speaking adolescents seeking mental health assistance. However, the analysis of language features suggests that findings might extend to a broader demographic group.

Clinical Implications and Implications for Future Research

Our transformer-based model has the potential to assist clinicians in identifying at-risk individuals and improving intervention prioritization. Continuous monitoring using such models is cost-effective and can detect cases that might otherwise be missed, facilitating referral to qualified human review and leading to more accurate and timely interventions. This research significantly contributes to prevention efforts addressing SIBs.

Despite the promise of such models, as indicated in our study, future research could enhance the model's accuracy and utility. Subsequent studies should aim to collect sample sizes that are adequate for the power requirements of ML. If faced with the common problem of class imbalance, researchers could opt for adjusting the cost function of the learner or using synthetic data generation through augmentation to mitigate this issue.

Other key priorities include validating the model across diverse clinical settings such as inpatient facilities, rural clinics, and school-based mental health services to enhance its generalizability. Additionally, evaluating the model's performance across different demographic groups and cultural contexts will be essential to ensure its effectiveness and identify potential biases.

Developing dynamic models that track changes over time and incorporate additional factors such as demographics could improve performance. This could be achieved by using retrained transformers in extracting longitudinal event data [101] or by incorporating explanations into the clinicians' counseling environment [102]. Integrating transformer embeddings into multimodal graph network models might refine precision and bolster prevention efforts by addressing the multifactorial nature of suicide.

Integrating multiple data types could provide a more holistic view of an individual's mental state. However, such approaches would require careful consideration of privacy concerns, data integration challenges, and the development of more complex models capable of processing diverse input types. Future studies could explore the relative contributions of different data modalities to suicide risk assessment and the feasibility of implementing multimodal systems in various clinical settings.

ML shows potential in mental health applications, but it should be considered a complementary tool to clinical expertise and evidence-based decision-making rather than a standalone solution. In suicidology, it is imperative to recognize that suicide risk emerges from complex life circumstances that require comprehensive social interventions, support systems, and targeted public policies. While ML algorithms can serve as diagnostic tools to detect linguistic markers associated with suicide risk in individuals, they cannot independently prevent or resolve the underlying personal challenges. Furthermore, the implementation of ML in this context necessitates careful ethical considerations to ensure that these tools augment human judgment without introducing or amplifying biases. Future research should focus on integrating ML approaches within broader, multifaceted suicide prevention strategies that address the full spectrum of risk factors and protective measures.

Conclusions

Suicide prevention represents a critical public health concern, particularly among adolescents. Timely identification of individuals displaying SIBs is crucial for early intervention. Our study introduces an explainable transformer-based ML model that outperforms a baseline word-vector approach in identifying SI and ASE in adolescents seeking help via a German crisis helpline. This model could be invaluable for clinicians prioritizing intervention cases. However, challenges remain in discerning finer-grained types of SI. This research underscores the potential of ML in detecting SI and represents a step toward more effective suicide prevention. Future work

XSL•FO
RenderX

should focus on enhancing model accuracy in longitudinal setups or incorporating multimodal features.

Acknowledgments

The publication was funded and supported by the Open Access Fund of Universität Trier and by the German Research Foundation (DFG).

Data Availability

This study used highly sensitive and confidential crisis helpline chat transcripts from vulnerable individuals, including minors. Due to the nature of these conversations, which frequently contain personal details and sensitive information regarding mental health and SI, the raw data cannot be made publicly available. This restriction is necessary to protect participant privacy and confidentiality and to comply with ethical guidelines and data protection regulations, including the General Data Protection Regulation. Even in an anonymized form, the risk of reidentification cannot be completely mitigated, thus precluding data sharing.

Conflicts of Interest

JT, AL, RW, and JS are or were employed and receive a salary from krisenchat. krisenchat did not influence aspects like data collection, analysis, and reporting of results. GM received funding from the Stanley Thomas Johnson Stiftung and Gottfried und Julia Bangerter-Rhyner-Stiftung (project PC 28/17 and PC 05/18), from Gesundheitsförderung Schweiz (project 18.191/K50001), from the Swiss Heart Foundation (project =FF21101), from the Research Foundation of the International Psychoanalytic University Berlin (project 5087 and 5217), from the Swiss National Science Foundation (SNSF; project 100014_135328), from the German Federal Ministry of Education and Research (budget item 68606) in the context of an evaluation project conducted among others in collaboration with krisenchat, from the Hasler Foundation (project 23004), in the context of a Horizon Europe project from the Swiss State Secretariat for Education, Research and Innovation (SERI; contract 22.00094), and from Wings Health in the context of a proof-of-concept study. GM is a cofounder, member of the board, and shareholder of Therayou AG, active in digital and blended mental health care. GM receives royalties from publishing companies as author, including a book published by Springer, and an honorarium from Lundbeck for speaking at a symposium. Furthermore, GM is compensated for providing psychotherapy to patients, acting as a supervisor, serving as a self-experience facilitator (“Selbsterfahrungsleiter”), and for postgraduate training of psychotherapists and supervisors.

Multimedia Appendix 1
Reporting guideline checklists.
[DOCX File , 37 KB - publichealth_v11i1e63809_app1.docx]

Multimedia Appendix 2
Supplementary tables and figures.
[DOCX File , 52 KB - publichealth_v11i1e63809_app2.docx]

Multimedia Appendix 3
A primer on neural networks and metrics.
[DOCX File , 256 KB - publichealth_v11i1e63809_app3.docx]

References

1. Hink AB, Killings X, Bhatt A, Ridings LE, Andrews AL. Adolescent suicide-understanding unique risks and opportunities for trauma centers to recognize, intervene, and prevent a leading cause of death. Curr Trauma Rep 2022;8(2):41-53 [FREE Full text] [doi: 10.1007/s40719-022-00223-7] [Medline: 35399601]
2. Becker M, Correll CU. Suicidality in Childhood and Adolescence. Dtsch Arztebl Int 2020 Apr 10;117(15):261-267. [doi: 10.3238/arztebl.2020.0261] [Medline: 32449889]
3. Shneidman ES. Definition of Suicide. New Jersey: Wiley; 1985.
4. Harmer B, Lee S, Rizvi A, Saadabadi A. Suicidal ideation. StatPearls. Florida: StatPearls Publishing URL: http://www.ncbi.nlm.nih.gov/books/NBK565877 [accessed 2024-07-29]
5. O'Carroll PW, Berman AL, Maris RW, Moscicki EK, Tanney BL, Silverman MM. Beyond the tower of babel: a nomenclature for suicidology. Suicide Life Threat Behav 1996;26(3):237-252. [doi: 10.1111/j.1943-278x.1996.tb00609.x]
6. Campisi SC, Carducci B, Akseer N, Zasowski C, Szatmari P, Bhutta ZA. Suicidal behaviours among adolescents from 90 countries: a pooled analysis of the global school-based student health survey. BMC Public Health 2020;20(1):1102 [FREE Full text] [doi: 10.1186/s12889-020-09209-z] [Medline: 32772922]

XSL•FO
RenderX

7. Turecki G, Brent DA, Gunnell D, O'Connor RC, Oquendo MA, Pirkis J, et al. Suicide and suicide risk. Nat Rev Dis Primers 2019;5(1):74 [FREE Full text] [doi: 10.1038/s41572-019-0121-0] [Medline: 31649257]
8. Benton TD, Muhrer E, Jones JD, Lewis J. Dysregulation and suicide in children and adolescents. Child Adolesc Psychiatr Clin N Am 2021;30(2):389-399. [doi: 10.1016/j.chc.2020.10.008] [Medline: 33743946]
9. Viner RM, Ozer EM, Denny S, Marmot M, Resnick M, Fatusi A, et al. Adolescence and the social determinants of health. Lancet 2012;379(9826):1641-1652. [doi: 10.1016/S0140-6736(12)60149-4] [Medline: 22538179]
10. Pelkonen M, Marttunen M. Child and adolescent suicide: epidemiology, risk factors, and approaches to prevention. Paediatr Drugs 2003;5(4):243-265. [doi: 10.2165/00128072-200305040-00004] [Medline: 12662120]
11. Richards D, Viganó N. Online counseling: a narrative and critical review of the literature. J Clin Psychol 2013;69(9):994-1011. [doi: 10.1002/jclp.21974] [Medline: 23630010]
12. Grimland M, Benatov J, Yeshayahu H, Izmaylov D, Segal A, Gal K, et al. Predicting suicide risk in real-time crisis hotline chats integrating machine learning with psychological factors: exploring the black box. Suicide Life Threat Behav 2024;54(3):416-424. [doi: 10.1111/sltb.13056] [Medline: 38345174]
13. Cunningham CE, Walker JR, Eastwood JD, Westra H, Rimas H, Chen Y, The Mobilizing Minds Research Group. Modeling mental health information preferences during the early adult years: a discrete choice conjoint experiment. J Health Commun 2014;19(4):413-440 [FREE Full text] [doi: 10.1080/10810730.2013.811324] [Medline: 24266450]
14. Bey R, Cohen A, Trebossen V, Dura B, Geoffroy P, Jean C, et al. Natural language processing of multi-hospital electronic health records for public health surveillance of suicidality. npj Ment Health Res 2024;3(1):6 [FREE Full text] [doi: 10.1038/s44184-023-00046-7] [Medline: 38609541]
15. Brody C, Star A, Tran J. Chat-based hotlines for health promotion: a systematic review. mHealth 2020;6:36 [FREE Full text] [doi: 10.21037/mhealth-2019-di-13] [Medline: 33437833]
16. Hoffberg AS, Stearns-Yoder KA, Brenner LA. The effectiveness of crisis line services: a systematic review. Front Public Health 2020;7:399 [FREE Full text] [doi: 10.3389/fpubh.2019.00399] [Medline: 32010655]
17. Lejeune A, Le Glaz A, Perron PA, Sebti J, Baca-Garcia E, Walter M, et al. Artificial intelligence and suicide prevention: a systematic review. Eur Psychiatry 2022;65(1):1-22 [FREE Full text] [doi: 10.1192/j.eurpsy.2022.8] [Medline: 35166203]
18. Bernert RA, Hilberg AM, Melia R, Kim JP, Shah NH, Abnousi F. Artificial intelligence and suicide prevention: a systematic review of machine learning investigations. Int J Environ Res Public Health 2020;17(16):5929 [FREE Full text] [doi: 10.3390/ijerph17165929] [Medline: 32824149]
19. Cohen J, Wright-Berryman J, Rohlfs L, Trocinski D, Daniel L, Klatt TW. Integration and validation of a natural language processing machine learning suicide risk prediction model based on open-ended interview language in the emergency department. Front Digit Health 2022;4:818705 [FREE Full text] [doi: 10.3389/fdgth.2022.818705] [Medline: 35187527]
20. Kusuma K, Larsen M, Quiroz JC, Gillies M, Burnett A, Qian J, et al. The performance of machine learning models in predicting suicidal ideation, attempts, and deaths: a meta-analysis and systematic review. J Psychiatr Res 2022;155:579-588. [doi: 10.1016/j.jpsychires.2022.09.050] [Medline: 36206602]
21. Nordin N, Zainol Z, Mohd Noor MH, Chan LF. Suicidal behaviour prediction models using machine learning techniques: a systematic review. Artif Intell Med 2022;132:102395. [doi: 10.1016/j.artmed.2022.102395] [Medline: 36207078]
22. Hill RM, Oosterhoff B, Kaplow JB. Prospective identification of adolescent suicide ideation using classification tree analysis: models for community-based screening. J Consult Clin Psychol 2017;85(7):702-711. [doi: 10.1037/ccp0000218] [Medline: 28414489]
23. Hopkins D, Rickwood DJ, Hallford DJ, Watsford C. Structured data vs. unstructured data in machine learning prediction models for suicidal behaviors: a systematic review and meta-analysis. Front Digit Health 2022;4:945006 [FREE Full text] [doi: 10.3389/fdgth.2022.945006] [Medline: 35983407]
24. Nock MK, Millner AJ, Ross EL, Kennedy CJ, Al-Suwaidi M, Barak-Corren Y, et al. Prediction of suicide attempts using clinician assessment, patient self-report, and electronic health records. JAMA Netw Open 2022;5(1):e2144373 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.44373] [Medline: 35084483]
25. Choi SB, Lee W, Yoon J, Won J, Kim DW. Ten-year prediction of suicide death using Cox regression and machine learning in a nationwide retrospective cohort study in South Korea. J Affect Disord 2018;231:8-14. [doi: 10.1016/j.jad.2018.01.019] [Medline: 29408160]
26. Chen Q, Zhang-James Y, Barnett EJ, Lichtenstein P, Jokinen J, D'Onofrio BM, et al. Predicting suicide attempt or suicide death following a visit to psychiatric specialty care: a machine learning study using Swedish national registry data. PLoS Med 2020;17(11):e1003416 [FREE Full text] [doi: 10.1371/journal.pmed.1003416] [Medline: 33156863]
27. Boudreaux ED, Rundensteiner E, Liu F, Wang B, Larkin C, Agu E, et al. Applying machine learning approaches to suicide prediction using healthcare data: overview and future directions. Front Psychiatry 2021;12:707916 [FREE Full text] [doi: 10.3389/fpsyt.2021.707916] [Medline: 34413800]
28. Broadbent M, Medina Grespan M, Axford K, Zhang X, Srikumar V, Kious B, et al. A machine learning approach to identifying suicide risk among text-based crisis counseling encounters. Front Psychiatry 2023;14:1110527 [FREE Full text] [doi: 10.3389/fpsyt.2023.1110527] [Medline: 37032952]

XSL•FO
RenderX

29. Roy K, Gaur M, Soltani M, Rawte V, Kalyan A, Sheth A. ProKnow: process knowledge for safety constrained and explainable question generation for mental health diagnostic assistance. Front Big Data 2023;5:1056728 [FREE Full text] [doi: 10.3389/fdata.2022.1056728] [Medline: 36700134]
30. Eloranta S, Boman M. Predictive models for clinical decision making: deep dives in practical machine learning. J Intern Med 2022;292(2):278-295 [FREE Full text] [doi: 10.1111/joim.13483] [Medline: 35426190]
31. Schafer KM, Kennedy G, Gallyer A, Resnik P. A direct comparison of theory-driven and machine learning prediction of suicide: a meta-analysis. PLoS One 2021;16(4):e0249833. [doi: 10.1371/journal.pone.0249833] [Medline: 33844698]
32. Selva Birunda S, Kanniga Devi R. A review on word embedding techniques for text classification. In: Raj JS, Iliyasu AM, Bestak R, Baig ZA, editors. Innovative Data Communication Technologies and Application. Lecture Notes on Data Engineering and Communications Technologies. Singapore: Springer; 2021:267-281.
33. Apidianaki M. From word types to tokens and back: a survey of approaches to word meaning representation and interpretation. Comput Linguist 2023;49(2):465-523. [doi: 10.1162/coli_a_00474]
34. Rosenbloom B, Hudson DA, Adeli E, Altman R, Arora S, Arx SV, et al. On the opportunities and risks of foundation models. arXiv Preprint posted on October 1, 2019. [doi: 10.5260/chara.21.2.8]
35. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, et al. Transformers: state-of-the-art natural language processing. 2020 Presented at: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations; October 2020; Virtual p. 38-45. [doi: 10.18653/v1/2020.emnlp-demos.6]
36. Malte A, Ratadiya P. Evolution of transfer learning in natural language processing. arXiv. Posted online on October 16, 2019 2019. [doi: 10.48550/arXiv.1910.07370]
37. Haroz EE, Grubin F, Goklish N, Pioche S, Cwik M, Barlow A, et al. Designing a clinical decision support tool that leverages machine learning for suicide risk prediction: development study in partnership with native American care providers. JMIR Public Health Surveill 2021;7(9):e24377 [FREE Full text] [doi: 10.2196/24377] [Medline: 34473065]
38. Ananthakrishnan G, Jayaraman AK, Trueman TE, Mitra SK, Murugappan AA. Suicidal intention detection in tweets using BERT-based transformers. 2022 Presented at: International Conference on Computing, Communication, and Intelligent Systems (ICCCIS); November 4-5, 2022; Greater Noida. [doi: 10.1109/icccis56430.2022.10037677]
39. Lin E, Sun J, Chen H, Mahoor MH. Data quality matters: suicide intention detection on social media posts using a RoBERTa-CNN model. arXiv. Posted online on February 3, 2024 2024. [doi: 10.48550/arXiv.2402.02262]
40. Alhamed F, Ive J, Specia L. Proceedings of the 9th Workshop on Computational Linguistics and Clinical Psycholog. In: Using Large Language Models (LLMs) to extract evidence from pre-annotated social media data. 2024 Presented at: Proceedings of the Eighth Workshop on Computational Linguistics and Clinical Psychology; 2024 July 29; Seattle, USA p. 232-327 URL: https://aclanthology.org/2024.clpsych-1.22 [doi: 10.18653/v1/2022.clpsych-1.23]
41. Tanaka R, Fukazawa Y. Integrating supervised extractive and generative language models for suicide risk evidence summarization. arXiv. Posted online on March 20, 2024 2024.
42. Devika SP, Pooja MR, Arpitha MS, Vinayakumar R. BERT-Based Approach for Suicide Depression Identification. Proc Third Int Conf Adv Comput Eng Commun Syst 2023;612:435-444. [doi: 10.1007/978-981-19-9228-5_36]
43. Sawhney R, Joshi H, Gandhi S, Shah R. A time-aware transformer based model for suicide ideation detection on social media. 2020 Presented at: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP); 2020 November 01; NA p. 7685-7697.
44. Akintoye O, Wei N, Liu Q. Suicide detection in tweets using LSTM and transformers. 2024 Presented at: 4th Asia Conference on Information Engineering (ACIE); 2024 January 26-28; Singapore, Singapore.
45. Chatterjee M, Kumar P, Samanta P, Sarkar D. Suicide ideation detection from online social media: a multi-modal feature based technique. Int J Inf Manage Data Insights 2022;2(2):100103. [doi: 10.1016/j.jjimei.2022.100103]
46. Pigoni A, Delvecchio G, Turtulici N, Madonna D, Pietrini P, Cecchetti L, et al. Machine learning and the prediction of suicide in psychiatric populations: a systematic review. Transl Psychiatry 2024;14(1):140 [FREE Full text] [doi: 10.1038/s41398-024-02852-9] [Medline: 38461283]
47. Goh ATC. Back-propagation neural networks for modeling complex systems. Artif Intell Eng 1995;9(3):143-151. [doi: 10.1016/0954-1810(94)00011-s]
48. Kokalj E, Škrlj B, Lavrač N, Pollak S, Robnik-Šikonja M. BERT meets Shapley: extending SHAP explanations to transformer-based classifiers. 2021 Presented at: Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation; April 1, 2021; Virtual p. 16-21 URL: https://aclanthology.org/2021.hackashop-1.3
49. Lundberg S, Lee SI. A unified approach to interpreting model predictions. arXiv. Posted online on November 25, 2017 2017;24.
50. Nordin N, Zainol Z, Mohd Noor MH, Chan LF. An explainable predictive model for suicide attempt risk using an ensemble learning and Shapley Additive Explanations (SHAP) approach. Asian J Psychiatr 2023;79:103316. [doi: 10.1016/j.ajp.2022.103316] [Medline: 36395702]
51. Malhotra A, Jindal R. XAI transformer based approach for interpreting depressed and suicidal user behavior on online social networks. Cogn Syst Res 2024;84(1):101186. [doi: 10.1016/j.cogsys.2023.101186]
52. Sandoval AA, Aguilar-Canto F, Jiménez D, Calvo H. Suicide tendency content detection with natural language processing and LIME explainer. Comput. Sci. Res 2023;152(12):143-152.

XSL•FO
RenderX

53. Mumenin N, Basar R, Hossain ABMK, Hossain MD, Hasan ML. Suicidal ideation detection from social media texts using an interpretable hybrid model. 2023 Presented at: 6th International Conference on Electrical Information and Communication Technology (EICT); 2023 December 07-09; Khulna, Bangladesh p. 1-6.
54. Naseem U, Khushi M, Kim J, Dunn AG. Hybrid text representation for explainable suicide risk identification on social media. IEEE Trans Comput Soc Syst 2024;11:4663-4672. [doi: 10.1109/TCSS.2022.3184984]
55. Thushari PD, Aggarwal N, Vajrobol V, Saxena GJ, Singh S, Pundir A. Identifying discernible indications of psychological well-being using ML: explainable AI in reddit social media interactions. Soc Netw Anal Min 2023;13(1):141. [doi: 10.1007/s13278-023-01145-1]
56. Verma A, Harper M, Assi S, Al-Hamid A, Yousif MG, Mustafina J, et al. Suicide ideation detection: a comparative study of sequential and transformer hybrid algorithms. In: Wah YB, Berry MW, Mohamed A, Al-Jumeily D, editors. Data Science and Emerging Technologies. DaSET 2022. Lecture Notes on Data Engineering and Communications Technologies, vol 165. Singapore: Springer; 2023.
57. Tang H, Miri Rekavandi A, Rooprai D, Dwivedi G, Sanfilippo FM, Boussaid F, et al. Analysis and evaluation of explainable artificial intelligence on suicide risk assessment. Sci Rep 2024;14(1):6163 [FREE Full text] [doi: 10.1038/s41598-024-53426-0] [Medline: 38485985]
58. Mahamud AH, Dey AK, Alam ANMS, Alam MGR, Zaman S. Implementation of Explainable AI in Mental Health Informatics: Suicide Data of the United Kingdom. 2022 Presented at: 12th International Conference on Electrical and Computer Engineering (ICECE); December 21-23, 2022; Dhaka, Bangladesh.
59. Kirtley OJ, van Mens K, Hoogendoorn M, Kapur N, de Beurs D. Translating promise into practice: a review of machine learning in suicide research and prevention. Lancet Psychiatry 2022;9(3):243-252. [doi: 10.1016/S2215-0366(21)00254-6] [Medline: 35183281]
60. Luk JW, Pruitt LD, Smolenski DJ, Tucker J, Workman DE, Belsher BE. From everyday life predictions to suicide prevention: clinical and ethical considerations in suicide predictive analytic tools. J Clin Psychol 2022;78(2):137-148. [doi: 10.1002/jclp.23202] [Medline: 34195998]
61. Van Calster B, Vickers AJ. Calibration of risk prediction models: impact on decision-analytic performance. Med Decis Making 2015;35(2):162-169. [doi: 10.1177/0272989X14547233] [Medline: 25155798]
62. Shah ND, Steyerberg EW, Kent DM. Big data and predictive analytics: recalibrating expectations. JAMA 2018;320(1):27-28. [doi: 10.1001/jama.2018.5602] [Medline: 29813156]
63. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014;35(29):1925-1931 [FREE Full text] [doi: 10.1093/eurheartj/ehu207] [Medline: 24898551]
64. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. BMJ 2020;369:m1328 [FREE Full text] [doi: 10.1136/bmj.m1328] [Medline: 32265220]
65. Prozess icons. Flaticon. 2024 Dec 13. URL: https://www.flaticon.com/de/kostenlose-icons/prozess [accessed 2024-12-13]
66. Automatisierung icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/automatisierung [accessed 2024-12-13]
67. Decision icons. Flaticon. URL: https://www.flaticon.com/free-icons/decision [accessed 2024-12-13]
68. Schicht icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/schicht [accessed 2024-12-13]
69. Etwas beachten icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/etwas-beachten [accessed 2024-12-13]
70. Algorithmus icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/algorithmus [accessed 2024-12-13]
71. Richtigkeit icons. Flaticon. URL: https://www.flaticon.com/de/kostenlose-icons/richtigkeit [accessed 2024-12-13]
72. Eckert M, Efe Z, Guenthner L, Baldofski S, Kuehne K, Wundrack R, et al. Acceptability and feasibility of a messenger-based psychological chat counselling service for children and young adults ("krisenchat"): A cross-sectional study. Internet Interv 2022;27:100508 [FREE Full text] [doi: 10.1016/j.invent.2022.100508] [Medline: 35242589]
73. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
74. Reeves KW, Vasconez G, Weiss SJ. Characteristics of suicidal ideation: a systematic review. Arch Suicide Res 2022;26(4):1736-1756. [doi: 10.1080/13811118.2021.2022551] [Medline: 35040371]
75. Joiner T. Why People Die by Suicide. U.S: Harvard University Press; 2005.
76. O'Connor RC, Kirtley OJ. The integrated motivational-volitional model of suicidal behaviour. Philos Trans R Soc Lond B Biol Sci 2018;373(1754):20170268 [FREE Full text] [doi: 10.1098/rstb.2017.0268] [Medline: 30012735]
77. Chu C, Buchman-Schmitt JM, Stanley IH, Hom MA, Tucker RP, Hagan CR, et al. The interpersonal theory of suicide: a systematic review and meta-analysis of a decade of cross-national research. Psychol Bull 2017;143(12):1313-1345. [doi: 10.1037/bul0000123] [Medline: 29072480]
78. Klonsky ED, May AM. The three-step theory (3ST): a new theory of suicide rooted in the “Ideation-to-Action” framework. Int J Cogn Ther 2015;8(2):114-129. [doi: 10.1521/ijct.2015.8.2.114]
79. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024;385:e078378 [FREE Full text] [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]

XSL•FO
RenderX

80. Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AICONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020;26(9):1364-1374 [FREE Full text] [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]
81. TensorFlow. 2023. URL: https://zenodo.org/records/10126399 [accessed 2023-12-06]
82. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature 2020;585(7825):357-362 [FREE Full text] [doi: 10.1038/s41586-020-2649-2] [Medline: 32939066]
83. GitHub - pandas-dev/pandas at v2.1.3. URL: https://github.com/pandas-dev/pandas/tree/v2.1.3 [accessed 2024-12-11]
84. GitHub - keras-team/keras: Deep learning for humans. URL: https://github.com/keras-team/keras [accessed 2024-12-11]
85. Matplotlib ? Visualization with Python. URL: https://matplotlib.org/ [accessed 2024-07-31]
86. Waskom M. seaborn: statistical data visualization. J Open Source Softw 2021;6(60):3021. [doi: 10.21105/joss.03021]
87. Andersson JM. JimmyMAndersson/StatKit. URL: https://github.com/JimmyMAndersson/StatKit [accessed 2024-07-31]
88. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. arXiv. Posted online on June 5, 2018 2018.
89. krisenchat/p91-suicide-model: This is a transformer-based text classification model which aims to predict and explain session-propensities for suicide. GitHub. URL: https://github.com/krisenchat/p91-suicide-model [accessed 2024-07-30]
90. Conneau A, Khandelwal K, Goyal N, Chaudhary V, Wenzek G, Guzmán F, et al. Unsupervised cross-lingual representation learning at scale. arXiv. Posted online on April 8, 2020 2020.
91. Elkind E, Rothe J. Cooperative game theory. In: Rothe J, editor. Economics and Computation. Springer Texts in Business and Economics. Berlin: Springer; 2016.
92. Merrick L, Taly A. The explanation game: explaining machine learning models using Shapley values. In: Machine Learning and Knowledge Extraction. CD-MAKE 2020. Lecture Notes in Computer Science. Switzerland: Springer, Cham; 2020.
93. Wysocki O, Davies JK, Vigo M, Armstrong AC, Landers D, Lee R, et al. Assessing the communication gap between AI models and healthcare professionals: explainability, utility and trust in AI-driven clinical decision-making. Artif Intell 2023;316:103839. [doi: 10.1016/j.artint.2022.103839]
94. Key Issue 5: Transparency Obligations - EU AI Act. URL: https://www.euaiact.com/key-issue/5 [accessed 2024-04-27]
95. Sierra G, Andrade-Palos P, Bel-Enguix G, Osornio-Arteaga A, Cabrera-Mora A, García-Nieto L, et al. Suicide risk factors: a language analysis approach in social media. J Lang Soc Psychol 2021;41(3):312-330 [FREE Full text] [doi: 10.1177/0261927X211036171]
96. Lao C, Lane J, Suominen H. Analyzing suicide risk from linguistic features in social media: evaluation study. JMIR Form Res 2022;6(8):e35563 [FREE Full text] [doi: 10.2196/35563] [Medline: 36040781]
97. Homan S, Gabi M, Klee N, Bachmann S, Moser A, Duri' M, et al. Linguistic features of suicidal thoughts and behaviors: a systematic review. Clin Psychol Rev 2022;95:102161 [FREE Full text] [doi: 10.1016/j.cpr.2022.102161] [Medline: 35636131]
98. Warner B, Misra M. Understanding neural networks as statistical tools. Am Stat 1996 Nov;50(4):284-293 [FREE Full text] [doi: 10.1080/00031305.1996.10473554]
99. Zhang Z, Rousson V, Lee WC, Ferdynus C, Chen M, Qian X, written on behalf of AME Big-Data Clinical Trial Collaborative Group. Decision curve analysis: a technical note. Ann Transl Med 2018 Aug;6(15):308 [FREE Full text] [doi: 10.21037/atm.2018.07.02] [Medline: 30211196]
100. Aldhyani THH, Alsubari SN, Alshebami AS, Alkahtani H, Ahmed ZAT. Detecting and analyzing suicidal ideation on social media using deep learning and machine learning models. Int J Environ Res Public Health 2022 Oct 03;19(19):12635 [FREE Full text] [doi: 10.3390/ijerph191912635] [Medline: 36231935]
101. Douglass RW, Gartzke E, Lindsay JR, Gannon JA, Scherer TL. What is escalation? Measuring crisis dynamics in international relations with human and LLM generated event data. arXiv. Posted online on January 18, 2024 2024.
102. Stern W, Goh SJ, Nur N, Aragon PJ, Mercer T, Bhattacharyya S, et al. Natural Language Explanations for Suicide Risk Classification Using Large Language Models. ceur-ws.org 2024;3649:1-10.

Abbreviations

AI: artificial intelligence
ASE: advanced suicidal engagement
AUC: area under the curve
BS: Brier score
C-SSRS: Columbia-Suicide Severity Rating Scale
CONSORT-AI: Consolidated Standards of Reporting Trials for Artificial Intelligence
DCA: decision curve analysis
LSTM: long short-term memory
ML: machine learning
NB: net benefit
ROC: receiver operating characteristic

XSL•FO
RenderX

SHAP: Shapley Additive Explanation
SI: suicidal ideation
SIB: suicidal ideation and behavior
tf-idf: term frequency-inverse document frequency
T-MLP: transformer-multilayer perceptron
W2V-MLP: word2vector-multilayer perceptron

Edited by A Mavragani; submitted 01.07.24; peer-reviewed by EU Kucuksille, GM Bairy, S Mao; comments to author 18.07.24; revised version received 30.08.24; accepted 07.11.24; published 29.01.25.

Please cite as:
Thomas J, Lucht A, Segler J, Wundrack R, Miché M, Lieb R, Kuchinke L, Meinlschmidt G
An Explainable Artificial Intelligence Text Classifier for Suicidality Prediction in Youth Crisis Text Line Users: Development and Validation Study
JMIR Public Health Surveill 2025;11:e63809
URL: https://publichealth.jmir.org/2025/1/e63809
doi:10.2196/63809
PMID:

©Julia Thomas, Antonia Lucht, Jacob Segler, Richard Wundrack, Marcel Miché, Roselind Lieb, Lars Kuchinke, Gunther Meinlschmidt. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Cloud-Based Platform for Harmonized COVID-19 Data: Design and Implementation of the Rapid Acceleration of Diagnostics (RADx) Data Hub

Marcos Martínez-Romero[1], PhD; Matthew Horridge[1], PhD; Nilesh Mistry[2], BS; Aubrie Weyhmiller[3], MS; Jimmy K Yu[1], PhD; Alissa Fujimoto[2], MA; Aria Henry[3], BA; Martin J O'Connor[1], MSc; Ashley Sier[2], MPH; Stephanie Suber[3], MS; Mete U Akdogan[1], PhD; Yan Cao[1], MS; Somu Valliappan[2], MS; Joanna O Mieczkowska[3], MA; Ashok Krishnamurthy[3], PhD; Michael A Keller[2], PhD; Mark A Musen[1], MD, PhD; RADx Data Hub Team[4]

[1]Stanford University, Stanford Center for Biomedical Informatics Research, Palo Alto, CA, United States

[2]Booz Allen Hamilton Inc., McLean, VA, United States

[3]University of North Carolina at Chapel Hill, Renaissance Computing Institute (RENCI), Chapel Hill, NC, United States

[4]see Authors' Contributions section

Corresponding Author:
Marcos Martínez-Romero, PhD
Stanford University
Stanford Center for Biomedical Informatics Research
3180 Porter Dr
Palo Alto, CA, 94304
United States
Phone: 1 (650) 724 3979
Email: marcosmr@stanford.edu

Abstract

Background: The COVID-19 pandemic exposed significant limitations in existing data infrastructure, particularly the lack of systems for rapidly collecting, integrating, and analyzing data to support timely and evidence-based public health responses. These shortcomings hampered efforts to conduct comprehensive analyses and make rapid, data-driven decisions in response to emerging threats. To overcome these challenges, the US National Institutes of Health launched the Rapid Acceleration of Diagnostics (RADx) initiative. A key component of this initiative is the RADx Data Hub—a centralized, cloud-based platform designed to support data sharing, harmonization, and reuse across multiple COVID-19 research programs and data sources.

Objective: We aim to present the design, implementation, and capabilities of the RADx Data Hub, a cloud-based platform developed to support findable, accessible, interoperable, reusable (FAIR) data practices and enable secondary analyses of the COVID-19–related data contributed by a nationwide network of researchers.

Methods: The RADx Data Hub was developed on a scalable cloud infrastructure, grounded in the FAIR data principles. The platform integrates heterogeneous data types—including clinical data, diagnostic test results, behavioral data, and social determinants of health—submitted by over 100 research organizations across 46 US states and territories. The data pipeline includes automated and manual processes for deidentification, quality validation, expert curation, and harmonization. Metadata standards are enforced using tools such as the Center for Expanded Data Annotation and Retrieval (CEDAR) Workbench and BioPortal. Data files are structured using a unified specification to support consistent representation and machine-actionable metadata.

Results: As of May 2025, the RADx Data Hub hosts 187 studies and over 1700 data files, spanning 4 RADx programs: RADx Underserved Populations (RADx-UP), RADx Radical (RADx-rad), RADx Tech, and RADx Digital Health Technologies (RADx DHT). The Study Explorer and Analytics Workbench components enable researchers to discover relevant studies, inspect rich metadata, and conduct analyses within a secure cloud-based environment. Harmonized data conforming to a core set of common data elements facilitate cross-study integration and support secondary use. The platform provides persistent identifiers (digital object identifiers) for each study and supports access to structured metadata that adhere to the CEDAR specification, available in both JSON and YAML formats for seamless integration into computational workflows.

Conclusions: The RADx Data Hub successfully addresses key data integration challenges by providing a centralized, FAIR-compliant platform for public health research. Its adaptable architecture and data management practices are designed to

XSL•FO
RenderX

support secondary analyses and can be repurposed for other scientific disciplines, strengthening data infrastructure and enhancing preparedness for future health crises.

(JMIR Public Health Surveill 2025;11:e72677) doi:10.2196/72677

KEYWORDS

COVID-19 surveillance; public health data infrastructure; data harmonization and integration; health disparities; FAIR data sharing; cloud-based data platform; pandemic response informatics; secondary data analysis; metadata standards; digital health research

Introduction

Background

The COVID-19 pandemic exposed significant limitations in the data infrastructures needed to support effective public health responses. Fragmented and noninteroperable systems—often operating in silos—hindered the timely collection, integration, and analysis of health data, delaying informed decision-making at both local and global levels [1-3]. As the volume and diversity of pandemic-related data grew, from clinical outcomes and diagnostic testing to genomic surveillance and social determinants, so too did the need for coordinated systems capable of managing, integrating, and analyzing these disparate sources.

In response, a range of national and international initiatives emerged to enable secure, scalable, and structured data sharing. The National COVID Cohort Collaborative in the United States aggregated electronic health records from more than 70 institutions into a centralized enclave for COVID-19 research, applying a common data model to facilitate cross-institutional analyses of clinical outcomes, risk factors, and treatment efficacy [4,5]. The European COVID-19 Data Portal [6], supported by the European Bioinformatics Institute, centralized access to multiomic, epidemiological, and imaging datasets, allowing researchers to explore genomic surveillance trends and viral evolution using standardized pipelines [7]. The ORCHESTRA (Connecting European Cohorts to Increase Common and Effective Response to SARS-CoV-2 Pandemic) project [8] built a federated infrastructure that harmonized data across longitudinal clinical cohorts in Europe, focusing on outcomes among diverse patient populations and informing evidence-based pandemic policy. In the United Kingdom, the OpenSAFELY platform [9] enabled secure, in situ analyses of electronic health records for over 58 million patients, providing rapid insights into COVID-19 risk factors and treatment outcomes. The US Centers for Disease Control and Prevention–led Data Modernization Initiative [10] aimed to transform national public health data infrastructure, catalyzed in part by COVID-19–related gaps in data access, interoperability, and surveillance. In parallel, domain-specific initiatives, such as the Infectious Diseases Data Observatory [11] and Global.health [12] addressed global data equity by supporting data sharing and outbreak analytics in low- and middle-income settings, including resources for contact tracing, genomic epidemiology, and real-time surveillance.

Despite these advances, implementations varied widely in scope, data quality, and adherence to data governance best practices. Many data-sharing efforts were created reactively, often without sustainable frameworks for data stewardship or unified governance mechanisms. While several platforms implemented elements of the findable, accessible, interoperable, reusable (FAIR) principles, such as assigning persistent identifiers or using standardized vocabularies, compliance was inconsistent and frequently deprioritized due to operational urgency [13]. The absence of common data models, interoperable metadata schemas, and cross-platform compatibility limited the potential for integrated analyses across datasets. Moreover, key aspects of reusability—such as provenance tracking, licensing, and semantic alignment—were often underdeveloped or absent altogether. These limitations led to duplication of effort, inefficient data discovery, and barriers to secondary analyses, particularly in domains that required rapid hypothesis testing and multicohort validation, such as diagnostics and health disparities research.

These challenges highlighted the need for a more coordinated and principled approach to data infrastructure—one that could enable consistent metadata standards, facilitate integration across studies, and support responsible data reuse at scale. To address gaps in data standardization and the lack of a centralized platform, the US National Institutes of Health (NIH) launched the Rapid Acceleration of Diagnostics (RADx) initiative in 2020 [14], with a key objective to develop such a centralized platform—the RADx Data Hub [15]. This platform aims to aggregate, harmonize, and provide access to data generated by various programs in the overall RADx initiative, enabling data reuse for secondary analyses and facilitating cross-study comparisons. The RADx Data Hub supports innovative research into diagnostic tools, implementation strategies, and health disparities.

At the core of the RADx Data Hub (hereafter referred to as the Data Hub) is a cloud-based repository grounded in the FAIR guiding principles for data sharing [16]. The machinery and the specifications of the Data Hub are designed to make data (1) findable, through curation, annotation, and support for advanced search and data exploration; (2) accessible, through established privacy and security protocols meant to protect sensitive data while supporting data reuse; (3) interoperable, by leveraging semantic technologies and by defining standard data and metadata models for Data Hub entities; and (4) reusable, by developing routines to validate data quality, techniques to harmonize heterogeneous data, and cloud-based infrastructure to support integrated data analysis on sensitive data.

While the Data Hub was developed specifically to host COVID-19 studies, its underlying technologies were intentionally designed to be repurposed for other domains. The architecture, data and metadata models, software components,

XSL•FO
RenderX

and protocols provide a reusable framework that can be adapted to support new systems addressing other areas of scientific research.

Objectives

We aim to present the Data Hub as a key resource for secondary analysis of COVID-19 data and provide details regarding the data management techniques and software tools developed to bring it to fruition as a reliable and accessible data platform.

Methods

Overview

The Data Hub's approach to supporting data reusability and interoperability is built upon 4 key components: (1) a structured model organizing study data, metadata, and documentation; (2) an end-to-end pipeline for data collection and analysis; (3) a scalable, cloud-based system architecture; and (4) a comprehensive data and metadata harmonization framework. These components are complemented by rigorous data deidentification measures to ensure compliance with established privacy standards, resulting in a flexible foundation that can be adapted for use beyond the COVID-19 context to support a wide range of scientific research domains. The following subsections detail each component and illustrate how they interconnect to form a cohesive, FAIR-compliant infrastructure for scientific data collection, harmonization, and analysis.

A Model to Organize Study Information

The Data Hub sources its data from 4 (C)DCCs (coordination and data collection centers), each aligned with a distinct NIH RADx program—RADx Radical (RADx-rad), RADx Underserved Populations (RADx-UP), RADx Tech, and RADx Digital Health Technologies (RADx DHT) [17]. The RADx-UP program includes a coordination function, whereas RADx-rad, RADx Tech, and RADx DHT focus solely on data collection. The abbreviation (C)DCCs is used to collectively refer to all of them. The (C)DCCs manage the collection, harmonization, and submission of study data to the Data Hub, supporting the mission and focus of their respective RADx programs. RADx-UP, the largest contributor, focuses on health disparities and community-level data collection from underserved populations. RADx-rad explores experimental (radical) technologies for nontraditional COVID-19 detection and surveillance settings. RADx Tech supports the development and validation of diagnostic technologies. RADx DHT generates digital health data from wearable devices, mobile apps, and real-time monitoring platforms. For RADx DHT studies, the Data Hub hosts rich metadata to enable data discovery and reuse, while the underlying data are stored in the Rapid AI Platform for Innovating Data Science repository [18], which is optimized for handling large-scale digital health data. Collectively, the Data Hub integrates diverse research efforts, enabling comprehensive analyses through harmonized study data and metadata.

The Data Hub stores information for each study, categorized into 3 overarching components: study data, study metadata, and study documentation (Figure 1). *Study data* constitute the core outputs of each study, capturing the essential variables and results generated during data collection activities. Study data are organized into one or several *file bundles*, each comprising 3 main components (Textbox 1): data file, data dictionary, and file metadata.

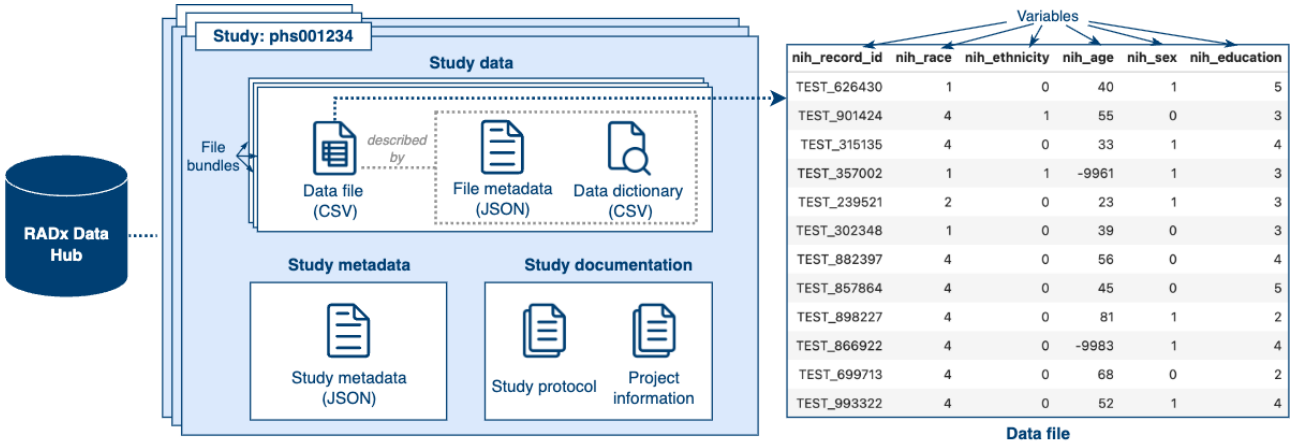


Figure 1. Components of Data Hub studies. This figure illustrates the organization and structure of study data within the Data Hub. On the left, the schematic shows how study data are organized into file bundles, each containing a data file (CSV), file metadata (JSON), and a data dictionary (CSV) that defines the structure and semantics of each variable. Each study is also accompanied by structured study metadata (in JSON format) and supporting study documentation, including protocols and project information. On the right, an example data file is shown, demonstrating a tabular format in which each row corresponds to a study participant and each column represents a standardized variable (eg, race, ethnicity, age, sex, education) harmonized using the Data Hub's common data elements.

XSL•FO
RenderX

Textbox 1. Components of the file bundles.

Data file

A tabular CSV document structured according to its corresponding data dictionary. *Variables*, defined in the data dictionary, represent measurable attributes or characteristics (eg, age, sex, and education) and are organized as columns in the data file. Each row corresponds to a data point, such as an experimental sample or study participant, with values recorded for one or more variables. A data dictionary is essential for interpreting the values in the data file.

Data dictionary

A CSV document in which each row unambiguously describes the specification of a variable measured or collected by a study or experiment. The columns of the data dictionary provide identifiers for a variable, including properties such as the variable name, label, unit of measurement, datatype, and other attributes relevant to specifying a variable's values. The complete specification of data dictionaries used by the Data Hub can be found in the Data Hub Data Dictionary Specification [19]. Data dictionaries are automatically validated against the data dictionary specification upon data submission, and compliance with the specification is strictly enforced. As a result, data dictionaries (and variable metadata, by extension) are amenable to uniform machine processing according to the specification.

File metadata

Each data file is paired with a metadata file that describes its key characteristics, including versioning information, creator details, and summary statistics.

Each study is also paired with a *study metadata* file, offering a comprehensive view of its scope, goals, and content. Study metadata includes details such as the study's geographic and demographic focus, methodologies used, key outcomes, funding sources, and contributing institutions. By providing a consolidated overview, study metadata facilitates the identification and evaluation of studies that are relevant to specific research questions, while enabling systematic reviews and meta-analyses.

In addition, each study is accompanied by *study documentation* that describes the design and execution processes. The study documentation includes study protocols, project summaries, manual of operations, and data-collection instruments. By capturing this rich contextual information, the study documentation helps end users understand the study and use the resulting data effectively. In addition, it offers critical insights into the rationale behind the study design, the measures taken to maintain data quality, and the study's adherence to regulatory and ethical standards, thereby building trust in the data and facilitating responsible data reuse.

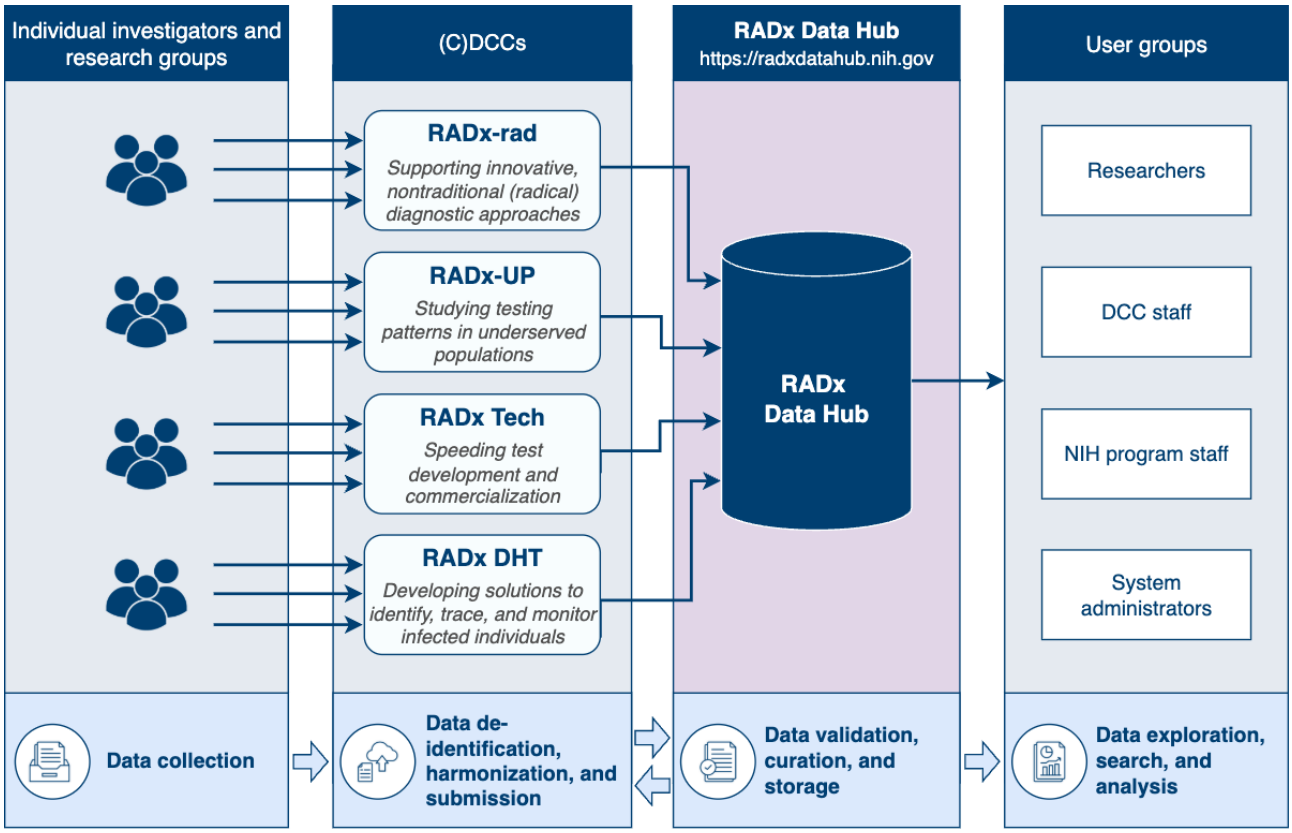
An End-to-End Pipeline for Data Collection and Analysis

Data Hub's end-to-end pipeline (Figure 2) streamlines the collection, deidentification, harmonization, validation, curation, storage, and analysis of the data. Data producers, including individual study investigators and research groups, transmit their study data to one of the 4 (C)DCCs. These (C)DCCs, each aligned with a specific focus within the RADx initiative, play a critical role in the pipeline by standardizing, harmonizing, and preparing the data for inclusion in the Data Hub.

After submission to the Data Hub, the data undergo automated validation to detect quality issues, such as nonadherence to standards, incorrect formats, or incomplete metadata. In addition, the validation process includes automated detection of potential protected health information (PHI) and personally identifiable information (PII) to ensure compliance with privacy and security regulations. These automated steps are followed by manual curation, enabling further refinement of the data. Once these processes are complete, the data are stored and made accessible to users.

XSL•FO
RenderX

Figure 2. Overview of the Data Hub's end-to-end data pipeline. Individual investigators and research groups collect and transmit data to the appropriate (C)DCC (coordination and data collection center). At the (C)DCCs, staff deidentify, harmonize, and submit the study data to the Data Hub. Any identified issues are communicated to data contributors, creating a feedback loop for iterative improvements and alignment with Data Hub standards. Once validated, the data are approved, curated, stored in the Data Hub, and then made accessible to users. The arrows reflect the flow of data through the pipeline and are not intended to represent the full range of user interactions with the system.



Overview

The Data Hub leverages the Center for Expanded Data Annotation and Retrieval (CEDAR) Workbench and the BioPortal ontology repository as foundational tools for metadata management and interoperability.

CEDAR [20] offers a suite of web-based tools designed to streamline the creation and management of high-quality metadata. At the core of CEDAR is the concept of metadata templates—structured forms that define the attributes and constraints necessary to describe specific types of data consistently. Templates incorporate controlled vocabularies, specifying permissible values for fields such as demographic characteristics, experimental conditions, and data collection methodologies. Researchers and data contributors use these templates to annotate data with standardized, machine-readable metadata, ensuring adherence to community standards and improving data FAIRness [21].

CEDAR uses the BioPortal ontology repository [22] to integrate ontology terms into metadata templates, ensuring fields are annotated with standardized, semantically rich vocabularies. BioPortal is a comprehensive repository that hosts a vast range of biomedical ontologies, providing standardized terms and definitions across diverse domains to support consistent semantic annotation of data. It provides access to over 1000 ontologies when building CEDAR templates, enabling users to select standardized terms and definitions to describe study variables.

By embedding CEDAR's structured metadata templates and BioPortal's rich vocabulary resources into its metadata workflow, the Data Hub provides a comprehensive framework for FAIR compliance. The following subsections detail how this framework was used to help ensure the findability, accessibility, interoperability, and reusability of RADx data.

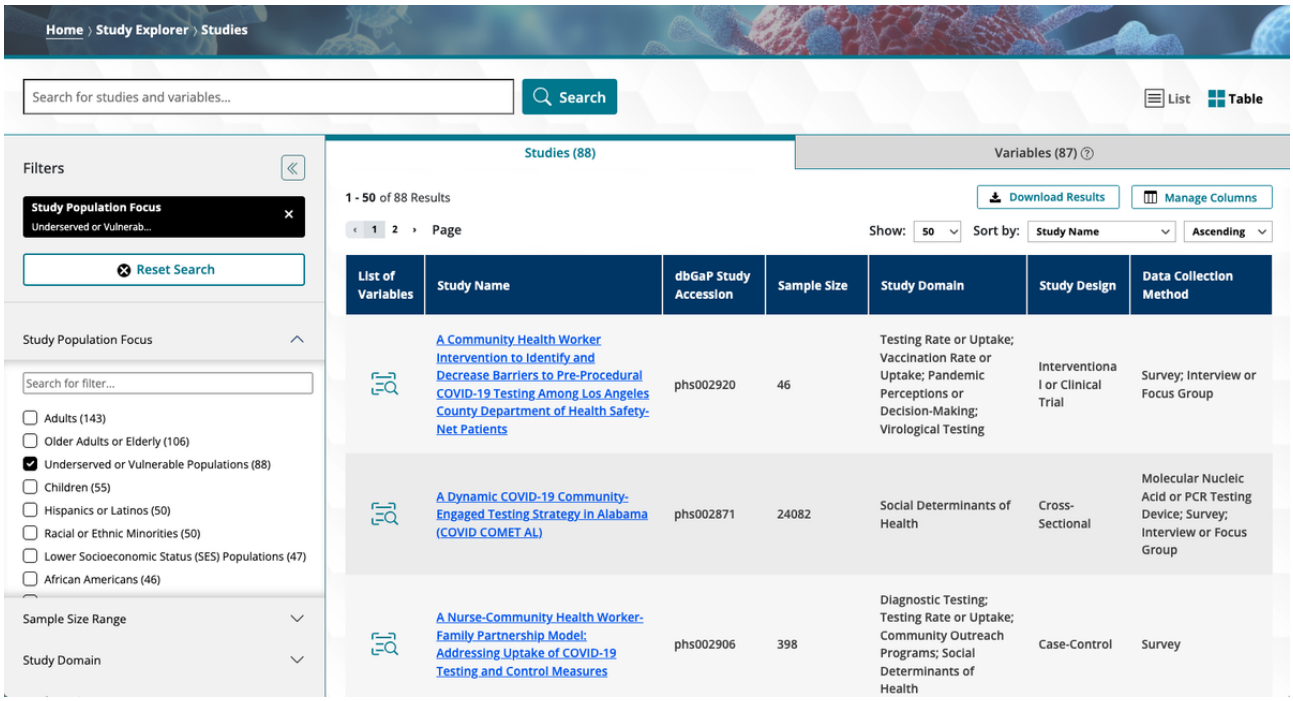
Findability: Exploring and Searching Study Data

One of the primary purposes of the Data Hub is to provide a platform for researchers to find study data appropriate for their secondary analyses. While access to study data is governed by the NIH Database of Genotypes and Phenotypes (dbGaP) [23] and requires formal approval, the Study Explorer component of the Data Hub (Figure 3) is openly available. It allows users to browse the list of studies, explore their characteristics, and assess their suitability before submitting a data access request through dbGaP. In addition to study-level exploration, the Study Explorer includes a Variables tab that enables users to search across all variables indexed in the Data Hub, supporting discovery of specific data elements of interest across multiple studies. The Study Explorer offers a range of search functionalities, including free-text search, conditional queries using multiple criteria, and options to sort, filter, and download results. Users can filter results using various facets, including study domain, sample size, and data collection method. Studies can also be located directly by querying for their title or other study attributes.

XSL•FO
RenderX

Upon locating a study of interest, users can view its details on the Study Overview page (Figure 4). This web page provides a concise summary of the study, offering access to the key components: study metadata, documentation, and data files, including file metadata and data dictionaries. Study metadata can be downloaded in machine-readable formats, such as JSON and YAML, for easy integration into research workflows. To simplify metadata viewing, the Data Hub provides an interactive Metadata Viewer (Figure 5) powered by CEDAR's Embeddable Editor [24], which includes links to ontology terms from the BioPortal ontology repository [22,25]. This tool allows users to inspect all metadata associated with a study and to verify that the study aligns with their research goals. In addition, the Study Overview page displays the variables used in each file, enabling researchers to identify variables relevant to their analyses or to search for related studies that share the same variables. Furthermore, each study on the Data Hub is assigned a unique DataCite digital object identifier (DOI), a globally persistent identifier that provides a permanent reference to the study and its metadata, enabling reliable citation and access from both within the Data Hub and external sources.

Figure 3. Study Explorer. The figure shows a screenshot of the Data Hub Study Explorer, with the "study population focus" filter applied to retrieve 88 studies on underserved or vulnerable populations. A search box at the top, equipped with an autocomplete feature, allows users to quickly find studies or variables by entering relevant keywords. The interface includes Studies and Variables tabs, enabling users to explore not only study-level metadata but also variable-level details across studies. The filtered results are displayed in a table that includes information such as study names, Database of Genotypes and Phenotypes accessions, sample sizes, study domains, study designs, and data collection methods.



XSL•FO
RenderX

Figure 4. Study Overview page. This figure provides an overview of the information available for a selected study within the Data Hub. The panel at left summarizes key metadata, such as the Database of Genotypes and Phenotypes (dbGaP) study accession, program, study domain, and data collection methods. The panel at right is divided into Study Documents and Data Files. The Study Documents section lists supporting files, such as study protocols and README files. The Data Files section contains the data files and their corresponding metadata, along with a button to request access to the data through the study's dbGaP page.

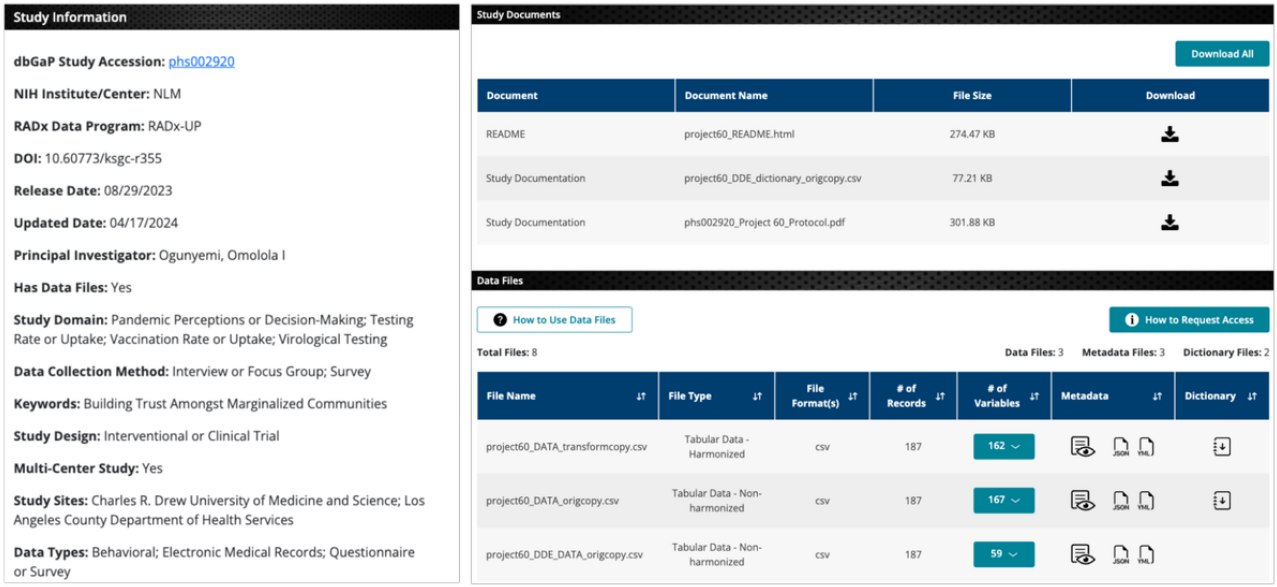
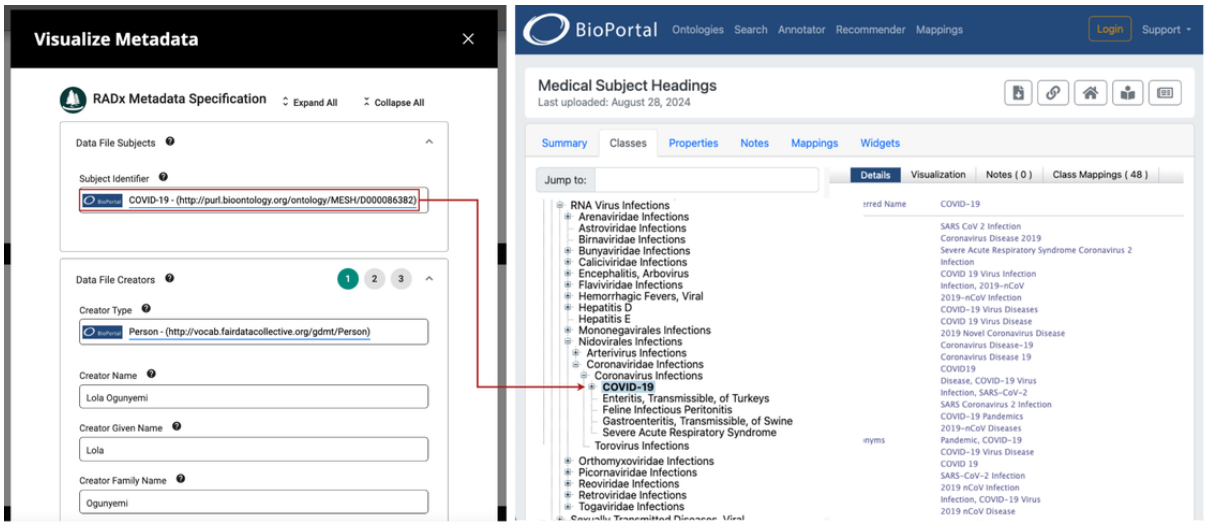


Figure 5. Metadata Viewer. The left side of the figure displays metadata for a specific data file, with terms constrained to ontology terms, linked to BioPortal for further exploration. The right side demonstrates a link from the "subject identifier" field to the corresponding term in the Medical Subject Headings (MeSH). Metadata are organized into expandable and collapsible sections for easy navigation, with help text providing guidance and context. The Metadata Viewer renders a Center for Expanded Data Annotation and Retrieval (CEDAR) template designed for Data Hub files, ensuring consistency and accessibility across metadata records.



Accessibility: Ensuring Data Availability and Secure Access

Protecting sensitive health-related data requires robust control mechanisms to ensure compliance with ethical standards, government regulations, and data use agreements. Access to files from the Data Hub is reviewed and authorized by the NIH Data Access Committee with an established data use agreement through dbGaP [23]. This agreement outlines specific terms for data use, including requirements for privacy protections and ethical conduct. Note that in dbGaP terminology, the document submitted to request access to controlled-access data is formally called a data use certification, which specifies the terms and conditions for data use. The term *data use agreement* is used more broadly across NIH to refer to the framework governing data access. In addition, users must follow the Data Hub user code of conduct [26], which sets standards for responsible data handling and site use.

To formally request access to study-level data, users must log into the Data Hub using their electronic Research Administration Commons or NIH credentials, locate the desired study (or

XSL•FO
RenderX

studies), and submit an access request. This requirement ensures that only verified researchers and authorized personnel can access controlled datasets, in compliance with NIH security and data governance policies. While this log-in process may pose a barrier for new users, it helps protect sensitive research data and aligns with NIH's data access standards. Any individual with principal investigator status can obtain an electronic Research Administration Commons log-in and dbGaP authorization. Once access is approved by the data access committee, principal investigators can manage permissions for their team members via dbGaP. Tutorials on the data request process are available on the Data Hub Resource Center web page [27] to assist users with each step.

Study metadata, summary statistics, and documentation remain openly accessible through the Study Explorer and Study Overview pages. This persistent availability allows users to evaluate the relevance of study data and variables without authentication, supporting transparency, discoverability, and reuse. Although these practices help maintain visibility into study-level details even when data files become unavailable, we acknowledge that a formal policy guaranteeing long-term metadata persistence has not yet been established. As part of our commitment to enhancing FAIRness, we plan to explore mechanisms to ensure long-term metadata retention.

Interoperability: Enforcing Data and Metadata Standards

The Data Hub establishes a foundation for interoperability by adopting globally recognized data formats and metadata standards, alongside a standardized set of NIH common data elements (CDEs). CDEs are defined as combinations of standardized questions (variables) paired with an enumerated set of possible responses (values) that are common across multiple studies [28]. Variables in data files are mapped to these CDEs to ensure consistent (ie, harmonized) interpretation across studies.

The CDEs used in the Data Hub are defined in the Data Hub Global Codebook [29]. This listing includes 133 core CDEs that provide a consistent framework for representing essential variables, such as demographic and participant health information, across all RADx studies. These CDEs are organized into 12 categories, including race, ethnicity, age, sex, education, domicile, employment, insurance status, disability status, medical history, symptoms, and health status. The Data Hub team is responsible for performing harmonization to these CDEs where possible by using a harmonization methodology detailed in the Data and Metadata Harmonization section of this paper. However, submitted studies are not limited by these CDEs and may include additional variables collected by the original researchers.

Beyond the provision of CDEs, the Data Hub uses standardized file formats and metadata specifications to further enhance interoperability. Data files are provided in CSV format, organized in a consistent structure where variables are represented as columns and rows as individual data points. While CSV files are not inherently self-descriptive or semantically annotated, their accompanying data dictionaries and metadata files include machine-readable definitions that support automated data interpretation and integration. Data dictionaries follow the RADx Data Dictionary Specification [19], also in CSV format, ensuring uniform interpretation of variables. Metadata for data files and studies are formatted in JSON for linked data, adhering to the CEDAR model specification [30]. In addition, study documentation is provided in standard formats, such as PDF and plain text, ensuring accessibility and compatibility for a wide range of users.

In practice, the (C)DCCs collaborate closely with the Data Hub team to ensure compliance with these global standards. The (C)DCCs conduct the initial harmonization, while the Data Hub team maintains the Global Codebook and performs additional post hoc harmonization as needed before studies are made publicly available. These efforts aim to enhance the interoperability of study data within the RADx ecosystem and with external data repositories and analytical platforms. This comprehensive standardization approach maximizes the data's utility for research and public health applications.

Reusability: Prioritizing Data Quality and Enabling Secondary Data Analysis

The Data Hub promotes study data reusability by focusing on 2 critical aspects: ensuring high data quality and providing an integrated platform for secondary data analysis.

Data Quality

To ensure study data utility and reliability, the Data Hub uses a rigorous quality assurance and quality control process. This process begins with close collaboration between the Data Hub team and (C)DCC staff during data submission, ensuring adherence to the data and metadata standards for metadata files, data dictionaries, and documentation described in previous sections. Quality assurance and quality control reviews are performed rigorously before data and metadata are made publicly findable, ensuring compliance with privacy policies, alignment with the Data Hub's standards, and readiness for harmonization and secondary use.

A key aspect of this process involves evaluating file bundle completeness, identifying and addressing missing values, and standardizing metadata formats to ensure consistency across data files and studies. While the quality of the data reflects the efforts of the original data producers, the Data Hub enhances value by applying harmonization practices during ingestion and by documenting any data and metadata gaps or limitations. These enhancements improve transparency, facilitate cross-study comparisons, and make the data more reliable for downstream analyses.

Data Analytics Workbench

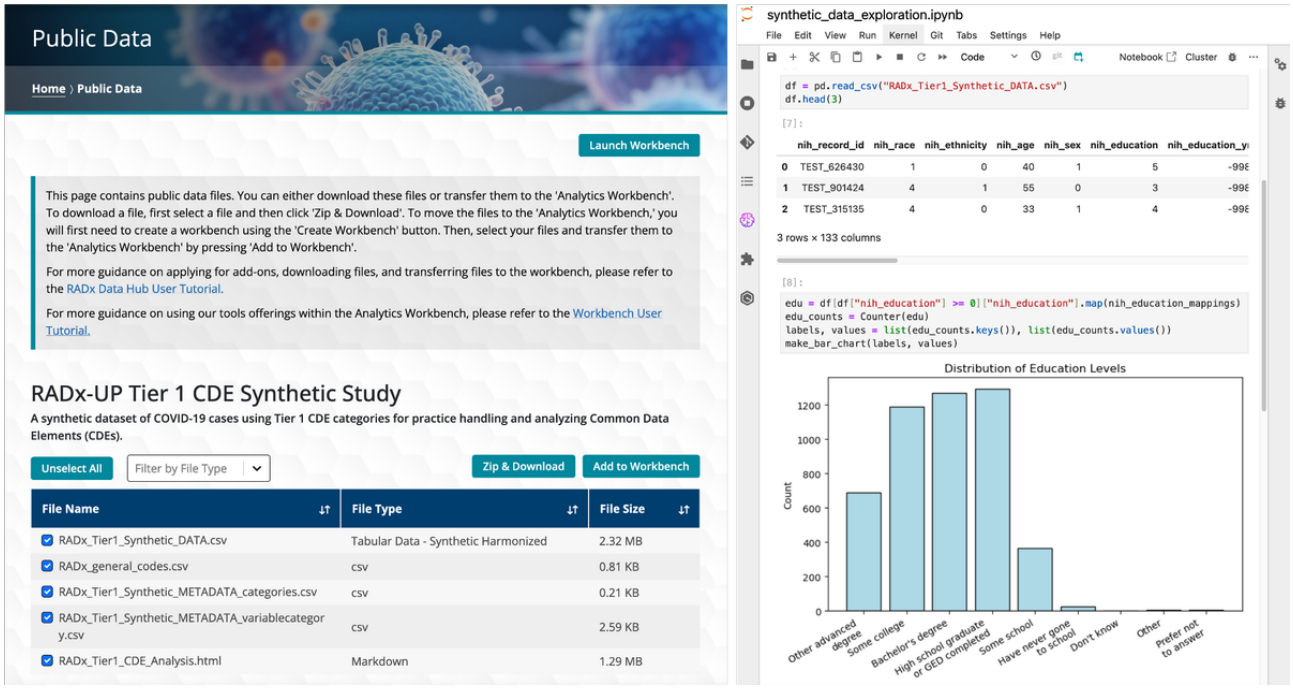
The Data Hub features a cloud-enabled platform called the Analytics Workbench as a key feature for promoting data reuse (Figure 6). This workbench is a versatile platform, powered by the Amazon Web Services (AWS) SageMaker Studio technology, that enables users to create personalized workspaces and to deploy advanced analytical tools. By leveraging this platform, researchers can securely access, manipulate, and analyze Data Hub data. The workbench provides users with standard data analysis tools, such as RStudio, Python, and JupyterLab, all readily accessible without additional setup. The

XSL•FO
RenderX

workbench also supports integration with external platforms, such as DockerHub and GitHub, enabling users to tailor their workflows to specific research needs.

The Public Data page provides access to synthetic data files, which are openly available to users for exploration, testing, and familiarization with the platform's tools. Users can download these data files directly or transfer them to the Analytics Workbench for further analysis. This feature allows researchers to explore the platform's tools and capabilities before requesting access to protected study data. The Analytics Workbench is shown on the right-hand side of Figure 6, where a Jupyter Notebook environment allows analysis of the distribution of the individual's education levels within a synthetic data file. This example highlights the Workbench's capacity to facilitate secure data analysis and visualization, offering a flexible environment equipped with various analytical tools.

Figure 6. Public data and Analytics Workbench. This figure demonstrates 2 key components of the Data Hub that support data accessibility and exploratory analysis. The left panel shows the Public Data page, where users can browse and download openly available, harmonized synthetic datasets or transfer them to the Analytics Workbench environment. Example files include harmonized data tables, codebooks, and metadata files. The right panel displays the Analytics Workbench in use, where a Jupyter Notebook is used to analyze education-level distributions using bar chart visualizations. This interface supports interactive, code-based data exploration using Python, R, and other tools, enabling secure, scalable, and reproducible analysis directly within the cloud platform.



FAIR Evaluation

To assess the RADx Data Hub's adherence to the FAIR principles, we conducted an internal evaluation using the Research Data Alliance (RDA) FAIR data maturity model [31]. This model defines a structured set of indicators and scoring criteria for evaluating findability, accessibility, interoperability, and reusability. The assessment was carried out by members of the Data Hub team with expertise in metadata standards and data governance. Each indicator was evaluated based on available system features and documentation. A summary of the evaluation results is presented in the Results section, with the complete assessment table provided in Multimedia Appendix 1.

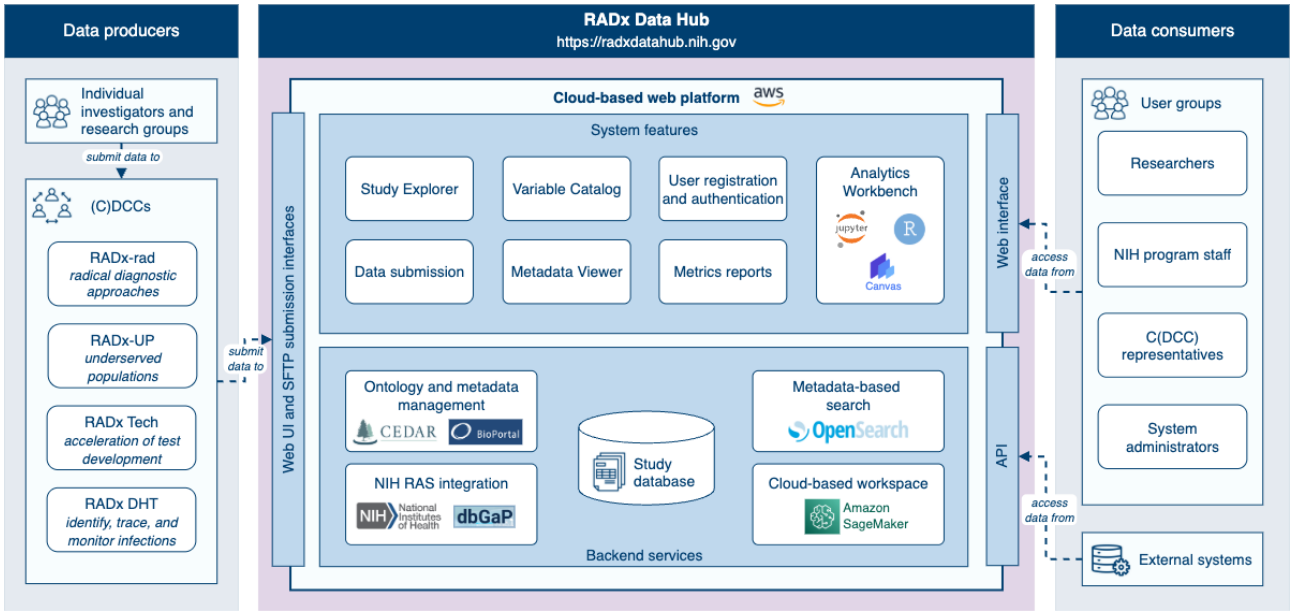
System Architecture

The Data Hub is a cloud-based platform designed to provide a scalable, secure, and FAIR-compliant infrastructure for collecting, harmonizing, and analyzing COVID-19 data. The Data Hub architecture supports the flow of diverse data types from data producers to a centralized data repository, enabling researchers and stakeholders to leverage data effectively. The architecture comprises 2 functional layers—the frontend and backend—both hosted on AWS cloud infrastructure and interconnected through secure application programming interfaces (Figure 7). This architecture is designed to scale with increasing data and user demands, while maintaining high performance and availability.

The frontend layer provides user-facing functionalities through a modern web interface built with React and Next.js. The frontend layer supports dynamic page rendering, device compatibility, and accessibility requirements (including compliance with Section 508 of the Rehabilitation Act [32]). Key features include the Study Explorer, Variables Catalog, and Metadata Viewer, which allow users to browse, search, and explore data and metadata, and access secure user registration and authentication systems. Additional capabilities include the Analytics Workbench, which supports advanced data analysis and visualization through interactive environments such as Jupyter for both R and Python. Complementing these tools, Metrics Reports provide actionable insights into data use, quality, and platform performance, supporting effective system monitoring.

XSL•FO
RenderX

Figure 7. Overview of the Data Hub's software architecture. This diagram presents the architecture of the Data Hub, highlighting its integration of backend services, web features, and data access pathways. On the left, data producers—including individual investigators and (C)DCCs (coordination and data collection centers)—submit study data via secure web or secure file transfer protocol interfaces. The core platform provides system features such as the Study Explorer, Variable Catalog, Metadata Viewer, and Analytics Workbench, with user authentication managed via the National Institutes of Health's Researcher Auth Service (RAS) and data access controlled by the Database of Genotypes and Phenotypes (dbGaP). Backend services support ontology-based metadata management (Center for Expanded Data Annotation and Retrieval [CEDAR], BioPortal), metadata-driven search (OpenSearch), and scalable analytics in a cloud-hosted workspace (Amazon SageMaker). On the right, data consumers—including researchers, National Institutes of Health staff, and system administrators—access harmonized study data through web interfaces or application programming interfaces (APIs) for exploration, analysis, and reuse. The modular architecture supports secure, interoperable, and reproducible digital health research across a wide range of public health applications.



The backend layer is implemented using a microservice architecture powered by Spring Boot and Python. It handles business logic and workflows essential for data ingestion, harmonization, and search. Core components include (1) submission services for validating and processing incoming data, (2) search services powered by AWS OpenSearch for efficient metadata-based retrieval, and (3) an integrated analytics workbench leveraging AWS SageMaker for scalable machine learning workflows. Security and compliance are maintained through robust identity-management systems, including NIH Researcher Auth Service integration and role-based access control.

The system's software architecture is designed for flexibility and scalability to accommodate additional COVID-19 study data, facilitate the addition of new features, and support the potential expansion into other health domains. A Postgres relational database, managed via the AWS Relational Database Service, stores metadata, while raw data files (without metadata) are housed in Amazon Simple Storage Service. This separation ensures efficient querying and modular scalability. AWS OpenSearch indexes metadata to provide fast and precise search capabilities. CEDAR and BioPortal enable metadata annotation and semantic interoperability, ensuring adherence to FAIR principles.

The cloud infrastructure is built on AWS, leveraging serverless technologies, such as AWS Fargate and Lambda for containerized microservices and tasks, and Elastic Load Balancers for high availability. Development, testing, and production environments are isolated to enhance security and streamline deployment workflows. Automated continuous integration and continuous delivery pipelines enable rapid and reliable updates to the system, maintaining operational stability and performance.

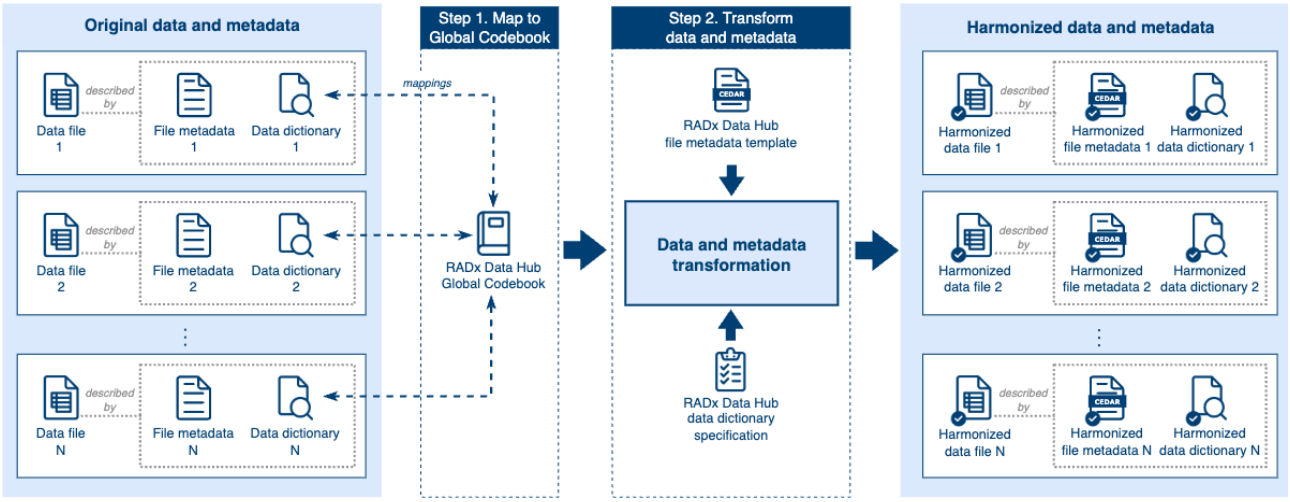
Data and Metadata Harmonization

Harmonization is a key component of the Data Hub's mission to standardize study data, enabling cross-study analyses and the re-exploration of data within individual studies. Harmonization can be approached in two ways: (1) *prospective harmonization*, where common data and metadata standards are defined and established ahead of data collection, and (2) *retrospective harmonization*, which enforces common data standards and ensures consistent representation of semantically equivalent data through careful curation and processing after the data have been collected [33-35]. Given the urgent, time-sensitive nature of the RADx initiative, comprehensive prospective harmonization was not feasible. Developing and implementing advanced standards before data collection would have delayed critical public health research efforts. Therefore, retrospective harmonization was adopted as a pragmatic solution, enabling the standardization of diverse information while ensuring timely data availability for the broader goals of the RADx initiative.

The harmonization process within the Data Hub (Figure 8), involves two main steps: (1) mapping data dictionaries to the Data Hub's Global Codebook [29] and (2) transforming data and metadata to align with standardized templates.

XSL•FO
RenderX

Figure 8. Harmonization workflow for data and metadata. This figure outlines the two-step process used to harmonize incoming data files and metadata within the Data Hub. The Data Hub receives original data files, accompanied by their respective metadata files and data dictionaries (left). In step 1, the original data dictionaries are mapped to the Data Hub Global Codebook for common data elements, standardizing variable definitions and allowable values across studies. In step 2, the original data and metadata are transformed into harmonized formats: data files are transformed using mappings from step 1; metadata files are aligned to a Center for Expanded Data Annotation and Retrieval (CEDAR) metadata template; and data dictionaries are converted to the Data Hub Data Dictionary Specification (center). This process ensures the resulting harmonized data files, metadata, and dictionaries are consistent and interoperable across the Data Hub (right).



The harmonization workflow begins with mapping original data dictionaries provided by investigators to the CDEs in the Global Codebook, which serve as a unified framework for standardizing variable definitions and values across studies. The identification and creation of these mappings are performed manually by the Data Hub team, leveraging their expertise to account for the variability and complexity of submitted data. The resulting mappings are stored in the Global Codebook, creating a centralized, reproducible resource that supports ongoing and future harmonization efforts.

Following the mapping process, data and metadata are transformed to achieve harmonization, ensuring consistency and interoperability across datasets. Data files are transformed by the (C)DCCs using the mappings identified previously. Study-level metadata (eg, principal investigator information, funding sources, and institutional affiliations) are aligned with a dedicated CEDAR metadata template [36]. Similarly, data file metadata (eg, file name, version, and creator) are structured using a CEDAR template [37] inspired by the DataCite Metadata Schema [38]. The DataCite Metadata Schema provides a widely recognized list of standard core metadata properties for citation and retrieval purposes, which the CEDAR template augments by incorporating additional fields tailored to specific needs, such as detailed funding information and file characteristics. By leveraging the strengths of the DataCite standard while extending its functionality, the CEDAR template ensures that metadata are both globally compatible and capable of meeting diverse research requirements. In addition, data dictionaries are converted to the Data Hub Data Dictionary Specification [19], ensuring uniformity in structure and representation.

Figure 9 illustrates an example of data harmonization for the variable *edu_years_of_school*, found in an original RADx-UP data file, representing a study participant's level of education. On the left, the original variable includes values and corresponding definitions outlined in the RADx-UP data dictionary. For example, a value of *3* represents *9th to 12th grade, no diploma*, whereas *4* corresponds to *High school graduate or GED completed*.

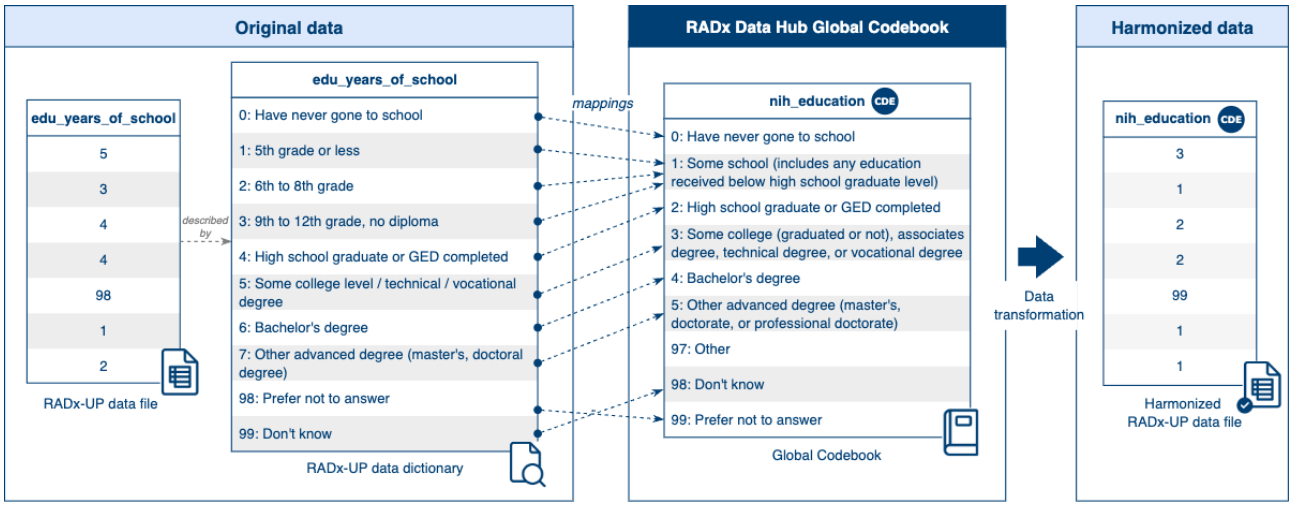
The center panel demonstrates how these original values and definitions are mapped to the corresponding CDE, *nih_education*, within the Data Hub Global Codebook. For instance, the value *4* in the original data file maps to *2: High school graduate or GED completed* in the Global Codebook, ensuring semantic consistency.

Finally, the right panel shows the harmonized data file, where the variable has been renamed *nih_education* and the values updated according to the mappings. This harmonized format ensures that the data are consistent and interoperable within the Data Hub, facilitating cross-study analysis and integration.

The harmonization process is supported by rigorous quality assurance measures conducted by the Data Hub team. Automated tools and structured workflows check for compliance with CDE standards, formatting specifications, and metadata guidelines. Any discrepancies or issues are communicated back to data contributors, creating a feedback loop that promotes iterative improvements and alignment with the Data Hub's standards (Figure 2). This quality assurance process ensures that harmonized data are reliable, consistent, and ready for reuse.

XSL•FO
RenderX

Figure 9. Data harmonization example. This figure illustrates the transformation of an original variable, "edu_years_of_school," from a RADx-UP study into a harmonized common data element (CDE) named nih_education using the Data Hub's Global Codebook. The left panel displays original data values and their definitions as captured in the study-specific data dictionary. In the center, these original values are mapped to standardized values in the Global Codebook, aligning semantically equivalent categories across studies. For instance, "6: Bachelor's degree" in the original dictionary maps to "4: Bachelor's degree" in the CDE. The right panel shows the resulting harmonized data, where the variable name and values have been transformed to conform with the CDE definition, ensuring consistency and interoperability for cross-study analyses.



To ensure semantic consistency across studies, harmonization mappings to CDEs are collaboratively developed by domain experts from both the (C)DCCs and the Data Hub team. Each mapping undergoes manual review to validate alignment with the Data Hub Global Codebook, ensuring standardized interpretation of variables and values. This dual-review approach provides iterative feedback and consensus-building to improve consistency and reproducibility. Future enhancements may include structured auditing protocols or interrater agreement assessments to further strengthen the harmonization quality.

The Data Hub stores both original and harmonized versions of data files. Harmonized data are recommended for cross-study analyses and data re-exploration due to their standardized format and semantic consistency. However, the original raw data are also retained and made accessible for users requiring unmodified data for specialized research needs. This dual offering supports diverse research applications while ensuring the integrity and reusability of the Data Hub's data resources.

Data Deidentification

To protect participant privacy, the (C)DCCs apply a comprehensive deidentification process [39] to study data before their submission to the Data Hub. This process ensures data are anonymized while preserving their scientific utility for secondary research. The (C)DCCs are tasked with balancing PHI and PII removal with the retention of sufficient data to maintain their research value. The Data Hub leverages Amazon Macie at the time of data submission to automatically detect and flag potential PHI and PII. Amazon Macie uses machine learning and pattern matching to identify sensitive data, allowing (C)DCCs to review and apply additional deidentification steps as needed. This automated screening process helps ensure that all data submitted to the Hub meet privacy and regulatory requirements.

The deidentification process follows a standard operating procedure that incorporates several key techniques. Direct identifiers, such as names and addresses, are redacted entirely to eliminate any risk of reidentification. In compliance with Health Insurance Portability and Accountability Act rules, zip codes are generalized by retaining only the first 3 digits, unless the area's population is 20,000 or fewer, in which case the zip code is replaced with *000*. Dates are shifted by a consistent interval to obscure precise temporal details while preserving temporal relationships within the data. Sites within studies are anonymized using random codes, and participant ages are altered: individuals aged <1 years are recorded as *0*, aged 21 to 89 years are modified by adding or subtracting 2 years, and aged >90 are top-coded as *90*.

These procedures are meticulously documented by the (C)DCCs to ensure traceability, replicability, and accountability within their centers. The Data Hub does not receive the internal linkage or detailed records used by the (C)DCCs to perform deidentification; it relies on the (C)DCCs' adherence to the Data Hub's deidentification guidance and on automated PHI and PII detection at submission time. In addition, each submitted file undergoes a manual and thorough deidentification review by the Data Hub team to ensure the absence of PII or PHI, and files are approved for release only if no issues are identified. By applying these deidentification standards, the Data Hub safeguards participant privacy while enabling meaningful secondary use of the data for public health research.

Ethical Considerations

The Data Hub hosts data originally collected by NIH-funded studies and submitted to the Data Hub by third-party (C)DCCs. These (C)DCCs gather data from studies conducted by original investigators, each of which obtained an institutional review board (IRB) approval or a documented waiver at the time of primary data collection. The (C)DCCs are responsible for managing participant consent, including documentation of informed consent forms and any applicable IRB waivers. The Data Hub does not collect data directly from participants. It receives only deidentified data that have been processed by the (C)DCCs in accordance with the Data Hub's deidentification

XSL•FO
RenderX

guidance. Only data from participants who consented to the sharing of their deidentified information are submitted. Because the Data Hub handles only deidentified secondary data for which consent and IRB oversight are managed upstream, no additional IRB approval is required for its operation or for most secondary analyses it supports. However, certain studies may include data elements that require IRB approval at the time of data access request, in accordance with the NIH dbGaP policies. Data access and use are governed by a Data Use Certification through dbGaP, ensuring compliance with data protection and ethical standards. Participant compensation, if applicable, is determined and administered by the original study teams in accordance with their IRB-approved protocols and is not processed or tracked by the Data Hub. The platform does not host or disseminate any identifiable personal information, multimedia, or supplementary materials.

Results

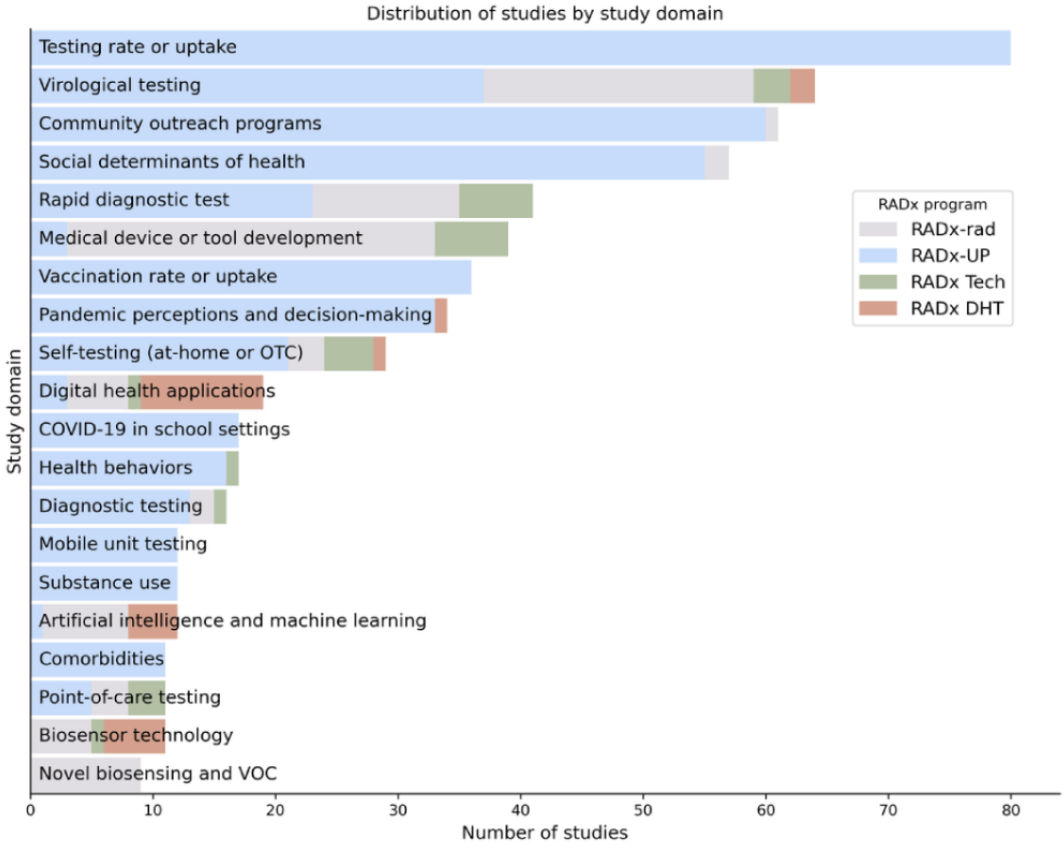
Overview

As of May 2025, the Data Hub hosts 187 studies and over 1700 data files contributed by researchers from more than 100 organizations across 46 US states and territories. The Data Hub represents one of the largest coordinated efforts to centralize COVID-19-related public health data, supporting both immediate and long-term responses to health crises. By adhering to the FAIR principles, the Data Hub enhances access to high-quality, standardized study data for secondary analyses and public health research.

The studies stored in the Data Hub cover a broad range of domains, including community-based research, diagnostic technologies, and surveillance in underserved populations. Metadata and documentation for these studies are openly accessible through the Study Explorer, enabling researchers to evaluate the relevance of datasets before requesting full access via dbGaP [23]. This governance ensures compliance with privacy protections and ethical standards while facilitating secure access to sensitive health data.

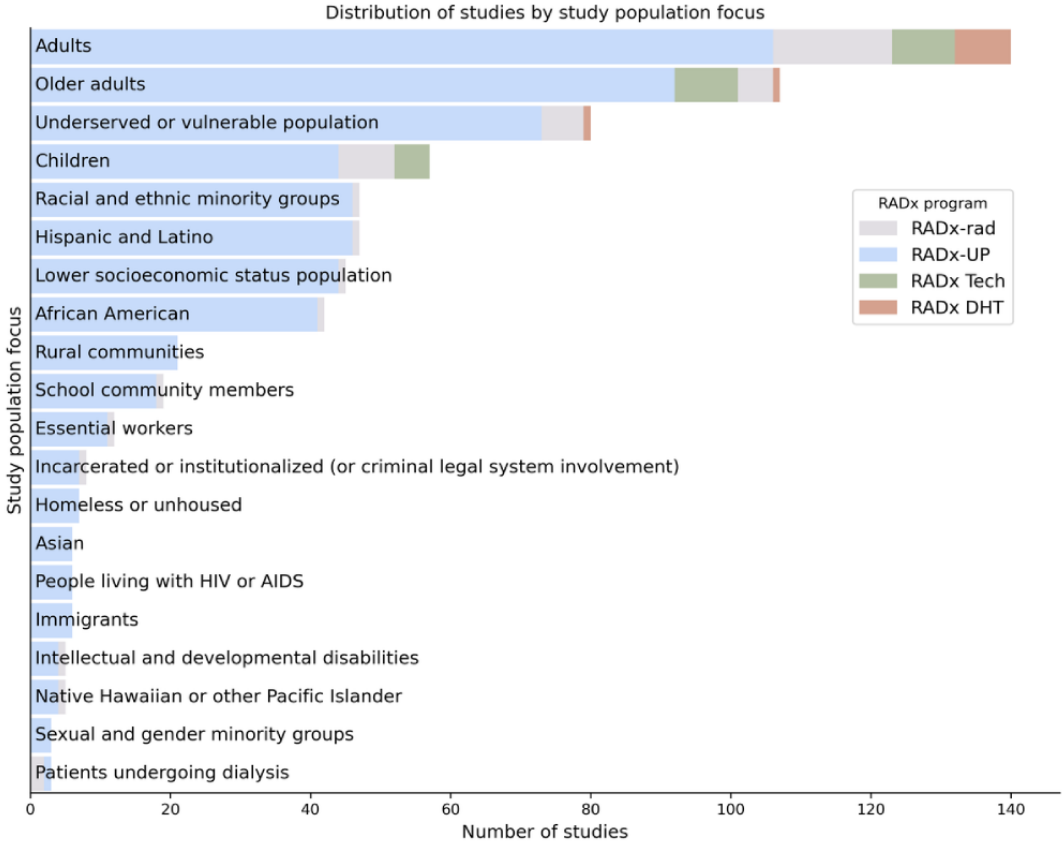
Figures 10 to 13 show the distribution of studies by domain, population focus, data collection method, and study design. Stacked bar charts highlight the contributions from the different RADx programs, illustrating the diverse focus areas and methodological approaches across RADx initiatives. The RADx-UP program, with its emphasis on health disparities, contributes a substantial portion of these studies. Many studies focus on underserved populations, including racial minority groups, older adults, and individuals with low socioeconomic status. Data collection methods, such as surveys, wearable device monitoring, and diagnostic testing, offer opportunities for cross-study integration. Most studies follow observational designs.

Figure 10. Distribution of Data Hub studies by study domain. This histogram illustrates the distribution of studies across various research domains, highlighting the diverse focus areas of RADx studies. RADx-UP accounts for the largest share of studies, particularly in domains such as community outreach, social determinants of health, and vaccination. The notable representation of RADx-rad in areas like rapid diagnostics and biosensing underscores the emphasis on novel and experimental approaches to COVID-19 detection and monitoring. This distribution highlights the cross-programmatic diversity and strategic focus areas of the RADx initiative. Note: Only the top 20 study domains are shown; additional study domains were excluded for clarity.



XSL•FO
RenderX

Figure 11. Distribution of Data Hub studies by population focus. This chart presents the breakdown of studies based on population focus. The figure underscores the RADx-UP program's emphasis on underserved and vulnerable populations, with substantial representation of studies involving older adults, racial and ethnic minorities, low socioeconomic status groups, rural communities, and incarcerated or institutionalized individuals. While many studies target specific at-risk groups—including children, essential workers, and people living with HIV/AIDS—a significant number also focus on general adult populations. This distribution highlights the wide demographic coverage of the Data Hub, supporting diverse and comparative analyses on health disparities, equity, and access in the context of the COVID-19 pandemic. Note: Only the top 20 population focus categories are shown; additional categories were excluded for clarity.



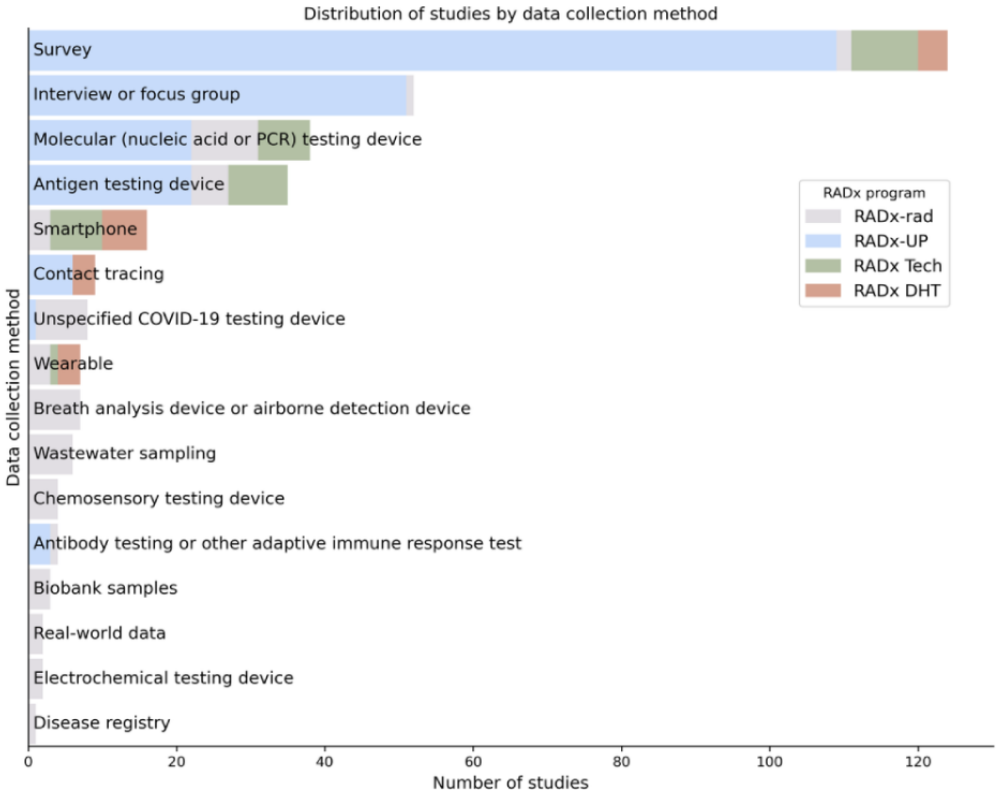
The Data Hub streamlines data discovery through tools such as the Study Explorer, Variables Catalog, and Metadata Viewer. Each study is assigned a persistent DOI to support long-term access and citation. The platform's metadata-driven infrastructure enables researchers to efficiently search for and explore studies and variables, reducing barriers to data reuse and collaboration.

The Analytics Workbench offers a secure, cloud-based environment where researchers can access harmonized datasets and perform analyses using tools such as Jupyter and RStudio. By enabling data manipulation, visualization, and advanced analyses directly on the platform, the Workbench reduces local resource demands and improves reproducibility. Researchers can analyze trends, compare study outcomes, and conduct large-scale investigations efficiently without managing their own infrastructure.

The Data Hub's ability to centralize and harmonize data is crucial in addressing key public health questions related to the COVID-19 pandemic. By enabling multistudy analyses, the Data Hub can facilitate research on health disparities, diagnostic effectiveness, and public health response strategies, reinforcing its role as a crucial resource for pandemic preparedness and response, with significant potential for long-term adaptation to future public health challenges.

XSL•FO
RenderX

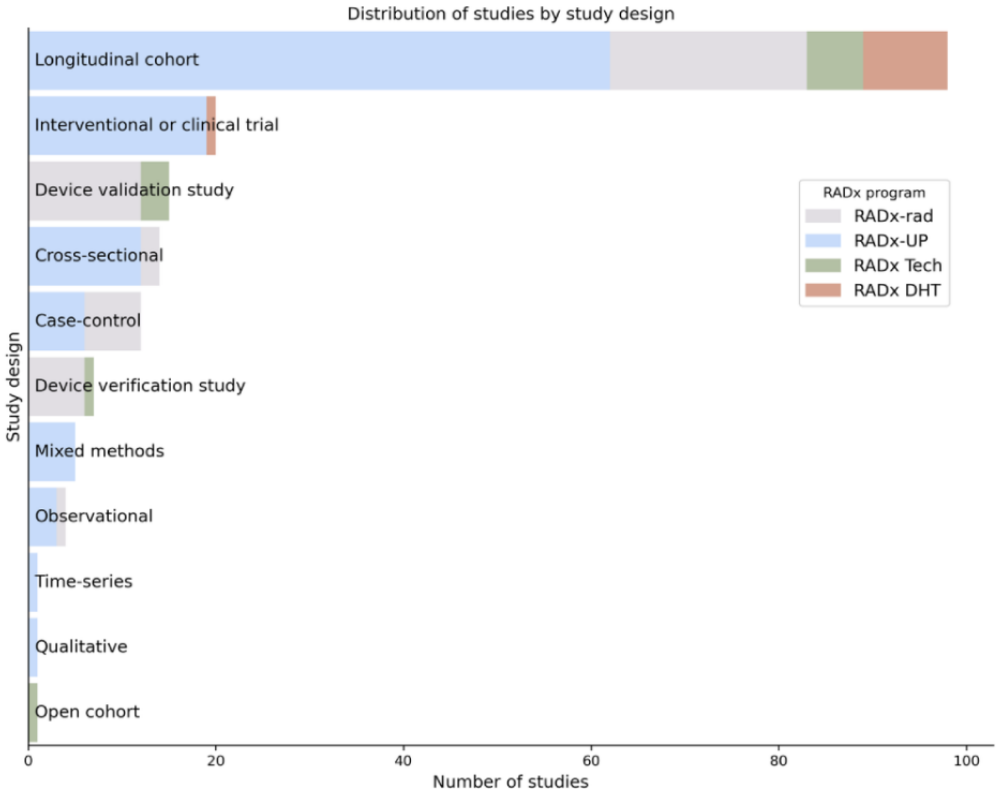
Figure 12. Distribution of Data Hub studies by data collection method. This chart categorizes studies based on their data collection approaches, stratified by RADx program. Survey-based and interview or focus group methods are predominant, particularly among RADx-UP studies, reflecting an emphasis on community engagement and participant-reported outcomes. RADx Tech and RADx DHT programs contribute more heavily to studies utilizing device-based methods, including molecular and antigen testing, wearables, smartphones, and contact tracing tools. Less common but innovative modalities—such as wastewater sampling, breath and chemosensory testing, and electrochemical diagnostics—demonstrate methodological breadth across the RADx portfolio. The heterogeneity of data sources supports integration and cross-study analyses for richer insights into pandemic-related health outcomes.



The Data Hub's value is evidenced by the substantial number of peer-reviewed publications derived from the studies it hosts. Since its inception, the RADx initiative has produced over 900 peer-reviewed publications across its constituent programs, addressing topics in diagnostics, health disparities, behavioral science, and digital health. The Data Hub serves as the central repository for the study data underlying many of these publications. Some examples of relevant RADx-supported studies include a qualitative investigation of vaccine hesitancy among staff in skilled nursing facilities [40], environmental surveillance of airborne SARS-CoV-2 particles [41], community-based testing strategies implemented in rural Alabama [42], an epidemiological study on the transition from Delta to Omicron variants [43], modeling responses to pathogens using laboratory-grown placenta models [44], and an analysis of school masking policies and secondary transmission of COVID-19 [45]. These publications exemplify the scientific breadth and translational relevance of RADx-funded research and highlight the Data Hub's pivotal role in enabling data-driven discovery and public health innovation.

XSL•FO
RenderX

Figure 13. Distribution of Data Hub studies by study design. This chart illustrates the frequency of study designs used across the Data Hub, categorized by program. Longitudinal cohort designs are by far the most common, particularly within RADx-UP, reflecting a focus on long-term, community-based public health research. Interventional or clinical trial designs also feature prominently, especially in RADx Tech and RADx DHT, which emphasize testing and digital health innovation. Additional designs—including cross-sectional, case-control, device validation and verification studies, and mixed methods—demonstrate methodological diversity across the platform. This variety supports a wide range of analyses, from exploratory to evaluative, across observational and experimental frameworks.



FAIR Evaluation Results

We evaluated the Data Hub's alignment with the FAIR principles using the RDA FAIR data maturity model [31], which provides a structured set of indicators across 4 key dimensions: findability, accessibility, interoperability, and reusability. The assessment revealed that the platform demonstrates strong foundational support for FAIR data practices, with several areas for continued improvement. A summary of these findings is provided in Table 1, and a full breakdown of the FAIR maturity assessment is provided in Multimedia Appendix 1.

Findability is supported through the assignment of persistent DOIs to each study. These DOIs resolve to public Study Overview pages that provide descriptive metadata, documentation, and links to related data artifacts. Metadata are structured using CEDAR templates and enriched with ontology terms from BioPortal, enabling semantic search and discovery. The Study Explorer interface further enhances findability by offering keyword search, filtering, and browsing tools.

Accessibility is facilitated through open access to metadata without requiring authentication. Metadata and public data files can be accessed via HTTPS, including support for automated download of synthetic data. For controlled-access data, each Study Overview page links directly to its dbGaP accession, and detailed instructions are provided for users to request access. While authorization is required for sensitive data, access protocols are standardized and widely supported.

Interoperability is enabled by the use of CEDAR templates that incorporate controlled vocabularies, as well as harmonization of many data fields using CDEs from the RADx Global Codebook. Data dictionaries define variable-level metadata and enable machine processing. These practices support semantic consistency across data files and improve the platform's capacity for integration and automated interpretation.

Reusability is supported through study overview pages that include data use limitations and metadata structured using CEDAR templates for machine readability. Provenance is tracked with the provenance, authoring, and versioning ontology, and metadata files include file-level details, such as contributor information, creation dates, and versioning. Harmonization to CDEs promotes consistent representation of core variables across studies.

XSL•FO
RenderX

Table 1. Summary of Rapid Acceleration of Diagnostics (RADx) Data Hub's adherence to the findable, accessible, interoperable, reusable (FAIR) principles based on the Research Data Alliance (RDA) FAIR data maturity model.

FAIR principle	Strengths	Areas for improvement
Findable	Each study is assigned a DOI[a] that resolves to a study overview page; detailed and structured metadata are available in machine-readable formats; easy-to-use search and filter tools	No persistent IDs for individual metadata or data records; some metadata lack direct links to data files
Accessible	Metadata include human-readable access information and are manually accessible without log-in; data can be manually accessed via dbGaP[b] or direct download for public data; study DOIs resolve to overview pages; metadata and data are accessed via HTTPS; public data support automated HTTPS access; metadata remain accessible even if data are not	Metadata lack machine-actionable access information; individual metadata and data files lack persistent IDs; no API[c] or token-based system for automated access to restricted data; no policy for long-term metadata availability
Interoperable	Metadata are partially annotated with controlled terms from BioPortal; data dictionaries enable machine processing; data fields are harmonized using RADx CDEs[d]	Some metadata fields still use free text and lack URIs[e]; CDEs do not cover all data fields; metadata cross-referencing is still under development; data files lack links to related data or metadata
Reusable	Each study overview page includes data use limitations; provenance is tracked using PAV[f] ontology; metadata follows CEDAR[g] specifications for machine readability; data use RADx standards and CDEs	Broader provenance standards and more controlled vocabularies are needed to improve machine understanding

[a]DOI: digital object identifier.

[b]dbGaP: Database of Genotypes and Phenotypes.

[c]API: application programming interface.

[d]CDE: common data element.

[e]URI: uniform resource identifier.

[f]PAV: provenance, authoring, and versioning.

[g]CEDAR: Center for Expanded Data Annotation and Retrieval.

While these features reflect strong FAIR alignment, several areas for improvement remain. For findability, individual data and metadata artifacts—such as specific data files and dictionaries—currently lack persistent identifiers, and DOIs do not resolve directly to stand-alone machine-readable metadata records. In the accessibility dimension, although metadata are openly available, access conditions for restricted data are not yet expressed in machine-actionable formats, and no application programming interface or token-based mechanism is available for automating controlled data access. Regarding interoperability, some key metadata fields, such as study design or domain, still accept free-text values or lack uniform resource identifier–based annotations, and metadata cross-referencing between study- and file-level records is still under development. For reusability, broader provenance standards (eg, PROV Ontology) are not yet implemented, and semantic alignment is incomplete for variables not covered by the Global Codebook.

Discussion

Principal Findings

The Data Hub provides a centralized platform to support rapid public health response to the COVID-19 pandemic, overcoming the challenges posed by fragmented public health data systems. By promoting data sharing, reuse, and analysis, the Data Hub enhances data usability and fosters collaboration among researchers, public health officials, and policy makers, ultimately strengthening the ability to respond swiftly and effectively to public health crises, such as the COVID-19 pandemic. The Data Hub uses modern integration techniques and robust data governance practices to ensure data security, privacy protection, and compliance with regulatory standards. The infrastructure and data management techniques developed as part of the Data Hub serve as invaluable tools to combat future public health events by maximizing the utility and impact of shared data within the public health and research communities.

Through rich, structured metadata, the Data Hub improves data findability and enables researchers to quickly assess whether the collected data aligns with their research interests and topic areas before initiating access requests. Moreover, the curation and standardization of data throughout the platform is intended to ensure data quality, resolve inconsistencies and heterogeneity where possible, and identify missing information. Standardizing data formats and metadata enhances interoperability and facilitates data integration, enabling the creation of aggregate data files for large-scale cross-study analyses. These capabilities are vital for evaluating and understanding the COVID-19 pandemic response, addressing disparities in intervention outcomes, advancing biomedical technologies such as testing and rapid detection, and preparing for future pandemics [46-48]. Importantly, the Analytics Workbench reduces common infrastructure barriers—such as the need for local high-performance computing resources, complex software installation, and advanced IT support—by providing a secure, cloud-based environment for reproducible analysis. This design enables more equitable access to advanced data analysis capabilities, particularly benefiting researchers in low-resource institutions or regions where such infrastructure may be limited.

XSL•FO
RenderX

Unlike some pandemic data platforms, the Data Hub prioritizes retrospective harmonization and metadata-driven interoperability. Harmonized data are supported by CDEs specified in the Global Codebook, structured metadata templates via CEDAR, and ontology annotations through BioPortal. These components work in concert with automated validation tools and collaborative curation workflows to ensure semantic consistency across studies. By aligning heterogeneous variables, resolving data quality issues, and enabling ready-to-use, interoperable study data, the platform strengthens the reusability and integration of public health data for secondary research.

Limitations

While the Data Hub has made significant strides in facilitating data sharing and reuse, its development and implementation have highlighted key challenges in managing a large-scale, multisource data platform. Retrospective harmonization remains a major challenge due to the diversity of study types—ranging from human participants to environmental monitoring—which often require extensive curation to align with evolving metadata models and the RADx Global Codebook. Although data stewards assist contributors during submission, the variability in data quality and standardization across studies demands ongoing refinement. Some datasets must be accepted with minimal curation to ensure repository completeness, emphasizing the need for sustained collaboration with (C)DCCs. In addition, while the data access process—governed by dbGaP and the Data Hub's user code of conduct—supports ethical and legal compliance, it can pose usability barriers. Researchers must evaluate the potential utility of data using only metadata and study overviews before initiating access requests.

Despite strong alignment with many aspects of the FAIR principles, our structured evaluation using the RDA FAIR data maturity model revealed several limitations that we are actively addressing. These include the lack of persistent identifiers for individual data and metadata artifacts, the continued use of free-text fields where standardized vocabularies would improve consistency, and incomplete links from metadata to the associated data files. We also plan to enhance machine-actionability by encoding access conditions in structured, reusable formats. Addressing these gaps is a key focus of our development road map for advancing FAIR compliance. A detailed assessment is provided in Multimedia Appendix 1.

Conclusions

The Data Hub addresses critical challenges in public health data management by providing a centralized, FAIR-compliant platform that integrates diverse datasets for secondary analyses. Through its scalable architecture, metadata-driven workflows, and robust privacy measures, the Data Hub has become a pivotal resource for facilitating cross-study research. By enabling rapid access to harmonized COVID-19 data, the platform advances research efforts on health disparities, diagnostic tools, and pandemic preparedness. The cloud-based Analytics Workbench further supports secure, reproducible analysis without local infrastructure requirements. Ongoing enhancements include a software framework to automate integration of heterogeneous data [49] and improvements in metadata quality and usability.

Beyond COVID-19, the Hub's modular design and governance model position it as a template for future public health data infrastructures. Features such as ontology-annotated metadata, Open Researcher and Contributor Identifier [50], and Research Organization Registry [51] integration, and improved metadata visualization are laying the groundwork for a more interoperable and user-friendly research ecosystem. Planned developments—such as a user-friendly data dictionary viewer, artifact-level persistent identifiers, and expanded use of standardized vocabularies and provenance—directly respond to gaps identified in our FAIR evaluation. These efforts aim to ensure that data remain findable, accessible, interoperable, and reusable across evolving needs. By enabling scalable, standards-based data reuse, the RADx Data Hub supports more agile, collaborative, and equitable research—critical for evidence-based responses to current and future public health challenges.

Acknowledgments

This work was supported by the National Institutes of Health under Other Transactions Authority award 1OT2DB000009. The authors are grateful to the Data Hub collaborators, including Rapid Acceleration of Diagnostics study investigators and the coordination and data collection centers, for their essential contributions. The authors also appreciate the support of the Database of Genotypes and Phenotypes, the Center for Expanded Data Annotation and Retrieval, and the BioPortal teams for facilitating data governance, metadata, and ontology management activities. They convey special thanks to the participants of Rapid Acceleration of Diagnostics-funded studies for their invaluable contributions. The authors acknowledge the members of the Data Hub User Advisory Board for their valuable insights and contributions: Liesl Jeffers-Francis, Oswaldo Alonso Lozoya, Nader Mehri, Thais Rivas, and Peter Rose. In addition, they thank Lucy L Hsu, Vivian Ota Wang, and Susan K Gregurick for their support. The authors are especially grateful to Mimi Duong-Quang, Sarah Lai, Tracy Maruo-Hanamoto, Ann Nguyen, Karen Stass, and Debra Vavreck, whose administrative and financial efforts were instrumental in ensuring the smooth operation of the project. They also acknowledge Josef Hardi and Rafael Gonçalves for their valuable contributions to technical discussions. Finally, the authors recognize the former members of the Data Hub team, whose dedication and efforts have been vital to the project's success: John Graybeal, Marian Mersmann, Alexandra Shlionskaya, and Stanley C Ahalt.

XSL•FO
RenderX

Data Availability

The datasets hosted in the Rapid Acceleration of Diagnostics (RADx) Data Hub are not publicly available due to participant privacy protections and data governance requirements but may be accessed by qualified researchers upon reasonable request. Instructions for submitting a data access request are provided on the RADx Data Hub Resource Center web page [27].

Authors' Contributions

MM-R, JKY, MJOC, and MUA led the conceptualization, drafting, and revision of the manuscript. The remaining authors reviewed the manuscript, provided critical feedback, and contributed to revisions to enhance its clarity and rigor. MAM, MAK, and AK supervised the project and provided strategic guidance. All authors reviewed and approved the final manuscript.

The Data Hub team represents the broader consortium of official project members who have actively participated in designing, implementing, maintaining, or promoting the Data Hub platform. The individual members of the Data Hub team are as follows, grouped by institution and listed alphabetically by surname. Stanford Center for Biomedical Informatics Research, Stanford University: MUA, Luna Baalbaki, YC, Michael Dorf, Attila L Egyedi, Crystal Han, MH, Christian Kindermann, MM-R, MAM, MJOC, Alex Skrenchuk, Sowmya Somasundaram, and JKY. Booz Allen Hamilton Inc: Mahesh Ahilandeswaran, Haven Barnes, Pierce Beckett, Johanna Brown, Beth DiGiulian, AF, Irina Gorelik, Lea Jih-Vieira, Krishnaveni Kaladi, MAK, Marquea King, Nino Koberidze, Chun Yee Lau, Patrick Lenhart, Li Li, Amy Lin, Anna Lu, Emma Luker, NM, Ashley Pryor, Bala Ramani, Joseph Schweikert, AS, Vivian Tran, and SV. Renaissance Computing Institute, University of North Carolina at Chapel Hill: Asiyah Ahmad, Ashley Evans, Brandy Farlow, Vicki Gardner, Suparna Goswami, AH, Robert Hubal, Yaphet Kebede, AK, Leah Mason, JOM, Keriayn Smith, Jasmine Snipe, SS, Josh Terrell, and AW.

Conflicts of Interest

None declared.

Multimedia Appendix 1

FAIR assessment results for the RADx Data Hub, conducted using the Research Data Alliance's FAIR Data Maturity Model. This spreadsheet details the platform's compliance with each FAIR subprinciple, including identified gaps and planned enhancements.

[XLSX File (Microsoft Excel File), 16 KB - publichealth_v11i1e72677_app1.xlsx]

References

1. Massoudi BL, Sobolevskaia D. Keep moving forward: health informatics and information management beyond the COVID-19 pandemic. Yearb Med Inform 2021 Aug;30(1):75-83. [doi: 10.1055/s-0041-1726499] [Medline: 34479380]
2. Bhatia S, Imai N, Watson OJ, Abbood A, Abdelmalik P, Cornelissen T, et al. Lessons from COVID-19 for rescalable data collection. Lancet Infect Dis 2023 Sep;23(9):e383-e388 [FREE Full text] [doi: 10.1016/S1473-3099(23)00121-4] [Medline: 37150186]
3. Galaitsi SE, Cegan JC, Volk K, Joyner M, Trump BD, Linkov I. The challenges of data usage for the United States' COVID-19 response. Int J Inf Manage 2021 Aug;59:102352 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2021.102352] [Medline: 33824545]
4. Haendel MA, Chute CG, Bennett TD, Eichmann DA, Guinney J, Kibbe WA, et al. The National COVID Cohort Collaborative (N3C): rationale, design, infrastructure, and deployment. J Am Med Inform Assoc 2021 Mar 01;28(3):427-443 [FREE Full text] [doi: 10.1093/jamia/ocaa196] [Medline: 32805036]
5. Pfaff ER, Girvin AT, Bennett TD, Bhatia A, Brooks IM, Deer RR, et al. Identifying who has long COVID in the USA: a machine learning approach using N3C data. Lancet Digit Health 2022 Jul;4(7):e532-e541 [FREE Full text] [doi: 10.1016/S2589-7500(22)00048-6] [Medline: 35589549]
6. The European COVID-19 data platform. COVID-19 Data Portal. URL: https://www.covid19dataportal.org/the-european-covid-19-data-platform [accessed 2025-07-29]
7. Harrison PW, Lopez R, Rahman N, Allen SG, Aslam R, Buso N, et al. The COVID-19 data portal: accelerating SARS-CoV-2 and COVID-19 research through rapid open access data sharing. Nucleic Acids Res 2021 Jul 02;49(W1):W619-W623 [FREE Full text] [doi: 10.1093/nar/gkab417] [Medline: 34048576]
8. Azzini AM, Canziani LM, Davis RJ, Mirandola M, Hoelscher M, Meyer L, et al. How European research projects can support vaccination strategies: the case of the ORCHESTRA project for SARS-CoV-2. Vaccines (Basel) 2023 Aug 14;11(8):1361 [FREE Full text] [doi: 10.3390/vaccines11081361] [Medline: 37631929]
9. Nab L, Schaffer AL, Hulme W, DeVito NJ, Dillingham I, Wiedemann M, et al. OpenSAFELY: a platform for analysing electronic health records designed for reproducible research. Pharmacoepidemiol Drug Saf 2024 Jun;33(6):e5815. [doi: 10.1002/pds.5815] [Medline: 38783412]
10. Data modernization initiative (DMI). Centers for Disease Control and Prevention. URL: https://www.cdc.gov/data-modernization/php/about/dmi.html [accessed 2025-05-22]

XSL•FO
RenderX

11. The Infectious Diseases Data Observatory homepage. The Infectious Diseases Data Observatory. URL: https://www.iddo.org/ [accessed 2025-05-19]
12. Global.health: a data science initiative. Global.health. URL: https://global.health [accessed 2025-05-19]
13. Maxwell L, Shreedhar P, Dauga D, McQuilton P, Terry RF, Denisiuk A, et al. FAIR, ethical, and coordinated data sharing for COVID-19 response: a scoping review and cross-sectional survey of COVID-19 data sharing platforms and registries. Lancet Digit Health 2023 Oct;5(10):e712-e736 [FREE Full text] [doi: 10.1016/S2589-7500(23)00129-2] [Medline: 37775189]
14. Tromberg BJ, Schwetz TA, Pérez-Stable EJ, Hodes RJ, Woychik RP, Bright RA, et al. Rapid scaling up of Covid-19 diagnostic testing in the United States - the NIH RADx initiative. N Engl J Med 2020 Sep 10;383(11):1071-1077 [FREE Full text] [doi: 10.1056/NEJMsr2022263] [Medline: 32706958]
15. NIH RADx Data Hub homepage. RADx Data Hub. URL: https://radxdatahub.nih.gov/ [accessed 2024-05-10]
16. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3:160018 [FREE Full text] [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
17. RADx programs. National Institutes of Health National Institutes of Health. URL: https://www.nih.gov/research-training/medical-research-initiatives/radx/radx-programs [accessed 2024-11-26]
18. RAPIDS. URL: https://rapids.ll.mit.edu/ [accessed 2025-08-05]
19. RADx data dictionary specification. GitHub. URL: https://github.com/bmir-radx/radx-data-dictionary-specification/blob/main/radx-data-dictionary-specification.md [accessed 2024-11-29]
20. Musen MA, Bean CA, Cheung KH, Dumontier M, Durante KA, Gevaert O, et al. The center for expanded data annotation and retrieval. J Am Med Inform Assoc 2015 Nov;22(6):1148-1152 [FREE Full text] [doi: 10.1093/jamia/ocv048] [Medline: 26112029]
21. Musen MA, O'Connor MJ, Schultes E, Martínez-Romero M, Hardi J, Graybeal J. Modeling community standards for metadata as templates makes data FAIR. Sci Data 2022 Nov 12;9(1):696 [FREE Full text] [doi: 10.1038/s41597-022-01815-3] [Medline: 36371407]
22. Noy NF, Shah NH, Whetzel PL, Dai B, Dorf M, Griffith N, et al. BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res 2009 Jul;37(Web Server issue):W170-W173 [FREE Full text] [doi: 10.1093/nar/gkp440] [Medline: 19483092]
23. Welcome to dbGaP. National Institutes of Health National Library of Medicine. URL: https://dbgap.ncbi.nlm.nih.gov/ [accessed 2024-11-28]
24. Cedar embeddable editor (CEE). GitHub. URL: https://github.com/metadatacenter/cedar-embeddable-editor [accessed 2024-12-03]
25. Whetzel PL, Noy NF, Shah NH, Alexander PR, Nyulas C, Tudorache T, et al. BioPortal: enhanced functionality via new web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic Acids Res 2011 Jul;39(Web Server issue):W541-W545 [FREE Full text] [doi: 10.1093/nar/gkr469] [Medline: 21672956]
26. RADx user code of conduct. RADx Data Hub. 2024. URL: https://radxdatahub.nih.gov/radx-s3-resources/RADxDataHubCodeOfConduct.pdf [accessed 2024-11-30]
27. NIH RADx data hub - resource center. RADx Data Hub. URL: https://radxdatahub.nih.gov/resourceCenter [accessed 2024-11-29]
28. Huser V, Amos L. Analyzing real-world use of research common data elements. AMIA Annu Symp Proc 2018;2018:602-608 [FREE Full text] [Medline: 30815101]
29. RADx data hub global codebook. Zenodo. URL: https://zenodo.org/records/16754495 [accessed 2024-11-30]
30. O'Connor MJ, Martínez-Romero M, Egyedi AL, Willrett D, Graybeal J, Musen MA. An open repository model for acquiring knowledge about scientific experiments. In: Proceedings of the 20th International Conference on Knowledge Engineering and Knowledge Management. 2016 Presented at: EKAW 2016; November 19-23, 2016; Bologna, Italy. [doi: 10.1007/978-3-319-49004-5_49]
31. Bahim C, Casorrán-Amilburu C, Dekkers M, Herczog E, Loozen N, Repanas K, et al. The FAIR data maturity model: an approach to harmonise fair assessments. Data Sci J 2020;19 [FREE Full text] [doi: 10.5334/dsj-2020-041]
32. Section508.gov homepage. Section508.gov. URL: https://www.section508.gov/ [accessed 2024-12-02]
33. Cheng C, Messerschmidt L, Bravo I, Waldbauer M, Bhavikatti R, Schenk C, et al. A general primer for data harmonization. Sci Data 2024 Jan 31;11(1):152 [FREE Full text] [doi: 10.1038/s41597-024-02956-3] [Medline: 38297013]
34. Fortier I, Raina P, Van den Heuvel ER, Griffith LE, Craig C, Saliba M, et al. Maelstrom research guidelines for rigorous retrospective data harmonization. Int J Epidemiol 2017 Mar 01;46(1):103-105 [FREE Full text] [doi: 10.1093/ije/dyw075] [Medline: 27272186]
35. Cheng C, Messerschmidt L, Bravo I, Waldbauer M, Bhavikatti R, Schenk C, et al. Harmonizing government responses to the COVID-19 pandemic. Sci Data 2024 Mar 14;11(1):204 [FREE Full text] [doi: 10.1038/s41597-023-02881-x] [Medline: 38355867]
36. RADx data hub study metadata template. Zenodo. 2025 Jan 11. URL: https://zenodo.org/records/14630797 [accessed 2025-07-29]

XSL•FO
RenderX

37. RADx data hub data file metadata template. Zenodo. URL: https://zenodo.org/records/14630748 [accessed 2025-07-29]
38. DataCite metadata schema documentation for the publication and citation of research data and other research outputs. DataCite. URL: https://datacite-metadata-schema.readthedocs.io/en/4.6/ [accessed 2024-12-02]
39. De-identification guidance. RADx Data Hub. URL: https://radxdatahub.nih.gov/radx-s3-resources/RADx_Data_Hub-DeIdentification_Guidance.pdf [accessed 2024-11-05]
40. Harrison J, Berry S, Mor V, Gifford D. "Somebody like me": understanding COVID-19 vaccine hesitancy among staff in skilled nursing facilities. J Am Med Dir Assoc 2021 Jun;22(6):1133-1137 [FREE Full text] [doi: 10.1016/j.jamda.2021.03.012] [Medline: 33861978]
41. Puthussery JV, Ghumra DP, McBrearty KR, Doherty BM, Sumlin BJ, Sarabandi A, et al. Real-time environmental surveillance of SARS-CoV-2 aerosols. Nat Commun 2023 Jul 10;14(1):3692 [FREE Full text] [doi: 10.1038/s41467-023-39419-z] [Medline: 37429842]
42. McCollum CG, Creger TN, Rana AI, Matthews LT, Baral SD, Burkholder GA, et al. COVID community-engaged testing in Alabama: reaching underserved rural populations through collaboration. Am J Public Health 2022 Oct;112(10):1399-1403. [doi: 10.2105/AJPH.2022.306985] [Medline: 35952331]
43. Fall A, Eldesouki RE, Sachithanandham J, Morris CP, Norton JM, Gaston DC, et al. The displacement of the SARS-CoV-2 variant Delta with Omicron: an investigation of hospital admissions and upper respiratory viral loads. EBioMedicine 2022 May;79:104008 [FREE Full text] [doi: 10.1016/j.ebiom.2022.104008] [Medline: 35460989]
44. Karvas RM, Khan SA, Verma S, Yin Y, Kulkarni D, Dong C, et al. Stem-cell-derived trophoblast organoids model human placental development and susceptibility to emerging pathogens. Cell Stem Cell 2022 May 05;29(5):810-25.e8 [FREE Full text] [doi: 10.1016/j.stem.2022.04.004] [Medline: 35523141]
45. Boutzoukas AE, Zimmerman KO, Inkelas M, Brookhart MA, Benjamin DK, Butteris S, et al. School masking policies and secondary SARS-CoV-2 transmission. Pediatrics 2022 Jun 01;149(6):e2022056687 [FREE Full text] [doi: 10.1542/peds.2022-056687] [Medline: 35260896]
46. Happel C, Peñalber-Johnstone C, Tagle DA. Pivoting novel exosome-based technologies for the detection of SARS-CoV-2. Viruses 2022 May 18;14(5):1083 [FREE Full text] [doi: 10.3390/v14051083] [Medline: 35632824]
47. Corbie G, D'Agostino EM, Knox S, Richmond A, Woods CW, Dave G, et al. RADx-UP coordination and data collection: an infrastructure for COVID-19 testing disparities research. Am J Public Health 2022 Nov;112(S9):S858-S863 [FREE Full text] [doi: 10.2105/AJPH.2022.306953] [Medline: 36194852]
48. Mannino RG, Nehl EJ, Farmer S, Peagler AF, Parsell MC, Claveria V, et al. The critical role of engineering in the rapid development of COVID-19 diagnostics: lessons from the RADx Tech Test Verification Core. Sci Adv 2023 Apr 07;9(14):eade4962 [FREE Full text] [doi: 10.1126/sciadv.ade4962] [Medline: 37027461]
49. Yu JK, Martínez-Romero M, Horridge M, Akdogan MU, Musen MA. A general-purpose data harmonization framework: supporting reproducible and scalable data integration in the RADx Data Hub. ArXiv. Preprint posted online on March 3, 2025 2025 [FREE Full text]
50. Open Researcher and Contributor Identifier (ORCID) homepage. Open Researcher and Contributor Identifier (ORCID). URL: https://orcid.org [accessed 2025-01-06]
51. Research Organization Registry (ROR) homepage. Research Organization Registry (ROR). URL: https://ror.org [accessed 2025-01-06]

Abbreviations

(C)DCC: coordination and data collection center
AWS: Amazon Web Services
CEDAR: Center for Expanded Data Annotation and Retrieval
CDE: common data element
dbGaP: Database of Genotypes and Phenotypes
DOI: digital object identifier
FAIR: findable, accessible, interoperable, reusable
IRB: institutional review board
NIH: National Institutes of Health
PHI: protected health information
PII: personally identifiable information
RADx: Rapid Acceleration of Diagnostics
RADx DHT: RADx Digital Health Technologies
RADx-rad: RADx Radical
RADx-UP: RADx Underserved Populations

XSL•FO
RenderX

Edited by A Mavragani; submitted 14.02.25; peer-reviewed by J Tenenbaum, CH Bernabe, E Mensah; comments to author 22.04.25; revised version received 13.06.25; accepted 04.07.25; published 20.08.25.

Please cite as:
Martínez-Romero M, Horridge M, Mistry N, Weyhmiller A, Yu JK, Fujimoto A, Henry A, O'Connor MJ, Sier A, Suber S, Akdogan MU, Cao Y, Valliappan S, Mieczkowska JO, Krishnamurthy A, Keller MA, Musen MA, RADx Data Hub Team
A Cloud-Based Platform for Harmonized COVID-19 Data: Design and Implementation of the Rapid Acceleration of Diagnostics (RADx) Data Hub
JMIR Public Health Surveill 2025;11:e72677
URL: https://publichealth.jmir.org/2025/1/e72677
doi:10.2196/72677
PMID:40834404

©Marcos Martínez-Romero, Matthew Horridge, Nilesh Mistry, Aubrie Weyhmiller, Jimmy K Yu, Alissa Fujimoto, Aria Henry, Martin J O'Connor, Ashley Sier, Stephanie Suber, Mete U Akdogan, Yan Cao, Somu Valliappan, Joanna O Mieczkowska, Ashok Krishnamurthy, Michael A Keller, Mark A Musen, RADx Data Hub Team. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prevalence, Risk Factors, Disease-Related Knowledge, and Vaccination Attitudes and Behaviors for Long COVID Among French Civil Servants: Cross-Sectional Survey

Florence Carrouel[1*], Prof Dr; Virginie-Eve Lvovschi[2,3*], MD, PhD; Benjamin du Sartz de Vigneulles[1], PhD; Maryem Rhanoui[1], PhD; Roger Salamon[4], Prof MD; Michel Lamure[1], PhD; Corélie Salque[1], MS; Romain Lan[1,5], Prof Dr; Claude Dussart[1,3], Prof Dr

[1]Laboratory "Health, Systemic, Process" (P2S), UR4129, University Claude Bernard Lyon 1, 11 rue Guillaume Paradin, Lyon, France

[2]Research on Healthcare Performance Laboratory (INSERM U1290), University Claude Bernard Lyon 1, Lyon, France

[3]Hospices Civils de Lyon, Lyon, France

[4]Institute of Public Health, Epidemiology and Development (ISPED), Inserm U1219, University of Bordeaux, Bordeaux, France

[5]Laboratory ADES, Aix Marseille University, CNRS, EFS, Marseille, France

[*]these authors contributed equally

Corresponding Author:
Florence Carrouel, Prof Dr
Laboratory "Health, Systemic, Process" (P2S), UR4129, University Claude Bernard Lyon 1, 11 rue Guillaume Paradin, Lyon, France

Abstract

Background: Long COVID affects millions worldwide, straining health systems and workforce stability. This first nationwide survey among French civil servants combines epidemiological assessment with a Knowledge, Attitudes, and Behaviors approach. Long COVID remains a diagnostic and epidemiological challenge with evolving symptoms and uncertain categorization, particularly among self-suspected cases. Beyond prevalence and risk factors, understanding behavioral dimensions is essential to developing prevention strategies and maintaining workforce resilience.

Objective: This study aimed to (1) assess the prevalence of long COVID among French civil servants; (2) identify associated sociodemographic, occupational, and health-related factors; (3) assess disease-related knowledge of long COVID and (4) examine attitudes and behaviors regarding COVID-19 vaccination.

Methods: This cross-sectional survey was conducted in 2024 among active or retired civil servants in France. A Knowledge, Attitudes, and Behaviors–validated questionnaire, based on World Health Organization guidelines, was used. Responses were compared across 4 COVID-19 status groups (no COVID, COVID-19 without long COVID, diagnosed long COVID, and suspected long COVID). Statistical analyses included univariate tests and multivariable logistic regressions to identify factors associated with diagnosed or suspected long COVID.

Results: Among 3962 eligible respondents, 61 (1.54%; 95% CI 1.20 - 1.97) reported a formal diagnosis of long COVID and 241 (6.08%; 95% CI 5.38 - 6.87) without diagnosis. Diagnosed long COVID was significantly associated with long-term sick leave (odds ratio [OR] 1.15, 95% CI 1.03 - 6.28; *P*=.04) and long-term illness coverage (OR 0.72, 95% CI 0.27 - 0.92; *P*=.03). Suspected long COVID was associated with being in a relationship (OR 1.65, 95% CI 1.08 - 2.52; *P*=.02), widowed (OR 2.25, 95% CI 1.18 - 4.31; *P*=.01), and uncertain (OR 1.90, 95% CI 1.32 - 2.74; *P*<.001) or incomplete COVID-19 vaccination status (OR 1.67, 95% CI 1.16 - 2.42; *P*=.01). Knowledge scores differed significantly across groups (ANOVA $F_{3,3476}$=24.31, *P*<.001; χ^2_6=54.92, *P*<.001), with diagnosed cases showing the highest proportion of high knowledge (13/61, 21%) compared to 12.4% in the non-COVID group. Among 61 diagnosed cases, 36 (59%; 95% CI 46.4 - 70.5) were vaccinated, 13 (21%; 95% CI 12.9 - 33.2) intended to get vaccinated, and 12 (20%; 95% CI 11.6 - 31.3) remained unvaccinated; among suspected cases, these proportions were 173 (71.8%; 95% CI 65.9 - 77.1), 30 (12.4%; 95% CI 8.8 - 17.3), and 38 (15.8%; 95% CI 11.6 - 21.0), respectively.

Conclusions: Unlike previous studies that examined the clinical or behavioral factors separately, this nationwide analysis linked epidemiological data with knowledge and vaccination behaviors. Among French civil servants, long COVID remains underdiagnosed, where absenteeism and sick leave threaten essential services. The study highlights disparities in disease-related knowledge, vaccination attitudes, and behaviors, underlining the importance of workplace health education and systematic screening. Vaccination is associated with lower odds of long COVID, reinforcing its preventive value. Thus, findings reveal organizational implications and support workplace-based prevention strategies integrating vaccination promotion, early detection, and health literacy to sustain the resilience of public services.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e83323) doi:10.2196/83323

KEYWORDS

long COVID; health literacy; Knowledge, Attitudes, and Behaviors; KAB; COVID-19 vaccination; public health

Introduction

Long COVID, or post-COVID-19 syndrome, is estimated to affect approximately 65 million people worldwide, underscoring its substantial and growing burden on global public health [1]. The World Health Organization (WHO) defined it in 2021 as a condition affecting people with a history of probable or confirmed SARS-CoV-2 infection, developing at least 3 months after the onset of COVID-19, with symptoms persisting for at least 2 months and not explained by another diagnosis [2,3]. This definition remains widely used in recent clinical and epidemiological research and has been complemented by reviews describing long COVID as a systemic, multiorgan disease involving viral persistence, immune dysregulation, endothelial dysfunction, and microthrombosis [1,4,5].

Epidemiological estimates of the proportion of people suffering from long COVID vary widely. A 2025 meta-analysis of 429 studies estimated a pooled global prevalence of around 36% among COVID-positive individuals, with higher estimates at longer follow-up [6]. In contrast, the Organization for Economic Co-operation and Development (OECD) survey data from primary care (Patient-Reported Indicator Surveys program) suggested that approximately 7% of adults across participating countries reported persistent symptoms consistent with long COVID [7], while national surveillance by the US Centers for Disease Control and Prevention continues to track variable prevalence across demographics and states [8]. In France, a large-scale population study estimated that approximately 4% of adults present manifestations compatible with long COVID, depending on the case definition applied [9]. Despite these differences, the burden is indisputable, and prevalence estimates vary because of inconsistent definitions (clinical diagnosis vs self-reported symptoms), thresholds for symptom duration, study populations, and data collection methods.

Symptom profiles are dominated by chronic fatigue, dyspnea, musculoskeletal pain, and cognitive complaints such as “brain fog” [1,10]. These symptoms often impair quality of life, social functioning, and the ability to work. Beyond physical manifestations, long COVID is strongly associated with psychological sequelae. Meta-analyses and OECD reports published in 2023 - 2024 indicate increased risks of anxiety, depression, and posttraumatic stress disorder in affected patients, as well as significant losses in quality-adjusted life years across OECD countries [1,10,11]. This corroborates earlier findings in European cohorts, where long COVID was associated with markedly higher odds of depression and anxiety compared to unaffected individuals [12].

The impact on work capacity and professional life represents another major challenge. Long COVID patients report a reduced ability to work (0.62 points; 0.30 - 0.95) than those without Long COVID [13]. While 60.9% of patients with long COVID return to work, a significant number have to modify their tasks or work schedules [14]. In addition, patients with long COVID have difficulty maintaining a normal social and family life due to fatigue and cognitive impairment [15].

Beyond its health-related consequences, long COVID also exerts a significant societal and economic burden, particularly by reducing workforce availability and increasing public spending [16]. Estimates indicate that the direct and indirect costs of long COVID range from US $864 billion to US $1.04 trillion in OECD countries [11]. It also leads to higher health care use, with treatment expenses reaching up to $9000 per patient annually [16]. Its impact on work dynamics is particularly worrying. One survey found that 44% of people with long COVID were out of the labor force, while 51% worked fewer hours. These workforce reductions exacerbate existing labor shortages and contribute to inflation [16]. The extent of workforce disorganization [17] represents a major challenge not only for the private sector, but also for public services. In France, where the public service accounts for a substantial share of national employment, these disruptions may impact the continuity and efficiency of essential services. Moreover, the public sector offers a particularly relevant context for studying the impact of long COVID on employment, as it includes a wide range of socioprofessional categories and age groups, making it broadly representative of the working population.

Vaccination continues to play a dual role in this context. Systematic reviews and meta-analyses published between 2023 and 2025 consistently demonstrate that vaccination prior to infection is associated with a lower risk of developing long COVID [18-20]. Some studies also suggest that vaccination after infection may reduce the risk or severity of persistent symptoms [19]. Attitudes towards vaccination among people with long COVID remain complex. While some patients report hesitancy or heightened concerns, others show increased acceptance, especially when informed about vaccination as a preventive measure against long-term complications. Recent evidence also highlights that vicarious experience—knowing someone affected by long COVID—is associated with higher perceived risk of COVID-19 and more favorable beliefs about vaccine efficacy [21,22].

To date, most studies have investigated either prevalence and risk factors or knowledge and attitudes separately. However, combining these dimensions in a single survey provides a more comprehensive perspective: epidemiological estimates quantify the burden of long COVID, while knowledge and attitudes provide insight into the behavioral and organizational responses that influence prevention, early detection, and workforce resilience. This integrated approach is particularly relevant for civil servants, where both individual health and the continuity of essential public services are at stake.

Thus, this study aims to (1) assess the prevalence of long COVID among French civil servants; (2) identify associated sociodemographic, occupational, and health-related factors; (3)

XSL•FO
RenderX

assess disease-related knowledge of long COVID; and (4) examine attitudes and behaviors regarding COVID-19 vaccination.

Methods

Study Design

This study was designed as a cross-sectional, anonymous questionnaire-based survey conducted between January 15, 2024, and February 12, 2024. It followed the CROSS (Checklist for Reporting of Survey Studies; Checklist 1) [23].

The study was supported by *Union Prévention Santé pour la Fonction publique* (UROPS), an entity operating under the supervision of public authorities, bringing together 11 mutual insurance organizations and overseeing the health coverage of approximately 1.7 million beneficiaries under the Morice Law of 1947 [24]. UROPS promotes preventive health care through the Compulsory Health Insurance Scheme. Regularly, it conducts the nationwide "Baromètre Santé" multisection survey to assess beneficiaries' Knowledge, Attitudes, and Behaviors (KAB) on health issues, guiding targeted prevention programs. Findings from previous editions were published [25]. This study analyzes the long COVID section introduced in the 2024 edition.

Study Population

The study population met the following inclusion criteria: (1) individuals aged ≥18 years, (2) active or retired civil servants, (3) affiliated with the UROPS insurance union for public sector workers, (4) registered on the public insurance platform [26] and having consented to be contacted, and (5) fluent in French; (6) willing to complete a questionnaire. Individuals were excluded if they were third-party beneficiaries of UROPS members or held nonpermanent employment status.

Sample Size

Based on 95% CI, a margin of error of 5%, and an assumed response rate of 50%, a conservative estimate for unknown proportions, the minimum required sample size was calculated to be at least 384 participants. This calculation was made using the standard formula for sample size estimation in survey studies [27], considering the total number of public service persons listed by the UROPS (836,307 mailing or email addresses in the database).

Sample Method

The population was recruited based on a convenience sampling methodology. Participants were recruited via a mail or email invitation sent by UROPS to its policyholders. UROPS covers approximately one-third of the State civil service workforce in France [24,28], including diverse occupational and regional categories. This framework provides access to a large and heterogeneous, although nonprobabilistic, sample of civil servants representative of multiple administrative branches.

Development and Validation of the KAB Questionnaire

Questionnaire Design and Validation

A structured French questionnaire was developed to evaluate disease-related knowledge of long COVID and attitudes and behaviors regarding COVID-19 vaccination (Multimedia Appendix 1), following WHO methodological recommendations [29]. Two public health researchers conducted a literature review to identify existing KAB tools, from which items were selected, adapted, or newly formulated. The initial version included 39 closed-ended questions.

Validation followed the approach of Andrade et al [30], combining face and content validation.

In the first validation phase, a multidisciplinary panel of experts (2 general practitioners, 1 microbiologist, and 1 public health specialist) evaluated each item for clarity and relevance. Using a standardized form, experts rated each item as 2=satisfactory or 1=unsatisfactory. The Item-level Content Validity Index (I-CVI) was calculated as the proportion of experts rating the item as "relevant." Following Polit and Beck [31], a minimum acceptable threshold of 0.75 (ie, at least 3 out of 4 experts in agreement) was applied: items with an I-CVI of ≥0.75 were retained, while those below this threshold were revised or removed. The Scale-level Content Validity Index (S-CVI/Ave) was then calculated as the average of the I-CVI across all items. The updated version was re-evaluated and approved by the expert panel. The second phase involved cognitive debriefing with 6 individuals not included in the main study. They reviewed the questionnaire and provided feedback on clarity and comprehension, leading to minor wording refinements.

The expert committee then validated this final version, which comprised 36 closed-ended questions: 12 questions on sociodemographic characteristics (Q1-Q12), 5 questions on health status (Q13-Q17), 10 questions to assess whether participants had been diagnosed with or likely experienced long COVID (Q18-Q27), 7 questions on long COVID knowledge (Q28-Q34), and 2 questions on attitudes and behavior regarding COVID-19 vaccination (Q35 and Q36), which explored participants' intentions to vaccinate and their motivations, providing information on at least 7 specific behaviors or attitudes via Q36.

Construction of the Knowledge Score and Analysis of Attitude and Behavior Answers

A scoring system was developed to assess respondents' knowledge of long COVID. From Q28 to Q34 (long COVID knowledge), a total of 9 correct answers distributed across the relevant items. Correct and incorrect answers were determined based on information published by the French High Health Authority (*Haute Autorité de la Santé* [HAS]) [32]. A binary scoring method was applied, assigning +1 point for each correct answer and 0 for incorrect ones. The resulting knowledge score ranged from 0 to 9 (9 items) and was categorized as follows: 0-3 (low), 4-6 (medium), and 7-9 (high).

For attitudes and behaviors related to COVID-19 vaccination, a nonscoring analysis was conducted.

Survey Administration

Participants could complete the paper questionnaire or the online questionnaire. For the paper questionnaire, respondents had to return it using a prepaid envelope. For the online questionnaire, the "Confirmit" software (SurveySparrow) with anonymized

XSL•FO
RenderX

responses was used, and to prevent multiple submissions, people contacted by email had an individual weblink that could only be used once. Participants could also pause and resume the questionnaire if needed. After validation, the web link was no longer accessible.

Definition of COVID-19 Status Groups

Long COVID was defined according to the HAS recommendations as the persistence of one or more symptoms beyond 4 weeks after acute infection, without another diagnosis explaining them [33]. This operational threshold is shorter than the WHO definition, which requires symptoms to persist for at least 12 weeks [34]. Therefore, caution is needed when comparing our results with international studies based on the WHO definition.

Participants were categorized into four mutually exclusive groups based on several answers to questions (Multimedia Appendix 2):

1. No COVID: never infected with SARS-CoV-2.
2. COVID-19 without long COVID: previous infection without persistent symptoms compatible with long COVID.
3. Diagnosed long COVID: formal medical diagnosis of long COVID following confirmed SARS-CoV-2 infection.
4. Suspected long COVID: persistent symptoms consistent with long COVID but without a formal medical diagnosis.

This approach distinguishes between medically confirmed cases and self-reported symptom persistence, but it also carries a risk of misclassification: undiagnosed individuals may underestimate their condition, while some self-reported cases could include symptoms attributable to other causes.

Statistical Analysis

Statistical analyses were performed using Python (version 3.10; Python Software Foundation) with *pandas*, *numpy*, and *statsmodels* libraries. Descriptive statistics summarized variables including gender, age, marital and parental status, region, employment, statutory category, education, chronic conditions, and administration type. Frequencies and percentages were used for categorical variables; medians and IQR for continuous and ordinal variables. Univariate analyses compared the 4 COVID-19 groups (no infection, infection without long COVID, suspected long COVID, and diagnosed long COVID) using chi-square test for categorical variables and Kruskal-Wallis or Mann-Whitney *U* tests for ordinal or continuous variables. The Spearman correlation assessed ordinal associations.

Multivariable logistic regressions identified factors associated with diagnosed or suspected long COVID. Covariates for multivariable logistic models were selected a priori based on clinical and epidemiological plausibility (age, gender, education, chronic disease, long-term illnesses coverage, vaccination status, and knowledge score K) and were not removed solely on the basis of univariate significance. All variables corresponded directly to questionnaire items, and no post hoc transformations or recoding were performed. Variables with $P<.1$ in univariate screening or strong subject-matter justification were retained. Prespecified interactions (chronic condition [Q13] × vaccination status [Q35], age [Q2] × gender [Q1], education [Q8] × score K, and group × score K) were assessed via likelihood-ratio tests; none met retention criteria, and final results are reported for main effects models. Results were expressed as odds ratios (OR) with 95% CI, significance set at $P<.05$.

No missing data were present in the final analytical dataset: the questionnaires retained for analysis were fully completed. Consequently, no imputation or missing data mechanism analyses (eg, missing completely at random tests) were required. As no missing data or imputed variables were involved, no sensitivity analyses were necessary. Similarly, no weighting procedures were applied, as analyses were conducted on the complete dataset of eligible respondents. No specific analysis of nonresponse bias was performed, as only aggregate demographic data of the target population were available.

The exposures included sociodemographic, occupational, and health-related characteristics, as well as COVID-19 vaccination status, while the outcomes were diagnosed long COVID and suspected long COVID. KAB responses were compared across groups: knowledge was analyzed with Kruskal-Wallis tests, and attitudes and behaviors were described but not modeled.

Ethical Considerations

UROPS regularly conducts the *Baromètre Santé* survey [25]. The research team designed a specific KAB module on long COVID, which was incorporated into the 2024 edition. As the survey generated fully anonymized data, no prior approval was required for its administration. Nevertheless, the research protocol was submitted to the Aix Marseille University ethics committee, which granted approval on July 17, 2025 (n° 2025-07-17-02), in accordance with national procedures for the secondary use of anonymized survey data. This study followed the Declaration of Helsinki and the reference methodology MR-004 of the French National Commission on Information Technology and Liberties. All participants provided informed consent prior to completing the questionnaire. Consent was obtained electronically for online respondents and in written form for postal respondents. Participation was voluntary, responses were anonymous, and data were used exclusively for research purposes. To ensure confidentiality, responses were completely anonymous at the time of collection and no personal identifiers were recorded. For the online questionnaire, access was provided via secure individual links that could only be used once, and returned paper questionnaires contained no identifying information. Participants did not receive any financial or nonfinancial compensation for their participation. No images or supplementary materials that could allow the identification of individual participants are included in this paper.

Results

Psychometric Properties of the Questionnaire

Content validity was assessed in two rounds. In the first round, 39 items were evaluated by 4 experts. I-CVI values ranged from 0.50 to 1.00, and the S-CVI/Ave was 0.87. Three items with I-CVI <0.75 were eliminated at this stage. In the second round, the revised questionnaire (36 items) was re-evaluated by the same panel. Final indices indicated strong agreement, with

XSL•FO
RenderX

I-CVI values ranging from 0.75 to 1.00 and an S-CVI/Ave of 0.92, confirming satisfactory content validity.

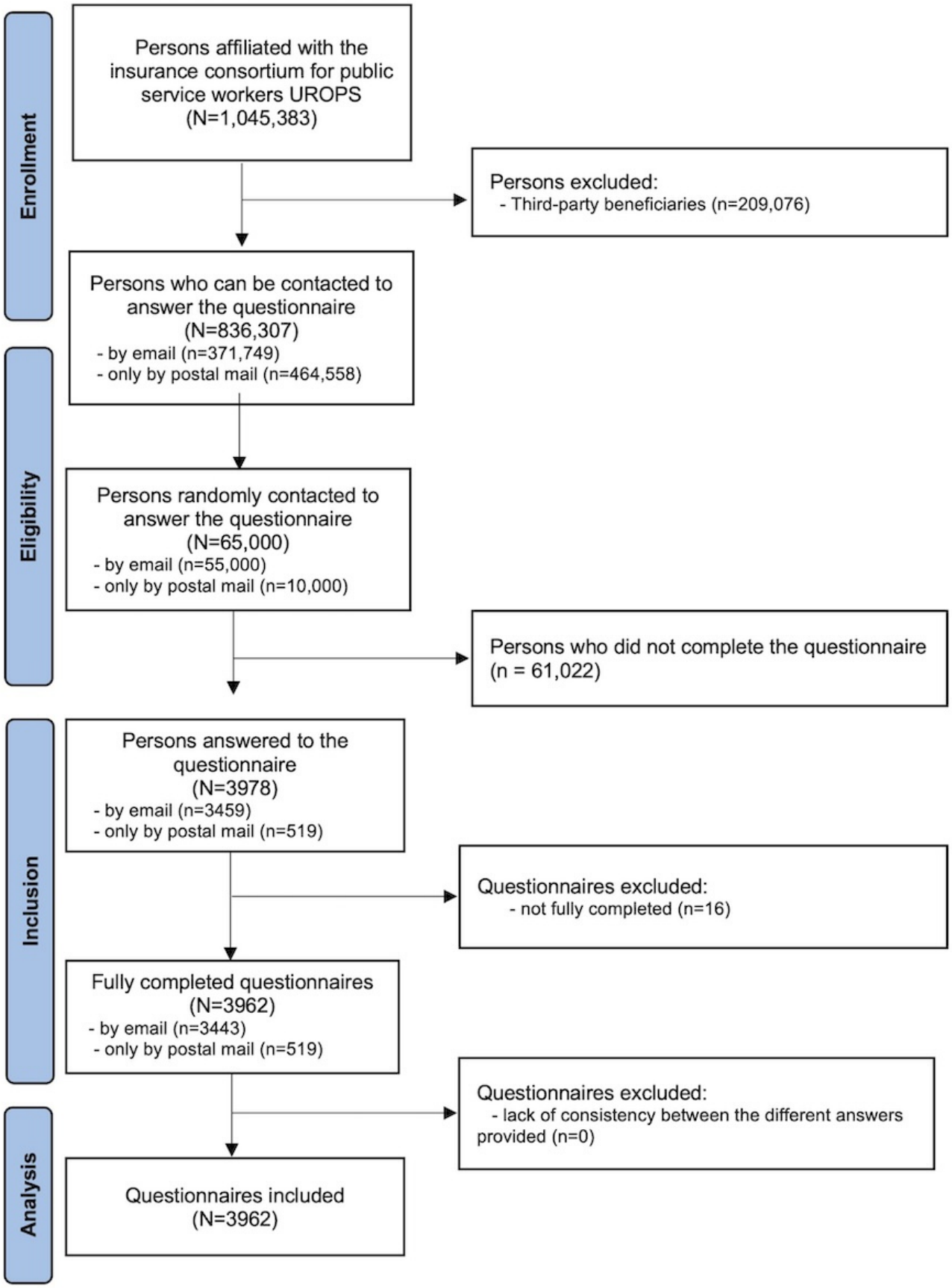
Sociodemographic and Health Characteristics of Respondents

Among the 1,045,383 individuals affiliated with UROPS in 2024, 209,076 were third-party beneficiaries (dependents of primary insured members) who could not be contacted directly because their consent for communication was not authorized. Among the remaining 836,307 eligible policyholders, 65,000 were invited to participate. 3978 (6.12%) insured participated in this survey. 3962 questionnaires were considered eligible, with all questions answered, whereas 16 were excluded because participants dropped out during the survey (Figure 1). No information was available on reasons for nonparticipation, as data collection was fully anonymized and participation was voluntary.

Among the 3962 respondents, 1091 (27.54%) reported never having had a COVID-19 infection, 2087 (52.68%) indicated a previous infection without subsequent long COVID symptoms, 61 (1.54%) had a confirmed diagnosis of long COVID, and 241 (6.08%) reported persistent symptoms consistent with long COVID despite not having received a formal medical diagnosis.

XSL•FO
RenderX

Figure 1. Flowchart of the study. UROPS: Union Prévention Santé Pour La Fonction Publique.



Sociodemographic, health, and occupational characteristics were described in Table 1 and Multimedia Appendix 3. Compared with national statistics on the French civil service [28], the sample displayed a broadly similar gender and age distribution. However, category B agents were slightly overrepresented (1466/3962, 37.00% vs 616,800/2,570,000,

XSL•FO
RenderX

24.00%) and category A agents underrepresented (1652/3962, 41.70% vs 1,440,000/2,570,000, 56.00%), suggesting moderate structural differences but overall demographic consistency.

XSL•FO
RenderX

Table . Sociodemographic and health characteristics of the respondents.

Variable	All, n (%; 95% CI)	No COVID-19[a], n (%; 95% CI)	COVID-19 without long COVID[b], n (%; 95% CI)	Diagnosed long COVID[c], n (%; 95% CI)	Suspected long COVID[d], n (%; 95% CI)
Gender					
Male	1823 (46.01; 44.46 - 47.56)	659 (47.8 ;45.2 - 50.5)	959 (46.0; 43.8 - 48.1)	22 (36; 24 - 48)	96 (39.8; 33.7 - 46.0)
Female	2139 (53.99; 52.44 - 55.54)	719 (52.2; 49.5 - 54.8)	1128 (54.0; 51.9 - 56.2)	39 (64; 52 - 76)	145 (60.2; 54.0 - 66.3
Age (y)					
18 - 29	27 (0.68; 0.43 - 0.94)	7 (0.5; 0.1 - 0.9)	14 (0.7; 0.3 - 1.0)	0 (0; 0 - 0)	3 (1.2; 0.0 - 2.6)
30 - 39	179 (4.52; 3.87 - 5.16)	40 (2.9; 2.0 - 3.8)	106 (5.1; 4.1 - 6.0)	7 (12; 4 - 20)	14 (5.8; 2.9 - 8.8)
40 - 49	512 (12.92; 11.88 - 13.97)	111 (8.1; 6.6 - 9.5)	333 (16.0; 14.4 - 17.5)	13 (21; 11 - 32)	27 (11.2; 7.2 - 15.2)
50 - 59	875 (22.08; 20.79 - 23.38)	252 (18.3; 16.2 - 20.3)	486 (23.3; 21.5 - 25.1)	17 (28; 17 - 39)	77 (32.0; 26.1 - 37.8)
60 - 69	1250 (31.55; 30.10 - 33.00)	458 (33.2; 30.7 - 35.7)	657 (31.5; 29.5 - 33.5)	16 (26; 15 - 37)	69 (28.6; 22.9 - 34.3)
70 - 79	974 (24.58; 23.24 - 25.92)	437 (31.7; 29.3 - 34.2)	436 (20.9; 19.1 - 22.6)	7 (12; 4 - 20)	45 (18.7; 13.8 - 23.6)
≥80	145 (3.66; 3.08 - 4.24)	73 (5.3; 4.1 - 6.5)	55 (2.6; 1.9 - 3.3)	1 (2; 0 - 5)	6 (2.5; 0.5 - 4.5)
Marital status					
Single	569 (14.36; 13.27 - 15.45)	232 (16.8; 14.9 - 18.8)	273 (13.1; 11.6 - 14.5)	16 (26; 15 - 37)	28 (11.6; 7.6 - 15.7)
Couple	2646 (66.78; 65.32 - 68.25)	842 (61.1; 58.5 - 63.7)	1477 (70.8; 68.8 - 72.7)	31 (51; 38 - 63)	168 (69.7; 63.9 - 75.5)
Divorced	495 (12.49; 11.46 - 13.52)	181 (13.1; 11.4 - 14.9)	243 (11.6; 10.3 - 13.0)	12 (20; 10 - 30)	26 (10.8; 6.9 - 14.7)
Widowed	252 (6.36; 5.60 - 7.12)	123 (8.9; 7.4 - 10.4)	94 (4.5; 3.6 - 5.4)	2 (3; 0 - 8)	19 (7.9; 4.5 - 11.3)
Current employment status					
Working	1724 (43.51; 41.97 - 45.06)	452 (32.8; 30.3 - 35.3)	1020 (48.9; 46.7 - 51.0)	36 (59; 47 - 71)	127 (52.7; 46.4 - 59.0)
Parental leave	4 (0.10; 0.00 - 0.20)	0 (0; 0.0 - 0.0)	4 (0.2; 0.0 - 0.4)	0 (0; 0 - 0)	0 (0.0; 0.0 - 0.0)
In school, training, and internship	11 (0.28; 0.11 - 0.44)	6 (0.4; 0.1 - 0.8)	4 (0.2; 0.0 - 0.4)	0 (0; 0 - 0)	1 (0.4; 0.0 - 1.2)
Retired	2118 (53.46; 51.90 - 55.01)	883 (64.1; 61.5 - 66.6)	1014 (48.6; 46.4 - 50.7)	18 (30; 18 - 41)	107 (44.4; 38.1 - 50.7)
On disability	16 (0.40; 0.21 - 0.60)	10 (0.7; 0.3 - 1.2)	4 (0.2; 0.0 - 0.4)	0 (0; 0 - 0)	1 (0.4; 0.0 - 1.2)
On sick leave for >3 months	89 (2.25; 1.78 - 2.71)	27 (2; 1.2 - 2.7)	41 (2; 1.4 - 2.6)	7 (12; 4 - 20)	5 (2.1; 0.3 - 3.9)
Last statutory category					
Category A+	661 (16.68; 15.52 - 17.84)	221 (16; 14.1 - 18.0)	390 (18.7; 17.0 - 20.4)	5 (8; 1 - 15)	31 (12.9; 8.6 - 17.1)
Category A	991 (25.01; 23.66 - 26.36)	318 (23.1; 20.9 - 25.3)	575 (27.6; 25.6 - 29.5)	16 (26; 15 - 37)	55 (22.8; 17.5 - 28.1)
Category B	1466 (37.00; 35.50 - 38.50)	511 (37.1; 34.5 - 39.6)	754 (36.1; 34.1 - 38.2)	26 (43; 30 - 55)	95 (39.4; 33.2 - 45.6)
Category C	685 (17.29; 16.11 - 18.47)	263 (19.1; 17.0 - 21.2)	297 (14.2; 12.7 - 15.7)	13 (21; 11 - 32)	47 (19.5; 14.5 - 24.5)
Contract	21 (0.53; 0.30 - 0.76)	7 (0.5; 0.1 - 0.9)	10 (0.5; 0.2 - 0.8)	0 (0; 0 - 0)	3 (1.2; 0.0 - 2.6)
Other	138 (3.48; 2.91 - 4.05)	58 (4.2; 3.1 - 5.3)	61 (2.9; 2.2 - 3.6)	1 (2; 0 - 5)	10 (4.1; 1.6 - 6.7)

XSL•FO
RenderX

Variable	All, n (%; 95% CI)	No COVID-19[a], n (%; 95% CI)	COVID-19 without long COVID[b], n (%; 95% CI)	Diagnosed long COVID[c], n (%; 95% CI)	Suspected long COVID[d], n (%; 95% CI)
Level of study					
Level I[e]	994 (25.09; 23.74 - 26.44)	401 (29.1; 26.7 - 31.5)	435 (20.8; 19.1 - 22.6)	15 (25; 14 - 35)	68 (28.2; 22.5 - 33.9)
Level II[f]	1596 (40.28; 38.76 - 41.81)	532 (38.6; 36.0 - 41.2)	855 (41; 38.9 - 43.1)	31 (51; 38 - 63)	94 (39.0; 32.8 - 45.2)
Level III[g]	1372 (34.63; 33.15 - 36.11)	445 (32.3; 29.8 - 34.8)	797 (38.2; 36.1 - 40.3)	15 (25; 14 - 35)	79 (32.8; 26.9 - 38.7)
Chronic disease or disability or health problem					
Yes	1840 (46.44; 44.89 - 47.99)	637 (46.2; 43.6 - 48.9)	958 (45.9; 43.8 - 48.0)	39 (64; 52 - 76)	107 (44.4; 38.1 - 50.7)
No	1976 (49.87; 48.32 - 51.43)	690 (50.1; 47.4 - 52.7)	1069 (51.2; 49.1 - 53.4)	19 (31; 20 - 43)	120 (49.8; 43.5 - 56.1)
Do not want to answer	0 (0.00; 0.00 - 0.00)	0 (0; 0.0 - 0.0)	0 (0; 0.0 - 0.0)	0 (0; 0 - 0)	0 (0.0; 0.0 - 0.0)
Long-term illnesses coverage[h]					
Yes	1175 (29.66; 28.23 - 31.08)	461 (33.5; 31.0 - 35.9)	553 (26.5; 24.6 - 28.4)	28 (46; 33 - 58)	63 (26.1; 20.6 - 31.7)
No	2706 (68.30; 66.85 - 69.75)	892 (64.7; 62.2 - 67.3)	1493 (71.5; 69.6 - 73.5)	31 (51; 38 - 63)	175 (72.6; 67.0 - 78.2)
Do not know	0 (0.00; 0.00 - 0.00)	0 (0; 0.0 - 0.0)	0 (0; 0.0 - 0.0)	0 (0; 0 - 0)	0 (0.0; 0.0 - 0.0)

[a]Participants who reported never having had a SARS-CoV-2 infection.

[b]Participants who reported a previous SARS-CoV-2 infection but no persistent symptoms compatible with long COVID.

[c]Participants who reported a formal medical diagnosis of long COVID following a confirmed SARS-CoV-2 infection.

[d]Participants who reported at least one persistent symptom consistent with long COVID beyond four weeks after the acute phase of infection but without a formal medical diagnosis.

[e]Level of study < “baccalaureate”.

[f]Level of study ≥ “baccalaureate” degree and ≤ “baccalaureate” degree +2 years.

[g]Level of study > “baccalaureate” degree >2 years.

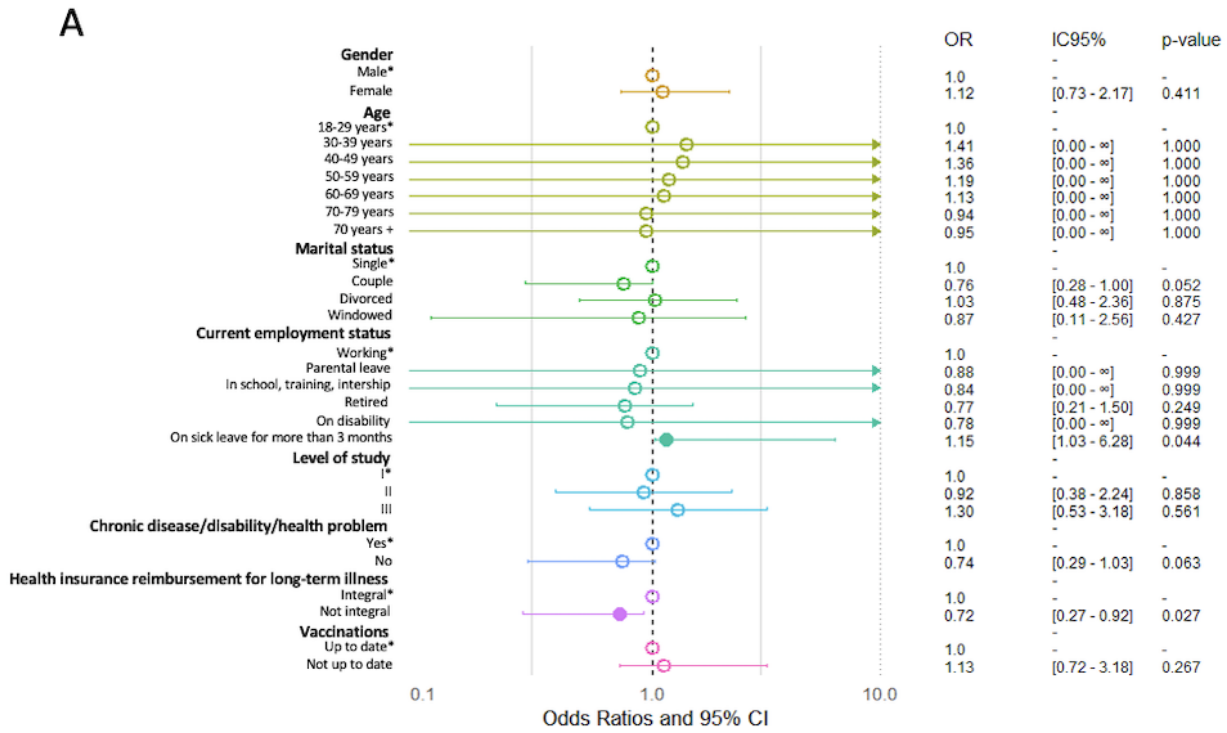
[h]Reimbursed at 100% by the health insurance for a long-term illness.

Determinants of Diagnosed or Suspected Long COVID Status in French Civil Servants

The results of the binary logistic regressions performed to evaluate the determinants diagnosed long COVID and suspected long COVID are presented in Figure 2.

XSL•FO
RenderX

Figure 2. Binary logistic regression performed to evaluate the determinants of (A) diagnosed or (B) suspected long COVID. Odds ratios are presented as points, with horizontal error bars representing 95% CIs. Level of study I: below baccalaureate degree; level of study II: greater than or equal to baccalaureate degree and less than or equal to baccalaureate degree plus 2 years; and level of study III: greater than or equal to baccalaureate degree plus 2 years.



In the multivariable models (Figure 2), diagnosed long COVID was significantly associated with employment-related variables. Participants on sick leave for >3 months had higher odds of receiving a diagnosis (OR 1.15, 1.03 - 6.28; *P*=.04), while those without full long-term illness reimbursement had lower odds (OR 0.72, 95% CI 0.27 - 0.92; *P*=.03). No associations were found with age, gender, education, or chronic conditions. For suspected long COVID, marital status and vaccination status remained significant. Compared to single individuals, those who were in a relationship (OR 1.65, 95% CI 1.08 - 2.52; *P*=.02) and widowed participants (OR 2.25, 95% CI 1.18 - 4.31; *P*=.01) had higher odds of reporting symptoms. Furthermore, being not up to date with COVID-19 vaccination (OR 1.67, 95% CI 1.16 - 2.42; *P*=.01) or uncertain about vaccination status

XSL•FO
RenderX

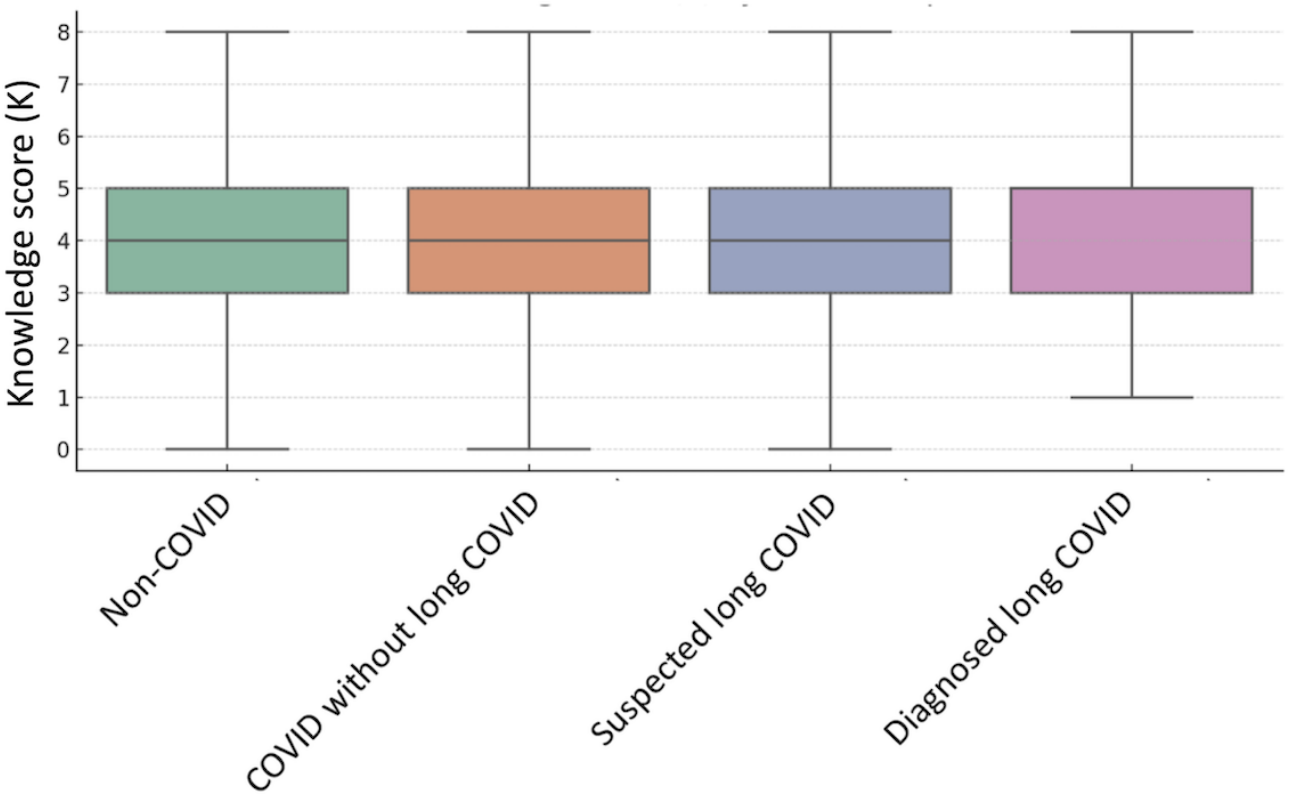
(OR 1.90, 95% CI 1.32 - 2.74; *P*<.001) was also significantly associated with suspected long COVID. Other covariates, including age, gender, occupational status, and long-term illness coverage, were not significant.

Analysis of KAB Regarding Long COVID Disease and Vaccination

Knowledge Score Related to Long COVID Among French Civil Servants

Knowledge scores varied according to participants' COVID-19 status (Figure 3). The diagnosed long COVID group reported the highest median score of 5.0 (IQR 3.0-5.0), followed by the suspected long COVID, COVID without long COVID, and non-COVID groups (all with a median 4.0, IQR 3.0-5.0 score). Despite similar medians across most groups, the distribution of scores differed.

Figure 3. Knowledge score regarding long COVID among French civil servants.



The Kruskal-Wallis test indicated a statistically significant difference in score distribution between groups (H=82.55; *P*<.001), suggesting that knowledge levels are associated with prior COVID-19 exposure and long COVID diagnosis. The diagnosed long COVID group had the highest proportion of high knowledge scores (21.3%; 95% CI 12.9 - 33.1), while the non-COVID groups showed a higher rate of low scores (23.6%; 95%CI 21.1 - 26.2; Multimedia Appendix 4). Post hoc chi-square pairwise comparisons showed that individuals in the non-COVID group had significantly lower knowledge scores compared to all other groups, including those with diagnosed long COVID (mean difference –0.72; *P*<.001), suspected long COVID (mean difference –0.35; *P*<.001), and COVID without long COVID (mean difference –0.47; *P*<.001).

Attitudes and Behavior Regarding Vaccination According to Long COVID Status

Attitudes and behaviors regarding vaccination diagnosed or suspected population with long COVID are presented in Table 2. A total of 13 (21%; 95%CI 13 - 33) respondents who intended to vaccinate against COVID-19 had been diagnosed with long COVID, while 30 (12.4%; 95% CI 8.8 - 17.3) reported suspected long COVID symptoms. In contrast, a significantly higher proportion of respondents who declared they were already vaccinated reported a history of long COVID, with 36 (59.0%; 95% CI 46.4 - 70.5) diagnosed and 173 (71.8%; 95% CI 65.9 - 77.1) suspected cases. Finally, among those who did not intend to get vaccinated and were not yet vaccinated, 12 (20%; 95% CI 12 - 31) had diagnosed long COVID and 38 (15.8%; 95% CI 11.6 - 21.0) had suspected long COVID.

XSL•FO
RenderX

Table . Analysis of attitudes and behavior related to COVID-19 vaccination among diagnosed (n=61) and suspected (n=241) population with long COVID.

	Diagnosed long COVID, n (%; 95% CI)	Suspected long COVID, n (%; 95% CI)
Do you intend to vaccinate against COVID-19?		
Yes	13 (21; 13 - 33)	30 (12.4; 8.8 - 17.3)
No, I'm already vaccinated	36 (59; 46 - 70)	173 (71.8; 65.9 - 77.1)
No, and I'm not vaccinated	12 (20; 12 - 31)	38 (15.8; 11.6 - 21.0)
For what reason(s) do you not intend to be vaccinated against COVID-19?		
Not effective in preventing me from getting sick	5 (19; 8 - 37)	8 (18.2; 9.5 - 32.1)
No confidence in vaccination, risk of side effects	4 (15; 6 - 33)	9 (21.6; 11.9 - 36.7)
No confidence in vaccine efficacy	4 (15; 6 - 33)	6 (14.8; 6.9 - 28.8)
I'm immune, already had COVID-19	2 (7; 2 - 23)	4 (9.1; 3.6 - 21.3)
It doesn't matter if I'm infected by this virus, the disease is not serious	1 (4; 1 - 18)	3 (6.8; 2.4 - 18.7)
Prefer other means of prevention, such as barrier gestures, wearing a mask or physical distancing	6 (22; 11 - 42)	6 (13.6; 5.9 - 28.0)
I don't like shots	1 (4; 1 - 18)	1 (2.3; 0.4 - 11.8)
Against vaccination in general	0 (0; 0-0)	1 (2.3; 0.4 - 11.8)
Too complicated to get vaccinated	0 (0; 0-0)	1 (2.3; 0.4 - 11.8)
There is currently little or no COVID-19	0 (0; 0-0)	3 (6.8; 2.4 - 18.7)
Other reasons	4 (15; 6 - 33)	2 (4.6;1.3 - 15.3)
For what reason(s) do you intend to be vaccinated against COVID-19?		
The COVID-19 vaccine has already protected me from a more serious form.	6 (16; 7 - 33)	10 (12.5; 6.9 - 21.6)
I need to protect myself from new variants	5 (14; 6 - 28)	16 (20; 12.7 - 30.1)
I need to boost my immunity	7 (19; 9 - 35)	14 (17.5; 10.7 - 27.4)
I need to protect my loved ones	9 (24; 14 - 40)	17 (21.3; 13.7 - 31.8)
I must follow the advice of my doctor/pharmacist.	7 (19; 9 - 35)	13 (16.3; 9.7 - 26.1)
I follow the recommendations of health authorities	2 (5; 2 - 18)	9 (11.3; 6.0 - 20.3)
Other reason	1 (3; 1 - 14)	1 (1.3; 0.2 - 6.9)

Discussion

Principal Results and Comparison With Prior Work

To our knowledge, this study is the first to analyze the prevalence and characteristics of long COVID among French civil servants. It explores the associations between long COVID status and sociodemographic, occupational, and behavioral factors, along with long COVID knowledge, attitudes, and behaviors related to COVID-19 vaccination.

This study differs from prior research by combining epidemiological analysis with a validated KAB approach in a large population of French civil servants—a group critical to maintaining public service continuity, but previously overlooked in long COVID research. This dual perspective highlights not only medical determinants but also behavioral and organizational dimensions essential for prevention and resilience strategies.

The findings highlight significant diagnostic disparities, knowledge gaps, attitudes, and behavioral trends, while also suggesting that long COVID may interfere with functioning and professional stability, even though the quality of life was not directly assessed.

Among 3962 respondents, 2087 (52.68%) indicated a COVID-19 infection without subsequent long COVID symptoms, 61 (1.54%) had a confirmed diagnosis of long COVID, and 241 (6.08%) reported persistent symptoms consistent with long COVID despite not having received a formal medical diagnosis. These results are consistent with French estimates: of the 48% of individuals who had been infected with SARS-CoV-2 for >3 months, 8% met the criteria for post-COVID-19 condition [35], and fall within the ranges reported by the WHO (10% - 20%), the US Centers for Disease Control and Prevention (20% - 25%) [36], and international meta-analyses [37]. The high proportion

XSL•FO
RenderX

of suspected undiagnosed cases highlights a possible underdetection, probably linked to barriers in access to care, poor recognition of symptoms, or reluctance to seek medical advice [38]. This underlines the importance of improving referral and screening in primary care.

Long COVID has a functional impact with a decline in well-being. Although the quality of life was not assessed through standardized tools (eg, 36-Item Short Form Health Survey and EuroQol 5 Dimensions), several questionnaire items can be interpreted as both risk factors and consequences for daily functioning. Participants with diagnosed long COVID were more often on long-term sick leave (11.5% vs 2.2% in the general population and 2% in the population of no long COVID), reflecting either preexisting vulnerability or reduced work capacity due to long COVID. Chronic conditions were more frequent among long COVID cases (46%), suggesting that underlying health status could influence both susceptibility and functional outcomes. These findings are consistent with previous studies indicating that long COVID reduces work capacity by 0.62 points on average [13], with 60% of patients needing workstation or schedule adaptations [39]. Mental health repercussions are also reported, with a higher prevalence of depression (OR 2.35, 95% CI 1.49 - 3.70) and anxiety (OR 2.53, 95% CI 1.76 - 3.61) [12]. In our study, respondents with diagnosed long COVID reported more comorbidities (respiratory diseases, musculoskeletal disorders, migraines, and mental health disorders such as depression, anxiety, or sleep disorders), reinforcing the hypothesis of multidimensional consequences [40].

Although mental health outcomes were not directly assessed in this study, the high prevalence of reported comorbidities such as depression, anxiety, and sleep disorders underscores the psychological burden of long COVID, which is well documented in the literature. Meta-analyses have shown increased risks of anxiety, depression, and cognitive impairment among affected individuals [41-43]. These manifestations can reduce productivity and slow return-to-work trajectories. These findings highlight the need to integrate mental health assessment and support into occupational reintegration programs in accordance with WHO recommendations for post-COVID management [44].

This study also supports evidence that vaccination is associated with lower odds of long COVID. While diagnosed or suspected cases were more frequent among vaccinated participants at the time of the survey (n=36, 59%; 46.4 - 70.5 and n=173, 71.8%; 65.9 - 77.1, respectively), this likely reflects postinfection vaccination, greater medical engagement, or higher health awareness, rather than causality. Conversely, unvaccinated individuals or those uncertain of their vaccination status had higher odds ratios of suspected long COVID (OR 1.67, 95% CI 1.16 - 2.41; P=.006 and OR 1.90, 95% CI 1.32 - 2.74; P<.001 respectively), aligning with studies reporting reduced incidence of long COVID after full vaccination [45,46].

The study reveals disparities in knowledge and health attitudes and behavior regarding long COVID. Knowledge was significantly higher among individuals who had experienced COVID-19, particularly those with a confirmed diagnosis, supporting evidence that personal illness experience enhances health-related engagement and information seeking [47,48]. In contrast, individuals who had never contracted COVID-19 scored lower, likely due to lower exposure and reduced perceived relevance [49]. Although knowledge scores differed significantly across groups, the absolute differences were small. Similar results were found in recent knowledge, attitude, and practice studies on long COVID and vaccination. He et al [50] found that modest knowledge score differences were nonetheless associated with greater adherence to preventive behaviors and vaccine acceptance, suggesting that even small awareness gaps may have behavioral implications. Lazarus et al [51] also showed that limited variations in COVID-19 knowledge and attitudes influenced vaccine perceptions and willingness to vaccinate. These results suggest that improving health literacy and risk perception remains essential. Strengthening disease literacy among unaffected populations is therefore crucial to encourage early recognition, care-seeking, and risk awareness. Workplace health education campaigns could help close this gap and foster preventive behaviors [47].

Health care inequities and structural barriers were underlined. Individuals with formal long-term illness coverage were more likely to have a confirmed long COVID diagnosis, suggesting that care access influences diagnostic likelihood. This may reflect better access to medical specialists, administrative support, or continuity of care within chronic disease networks [52,53]. Conversely, many suspected cases may remain undiagnosed due to logistical, economic, or informational barriers [54,55].

Diagnosing long COVID remains challenging, with many cases potentially undetected due to health care inequities, making SARS-CoV-2 monitoring essential. Between May 26, 2025, and June 1, 2025, WHO recorded 60,216 SARS-CoV-2 tests from 80 countries, with 4.1% (2471 samples) positive [34]. While worldwide COVID-19 activity was low, localized increases appeared in many countries. These data underscore ongoing risks of post-COVID conditions including in France and the need for proactive workplace prevention strategies, particularly within the public sector, where long COVID can disrupt essential services. Recommended interventions include occupational health screening for post-COVID symptoms, tailored reintegration plans for affected employees, awareness campaigns, flexible work policies, and accommodations for cognitive or physical impairments and mental health support.

In France, implementing these interventions would align with existing recommendations from the HAS and the *Conseil scientifique* COVID-19 (state scientific council), which emphasize early identification and personalized return-to-work pathways for post-COVID employees [56,57]. Occupational health services within the public sector could play a central role in this process by reinforcing health surveillance, facilitating multidisciplinary coordination, and adapting workplace accommodations. The WHO also recommends integrating long COVID into occupational health monitoring and risk assessment systems to improve prevention and equity [44]. Such structured reintegration and follow-up programs would strengthen the resilience of public institutions while supporting long-term workforce health. Such strategies protect both individual health

XSL•FO
RenderX

and institutional productivity and resilience. OECD estimates long COVID costs between $864 billion and $1.04 trillion across high-income countries [58], mainly from health care expenditures and indirect productivity losses.

Limitations

This study had several limitations. First, the cross-sectional design limits causal interpretations, as associations cannot be interpreted as temporal or causal associations. Second, outcomes relied on self-reported data, which may introduce bias at several levels: recall bias (participants may not accurately remember or report their COVID-19 history or symptom duration), misclassification bias (symptoms could be wrongly attributed to or excluded from long COVID), and reporting bias (participants may over- or under-report sensitive information such as vaccination status or comorbidities). Third, no validated quality of life instruments were used, limiting conclusions regarding functional and psychosocial impact. Fourth, the sampling strategy may also affect representativeness and generalizability: participants were recruited through convenience sampling of UROPS members and voluntary participation, which introduces self-selection bias. Individuals with a stronger interest in COVID-19 or more health concerns may have been more likely to participate, leading to overrepresentation of certain profiles and underrepresentation of others. In addition, the overall response rate was low (6.12%), which may further accentuate selection bias. Moreover, generalizability is limited: although seemingly applicable to all 5.67 million French public servants, the study focused only on 1.46 million state agents, excluding 1.06 million in national education [28]. Fifth, the questionnaire did not assess whether vaccination occurred before or after SARS-CoV-2 infection, preventing interpretation of the temporal sequence and limiting conclusions. Finally, psychometric properties other than the Content Validity Index, such as internal consistency or test-retest reliability, were not assessed, as the knowledge items measured factual information rather than a latent construct, and knowledge levels may legitimately change over time with exposure to new information. In addition, the cognitive debriefing involved only six individuals; although this number is consistent with standard practice, it may still limit robustness. The knowledge score was validated internally but not benchmarked against external instruments or tested in known-groups comparisons, which may limit external validity. Future studies should consider using probability sampling strategies or weighting procedures to improve representativeness and strengthen external validity.

Conclusions

In a context of widespread SARS-CoV-2 infection, most French civil servants had experienced a past infection and a smaller but substantial proportion reported long COVID, either diagnosed or suspected. Disparities in knowledge about long COVID, attitudes, and behaviors toward COVID-19 vaccination were observed. Importantly, vaccination was found to be associated with a lower odds ratio of developing long COVID. These findings have implications for workplace prevention suggesting the value of health education, post-COVID screening for persistent symptoms, tailored reintegration plans, vaccination promotion, and mental health support. Enhancing disease literacy and occupational health programs could help strengthen the resilience of essential public services.

Acknowledgments

The authors would like to thank all the participants that answered the questionnaire. The authors also would like to thank Union Prévention Santé pour la Fonction publique, a French player in preventive health care as manager of the Public Service's compulsory health insurance scheme. Union Prévention Santé pour la Fonction publique, promoter of the initial survey and holder of the data, gave the authors permission to use the data. More particularly, the authors would like to thank Marina Molin Orru, David Ollivier-Lannuzel, José de Freitas, Adeline Stawinski, and Caroline Lefevre.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Final validated study questionnaire.
[DOCX File, 38 KB - publichealth_v11i1e83323_app1.docx]

Multimedia Appendix 2
Classification of participants by COVID-19 and long COVID status.
[DOCX File, 26 KB - publichealth_v11i1e83323_app2.docx]

Multimedia Appendix 3
Sociodemographic and health characteristics of the respondents.
[DOCX File, 36 KB - publichealth_v11i1e83323_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Analysis of long COVID knowledge.
[DOCX File, 27 KB - publichealth_v11i1e83323_app4.docx]

Checklist 1
CROSS checklist.
[DOCX File, 43 KB - publichealth_v11i1e83323_app5.docx]

References

1. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: Major findings, mechanisms and recommendations. Nat Rev Microbiol 2023 Mar;21(3):133-146. [doi: 10.1038/s41579-022-00846-2] [Medline: 36639608]
2. Post COVID-19 condition (long COVID). World Health Organization. URL: https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition [accessed 2025-04-24]
3. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV, WHO Clinical Case Definition Working Group on Post-COVID-19 Condition. A clinical case definition of post-COVID-19 condition by a Delphi consensus. Lancet Infect Dis 2022 Apr;22(4):e102-e107. [doi: 10.1016/S1473-3099(21)00703-9] [Medline: 34951953]
4. Altmann DM, Whettlock EM, Liu S, Arachchillage DJ, Boyton RJ. The immunology of long COVID. Nat Rev Immunol 2023 Oct;23(10):618-634. [doi: 10.1038/s41577-023-00904-7] [Medline: 37433988]
5. Greenhalgh T, Sivan M, Perlowski A, Nikolich J. Long COVID: A clinical update. Lancet 2024 Aug 17;404(10453):707-724. [doi: 10.1016/S0140-6736(24)01136-X] [Medline: 39096925]
6. Hou Y, Gu T, Ni Z, Shi X, Ranney ML, Mukherjee B. Global prevalence of long COVID, its subtypes, and risk factors: an updated systematic review and meta-analysis. Open Forum Infect Dis 2025 Sep;12(9):ofaf533. [doi: 10.1093/ofid/ofaf533] [Medline: 41018705]
7. Organization for Economic Co-operation and Development. The prevalence and impact of long COVID in the primary care population. : Organization for Economic Co-operation and Development; 2025. [doi: 10.1787/119b0e8f-en]
8. Tracking long COVID. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/long-covid/php/surveillance/index.html [accessed 2025-11-20]
9. Coste J, Delpierre C, Richard JB, et al. Prevalence of long COVID in the general adult population according to different definitions and sociodemographic and infection characteristics. A nationwide random sampling survey in France in autumn 2022. Clin Microbiol Infect 2024 Jul;30(7):924-929. [doi: 10.1016/j.cmi.2024.03.020] [Medline: 38527615]
10. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. Nat Med 2021 Apr;27(4):626-631. [doi: 10.1038/s41591-021-01292-y] [Medline: 33692530]
11. The impacts of long COVID across OECD countries. : Organization for Economic Co-operation and Development; 2024 URL: https://www.oecd.org/en/publications/the-impacts-of-long-covid-across-oecd-countries_8bd08383-en.html [accessed 2025-11-28] [doi: 10.1787/8bd08383-en]
12. Engelmann P, Reinke M, Stein C, et al. Psychological factors associated with long COVID: a systematic review and meta-analysis. EClinicalMedicine 2024 Aug;74:102756. [doi: 10.1016/j.eclinm.2024.102756] [Medline: 39764180]
13. Kerksieck P, Ballouz T, Haile SR, et al. Post COVID-19 condition, work ability and occupational changes in a population-based cohort. Lancet Reg Health Eur 2023 Jun 23;31:100671. [doi: 10.1016/j.lanepe.2023.100671] [Medline: 37366496]
14. Ottiger M, Poppele I, Sperling N, Schlesinger T, Müller K. Work ability and return-to-work of patients with post-COVID-19: a systematic review and meta-analysis. BMC Public Health 2024 Jul 7;24(1):1811. [doi: 10.1186/s12889-024-19328-6] [Medline: 38973011]
15. Vartanian K, Fish D, Kenton N, Gronowski B, Wright B, Robicsek A. Integrating patient-reported physical, mental, and social impacts to classify long COVID experiences. Sci Rep 2023 Sep 28;13(1):16288. [doi: 10.1038/s41598-023-43615-8] [Medline: 37770554]
16. Cutler DM. The costs of long COVID. JAMA Health Forum 2022 May 6;3(5):e221809. [doi: 10.1001/jamahealthforum.2022.1809] [Medline: 36219031]
17. Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. EClinicalMedicine 2021 Aug;38:101019. [doi: 10.1016/j.eclinm.2021.101019] [Medline: 34308300]
18. Ceban F, Ling S, Lui LMW, et al. Fatigue and cognitive impairment in post-COVID-19 syndrome: a systematic review and meta-analysis. Brain Behav Immun 2022 Mar;101:93-135. [doi: 10.1016/j.bbi.2021.12.020] [Medline: 34973396]
19. Peine C, Stoliaroff-Pepin A, Reinacher U, et al. Effectiveness of COVID-19 vaccines against post-COVID-19 condition/long COVID: systematic review and meta-analysis. Clin Microbiol Infect 2025 Aug 7;S1198-743X(25):00367-00362. [doi: 10.1016/j.cmi.2025.07.026] [Medline: 40754067]
20. Chow NKN, Tsang CYW, Chan YH, et al. The effect of pre-COVID and post-COVID vaccination on long COVID: a systematic review and meta-analysis. J Infect 2024 Dec;89(6):106358. [doi: 10.1016/j.jinf.2024.106358] [Medline: 39580033]

XSL•FO
RenderX

21. MacEwan SR, Rahurkar S, Tarver WL, et al. COVID-19 vaccination perspectives among patients with long COVID: a qualitative study. Hum Vaccin Immunother 2024 Dec 31;20(1):2327663. [doi: 10.1080/21645515.2024.2327663] [Medline: 38532547]
22. Eitze S, Sprengholz P, Korn L, Shamsrizi P, Felgendreff L, Betsch C. Vicarious experiences of long COVID: a protection motivation theory analysis for vaccination intentions. Vaccine X 2024 Jan;16:100417. [doi: 10.1016/j.jvacx.2023.100417] [Medline: 38192617]
23. Sharma A, Minh Duc NT, Luu Lam Thang T, et al. A consensus-based Checklist for Reporting of Survey Studies (CROSS). J Gen Intern Med 2021 Oct;36(10):3179-3187. [doi: 10.1007/s11606-021-06737-1] [Medline: 33886027]
24. Loi n°47-649 du 9 avril 1947 dite morice portant ratification du décret 462971 du 31-12-1946 relatif à l'institution du régime de sécurité sociale des fonctionnaires sous réserve de l'art 2 dudit décret qui se trouve modifie [Website in French]. Légifrance. URL: https://www.legifrance.gouv.fr/jorf/id/JORFTEXT000000865022 [accessed 2022-10-06]
25. Lvovschi VE, Carrouel F, du Sartz de Vigneulles B, et al. Knowledge, attitudes and practices related to medication, antibiotics, and vaccination among public service population: national survey conducted in France. Int J Environ Res Public Health 2022 Oct 28;19(21):14044. [doi: 10.3390/ijerph192114044] [Medline: 36360926]
26. ameli.fr [Website in French]. URL: https://www.ameli.fr/ [accessed 2022-10-06]
27. Cochran W. Sampling Techniques: John Wiley & Sons, Scientific Research Publishing; 1977. URL: https://www.scirp.org/reference/ReferencesPapers?ReferenceID=1390266 [accessed 2025-07-13]
28. Rapport annuel sur l'état de la fonction publique—édition 2023. : Portail Fonct Publique; 2025 URL: https://www.fonction-publique.gouv.fr/toutes-les-publications/rapport-annuel-sur-letat-de-la-fonction-publique-edition-2025 [accessed 2025-09-22]
29. Global Programme on Tuberculosis and Lung Health (GTB). Advocacy, Communication and Social Mobilization for TB Control: A Guide to Developing Knowledge, Attitude and Practice Surveys: World Health Organization; 2008. URL: https://www.who.int/publications/i/item/9789241596176
30. Andrade ALM, Scatena A, Martins GDG, et al. Validation of Smartphone Addiction Scale–Short Version (SAS-SV) in Brazilian adolescents. Addict Behav 2020 Nov;110:106540. [doi: 10.1016/j.addbeh.2020.106540] [Medline: 32682269]
31. Polit DF, Beck CT. The Content Validity Index: Are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 2006 Oct;29(5):489-497. [doi: 10.1002/nur.20147] [Medline: 16977646]
32. Haute Autorité de Santé. Methodology for guideline development. Haute Aut Santé. 2015. URL: https://www.has-sante.fr/jcms/c_2040454/en/methodology-for-guideline-development [accessed 2025-06-24]
33. Symptômes prolongés suite à une covid-19 de l'adulte - diagnostic et prise en charge [Website in French]. Haute Aut Santé. URL: https://www.has-sante.fr/jcms/p_3237041/fr/symptomes-prolonges-suite-a-une-covid-19-de-l-adulte-diagnostic-et-prise-en-charge [accessed 2025-10-07]
34. WHO COVID-19 dashboard. World Health Organization. 2025. URL: https://data.who.int/dashboards/covid19/summary [accessed 2025-06-23]
35. Santé Publique France. Point épidémiologique COVID-19 du 10 mars. Ralentissement moins marqué de la circulation du SARS-CoV-2 au niveau national [Website in French]. 2022. URL: https://www.santepubliquefrance.fr/presse/2022/point-epidemiologique-covid-19-du-10-mars-2022.-ralentissement-moins-marque-de-la-circulation-du-sars-cov-2-au-niveau-national [accessed 2025-11-20]
36. Long COVID basics. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/covid/long-term-effects/index.html [accessed 2025-06-23]
37. Groff D, Sun A, Ssentongo AE, et al. Short-term and long-term rates of postacute sequelae of SARS-CoV-2 infection: a systematic review. JAMA Netw Open 2021 Oct 1;4(10):e2128568. [doi: 10.1001/jamanetworkopen.2021.28568] [Medline: 34643720]
38. Gamillscheg P, Łaszewska A, Kirchner S, Hoffmann K, Simon J, Mayer S. Barriers and facilitators of healthcare access for long COVID-19 patients in a universal healthcare system: qualitative evidence from Austria. Int J Equity Health 2024 Oct 23;23(1):220. [doi: 10.1186/s12939-024-02302-4] [Medline: 39443935]
39. Ayoubkhani D, Zaccardi F, Pouwels KB, et al. Employment outcomes of people with long Covid symptoms: community-based cohort study. Eur J Public Health 2024 Jun 7;34(3):489-496. [doi: 10.1093/eurpub/ckae034] [Medline: 38423541]
40. Zawilska JB, Kuczyńska K. Psychiatric and neurological complications of long COVID. J Psychiatr Res 2022 Dec;156:349-360. [doi: 10.1016/j.jpsychires.2022.10.045] [Medline: 36326545]
41. Seighali N, Abdollahi A, Shafiee A, et al. The global prevalence of depression, anxiety, and sleep disorder among patients coping with Post COVID-19 syndrome (long COVID): a systematic review and meta-analysis. BMC Psychiatry 2024 Feb 6;24(1):105. [doi: 10.1186/s12888-023-05481-6] [Medline: 38321404]
42. Fancourt D, Steptoe A, Bu F. Psychological consequences of long COVID: comparing trajectories of depressive and anxiety symptoms before and after contracting SARS-CoV-2 between matched long- and short-COVID groups. Br J Psychiatry 2023 Feb;222(2):74-81. [doi: 10.1192/bjp.2022.155] [Medline: 36458509]
43. Goodman ML, Molldrem S, Elliott A, Robertson D, Keiser P. Long COVID and mental health correlates: a new chronic condition fits existing patterns. Health Psychol Behav Med 2023;11(1):2164498. [doi: 10.1080/21642850.2022.2164498] [Medline: 36643576]

XSL•FO
RenderX

44. Post COVID-19 condition (long COVID). World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/post-covid-19-condition-(long-covid) [accessed 2025-10-04]
45. Català M, Mercadé-Besora N, Kolde R, et al. The effectiveness of COVID-19 vaccines to prevent long COVID symptoms: staggered cohort study of data from the UK, Spain, and Estonia. Lancet Respir Med 2024 Mar;12(3):225-236. [doi: 10.1016/S2213-2600(23)00414-9] [Medline: 38219763]
46. Yang J, Rai KK, Alfred T, et al. The impact of COVID vaccination on incidence of long COVID and healthcare resource utilisation in a primary care cohort in England, 2021-2022. BMC Infect Dis 2025 Feb 13;25(1):214. [doi: 10.1186/s12879-024-10097-6] [Medline: 39948466]
47. Binder AR, May K, Murphy J, Gross A, Carlsten E. Environmental health literacy as knowing, feeling, and believing: analyzing linkages between race, ethnicity, and socioeconomic status and willingness to engage in protective behaviors against health threats. Int J Environ Res Public Health 2022 Feb 25;19(5):2701. [doi: 10.3390/ijerph19052701] [Medline: 35270393]
48. Vo HT, Dao TV, Do TX, et al. Association between underlying health conditions and long COVID among non-hospitalized and hospitalized individuals as modified by health literacy: A multi-center study. Public Health (Fairfax) 2025 Feb;239:87-93. [doi: 10.1016/j.puhe.2024.12.032] [Medline: 39798220]
49. Ling J, Jiang H, Wang X, Rao H. Health literacy affected the residents' knowledge, attitude, practice for prevention and control of COVID-19 in Shanxi Province, China. Sci Rep 2023 Mar 2;13(1):3567. [doi: 10.1038/s41598-023-30730-9] [Medline: 36864261]
50. He N, Tunala S, Wu R, Zhao C, Wuren T, Hushuud C. Knowledge, attitudes, and practices among community populations toward long COVID in China. Int J Gen Med 2024;17:4455-4463. [doi: 10.2147/IJGM.S474636] [Medline: 39372136]
51. Lazarus JV, Wyka K, White TM, et al. A survey of COVID-19 vaccine acceptance across 23 countries in 2022. Nat Med 2023 Feb;29(2):366-375. [doi: 10.1038/s41591-022-02185-4] [Medline: 36624316]
52. Ballering AV, van Zon SKR, Olde Hartman TC, Rosmalen JGM, Lifelines Corona Research Initiative. Persistence of somatic symptoms after COVID-19 in the Netherlands: an observational cohort study. Lancet 2022 Aug 6;400(10350):452-461. [doi: 10.1016/S0140-6736(22)01214-4] [Medline: 35934007]
53. Menges D, Ballouz T, Anagnostopoulos A, et al. Burden of post-COVID-19 syndrome and implications for healthcare service planning: a population-based cohort study. PLoS ONE 2021;16(7):e0254523. [doi: 10.1371/journal.pone.0254523] [Medline: 34252157]
54. Matta J, Wiernik E, Robineau O, et al. Association of self-reported COVID-19 infection and SARS-CoV-2 serology test results with persistent physical symptoms among French adults during the COVID-19 pandemic. JAMA Intern Med 2022 Jan 1;182(1):19-25. [doi: 10.1001/jamainternmed.2021.6454] [Medline: 34747982]
55. Walker AJ, MacKenna B, Inglesby P, et al. Clinical coding of long COVID in English primary care: a federated analysis of 58 million patient records in situ using OpenSAFELY. Br J Gen Pract 2021 Nov;71(712):e806-e814. [doi: 10.3399/BJGP.2021.0301] [Medline: 34340970]
56. Conseil scientifique Covid-19. Vivre avec les variants: la pandémie n'est pas terminée - mieux anticiper [Website in French]. Ministère Trav Santé Solidar Fam. URL: https://sante.gouv.fr/actualites-presse/presse/dossiers-de-presse/article/conseil-scientifique-covid-19 [accessed 2025-10-04]
57. Décision n° 20220070/DP/SBP du 16 février 2022 de la présidente de la haute autorité de santé portant actualisation du document intitulé « réponses rapides dans le cadre de la covid-19 – prise en charge de premier recours des patients atteints de covid-19 [Website in French]. Haute Aut Santé. URL: https://tinyurl.com/mr25pyk9 [accessed 2025-10-04]
58. Organization for Economic Co-operation and Development. OECD policy responses to coronavirus (COVID-19). 2025. URL: https://www.oecd.org/en/publications/oecd-policy-responses-to-coronavirus-covid-19_5b0fd8cd-en.html [accessed 2025-06-23]

Abbreviations

CROSS: Checklist for Reporting of Survey Studies
HAS: Haute Autorité de la Santé
I-CVI: Item-level Content Validity Index
KAB: Knowledge, Attitudes, and Behaviors
OECD: Organization for Economic Co-operation and Development
OR: odds ratio
S-CVI/Ave: Scale-level Content Validity Index
UROPS: Union Prévention Santé pour la Fonction publique
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Mavragani, S Brini; submitted 31.08.25; peer-reviewed by E Oliveira, GY Rochefort; revised version received 28.10.25; accepted 05.11.25; published 05.12.25.

Please cite as:
Carrouel F, Lvovschi VE, du Sartz de Vigneulles B, Rhanoui M, Salamon R, Lamure M, Salque C, Lan R, Dussart C
Prevalence, Risk Factors, Disease-Related Knowledge, and Vaccination Attitudes and Behaviors for Long COVID Among French Civil Servants: Cross-Sectional Survey
JMIR Public Health Surveill 2025;11:e83323
URL: https://publichealth.jmir.org/2025/1/e83323
doi:10.2196/83323

© Florence Carrouel, Virginie-Eve Lvovschi, Benjamin du Sartz de Vigneulles, Maryem Rhanoui, Roger Salamon, Michel Lamure, Corélie Salque, Romain Lan, Claude Dussart. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 5.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using the Health Belief Model to Examine Parental Knowledge and Health Beliefs About Human Papilloma Virus (HPV) and iHPV Vaccine in Kuwait: Cross-Sectional Survey Study

Ahmad Abuzoor[1], PhD; Md Shafiqur Rahman Jabin[1,2], PhD; Cyril Eshareturi[3], BSc, MPH, PhD; Rae Nesbitt[4], MSc, RM; Cor Jonker[1], Drs

[1]Department of Public Health, Faculty of Health Studies, University of Bradford, Bradford, United Kingdom

[2]Department of Medicine and Optometry, Linnaeus University, Pedalstråket 11, Kalmar, Sweden

[3]Public Health Principal, London Borough of Ealing, London, United Kingdom

[4]Department of Public Health, Faculty of Health Studies, University of Bradford, Bradford, United Kingdom

Corresponding Author:
Md Shafiqur Rahman Jabin, PhD
Department of Public Health, Faculty of Health Studies, University of Bradford, Bradford, United Kingdom

Abstract

Background: Cervical cancer (CC) is a major public health issue, accounting for approximately 350,000 deaths, around 7.5% of all female cancer deaths worldwide, in 2018. Human papillomavirus (HPV) is the most common virus infecting the reproductive system. Despite the high number of diagnosed cases of CC globally, prevention is possible. Vaccination against HPV is considered to be a primary prevention strategy, while cervical screening can also play a secondary prevention role.

Objective: This study aimed to examine the knowledge and health beliefs of parents in Kuwait towards HPV and HPV vaccination in order to prepare for the development of a national policy on CC.

Methods: A cross-sectional survey was conducted among a representative multistage sample of 538 parents and guardians of eligible children aged 12–17 years in Kuwait, yielding a response rate of 89%. The survey was structured using the health belief model. Analysis showed statistically significant links between knowledge, health beliefs, concepts, and vaccination intention.

Results: Knowledge of HPV and HPV vaccination was low in our study population: 55.6% (n=297), 24.9% (n=133), and 19.5% (n=104) for poor, fair, and good knowledge, respectively. Parents of daughters scored lower on perceived susceptibility to HPV and were more likely to have a higher perception of barriers to HPV vaccination, even though fathers were more likely to believe their daughters were at risk. HPV vaccination has the stigma of promiscuity attached, even though half of the parents are willing to accept HPV vaccination if that recommendation comes from Uhealth officials or relatives. A greater proportion of parents with female children had a low perception of the severity of HPV infection compared to those with male children (n=154, 58.6% vs n=134, 49.4%; *P*=.043). Around 52% (n=278) of parents perceived a high benefit of HPV vaccination. Parents with a female child had a lower perception of HPV vaccine benefits compared to parents with a male child. The findings demonstrated that parents with higher levels of education were better informed about the use of HPV vaccines in controlling the illness. Parents with female children were 1.34 times more likely to act on the recommendation for HPV vaccination compared to parents with male children after a recommendation from an official source, such as doctors or healthcare professionals.

Conclusions: Recommendations for a Kuwaiti vaccination policy for HPV must take into consideration different knowledge levels of parents for groups with different educational levels, as well as the stigma of promiscuity and other barriers, and various health beliefs regarding susceptibility for daughters and sons, respectively.

(JMIR Public Health Surveill 2025;11:e75818) doi:10.2196/75818

KEYWORDS

human papillomavirus; human papillomavirus vaccine; health belief model; parents.

Introduction

Human papillomavirus (HPV), a sexually transmitted virus, accounts for 99% of cervical cancer (CC) cases worldwide [1]. Also, HPV is responsible for around 70% of vulvar and vaginal cancers and 60% of penile cancers. Approximately 80% of men and women are predicted to contract HPV over their lives, typically as a result of sexual activity [2]. The best prevention against HPV infection and, therefore, CC is the HPV vaccine. Getting HPV vaccination at younger ages is medically advantageous in various respects. HPV vaccines demonstrated over 99% effectiveness when given to women without previous

XSL•FO
RenderX

HPV exposure [2]. There are over 100 species, with 30 - 40 primarily found in the human genital tract. Sero-risk strains cause cervical, vulvar, vaginal, and anal cancers [3]. HPV 16 and 18 are implicated in 70% of all CC globally [4,5]. The vaccine is generally given in a series of 3 doses over 6 months to individuals aged 11 to 26, demonstrating considerable efficacy in preventing CC and decreasing the incidence of several HPV-related diseases [6]. In this study, a quantitative survey was conducted to evaluate parents' knowledge and perceptions of HPV and its vaccination, aiming to determine demographic features of parents in Kuwait who have a child between the ages of 12 and 17 and parental awareness and health behaviors as measured by the health belief model (HBM), to create a comprehensive health promotion program for Kuwait.

Toward this end, the study has the following specific objectives: to determine Kuwaiti parents' health beliefs and levels of knowledge on HPV and its vaccine. To examine health beliefs within the HBM about HPV and its vaccination. To evaluate the differences in health belief scores based on demographic characteristics and identify distinctions across different societal groups. To develop recommendations for health promotion to vaccinate against HPV. The lack of knowledge and awareness is a significant barrier to the uptake of HPV vaccination in Middle Eastern countries [7]. Previous research found higher acceptance rates among better-educated parents and those living in cities, which supports this pattern. To increase vaccination uptake in these populations, targeted interventions are required to meet the unique needs and concerns of rural and less educated populations [8].

Kuwait has lower rates of CC than other high-income nations like the United States and the United Kingdom, but its age-standardized death rate of 2.99 per 100,000 women is higher than the United States (2.77) and the United Kingdom (2.01). This gap in treatment and prevention highlights a significant gap in CC prevention. The rising prevalence of CC and deaths in Kuwait highlights the need for stronger preventive measures and better access to early detection and treatment services.

Mass vaccination of boys and girls could potentially eliminate HPV, the leading cause of CC, and contribute to the WHO's universal campaign to end CC as a public health issue. However, recent research suggests that gender-neutral programs, including HPV shots for both genders, could be more effective. Gender norms may also affect vaccine acceptability, with females potentially facing greater challenges related to vaccination [9]. One study in Saudi Arabia mentioned that only 11% of parents were aware of HPV [10]. Another study in Turkey found that 24.5% of mothers and 21.2% of fathers were mindful of HPV vaccination [11]. In Iran, about 76% of parents were unaware of the seriousness of HPV infection. 76% of parents were unaware of the seriousness of HPV infection [12]. On the other hand, a high level of HPV vaccine knowledge was the most important factor influencing HPV vaccine acceptability, as evidenced by various studies. In a study of female university students studying in health sciences in Fujian Province, the results showed that the most important factor related to the intention to obtain the HPV vaccine was high HPV vaccination knowledge [13]. The most common ways of HPV vaccine promotion were hospital or school health education, as well as doctor or nurse recommendations [14].

The HPV vaccine is not part of Kuwait's national vaccination program but is available in private medical care through the private health system [15]. Recently, at the end of 2024, it became optional in the health care center government. The absence of a national cervical screening program and HPV vaccination program is a significant obstacle to preventing CC in Kuwait. One of the essential motivations for this thesis is to reduce morbidity and mortality of CC by introducing the HPV vaccine (primary prevention) in the national immunization program in Kuwait. Also, the cost of treatment for cancer is more expensive than prevention by the HPV vaccine or early diagnosis by regular CC screening. New Kuwaiti studies on children's vaccinations should consider parental perspectives before, during, and after HPV vaccination and integrate HPV vaccination into the national immunization program. Because vaccinations are given at a young age, parents are generally the primary decision makers, and their consent is required for immunization. Consequently, educational HPV campaigns have primarily concentrated on educating parents rather than children to enhance vaccination rates [16]

Methods

Overview

The cross-sectional descriptive questionnaire was conducted in Kuwait in May 2021, with 6 educational regions: the Asimah, Hawalli, Farwaniya, Ahmadi, Mubarak Al Kabeer, and Al-Jahra Governorates. The parents of individuals aged 12 - 17 years in the intermediate stage in government schools have been chosen as eligible study participants. The sample size is calculated using a formula for estimating continuous data, as the study involves calculating knowledge and perception scores for health beliefs. The aggregate scores are measured through knowledge questions, where parents earn points for correct answers and have perception scores on a Likert scale. An 88-item questionnaire was designed for the study's purpose. This consisted of 53 items based on HBM, 25 questions about HPV knowledge, the HPV vaccine, and HPV-associated cancers, and one additional question, "Would you like to participate in an interview?"

The study aimed to measure participants' awareness of HPV and the HPV vaccine, as well as their health beliefs and experiences about vaccination. Closed-ended questions were used for statistical analysis, with each respondent given a choice of "True," "False," or "I do not know." A 5-point Likert scale was used to measure parental perceptions and health beliefs, according to the HBM constructs.

The sample comprised equal numbers of boys and girls, with a sex ratio of 1:1. The CI is 95%, and the power is 80%. No previous studies in the literature have found anything that could help in predicting the scores. The sample size calculation is based on the perceived barriers score, a summation score ranging from 12 to 60 points. We assume that the means are 45 and 43 in the girls' and boys' groups, respectively. In addition, because no previous studies were found from which we could derive an

XSL•FO
RenderX

SD, we are assuming an SD of 7 for each group, resulting in a sample size of 386. The margin of error is assumed to be 10%, and 35% of the questionnaires are expected to lack responses or be unreturned [17]. To compensate, the total sample size should be 573.21; rounding up, there will be 600 questionnaires—sample size calculation based on the Open Epi formula calculation [18].

A multistage sampling technique was used to select students for the study. In the first stage, 3 out of the 6 educational regions in Kuwait were chosen using simple random sampling. In the second stage, 20 schools were selected by ballot, without replacement, from the list of all schools in the region. An equal number of students (30 each) were recruited consecutively from each of the 20 schools in the final stage, for a total of 600 participants.

The questionnaire, originally developed in English, was then translated into Arabic, Kuwait's official language, by a licensed translation center to ensure cultural and linguistic appropriateness for the target audience. This process was rigorous, designed to build an accessible, accurate data-collection instrument for the research goals. The questionnaire was administered in 2 stages: the pilot study and the final study. The survey was developed and presented in Arabic to participants. Before the pilot test, the researcher and supervisors reviewed the face validity of each item. The instrument was validated in a pilot test with 15 parents who did not participate in the final analysis. Following the pilot test, no further revisions were made to the questionnaire. The printed questionnaires and information sheets were given to students, who took them home to their parents and, upon completion, returned them to the schools, from which they were collected. During data collection, the vaccine was only available in private health sectors at a cost during the study period. Recently, at the end of 2024, it became optional in the government health care centers.

Data Analysis

The study used a quantitative method to analyze the knowledge and beliefs about HPV vaccines among parents in Kuwait. The data was collected through questionnaires and coded, cleaned, and verified using SPSS version 27 software (IBM Corp). The data analysis was structured using questionnaire sections aligned with the HBM constructs. Descriptive statistics were used to represent the data clearly and easily, while graphs, tables, and statistical models were used to interpret and present numerical data [19].

The HBM scores were compared with demographic variables to discover patterns in responses. Descriptive and inferential statistics were used to compare trends in parental knowledge and health beliefs about the HPV vaccine. The descriptive analysis focused on central tendency and dispersion, while the statistical analysis aimed to predict behavioral effects associated with HPV vaccination uptake by finding meaningful patterns and correlations in the data.

The research question related to HBM was analyzed using a binary categorization of knowledge score (high or low) and HBM constructs based on the median. Both the Knowledge score and HBM variables were used in logistic regression to calculate the odds ratio [20]. Chi-square tests and logistic regression analyses were used to evaluate the relationships between categorical variables, such as parents' interests in vaccinating their children against HPV [21]. Inferential analyses investigated the knowledge about the HPV vaccine and each HBM construct over different demographic groups. The *t* tests and ANOVA or Mann-Whitney *U* tests were applied to define mean and median scores between two or more groups.

Using unadjusted and adjusted odds ratios is shown to random imbalances. When using an adjusted method, there is a chance that the nature of the link between the covariate and outcome will be misspecified for continuous covariates. The disparities between parents, including age, sex, parent-child sex, and other participant variables, are disregarded. With these "baseline variables" adjusted, we can more accurately predict a response by taking these parent-to-parent differences into account. As a result, the analysis's statistical power is increased. In conclusion, the quantitative method was chosen because it does not involve a direct relationship with participants and can improve response and decrease bias [22].

Ethical Considerations

The study received ethical approval from the University of Bradford which had reference number (E843), the Ministry of Health in Kuwait with reference number (1345/2020), and the Ministry of Education in Kuwait with reference number (259006 and 00010918). Participants were informed about the research's purpose, methodology, and confidentiality policy, ensuring anonymity and privacy. The researcher prioritized neutrality and objectivity for a credible data collection process. Participants signed informed consent forms, which explained the study's purpose, methodology, and confidentiality policy. Parents provided informed consent for the pilot study. The questionnaire provided an explanation of the study's purpose, methodology, and confidentiality policy. Participants indicated their consent by returning completed questionnaires, which were treated as implied consent. Data will be preserved for five years and archived within the University of Bradford, and no medical experiments, medicines, biopsies, or other interventional testing will be conducted in the future.

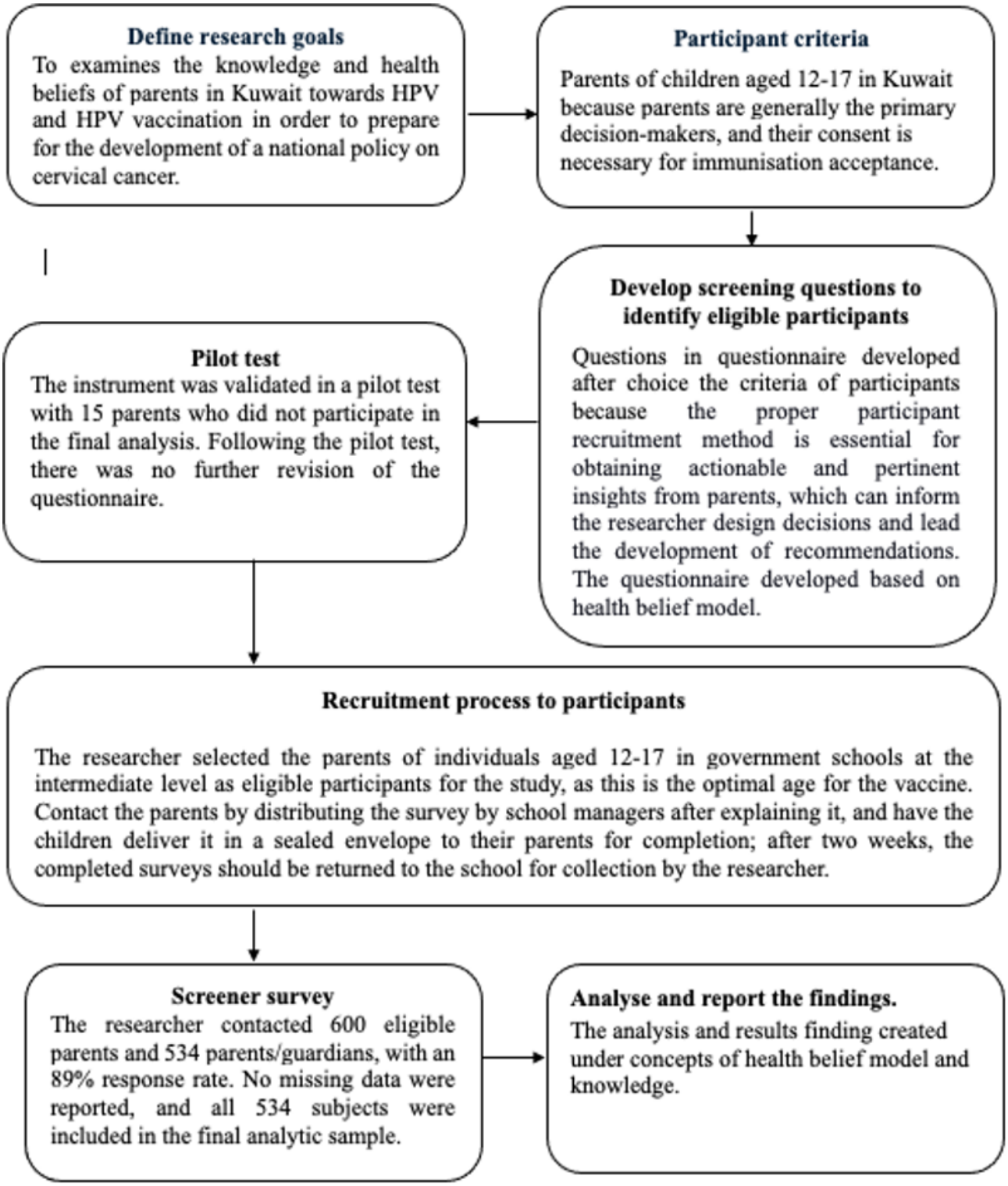
Results

Overview

The researcher contacted 600 eligible parents and 534 parents and guardians, with an 89% response rate. No missing data were reported, and all 534 individuals were included in the final analytic sample. Figure 1 shows a flow diagram for the research study recruitment process.

XSL•FO
RenderX

Figure 1. A flow diagram for research study recruitment process.



Demographic Characteristics

The study's data quality in Kuwait indicates that the participants' age is similar to the general population, with a mean age of 42.78 years (SD 5.71; range, 24 - 60 y), identical to the average Kuwaiti age for marriage. The HPV vaccine is given to children aged 12 - 16 years, considered the Intermediate School period in Kuwait. The mean age of children surveyed was 13.45 years (SD 1.23, range 12 - 17 y), with almost as many boys as girls. Females filled in survey forms twice as many times as fathers, possibly due to the mother's higher education and refusal to allow the father to fill in the form.

The employed versus unemployed ratios in Kuwait are similar to those of the general population, with many Kuwaitis participating. Almost half of the sampled population holds a Bachelor's degree, similar to the general population. However, the females had higher education compared to males, which might be because educated females do not let the child's father fill out the forms. Nearly half of the participants (n=247, 46.4%) have a monthly family income of 700-1500 KD. The data

XSL•FO
RenderX

represent the Kuwaiti population well in all aspects, but there is evidence that females are slightly more educated than the general population, resulting in more participation in the survey.

Previous Knowledge of HPV and Vaccine Among the Study Parents

The study found low knowledge of HPV and its vaccine among the population, with significant associations between demographic characteristics such as nationality, monthly family income, and educational level. Non-Kuwaiti parents of eligible children or adolescents had three times higher HPV knowledge compared to Kuwaitis. Parents with monthly family incomes of between US $2260 and US $4,800 were more likely to have a better understanding than those with lower incomes. Parents of eligible children with advanced degrees were 3.6 times more likely to have higher HPV knowledge.

Health Belief Concepts

HBM builds on behavioral and psychological models that elaborate on two primary motivations for health behavior: (1) wanting to avoid illness or to recover from disease if already ill, and (2) believing that an action in the health world will prevent or cure illness. In most cases, a person makes a choice based on how they think about the advantages and disadvantages of health behavior.

Perceived Susceptibility

The mean score of perceived susceptibility was 29.06 (SD 4.25), with scores ranging from 18.00 to 46.00. Female participants (n=363) outnumbered male participants (n=171) in the analysis (Multimedia Appendix 1).

Out of 534 participants, 12.0% (n=64) believed their adolescent sons would get anal cancer in their lifetime, compared to 12.9% (n=69) individuals who believed that their daughters would have CC. More men, compared to women, believed that their daughters were at risk of HPV during their lifetime (n=53, 31.0% vs n=66, 18.2%; *P*=.001; Multimedia Appendix 2).

Statistically significant relationships were observed between perceived susceptibility to HPV infection and a number of sociodemographic characteristics of the study participants.

Table 1 showed a slightly greater proportion of participants with a bachelor's degree had an increased perception of susceptibility to HPV compared to their counterparts of various educational levels (*P*<.001).

Table . Comparison of levels of perceived susceptibility (high vs low) versus demographic characteristics.

Demographic characteristics	Low (n=244)	High (n=290)	*P* value
Sex (Female), n (%)	140 (57.4)	148 (51.0)	.17
Parent sex (Male), n (%)	76 (31.1)	95 (32.8)	.76
Child age, mean (SD)	013.55 (1.27)	013.36 (1.18)	.06
Parent age, mean (SD)	43.07 (6.35)	42.54 (5.12)	.29
Income, n (%)			<.001[a, b]
Less than 700 KD	41 (16.8)	30 (10.3)	
700 - 1500 KD	132 (54.1)	117 (40.3)	
More than 1500 KD	71 (29.1)	143 (49.3)	
Marital status, n (%)			.005[a, b]
Single	11 (4.5)	5 (1.7)	
Married	189 (77.5)	257 (88.6)	
Divorced	36 (14.8)	25 (8.6)	
Widowed	8 (3.3)	3 (1.0)	
Employment (unemployment), n (%)	28 (011.50)	31 (10.7)	.88
Nationality (non-Kuwait), n (%)	18 (7.4)	21 (7.2)	≥.99
Education, n (%)			<.001[a, b]
Primary education	9 (3.7)	2 (0.7)	
High school education	52 (21.3)	31 (10.7)	
Diploma	53 (21.7)	64 (22.1)	
Bachelor's degree	122 (50.0)	152 (52.4)	
Master's or PhD	8 (3.3)	41 (14.1)	

[a]Indicates statistical significance.

[b]indicates the χ^2 test.

XSL•FO
RenderX

Table 2 shows unadjusted and adjusted results for perceived susceptibility to HPV, categorized by a median score from our logistic regression models. In our unadjusted logistic regression models, participants' educational level and monthly family income showed a statistically significant and positive correlation with perceived susceptibility to HPV.

Table . Logistic regression models illustrate the relationship between participants' characteristics and their perceived susceptibility to human papillomavirus (HPV) infection among adolescents. Adjusted odds ratios are corrected for the other variables in the model and used for multivariate analysis, whereas unadjusted odds ratios are used for univariate analysis.

Characteristics	Perceived susceptibility to infection of HPV			
	Unadjusted odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Age	0.880 (0.760 - 1.010)	.06	0.870 (0.750 - 1.020)	.82
Child adolescent sex				
Male (reference)	__[a]	—	—	—
Female	0.770 (0.550 - 1.090)	.14	0.620 (0.420 - 0.910)	.03[b]
Parent characteristics				
Age of respondent	0.98 (0.95 - 1.01)	.29	0.99 (0.97 - 1.03)	.55
Parent sex				
Female (reference)	—	—	—	—
Male	1.080 (0.750 - 1.550)	.69	0.96 (0.62 - 1.50)	.85
Nationality				
Kuwaiti (reference)	—	—	—	—
Non-Kuwaiti	0.98 (0.51 - 1.90)	.95	0.97 (0.48 - 1.99)	.94
Education				
Primary education (reference)	—	—	—	—
High school education	2.68 (0.64 - 18.35)	.23	2.73 (0.57 - 20.10)	.25
Diploma	5.43 (1.33 - 36.67)	.04	6.06 (1.27 - 44.69)	.04[b]
Bachelor's degree	5.61 (1.41 - 37.22)	.03	5.78 (1.23 - 42.32)	.04[b]
Masters or PhD	23.06 (4.88 - 171.98)	<.001[b]	21.31 (3.77 - 178.35)	.001[b]
Monthly family income				
Less than 700 KD (reference)	—	—	—	—
700 - 1500 KD	1.21 (0.71 - 2.08)	.48	1.04 (0.58 - 1.89)	.62
More than 1500 KD	2.75 (1.59 - 4.80)	<.001[b]	1.73 (0.92 - 3.26)	.91
Employment status				
Employment (reference)	—	—	—	—
Unemployment	0.920 (0.540 - 1.590)	.77	1.780 (0.880 - 3.700)	.74
Marital status				
Single (reference)	—	—	—	—
Married	2.99 (1.07 - 9.63)	.05	2.06 (0.68 - 7.05)	.22
Divorced	1.53 (0.49 - 5.35)	.48	1.35 (0.40 - 5.05)	.64
Widowed	0.83 (0.14 - 4.43)	.82	0.42 (0.06 - 2.49)	.34

[a]Not applicable.

[b]Indicates statistical significance.

Approximately half of the parents reported that HPV and its related diseases might affect their children. After adjusting for other parent and child or adolescent factors, significant associations were observed between education level, the child's

XSL•FO
RenderX

gender, and perceived susceptibility to HPV. Additionally, parents of a girl child did not perceive vulnerability (susceptibility) to HPV more readily than parents of a boy child.

Perceived Severity

Multimedia Appendix 3 provides descriptive statistics for the overall perceived severity score of HPV across child participants, classified by gender. The study had more female (n=363) than male (n=171) participants. The perceived severity score of men and fathers was greater (mean 31.05, SD 9.53) than that of mothers and female guardians (mean 30.17, SD 9.24). The overall median perceived severity score for HPV was 30.46 (SD 9.34), ranging from 9.00 to 45.00.

A total of 27.2% (n=145) parents believed that infection with HPV might lead to morbidity (Multimedia Appendix 4). Although male and female participants generally held similar perceptions of HPV severity, many significant differences were observed. A higher percentage of men, compared to women, believed that HPV may result in serious health complications for their children. Moreover, a more significant number of men (n=88, 51.50%), compared to women (n=151, 41.60%; *P*=.04) believed that infection with HPV could have a negative impact on their children's education, interruption of life (fathers: n=99, 57.90% vs mothers: n=165, 45.50%; *P*=.01), or impede their girls' chances of marriage (fathers: n=94, 55.00% vs mothers: n=144, 39.70%; *P*=.001).

We found statistically significant relationships between the perceived severity of HPV infection and a number of baseline sociodemographic characteristics of the study participants (Table 3). A more significant proportion of parents with female children (n=154, 58.6%) compared to male children (n=134, 49.4%; *P*=.04) had a low perception of the severity of HPV infection. Our analysis showed a significant percentage difference in perceived HPV infection severity among respondents of non-Kuwaiti nationality (low severity: n=9, 3.4% vs high severity: n=30, 11.1%; *P*=.001).

Table . Comparisons of levels of perceived severity versus demographic characteristics.

Demographic characteristics	Low (n=263)	High (n=271)	*P* value[a]
Sex (Female), n (%)	154 (58.6)	134 (49.4)	.04[b]
Sex of parent (Male), n (%)	75 (28.5)	96 (35.4)	.11
Child age, mean (SD)	013.42 (1.19)	013.47 (1.26)	.66
Parents' age, mean (SD)	43.47 (5.52)	042.11 (5.82)	.006[b]
Income monthly, n (%)			.61
Under 700 KD	37 (14.1)	34 (12.5)	
700 - 1500 KD	117 (44.5)	132 (48.7)	
More than 1500 KD	109 (41.4)	105 (38.7)	
Marital status, n (%)			.32
Single	6 (2.30)	10 (3.70)	
Married	215 (81.7)	231 (85.2)	
Divorced	36 (13.7)	25 (9.2)	
Widowed	6 (2.30)	5 (1.80)	
Employment (unemployment), n (%)	30 (11.4)	29 (10.7)	.90
Nationality (non-Kuwait), n (%)	9 (3.4)	30 (11.1)	.001[b]
Education, n (%)			.054[a]
Primary education	9 (3.4)	2 (0.7)	
High school education	39 (14.8)	44 (16.2)	
Diploma	59 (22.4)	58 (21.4)	
Bachelor's degree	139 (52.9)	135 (49.8)	
Masters or PhD	17 (6.5)	32 (11.8)	

[a]indicates the χ^2 test.

[b]Indicates statistical significance.

Table 4 shows unadjusted and adjusted results for the perceived severity of HPV, categorized by median severity score from our logistic regression models. In the unadjusted logistic regressions, child sex (female), age, education, and nationality were statistically associated with perceived HPV severity. Parents of a female child had 29% less empathy for the severity of HPV than parents of a male child (OR: 0.69, 95% CI 0.49 - 0.97, *P*=.04). For each year of age increase, participants

XSL•FO
RenderX

perceived 4% less severe HPV infection (OR: 0.96, 95% CI 0.93 - 0.99, *P*=.006).

Table . The association between the characteristics of participants and their perceived severity of infection of human papillomavirus (HPV) among adolescents is illustrated by logistic regression models.

Characteristics of adolescents	Perceived severity of the infection of HPV			
	Unadjusted odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Child age	1.03 (0.89 - 1.19)	.66	1.09 (0.94 - 1.27)	.23
Child adolescent sex				
Male (reference)	__[a]	—	—	—
Female	0.69 (0.49 - 0.97)	.04[b]	0.67 (0.46 - 0.98)	.04[b]
Parent characteristics				
Age of respondent	0.96 (0.93 - 0.99)	.006[b]	0.95 (0.92 - 0.98)	.004
Parent sex				
Female (Reference)	—	—	—	—
Male	1.38 (0.96 - 1.99)	.09	1.30 (0.85 - 1.99)	.23
Nationality				
Kuwaiti (reference)	—	—	—	—
Non-Kuwaiti	3.51 (1.70 - 7.99)	.001[b]	3.61 (1.67 - 8.52)	.002[b]
Education				
Primary education (reference)	—	—	—	—
High school education	5.08 (1.22 - 34.65)	.05[b]	3.89 (0.82 - 28.44)	.12
Diploma	4.42 (1.08 - 29.84)	.06	3.74 (0.81 - 27.09)	.12
Bachelor’s degree	4.37 (1.10 - 29.01)	.06	4.28 (0.94 - 30.63)	.09
Masters or PhD	8.47 (1.92 - 59.87)	.01[b]	7.89 (1.54 - 61.01)	.02[b]
Monthly household income				
Less than US $2260	—	—	—	—
Income US $2260-4800	1.23 (0.72 - 2.09)	.45	1.32 (0.74 - 2.38)	.35
More than US $4800	1.05 (0.61 - 1.79)	.86	0.97 (0.52 - 1.81)	.93
Employment status				
Employment (Reference)	—	—	—	—
Unemployment	0.93 (0.540 - 1.60)	.79	1.14 (0.59 - 2.21)	.69
Marital status				
Single (Reference)	—	—	—	—
Married	0.65 (0.22 - 1.77)	.40	0.64 (0.20 - 1.87)	.42
Divorced	0.42 (0.13 - 1.27)	.13	0.44 (0.12 - 1.44)	.18
Widowed	0.50 (0.09 - 2.36)	.38	0.53 (0.10 - 2.65)	.44

[a]Not appilcable.

[b]Indicates statistical significance.

In total, 50% (n=267) of the children’s parents had a high perception of HPV severity. A greater proportion of parents with female children had a low perception of the severity of HPV infection (n=154, 58.6% vs n=134 , 49.4%; *P*=.04), compared to those with male children. We also found that Kuwaiti nationals, parents of male children, and parents of younger children had significantly higher odds of reporting high perceived severity of HPV. This suggests that these groups may have more awareness or concern about the potential consequences of HPV infection.

XSL•FO
RenderX

Perceived Benefits

Multimedia Appendix 5 presents the descriptive statistics for the perceived benefit scores of HPV vaccination among child parents, categorized by gender. The analysis includes more female participants (n=363) than male participants (n=171). Fathers and male guardians had a slightly higher mean of the total benefit score (mean 28.96, SD 8.57) than mothers and female guardians (mean 27.41, SD 8.01). These data reveal a slight difference in perceptions of HPV vaccine benefits for male and female guardians. In the sample, 51% of parents recognized HPV vaccination as a significant means for preventing cervical and anal cancer. The overall mean perceived score of HPV was mean 27.9 (SD 8.22, range: 8.00 - 40.00).

Multimedia Appendix 6 showed that more than half of the parents believed that preventing cancer diseases by HPV vaccination is better than curing them, 51.8% (n=282). However, less than half (44.4%) believed the HPV vaccine is successful in preventing HPV infection for their daughters or sons. Some items of the perceived benefit of the HPV vaccine differed by parent sex. Differences were also noted in comparison to male parents. For instance, a smaller percentage of female parents agree that an HPV vaccine for their sons would help them stop worrying about anal and penile cancer compared to their male counterparts. Furthermore, a much smaller proportion of female parents felt the HPV vaccine was safe and prevented disease compared to their male counterparts.

The researcher found statistically significant relationships between the perceived benefit of the HPV vaccine or vaccination and a number of baseline sociodemographic characteristics of the study participants (refer to Table 5). The perceived benefit of the HPV vaccine or vaccination differed according to parents' sex and level of education. There was a greater proportion of male respondents with a high perception of HPV vaccine or vaccination benefits than females (n=69, 26.8% vs n=102, 36.8%; *P*=.02). However, there were more respondents with a bachelor's degree among those in the group with a high perception of the benefits of the HPV vaccine compared with those with a lower perception of benefit (n=133, 51.8% vs n=141, 50.9%; *P*=.02).

Table . Comparison between perceived benefits of HPV vaccination and demographic characteristics.

Demographic characteristics	Low (n=257)	High (n=277)	*P* value
Sex (Female), n (%)	148 (57.6)	140 (50.5)	.12
Parent sex (Male), n (%)	69 (26.8)	102 (36.8)	.02[a]
Child age, mean (SD)	13.42 (01.20)	13.47 (01.25)	.59
Parent age, mean (SD)	42.34 (05.27)	43.19 (6.08)	.09
Income, n (%)			.07
Less than US $2260	43 (16.7)	28 (10.1)	
US $2260 - 4800	118 (45.9)	131 (47.3)	
More than US $4800	96 (37.4)	118 (42.6)	
Marital status, n (%)			.21
Single	4 (1.6)	12 (4.3)	
Married	214 (83.3)	232 (83.8)	
Divorced	33 (12.8)	28 (10.1)	
Widowed	6 (2.30)	5 (1.80)	
Employment unemployment, n (%)	27 (10.50)	32 (11.60)	.80
Nationality (non-Kuwait), n (%)	16 (6.20)	23 (8.30)	.45
Education, n (%)			.02[a]
Primary education	9 (3.5)	2 (0.7)	
High school education	40 (15.6)	43 (15.5)	
Diploma	60 (23.3)	57 (20.6)	
Bachelor's degree	133 (51.8)	141 (50.9)	
Master's or PhD	15 (5.8)	34 (12.3)	

[a]Indicates statistical significance.

An independent *t* test—Test of Levene's (*P* >.05) was not significant; hence, equal variances are assumed for the independent *t* test—The results, categorized by gender based on total perceived benefit scores, were presented in Multimedia Appendix 7. This test has a *P* value less than the α-value of .05, so there is sufficient data to infer a significant difference between the benefit scores of male (fathers and male guardians) and female (mothers and female guardians) individuals (t_{532}=-2.045; *P*=.04).

XSL•FO
RenderX

Table 6 demonstrates unadjusted and adjusted results for the perceived benefit of HPV, categorized by median benefit score from logistic regression models.

Table . Logistic regression models illustrate the relationship between participants' characteristics and their perceived benefit from human papillomavirus (HPV) infection among adolescents.

Characteristics of adolescents	Perceived benefit of HPV vaccine or vaccination			
	Unadjusted odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Child age	1.04 (0.90 - 1.19)	.59	1.04 (0.89 - 01.21)	.57
Child or adolescent sex				
Male (reference)	__[a]	—	—	—
Female	0.75 (0.53 - 1.06)	.10	0.68 (0.47 - 0.99)	.04
Parent characteristics				
Age of respondent	1.03 (0.99 - 1.06)	.09	1.02 (0.99 - 1.06)	.17
Parent sex				
Female (reference)	—	—	—	—
Male	1.59 (1.100 - 02.30)	.01	1.37 (0.89 - 02.09)	.15
Nationality				
Kuwaiti (reference)	—	—	—	—
Non-Kuwaiti	1.36 (0.71 - 2.69)	.36	1.52 (0.75 - 3.15)	.25
Education				
Primary education (reference)	—	—	—	—
High school education	4.84 (1.16 - 33.01)	.05	3.50 (0.73 - 25.84)	.15
Diploma	4.28 (1.05 - 28.84)	.07	4.41 (0.93 - 32.55)	.09
Bachelor's degree	4.77 (1.20 - 31.67)	.05	4.96 (1.06 - 36.28)	.06
Masters or PhD	10.20 (2.29 - 72.48)	.006	9.47 (1.80 - 74.75)	.01
Monthly family income				
Less than 700 KD (reference)	—	—	—	—
700 - 1500 KD	1.70 (1.00 - 2.94)	.05	1.66 (0.92 - 3.03)	.09
More than 1500 KD	1.89 (1.09 - 3.29)	.02	1.65 (0.88 - 3.13)	.12
Employment status				
Employment (reference)	—	—	—	—
Unemployment	1.11 (0.65 - 01.93)	.70	01.41 (0.72 - 2.78)	.32
Marital status				
Single (reference)	—	—	—	—
Married	0.36 (0.10 - 1.06)	.08	0.27 (0.07 - 0.84)	.03
Divorced	0.28 (0.07 - 0.91)	.05	0.256 (0.06 - 0.89)	.04
Widowed	0.28 (0.05 - 1.38)	.13	0.17 (0.03 - 0.89)	.04

[a]Not applicable.

Approximately 52.0% (n=278) of parents perceived a high benefit of HPV vaccination. Parents with a female child demonstrated a lesser perception of the benefits of the HPV vaccine compared to those with a male child. Additionally, higher levels of parental education were strongly associated with greater perceived benefits.

Perceived Barriers

About 72.0% (n=388) of parents perceived that there are barriers to HPV vaccination for cervical and anal cancer prevention. The mean score for the perceived barrier to HPV vaccination

XSL•FO
RenderX

was 37.13, SD 8.22 (range: 12.00 - 60.00; Multimedia Appendix 8).

In Multimedia Appendix 9, up to 20% of parents who completed the survey believed that the stigma associated with receiving HPV vaccination and the perception that receiving the HPV vaccine is synonymous with promiscuity, according to the Kuwaiti culture, are key barriers to the uptake of the vaccine. Also, only 13% - 28% of participants acknowledged restrictions from their religion, lack of vaccine availability in public hospitals, affordability, unwillingness to pay for the vaccines, or fear of needles on the part of children as barriers to HPV vaccine uptake. There were no statistically significant differences in responses between male and female participants on these statements. However, they differed on the need for reassurance about the effectiveness and safety of the HPV vaccine. More males thought they needed reassurance about effectiveness (P=.006) and safety (P=.001).

Perceived barriers to the HPV vaccine or vaccination differed by sex of child, employment status, monthly family income, and level of education (see Table 7). We found a more significant proportion of parents with female children with the higher perception of barriers to the HPV vaccine or vaccination than those with male children (n=67, 45.9 vs n=221, 57.0; P=.03). Similar trends were observed for family income; monthly family income significantly differed between participants with high versus low perceived barriers to HPV infection (P<.01). Since the P value was more significant than the significance level α-value of 0.05, it can be concluded that there is no statistically significant difference in the barrier scores between male participants (fathers and male guardians) and female participants (mothers and female guardians). The results of the test were t_{532}=–1.362; P=.17. This suggests that gender does not significantly influence the perceived barriers related to HPV vaccination.

Table . Comparison between perceived barriers and demographic characteristics.

Demographic characteristics	Low (n=146)	High (n=388)	P value[a]
Sex (Female), n (%)	67 (45.9)	221 (57.0)	.03[b]
Parent sex (Male), n (%)	38 (26.0)	133 (34.3)	.09
Child age, mean (SD)	013.48 (1.21)	013.43 (1.230)	.70
Parent age, mean (SD)	64 (6.45)	84 (05.42)	.72
Monthly family income, n (%)			.01[b]
Under 700 KD	28 (19.2)	43 (11.1)	
700 - 1500 KD	71 (48.6)	178 (45.9)	
More than 1500 KD	47 (32.2)	167 (43.0)	
Marital status, n (%)			<.001[b]
Single	12 (8.2)	4 (1.0)	
Married	99 (67.8)	347 (89.4)	
Divorced	27 (18.5)	34 (8.8)	
Widowed	8 (5.5)	3 (0.8)	
Employment (unemployed), n (%)	29 (19.9)	30 (7.7)	<.001[b]
Nationality (non-Kuwaiti), n (%)	7 (4.8)	32 (8.2)	.24
Education, n (%)			<.001[b]
Primary education	7 (4.8)	4 (1.0)	
High school education	38 (26.0)	45 (11.6)	
Diploma	26 (17.8)	91 (23.5)	
Bachelor's degree	64 (43.8)	210 (54.1)	
Master's or PhD	11 (7.5)	38 (9.8)	

[a]indicates χ^2 test.

[b]Indicates statistical significance.

Table 8 demonstrates unadjusted and adjusted results for a perceived barrier of HPV from our logistic regression models. The sex of the child (female), parents' level of education, employment status, monthly family income, and marital status demonstrated a statistically significant correlation with a perceived barrier to HPV vaccination in the unadjusted logistic regression models. Parents with female children are likely to have a 1.6 times higher perception of the barrier to HPV vaccination compared to parents with male children (OR: 1.56, 95% CI 1.07 - 2.29; P=.02).

XSL•FO
RenderX

Table . Logistic regression models illustrate the relationship between participants' characteristics and their perceived barriers to human papillomavirus (HPV) infection among adolescents.

Characteristics of adolescents	Perceived barriers to HPV vaccination			
	Unadjusted odds ratio (95% CI)	*P* value	Adjusted odds ratio (95% CI)	*P* value
Child age	0.97 (0.83 - 01.13)	.70	0.98 (0.82 - 01.168)	.82
Child or adolescent sex				
Male (reference)				
Female	1.560 (1.070 - 02.29)	.02[a]	1.60 (1.03 - 2.47)	.04[a]
Parent characteristics				
Age of respondent	1.01 (0.97 - 1.04)	.72	1.01 (0.97 - 1.05)	.55
Parent sex				
Sex female (reference)				
Sex male	1.48 (0.98 - 2.29)	.06	1.89 (1.13 - 3.21)	.02[a]
Nationality				
Kuwaiti (reference)				
Non-Kuwaiti	1.78 (0.81 - 4.49)	.18	1.69 (0.70 - 4.67)	.27
Education				
Primary education (reference)				
High school education	2.07 (0.58 - 8.42)	.27	0.80 (0.18 - 3.74)	.77
Diploma	6.13 (1.72 - 24.93)	.006[a]	2.69 (0.63 - 12.57)	.19
Bachelor's degree	5.74 (1.68 - 22.51)	.007[a]	2.05 (0.49 - 9.31)	.32
Masters or PhD	6.05 (1.55 - 26.91)	.01[a]	2.15 (0.45 - 11.22)	.34
Monthly family income				
Less than 700 KD (reference)				
700 - 1500 KD	1.63 (0.94 - 2.82)	.08[a]	1.18 (0.60 - 2.26)	.62
More than 1500 KD	2.31 (1.30 - 4.11)	.004[a]	1.04 (0.50 - 2.11)	.91
Status of employment				
Employment (reference)				
Unemployment	0.34 (0.190 - 0.59)	<.001[a]	0.89 (0.44 - 1.83)	.74
Marital status				
Single (reference)				
Married	10.52 (3.58 - 38.27)	<.001[a]	9.39 (2.86 - 37.50)	<.001[a]
Divorced	3.78 (1.17 - 14.74)	.04[a]	3.94 (1.07 - 17.12)	.04[a]
Widowed	1.13 (0.18 - 6.52)	.89	1.03 (0.15 - 6.72)	.97

[a]Indicates statistical significance.

More than two-thirds of parents who responded to the survey alluded to barriers to HPV vaccination. The sex of the child and marital status are the factors that demonstrated significant associations with perceived barriers to HPV vaccination.

Perceived Self-Effect

Multimedia Appendix 10 provides descriptive statistics for the overall perceived self-efficacy score related to HPV among child participants. The mean self-efficacy score for all individuals was 23.51 (SD 6.03). The study had more women (n=363) than men (n=171). Dads and male caregivers scored

XSL•FO
RenderX

higher on the overall self-efficacy score (mean 25.09, SD 5.32) than moms and female caregivers (mean 22.77, SD 6.21).

Multimedia Appendix 11 shows that more than 40% of parents affirmed complete control over giving their son or daughter the HPV vaccine against anal, penile, or CC. However, about 33.7% (n=180) had good enough knowledge to decide on HPV vaccine uptake for their children. In addition, less than 40% (n=210) felt they could do something about their health by getting vaccinated against HPV. More than 40% (n=210) of parents affirmed complete control over giving their son or daughter HPV vaccine against anal, penile, or CC. However, about 33.7% (n=180) had good enough knowledge to decide on HPV vaccine uptake for their children. In addition, less than 40% (n=210) felt they could do something about their health by getting vaccinated against HPV.

Table 9 shows that parents with female children had a slightly greater percentage of high perceived self-efficacy than those with low self-efficacy (n=154, 59.0% vs n=134, 49.1%; *P*=.03). We found a greater proportion of male respondents with a high perception of perceived self-efficacy than females (n=58, 22.2 % vs n=113, 41.1%;, *P*<.001). Our analysis revealed a significant percentage difference in perceived self-efficacy among respondents of non-Kuwaiti nationality (n=11, 4.2% vs n=28, 10.3%; *P*=.01).

Table . Self-efficacy and baseline characteristics of respondents and children.

Demographic characteristics	Low (n=261)	High (n=273)	*P* value[a]
Sex of child, (Female), n (%)	154 (59.00)	134 (49.10)	.03[b]
Sex of parent, (Male), n (%)	58 (22.20)	113 (41.40)	<.001[b]
Child age, mean (SD)	13.47 (1.230)	13.42 (1.230)	.64
Respondent age, mean (SD)	42.57 (5.730)	42.99 (5.700)	.40
Income monthly, n (%)			.08
Under US $ 2260 $	36 (13.80)	35 (12.80)	
Between US $2260 and US $4800	133 (51.00)	116 (42.50)	
Over US $4800	92 (35.20)	122 (44.70)	
Marital status, n (%)			.20
Unmarried	6 (2.30)	10 (3.70)	
Wedded	212 (81.20)	234 (85.70)	
Divorce	37 (4.20)	24 (8.80)	
Widow	6 (2.30)	5 (1.80)	
Employment (unemployment), n (%)	30 (11.50)	29 (10.60)	.86
Nationality (non-Kuwaiti), n (%)	11 (4.20)	28 (10.30)	.01[b]
Level of education, n (%)			.001[b]
Primary education	11 (4.20)	0 (0.00)	
High school education	42 (6.1)	41 (15.0)	
Diploma	57 (21.8)	60 (22.0)	
Bachelor’s degree	136 (52.1)	138 (50.5)	
Masters or PhD	15 (5.7)	34 (12.5)	

[a]indicates χ^2 test.

[b]Indicates statistical significance.

Unadjusted logistic regression models, the sex of the parent (male) was found to have a 2.47 times higher perception of self-efficacy for HPV vaccination compared to female parents (OR: 2.47; 95% CI (1.70 - 3.63), *P*<.001). Adjusted models showed a significant association of sex of parent (male, 2.23; 95% CI (1.49 - 3.39), *P*<.001). Nationality with perceived self-efficacy to HPV vaccination in unadjusted logistic regression models showed significant predictors of perceived self-efficacy of HPV vaccination (2.60; 95% CI 1.29 - 5.55, *P*=.009), and in adjusted logistic regression models showed (3.11; 95% CI 1.50 - 6.86, *P*=.003) (Table 10)

XSL•FO
RenderX

Table . Logistic regression models illustrate the relationship between participants' characteristics and their perceived self-efficacy of human papillomavirus (HPV) vaccination to prevent HPV infection and its consequences among children.

Characteristics of adolescents	Perceived self-efficacy of vaccination to prevent HPV infection			
	Unadjusted odds ratio (95% CI)	*P* values	Adjusted odds ratio (95% CI)	*P* values
Child age	0.97 (0.84 - 01.11)	.64	0.98 (0.85 - 01.14)	.83
Child or adolescent age				
Male (Reference)	__[a]	—	—	—
Female	0.67 (0.48 - 0.94)	.02	0.69 (0.47 - 0.99)	.04
Parent characteristics				
Age of respondent	1.01 (0.98 - 1.04)	.39	1.00 (0.97 - 1.04)	.77
Parent sex				
Sex female (reference)	—	—	—	—
Sex male	2.47 (1.70 - 3.63)	<.001	2.23 (1.49 - 3.39)	<.001
Nationality				
Kuwaiti (reference)	—	—	—	—
Non-Kuwaiti	2.60 (1.29 - 5.55)	.009	3.11 (1.50 - 6.86)	.003
Monthly family income				
Less than US $2260 (reference)	—	—	—	—
US $2260 - 4800	0.89 (0.53 - 1.53)	.69	0.90 (0.51 - 1.59)	.72
More than US $4800	1.36 (0.79 - 2.34)	.26	1.39 (0.77 - 2.51)	.28
Employment status				
Employment (reference)	—	—	—	—
Unemployment	0.920 (0.53 - 01.57)	.75	01.09 (0.59 - 1.98)	.78
Marital status				
Single (reference)	—	—	—	—
Married	0.66 (0.22 - 1.81)	.43	0.57 (0.18 - 1.67)	.32
Divorced	0.39 (0.12 - 1.19)	.10	0.42 (0.13 - 1.34)	.15
Widowed	0.50 (0.10 - 2.36)	.38	0.46 (0.09 - 2.29)	.34

[a]Not applicable.

There was a significant difference in the perceived self-efficacy scores for fathers and male guardians and the mothers and female guardians, which means that males have more confidence in themselves as compared to females. The sex of the child (son for focus more than female daughter) and nationality were the significant predictors of perceived self-efficacy to HPV vaccination.

Cue to Action

Multimedia Appendix 12 provides descriptive statistics for the total cue-to-action score related to HPV vaccination with the parents of children. The overall mean score for perceived self-efficacy of the HPV vaccine was 23.34 (SD=7.19). This study was conducted with more women (n=363) and men (n=171). Fathers and male guardians scored less cue-to-action on average (mean 22.54, SD 7.32) than mothers and female guardians, with a higher mean score (mean 23.93, SD 7.07).

Multimedia Appendix 13 shows that up to 50% of parents believed that stimuli such as recommendations from health professionals, relatives, and health promotion by MOH would trigger the decision-making process to accept HPV vaccination for their children. More than half of respondents said they would get vaccinated if health professionals recommended the HPV vaccine for their children.

Parents' likelihood to cue to action for deciding to take the HPV vaccine differed by the child's sex, nationality, marital status, and education (Table 11). We found a slightly more significant percentage of parents with female children with a higher perception of cue to action for HPV vaccination than those with a lower perception of cue to action (n=123, 48.6% vs n=165, 58.7%; *P*=.02). Married participants were more likely to take action against HPV vaccination than participants with other marital statuses (*P*=.04).

XSL•FO
RenderX

Table . Comparison between chances to cue to action for human papillomavirus (HPV) vaccination and demographic characteristics.

Demographic characteristics	Low (n=253)	High (n=281)	*P* value[a]
Child sex (female), n (%)	123 (48.6)	165 (58.7)	.02[b]
Parent sex (Male), n (%)	86 (34.0)	85 (30.2)	<.001[b]
Child age, mean (SD)	13.51 (0.20)	13.39 (01.25)	.25
Parent age, mean (SD)	42.83 (05.28)	42.74 (6.09)	.85
Income monthly, n (%)			.68
Under US $2260	32 (12.6)	39 (13.9)	
US $2260 - 4800	123 (48.6)	126 (44.8)	
More than US $4800	98 (38.7)	116 (41.3)	
Marital status, n (%)			.04[b]
Single	4 (1.6)	12 (4.3)	
Married	210 (83.0)	236 (84.0)	
Divorced	36 (14.2)	25 (8.9)	
Widowed	3 (1.2)	8 (2.8)	
Employment (unemployed), n (%)	28 (11.1)	31 (11.0)	1
Nationality (non-Kuwait), n (%)	11 (4.3)	28 (10.0)	.02[b]
Education level (%)			<.001[b]
Primary education	9 (3.6)	2 (0.7)	
High school education	50 (19.8)	33 (11.7)	
Diploma	52 (20.6)	65 (23.1)	
Bachelor's degree	130 (51.4)	144 (51.2)	
Masters or PhD	12 (4.7)	37 (13.2)	

[a]indicates the χ^2 test.

[b]Indicates statistical significance.

Table 12 demonstrates unadjusted and adjusted results related to the likelihood of cue-to-action for uptake of HPV vaccination, categorized by a median cue-to-action score from logistic regression models. Our unadjusted logistic regression models showed that the sex of the child (female), level of education, nationality, and marital status demonstrated a statistically significant correlation with a child's (female), level of education, nationality, and marital status demonstrated a statistically significant correlation with a perceived cue to action to HPV vaccination uptake. Parents with female children were 1.5 times more likely to cue to action for HPV vaccination than parents with male children (OR: 1.50; 95% CI 1.07 - 2.12, *P*=.02). Similar trends of positive correlation were observed in the analysis. We found a significant increase in likelihood to cue to action for uptake of HPV vaccination for each step of education achieved among respondents compared to primary education (High School; OR: 4.84, 95% CI 1.16 - 33.01, *P*=.05; Diploma; OR: 5.63; 95%CI 1.38 - 37.96, *P*=.03; Bachelor's Degree; OR: 4.98; 95%CI 1.26 - 33.09, *P*=.04; Masters or PhD; 13.88 3.07 - 99.86, *P*=.002). Married respondents were less likely to cue to action for HPV vaccination uptake for their children and themselves (OR: 0.23; 95% CI 0.06 - 0.75, *P*=.02). Respondents of non-Kuwaiti nationality were more likely to cue to action for HPV vaccination uptake compared to Kuwaiti (OR: 2.43; 95% CI 1.22 - 5.21, *P*=.02).

XSL•FO
RenderX

Table . Logistic regression models illustrate the relationship between participants' characteristics and their cue to action regarding human papillomavirus vaccination in children.

Characteristics of adolescent	Cue to action			
	Unadjusted odds ratio (95% CI)	*P* values	Adjusted odds ratio (95% CI)	*P* values
Child age	0.92 (0.80 - 01.06)	.25	0.92 (0.79 - 01.07)	.27
Child or adolescent sex				
Male (Reference)	—	—	—	—
Female	1.50 (1.07 - 2.12)	.02[a]	1.34 (0.92 - 1.95)	.13
Parent characteristics				
Age of respondent	0.99 (0.97 - 1.03)	.85	1.01 (0.98 - 1.05)	.41
Parent sex				
Female (reference)	—	—		—
—— Male.	0.84 (0.58 - 01.21)	.36	0.87 (0.57 - 01.35)	.54
Nationality				
Kuwaiti (reference)	—	—	—	—
Non-Kuwaiti	2.43 (1.22 - 5.21)	.015[a]	1.90 (0.91 - 4.20)	.01[a]
Education				
Primary education (reference)	—	—	—	—
High school education	2.97 (0.71 - 20.29)	.18	2.08 (0.44 - 15.29)	.39
Diploma	5.63 (1.38 - 37.96)	.03[a]	5.06 (1.09 - 36.79)	.05[a]
Bachelor’s degree	4.98 (1.26 - 33.09)	.04[a]	4.65 (1.02 - 33.43)	.03[a]
Masters or PhD	13.88 (3.07 - 99.86)	.002[a]	15.58 (2.99 - 122.55)	.003[a]
Monthly family income				
Less than US $2260 (reference)				
US $2260 - 4800	0.84 (0.49 - 1.43)	.52	0.62 (0.34 - 1.12)	.12
More than US $4800	0.97 (0.56 - 1.66)	.92	0.51 (0.27 - 0.95)	.04[a]
Employment status				
Employment (reference)	—	—	—	—
Unemployment	0.99 (0.58 - 1.72)	.99	1.26 (0.64 - 2.50)	.51
Marital status				
Single (reference)	—	—	—	—
Married	0.37 (0.10 - 1.09)	.09	0.33 (0.09 - 1.03)	.07
Divorced	0.23 (0.06 - 0.75)	.02[a]	0.20 (0.05 - 0.69)	.02[a]
Widowed	0.89 (0.15 - 5.54)	.89	0.74 (0.12 - 4.82)	.74

[a]Indicates statistical significance.

The researcher observed significant correlations between level of education and marital status and respondents’ likelihood of acting on the cue to take up HPV vaccination. About 53% (n=283) of parents were more likely to act on the recommendation for HPV vaccination uptake after a recommendation from health care professionals.

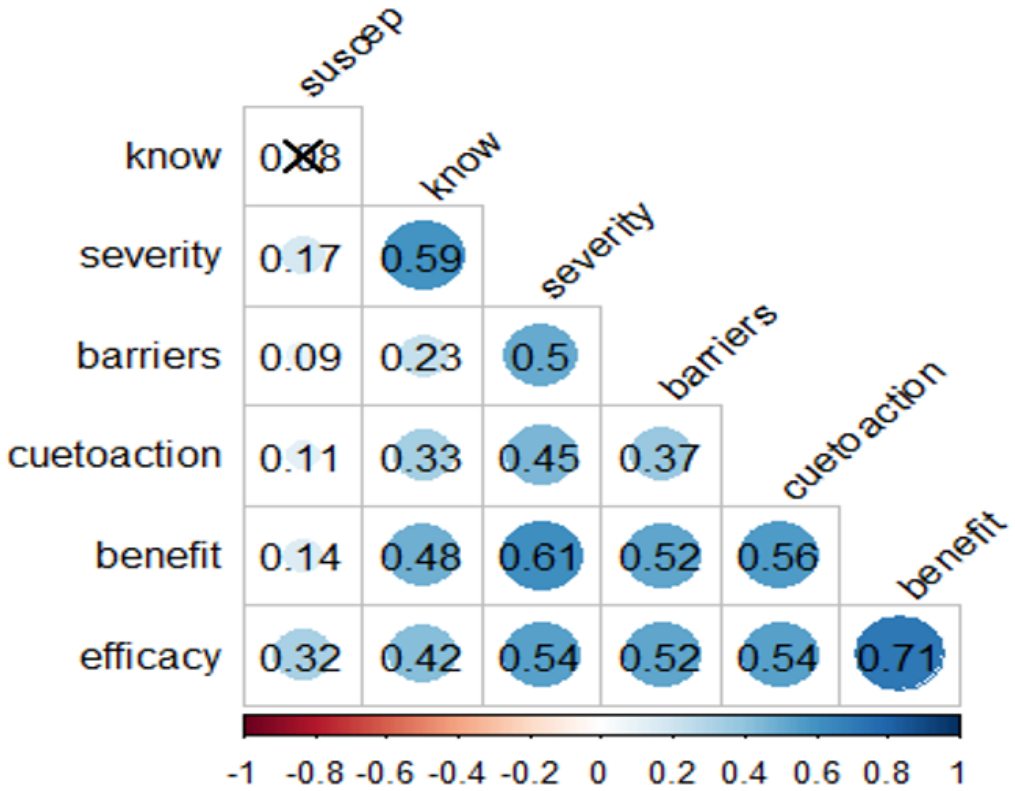
Correlation Between Participants and Knowledge

There was a positive correlation between all 6 constructs and knowledge of HPV, as the *P* value for all the correlation coefficients was less than 0.05. As one variable increases, the other also increases, and the degree of the relationship was statistically significant for all correlations except for susceptibility to HPV and knowledge of HPV. The strongest

XSL•FO
RenderX

correlation was seen between the benefit of HPV vaccination and self-efficacy to recommend or make decisions for HPV vaccination for children (r=0.71, P<.001). The significant but weakest correlation was barriers to HPV vaccination and perceived susceptibility to HPV (r=0.09, P<.05), as seen in Figure 2.

Figure 2. The matrix of correlation between the total knowledge score and the overall scores for each of the six HBM components.



Discussion

Principal Findings

The survey findings significantly contribute to the study's overall aim by providing insights into the factors that affect parents' acceptance of HPV vaccination in Middle Eastern countries. It is pertinent to mention that although the aim of the current thesis was specific to the State of Kuwait. The researcher targeted available literature for all Middle Eastern countries because Kuwait and other Middle Eastern countries share several common characteristics.

Our findings were consistent with other research using the HBM to assess the level of knowledge, beliefs, and acceptability of the HPV vaccine among parents of their children. For instance, one study stated that Kuwaiti university students were lagging in their knowledge of HPV and the HPV vaccine. In addition, the research reported that vaccine conspiracy beliefs were negatively correlated with their willingness to receive vaccination, demonstrating the role of misinformation and beliefs in decision-making around vaccination [23]. These findings agreed with our observations and underscore the need to correct information asymmetry and misconceptions to enhance HPV vaccination adoption. From this similarity, we can infer that the same problem in the Kuwait community, whether among university students or parents, a lack of knowledge is one of the obstacles that leads to decreased uptake of the HPV vaccine. On the other hand, in this quantitative study, when health professionals explain or advise parents to take the vaccine, they are more likely to have their children vaccinated.

Jordanian female undergraduates at health schools had a relatively high level of knowledge and attitudes toward HPV vaccination [24]. Their intention to vaccinate was derived directly from perceived susceptibility, perceived benefit, perceived barriers, and cues to action, which were consistent with the HBM. For example, in the same study, the participants explained the perceived severity of the vaccine. The study emphasized how the COVID-19 pandemic could have heightened concerns about vaccine safety and, in turn, further fueled vaccine hesitancy. Because of the high level of knowledge in Jordanian female undergraduates, which is different from our study, and the same intention regarding perceived severity in our study, even though there are several common characteristics between the 2 different countries, as both are Arabic countries, there are similarities and differences in HPV knowledge and perceived severity among different populations, respectively [24].

However, similar frameworks to explore parents' perspectives on HPV vaccination. For instance, one study found that Arabic-speaking mothers in Australia had low understanding and awareness of HPV and its vaccine, and that cultural and religious beliefs, perceived severity, perceived benefits,

XSL•FO
RenderX

perceived barriers, self-efficacy, and social norms influenced their acceptance of the vaccine [25]. For example, results found that cost, safety concerns, negative feelings, and knowledge gaps were the most common reasons given for the decline in HPV vaccine acceptability [25]. This explains that even some Arab participants living in other communities have the same perceived barriers in our study regarding the HPV vaccine, as well as low levels of knowledge. This means stigma, Kuwaiti culture is a key barrier, and low levels of expertise are similar for other Arabic participants, even in different communities.

Other studies found that Syrian mothers had low knowledge and negative attitudes toward HPV vaccination for their teen girls. Their main barriers were a lack of information, fear of side effects, and distrust of health authorities [26]. Also, the Syrian mothers are similar to the sociocultural context of the Kuwait mother community, sharing the same language and religion, which explains why the lack of knowledge and cultural barriers should be addressed to increase uptake of the HPV vaccine [26].

To provide a wider explanation regarding HBM in different areas of Middle Eastern culture and community, one study in Thailand showed Thai parents had a limited to moderate understanding of HPV and the HPV vaccination, and their acceptance of the vaccine was significantly influenced by perceived susceptibility, severity, advantages, hurdles, provocations, and values. The findings, along with similar research, suggest that parents' understanding, views, and preferences regarding HPV vaccination for their children vary significantly across nations and cultures. This necessitates treatments customized to the needs, interests, and cultural attributes of specific people [27]. That means, in some studies across different community cultures, we can interpret that when parents have high knowledge, they will have a high degree of understanding of HBM concepts in various ways, such as through interventions from the MOH or recommendations from health professionals. So, immunization in any community will increase directly because all parents will gain more information about the perceived benefits, severity, and other HBM concepts regarding the HPV vaccine, which will ultimately lead to herd immunity. One of the limitations of this study is that it was conducted only in government schools and did not collect data from private schools, with most of the students being non-Kuwaiti citizens. Therefore, these results are unlikely, given the knowledge and health beliefs of all parents in Kuwait.

Conclusions

Our study shares similarities and differences with other research that used the HBM or similar frameworks to investigate parents' knowledge, beliefs, and acceptance of the HPV vaccine for their children in various countries and cultures. One notable similarity is the consistent finding that parents' knowledge of HPV and the HPV vaccine tends to be low. In addition, the higher the knowledge level, the greater the perceived susceptibility, severity, benefits, and acceptance of the vaccine. Such a correlation highlights the need for interventions to build parents' knowledge and awareness about HPV and the HPV vaccine everywhere because it is a crucial component of improving vaccine acceptance and absorption. Based on our findings, some preliminary conclusions and implications for health promotion and policy formulation regarding HPV vaccination in Kuwait. To start with, it is crucial to address parents' lack of understanding of HPV and the HPV vaccine, especially among the less educated, lower-income, and Kuwaiti nationals. Comprehensive and culturally appropriate education efforts should be undertaken through schools, physicians, the media, and community organizations to create awareness and foster a better understanding of HPV and the importance of vaccination.

Acknowledgments

The authors would like to thank the librarians of the University of Bradford for their guidance and assistance in this systematic review. No generative AI was used in any portion of the manuscript generation. This work has been a part of the Doctoral Degree at the University of Bradford. However, a publishing grant was received from Linnaeus University, Sweden, as part of the University Library's research support.

Authors' Contributions

Conceptualization: AA, CE, CJ
Formal analysis: AA, CJ
Methodology: AA, CJ
Project administration: AA, CJ
Writing – original draft: AA
Writing – review & editing: AA, CJ, MSRJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Descriptive for total susceptibility score for participants (male/female).
[DOCX File, 14 KB - publichealth_v11i1e75818_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Responses to perceived susceptibility to HPV infection and related conditions.
[DOCX File, 16 KB - publichealth_v11i1e75818_app2.docx]

Multimedia Appendix 3
Descriptive for total severity score for participants (male and female).
[DOCX File, 15 KB - publichealth_v11i1e75818_app3.docx]

Multimedia Appendix 4
Response to the perceived severity of infection of HPV and associated conditions.
[DOCX File, 16 KB - publichealth_v11i1e75818_app4.docx]

Multimedia Appendix 5
Descriptive total benefit score for participants (male and female).
[DOCX File, 15 KB - publichealth_v11i1e75818_app5.docx]

Multimedia Appendix 6
Descriptive statistics for perceived benefits items.
[DOCX File, 15 KB - publichealth_v11i1e75818_app6.docx]

Multimedia Appendix 7
Difference between male and female participants' benefit score.
[DOCX File, 14 KB - publichealth_v11i1e75818_app7.docx]

Multimedia Appendix 8
Descriptive for total barrier score for participants (male and female).
[DOCX File, 15 KB - publichealth_v11i1e75818_app8.docx]

Multimedia Appendix 9
Descriptive statistics for perceived barrier items.
[DOCX File, 16 KB - publichealth_v11i1e75818_app9.docx]

Multimedia Appendix 10
Descriptive for total self-efficacy score for participants (male and female).
[DOCX File, 15 KB - publichealth_v11i1e75818_app10.docx]

Multimedia Appendix 11
Descriptive statistics for perceived self-efficacy items.
[DOCX File, 15 KB - publichealth_v11i1e75818_app11.docx]

Multimedia Appendix 12
Descriptive for total cue to action score for participants (male and female).
[DOCX File, 15 KB - publichealth_v11i1e75818_app12.docx]

Multimedia Appendix 13
Descriptive for cue to action items.
[DOCX File, 16 KB - publichealth_v11i1e75818_app13.docx]

References

1. Alexander KA, Giuliano AR. HPV—beyond cervical cancer (online resource center). Am J Med 2012 Jul;125(7):S1. [doi: 10.1016/j.amjmed.2012.03.005] [Medline: 22727241]
2. Basic information about HPV and cancer. Centers for Disease Control and Prevention (CDC). 2024. URL: https://www.cdc.gov/cancer/hpv/basic-information.html [accessed 2025-10-08]
3. Muñoz N, Bosch FX, de Sanjosé S, et al. Epidemiologic classification of human papillomavirus types associated with cervical cancer. N Engl J Med 2003 Feb 6;348(6):518-527. [doi: 10.1056/NEJMoa021641] [Medline: 12571259]

XSL•FO
RenderX

4. WHO guideline for screening and treatment of cervical pre-cancer lesions for cervical cancer prevention, second edition [internet]. World Health Organization (WHO). 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK572321/#ch1.s1 [accessed 2021-11-10]
5. The HPV vaccine: access and use in the US. Women's health policy, KFF. 2021. URL: https://www.kff.org/womens-health-policy/fact-sheet/the-hpv-vaccine-access-and-use-in-the-u-s/ [accessed 2025-12-04]
6. Alshahrani NZ, Alshahrani JA, Almushari BS, et al. Parental perspectives on human papillomavirus (HPV) vaccination in gulf cooperation council countries: a systematic review. Medicine (Baltimore) 2024 Oct 18;103(42):e40124. [doi: 10.1097/MD.0000000000040124] [Medline: 39432651]
7. Lubeya MK, Zekire Nyirenda JC, Chanda Kabwe J, Mukosha M. Knowledge, attitudes and practices towards human papillomavirus vaccination among medical doctors at a tertiary hospital: a cross sectional study. Cancer Control 2022;29:10732748221132646. [doi: 10.1177/10732748221132646] [Medline: 36214063]
8. Fernandes Q, Allouch S, Gupta I, et al. Human papillomaviruses-related cancers: an update on the presence and prevention strategies in the Middle East and North African Regions. Pathogens 2022 Nov 19;11(11):1380. [doi: 10.3390/pathogens11111380] [Medline: 36422631]
9. Rajkhowa P, Pundir P, Dsouza SM, Patil DS, Narayanan P, Brand H. Factors influencing HPV vaccine implementation in South Asia: a scoping review protocol. Syst Rev 2024 Jan 5;13(1):14. [doi: 10.1186/s13643-023-02415-x] [Medline: 38178265]
10. Tobaiqy MA, Mehdar SA, Altayeb TI, Saad TM, Alqutub ST. Parental knowledge, views, and perceptions of human papilloma virus infection and vaccination-cross-sectional descriptive study. J Family Med Prim Care 2023 Mar;12(3):556-560. [doi: 10.4103/jfmpc.jfmpc_1673_22] [Medline: 37122659]
11. Seven M, Güvenç G, Şahin E, Akyüz A. Attitudes to HPV vaccination among parents of children aged 10 to 13 Years. J Pediatr Adolesc Gynecol 2015 Oct;28(5):382-386. [doi: 10.1016/j.jpag.2014.11.005] [Medline: 26209868]
12. Ghojazadeh M, Naghavi-Behzad M, Azar ZF, Saleh P, Ghorashi S, Pouri AA. Parental knowledge and attitudes about human papilloma virus in Iran. Asian Pac J Cancer Prev 2012;13(12):6169-6173. [doi: 10.7314/apjcp.2012.13.12.6169] [Medline: 23464425]
13. Lin Y, Lin Z, He F, et al. Factors influencing intention to obtain the HPV vaccine and acceptability of 2-, 4- and 9-valent HPV vaccines: A study of undergraduate female health sciences students in Fujian, China. Vaccine (Auckl) 2019 Oct;37(44):6714-6723. [doi: 10.1016/j.vaccine.2019.09.026]
14. Dapari R, Li M, Chen X, et al. Factors influencing HPV vaccine acceptance among females in mainland China: A systematic review. Clin Epidemiol Glob Health 2024 Mar;26:101514. [doi: 10.1016/j.cegh.2024.101514] [Medline: 39345353]
15. Schedule of essential vaccination by age, state of kuwait, 2008. Kuwait Government Online. 2021. URL: https://e.gov.kw/sites/kgoenglish/Pages/CitizensResidents/Health/BasicVaccinationsInKuwait.aspx [accessed 2021-09-16]
16. Elissa N, Charbel H, Marly A, Ingrid N, Nadine S, Rachel A. Knowledge and perception of HPV vaccination among Lebanese mothers of children between nine and 17 years old. Reprod Health 2024 Mar 27;21(1):40. [doi: 10.1186/s12978-024-01764-7] [Medline: 38539219]
17. Al-Khaledi M, Lincoln M, McCabe P, Packman A, Alshatti T. The attitudes, knowledge and beliefs of Arab parents in Kuwait about stuttering. J Fluency Disord 2009 Mar;34(1):44-59. [doi: 10.1016/j.jfludis.2009.02.003] [Medline: 19500714]
18. Barath E, Rosner B. Fundamentals of Biostatistics: Biometric Society; 1992. [doi: 10.2307/2532371]
19. Saunders M, Lewis P, Thornhill A. Research Methods for Business Students: Pearson Education limited; 2019.
20. Potts J, Southard E. Teaching it forward: educating parents about HPV/HPV vaccine. J Dr Nurs Pract 2019 Apr 1;12(1):46-58. [doi: 10.1891/2380-9418.12.1.46] [Medline: 32745055]
21. Al Saad M, Ali Jadoo SA. Syrian mothers, why to accept or to refuse HPV vaccine for their teen girls. Journal of ideas in health (Online) 2018;1(1):7-13. [doi: 10.47108/jidhealth.Vol1.Iss1.3]
22. Allen JD, Coronado GD, Williams RS, et al. A systematic review of measures used in studies of human papillomavirus (HPV) vaccine acceptability. Vaccine (Auckl) 2010 May 28;28(24):4027-4037. [doi: 10.1016/j.vaccine.2010.03.063] [Medline: 20412875]
23. Alsanafi M, Salim NA, Sallam M. Willingness to get HPV vaccination among female university students in Kuwait and its relation to vaccine conspiracy beliefs. Hum Vaccin Immunother 2023 Jan 2;19(1):2194772. [doi: 10.1080/21645515.2023.2194772] [Medline: 37005342]
24. Sallam M, Al-Mahzoum K, Eid H, et al. Attitude towards HPV vaccination and the intention to get vaccinated among female University Students in Health Schools in Jordan. Vaccines (Basel) 2021 Dec 3;9(12):1432. [doi: 10.3390/vaccines9121432] [Medline: 34960177]
25. Netfa F, King C, Davies C, et al. Perceived facilitators and barriers to the uptake of the human papillomavirus (HPV) vaccine among adolescents of Arabic-speaking mothers in NSW, Australia: a qualitative study. Vaccine: X 2023 Aug;14:100335. [doi: 10.1016/j.jvacx.2023.100335]
26. Al Sad S, Pandit R, Alhashim N, Abdel-Rasoul M. Primary care providers' approaches to cervical cancer screening in Muslim females. Prev Med Rep 2023 Apr;32:102126. [doi: 10.1016/j.pmedr.2023.102126] [Medline: 36852309]

XSL•FO
RenderX

27. Grandahl M, Paek SC, Grisurapong S, Sherer P, Tydén T, Lundberg P. Parents' knowledge, beliefs, and acceptance of the HPV vaccination in relation to their socio-demographics and religious beliefs: a cross-sectional study in Thailand. PLoS ONE 2018;13(2):e0193054. [doi: 10.1371/journal.pone.0193054] [Medline: 29447271]

Abbreviations

CC: cervical cancer
HBM: health belief model
HPV: human papillomavirus
OR: odds ratio

Edited by A Mavragani, T Sanchez; submitted 11.04.25; peer-reviewed by K Li, N Rancic; revised version received 15.10.25; accepted 20.10.25; published 09.12.25.

Please cite as:
Abuzoor A, Jabin MSR, Eshareturi C, Nesbitt R, Jonker C
Using the Health Belief Model to Examine Parental Knowledge and Health Beliefs About Human Papilloma Virus (HPV) and iHPV Vaccine in Kuwait: Cross-Sectional Survey Study
JMIR Public Health Surveill 2025;11:e75818
URL: https://publichealth.jmir.org/2025/1/e75818
doi:10.2196/75818

© Ahmad Abuzoor, Md Shafiqur Rahman Jabin, Cyril Eshareturi, Rae Nesbitt, Cor Jonker. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 9.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Social Media Marketing to Improve Retention of Children in the Special Supplemental Nutrition Program for Women, Infants, and Children: Implementation Study

Eriko M Robinson[1], RDN, MS-MPH; Carla J Sabugo[1], RDN, MPH; Carolina Crisafi[1], BA, IBCLC; Susan M Gross[2], RDN, PhD; Yunhee Kang[3], PhD; Elisabet V Eppes[2], MPH; Marycatherine Augustyn[2], PhD; Karen Castellanos-Brown[4], PhD; David M Paige[2], MD; Laura E Caulfield[3], PhD

[1]Florida Department of Health in Miami-Dade County, WIC and Nutrition Program, Miami, FL, United States

[2]Department of Population, Family and Reproductive Health, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD, United States

[3]Department of International Health, Center for Human Nutrition, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Room W2041, Baltimore, MD, United States

[4]Supplemental Nutrition and Safety Program Research and Analysis Evaluation Branch, Food and Nutrition Service, USDA, Alexandria, VA, United States

Corresponding Author:
Laura E Caulfield, PhD
Department of International Health, Center for Human Nutrition, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Room W2041, Baltimore, MD, United States

Abstract

Background: Many eligible infants and children do not participate in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC); coverage declines throughout the preschool period of eligibility. National and state-level social marketing campaigns promote the value of WIC and increase enrollment and participation. Local contextualization and targeting of materials may increase effectiveness, considering the diversity of families eligible for the program. However, there are few examples of such approaches and their impact.

Objective: This study evaluated the impact on child retention of a locally contextualized and targeted social media marketing campaign directed to WIC-eligible families living in the minority-majority population of Miami-Dade County, Florida.

Methods: The digital marketing campaign geographically targeted low-income families with young children with customized static image and video advertisements on Facebook and Instagram, and a bilingual Google Ads campaign. It was implemented in 2 of 15 clinics operated by the Miami-Dade WIC local agency from May 2020 through April 2021. A before and after evaluation used program administrative data to compare the outcomes for infants and children in 2 innovation clinics (n=6162) with 11 comparison clinics (n=41,074) during a baseline period (2019 calendar year) and the implementation period (n=5636 and n=38,241, respectively). Outcome measures included recertification (re-enrollment during a period), retention (active in the program at the end of a period), and participation (household continuous benefit issuance defined as 11 out of 12 mo). Impact was assessed following cluster-adjusted propensity score weighting and difference-in-difference modeling. Household continuous benefit issuance was estimated in households with only an infant or a child.

Results: Overall, 1,994,170 people were exposed to the campaign advertisements; 16.68% engaged with an advertisement. There were 22,983 unique visits to the local program website, 69.6% of which were acquired directly from the campaign. Four of the 5 top-performing advertisements were locally tailored messages and in Spanish. The change in recertification over time was 5.2% points (95% CI 3.4%-7.1%), greater for those in the innovation group than those in the comparison group. For retention and continuous benefit issuance, the absolute difference in change was 5.5% points (95% CI 3.7%-7.3%), and 6.6% points (95% CI 3.5%-9.7%), respectively. Differences in change over time associated with the innovation were qualitatively stronger for infants than for children; the difference in change for recertification was 7.6% points (95% CI 5.1%-10.1%) for infants and 4.0% points (95% CI 2.2%-5.9%) for children.

Conclusions: Engaging low-income families with young children through a locally contextualized targeted media marketing campaign can improve retention of children in WIC.

(JMIR Public Health Surveill 2025;11:e77172) doi:10.2196/77172

XSL•FO
RenderX

KEYWORDS

social media; infants; children; Special Supplemental Nutrition Program for Women, Infants, and Children; WIC; program evaluation

Introduction

The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) was established by the US Department of Agriculture (USDA) to improve the health of low-income women and children by providing nutritious supplemental foods, education, breastfeeding support, and referrals to health and social service programs [1]. Among income-eligible individuals, participation in WIC is associated with improved birth outcomes, child nutrition status, and child health outcomes, all of which have impacts lasting into adulthood [2,3].

Despite the documented benefits, participation and retention in the WIC program has been a national concern, particularly for children. Nationally, in 2019, it was estimated that 83.8% of eligible infants participated in WIC, but participation declined after the age of 1 year to 56.6%, with declines to 42.8% and 38% to reach only 22% of eligible children at the age of 4 years [4]. Research has identified common barriers to program enrollment and retention, which include lack of knowledge of the program and of eligibility requirements, social stigma, language barriers, as well as issues in service quality [5-10]. To address service quality barriers, the USDA has launched initiatives to modernize WIC, and a national media campaign was introduced to promote enrollment and retention [11,12].

Beyond a national campaign, there is a need to design, implement, and evaluate locally contextualized marketing approaches to enhance WIC participation and retention considering the diversity of families eligible for WIC [4]. This project was conducted by the Miami-Dade WIC local agency, which serves Miami-Dade County, a large and ethnically diverse metropolitan county in Florida. According to the US census, in 2020, there were 2,701,767 residents, of whom 14.1% were living in poverty [13]. Overall, 69% of the Miami-Dade County population identifies as Hispanic or Latino, and 3.7% identify as Haitian or Kreyol. Approximately 55% are foreign-born, and of those, 43% do not hold US citizenship. Three-quarters of the Miami-Dade population speak a language other than English in their home. Building on the national campaign and the USDA-supported opportunity provided by the Hopkins Participant Research Innovation Laboratory for Enhancing WIC Services, we designed and tested an integrated media marketing campaign that digitally and geographically targeted WIC-eligible families with customized static image and video advertisements on Facebook and Instagram, as well as a bilingual Google Ads campaign. Based on consumer buying behavior theory [14], it was hypothesized that a digital media campaign aimed at driving users to the Miami-Dade WIC website and appointment call center and increasing engagement and brand awareness could have a positive impact on child retention by "converting" the passive WIC enrollee or family that does not participate in benefit issuance, is late for recertification, or is terminated from the program after failure to recertify. Here we report on the implementation of the integrated media marketing campaign and the evaluation of its impact on the child recertification, benefit issuance, and retention in WIC.

Methods

Setting: The Florida Department of Health in Miami-Dade WIC Program

The Miami-Dade WIC Program has 15 service locations throughout the county with a total caseload of more than 50,000. Most clients live in the 4 to 5 zip codes surrounding the physical location of each clinic. For this project, 2 innovation clinics were selected that shared clients from the same zip codes and were similar in sociodemographic make-up. Total enrollment in WIC for these 2 sites in February 2019 was approximately 10,250. A total of 9 zip codes represented >75% of the caseload for these 2 sites, and of these, 7 were specific to the innovation sites and had limited overlap with clients that received services at other Miami-Dade WIC sites. These 7 ZIP codes were selected as target zip codes. Of the 13 remaining clinics, 2 clinics served a substantial number of clients from the same target zip codes; to reduce the likelihood of client crossover, these clinics were removed from the study. Because only small numbers of clients at the remaining 11 WIC clinics resided in the targeted zip codes, these clinics formed the comparison group.

Social Media Marketing Framework

The integrated digital marketing campaign was developed as an innovative tool to retain and recruit potentially eligible families with infants and children aged 1 to 3 years. The overall design process and campaign materials are described elsewhere [15], and key aspects of the approach are described in this section. The framework integrated different methods to reach the target audience across different social media platforms in a coordinated fashion over time. The target audience received the campaign via four methods: (1) behavioral targeting, (2) social media marketing, (3) search engine marketing, and (4) mobile geo-precise strategy. Behavioral targeting refers to the use of historical behavior to customize the presentation of advertisements. For example, previous browsing history and interaction on advertisements for "picky eating" toddler play, toys for children aged 1 year, and potty training meet the target profile for parents of children aged 1 to 3 years. Layering income and sociodemographic data with behavioral data helps to identify existing and potential WIC clients. Marketing through social media platforms such as YouTube, Facebook, and Instagram targets users and their social circles, and the opportunity for sharing and "likes" increases engagement and allows for connections with other digital users, substantially increasing the reach of the campaign. This project used paid search advertisements on Google; we tested and used unique keywords specific to the project target audience (caregivers and parents of children aged 1-3 years, cultural and Spanish language keywords, and keywords specific to the geographic community and location). Finally, mobile geotargeting involves the use of privacy-compliant location data from cell phones, iPads, or

XSL•FO
RenderX

computer IP addresses to present advertising. Users who lived, worked, or moved throughout the geotargeted areas were presented with WIC retention advertisements if they met all other criteria (either via social media or search engine).

Changes were made to the Miami-Dade WIC website and content to increase visibility and improve rankings in search engine results. All website content in English was professionally translated to both Spanish and Creole, the primary languages spoken in Miami-Dade County. Priority was placed on ensuring ease of access to the website (a short domain was purchased in lieu of the long state URL) as well as ease of use while on the website (appropriate reading level, content placement, navigation, accessibility of downloads, and contact information).

Media Content

The Miami-Dade WIC Program obtained the rights to use and curate digital media content from the existing national campaign and developed custom creative content for the local campaign. A local WIC Advisory Committee made up of a diverse group of staff was established to provide insights, guidance, and recommendations for selected campaign messaging and content, to ensure the digital media campaign would be culturally, visually, and linguistically sensitive and meet the unique needs of the Miami-Dade population. A guide was created to provide direction on language, relatable imagery, and messaging needs, and further defined culturally acceptable material specific to the Miami-Dade community. All selected social media campaign material was also formally vetted and approved by USDA prior to posting. The Google Ads campaign keywords were developed with input from the vendor, existing keywords from the national campaign, and the WIC Advisory Committee.

Implementation Timeline

The social media campaign was implemented over a 1-year period, beginning in May 2020, in 4 batches, with each batch consisting of 6 static advertisement posts and 2 videos. Batches 1 and 3 presented custom creative digital content, whereas batches 2 and 4 presented customized digital content from the existing national campaign. Performance was reviewed after each batch and edits to the content and advertisement strategy were made accordingly. For example, after reviewing language performance in batch 1, all new posts included Spanish translation in the social media caption incorporating a Hispanic dialect specific to the Caribbean and Central American population served. The Google Ads campaign began in May 2020, and keyword performance was monitored monthly. Facebook and Instagram advertisements were launched in mid-June 2020. The marketing vendor used software, internet platforms, and algorithms to behaviorally target advertisements. Targeting was also accomplished by testing various objectives, based on the desired outcome or goal for an advertisement. The campaign also tested various calls to action on the advertisement posts, which encouraged users to take specific actions on a social media post. After clicking the ad, the user would either be redirected to the Miami-Dade WIC website *or* be prompted to call the Miami-Dade WIC call center (click to call). In addition, performance based on content language, social media placement, run schedules, and creatives was assessed for each batch. Strategies that performed better were implemented in the upcoming batch.

The pandemic began between March 2020 and April 2020, and all Miami-Dade WIC clinics physically closed due to the COVID-19 response. Select sites remained open for limited in-person needs such as benefit card issuance, direct distribution issuance, and services for clients without access to email or phones. All other services were shifted to a virtual model which allowed Miami-Dade WIC to meet capacity and the needs of 100% of scheduled clients during the closures. In October 2020, all Miami-Dade WIC clinics reopened for in-person services as a complement to virtual services which continued to be an option for clients. The USDA federal waivers for COVID-19 remained in place through the end of the national emergency declaration (May 2023).

Evaluation Design and Methods

The evaluation was designed to assess the performance of the innovative tool, and to assess the impact of the tool on relevant child outcomes in WIC.

Process Evaluation

Marketing vendor data were used to assess the campaign's reach, engagement, and conversion as key performance indicators (KPIs) [16]. Reach was defined as the number of unique users the advertisement is exposed to, and impression was defined by the number of views. Engagement occurs when a user likes, saves, shares, or comments on the post, and this is measured as a rate (engagements over impressions). When a user responded to a call-to-action and clicked on an advertisement either on social media or in the Google search engine, they were redirected either to the local agency website or the Miami-Dade WIC call center (depending on the predetermined media strategy). Total clicks, the click-through rate (CTR), and the share of website traffic attributable to the campaign measure conversion.

The original process evaluation plan also included in-clinic exit surveys to evaluate direct WIC client engagement with the campaign. When WIC services transitioned to a virtual service model due to the pandemic, the client engagement survey was necessarily moved to a virtual client-facing survey platform. The survey was administered in Spanish, Creole, or English, and clients were asked about where they had heard about WIC and whether they were familiar with Miami-Dade WIC social media pages. They were also asked if they had seen WIC advertisements on their cell phones or computers in the last month, and if so, they were asked to recall the platforms where they saw them (social media and Google). A competition was used to motivate staff to promote the survey during virtual certification appointments (via email link), but no client incentive was offered. Because the overall participant response rate was <5%, survey results are not presented here.

Impact Evaluation

The impact evaluation focused on whether the integrated marketing tool increased the retention of enrolled infants and children aged 1 to 3 years. Program administrative data were obtained on cohorts of child WIC participants in innovation and

XSL•FO
RenderX

comparison sites during a baseline period (prior to the innovation) and during the period when the innovation was implemented. This design, as previously described in Chancay et al [17], allowed for the evaluation of differences in outcomes between a baseline period and the implementation period, adjusting for changes over time common to both groups and group differences that do not vary over time. The baseline period was the 2019 calendar year (designated T1) and the implementation period was May 2020 through April 2021 (T2). Participant characteristics were also obtained to understand group differences and changes over time, and for use as covariates during analyses and included child age, sex, race and ethnicity, household size, number of children in WIC, primary language other than English, need for translation services, participation in other services (Medicaid, Supplemental Nutrition Assistance Program [SNAP], and Temporary Assistance for Needy Families).

Three outcomes were examined: recertification, retention, and benefit issuance. Data were obtained on WIC visits, recertification, and benefit issuance, for all infants (<12 mo) and children aged 12 to 47 months who were active in WIC in the innovation and comparison groups at the beginning of each period. Over a year-long period, each participant would have multiple visits, and benefits would be issued to them by WIC. In addition, the certification period for each child would end, and they would need to be recertified to continue in the WIC program. Through this approach, we could assess WIC recertification as an indicator of retention in cohorts of children and evaluate changes in the outcomes over time for the innovation group as compared to changes for the comparison group. To prevent censoring of outcomes for children whose certification period ended during months 11 or 12 of a period, data were obtained for an additional 2 months (that is, 14 mo of data were collected for each period), to identify those children who did recertify in a timely fashion, but after the end of the period. The WIC State agency also identified the participation status of each child at the end of a 12-month period as active, moved to another agency, inactive, or terminated; this status variable was used to define retention. Household WIC participation was also evaluated by continuous benefit issuance, defined as 11 to 12 months of benefit issuance over each 12-month period.

Statistical Analysis

Descriptive analyses were conducted to describe the digital marketing campaign KPIs and to assess the comparability of participants in the innovation and the comparison groups within each period. Differences in performance and in participant characteristics were evaluated using 1- or 2-tailed *t* tests or chi-square tests as appropriate. We documented characteristics with a significant percentage of missing values (>10%), which might limit their usefulness during analysis. No values were imputed because we deemed it not necessary.

A difference-in-difference (DID) approach was used to assess program impact, which involved estimation of the changes over time in each outcome in the innovation versus the comparison group. Analyses were conducted for the overall sample as well as separately for infants and children. Propensity score weighting (PSW) was used to support adjustment for differences in participant characteristics between the innovation and comparison groups at each period (T1 and T2) as well as differences across the 2 periods. Weights were estimated using multinomial logistic regression in which observations were weighted as compared to the characteristics of those in the innovation group during T1 [17]. To evaluate the balance in participant characteristics achieved through weighting, we compared the absolute standardized differences (ASD) for the means of each participant characteristic before and after the weighting. This involved comparing the balance achieved through weighting for the innovation group over time (at T1 and T2), the innovation group at T1 and the comparison group at T1, and the innovation group at T1 with the comparison group at T2. These steps were repeated with samples restricted to infants and children. Final adjusted models for recertification and retention were also adjusted for clustering of households within clinics and of children within households. To fully present the results, we provide the crude and regression-adjusted means (with 95% CIs) for each outcome and the impact estimate as percentage points from each of two DID models: (1) DID (crude, unweighted) and (2) PSW-DID (adjusted) [18,19]. To evaluate household continuous benefit issuance, we restricted analyses to households with only 1 child WIC participant. This allowed us to consider results overall, for households with an infant, and for those with a child. In these analyses, adjustment was also made for clustering of households within clinics. The PSW step focused on household or caregiver characteristics.

Ethical Considerations

The study design and methods were evaluated and determined to be public health practice and not research by the Florida State Department of Health Ethics and Human Research Protection Program. Deidentified administrative data were sent to Hopkins Participant Research Innovation Laboratory for Enhancing WIC Services for the impact evaluation analyses, which was deemed not to be research by the institutional review board at the Johns Hopkins Bloomberg School of Public Health.

Results

Process Evaluation

A total of 1,994,170 people were exposed to campaign advertisements, and 332,580 (16.68%) engaged with an advertisement (Google Analytics). There were 22,983 unique visits to the local WIC website, 15,996 (69.60%) of which were acquired directly from the campaign and represent the conversion rate. According to Google Analytics, the Google Ads campaign drove the highest number of users to the WIC website (13,149/15,996, 82.20%) followed by Facebook (2367/15,996, 14.80%). The mean number of impressions per month for the implementation period was 2330 (SD 209), and the average CTR was 20.95% (SD 1.6%). The top 5 performing key words (in terms of clicks to the website) in the Google Ads campaign were (1) “food stamps” (1243/8235, CTR=15%), (2) “WIC program” (909/2472, CTR=37%), (3) “wic” (713/2357, CTR=30%), (4) “comida gratis” (free food; 366/2324, CTR=16%), and (5) “EBT” (710/5283, CTR=13%).

XSL•FO
RenderX

Across all campaigns (Table 1), a greater share of impressions, views, and clicks were attributable to Facebook as compared to other social media platforms (Instagram, YouTube, and Audience Network 3rd party advertisements). Among the subset of advertisements run to direct traffic to the website or with a click to call or action, Facebook campaigns yielded a significantly higher overall CTR of 0.39% (4194/108,422) compared to 0.19% (1670/872,854) for all other platforms (P<.001). The average CTR for Facebook did not differ from the other platforms (P=.06).

Table . Overall performance of social media platforms during the campaign targeted to Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)-eligible families in Miami-Dade, Florida, May 2020-April 2021. Results are presented by share of impressions, clicks and views, and average click-through rate.

Platform	Impressions (n=1,959,276)	Clicks (n=5864)	Engagement or views (n=319,863)	Average CTR[ab] (SD)
Facebook, n (%)	1,086,422 (55.45)	4194 (71.52)	130,171 (40.70)	0.51 (0.46)[c]
Instagram, n (%)	421,156 (21.50)	933 (15.91)	10,615 (3.32)	0.39 (0.22)
YouTube, n (%)	450,370 (22.99)	734 (12.52)	177,739 (55.56)	0.17 (0.05)
Audience Network, n (%)	1328 (0.07)	3 (0.05)	1338 (0.42)	0.23[d]

[a]CTR: click-through rate. This is the average CTR (%) for advertisements running in traffic objective and click-to-call action on each platform.

[b]The average CTR for all campaigns is 0.47 (SD 0.41).

[c]Difference from other platforms (P=.06).

[d]This is the CTR and not average CTR.

Custom creative advertisements performed significantly better than existing national campaign advertisements (CTR of 0.38% vs 0.23% CTR; P<.001). Eligibility-themed content performed best with an average CTR of 0.49%, which was significantly greater than the other 3 themes (0.22% - 0.27% CTR; Table 2).

Four of the 5 top-performing advertisements ran under the click-to-call or traffic objective, were custom creative type, and ran in the Spanish language. All the top-performing advertisements ran on Facebook (Table 3).

Table . Social media posts by theme, total clicks, and average click-through rate (CTR) as part of a social media campaign targeted to Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)-eligible families in Miami-Dade, Florida, May 2020-April 2021.

Theme	Posts, n (%)[a]	Total clicks, n (%)	Average CTR (SD)
WIC added-value	46 (37.1)	2556 (43.59)	0.25 (0.20)
COVID-19	4 (3.2)	112 (1.91)	0.22 (0.12)
WIC eligibility	16 (12.9)	2027 (34.57)	0.49 (0.56)[b]
WIC misconceptions	18 (14.5)	1169 (19.93)	0.27 (0.21)

[a]Count includes individual advertisements run on multiple platforms and objectives as separate “posts.”

[b]Different from average CTR from other posts (P=.003).

XSL•FO
RenderX

Table . Top performing ads on Facebook during the campaign targeted to Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)-eligible families in Miami-Dade, Florida, May 2020-April 2021. Performance is evaluated by click-through rate (CTR).

Advertisement title	Batch	Objective	Type	Theme	Language	Clicks/impressions	CTR (%)
Medicaid or SNAP[a]	3	Click to call	CC[b]	Eligibility	Spanish	786/44,454	1.77
Medicaid or SNAP—RE-RUN	4	Click to call	CC	Eligibility	Spanish	329/21,064	1.56
Welcome to WIC Video	3	Engagement	US[c]	Eligibility	English	340/23,995	1.42
Missing Meals	1	Traffic	CC	Added value	Spanish	168/15,566	1.08
Stay on WIC until your child is 5	1	Traffic	CC	Misconception	Spanish	217/26,056	0.83

[a]SNAP: Supplemental Nutrition Assistance Program

[b]CC: custom creative.

[c]US refers to the national campaign.

Impact Evaluation

Shown in Table 4 are characteristics of the WIC participants in the innovation and comparison clinics during the baseline period (T1) and implementation period (T2). There were 3 potentially important differences between the innovation and comparison groups in Miami-Dade WIC. The innovation clinics at both T1 and T2 had fewer Black or African American participants (430/6162, 6.98% at T1 and 350/5636, 6.21% at T2) than the comparison clinics (12,793/41,074, 31.15% at T1 and 11,736/38,241, 30.69%), a greater number of White participants (5903/6162, 95.80% at T1 and 5411/5636, 95.97% at T2) than the comparison clinics (28,885/41074, 70.32% at T1 and 27,091/38,241, 70.84% at T2), and a greater number of Hispanic participants (5941/6162, 96.41% at T1 and 5496/5636, 96.10% at T2 in innovation clinics vs 28,625/41,074, 69.69% at T1 and 69.87 at T2 (26,719/38,241) in comparison clinics). There was a higher proportion of participants whose household language was something other than English (5031/6162, 81.64% at T1 and 4419/5636, 78.41% at T2 in the innovation clinics vs 20,823/41,074, 50.70% at T1 and 19,026/38,241, 49.75% at T2 in the comparison clinics). The only notable difference in participant characteristics within each group over time was a decline in SNAP participation. At T1, 49.60% (20,374/41,074) of the comparison group and 57.43% (3539/6162) of the innovation group participated in SNAP, but at T2, 37.95% (14,513/38,241) of the comparison group and 44.77% (2523/5636) of the innovation group were SNAP participants.

XSL•FO
RenderX

Table . Demographic characteristics of children 0 - 3 at Miami-Dade Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) comparison and innovation clinics at baseline period (T1) and implementation period (T2).

Characteristics	T1			T2		
	Comparison (n=41,074), n (%)	Innovation (n=6162), n (%)	*P* value[a]	Comparison (n=38,241), n (%)	Innovation (n=5636), n (%)	*P* value
Participant category			<.001			<.001
IBE[b]	1491 (3.63)	256 (4.15)		1401 (3.66)	229 (4.06)	
IBP[c]	5750 (14.00)	1048 (17.01)		5588 (14.61)	1121 (19.89)	
IFF[d]	7256 (17.67)	771 (12.51)		6624 (17.32)	624 (11.07)	
C1[e]	11,055 (26.92)	1643 (26.66)		10,162 (26.57)	1503 (26.67)	
C2[f]	8367 (20.37)	1388 (22.53)		7833 (20.48)	1172 (20.80)	
C3[g]	7155 (17.42)	1056 (17.14)		6633 (17.35)	987 (17.51)	
WIC participants			<.001			<.001
1	20,873 (50.82)	3327 (53.99)		20,576 (53.81)	3325 (59.00)	
2	6579 (16.02)	798 (12.95)		6673 (17.45)	964 (17.10)	
≥3	658 (1.60)	68 (1.10)		693 (1.81)	58 (1.03)	
Missing	12,964 (31.56)	1969 (31.95)		10,299 (26.93)	1289 (22.87)	
Race[h]						
American Indian or Alaska Native	58 (0.14)	14 (0.23)	.107	76 (0.20)	9 (0.16)	.534
Asian	383 (0.93)	32 (0.52)	.001	323 (0.85)	26 (0.46)	.002
Black or African American	12,793 (31.15)	430 (6.98)	<.001	11,736 (30.69)	350 (6.21)	<.001
Hispanic	28,625 (69.69)	5941 (96.41)	<.001	26,719 (69.87)	5416 (96.10)	<.001
Native Hawaiian or Other Pacific Islander	84 (0.20)	2 (0.03)	.003	82 (0.21)	0 (0.00)	.001
White	28,885 (70.32)	5903 (95.80)	<.001	27,091 (70.84)	5409 (95.97)	<.001
Enrolled						
TANF[ii]	569 (1.39)	63 (1.02)	.021	423 (1.11)	47 (0.83)	.064
SNAP[j]	20,374 (49.60)	3539 (57.4)	<.001	14,513 (38.0)	2523 (44.77)	<.001
Medicaid	38,362 (93.40)	5919 (96.1)	<.001	35,050 (91.7)	5296 (93.97)	<.001
Primary language other than English	20,823 (50.70)	5031 (81.7)	<.001	19,026 (49.8)	4419 (78.41)	<.001
Ever breastfed			<.001			<.001
Yes	32,042 (78.01)	5074 (82.34)		32,617 (85.3)	4918 (87.26)	
No	5077 (12.36)	596 (9.7)		4730 (12.4)	551 (9.78)	
Missing	3955 (9.64)	492 (7.98)		894 (2.34)	167 (2.96)	
Household size			<.001			<.001
0 - 4	27,778 (67.63)	4441 (72.07)		26,505 (69.31)	4084 (72.46)	
≥5	12,982 (31.61)	1694 (27.49)		11,422 (29.87)	1511 (26.81)	
Missing	314 (0.76)	27 (0.44)		314 (0.82)	41 (0.73)	

[a]Difference between groups within a time period.

[b]IBE: infant, exclusive breastfeeding.

[c]IBP: infant, partial breastfeeding.

[d]IFF: infant, formula feeding.

XSL•FO
RenderX

[e]C1: child category 1 (aged 1 year).

[f]C2: child category 2.

[g]C3: child category 3.

[h]Participants can respond to more than 1 category, so the total percentage may be >100.

[i]TANF: Temporary Assistance for Needy Families.

[j]SNAP: Supplemental Nutrition Assistance Program.

PSW was successful in balancing differences between innovation and comparison groups and differences over time (Table S1 in Multimedia Appendix 1). After PSW, the ASDs were all reduced below 0.10, and the 3 mean ASDs were reduced to 0.01. This was also true for the analyses at the household level to evaluate continuous benefit issuance.

As shown in Table 5, the integrated media marketing campaign was associated with greater recertification and retention of children in WIC. Overall, the change in both recertification and retention was 5.2% to 5.5% points significantly greater among those exposed to the campaign than those in the comparison group, with significant differences in the change for these outcomes of 7.6% to 7.8% points among infants, and 4% to 4.4% points among children aged 1 to 3 years.

XSL•FO
RenderX

Table . Recertification and retention of infants and children in the innovation and comparison groups in the baseline and innovation periods using crude and adjusted models overall and for infants and children, and difference in difference (DID) estimate of innovation effectiveness.

Outcome	Baseline (T1)[a]		Implementation (T2)		DID
	Comparison	Innovation	Comparison	Innovation	
Overall					
Recertification (%)					
Observed[b], n (%)	25,405 (63.23)	3826 (62.98)	26,793 (70.99)	4200 (75.39)	—[c]
Crude[d], cluster adjusted[e]	63.9 (62.6, 65.2)	63.9 (60.6, 67.1)	64.7 (63.4, 66.0)	69.3 (66.0, 72.6)	4.6[f] (3.0, 6.2)
PSW[d], cluster adjusted	66.2 (65.2, 67.2)	62.9 (62.2, 63.6)	73.5 (71.7, 75.4)	75.4 (75.3, 75.5)	5.2[f] (3.4, 7.1)
Retention (%)					
Observed, n (%)	24,822 (61.78)	3721 (61.25)	26,903 (71.29)	4203 (75.44)	—
Crude, cluster adjusted	62.6 (61.4, 63.9)	62.3 (59.1, 65.5)	64.6 (63.3, 65.8)	68.9 (65.7, 72.1)	4.6[f] (3.0, 6.2)
PSW, cluster adjusted	64.9 (63.7, 66.2)	61.2 (61.0, 61.4)	73.8 (72.1, 75.5)	75.5 (75.5, 75.6)	5.5[f] (3.7, 7.3)
Infants					
Recertification (%)					
Observed, n (%)	9579 (67.84)	1413 (69.13)	10,016 (74.73)	1557 (80.13)	—
Crude, cluster adjusted	67.8 (66.2, 69.3)	68.8 (64.8, 72.8)	73.3 (71.7, 74.9)	78.6 (74.6, 82.6)	4.3[h] (1.3, 7.2)
PSW, cluster adjusted	73.9 (72.4, 75.4)	68.2 (64.8, 71.7)	78.2 (75.5, 80.9)	80.1 (75.8, 84.3)	7.6[f] (5.1, 10.1)
Retention (%)					
Observed, n (%)	9283 (65.74)	1369 (66.98)	10,027 (74.81)	1552 (79.88)	—
Crude, cluster adjusted	65.8 (64.4, 67.3)	66.8 (63.0, 70.5)	73.4 (71.9, 74.9)	78.5 (74.7, 82.2)	4.1[i] (1.2, 7.1)
PSW, cluster adjusted	72.2 (70.4, 74.0)	66.2 (63.5, 68.9)	78.2 (75.5, 80.8)	80.0 (75.8, 84.2)	7.8[i] (5.1, 10.6)
Children					
Recertification (%)					
Observed, n (%)	15,826 (60.73)	2413 (59.84)	16,777 (68.94)	2643 (72.85)	—
Crude, cluster adjusted	60.9 (59.4, 62.4)	60.4 (56.6, 64.2)	63.5 (62.0, 65.0)	67.5 (63.7, 71.2)	4.4[f] (2.4, 6.5)
PSW, cluster adjusted	62.5 (61.4, 63.6)	60.0 (59.2, 60.7)	71.5 (69.3, 73.7)	73.0 (72.0, 73.9)	4.0[f] (2.2, 5.9)
Retention (%)					
Observed, n (%)	15,539 (59.63)	2352 (58.35)	16,876 (69.34)	2651 (73.07)	—
Crude, cluster adjusted	59.9 (58.4, 61.3)	59.1 (55.4, 62.8)	63.6 (62.1, 65.1)	67.3 (63.6, 71.1)	4.6[f] (2.5, 6.6)
PSW, cluster adjusted	61.5 (60.3, 62.8)	58.5 (57.2, 59.7)	71.9 (69.8, 73.9)	73.2 (72.4, 74.1)	4.4[f] (2.6, 6.2)

[a]Sample sizes for overall analyses are 40,181 and 6075 for comparison and innovation, respectively, at T1 and 37,740 and 5571, respectively, at T2. For infants, the corresponding sample sizes are 14,121 and 2044 at T1 and 10,016 and 1557 at T2, and for children, they are 26,060 and 4031 at T1 and 24,337 and 3,628 at T2.

[b]Observed values do not have applicable DID results.

[c]—: not applicable.

[d]Models were adjusted for clustering of children within households and households within clinics.

XSL•FO
RenderX

[e]The observed values are crude percentages. The crude and PSW are predicted values from regression models and their 95% confidence intervals.

[f]*P*<.001.

[g]PSW: propensity score weighting.

[h]*P*=.004.

[i]*P*=.006.

Among 67,198 households with only 1 child and using the PSW-DID adjusted model, the campaign was associated with 6.6% points greater change in continuous benefit issuance overall, with 7% points greater for households with an infant, and 5.9% points greater issuance for those with a child (Table 6).

XSL•FO
RenderX

Table . Continuous benefit issuance for households with only 1 child in the innovation and comparison groups in the baseline and innovation periods using crude and adjusted models overall, and for infants and children, and difference-in-difference (DID) estimate of innovation effectiveness.

Outcome	Baseline (T1)		Implementation (T2)		DID
	Comparison	Innovation	Comparison	Innovation	
Overall[a]					
Continuous benefit					
Issuance (%)					
Observed, n (%)[b]	14,675 (48.98)	2396 (50.22)	16,873 (59.98)	2882 (66.53)	—[c]
Crude, cluster adjusted[d]	48.9 (47.0, 51.0)	50.2 (49.2, 51.2)	60.0 (56.7, 63.2)	66.5 (66.2, 66.8)	5.3[e] (2.6, 8.0)
PSW[f], cluster adjusted	52.9 (51.8, 54.0)	50.2 (49.4, 51.0)	62.6 (59.3, 66.0)	66.6 (66.3, 67.0)	6.6[g] (3.5, 9.7)
Infants[h]					
Continuous benefit					
Issuance (%)					
Observed, n (%)	5602 (51.19)	921 (56.16)	6530 (63.17)	1108 (71.76)	—
Crude, cluster adjusted	51.2 (49.3, 53.1)	56.2 (52.9, 59.4)	63.2 (60.4, 65.9)	71.8 (67.5, 76.0)	3.6[i] (1.4, 5.9)
PSW, cluster adjusted	57.5 (56.1, 58.9)	56.2 (52.8, 59.6)	64.7 (61.8, 67.5)	70.3 (66.6, 74.2)	7.0[g] (4.3, 9.7)
Children[j]					
Continuous benefit					
Issuance (%)					
Observed, n (%)	9073 (47.70)	1475 (47.11)	10,343 (58.12)	1744 (62.55)	—
Crude, cluster adjusted	47.7 (45.4, 50.0)	47.1 (43.8, 50.4)	58.1 (54.2, 62.0)	63.6 (61.5, 65.8)	6.1[e] (3.1, 9.2)
PSW, cluster adjusted	50.5 (49.1, 51.9)	47.1 (44.0, 50.2)	61.1 (57.0, 65.1)	63.5 (61.9, 65.2)	5.9[k] (2.4, 9.4)

[a]Overall household sample sizes at T1 for innovation and comparison are 4771 and 29,963, respectively, and for T2 are 4332 and 28,132, respectively.

[b]Observed values do not have applicable DID results.

[c]Not applicable.

[d]Models were adjusted for clustering of households within clinics.

[e]*P*=.001.

[f]PSW: propensity score weighting.

[g]*P*<.001.

[h]For households with 1 infant only, sample sizes at T1 for innovation and comparison are 1640 and 10,944 respectively, and for T2 are 1544 and 10,337 respectively.

[i]*P*=.004.

[j]For households with 1 child only, sample sizes at T1 for innovation and comparison are 3131 and 19,019 respectively, and for T2 are 2788 and 17,795 respectively.

[k]*P*=.003.

Discussion

Principal Results

To our knowledge, this is the first report of the impact of a multicomponent targeted social media campaign implemented by a local WIC agency to increase retention and participation of children in WIC. The digital marketing campaign reached a large, targeted audience, and the advertisements tailored to the local community had a higher engagement and drove a higher percentage of those who viewed it to the local agency site than the content created for a national audience. These findings add to the literature on the value of locally tailored messaging

XSL•FO
RenderX

[20,21]. The results indicate that the Miami-Dade WIC social marketing campaign was associated with positive changes in recertification, retention, and continuous benefit issuance. In the adjusted analyses, overall recertification was 5.2% higher, with impacts on infants and children of 4.0% to 7.6%, and the impacts on retention were of similar magnitude. Household WIC participation as measured by continuous benefit issuance was 6.6% higher for the group exposed to the campaign than the comparison group. Together, these results indicate that targeted messaging tailored to the local community can successfully direct users to designated websites with trusted information or services and significantly impact continued participation in a program supporting maternal and child health.

There are few studies in the literature of media campaigns related to maternal and child health or related programs presenting results on digital KPIs of reach, engagement, or CTR. In an observational study of a media campaign to direct Canadian pregnant women to guidance on gestational weight gain [22], top keywords for clicking had CTR of 8% to 18%, while other top keywords with fewer clicks showed higher CTR of 18% to 25%. We also observed this pattern of CTR, but did also identify keywords with high numbers of clicks and a high CTR. We found the CTR for our top advertisements between 0.8% to 1.8%, with higher CTR for engagement and CTA objectives; these rates are comparable to those reported by Graham et al [22] among Canadian women (0.3% - 2.6%). Our conversion rate of 69.6% is lower but comparable to their reported conversion rate of 78.9%.

Limitations

There are limitations to both the campaign and the evaluation. First, the social media campaign was targeted toward the innovation clinic geographic areas and WIC participant demographic characteristics, and therefore participant exposure to the innovation cannot be measured directly. We intended to use an in-person client survey to test recognition of campaign components, but this was not feasible with the transition to virtual service provision, and the low response rate limited its usefulness. In terms of the campaign itself, due to the strict 1-year timeline and the required content deliverables, the social media campaign was forced to test several strategies at one time (language testing, objective testing, and audience targeting), which diluted the accuracy of the results. Thoughtful testing of individual strategies and inclusion of WIC participants in that process would have provided valuable insights for future campaigns. The campaign was implemented in quarters, and analyses of performance metrics suggested stronger performance during some quarters than others. On the basis of the overall evaluation approach using administrative data to estimate annual rates, we could not assess or compare the impact by quarter. The focus of the evaluation was on retention of current WIC participants and was not designed to evaluate impact on enrollment in WIC. Finally, with only 1 year of preinnovation data, we are unable to evaluate the parallel trends assumption for the evaluation analyses, and despite our PSW adjustment, the possibility of unmeasured differences remains, thereby affecting our results. Related to this, we have no way of knowing whether or how the COVID-19 pandemic impacted the results.

Conclusions

The study provides novel evidence that a locally tailored targeted social media outreach to WIC-participating families led to absolute improvements on the order of 4% to 7% in household continuous benefit issuance, rates of recertification, and the retention of children in WIC. Although the effect sizes are generally small, we would argue that these differences are important programmatically, given the importance of improving the retention of children in WIC, and findings from the published literature [23]. We provide initial estimates for planning future social media approaches to further test effectiveness on these outcomes. Future research should also evaluate these approaches to increase WIC coverage through new or re-enrollment in WIC. It would also be important to evaluate the impact on outcomes when implemented as a WIC State agency-level strategy.

Funding

This work was supported by a subaward to the Florida Department of Health in Miami-Dade WIC Program from the Hopkins or US Department of Agriculture (USDA) Participant Research Innovation Laboratory for Enhancing WIC Services (HPRIL), a cooperative agreement between the USDA Food and Nutrition Service and the Johns Hopkins Bloomberg School of Public Health (cooperative agreement #OPS-WIC-PRIL-JH-2018). The findings and conclusions in this publication are those of the author(s) and should not be construed to represent any official USDA or US Government determination or policy.

Data Availability

The data used in this study are not publicly available. Data on the performance of social media may be available through request to the first author (ER) and is subject to approval from the Florida Department of Health. Restrictions apply to the access of WIC administrative data, and access would require approval and a use agreement with the Florida Department of Health.

Authors' Contributions

ER, CS, CC, SG, YK, EE, MA, KCB, DP, and LC conceptualized the project. ER, CS, CC collected the data. LC, YK, SG, EE, and CC analyzed the data. EE and SG wrote the first draft with contributions from ER, CS, CC, and KCB. All authors contributed to and reviewed subsequent drafts of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Absolute standardized differences in characteristics.
[DOCX File, 22 KB - publichealth_v11i1e77172_app1.docx]

References

1. WIC at a glance. US Department of Agriculture Food and Nutrition Service. 2013. URL: https://www.fns.usda.gov/wic/about-wic-glance [accessed 2025-04-30]
2. Caulfield LE, Bennett WL, Gross SM, et al. Maternal and child outcomes associated with the special supplemental nutrition program for women, infants, and children (WIC) comparative effectiveness review no 253. : Rockville (MD): Agency for Healthcare Research and Quality (US) Report No.: 22-EHC019. [doi: 10.23970/AHRQEPCCER253] [Medline: 35503870]
3. US Department of Agriculture, Food and Nutrition Service, Office of Research and Analysis. Effects of the special supplemental nutrition program for women, infants, and children (WIC): a review of recent research. : Alexandria, VA: USDA; 2012 Report WIC-12-WM URL: https://www.fns.usda.gov/research/effects-special-supplemental-nutrition-program-women-infants-and-children-wic-review-recent-research [accessed 2025-04-30]
4. Farson Gray K, Balch-Crystal E, Giannarelli L, Johnson P. National- and state-level estimates of WIC eligibility and WIC program reach in 2019.: Alexandria, VA: USDA; 2022. URL: https://fns-prod.azureedge.us/sites/default/files/resource-files/WICEligibles2019-Volume1-revised.pdf [accessed 2025-04-30]
5. Pati S, Siewert E, Wong AT, Bhatt SK, Calixte RE, Cnaan A. The influence of maternal health literacy and child's age on participation in social welfare programs. Matern Child Health J 2014 Jul;18(5):1176-1189. [doi: 10.1007/s10995-013-1348-0] [Medline: 23990157]
6. Whaley SE, Whaley M, Au LE, Gurzo K, Ritchie LD. Breastfeeding Is associated with higher retention in WIC after age 1. J Nutr Educ Behav 2017;49(10):810-816. [doi: 10.1016/j.jneb.2017.07.003] [Medline: 28890264]
7. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Food and Nutrition Board; Committee to Review WIC Food Packages. Review of WIC food packages: improving balance and choice: final report. : National Academies Press (US); 2017. [doi: 10.17226/23655]
8. Lora KR, Hodges L, Ryan C, Ver Ploeg M, Guthrie J. Factors that influence children's exits from the Special Supplemental Nutrition Program for Women, Infants, and Children: a systematic review. Nutrients 2023 Feb 2;15(3):766. [doi: 10.3390/nu15030766] [Medline: 36771472]
9. Third National Survey of WIC Participants (NSWP-111): Brief report 4: retention and potential barriers to participation among WIC participants. US Department of Agriculture Food and Nutrition Service. 2021. URL: https://www.fns.usda.gov/wic/third-national-survey-wic-participants [accessed 2025-04-29]
10. Making WIC work better: strategies to reach more women and children and strengthen benefits use. Food Research and Action Center. 2019. URL: https://frac.org/wp-content/uploads/Making-WIC-Work-Better-Full-Report.pdf [accessed 2025-04-29]
11. WIC modernization. US Department of Agriculture Food and Nutrition Service. URL: https://www.fns.usda.gov/wic/modernization [accessed 2025-11-13]
12. NWA's efforts to recruit and retain WIC participants. National WIC Association. 2019. URL: https://s3.amazonaws.com/aws.upl/nwica.org/caseload-one-pager.pdf [accessed 2025-04-29]
13. Miami-Dade County, Florida. US Census Bureau. 2020. URL: https://data.census.gov/profile/Miami-Dade_County,_Florida?g=050XX00US12086#populations-and-people [accessed 2025-09-20]
14. Lemon KN, Verhoef PC. Understanding customer experience throughout the customer journey. J Mark 2016 Nov;80(6):69-96. [doi: 10.1509/jm.15.0420]
15. Robinson R, Sabugo C, Crisafi C. Targeted digital marketing for WIC child retention. : Florida Department of Health Miami Dade WIC Program; 2022 URL: https://publichealth.jhu.edu/sites/default/files/2023-03/04-miami-dade-wic-hpril-grant-final-reportupdated-dec-2022508-compliant.pdf [accessed 2025-09-20]
16. Neiger BL, Thackeray R, Van Wagenen SA, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. Health Promot Pract 2012 Mar;13(2):159-164. [doi: 10.1177/1524839911433467] [Medline: 22382491]
17. Chancay J, Gross SM, Koegel K, et al. A WIC agency drip marketing campaign is well received and improves program participation and retention: a quasi-experimental study. J Acad Nutr Diet 2025 Mar;125(3):323-334. [doi: 10.1016/j.jand.2024.07.163] [Medline: 39089622]
18. Stuart EA, Huskamp HA, Duckworth K, et al. Using propensity scores in difference-in-differences models to estimate the effects of a policy change. Health Serv Outcomes Res Methodol 2014 Dec 1;14(4):166-182. [doi: 10.1007/s10742-014-0123-z] [Medline: 25530705]

XSL•FO
RenderX

19. Villa JM. Diff: Simplifying the estimation of difference-in-differences treatment effects. Stata J 2016 Mar;16(1):52-71. [doi: 10.1177/1536867X1601600108]
20. Lapinski MK, Oetzel JG, Park S, Williamson AJ. Cultural tailoring and targeting of messages: a systematic literature review. Health Commun 2025 May;40(5):808-821. [doi: 10.1080/10410236.2024.2369340] [Medline: 38961665]
21. Gaysynsky A, Heley K, Chou WYS. An overview of innovative approaches to support timely and agile health communication research and practice. Int J Environ Res Public Health 2022 Nov 16;19(22):15073. [doi: 10.3390/ijerph192215073] [Medline: 36429796]
22. Graham JE, Moore JL, Bell RC, Miller T. Digital marketing to promote healthy weight gain among pregnant women in Alberta: an implementation study. J Med Internet Res 2019 Feb 1;21(2):e11534. [doi: 10.2196/11534] [Medline: 30707100]
23. Davis RA, Leavitt HB, Chau M. A review of interventions to increase WIC enrollment and participation. J Community Health 2022 Dec;47(6):990-1000. [doi: 10.1007/s10900-022-01131-2] [Medline: 35962868]

Abbreviations

ASD: absolute standardized difference
CTR: click-through rate
DID: difference in difference
KPI: key performance indicator
PSW: propensity score weighting
SNAP: Supplemental Nutrition Assistance Program
USDA: US Department of Agriculture
WIC: Special Supplemental Nutrition Program for Women, Infants, and Children

Edited by A Mavragani, T Sanchez; submitted 08.05.25; peer-reviewed by C Gago, L Olsho, Q Zhang; revised version received 03.11.25; accepted 04.11.25; published 10.12.25.

Please cite as:
Robinson EM, Sabugo CJ, Crisafi C, Gross SM, Kang Y, Eppes EV, Augustyn M, Castellanos-Brown K, Paige DM, Caulfield LE Using Social Media Marketing to Improve Retention of Children in the Special Supplemental Nutrition Program for Women, Infants, and Children: Implementation Study
JMIR Public Health Surveill 2025;11:e77172
URL: https://publichealth.jmir.org/2025/1/e77172
doi:10.2196/77172

© Eriko M Robinson, Carla J Sabugo, Carolina Crisafi, Susan M Gross, Yunhee Kang, Elisabet V Eppes, Marycatherine Augustyn, Karen Castellanos-Brown, David M Paige, Laura E Caulfield. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clustering of Lifestyle Behaviors and Their Association With Risk of Metabolic Syndrome Among Adults in Taiwan: Nationwide Cross-Sectional Study

Ya-Hui Chang[1,2*], PhD; Chung-Yi Li[3,4,5], PhD; Hon-Ping Ma[1,2,6*], MS, MD; Chien-Yuan Wu[7], MS; Yann-Yuh Jou[7], MSN; Chiachi Bonnie Lee[8], PhD

[1]Graduate Institute of Injury Prevention and Control, College of Public Health, Taipei Medical University, Taipei, Taiwan

[2]Emergency Department, Shuang Ho Hospital, Taipei Medical University, New Taipei City, Taiwan

[3]Department of Public Health, College of Medicine, National Cheng Kung University, Tainan, Taiwan

[4]Department of Public Health, College of Public Health, China Medical University, Taichung, Taiwan

[5]Department of Healthcare Administration, College of Medical and Health Science, Asia University, Taichung, Taiwan

[6]Department of Emergency, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

[7]Aging and Chronic Disease Control Division, Health Promotion Administration, Ministry of Health and Welfare, Taipei, Taiwan

[8]Department of Health Services Administration, College of Public Health, China Medical University, 100 Section 1, Jingmao Road, Beitun District, Taichung, Taiwan

[*]these authors contributed equally

Corresponding Author:
Chiachi Bonnie Lee, PhD
Department of Health Services Administration, College of Public Health, China Medical University, 100 Section 1, Jingmao Road, Beitun District, Taichung, Taiwan

Abstract

Background: Metabolic syndrome (MetS) is a multifaceted health condition influenced by physiological and lifestyle factors, leading to increased risks of cardiovascular disease and other chronic health issues. Lifestyle behaviors often manifest in various clustering patterns, and evidence of their impact on MetS remains limited.

Objective: This study explores the relationship of latent classes of lifestyle behaviors with the risk of MetS and its components.

Methods: This cross-sectional study used data from Taiwan's 2020 - 2022 Adult Preventive Health Services Database, which was linked to 2020 - 2022 National Health Insurance claim data. The study included 241,156 adults aged 40 years and older who participated in adult preventive health services between 2020 and 2022. Lifestyle behaviors were assessed through smoking, alcohol consumption, betel quid chewing, and physical activities. Latent class analysis was used to identify lifestyle behavior patterns, while binary logistic regression examined the association of these patterns with MetS risk and its components.

Results: The latent class analysis identified 5 distinct lifestyle behavior patterns, with an overall MetS prevalence of 35.72% (86,143/241,156). Compared to the "healthy lifestyle" group (27,465/241,156, 11.39% prevalence), the "insufficiently physically active (IPA)" group (182,101/241,156, 75.51%, adjusted odds ratio [aOR] 1.41, 95% CI 1.37 - 1.45; *P*<.001), the "occasional drinking but physically active" group (18,244/241,156, 7.57%, aOR 1.27, 95% CI 1.21 - 1.32; *P*<.001), the "occasional drinking and regular smoking with IPA" group (9539/241,156, 3.96%, aOR 2.38, 95% CI 2.26 - 2.50; *P*<.001), and the "unhealthy in all behaviors" group (3807/241,156, 1.58%, aOR 2.38, 95% CI 2.22 - 2.55; *P*<.001) showed significantly higher odds of developing MetS. Compared to the "healthy lifestyle" group, all other lifestyle patterns were also associated with significantly higher odds of central obesity (*P*<.001), elevated blood pressure (*P*<.001), elevated fasting blood glucose (*P*<.001), elevated fasting triglycerides (*P*<.001), and reduced high-density lipoprotein cholesterol (*P*<.001), with the most potent effects observed in the "occasional drinking and regular smoking with IPA" group and the "unhealthy in all behaviors" group. An exception was noted for the "occasional drinking but physically active" group, which showed a significantly lower likelihood of reduced high-density lipoprotein cholesterol (aOR 0.90, 95% CI 0.85 - 0.94; *P*<.001).

Conclusions: Engaging in sufficient physical activity and adopting multibehavior interventions tailored to specific lifestyle patterns are crucial for effectively preventing MetS in adults.

(JMIR Public Health Surveill 2025;11:e73114) doi:10.2196/73114

KEYWORDS

health promotion; latent class analysis; lifestyle patterns; metabolic syndrome; physical activity

XSL•FO
RenderX

Introduction

Metabolic syndrome (MetS) is a multifactorial disorder marked by a cluster of interrelated metabolic abnormalities, including central obesity, insulin resistance, and elevated blood pressure [1]. These conditions synergistically elevate the risk of type 2 diabetes, cardiovascular disease, stroke, and cognitive impairment, making MetS a critical global public health issue [2]. MetS has become a significant public health concern in Taiwan, with approximately 34.2% of adults aged over 50 years affected by this condition [3]. This high prevalence underscores the urgent need for effective prevention and intervention strategies.

The etiology of MetS extends beyond biological predispositions, involving a complex interplay of lifestyle and environmental exposure. In East and Southeast Asia, including Taiwan, behavioral risk factors such as tobacco use, alcohol consumption, physical inactivity, unhealthy dietary habits, and betel quid chewing are especially prevalent and contribute substantially to cardiometabolic disease burden [4,5]. In Taiwan, the prevalence of smoking among adults aged 18 years and over was reported at 12.8% in 2024, with higher rates observed among men [6]. Betel quid chewing remains a notable concern, particularly among men, with prevalence rates ranging from 1.6% to 15.4% across various regions [7]. Betel quid chewing has been linked to oxidative stress, inflammation, and insulin resistance, yet it remains under-investigated in MetS research [8].

Emerging research has shifted from evaluating isolated lifestyle factors to examining how unhealthy behaviors cluster within individuals [9,10]. This clustering effect is known to amplify health risks beyond the additive effects of single behaviors. Some studies have applied healthy lifestyle indices or scores to quantify these risks, such as composite metrics that often dichotomize behaviors (eg, active vs inactive), potentially obscuring nuanced behavior patterns and underestimating outcomes variability [10-12].

Latent class analysis (LCA) offers a more sophisticated approach to uncover underlying subgroups of individuals who share similar patterns of lifestyle behaviors [13]. This person-centered methodology can reveal heterogeneity in health risk profiles that traditional methods overlook. However, few population-based studies have applied LCA to investigate lifestyle behavior patterns, including culturally specific practices, like betel quid chewing, related to MetS and its components among Taiwanese adults. Therefore, this study aimed to identify lifestyle behavior patterns using LCA and assess their associations with MetS and its components in a population-based study of Taiwanese adults, with particular attention to the role of betel quid chewing.

Methods

Study Design

This nationwide cross-sectional study analyzed data from the Adult Preventive Health Services Database (APHSD), which was linked to National Health Insurance (NHI) claims data from 2020 to 2022. The study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology guidelines for cross-sectional research (see Checklist 1).

Data Source and Study Population

This study analyzed data from the APHSD, obtained from the Health Promotion Administration, Ministry of Health and Welfare (MOHW), Taiwan, which were linked to NHI claims data from 2020 to 2022 managed by the Health and Welfare Data Science Center (HWDC), MOHW. The service offers free health checkups for residents aged 40 years and older and individuals with a history of polio aged 35 years and older at clinics and hospitals in Taiwan. The study included individuals aged 40 years and older who accessed adult preventive health services during this period. Only the most recent record was retained for those with multiple visits, resulting in an initial dataset of 4,502,104 individuals. After excluding individuals with polio (*International Classification of Diseases, Tenth Revision, Clinical Modification* [*ICD-10-CM*]: A80.0-A80.2, A80.30-A80.39), individuals with missing data, or those identified as outliers (483,051 individuals), a final pool of 4,019,053 participants was established. A stratified random sampling method was applied, selecting 6% of that final participant pool based on gender and 5-year age intervals, leading to a final sample size of 241,156 individuals. As shown in Table S1 in Multimedia Appendix 1, there were no significant differences between the sample population and the overall population.

Lifestyle Variables

This study included four lifestyle behaviors for this analysis: physical activity, cigarette smoking, alcohol drinking, and betel quid chewing status in the past 6 months. Cigarette smoking status was collected by answering the question, “In the past six months, what is your smoking status?” (Answer options: no smoking, smoking only during social occasions, smoking one or less pack per day, and smoking more than one pack per day). Alcohol consumption was measured by the question, “In the past six months, what is your alcohol drinking status?” (Answer options: no drinking, drinking only during social occasions, and regular drinking). Participants reported their betel quid chewing status through the question, “In the past six months, what is your betel quid chewing status?” (Answer options: no chewing, chewing only during social occasions, and regular chewing). Physical activity was measured with the following question: “In the last two weeks, have you done any physical activity for more than 150 minutes per week?” (Answer options: no; yes, but less than 150 minutes per week; yes, and more than 150 minutes per week). We categorized the status of cigarette smoking, alcohol consumption, and betel quid chewing into three levels: no, occasional smoking/drinking/chewing, and regular smoking (smoking one or more than one pack per day)/drinking/chewing. Data on physical activity were also defined as three levels: no, moderate, and sufficient activity. Each behavior (smoking, drinking, betel quid chewing, and physical activity) was a mutually exclusive categorical variable.

XSL•FO
RenderX

MetS and Its Risk Components

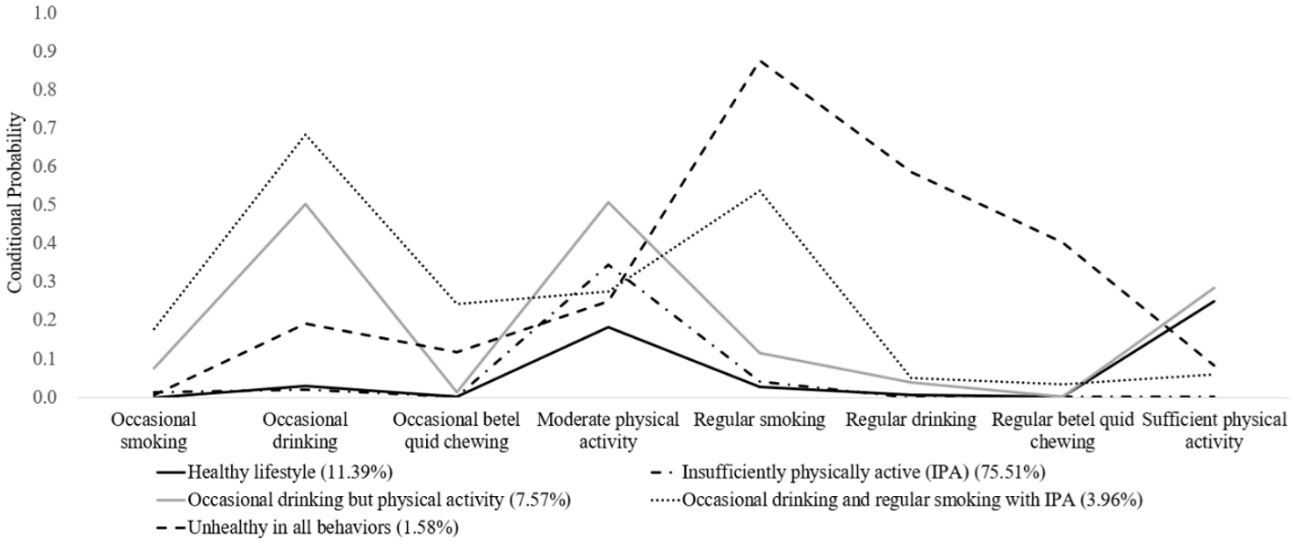
The participants underwent a physical examination and routine blood and urine laboratory tests to measure waist circumference, systolic blood pressure, diastolic blood pressure, fasting plasma glucose, high-density lipoprotein cholesterol (HDL-C), and triglycerides. The MetS risk components included central obesity, elevated blood pressure, elevated fasting glucose, elevated fasting triglycerides, and reduced HDL-C [14]. MetS was defined as the presence of 3 or more of its risk components [14]. Table S2 in Multimedia Appendix 1 defines those MetS risk components.

Statistical Analysis

LCA was performed to categorize patterns of lifestyle. The optimal number of lifestyle patterns was identified by selecting the model with the lowest values of Akaike information criteria, Bayesian information criteria (BIC), consistent Akaike information criterion, and sample-size adjusted BIC, indicating superior model fit [15,16]. The Bayesian factor (BF) is a ratio comparing the likelihood of 2 competing models, where 1<BF<3 indicates weak support, 3<BF<20 indicates moderate support, and BF>20 indicates strong support for the model with K classes compared to a model with a K+1 class [17]. Correct model probability (cmP) estimates the likelihood that each model is the "correct" one among those considered, with the model having the highest cmP being selected [16]. Fit statistics from the 1- to 6-class models to identify appropriate lifestyle patterns are shown in Table S3 in Multimedia Appendix 1. Akaike information criteria, BIC, consistent Akaike information criterion, sample size–adjusted BIC, BF, and cmP values favored the 5-class model. The 5-class model was selected to define patterns of lifestyle in this study. A model with 5 latent classes was chosen: "healthy lifestyle" group, "insufficiently physically active (IPA)" group, "occasional drinking but physically active" group, "occasional drinking and regular smoking with IPA" group, and "unhealthy in all behaviors" group (Figure 1).

This study used the χ^2 test to examine the similarity in the distribution of demographics, lifestyle, and MetS and its risk components between the sampled population and the general population, as well as the differences in demographics and lifestyle across different lifestyle clusters. Binary logistic regression was used to assess the association between lifestyle patterns and MetS risk. We also considered gender and age covariates to adjust in the multivariable regression. The analysis used SAS version 9.4 (SAS Institute Inc). A 2-tailed *P* value of .05 was considered statistically significant.

Figure 1. Latent profiles of lifestyle patterns identified by a 5-class model based on latent class analysis using Taiwan's nationwide claims data from 2020 to 2022.



Ethical Considerations

This study analyzed secondary administrative data from the APHSD (provided by the Health Promotion Administration, MOHW, Taiwan) and NHI claims data (managed by the HWDC, MOHW, Taiwan). The protocol was approved by the institutional review board of National Cheng Kung University (approval no. 112 - 595). All data were fully deidentified and could not be traced to individual identities; thus, informed consent was waived. Analyses were conducted within the secure HWDC environment, with no access to direct identifiers. Only aggregated results were exported, ensuring confidentiality. As no direct contact with participants occurred, no compensation was provided. The manuscript and supplementary materials include no identifiable images, so no image-use consent forms were required.

Results

Table 1 summarizes the descriptive statistics and bivariate analyses of adults based on LCA. Among the study participants, 57.29% (138,149/241,156) were female, and 46.89% (113,075/241,156) were between 45 and 64 years old. Regarding lifestyle behaviors, 8.34% (20,111/241,156) were regular smokers, 10.57% (25,490/241,156) occasionally consumed alcohol, 2.74% (6604/241,156) occasionally or regularly chewed betel quid, and 57.89% (139,608/241,156) reported no physical activity.

XSL•FO
RenderX

Table . Demographic and lifestyle profiles of 5 subgroups identified by latent class analysis using Taiwan's nationwide claims data from 2020 to 2022.

Variables	Total sample (N=241,156)	Healthy lifestyle (n=27,465)	IPA[a] (n=182,101)	Occasional drinking but physically active (n=18,244)	Occasional drinking and regular smoking with IPA (n=9539)	Unhealthy in all behaviors (n=3807)	*P* value
Gender, n (%)							<.001
Female	138,149 (57.29)	16,429 (59.82)	113,568 (62.37)	5854 (32.09)	1756 (18.41)	542 (14.24)	
Male	103,007 (42.71)	11,036 (40.18)	68,533 (37.63)	12,390 (67.91)	7783 (81.59)	3265 (85.76)	
Age (y), n (%)							<.001
40 - 44	24,483 (10.15)	1377 (5.01)	18,416 (10.11)	2356 (12.91)	1684 (17.65)	650 (17.07)	
45 - 54	53,087 (22.01)	4197 (15.28)	39,315 (21.59)	4982 (27.31)	3310 (34.70)	1283 (33.70)	
55 - 64	59,988 (24.88)	6660 (24.25)	44,398 (24.38)	5115 (28.04)	2690 (28.20)	1125 (29.55)	
65 - 74	65,711 (27.25)	10,440 (38.01)	48,810 (26.80)	4363 (23.91)	1502 (15.75)	596 (15.66)	
75 - 84	28,442 (11.79)	4071 (14.82)	22,711 (12.47)	1219 (6.68)	308 (3.23)	133 (3.49)	
85+	9445 (3.92)	720 (2.62)	8451 (4.64)	209 (1.15)	45 (0.47)	20 (0.53)	
Lifestyle, n (%)							
Cigarette smoking							<.001
No	215,852 (89.51)	26,410 (96.16)	171,551 (94.21)	16,270 (89.18)	1496 (15.68)	125 (3.28)	
Occasional	5193 (2.15)	0 (0)	2155 (1.18)	1370 (7.51)	1655 (17.35)	13 (0.34)	
Regular	20,111 (8.34)	1055 (3.84)	8395 (4.61)	604 (3.31)	6388 (66.97)	3669 (96.38)	
Alcohol consumption							<.001
No	210,877 (87.44)	267,67 (97.46)	182,101 (100)	314 (1.72)	1282 (13.44)	413 (10.85)	
Occasional	25,490 (10.57)	0 (0)	0 (0)	16,973 (93.03)	8064 (84.54)	453 (11.90)	
Regular	4789 (1.99)	698 (2.54)	0 (0)	957 (5.25)	193 (2.02)	2941 (77.25)	
Betel quid chewing							<.001
No	234,552 (97.26)	27,442 (99.92)	181,922 (99.90)	18,164 (99.56)	5543 (58.11)	1481 (38.90)	
Occasional	4194 (1.74)	0 (0)	0 (0)	68 (0.37)	3763 (39.45)	363 (9.54)	
Regular	2410 (1.00)	23 (0.08)	179 (0.10)	12 (0.07)	233 (2.44)	1963 (51.56)	
Physical activity							<.001
No	139,608 (57.89)	698 (2.54)	124,138 (68.17)	5880 (32.23)	6413 (67.23)	2479 (65.12)	
Occasional	69,285 (28.73)	0 (0)	57,963 (31.83)	7455 (40.86)	2886 (30.25)	981 (25.77)	
Regular	32,263 (13.38)	26,767 (97.46)	0 (0)	4909 (26.91)	240 (2.52)	347 (9.11)	

[a]IPA: insufficiently physically active.

Each pattern's demographic and lifestyle characteristics varied significantly (Figure 1; Table 1). The class prevalence of the IPA group was 75.51% (182,101/241,156), indicating suboptimal health among participants who had low likelihood of occasional or regular smoking, consuming alcohol, or chewing betel quid, as well as a low probability of sufficient physical activity. Approximately 11.39% (27,465/241,156) of the study participants, classified as the "healthy lifestyle" group, had a very low likelihood of smoking, alcohol consumption, and betel quid chewing and a higher likelihood of engaging in sufficient physical activity. The "occasional drinking but physically active" group, with the second highest probability of occasional drinking and the highest probabilities of moderate and sufficient physical activity, represented 7.57% (18,244/241,156) of the study population. The "occasional drinking and regular smoking with IPA" group comprised 3.96% (9539/241,156) of the study population. Only a few participants (1.58%, 3807/241,156) were categorized as the "unhealthy in all behaviors" group, characterized by high likelihoods of regular drinking, smoking, betel quid chewing, and insufficient physical

XSL•FO
RenderX

activity. Males and middle-aged individuals were significantly overrepresented in the "occasional drinking and regular smoking with IPA" group, as well as the "unhealthy in all behaviors" group. Those in the "occasional drinking and regular smoking with IPA" group and the "unhealthy in all behaviors" group had more unhealthy behaviors in various domains of lifestyle (P<.001).

Among the total sample of 241,156 participants, the prevalence of MetS was 35.72% (86,143/241,156) (Table 2). The prevalence of MetS varied across latent classes, being the lowest in the "healthy lifestyle" group at 29.84% (8195/27,465) and the highest in the "unhealthy in all behaviors" group at 47.10% (1793/3807). Central obesity was present in 48.69% (117,425/241,156) of the total sample and was most common in the "unhealthy in all behaviors" group (49.57%, 1887/3807). The "unhealthy in all behaviors" group exhibited the highest prevalence of elevated blood pressure, with 65.85% (2507/3807) affected. Elevated fasting blood glucose was prevalent in 41.17% (99,278/241,156) and was most common in the "unhealthy in all behaviors" group (45.50%, 1732/3807) and in the "occasional drinking and regular smoking with IPA" group (44.17%, 4213/9539). Elevated fasting triglycerides were prevalent in 29.47% (71,069/241,156) of the sample, notably increasing in the "unhealthy in all behaviors" group to 54.40% (2071/3807). Additionally, reduced HDL-C levels were present in 26.40% (63,672/241,156) of the total population, with the highest prevalence seen in the "occasional drinking and regular smoking with IPA" group (30.49%, 2908/9539) and the "IPA" group (27.83%, 50,677/182,101). Overall, MetS and its risk components were significantly higher in the "occasional drinking and regular smoking with IPA" group and the "unhealthy in all behaviors" group.

Table . Metabolic syndrome and its components in 5 subgroups by latent class analysis using Taiwan's nationwide claims data from 2020 to 2022.

Variables	Total sample (N=241,156)	Healthy lifestyle (n=27,465)	IPA[a] (n=182,101)	Occasional drinking but physically active (n=18,244)	Occasional drinking and regular smoking with IPA (n=9539)	Unhealthy in all behaviors (n=3807)	*P* value
Metabolic syndrome, n (%)							<.001
No	155,013 (64.28)	19,270 (70.16)	116,454 (63.95)	12,201 (66.88)	5074 (53.19)	2014 (52.9)	
Yes	86,143 (35.72)	8195 (29.84)	65,647 (36.05)	6043 (33.12)	4465 (46.81)	1793 (47.1)	
Central obesity, n (%)							<.001
No	123,731 (51.31)	15,436 (56.20)	92,093 (50.57)	9695 (53.14)	4587 (48.09)	1920 (50.43)	
Yes	117,425 (48.69)	12,029 (43.80)	90,008 (49.43)	8549 (46.86)	4952 (51.91)	1887 (49.57)	
Elevated blood pressure, n (%)							<.001
No	104,144 (43.19)	11,724 (42.69)	79,800 (43.82)	7698 (42.19)	3622 (37.97)	1300 (34.15)	
Yes	137,012 (56.81)	15,741 (57.31)	102,301 (56.18)	10,546 (57.81)	5917 (62.03)	2507 (65.85)	
Elevated fasting blood glucose, n (%)							<.001
No	141,878 (58.83)	16,521 (60.15)	107,013 (58.77)	10,943 (59.98)	5326 (55.83)	2075 (54.50)	
Yes	99,278 (41.17)	10,944 (39.85)	75,088 (41.23)	7301 (40.02)	4213 (44.17)	1732 (45.50)	
Elevated fasting triglycerides, n (%)							<.001
No	170,087 (70.53)	21,298 (77.55)	129,605 (71.17)	12,675 (69.47)	4773 (50.04)	1736 (45.60)	
Yes	71,069 (29.47)	6167 (22.45)	52,496 (28.83)	5569 (30.53)	4766 (49.96)	2071 (54.40)	
Reduced HDL-C[b], n (%)							<.001
No	177,484 (73.60)	21,757 (79.22)	131,424 (72.17)	14,915 (81.75)	6631 (69.51)	2757 (72.42)	
Yes	63,672 (26.40)	5708 (20.78)	50,677 (27.83)	3329 (18.25)	2908 (30.49)	1050 (27.58)	

[a]IPA: insufficiently physically active.

[b]HDL-C: high-density lipoprotein cholesterol.

MetS and its risk components significantly varied across the groups with different lifestyle patterns (Tables 3 and 4). After adjusting for age and gender, the highest likelihood of MetS and its risk components was observed in participants who regularly smoked, drank alcohol, chewed betel quid, and had insufficient physical activity. Specifically, those in the "unhealthy in all behaviors" group had significantly elevated odds of high fasting triglycerides (adjusted odds ratio [aOR] 3.16, 95% CI 2.95 - 3.40; P<.001) and MetS overall (adjusted OR 2.38, 95% CI 2.22 - 2.55; P<.001) compared to the "healthy lifestyle" group. The next highest risk group was those with occasional drinking, regular smoking, and IPA, with an odds ratio of 2.70 (95% CI 2.56 - 2.84; P<.001) for elevated triglycerides and 2.38 (95% CI 2.26 - 2.50; P<.001) for MetS. Compared with the "healthy lifestyle" group, the IPA group also consistently displayed elevated risks across all MetS and

XSL•FO
RenderX

its risk components, particularly for MetS (aOR 1.41; 95% CI 1.37-1.45; *P*<.001), central obesity (aOR 1.30; 95% CI 1.27-1.34; *P*<.001), elevated triglycerides (aOR 1.40; 95% CI 1.36-1.45; *P*<.001), and reduced HDL-C (aOR 1.46; 95% CI 1.42-1.51; *P*<.001). The "occasional drinking but physically active" group also presents significantly heightened risks for MetS and its components, though with some variation, including a reduced likelihood of low HDL-C (aOR 0.90; 95% CI 0.85-0.94 *P*<.001).

Table . The likelihood of metabolic syndrome and its risk components (central obesity and elevated blood pressure) between each lifestyle subgroup based on Taiwan's nationwide claims data from 2020 to 2022.[a]

Variables	Metabolic syndrome, aOR[b] (95% CI)	Central obesity, aOR (95% CI)	Elevated blood pressure, aOR (95% CI)
Lifestyle subgroup			
Healthy lifestyle	Ref	Ref	Ref
IPA[c]	1.41 (1.37 - 1.45)	1.30 (1.27 - 1.34)	1.06 (1.04 - 1.09)
Occasional drinking but physically active	1.27 (1.21 - 1.32)	1.40 (1.34 - 1.45)	1.13 (1.09 - 1.18)
Occasional drinking and regular smoking with IPA	2.38 (2.26 - 2.50)	1.91 (1.82 - 2.00)	1.44 (1.37 - 1.51)
Unhealthy in all behaviors	2.38 (2.22 - 2.55)	1.76 (1.64 - 1.89)	1.66 (1.55 - 1.79)
Gender			
Female	Ref	Ref	Ref
Male	1.13 (1.11-1.15)	0.66 (0.65-0.67)	1.45 (1.42 - 1.47)

[a]*P*<.001 for all comparisons with the healthy lifestyle group and female group.

[b]aOR: adjusted odds ratio; age of participants was adjusted in the binary logistic regression model as well.

[c]IPA: insufficiently physically active.

Table . The likelihood of metabolic syndrome risk components (elevated fasting blood glucose, elevated fasting triglycerides, and reduced HDL-C[a]) between each lifestyle subgroup based on Taiwan's nationwide claims data from 2020 to 2022.[b]

Variables	Elevated fasting blood glucose, aOR[c] (95% CI)	Elevated fasting triglycerides, aOR (95% CI)	Reduced HDL-C, aOR (95% CI)
Lifestyle subgroup			
Healthy lifestyle	Ref	Ref	Ref
IPA[d]	1.19 (1.15 - 1.22)	1.40 (1.36 - 1.45)	1.46 (1.42 - 1.51)
Occasional drinking but physically active	1.12 (1.07 - 1.16)	1.28 (1.23 - 1.34)	0.90 (0.85 - 0.94)
Occasional drinking and regular smoking with IPA	1.43 (1.37 - 1.51)	2.70 (2.56 - 2.84)	1.81 (1.71 - 1.91)
Unhealthy in all behaviors	1.48 (1.38 - 1.59)	3.16 (2.95 - 3.40)	1.57 (1.46 - 1.70)
Gender			
Female	Ref	Ref	Ref
Male	1.42 (1.39 - 1.44)	1.58 (1.55 - 1.61)	0.91 (0.89 - 0.92)

[a]HDL-C: high-density lipoprotein cholesterol.

[b]*P*<.001 for all comparisons with the healthy lifestyle group and female group.

[c]aOR: adjusted odds ratio; age of participants was adjusted in the binary logistic regression model as well.

[d]IPA: insufficiently physically active.

Discussion

Principal Findings

To the best of our knowledge, this research is the first population-based study with a large sample size to examine the associations of MetS and its risk components with latent classes of lifestyle behavior patterns, using a 3-level ordinal measure for each behavior among adults. The findings of this study are generalizable to the adult population in Taiwan. This study identified 5 distinct lifestyle behavior patterns among Taiwanese adults using an LCA approach. Compared to the "healthy

XSL•FO
RenderX

lifestyle" group, the IPA group, the "occasional drinking but physically active" group (except for a reduced likelihood of reduced HDL-C), the "occasional drinking and regular smoking with IPA" group, and the "unhealthy in all behaviors" group showed significantly higher odds of developing MetS and its components, with the most potent effects observed in the "occasional drinking and regular smoking with IPA" group and the "unhealthy in all behaviors" group.

Our study showed that the prevalence of MetS was 35.72%, aligning with previous findings of 33% [18]. Chang et al [19] reported a prevalence of 37.9% among individuals over 20 years old in northern Taiwan, while Loke et al [20] observed a prevalence of 34.2% in Taiwanese residents undergoing health examinations at a tertiary medical care facility in southern Taiwan. These results suggest that the prevalence of MetS in Taiwan is slightly higher than in other regions, such as Europe (24.3%) and Japan (23.3% in men, 8.7% in women) [21,22].

Our study revealed clustering patterns of lifestyle behaviors among adults. Notably, approximately 1.58% of participants fell into the "unhealthy in all behaviors" group, characterized by frequent smoking, alcohol consumption, betel quid chewing, and insufficient physical activity. Additionally, 3.96% of participants exhibited occasional drinking and regular smoking, along with insufficient physical activity. These two groups showed the highest odds of MetS and its risk components. Our findings corroborated with previous research, which suggested that the lowest adherence to a healthy lifestyle significantly increased MetS risk [10]. Furthermore, this study presents notable evidence that adopting multiple healthy behavior interventions—including alcohol cessation, smoking cessation, betel nut chewing cessation, and increased physical activity—could be associated with an improved cardiometabolic risk profile and may significantly reduce the risk of MetS. A recent meta-analytic review emphasized the limited impact of single-behavior interventions, which typically yield small-to-medium effect sizes in achieving behavioral or clinical changes, and provided compelling evidence that behavioral change is linearly enhanced as the number of recommendations increases, particularly when the recommendations range between 0 and 4 [12]. Therefore, implementing multibehavior interventions tailored to specific lifestyle behavior patterns is essential for effectively preventing MetS among adults.

Our study found a significant risk of regular smoking on MetS and its risk components, which corroborated with those of previous research. In the Netherlands, Slagter et al [23] demonstrated a positive correlation between smoking and MetS across both genders, irrespective of BMI. Additionally, studies in various demographics have consistently shown that smoking (both current and past smokers) is associated with increased MetS risk [24,25]. The mechanisms linking smoking to MetS are complex. Active smoking disrupts hormonal balance, increases appetite, and gradually reduces metabolic rate [26]. Nicotine exposure is associated with insulin resistance and elevated cortisol, which contributes to abdominal obesity. Smoking also triggers inflammatory responses, as shown by elevated C-reactive protein levels in smokers, a factor linked to type 2 diabetes [27]. Moreover, smoking adversely impacts the lipid profile by increasing total cholesterol and triglycerides and decreasing HDL-C [28]. Notably, a recent meta-analytic review revealed that in the early stages of smoking cessation, the risk of MetS may initially increase compared to active smoking, but it diminishes over time, emphasizing the importance of quitting smoking promptly [29]. Post-cessation obesity may lead to insulin resistance and weight gain, highlighting the need for smoking cessation programs focused on managing weight gain [30].

Moreover, our study identified a link between regular betel quid chewing and increased risks of MetS and its components, consistent with research linking general/central obesity and betel nut chewing [31], which may explain why betel quid chewing was associated with a higher risk of elevated fasting triglycerides and MetS. A longitudinal study by Huang et al [8] using the Taiwan Biobank confirmed that a history of betel nut chewing is associated with MetS and its components, including abdominal obesity and hypertriglyceridemia. This phenomenon may be due to areca alkaloids, such as arecoline, arecaidine, guvacoline, and guvacine, that inhibit the γ-aminobutyric acid receptor, potentially increasing appetite and contributing to obesity [32]. Additionally, betel quid chewing is often accompanied by cigarette smoking and alcohol drinking, and those combinations might also increase the risk of having a cardiometabolic profile.

Notably, a higher likelihood of MetS and its risk components was observed in the IPA group and the "occasional drinking but physically active" group (except for low HDL-C) compared to the "healthy lifestyle" group. Key differences included physical activity levels and occasional drinking prevalence: 97.47% of the "healthy lifestyle" group engaged in sufficient activity, while in the IPA group, none achieved sufficient activity, 31.83% engaged in moderate activity, and 68.17% were inactive. In the "occasional drinking but physically active" group, 26.91% achieved sufficient activity, 40.86% engaged in moderate activity, 32.23% were inactive, and 93.03% reported occasional drinking. Our findings responded to the World Health Organization's recommendation of at least 150-300 minutes of moderate aerobic activity per week for adults [33]. Physical activity provides protective benefits against MetS. It enhances plasma lipid concentration by raising HDL-C and lowering triglycerides [34,35]. Physical activity also improves glucose tolerance and insulin sensitivity [36], potentially reducing the risk of type 2 diabetes [37]. The IPA group, comprising 75.51% of the study participants, could reduce the likelihood of MetS and its risk components by engaging in sufficient physical activity. Furthermore, our study observed a reduced HDL-C risk in the "occasional drinking but physically active" group. A study in Korea found higher risks of low HDL-C in males who drank alcohol once or less per month but lower risks in females who drank 2 - 4 times per month [38]. However, the relationship between occasional drinking and reduced HDL-C remains inconclusive, highlighting the need for further research to explore this association by gender.

Limitations

This study has some limitations. First, the questionnaire on risk factors underwent validation; however, some measurement errors are inevitable. Second, lifestyle data were obtained from

XSL•FO
RenderX

the Adult Preventive Health Services administrative database, which limited the selection of measurement tools; therefore, alternative and potentially more robust instruments could not be employed. Third, sedentary and dietary lifestyles are essential for metabolic health, but they were not accounted for in this study. Additionally, sociodemographic factors, which might influence the likelihood of MetS prevalence, were not considered in this analysis. However, from a health promotion perspective, these factors are generally less modifiable than lifestyle behaviors. Lastly, the study's cross-sectional design limits the interpretation of results, as it does not establish directionality or causality in the observed relationships.

Conclusions

This study identified 5 distinct lifestyle behavior patterns, with a MetS prevalence of 35.72%. Compared to the "healthy lifestyle" group, the IPA group, the "occasional drinking but physically active" group, the "occasional drinking and regular smoking with IPA" group, and the "unhealthy in all behaviors" group showed significantly higher odds of developing MetS, with the strongest associations observed in the "occasional drinking and regular smoking with IPA" and "unhealthy in all behaviors" groups. Similarly, all other lifestyle patterns were significantly associated with increased odds of central obesity, elevated blood pressure, elevated fasting blood glucose, elevated fasting triglycerides, and reduced HDL-C relative to the "healthy lifestyle" group. An exception was found for the "occasional drinking but physically active" group, demonstrating a significantly lower HDL-C likelihood. Our results highlight the importance of engaging in sufficient physical activity and implementing multibehavior interventions tailored to specific lifestyle behavior patterns among adults to prevent MetS effectively.

Acknowledgments

This work was funded by the Health Promotion Administration, Ministry of Health and Welfare (D1120515-113/MOHW 113-HPA-H-113 - 112101). The content of this research may not represent the opinion of the Health Promotion Administration, Ministry of Health and Welfare. We used the generative artificial intelligence tools ChatGPT Plus and Grammarly to assist with grammar correction and writing enhancement during the preparation and revision of this manuscript. All content was subsequently reviewed and edited by the authors, who assume full responsibility for the final version.

Data Availability

An anonymized dataset supporting the results of this study is available upon request.

Authors' Contributions

Conceptualization: CBL, CYL, HPM, YHC
Data curation: CBL, YHC
Formal analysis: CBL, YHC
Investigation: CYL, HPM,
Validation: CBL, CYW, HPM, YHC, YYJ
Writing – original draft: CBL, HPM, YHC
Writing – review & editing: CBL, CYL, CYW, HPM, YHC, YYJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables with further data on comparisons between the population and the study sample, definitions, and latent cluster analysis.
[DOCX File, 50 KB - publichealth_v11i1e73114_app1.docx]

Checklist 1
STROBE checklist for cross-sectional studies.
[DOCX File, 52 KB - publichealth_v11i1e73114_app2.docx]

References

1. Swarup S, Ahmed I, Grigorova Y, Zeltser R. Metabolic syndrome. In: StatPearls [Internet]: StatPearls Publishing; 2025. URL: https://www.ncbi.nlm.nih.gov/books/NBK459248 [accessed 2025-06-15]
2. Wang Z, Chen J, Zhu L, Jiao S, Chen Y, Sun Y. Metabolic disorders and risk of cardiovascular diseases: a two-sample mendelian randomization study. BMC Cardiovasc Disord 2023 Oct 31;23(1):529. [doi: 10.1186/s12872-023-03567-3] [Medline: 37907844]

XSL•FO
RenderX

3. Yen HY, Lin YH, Wang YF, et al. The association between metabolic syndrome components and cognitive function in community-dwelling middle-aged and older adults: the first wave result of a cohort study. J Health Popul Nutr 2025 Mar 29;44(1):94. [doi: 10.1186/s41043-025-00824-3] [Medline: 40158120]
4. Aung AA, Soe Zin SN, Ko Ko A, Thet AC. The association between betel quid chewing and metabolic syndrome among urban adults in Mandalay district of Myanmar. J ASEAN Fed Endocr Soc 2023;38(2):50-56. [doi: 10.15605/jafes.038.02.06] [Medline: 38045670]
5. Rai RK, Bromage S, De Neve JW, Barik A. Lifestyle risk factors for overweight and obesity among rural Indian adults: a community-based prospective cohort study. J Nutr Sci 2025;14:e18. [doi: 10.1017/jns.2025.4] [Medline: 40028378]
6. Adult smoking behavior surveillance system (ASBS). Health Promotion Administration, Ministry of Health and Welfare. URL: https://www.hpa.gov.tw/EngPages/Detail.aspx?nodeid=1718&pid=9913 [accessed 2025-06-15]
7. 2021 National Health Interview Survey (NHIS). Health Promotion Administration, Ministry of Health and Welfare. URL: https://www.hpa.gov.tw/Pages/Detail.aspx?nodeid=364&pid=18514 [accessed 2025-06-15]
8. Huang YC, Geng JH, Wu PY, et al. Betel nut chewing increases the risk of metabolic syndrome and its components in a large Taiwanese population follow-up study category: original investigation. Nutrients 2022 Feb 28;14(5):1018. [doi: 10.3390/nu14051018] [Medline: 35267993]
9. Santos L. The impact of nutrition and lifestyle modification on health. Eur J Intern Med 2022 Mar;97:18-25. [doi: 10.1016/j.ejim.2021.09.020] [Medline: 34670680]
10. Vajdi M, Karimi A, Farhangi MA, Ardekani AM. The association between healthy lifestyle score and risk of metabolic syndrome in Iranian adults: a cross-sectional study. BMC Endocr Disord 2023 Jan 17;23(1):16. [doi: 10.1186/s12902-023-01270-0] [Medline: 36647030]
11. Park YS, Kang SH, Jang SI, Park EC. Association between lifestyle factors and the risk of metabolic syndrome in the South Korea. Sci Rep 2022 Aug 3;12(1):13356. [doi: 10.1038/s41598-022-17361-2] [Medline: 35922546]
12. Zhang AL, Liu S, White BX, et al. Health-promotion interventions targeting multiple behaviors: a meta-analytic review of general and behavior-specific processes of change. Psychol Bull 2024 Jul;150(7):798-838. [doi: 10.1037/bul0000427] [Medline: 38913732]
13. de Mello GT, Thirunavukkarasu S, Jeemon P, Thankappan KR, Oldenburg B, Cao Y. Clustering of health behaviors and their associations with cardiometabolic risk factors among adults at high risk for type 2 diabetes in India: a latent class analysis. J Diabetes 2024 May;16(5):e13550. [doi: 10.1111/1753-0407.13550] [Medline: 38708436]
14. Alberti KGMM, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome. Circulation 2009 Oct 20;120(16):1640-1645. [doi: 10.1161/CIRCULATIONAHA.109.192644]
15. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. Prev Sci 2013 Apr;14(2):157-168. [doi: 10.1007/s11121-011-0201-1] [Medline: 21318625]
16. Nylund-Gibson K, Garber AC, Carter DB, et al. Ten frequently asked questions about latent transition analysis. Psychol Methods 2023 Apr;28(2):284-300. [doi: 10.1037/met0000486] [Medline: 35834194]
17. Kass RE, Raftery AE. Bayes factors. J Am Stat Assoc 1995 Jun;90(430):773-795. [doi: 10.1080/01621459.1995.10476572]
18. Khodarahmi M, Asghari-Jafarabadi M, Abbasalizad Farhangi M. A structural equation modeling approach for the association of a healthy eating index with metabolic syndrome and cardio-metabolic risk factors among obese individuals. PLoS One 2019;14(7):e0219193. [doi: 10.1371/journal.pone.0219193] [Medline: 31260504]
19. Chang SH, Chang YY, Wu LY. Gender differences in lifestyle and risk factors of metabolic syndrome: do women have better health habits than men? J Clin Nurs 2019 Jun;28(11-12):2225-2234. [doi: 10.1111/jocn.14824] [Medline: 30786102]
20. Loke SS, Chang HW, Li WC. Association between metabolic syndrome and bone mineral density in a Taiwanese elderly population. J Bone Miner Metab 2018 Mar;36(2):200-208. [doi: 10.1007/s00774-017-0826-7] [Medline: 28303402]
21. Scuteri A, Laurent S, Cucca F, et al. Metabolic syndrome across Europe: different clusters of risk factors. Eur J Prev Cardiol 2015 Apr;22(4):486-491. [doi: 10.1177/2047487314525529] [Medline: 24647805]
22. Hattori T, Konno S, Munakata M. Gender differences in lifestyle factors associated with metabolic syndrome and preliminary metabolic syndrome in the general population: the Watari study. Intern Med 2017 Sep 1;56(17):2253-2259. [doi: 10.2169/internalmedicine.8578-16] [Medline: 28794374]
23. Slagter SN, van Vliet-Ostaptchouk JV, Vonk JM, et al. Associations between smoking, components of metabolic syndrome and lipoprotein particle size. BMC Med 2013 Sep 3;11:1-15. [doi: 10.1186/1741-7015-11-195] [Medline: 24228807]
24. Hwang GY, Cho YJ, Chung RH, Kim SH. The relationship between smoking level and metabolic syndrome in male health check-up examinees over 40 years of age. Korean J Fam Med 2014 Sep;35(5):219-226. [doi: 10.4082/kjfm.2014.35.5.219] [Medline: 25309702]
25. Wada T, Urashima M, Fukumoto T. Risk of metabolic syndrome persists twenty years after the cessation of smoking. Intern Med 2007;46(14):1079-1082. [doi: 10.2169/internalmedicine.46.0026] [Medline: 17634703]
26. Audrain-McGovern J, Benowitz NL. Cigarette smoking, nicotine, and body weight. Clin Pharmacol Ther 2011 Jul;90(1):164-168. [doi: 10.1038/clpt.2011.105] [Medline: 21633341]
27. Emerging Risk Factors Collaboration, Kaptoge S, Di Angelantonio E, et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. Lancet 2010 Jan 9;375(9709):132-140. [doi: 10.1016/S0140-6736(09)61717-7] [Medline: 20031199]

XSL•FO
RenderX

28. Gokulakrisnan A, Jayachandran Dare B, Thirunavukkarasu C. Attenuation of the cardiac inflammatory changes and lipid anomalies by (-)-epigallocatechin-gallate in cigarette smoke-exposed rats. Mol Cell Biochem 2011 Aug;354(1-2):1-10. [doi: 10.1007/s11010-011-0785-6] [Medline: 21633901]
29. Kim HJ, Cho YJ. Smoking cessation and risk of metabolic syndrome: a meta-analysis. Medicine (Baltimore) 2024 May 31;103(22):e38328. [doi: 10.1097/MD.0000000000038328] [Medline: 39259087]
30. Harris KK, Zopey M, Friedman TC. Metabolic effects of smoking cessation. Nat Rev Endocrinol 2016 May;12(5):299-308. [doi: 10.1038/nrendo.2016.32] [Medline: 26939981]
31. Guh JY, Chuang LY, Chen HC. Betel-quid use is associated with the risk of the metabolic syndrome in adults. Am J Clin Nutr 2006 Jun;83(6):1313-1320. [doi: 10.1093/ajcn/83.6.1313] [Medline: 16762942]
32. Giri S, Idle JR, Chen C, Zabriskie TM, Krausz KW, Gonzalez FJ. A metabolomic approach to the metabolism of the areca nut alkaloids arecoline and arecaidine in the mouse. Chem Res Toxicol 2006 Jun;19(6):818-827. [doi: 10.1021/tx0600402] [Medline: 16780361]
33. WHO guidelines on physical activity and sedentary behaviour: at a glance. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/9789240015128 [accessed 2024-12-20]
34. Leon AS, Sanchez OA. Response of blood lipids to exercise training alone or combined with dietary intervention. Med Sci Sports Exerc 2001 Jun;33(6 Suppl):S502 S515. [doi: 10.1097/00005768-200106001-00021] [Medline: 11427777]
35. Wilmore JH, Green JS, Stanforth PR, et al. Relationship of changes in maximal and submaximal aerobic fitness to changes in cardiovascular disease and non-insulin-dependent diabetes mellitus risk factors with endurance training: the HERITAGE Family Study. Metab Clin Exp 2001 Nov;50(11):1255-1263. [doi: 10.1053/meta.2001.27214] [Medline: 11699041]
36. Arciero PJ, Vukovich MD, Holloszy JO, Racette SB, Kohrt WM. Comparison of short-term diet and exercise on insulin action in individuals with abnormal glucose tolerance. J Appl Physiol (1985) 1999 Jun;86(6):1930-1935. [doi: 10.1152/jappl.1999.86.6.1930] [Medline: 10368358]
37. Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. Arch Intern Med 2001 Jun 25;161(12):1542-1548. [doi: 10.1001/archinte.161.12.1542] [Medline: 11427103]
38. Oh SS, Kim W, Han KT, Park EC, Jang SI. Alcohol consumption frequency or alcohol intake per drinking session: which has a larger impact on the metabolic syndrome and its components? Alcohol 2018 Sep;71:15-23. [doi: 10.1016/j.alcohol.2018.01.005] [Medline: 29929088]

Abbreviations

aOR: adjusted odds ratio
APHSD: Adult Preventive Health Services Database
BF: Bayesian factor
BIC: Bayesian information criteria
cmP: correct model probability
HDL-C: high-density lipoprotein cholesterol
HWDC: Health and Welfare Data Science Center
***ICD-10-CM*:** *International Classification of Diseases, Tenth Revision, Clinical Modification*
IPA: insufficiently physically active
LCA: latent class analysis
MetS: metabolic syndrome
MOHW: Ministry of Health and Welfare
NHI: National Health Insurance

Edited by A Mavragani; submitted 25.02.25; peer-reviewed by J Chen, R Almendra-Pegueros; revised version received 07.07.25; accepted 24.07.25; published 18.09.25.

Please cite as:
Chang YH, Li CY, Ma HP, Wu CY, Jou YY, Lee CB
Clustering of Lifestyle Behaviors and Their Association With Risk of Metabolic Syndrome Among Adults in Taiwan: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e73114
URL: https://publichealth.jmir.org/2025/1/e73114
doi:10.2196/73114

XSL•FO
RenderX

© Ya-Hui Chang, Chung-Yi Li, Hon-Ping Ma, Chien-Yuan Wu, Yann-Yuh Jou, Chiachi Bonnie Lee. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Temporal Trends in Cervical Human Papillomavirus Prevalence Among Females in Xiamen, China (2016-2023): Cross-Sectional Study

Zhuju Chen[1], BMed; Erjuan Lin[2,3]; Mo Chen[4]; Shihan Wang[2,3], MM; Qiuyuan Lin[2,3], PhD; Qingquan Chen[2,3], MM; Heng Xue[5], MM; Guanbin Zhang[2,3,6], PhD

[1]Center of Clinical Laboratory, Zhongshan Hospital Xiamen University, School of Medicine, Xiamen University, Xiamen, China

[2]Department of Laboratory Medicine, Medical Technology Experimental Teaching Center, School of Medical Technology and Engineering, Fujian Medical University, 1 North Xuefu Road, Daxue New District, Fuzhou, Fujian, China

[3]Key Laboratory of Clinical Laboratory Technology for Precision Medicine (Fujian Medical University), Fujian Province University, Fuzhou, China

[4]Grade 2023, Stomatology, School and Hospital of Stomatology, Fujian Medical University, Fuzhou, China

[5]Department of Laboratory Medicine, Fujian Maternity and Child Health Hospital College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[6]School of Intelligent Medicine, Chengdu University of Traditional Chinese Medicine, Chengdu, China

Corresponding Author:
Guanbin Zhang, PhD
Department of Laboratory Medicine, Medical Technology Experimental Teaching Center, School of Medical Technology and Engineering, Fujian Medical University, 1 North Xuefu Road, Daxue New District, Fuzhou, Fujian, China

Abstract

Background: Human papillomavirus (HPV) is a primary causative agent of cervical cancer, accounting for more than 90% of cases worldwide. Epidemiological data on regional HPV prevalence and genotype distribution are critical for tailoring targeted cervical cancer prevention strategies, particularly in regions with limited population-based studies.

Objective: This study aimed to investigate temporal trends in the prevalence of overall HPV infection and vaccine-targeted HPV genotypes among females in Xiamen between 2016 and 2023 using annual cross-sectional analyses.

Methods: We analyzed retrospective deidentified data from 63,553 females who underwent HPV genotyping of cervical exfoliated cells at Zhongshan Hospital affiliated with Xiamen University from 2016 to 2023. Data on HPV genotyping, age, and detection time were collected from the hospital's electronic information system. For each year, we conducted a cross-sectional assessment of HPV infection status to calculate annual HPV prevalence. Temporal trends of HPV prevalence were analyzed across 3 pandemic periods (prepandemic: 2016 - 2019, pandemic: 2020 - 2022, and postpandemic: 2023) and by age groups.

Results: The overall HPV prevalence was 25.24% (16,039/63,553), comprising high-risk human papillomavirus (HR-HPV) at 19.26% (12,242/63,553) and low-risk human papillomavirus (LR-HPV) at 10.08% (6409/63,553). Vaccine-targeted HPV prevalence rates were bivalent human papillomavirus at 3.56% (2264/63,553), quadrivalent human papillomavirus at 5.89% (3746/63,553), and nine-valent human papillomavirus at 13.64% (8666/63,553), respectively. Notably, the number of non–vaccine-targeted HPV genotypes accounted for 16.01% (10,177/63,553) of all tested females and 63.45% (10,177/16,039) of HPV-positive cases. The top 5 HR-HPV genotypes were HPV52 (3000/63,553, 4.72%), HPV58 (1895/63,553, 2.98%), HPV53 (1582/63,553, 2.49%), HPV16 (1461/63,553, 2.30%), and HPV39 (1116/63,553, 1.76%), while HPV81 (1407/63,553, 2.21%), HPV61 (1268/63,553, 2%), and HPV6 (1101/63,553, 1.73%) were the most prevalent LR-HPV genotypes. Temporal analysis revealed significant declines in the prevalence of overall HPV, HR-HPV, LR-HPV, bivalent human papillomavirus, quadrivalent human papillomavirus, nine-valent human papillomavirus, and specific genotypes (HPV52, HPV58, HPV16, HPV39, and HPV6) from 2016 to 2019 to 2023 (all $P<.001$). Conversely, HPV81 prevalence increased significantly in 2023 compared to 2020 - 2022 (2.44% vs 1.96%; $P<.001$). Age-stratified analysis of HPV prevalence showed a significant declining trend with increasing age ($P<.001$), with peak prevalence observed in the ≤20-year age group.

Conclusions: Cervical HPV infection, particularly non–vaccine-targeted genotypes, remains a substantial public health burden in Xiamen, highlighting the urgency to develop broader spectrum vaccines, to enhance cervical cancer screening programs, and to implement age-specific interventions, specifically for females aged ≤20 years. Long-term surveillance of emerging HPV genotypes and vaccination coverage is recommended.

(JMIR Public Health Surveill 2025;11:e70507) doi:10.2196/70507

XSL•FO
RenderX

KEYWORDS

human papillomavirus; prevalence; cervical cancer; vaccine; COVID-19 pandemic

Introduction

Cervical cancer is the fourth most common cancer and the fourth leading cause of cancer-related death among women, and there were approximately 604,000 cases and 342,000 cancer-related deaths in 2020 globally [1]. In China, it was reported that there were 111,820 new cases and 61,579 deaths of cervical cancer in 2022 [2]. The estimated number of deaths from cervical cancer could even reach 133,000 by 2035 [3]. Therefore, the burden of cervical cancer is still relatively high in China.

Persistent infection with oncogenic human papillomavirus (HPV) genotypes is one of the main causes of cervical cancer [4]. HPV is a circular nonenveloped double-stranded DNA virus, containing approximately 8000 bases. There are more than 200 HPV genotypes that have been identified [5,6]. Conventionally, high-risk human papillomavirus (HR-HPV) includes 12 genotypes classified as Group 1 carcinogens (definitely carcinogenic to humans: HPV16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59) and 1 genotype classified as Group 2A (probably carcinogenic to humans: HPV68) [7]. In addition, HPV26, 53, 66, 67, 70, 73, and 82 were classified as agents with limited evidence in humans [8]. A recent study found that there was a causal relationship between invasive cervical cancer and 17 HPV genotypes (HPV16, 18, 45, 33, 31, 58, 52, 59, 26, 69, 35, 39, 73, 68, 56, 82, and 51) [9]. The HPV prevalence of patients with invasive cervical cancer was 48.3 to 1.4 times higher than that of the normal cervical cytology examination population [9], which indicates that the potential carcinogenicity for cervical cancer varied greatly among different HPV genotypes. Therefore, in this study, we analyzed an expanded panel of 17 HR-HPV genotypes, including the 13 conventionally defined HR-HPV genotypes plus 4 additional genotypes (HPV26, 53, 66, and 82) that were frequently included in clinical genotyping assays in our region.

HPV vaccination, as well as cervical cancer screening and treatment, is the main measure to control cervical cancer. Currently, licensed prophylactic HPV vaccines include bivalent vaccines, which target HPV16 and 18 (2V-HPV), quadrivalent vaccines, which target HPV6, 11, 16, and 18 (4V-HPV), and nine-valent vaccines, which cover HPV6, 11, 16, 18, 31, 33, 45, 52, and 58 (9V-HPV). A significant reduction in the prevalence of HPV vaccine genotype and incidence of other HPV-associated diseases had been found in countries with high HPV vaccination [10-12]. However, the HPV vaccine mainly prevents vaccine genotype infection, and the preventive effect on genotypes not covered by the vaccine is quite limited. Therefore, it is necessary to monitor the prevalence and distribution of the HPV genotype dynamically for a long time to provide baseline information for estimating the effectiveness of HPV vaccination, implementing appropriate HPV vaccination strategies, and developing new HPV vaccines.

The prevalence and genotype distribution of HPV exhibited significant differences among countries [13,14], and even varied among different areas within a country [15-19]. In addition, the trend of HPV prevalence also varied among different areas within a province [17,18]. Within the same area, the prevalence and distribution of HPV genotypes also varied in different periods [20]. The prevalence of HPV was also affected by age and the spread of other viruses, with higher prevalence in younger women (<25 years old) [15,21] and older women (≥60 years old) [22].

The COVID-19 epidemic broke out at the end of 2019. To quickly control the epidemic, stringent restrictions on human activities, such as physical distancing and cessation of nonessential activities, were implemented by many countries. The restriction may have an impact on cervical cancer screening, HPV vaccination, and HPV transmission. Recent studies showed that the pandemic had an impact on the prevalence of HPV [21,23]. However, to date, data on HPV prevalence and the distribution of vaccine-targeted genotypes remain limited in Xiamen, particularly across the pre-, intra-, and post–COVID-19 pandemic periods. A global goal of eliminating cervical cancer by 2030 has been set by the World Health Organization. In order to provide more accurate HPV prevalence information, for improving cervical cancer prevention and HPV screening strategies, a comprehensive analysis of HPV prevalence characteristics was conducted in Xiamen from 2016 to 2023.

Methods

Study Setting and Data Collection

This study used a cross-sectional design to estimate cervical HPV prevalence and its temporal changes among females at Zhongshan Hospital affiliated with Xiamen University, a tertiary-care academic medical center located in Xiamen, China, from January 2016 to December 2023. Each calendar year (2016 - 2023) served as a discrete "time point" for cross-sectional data collection. Annual cross-sectional data were aggregated to evaluate temporal changes in HPV prevalence. The study period was stratified into 3 phases based on the COVID-19 pandemic timeline: prepandemic (2016 - 2019), pandemic (2020 - 2022), and postpandemic (2023).

Data on HPV genotyping, age, and detection time—originating from the hospital's health examination centers, outpatient departments, and inpatient wards—were extracted from its electronic information system. To avoid duplicate counting and minimize bias, only data from participants' first HPV test were included in the analysis, as some individuals underwent multiple tests during the study period. Thus, each year's dataset represented an independent sample of females tested in that year, with no overlap in the study population across years.

This study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement guidelines for reporting observational studies (STROBE Checklist). A total of 63,553 females who underwent HPV genotyping of cervical exfoliated cells were included in this study. Analyses were stratified by data source and age to minimize bias in HPV prevalence.

XSL•FO
RenderX

The inclusion criteria were as follows: (1) specimens were cervical exfoliated cells, and (2) the period of sample collection was from January 2016 to December 2023. The exclusion criteria were females with incomplete information, such as HPV genotyping results, age, specimen type, and detection time. No age restrictions were applied, and participants were included regardless of their clinical presentation, history of cervical pathology follow-up, pregnancy status, or immunological condition.

Human Papillomavirus Genotyping

Cervical exfoliated cells were collected by a gynecologist using a cytobrush and stored in a preservation solution for HPV detection. HPV genotyping was performed with the HPV genotyping detection kit (flow-through fluorescence hybridization, Tellgen). The kit integrates multiplex polymerase chain reaction and flow cytometry–based fluorescence hybridization to simultaneously detect 27 HPV genotypes, including 17 HR-HPV (16, 18, 26, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 59, 66, 68, and 82), and 10 low-risk human papillomavirus (LR-HPV; 6, 11, 40, 42, 43, 44, 55, 61, 81, and 83) genotypes. Total DNA was extracted from cervical cells using the nucleic acid lysis reagent provided in the kit. This kit used the human β-globin as an internal control to monitor sampling adequacy and amplification efficiency. The assay was conducted on a Luminex200 analyzer (Luminex), which quantified fluorescence intensity for each microsphere. Each genotype was interrogated 100 times, and mean values were calculated to determine positivity. All the experimental procedures strictly followed the manufacturer's instructions.

Statistical Analysis

Statistical analyses were performed using SPSS v18.0 (SPSS Inc.), GraphPad Prism v5.0 (GraphPad Software), and Python v3.9.7 (Python Software Foundation). Descriptive statistics for categorical variables were expressed as frequencies and percentages, and continuous variables were presented as median (IQR) values. Annual HPV prevalence was calculated as ([number of individuals with positive HPV test results in a given year]/[total number of eligible individuals tested in that year]×100). The Mann-Kendall test was applied to assess temporal trends in HPV prevalence across the continuous years (2016 - 2023). The Cochran-Armitage trend test was used to assess temporal trends in HPV prevalence across 3 periods (2016 - 2019, 2020 - 2022, and 2023) and by age group. The chi-square test was used to evaluate differences in HPV prevalence between groups. A 2-sided *P* value less than .05 was considered statistically significant.

Ethical Considerations

The study was conducted following the Declaration of Helsinki principles. Ethical approval was obtained from the Ethics Committee of Zhongshan Hospital affiliated with Xiamen University (approval number 2024 - 062). This research constituted a retrospective analysis of previously test data, and all data have been anonymized. In compliance with ethical guidelines, we applied for and were granted a waiver of informed consent, thereby eliminating the need to obtain additional consent from the participants involved. All study data were anonymized and subsequently stored and analyzed within a secure, password-protected database. Information about individuals was removed from all records, and the data were aggregated to prevent the reidentification of any individual. No identifiable information was included in the study or supplementary materials. This study only involved analyzing deidentified data and did not impose any extra financial or other burdens on the participants. Consequently, no compensation, financial or otherwise, was provided to participants included in the study. Furthermore, the study and supplementary files did not contain any images of participants or identifiable materials. Therefore, there was no risk of individual identification in this study.

Results

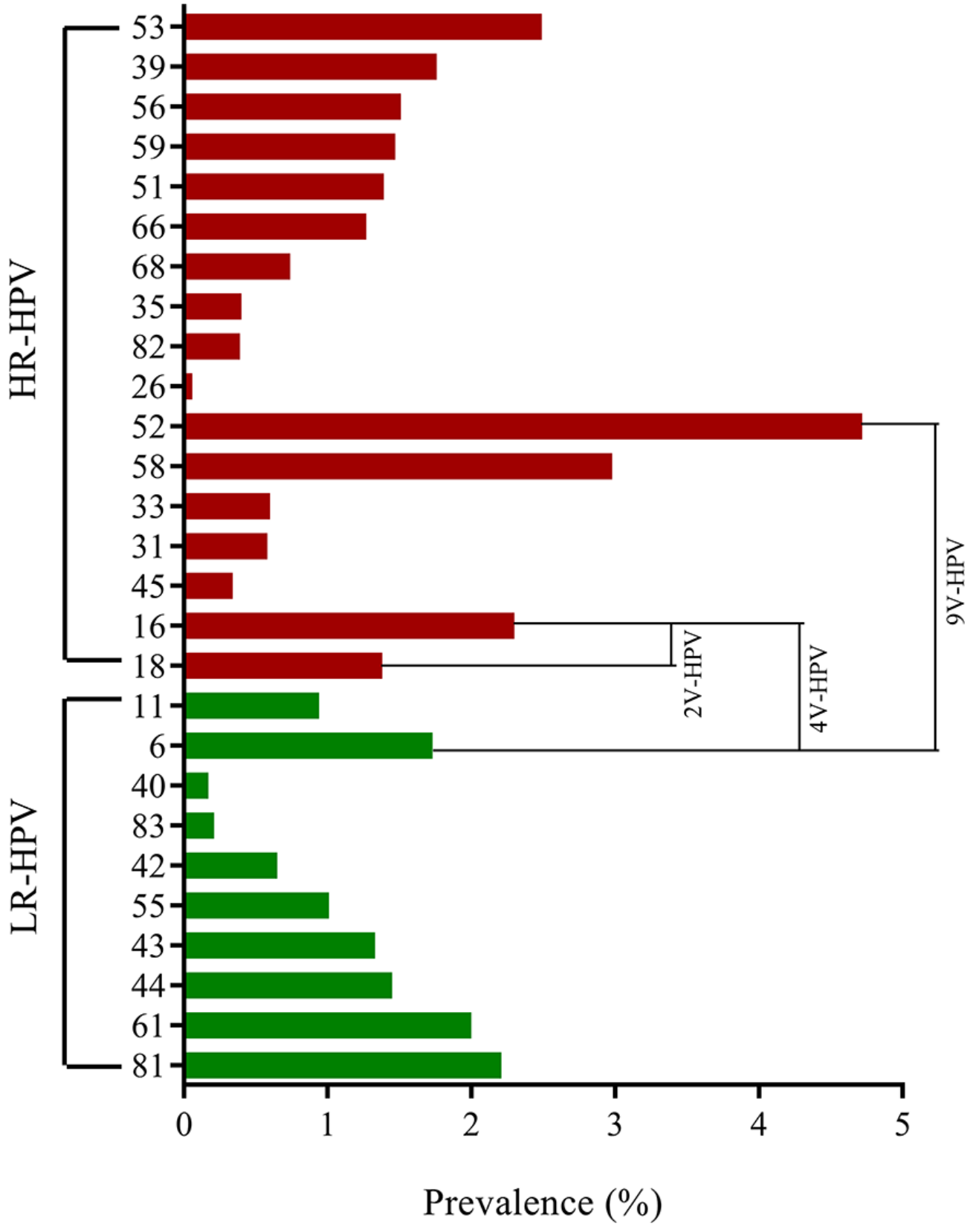
Overall Human Papillomavirus Genotype Distribution (2016-2023)

A total of 63,553 females undergoing cervical exfoliated cell HPV genotyping at Zhongshan Hospital affiliated with Xiamen University from January 2016 to December 2023 were included. Between 2016 and 2023, the average prevalence rates of overall HPV, HR-HPV, LR-HPV, 2V-HPV, 4V-HPV, and 9V-HPV were 25.24% (16,039/63,553), 19.26% (12,242/63,553), 10.08% (6409/63,553), 3.56% (2264/63,553), 5.89% (3746/63,553), and 13.64% (8666/63,553), respectively (Table S1 in Multimedia Appendix 1). Notably, the prevalence of non–vaccine-targeted HPV genotypes was 16.01% (10,177/63,553), accounting for 63.45% (10,177/16,039) of the total positive cases.

The prevalence of specific HPV genotypes among females from 2016 to 2023 is shown in Figure 1. The 5 most common HR-HPV genotypes were HPV52 (3000/63,553, 4.72%), HPV58 (1895/63,553, 2.98%), HPV53 (1582/63,553, 2.49%), HPV16 (1461/63,553, 2.30%), and HPV39 (1116/63,553, 1.76%). The top 3 LR-HPV genotypes were HPV81 (1407/63,553, 2.21%), HPV61 (1268/63,553, 2%), and HPV6 (1101/63,553, 1.73%). In particular, among the positive cases, HR-HPV infections accounted for 76.79% (12,317/16,039), while LR-HPV positive cases represented 39.37% (6314/16,039) of the total infections.

XSL•FO
RenderX

Figure 1. The prevalence of specific human papillomavirus genotypes among females from 2016 to 2023. HR-HPV, LR-HPV, 2V-HPV genotypes (HPV16 and 18) targeted by the bivalent vaccine, 4V-HPV genotypes (HPV16, 18, 6, and 11) targeted by the quadrivalent vaccine, and 9V-HPV genotypes (HPV16, 18, 6, 11, 31, 33, 45, 52, and 58) targeted by the nine-valent vaccine. 2V-HPV: bivalent human papillomavirus; 4V-HPV: quadrivalent human papillomavirus; 9V-HPV: nine-valent human papillomavirus; HPV: human papillomavirus; HR-HPV: high-risk human papillomavirus; LR-HPV: low-risk human papillomavirus.

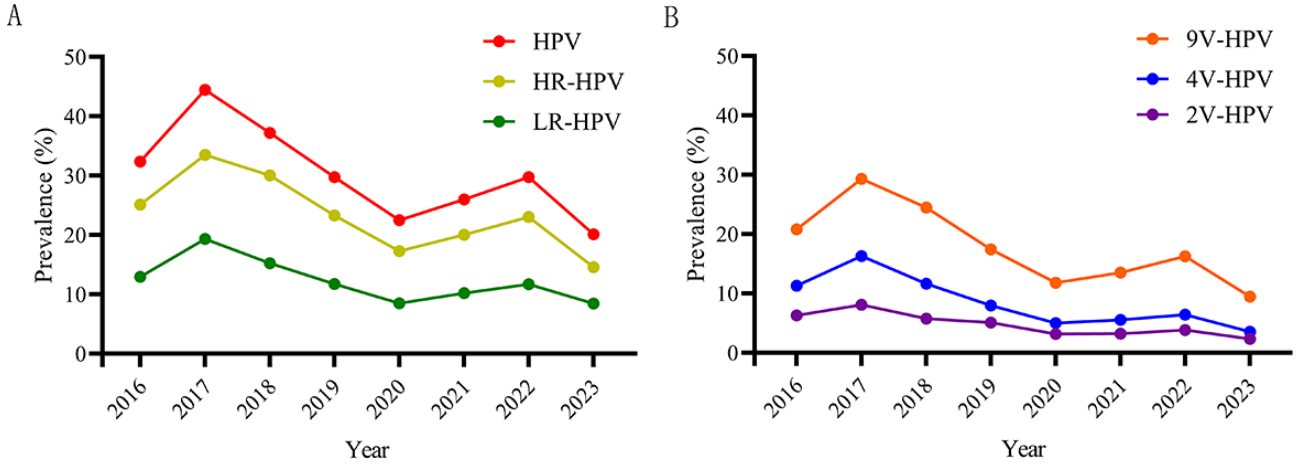


XSL•FO
RenderX

Temporal Trends in the Prevalence of Overall HPV, Vaccine-Targeted HPV, and Specific HPV Genotype (2016-2023)

The annual prevalence trends of overall HPV, HR-HPV, and LR-HPV demonstrated a declining trajectory (overall HPV: *P*=.03; HR-HPV: *P*=.04; LR-HPV: *P*=.02), generally exhibiting a distinct “M-shaped” pattern. This pattern featured 2 peaks in 2017 and 2022, along with 3 troughs in 2016, 2020, and 2023. In 2023, the prevalence rates of overall HPV, HR-HPV, and LR-HPV were 20.15% (3994/19,826), 14.61% (2897/19,826), and 8.46% (1678/19,826), respectively (Figure 2A and Table S1 in).

Figure 2. Annual cross-sectional estimates of HPV prevalence among females from 2016 to 2023. (**A**) Overall HPV prevalence, including the prevalence of both HR-HPV and LR-HPV genotypes. (**B**) Prevalence of vaccine-targeted HPV genotypes. 2V-HPV genotypes (HPV16 and 18) targeted by the bivalent vaccine, 4V-HPV genotypes (HPV16, 18, 6, and 11) targeted by the quadrivalent vaccine, and 9V-HPV genotypes (HPV16, 18, 6, 11, 31, 33, 45, 52, and 58) targeted by the nine-valent vaccine. The Mann-Kendall test was applied to assess temporal trends in HPV prevalence across 2016 - 2023. 2V-HPV: bivalent human papillomavirus; 4V-HPV: quadrivalent human papillomavirus; 9V-HPV: nine-valent human papillomavirus; HPV: human papillomavirus; HR-HPV: high-risk human papillomavirus; LR-HPV: low-risk human papillomavirus.

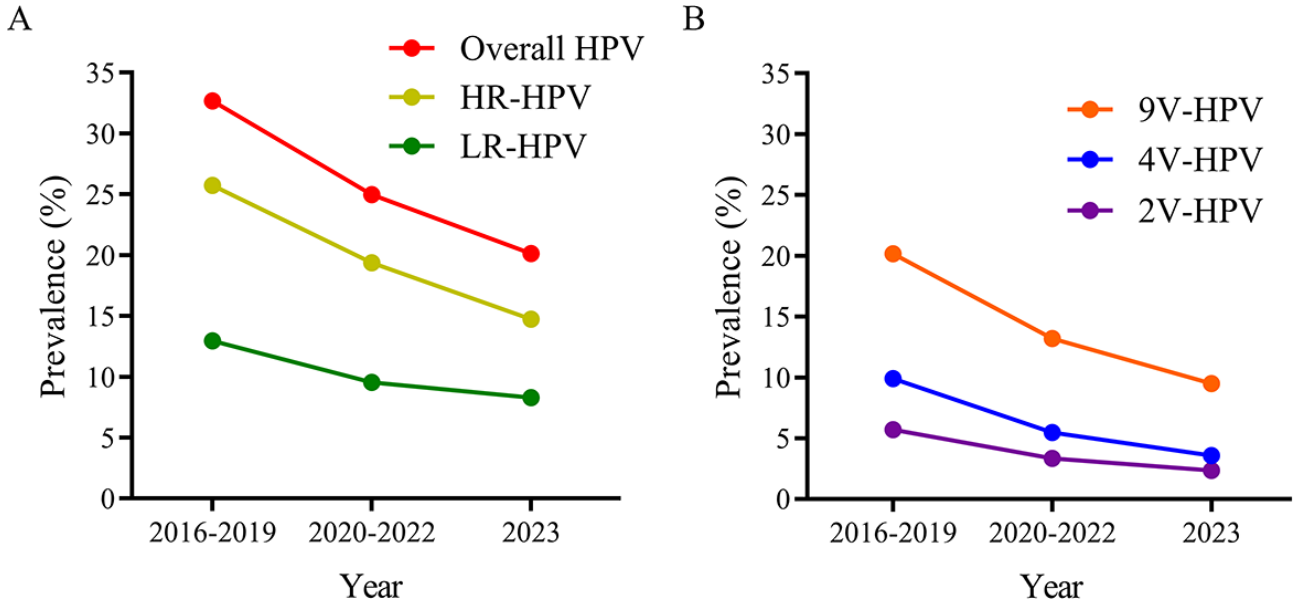


All vaccine-targeted HPV genotypes (2V-HPV, 4V-HPV, and 9V-HPV) exhibited significant declining trends from 2016 to 2023 (2V-HPV: *P*=.04; 4V-HPV: *P*=.02; 9V-HPV: *P*=.01) (Figure 2B; Table S1 in). The prevalences of 9V-HPV and 4V-HPV showed greater fluctuations, whereas 2V-HPV prevalence remained relatively stable. Notably, the first prevalence peak in 2017 occurred shortly after the first HPV vaccine received regulatory approval in China in 2016, potentially reflecting increased HPV testing following the formal clinical introduction of vaccines. The subsequent rise from the 2020 trough to the 2022 peak overlapped with the COVID-19 pandemic.

Stratifying by time periods (2016 - 2019, 2020 - 2022, and 2023) revealed consistent downward trends (all *P*<.001) in overall and vaccine-targeted HPV prevalence (Figure 3A and B; Table S2 in). Compared to 2016 - 2019, overall HPV prevalence in 2023 decreased by 38.34% (32.68% vs 20.15%), while HR-HPV and LR-HPV prevalence decreased by 42.89% (25.58% vs 14.61%) and 35.81% (13.18% vs 8.46%), respectively. For vaccine-targeted genotypes, 2V-HPV prevalence decreased by 58.88% (5.69% vs 2.34%), 4V-HPV by 63.84% (9.90% vs 3.58%), and 9V-HPV by 52.97% (20.18% vs 9.49%) over the same period. These findings confirm significant reductions in HPV burden during and after the COVID-19 pandemic.

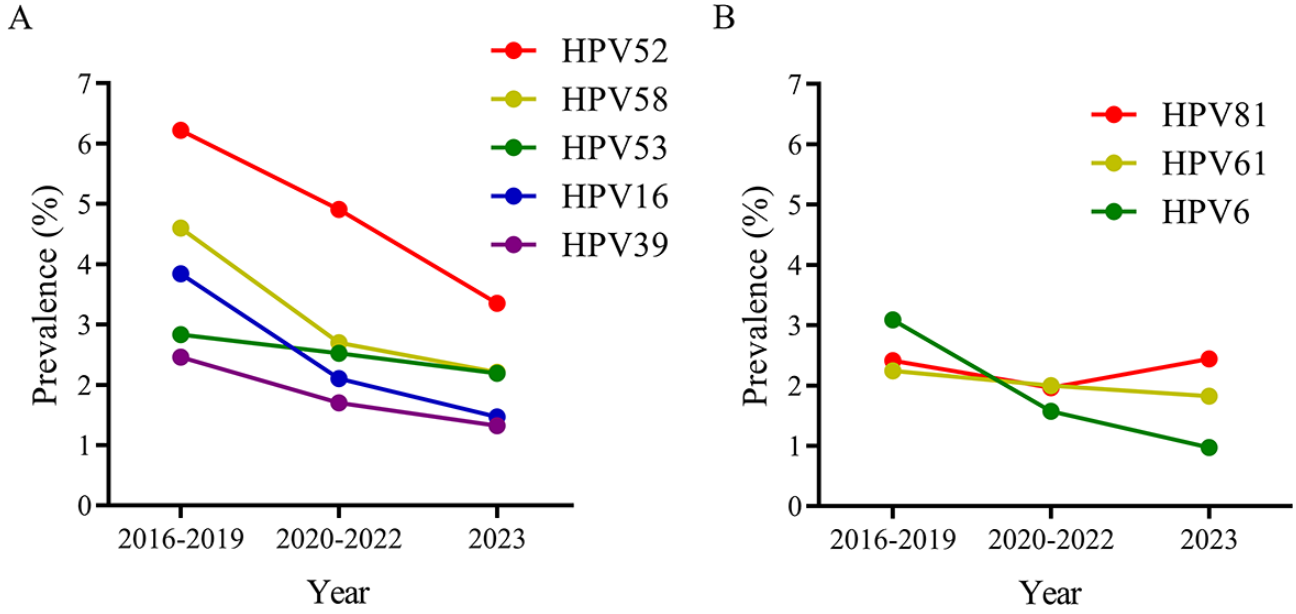
XSL•FO
RenderX

Figure 3. HPV prevalence among females stratified by time periods: 2016 - 2019, 2020 - 2022, and 2023. (**A**) Overall HPV prevalence, including the prevalence of both HR-HPV and LR-HPV genotypes. (**B**) Prevalence of vaccine-targeted HPV genotypes. 2V-HPV genotypes (HPV16 and 18) targeted by the bivalent vaccine, 4V-HPV genotypes (HPV16, 18, 6, and 11) targeted by the quadrivalent vaccine, and 9V-HPV genotypes (HPV16, 18, 6, 11, 31, 33, 45, 52, and 58) targeted by the nine-valent vaccine. The Cochran-Armitage trend test was applied to assess the differences in HPV prevalence among the 2016 - 2019, 2020 - 2022, and 2023 groups. 2V-HPV: bivalent human papillomavirus; 4V-HPV: quadrivalent human papillomavirus; 9V-HPV:nine-valent human papillomavirus; HPV: human papillomavirus; HR-HPV: high-risk human papillomavirus; LR-HPV: low-risk human papillomavirus.



We further examined how the most prevalent HR-HPV and LR-HPV genotypes changed across the 3 time periods (2016 - 2019, 2020 - 2022, and 2023) to align with broader trends (Figure 4A and B; Table S3 in). All 5 top HR-HPV genotypes showed significant downward trends (all $P<.001$): HPV52 (46.14% reduction), HPV58 (51.96% reduction), HPV53 (22.61% reduction), HPV16 (61.72% reduction), and HPV39 (46.34% reduction) from 2016 - 2019 to 2023. Notably, HPV39, which is not targeted by current vaccines, also exhibited a marked decline. Among the top LR-HPV genotypes, HPV61 and HPV6 also showed significant decreases from 2016 - 2019 to 2023 (18.75% reduction for HPV61; $P=.003$; and 68.61% reduction for HPV6; $P<.001$). In contrast, HPV81 displayed a V-shaped trend with no overall significant temporal pattern ($P=.46$); however, its prevalence increased from 1.96% in 2020 - 2022 to 2.44% in 2023 ($P<.001$).

Figure 4. Prevalence of the top 5 high-risk HPV and top 3 low-risk HPV genotypes among females, stratified by time periods: 2016 - 2019, 2020 - 2022, and 2023. (**A**) The top 5 high-risk HPV genotypes. (**B**) The top 3 low-risk HPV genotypes. The Cochran-Armitage trend test was applied to assess the differences in HPV prevalence among the 2016 - 2019, 2020 - 2022, and 2023 groups. The chi-square test also revealed a significant difference in HPV81 prevalence between the 2023 and 2020 - 2022 groups (2.44% vs 1.96%; $P<.001$). HPV: human papillomavirus.



Collectively, these findings highlight dynamic shifts in HPV epidemiology over the 8-year period, with widespread declines

XSL•FO
RenderX

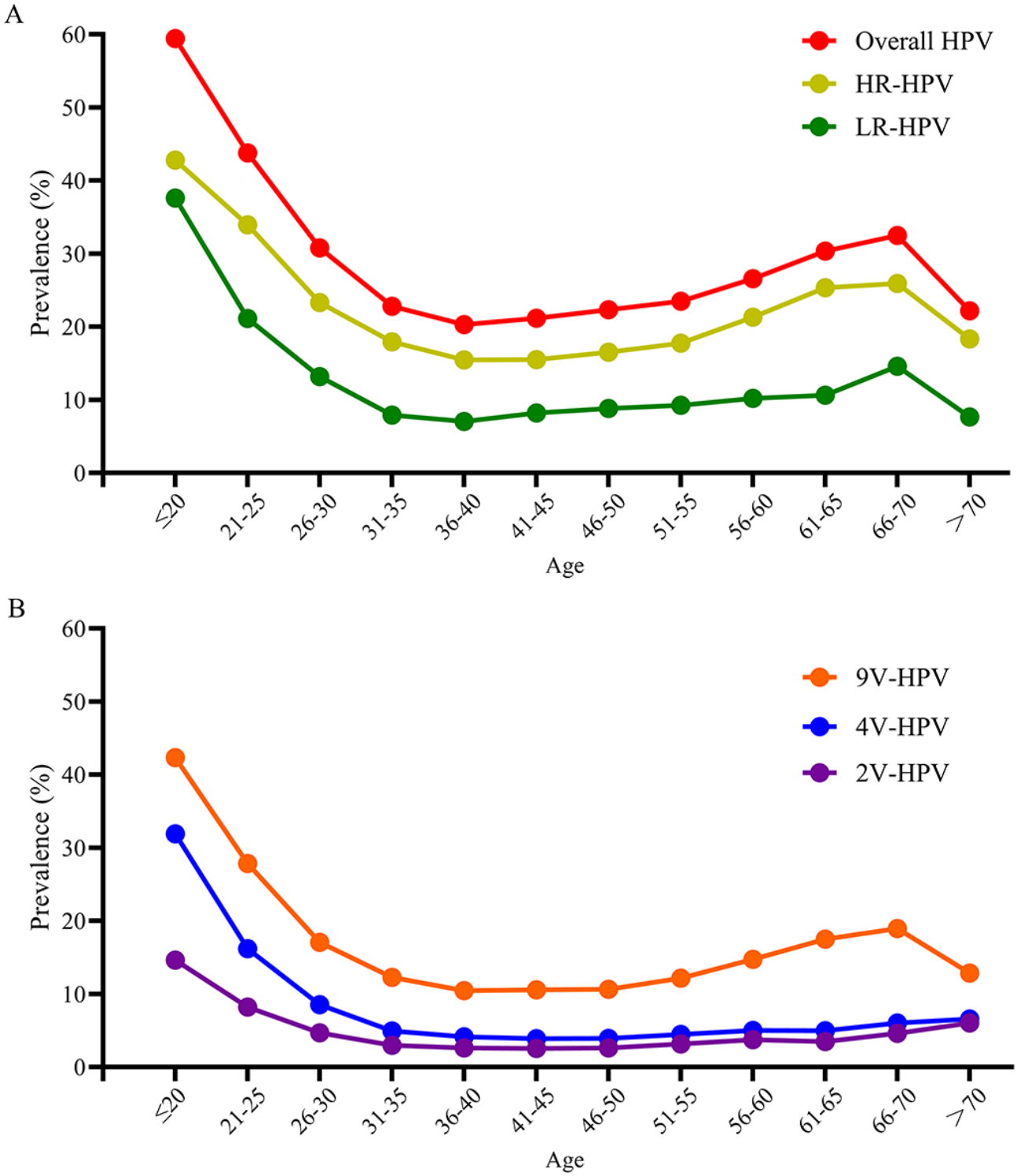
in the prevalence of overall HPV, HR-HPV, LR-HPV, and vaccine-targeted HPV genotypes, alongside distinct trends in specific genotypes.

Prevalence of Overall HPV and Vaccine-Targeted Genotypes by 5-Year Age Intervals (2016–2023)

To investigate the influence of age on HPV infection, we assessed the prevalence of overall HPV and vaccine-targeted HPV genotypes among females stratified by 5-year age intervals. Overall, the prevalence of HPV (including overall HPV, HR-HPV, LR-HPV, 2V-HPV, 4V-HPV, and 9V-HPV) showed a significant declining trend with increasing age (P<.001). Specifically, HPV prevalence varied significantly by age, with distinct patterns observed across 5-year intervals (Figure 5A and B; Table S4 in). The prevalence of overall HPV, HR-HPV, LR-HPV, and 9V-HPV peaked in the ≤20 years age group, with prevalence of 59.45% (365/614) for overall HPV, 42.83% (263/614) for HR-HPV, 37.62% (231/614) for LR-HPV, and 42.35% (260/614) for 9V-HPV, respectively, followed by a gradual decline in the 21-25 and 26- to 30-year age groups. Prevalence remained relatively stable from 31 - 35 to 51 - 55 years, then increased slowly starting at 56 - 60 years, with a smaller secondary peak observed at 66 - 70 years before declining again in the >70-year group. For 2V-HPV and 4V-HPV, prevalence also peaked in the ≤20-year age group at 14.66% (90/614) for 2V-HPV and 31.92% (196/614) for 4V-HPV, respectively, followed by a gradual decline in the 21 - to 25-year age group and stable prevalence thereafter (26 - 30 to >70 years). Stratification by 5-year intervals revealed that the youngest age group (≤20 years) consistently exhibited the highest prevalence of both overall HPV and vaccine-targeted genotypes, underscoring the importance of early screening and vaccination initiatives in this population.

XSL•FO
RenderX

Figure 5. Prevalence of overall HPV and vaccine-targeted HPV genotypes by 5-year age intervals (2016 - 2023). (**A**) Overall HPV prevalence, including the prevalence of both HR-HPV and LR-HPV genotypes. (**B**) Prevalence of vaccine-targeted HPV genotypes. 2V-HPV genotypes (HPV16 and 18) targeted by the bivalent vaccine, 4V-HPV genotypes (HPV16, 18, 6, and 11) targeted by the quadrivalent vaccine, and 9V-HPV genotypes (HPV16, 18, 6, 11, 31, 33, 45, 52, and 58) targeted by the nine-valent vaccine. The Cochran-Armitage trend test was used to assess the trend of HPV prevalence changes across age groups. 2V-HPV: bivalent human papillomavirus; 4V-HPV: quadrivalent human papillomavirus; 9V-HPV: nine-valent human papillomavirus; HPV: human papillomavirus; HR-HPV: high-risk human papillomavirus; LR-HPV: low-risk human papillomavirus.



Discussion

Principal Findings

The infection of HPV in the cervical area, especially the persistent infection of HR-HPV, is closely related to cervical cancer. The high prevalence of HPV poses a challenge to the prevention and control of cervical cancer. HPV prevalence is an important basis for analyzing and formulating cervical cancer prevention and control strategies. The prevalence and genotype distribution of HPV varied in different areas, even in the same areas during different periods. Our results showed that the prevalence of overall HPV was 25.24% (16,039/63,553), which aligned with findings from Zhejiang (22.3% in all participants

XSL•FO
RenderX

and 27.3% in the gynecological clinic subgroup) [24]. However, our HPV prevalence was higher than the 18.4% reported in Xiamen [17] and the 8.60% observed in Putian, Fujian Province [18]. Several factors may account for the observed differences. First, the number of HPV genotypes included in the testing differed. Our study analyzed 27 HPV genotypes, whereas the Xiamen [17] and Putian [18] studies only included 21 genotypes. Second, population composition significantly influences HPV prevalence. Inpatient and outpatient populations often include individuals with cervical cancer or precancerous lesions, who tend to have higher HPV prevalence. In contrast, health examination populations primarily consist of asymptomatic individuals undergoing routine screening, resulting in lower HPV prevalence. Our study included participants from health examination centers, outpatient departments, and inpatient wards. Consequently, the overall HPV prevalence was similar to that reported in the study by Yan et al [24] for a similar population group and fell between the rates observed in health examination and outpatient groups. Third, temporal variations in HPV prevalence may influence prevalence differences. Our study spanned 2016 - 2017, during which the HPV prevalence was relatively high compared to that reported by Lin et al [18]. Consistent with Yao et al [17], our data also showed a downward trend in the prevalence of overall HPV, HR-HPV, and LR-HPV from 2016 to 2023. Our findings indicate that the HPV prevalence among females in Xiamen is still relatively high, although showing a downward trend over the years.

Several studies have identified the top 5 HR-HPV genotypes as HPV52, 16, 58, 51, and 39 in Beijing [25], Luoyang of Henan Province [26], Xiamen [17], and Putian [18] of Fujian Province. Notably, our study and a recent investigation in Chengdu-Chongqing (2014 - 2023) [27] consistently observed HPV52, 58, 53, 16, and 39 as the predominant genotypes, with HPV53 emerging as a novel top-ranked genotype across regions. HPV53 is not targeted by authorized licensed HPV vaccines currently. Even the classification of HPV53 is still unclear. The International Agency for Research on Cancer has not yet classified it as HR-HPV, while an increasing number of studies [16,18,19,26-28] have done so. Therefore, it is essential to enhance HPV53 detection and research on its association with cervical cancer.

While LR-HPV genotypes such as HPV6 and HPV11 are mainly associated with genital warts [29,30] with low malignant potential, persistent infections impose significant burdens on the patient's appearance, psychology, and finances [27,31]. HPV6 and HPV11 also have long been considered the most common LR-HPV genotypes. However, recent studies have found that HPV81 is the most common LR-HPV genotype in exfoliated cervical cells [19,27], a finding corroborated by our study. Notably, HPV81 exhibited a rising prevalence trend (1.96% in 2020 - 2022 vs 2.44% in 2023), contrasting with the declining prevalence of HPV6. This upward trajectory aligns with a recent study showing HPV81 as the most frequently detected LR-HPV genotype in exfoliated cervical cells, with a statistically significant increasing trend [27]. Given that HPV81 exhibits strong survival competitiveness and is not included in any licensed HPV vaccines, its prevalence trend and pathogenicity require urgent attention.

HPV vaccination is a primary strategy for preventing HPV infection and associated diseases. A meta-analysis involving 60 million individuals showed an 80% reduction in the prevalence of HPV16 and HPV18 more than 8 years postvaccination [32]. Currently, licensed prophylactic HPV vaccines mainly include the bivalent, quadrivalent, and nine-valent vaccines. Our research shows that the prevalence rates of 2V-HPV, 4V-HPV, and 9V-HPV among females in Xiamen were 3.56% (2264/63,553), 5.89% (3746/63,553), and 13.64% (8666/63,553), respectively, with a downward trend observed from 2016 to 2023. While vaccine coverage expansion likely contributed to declining vaccine-genotype HPV prevalence, this trend cannot be exclusively attributed to immunization programs. Confounding factors, including shifts in screening practices, sexual behavior dynamics, and demographic changes over the study period, may also influence these outcomes. Notably, the relatively high persistence of 9V-HPV prevalence underscores the continued importance of expanding vaccination coverage as a key component of comprehensive strategies to prevent and control HPV infection and its associated sequelae, including cervical cancer.

HPV prevalence exhibited an age-specific pattern, with a significant age-related decline observed across age groups. In our study, the highest prevalence was observed among females aged ≤20 years, aligning with findings from previous studies [22,26]. This elevated prevalence in younger females may be attributed to frequent sexual activity and lack of immunity to HPV [33]. Thus, targeted interventions are warranted, including the development of age-tailored vaccination strategies and prioritization of HPV immunization for females aged ≤20 years.

As a sexually transmitted virus, HPV prevalence is influenced not only by sexual behavior patterns and activity levels [34] but also by a complex interplay of behavioral, socioeconomic, and environmental factors. Human activity has been profoundly changed due to the emergence of the COVID-19 pandemic. Several studies found that the COVID-19 pandemic harmed cancer screening and vaccination [35,36]. Our analysis identified a notable decline in the prevalence of overall HPV, HR-HPV, LR-HPV, 2V-HPV, 4V-HPV, and 9V-HPV; and specific genotypes, such as HPV52, HPV58, HPV16, HPV39, and HPV6 during and postpandemic compared to prepandemic periods—a finding corroborated by other regional studies [19,21]. Our study found that HPV testing volumes peaked at 19,826 in 2023 and 14,886 in 2020 during the study period, likely driven by policy interventions: a free domestic bivalent HPV vaccination program launched in September 2020 and free cervical cancer screening initiatives introduced in October 2022. These policies probably motivated greater participation in HPV testing, including asymptomatic or healthy women. The COVID-19 pandemic also influenced testing uptake indirectly. Thus, the observed decline in HPV prevalence reflects synergistic effects of pandemic-related behavioral changes, heightened health awareness, and expanded access to preventive services.

Limitations

This study has several limitations. First, detailed information regarding HPV vaccination for each subject was not obtained, and an in-depth investigation into the influence of economic

XSL•FO
RenderX

and cultural factors on HPV prevalence was lacking. Second, selection bias may exist as hospital-based data primarily reflect testing behaviors rather than the true prevalence at the population level. Third, we did not document participants' symptom status, cervical conditions, reproductive status, or immunological status. Consequently, it was impossible to differentiate between participants seeking care due to symptoms or undergoing follow-up for cervical conditions and those undergoing routine screening. This prevented us from exploring the relationship between HPV prevalence and the incidence rates of cervical cancer or cervical epithelial cell lesions. Pregnant women and immunocompromised individuals may exhibit distinct HPV infection patterns. We could neither exclude nor characterize these subgroups due to incomplete clinical documentation. This may introduce heterogeneity into the study population, as HPV prevalence and clearance rates could be influenced by hormonal changes during pregnancy or immune dysfunction. Fourth, detailed data on residential location, sociodemographic factors, behaviors, and clinical covariates were unavailable, hindering stratified analyses of their impact on HPV prevalence. These limitations underscore the necessity for future studies to proactively collect comprehensive data on clinical, geographic, demographic, and behavioral aspects to enable stratified analyses of vulnerable subgroups.

Conclusions

HPV prevalence among females in Xiamen remains high, declining from 2016 to 2023. HPV52, HPV58, HPV53, HPV16, and HPV39 were the top 5 HR-HPV genotypes in Xiamen, whereas HPV81 was the most prevalent LR-HPV genotype, with an upward trend after the COVID-19 pandemic. Non–vaccine-targeted HPV genotypes demonstrated high prevalence, highlighting the need for developing and implementing broader spectrum HPV vaccines. In addition, further attention should be paid to the prevention and control of HPV infection in females aged ≤20 years.

Disclaimer

This research received no external funding and no generative artificial intelligence was used in the creation of this article.

Data Availability

The data associated with this study have not been deposited into a publicly available repository and will be made available upon reasonable request.

Authors' Contributions

GZ and QC conceptualized and designed the study. ZC and HX collected and interpreted the data. SW, EL (undergraduate student), and MC (undergraduate student) analyzed the data. QC, QL, and ZC wrote the original article. GZ reviewed the article. Authors QC, HX, and GZ are co-corresponding authors for this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data on human papillomavirus prevalence.
[DOC File, 128 KB - publichealth_v11i1e70507_app1.doc]

Checklist 1
Strengthening the Reporting of Observational Studies in EpidemiologyReporting of Observational Studies in Epidemiology statement and data on human papillomavirus prevalence.
[DOC File, 60 KB - publichealth_v11i1e70507_app2.doc]

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
2. Xia C, Dong X, Li H, et al. Cancer statistics in China and United States, 2022: profiles, trends, and determinants. Chin Med J (Engl) 2022 Feb 9;135(5):584-590. [doi: 10.1097/CM9.0000000000002108] [Medline: 35143424]
3. Pei W, Li J, Lei S, Nie S, Liu L. Burden of major cancers in China attributable to modifiable risk factors: predictions from 2012 to 2035. Int J Cancer 2025 Apr 1;156(7):1369-1379. [doi: 10.1002/ijc.35233] [Medline: 39503513]
4. Lin S, Gao K, Gu S, et al. Worldwide trends in cervical cancer incidence and mortality, with predictions for the next 15 years. Cancer 2021 Nov 1;127(21):4030-4039. [doi: 10.1002/cncr.33795] [Medline: 34368955]

XSL•FO
RenderX

5. Arroyo Mühr LS, Lagheden C, Hassan SS, Eklund C, Dillner J. The International Human Papillomavirus Reference Center: standardization, collaboration, and quality assurance in HPV research and diagnostics. J Med Virol 2023 Dec;95(12):38115556. [doi: 10.1002/jmv.29332] [Medline: 38115556]
6. McBride AA. Human papillomaviruses: diversity, infection and host interactions. Nat Rev Microbiol 2022 Feb;20(2):95-108. [doi: 10.1038/s41579-021-00617-5] [Medline: 34522050]
7. Bouvard V, Baan R, Straif K, et al. A review of human carcinogens--Part B: biological agents. Lancet Oncol 2009 Apr;10(4):321-322. [doi: 10.1016/s1470-2045(09)70096-8] [Medline: 19350698]
8. Cogliano VJ, Baan R, Straif K, et al. Preventable exposures associated with human cancers. J Natl Cancer Inst 2011 Dec 21;103(24):1827-1839. [doi: 10.1093/jnci/djr483] [Medline: 22158127]
9. Wei F, Georges D, Man I, Baussano I, Clifford GM. Causal attribution of human papillomavirus genotypes to invasive cervical cancer worldwide: a systematic analysis of the global literature. Lancet 2024 Aug 3;404(10451):435-444. [doi: 10.1016/S0140-6736(24)01097-3] [Medline: 39097395]
10. Arroyo Mühr LS, Gini A, Yilmaz E, et al. Concomitant human papillomavirus (HPV) vaccination and screening for elimination of HPV and cervical cancer. Nat Commun 2024 May 1;15(1):3679. [doi: 10.1038/s41467-024-47909-x] [Medline: 38693149]
11. Patel C, Brotherton JM, Pillsbury A, et al. The impact of 10 years of human papillomavirus (HPV) vaccination in Australia: what additional disease burden will a nonavalent vaccine prevent? Euro Surveill 2018 Oct;23(41):30326995. [doi: 10.2807/1560-7917.ES.2018.23.41.1700737] [Medline: 30326995]
12. Kavanagh K, Pollock KG, Cuschieri K, et al. Changes in the prevalence of human papillomavirus following a national bivalent human papillomavirus vaccination programme in Scotland: a 7-year cross-sectional study. Lancet Infect Dis 2017 Dec;17(12):1293-1302. [doi: 10.1016/S1473-3099(17)30468-1] [Medline: 28965955]
13. Du J, Ährlund-Richter A, Näsman A, Dalianis T. Human papilloma virus (HPV) prevalence upon HPV vaccination in Swedish youth: a review based on our findings 2008-2018, and perspectives on cancer prevention. Arch Gynecol Obstet 2021 Feb;303(2):329-335. [doi: 10.1007/s00404-020-05879-7] [Medline: 33247317]
14. Oyervides-Muñoz MA, Pérez-Maya AA, Sánchez-Domínguez CN, et al. Multiple HPV infections and viral load association in persistent cervical lesions in Mexican women. Viruses 2020 Mar 31;12(4):380. [doi: 10.3390/v12040380] [Medline: 32244347]
15. Li K, Li Q, Song L, Wang D, Yin R. The distribution and prevalence of human papillomavirus in women in mainland China. Cancer 2019 Apr 1;125(7):1030-1037. [doi: 10.1002/cncr.32003] [Medline: 30748006]
16. Chen Q, Qu W, Zhao Y, Shu L, Wang Y, Chen X. The prevalence of HPV among 164,137 women in China exhibited some unique epidemiological characteristics. Infect Agent Cancer 2023 Nov 10;18(1):72. [doi: 10.1186/s13027-023-00553-4] [Medline: 37950328]
17. Yao X, Li Q, Chen Y, et al. Epidemiology of human papillomavirus infection in women from Xiamen, China, 2013 to 2023. Front Public Health 2013;12(1332696):38590815. [doi: 10.3389/fpubh.2024.1332696]
18. Lin K, Hong Q, Fu Y, et al. Cervical HPV infection and related diseases among 149,559 women in Fujian: an epidemiological study from 2018 to 2023. Front Microbiol 2024;15:1418218. [doi: 10.3389/fmicb.2024.1418218] [Medline: 38962121]
19. Lin B, Zhang F, Liu F, et al. The prevalence and genotype distribution of human papillomavirus in central Fujian Province during the COVID-19 pandemic. Virol J 2024 Jun 5;21(1):129. [doi: 10.1186/s12985-024-02393-z] [Medline: 38840267]
20. Wei L, Ma L, Qin L, Huang Z. The prevalence and genotype distribution of human papillomavirus among women in Guangxi, southern China. Infect Agent Cancer 2022 Apr 21;17(1):19. [doi: 10.1186/s13027-022-00431-5] [Medline: 35449004]
21. Zhang H, Li X, Yang Z, et al. Influence of COVID-19 pandemic on prevalence and genotype distribution of HPV in cervical cancer screening population. Virol J 2024 Oct 23;21(1):261. [doi: 10.1186/s12985-024-02497-6] [Medline: 39443958]
22. Shen Y, Huang Y, Wang W, et al. Prevalence and genotype distribution of HPV infection among women in Xiamen, China. Front Microbiol 2023;14:1130226. [doi: 10.3389/fmicb.2023.1130226] [Medline: 37323886]
23. Ma Y, Xia X, Zheng W, Dai Y, Zhuang X. HPV prevalence and genotype distribution among women in eastern China during the Covid-19 pandemic. Hum Vaccin Immunother 2023 Dec 31;19(1):2212571. [doi: 10.1080/21645515.2023.2212571] [Medline: 37226673]
24. Yan X, Shen L, Xiao Y, Wang Q, Li F, Qian Y. Prevalence, characteristics, and distribution of HPV genotypes in women from Zhejiang Province, 2016-2020. Virol J 2021 Oct 20;18(1):208. [doi: 10.1186/s12985-021-01676-z] [Medline: 34670576]
25. Zhang W, Guo N, Li B, et al. Prevalence and genotype distribution of human papillomavirus infections in Beijing, China between 2016 and 2020. Virol J 2023 Jan 18;20(1):11. [doi: 10.1186/s12985-023-01959-7] [Medline: 36653807]
26. Wang Y, Han S, Wang X, Song S, Wang X. Characteristics of human papillomavirus infection among females and the genetic variations of HPV18 and HPV58 in Henan Province, China. Sci Rep 2023 Feb 8;13(1):2252. [doi: 10.1038/s41598-022-24641-4] [Medline: 36755053]
27. Chen Z, Li Q, Xu Y, et al. Study on the detection rate, genetic polymorphism, viral load, persistent infection capacity, and pathogenicity of human papillomavirus type 81. J Med Virol 2024 Jun;96(6):38895800. [doi: 10.1002/jmv.29753] [Medline: 38895800]

XSL•FO
RenderX

28. Huang Z, Yao S, Zou L, et al. Evaluation of HPV infection and presence of licensed HPV vaccine genotypes among genital warts in Foshan, China. Front Microbiol 2024;15(1376141):38699478. [doi: 10.3389/fmicb.2024.1376141] [Medline: 38699478]
29. Steben M, Garland SM. Genital warts. Best Pract Res Clin Obstet Gynaecol 2014 Oct;28(7):1063-1073. [doi: 10.1016/j.bpobgyn.2014.07.002] [Medline: 25155525]
30. Yuan H, Li R, Lv J, et al. Epidemiology of human papillomavirus on condyloma acuminatum in Shandong Province, China. Hum Vaccin Immunother 2023 Dec 31;19(1):2170662. [doi: 10.1080/21645515.2023.2170662] [Medline: 36919446]
31. Vorsters A, Bosch FX, Poljak M, et al. HPV prevention and control - The way forward. Prev Med 2022 Mar;156(106960):106960. [doi: 10.1016/j.ypmed.2022.106960] [Medline: 35065979]
32. Drolet M, Bénard É, Pérez N, Brisson M, HPV Vaccination Impact Study Group. Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis. Lancet 2019 Aug 10;394(10197):497-509. [doi: 10.1016/S0140-6736(19)30298-3] [Medline: 31255301]
33. Mai Q, Yang X, Cheng H, Wu G, Wu Z. Prevalence and genotype distribution of human papillomavirus among women with cervical lesions in Shenzhen city, China. Hum Vaccin Immunother 2021 Apr 3;17(4):965-971. [doi: 10.1080/21645515.2020.1805993] [Medline: 32961080]
34. Zou K, Huang Y, Li Z. Prevention and treatment of human papillomavirus in men benefits both men and women. Front Cell Infect Microbiol 2022;12:1077651. [doi: 10.3389/fcimb.2022.1077651] [Medline: 36506029]
35. Shet A, Carr K, Danovaro-Holliday MC, et al. Impact of the SARS-CoV-2 pandemic on routine immunisation services: evidence of disruption and recovery from 170 countries and territories. Lancet Glob Health 2022 Feb;10(2):e186-e194. [doi: 10.1016/S2214-109X(21)00512-X] [Medline: 34951973]
36. Ortiz AP, Gierbolini-Bermúdez A, Ramos-Cartagena JM, et al. Cervical cancer screening among medicaid patients during natural disasters and the COVID-19 pandemic in Puerto Rico, 2016 to 2020. JAMA Netw Open 2021 Oct 1;4(10):e2128806. [doi: 10.1001/jamanetworkopen.2021.28806] [Medline: 34652451]

Abbreviations

2V-HPV: bivalent human papillomavirus
4V-HPV: quadrivalent human papillomavirus
9V-HPV: nine-valent human papillomavirus
HPV: human papillomavirus
HR-HPV: high-risk human papillomavirus
LR-HPV: low-risk human papillomavirus
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani; submitted 23.12.24; peer-reviewed by A Ouyahia, X Yao; revised version received 03.09.25; accepted 25.09.25; published 16.10.25.

Please cite as:
Chen Z, Lin E, Chen M, Wang S, Lin Q, Chen Q, Xue H, Zhang G
Temporal Trends in Cervical Human Papillomavirus Prevalence Among Females in Xiamen, China (2016-2023): Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e70507
URL: https://publichealth.jmir.org/2025/1/e70507
doi:10.2196/70507

© Zhuju Chen, Erjuan Lin, Mo Chen, Shihan Wang, Qiuyuan Lin, Qingquan Chen, Heng Xue, Guanbin Zhang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of Machine Learning–Based Risk Prediction Models to Predict Rapid Weight Gain in Infants: Analysis of Seven Cohorts

Miaobing Zheng[1,2], PhD; Yuxin Zhang[1], PhD; Rachel A Laws[1], PhD; Peter Vuillermin[3], PhD; Jodie Dodd[4], PhD; Li Ming Wen[5], PhD; Louise A Baur[5], PhD; Rachael Taylor[6], PhD; Rebecca Byrne[7], PhD; Anne-Louise Ponsonby[8], PhD; Kylie D Hesketh[1], PhD

[1]Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Geelong, Australia

[2]School of Health Sciences, Faculty of Health & Medicine, UNSW Sydney, Wallace Wurth Building, Kensington, Australia

[3]Barwon Health, Geelong, Australia

[4]Discipline of Obstetrics and Gynaecology, The Robinson Research Institute, The University of Adelaide, Adelaide, Australia

[5]School of Public Health and Sydney Medical School, The University of Sydney, Sydney, Australia

[6]Department of Medicine, University of Otago, Dunedin, New Zealand

[7]School of Exercise and Nutrition Sciences, Faculty of Health, Queensland University of Technology, Kelvin Grove, Australia

[8]The Florey Institute of Neuroscience and Mental Health, Murdoch Children's Research Institute, Royal Children's Hospital, The University of Melbourne, Parkville, Australia

Corresponding Author:
Miaobing Zheng, PhD
Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Geelong, Australia

Abstract

Background: Rapid weight gain (RWG) during infancy, defined as an upward crossing of one centile line on a weight growth chart, is highly predictive of subsequent obesity risk. Identification of infant RWG could facilitate obesity risk assessment from infancy.

Objective: Leveraging machine learning (ML) algorithms, this study aimed to develop and validate risk prediction models to identify infant RWG by the age of 1 year.

Methods: Data from 7 Australian and New Zealand cohorts were pooled for risk model development and validation (n=5233). A total of 8 ML algorithms predicted infant RWG using routinely available prenatal and early postnatal factors, including maternal prepregnancy weight status, maternal smoking during pregnancy, gestational age, parity, infant sex, birth weight, any breastfeeding and timing of solids introduction at the age of 6 months. Pooled data were randomly split into a training dataset (70%) and a test dataset (30%) for model training and validation, respectively. Model consistency was evaluated using 5-fold cross-validation. Model predictive performance was evaluated by area under the receiver operating characteristic (ROC) curve (AUC), accuracy, precision, sensitivity, specificity, and Cohen κ.

Results: The average prevalence of infant RWG was 27%. In the training dataset, all ML algorithms showed acceptable to excellent discrimination with AUCs ranging from 0.75 to 0.86. Accuracy, which indicates the overall correctness of the model, ranged from 0.69 to 0.78. Precision, which measures the model's ability to avoid false positives, ranged from 0.68 to 0.77. The spread of sensitivity, specificity, and Cohen κ of all models was 0.68 - 0.80, 0.65 - 0.78, and 0.38 - 0.56, respectively. Of the 8 algorithms, the Gradient Boosting model showed the most favorable predictive accuracy. Validation of the Gradient Boosting model in the testing dataset exhibited excellent discrimination (AUC 0.3 - 0.6) and good ability to make accurate predictions, particularly true positive cases (with accuracy and sensitivity>0.75), but modest performance for precision (0.57 - 0.60) and Cohen κ (0.47 - 0.52).

Conclusions: This study developed the first set of ML-based risk prediction models to identify infants' risk of experiencing RWG by the age of 1 year with acceptable accuracy. The models could be feasibly integrated into routine child growth monitoring and may facilitate population-wide early obesity risk assessment in primary health care.

(JMIR Public Health Surveill 2025;11:e69220) doi:10.2196/69220

KEYWORDS

infants; rapid weight gain; risk prediction; machine learning; childhood obesity; pooled analysis

XSL•FO
RenderX

Introduction

Overweight and obesity are an intractable global challenge with substantial health consequences [1]. Compelling evidence suggests that obesity risk originates early in life and is hard to reverse once developed [2]. World Health Organization (WHO) global estimates indicate that a high prevalence of overweight and obesity is already evident in children younger than 5 years, with 37 million cases being reported in 2022 [3]. The WHO has thus recognized "Lowering the prevalence of overweight in children under the age of five" as one of the key outcome targets in the acceleration plan to stop obesity [4]. Identification of obesity risk from early life along with timely delivery of tailored interventions, are therefore strongly recommended to achieve effective and sustained impact.

Extensive research has attempted to develop obesity risk prediction models to identify high-risk populations [5], mostly in older children and adults when obesity and modifiable risk factors are already established [6,7]. Despite some models having been developed to predict early childhood obesity risk, these models are largely derived from a single cohort using statistical approaches and are often constrained to linear relationship assumptions, resulting in suboptimal model prediction accuracy and low generalisability [8]. A key consideration for developing clinically useful risk prediction models is clinical impact and ease of clinical integration [9]. It is not possible to evaluate the clinical impact of these models due to the costs and ethical reasons associated with long-term follow-up and allowing children to develop obesity without intervention. Moreover, widespread obesity stigma and the reluctance to label children as having obesity further impede the clinical integration of existing obesity risk models [10,11]. Thus, novel approaches to facilitate early obesity risk identification, communication, and model integration in clinical practice are needed.

Convincing evidence suggests that accelerated physical growth in early life is highly predictive of later obesity and cardiometabolic outcomes [12,13]. Several systematic reviews concluded that rapid weight gain (RWG) from birth to the age of 1 year, defined as the upward centile crossing in a weight growth chart, was associated with a fourfold higher risk of later obesity than those without infant RWG [14,15]. As such, infant RWG is considered a sensitive proxy marker that denotes future obesity risk. A risk prediction model to predict infant RWG could facilitate early obesity risk identification from infancy and the timely delivery of interventions before adverse health behaviors are established. Furthermore, such a model could be easily evaluated and integrated in primary health care where infant growth is already routinely monitored. To improve risk prediction accuracy, we propose the use of machine learning (ML), a type of artificial intelligence technique that is powerful at making accurate predictions. Reviews have indicated that ML models outperform statistical models for obesity risk prediction [16,17]. Using data from 7 Australian and New Zealand cohorts, we used ML to develop and validate risk prediction models to identify infants at risk of experiencing RWG by the age of 1 year.

Methods

Data Source and Variables

Data from 7 Australian and New Zealand cohorts: Barwon Infant Study (BIS) [18], Healthy Beginnings [19,20], Infant Feeding Activity and Nutrition Trial (InFANT) [21,22], InFANT Extend [23], LIMIT [24,25], NOURISH [26,27], and Prevention of Overweight in Infancy (POI) were used [28,29]. Baseline recruitment of all studies was undertaken either antenatally (15 - 34 wk gestation) or postnatally (between ages 3 - 4 m) from 2007 to 2012. The 7 studies include 6423 parent-infant dyads from diverse geographical areas at the study baseline. BIS is a prospective birth cohort. The remaining studies are intervention studies targeting nutrition and physical activity in infants or pregnant women. Detailed study characteristics of the cohorts were published elsewhere [30,31]. None of the intervention studies showed differences in infant growth outcomes between intervention and control groups by the age of 1 year [32], and thus groups were combined to study prognosis [33].

Rapid Weight Gain During Infancy

All cohorts collected child anthropometrics at birth and around 1 year of age. Child length and weight at birth were transcribed from child health records. Child length and weight around the age of 1 year were measured by trained staff. WHO growth standards were used to calculate age- and sex-specific weight-for-age *z* scores at birth and around the age of 1 year, and the difference between the two time points was calculated. RWG during infancy was defined as a change in weight-for-age *z* score>0.67, which is clinically equivalent to upward crossing one centile line in a weight growth chart [14,15].

Selection of Predictor Variables

A range of prenatal and early postnatal factors were collected across all the cohorts. Prenatal factors include maternal country of birth (native-born vs overseas-born), an education level (tertiary vs nontertiary), smoking during pregnancy (smoker vs nonsmoker), and pre or early-pregnancy BMI (kg/m^2). Child early postnatal factors included sex (boys vs girls), birth weight (kg), gestational age (weeks), parity (≥1 sibling vs no sibling), any breastfeeding (yes vs no) at age 6 months, and whether introduced solids (yes vs no) at age 6 months.

The selection of predictor variables for risk prediction model development was informed by previous findings on determinants of infant RWG, the availability of data or ease of data collection at primary health care visits, and the level of data missingness. Our previous statistical findings in the 7 cohorts assessed maternal prepregnancy BMI, smoking during pregnancy, country of birth, gestational age, parity, infant sex, birth weight, breastfeeding, and solids introduction at age 6 months as potential predictors of RWG by the age of 1 year [30], and these were also considered as potential predictors for the current risk prediction model development. Continuous variables capture more nuanced information compared with categorical variables, hence, continuous forms of maternal prepregnancy BMI, gestational age, and birth weight were used for risk prediction whenever possible. Maternal country of birth was not included

XSL•FO
RenderX

as a predictor as it had missing data >40%, and missing data imputation would not be reliable. A total of 3 models were purposely chosen. Model 1 is limited to variables that are collected as part of the antenatal and child health records in Australia or can be easily collected at the primary health care visits: maternal prepregnancy BMI, smoking during pregnancy, gestational age, parity, infant sex, and birth weight. Infant feeding at age 6 months is not routinely collected in Australia. However, given the vital contribution of infant feeding, especially breastfeeding, in determining the risk of infant RWG, we conducted 2 additional models to further incorporate the 2 infant feeding variables. Model 2 included model 1 variables plus any breastfeeding at the age of 6 months. Model 3 additionally included solids introduction at age 6 months, upon model 2. The 3 models also enable the comparison of risk prediction accuracy at birth versus at 6 months postnatal.

Model Development and Validation

Detailed model development and validation processes are summarised in Supplementary Figure 1 in Multimedia Appendix 1. All ML models were trained and validated using Python 3.11.8 and Scikit Learning Tool Kit library 1.6.1. Data from 7 cohorts were pooled and randomly split into a training dataset (70%) and a testing dataset (30%) [34]. The training dataset was used for model development and internal validation, from which the best-performing model was selected. The testing dataset was reserved for external validation to assess how the best-performing model performed with the new data. Missing data imputation was conducted to impute predictors with missing data: maternal prepregnancy BMI, smoking, gestational age, breastfeeding, and solids introduction. Maternal smoking during pregnancy (yes vs no) was imputed by mode imputation [35], whereby missing values were replaced with the most frequent category observed in the dataset. Median imputation was used to impute missing data on maternal prepregnancy BMI and gestational age. For breastfeeding and solids introduction, we adopted ML algorithms to impute missing values using infant sex, birth weight, and parity as predictors and those with complete data as the training dataset for missing data prediction. Detailed missing data imputation methods are provided in Supplementary Table 1 in Multimedia Appendix 1. All predictor variables were standardized before risk prediction, ensuring all variables were in the same scale to facilitate convergence speed and interpretability.

For model development in the training dataset, 8 ML algorithms widely used in study prognosis in health research were explored: Logistic Regression, Decision Tree, Random Forest, AdaBoost, Gradient Boosting, K-Neighbours, Support Vector, and Multi-Layer Perceptron [16]. The hyperparameter (configuration setting) of each model were optimized using the grid search cross-validation technique, which systematically finds the best combination of hyperparameter for a ML model [36]. The grid search cross-validation technique reduces overfitting and provides more reliable hyperparameter selection than manual tuning. Detailed model hyperparameters are presented in Supplementary Table 2 in Multimedia Appendix 1. For model training and internal validation, 5-fold cross-validation was conducted, where the training dataset was randomly split into 5 blocks of equal size (fold), and 5 independent iterations of training and internal validation were conducted, so every block of data was used for both training and internal validation [37]. Imbalanced outcome data (ie, higher proportion of infants without RWG and vs with RWG) can result in biased risk prediction models with better performance in predicting infants with no RWG. We used the Synthetic Minority Over-sampling Technique (SMOTE) [38], a commonly used method in ML, to address imbalanced outcome data. SMOTE generates synthetic samples of infants with RWG that are combinations of existing RWG infants, thereby balancing the outcome distribution and improving the prediction of infants with RWG. SMOTE was exclusively performed on the training folds, which prevents data leakage, eliminates the risk of overestimating the model's performance, and maintains a proper evaluation of the model's generalisability (Supplementary Figure 2 in Multimedia Appendix 1). The performance metrics of five iterations were averaged to obtain the mean performance metrics of each ML model. The average model performance of 8 ML models was compared, from which the best-performing models were selected for subsequent external validation in the testing dataset.

The predictive performance of the models against the observed prevalence of infant RWG in both internal and external validation was assessed by various metrics, including receiver operating characteristic (ROC) plot, precision-recall curve (PRC) plot, accuracy, precision, sensitivity, specificity, F_1-score, and Cohen κ. A ROC curve plots false positive rate (1-specificity) as the x-axis against true positive rate (sensitivity) as the y-axis at various thresholds (defined as the cutoff value at which the model determines a positive or negative prediction). The area under the ROC curve (AUC) summarizes the model's overall ability to discriminate between infants with or without RWG across various thresholds. An AUC of 0.70 - 0.79 and ≥0.80 is considered acceptable and excellent discrimination, respectively [39]. Accuracy evaluates the overall correctness of the model (both true positives and negatives). Precision examines the model's ability to predict true positives while minimizing false positives. Sensitivity (recall) assesses the model's ability to correctly classify true positives. The PRC plots recall (y-axis) against precision (x-axis) over different thresholds, which is particularly useful for a classification model with imbalanced outcome data. The closer the curve toward the top right corner (high precision and high recall) or the larger area under the PRC (AUPRC), the better model performance with >0.7 is generally considered as decent model performance [40]. Likewise, F_1-score evaluates the balanced performance of the model that accounts for both precision and sensitivity, which is particularly useful for risk prediction with the outcome that has an uneven distribution. Specificity measures the model's ability to correctly identify true negatives. Higher values of these metrics indicate better model performance. Cohen κ assesses agreement between predicted and observed outcomes, accounting for agreement by chance, and is also useful for imbalanced outcome data (0.61 - 1.0: perfect agreement; 0.41 - 0.60 moderate agreement; 0.21 - 0.40 fair agreement; ≤0.20 slight or poor agreement). Detailed interpretation and calculation of these metrics are provided in Supplementary Table 3 in Multimedia Appendix 1. To enhance model transparency and interpretability, additional Shapley Additive Explanations (SHAP) analyses were

XSL•FO
RenderX

performed to quantify the contribution of each predictor in the risk prediction of infant RWG.

Ethical Considerations

Each study received ethical approval from relevant institutional ethics committees and informed written consent was obtained from participants. The original consent covers secondary analysis without additional consent. The current analysis uses deidentified data and received an ethics exemption from the Deakin University Human Research Ethics Committee (2023-033).

Results

Of 6423 parent-infant dyads from the 7 cohorts at baseline, 81% (n=5233) with all required data were included in analyses. The mean prepregnancy BMI was 27.7 kg/m^2. The majority of women did not smoke during pregnancy (89.1%). The mean gestational age and infant birth weight were 39 weeks and 3.5 kg, respectively. The pooled sample included an even proportion of males and females with 64.1% being first-born child. Most infants were breastfed beyond age 6 months (59.5%) and started solids at or after age 6 months (79.9%). Across the 7 cohorts, the mean age (SD) at RWG assessment ranged from 7.0 (1.9) to 13.1 (1) months [30]. The average prevalence of RWG across 7 cohorts spanned from 21.6% to 41.3% [30]. The mean age and prevalence of RWG of 7 cohorts were 10.3 (SD 3) months and 27.4%, respectively (Table 1).

Table . Characteristics of the pooled sample from 7 cohorts (N=5233).

Maternal and infant characteristics	Values
Maternal prepregnancy BMI (kg/m^2), n; mean (SD)	5065; 27.7 (6.5)
Maternal smoking during pregnancy	
Smoker, n (%)	466 (10.9)
Nonsmoker, n (%)	3818 (89.1)
Gestational age (weeks), n; mean (SD)	4185; 39.3 (1.8)
Infant sex	
Males, n (%)	2649 (50.6)
Females, n (%)	2584 (49.4)
Infant birth weight (kg), n; mean (SD)	5233; 3.5 (0.5)
Parity, n (%)	
No sibling	3352 (64.1)
≥1 sibling	1877 (35.9)
Breastfeeding duration, n (%)	
<6 months	1975 (40.5)
≥6 months	2904 (59.5)
Timing of solids introduction, n (%)	
Before 6 months	695 (20.1)
At or after 6 months	2615 (79.9)
Rapid weight gain around the age of 1 year, n (%)	
Yes	1433 (27.4)
No	3800 (72.6)

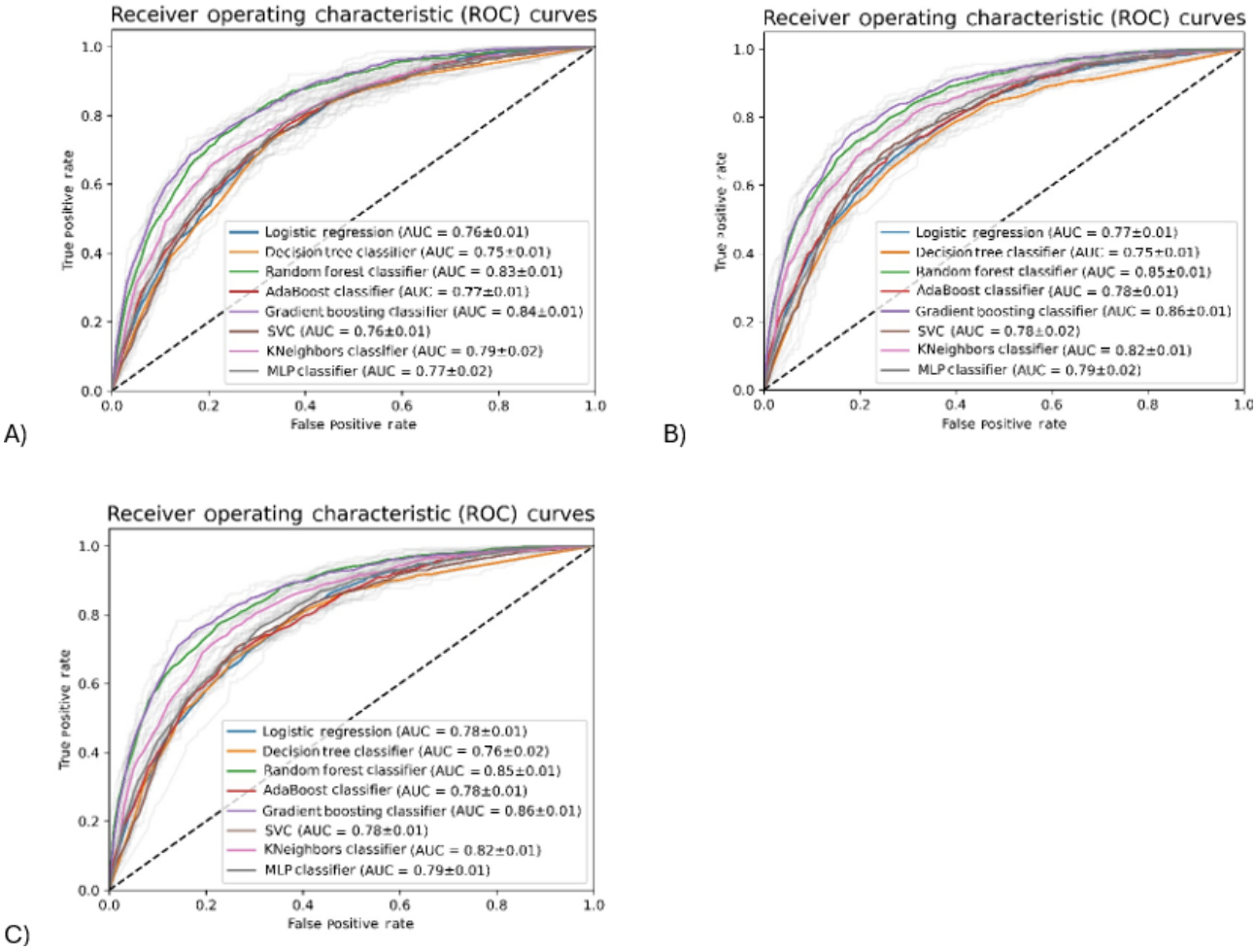
Internal Validation

Visual inspection showed that ROC curves for all models (Figure 1) deviated from the diagonal line and were arched toward the top-left corner of the plot, indicating good discriminative ability. All 3 models showed similar AUCs with values ranging from 0.76 - 0.85 (model 1), 0.75 - 0.86 (model 2), and 0.76 - 0.86 (model 3). Among the 8 ML algorithms, the Gradient Boosting classifier showed the best performance in terms of AUC (0.84 - 0.86) in all models, suggesting an "excellent" ability to distinguish between those with and without RWG. Other predictive performance metrics of the models are shown in Table 2. The accuracy of models 1, 2, and 3 ranged from 0.69 - 0.76, 0.70 - 0.77, and 0.70 - 0.78, respectively, which indicates that the models correctly predicted the outcomes for 69% - 78% of infants. Likewise, the 3 sets of models demonstrated similar precision of 0.68 - 0.75 (model 1), 0.68 - 0.76 (model 2), and 0.67 - 0.77 (model 3), indicating that the models correctly predicted 67% - 77% of positive cases while showing good ability to minimize false positives. In addition, models showed good ability to identify infants with RWG as indicated by sensitivity of 0.68 - 0.79 (model 1),

XSL•FO
RenderX

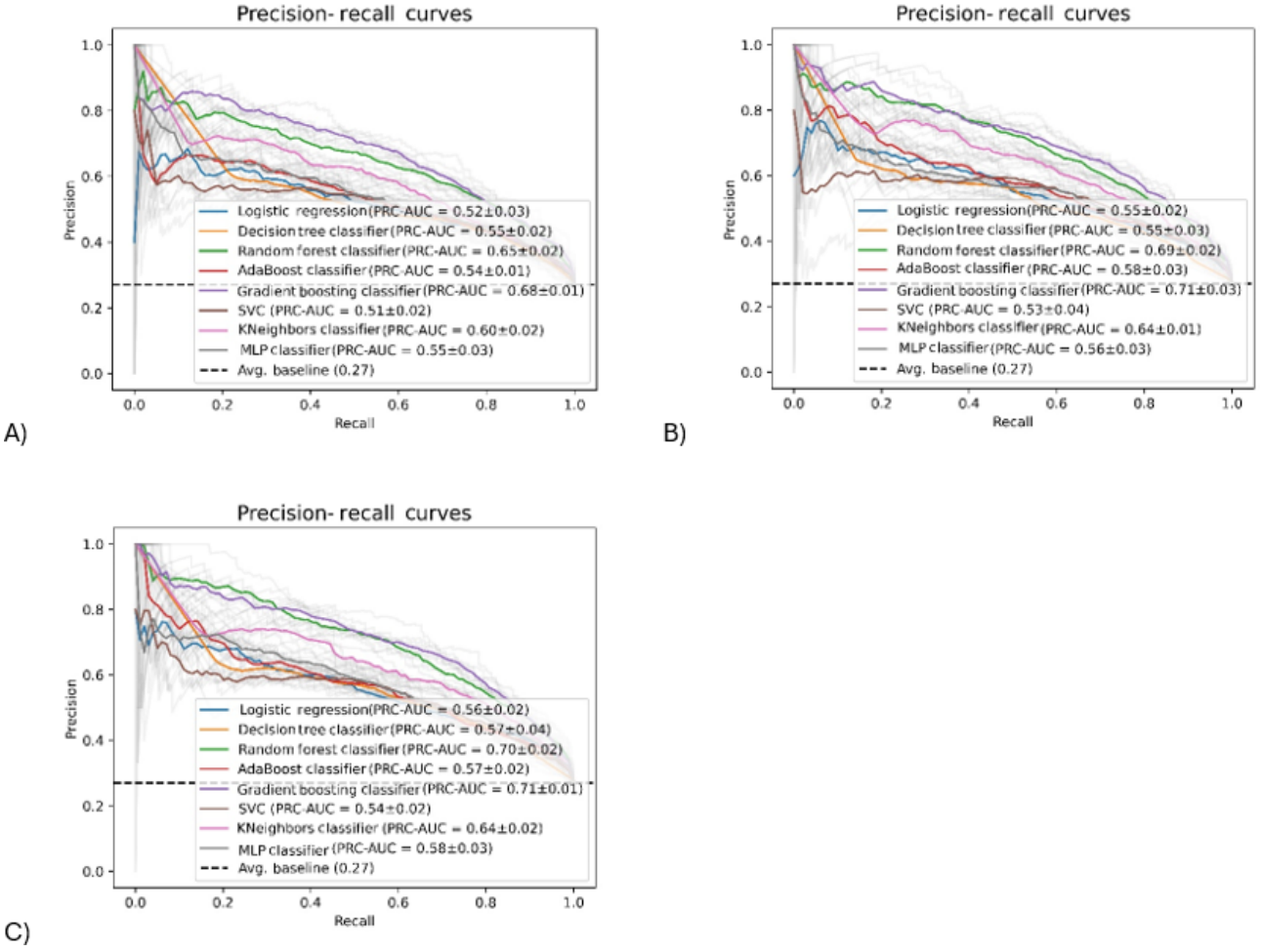
0.69 - 0.80 (model 2), 0.69 - 0.80 (model 3). The AUC range of PRC for models 1 to 3 was 0.51 - 0.68, 0.53 - 0.71, and 0.54 - 0.71, respectively (Figure 2). For F_1-score, the respective range for all models was 0.68 - 0.76 (model 1), 0.69 - 0.77 (model 2), and 0.69 - 0.78 (model 3), indicating that the models exhibited a good balance between precision and sensitivity and suggesting good overall model performance with low prediction errors. Models 1, 2, and 3 had a specificity of 0.65 - 0.75, 0.67 - 0.76, and 0.65 - 0.78, respectively, indicating a good ability to identify infants with no RWG. The range of Cohen κ for all models was 0.38 - 0.53 (model 1), 0.39 - 0.55 (model 2), and 0.40 - 0.56 (model 3). The addition of infant feeding variables at age 6 months in models 2 and 3 showed 0.01 - 0.03 higher increase across all predictive performance metrics when compared with model 1. Consistent with AUC results, the Gradient Boosting classifier showed the best performance in terms of accuracy, precision, sensitivity, AUPRC, F_1-score, specificity, and Cohen κ.

Figure 1. Receiver operating characteristics (ROC) curves of eight machine learning algorithms in predicting infants at risk of rapid weight gain by the age of 1 year in the training dataset (internal validation). (**A**) Model 1 included maternal prepregnancy BMI, smoking during pregnancy, gestational age, parity, infant sex, and birth weight as predictors. (**B**) Model 2 included model 1 predictors plus any breastfeeding (yes or no) at age 6 months. (**C**) Model 3 included model 2 predictors plus solids introduction (yes or no) at age 6 months. AUC: area under the receiver operating characteristic curve; MLP: multilayer perceptron; SVC: support vector classifier.



XSL•FO
RenderX

Figure 2. Precision recall curves (PRC) of 8 machine learning algorithms in predicting infants at risk of rapid weight gain by the age of 1 year in the training dataset (internal validation). **(A)** Model 1 included maternal prepregnancy BMI, smoking during pregnancy, gestational age, parity, infant sex, and birth weight as predictors. (**B**) Model 2 included model 1 predictors plus any breastfeeding (yes or no) at age 6 months. (**C**) Model 3 included model 2 predictors plus solids introduction (yes or no) at age 6 months. AUC: area under the receiver operating characteristic curve; MLP: multilayer perceptron; SVC: support vector classifier.



XSL•FO
RenderX

Table . Predictive performance metrics of 8 machine learning algorithms in predicting infants at risk of rapid weight gain by the age of 1 year in the training dataset (internal validation).

Machine learning algorithms	Accuracy	Precision	Sensitivity	F_1-score	Specificity	Cohen κ
Model 1[a], mean (SD)[b]						
Logistic regression	0.69 (0.01)	0.69 (0.03)	0.68 (0.04)	0.68 (0.01)	0.70 (0.03)	0.38 (0.02)
Decision tree classifier	0.70 (0.01)	0.68 (0.02)	0.75 (0.03)	0.71 (0.01)	0.65 (0.02)	0.40 (0.01)
Random forest classifier	0.76 (0.01)	0.74 (0.02)	0.77 (0.03)	0.76 (0.02)	0.74 (0.01)	0.51 (0.02)
AdaBoost classifier	0.69 (0.01)	0.67 (0.03)	0.73 (0.03)	0.70 (0.01)	0.65 (0.04)	0.38 (0.02)
Gradient boosting classifier	0.76 (0.01)	0.75 (0.02)	0.78 (0.03)	0.77 (0.02)	0.75 (0.02)	0.53 (0.03)
Support vector classifier	0.70 (0.01)	0.68 (0.02)	0.74 (0.02)	0.71 (0.01)	0.67 (0.02)	0.41 (0.02)
K-neighbors classifier	0.73 (0.01)	0.70 (0.02)	0.79 (0.03)	0.74 (0.02)	0.68 (0.02)	0.47 (0.02)
Multilayer perception classifier	0.71 (0.01)	0.69 (0.02)	0.74 (0.03)	0.71 (0.01)	0.68 (0.03)	0.42 (0.02)
Model 2[c], mean (SD)						
Logistic regression	0.70 (0.01)	0.70 (0.02)	0.69 (0.02)	0.69 (0.01)	0.71 (0.02)	0.40 (0.01)
Decision tree classifier	0.70 (0.01)	0.68 (0.02)	0.72 (0.02)	0.70 (0.01)	0.67 (0.02)	0.39 (0.02)
Random forest classifier	0.76 (0.01)	0.75 (0.02)	0.77 (0.03)	0.76 (0.01)	0.76 (0.02)	0.52 (0.02)
AdaBoost classifier	0.71 (0.01)	0.69 (0.03)	0.73 (0.03)	0.71 (0.01)	0.68 (0.04)	0.41 (0.02)
Gradient boosting classifier	0.77 (0.01)	0.77 (0.02)	0.78 (0.02)	0.77 (0.01)	0.77 (0.02)	0.55 (0.01)
Support vector classifier	0.72 (0.01)	0.70 (0.03)	0.75 (0.02)	0.72 (0.02)	0.69 (0.02)	0.44 (0.03)
K-neighbors classifier	0.75 (0.01)	0.72 (0.02)	0.80 (0.03)	0.76 (0.01)	0.70 (0.02)	0.50 (0.02)
Multilayer perception classifier	0.73 (0.01)	0.71 (0.02)	0.74 (0.02)	0.73 (0.02)	0.71 (0.01)	0.46 (0.02)
Model 3[d], mean (SD)						
Logistic regression	0.70 (0.01)	0.70 (0.02)	0.69 (0.02)	0.69 (0.01)	0.71 (0.02)	0.40 (0.01)
Decision tree classifier	0.70 (0.01)	0.67 (0.02)	0.74 (0.02)	0.71 (0.01)	0.65 (0.01)	0.40 (0.01)
Random forest classifier	0.77 (0.00)	0.76 (0.02)	0.77 (0.02)	0.76 (0.01)	0.76 (0.02)	0.53 (0.01)
AdaBoost classifier	0.71 (0.01)	0.69 (0.03)	0.74 (0.02)	0.71 (0.01)	0.68 (0.04)	0.41 (0.02)
Gradient boosting classifier	0.78 (0.00)	0.77 (0.02)	0.79 (0.01)	0.78 (0.01)	0.77 (0.01)	0.56 (0.01)

XSL•FO
RenderX

Machine learning algorithms	Accuracy	Precision	Sensitivity	F_1-score	Specificity	Cohen κ
Support vector classifier	0.71 (0.01)	0.69 (0.02)	0.75 (0.02)	0.72 (0.02)	0.68 (0.02)	0.43 (0.03)
K-neighbors classifier	0.74 (0.01)	0.71 (0.02)	0.80 (0.03)	0.75 (0.01)	0.69 (0.02)	0.49 (0.02)
Multilayer perception classifier	0.73 (0.01)	0.72 (0.02)	0.73 (0.02)	0.72 (0.01)	0.72 (0.02)	0.45 (0.02)

[a]Model 1 included maternal prepregnancy BMI, smoking during pregnancy, gestational age, parity, infant sex, and birth weight as predictors.

[b]Values are presented as mean (SD) and are calculated as the average performance metrics of five iterations from the 5-fold cross validation

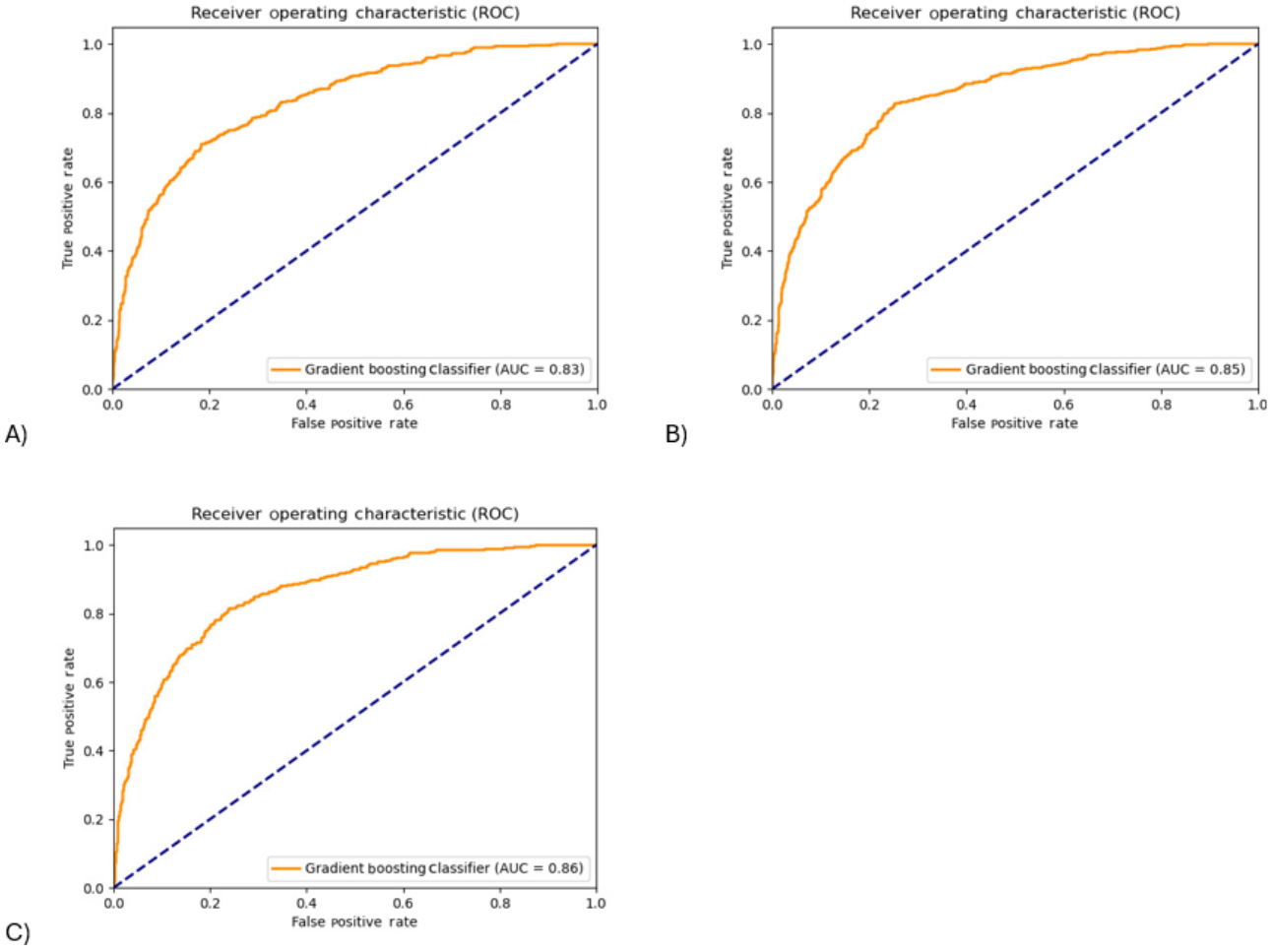
[c]Model 2 included model 1 predictors plus breastfeeding (yes or no) at age 6 months.

[d]Model 3 included model 2 predictors plus solids introduction (yes or no) at age 6 months.

External Validation

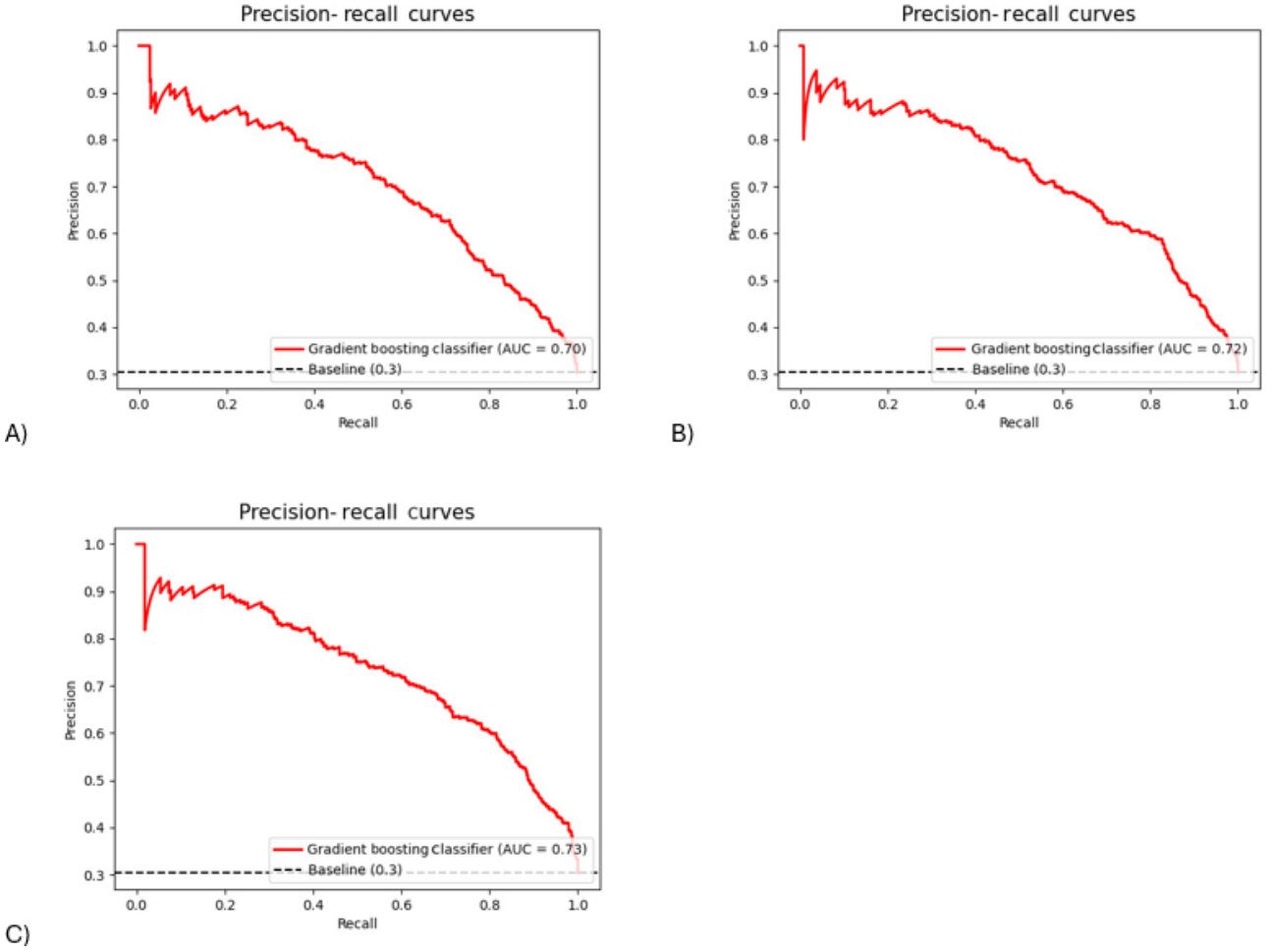
The Gradient Boosting classifier was used to predict infant RWG in the testing dataset. The ability of the Gradient Boosting classifier to distinguish between infants with or without RWG was excellent in the testing dataset with ROC AUC of 0.83 - 0.86 for 3 models, respectively (Figure 3). The AUPRC of 3 models spanned from 0.70 to 0.73, indicating a decent model performance (Figure 4). All 3 models (model 1, 2, 3) performed well in terms of accuracy (0.75, 0.78, and 0.78), sensitivity (0.75, 0.78, and 0.80) and specificity (0.76, 0.77, and 0.77), but performed modestly for precision (0.57, 0.60, and 0.60), F_1-score (0.65, 0.68, and 0.69), and Cohen κ (0.47, 0.51, and 0.52). Consistent with the results of the internal validation, model 2 with the addition of breastfeeding improved the predictive performance of model 1 by 0.01-0.04. However, the performance of models 2 and 3 was similar.

Figure 3. Receiver operating characteristics (ROC) curves of Gradient boosting classifier (identified in the internal validation) in predicting infants at risk of rapid weight gain by the age of 1 year in the testing dataset (external validation). (**A**) Model 1 included maternal prepregnancy BMI, smoking during pregnancy, gestational age, parity, infant sex, and birth weight as predictors. (**B**) Model 2 included model 1 predictors plus any breastfeeding (yes or no) at age 6 months. (**C**) Model 3 included model 2 predictors plus solids introduction (yes or no) at age 6 months. AUC: area under the curve.



XSL•FO
RenderX

Figure 4. Precision recall curves (PRC) of Gradient Boosting classifier (identified in the internal validation) in predicting infants at risk of rapid weight gain by the age of 1 year in the testing dataset (external validation). (**A**) Model 1 included maternal prepregnancy BMI, smoking during pregnancy, gestational age, parity, infant sex, and birth weight as predictors. (**B**) Model 2 included model 1 predictors plus any breastfeeding (yes or no) at age 6 months. (**C**) Model 3 included model 2 predictors plus solids introduction (yes or no) at age 6 months.



Additional Analyses

The SHAP analyses revealed that birth weight was the most influential factor underlying the prediction of infant RWG, followed by breastfeeding, gestational age, and maternal prepregnancy BMI. In contrast, parity, infant sex, timing of solids introduction, and maternal smoking had relatively less impact on the prediction (Supplementary Figure 3 in Multimedia Appendix 1).

Discussion

Principal Findings

Harnessing high-quality prenatal and early postnatal data from 7 Australian and New Zealand cohorts, we developed the first ML models to predict the risk of infant RWG by the age of 1 year (a proxy marker for later obesity risk) using only measures routinely or feasibly collected in primary health care. Our models enable risk prediction in infancy from birth to age 6 months, a critical window to intervene before the development of modifiable obesity-contributing behaviors such as unhealthy dietary intake and low physical activity. The resulting models demonstrated acceptable performance in predicting the risk of infant RWG.

Given the strong association between infant RWG and elevated obesity risk [12-15], early identification of infants at risk of RWG is a crucial step toward early obesity prevention. While research has assessed factors associated with infant RWG [30,41], here we present the first risk prediction models to predict RWG by the age of 1 year. The Gradient Boosting classifier exhibited excellent discrimination (>80%) to distinguish between infants with or without RWG, and performed well in making accurate predictions, particularly true positive cases (with accuracy and sensitivity all exceeding 75% confidence). However, the performance of the model in minimizing false positives was modest as indicated by accuracy, F_1-score, AUPRC, and Cohen κ. Notably, the addition of breastfeeding at age 6 months in risk prediction (model 2) improved model predictive accuracy by 3% - 4% across all performance metrics. However, the inclusion of solids introduction showed minimal impact on model performance. This suggests that acceptable risk prediction can be obtained at birth with very few routinely collected prenatal and birth variables, but the addition of postnatal breastfeeding would further enhance the risk prediction.

Comparison With Previous Work

Several statistically based risk prediction models have been developed to identify early obesity risk at ages 2 - 5 years using

XSL•FO
RenderX

predictors collected from birth to age 2 years among children from the United Kingdom [42-45], the United States [46], Israel [47], and New Zealand [48]. Furthermore, various ML-based early childhood obesity risk prediction models have also been developed [49-51]. However, the clinical utility of these models is questionable. For example, the statistical-based models were often derived from a small or a single dataset with external validation being conducted in only 1 or 2 datasets [42-44,48], both of which limit the model predictive accuracy and generalizability. For previous ML-based childhood obesity risk prediction models [49-51], although large eHealth data were used, these models are not clinically feasible as they require over a hundred predictors for risk prediction, undermining utility in primary health care settings.

Studies to evaluate the real-world impact of risk prediction models are the essential next step in the translation pathways [33]. There are, however, important economic, ethical, and clinical challenges for impact evaluation of existing obesity risk prediction models [7]. We proposed a novel and pragmatic solution by using infant RWG as a more acceptable and implementable alternative marker for obesity risk. Apart from systematic reviews [14,15], childhood obesity risk prediction models also consistently found that infant RWG is the most influential predictor of obesity outcomes [43,44,46], providing further evidence to support infant RWG as a valid infancy marker of future obesity risk. Our previous analysis in the current 7 cohorts showed that infant RWG was associated with 4.5 times increased risk of child overweight or obesity by the age of 5 years [31]. This translates to a population attributable fraction of 48.6%, suggesting that about half of child overweight or obesity were attributed to infant RWG with the prevalence of 27%.

Predicting the risk of RWG in infancy offers a pragmatic and low-cost solution for conducting ethically acceptable impact studies to assess clinical impact, as child growth indicators are routinely collected. Our models can be feasibly integrated into routine growth monitoring. Clinicians providing infant care are familiar with the routine use of risk prediction instruments. For example, the neonatal sepsis calculator has been successfully implemented around the world to improve antibiotic stewardship over the past decade [52]. Assessing the risk of infant RWG will raise the awareness of both primary health care professionals and parents about the vital role of early growth in shaping later child health. Research has shown that primary health professionals require support to monitor growth effectively despite growth monitoring being part of the routine infant care [53]. Moreover, both health professionals and parents have poor awareness on the concept and adverse health impact of infant RWG [10,54]. Embedding risk assessment of RWG in routine growth monitoring may motivate families with infants to make behavioral changes and establish positive health behaviors from the beginning of life.

Past studies primarily focused on exploring factors associated with childhood obesity but not infant RWG. Our study is novel in leveraging SHAP analyses to understand the relative contribution of prenatal and postnatal factors in infant RWG prediction. Our finding that birth weight was the most influential factor contributing to infant RWG prediction is supported by postnatal catch-up growth in body weight commonly observed in low-birth-weight infants [55]. It is well known that birth weight is influenced by a range of maternal factors from preconception to pregnancy such as maternal prepregnancy BMI, gestational weight gain, maternal health status, maternal nutrition, and other lifestyle behaviors during pregnancy [56]. Furthermore, our analyses showed that maternal prepregnancy BMI, gestational age, and breastfeeding also played a crucial role in infant RWG prediction, which further highlights the importance of targeting women during preconception and pregnancy to prevent infant RWG.

Leveraging statistical findings, ML, and high-quality cohort data for risk prediction model development is a major strength of the current study, ensuring predictive accuracy, generalisability, and clinical interpretability of the resulting risk prediction models. Statistics focuses on inference, whereas ML focuses on making accurate predictions, it has been increasingly advocated that statistics and ML could be combined to build reliable and clinically useful risk prediction models [57]. The use of statistical inferences to guide our predictor selection improves the clinical interpretability and utility of ML based risk prediction models. On the other hand, ML enables accurate predictions by discovering nuanced relationships between variables and data dimensions beyond “linear” relationship that is often assumed in statistical prediction. Another strength of our study was the use of diverse high-quality data from 7 cohorts, which enhances the model generalisability. We also considered clinical integration during model development by purposely selecting predictor variables readily collectable in primary health care, promoting future clinical integration. Moreover, ML-based risk prediction models can be easily transformed into digital tools that allow automatic risk prediction and convenient integration with electronic management systems in primary health care. Self-assessment by families with newborns or infants is also possible.

Limitations

Notwithstanding, the limitations of our study should be noted. Some predictors had high levels of missing data. Although missing data imputation was conducted, further model validation with high quality data will be desirable to further enhance the predictive accuracy, generalisability, and clinical utility of our models. Moreover, further research incorporating eHealth records or clinical data for model calibration is needed before clinical integration. However, notably, the use of high-quality cohort data also has its benefit in building accurate models. Our models will serve as a foundation for future validation and calibration with more robust cohort and clinical data. The prevalence of infant RWG across 7 cohorts is comparable to those reported in previous research [14,15]. However, the utility of the resulting model in countries other than Australia and New Zealand requires further investigation. Finally, our model did not capture other genetic or environmental predictors of infant RWG, but it is important to note that we aim to build simple and practical models for clinical integration instead of complex models with superior predictive accuracy but limited clinical utility.

XSL•FO
RenderX

Conclusions

We developed the first set of ML-based risk prediction models to identify infant RWG, a potent proxy marker denoting obesity risk later in life. The resulting models using routinely collected prenatal and postnatal factors showed good predictive accuracy as evidenced by various predictive performance metrics with gradient boosting classifiers exhibiting the best performance. Our next step is to further calibrate the resulting models using routine clinical data and convert our ML models into a digital tool for use in primary health care or enable self-assessment for families with newborn or infants. Future clinical trials will be conducted to evaluate the real-world utility of our models in identifying RWG and assess the impact of our models along with early interventions in prevention of infant RWG and, in turn, long-term overweight, and obesity.

Data Availability

The datasets analyzed during this study are not publicly available due to ethical restrictions but are available from the corresponding authors along with relevant study chief investigators on reasonable request.

Authors' Contributions

MZ conceived the study and wrote the original draft. YZ developed the study methodology and conducted the formal analysis. KDH, RL, PV, JD, LMW, LAB, RT, RB, and ALP collected and curated the data. All authors critically reviewed and approved the final version of the paper. MZ obtained financial support for the project and MZ is supported by Australian Research Council Discovery Early Career Researcher Award (DE240100635).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional tables and figures.
[DOCX File, 495 KB - publichealth_v11i1e69220_app1.docx]

References

1. Report of the commission on ending childhood obesity. : World Health Organisation; 2016 URL: https://www.who.int/publications/i/item/9789241510066 [accessed 2025-06-10]
2. Simmonds M, Llewellyn A, Owen CG, et al. Predicting adult obesity from childhood obesity: a systematic review and meta-analysis. Obes Rev 2016 Feb;17(2):95-107. [doi: 10.1111/obr.12334] [Medline: 26696565]
3. Obesity and overweight. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2025-05-19]
4. WHO acceleration plan to stop obesity. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240075634 [accessed 2025-05-31]
5. Gou H, Song H, Tian Z, et al. Prediction models for children/adolescents with obesity/overweight: a systematic review and meta-analysis. Prev Med 2024 Feb;179:107823. [doi: 10.1016/j.ypmed.2023.107823] [Medline: 38103795]
6. Seyednasrollah F, Mäkelä J, Pitkänen N, et al. Prediction of adulthood obesity using genetic and childhood clinical risk factors in the cardiovascular risk in Young Finns study. Circ Cardiovasc Genet 2017 Jun;10(3):e001554. [doi: 10.1161/CIRCGENETICS.116.001554] [Medline: 28620069]
7. Butler É, Derraik JGB, Taylor RW, et al. Prediction models for early childhood obesity: applicability and existing Issues. Horm Res Paediatr 2018;90(6):358-367. [doi: 10.1159/000496563] [Medline: 30739117]
8. Butler É, Derraik JGB, Taylor RW, et al. Childhood obesity: how long should we wait to predict weight? J Pediatr Endocrinol Metab 2018 Apr 25;31(5):497-501. [doi: 10.1515/jpem-2018-0110] [Medline: 29668465]
9. Redsell SA, Rose J, Weng S, et al. Digital technology to facilitate proactive assessment of obesity risk during Infancy (ProAsk): a feasibility study. BMJ Open 2017 Sep 6;7(9):e017694. [doi: 10.1136/bmjopen-2017-017694] [Medline: 28882926]
10. Redsell SA, Atkinson P, Nathan D, et al. Parents' beliefs about appropriate infant size, growth and feeding behaviour: implications for the prevention of childhood obesity. BMC Public Health 2010 Nov 18;10:711. [doi: 10.1186/1471-2458-10-711] [Medline: 21087482]
11. Redsell SA, Atkinson PJ, Nathan D, et al. Preventing childhood obesity during infancy in UK primary care: a mixed-methods study of HCPs' knowledge, beliefs and practice. BMC Fam Pract 2011 Jun 23;12:54. [doi: 10.1186/1471-2296-12-54] [Medline: 21699698]
12. Singhal A. Long-term adverse effects of early growth acceleration or catch-up growth. Ann Nutr Metab 2017;70(3):236-240. [doi: 10.1159/000464302] [Medline: 28301849]

XSL•FO
RenderX

13. Arisaka O, Ichikawa G, Koyama S, et al. Childhood obesity: rapid weight gain in early childhood and subsequent cardiometabolic risk. Clin Pediatr Endocrinol 2020;29(4):135-142. [doi: 10.1297/cpe.29.135] [Medline: 33088012]
14. Zheng M, Lamb KE, Grimes C, et al. Rapid weight gain during infancy and subsequent adiposity: a systematic review and meta-analysis of evidence. Obes Rev 2018 Mar;19(3):321-332. [doi: 10.1111/obr.12632] [Medline: 29052309]
15. Ong KK, Loos RJF. Rapid infancy weight gain and subsequent obesity: systematic reviews and hopeful suggestions. Acta Paediatr 2006 Aug;95(8):904-908. [doi: 10.1080/08035250600719754] [Medline: 16882560]
16. An R, Shen J, Xiao Y. Applications of artificial intelligence to obesity research: scoping review of methodologies. J Med Internet Res 2022 Dec 7;24(12):e40589. [doi: 10.2196/40589] [Medline: 36476515]
17. Colmenarejo G. Machine learning models to predict childhood and adolescent obesity: a review. Nutrients 2020 Aug 16;12(8):2466. [doi: 10.3390/nu12082466] [Medline: 32824342]
18. Vuillermin P, Saffery R, Allen KJ, et al. Cohort profile: the Barwon infant study. Int J Epidemiol 2015 Aug;44(4):1148-1160. [doi: 10.1093/ije/dyv026] [Medline: 25829362]
19. Wen LM, Baur LA, Simpson JM, et al. Effectiveness of an early intervention on infant feeding practices and “tummy time”: a randomized controlled trial. Arch Pediatr Adolesc Med 2011 Aug;165(8):701-707. [doi: 10.1001/archpediatrics.2011.115] [Medline: 21810633]
20. Wen LM, Baur LA, Simpson JM, et al. Effectiveness of home based early intervention on children’s BMI at age 2: randomised controlled trial. BMJ 2012 Jun 26;344:e3732. [doi: 10.1136/bmj.e3732] [Medline: 22735103]
21. Campbell K, Hesketh K, Crawford D, et al. The INFANT Feeding Activity and Nutrition Trial (INFANT) an early intervention to prevent childhood obesity: cluster-randomised controlled trial. BMC Public Health 2008 Mar 31;8:103. [doi: 10.1186/1471-2458-8-103] [Medline: 18373877]
22. Campbell KJ, Lioret S, McNaughton SA, et al. A parent-focused intervention to reduce infant obesity risk behaviors: a randomized trial. Pediatrics 2013 Apr;131(4):652-660. [doi: 10.1542/peds.2012-2576] [Medline: 23460688]
23. Campbell KJ, Hesketh KD, McNaughton SA, et al. The extended infant feeding, activity and nutrition trial (InFANT Extend) program: a cluster-randomized controlled trial of an early intervention to prevent childhood obesity. BMC Public Health 2016 Feb 18;16:166. [doi: 10.1186/s12889-016-2836-0] [Medline: 26888759]
24. Dodd JM. Dietary and lifestyle advice for pregnant women who are overweight or obese: the LIMIT randomized trial. Ann Nutr Metab 2014;64(3-4):197-202. [doi: 10.1159/000365018] [Medline: 25300260]
25. Dodd JM, Turnbull D, McPhee AJ, et al. Antenatal lifestyle advice for women who are overweight or obese: LIMIT randomised trial. BMJ 2014 Feb 10;348:g1285. [doi: 10.1136/bmj.g1285] [Medline: 24513442]
26. Daniels LA, Magarey A, Battistutta D, et al. The NOURISH randomised control trial: positive feeding practices and food preferences in early childhood - a primary prevention program for childhood obesity. BMC Public Health 2009 Oct 14;9:387. [doi: 10.1186/1471-2458-9-387] [Medline: 19825193]
27. Daniels LA, Mallan KM, Battistutta D, Nicholson JM, Perry R, Magarey A. Evaluation of an intervention to promote protective infant feeding practices to prevent childhood obesity: outcomes of the NOURISH RCT at 14 months of age and 6 months post the first of two intervention modules. Int J Obes (Lond) 2012 Oct;36(10):1292-1298. [doi: 10.1038/ijo.2012.96] [Medline: 22710926]
28. Taylor BJ, Heath ALM, Galland BC, et al. Prevention of Overweight in Infancy (POI.nz) study: a randomised controlled trial of sleep, food and activity interventions for preventing overweight from birth. BMC Public Health 2011 Dec 19;11:942. [doi: 10.1186/1471-2458-11-942] [Medline: 22182309]
29. Taylor RW, Heath ALM, Galland BC, et al. Three-year follow-up of a randomised controlled trial to reduce excessive weight gain in the first two years of life: protocol for the POI follow-up study. BMC Public Health 2016 Aug 11;16(1):771. [doi: 10.1186/s12889-016-3383-4] [Medline: 27514714]
30. Zheng M, Hesketh KD, Vuillermin P, et al. Determinants of rapid infant weight gain: a pooled analysis of seven cohorts. Pediatr Obes 2022 Oct;17(10):e12928. [doi: 10.1111/ijpo.12928] [Medline: 35510714]
31. Zheng M, Hesketh KD, Vuillermin P, et al. Understanding the pathways between prenatal and postnatal factors and overweight outcomes in early childhood: a pooled analysis of seven cohorts. Int J Obes (Lond) 2023 Jul;47(7):574-582. [doi: 10.1038/s41366-023-01301-9] [Medline: 37012427]
32. Askie LM, Espinoza D, Martin A, et al. Interventions commenced by early infancy to prevent childhood obesity-the EPOCH collaboration: an individual participant data prospective meta-analysis of four randomized controlled trials. Pediatr Obes 2020 Jun;15(6):e12618. [doi: 10.1111/ijpo.12618] [Medline: 32026653]
33. Moons KGM, Royston P, Vergouwe Y, et al. Prognosis and prognostic research: what, why, and how? BMJ 2009 Feb 23;338:b375. [doi: 10.1136/bmj.b375] [Medline: 19237405]
34. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Presented at: International Joint Conference on Artificial Intelligence; Jul 2009; São Paulo, Brazil p. 1137-1145.
35. Zhang Z. Missing data imputation: focusing on single imputation. Ann Transl Med 2016;4(1):9. [doi: 10.21037/atm.2016.07.17]
36. Belete DM, Huchaiah MD. Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. Int J Comput Appl 2022 Sep 2;44(9):875-886. [doi: 10.1080/1206212X.2021.1974663]

XSL•FO
RenderX

37. Sampath R, Indumathi J. Earlier detection of Alzheimer disease using N-fold cross validation approach. J Med Syst 2018 Oct 2;42(11):217. [doi: 10.1007/s10916-018-1068-5] [Medline: 30280260]
38. Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: Synthetic minority over-sampling technique. J Artif Intell 2002;16:321-357. [doi: 10.1613/jair.953]
39. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: Users' guides to the medical literature. JAMA 2017 Oct 10;318(14):1377-1384. [doi: 10.1001/jama.2017.12126] [Medline: 29049590]
40. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 2015;10(3):e0118432. [doi: 10.1371/journal.pone.0118432] [Medline: 25738806]
41. Mihrshahi S, Battistutta D, Magarey A, Daniels LA. Determinants of rapid weight gain during infancy: baseline results from the NOURISH randomised controlled trial. BMC Pediatr 2011 Nov 7;11:99. [doi: 10.1186/1471-2431-11-99] [Medline: 22054415]
42. Levine RS, Dahly DL, Rudolf MCJ. Identifying infants at risk of becoming obese: can we and should we? Public Health 2012 Feb;126(2):123-128. [doi: 10.1016/j.puhe.2011.10.008] [Medline: 22177581]
43. Santorelli G, Petherick ES, Wright J, et al. Developing prediction equations and a mobile phone application to identify infants at risk of obesity. PLoS One 2013;8(8):e71183. [doi: 10.1371/journal.pone.0071183] [Medline: 23940713]
44. Weng SF, Redsell SA, Nathan D, et al. Estimating overweight risk in childhood from predictors during infancy. Pediatrics 2013 Aug;132(2):e414-e421. [doi: 10.1542/peds.2012-3858] [Medline: 23858427]
45. Ziauddeen N, Wilding S, Roderick PJ, et al. Predicting the risk of childhood overweight and obesity at 4-5 years using population-level pregnancy and early-life healthcare data. BMC Med 2020 May 11;18(1):105. [doi: 10.1186/s12916-020-01568-z] [Medline: 32389121]
46. Robson JO, Verstraete SG, Shiboski S, et al. A risk score for childhood obesity in an urban Latino cohort. J Pediatr 2016 May;172:29-34. [doi: 10.1016/j.jpeds.2016.01.055] [Medline: 26952117]
47. Rossman H, Shilo S, Barbash-Hazan S, et al. Prediction of childhood obesity from nationwide health records. J Pediatr 2021 Jun;233:132-140. [doi: 10.1016/j.jpeds.2021.02.010] [Medline: 33581105]
48. Butler É, Pillai A, Morton SMB, et al. A prediction model for childhood obesity in New Zealand. Sci Rep 2021 Mar 18;11(1):6380. [doi: 10.1038/s41598-021-85557-z] [Medline: 33737627]
49. Dugan TM, Mukhopadhyay S, Carroll A, et al. Machine learning techniques for prediction of early childhood obesity. Appl Clin Inform 2015;6(3):506-520. [doi: 10.4338/ACI-2015-03-RA-0036] [Medline: 26448795]
50. Lingren T, Thaker V, Brady C, et al. Developing an algorithm to detect early childhood obesity in two tertiary Pediatric Medical Centers. Appl Clin Inform 2016 Jul 20;7(3):693-706. [doi: 10.4338/ACI-2016-01-RA-0015] [Medline: 27452794]
51. Hammond R, Athanasiadou R, Curado S, et al. Predicting childhood obesity using electronic health records and publicly available data. PLOS ONE 2019;14(4):e0215571. [doi: 10.1371/journal.pone.0215571] [Medline: 31009509]
52. Deshmukh M, Mehta S, Patole S. Sepsis calculator for neonatal early onset sepsis - a systematic review and meta-analysis. J Matern Fetal Neonatal Med 2021 Jun;34(11):1832-1840. [doi: 10.1080/14767058.2019.1649650] [Medline: 31352846]
53. Rossiter C, Cheng H, Denney-Wilson E. Primary healthcare professionals' role in monitoring infant growth: a scoping review. J Child Health Care 2024 Dec;28(4):880-897. [doi: 10.1177/13674935231165897] [Medline: 36963017]
54. Pesch MH, Levitt KJ, Danziger P, et al. Pediatrician's beliefs and practices around rapid infant weight gain: a qualitative study. Glob Pediatr Health 2021;8:2333794X21992164. [doi: 10.1177/2333794X21992164] [Medline: 33614855]
55. Martin A, Connelly A, Bland RM, et al. Health impact of catch-up growth in low-birth weight infants: systematic review, evidence appraisal, and meta-analysis. Matern Child Nutr 2017 Jan;13(1). [doi: 10.1111/mcn.12297] [Medline: 27002681]
56. Yu Z, Han S, Zhu J, et al. Pre-pregnancy body mass index in relation to infant birth weight and offspring overweight/obesity: a systematic review and meta-analysis. PLoS ONE 2013;8(4):e61627. [doi: 10.1371/journal.pone.0061627]
57. Finlayson SG, Beam AL, van Smeden M. Machine learning and statistics in clinical research articles-moving past the false dichotomy. JAMA Pediatr 2023 May 1;177(5):448-450. [doi: 10.1001/jamapediatrics.2023.0034] [Medline: 36939696]

Abbreviations

AUC: area under the receiver operating characteristic curve
AUPRC: area under the precision-recall curve
BIS: Barwon Infant Study
InFANT: Infant Feeding Activity and Nutrition Trial
ML: machine learning
POI: Prevention of Overweight in Infancy
PRC: precision-recall curve
ROC: receiver operating characteristic
RWG: rapid weight gain
SHAP: Shapley Additive Explanations
SMOTE: Synthetic Minority Over-sampling Technique
WHO: World Health Organization

XSL•FO
RenderX

Edited by OL Neto; submitted 25.11.24; peer-reviewed by C Mayadewi, T Hase; revised version received 18.03.25; accepted 25.03.25; published 18.06.25.

Please cite as:
Zheng M, Zhang Y, Laws RA, Vuillermin P, Dodd J, Wen LM, Baur LA, Taylor R, Byrne R, Ponsonby AL, Hesketh KD
Development of Machine Learning–Based Risk Prediction Models to Predict Rapid Weight Gain in Infants: Analysis of Seven Cohorts
JMIR Public Health Surveill 2025;11:e69220
URL: https://publichealth.jmir.org/2025/1/e69220
doi:10.2196/69220

© Miaobing Zheng, Yuxin Zhang, Rachel A Laws, Peter Vuillermin, Jodie Dodd, Li Ming Wen, Louise A Baur, Rachael Taylor, Rebecca Byrne, Anne-Louise Ponsonby, Kylie D Hesketh. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prevalence of Chronic Back Pain and Associated Factors in Children and Adolescents: Secondary Analysis of the 2001–2019 Health Behavior in School-Aged Children Study

Camila Miño[1], MPH; Héctor Gutiérrez-Espinoza[2], PhD; Jorge Olivares-Arancibia[3], PhD; Rodrigo Yáñez-Sepúlveda[4], PhD; Daniel Duclos-Bastías[5], PhD; Felipe Araya-Quintanilla[6], PhD; Lee Smith[7,8], PhD; José Francisco López-Gil[1,9], PhD

[1]School of Medicine, Universidad Espíritu Santo, Av Samborondón 320, Sector La Puntilla, Samborondón, Ecuador

[2]Faculty of Education, Universidad Autónoma de Chile, Santiago, Chile

[3]AFySE Group, Research in Physical Activity and School Health, School of Physical Education, Faculty of Education, Universidad de las Américas, Santiago, Chile

[4]Faculty Education and Social Sciences, Universidad Andres Bello, Viña del Mar, Chile

[5]Investigación en Gestión Deportiva y Estudios Olímpicos, School of Physical Education, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile

[6]Escuela de Kinesiología, Facultad de Ciencias de la Rehabilitación y Calidad de Vida, Universidad San Sebastián, Santiago, Chile

[7]Centre for Health Performance and Wellbeing, Anglia Ruskin University, Cambridge, United Kingdom

[8]Department of Public Health, Faculty of Medicine, Biruni University, Istanbul, Turkey

[9]Vicerrectoría de Investigación y Postgrado, Universidad de Los Lagos, Osorno, Chile

Corresponding Author:
José Francisco López-Gil, PhD
School of Medicine, Universidad Espíritu Santo, Av Samborondón 320, Sector La Puntilla, Samborondón, Ecuador

Abstract

Background: Backache, which is considered a leading cause of adult disability worldwide, is seldom studied in children and adolescents despite it being associated with multiple detrimental health outcomes in this age group.

Objective: The aim of this study is to explore the prevalence, time trends, and correlates of chronic backache among children and adolescents.

Methods: Data were extracted from the cross-sectional Health Behavior in School-Aged Children (HBSC) study, which collected data from 2001 to 2019. The final sample size for this study consisted of 1,011,368 children and adolescents (515,899/1,011,368 or 51.01% were girls). Prevalence estimates and time trends were calculated using weighted proportions. Generalized linear mixed models were conducted to determine whether certain correlates (ie, age group, sex, socioeconomic status, excess weight status, or year of data collection) were associated with higher odds of reporting chronic backache in the population of the HBSC. Additionally, we analyzed country-level changes in chronic backache trends over time.

Results: The weighted global prevalence of chronic backache among children and adolescents was 6.6% (n=12,641; wave 2017-2019). The time trend indicated a general increase in the prevalence of chronic backache in children and adolescents over time (from 2001 to 2019). Indeed, Northern and Eastern Europe, as well as parts of North America, presented relatively higher weighted prevalence rates of chronic backache, ranging from 7% to 9%. This corresponded to approximately 70,796 to 91,023 cases of 1,011,368 participants. Regarding correlates, older age (aged 12.5 - 14.5 y: odds ratio [OR] 1.10, 95% confidence interval [CI] 1.05 - 1.15; *P*<.001; aged 14.5 - 17 y: OR 1.27, 95% CI 1.22 - 1.32; *P*<.001), female sex (OR 1.22, 95% CI 1.17 - 1.28; *P*<.001), excess weight (OR 1.14, 95% CI 1.11 - 1.17; *P*<.001), and later years of data collection (OR 1.04, 95% CI 1.04 - 1.05; *P*<.001) were associated with higher odds of reporting chronic backache. In contrast, medium (OR 0.83, 95% CI 0.81 - 0.85; *P*<.001) and high (OR 0.90, 95% CI 0.86 - 0.93; *P*<.001) socioeconomic status showed lower odds of experiencing chronic backache than low socioeconomic status.

Conclusions: This study revealed a sustained upward trend in the prevalence of backache in children and adolescents over the years (2001 - 2019), especially in North America and Northern and Eastern Europe. In addition, older age, female sex, excess weight, and later years of data collection were identified as correlates of experiencing chronic backache. Moreover, middle and high socioeconomic status were associated with lower odds of reporting chronic backache, potentially indicating differences in lifestyle, access to health care, or other protective factors.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e67960) doi:10.2196/67960

KEYWORDS

backache; pediatric population; cross-national trends; socioeconomic status; linear mixed model; cross-sectional study; health outcomes; prevalence; time trends; risk factor; lifestyle; health care access; overweight; obesity

Introduction

Backache, generally characterized by pain and discomfort in the lumbar region, can be categorized based on duration [1]. Almost everyone experiences acute back pain at least once in their lifetime, making it the leading cause of disability worldwide and imposing a significant burden on both individuals and society [2,3]. According to a Global Burden of Disease study, backache affected almost 619 million individuals worldwide in 2020; however, 843 million back pain cases are expected to occur by 2050 [2]. Approximately 5% of children may experience backache at any given moment, with prevalence rates ranging from 1% to 6% in children and between 18% and 51% in adolescents [4,5]. Furthermore, back pain is more common in females, and its occurrence tends to increase with age [2,4]. Backache can diminish the overall perception of health-related quality of life, negatively impacting both physical and emotional well-being [6]. Indeed, it contributes to disability and disrupts daily activities for approximately 10% to 40% of adolescents, representing a critical public health challenge, as it has been associated with a greater probability of chronic backache later in life [6,7].

The development of back pain is influenced by a variety of multifactorial causes and risk factors [8]. Nevertheless, the biopsychosocial model posits that backache arises from interactions among biological elements, such as intervertebral disc degeneration, which are linked to aging, obesity, genetic factors, and physical strain, as well as psychosocial influences, which encompass conditions such as depression, anxiety, and social isolation [3,9]. Most importantly, engagement in competitive sports and older age groups are persistently associated with backache in children and adolescents [9]. Moreover, environmental factors, including socioeconomic status, significantly influence outcomes, as families with more resources may be more likely to offer their children superior nutrition and increased opportunities for engagement in extracurricular activities [10]. Likewise, sedentary behaviors, such as spending more than 15 hours per week on electronic devices, have been associated with backache [7]. Specifically, prolonged screen time often results in poor posture, leading to musculoskeletal strain, muscle spasms, fatigue, and potential intervertebral disc degeneration [3]. These lifestyle habits can weaken muscles, reduce flexibility, impair proprioception, and destabilize the spine, increasing the risk of injury and pain [3,7,11].

The Health Behavior in School-Aged Children (HBSC) study is a collaborative cross-sectional international study conducted by the World Health Organization that has been conducted every 4 years since 1983. This study aims to monitor and enhance the understanding of health, health behaviors, and the contextual factors influencing the lives of children and adolescents [12]. Previous research has reported cross-national trends in back pain in the adolescent population using results from the HBSC study. For example, Roy et al [13] reported a greater prevalence of back pain from 2001 - 2002 to 2013 - 2014, especially among older girls from the HBSC study. Additionally, chronic back pain has been associated with early menarche in girls, sleep difficulties and psychological symptoms, and increased screen time and obesity in adolescents according to the HBSC study [14-16]. In this context, by understanding how widespread backache is in younger populations, interventions can be implemented early to prevent the condition from worsening or becoming chronic in adulthood. Moreover, since backache is a leading cause of disability, identifying its prevalence in younger populations can help allocate resources, raise awareness, and prioritize public health initiatives aimed at reducing its impact not just in younger populations but across the lifespan. Nevertheless, updated analyses of the prevalence and temporal trends in this population group are scarce. This study aimed to explore the prevalence, time trends, and correlates (ie, age, sex, socioeconomic status, excess weight status, and year of data collection) of chronic backache among children and adolescents from 45 countries between 2001 and 2019.

Methods

Study Design and Population

This research extracted data from the 2001 - 2002, 2005 - 2006, 2009 - 2010, 2013 - 2014, and 2017 - 2019 waves of the HBSC study's open access repository. In each round of the survey, various nationally representative samples of students aged 10 - 17 years participated in an internationally standardized questionnaire administered in school settings. The children and adolescents were chosen randomly from their respective schools to participate in the study and anonymously completed a standardized test administered in their native language. The students had the option to leave questions unanswered. The study sample included children and adolescents aged between 10 and 17 years from different countries: Albania, Armenia, Austria, Azerbaijan, Belgium (both Flanders and Wallonia), Bulgaria, Canada, Croatia, the Czech Republic, Denmark, England, Estonia, Finland, France, Georgia, Germany, Greece, Greenland, Hungary, Iceland, Israel, Italy, Kazakhstan, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, North Macedonia, Norway, Poland, Portugal, the Republic of Moldova, Romania, the Russian Federation, Scotland, Serbia, Slovenia, Spain, Sweden, Ukraine, and Wales [12]. In the 2001 - 2002 wave, data were collected from 35 countries, whereas in 2005 - 2006, 41 countries participated. Similarly, in 2009 - 2010 and 2013 - 2014, data were gathered from 40 and 41 countries, respectively. In the most recent wave, 2017 - 2019, participation expanded to 45 countries.

After countries and participants with missing data were excluded, the total sample size for this study comprised 1,011,368 children and adolescents. During the 2001 - 2002

XSL•FO
RenderX

wave, 162,305 children and adolescents were analyzed, consisting of 78,936 boys (48.6%) and 83,369 girls (51.4%). In the 2005 - 2006 wave, the study included 205,938 adolescents, with 101,047 (49.1%) boys and 104,891 (50.9%) girls. For the wave of 2009 - 2010, a total of 213,595 children and adolescents were included, with 105,099 (49.2%) boys and 108,496 (50.8%) girls. Similarly, in the 2013 - 2014 wave, 214,080 children and adolescents participated, comprising 105,414 (49.2%) boys and 108,666 (50.8%) girls. Finally, in the 2017 - 2019 wave, the analysis included 240,951 children and adolescents, with 118,788 (49.3%) boys and 122,163 (50.7%) girls.

Ethical Considerations

The original data collection was approved by an institutional review board in each participating country, and written informed consent was obtained from schools and adolescents and their parents or legal guardians. As this study involved a secondary analysis of fully anonymized and publicly available data, no additional ethical approval or permission was required. The HBSC survey waves used in this analysis (2001 - 2002, 2005 - 2006, 2009 - 2010, 2013 - 2014, and 2017 - 2018) are freely accessible for research purposes. Further details on the HBSC study design and procedures are available elsewhere [17,18].

Procedures

Backache

In all waves, participants were asked the following question to report backache frequency: "How often have you had backache in the last six months?" (response options: 1=approximately every day, 2=more than once a week, 3=approximately every week, 4=approximately every month, 5=rarely or never). These scores were subsequently classified into two categories: chronic backache and no chronic backache. Chronic backache was defined as "approximately every day" backache. For further analysis, the data were categorized by wave and country. This involved recording the total number of children and adolescents with chronic backache identified in the waves from 2001 - 2002, 2005 - 2006, 2009 - 2010, 2013 - 2014, and 2017 - 2019 by country.

Sociodemographic and Socioeconomic Data

Sociodemographic data encompassed both sex (males and females) and age categories, specifically, 10 - 12.5, 12.5 - 14.5, and 14.5 - 17 years. The socioeconomic status of children and adolescents was assessed through 4 measures related to family and local socioeconomic position. For instance, in the study waves of 2001-2002, 2005-2006, and 2009-2010, a series of questions were used to assess socioeconomic status, such as "Does your family own a car, van, or truck?" "Do you have your own bedroom for yourself?" "During the past 12 months, how many times did you travel away on holiday with your family?" "How many computers do your family own?" "How well off do you think your family is?" and "Some young people go to school or to bed hungry because there is not enough food at home. How often does this happen to you?" The occupational socioeconomic status of parents was evaluated with various socioeconomic status levels, job-related statuses, and other categories. Additionally, in the 2013-2014 wave, two questions were added: "Does your father/mother have a job?" and "Father/mother occupation socioeconomic status" [17]. Finally, during the 2017-2019 study survey, the Family Affluence Scale-III was used to evaluate socioeconomic status [19]. The scale comprises 6 items and calculates an aggregate index ranging from 0 to 13 points, and each item is scored between 0 and 3 (low socioeconomic status=0 - 7 points, medium socioeconomic status=8 - 11 points, and high socioeconomic status=12 - 13 points). The individual sum scores were converted into proportional ranks, reflecting the relative family affluence of participants within their respective countries. These scores were subsequently classified into 3 categories: the lowest 20%, the middle 60%, and the highest 20% for each country.

Anthropometric Variables

Across all waves, children and adolescents provided self-reported data on their body weight and height, from which body mass index (BMI) was derived (kg/m^2). It should be mentioned, in the 2017 - 2019 wave, BMI classification adhered to the standards outlined by the International Obesity Task Force [20] (1=thinness, 2=normal weight, 3=overweight, 4=obesity). For further analysis, excess weight (ie, overweight or obesity) was determined as follows: 0=no excess weight (ie, thinness or normal weight); 1=excess weight (ie, overweight or obesity).

Statistical Analysis

All the statistical analyses were performed via R statistical software (version 4.4.0; R Core Team) along with RStudio (2024.04.1+748; Posit PBC). For the main analyses, missing data were handled using the listwise deletion method, in which only cases with complete data for all variables included in the model were retained. To adjust for the complex survey design of the HBSC study, we applied the provided survey weights as probability weights. These weights account for unequal probabilities of selection, nonresponse, and poststratification adjustments. For descriptive purposes, the data in this research were expressed in terms of weighted number and percentages. Prevalence estimates and time trends were calculated using weighted proportions. For analytical purposes, a generalized linear mixed model (GLMM) was fitted using the *glmer* function from the *lme4* package to examine the sociodemographic and anthropometric correlates associated with chronic backache. This model included country as a random intercept to account for between-country variability, while age group, sex, socioeconomic status, BMI status, and year of data collection were included as fixed effects. From this model, odds ratios (ORs) and their corresponding 95% confidence intervals (CIs) were computed to determine the strength and precision of these associations. Given the cross-sectional design, ORs from the GLMM are interpreted as measures of association rather than causal inference. The significance of the additional country variance in the multilevel model was determined via the likelihood ratio test. A *P* value below .05 was considered to indicate statistical significance, and 95% CIs are presented. Multicollinearity was assessed using variance inflation factors (VIF), which measure the ratio of the variance of a parameter estimate in a full model to its variance in a single-parameter model.

XSL•FO
RenderX

Interaction terms (age group × sex, socioeconomic status × age group, and socioeconomic status × sex) were tested to assess their contribution to the model. Model performance was evaluated using the Akaike information criterion (AIC), the Bayesian information criterion (BIC), log-likelihood, and deviance. Model selection was based on likelihood ratio tests and AIC/BIC criteria, retaining only interactions that significantly improved model fitness. Random effect variance was also examined to assess the contribution of country-level variability to the model. To mitigate the risk of bias arising from missing data, we performed supplementary analyses using multiple imputation techniques. We used multivariate imputation through chained equations to evaluate the presence of any missing values in the dataset, under the assumption that the missing data occurred at random [21]. The *mice* package was used to replace missing values using chained equations, resulting in the creation of 42 data sets with multiple imputations. This approach adhered to the guideline of setting the number of imputations (m) to exceed 100 times the maximum percentage of missing data [22].

Results

Table 1 describes the weighted characteristics of the study participants classified by chronic backache and no chronic backache. In general, participants aged 14.5 - 17 years had a greater prevalence of chronic backache (23,051/333,419, 6.9%) than did those between 10 - 12.5 years (14,865/325,849, 4.6%) and 12.5 - 14.5 years (19,415/344,552, 5.6%). Furthermore, more girls had chronic backache (34,348/516,442, 6.7%) than did boys (23,518/494,677, 4.8%). Similarly, 6.4% (17,153/266,985) of children and adolescents with low socioeconomic status had chronic backache, whereas 5.3% (30,286/566,998) and 5.7% (7282/127,488) of those with middle and high socioeconomic status, respectively, had chronic backache. Moreover, 6.2% (11,660/188,106) of participants with excess weight status had chronic backache, whereas 5.5% (35,710/644,951) of individuals with normal weight status reported chronic backache. The weighted prevalence of chronic backache among participants increased from 2001-2002 (7774/159,664, 4.9%) to 2017-2019 (12,641/192,396, 6.6%). The descriptive data of the study participants using a multiple imputation analysis can be found in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Weighted prevalence of chronic and nonchronic backache across demographic and anthropometric characteristics (Health Behavior in School-Aged Children study 2001 - 2019; N=1,011,368).[a]

Variable and category	No chronic backache, n (%)	Chronic backache, n (%)
Age group (total=1,004,010; missing=7358)		
10 - 12.5 y	310,984 (95.4)	14,865 (4.6)
12.5 - 14.5 y	325,137 (94.4)	19,415 (5.6)
14.5 - 17 y	310,368 (93.1)	23,051 (6.9)
Sex (total=1,011,368; missing=0)		
Male	471,159 (95.2)	23,518 (4.8)
Female	482,094 (93.3)	34,348 (6.7)
Socioeconomic status (total=959,784; missing=51,584)		
Low	249,832 (93.6)	17,153 (6.4)
Medium	536,712 (94.7)	30,286 (5.3)
High	120,206 (94.3)	7282 (5.7)
Excess weight status [b] (total=831,690; missing=179,678)		
No[c]	609,241 (94.5)	35,710 (5.5)
Yes[d]	176,446 (93.8)	11,660 (6.2)
Year of data collection (total=799,692; missing=211,676)		
2001-2002	151,890 (95.1)	7774 (4.9)
2005-2006	186,229 (94.5)	10,799 (5.5)
2009-2010	194,662 (94.3)	11,671 (5.7)
2013-2014	41,646 (94.8)	2275 (5.2)
2017-2019	179,755 (93.4)	12,641 (6.6)

[a]Values are weighted to account for the complex sampling design. Percentages represent the prevalence of chronic backache and no chronic backache within each subgroup.

[b]According to the cutoff points for body mass index by the International Obesity Task Force [20].

[c]No excess weight includes participants with thinness or normal weight.

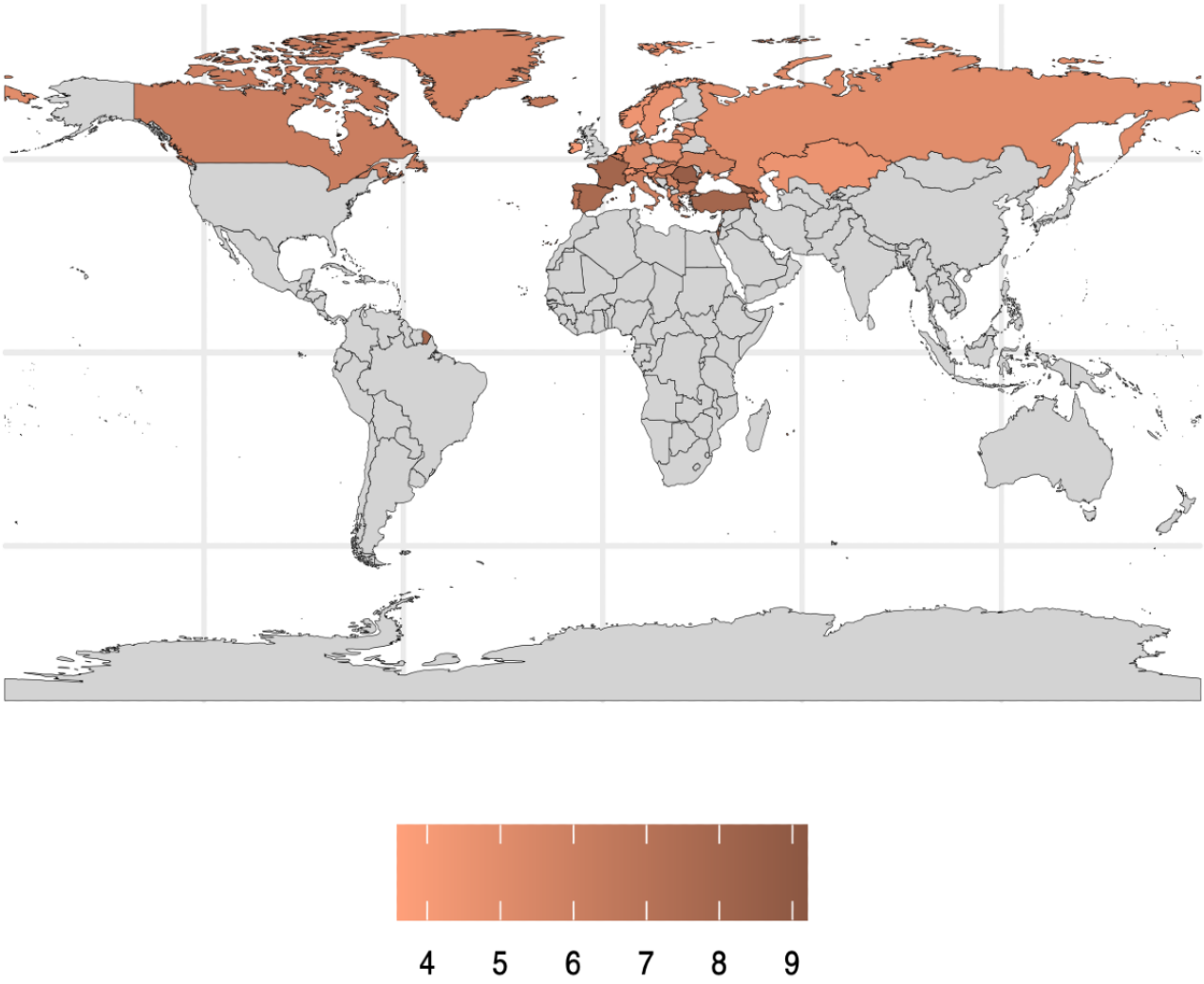
[d]Excess weight includes participants with overweight or obesity.

Figure 1 illustrates the prevalence of chronic backache among 1,011,368 participants from the HBSC study between 2001 and 2019, with data from multiple countries worldwide. Countries in Northern and Eastern Europe, as well as parts of North America, presented a relatively high prevalence of chronic backache, with values ranging from 7% to 9% (70,796/1,011,368 and 91,023/1,011,368, respectively). In comparison, some parts of Western Europe and a few countries in Asia presented moderate prevalence rates, ranging from approximately 5% to 7% (50,568/1,011,368 and 70,796/1,011,368, respectively). Multimedia Appendix 2 provides the prevalence rates of chronic and no chronic backache by country.

Figure 2 displays a detailed breakdown of the temporal trends in chronic backache among children and adolescents across various countries from 2001 to 2019. This trend indicates a general increase in the prevalence of chronic backache in children and adolescents over time, especially from 2001 - 2002 to 2017 - 2018. In 2017 - 2019, the highest prevalence was observed in Georgia (1176/12,641, 9.3%), France (1163/12,641, 9.2%), and Malta (1150/12,641, 9.1%). France and Israel presented significant increases in chronic backache prevalence across the years. Norway (354/12,641, 2.8%), Macedonia (366/12,641, 2.9%), and Lithuania (379/12,641, 3.0%) presented the lowest prevalence rates in 2017-2019. Countries such as Ireland and Switzerland remained consistently low throughout the observed period. Additionally, some countries showed a steady increase (eg, Israel, Romania, Portugal), whereas others had more fluctuations (eg, Belgium, Spain). In comparison, a few countries (eg, Croatia, Poland) remained relatively stable, with slight increases in later years. Multimedia Appendix 3 provides the chronic backache temporal trends stratified by country and wave.

XSL•FO
RenderX

Figure 1. Prevalence of chronic backache in the 10- to 17-year-old population of the Health Behavior in School-Aged Children cross-sectional study (N=1,011,368; 2001 - 2019).



XSL•FO
RenderX

Figure 2. Temporal trends of chronic backache in the 10- to 17-year-old population of the Health Behavior in School-Aged Children cross-sectional study (2001 - 2019) stratified by country and wave (N=1,011,368).

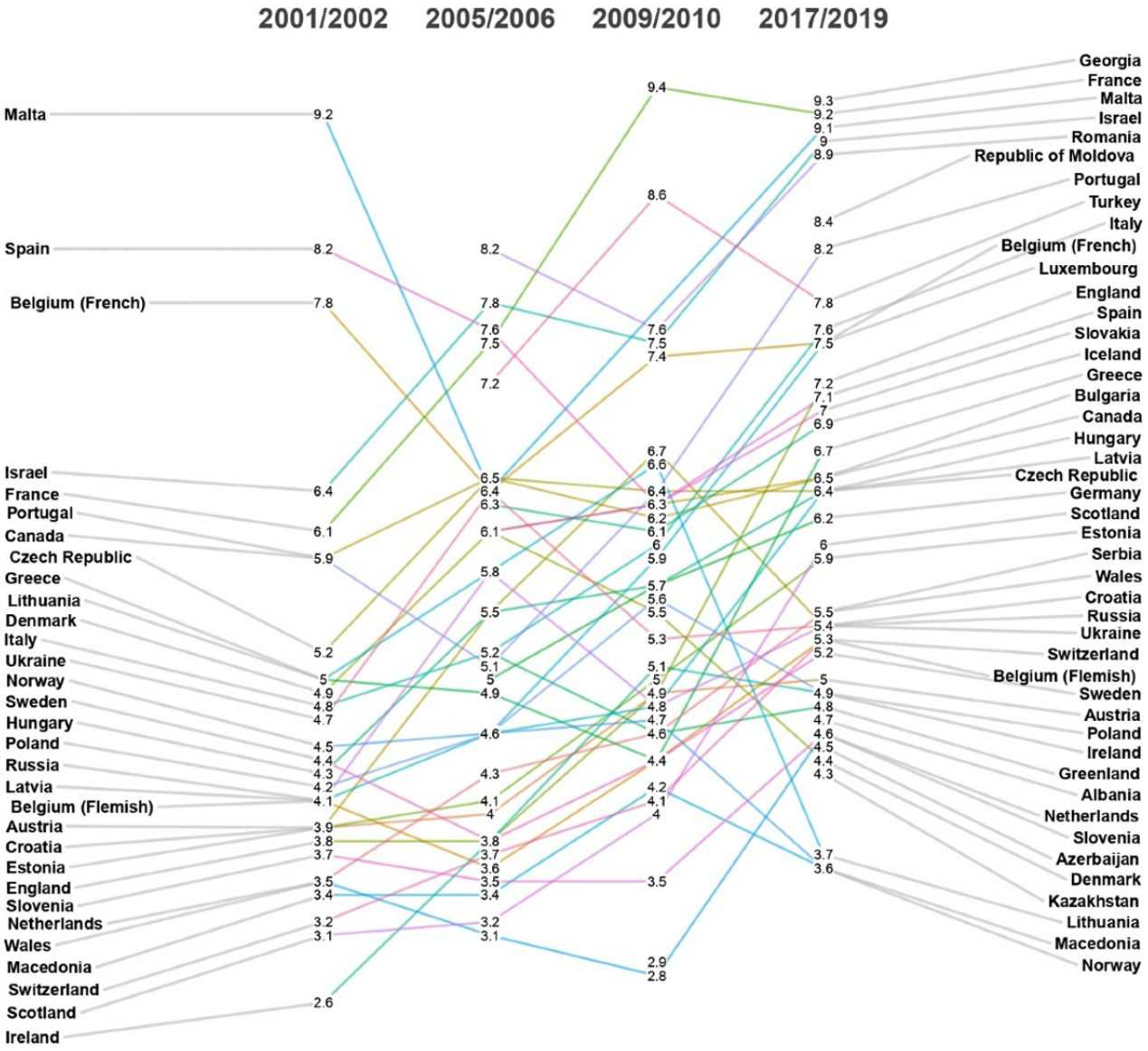


Table 2 presents the results of a GLMM evaluating factors associated with the probability of having chronic backache in the 10- to 17-year-old population of the HBSC study data from 2001 - 2019, including the interaction between age group and sex. We tested interaction terms to evaluate their contribution to the model fit. The inclusion of the age group × sex interaction significantly improved the model (χ^2_2=85.37; *P*<.001), reducing AIC/BIC values compared to the base model. In contrast, neither the model without any interaction (Multimedia Appendix 4) nor those including age group × socioeconomic status (Multimedia Appendix 5) or sex × socioeconomic status (Multimedia Appendix 6) resulted in a better fit. Therefore, only the age group × sex interaction was retained in the final model. Multicollinearity was checked using VIF, and all variables showed acceptable values (VIF <5), indicating no serious collinearity issues. Compared with participants aged 10.5 - 12.5 years, adolescents aged 14.5 - 17 years were 27% more likely to experience chronic backache (OR 1.27, 95% CI 1.22 - 1.32; *P*<.001), whereas participants aged 12.5 - 14.5 years were 10% more likely to experience chronic backache than individuals aged 10 - 12.5 years (OR 1.10, 95% CI 1.05 - 1.15; *P*<.001). In terms of sex, girls were 22% more likely to experience chronic backache than boys (OR 1.22; 95% CI 1.17 - 1.28; *P*<.001). Moreover, middle socioeconomic status was associated with lower odds of chronic backache than low socioeconomic status (OR 0.83, 95% CI 0.81 - 0.85; *P*<.001). Similarly, participants with high socioeconomic status had a 10% lower probability of chronic backache (OR 0.90, 95% CI 0.86 - 0.93; *P*<.001) than did participants with low socioeconomic status. In addition, children and adolescents with overweight were more likely to experience chronic backache than were those with normal weight (OR 1.14, 95% CI 1.11 - 1.17; *P*<.001). Additionally, it was identified that the later the year of data collection, the higher the odds of having chronic backache (OR 1.04, 95% CI 1.04 - 1.05; *P*<.001). Multimedia Appendix 7 presents the full GLMM estimating the odds of having chronic backache, including the interaction between age group and sex, based on a multiple imputation analysis. Notably, the results remained consistent regardless of the method used to handle missing data (listwise deletion method

XSL•FO
RenderX

versus multiple imputation method). Additionally, the complete results of the model without interaction, as well as models including the interactions between age group and socioeconomic status and between sex and socioeconomic status, based on multiple imputation analysis, are provided in the supplementary material (Multimedia Appendices 8-10, respectively).

Table . Generalized linear mixed model to evaluate the odds of having chronic backache in the 10- to 17-year-old population of the Health Behavior in School-Aged Children cross-sectional study (2001 - 2019), including the interaction between age group and sex.

Predictor	Odds ratio[a]	95% CI	*P* value
Age group			
10 - 12.5 y	Reference		
12.5 - 14.5 y	1.10	1.05-1.15	<.001
14.5 - 17 y	1.27	1.22-1.32	<.001
Sex			
Male	Reference		
Female	1.22	1.17-1.28	<.001
Socioeconomic status			
Low	Reference		
Medium	0.83	0.81-0.85	<.001
High	0.90	0.86-0.93	<.001
Excess weight status			
No excess weight	Reference		
Excess weight	1.14	1.11-1.17	<.001
Year of data collection (per one year)	1.04	1.04-1.05	<.001
Age group × sex			
12.5 - 14.5 y × female	1.19	1.12-1.26	<.001
14.5 - 17 y × female	1.31	1.24-1.38	<.001

[a]Model fit was assessed using the Akaike information criterion (452,439) and the Bayesian information criterion (452,569). The model's log-likelihood was –226,208, with a deviance of 452416.8. This analysis includes only cases with complete data for all variables in the analysis (n=599,553).

Discussion

This study presented the prevalence, time trends, and associated factors (age, sex, socioeconomic status, and excess weight) of chronic backache among a large sample of children and adolescents from the HBSC study in 45 countries. Currently, there is limited evidence reporting the epidemiology and associated factors of chronic backache in younger populations using data from the HBSC study from 2001 to 2019. Nevertheless, Roy et al [13] reported increased trends in chronic back pain among females and older adolescents using data from the HBSC study from 2001 to 2014; however, they did not include the related variables explored in our analysis. Other previous analyses used HBSC data to analyze back pain. For example, Roman-Juan et al [14] reported that screen time and obesity were associated with chronic back pain in adolescents in the HBSC study between 2002 and 2014. Similarly, findings from the HBSC study between 2002 and 2018 indicated a general increase of 0.5% in the prevalence of backache in adolescents, which was partially influenced by sleep difficulties and psychological symptoms [15]. Moreover, insights from the HBSC study (2002 - 2014) revealed an association between an increase in early menarche and the prevalence of chronic back pain among girls [16].

Overall, our findings indicated a greater prevalence of chronic backache in children and adolescents residing in Northern and Eastern Europe, which can be attributed to a combination of cultural, psychosocial, environmental, health care access, and socioeconomic factors. In fact, the literature highlights that socioeconomic differences between European countries may play a role in the prevalence of back pain among adolescents, with higher prevalence rates observed in countries with less risk of poverty or social exclusion [23,24]. Furthermore, cultural differences in the perception and reporting of pain may play a role in the variation in prevalence rates among European countries [25]. Besides, the colder climate in Northern and Eastern Europe can lead to muscle stiffness and reduced physical activity, which are risk factors for chronic musculoskeletal complaints such as back pain [26,27]. In addition, older age, female sex, excess weight, lower socioeconomic status, and later years of data collection were associated with a greater likelihood of chronic backache. These results are in line with those of previous studies [9,13,23,28-33]. Moreover, the etiology of backache in children and adolescent populations can be attributed to a wide range of medical conditions across various

XSL•FO
RenderX

disease categories, regardless of age [34]. This is particularly evident in cases involving infectious diseases, oncological disorders, and congenital anomalies affecting spinal structure [3,34]. However, the specific mechanisms by which nonspecific back pain may develop in children and adolescents are diverse and not fully understood (eg, genetics, hormonal and biological changes, diet, physical activity and sedentary behaviors, psychological factors, and socioenvironmental factors) [3,9,34]. Consequently, it is important to understand the mechanisms by which they may promote back pain in this population group because it may be regarded as a red flag for underlying medical conditions or chronic backache in the future. Despite this, several potential explanations could account for our results.

Adolescence is a period of rapid physical growth; therefore, sudden changes in height can stress the developing spine [3,35]. Research has indicated that the musculoskeletal systems of children insufficiently develop during peak height velocity in puberty to withstand abrupt mechanical load variations in the spine [34,36]. This was attributed to the disparity in growth rates between the legs and the trunk, where the long bones of the legs undergo a growth spurt prior to the shorter bones of the trunk [36]. Likewise, the higher prevalence of back pain in girls than in boys could be due to hormonal and biological changes during puberty [34,35]. The onset of puberty resulted in an increase in height and muscle mass among boys, whereas girls experienced an increase in body fat [35]. Additionally, estrogen and progesterone can influence muscle and joint health, leading to increased vulnerability to pain, as well as early menarche [16,37]. International cross-sectional studies have reported significant differences between boys and girls, with a higher incidence of backache among girls (P<.001) [28,30,33].

Obesity is increasingly recognized as a significant public health concern, impacting 1 in 3 individuals below the age of 18 years [38]. Excess weight places additional strain on the lumbar spine and musculoskeletal system, which may be linked to misalignment of vertebrae, discs, and surrounding muscles [37]. Additionally, obesity is associated with chronic inflammation and metabolic changes that can affect musculoskeletal health and are associated with pain [3,38]. Similarly, obesity is often linked to lower levels of physical activity, which could weaken the core muscles that support the spine [39,40]. Despite these possible explanations, the association of excess weight with back pain has been controversial in the literature. Some studies have not identified an association between pediatric backache and physical factors (such as height, weight, or BMI) [9,41]. However, the meta-analysis of García-Moreno et al [42] reported a significant association between overweight-obesity and back pain (OR 1.18, 95% CI 1.14 - 1.23) compared with normal weight in children and adolescents.

A growing body of evidence suggests that psychological and psychosocial factors significantly contribute to the etiology of backache [3,7]. Stress has been associated with muscle tension and reduced physical activity and affects pain perception, increasing the prevalence of back pain in this age group [3]. In particular, psychological distress in childhood has been related to a higher likelihood of developing spinal pain in adolescence by 33% (OR 1.33, 95% CI 1.01 - 1.76) [11]. Additionally, Batley et al [43] reported that low mood, bad mood, nervousness, difficulty sleeping, and loneliness were related to spinal pain among Danish children and adolescents aged 11 - 13 years. In addition, adverse psychological factors are associated with a reduced range of motion in the spine and increased trunk muscle activation in individuals experiencing back pain according to a systematic review and meta-analysis [44]. Adolescents from lower socioeconomic status backgrounds may have reduced access to physical activity opportunities, proper health care, or ergonomic school environments [45]. This lack of resources can result in weaker muscles, reduced flexibility, and poor posture, all of which are related to back pain [3,45]. Hestbaek et al [46] reported that a favorable social background during adolescence served as a protective factor against chronic backache in adulthood, whereas no correlation was found with any form of back pain. Furthermore, Huang et al [47] reported that although most backache cases are resolved acutely, approximately 30% progress to a chronic form. In this context, socioeconomic status is one of the most significant determining factors in the transition to chronic backache compared with traditional biomedical descriptors [47]. In addition, back pain represents an important burden to global public health in younger populations. According to the Global Burden of Disease study, in 2019, the disability-adjusted life year rate of back pain was 627.66 (95% CI 419.71 - 866.97) among individuals aged 15 - 39 years [48].

The findings in our study have important implications for clinical practice. These results suggest that chronic backache is a widespread complaint in children and adolescents, and its prevalence varies across populations and geographic regions. In addition, the prevalence rates have increased over time, in part due to heightened awareness and detection of chronic pain in children and recognition of its impact. Identification of the factors that contribute to the onset and persistence of back pain in children is important to inform preventive strategies and potential treatments. Although beyond the scope of this study, besides the educational and social aspects, backache in children and adolescents carries economic implications, as it has been associated with increased rates of medical visits.

The present analysis has several limitations that should be considered when the findings are interpreted. Importantly, anthropometric data (weight and height) and backache were self-reported through questionnaires rather than objective methods, which could introduce recall or social desirability bias into the findings. Besides, our study lacks the analysis of national-level covariates (such as the Gini coefficient or the proportion of health care expenditure) to explain regional differences. In addition, the cross-sectional design used in the study limited the capacity to determine a causal relationship between the observed outcomes. Consequently, future prospective cohort observational studies are needed to provide valuable insights into how different exposures, behaviors, and conditions affect backache in children and adolescent populations over time, helping to identify preventive measures, high-risk groups, and the effectiveness of existing interventions. A significant advantage of the study lies in its utilization of a substantial and representative sample of children and adolescents across 45 countries, which consequently strengthens the external validity of the results. Furthermore, we performed a multiple

XSL•FO
RenderX

imputation analysis to assess the robustness and enhance the generalizability of our results.

This study revealed a sustained upward trend in the prevalence of chronic backache in children and adolescents over the years (2001‐2019), especially in North America and Northern and Eastern Europe. In addition, older age, female sex, excess weight, and later years of data collection were identified as correlates of experiencing chronic backache. Additionally, a higher socioeconomic status was associated with a lower likelihood of chronic backache, potentially indicating differences in lifestyle, access to health care, or other protective factors. Given that backache is a primary contributor to disability worldwide, assessing its prevalence among younger demographics is important for identifying potential correlates and informing prevention or intervention strategies specifically designed for younger populations. Nevertheless, the causal nature of these associations remains unclear and warrants further prospective research to elucidate potential mechanisms. Moreover, unmeasured factors may also contribute to the prevalence of chronic backache among children and adolescents.

Acknowledgments

Open access was provided by Universidad Espíritu Santo (Samborondón, Ecuador).

Data Availability

The datasets used in this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: CM and JFL-G; validation: JFL-G; formal analysis: JFLG; writing – original draft preparation: CM; writing – review and editing: HG-E, JO-A, RY-S, DD-B, FA-Q, LS, and JFL-G; supervision: JFL-G. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table S1. Descriptive data of the study participants using a multiple imputation analysis (N=1,036,869).
[DOCX File, 17 KB - publichealth_v11i1e67960_app1.docx]

Multimedia Appendix 2
Table S2. Prevalence rates of chronic and no chronic backache by country in the 10- to 17-year-old population of the Health Behavior in School-Aged Children study (2001-2019).
[DOCX File, 18 KB - publichealth_v11i1e67960_app2.docx]

Multimedia Appendix 3
Table S3. Chronic backache temporal trend rates in the 10- to 17-year-old population of the Health Behavior in School-Aged Children study (2001-2019) stratified by country and wave.
[DOCX File, 33 KB - publichealth_v11i1e67960_app3.docx]

Multimedia Appendix 4
Table S4. Generalized linear mixed model to evaluate the probability of having chronic backache in the 10- to 17-year-old population of the Health Behavior in School-Aged Children study (2001-2019).
[DOCX File, 17 KB - publichealth_v11i1e67960_app4.docx]

Multimedia Appendix 5
Table S5. Generalized linear mixed model to evaluate the probability of having chronic backache in the 10- to 17-year-old population of the Health Behavior in School-Aged Children study (2001-2019), including the interaction between age group and socioeconomic status.
[DOCX File, 18 KB - publichealth_v11i1e67960_app5.docx]

Multimedia Appendix 6
Table S6. Generalized linear mixed model to evaluate the probability of having chronic backache in the 10- to 17-year-old population of the Health Behavior in School-Aged Children study (2001-2019), including the interaction between sex and socioeconomic status.

XSL•FO
RenderX

[DOCX File, 17 KB - publichealth_v11i1e67960_app6.docx]

Multimedia Appendix 7
Table S7. Generalized linear mixed model evaluating the probability of having chronic backache among 10- to 17-year-olds in the Health Behavior in School-Aged Children cross-sectional study (2001-2019), including an interaction between age group and sex. The model was estimated using multiple imputation.
[DOCX File, 19 KB - publichealth_v11i1e67960_app7.docx]

Multimedia Appendix 8
Table S8. Generalized linear mixed model to evaluate the probability of having chronic backache in the 10- to 17-year-old population of the Health Behavior in School-Aged Children study (2001-2019). The model was estimated using multiple imputation.
[DOCX File, 18 KB - publichealth_v11i1e67960_app8.docx]

Multimedia Appendix 9
Table S9. Generalized linear mixed model evaluating the probability of having chronic backache among 10- to 17-year-olds in the Health Behavior in School-Aged Children cross-sectional study (2001-2019), including an interaction between age group and socioeconomic status. The model was estimated using multiple imputation.
[DOCX File, 19 KB - publichealth_v11i1e67960_app9.docx]

Multimedia Appendix 10
Table S10. Generalized linear mixed model evaluating the probability of having chronic backache among 10- to 17-year-olds in the Health Behavior in School-Aged Children cross-sectional study (2001-2019), including an interaction between sex and socioeconomic status. The model was estimated using multiple imputation.
[DOCX File, 19 KB - publichealth_v11i1e67960_app10.docx]

References

1. van Tulder M, Becker A, Bekkering T, et al. Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. Eur Spine J 2006 Mar;15 Suppl 2(Suppl 2):S169-S191. [doi: 10.1007/s00586-006-1071-2] [Medline: 16550447]
2. GBD 2021 Low Back Pain Collaborators. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol 2023 May 22;5(6):e316-e329. [doi: 10.1016/S2665-9913(23)00098-X] [Medline: 37273833]
3. Knezevic NN, Candido KD, Vlaeyen JWS, Van Zundert J, Cohen SP. Low back pain. The Lancet 2021 Jul;398(10294):78-92. [doi: 10.1016/S0140-6736(21)00733-9]
4. Michaleff ZA, Kamper SJ, Maher CG, Evans R, Broderick C, Henschke N. Low back pain in children and adolescents: a systematic review and meta-analysis evaluating the effectiveness of conservative interventions. Eur Spine J 2014 Oct;23(10):2046-2058. [doi: 10.1007/s00586-014-3461-1] [Medline: 25070788]
5. Jones GT, Macfarlane GJ. Epidemiology of low back pain in children and adolescents. Arch Dis Child 2005 Mar;90(3):312-316. [doi: 10.1136/adc.2004.056812] [Medline: 15723927]
6. Potaczek T, Jasiewicz B. How back pain influences daily activities and quality of life: incidence of back pain related to age. J Child Orthop 2023 Dec;17(6):505-511. [doi: 10.1177/18632521231215855] [Medline: 38050592]
7. Minghelli B. Low back pain in childhood and adolescence phase: consequences, prevalence and risk factors - a revision. J Spine 2017;06(1). [doi: 10.4172/2165-7939.1000351]
8. Wall J, Meehan WP 3rd, Trompeter K, et al. Incidence, prevalence and risk factors for low back pain in adolescent athletes: a systematic review and meta-analysis. Br J Sports Med 2022 Nov;56(22):1299-1306. [doi: 10.1136/bjsports-2021-104749] [Medline: 36150752]
9. Calvo-Muñoz I, Kovacs FM, Roqué M, Gago Fernández I, Seco Calvo J. Risk factors for low back pain in childhood and adolescence: a systematic review. Clin J Pain 2018 May;34(5):468-484. [doi: 10.1097/AJP.0000000000000558] [Medline: 28915154]
10. Musić Milanović S, Buoncristiano M, Križan H, et al. Socioeconomic disparities in physical activity, sedentary behavior and sleep patterns among 6- to 9-year-old children from 24 countries in the WHO European region. Obes Rev 2021 Nov;22 Suppl 6:e13209. [doi: 10.1111/obr.13209] [Medline: 34235843]
11. Dario AB, Kamper SJ, Williams C, et al. Psychological distress in early childhood and the risk of adolescent spinal pain with impact. Eur J Pain 2022 Feb;26(2):522-530. [doi: 10.1002/ejp.1878] [Medline: 34695280]
12. Inchley J, Currie D. Health Behaviour in School-Aged Children (HBSC) Study Protocol: Background, Methodology and Mandatory Items for the 2017/18 Survey: CAHRU; 2018. URL: https://drive.google.com/file/d/1M3X_5j66-TfuzfCq0-SLle2XW5X0K3FO/view [accessed 2025-07-30]

XSL•FO
RenderX

13. Roy R, Galán S, Sánchez-Rodríguez E, et al. Cross-national trends of chronic back pain in adolescents: results from the HBSC study, 2001-2014. J Pain 2022 Jan;23(1):123-130. [doi: 10.1016/j.jpain.2021.07.002] [Medline: 34339858]
14. Roman-Juan J, Roy R, Jensen MP, Miró J. The explanatory role of sedentary screen time and obesity in the increase of chronic back pain amongst European adolescents: the HBSC study 2002-2014. Eur J Pain 2022 Sep;26(8):1781-1789. [doi: 10.1002/ejp.2003] [Medline: 35761772]
15. Roman-Juan J, Jensen MP, Miró J. Increases in sleep difficulties and psychological symptoms are associated with the increase of chronic back pain in adolescents: the HBSC study 2002 to 2018. J Pain 2024 Feb;25(2):407-417. [doi: 10.1016/j.jpain.2023.09.004] [Medline: 37690475]
16. Roman-Juan J, Jensen MP, Miró J. The increase in early menarche is associated with the increase in chronic back pain in female adolescents: the Health Behavior in School-aged Children study 2002-2014. Clin J Pain 2024 Nov 1;40(11):684-689. [doi: 10.1097/AJP.0000000000001247] [Medline: 39285799]
17. Currie C, Inchley J, Molcho M, Lenzi M, Veselska Z, Wild F. Health behaviour in school-aged children (HBSC) study protocol: background, methodology and mandatory items for the 2013/14 survey. : Child and Adolescent Health Research Unit; 2014 URL: https://drive.google.com/file/d/1FZ8c2Xa_FcZ5Yx5gPXkPtlbV545NKhnx/view [accessed 2025-07-30]
18. Roberts C, Freeman J, Samdal O, et al. The Health Behaviour in School-aged Children (HBSC) study: methodological developments and current tensions. Int J Public Health 2009 Sep;54 Suppl 2(Suppl 2):140-150. [doi: 10.1007/s00038-009-5405-9] [Medline: 19639259]
19. Torsheim T. HBSC Family Affluence Scale Coding Guidance (V1): HBSC Methods Note 1, HBSC Data Management Centre, University of Bergen, Bergen. 2018. URL: https://drive.google.com/file/d/1oN8PHrBGhG-DEG316-GBMbFveQShG4vA/view [accessed 2025-07-30]
20. Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatr Obes 2012:284-294. [doi: 10.1111/j.2047-6310.2012.00064.x] [Medline: 22715120]
21. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011 Dec 12;45(3):1-67. [doi: 10.18637/jss.v045.i03]
22. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-399. [doi: 10.1002/sim.4067] [Medline: 21225900]
23. Holstein BE, Damsgaard MT, Madsen KR, Pedersen TP, Toftager M. Chronic backpain among adolescents in Denmark: trends 1991-2018 and association with socioeconomic status. Eur J Pediatr 2022 Feb;181(2):691-699. [doi: 10.1007/s00431-021-04255-0] [Medline: 34529135]
24. Joergensen AC, Hestbaek L, Andersen PK, Nybo Andersen AM. Epidemiology of spinal pain in children: a study within the Danish National Birth Cohort. Eur J Pediatr 2019 May;178(5):695-706. [doi: 10.1007/s00431-019-03326-7] [Medline: 30788593]
25. Kędra A, Kolwicz-Gańko A, Sitarski D, Kędra P, Czaprowski D. Prevalence of back pain and the knowledge of preventive measures in a cohort of 11619 Polish school-age children and youth—an epidemiological study. Medicine (Abingdon) 2019;98(22):e15729. [doi: 10.1097/MD.0000000000015729]
26. Lewis C, Stjernbrandt A, Wahlström J. The association between cold exposure and musculoskeletal disorders: a prospective population-based study. Int Arch Occup Environ Health 2023 May;96(4):565-575. [doi: 10.1007/s00420-022-01949-2] [Medline: 36592178]
27. Stjernbrandt A, Hoftun Farbu E. Occupational cold exposure is associated with neck pain, low back pain, and lumbar radiculopathy. Ergonomics 2022 Sep;65(9):1276-1285. [doi: 10.1080/00140139.2022.2027030] [Medline: 35023451]
28. Fabricant PD, Heath MR, Schachne JM, Doyle SM, Green DW, Widmann RF. The epidemiology of back pain in American children and adolescents. Spine 2020 Aug 15;45(16):1135-1142. [doi: 10.1097/BRS.0000000000003461] [Medline: 32097269]
29. Bento TPF, Cornelio GP, Perrucini PDO, Simeão S, de Conti MHS, de Vitta A. Low back pain in adolescents and association with sociodemographic factors, electronic devices, physical activity and mental health. J Pediatr (Rio J) 2020;96(6):717-724. [doi: 10.1016/j.jped.2019.07.008] [Medline: 31580844]
30. Martínez-Romero MT, Cejudo A, Sainz de Baranda P. Prevalence and characteristics of back pain in children and adolescents from the region of Murcia (Spain): ISQUIOS programme. Int J Environ Res Public Health 2022 Jan 15;19(2):946. [doi: 10.3390/ijerph19020946] [Medline: 35055768]
31. Palmer K, Robbins LB, Ling J, Kao TSA, Voskuil VR, Smith AL. Adolescent autonomous motivation for physical activity: a concept analysis. J Pediatr Nurs 2020;54:e36-e46. [doi: 10.1016/j.pedn.2020.04.020] [Medline: 32451168]
32. Silva M, Badaró AFV, Dall'Agnol MM. Low back pain in adolescent and associated factors: a cross sectional study with schoolchildren. Braz J Phys Ther 2014;18(5):402-409. [doi: 10.1590/bjpt-rbf.2014.0051] [Medline: 25372002]
33. Vidal-Conti J, Borràs PA, Palou P, Muntaner-Mas A. Prevalence of low back pain among school-aged children between 10 and 12 years. Sustainability 2021 Nov 6;13(21):12254. [doi: 10.3390/su132112254]
34. Frosch M, Mauritz MD, Bielack S, et al. Etiology, risk factors, and diagnosis of back pain in children and adolescents: evidence- and consensus-based interdisciplinary recommendations. Children 2022 Feb 2;9(2):192. [doi: 10.3390/children9020192] [Medline: 35204913]

XSL•FO
RenderX

35. Best O, Ban S. Adolescence: physical changes and neurological development. Br J Nurs 2021 Mar 11;30(5):272-275. [doi: 10.12968/bjon.2021.30.5.272] [Medline: 33733842]
36. Lloyd RS, Oliver JL, Faigenbaum AD, Myer GD, De Ste Croix MBA. Chronological age vs. biological maturation: implications for exercise programming in youth. J Strength Cond Res 2014 May;28(5):1454-1464. [doi: 10.1519/JSC.0000000000000391] [Medline: 24476778]
37. Myszko O, Al-Husayni N, Talib HJ. Painful periods in the adolescent girl. Pediatr Ann 2020 Apr 1;49(4):e176-e182. [doi: 10.3928/19382359-20200318-01] [Medline: 32275762]
38. Nittari G, Scuri S, Petrelli F, Pirillo I, di Luca NM, Grappasonni I. Fighting obesity in children from European World Health Organization member states. Epidemiological data, medical-social aspects, and prevention programs. Clin Ter 2019;170(3):e223-e230. [doi: 10.7417/CT.2019.2137] [Medline: 31173054]
39. Ambrosio L, Mazzuca G, Maguolo A, et al. The burden of low back pain in children and adolescents with overweight and obesity: from pathophysiology to prevention and treatment strategies. Therapeutic Advances in Musculoskeletal 2023 Jan;15:1759720X231188831. [doi: 10.1177/1759720X231188831]
40. Kędra A, Plandowska M, Kędra P, Czaprowski D. Physical activity and low back pain in children and adolescents: a systematic review. Eur Spine J 2021 Apr;30(4):946-956. [doi: 10.1007/s00586-020-06575-5] [Medline: 32845380]
41. Kamper SJ, Yamato TP, Williams CM. The prevalence, risk factors, prognosis and treatment for back pain in children and adolescents: an overview of systematic reviews. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1021-1036. [doi: 10.1016/j.berh.2017.04.003] [Medline: 29103547]
42. García-Moreno JM, Calvo-Muñoz I, Gómez-Conesa A, López-López JA. Obesity and overweight as risk factors for low back pain in children and adolescents: a meta-analysis. Int J Obes 2024 May;48(5):612-625. [doi: 10.1038/s41366-024-01475-w]
43. Batley S, Aartun E, Boyle E, Hartvigsen J, Stern PJ, Hestbæk L. The association between psychological and social factors and spinal pain in adolescents. Eur J Pediatr 2019 Mar;178(3):275-286. [doi: 10.1007/s00431-018-3291-y] [Medline: 30465273]
44. Christe G, Crombez G, Edd S, Opsommer E, Jolles BM, Favre J. Relationship between psychological factors and spinal motor behaviour in low back pain: a systematic review and meta-analysis. Pain 2021 Mar 1;162(3):672-686. [doi: 10.1097/j.pain.0000000000002065] [Medline: 33591109]
45. Karran EL, Grant AR, Moseley GL. Low back pain and the social determinants of health: a systematic review and narrative synthesis. Pain 2020 Nov;161(11):2476-2493. [doi: 10.1097/j.pain.0000000000001944] [Medline: 32910100]
46. Hestbaek L, Korsholm L, Leboeuf-Yde C, Kyvik KO. Does socioeconomic status in adolescence predict low back pain in adulthood? A repeated cross-sectional study of 4,771 Danish adolescents. Eur Spine J 2008 Dec;17(12):1727-1734. [doi: 10.1007/s00586-008-0796-5] [Medline: 18830719]
47. Huang Z, Guo W, Martin JT. Socioeconomic status, mental health, and nutrition are the principal traits for low back pain phenotyping: data from the osteoarthritis initiative. JOR SPINE 2023 Jun;6(2):e1248. [doi: 10.1002/jsp2.1248] [Medline: 37361325]
48. Wu Z, Huang G, Ai J, Liu Y, Pei B. The burden of low back pain in adolescents and young adults. J Back Musculoskelet Rehabil 2024;37(4):955-966. [doi: 10.3233/BMR-230215] [Medline: 38517768]

Abbreviations

AIC: Akaike information criterion
BIC: Bayesian information criterion
BMI: body mass index
GLMM: generalized linear mixed model
HBSC: Health Behavior in School-Aged Children
OR: odds ratio
VIF: variance inflation factors

Edited by A Mavragani, T Sanchez; submitted 24.10.24; peer-reviewed by A Zaborskis, H Li, M Aryafar; revised version received 15.04.25; accepted 19.05.25; published 06.08.25.

Please cite as:
Miño C, Gutiérrez-Espinoza H, Olivares-Arancibia J, Yáñez-Sepúlveda R, Duclos-Bastías D, Araya-Quintanilla F, Smith L, López-Gil JF
Prevalence of Chronic Back Pain and Associated Factors in Children and Adolescents: Secondary Analysis of the 2001–2019 Health Behavior in School-Aged Children Study
JMIR Public Health Surveill 2025;11:e67960
URL: https://publichealth.jmir.org/2025/1/e67960
doi:10.2196/67960

XSL•FO
RenderX

© Camila Miño, Héctor Gutiérrez-Espinoza, Jorge Olivares-Arancibia, Rodrigo Yáñez-Sepúlveda, Daniel Duclos-Bastías, Felipe Araya-Quintanilla, Lee Smith, José Francisco López-Gil. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 6.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

How Healthy and Unhealthy Lifestyle Behaviors Affect Cognitive Function—Evidence From Older Adults in Chinese Communities: Cross-Sectional Study

Huixiu Hu[1,2*], PhD; Yajie Zhao[3*], MM; Yuqing Hao[1,4*], MM; Huanhuan Luo[1,4], MM; Lanying Xie[5], MM; Chao Sun[1], PhD

[1]Department of Nursing, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, No.1 Dahua, Dongcheng District, Beijing, China

[2]Department of Clinical Nutrition, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China

[3]Department of Cardiology, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, Beijing, China

[4]Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

[5]School of Nursing, Beijing University of Chinese Medicine, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Chao Sun, PhD
Department of Nursing, Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, No.1 Dahua, Dongcheng District, Beijing, China

Abstract

Background: Many lifestyle behaviors—including smoking, alcohol consumption, and engagement in physical activity and social activity—have been identified as potential determinants of the risk of cognitive impairment. Understanding how those lifestyle behavior patterns in older adults affect cognitive function is crucial for developing targeted interventions.

Objective: This study examined the lifestyle behavior patterns of Chinese community-dwelling older adults and their associations with cognitive impairment.

Methods: A cross-sectional study was conducted with 2060 community-dwelling older adults in Beijing, China. Latent class analysis identified distinct lifestyle behavior patterns based on unhealthy lifestyle behaviors (smoking and alcohol consumption) and healthy behaviors (physical activity and social activity). Cognitive function was evaluated using the Mini-Mental State Examination. Multiple logistic regression was conducted to examine the associations between lifestyle behavior patterns and cognitive impairment.

Results: Three distinct lifestyle behavior patterns emerged: (1) high control-high engagement (685/2060, 33.3%), (2) high control-low engagement (1210/2060, 58.7%), and (3) low control-low engagement (165/2060, 8.0%). The high control-high engagement group, characterized by non-smoking, low-to-moderate alcohol consumption, and frequent engagement in physical and social activities, exhibited the lowest risk of cognitive impairment. In contrast, participants in the high control-low engagement group (OR 1.852, 95% CI 1.314-2.655) and low control-low engagement group (OR 2.905, 95% CI 1.670-5.001) exhibited significantly higher risks. Subgroup analyses revealed that males and hypertensive individuals within the high control-low engagement group were at an even greater risk.

Conclusions: Our findings revealed that both avoiding harmful behaviors and actively engaging in health-promoting activities are important for cognitive health in older adults. Based on the results, we propose adopting a dual-pathway intervention model in policy making, simultaneously optimizing risk behaviors management and healthy behaviors promotion mechanisms.

(JMIR Public Health Surveill 2025;11:e73398) doi:10.2196/73398

KEYWORDS

lifestyle behaviors; cognitive function; latent class analysis; health behavior promotion; behavioral synergy

XSL•FO
RenderX

Introduction

With the global population aging, dementia has emerged as a critical public health challenge, with annual global costs exceeding USD 1.3 trillion [1]. China, in particular, is facing a mounting burden. Of the estimated 57.4 million individuals living with dementia worldwide in 2019, over one-quarter (15.3 million) resided in China [2]. Dementia-related deaths totaled 1.6 million globally that year, with China accounting for approximately 0.3 million deaths (19.8% of the global total) [3]. This demographic trend poses significant threats to the sustainability of health care systems unless cost-effective preventive strategies are implemented.

Given the limitations of pharmacological treatments, including high costs, potential side effects, and uncertain long-term safety and efficacy, identifying modifiable risk factors and implementing early interventions are critical for dementia prevention. The Lancet Commission [4] emphasized that up to 40% of dementia cases could be delayed or prevented by addressing modifiable risk factors through lifestyle interventions. Among the most modifiable are physical activity, social engagement, alcohol consumption, and smoking, each of which has well-established intervention frameworks. Evidence suggests that behavioral modifications targeting these domains not only enhance cognitive function in older adults but also reduce dementia incidence [5-8].

The evidence for cognitive health benefits of changing individual healthy lifestyle behaviors is robust. Many lifestyle behaviors—including smoking, alcohol consumption, and engagement in physical activity and social activity—have been identified as potential determinants of the risk of cognitive impairment. Moreover, there is some evidence that healthy behaviors often co-occur or cluster [9]. Existing studies frequently adopt a “one-size-fits-all” approach by aggregating multiple lifestyle factors into a single score [10-12]. Such an approach may obscure meaningful subgroup heterogeneity, limiting the precision of policy and intervention strategies [13]. Latent class analysis offers an alternative person-centered approach that classifies individuals into distinct classes based on shared characteristics, even when they appear similar in composite scores [14]. By maximizing between-group variation while minimizing within-group differences, LCA enables more nuanced insights into complex behavior-cognition relationships and supports the identification of distinct risk profiles in diverse older adults.

Recently, several studies have used latent class or profile analysis to identify health lifestyle-related behavior profiles and investigate the association with cognitive function [15-22]. Among these, 5 studies have focused exclusively on healthy lifestyle behaviors. However, simultaneously examining both health-promoting and health-risk behaviors allows for a more comprehensive and ecologically valid understanding of the full lifestyle continuum [15-20]. Notably, Norton et al [21] and Dingle et al [22] investigated both healthy and unhealthy behaviors among American and Australian adults, respectively. Importantly, the majority of existing evidence is derived from Western populations, whose lifestyle norms differ substantially from those in collectivist societies such as China [23]. For instance, in China, behaviors such as alcohol consumption and participation in physical activities are often embedded within social contexts that diverge markedly from those in Western cultures. As such, the generalizability of prior findings to the Chinese population remains uncertain and warrants further investigation.

In this study, we performed LCA to derive distinct lifestyle behavior classes based on both unhealthy lifestyle behaviors (smoking and alcohol consumption) and healthy behaviors (physical activity and social activity). We then examined the associations between these classes and the risk of cognitive impairment. Our findings may support the identification of older adult subgroups that would benefit most from targeted behavioral interventions aimed at reducing cognitive decline.

Methods

Study Design and Participants

A cross-sectional study was conducted from January to June 2023 in the 2 communities located in different districts of Beijing, China. Participants were recruited through community health service centers, which serve as primary care hubs in Chinese neighborhoods. Eligible older adults were approached by trained research assistants during routine medical visits or physical examinations. After receiving a detailed explanation of the study’s aims and procedures, individuals who provided informed consent were enrolled. Additional recruitment strategies included referrals by health care providers and digital outreach through WeChat-based community groups. All participants were screened and enrolled based on predefined inclusion and exclusion criteria. For older adults, the inclusion criteria were (1) aged 60 years or older and (2) having lived in this community for 6 months or longer. The residents were excluded if they had (1) vision or hearing impairment, preventing the completion of data collection and (2) mental disorders, such as schizophrenia and depression.

Measurement Indicators

Lifestyle Behaviors

Based on a previous study [12], this study focused on unhealthy lifestyle behaviors (smoking and alcohol consumption) and healthy behaviors (physical activity and social activity). Each behavior was dichotomized and scored as either healthy (score=1) or unhealthy (score=0). Smoking was categorized into a current smoker (smoking) and not a current smoker (non-smoking). The score for non-smoking status was regarded as 1; otherwise, 0. Alcohol consumption was categorized into no-to-moderate consumption (up to 1.5 standard drinks per day) or heavy consumption. The score of no-to-moderate consumption was regarded as 1; otherwise, it was 0. Physical activity was classified into 2 categories: regular engagement in physical activity, defined as engaging in physical activity for at least 3 days per week with each session lasting ≥30 minutes, and less frequent physical activity. Examples of physical activities included brisk walking, jogging, Tai Chi, dancing, and aerobics. Regular physical activity was assigned a score of 1, while less frequent physical activity was scored 0. The score

XSL•FO
RenderX

of weekly physical activity was regarded as 1; otherwise, it was 0. Finally, social engagement was categorized as active social engagement (participating in social activities at least 3 days per week (score=1) versus less frequent or no social activities (score=0). Social activities encompassed interacting with friends, playing mahjong, chess, or cards, visiting a community club, and attending sports events.

Cognitive Function

Cognitive function was assessed using the Chinese version of the Mini-Mental State Examination (MMSE). The MMSE evaluates 5 domains: orientation, memory, attention and calculation, recall, and language. The total score ranges from 0 to 30, with higher scores indicating better cognitive function. Cutoff points for cognitive impairment were determined based on education level, according to the latest normative and validation studies of the MMSE in the Chinese population. Cognitive impairment was defined as a score of ≤17 for individuals with no formal education, ≤20 for those with 1-6 years of education, and ≤24 for those with more than 6 years of education [24].

Covariates

Covariates were selected on the basis of previous evidence of associations with lifestyle and cognitive function and were ascertained by self-report. Sociodemographic covariates included gender (male or female), age in years, and marital status (married; or single, divorced, or widowed). Socioeconomic covariates included years of schooling (≤6 years, 7 - 9 years, 10 - 12 years, >12 years of education) and family economic status (poor, average, and rich). Chronic conditions were ascertained based on self-report of clinical diagnosis and included coronary heart disease, hypertension, diabetes, and hyperlipidemia. Self-report of clinical diagnosis of chronic conditions has been shown to have good agreement with ascertainment based on medical records [25]. The health status covariates included BMI (underweight, normal weight, overweight, and obesity) and annual physical examination (yes or no).

Data Collection and Quality Control

Registered nurses in community health service centers who had more than 1 year of experience were considered research assistants. Unified training in administering the questionnaire and data collection was provided to the research assistants by researchers. Questionnaires were collected in the community health center. Researchers obtained written informed consent from the patients willing to participate and told them they could withdraw from the study at any point. Well-trained research assistants collected the information through face-to-face, one-on-one interviews.

Data Analysis

Initially, LCA was performed to identify lifestyle behavior classes at baseline. Starting with a 1-category model, we systematically increased the number of classes, determining the optimal model by considering model fitting indices and the real-world significance of each category. The model fitting indices comprised (1) Akaike information criteria (AIC), Bayesian information criteria (BIC), and adjusted Bayesian information criteria, assessing model fit, with lower values indicating better fit; (2) entropy, reflecting the model's classification quality, ranging from 0 to 1, with values closer to 1 indicating higher precision; and (3) likelihood ratio test and bootstrapped likelihood ratio test for comparing consecutive models, where a *P* value less than .05 indicated the Kth model had a better fit than the K–1th model.

To verify the relationship between lifestyle behavior classes and the risk of cognitive impairment, we used binary logistic regression analysis. The outcome variable of interest was the occurrence of cognitive impairment, with the independent variable being the lifestyle behavior classes. We performed a univariate analysis, followed by a comprehensive multivariate analysis that adjusted for all relevant covariates. Subsequently, we performed binary logistic regression analyses within each covariate-defined subgroup to investigate the association between health behavior profiles and the risk of cognitive impairment specific to those subgroups. Furthermore, we examined the joint effects of individual covariates and health behavior profiles on the likelihood of cognitive impairment, providing a nuanced understanding of the variables' combined impact.

LCA was conducted in Mplus version 8.3 (Muthen & Muthen). The other statistical analyses were conducted using R (version 4.4.0; R Foundation for Statistical Computing), with binary logistic regression analysis leveraging the glm function. Moreover, additional visualizations were created using the versatile ggplot2 package. All cases with missing values have been deleted. The significance level for the 2-tailed test was established at *P* less than .05.

Ethical Considerations

Ethics approval for this study was obtained from the Ethics Committee of Beijing Hospital (approval number: 2023BJYYEC-446 - 02). Written informed consent was obtained from all participants prior to their participation. All the study data are anonymous. None of the participants received compensation. No images in the manuscript or supplementary materials contain any identifiable information of participants.

Results

Baseline Characteristics

The study screened 2434 older adults, ultimately including 2060 participants. Participants ranged in age from 60 to 94, with a mean age of 70.86 (SD 6.40). The 1203 women made up 58.4% of the population, with 857 men making up the remaining 41.6%. The 228 (11.0%) older adults had cognitive scores suggesting cognitive impairment, with an MMSE score of 28.44 (SD 2.59). The detailed results are presented in Table 1.

XSL•FO
RenderX

Table . General information of included community-dwelling older adults (n=2060).

Variable	n (%)
Gender	
Male	857 (41.6)
Female	1203 (58.4)
Age (years)	
<75	1534 (74.5)
≥75	526 (25.5)
Marital status	
Married	1938 (94.1)
Single, divorced, or widowed	122 (5.9)
Years of schooling (years)	
≤6	140 (6.8)
7 - 9	448 (21.7)
10 - 12	755 (36.7)
>12	717 (34.8)
Family economic status	
Poor	146 (7.1)
General	1337 (64.9)
Rich	577 (28.0)
Coronary heart disease	
No	1469 (71.3)
Yes	591 (28.7)
Hypertension	
No	697 (33.8)
Yes	1363 (66.2)
Diabetes	
No	1347 (65.4)
Yes	713 (34.6)
Hyperlipidemia	
No	1805 (87.6)
Yes	255 (12.4)
BMI	
Underweight	29 (1.4)
Normal	914 (44.4)
Overweight	857 (41.6)
Obesity	260 (12.6)
Annual physical examination	
No	359 (17.4)
Yes	1701 (82.6)

Latent Class Analysis of Lifestyle Behaviors

We derived 4 models via the LCA, differing in terms of the number of lifestyle behaviors classes. Table 2 presents the fit indices for the latent classes. The Akaike information criteria, Bayesian information criteria, and adjusted Bayesian information criteria values kept decreasing from the first to the third model. The LMR and bootstrapped likelihood ratio test remained

XSL•FO
RenderX

significant (P<.05) up to the 3-class model and then became non-significant. The entropy value was the highest in the 2-class model. Moreover, the conceptual sense of the 3-class model in the clinic was more significant than the 2-class model. Taking into account results obtained, we chose the solution with 3 classes.

Assigning the name for latent classes was based on a pattern of probability response to items of each latent class and comparison of the probability scale of grouping variables for each latent class. The first class consisted of 8% (165 older adults) and was characterized by low values for all 4 domains. The older adults reported a very low degree of probability in no-smoking and no-drinking domains. Also, older adults reported a low degree of probability in physical activity and social activity domains. Overall, the older adults in this class exhibited lower control levels of smoking and alcohol consumption while demonstrating low engagement in physical and social activities. We called this class "low control-low engagement." Similarly, class 2 (n=1210, 58.7%) was labeled as "high control-low engagement." Class 3 (n=685, 33.3%) was labeled as "high control-high engagement." Figure 1 shows the distribution of lifestyle behavior characteristics across the 3 classes.

Table . Fit indices for latent class analysis of health lifestyle-related behaviors among community-dwelling older adults.

Model	H0 likelihood value	AIC[a]	BIC[b]	aBIC[c]	Entropy	LMR[d] P value	BLRT[e] P value	Category probability (%)
1	−4009.695	8027.389	8049.911	8037.203	-	-	-	-
2	−3866.229	7750.459	7801.133	7772.539	1.000	<.001	<.001	92.3/7.7
3	−3776.560	7581.120	7659.947	7615.467	0.921	<.001	<.001	8.0/58.7/33.3
4	−3772.757	7583.514	7690.492	7630.128	0.762	.0562	.0505	7.0/2.0/32.2/58.7

[a]AIC: Akaike information criterion.

[b]BIC: Bayesian information criterion.

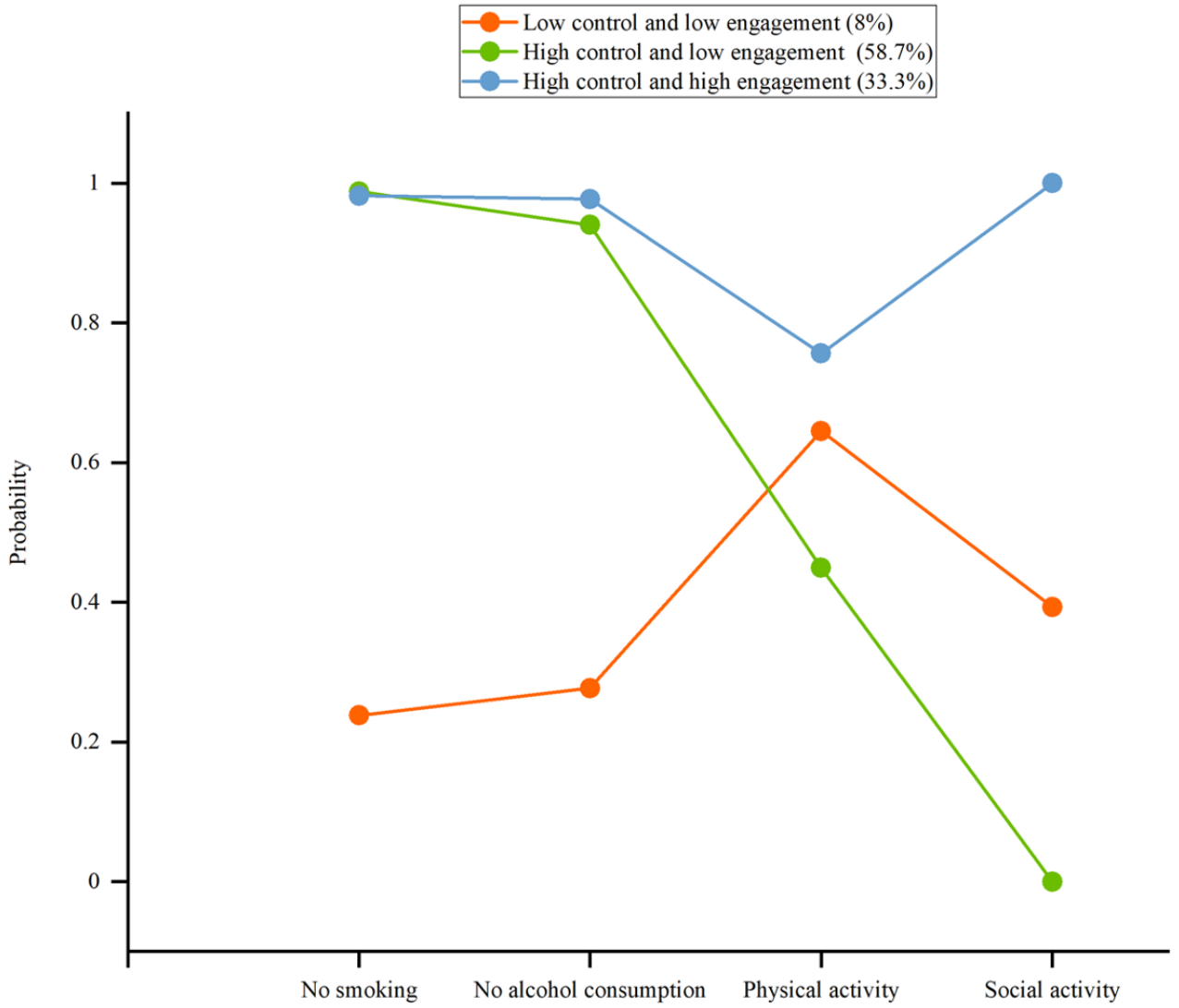
[c]aBIC: sample size-adjusted Bayesian information criterion.

[d]LMR: Lo-Mendell-Rubin adjusted likelihood ratio tests.

[e]BLRT: bootstrap likelihood ratio tests.

XSL•FO
RenderX

Figure 1. Probability distributions of lifestyle behaviors in the 3 latent classes among community dwelling older adults.



Association Between Lifestyle Behavior Classes and Cognitive Function

In the crude model, both high control-low engagement group (OR 2.026, 95% CI 1.448-2.885) and low control-low engagement group (OR 2.839, 95% CI 1.698-4.677) showed significantly higher risks of cognitive impairment compared to the high control-high engagement group.

In our fully adjusted model, using the high control-high engagement group as the reference, older participants in the high control-low engagement group were more likely to develop cognitive impairment (OR 1.852, 95% CI 1.314-2.655), while those in the low control-low engagement group were also significant (OR 2.905, 95% CI 1.670-5.001). Detailed results are presented in Table 3.

XSL•FO
RenderX

Table . Association of lifestyle behavior classes and risk of cognitive impairment among community-dwelling older adults.

Variable	Crude			Adjust[b]		
	OR[a]	95% CI	*P* value	OR	95% CI	*P* value
Lifestyle behaviors						
High control-low engagement	2.026	1.448-2.885	<.001	1.852	1.314-2.655	<.001
Low control-low engagement	2.839	1.698-4.677	<.001	2.905	1.670-5.001	<.001
High control-high engagement	1 (Ref.)		8.072	1 (Ref.)		

[a]OR: odds ratio.

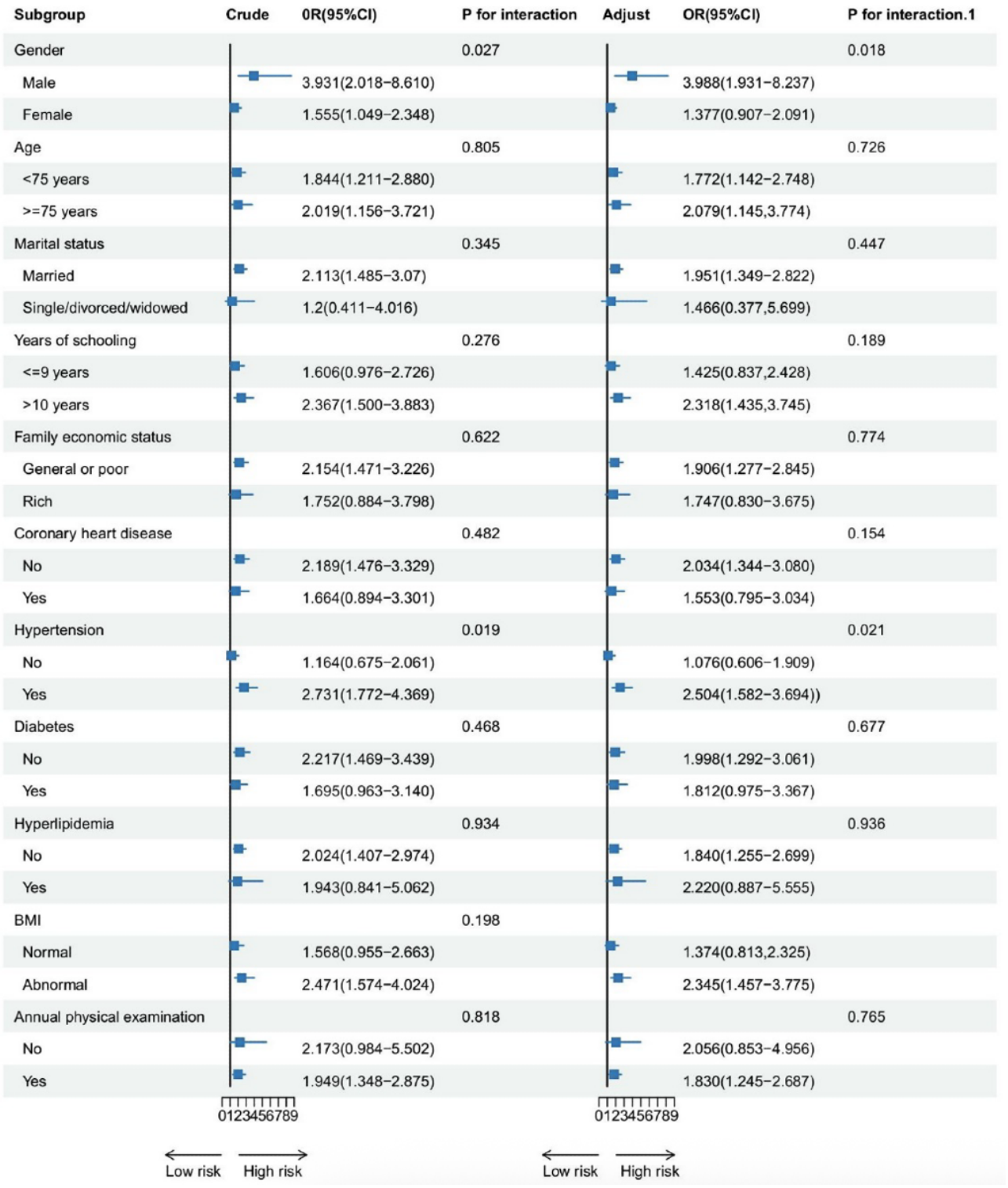
[b]Adjust model: Logistic regression model adjusted for gender, age, marital status, years of schooling, family economic status, history of coronary heart disease, history of hypertension, history of diabetes, history of hyperlipidemia, BMI, and annual physical examination.

Subgroup Analysis

Subgroup analysis showed that there were no statistically significant differences among subgroups in terms of the comparison between low control-low engagement group and high control-high engagement group. Compared to older female adults (OR 1.377, 95% CI 0.907-2.091), high control-low engagement group posed a higher risk of cognitive impairment for older male adults (OR 3.988, 95% CI 1.931 - 8.237; *P*=0.018). Compared to older adults without hypertension (OR 1.076, 95%CI 0.606 - 1.909), the high control-low engagement group posed a higher risk of cognitive impairment for older adults with hypertension (OR 2.504, 95%CI 1.582 - 3.694; *P*=0.021). Detailed results are presented in Figures 2 and 3.

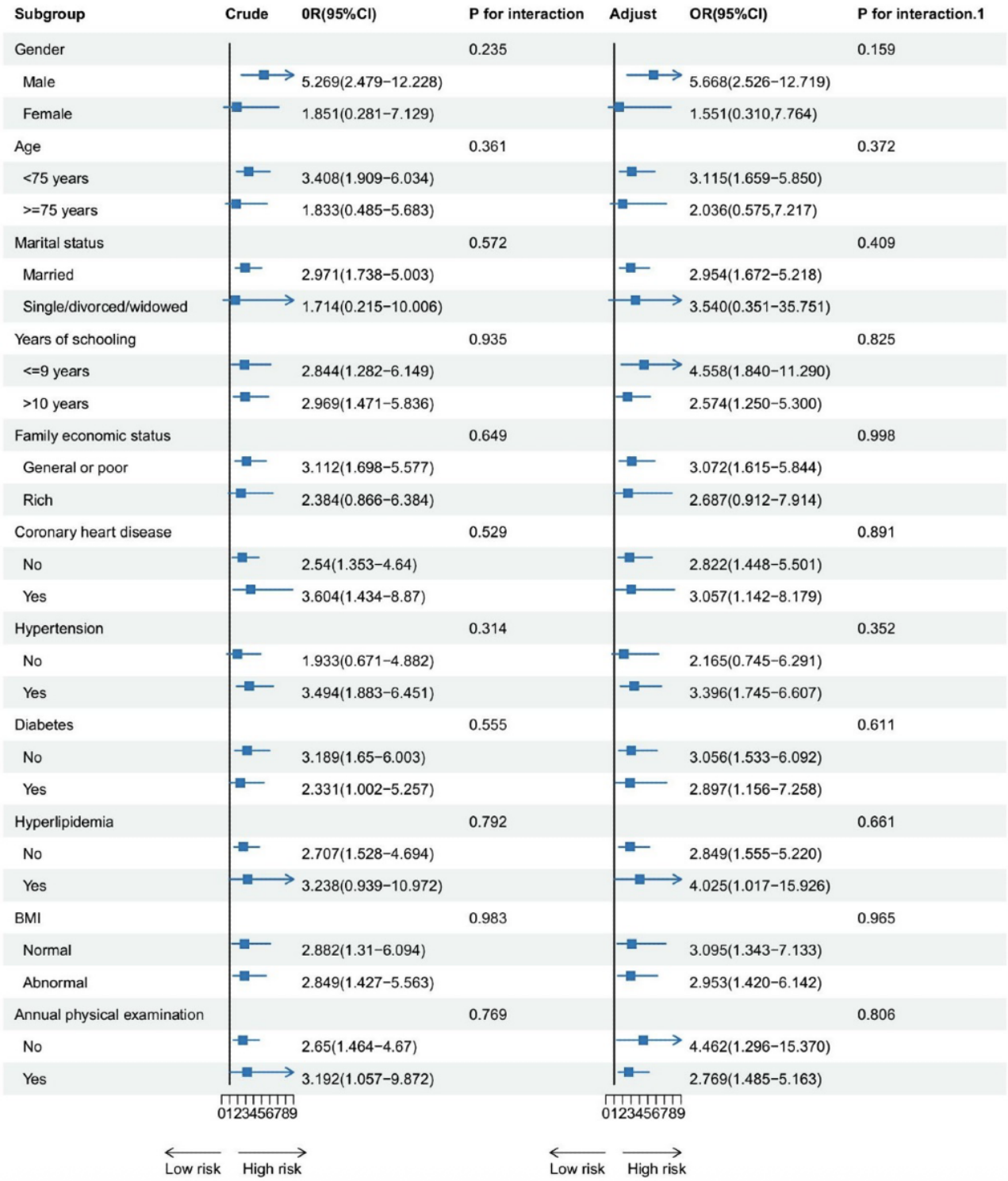
XSL•FO
RenderX

Figure 2. Subgroup analysis of the association between lifestyle classes and risk of cognitive impairment among community-dwelling older adults (high control-low engagement group vs high control-high engagement group). Adjust: model adjusted for gender, age, marital status, years of schooling, family economic status, history of coronary heart disease, history of hypertension, history of diabetes, history of hyperlipidemia, BMI, and annual physical examination. P for interaction refers to the crude model, while P for interaction.1 refers to the adjusted model.



XSL•FO
RenderX

Figure 3. Subgroup analysis of the association between lifestyle classes and risk of cognitive impairment among community-dwelling older adults (low control-low engagement group vs high control-high engagement group). Adjust: model adjusted for gender, age, marital status, years of schooling, family economic status, history of coronary heart disease, history of hypertension, history of diabetes, history of hyperlipidemia, BMI, and annual physical examination. P for interaction refers to the crude model, while P for interaction.1 refers to the adjusted model.



Discussion

Principal Findings

Using a large sample of community-dwelling older adults in China, this study used LCA to identify 3 distinct lifestyle behavior classes: high control-low engagement, low control-low engagement, and high control-high engagement group. Individuals in the high control-high engagement group, characterized by non-smoking, moderate alcohol consumption, and frequent physical and social activities, exhibited the lowest risk of cognitive impairment. In contrast, the high control-low engagement and low control-low engagement groups were associated with significantly higher risks (OR 1.852, 95% CI 1.314-2.655 and OR 2.905, 95% CI 1.670-5.001, respectively). These findings underscore the synergistic importance of both

XSL•FO
RenderX

avoiding harmful behaviors and actively engaging in health-promoting activities for cognitive health in later life.

Our findings align with existing literature emphasizing the protective effects of multidomain healthy lifestyles on cognitive function [22]. However, this study advances the field by revealing a notable divergence between self-control behaviors (eg, smoking cessation and alcohol moderation) and active engagement behaviors (eg, physical and social activities). While 58.7% of participants exhibited strong self-control but limited engagement, only 33.3% combined both domains optimally. This dichotomy suggests that public health campaigns focusing solely on risk avoidance (eg, anti-smoking policies) may insufficiently address cognitive health unless paired with strategies to promote active engagement. Such insights resonate with Katayama et al [26], who highlighted that holistic lifestyle modifications—rather than isolated behavioral changes—are critical for reversing mild cognitive impairment. Given the challenges of modifying unhealthy behaviors such as smoking and alcohol consumption, encouraging positive behaviors such as increased physical and social activity may be more achievable and impactful. Therefore, health care providers may consider prioritizing interventions that foster physical and social engagement as a foundation, gradually encouraging broader lifestyle improvements over time.

The co-occurrence of smoking and alcohol consumption observed in the low control groups may reflect shared neurobiological pathways, including genetic predispositions to addiction and psychosocial coping mechanisms for stress. Genetically, smoking and alcohol dependence share common genetic markers, increasing the likelihood of co-occurrence [27]. Psychosocially, both behaviors often serve as coping mechanisms for stress and negative emotions, further reinforcing their association [28]. What's more, the co-occurrence of smoking and alcohol consumption may reflect culturally shaped behavioral synergies in Chinese collectivist contexts. Unlike Western individualistic settings where substance use often occurs in isolation [23], these behaviors in China are deeply embedded in social rituals, for instance, cigarette sharing as a bonding gesture during family gatherings or Chinese liquor consumption as a ceremonial act in community banquets. Conversely, the protective effects observed in the high control-high engagement group may stem from culturally reinforced behavioral synergy. Group-based exercises like morning Tai Chi in parks or evening square dancing serve dual roles as both physical training and social ritual—a phenomenon seldom captured in Western studies [29]. And these activities also offer psychological benefits, such as reducing loneliness, which may further encourage simultaneous participation [30,31]. These findings challenge the direct applicability of Western-derived lifestyle interventions in China. Where previous approaches emphasize individual goal-setting, our results suggest that culturally adapted, group-oriented strategies may be more effective. For instance, redesigning alcohol control campaigns as family-level challenges or framing physical activity as community duty rather than individual choice.

Subgroup analyses further revealed that males in the high control-low engagement group faced a higher risk (OR 3.988, 95% CI 1.931 - 8.237) compared to females (OR 1.377, 95% CI 0.907-2.091). This suggests that, under similar conditions of effective control over harmful behaviors (eg, smoking, alcohol consumption), low engagement in physical or social activities presents a greater threat to cognitive health in men. This gender disparity may be attributed to differences in activity intensity, as men are more likely to engage in moderate-to-vigorous physical activities, which confer greater neuroprotective benefits [32]. Additionally, hypertensive individuals in this group exhibited heightened vulnerability, potentially due to the compounding effects of vascular dysfunction and insufficient engagement in cardioprotective activities [33-35]. These findings highlight the need to integrate conventional hypertension management (eg, pharmacotherapy) with behavioral activation strategies aimed at enhancing physical and social engagement.

These findings support the implementation of an LCA-informed stratification framework for managing lifestyle behaviors among China's aging population. For instance, for the high control-low engagement group, integrate behavioral activation into the National Basic Public Health Service Package. Community nurses could prescribe "Social Exercise Credits"—requiring ≥3 weekly activities (eg, square dancing, Tai Chi clubs), meanwhile, targeting the low control-low engagement group through the Family Doctor Contracted Services and deploying intergenerational health pledges to reduce behavioral risks. Older adults in this group receive monthly home visits jointly conducted by family doctors and their grandchildren. This tiered approach transcends conventional "one-size-fits-all" interventions by precision-targeting latent behavioral profiles while capitalizing on China's unique familial and policy architectures.

Limitations

Several limitations warrant consideration. First, the cross-sectional design precludes causal inferences between lifestyle patterns and cognitive outcomes. Second, reliance on self-reported data introduces potential recall bias, particularly for alcohol consumption and social activity metrics. Third, the convenience sampling method and geographic restriction to Beijing may limit generalizability to broader populations. Future studies should use longitudinal designs with objective measures (eg, accelerometry for physical activity, biomarkers for smoking or alcohol use) to validate these associations. Expanding recruitment to diverse regions and incorporating neuroimaging biomarkers could further elucidate the neurovascular mechanisms linking lifestyle behaviors to cognitive decline.

Conclusions

This study examined the influence of both unhealthy behaviors (smoking and alcohol consumption) and healthy behaviors (frequent engagement in physical and social activities) on cognitive health among older adults. Analytical results demonstrated that individuals in the high control-high engagement group—characterized by non-smoking, moderate alcohol consumption, and frequent physical and social activities—exhibited the lowest risk of cognitive impairment. These findings suggest that public health strategies focused exclusively on risk-reduction interventions (eg, tobacco control policies) may prove insufficient for cognitive protection unless

XSL•FO
RenderX

integrated with proactive engagement initiatives, such as the development of community-based physical activity infrastructure and programs to promote social participation among older adults. We propose adopting a dual-pathway intervention model in policy making, simultaneously optimizing risk behaviors management and healthy behaviors promotion mechanisms.

Acknowledgments

This work was supported by the National High Level Hospital Clinical Research Funding (BJ-2024-199, BJ-2024-167, BJ-2024-250), the National Natural Science Foundation of China (72304046, 72474032), the Beijing Research Ward Excellence Program (BRWEP2024W234050108), and the National Key R&D Program of China (2020YFC2008504).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

HH, YZ, and YH designed the study, performed the statistical analysis and wrote the manuscript. HL and LX engaged in data management and data quality supervision. CS designed the study and revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Wimo A, Seeher K, Cataldi R, et al. The worldwide costs of dementia in 2019. Alzheimers Dement 2023 Jul;19(7):2865-2873. [doi: 10.1002/alz.12901] [Medline: 36617519]
2. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
3. Ren R, Qi J, Lin S, et al. The China Alzheimer Report 2022. Gen Psychiatr 2022;35(1):e100751. [doi: 10.1136/gpsych-2022-100751] [Medline: 35372787]
4. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 8;396(10248):413-446. [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
5. Deal JA, Power MC, Palta P, et al. Relationship of cigarette smoking and time of quitting with incident dementia and cognitive decline. J Am Geriatr Soc 2020 Feb;68(2):337-345. [doi: 10.1111/jgs.16228] [Medline: 31675113]
6. Falck RS, Davis JC, Liu-Ambrose T. What is the association between sedentary behaviour and cognitive function? A systematic review. Br J Sports Med 2017 May;51(10):800-811. [doi: 10.1136/bjsports-2015-095551]
7. Kuiper JS, Zuidersma M, Oude Voshaar RC, et al. Social relationships and risk of dementia: a systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev 2015 Jul;22:39-57. [doi: 10.1016/j.arr.2015.04.006] [Medline: 25956016]
8. Lao Y, Hou L, Li J, Hui X, Yan P, Yang K. Association between alcohol intake, mild cognitive impairment and progression to dementia: a dose–response meta-analysis. Aging Clin Exp Res 2021 May;33(5):1175-1185. [doi: 10.1007/s40520-020-01605-0]
9. Spring B, Moller AC, Coons MJ. Multiple health behaviours: overview and implications. J Public Health (Oxf) 2012 Mar;34 Suppl 1(Suppl 1):i3-10. [doi: 10.1093/pubmed/fdr111] [Medline: 22363028]
10. Liu T, Luo H, Tang JY, Wong GH. Does lifestyle matter? Individual lifestyle factors and their additive effects associated with cognitive function in older men and women. Aging Ment Health 2020 Mar 3;24(3):405-412. [doi: 10.1080/13607863.2018.1539833]
11. Wang Z, Pang Y, Liu J, Wang J, Xie Z, Huang T. Association of healthy lifestyle with cognitive function among Chinese older adults. Eur J Clin Nutr 2021 Feb;75(2):325-334. [doi: 10.1038/s41430-020-00785-2]
12. Bloomberg M, Muniz-Terrera G, Brocklebank L, Steptoe A. Healthy lifestyle and cognitive decline in middle-aged and older adults residing in 14 European countries. Nat Commun 2024 Jun 27;15(1):5003. [doi: 10.1038/s41467-024-49262-5] [Medline: 38937442]
13. Scotto Rosato N, Baer JC. Latent class analysis: a method for capturing heterogeneity. Soc Work Res 2012 Mar 1;36(1):61-69. [doi: 10.1093/swr/svs006]
14. Mori M, Krumholz HM, Allore HG. Using latent class analysis to identify hidden clinical phenotypes. JAMA 2020 Aug 18;324(7):700-701. [doi: 10.1001/jama.2020.2278] [Medline: 32808993]
15. Zhang J, Zou L, Jiao C, et al. Cognitive benefits of activity engagement among 12,093 adults aged over 65 years. Brain Sci 2020 Dec 10;10(12):967. [doi: 10.3390/brainsci10120967] [Medline: 33321918]

XSL•FO
RenderX

16. Katayama O, Lee S, Bae S, et al. Lifestyle activity patterns related to physical frailty and cognitive impairment in urban community-dwelling older adults in Japan. J Am Med Dir Assoc 2021 Mar;22(3):583-589. [doi: 10.1016/j.jamda.2020.05.031] [Medline: 32690365]
17. Halloway S, Wagner M, Tangney C, et al. Profiles of lifestyle health behaviors and cognitive decline in older adults. Alzheimers Dement 2024 Jan;20(1):472-482. [doi: 10.1002/alz.13459] [Medline: 37676928]
18. Moored KD, Bandeen-Roche K, Snitz BE, et al. Risk of dementia differs across lifestyle engagement subgroups: a latent class and time-to-event analysis in community-dwelling older adults. J Gerontol B Psychol Sci Soc Sci 2022 May 5;77(5):872-884. [doi: 10.1093/geronb/gbab152] [Medline: 34387336]
19. Paolillo EW, Saloner R, VandeBunte A, Lee S, Bennett DA, Casaletto KB. Multimodal lifestyle engagement patterns support cognitive stability beyond neuropathological burden. Alzheimers Res Ther 2023 Dec 18;15(1):221. [doi: 10.1186/s13195-023-01365-9] [Medline: 38111051]
20. Hu H, Zhao Y, Guo D, et al. Cognitive function differs across healthy lifestyle behavior profiles: a 10-year population-based prospective cohort study. J Nutr Health Aging 2025 Feb;29(2):100487. [doi: 10.1016/j.jnha.2025.100487] [Medline: 39818010]
21. Norton MC, Dew J, Smith H, et al. Lifestyle behavior pattern is associated with different levels of risk for incident dementia and Alzheimer's disease: the Cache County study. J Am Geriatr Soc 2012 Mar;60(3):405-412. [doi: 10.1111/j.1532-5415.2011.03860.x] [Medline: 22316091]
22. Dingle SE, Bowe SJ, Bujtor M, et al. Associations between data-driven lifestyle profiles and cognitive function in the AusDiab study. BMC Public Health 2022 Nov 29;22(1):1990. [doi: 10.1186/s12889-022-14379-z] [Medline: 36443724]
23. Zhu W, Li Y. On the differences between Chinese and western cultures from the perspective of cultural dimension theory. Stud Literature Lang 2019 Feb;18:22-26. [doi: 10.3968/10811]
24. Li H, Jia J, Yang Z. Mini-Mental State Examination in Elderly Chinese: a population-based normative study. J Alzheimers Dis 2016 May 7;53(2):487-496. [doi: 10.3233/JAD-160119] [Medline: 27163822]
25. Steinkirchner AB, Zimmermann ME, Donhauser FJ, et al. Self-report of chronic diseases in old-aged individuals: extent of agreement with general practitioner medical records in the German AugUR study. J Epidemiol Community Health 2022 Aug 26;76(11):931-938. [doi: 10.1136/jech-2022-219096] [Medline: 36028306]
26. Katayama O, Lee S, Bae S, et al. Lifestyle changes and outcomes of older adults with mild cognitive impairment: a 4-year longitudinal study. Arch Gerontol Geriatr 2021;94:104376. [doi: 10.1016/j.archger.2021.104376] [Medline: 33588157]
27. Grucza RA, Bierut LJ. Co-occurring risk factors for alcohol dependence and habitual smoking: update on findings from the Collaborative Study on the Genetics of Alcoholism. Alcohol Res Health 2006;29(3):172-178. [Medline: 17373405]
28. Ruggles KV, Fang Y, Tate J, et al. What are the patterns between depression, smoking, unhealthy alcohol use, and other substance use among individuals receiving medical care? A longitudinal study of 5479 participants. AIDS Behav 2017 Jul;21(7):2014-2022. [doi: 10.1007/s10461-016-1492-9] [Medline: 27475945]
29. Luo M, Ding D, Bauman A, Negin J, Phongsavan P. Social engagement pattern, health behaviors and subjective well-being of older adults: an international perspective using WHO-SAGE survey data. BMC Public Health 2020 Jan 23;20(1):99. [doi: 10.1186/s12889-019-7841-7] [Medline: 31973695]
30. Shvedko A, Whittaker AC, Thompson JL, Greig CA. Physical activity interventions for treatment of social isolation, loneliness or low social support in older adults: a systematic review and meta-analysis of randomised controlled trials. Psychol Sport Exerc 2018 Jan;34:128-137. [doi: 10.1016/j.psychsport.2017.10.003]
31. Dahlberg L, Andersson L, Lennartsson C. Long-term predictors of loneliness in old age: results of a 20-year national study. Aging Ment Health 2018 Feb;22(2):190-196. [doi: 10.1080/13607863.2016.1247425]
32. Williams DM. Increasing fitness is associated with fewer depressive symptoms during successful smoking abstinence among women. Int J Fit 2008 Jan 1;4(1):39-44. [Medline: 20072720]
33. Ungvari Z, Toth P, Tarantini S, et al. Hypertension-induced cognitive impairment: from pathophysiology to public health. Nat Rev Nephrol 2021 Oct;17(10):639-654. [doi: 10.1038/s41581-021-00430-6] [Medline: 34127835]
34. Lachman S, Boekholdt SM, Luben RN, et al. Impact of physical activity on the risk of cardiovascular disease in middle-aged and older adults: EPIC Norfolk prospective population study. Eur J Prev Cardiol 2018 Jan;25(2):200-208. [doi: 10.1177/2047487317737628] [Medline: 29161890]
35. Han SH, Tavares JL, Evans M, Saczynski J, Burr JA. Social activities, incident cardiovascular disease, and mortality. J Aging Health 2017 Mar;29(2):268-288. [doi: 10.1177/0898264316635565] [Medline: 26944804]

Abbreviations

LCA: latent class analysis
MMSE: Mini-Mental State Examination

XSL•FO
RenderX

Edited by A Mavragani; submitted 03.03.25; peer-reviewed by O Abasilim, S Hughes; revised version received 24.06.25; accepted 26.06.25; published 12.08.25.

Please cite as:
Hu H, Zhao Y, Hao Y, Luo H, Xie L, Sun C
How Healthy and Unhealthy Lifestyle Behaviors Affect Cognitive Function—Evidence From Older Adults in Chinese Communities: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e73398
URL: https://publichealth.jmir.org/2025/1/e73398
doi:10.2196/73398

© Huixiu Hu, Yajie Zhao, Yuqing Hao, Huanhuan Luo, Lanying Xie, Chao Sun. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Tenofovir to Prevent HIV Infection in Western China: Pragmatic Randomized Controlled Trial

Yi Tao[1*], PhD; Yan Zhang[2*], Prof Dr; Bin Peng[2*], Prof Dr; Aizhong Zeng[3], Prof Dr; Jianghong Dai[4], Prof Dr; Hao Liang[5], Prof Dr; Juying Zhang[6], Prof Dr; Huarui Shao[2,7], MS; Shihan Feng[2], MS; Xiaoni Zhong[2*], Prof Dr; Ailong Huang[2*], Prof Dr

[1]Phase I Clinical Trial Ward, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

[2]School of Public Health, Chongqing Medical University, Yixue Rd1#, Yuzhong district, Chongqing, China

[3]The First Affiliated Hospital of Chongqing Medical University, Chongqing, China

[4]School of Public Health, Xinjiang Medical University, Wulumuqi, China

[5]School of Public Health, Guangxi Medical University, Nanning, China

[6]West China School of Public Health, Sichuan University, Chengdu, China

[7]Department of Pharmacy, Bishan Hospital of Chongqing Medical University, Yixue Rd 1, Yuzhong District, Chongqing, China

[*]these authors contributed equally

Corresponding Author:
Ailong Huang, Prof Dr
School of Public Health, Chongqing Medical University, Yixue Rd1#, Yuzhong district, Chongqing, China

Abstract

Background: Pre-exposure prophylaxis (PrEP) programs have been implemented in multiple countries. Evidence from clinical trials and cohort studies has proven the safety and effectiveness of PrEP. However, minimizing drug-related adverse effects and cost should be primarily considered in PrEP. Most trials used tenofovir combined with emtricitabine as the intervention; yet, the use of tenofovir disoproxil fumarate (TDF) (ie, Tenofovir) alone has not been thoroughly evaluated. Furthermore, the medication regimen in most trials was used every day, with a few studies proposing an optimal medication regimen for PrEP.

Objective: This study was designed to systematically evaluate the preventive efficacy and safety profile of TDF-based PrEP in the Chinese population. We also aimed to explore medication compliance, changes in sexual behavior, and hazard factors of HIV infection.

Methods: We conducted a pragmatic randomized controlled trial (RCT) to evaluate the effectiveness and safety of TDF for HIV PrEP. Participants were randomly assigned (1:1:1) to a time-driven group (TDF 300 mg administered orally once daily), an event-driven group (TDF 300 mg administered orally 24 to 48 h before sexual activity and 2 hours after sexual activity, not exceeding 300 mg within 24 h), or an untreated control group. The primary outcomes were the effectiveness and safety of TDF during periods of PrEP use. Secondary outcomes focused on the effectiveness of TDF among participants with good compliance during PrEP use. Tertiary outcomes included the risk factors of HIV infection and behavioral changes from PrEP initiation to the last visit. For ethical reasons, all participants received condoms and health education. This study was registered with the Chinese Clinical Trial Registry (ChiCTR-TRC-13003849).

Results: A total of 1914 participants underwent randomization. During the follow-up of 3513.5 person-years from June 2013 to May 2016, HIV seroconversion was observed in 30 persons (2.02 per 100 person-years) in the time-driven group (time-driven vs control group: hazard ratio [HR] 0.93, 95% CI 058 - 1.51; *P*=.78), 35 (1.73 per 100 person-years) in the event-driven group (event-driven group vs control group: HR 0.83, 95% CI 0.52 - 1.31; *P*=.42), and 37 (2.06 per 100 person-years) in the control group. Post hoc analysis showed that participants with good medication compliance reduced their HIV infection risk by 53% (*P*=.01) and event-driven medication with good compliance reduced the risk by 62% (*P*=.009). We recorded no severe adverse events during the trial. For tertiary outcomes, low medication compliance, sexual role, no condom use, and more number of sexual partners remained significantly associated with HIV risk.

Conclusions: The TDF-based PrEP is ineffective without good adherence. However, when medication compliance is achieved, event-driven dosing is recommended as an optimal PrEP regimen.

Trial Registration: ChiCTR-TRC-13003849; https://www.chictr.org.cn/showproj.html?proj=5716

(JMIR Public Health Surveill 2025;11:e71494) doi:10.2196/71494

XSL•FO
RenderX

KEYWORDS

pre-exposure prophylaxis; HIV; tenofovir; medication compliance; prep regimen; MSM

Introduction

While the global AIDS epidemic remains a persistent public health threat, advancements in antiretroviral therapy and prevention strategies have substantially reduced mortality and improved treatment access, albeit unevenly across regions and populations. In 2023, there were approximately 1.3 million new HIV infections and 630,000 AIDS-related deaths globally, reflecting reductions in both new infections and mortality compared to previous years [1]. HIV remains a major public health issue in China, with an estimated 570,576 HIV infections between 2005 and 2020, particularly concentrated in western China [2]. In 2023, data released by China's National Administration of Disease Control and Prevention and the Chinese Center for Disease Control and Prevention (China CDC) provided a clear overview of the HIV/AIDS epidemic in mainland China. Notably, the provinces of Sichuan, Guangxi, and Chongqing reported the highest HIV infection rates nationwide. Additionally, regions such as Xinjiang, Yunnan, and Guizhou also emerged as high-prevalence areas, with the epidemic primarily concentrated in western China [3]. In 2023, men who have sex with men (MSM), bisexual individuals, sex workers, and people who inject drugs remained the highest-risk groups for HIV, with sexual transmission accounting for more than 95% of new infections in China [1,4]. MSM continue to represent the group with the highest HIV incidence rate globally, with infection rates rising nearly 15-fold from 1.5% in 2006 to 25.6% in 2022 [5,6].

In the absence of an effective vaccine, core strategies for AIDS prevention include HIV testing [7], circumcision [8], condom use, and health education [9]. However, these measures alone are insufficient. The introduction of tenofovir-based pre-exposure prophylaxis (PrEP) has revolutionized HIV prevention globally [10]. PrEP involves taking antiretroviral drugs before potential exposure to HIV. It is now widely available as a biomedical prevention method to reduce the risk of HIV infection. The coformulation of two nucleoside or nucleotide reverse transcriptase inhibitors has proven highly effective as PrEP in various populations worldwide [11]. The compound tablet emtricitabine (ETC)/tenofovir disoproxil fumarate TDF (Truvada) was licensed by the US Food and Drug Administration in 2012 as the first HIV PrEP drug, with studies showing it can reduce HIV incidence by 44% [12]. However, PrEP's effectiveness is closely linked to adherence [13]. A notable study has demonstrated that cisgender women consistently adhering to a daily or high-adherence regimen of PrEP using ETC and TDF experienced significantly low HIV incidence [14]. Regarding the optimal dosing regimen, recent research suggests no significant difference in efficacy between daily and event-driven PrEP regimens [15,16]. A study conducted in sub-Saharan Africa indicated that event-driven PrEP regimen provided protection across foreskin tissue, highlighting the need for further clinical evaluation of precoital PrEP, specifically for insertive sex [16].

Despite substantial investments in HIV prevention over the past decades, this field continues to struggle with inadequate long-term investment strategies and resource disparities [17]. PrEP has been implemented as a national health program in multiple countries [18]. Acknowledging its substantial cost, ensuring cost-effectiveness and demonstrating potential efficacy are paramount considerations in the implementation of PrEP [19-21]. The majority of PrEP studies have focused on coformulation of two nucleoside or nucleotide reverse transcriptase inhibitors [12,21-23]. Nevertheless, the exploration of PrEP, particularly the use of TDF monotherapy, remains inconclusive. TDF alone is associated with a lower cost, making it a more accessible option in resource-limited settings where budget constraints are critical [24]. Additionally, by using TDF alone, the regimen may potentially reduce the incidence of certain side effects such as gastrointestinal discomfort or other FTC-related adverse reactions, which could improve patient adherence and overall quality of life [23]. If TDF is proven to be both effective and safe, it should be regarded as a pivotal public health intervention.

We designed a randomized controlled trial in western China with the primary objective of assessing the effectiveness and safety of TDF use alone among MSM. Furthermore, we aimed to gain insights into its efficacy and safety when administered with high compliance rates. Additionally, we aimed to delve into identifying the optimal medication regimen, hazard factors contributing to HIV infection, as well as examining changes in sexual behavior patterns.

Methods

Study Design

This trial was a multiprovince, open-label, randomized controlled trial conducted in western China, provinces including Chongqing, SiChuan, GuangXi, and XinJiang.

Participants

A total of 1500 participants were planned for recruitment at four sites, including Chongqing, Guangxi, Xinjiang, and Sichuan, using a combination of competitive enrollment and snowball sampling methods. By leveraging existing cohorts in each province, researchers streamlined the recruitment process. Eligible participants were MSM aged 18 to 65 years, who reported engaging in sexual activity at least once every two weeks on average and having had at least one sexual encounter with a same-sex partner in the month preceding the trial. Exclusion criteria included participants who were HIV-, HBV-, or HCV-positive; had severe underlying medical conditions; had a history of alcoholism or drug dependence within the past year prior to screening; or had a history of severe allergies.

Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki and ICH-GCP (International Council for Harmonisation Good Clinical Practice guidelines). All volunteers provided written informed consent prior to the start

XSL•FO
RenderX

of the study. Participants were informed that study data would be deidentified and their participation was voluntary. They were assured that their information would be used solely for this study, and their questionnaire responses would not affect their careers. The study covered the medical expenses of any drug-related adverse reactions or adverse events. No additional financial compensation was provided to participants beyond the standard study procedures. However, participants assigned to the intervention group received free TDF medication, and all participants had access to free condoms and comprehensive health education. In case of irreversible damage, corresponding economic compensation was provided in accordance with national laws. The trial protocol was approved by the Ethics Committee of Chongqing Medical University and registered with the Chinese Clinical Trial Registry (ChiCTR-TRC-13003849).

Randomization and Masking

Randomization was performed using a central, computerized system using block randomization, stratified by hospital. Except for questionnaire administrators, neither patients nor other researchers were blinded to group allocation. Participants were randomly assigned in a 1:1:1 ratio to the time-driven group (TDF 300 mg administered orally once daily), the event-driven group (Tenofovir 300 mg administered orally 24 - 48 h before and 2 h after sexual activity, not exceeding 300 mg within 24 h), or the untreated control group.

Procedures

Baseline information encompassing demographic data, medical histories, and sexual behavior patterns was meticulously documented (see Multimedia Appendix 1). Comprehensive laboratory assessments, including tests for transfusion-associated infections such as HIV, HBV, and HCV, along with electrocardiograms, were conducted, with all participants undergoing an HIV test on the day of randomization. Interventions were tailored to the standard operating procedures at each study site. Follow-up visits were scheduled at regular intervals of 12 weeks, extending up to 96 weeks, and encompassed a range of evaluations. These included repeat HIV testing, routine laboratory investigations, assessments of medication adherence, inquiries into sexual behavior and sexually transmitted infections (STIs), collection of medication history, and comprehensive physical examinations. Adherence to these follow-up visits was mandatory for all participants.

Outcomes

The primary outcomes were effectiveness and safety of TDF during individuals' periods of PrEP use. Secondary outcomes assessed the effectiveness of TDF for participants with good compliance during periods of PrEP use. Tertiary outcomes included the risk factors of HIV infection and behavioral change from PrEP initiation to the last visit.

Statistical Analysis

Categorical data were presented as numbers (frequencies) and analyzed using 2-tailed χ^2 tests or Fisher exact test. Continuous variables were presented as means and SD for normally distributed data and median and IQR for skewed data. Groups were compared using one-way ANOVA or the Kruskal-Wallis test. Overall incidence rates per 100 person-years during the follow-up period were calculated for all adverse drug reactions. The risk of HIV infection was analyzed using Cox proportional hazards models, with hazard ratios, corresponding 95% CIs, and associated *P* values.

The 'survcutpoint' function from the R package *survminer* was used to determine the cutoff value for medication compliance, which was estimated based on pill counts and self-reports. The optimal critical value for good medication compliance was established by combining data from previous studies [23]. The Kaplan-Meier method and log-rank tests were used to analyze the incidences of HIV seroconversion and medication compliance rates among the groups. The multiple imputation method was used to address missing data on the number of sexual partners, sexual acts, and condom use, with five imputations conducted.

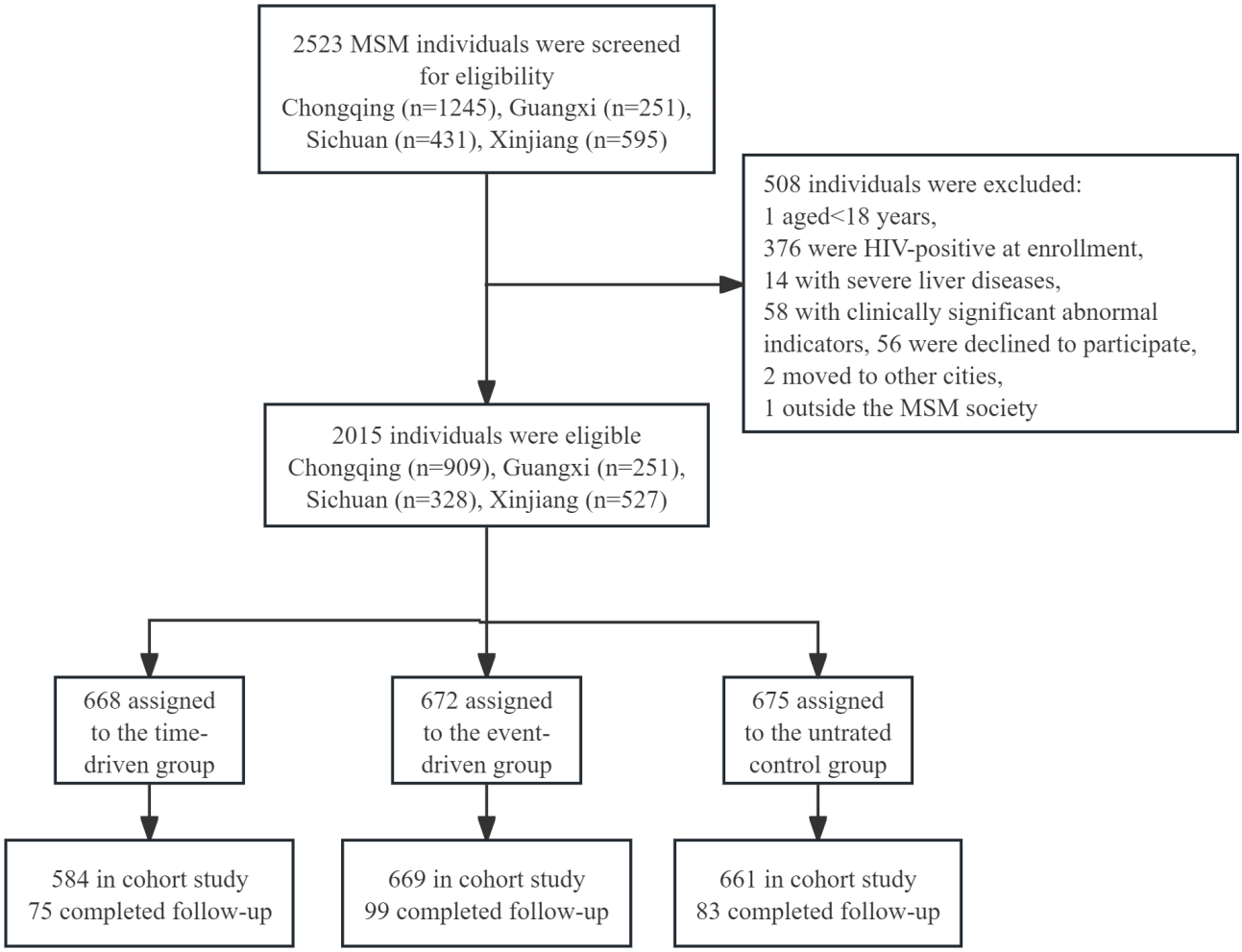
Statistical analyses were conducted using R software (version 4.2.1; Foundation for Statistical Computing) and Stata software (version 17; StataCorp). A two-sided *P* value <.05 was considered statistically significant. For comparison between the three groups, the significance level was adjusted using Bonferroni adjustment.

Results

From June 2013 to May 2016, 2015 participants were enrolled across four sites: Chongqing (n=909), Guangxi (n=251), Xinjiang (n=328), and Sichuan (n=527), among whom 1914 individuals took at least one medication and participated in the study. This cohort was subsequently followed for a total of 3513.5 person-years, with a median follow-up period of 1.5 years per participant (Figure 1). Baseline characteristics were balanced among the three groups: time-driven (n=584; mean 29, SD 8.4 years), event-driven (n=669; mean age of 30 SD 8.4 years), and untreated control group (n=661; mean age of 30.5 SD 9.2 y) (Table 1).

XSL•FO
RenderX

Figure 1. Participants flow chart. MSM: men who have sex with other men.



XSL•FO
RenderX

Table . Baseline characteristics of participants (N=1914).

Characteristics	Time-driven group (n=584)	Event-driven group (n=669)	Untreated control group (n=661)
Age (years), median (IQR)	29.02 (8.43)	29.98 (8.38)	30.48 (9.23)
Nationality (n=1913), n (%)			
Han	541 (92.80)	617 (92.23)	619 (93.65)
Uygur	3 (0.51)	7 (1.05)	6 (0.91)
Hui	7 (1.20)	10 (1.49)	11 (1.66)
Other	32 (5.49)	35 (5.49)	25 (3.78)
Degree of education (n=1912), n (%)[a]			
High school and below	225 (38.53)	239 (35.72)	288 (43.57)
Colleges	142 (24.36)	155 (23.20)	168 (25.42)
Bachelor's degree or above	216 (37.05)	274 (41.02)	205 (31.01)
Registered residence (n=1908), n (%)			
Urban	415 (71.31)	497 (74.40)	456 (69.30)
Rural	167 (28.69)	171 (25.60)	202 (30.70)
Marriage, n (%)[a]			
Unmarried	458 (78.42)	491 (73.39)	478 (72.31)
Married	87 (14.90)	110 (16.44)	130 (19.67)
Other	39 (6.68)	68 (10.16)	53 (8.02)
Occupation (n=1906)			
Public institutions and enterprises	65 (11.21)	86 (12.89)	73 (11.08)
Professional work	128 (22.07)	182 (27.29)	142 (21.55)
Physical work	30 (5.17)	43 (6.45)	49 (7.44)
Commercial work	82 (14.14)	89 (13.34)	86 (13.05)
Service Industry	89 (15.34)	79 (11.84)	95 (14.42)
Farmers and herdsmen	3 (0.52)	2 (0.30)	3 (0.46)
Students	91 (15.58)	94 (14.05)	83 (12.59)
Unemployed	54 (9.25)	49 (7.32)	76 (11.53)
Other	38 (6.51)	43 (6.43)	52 (7.89)
Monthly disposable income (Chinese Yuan)			
≤3000	309 (52.91)	326 (48.73)	366 (56.22)
3000 - 5000	192 (32.88)	248 (37.07)	212 (32.57)
≥5000	75 (12.84)	88 (13.15)	73 (11.21)

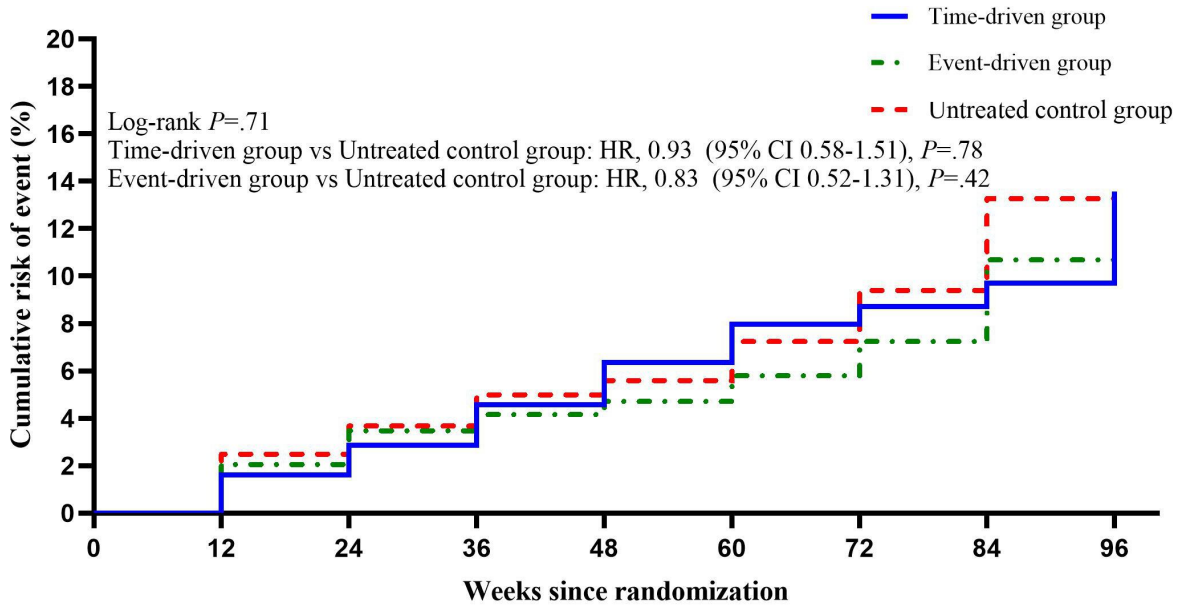
[a]Statistical differences among the three groups.

Primary Outcomes

HIV seroconversion was observed in 30 participants (incidence rate [IR] 5.1%, 2.02 per 100 person-years) in the time-driven group (time-driven group vs untreated control group, hazard ratio [HR] 0.93; 95% CI, 0.58 - 1.51, *P*=.78), 35 participants in the event-driven group (IR 5.2%, 1.73 per 100 person-years) (event-driven group vs untreated control group: HR 0.83; 95% CI, 0.52 - 1.31, *P*=.42) and 37 participants in the untreated control group (IR 5.6%, 2.06 per 100 person-years) (Figure 2).

XSL•FO
RenderX

Figure 2. Kaplan-Meier estimates of HIV seroconversion in the time-driven, event-driven, and untreated control groups.

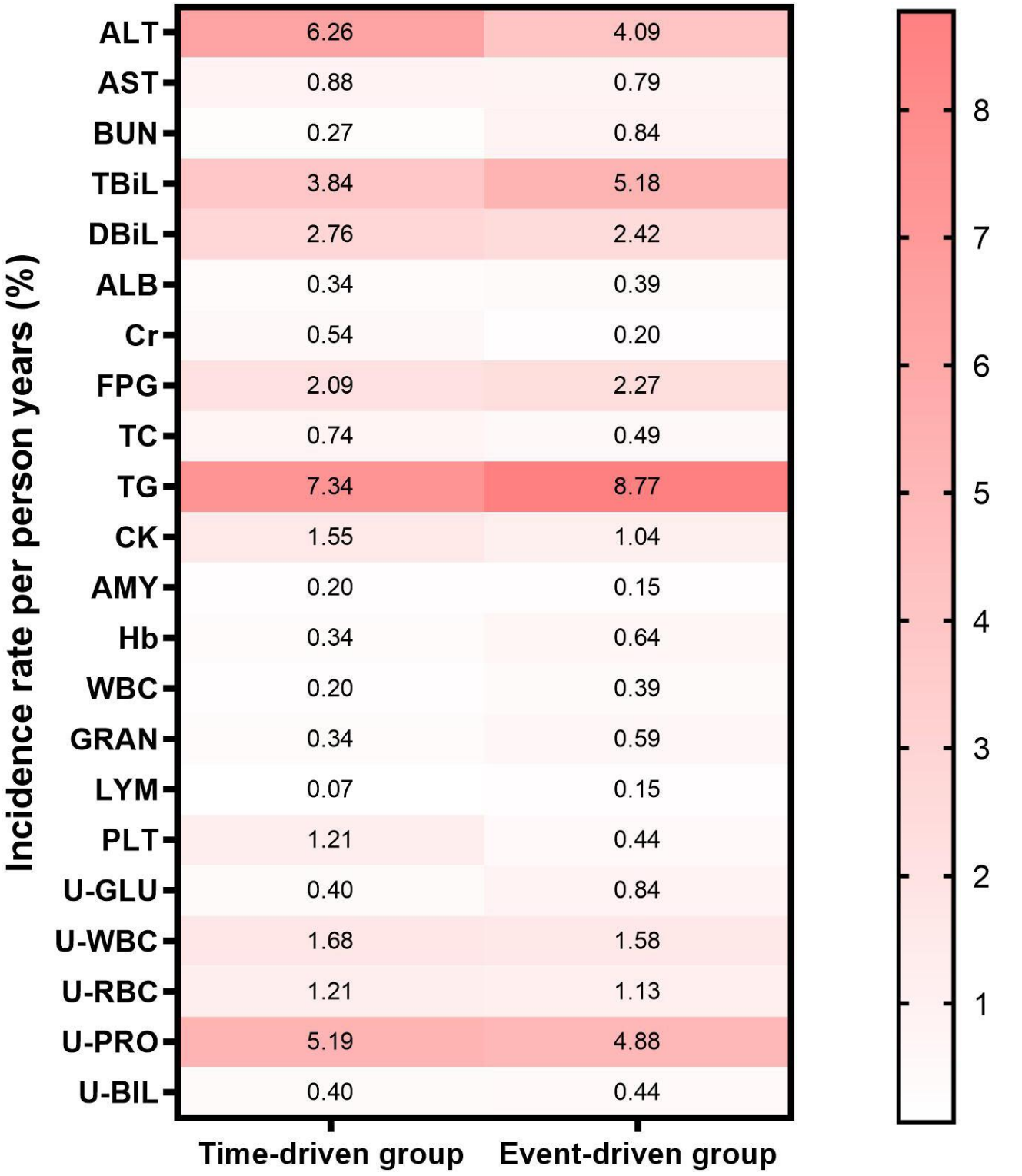


No. at risk/weeks									
Weeks	0	12	24	36	48	60	72	84	96
Time-driven group	584	437	389	340	269	174	123	62	47
Event-driven group	669	536	484	420	349	264	195	135	65
Blank group	661	484	408	367	317	229	172	117	46

No severe adverse events were reported. For dosing as a precautionary medication, laboratory test indicators exceeding 1.5 to 2 times the upper limit of the normal range were regarded as abnormal. The incidence of composite drug-related adverse events was 39.33%. Rates of abnormality in laboratory indicators in the time-driven and event-driven groups were 37.9% and 37.7%, respectively. The toped three abnormalities were elevated triglycerides (8.2%), alanine aminotransferase (ALT; 5.0%), and urine protein (5.0%). Symptomatic adverse events were reported in 1.5% and 1.6% of participants, respectively, with the top-ranked three abnormalities being diarrhea, vomiting, and nausea (Figure 3).

XSL•FO
RenderX

Figure 3. Adverse events in time-driven group and event-driven groups. ALT: Alanine Aminotransferase; AST: Aspartate Aminotransferase; BUN: Blood Urea Nitrogen; TBiL: Total Bilirubin; DBiL: Direct Bilirubin (also known as Conjugated Bilirubin); ALB: Albumin; Cr: Creatinine; FPG: Fasting Plasma Glucose; TC: Total Cholesterol; CK: Creatine Kinase; AMY: Amylase; Hb: Hemoglobin; WBC: White Blood Cell Count; GRAN: Granulocyte Percentage; LYM: Lymphocyte Percentage; PLT: Platelet Count; U-GLU: Urine Glucose; U-WBC: Urine White Blood Cells; U-RBC: Urine Red Blood Cells; U-PRO: Urine Protein; U-BIL: Urine Bilirubin.



The cohort was followed for 3513.5 person-years with a variable duration of visits (median 1.5 y). HIV seroconversion was observed in 30 participants (HR, 5.1%, 2.02 per 100 person-years) in the time-driven group (time-driven group vs blank group, HR, 0.93; 95% CI, 0.58 - 1.51, *P*=.78), 35 (5.2%, 1.73 per 100 person-years) in the event-driven group (event-driven group vs blank group, hazard ratio, 0.83; 95% CI, 0.52 - 1.31, *P*=.420) and 37 (5.6%, 2.06 per 100 person-years) in the blank group.

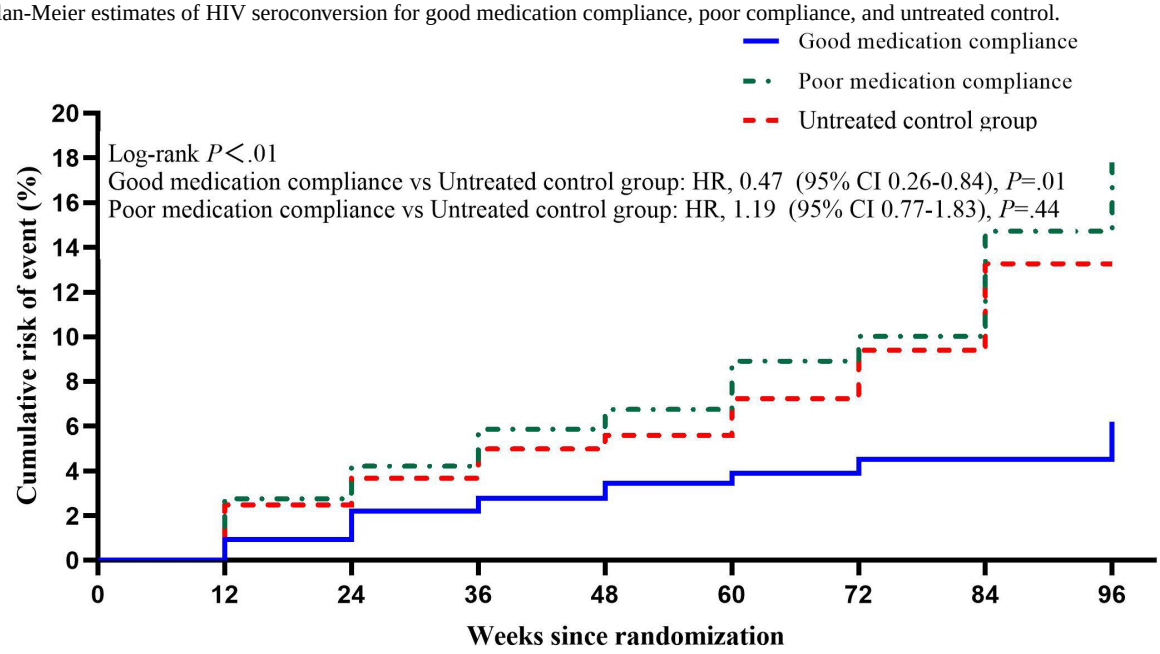
Secondary Outcome

The overall average medication compliance rate was 68.4%; event-driven group (77.0%) revealed higher adherence than in

XSL•FO
RenderX

the time-event group (57.4%). The cutoff value calculated by *R* software was 40%, double the cutoff value and integrated with previous studies, we defined the optimal critical value of good medication compliance as 80% [25]. Kaplan-Meier estimates showed that participants with good medication compliance had a 53% reduced hazard of HIV infection compared with the untreated group (P=.01) (Figure 4). Additionally, time-driven PrEP reduced HIV infection hazard by 62% among those with ≥80% compliance (P<.01) (Figure 5).

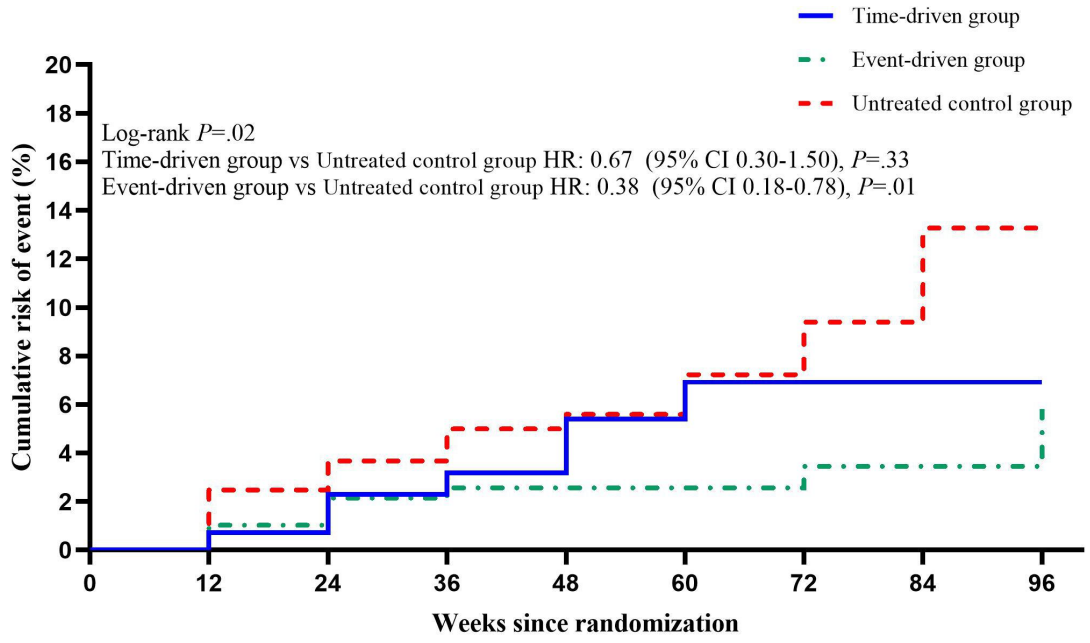
Figure 4. Kaplan-Meier estimates of HIV seroconversion for good medication compliance, poor compliance, and untreated control.



No. at risk/weeks									
Weeks	0	12	24	36	48	60	72	84	96
Time-driven group	439	434	389	342	286	218	151	110	57
Event-driven group	814	509	468	404	321	215	164	115	55
Blank group	661	484	408	367	317	229	172	117	46

XSL•FO
RenderX

Figure 5. Kaplan-Meier estimates for HIV seroconversion in time-driven group, event-driven group, and untreated control groups with medication compliance rate ≥80%.



No. at risk/weeks									
Weeks	0	12	24	36	48	60	72	84	96
Time-driven group	141	140	125	110	88	62	41	28	16
Event-driven group	298	294	264	232	198	156	110	82	41
Blank group	661	484	408	368	319	231	175	125	46

After adjustment for medication compliance rate ≥ 80%, good medication compliance reduced the hazard of HIV infection by 53% (*P*=.01).

Similarly, event-driven dosing reduced the hazard by 62% (*P*=.01).

Tertiary Outcome

Both unadjusted and adjusted HRs for factors associated with HIV infection during follow-up are shown in Table 2. Bivariable Cox regression analyses identified several characteristics significantly associated with HIV infection, including medication compliance, residence, sex role, condom use, number of sexual partners, casual sexual behavior, STI diagnosis, commercial sex, and drug addiction. In the multivariable model, medication compliance, sex role, condom use, and number of sex partners remained significantly associated with HIV risk (Table 2).

XSL•FO
RenderX

Table . Unadjusted and adjusted hazard ratios for factors associated with HIV diagnosis among participants during follow-up n=1914. Note: 1. Sex role, “0” represents a receptive/bottom role; “1” represents a insertive/top role; “0.5” represents a versatile/switch role. 2. “Drug” refers to recreational substances (eg, alcohol, cannabis, ecstasy, poppers).

Characteristics	Number of HIV seroconversion (%)	Bivariable analysis		Multivariable analysis	
		Unadjusted hazard ratio (95% CI)	*P* value	Adjusted hazard ratio (95% CI)	*P* value
Group					
Untreated control group (n=661)	37 (5.60)	1 (Reference)	—[b]	—	—
Time-driven group (n=584)	30 (5.14)	0.93 (0.58 - 1.51)	.78	—	—
Event-driven group (n=669)	35 (5.23)	0.83 (0.52 - 1.31)	.42	—	—
Dosing			.51		—
No (n=661)	37 (5.60)	1 (Reference)			
Yes (n=1253)	65 (5.19)	0.87 (0.58 - 1.31)			
PrEP compliance			.001	—	.001
≥80% (n=434)	16 (3.69)	1 (Reference)		1 (Reference)	
<80% (n=509)	49 (9.63)	2.55 (1.44 - 4.51)		2.97 (1.59 - 5.54)	
Age	—	0.98 (0.95 - 1.00)	.07	—	—
Residence	—	—	.04	—	.07
Urban (n=1368)	65 (4.75)	1 (Reference)		1 (Reference)	
Rural (n=540)	37 (6.85)	0.66 (0.44 - 0.98)		1.67 (0.95 - 2.93)	
Education					
High school and below (n=236)	14 (5.93)	1 (Reference)		—	—
College degree (n=981)	60 (6.12)	1.07 (0.60 - 1.92)	.82	—	—
Bachelor’s and above (n=695)	28 (4.03)	0.71 (0.37 - 1.35)	.30	—	—
Career					
Civil service institutions (n=109)	4 (3.67)	1 (Reference)	—	—	—
Corporation (n=824)	39 (4.73)	0.53 (0.18 - 1.56)	.25	—	—
Service industry (n=263)	18 (6.84)	0.68 (0.39 - 1.19)	.18	—	—
Physical labor (n=263)	18 (6.84)	0.96 (0.50 - 1.85)	.91	—	—
Marriage			.77	—	—
Unmarried (n=1587)	83 (5.23)	1 (Reference)			
Married (n=327)	19 (5.81)	1.04 (0.78 - 1.40)			
Monthly income					
≤3000 (n=1001)	58 (5.79)	1 (Reference)	—	—	—
3000-5000 (n=652)	34 (5.21)	0.95 (0.63 - 1.46)	.83	—	—
≥5000 (n=236)	9 (3.81)	0.79 (0.39 - 1.60)	.51	—	—
Sexual roles					
1 (n=923)	43 (4.66)	1 (Reference)	—	1 (Reference)	—
0 (n=485)	38 (7.84)	1.72 (1.11 - 2.66)	.02	2.22 (1.25 - 3.95)	.007

XSL•FO
RenderX

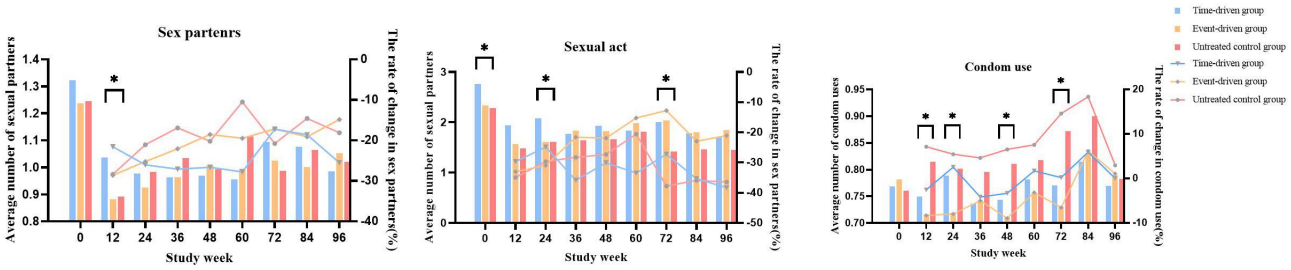
Characteristics	Number of HIV seroconversion (%)	Bivariable analysis		Multivariable analysis	
		Unadjusted hazard ratio (95% CI)	*P* value	Adjusted hazard ratio (95% CI)	*P* value
0.5 (n=498)	21 (4.22)	0.94 (0.56 - 1.59)	.83	0.82 (0.38 - 1.77)	.61
Condom use in last 6 months (%)			.02		.02
Always (n=982)	45 (4.58)	1 (Reference)		1 (Reference)	
Sometimes or never (n=757)	43 (6.74)	1.65 (1.10 - 2.46)		1.88 (1.11 - 3.20)	
Sex partners in 6 months (%)					
1-5 (n=1729)	82 (4.74)	1 (Reference)		1 (Reference)	
>5 (n=116)	18 (15.52)	3.54 (2.12 - 5.90)	<.001	2.60 (1.28 - 5.30)	.008
STI[a] diagnosis			.005		.15
Ever diagnosed with STI (n=156)	17 (10.90)	1 (Reference)		1 (Reference)	
Never diagnosed with STI (n=1753)	84 (4.79)	0.47 (0.28 - 0.79)		0.56 (0.26 - 1.24)	
Commercial sexual behavior			.005		.06
Yes (n=107)	12 (11.21)	1 (Reference)		1 (Reference)	
No (n=1801)	89 (4.94)	0.42 (0.23 - 0.77)		0.44 (0.18 - 1.04)	
Drug contacts			.05		.31
Never contacted with drugs (n=1830)	96 (5.25)	1 (Reference)		1 (Reference)	
Ever contacted with drugs (n=54)	5 (9.26)	2.46 (1.00 - 6.07)		1.86 (0.56 - 6.18)	

[a]STI: sexually transmitted infection.

[b]Not applicable.

Regarding shifts in sexual behaviors, the number of sexual partners exhibited a fluctuating pattern, initially decreasing across all groups when compared to the baseline, but subsequently increased later during follow-up. Participants in the time-driven group reported the highest number of sexual acts, followed by the event-driven and untreated control groups. However, a positive correlation emerged between condom use and the follow-up period. Notably, participants in the untreated control group reported the highest frequency of condom use, followed by the time-driven and event-driven groups reporting the lowest levels (Figure 6).

Figure 6. Changes in sexual practice of MSM during the follow-up. MSM: men who have sex with men.



The event-driven group exhibits a higher total number of symptom abnormalities (n=33 adverse events) compared to the time-driven group (n=22). Specifically, the number of diarrhea abnormalities was 6 in the time-driven group and 11 in the event-driven group; for vomiting, 1 versus 5; for dizziness, 1 versus 4; and for nausea, 2 versus 3, in the time-driven and event-driven groups, respectively. Conversely, the time-driven group showed a higher count in the "Others" category, with 10 abnormalities compared to 5 in the event-driven group.

XSL•FO
RenderX

Discussion

Principal Findings

Although the data from this study are nearly a decade old, our findings remain highly relevant in the field of PrEP, particularly in China. To our knowledge, this study represents the most extensive PrEP trial conducted in China to date, characterized by its comprehensive scope and substantial sample size, which collectively enhance the robustness and generalizability of our findings. Through rigorous statistical analysis, we evaluated the feasibility, efficacy, and safety of TDF as a standalone PrEP intervention. Our analysis demonstrates that while TDF administered as standalone PrEP is safe, it does not provide significant protection against HIV. However, under conditions of high adherence (≥80% dosing compliance), TDF exhibits a substantial 53% reduction in HIV infection risk, underscoring the critical role of medication consistency in achieving prophylactic efficacy. Furthermore, we advocate for an event-driven medication regimen to balance safety with cost-effectiveness. Our analysis also highlights several risk factors for HIV acquisition, including poor PrEP compliance, receptive sexual role, occasional condom use, and having multiple sex partners (>5). Notably, contrary to early studies, we observed sexual behavioral disinhibition during the follow-up period, further underscoring the complexity of sexual practices and the need for targeted interventions [12,26].

The combination of tenofovir and emtricitabine has demonstrated efficacy in preventing HIV acquisition, which was licensed as PrEP in China in August 2020 [12]. However, due to the high incidence of adverse drug reactions observed in pilot trials and the barrier of expensive drug prices, we ultimately decided to explore the efficacy and safety of using TDF alone as PrEP. Initiating treatment with a foundational drug was pivotal in mitigating the potential risk of HIV drug resistance associated with PrEP utilization. While individuals engaging in behaviors associated with heightened HIV risk may experience disparities in certain health indicators, epidemiological data suggest this population often demonstrates baseline physiological resilience. This relative health stability, coupled with tenofovir's well-documented safety profile in diverse cohorts, supports its consideration as a sustainable PrEP option for long-term use, particularly when integrated with adherence support systems. Furthermore, given the good medication compliance observed in pilot trials and the relatively easier recruitment of individuals, MSM were chosen as participants for the pivotal trials.

The HIV incidence reported in our study was notably lower than that observed in previous investigations [4,12], which may be attributable to the comprehensive health education and consistent condom distribution implemented at every study visit. No statistically significant difference emerged between the medication groups and the untreated control group. However, tenofovir monotherapy was found to confer an additional 53% reduction in the hazard of HIV infection. Furthermore, when medication was administered on an event-driven basis, this protective effect was even more pronounced, with a 62% reduction in the hazard of HIV infection. These findings highlight the importance of adherence and tailored dosing schedules in maximizing the preventative potential of tenofovir as PrEP [27,28].

After comprehensive multivariable adjustment, only poor PrEP compliance was significantly related to sexual practices among the five major hazard components (HR > 2). Adherence to daily PrEP dosing was influenced by several factors, notably diminished motivation when individuals regarded PrEP as a situational preventive tool rather than an integral part of their routine health management. This perception often undermines consistent pill-taking, as prophylactic measures are frequently viewed as temporary responses to risk rather than sustained health practices [29]. Echoing earlier findings from the pioneering PrEP clinical trials and the landmark PROUD study, our results highlight the paramount importance of high medication adherence in maximizing PrEP's effectiveness in preventing HIV seroconversion [12].

While sexual practices are well-established risk factors for HIV infection [30], our study observed a phenomenon of sexual behavioral disinhibition during the follow-up period. This aligns with a systematic review of 17 open-label PrEP studies, which consistently reported a rise in sexual practice and bacterial STIs among PrEP users [31]. HIV prevention measures encompass HIV testing, the free distribution of condoms, health education, and HIV counseling. After initiating a PrEP scheme, there may be a reduction in condom use and the frequency of HIV testing, or an increase in sexually transmitted infections (STIs). The medication group's slightly disinhibited behavior could be attributed to a multifaceted interplay of psychological, behavioral, and biological factors. First, risk compensation psychology suggests that individuals may subconsciously adjust their behavior in response to perceived reductions in HIV risk, a cognitive bias where decreased vigilance towards HIV acquisition risks aligns with Prospect Theory—perceived risk reduction increases tolerance for riskier choices [32]. Second, miscalibration of PrEP efficacy may play a role, as overestimation of PrEP's protective effects (eg, ignoring the need for adherence) or underestimating the persistence of other STIs can fuel riskier practices; notably, a study in JAMA Network Open found that 30% of PrEP users overestimated its efficacy against STIs [13]. Third, social norms and sexual scripting within MSM communities may inadvertently normalize riskier behaviors if narratives around "treatment as prevention" (TasP) or "undetectable=untransmittable" are decoupled from broader sexual health education [31]. Finally, biological factors might contribute, as reduced anxiety about HIV acquisition could heighten sexual arousal and motivation, indirectly elevating behavioral frequency or intensity, even though PrEP itself does not increase libido[加5]. Together, these elements create a complex dynamic that may underpin the observed behavioral disinhibition in the medication group. This indicates the need for comprehensive risk reduction counseling and ongoing education to mitigate potential behavioral changes that may undermine the protective benefits of PrEP. Experience from the USA, UK, and Australia suggests that to address these issues, it is important to enhance STI testing, establish a robust surveillance system, and base prescriptions on both HIV and STI test results [4,6].

XSL•FO
RenderX

Implementing PrEP as a nationwide program is imperative, given the compelling evidence that targeting high-risk individuals with PrEP is not only effective but also cost-efficient [19]. While the adoption of PrEP holds significant promise, it's crucial to acknowledge a potential drawback: the development of long-term "tolerance," which could potentially narrow down treatment avenues in the event of an acute HIV infection [33]. As a fundamental antiviral, tenofovir stands out for its ability to circumvent drug resistance concerns that often complicate treatment regimens. Furthermore, its cost-effectiveness is a significant advantage, alleviating financial burdens for both governmental bodies and individual patients alike. By embracing tenofovir-based PrEP, we can forge a path towards a more accessible and sustainable approach to HIV prevention. An event-driven medication regimen was recommended for PrEP. The IPERGAY trial demonstrated that an event-driven regimen of tenofovir disoproxil fumarate-emtricitabine reduced incidence of HIV infection by 86% [34]. Pharmacokinetically, tenofovir reaches its maximum concentration (T_{max}) in approximately 1 hour and has a half-life ($T_{1/2}$) of around 18 hours. This event-driven regimen maintains drug concentration at a high level in vivo. Additionally, medication compliance was higher with the event-driven regimen compared to daily dosing. In the event-driven group, the median number of pills used was four per month, compared to an average of sixteen pills reported in the IPERGAY trial for event-driven use, making it significantly less costly than daily medication.

The study has several limitations. First, the absence of a placebo and blinding may introduce bias in the safety assessment, as participants in the time-event group were unwilling to take the medication daily [34]. The use of a placebo or blinding could lead to unblinding, potentially underestimating compliance. Second, to closely mimic real-world conditions, no additional interventions were implemented to maintain the cohort or improve medication compliance during the trial, resulting in a high drop-out rate and poor adherence. The reliance on self-reported adherence and pill counts may also introduce bias, as these methods can be subjective and prone to overestimation. Finally, while the participants in this study may not represent all MSM in China, the large sample size helped mitigate this concern to some extent [35].

Conclusion

In conclusion, the TDF is ineffective without good adherence. When reliable medication compliance is attainable, event-driven dosing emerges as a highly recommended and effective PrEP strategy. These findings resonate profoundly with the practical challenges and opportunities within China, compellingly underscoring the critical urgency of integrating PrEP into national healthcare policies. By prioritizing and implementing such policies, we can fortify our defenses against the spread of HIV, thereby safeguarding the health and well-being of all citizens and fostering a more resilient public health infrastructure.

Acknowledgments

This study was funded by The National Major Science and Technology Project (AIDS and hepatitis, and other major infectious disease control and prevention) (2008ZX10001-016).

ALH was awarded a National Major Science and Technology project (AIDS and hepatitis, and other major infectious disease control and prevention) to support the implementing of this study; he is dedicated to infectious diseases research for two decades. We thank the participants and the dedication of the team members.

The funder of the study had no role in study design, data collection, data analysis, data interpretation, and writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Generative AI was not used in any portion of the manuscript writing.

Data Availability

The datasets generated or analyzed during this study are not publicly available due the sensitive nature of the data and the terms outlined in the informed consent, but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: ALH (lead), XNZ (equal)
Data curation: SHF, HRS
Formal analysis: YT (lead), YZ (supporting)
Funding acquisition: ALH
Investigation: YZ, BP, AZZ, JHD, HL, JYZ
Methodology: YT
Project administration: AZZ, JHD, HL, JYZ
Resources: ALH, XNZ
Supervision: ZXN
Validation: YZ, BP

XSL•FO
RenderX

Visualization: ALH (lead), XNZ (equal)
Writing – original draft: YT (lead), BP and YZ (supporting)
Writing – review & editing: ALH, XNZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study questionnaire
[DOCX File, 81 KB - publichealth_v11i1e71494_app1.docx]

Checklist 1
CONSORT checklist
[PDF File, 219 KB - publichealth_v11i1e71494_app2.pdf]

References

1. Laurence J. “AIDS at a Crossroads:” Highlights from the 2024 UNAIDS Report. AIDS Patient Care STDS 2024 Nov 1;38(11):493-494. [doi: 10.1089/apc.2024.0172]
2. Piri Sahragard H, Karami P. Spatiotemporal analysis of seasonal trends in land surface temperature within the distribution range of Moringa peregrina (Forssk.) in Southern and Southeastern Iran. PLoS ONE 2024;19(7):e0306642. [doi: 10.1371/journal.pone.0306642]
3. Cai C, Tang H, Li D, et al. Evolution of HIV epidemic and emerging challenges - China, 1989-2023. China CDC Wkly 2024 Nov 29;6(48):1251-1256. [doi: 10.46234/ccdcw2024.251] [Medline: 39698324]
4. McCormack S, Dunn DT, Desai M, et al. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. Lancet 2016 Jan 2;387(10013):53-60. [doi: 10.1016/S0140-6736(15)00056-2] [Medline: 26364263]
5. Cohen J, Lo YR, Caceres CF, Klausner JD, Grp WGW. WHO guidelines for HIV/STI prevention and care among MSM and transgender people: implications for policy and practice. Sex Transm Infect 2013 Nov;89(7):536-538. [doi: 10.1136/sextrans-2013-051121] [Medline: 24123866]
6. Traeger MW, Guy R, Asselin J, et al. Real-world trends in incidence of bacterial sexually transmissible infections among gay and bisexual men using HIV pre-exposure prophylaxis (PrEP) in Australia following nationwide PrEP implementation: an analysis of sentinel surveillance data. Lancet Infect Dis 2022 Aug;22(8):1231-1241. [doi: 10.1016/S1473-3099(22)00175-X] [Medline: 35643090]
7. Lockman S, Holme MP, Makhema J, et al. Implementation of universal HIV testing and treatment to reduce HIV incidence in Botswana: the Ya Tsie study. Curr HIV/AIDS Rep 2020 Oct;17(5):478-486. [doi: 10.1007/s11904-020-00523-0] [Medline: 32797382]
8. Weidner W. Re: Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial. Eur Urol 2007 Aug;52(2):605-606. [doi: 10.1016/j.eururo.2007.04.090]
9. Faust L, Yaya S. The effect of HIV educational interventions on HIV-related knowledge, condom use, and HIV incidence in sub-Saharan Africa: a systematic review and meta-analysis. BMC Public Health 2018 Nov 13;18(1):1254. [doi: 10.1186/s12889-018-6178-y] [Medline: 30424761]
10. Abdool Karim SS, Baxter C, Abdool Karim Q. Advancing HIV prevention using tenofovir-based pre-exposure prophylaxis. Antivir Ther (Lond) 2022 Apr;27(2). [doi: 10.1177/13596535211067589]
11. Ambrosioni J, Petit E, Liegeon G, Laguno M, Miró JM. Primary HIV-1 infection in users of pre-exposure prophylaxis. Lancet HIV 2021 Mar;8(3):e166-e174. [doi: 10.1016/S2352-3018(20)30271-X] [Medline: 33316212]
12. Berger RE. Re.: Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. J Urol 2011 May;185(5):1729-1730. [doi: 10.1016/S0022-5347(11)60198-5] [Medline: 22088703]
13. Ong JJ, Baggaley RC, Wi TE, et al. Global epidemiologic characteristics of sexually transmitted infections among individuals using preexposure prophylaxis for the prevention of HIV infection: a systematic review and meta-analysis. JAMA Netw Open 2019 Dec 2;2(12):e1917134. [doi: 10.1001/jamanetworkopen.2019.17134] [Medline: 31825501]
14. Marrazzo J, Tao L, Becker M, et al. HIV preexposure prophylaxis with emtricitabine and tenofovir disoproxil fumarate among cisgender women. JAMA 2024 Mar 19;331(11):930. [doi: 10.1001/jama.2024.0464]
15. Molina JM, Ghosn J, Assoumou L, et al. Daily and on-demand HIV pre-exposure prophylaxis with emtricitabine and tenofovir disoproxil (ANRS PREVENIR): a prospective observational cohort study. Lancet HIV 2022 Aug;9(8):e554-e562. [doi: 10.1016/S2352-3018(22)00133-3] [Medline: 35772417]
16. Herrera C, Serwanga J, Else L, et al. Dose finding study for on-demand HIV pre-exposure prophylaxis for insertive sex in sub-Saharan Africa: results from the CHAPS open label randomised controlled trial. EBioMedicine 2023 Jul;93:104648. [doi: 10.1016/j.ebiom.2023.104648] [Medline: 37327677]

XSL•FO
RenderX

17. Du AB, Head M, Brede M. Integration vs segregation: network analysis of interdisciplinarity in funded and unfunded research on infectious diseases. J Informetr 2025 Feb;19(1):101634. [doi: 10.1016/j.joi.2024.101634]
18. Grulich AE, Guy R, Amin J, et al. Population-level effectiveness of rapid, targeted, high-coverage roll-out of HIV pre-exposure prophylaxis in men who have sex with men: the EPIC-NSW prospective cohort study. Lancet HIV 2018 Nov;5(11):e629-e637. [doi: 10.1016/S2352-3018(18)30215-7] [Medline: 30343026]
19. Bórquez A, Guanira JV, Revill P, et al. The impact and cost-effectiveness of combined HIV prevention scenarios among transgender women sex-workers in Lima, Peru: a mathematical modelling study. Lancet Public Health 2019 Mar;4(3):e127-e136. [doi: 10.1016/S2468-2667(18)30236-6] [Medline: 30683587]
20. Wu L, Kaftan D, Wittenauer R, et al. Health impact, budget impact, and price threshold for cost-effectiveness of lenacapavir for HIV pre-exposure prophylaxis in eastern and southern Africa: a modelling analysis. Lancet HIV 2024 Nov;11(11):e765-e773. [doi: 10.1016/S2352-3018(24)00239-X] [Medline: 39312933]
21. Desai M, Field N, Grant R, McCormack S. Recent advances in pre-exposure prophylaxis for HIV. BMJ 2017 Dec 11;359:j5011. [doi: 10.1136/bmj.j5011] [Medline: 29229609]
22. Ngure K, Ortblad KF, Mogere P, et al. Efficiency of 6-month PrEP dispensing with HIV self-testing in Kenya: an open-label, randomised, non-inferiority, implementation trial. Lancet HIV 2022 Jul;9(7):e464-e473. [doi: 10.1016/S2352-3018(22)00126-6] [Medline: 35777410]
23. Kearney BP, Flaherty JF, Shah J. Tenofovir disoproxil fumarate: clinical pharmacology and pharmacokinetics. Clin Pharmacokinet 2004;43(9):595-612. [doi: 10.2165/00003088-200443090-00003] [Medline: 15217303]
24. Wu B, Shen JF, Cheng HF. Cost-effectiveness analysis of different rescue therapies in patients with lamivudine-resistant chronic hepatitis B in China. BMC Health Serv Res 2012 Nov 8;12:12. [doi: 10.1186/1472-6963-12-385] [Medline: 23137013]
25. Pampati S, Emrick K, Siegler AJ, Jones J. Changes in sexual behavior, PrEP adherence, and access to sexual health services because of the COVID-19 pandemic among a cohort of PrEP-using MSM in the South. J Acquir Immune Defic Syndr 2021 May 1;87(1):639-643. [doi: 10.1097/QAI.0000000000002640] [Medline: 33512848]
26. Molina JM, Capitant C, Spire B, et al. On-Demand preexposure prophylaxis in men at high risk for HIV-1 infection. N Engl J Med 2015 Dec 3;373(23):2237-2246. [doi: 10.1056/NEJMoa1506273] [Medline: 26624850]
27. Liu AY, Cohen SE, Vittinghoff E, et al. Preexposure prophylaxis for HIV infection integrated with municipal- and community-based sexual health services. JAMA Intern Med 2016 Jan;176(1):75-84. [doi: 10.1001/jamainternmed.2015.4683] [Medline: 26571482]
28. Grulich AE, Jin F, Bavinton BR, et al. Long-term protection from HIV infection with oral HIV pre-exposure prophylaxis in gay and bisexual men: findings from the expanded and extended EPIC-NSW prospective implementation study. Lancet HIV 2021 Aug;8(8):e486-e494. [doi: 10.1016/S2352-3018(21)00074-6] [Medline: 34217426]
29. Wahome EW, Graham SM, Thiong'o AN, et al. PrEP uptake and adherence in relation to HIV-1 incidence among Kenyan men who have sex with men. EClinicalMedicine 2020 Sep;26:100541. [doi: 10.1016/j.eclinm.2020.100541] [Medline: 33089128]
30. Rodger AJ, Cambiano V, Bruun T, et al. Sexual activity without condoms and risk of HIV transmission in serodifferent couples when the hiv-positive partner is using suppressive antiretroviral therapy. JAMA 2016 Jul 12;316(2):171-181. [doi: 10.1001/jama.2016.5148] [Medline: 27404185]
31. Traeger MW, Schroeder SE, Wright EJ, et al. Effects of pre-exposure prophylaxis for the prevention of Human Immunodeficiency Virus infection on sexual risk behavior in men who have sex with men: a systematic review and meta-analysis. Clin Infect Dis 2018 Aug 16;67(5):676-686. [doi: 10.1093/cid/ciy182] [Medline: 29509889]
32. Spees LP, Wirth KE, Mawandia S, Bazghina-Werq S, Ledikwe JH. Sexual risk compensation following voluntary medical male circumcision: results from a prospective cohort study amongst human immunodeficiency virus-negative adult men in Botswana. South Afr J HIV Med 2020;21(1):1157. [doi: 10.4102/sajhivmed.v21i1.1157] [Medline: 33391832]
33. van de Vijver DAMC, Nichols BE, Abbas UL, et al. Preexposure prophylaxis will have a limited impact on HIV-1 drug resistance in sub-Saharan Africa: a comparison of mathematical models. AIDS 2013 Nov 28;27(18):2943-2951. [doi: 10.1097/01.aids.0000433237.63560.20] [Medline: 23939237]
34. Goldwirt L, Bauer R, Liegeon G, et al. Estimated pill intake with on-demand PrEP with oral TDF/FTC using TFV-DP concentration in dried blood spots in the ANRS IPERGAY trial. J Antimicrob Chemother 2021 Sep 15;76(10):2675-2680. [doi: 10.1093/jac/dkab253] [Medline: 34278433]
35. Dong MJ, Peng B, Liu ZF, et al. The prevalence of HIV among MSM in China: a large-scale systematic analysis. BMC Infect Dis 2019 Nov 27;19(1):1000. [doi: 10.1186/s12879-019-4559-1] [Medline: 31775654]

Abbreviations

MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
STI: sexually transmitted infection
TDF: tenofovir disoproxil fumarate

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 20.01.25; peer-reviewed by S Chen, X Su; revised version received 01.04.25; accepted 28.04.25; published 20.08.25.

Please cite as:
Tao Y, Zhang Y, Peng B, Zeng A, Dai J, Liang H, Zhang J, Shao H, Feng S, Zhong X, Huang A
Tenofovir to Prevent HIV Infection in Western China: Pragmatic Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e71494
URL: https://publichealth.jmir.org/2025/1/e71494
doi:10.2196/71494

© Yi Tao, Yan Zhang, Bin Peng, Aizhong Zeng, Jianghong Dai, Hao Liang, Juying Zhang, Huarui Shao, Shihan Feng, Xiaoni Zhong, Ailong Huang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Knowledge, Attitudes, and Practices to Periodontal Health of the Northeast Chinese Public: Cross-Sectional Study

Chengde Jin[1*]; Ming Chi[1*], MD; Chunxiao Li[1]; Yushan Li[1]; Wanting Wang[1]; Yushi Zhang[1], MSc; Qiao Liu[2], MD

[1]China Medical University, Shenyang City, China

[2]Huludao Central Hospital, Huludao, Liaoning Province, China

[*]these authors contributed equally

Corresponding Author:
Qiao Liu, MD
Huludao Central Hospital, Huludao, Liaoning Province, China

Abstract

Background: Periodontitis affects approximately 50% of adults in China and is a leading cause of tooth loss in this population. However, there is a notable paucity of research on knowledge, attitudes, and practices (KAP) related to periodontitis among patients in Northeast China.

Objective: This study aimed to investigate the KAP regarding periodontitis among populations in Northeast China, focusing on 5 demographic factors: gender, age, income, education level, and region.

Methods: A cross-sectional survey was conducted by convenience sampling over a period of 1 week. A structured questionnaire was used to collect detailed responses on periodontitis-related KAP. Descriptive statistics (means, SDs, frequencies, and percentages) were used. Normality was assessed by Shapiro-Wilk tests. Nonparametric Kruskal-Wallis and multivariate regression analyses were conducted to examine demographic influences and interactions on periodontal KAP scores. Statistical significance was defined at $P<.05$.

Results: A total of 619 questionnaires were distributed, resulting in 562 valid responses comprising 242 (43.06%) males and 320 (56.94%) females, with a mean participant age of 41.27 (95% CI 37.4 - 45.1) years. The overall awareness of periodontal disease was relatively low in Northeast China, with the mean KAP scores being 3.88/8, 5.28/7, and 5.19/11. Age and educational level were both significantly associated with individuals' KAP regarding periodontitis ($P<.05$), whereas gender showed a significant association with knowledge only ($P<.05$). Regional and income-related differences were generally significant, with only a variable showing marginal effects ($P=.05$). Multiple regression analysis indicated that both knowledge and attitude scores tend to increase with age. The increase in knowledge was most pronounced in the age group of 41 - 50 (Nagelkerke Pseudo Coefficient of Determination (Nagelkerke R^2)=1.58, 95% CI 0.83 - 2.33; $P<.01$) years. Attitude scores exhibited a more consistent upward trend across all age groups. In contrast, practice scores declined with age. In terms of interactions, young females exhibited significantly higher awareness than males, whereas no significant gender differences were observed among older populations. In addition, higher education levels and economic status were strongly associated with improved awareness. Notably, the presence of gingival bleeding significantly enhanced public awareness of periodontitis, especially knowledge score (Nagelkerke R^2 =1.07, 95% CI 0.69 - 1.44; $P<.01$).

Conclusions: This study provides a comprehensive understanding of the KAP regarding periodontitis among populations in Northeast China. The findings offer valuable insights for the formulation of targeted policies and underscore the importance of improving periodontal KAP in the region.

(JMIR Public Health Surveill 2025;11:e72069) doi:10.2196/72069

KEYWORDS

KAP; sociodemographic factors; gingival bleeding; Knowledge, Attitudes, and Practices; periodontal health awareness

Introduction

Periodontitis is the sixth most prevalent chronic noncommunicable disease worldwide [1], characterized by inflammation and destruction of the 4 periodontal supporting tissues [2]. This condition can cause tooth loosening, displacement, and eventual loss, resulting in impaired chewing and speaking abilities [3]. Furthermore, periodontitis is implicated in the development of various systemic conditions, including cardiovascular disease, type 2 diabetes, rheumatoid arthritis, and inflammatory bowel disease [4]. These complications significantly diminish patients' quality of life.

Globally, periodontitis affects approximately 12.5% of the population and ranks as the second most common oral disease

XSL•FO
RenderX

after dental caries in permanent teeth, impairing oral health for over 1 billion people worldwide [5,6]. A recent meta-analysis estimated that about 62% of dentate adults have periodontitis, with approximately 23.6% classified as severe cases, marking a significant increase compared to prevalence rates from 1990 to 2010 [7]. The situation in China is similarly concerning, as gingival inflammation affects as many as 61.0% of adolescents aged 14 years, and over half of the individuals aged 35 years and older have periodontal disease, which notably deteriorates with increasing age [8,9]. Despite the widespread prevalence of periodontal disease, it remains insufficiently addressed due to its silent progression and low public awareness, with patients typically seeking treatment only at advanced stages [2,10,11]. Consequently, this delay significantly exacerbates the burden associated with periodontitis.

The primary risk factor for periodontitis is the accumulation of microbial plaque within the oral cavity [12]; thus, maintaining good oral hygiene is essential for the effective prevention of periodontal disease. Effective plaque control through proper brushing techniques, flossing or interdental cleaning, and using fluoride-containing toothpaste and mouthwash is critical to maintaining oral hygiene [13]. Although preventive measures appear straightforward theoretically, their implementation at both individual and public health levels often presents significant challenges [14]. Studies indicate that the high prevalence of periodontal disease is strongly associated with inadequate public awareness and understanding of the disease [15]. Therefore, investigating public awareness regarding periodontitis and formulating corresponding policies are crucial steps toward reducing its prevalence and burden.

Knowledge, Attitudes, and Practices (KAP) surveys represent a structured research method designed to assess individuals' understanding, perceptions, and behaviors related to specific health conditions or preventive practices. These studies aim to identify obstacles hindering the adoption of desirable health behaviors, highlight specific health concerns within target and broader populations, and inform the implementation of targeted oral health education policies based on the findings, ultimately enhancing community health awareness and easing public health burdens [16-18].

Understanding the KAP of the population regarding periodontitis is vital for developing targeted oral health education policies tailored to different demographic groups. It also facilitates the creation of referral pathways for patients with periodontal disease and comorbid medical conditions, thereby supporting the integration and transition of care. These efforts are essential for maintaining oral health in older adults and achieving the "80 - 20" goal, which aims for individuals aged 80 years to retain at least 20 functional teeth. Although nationwide studies have investigated periodontal knowledge and behaviors in China [19], or have been limited to surveys of individuals with periodontitis [20], no research has specifically addressed the northeastern population. Crucially, prior studies have neither systematically examined how sociodemographic variables interact to shape periodontal KAP in this region nor have they considered the role of gingival bleeding as a contributing factor. This study addresses this gap using a representative sample from northeastern China to evaluate the multivariate and combined effects of 5 key demographic factors, such as gender, age, education, income, and geographic classification, on periodontal KAP. It identifies high-priority subpopulations for policy intervention and offers insights potentially generalizable to other regions with similar demographic profiles.

Method

Questionnaire

The questionnaire consisted of 23 items divided into 5 sections, with all questions carefully designed to avoid leading or ambiguous language. The first section collected basic demographic information. The second section assessed participants' knowledge of periodontitis through 8 multiple-response questions. The third section evaluated attitudes toward periodontitis with 3 questions. The fourth section measured participants' practices with 5 questions, focusing on daily oral care and professional dental care. The fifth section included a question on gingival bleeding, which served as an indicator of potential periodontal disease. As supplementary information to the questionnaire, the title clearly outlined the study objectives, emphasized voluntary participation, detailed privacy protection measures, and informed participants that the collected data would serve as raw data for publication. A statement, "I have read and understood the information provided in the title and voluntarily agree to participate in this survey,"was included as a question item, with "Agree" or "Disagree" as response options. In addition, if participants selected an age between 12 and 17 years, an additional item querying guardian consent ("Yes" or "No") was presented.

Sample Method and Recruitment

A cross-sectional survey using a convenience sampling method was conducted between December 31, 2024 and January 7, 2025, to evaluate KAP regarding periodontitis in Northeast China. Prior to questionnaire distribution, a pilot study involving these 42 individuals was conducted to ensure clarity and comprehensibility of all questions.

This cross-sectional study was conducted across 3 provinces in Northeast China (Liaoning, Jilin, and Heilongjiang), encompassing both economically higher-income and lower-income regions to ensure demographic and socioeconomic diversity. From each province, 1 economically developed and 1 underdeveloped city were randomly selected, resulting in 6 cities. Within each city, 1 urban district and 1 rural township were randomly chosen, yielding a total of 12 distinct regions for participant recruitment.

Recruitment was conducted via a community-based approach. Local community health centers, secondary schools, and residential committees in the selected districts and townships collaborated in disseminating the questionnaire link to potential participants through WeChat (Tencent; a widely used communication platform in China). Volunteers from local health institutions, including public health nurses and dental practitioners, assisted in explaining the survey and encouraging participation. They also ensured that participants understood the purpose of the study and were available to address any

XSL•FO
RenderX

questions. The online questionnaire was distributed via the Wenjuanxing (Changsha Ranxing Information Technology Co Ltd) platform, and participation was entirely voluntary.

Evaluation Criterion

The scoring criteria were as follows: Questionnaire items were divided into 3 categories. For single-choice questions with 1 correct answer, responses were scored as either correct (1) or incorrect (0). Likert-scale items were scored on a scale from −2 to 2. Items assessing daily behaviors were scored from 1 to 3 based on the extent to which responses aligned with established behavioral standards.

The criteria for inclusion and exclusion were as follows: responses were excluded if participants did not explicitly consent to the informed consent and privacy policy. For participants aged between 12 and 17 years, additional consent from a guardian was required. Questionnaire responses relating to KAP that demonstrated clear irregularities, such as selecting identical answers for all items, were also excluded. As the questionnaire included prompts for unanswered questions, there were no missing values. All remaining data were retained for subsequent analyses.

Reliability Analysis and Validity Assessment

To assess the internal consistency of the questionnaire, Cronbach α was calculated for each KAP dimension. The results were as follows: Knowledge (α=0.437), Attitude (α=0.417), and Practice (α=0.319). These values suggest moderate to low internal consistency, which may be attributed to the multifaceted nature of the questionnaire. Since the study primarily focuses on descriptive statistical analysis and regression modeling rather than scale validation, no further item reduction or confirmatory factor analysis was performed. To evaluate the suitability of the dataset for correlation and regression analyses, Bartlett test for sphericity was conducted. The test yielded a significant result (χ^2_{153}=1159.3; P<.001), indicating sufficient correlation among variables.

Statistical Methods

All statistical analyses were performed using IBM SPSS Statistics 27.0 (IBM Corp), RStudio (Posit, PBC), and WPS Office (Kingsoft Office Software). Data were first checked for normality using the Shapiro-Wilk and D'Agostino-Pearson tests, which confirmed that the KAP scores were non-normally distributed (P<.05). Accordingly, nonparametric statistical methods were used for comparisons. Given the categorical nature of sociodemographic variables (eg, gender, age, region, income, and education level) and the continuous nature of KAP scores, Kruskal-Wallis tests were used to compare multiple groups. For pairwise comparisons, the Dunn test with Benjamini-Hochberg correction was used to control the false discovery rate. In cases involving multiple hypothesis testing (eg, age groups with 7 categories), Holm correction was applied to maintain the conservativeness of the results. Statistical significance was set at P<.05 for all analyses.

For multivariate analysis, multiple regression models were used to examine the effects of sociodemographic factors on KAP scores. Interaction effects between variables (eg, age×gender and income×education) were explored to assess their impact on knowledge, attitude, and practice scores. Data visualization was performed using grouped bar charts, line graphs, and regression coefficient plots to provide an intuitive representation of differences across demographic groups.

Ethical Considerations

The study received ethical approval from the Medical Ethics Committee of the Affiliated Stomatological Hospital, China Medical University (2025002). Prior to participation, the questionnaire clearly outlined the study's objectives and content and assured participants that all data would be used solely for research purposes with strict confidentiality. Participants were informed of their right to withdraw from the study at any time without penalty or consequence. All data were collected with informed consent from participants, including additional parental consent for minors aged 12 - 17 years. All data collected in this study were anonymized and stored securely. No personally identifiable information was retained or analyzed. Participants did not receive any monetary or material compensation for their involvement in this research. No images of participants are included in the paper or supplementary materials. Therefore, there is no risk of participant identification. Otherwise, respondents were required to read an information statement embedded in the questionnaire title explaining the purpose, confidentiality, and voluntariness of the study. Informed consent was obtained digitally before submission. For respondents aged 12–17 years, a mandatory additional item was included to verify guardian consent before their data were considered valid.

Results

Basic Information

A total of 619 questionnaires were distributed. Among these, 562 were valid responses, with 242 (43.06%) from males and 320 (56.94%) from females. The average age of the participants was recorded as 41.27 (95% CI 37.4-45.1).

Knowledge

The average knowledge score of participants regarding periodontitis was 3.88 out of 8 (Table 1). Normality tests using the Shapiro-Wilk and D'Agostino-Pearson tests (both P<.05), combined with a skewness of 0.06 and kurtosis of −0.49 (Figure 1), indicated that the knowledge scores were primarily concentrated at the lower end, with fewer extreme low and high values. This distribution reflects generally poor awareness of periodontitis among the population.

XSL•FO
RenderX

Table . Characteristics of the participants and knowledge, attitude, and practice scores and proportion of gingival bleeding on periodontitis among participants in Northeast China (Dec 31, 2024–Jan 7, 2025).

Characteristics	Value (n)	Knowledge score		Attitude score		Practice score		Gingival bleeding rate (%)
		Mean (SD; 95% CI)	*P* value	Mean (SD; 95% CI)	*P* value	Mean (SD; 95% CI)	*P* value	
All	562	3.88 (1.55; 3.76-4.01)	<.01[b]	5.28 (1.80; 5.13-5.43)	<.01[b]	5.19 (2.32; 5.00-5.38)	.14	51.42
Sex			.03[c]		.23		.73	
Man	242	3.72 (1.44; 3.54-3.90)		5.19 (1.81; 4.97-5.42)		5.14 (2.26; 4.85-5.42)		53.31
Women	320	4.01 (1.62; 3.83-4.18)		5.34 (1.79; 5.14-5.54)		5.22 (2.37; 4.96-5.48)		50.00
Age (years)			.02[c]		.03[c]		.04[c]	
12-17	19	3.26 (1.91; 2.34-4.18)		3.68 (2.16; 2.64-4.73)		6.21 (2.35; 5.08-7.34)		57.89
18-25	89	4.19 (1.43; 3.89-4.49)		5.69 (1.25; 5.42-5.95)		4.67 (2.53; 4.14-5.21)		51.69
26-30	65	3.14 (1.29; 2.82-3.46)		4.26 (2.25; 3.7-4.82)		5.77 (2.04; 5.26-6.28)		50.77
31-40	167	3.78 (1.46; 3.56-4)		5.45 (1.54; 5.21-5.68)		5.49 (2.18; 5.15-5.82)		56.29
41-50	99	4.63 (1.58; 4.31-4.94)		6.08 (1.25; 5.83-6.33)		4.8 (2.38; 4.32-5.27)		46.46
51-60	53	3.66 (1.39; 3.28-4.04)		5.79 (1.67; 5.33-6.25)		4.72 (2.13; 4.13-5.31)		32.08
more than 60	70	3.73 (1.63; 3.34-4.12)		4.2 (2.03; 3.72-4.68)		5.2 (2.43; 4.62-5.78)		60.00
Region			.05		.02[c]		.03[c]	
Rural areas	259	4.36 (1.42; 4.19-4.53)		5.71 (1.78; 5.49-5.92)		4.98 (2.44; 4.68-5.28)		47.76
Small cities	119	3.76 (1.46; 3.49-4.02)		5.26 (1.42; 5-5.52)		5.97 (2.13; 5.58-6.35)		52.14
Mid-sized cities	117	3.17 (1.63; 2.87-3.47)		4.52 (1.79; 4.19-4.85)		5.14 (1.91; 4.79-5.49)		63.87
Metropolitan areas	67	3.52 (1.47; 3.16-3.88)		4.97 (2.01; 4.48-5.46)		4.67 (2.57; 4.05-5.3)		46.33
Education			.03[c]		.04[c]		.01[b]	
High school or below	143	3.88 (1.47; 3.64-4.12)		5.12 (1.76; 4.83-5.41)		5.25 (2.28; 4.88-5.63)		60.84
Associate	128	3.55 (1.61; 3.27-3.84)		5.55 (1.46; 5.3-5.81)		5.57 (2.21; 5.18-5.96)		44.53
Bachelor	217	4.07 (1.49; 3.87-4.27)		5.52 (1.81; 5.27-5.76)		5.06 (2.22; 4.77-5.36)		53.46
Master or above	74	3.92 (1.71; 3.52-4.32)		4.41 (2.09; 3.92-4.89)		4.74 (2.77; 4.10-5.39)		39.19
Annual income[a]			.04[c]		.01[b]		.05	
Low	62	4.31 (1.30; 3.98-4.64)		4.98 (1.49; 4.61-5.36)		4.45 (1.79; 4.00-4.91)		40.32
Low-middle	201	3.61 (1.65; 3.38-3.84)		5.35 (1.57; 5.14-5.57)		5.32 (2.43; 4.98-5.66		55.22
Middle	194	3.85 (1.33; 3.66-4.04)		5.47 (1.83; 5.22-5.73)		5.48 (2.28; 5.16-5.81)		55.67
High-middle	68	4.09 (1.58; 3.71-4.47)		5.21 (1.96; 4.73-5.68)		4.76 (2.18; 4.24-5.29)		51.47
High	37	4.49 (2.02; 3.81-5.16)		4.46 (2.63; 3.58-5.34)		4.89 (2.71; 3.99-5.79)		27.03

[a]Annual income five groups: less than 30,000 (US $4225), 30,000-80,000 (US $4225-US $11,268), 80,000-150,000 (US $11,268-US $21,127),

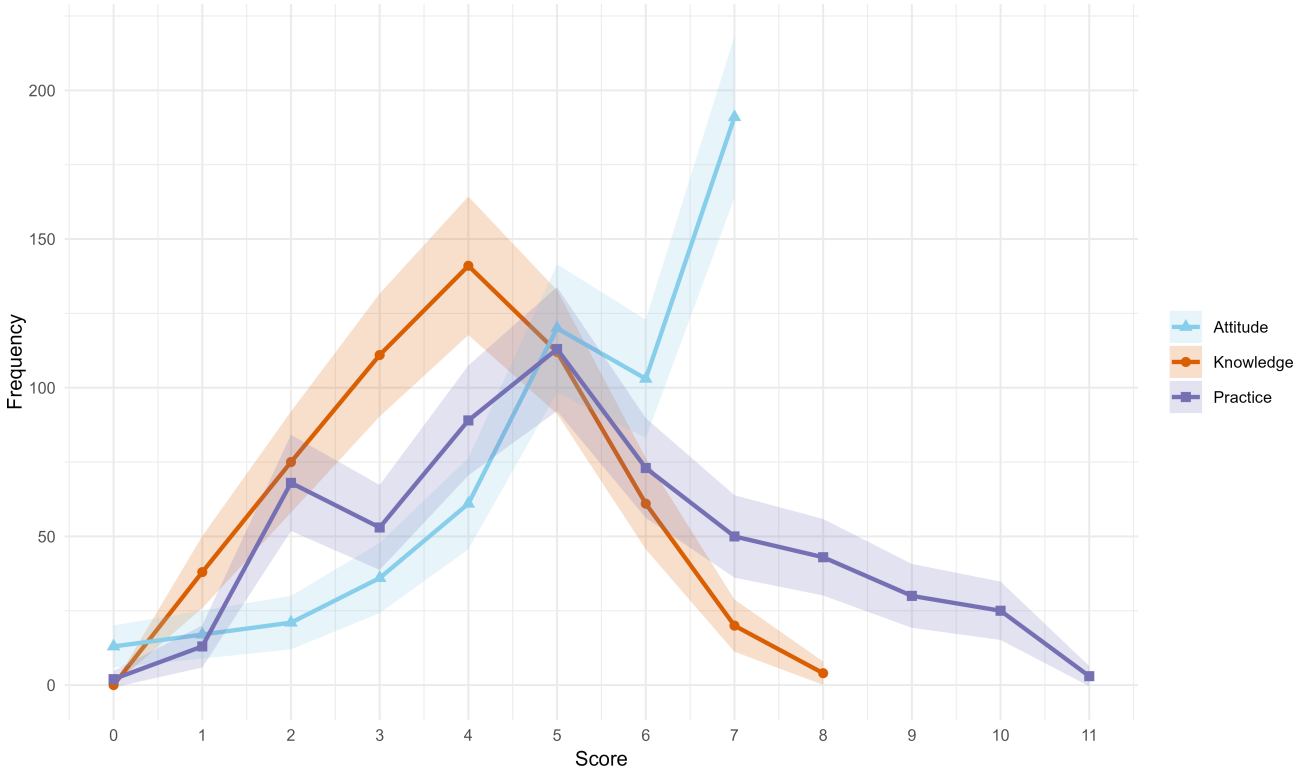
XSL•FO
RenderX

150,000-300,000 (US $21,127-US $42,254), and more than 300,000 (more than US $42,254).

[b]**P<0.01.

[c]*P<0.05.

Figure 1. Frequency distribution of scores across Knowledge, Attitude, and Practice categories on periodontitis (Northeast China, Dec 31, 2024 – Jan 7, 2025). This line chart illustrates the frequency distribution of scores for Knowledge (orange), Attitude (blue), and Practice (purple) regarding periodontitis prevention and awareness. The x-axis represents the score values while the y-axis indicates frequency, measured as the number of participants who achieved each score. The maximum possible scores vary by category: Knowledge (0 - 8), Attitude (0 - 7), and Practice (0 - 11). Shaded areas represent 95% CIs, reflecting variability in score distributions across the population. Statistical tests (Shapiro-Wilk and D'Agostino-Pearson) showed significant deviations from normality in all 3 domains (P<.01).

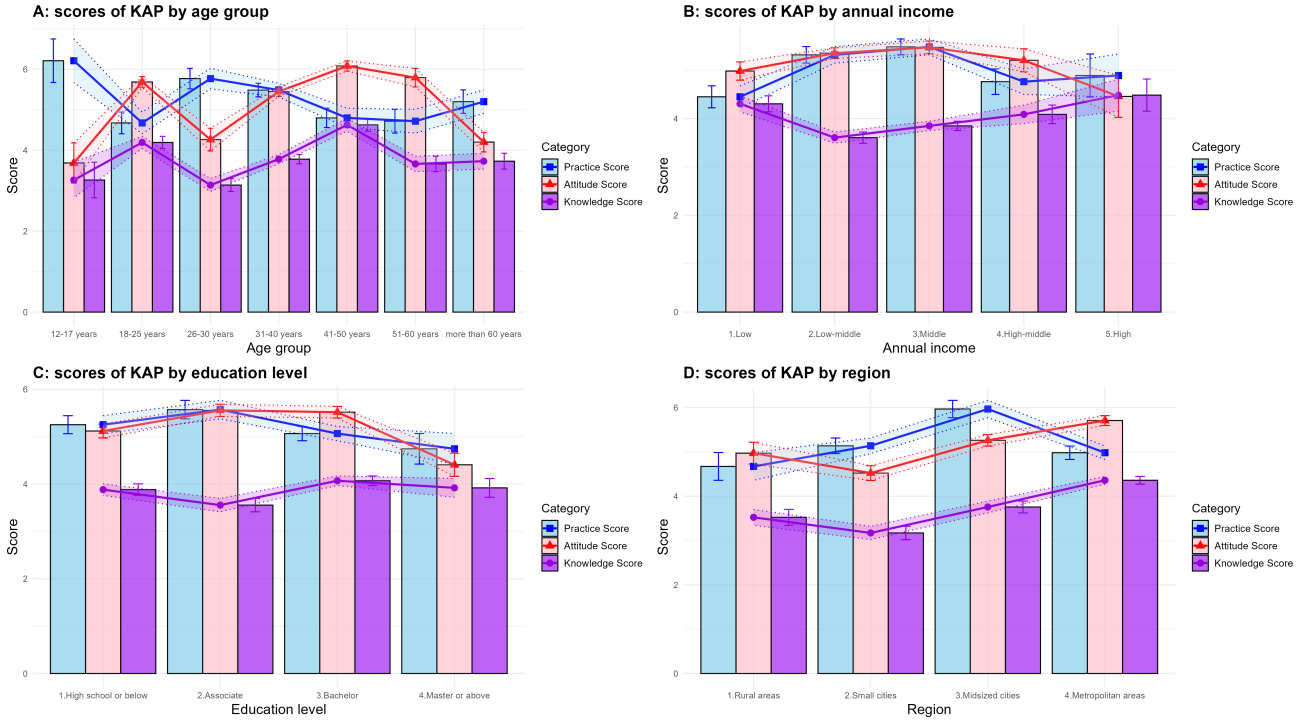


Variable	Shapiro.Wilk.p.value	D.Agostino.Pearson.p.value	Skewness	Kurtosis
Knowledge	<0.01	<0.01	0.061	-0.494
Attitude	<0.01	<0.01	-1.080	0.591
Practice	<0.01	<0.01	0.315	-0.524

In terms of demographic characteristics, gender, age, region, education level, and annual income all had significant effects on knowledge scores (P<.05). Group analysis revealed that females had a mean knowledge score of 4.01 (SD 1.62; 95% CI 3.83 - 4.18), significantly higher than males, whose mean score was 3.72 (SD 1.44; 95% CI 3.54 - 3.90). Age-specific analysis showed 2 peaks in knowledge scores: 4.19 (SD 1.43; 95% CI 3.89 - 4.49) in the age group of 18 - 25 years and 4.63 (SD 1.58; 95% CI 4.31 - 4.94) in the age group of 41 - 50 years. Participants with higher annual incomes demonstrated higher knowledge scores. However, unexpectedly high scores were also observed in low-income regions, which may be attributed to variations in regional development levels. Education level was positively correlated with knowledge scores (P<.05), with higher education levels associated with better scores (Table 1, Figure 2). However, the differences in scores between educational levels were relatively small, with a maximum intergroup difference of only 0.11.

XSL•FO
RenderX

Figure 2. Mean scores of Knowledge, Attitude, and Practice (KAP) on Periodontitis Across Demographic Groups in Northeast China (Dec 31, 2024 – Jan 7, 2025). This figure shows the mean scores of Knowledge (purple line), Attitude (red line), and Practice (blue line) related to periodontitis. Error bars represent SEs of the mean. Statistical significance of differences between groups can be inferred from nonoverlapping error bars.



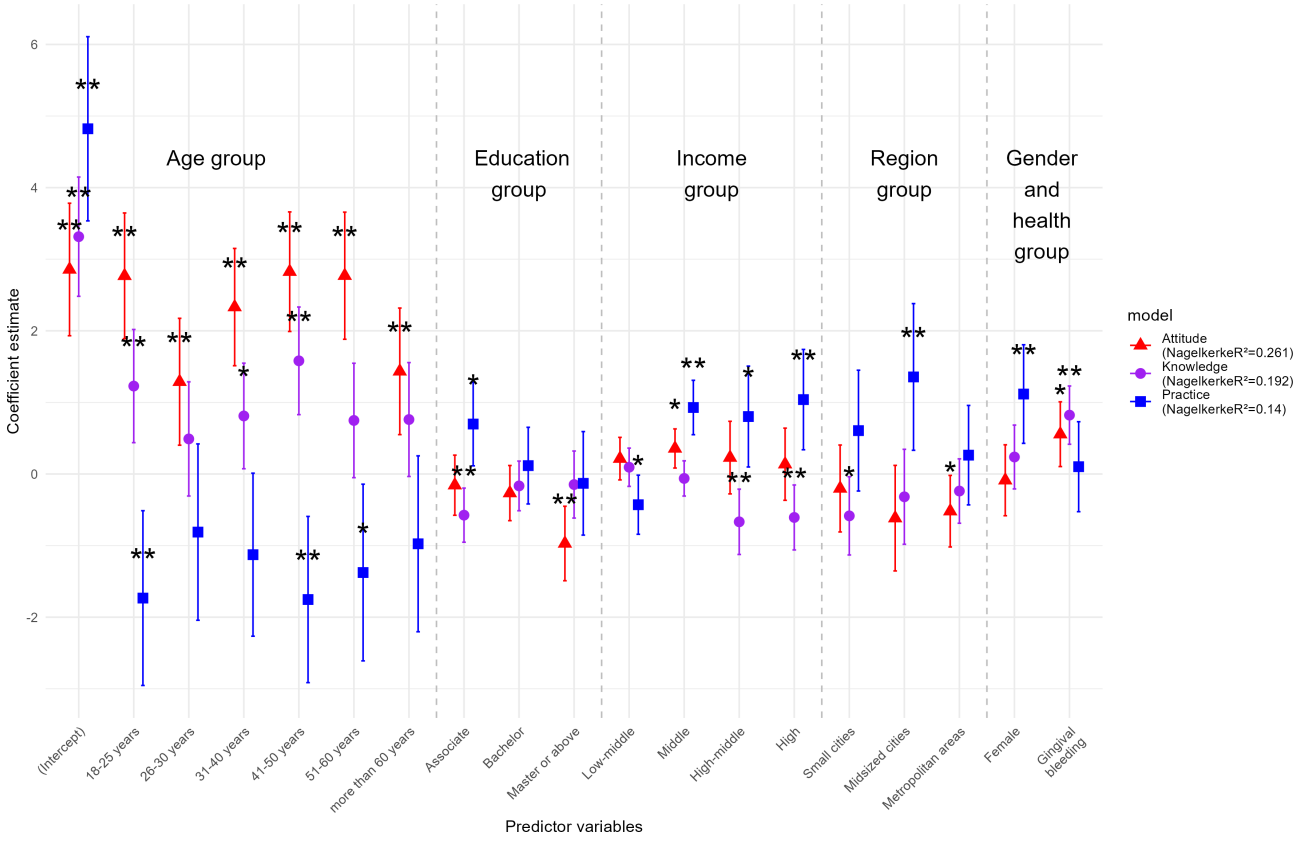
Attitude

The overall attitude scores were relatively high, with a mean score of 5.28 out of 7 (Table 1). The skewness was −1.08, and the kurtosis was 0.59 (Figure 1), indicating a generally positive attitude toward periodontitis among the population.

Apart from gender (P=.23), all other sociodemographic characteristics showed significant differences in attitude scores. Group analysis revealed that attitude scores across different age groups and regions followed a trend similar to the variations observed in knowledge scores. However, among income groups, a significant decline in attitude scores was observed in the high-income population. This decline may be related to the lower prevalence of gingival bleeding in the high-income group (27.03%) compared to other groups (40.32% - 55.67%). Further analysis demonstrated that participants with gingival bleeding had significantly higher attitude scores than those without gingival bleeding (P<.05). In addition, attitude scores across different educational levels were also influenced by the presence of gingival bleeding (Table 1, Figure 3).

XSL•FO
RenderX

Figure 3. Regression coefficients for the Knowledge, Attitude, and Practice scores across various predictor variables (unstandardized estimates). Error bars represent the 95% CIs. **P<.05 and **P<.01.



Practice

Practice scores showed a peak around 5 but displayed a long-tailed distribution on the right, indicating substantial variability in the data (Figure 1). Despite the right-skewed trend, the mean practice score was 5.19 out of 11 (Table 1). This suggests that, while the overall practice level was relatively low, a small proportion of individuals achieved higher scores.

Significant differences in practice scores were observed across all demographic groups except gender (*P*<.05). However, univariate analysis revealed a significant correlation between gender and practice scores, with females outperforming males (*P*<.01). Age group analysis indicated a peak in practice scores among participants aged 12 - 17 years, with a range of 5.08 - 7.34. Regionally, participants from medium-sized cities exhibited the highest practice scores, which may be linked to the higher prevalence of gingival bleeding in this group (63.87%) compared to others (46.33% - 52.14%). However, further analysis revealed that participants with gingival bleeding had no significantly higher practice scores than those without gingival bleeding (*P*>.05) (Table 1, Figure 3).

Nagelkerke Pseudo Coefficient of Determination

Multivariate regression analysis revealed that gender had no significant effect on KAP scores, while age was significantly associated with all dimensions of KAP. Knowledge and attitude scores were positively correlated with age, whereas practice scores showed a negative correlation. However, the influence of age diminished with increasing age, indicating that in older populations, age had minimal impact on scores. Specifically, knowledge scores significantly increased in all age groups compared to the age group of 12 - 17 years, except for the age group of 26 - 30 years. The age group of 41 - 50 years showed the highest increase (Nagelkerke R^2=1.58, 95% CI 0.83 - 2.33; *P*<.01). Attitude scores exhibited a trend similar to knowledge score,s while practice scores were negatively correlated with age, with lower significance observed in the age groups of 26 - 30 years and more than 60 years. Annual income had a negative correlation with knowledge scores, with high-income groups scoring significantly lower than low-income groups. No significant differences were observed in attitude scores among income groups. However, medium- and high-income groups exhibited a significant positive influence on practice scores. The effects of education level and region on KAP scores varied across specific populations (Figure 3).

A significant interaction between gender and age was observed in practice behaviors, with women more than 60 years scoring significantly higher than their male counterparts (interaction coefficient=3.70; *P*<.01), suggesting that older women tend to engage more actively in oral health practices. In contrast, the interaction between gender and age was not statistically significant for knowledge or attitude, although a slight upward trend in attitude was observed among women aged 31 - 40 years (interaction coefficient=1.50; *P*=.06). The interaction between educational attainment and annual income showed a predominantly negative association with attitude scores, particularly among higher-educated individuals within the same income bracket. Notably, individuals with high income and a bachelor's degree exhibited significantly higher attitude scores compared to other groups (interaction coefficient=3.13; *P*<.05).

XSL•FO
RenderX

No significant differences were observed in knowledge or practice dimensions.

Discussion

Principal Findings

Drawing on a convenience sample from Northeast China, this study uses descriptive statistics and multivariate regression analyses, including interaction terms, to examine the relationships between sociodemographic variables and periodontal KAP scores. The study's novelty lies in addressing the absence of KAP research on periodontitis in Northeast China and in uncovering nuanced sociodemographic influences through interaction regression analysis. By extending the analytical framework beyond descriptive statistics, this research offers generalizable methodologies and insights that can guide oral health policy development and educational interventions in other regions experiencing similar disparities. Furthermore, this study included adolescents aged 12 - 17 years and incorporated gingival bleeding as an early marker of periodontal disease, thereby enhancing the comprehensiveness of population-level assessment.

Most participants lacked sufficient awareness of periodontitis and effective measures to maintain periodontal health. This suggests that awareness and oral care practices related to periodontitis in Northeast China have not significantly improved over the past decade, indicating persistent inadequacies in public health policies related to periodontal health. The findings revealed poor overall knowledge and practices, consistent with previous studies. While the findings of this study bear some similarity to previous research [19,21], a key distinction lies in the inclusion of the age group 12 - 17 years for the first time. Results indicate that this age group had significantly higher practice scores and a higher detection rate of gingival bleeding; however, markedly lower scores in knowledge and attitude. This may be attributed to parental supervision, which promotes daily periodontal maintenance and thereby elevates practice scores in this group [22,23]. However, such externally imposed routines may not foster genuine improvements in knowledge or attitudes toward periodontal disease. Thus, motivational oral health education may be more effective in enhancing adolescents' understanding of periodontal disease.

However, it is encouraging that attitudes toward periodontitis have shown significant improvement, which may be attributed to several factors. First, there appears to be a shift in public perception of the disease. Recent research has highlighted the association between periodontitis and various systemic diseases [24], demonstrating that it is not merely an isolated oral condition. This connection may have heightened public concern, thereby improving attitudes toward the disease. Second, the influence of social media should not be overlooked. Social media platforms have played a key role in increasing public awareness of oral health [25]. Although much of the evidence supporting this comes from international studies, the impact of media on raising health awareness is undeniable. Finally, the rapid aging of China's population over the past 40 years [26] has also contributed to heightened awareness. The prevalence and severity of periodontitis increase with age [27], leading to higher dental consultation rates among older populations. In addition, interventions by health care professionals have had a positive effect on improving attitudes toward periodontal health [28].

In recent decades, China's economic growth [29] and significant improvements in education levels have likely contributed to improved attitudes toward periodontitis. This study found that increased income and education levels positively correlated with higher KAP scores regarding periodontitis. This finding aligns with previous research but places particular emphasis on older male populations [30]. In addition, studies have shown that respondents in lower income categories are approximately 3 times more likely to report poor self-reported oral health compared to those in the highest income categories. Similarly, individuals with less than a high school education are about 1.5 times more likely to report poor self-reported oral health compared to those with higher education levels [31].

Age had a notable impact on KAP scores, with stronger correlations observed in younger populations. Among older individuals, the relationship was less pronounced, likely due to cognitive decline associated with aging, increased difficulty in acquiring new knowledge, and limited access to periodontal health information and resources [32]. Interestingly, gender demonstrated an age-dependent effect on KAP scores. While gender alone did not show a significant impact on KAP, females consistently outperformed males in practice. Furthermore, this trend rose with increasing age. Gender differences are often overlooked in clinical practice [33], underscoring the need to prioritize gender-based considerations in the clinical diagnosis and treatment of periodontitis. Oral health education campaigns should also address gender differences to deliver more targeted and effective interventions.

This study also found that individuals with gingival bleeding exhibited slightly better overall KAP scores, consistent with previous research [20]. Gingival bleeding is an early symptom of gingivitis [34]. Although self-reported gingival bleeding on brushing has low sensitivity for diagnosing periodontitis, it serves as an important sentinel sign of gingival inflammation and is visually detectable even with minimal blood loss [35]. Consequently, timely dental visits following the detection of gingival bleeding, coupled with health education and professional intervention, are associated with improved KAP scores [28]. However, disappointingly, a cross-sectional study on periodontal health KAP among Chinese adults reported that only 6.4% of individuals with gingival bleeding sought professional dental care [19]. Most people choose not to seek professional help even after noticing symptoms, which may be attributed to low oral health awareness, dental anxiety, or other related factors [36].

Studies have shown that oral health education and promotional interventions can significantly increase patients' willingness to seek dental care and positively influence their attitudes and practices [37]. Community-based health education programs have been identified as effective strategies to raise public awareness of oral health and improve periodontal KAP [38,39], especially in enhancing the oral health knowledge of children and mothers [39]. Therefore, community-based oral health

XSL•FO
RenderX

prevention institutions should strengthen efforts to promote oral health education and training, teaching the public how to respond appropriately to gingival bleeding and ultimately increasing dental visit rates.

Limitations

This study has several limitations. First, there was some bias in the representativeness of the sample, as sample sizes were not evenly distributed across all groups. However, significance testing was performed for all categories to ensure the statistical reliability of the data. Second, all collected data were self-reported by participants, which may have introduced potential biases, including social desirability and recall biases. As a result, scores for attitudes and practices might have been affected. Third, the use of convenience sampling likely introduced selection bias, since participants were recruited based on availability and willingness rather than through a fully randomized process. Fourth, participants' voluntary enrollment in the survey created a self-selection bias: those who chose to respond may differ systematically in their periodontal knowledge, attitudes, or practices from those who did not participate. Nevertheless, the large sample size allowed the study to capture specific characteristics and distributions of periodontal KAP in Northeast China.

Conclusion

In conclusion, this study used a large, community-representative sample to assess KAP concerning periodontitis among populations in Northeast China. The KAP concerning periodontitis in Northeast China has improved to some extent but remains relatively low overall. Furthermore, regional, income, educational, and gender differences in KAP regarding periodontitis demonstrate complex interactions, highlighting the need for additional evaluative factors in future research. Gingival bleeding significantly influences KAP and underscores the critical role of oral health care professionals in enhancing public awareness of periodontitis through clinical education programs.

Therefore, policy interventions should prioritize community-based oral health education, particularly in economically disadvantaged and rural regions. Such programs should focus on high-risk populations and leverage early indicators, such as gingival bleeding, to prompt timely professional consultation. Efforts to enhance oral hygiene practices and motivation among older men may also contribute to improving the population's overall KAP regarding oral health. Addressing attitude-related misalignments among highly educated individuals is also a pressing concern. Moreover, policies aimed at minors should emphasize motivational rather than mandatory oral health education, empowering parents to actively support children's understanding and adoption of periodontal care practices. Integrating oral health education into school curricula and workplace wellness programs could further reinforce preventive behaviors from an early stage. Future research should explore underlying determinants to more effectively advance population-level oral health literacy and behavior.

Acknowledgments

We express our gratitude to our teachers and participants for their valuable contributions to this teaching and research endeavor.

Data Availability

The datasets analyzed in the current study are available from the corresponding author on reasonable request.

Authors' Contributions

CJ, CL, ZZ, WW, YZ designed the research. MC, YL, YZ, WW collected the data and verified the accuracy of the data. CJ, MC, CL, YL, YZ contributed to data interpretation. CJ, MC performed the statistical analysis and visualization. CJ, MC wrote the manuscript. CJ, MC, YL prepared the figures. CL prepared the table. CJ and MC contributed equally to this work.

Conflicts of Interest

None declared.

References

1. Frencken JE, Sharma P, Stenhouse L, Green D, Laverty D, Dietrich T. Global epidemiology of dental caries and severe periodontitis - a comprehensive review. J Clin Periodontol 2017 Mar;44(Suppl 18):S94-S105. [doi: 10.1111/jcpe.12677] [Medline: 28266116]
2. Kalhan AC, Wong ML, Allen F, Gao X. Periodontal disease and systemic health: an update for medical practitioners. Ann Acad Med Singap 2022 Sep;51(9):567-574. [doi: 10.47102/annals-acadmedsg.2021503] [Medline: 36189701]
3. Kwon T, Lamster IB, Levin L. Current concepts in the management of periodontitis. Int Dent J 2021 Dec;71(6):462-476. [doi: 10.1111/idj.12630] [Medline: 34839889]
4. Hajishengallis G, Chavakis T. Local and systemic mechanisms linking periodontal disease and inflammatory comorbidities. Nat Rev Immunol 2021 Jul;21(7):426-440. [doi: 10.1038/s41577-020-00488-6] [Medline: 33510490]

XSL•FO
RenderX

5. Nascimento GG, Alves-Costa S, Romandini M. Burden of severe periodontitis and edentulism in 2021, with projections up to 2050: The Global Burden of Disease 2021 study. J Periodontal Res 2024 Oct;59(5):823-867. [doi: 10.1111/jre.13337] [Medline: 39192495]
6. GBD 2021 Oral Disorders Collaborators. Trends in the global, regional, and national burden of oral conditions from 1990 to 2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2025 Mar;405(10482):897-910. [doi: 10.1016/s0140-6736(24)02811-3]
7. Trindade D, Carvalho R, Machado V, Chambrone L, Mendes JJ, Botelho J. Prevalence of periodontitis in dentate people between 2011 and 2020: a systematic review and meta-analysis of epidemiological studies. J Clin Periodontol 2023 May;50(5):604-626. [doi: 10.1111/jcpe.13769] [Medline: 36631982]
8. Chen X, Ye W, Zhan JY, et al. Periodontal Status of Chinese Adolescents: Findings from the 4th National Oral Health Survey. Chin J Dent Res 2018;21(3):195-203. [doi: 10.3290/j.cjdr.a41075] [Medline: 30255170]
9. Jiao J, Jing W, Si Y, et al. The prevalence and severity of periodontal disease in Mainland China: Data from the Fourth National Oral Health Survey (2015-2016). J Clin Periodontol 2021 Feb;48(2):168-179. [doi: 10.1111/jcpe.13396] [Medline: 33103285]
10. Jin L. Group E. Initiator paper. Interprofessional education and multidisciplinary teamwork for prevention and effective management of periodontal disease. J Int Acad Periodontol 2015 Jan;17(1 Suppl):74-79. [Medline: 25764596]
11. Tonetti MS, Jepsen S, Jin L, Otomo-Corgel J. Impact of the global burden of periodontal diseases on health, nutrition and wellbeing of mankind: a call for global action. J Clin Periodontol 2017 May;44(5):456-462. [doi: 10.1111/jcpe.12732] [Medline: 28419559]
12. Di Stefano M, Polizzi A, Santonocito S, Romano A, Lombardi T, Isola G. Impact of oral microbiome in periodontal health and periodontitis: a critical review on prevention and treatment. Int J Mol Sci 2022 May 5;23(9):5142. [doi: 10.3390/ijms23095142] [Medline: 35563531]
13. Kumar S. Evidence-based update on diagnosis and management of gingivitis and periodontitis. Dent Clin North Am 2019 Jan;63(1):69-81. [doi: 10.1016/j.cden.2018.08.005] [Medline: 30447793]
14. Scannapieco FA, Gershovich E. The prevention of periodontal disease-An overview. Periodontol 2000 2020 Oct;84(1):9-13. [doi: 10.1111/prd.12330] [Medline: 32844421]
15. Varela-Centelles P, Diz-Iglesias P, Estany-Gestal A, Seoane-Romero JM, Bugarín-González R, Seoane J. Periodontitis awareness amongst general public: a critical systematic review to identify gaps of knowledge. J Periodontol 2016 Apr;87(4):403-415. [doi: 10.1902/jop.2015.150458] [Medline: 26545044]
16. He Q, Wang G, He J, et al. Knowledge, attitude and practice regarding occupational protection against COVID-19 among midwives in China: a nationwide cross-sectional study. Int J Disaster Risk Reduct 2022 Sep;79:103184. [doi: 10.1016/j.ijdrr.2022.103184] [Medline: 35859908]
17. Andrade C, Menon V, Ameen S, Kumar Praharaj S. Designing and conducting knowledge, attitude, and practice surveys in psychiatry: practical guidance. Indian J Psychol Med 2020 Sep;42(5):478-481. [doi: 10.1177/0253717620946111] [Medline: 33414597]
18. Zarei F, Dehghani A, Ratansiri A, et al. ChecKAP: A Checklist for Reporting a Knowledge, Attitude, and Practice (KAP) Study. Asian Pac J Cancer Prev 2024 Jul 1;25(7):2573-2577. [doi: 10.31557/APJCP.2024.25.7.2573] [Medline: 39068593]
19. Zhao Q, Wang SB, Xu G, et al. Periodontal health: a national cross-sectional study of knowledge, attitudes and practices for the public oral health strategy in China. J Clin Periodontol 2019 Apr;46(4):406-419. [doi: 10.1111/jcpe.13082] [Medline: 30768801]
20. Sun J, Tong D, Sun C, et al. Knowledge, attitude, and practice toward self-control of dental plaque among patients with periodontal diseases: a cross-sectional study. BMC Oral Health 2023 Sep 2;23(1):628. [doi: 10.1186/s12903-023-03352-w]
21. Wang QT, Wu ZF, Wu YF, Shu R, Pan YP, Xia JL. Epidemiology and preventive direction of periodontology in China. J Clin Periodontol 2007 Nov;34(11):946-951. [doi: 10.1111/j.1600-051X.2007.01139.x] [Medline: 17877744]
22. Pakpour AH, Hidarnia A, Hajizadeh E, Kumar S, Fridlund B. Why Iranian adolescents do not brush their teeth: a qualitative study. Int J Dent Hyg 2012 May;10(2):86-90. [doi: 10.1111/j.1601-5037.2011.00513.x] [Medline: 21777380]
23. Elkhodary HM, Abdelnabi MH, Swelem AA, et al. Individual, familial and country-level factors associated with oral hygiene practices in children: an international survey. BMC Oral Health 2023 Jan 30;23(1):50. [doi: 10.1186/s12903-023-02746-0] [Medline: 36710323]
24. Botelho J, Mascarenhas P, Viana J, et al. An umbrella review of the evidence linking oral health and systemic noncommunicable diseases. Nat Commun 2022 Dec 9;13(1):7614. [doi: 10.1038/s41467-022-35337-8] [Medline: 36494387]
25. Jayasinghe YA, Kanmodi KK, Jayasinghe RM, Jayasinghe RD. Assessment of patterns and related factors in using social media platforms to access health and oral health information among Sri Lankan adults, with special emphasis on promoting oral health awareness. BMC Public Health 2024 Jun 1;24(1):1472. [doi: 10.1186/s12889-024-19008-5] [Medline: 38824505]
26. Chen X, Giles J, Yao Y, et al. The path to healthy ageing in China: a Peking University–Lancet Commission. Lancet 2022 Dec 3;400(10367):1967-2006. [doi: 10.1016/S0140-6736(22)01546-X] [Medline: 36423650]
27. Demetriou N, Parashis A, Tsami-Pandi A. Relationship between age and clinical symptoms of periodontal disease. Stomatologia (Athenai) 1990;47(4):231-241. [Medline: 2131686]

XSL•FO
RenderX

28. Tahani B, Asgari I, Golkar S, Ghorani A, Hasan Zadeh Tehrani N, Arezoo Moghadam F. Effectiveness of an integrated model of oral health-promoting schools in improving children's knowledge and the KAP of their parents, Iran. BMC Oral Health 2022 Dec 12;22(1):599. [doi: 10.1186/s12903-022-02644-x] [Medline: 36510207]
29. Guo Y. Evolution and stages of China's economic inequality from 1978 to 2018. PLoS ONE 2023;18(7):e0288873. [doi: 10.1371/journal.pone.0288873]
30. Hernández-Palacios RD, Ramírez-Amador V, Jarillo-Soto EC, Irigoyen-Camacho ME, Mendoza-Núñez VM. Relationship between gender, income and education and self-perceived oral health among elderly Mexicans. An exploratory study. Cien Saude Colet 2015 Apr;20(4):997-1004. [doi: 10.1590/1413-81232015204.00702014] [Medline: 25923612]
31. Farmer J, Phillips RC, Singhal S, Quiñonez C. Inequalities in oral health: understanding the contributions of education and income. Can J Public Health 2017 Sep 14;108(3):e240-e245. [doi: 10.17269/CJPH.108.5929] [Medline: 28910244]
32. Jongsiriyanyong S, Limpawattana P. Mild cognitive impairment in clinical practice: a review article. Am J Alzheimers Dis Other Demen 2018 Dec;33(8):500-507. [doi: 10.1177/1533317518791401] [Medline: 30068225]
33. Michelson C, Al-Abedalla K, Wagner J, Swede H, Bernstein E, Ioannidou E. Lack of attention to sex and gender in periodontitis-related randomized clinical trials: a meta-research study. J Clin Periodontol 2022 Dec;49(12):1320-1333. [doi: 10.1111/jcpe.13707] [Medline: 35924761]
34. Bhagat S, Jaiswal P, Kotecha SV. Clinical features, etiopathogenesis, and therapeutic approaches of acute gingival lesions: a narrative review. Cureus 2024 Aug;16(8):e67129. [doi: 10.7759/cureus.67129] [Medline: 39290940]
35. Deng K, Pelekos G, Jin L, Tonetti MS. Gingival bleeding on brushing as a sentinel sign of gingival inflammation: a diagnostic accuracy trial for the discrimination of periodontal health and disease. J Clin Periodontol 2021 Dec;48(12):1537-1548 [FREE Full text] [doi: 10.1111/jcpe.13545]
36. Aburas S, Pfaffeneder-Mantai F, Hofmann A, Meller O, Schneider B, Turhani D. Dentophobia and dental treatment: an umbrella review of the published literature. Spec Care Dentist 2023 Mar;43(2):163-173. [doi: 10.1111/scd.12749] [Medline: 35700448]
37. Ghaffari M, Rakhshanderou S, Ramezankhani A, Noroozi M, Armoon B. Oral Health Education and Promotion Programmes: Meta-Analysis of 17-Year Intervention. Int J Dent Hyg 2018 Feb;16(1):59-67. [doi: 10.1111/idh.12304] [Medline: 28836347]
38. Nakre PD, Harikiran AG. Effectiveness of oral health education programs: A systematic review. J Int Soc Prev Community Dent 2013 Jul;3(2):103-115. [doi: 10.4103/2231-0762.127810] [Medline: 24778989]
39. Nghayo HA, Palanyandi CE, Ramphoma KJ, Maart R. Oral health community engagement programs for rural communities: A scoping review. PLOS ONE 2024;19(2):e0297546. [doi: 10.1371/journal.pone.0297546] [Medline: 38319914]

Abbreviations

KAP: Knowledge, Attitudes, and Practices

Edited by A Mavragani, T Sanchez; submitted 03.02.25; peer-reviewed by MB Amin, S Hansun; revised version received 04.06.25; accepted 12.06.25; published 22.08.25.

Please cite as:
Jin C, Chi M, Li C, Li Y, Wang W, Zhang Y, Liu Q
Knowledge, Attitudes, and Practices to Periodontal Health of the Northeast Chinese Public: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e72069
URL: https://publichealth.jmir.org/2025/1/e72069
doi:10.2196/72069

© Chengde Jin, Ming Chi, Chunxiao Li, Yushan Li, Wanting Wang, Yushi Zhang, Qiao Liu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 22.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Gender-Specific Moderating Role of Physical Fitness and Cardiovascular Risk Factors in the Relationship Between BMI and C-Reactive Protein: Cross-Sectional Study

Chunyuan Lai[1], BEd; Yunqing Liu[2*], MM; Yue Yan[3*], BEd; Boli Cheng[3], BEd; Yakun Liu[3], BEd; Fanshu Yu[3], BEd; Chengyuan Hu[1], BEd; Liqiang Su[1*], PhD; Fanghui Li[3*], PhD

[1]School of Physical Education, Jiangxi Normal University, Nanchang, China

[2]Changzhou Sports Hospital, Affiliated Sports Hospital of Nanjing Normal University, Changzhou, China

[3]School of Sport Sciences, Nanjing Normal University, Xianlin Campus, Xianlin Street, Qixia District, Nanjing, Jiangsu Province, China

[*]these authors contributed equally

Corresponding Author:
Fanghui Li, PhD
School of Sport Sciences, Nanjing Normal University, Xianlin Campus, Xianlin Street, Qixia District, Nanjing, Jiangsu Province, China

Abstract

Background: Physical fitness and cardiovascular risk factors may affect C-reactive protein (CRP) levels by influencing BMI. However, the extent to which these factors influence CRP through BMI remains unclear.

Objective: This study aims to explore how indicators related to physical fitness and cardiovascular risk factors affect the relationship between BMI and CRP levels.

Methods: This study selected 453 participants from 518 individuals at a sports hospital in Jiangsu Province, including 231 (51%) males and 222 (49%) females. Using interaction effect analyses, CRP served as the dependent variable and BMI as the independent variable. Physical fitness and cardiovascular risk factors (body morphology, blood pressure, blood lipids, and blood glucose) were categorized into quartiles from the lowest (Q1) to the highest (Q4) and treated as effect modifiers for interaction effect analyses. Sex-specific analyses were conducted to explore the influence of BMI on CRP across different genders, based on physical fitness and cardiovascular risk factor-related indicators.

Results: BMI was positively associated with CRP (β=0.203, 95% CI 0.163 - 0.243; *P*<.001). After adjusting for confounding factors, this association remained significant (β=0.206, 95% CI 0.164 - 0.248; *P*<.001). In the stratified gender analysis, interaction effect analyses showed that in females, waist circumference (median 110.5, IQR 93 - 128; β=0.342, 95% CI 0.232 - 0.452; *P*<.001), hip circumference (median 120, IQR 104 - 136; β=0.361, 95% CI 0.260 - 0.462; *P*<.001), waist-hip ratio (median 1.08, IQR 0.9 - 1.25; β=0.291, 95% CI 0.204 - 0.378; *P*<.001), diastolic blood pressure (median 101, IQR 85 - 117; β=0.344, 95% CI 0.252 - 0.436; *P*<.001), triglycerides (median 1.41, IQR 1.13 - 1.68; β=0.313, 95% CI 0.216 - 0.409; *P*<.001), triglyceride-rich lipoprotein cholesterol (median 0.64, IQR 0.51 - 0.76; β=0.365, 95% CI 0.249 - 0.481; *P*<.001), 20 - 49 years one-leg standing time with closed eyes (median 3, IQR 1 - 5; β=0.371, 95% CI 0.276 - 0.466; *P*<.001), and 20 - 49 years grip strength (median 6.5, IQR 6 - 7; β=0.369, 95% CI 0.270 - 0.467; *P*<.001) significantly influenced the effect of BMI on CRP. In males, etotal cholesterol, glucose, one-leg standing time with closed eyes, sit-and-reach, and grip strength demonstrated trends in the effect of BMI on CRP (*P*>.05).

Conclusions: This study demonstrates a positive correlation between BMI and CRP with gender-specific characteristics. In females, multiple body morphology, physiological, and physical fitness indicators significantly influence the effect of BMI on CRP, while some indicators in males also exhibit trends. These findings suggest the necessity of developing health management strategies tailored to different genders, particularly for female populations. Such strategies should consider body morphology indicators such as waist circumference and hip circumference, physiological indicators including blood pressure and blood lipids, and physical fitness metrics such as one-leg standing time with closed eyes. This comprehensive approach can better regulate the impact of BMI on CRP and promote overall health.

(JMIR Public Health Surveill 2025;11:e76485) doi:10.2196/76485

KEYWORDS

BMI; C-reactive protein; physical fitness; prevention of cardiovascular diseases; gender differences

XSL•FO
RenderX

Introduction

Cardiovascular disease (CVD) remains the leading cause of death globally, with the Global Burden of Disease Study revealing 612 million cases of CVD in 2021, accounting for 26.8% of all deaths worldwide [1]. Since 1990, the prevalence of unhealthy lifestyles among Chinese residents, combined with a large population at risk for CVD and accelerated population aging, has made CVD one of the most significant threats to public health in China. In the composition of disease-related deaths among urban and rural residents in China, CVD ranks first, accounting for 48.98% and 47.35% of deaths in rural and urban areas, respectively, in 2021. Among every 5 deaths, 2 are attributed to CVD [2]. Research indicates that obesity is a major risk factor for global severe morbidity and all-cause mortality, significantly increasing the risk of CVDs such as hypertension, type 2 diabetes, dyslipidemia, coronary heart disease, and myocardial infarction [3].

Currently, BMI is recognized as the standard metric for evaluating obesity and is closely associated with inflammation and the occurrence of multiple CVDs [4,5]. However, in recent years, the use of BMI as an indicator for assessing obesity and CVD has been questioned [6]. Some studies have pointed out that BMI overlooks important factors affecting health status, such as age, gender, and ethnicity, and thus cannot comprehensively measure an individual's health level [7,8]. The American Heart Association stated in a scientific statement that BMI has limitations in predicting individual body fat [9]. Research has found that using BMI in conjunction with lipid metabolism indicators can more accurately predict type 2 diabetes [10]. Additionally, one study indicated that the waist-hip ratio (WHR) is more effective than BMI in predicting mortality [11]. However, other studies argue that BMI is a direct factor in the onset of various diseases and is an independent risk factor for inflammation [12,13].

Although BMI has a certain effect in evaluating inflammation, it is influenced by various factors such as physical fitness and cardiovascular risk factors. Studies indicate a subtle association between adult body morphology, fat distribution, and inflammatory manifestations. Some research has elucidated that inflammation is linked to lipid levels, specifically demonstrating that elevated plasma C-reactive protein (CRP) levels are often accompanied by increased concentrations of total cholesterol (TC), triglycerides (TG), low-density lipoprotein cholesterol (LDL-C), and triglyceride-rich lipoprotein cholesterol (TRL-C), as well as decreased concentrations of high-density lipoprotein cholesterol (HDL-C) [14,15]. Notably, inflammation has been identified as a significant driver of insulin resistance, metabolic syndrome, and type 2 diabetes, suggesting an inseparable relationship between inflammation and blood glucose levels [16]. Additionally, low levels of physical fitness are associated with elevated plasma CRP levels, while reduced inflammation levels are closely related to improved physical health, which may be linked to decreased BMI [17]. Based on the above evidence, this study considers physical fitness and cardiovascular risk factors as key factors, with body morphology, blood pressure, blood lipids, and blood glucose as indicators of cardiovascular risk factors.

In early research, the critical variable of gender was often overlooked, primarily due to the substantial underrepresentation of female participants in both preclinical and clinical studies [18,19]. The National Institutes of Health implemented a policy requiring all National Institutes of Health–funded research projects to include both male and female participants [18]. This policy has driven in-depth exploration of gender differences across various research fields, and obesity research is no exception. Currently, significant evidence demonstrates that the pathogenesis, risk factors, and clinical implications of obesity all exhibit substantial gender-specific differences [12,20,21]. Therefore, systematically investigating gender-specific outcomes in research is fully justified by scientific evidence and necessity.

Therefore, this study aims to investigate the gender differences in the relationship between BMI and CRP under the influence of physical fitness and multiple cardiovascular risk factors, while exploring more comprehensive evaluation methods and intervention strategies. The findings hold significant practical implications for the prevention and management of inflammation induced by obesity and its associated risks, providing a scientific basis for the widespread implementation of public health education and the precise prevention and treatment of CVDs.

Methods

Participants and Study Design

This cross-sectional study was conducted at a sports hospital in Changzhou City, Jiangsu Province, and 518 participants aged 20 years and older were recruited using convenience sampling. Data were collected between December 2022 and June 2023, and all participants underwent a systematic physical examination by a health care professional and signed an informed consent form. The examination included a physical examination, blood pressure and blood biochemical tests (including fasting blood glucose, blood lipids, and CRP). Based on the exclusion criteria, the following participants were excluded: age less than 20 years (n=15), long-term medication use (n=14), and lack of health screening data (n=36). After these exclusion criteria, the final effective sample size was 453.

Inclusion and Exclusion Criteria

Since the proportion of individuals under 20 years old with CVD is relatively low, and there are many physiological and behavioral variables in minors that are difficult to control in the study, this study selected individuals aged 20 years and above as the study population [22-24].

Inclusion criteria are as follows: (1) age ≥20 years; (2) no acute inflammatory diseases; (3) presence of at least one cardiovascular risk factor (eg, hypertension, hyperlipidemia, hyperglycemia, smoking history, and obesity); (4) no psychological or consciousness disorders; and (5) complete baseline information, body morphology, blood pressure, blood lipids, blood glucose, and physical fitness–related health examination data.

Exclusion criteria are as follows: (1) individuals with mental disorders or severe intellectual disabilities; (2) pregnant or lactating females; (3) patients with a history of major cardiovascular events (eg, heart failure, myocardial infarction,

XSL•FO
RenderX

unstable angina, and stroke); (4) diseases or treatments that may affect body composition and inflammatory status, such as corticosteroid use, advanced renal failure, bariatric surgery, severe fluid retention syndrome, and cancer; and (5) missing health examination data.

Ethical Considerations

This study has obtained approval from the Biomedical Research Ethics Committee of Nanjing Normal University (NNU202310007). All participants provided informed consent prior to their involvement in the study, and they were explicitly informed of their right to withdraw at any point without any negative consequences. All data were coded to ensure the confidentiality of participant information, and additional measures, such as secure data storage and restricted access, were implemented to further protect privacy. No compensation was provided to the participants, as their involvement was entirely voluntary and based on their willingness to contribute to the research.

Observation Indicators

The outcome measures in this study include CRP, body morphology, blood pressure, blood lipids, blood glucose, and physical fitness [25-27]. In 1999, Yudkin et al [28] proposed that obesity induces a systemic, persistent low-level chronic inflammatory state mediated by various inflammatory factors, based on the acute inflammation symptoms of "redness, swelling, heat, and pain." Inflammation is generally assessed using inflammatory markers such as CRP and tumor necrosis factor-alpha [28]. Moreover, studies have shown that elevated CRP levels are an indicator of inflammation, a major risk factor for CVD [29-31]. Therefore, in this study, CRP was used as the inflammatory marker.

In addition, studies indicate a significant positive correlation between body morphology and CVD mortality, highlighting the importance of body morphology in CVD prevention [32,33]. Blood pressure, blood lipids, blood glucose, and inflammation are recognized as major CVD risk factors, and abnormalities in these indicators directly increase the risk of cardiovascular events [34]. Strong evidence demonstrates that improving physical fitness not only reduces the incidence of CVD events but also extends life span, underscoring the critical role of physical fitness in CVD prevention [35]. Additionally, recent research has found that cholesterol in TRL-C, as a product of their hydrolysis, is closely associated with diabetes and CVD [36]. Based on a comprehensive consideration of extensive research evidence and clinical practice, this study incorporates body morphology, blood pressure, blood lipids, blood glucose, and physical fitness as covariates to validate their roles in the effect of BMI on CRP.

Body morphology includes height, weight, waist circumference (WC), hip circumference (HC), WHR, and BMI. Blood pressure includes systolic blood pressure (SBP) and diastolic blood pressure (DBP). Blood lipids include TC, TG, TRL-C, LDL-C, and HDL-C. Blood glucose includes glucose (GLU) and glycated hemoglobin (HbA_{1c}). Physical fitness includes one-leg standing time with closed eyes, sit-and-reach, and grip strength (see Multimedia Appendix 1 for physical fitness test methodology [37,38]), assessed based on the National Physical Fitness Test Standards Manual and Standards (Adults) and the National Physical Fitness Test Standards Manual (Elderly), published by the General Administration of Sport of China in 2023 [37]. The calculation method for TRL-C is as follows: when TG<4.5 mmol/L, LDL-C is calculated using the Friedewald formula (LDL-C = TC – HDL-C – TG/2.2); when TG≥4.5 mmol/L, LDL-C is directly measured; subsequently, the TRL-C concentration is calculated using the formula (TRL-C = TC – HDL-C – LDL-C) proposed by Varbo et al [36]. The indicators were divided into quartiles from lowest (Q1) to highest (Q4).

Statistical Analysis

Data were processed and analyzed using R version 4.3.1 (The R Foundation for Statistical Computing) and SPSS version 26.0 (IBM Corp). For the basic characteristics of the study participants, participants were grouped and described by gender. Differences among groups in body morphology indicators, blood pressure, blood lipids, blood glucose, physical fitness, and CRP were compared. Continuous variables that follow a normal distribution are expressed as mean (SD), and group comparisons are performed using independent samples *t* tests. Continuous variables that do not follow a normal distribution are expressed as median (IQR), and group comparisons are performed using the Kruskal-Wallis test. If the Kruskal-Wallis test reveals significant differences, post hoc analysis is subsequently conducted to identify specific groups between which significant differences exist.

To verify whether BMI is an independent risk factor for inflammation, linear regression models with adjusted variables were used. In these models, CRP was set as the dependent variable, and BMI as the independent variable. Age, gender, smoking, and alcohol consumption were included as covariates for adjustment. The effects of BMI on CRP were observed both before and after adjustment to determine if BMI independently influences CRP. To address the issue of multiple comparisons, the Bonferroni correction method is applied in the interaction tests, adjusting the significance level to reduce the risk of false positives.

To assess the interaction effects of body morphology indicators, blood pressure, blood lipids, blood glucose, and physical fitness on the relationship between BMI and CRP, CRP was used as the dependent variable and BMI was set as the independent variable. The aforementioned covariates were divided into quartiles from lowest (Q1) to highest (Q4) to serve as interaction effect modifiers, and age, gender, smoking, and alcohol consumption were included as covariates for adjustment. This analysis aimed to determine whether there are interaction effects between the covariates divided into quartiles and BMI on CRP.

The study was stratified by gender into male and female groups for separate interaction effect analyses. Within each gender group, CRP was set as the dependent variable and BMI as the independent variable. The selected covariates were divided into quartiles from lowest (Q1) to highest (Q4) and used as interaction effect modifiers to analyze whether there are interaction effects between the covariates divided into quartiles and BMI on CRP in different genders, and age, smoking, and

XSL•FO
RenderX

alcohol consumption were included as covariates for adjustment. Statistical significance was defined as $P<.05$, with $P<.001$ indicating highly significant.

Results

Description of Basic Characteristics of the Study Population

As shown in Table 1 and Figure 1, a total of 518 participants were initially enrolled in the study and screened according to the inclusion and exclusion criteria, of which 453 participants including 231 males and 222 females were finally enrolled in this study. The mean age of females was 42.1 (SD 10.6) years, while that of males was 44.2 (SD 10.7) years. In the comparison between males and females, all other indicators showed statistically significant differences ($P<.05$) except for TC, HbA_{1c}, and CRP.

XSL•FO
RenderX

Table . Basic characterization of the study population by gender grouping.

Indicator	Male (n=231)	Female (n=222)	*t* test (*df*)	*P* value
Age (years), mean (SD)	44.2 (10.7)	42.1 (10.6)	–2.06 (451)	.04
Height (cm), mean (SD)	172.8 (6.4)	160.6 (5.7)	–21.4 (451)	<.001
Weight (kg), mean (SD)	83.2 (14.8)	66.1 (13.5)	–12.9 (451)	<.001
WC[a] (cm), mean (SD)	96.6 (10.4)	85.7 (11.9)	–10.4 (451)	<.001
HC[b] (cm), mean (SD)	103.2 (7.8)	99.0 (9.1)	–5.22 (451)	<.001
WHR[c] (cm), mean (SD)	0.93 (0.05)	0.86 (0.07)	–11.9 (451)	<.001
SBP[d] (mm Hg), mean (SD)	134 (14)	123 (17)	–7.38 (451)	<.001
DBP[e] (mm Hg), mean (SD)	86 (11)	78 (11)	–7.67 (451)	<.001
LDL-C[f] (mmol/L), mean (SD)	1.58 (0.75)	1.39 (0.74)	–2.73 (451)	.007
HDL-C[g] (mmol/L), mean (SD)	1.21 (0.30)	1.42 (0.35)	6.93 (451)	<.001
TG[h] (mmol/L), median (IQR)	2.06 (1.26)	1.40 (1.1)	–6.01 (451)	<.001
TC[i] (mmol/L), median (IQR)	5.02 (0.96)	4.88 (0.94)	–1.59 (451)	.11
TRL-C[j] (mmol/L), mean (SD)	0.94 (0.61)	0.62 (0.48)	–6.19 (451)	<.001
HbA_{1c}[k] (%), median (IQR)	5.30 (0.70)	5.18 (0.64)	–1.87 (451)	.06
GLU[l] (mmol/L), median (IQR)	5.95 (1.6)	5.53 (1.4)	–3.07 (451)	.002
One-leg standing time with closed eyes (seconds), mean (SD)	15.6 (19.1)	20.0 (22.5)	2.24 (451)	.03
Sit-and-reach (cm), mean (SD)	–2.39 (9.7)	3.54 (9.3)	6.66 (451)	<.001
Grip strength (kg), mean (SD)	40.8 (7.8)	23.8 (5.0)	–27.5 (451)	<.001
BMI (kg/m^2), mean (SD)	27.76 (4.1)	25.6 (4.7)	–5.28 (451)	<.001
CRP[m] (mg/L), mean (SD)	2.72 (2.3)	2.40 (2.1)	–1.57 (451)	.12

[a]WC: waist circumference.

[b]HC: hip circumference.

[c]WHR: waist-hip ratio.

[d]SBP: systolic blood pressure.

[e]DBP: diastolic blood pressure.

[f]LDL-C: low-density lipoprotein cholesterol.

[g]HDL-C: high-density lipoprotein cholesterol.

[h]TG: triglycerides.

[i]TC: total cholesterol.

[j]TRL-C: triglyceride-rich lipoprotein cholesterol.

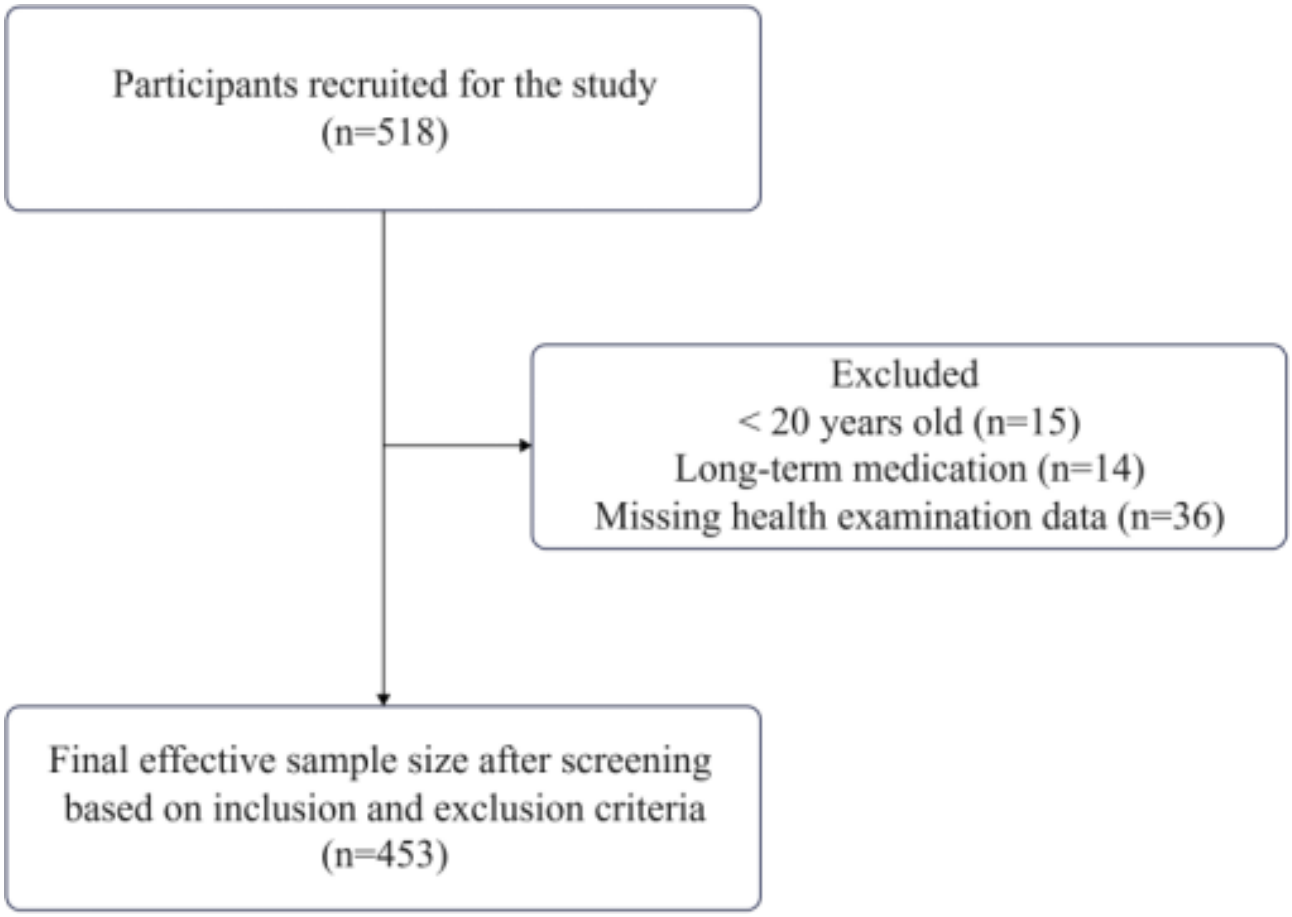
[k]HbA_{1c}: glycated hemoglobin.

[l]GLU: glucose.

[m]CRP: C-reactive protein.

XSL•FO
RenderX

Figure 1. Participant flowchart.



Relationship Between BMI and CRP

A linear regression analysis was conducted with CRP as the dependent variable and BMI as the independent variable. The results indicated that in the unadjusted model, each unit increase in BMI was associated with an increase in CRP concentration of 0.203 mg/L, and this effect was statistically significant (β=0.203, 95% CI 0.163 - 0.243; *P*<.001). Similarly, after adjusting for age, gender, smoking, and alcohol consumption as confounding variables, the effect of BMI on CRP slightly increased to 0.206 mg/L, meaning that each unit increase in BMI was associated with an increase in CRP concentration of 0.206 mg/L, and this effect remained statistically significant (β=0.206, 95% CI 0.164 - 0.248; *P*<.001; Table 2).

Table . Linear regression relationship between BMI and CRP[a,b].

	Model 1[c]		Model 2[d]	
Variant	β (95% CI)	*P* value	β (95% CI)	*P* value
BMI	0.203 (0.163-0.243)	<.001	0.206 (0.164-0.248)	<.001

[a]CRP: C-reactive protein.

[b]CRP as the dependent variable and BMI as the independent variable.

[c]Model 1: unadjusted model.

[d]Model 2: adjusted for age, gender, smoking, and alcohol consumption.

The Role of Cardiovascular Risk Factors in the Effect of BMI on CRP

Using CRP as the dependent variable and BMI as the independent variable, the adjustment for age, gender, smoking, and alcohol consumption had little effect on the β coefficient (Table 3).

XSL•FO
RenderX

Table . Interaction analysis of covariates (n=453).

Indicator	n (%)	Model 1[a]		*P* value	Model 2[b]		*P* value
		β (95% CI)	*P* value		β (95% CI)	*P* value	
WC[c] (cm)				.08			.13
Q1	120 (26.5)	0.131 (–0.017 to 0.279)	.08		0.138 (–0.017 to 0.294)	.08	
Q2	115 (25.4)	0.162 (0.006 to 0.318)	.04		0.168 (0.003 to 0.333)	.047	
Q3	105 (23.2)	0.083 (–0.068 to 0.233)	.28		0.090 (–0.066 to 0.247)	.26	
Q4	113 (24.9)	0.285 (0.192 to 0.378)	<.001		0.285 (0.190 to 0.380)	<.001	
HC[d] (cm)				.15			.32
Q1	131 (28.9)	0.161 (0.024 to 0.298)	.02		0.189 (0.044 to 0.333)	.011	
Q2	102 (22.5)	0.124 (–0.030 to 0.278)	.12		0.144 (–0.018 to 0.307)	.082	
Q3	112 (24.7)	0.199 (0.061 to 0.337)	.005		0.213 (0.072 to 0.355)	.003	
Q4	108 (23.8)	0.296 (0.209 to 0.384)	<.001		0.297 (0.209 to 0.385)	<.001	
WHR[e] (cm)				.37			.37
Q1	115 (25.4)	0.123 (0.024 to 0.222)	.02		0.120 (0.019 to 0.221)	.02	
Q2	112 (24.7)	0.167 (0.048 to 0.285)	.006		0.162 (0.041 to 0.283)	.009	
Q3	112 (24.7)	0.171 (0.075 to 0.266)	<.001		0.161 (0.062 to 0.261)	.002	
Q4	114 (25.2)	0.230 (0.157 to 0.304)	<.001		0.225 (0.150 to 0.301)	<.001	
SBP[f] (mm Hg)				.16			.2
Q1	122 (26.9)	0.094 (–0.010 to 0.198)	.08		0.103 (–0.004 to 0.209)	.059	
Q2	107 (23.6)	0.235 (0.140 to 0.329)	<.001		0.238 (0.141 to 0.334)	<.001	
Q3	116 (25.6)	0.227 (0.153 to 0.301)	<.001		0.228 (0.153 to 0.303)	<.001	
Q4	108 (23.8)	0.211 (0.134 to 0.289)	<.001		0.214 (0.134 to 0.295)	<.001	
DBP[g] (mm Hg)				.31			.22
Q1	124 (27.4)	0.171 (0.090 to 0.252)	<.001		0.170 (0.088 to 0.251)	<.001	
Q2	108 (23.8)	0.149 (0.074 to 0.225)	.001		0.144 (0.067 to 0.221)	<.001	
Q3	112 (24.7)	0.236 (0.148 to 0.324)	<.001		0.248 (0.159 to 0.338)	<.001	
Q4	109 (24.1)	0.238 (0.152 to 0.324)	<.001		0.234 (0.147 to 0.322)	<.001	
LDL-C[h] (mmol/L)				.9			.9

XSL•FO
RenderX

Indicator	n (%)	Model 1[a]		*P* value	Model 2[b]		*P* value
		β (95% CI)	*P* value		β (95% CI)	*P* value	
Q1	114 (25.2)	0.184 (0.113 to 0.256)	<.001		0.187 (0.114 to 0.261)	<.001	
Q2	114 (25.2)	0.220 (0.137 to 0.303)	<.001		0.224 (0.140 to 0.308)	<.001	
Q3	111 (24.5)	0.219 (0.136 to 0.301)	<.001		0.223 (0.139 to 0.306)	<.001	
Q4	114 (25.2)	0.201 (0.112 to 0.290)	<.001		0.206 (0.116 to 0.296)	<.001	
HDL-C[i] (mmol/L)				.52			.54
Q1	115 (25.4)	0.225 (0.133 to 0.318)	<.001		0.219 (0.125 to 0.313)	<.001	
Q2	114 (25.2)	0.237 (0.151 to 0.323)	<.001		0.241 (0.155 to 0.328)	<.001	
Q3	113 (24.9)	0.155 (0.070 to 0.239)	<.001		0.156 (0.071 to 0.242)	<.001	
Q4	111 (24.5)	0.187 (0.114 to 0.259)	<.001		0.192 (0.117 to 0.267)	<.001	
TG[j] (mmol/L)				.21			.28
Q1	115 (25.4)	0.124 (0.035 to 0.213)	.006		0.132 (0.042 to 0.223)	.004	
Q2	112 (24.7)	0.247 (0.165 to 0.328)	<.001		0.248 (0.164 to 0.331)	<.001	
Q3	113 (24.9)	0.223 (0.143 to 0.304)	<.001		0.222 (0.139 to 0.304)	<.001	
Q4	113 (24.9)	0.196 (0.115 to 0.278)	<.001		0.194 (0.112 to 0.277)	<.001	
TC[k] (mmol/L)				.50			.47
Q1	114 (25.2)	0.158 (0.071 to 0.244)	<.001		0.162 (0.074 to 0.249)	<.001	
Q2	114 (25.2)	0.233 (0.142 to 0.324)	<.001		0.243 (0.151 to 0.335)	<.001	
Q3	113 (24.9)	0.187 (0.104 to 0.270)	<.001		0.186 (0.102 to 0.270)	<.001	
Q4	112 (24.7)	0.234 (0.164 to 0.303)	<.001		0.235 (0.163 to 0.307)	<.001	
TRL-C[l] (mmol/L)				.18			.23
Q1	113 (24.9)	0.119 (0.030 to 0.208)	.009		0.128 (0.037 to 0.218)	.006	
Q2	114 (25.2)	0.249 (0.168 to 0.330)	<.001		0.250 (0.167 to 0.333)	<.001	
Q3	118 (26.0)	0.219 (0.134 to 0.303)	<.001		0.219 (0.133 to 0.306)	<.001	
Q4	108 (23.8)	0.203 (0.124 to 0.282)	<.001		0.199 (0.118 to 0.279)	<.001	
HbA_{1c}[m] (%)				.04			.06
Q1	120 (26.5)	0.130 (0.044 to 0.215)	.003		0.133 (0.045 to 0.221)	.003	

XSL•FO
RenderX

Indicator	n (%)	Model 1[a]		*P* value	Model 2[b]		*P* value
		β (95% CI)	*P* value		β (95% CI)	*P* value	
Q2	132 (29.1)	0.129 (0.053 to 0.206)	.001		0.119 (0.040 to 0.198)	.003	
Q3	104 (23.0)	0.256 (0.160 to 0.352)	<.001		0.244 (0.147 to 0.341)	<.001	
Q4	97 (21.4)	0.245 (0.173 to 0.317)	<.001		0.235 (0.162 to 0.309)	<.001	
GLU[n] (mmol/L)				.13			.11
Q1	115 (25.4)	0.148 (0.059 to 0.237)	.001		0.141 (0.049 to 0.233)	.003	
Q2	112 (24.7)	0.212 (0.125 to 0.300)	<.001		0.217 (0.126 to 0.307)	<.001	
Q3	112 (24.7)	0.137 (0.053 to 0.220)	.001		0.129 (0.043 to 0.214)	.004	
Q4	114 (25.2)	0.252 (0.181 to 0.322)	<.001		0.246 (0.174 to 0.317)	<.001	

[a]Model 1: unadjusted model.

[b]Model 2: adjusted for age, gender, smoking, and alcohol consumption.

[c]WC: waist circumference.

[d]HC: hip circumference.

[e]WHR: waist-hip ratio.

[f]SBP: systolic blood pressure.

[g]DBP: diastolic blood pressure.

[h]LDL-C: low-density lipoprotein cholesterol.

[i]HDL-C: high-density lipoprotein cholesterol.

[j]TG: triglycerides.

[k]TC: total cholesterol.

[l]TRL-C: triglyceride-rich lipoprotein cholesterol.

[m]HbA_{1c}: glycated hemoglobin.

[n]GLU: glucose.

In females, WC (median 110.5, IQR 93 - 128; β=0.342, 95% CI 0.232 - 0.452; *P*<.001), HC (median 120, IQR 104 - 136; β=0.361, 95% CI 0.260 - 0.462; *P*<.001), WHR (median 1.08, IQR 0.9 - 1.25; β=0.291, 95% CI 0.204 - 0.378; *P*<.001), DBP (median 101, IQR 85 - 117; β=0.344, 95% CI 0.252 - 0.436; *P*<.001), TG (median 1.41, IQR 1.13 - 1.68; β=0.313, 95% CI 0.216 - 0.409; *P*<.001), TRL-C (median 0.64, IQR 0.51 - 0.76; β=0.365, 95% CI 0.249 - 0.481; *P*<.001), 20 - 49 years one-leg standing time with closed eyes (median 3, IQR 1 - 5; β=0.371, 95% CI 0.276 - 0.466; *P*<.001), and 20 - 49 years grip strength (median 6.5, IQR 6 - 7; β=0.369, 95% CI 0.270 - 0.467; *P*<.001) significantly influenced the effect of BMI on CRP. Additionally, in males, TC, GLU, one-leg standing time with closed eyes, sit-and-reach, and grip strength demonstrated trends in influencing the effect of BMI on CRP (*P*>.05; Table 4). For interaction plots, refer to Figure S1 in Multimedia Appendix 2.

XSL•FO
RenderX

Table . Interaction analysis of covariates by gender. Model: adjusted for age, smoking, and alcohol consumption.

Indicator	Male (n=231)				Female (n=222)			
	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value
WC[a] (cm)				.86				.002
Q1	81.5 (73-90)	63 (27)	0.141 (–0.116 to 0.399); .28		69 (60-78)	57 (26)	–0.004 (–0.228 to 0.220); .97	
Q2	93 (90-96)	60 (26)	0.137 (–0.184 to 0.459); .4		80.5 (78-83)	55 (25)	0.113 (–0.110 to 0.337); .32	
Q3	99 (96-102)	50 (22)	0.168 (–0.191 to 0.528); .36		88 (83-93)	55 (25)	0.018 (–0.155 to 0.190); .84	
Q4	118 (102-134)	58 (25)	0.248 (0.091 to 0.404); .002		110.5 (93-128)	55 (24)	0.342 (0.232 to 0.452);<.001	
HC[b] (cm)				.75				.003
Q1	93.5 (89-98)	66 (29)	0.326 (0.058 to 0.593); .02		82.8 (71.5 - 94)	65 (30)	0.011 (–0.158 to 0.181); .9	
Q2	100 (98-102)	50 (21)	0.092 (–0.248 to 0.431); .6		96 (94-98)	52 (23)	0.132 (–0.079 to 0.342); .22	
Q3	104.5 (102-107)	60 (26)	0.215 (–0.043 to 0.473); .1		101 (98-104)	52 (23)	0.296 (0.132 to 0.461); .001	
Q4	122 (107-137)	55 (24)	0.238 (0.087 to 0.389); .002		120 (104-136)	53 (24)	0.361 (0.260 to 0.462);<.001	
WHR[c] (cm)				.58				.045
Q1	0.85 (0.79 - 0.9)	59 (25)	0.011 (–0.184 to 0.206); .91		0.75 (0.68 - 0.82)	56 (25)	0.164 (0.057 to 0.271); .003	
Q2	0.92 (0.9 - 0.93)	57 (25)	0.122 (–0.072 to 0.315); .22		0.84 (0.82 - 0.86)	55 (25)	0.103 (–0.048 to 0.255); .18	
Q3	0.95 (0.93 - 0.97)	57 (25)	0.165 (–0.022 to 0.353); .09		0.88 (0.86 - 0.9)	55 (25)	0.130 (0.016 to 0.244); .03	
Q4	1.05 (0.97 - 1.12)	58 (25)	0.160 (0.035 to 0.286); .01		1.08 (0.9 - 1.25)	56 (25)	0.291 (0.204 to 0.378);<.001	
SBP[d] (mm Hg)				.82				.02
Q1	109.5 (95-124)	58 (25)	0.109 (–0.082 to 0.300); .27		99.5 (87-112)	64 (29)	0.060 (–0.064 to 0.185); .34	
Q2	129 (124-134)	60 (26)	0.170 (0.018 to 0.321); .03		116 (112-120)	47 (21)	0.311 (0.192 to 0.430);<.001	
Q3	138.5 (134-143)	61 (26)	0.175 (0.064 to 0.287); .002		126 (120-132)	55 (25)	0.275 (0.170 to 0.379);<.001	
Q4	159 (143-175)	52 (23)	0.093 (–0.077 to 0.263); .28		153 (132-174)	56 (25)	0.264 (0.182 to 0.346);<.001	
DBP[e] (mm Hg)				.08				.008
Q1	67.5 (57-78)	61 (26)	0.139 (0.014 to 0.263); .03		60 (50-70)	63 (28)	0.187 (0.083 to 0.291); .001	

XSL•FO
RenderX

Indicator	Male (n=231)				Female (n=222)			
	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value
Q2	81.5 (78-85)	59 (26)	0.144 (0.018 to 0.270); .03		73 (70-76)	49 (22)	0.121 (0.024 to 0.219); .02	
Q3	88.5 (85-92)	56 (24)	0.258 (0.118 to 0.398);<.001		81 (76-85)	56 (25)	0.264 (0.152 to 0.377); <.001	
Q4	108.5 (92-125)	55 (24)	–0.027 (–0.199 to 0.145); .76		101 (85-117)	54 (25)	0.344 (0.252 to 0.436);<.001	
LDL-C[f] (mmol/L)				.77				.8
Q1	0.66 (0.34 - 0.97)	58 (25)	0.126 (0.009 to 0.244); .04		0.55 (0.26 - 0.84)	56 (25)	0.247 (0.155 to 0.340); <.001	
Q2	1.17 (0.97 - 1.36)	58 (25)	0.221 (0.074 to 0.368); .004		1 (0.84 - 1.16)	56 (25)	0.196 (0.101 to 0.290); <.001	
Q3	1.77 (1.36 - 2.18)	57 (25)	0.187 (0.043 to 0.330); .01		1.51 (1.16 - 1.86)	54 (25)	0.257 (0.157 to 0.356); <.001	
Q4	5.6 (2.18 - 3.41)	58 (25)	0.163 (0.012 to 0.314); .04		2.71 (1.86 - 3.56)	56 (25)	0.247 (0.137 to 0.357); <.001	
HDL-C[g] (mmol/L)				.27				.26
Q1	0.76 (0.52 - 1)	58 (25)	0.022 (–0.146 to 0.190); .8		0.93 (0.68 - 1.18)	57 (26)	0.312 (0.209 to 0.416); <.001	
Q2	1.11 (1 - 1.21)	60 (26)	0.194 (0.051 to 0.337); .008		1.29 (1.18 - 1.40)	54 (24)	0.242 (0.139 to 0.346); <.001	
Q3	1.31 (1.21 - 1.41)	58 (25)	0.121 (–0.027 to 0.269); .11		1.52 (1.40 - 1.64)	55 (25)	0.192 (0.093 to 0.291); .002	
Q4	1.78 (1.41 - 2.14)	55 (24)	0.203 (0.096 to 0.309);<.001		2.29 (1.64 - 2.94)	56 (25)	0.160 (0.016 to 0.303); .03	
TG[h] (mmol/L)				.27				.01
Q1	0.84 (0.39 - 1.28)	58 (25)	0.158 (0.015 to 0.302); .03		0.49 (0.23 - 0.74)	57 (26)	0.065 (–0.052 to 0.182); .28	
Q2	1.51 (1.28 - 1.74)	58 (25)	0.249 (0.102 to 0.396); .001		0.94 (0.74 - 1.13)	54 (24)	0.242 (0.140 to 0.345); <.001	
Q3	2.1 (1.74 - 2.45)	58 (25)	0.053 (–0.086 to 0.191); .46		1.41 (1.13 - 1.68)	55 (25)	0.313 (0.216 to 0.409); <.001	
Q4	3.88 (2.45 - 5.31)	57 (25)	0.165 (0.041 to 0.289); .01		3.47 (1.68 - 5.26)	56 (25)	0.256 (0.152 to 0.360); <.001	
TC[i] (mmol/L)				.26				.8
Q1	3.31 (2.21 - 4.41)	58 (25)	0.063 (–0.082 to 0.207); .39		3.28 (2.31 - 425)	56 (25)	0.260 (0.154 to 0.367); <.001	

XSL•FO
RenderX

Indicator	Male (n=231)				Female (n=222)			
	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value
Q2	4.72 (4.41 - 5.02)	58 (25)	0.113 (–0.067 to 0.292); .22		4.53 (4,25 - 4.81)	56 (25)	0.266 (0.162 to 0.371); <.001	
Q3	5.32 (5.02 - 5.61)	57 (25)	0.161 (0.015 to 0.306); .03		5.18 (4.81 - 5.54)	56 (25)	0.204 (0.108 to 0.300); <.001	
Q4	7.06 (5.61 - 8.51)	58 (25)	0.236 (0.124 to 0.347); <.001		6.87 (5.54 - 8.19)	54 (25)	0.231 (0.134 to 0.328); <.001	
TRL-C[j] (mmol/L)				.29				.002
Q1	0.38 (0.18 - 0.58)	58 (25)	0.159 (0.016 to 0.303); .03		0.22 (0.1 - 0.34)	55 (25)	0.044 (–0.073 to 0.161); .46	
Q2	0.69 (0.58 - 0.79)	58 (25)	0.250 (0.103 to 0.397); .001		0.43 (0.34 - 0.51)	56 (25)	0.255 (0.155 to 0.354); <.001	
Q3	0.95 (0.79 - 1.11)	63 (27)	0.064 (–0.065 to 0.193); .33		0.64 (0.51 - 0.76)	55 (25)	0.365 (0.249 to 0.481); <.001	
Q4	1.73 (1.11 - 2.34)	52 (23)	0.165 (0.033 to 0.298); .015		1.43 (0.76 - 2.1)	56 (25)	0.242 (0.151 to 0.334); <.001	
HbA_{1c}[k] (%)				.5				.06
Q1	4.6 (4.3 - 4.9)	62 (27)	0.071 (–0.067 to 0.209); .31		4.4 (4 - 4.8)	58 (26)	0.174 (0.048 to 0.300); .007	
Q2	5.1 (4.9 - 5.2)	72 (31)	0.116 (–0.013 to 0.245); .08		5 (4.8 - 5.1)	60 (27)	0.106 (0.007 to 0.205); .04	
Q3	5.4 (5.2 - 5.5)	48 (21)	0.207 (0.044 to 0.371); .01		5.3 (5.1 - 5.4)	56 (25)	0.272 (0.159 to 0.385);<.001	
Q4	7.9 (5.5 - 10.3)	49 (21)	0.185 (0.054 to 0.315); .006		6.8 (5.4 - 8.1)	48 (22)	0.257 (0.175 to 0.339);<.001	
GLU[l] (mmol/L)				.12				.35
Q1	4.65 (3.99 - 5.3)	59 (25)	0.042 (–0.101 to 0.185); .57		4.32 (3.82 - 4.81)	56 (25)	0.174 (0.043 to 0.305); .01	
Q2	5.5 (5.3 - 5.7)	57 (25)	0.198 (0.056 to 0.339); .007		5.11 (4.81 - 5.41)	55 (25)	0.314 (0.181 to 0.447);<.001	
Q3	5.95 (5.7 - 6.2)	57 (25)	0.069 (–0.081 to 0.219); .37		5.63 (5.41 - 5.84)	55 (25)	0.182 (0.085 to 0.280);<.001	
Q4	9.35 (6.2 - 12.5)	58 (25)	0.229 (0.109 to 0.349);<.001		8.92 (5.84 - 12)	56 (25)	0.237 (0.154 to 0.321);<.001	
One-leg standing time with closed eyes (score)								

XSL•FO
RenderX

Indicator	Male (n=231)				Female (n=222)			
	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value
20 - 49 y (male: n=159; female: n=162)				.37				.03
Q1	3 (1-5)	62 (27)	0.194 (0.075 to 0.314); .002		3 (1-5)	50 (23)	0.371 (0.276 to 0.466);<.001	
Q2	5.25 (5 - 5.5)	32 (14)	0.162 (–0.021 to 0.344); .08		5.65 (5 - 6.3)	31 (14)	0.252 (0.128 to 0.376); .001	
Q3	6.75 (5.5 - 8)	38 (16)	0.002 (–0.185 to 0.189); .98		7.15 (6.3 - 8)	49 (22)	0.187 (0.073 to 0.302); .002	
Q4	9 (8-10)	27 (12)	0.138 (–0.079 to 0.355); .22		9 (8-10)	32 (14)	0.167 (0.047 to 0.288); .007	
50 - 59 y (male: n=50; female: n=47)				.37				.55
Q1	3 (1-5)	20 (9)	0.332 (–0.012 to 0.676); .07		3 (1-5)	13 (6)	–0.093 (–0.357 to 0.172); .5	
Q2	5.63 (5 - 6.25)	7 (3)	0.152 (–0.260 to 0.564); .47		5.75 (5 - 6.5)	11 (5)	–0.092 (–0.401 to 0.216); .56	
Q3	7.38 (6.25 - 8.5)	12 (5)	0.188 (–0.366 to 0.741); .51		7.65 (6.5 - 8.8)	11 (5)	0.043 (–0.393 to 0.478); .85	
Q4	9.25 (8.5 - 10)	11 (5)	–0.104 (–0.552 to 0.344); .65		9.4 (8.8 - 10)	12 (5)	0.132 (–0.142 to 0.407); .35	
60 - 79 y (male: n=22; female: n=13)				.06				.2
Q1	3 (1-5)	7 (3.)	0.606 (0.111 to 1.101); .04		3 (1-5)	5 (2)	0.063 (–0.756 to 0.881); .89	
Q2	5.88 (5 - 6.75)	3 (1)	0.632 (–0.615 to 1.880); .34		5.5 (5-6)	3 (1)	0.744 (–0.336 to 1.824); .24	
Q3	7.13 (6.75 - 7.5)	9 (4)	0.127 (–0.240 to 0.494); .51		7 (6-8)	1 (1)	0.123 (–0.398 to 0.643); .66	
Q4	8.75 (7.5 - 10)	3 (1)	–0.034 (–0.840 to 0.771); .94		8.8 (8 - 9.5)	4 (2)	–0.005 (–0.687 to 0.676); .99	
Sit-and-reach (score)								
20 - 49 y (male: n=159; female: n=162)				.17				.2
Q1	2 (1-3)	83 (36)	0.126 (0.018 to 0.234); .02		3 (1-5)	66 (30)	0.303 (0.225 to 0.381);<.001	
Q2	4 (3-5)	18 (8)	0.174 (–0.119 to 0.467); .25		5.25 (5 - 5.5)	20 (9)	0.259 (0.125 to 0.394);<.001	

XSL•FO
RenderX

Indicator	Male (n=231)				Female (n=222)			
	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value
Q3	5.75 (5 - 6.5)	26 (11)	0.289 (0.123 to 0.455); .001		6.25 (5.5 - 7)	34 (15)	0.245 (0.110 to 0.379);<.001	
Q4	8.25 (6.5 - 10)	32 (14)	–0.018 (–0.255 to 0.218); .88		8.5 (7-10)	42 (19)	0.142 (0.014 to 0.271); .03	
50 - 59 y (male: n=50; female: n=47)				.56				.86
Q1	4.5 (1.5 - 7.5)	17 (7)	0.247 (–0.053 to 0.547); .11		4.5 (1.5 - 7.5)	13 (6)	–0.058 (–0.309 to 0.193); .66	
Q2	8.45 (7.5 - 9.4)	13 (5)	0.085 (–0.354 to 0.524); .71		8.65 (7.5 - 9.8)	14 (6)	–0.040 (–0.308 to 0.227); .77	
Q3	10.7 (9.4 - 12)	11 (5)	0.501 (0.038 to 0.965); .04		10.7 (9.8 - 11.6)	11 (5)	0.086 (–0.227 to 0.398); .59	
Q4	13.1 (12 - 14.2)	9 (4)	0.158 (–0.619 to 0.935); .69		13.3 (11.6 - 15)	9 (4)	0.069 (–0.363 to 0.501); .76	
60 - 79 y (male: n=22; female: n=13)				.39				.69
Q1	3 (1-5)	7 (3)	0.451 (–0.162 to 1.065); .18		4..25 (3 - 5.5)	3 (1)	0.171 (–0.621 to 0.964); .7	
Q2	5.4 (5 - 5.8)	6 (3)	0.055 (–0.463 to 0.573); .84		6 (5.5 - 6.5)	3 (1)	0.202 (–0.204 to 0.608); .37	
Q3	6.85 (5.8 - 7.9)	3 (1)	0.070 (–0.618 to 0.758); .85		6.75 (6.5 - 7)	6 (3)	–0.038 (–0.812 to 0.736); .93	
Q4	8.7 (7.9 - 9.5)	6 (3)	0.188 (–0.345 to 0.721); .5		8 (7-9)	1 (1)	0.104 (–0.670 to 0.879); .8	
Grip strength (score)								
20 - 49 y (male: n=159; female: n=162)				.74				.045
Q1	3 (1-5)	42 (18)	0.133 (–0.053 to 0.319); .16		3 (1-5)	46 (21)	0.190 (0.081 to 0.299);<.001	
Q2	5.75 (5 - 6.5)	35 (15)	0.204 (0.024 to 0.383); .03		5.5 (5-6)	45 (20)	0.308 (0.196 to 0.420);<.001	
Q3	7.25 (6.5 - 8)	49 (21)	0.182 (0.055 to 0.309); .006		6.5 (6-7)	51 (23)	0.369 (0.270 to 0.467);<.001	
Q4	9 (8-10)	33 (14)	0.080 (–0.100 to 0.259); .39		8.5 (7-10)	19 (9)	0.198 (0.074 to 0.321); .002	
50 - 59 y (male: n=50; female: n=47)				.62				.68

XSL•FO
RenderX

Indicator	Male (n=231)				Female (n=222)			
	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value	Indicator, median (IQR)	n (%)	Model: β (95% CI); *P* value	*P* value
Q1	4.9 (1.5 - 8.3)	6 (3)	–0.015 (–0.533 to 0.503); .96		4.5 (1.5 - 7.5)	19 (9)	0.063 (–0.240 to 0.366); .69	
Q2	9.1 (8.3 - 9.8)	18 (8)	0.139 (–0.252 to 0.531); .49		7.9 (7.5 - 8.3)	14 (6)	0.098 (–0.215 to 0.411); .54	
Q3	10.5 (9.8 - 11.1)	13 (6)	0.155 (–0.286 to 0.595); .5		9.4 (8.3 - 10.5)	5 (2)	0.144 (–0.278 to 0.565); .51	
Q4	13.1 (11.1 - 15)	13 (6)	0.329 (–0.021 to 0.679); .07		12.4 (10.5 - 14.2)	10 (5)	–0.066 (–0.292 to 0.160); .57	
60 - 79 y (male: n=22; female: n=13)				.13				.53
Q1	5.65 (1.5 - 9.8)	5 (2)	0.410 (–0.273 to 1.093); .26		7.9 (7.5 - 8.3)	5 (2)	0.192 (–0.781 to 1.165); .72	
Q2	10.2 (9.8 - 10.5)	7 (3)	0.001 (–0.668 to 0.671); .99		9.1 (8.3 - 9.8)	3 (1)	0.379 (–0.108 to 0.865); .19	
Q3	10.8 (10.5 - 11.1)	5 (2)	0.457 (–0.080 to 0.994); .12		11.7 (9.8 - 13.5)	1 (1)	0.221 (–1.037 to 1.480); .75	
Q4	13.1 (11.1 - 15)	5 (2)	–0.009 (–0.717 to 0.699); .98		13.9 (13.5 - 14.2)	4 (1)	0.765 (–0.493 to 2.022); .29	

[a]WC: waist circumference.

[b]HC: hip circumference.

[c]WHR: waist-hip ratio.

[d]SBP: systolic blood pressure.

[e]DBP: diastolic blood pressure.

[f]LDL-C: low-density lipoprotein cholesterol.

[g]HDL-C: high-density lipoprotein cholesterol.

[h]TG: triglycerides.

[i]TC: total cholesterol.

[j]TRL-C: triglyceride-rich lipoprotein cholesterol.

[k]HbA_{1c}: glycated hemoglobin.

[l]GLU: glucose.

Discussion

Principal Findings

The findings of this study indicate that BMI is an independent risk factor for CRP, and its association with CRP is significantly influenced by body morphology, blood pressure, blood lipids, blood glucose, and physical fitness levels. Notably, the impact of BMI on CRP exhibits gender-specific differences. In the female population, WC, HC, WHR, SBP, DBP, TG, TRL-C, one-leg standing time with closed eyes in the 20 - 49 years age group, and grip strength are significantly correlated with the influence of BMI on CRP. In contrast, in the male population, although no significant differences were observed, analysis across different IQR intervals reveals that most cardiovascular risk factors exhibit variations in their modulating effects on the BMI-CRP association. These findings underscore the importance of gender, physical fitness, and multiple cardiovascular risk factors in the relationship between BMI and CRP, providing new perspectives for further research into the potential application value of these risk indicators in CVD prevention.

The interaction analysis stratified by gender revealed significant gender differences in the impact of BMI on CRP, with WC, HC, and WHR identified as important risk factors for increased CRP in females ($P<.05$). A study investigating the relationship between novel and traditional anthropometric measures and inflammation found that traditional body morphology indicators (such as weight, BMI, WC, and WHR) were positively associated with CRP in females [39]. This finding aligns with the results of the current study, further supporting the notion that WC, HC, and WHR can serve as significant indicators of cardiovascular risk in females. Additionally, another study demonstrated that trunk fat ratio was positively associated with hypersensitive C-reactive protein (hsCRP), while lower limb fat ratio was negatively associated with hsCRP, with this

XSL•FO
RenderX

association being statistically significant only in females [40]. The discrepancy between these findings and those of the current study may stem from differences in sample characteristics, as the female participants in that study were in the initial stages of menopause, a period characterized by the redistribution of fat from lower limb subcutaneous adipose tissue to the abdominal region, resulting in a more proinflammatory fat distribution pattern. Consequently, in females, health interventions should particularly focus on the management of WC, HC, and WHR. Although the impact of BMI on CRP is less pronounced in males than in females, attention to BMI management remains crucial, especially for overweight and obese males.

In females, SBP and DBP significantly interact with the effect of BMI on CRP (P<.05). In the Female Health Study, higher levels of CRP were associated with incident hypertension in females aged 45 years or older [41]. However, in the Physician's Health Study performed in male only, CRP was not associated with a higher risk of hypertension among individuals aged 40 - 84 years [42]. Additionally, a cohort study revealed a gender-specific association between circulating inflammatory marker levels and the incidence of blood pressure and hypertension. The results showed that higher levels of inflammation were significantly associated with elevated SBP and DBP in females, but not in males [43]. These findings are consistent with the results of the study, which further provides novel evidence underscoring the importance of controlling blood pressure for the prevention and treatment of inflammation-related CVD in females.

The effect of BMI on CRP is influenced by certain blood lipid indicators. In females, TG and TRL-C significantly interact with the relationship between BMI and CRP (P<.05), whereas in males, TC poses a greater risk for increases in CRP. The National Health and Nutrition Examination Survey, a population-based study, revealed an association between the systemic immune-inflammation index and hyperlipidemia, further exploring gender differences in the threshold effects of hyperlipidemia [15]. However, this study did not analyze the gender-specific associations between specific blood lipid indicators and CRP or reveal a complex interaction effect among BMI, CRP, and dyslipidemia. In contrast, this study elucidates this effect and provides gender-specific insights for personalized health guidance. It was not only found that lower TG and TRL-C levels provide a certain protective effect for females, helping to mitigate the negative impact of BMI on CRP, but also indicated that in males, greater attention should be paid to the influence of TC levels on inflammatory responses, thereby offering more possibilities for personalized health guidance.

In obese populations, the accumulation of visceral and subcutaneous fat activates the immune system, triggering inflammatory responses and leading to elevated circulating levels of proinflammatory cytokines [44]. This chronic inflammatory environment alters insulin signaling, thereby exacerbating insulin resistance. Prolonged insulin resistance can lead to carcinogenesis, CVD, and type 2 diabetes [45]. Additionally, a large-scale cohort study demonstrated a positive correlation between CRP levels and the risk of developing type 2 diabetes, revealing gender differences [46]. These findings are consistent with the results of the study, which provides additional evidence through gender-stratified analysis of specific blood glucose indicators, offering deeper insights into the relationship between blood glucose and inflammation. Although statistical analysis results showed no significant interaction between GLU and CRP (P>.05), the results indicated that in the male population, there is a trend where GLU influences the effect of BMI on CRP to some extent. This suggests that maintaining appropriate GLU levels is important for mitigating the potential adverse effects of BMI on CRP in males.

Static exercises (one-leg standing time with closed eyes, sit-and-reach, and grip strength) demonstrate varying protective effects on the relationship between BMI and CRP across different age groups and genders. Among females aged 20 - 49 years, one-leg standing time with closed eyes and grip strength were negatively associated with CRP (P<.05). In males, static exercises were negatively associated with CRP, except for grip strength, which was positively associated with CRP in the 50 - 59 years age group. A cross-sectional study demonstrated that higher absolute strength levels in older individuals are associated with lower inflammation levels, playing a key role in reducing the risk of cardiovascular mortality [17]. However, this study only used grip strength tests to assess muscle strength. Another cross-sectional study found a significant correlation between good physical fitness levels and lower plasma proinflammatory marker levels in the older population. Additionally, through gender-specific analysis, it further confirmed that physical activity has a positive effect on reducing inflammation levels in females [47]. The above studies focused on the association between different intensities of physical activity and inflammation but did not delve into specific quantitative indicators of physical fitness. This study, by examining multiple specific indicators of physical fitness, not only provides a more comprehensive new perspective on the age and gender differences in the impact of static exercises on CRP levels but also demonstrates that high levels of static exercise ability effectively modulate the influence of BMI on CRP in males.

This study reveals significant gender differences in physical fitness and cardiovascular risk factors, indicating that BMI-related modulation of CRP may be more pronounced in females than in males. These findings are likely attributed to biological disparities in fat metabolism and hormone levels between genders. For instance, females tend to store lipids in subcutaneous fat, while males are more prone to visceral fat accumulation. Such distinctions may render females more susceptible to BMI-associated inflammatory responses [48]. Furthermore, gender-specific factors, including chromosomal variations, sex hormones, metabolic profiles, social determinants, and lifestyle elements, could also contribute to these differences [49,50]. Consequently, the regulatory role of gender-specific factors in physical fitness and cardiovascular risk warrants considerable attention.

In this study, despite the implementation of stringent inclusion criteria and adjustment for multiple covariates (eg, age, gender, smoking, and alcohol consumption), several limitations may influence the interpretation of the results. First, the cross-sectional design precludes the establishment of causal

XSL•FO
RenderX

relationships, providing only associations between variables. Second, this study used convenience sampling to recruit healthy residents who underwent self-assessments at sports hospitals as research subjects, potentially introducing selection bias, as individuals with better health status or higher health awareness may be more inclined to participate, thereby limiting the generalizability of the findings. Additionally, although gender-stratified interaction tests were conducted, potential confounding factors (eg, dietary habits and psychological stress) may not have been fully controlled, potentially affecting the interpretation of the relationship between BMI and CRP. Finally, despite comprehensive assessments (eg, physical examinations, blood biochemical tests, and imaging studies), unmeasured biomarkers or environmental factors could still exert potential influences on the results. Despite these limitations, the large sample size, rigorous standards, multivariable adjustments, and standardized procedures minimized bias and enhanced the reliability of the findings. Future studies should adopt longitudinal designs and incorporate additional potential confounders to further validate and extend the results of this study. This study exhibits several notable strengths. First, the adoption of a multidimensional and multilevel analytical framework provides an in-depth exploration of how physical fitness and cardiovascular risk factors influence CRP through BMI, offering novel perspectives and theoretical foundations for research in this field. Second, the implementation of gender-specific analytical methods reveals significant gender differences in the impact of BMI on CRP, which holds substantial practical implications for improving clinical practice and informing policy development.

Limitations

This study also presents several limitations. First, the use of cross-sectional data precludes the inference of causal relationships between variables. The analytical results merely reflect associations between variables, and future studies should consider using longitudinal or experimental designs to further explore potential causal mechanisms. Second, the use of convenience sampling of participants recruited from sports hospitals for cross-sectional surveys limits the generalizability of the results and, as with most cross-sectional studies, there may be a selection bias that may render the sample unrepresentative, but we have tried to control bias as much as possible by clarifying the inclusion and exclusion criteria, tightly controlling the stage of data collection to improve the representativeness of the sample and the trustworthiness of the study, and so on. Subsequent studies may mitigate these biases by expanding the sampling scope, conducting multistage sampling, and others, to provide wider generalization value.

Conclusions

BMI exerts a significant influence on CRP, with distinct cardiovascular risk factors, including body morphology, blood pressure, blood lipids, and blood glucose, and physical fitness modulating the BMI-CRP relationship differently between genders. Specifically, in females, body morphology, physiological parameters, and physical fitness significantly influence the relationship between BMI and CRP. In males, certain indicators also show similar trends, but the extent of the influence differs by gender. Therefore, when managing the BMI-CRP relationship in the female population, high-risk cardiovascular risk factors should be considered in combination with age in order to optimize intervention strategies. This finding provides an important scientific basis for the development of gender-specific health management strategies, which is of great clinical significance for reducing CVD risk and improving health outcomes.

Acknowledgments

The authors sincerely thank all personnel at Changzhou Sports Hospital who participated in the physical examinations for their kindness. They also acknowledge the enthusiasm and perseverance of all research team members and the assistance and support provided by the staff of Changzhou Sports Hospital involved in the physical examinations. This work was supported by the National Natural Science Foundation of China (grants 32371180 and 31971099), supported by the training targets of young and middle-aged academic leaders in the Jiangsu Province Blue Project, the Major Sports Research Project of the Jiangsu Provincial Sports Bureau in 2023 (ST231208), the Major Science and Technology Demonstration Project for Social Development in Jiangsu Province (BE2018752), the National Social Science Foundation of China (23CTY018), and the Jiangsu Province Social Science Foundation Program (21TYD004).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

CL contributed to the conception and editing of the manuscript. Yunqing Liu, YY, BC, Yakun Liu, and FY were involved in data collection. CH performed statistical analyses. Fanghui Li and Liqiang Su contributed to the conception, study design, and data interpretation. CL, Yunqing Liu, and YY made equal contributions in their respective roles, and Fanghui Li and Liqiang Su are co-corresponding authors.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Physical fitness test methodology.
[DOCX File, 18 KB - publichealth_v11i1e76485_app1.docx]

Multimedia Appendix 2
Interaction plots and interaction analysis of covariates.
[DOCX File, 116 KB - publichealth_v11i1e76485_app2.docx]

References

1. Tan SCW, Zheng BB, Tang ML, Chu H, Zhao YT, Weng C. Global burden of cardiovascular diseases and its risk factors, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. QJM 2025 Jan 23:hcaf022. [doi: 10.1093/qjmed/hcaf022] [Medline: 39847534]
2. National Center for Cardiovascular Diseases, China. Report on cardiovascular health and diseases in China 2024:an updated summary. Chin Circ J 2025;40(6):521-559. [doi: 10.3969/j.issn.1000-3614.2025.06.001]
3. Rocha T, Melson E, Zamora J, Fernandez-Felix BM, Arlt W, Thangaratinam S. Sex-specific obesity and cardiometabolic disease risks in low- and middle-income countries: a meta-analysis involving 3 916 276 individuals. J Clin Endocrinol Metab 2024 Mar 15;109(4):1145-1153. [doi: 10.1210/clinem/dgad599] [Medline: 37930879]
4. Csige I, Ujvárosy D, Szabó Z, et al. The impact of obesity on the cardiovascular system. J Diabetes Res 2018;2018(3407306):3407306. [doi: 10.1155/2018/3407306] [Medline: 30525052]
5. Koliaki C, Liatis S, Kokkinos A. Obesity and cardiovascular disease: revisiting an old relationship. Metab Clin Exp 2019 Mar;92:98-107. [doi: 10.1016/j.metabol.2018.10.011] [Medline: 30399375]
6. Aleksandrova K, Mozaffarian D, Pischon T. Addressing the perfect storm: biomarkers in obesity and pathophysiology of cardiometabolic risk. Clin Chem 2018 Jan;64(1):142-153. [doi: 10.1373/clinchem.2017.275172] [Medline: 29138271]
7. Heymsfield SB, Peterson CM, Thomas DM, Heo M, Schuna JM Jr. Why are there race/ethnic differences in adult body mass index-adiposity relationships? A quantitative critical review. Obes Rev 2016 Mar;17(3):262-275. [doi: 10.1111/obr.12358] [Medline: 26663309]
8. Prillaman M. Why BMI is flawed and how to redefine obesity. Nature New Biol 2023 Oct;622(7982):232-233. [doi: 10.1038/d41586-023-03143-x] [Medline: 37821595]
9. Powell-Wiley TM, Poirier P, Burke LE, et al. Obesity and cardiovascular disease: a scientific statement from the American Heart Association. Circulation 2021 May 25;143(21):e984-e1010. [doi: 10.1161/CIR.0000000000000973] [Medline: 33882682]
10. Beyene HB, Giles C, Huynh K, et al. Metabolic phenotyping of BMI to characterize cardiometabolic risk: evidence from large population-based cohorts. Nat Commun 2023 Oct 7;14(1):6280. [doi: 10.1038/s41467-023-41963-7] [Medline: 37805498]
11. Harris E. Study: waist-to-hip ratio might predict mortality better than BMI. JAMA 2023 Oct 24;330(16):1515-1516. [doi: 10.1001/jama.2023.19205]
12. Liqiang S, Fang-Hui L, Minghui Q, Haichun C. Threshold effect and sex characteristics of the relationship between chronic inflammation and BMI. BMC Endocr Disord 2023 Aug 16;23(1):175. [doi: 10.1186/s12902-023-01396-1] [Medline: 37582770]
13. Nowak MM, Niemczyk M, Gołębiewski S, Pączek L. Impact of body mass index on all-cause mortality in adults: a systematic review and meta-analysis. J Clin Med 2024 Apr 16;13(8):2305. [doi: 10.3390/jcm13082305] [Medline: 38673577]
14. Vigna L, Tirelli AS, Gaggini M, et al. Insulin resistance and cardiometabolic indexes: comparison of concordance in working-age subjects with overweight and obesity. Endocrine 2022 Aug;77(2):231-241. [doi: 10.1007/s12020-022-03087-8] [Medline: 35665880]
15. Mahemuti N, Jing X, Zhang N, et al. Association between systemic immunity-inflammation index and hyperlipidemia: a population-based study from the NHANES (2015-2020). Nutrients 2023 Feb 26;15(5):1177. [doi: 10.3390/nu15051177] [Medline: 36904176]
16. Elimam H, Abdulla AM, Taha IM. Inflammatory markers and control of type 2 diabetes mellitus. Diabetes Metab Syndr 2019 Jan;13(1):800-804. [doi: 10.1016/j.dsx.2018.11.061]
17. Silva B, Camões M, Simões M, Bezerra P. Obesity, physical fitness and inflammation in the elderly. Geriatrics (Basel) 2017 Sep 21;2(4):30. [doi: 10.3390/geriatrics2040030] [Medline: 31011040]
18. Garcia-Sifuentes Y, Maney DL. Reporting and misreporting of sex differences in the biological sciences. Elife 2021 Nov 2;10:e70817. [doi: 10.7554/eLife.70817] [Medline: 34726154]
19. Allegra S, Chiara F, Di Grazia D, Gaspari M, De Francia S. Evaluation of sex differences in preclinical pharmacology research: how far is left to go? Pharmaceuticals (Basel) 2023 May 24;16(6):786. [doi: 10.3390/ph16060786] [Medline: 37375734]

XSL•FO
RenderX

20. Hu C, Lv Z, Zhu J, et al. The impact of digital technology-based exercise combined with dietary intervention on body composition in college students with obesity: prospective study. J Med Internet Res 2025 Jun 2;27:e65868. [doi: 10.2196/65868] [Medline: 40455565]
21. Cooper AJ, Gupta SR, Moustafa AF, Chao AM. Sex/gender differences in obesity prevalence, comorbidities, and treatment. Curr Obes Rep 2021 Dec;10(4):458-466. [doi: 10.1007/s13679-021-00453-x] [Medline: 34599745]
22. Benjamin EJ, Virani SS, Callaway CW, et al. Heart disease and stroke statistics—2018 update: a report from the American Heart Association. Circulation 2018 Mar 20;137(12):e67-e492. [doi: 10.1161/CIR.0000000000000558] [Medline: 29386200]
23. Sirisena A, Okeahialam B. Association of obesity anthropometric indices with hypertension, diabetes mellitus and hypertriglyceridemia in apparently healthy adult Nigerian population. World J Cardiol 2022 Jun 26;14(6):363-371. [doi: 10.4330/wjc.v14.i6.363] [Medline: 35979178]
24. Chowdhury MZI, Haque MA, Farhana Z, et al. Prevalence of cardiovascular disease among Bangladeshi adult population: a systematic review and meta-analysis of the studies. Vasc Health Risk Manag 2018;14:165-181. [doi: 10.2147/VHRM.S166111] [Medline: 30174432]
25. Vázquez-Aguilar A, Rueda-Robles A, Rivas-García L, et al. Association of dietary patterns with cardiovascular disease risk factors in Mexican adults: insights from a cross-sectional descriptive study. Nutrients 2024 Mar 12;16(6):804. [doi: 10.3390/nu16060804] [Medline: 38542715]
26. Relja A, Miljković A, Gelemanović A, et al. Nut consumption and cardiovascular risk factors: a cross-sectional study in a Mediterranean population. Nutrients 2017 Nov 28;9(12):1296. [doi: 10.3390/nu9121296] [Medline: 29182576]
27. Zou Q, Su C, Du W, et al. The mediation and moderation effect association among physical activity, body-fat percentage, blood pressure, and serum lipids among Chinese adults: findings from the China Health and Nutrition Surveys in 2015. Nutrients 2023 Jul 12;15(14):3113. [doi: 10.3390/nu15143113] [Medline: 37513531]
28. Yudkin JS, Stehouwer CD, Emeis JJ, Coppack SW. C-reactive protein in healthy subjects: associations with obesity, insulin resistance, and endothelial dysfunction: a potential role for cytokines originating from adipose tissue? Arterioscler Thromb Vasc Biol 1999 Apr;19(4):972-978. [doi: 10.1161/01.atv.19.4.972] [Medline: 10195925]
29. Kuppa A, Tripathi H, Al-Darraji A, Tarhuni WM, Abdel-Latif A. C-reactive protein levels and risk of cardiovascular diseases: a two-sample bidirectional Mendelian randomization study. Int J Mol Sci 2023 May 23;24(11):9129. [doi: 10.3390/ijms24119129] [Medline: 37298077]
30. Wang Z, Nakayama T. Inflammation, a link between obesity and cardiovascular disease. Mediators Inflamm 2010;2010:535918. [doi: 10.1155/2010/535918] [Medline: 20847813]
31. Ni P, Yu M, Zhang R, et al. Dose-response association between C-reactive protein and risk of all-cause and cause-specific mortality: a systematic review and meta-analysis of cohort studies. Ann Epidemiol 2020 Nov;51:20-27. [doi: 10.1016/j.annepidem.2020.07.005] [Medline: 32702432]
32. Pan M, Tu R, Gu J, et al. Associations of socioeconomic status and physical activity with obesity measures in rural Chinese adults. Front Public Health 2020;8:594874. [doi: 10.3389/fpubh.2020.594874] [Medline: 33490019]
33. Zhou W, Zhu L, Yu Y, Yu C, Bao H, Cheng X. A body shape index is positively associated with all-cause and cardiovascular disease mortality in the Chinese population with normal weight: a prospective cohort study. Nutr Metab Cardiovasc Dis 2023 Sep;33(9):1702-1708. [doi: 10.1016/j.numecd.2023.05.016] [Medline: 37414663]
34. Roth GA, Mensah GA, Johnson CO, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019: update from the GBD 2019 study. J Am Coll Cardiol 2020 Dec 22;76(25):2982-3021. [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
35. Yang J, Christophi CA, Farioli A, et al. Association between push-up exercise capacity and future cardiovascular events among active adult men. JAMA Netw Open 2019 Feb 15;2(2):e188341. [doi: 10.1001/jamanetworkopen.2018.8341]
36. Xuyang L. Impact of triglyceride-rich lipoprotein-cholesterol on cardiovascular disease and all-cause mortality. Chin Circ J 2022;37(9):920-927. [doi: 10.3969/j.issn.1000-3614.2022.09.009]
37. National physical fitness measurement standards. General Administration of Sport of China. 2023 Aug 10. URL: https://www.sport.gov.cn/n315/n20001395/c25880704/content.html [accessed 2025-8-9]
38. Jie Z, Chen C, Hao L, et al. Life history recorded in the vagino-cervical microbiome along with multi-omes. Genomics Proteomics Bioinformatics 2022 Apr 1;20(2):304-321. [doi: 10.1016/j.gpb.2021.01.005]
39. González-Gil EM, Peruchet-Noray L, Sedlmeier AM, et al. Association of body shape phenotypes and body fat distribution indexes with inflammatory biomarkers in the European Prospective Investigation into Cancer and Nutrition (EPIC) and UK Biobank. BMC Med 2024 Aug 15;22(1):334. [doi: 10.1186/s12916-024-03544-3] [Medline: 39148045]
40. Wu B, Huang J, Fukuo K, Suzuki K, Yoshino G, Kazumi T. Different associations of trunk and lower-body fat mass distribution with cardiometabolic risk factors between healthy middle-aged men and women. Int J Endocrinol 2018;2018:1289485. [doi: 10.1155/2018/1289485] [Medline: 29531527]
41. Sesso HD, Buring JE, Rifai N, Blake GJ, Gaziano JM, Ridker PM. C-reactive protein and the risk of developing hypertension. JAMA 2003 Dec 10;290(22):2945-2951. [doi: 10.1001/jama.290.22.2945] [Medline: 14665655]
42. Sesso HD, Jiménez MC, Wang L, Ridker PM, Buring JE, Gaziano JM. Plasma inflammatory markers and the risk of developing hypertension in men. J Am Heart Assoc 2015 Sep 21;4(9):e001802. [doi: 10.1161/JAHA.115.001802] [Medline: 26391130]

XSL•FO
RenderX

43. Kringeland E, Gerdts E, Ulvik A, et al. Inflammation, sex, blood pressure changes and hypertension in midlife: the Hordaland health study. J Hum Hypertens 2023 Aug;37(8):718-725. [doi: 10.1038/s41371-022-00772-z] [Medline: 36400946]
44. Silva FM, Duarte-Mendes P, Teixeira AM, Soares CM, Ferreira JP. The effects of combined exercise training on glucose metabolism and inflammatory markers in sedentary adults: a systematic review and meta-analysis. Sci Rep 2024 Jan 22;14(1):1936. [doi: 10.1038/s41598-024-51832-y] [Medline: 38253590]
45. McLeod A, Schiffer L, Castellanos K, et al. Impact of physical activity and weight loss on fat mass, glucose metabolism, and inflammation in older African Americans with osteoarthritis. Nutrients 2020 Oct 28;12(11):3299. [doi: 10.3390/nu12113299] [Medline: 33126555]
46. Kanmani S, Kwon M, Shin MK, Kim MK. Association of c-reactive protein with risk of developing type 2 diabetes mellitus, and role of obesity and hypertension: a large population-based Korean cohort study. Sci Rep 2019 Mar 14;9(1):4573. [doi: 10.1038/s41598-019-40987-8] [Medline: 30872696]
47. Elosua R, Bartali B, Ordovas JM, et al. Association between physical activity, physical performance, and inflammatory biomarkers in an elderly population: the InCHIANTI study. J Gerontol A Biol Sci Med Sci 2005 Jun;60(6):760-767. [doi: 10.1093/gerona/60.6.760] [Medline: 15983180]
48. Palmisano BT, Zhu L, Eckel RH, Stafford JM. Sex differences in lipid and lipoprotein metabolism. Mol Metab 2018 Sep;15(45-55):45-55. [doi: 10.1016/j.molmet.2018.05.008]
49. Appelman Y, van Rijn BB, Ten Haaf ME, Boersma E, Peters SAE. Sex differences in cardiovascular risk factors and disease prevention. Atherosclerosis 2015 Jul;241(1):211-218. [doi: 10.1016/j.atherosclerosis.2015.01.027] [Medline: 25670232]
50. Lu J, Li M, He J, et al. Association of social determinants, lifestyle, and metabolic factors with mortality in Chinese adults: a nationwide 10-year prospective cohort study. Cell Rep Med 2024 Aug 20;5(8):101656. [doi: 10.1016/j.xcrm.2024.101656] [Medline: 39067445]

Abbreviations

CRP: C-reactive protein
CVD: cardiovascular disease
DBP: diastolic blood pressure
GLU: glucose
HbA_{1c}: glycated hemoglobin
HC: hip circumference
HDL-C: high-density lipoprotein cholesterol
hsCRP: hypersensitive C-reactive protein
LDL-C: low-density lipoprotein cholesterol
SBP: systolic blood pressure
TC: total cholesterol
TG: triglycerides
TRL-C: triglyceride-rich lipoprotein cholesterol
WC: waist circumference
WHR: waist-hip ratio

Edited by A Mavragani, T Sanchez; submitted 24.04.25; peer-reviewed by F Zhu, S Sahai; revised version received 10.07.25; accepted 11.07.25; published 27.08.25.

Please cite as:
Lai C, Liu Y, Yan Y, Cheng B, Liu Y, Yu F, Hu C, Su L, Li F
Gender-Specific Moderating Role of Physical Fitness and Cardiovascular Risk Factors in the Relationship Between BMI and C-Reactive Protein: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e76485
URL: https://publichealth.jmir.org/2025/1/e76485
doi:10.2196/76485

© Chunyuan Lai, Yunqing Liu, Yue Yan, Boli Cheng, Yakun Liu, Fanshu Yu, Chengyuan Hu, Liqiang Su, Fanghui Li. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Epidemiological Features, Clinical Symptoms, and Environmental Risk Factors for Notifiable Japanese Encephalitis in Taiwan From 2008 to 2020: Retrospective Study

Fu-Huang Lin[1], PhD; Yu-Ching Chou[1], PhD; Chi-Jeng Hsieh[2], PhD; Chia-Peng Yu[1], PhD

[1]School of Public Health, National Defense Medical Center, Taipei City, Taiwan

[2]Department of Health Care Administration, Asia Eastern University of Science and Technology, New Taipei City, Taiwan

Corresponding Author:
Chia-Peng Yu, PhD
School of Public Health, National Defense Medical Center, Taipei City, Taiwan

Abstract

Background: Japanese encephalitis (JE) is a zoonotic parasitic disease caused by the Japanese encephalitis virus (JEV), and may cause fever, nausea, headache, or meningitis. It is currently unclear whether the epidemiological characteristics of the JEV have been affected by the extreme climatic conditions that have been observed in recent years.

Objective: This study aimed to examine the epidemiological characteristics, trends, and potential risk factors of JE in Taiwan from 2008 to 2020. Specifically, the study focused on gender, age, season, residential area, clinical manifestations, high-risk areas, and the impact of environmental and climate factors.

Methods: This study reviewed publicly available annual summary data on reported JE cases in the Taiwan Centers for Diseases Control between 2008 and 2020.

Results: This study collected 309 confirmed domestic patients and 4 patients with imported JE. There was an increasing trend in the incidence of JE, 0.69 - 1.57 cases per 1,000,000 people, peaking in 2018. Case fatality rate was 7.7% (24/313). Comparing sex, age, season, and place of residence, the incidence rate was highest in males, 40 - to 59-year-old patients, summer, and the Eastern region, with 1.89, 3.27, 1.25, and 12.2 cases per million people, respectively. The average coverage rate of the JE vaccine for children in Taiwan is 94.9%. Additionally, the major clinical manifestations of the cases included fever, unconsciousness, headache, stiff necks, psychological symptoms, vomiting, and meningitis. The major occurrence places of JE included paddy fields, pig farms, pigeon farms, poultry farms, and ponds. For air pollution factors, linear regression analysis showed that SO_2 (ppb) concentration was positively associated with JE cases (β=2.184, *P*=.02), but O_3 (ppb) concentration was negatively associated with them (β=−0.157, *P*=.01). For climate factors, relative humidity (%) was positively associated with JE cases (β=.380, *P*=.02).

Conclusions: This study is the first to report confirmed cases of JE from the surveillance data of the Taiwan Centers for Diseases Control between 2008 and 2020. It identified residence, season, and age as risk factors for JE in Taiwan. Air pollution and climatic factors also influenced the rise in JE cases. This study confirmed that JE remains a prevalent infectious disease in Taiwan, with its epidemic gradually increasing in severity. These findings empower clinicians and health care providers to make informed decisions, guiding their care and resource allocation for patients with JE, a disease that significantly impacts the health and well-being of the Taiwanese population.

(JMIR Public Health Surveill 2025;11:e63053) doi:10.2196/63053

KEYWORDS

epidemiology; Japanese encephalitis virus; domestic; environmental factor; retrospective study

Introduction

The Japanese encephalitis virus (JEV) is a significant public health concern in Asia, causing an estimated 68,000 clinical cases annually, primarily in East Asia and Southeast Asia [1]. Although symptomatic Japanese encephalitis (JE) occurs infrequently, it has a serious case fatality rate, with up to 30% of encephalitis cases resulting in death. Encephalitis survivors have a high likelihood of developing permanent neurological or psychiatric sequelae, with an estimated occurrence rate of 30%-50% of cases. The endemic transmission of JEV poses a health threat to over 3 billion people in a substantial portion of the World Health Organization Southeast Asia and Western Pacific regions [1]. The disease currently lacks specific antiviral treatment, necessitating supportive care measures like managing symptoms and providing respiratory and nutritional support to help patients overcome the infection [2]. Vaccination stands as the effective strategy to prevent the disease, with World Health Organization recommending the integration of JE vaccination

XSL•FO
RenderX

into national immunization programs in all regions where the disease poses a significant public health threat [1,3].

JE is a mosquito-borne viral infection caused by the JEV, a member of the *Flaviviridae* family closely related to dengue, yellow fever, and West Nile viruses [4]. The primary means of transmission of JEV from animals to humans is through the bite of infected mosquitoes, particularly Culex species, with *Culex tritaeniorhynchus* (*Cx. tritaeniorhynchus*) being of particular importance. Pigs and ardeid birds serve as the main amplifying hosts for the virus [5]. The clinical presentation of JE in humans varies widely, encompassing a spectrum ranging from mild or asymptomatic infection to severe encephalitis. This severe form of the disease is characterized by significant mortality, with a high death rate among infected individuals. Additionally, survivors of severe JE often experience sequelae within the central nervous system [6].

In recent years, air pollution has been identified as a significant public health concern in China, with an estimated 1 million deaths annually attributed to its effects. This has underscored the urgency for the implementation of effective prevention strategies. However, future climate change is likely to increase the frequency and duration of severe weather conditions, thereby promoting exposure to air pollution and exacerbating the impacts on human health [7]. A previous study indicated that tropical and subtropical populations of *Cx. tritaeniorhynchus* exhibited higher genetic diversity and that the effects of environmental factors (climatic conditions) provide significant advantages for the establishment and expansion of *Cx. tritaeniorhynchus* [8]. An understanding of the impact of environmental factors on JE can provide valuable insights into its morbidity and mortality, which remain unclear.

JE poses a significant public health threat, particularly in Asian countries, especially in South Asia, Southeast Asia, and East Asia [9]. Several countries in these regions have successfully implemented various intervention measures to reduce JE morbidity, including early diagnosis, prompt treatment, national immunization programs, and effective vector control strategies [10,11]. Despite these efforts, over 3 billion people still reside in JE-endemic countries, with an estimated 67,900 cases occurring annually [8]. The severity of JE is evident in its high mortality rate, with approximately 20% - 30% of cases resulting in fatalities. Additionally, 30% - 50% of JE survivors experience significant long-term neurological sequelae [9,12]. Previous studies have also highlighted the ongoing public health concern posed by JE in Taiwan [13].

Nestled at 23°4' north latitude and 121°0' east longitude, Taiwan enjoys a subtropical climate with average monthly temperatures ranging from a comfortable 16 to 29 °C. Average monthly relative humidity hovers between 75% and 90%. Despite its status as a developed nation with a per capita gross domestic product of US$35,244 [14], Taiwan continues to grapple with the challenges posed by JE. Despite occasional imported cases of JE in recent years [15], local infections persist in Taiwan, indicating limitations in the effectiveness of preventive measures in controlling or eradicating the disease. The number of JE cases in Taiwan ranges from sporadic confirmed cases to cluster epidemics, with the middle-aged population exhibiting the highest rate of JEV. Addressing the disease burden caused by JEV is imperative, and effective public health measures should be adopted in Taiwan. Although some studies provided epidemiological data on JE in Taiwan [16], their study did not delve into the clinical symptoms of JE patients over the past decade nor investigate the potential association between climate and air pollution factors and the number of cases [17]. The objective of this study is to conduct a comprehensive investigation of the epidemiological, clinical, and environmental characteristics of JE in Taiwan. To this end, population surveillance data gathered over a 13-year period will be used.

Methods

Study Design and Population

Study Area

Taiwan is an island in East Asia located between 21°45' N and 25°56' N. The Tropic of Cancer (23.5° N) passes through Chiayi City, situated in south-central Taiwan, dividing the entire island into 2 climate zones. Taiwan has a total land area of 35,980 km² and a population of approximately 23 million people, resulting in an average population density of 635 individuals per km². The northern part of Taiwan belongs to the subtropical climate zone, whereas the southern part belongs to the tropical climate zone. Consequently, the weather in Taiwan is relatively warm, with high humidity occurring throughout the year [18,19].

Case Definitions

A clinical case was defined as an individual of any age experiencing an acute onset of fever accompanied by a change in mental status and new onset of seizures (excluding simple febrile seizures) at any time of the year [9,13]. A confirmed case was defined as a clinical case with positive laboratory test results, including the presence of IgM antibodies specific to the JE virus in a single sample of cerebrospinal fluid (CSF) or serum; a 4-fold increase in IgG antibodies; the detection of JE virus antigens in tissue via immunohistochemistry; or the detection of the JE virus genome in serum, plasma, blood, CSF, or tissue samples. Additionally, cases meeting the clinical case definition and epidemiologically linked to a confirmed case were also considered confirmed [20,21].

Data Collection

Surveillance for JE Infection

JE has been categorized as a notifiable infectious disease since 1955. Physicians are required to report all cases that meet the case definition of JE, collect samples, and send them to the Taiwan Centers for Disease Control (TCDC) within one week of the case being reported for examination [20]. We collected data from all JE-confirmed cases reported to the TCDC from January 2008 to December 2020. The reported information included patient age, sex, area of residence, geographic location, seasonal variation of exposure, and date of JE onset.

Laboratory Tests

In 1998, the TCDC developed an E/M-specific capture IgM and IgG enzyme-linked immunosorbent assay (E/M-specific IgM/IgG ELISA) for JE and dengue fever (DF) [13,22,23].

XSL•FO
RenderX

Hemagglutination inhibition (HI) and E/M-specific IgM/IgG ELISA are both employed as screening tools for detecting antibodies against JE and DF [24-26]. This study employed a previously described differential testing algorithm [26]. Briefly, positive samples on the JE ELISA underwent further testing using DF ELISA. The combined results from both assays informed the final interpretation. Positive JE ELISA and negative DF ELISA indicated the presence of JEV antibodies only, while positive results in both tests suggested a false-positive JE result. Long-term evaluation demonstrated high sensitivity and specificity of the E/M-specific IgM/IgG ELISA, effectively differentiating JEV infections from dengue virus infections [26]. Since 2001, real-time polymerase chain reaction (PCR) analysis of acute-phase serum samples collected within 7 days of symptom onset and CSF from individuals diagnosed with JE has been employed for diagnostic purposes [27]. Nevertheless, enzyme-linked immunosorbent assay (ELISA) remains the primary diagnostic testing method, with the E/M-specific IgM/IgG ELISA being the most widely used ELISA method.

The JE laboratory has confirmed that clinical cases meeting one of the following specific laboratory criteria should be defined as confirmed cases of JE: (1) an HI titer of the convalescent serum of 1:160 and at least a 4-fold increase between the HI titers of convalescent and acute sera [24,25]; (2) an HI titer of 1:320 derived from a single serum sample [24,25]; (3) an IgM-positive serum sample, as determined by an ELISA test, or a fourfold increase in IgG levels between paired serum samples [25]; (4) a sample exhibiting a positive real-time PCR [25]; (5) a sample that has been found to be positive for indirect immunofluorescence antibody staining following the isolation of the virus from CSF [25,28]; (6) virus isolation can be accomplished through cell culture techniques using either the mosquito C6/36 cell line or plaque assays with the BHK-21 cell line [26]; and (7) an alternative approach is the extraction of viral RNA from JEV-infected culture medium using the QIAamp Viral RNA Mini Kit (Qiagen, Germantown, MD) [15].

Surveillance of Environmental Factors

The dataset comprised monthly air pollutant data from 2008 to 2020, obtained from the air quality monitoring network of the Environmental Protection Agency. The pollutants analyzed included total suspended particulates, particulate matter (PM) 2.5 and 10 (PM2.5 and PM10), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), carbon monoxide (CO), ozone (O_3), dust fall volume, and lead concentration [29]. Furthermore, monthly weather data from the Meteorological Bureau of the Ministry of Communications for the same period were analyzed. The meteorological variables included temperature, rainfall, relative humidity, atmospheric pressure, the number of days with precipitation, and the duration of sunshine [30]. Statistical analyses and correlation tests were employed to assess the temporal and spatial trends in air pollutants and meteorological factors, as well as their potential association with JE case numbers. The data collected for this study were stored in the supplementary file.

Statistical Analysis

This study presents a retrospective historical analysis of all cases of JE diagnosed in Japan since 2008, including both domestic and imported cases. The number of individuals diagnosed with JE from 2008 to 2020 was confirmed, and the epidemiological characteristics of these cases were examined. These characteristics included sex (male and female), age groups (<20, 20 - 39, 40 - 59, and ≥60 years old), time of diagnosis (spring, summer, fall, and winter), and area of residence (northern, central, southern, and eastern Taiwan). Disparities and outcomes were analyzed. Our analysis then focused on the variables of sex, age, time of diagnosis, changes in living area, trends, and related outcomes for cases of JE from 2011 to 2020. Descriptive data are presented as means and summaries where appropriate. Categorical variables were compared using chi-squared tests. Odds ratios were calculated using logistic regression, with 95% confidence intervals estimated using parameter estimation. All statistical analyses were performed using SPSS (IBM SPSS version 21; Asia Analytics Taiwan Ltd). All statistical tests were 2-tailed, with an α value of .05. A *P* value less than .05 was considered statistically significant.

Ethical Considerations

The ethical considerations of this study were consistent with those of Holland et al [31] in *JAMA Psychiatry* and Pan et al [32] in *BMC Infectious Diseases*. All case data are accessible via the internet, with information regarding JE cases available on the TCDC website. This study was deemed exempt from ethical approval according to the Communicable Disease Control Act of Taiwan [33] because the datasets used were deidentified and the study did not involve any of the datasets in a way that could potentially lead to the identification of individuals. Along with the absence of personal identifiers and the absence of any potential harm to individuals, the study was conducted in accordance with institutional research guidelines to ensure scientific rigor and ethical standards.

Results

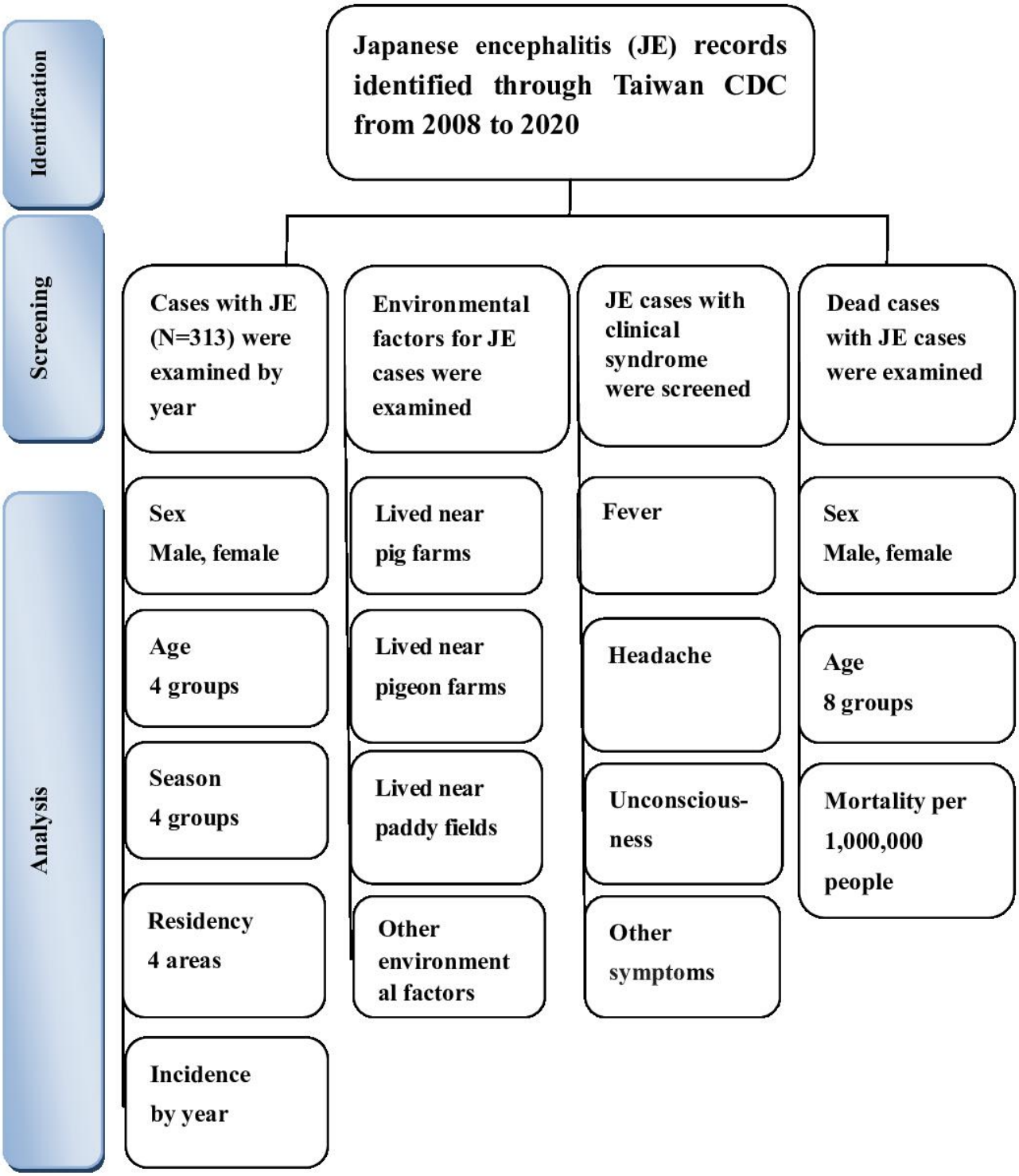
Demographic Data

The figures and tables selected for analysis in this study are representative of an examination of demographic characteristics (incidence rates), patient clinical symptoms, environmental factors, air pollution, and climate factors on JE in Taiwan over the recent 13-year period. The flowchart of this study is shown in Figure 1. There were 313 confirmed cases of local and imported infection, consisting of 309 (98.7%) confirmed local cases, 195 (62.3%) male patients, 182 (58.1%) patients aged 40 - 59 years, 245 (78.3%) cases in summer, and 145 (46.3%) cases in the southern region. There were 4 (1.3%) confirmed cases of imported infection (Table 1). There was no statistically significant difference between the number of confirmed cases in 2008 and 2020 and the relationship between different identities (local or imported), sex, age group, season, and place of residence (all $P>.05$) (Table 1). There was a statistically significant difference between the number of confirmed cases in 2008 and 2020 and the relationship between different seasons and place of residence (all $P<.05$) (Table 1). The main environmental factors associated with JE were paddy fields

XSL•FO
RenderX

(160/582, 27.5%), pig farms (158/582, 27.1%), pigeon farms (125/582, 21.5%), poultry farms (65/582, 11.2%), ponds (29/582, 5%), and domestic pets (18/582, 3.1%), as indicated in Table 2.

Figure 1. Flowchart of the study sample selection from the Taiwan Centers for Disease Control Database (CDC) in Taiwan from January 2008 to December 2020.



XSL•FO
RenderX

Table . Epidemiological features of Japanese encephalitis cases in Taiwan from 2008 to 2020.

Variable	Total	Year													*P* value
Domestic		2008 (n=17)	2009 (n=18)	2010 (n=33)	2011 (n=22)	2012 (n=32)	2013 (n=16)	2014 (n=18)	2015 (n=30)	2016 (n=23)	2017 (n=25)	2018 (n=37)	2019 (n=21)	2020 (n=21)	
Individual, n															.12
Domestic	309	17	17	32	22	32	16	18	30	23	25	37	19	21	
Imported	4	0	1	1	0	0	0	0	0	0	0	0	2	0	
Sex, n															.90
Male	195	12	13	22	14	17	12	12	19	12	15	20	13	14	
Female	118	5	5	11	8	15	4	6	11	11	10	17	8	7	
Age group, n															.26
<20	10	3	1	1	1	0	1	0	0	1	0	1	0	1	
20 - 39	68	5	3	11	8	9	5	3	7	2	4	6	3	2	
40 - 59	182	9	9	16	9	17	9	10	20	16	14	24	14	15	
≥60	53	0	5	5	4	6	1	5	3	4	7	6	4	3	
Season, n															<.001
Spring	40	1	2	2	0	10	2	0	0	2	3	10	7	1	
Summer	245	15	14	29	20	19	14	15	28	14	19	25	13	20	
Fall	28	1	2	2	2	3	0	3	2	7	3	2	1	0	
Winter	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Residency, n															.008
Northern	78	2	5	10	9	4	5	2	7	4	8	10	5	7	
Central	65	4	2	6	5	5	3	5	4	5	5	9	5	7	
Southern	145	9	4	14	4	21	8	11	18	11	10	17	11	7	
Eastern	25	2	7	3	4	2	0	0	1	3	2	1	0	0	

XSL•FO
RenderX

Table . Environmental factors for Japanese encephalitis cases in Taiwan between 2008 and 2020.

Variable (lived nearby)	Year													
	Total, n	2008, n	2009, n	2010, n	2011, n	2012, n	2013, n	2014, n	2015, n	2016, n	2017, n	2018, n	2019, n	2020, n
Pig farms	158	4	6	8	10	23	10	11	13	11	13	25	13	11
Pigeon farms	125	7	4	14	2	17	5	7	9	11	10	19	8	12
Paddy fields	160	9	6	17	9	16	9	11	14	11	8	26	14	10
Egret nests	6	1	—[a]	1	—	4	—	—	—	—	—	—	—	—
Hen-houses	2	1	1	—	—	—	—	—	—	—	—	—	—	—
Pond	29	2	2	3	1	7	3	2	3	1	1	3	1	—
Clam breeding farms	1	—	1	—	—	—	—	—	—	—	—	—	—	—
Poultry farms	65	—	—	9	6	11	3	—	5	6	5	12	4	4
Goat farms	4	—	—	2	—	1	—	—	—	—	1	—	—	—
Domestic pets	18	—	—	1	5	5	3	—	2	1	—	—	—	1
Farmlands	6	—	—	—	3	3	—	—	—	—	—	—	—	—
Orchards	8	—	—	—	3	5	—	—	—	—	—	—	—	—

[a]Not applicable

Clinical Symptoms

The main clinical symptoms of JE were fever (280/866, 32.3%), unconsciousness (162/866, 18.7%), headache (136/866, 15.7%), stiff neck (63/866, 7.3%), psychiatric symptoms (59/866, 6.8%), and vomiting (49/866, 5.7%), as shown in Table 3.

There were 24 deaths from JE, of which 15 (62.5%) were male and 9 (37.5%) were aged 50 - 60 years. The mortality rate (per 1,000,000 people) ranged from 0.04 to 0.17 between 2008 and 2020 (Table 4).

XSL•FO
RenderX

Table . Clinical symptoms of Japanese encephalitis cases in Taiwan from 2008 to 2020.

Variable		Year												
	Total	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Fever, n	280	16	18	30	18	26	15	14	28	21	21	33	20	20
Headache, n	136	13	5	19	13	15	8	7	9	9	8	14	6	10
Stiff necks, n	63	1	6	3	6	6	3	5	3	5	7	11	3	4
Muscle cramps, n	35	7	2	5	—[a]	—	3	2	6	2	3	2	—	3
Vomiting, n	49	4	2	11	4	5	3	1	2	1	3	8	4	1
Speaking difficulty, n	9	—	2	—	3	4	—	—	—	—	—	—	—	—
Coma, n	30	4	5	12	7	2	—	—	—	—	—	—	—	—
Unconscious, n	162	5	6	21	11	9	11	9	13	17	14	20	12	14
Encephalitis, n	3	—	—	—	1	2	—	—	—	—	—	—	—	—
Paralysis, n	1	—	—	—	1	—	—	—	—	—	—	—	—	—
Psychological symptoms, n	59	—	—	—	—	9	—	5	11	5	5	15	6	3
Paralyzed limbs, n	3	—	—	—	—	3	—	—	—	—	—	—	—	—
Diarrhea, n	2	—	—	—	—	2	—	—	—	—	—	—	—	—
Muscular ache, n	2	—	—	—	—	2	—	—	—	—	—	—	—	—
Encephalitic stimulation symptoms, n	10	—	—	—	—	—	—	1	3	2	1	2	—	1
Dystonia, n	16	—	—	—	—	—	—	—	4	4	2	1	2	3
Meningitis symptoms, n	6	—	—	—	—	—	—	—	—	—	—	—	—	3

XSL•FO
RenderX

Variable	Year												
	Total	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Aseptic meningitis, n	—	—	—	—	—	—	—	—	—	—	—	—	—	3

[a]Not applicable.

Table . The mortality rate per 1,000,000 people and the characteristics of deceased cases with Japanese encephalitis from 2008 to 2020[a] in Taiwan.

Variable	2008 (n=2)	2009 (n=2)	2010 (n=2)	2011 (n=1)	2013 (n=1)	2014 (n=3)	2015 (n=1)	2016 (n=4)	2017 (n=3)	2018 (n=2)	2019 (n=3)
Mortality rate per 1,000,000 people	0.09	0.09	0.09	0.04	0.04	0.13	0.04	0.17	0.13	0.08	0.13
Sex, n											
Male	1	2	0	1	0	3	0	3	2	2	1
Female	1	0	2	0	1	0	1	1	1	0	2
Age group, n											
<20	—[b]	—	—	—	—	—	—	—	—	—	—
20 - 39	—	—	1	—	—	1	—	—	—	—	—
40 - 59	1	1	1	—	1	2	1	3	1	—	3
≥60	1	1	—	1	—	—	—	1	2	2	—

[a]There were no reported deaths due to Japanese encephalitis in 2012 and 2020.

[b]Not applicable

Regression Analysis

Linear regression analysis showed that air pollution factors were associated with JE cases: SO_2 concentration (β=2.184, SE 0.922; *P*=.02) and O_3 concentration (β=−0.157, SE 0.063; *P*=.01) (Table 5). Linear regression analysis showed that climatic factors were associated with JE cases (R^2=0.344, $F_{6,137}$=11.980, *P*<.001) (Table 6).

The incidence of confirmed cases of JE per million people was 0.69 - 1.57 (Figure 2A); 1.02 - 1.89 in males (Figure 2B); 1.25 - 3.27 in the 40 - 59 years age group (Figure 2C); 0.55 - 1.25 in summer (Figure 2D); and 0 - 12.2 in the northern region (Figure 2E), during 2008 - 2020. The national coverage of JE vaccine in Taiwan is shown in Figure 3.

XSL•FO
RenderX

Table . Association between air pollution factors and Japanese encephalitis cases examined through multiple linear regression analysis. R^2=0.417; *F* test (*df*) –> 6.824 (9, 86); *P* <.001; n=96.

Variables	Nonstandardization coefficient		*P* value
	β value	Standard error	
TSP[a] (μg/m^3)	−0.028	0.055	.61
PM[b] 2.5 (μg/m^3)	0.022	0.177	.90
PM 10 (μg/m^3)	0.007	0.124	.95
SO_2 (ppb)	2.184	0.922	.02
CO (ppm)	−17.244	19.092	.37
NO_2 (ppb)	−0.342	0.564	.55
O_3 (ppb)	−0.157	0.063	.01
Dustfall volume (tonne/km^2/month)	0.029	0.509	.96
Lead (μg/m^3)	43.751	72.026	.55

[a]TSP: total suspended particulate.

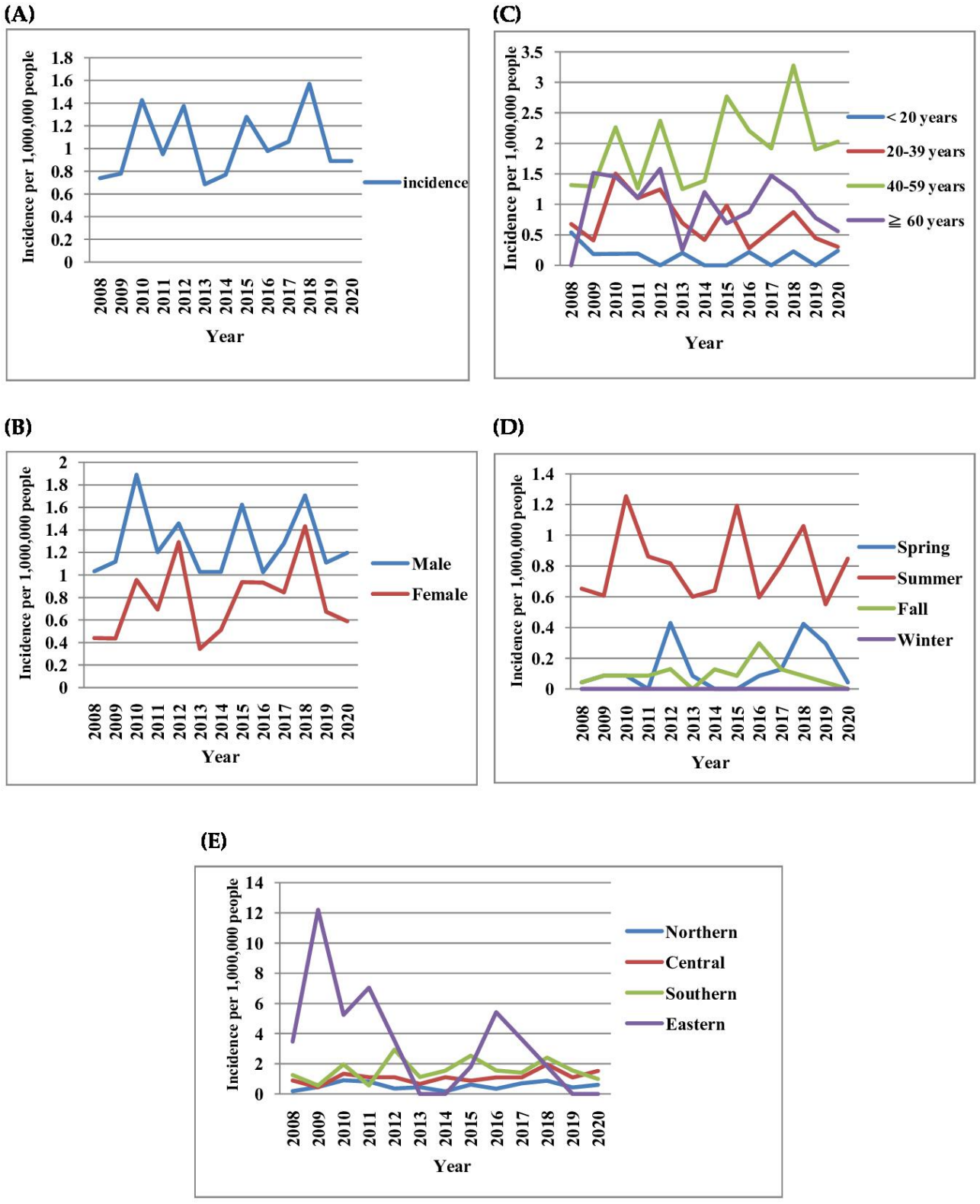
[b]PM: particulate matter.

Table . Relationship between climate factors and Japanese encephalitis cases analyzed using multiple linear regression analysis. R^2=0.344; *F* test (*df*) –> 11.980 (6, 137); *P*<.001; n=144.

Variables	Nonstandardization coefficient		*P* value
	β value	Standard error	
Temperature (˚C)	−0.025	0.192	.90
Precipitation (mm)	−0.004	0.004	.30
Relative humidity (%)	0.380	0.163	.02
Mean pressure (hPa)	−0.247	0.155	.12
Number of days with precipitation ≥0.1 mm (day)	−0.017	0.199	.93
Sunshine duration (h)	0.029	0.016	.08

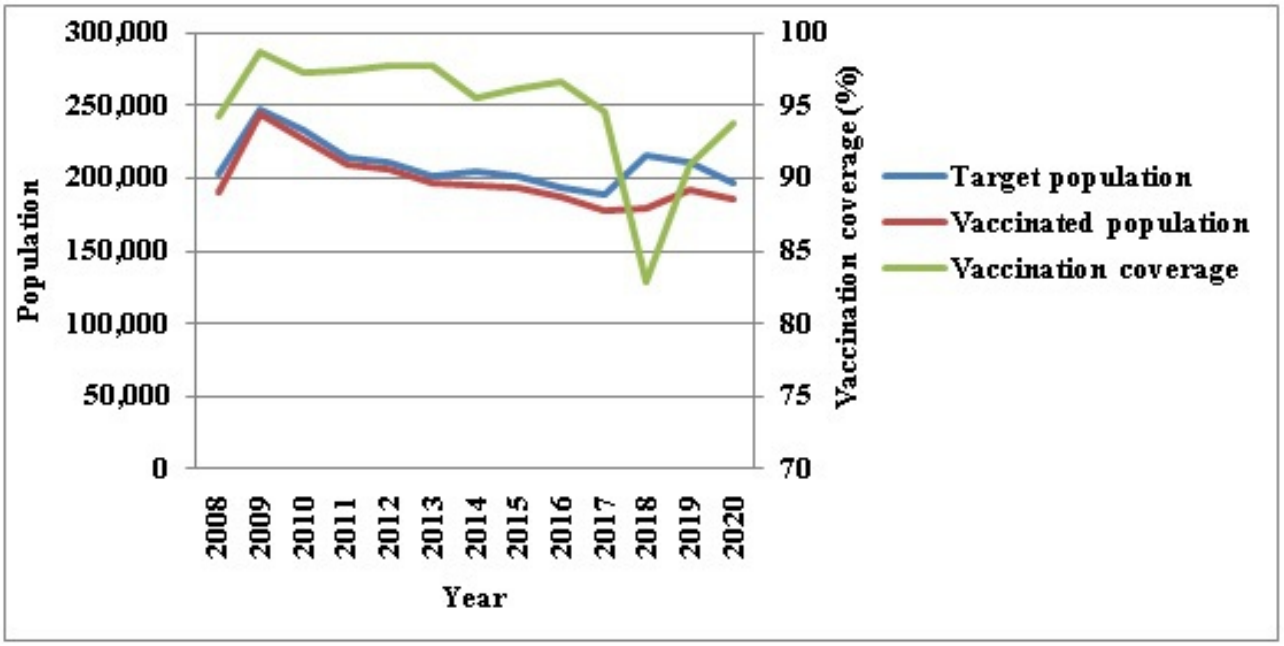
XSL•FO
RenderX

Figure 2. Incidence of confirmed JEV among patients in Taiwan according to (**A**) population, (**B**) gender, (**C**) age, (**D**) season, and (**E**) region of residence by year from 2008 to 2020.



XSL•FO
RenderX

Figure 3. National target population, vaccinated population, and vaccination coverage for Japanese encephalitis in Taiwan from 2008 to 2020.



Discussion

Principal Findings

JE is the most prevalent cause of viral encephalitis in Asia, with transmission occurring through mosquito bites. It is endemic in most of South and Southeast Asia, but the number of cases can vary significantly between regions [34]. The study yielded several significant findings. With regard to epidemiological characteristics, the data indicated that men, middle-aged individuals, those residing in summer, and residents of southern Taiwan are at greater risk. The primary clinical symptoms observed were fever, headache, and unconsciousness. An increase in the concentration of the air pollutant SO_2 and the relative humidity percentage, when considered as climatic factors, was found to correlate with an increase in the number of JE cases.

Epidemiology examines the epidemiological characteristics associated with disease transmission, presentation, and outcome. This discipline has consistently been influenced by policy interventions and disease prevention measures [35,36]. From 2008 to 2020, a total of 313 cases of JE were confirmed in Taiwan. Of these cases, 1.3% were imported, while 98.7% were indigenous, indicating that JE is endemic in Taiwan. The high incidence rate among men is consistent with previous studies [37-39]. It can be postulated that the underlying cause may be differences in personal hygiene habits. Furthermore, most cases occur in individuals aged 40-59 years. Patients in this age group exhibited a higher incidence compared to those younger than 20 years (1.94 vs 0.73), which is consistent with previous studies [13,40,41]. The government of Taiwan can use the similarities and differences in epidemiological characteristics as a foundation for developing and implementing its epidemic prevention policy or strategy.

Among all epidemiological characteristics, seasonal variation is the most susceptible to the distribution of all confirmed cases. In essence, seasonal variation is sufficient to affect the significant increase or decrease of confirmed cases, thereby indicating that the disease in South Asia, Southeast Asian countries, or even Taiwan is endemic [42]. The disease in question typically manifests during the summer months in Taiwan, as evidenced by the findings of previous studies [43-46]. According to the Central Weather Administration of Taiwan, the average summer temperature in Taiwan is 26.4 °C [47]. Consequently, the findings indicate that temperature is a pivotal factor influencing the vector competence and survival of infected mosquitoes. Moreover, southern Taiwan is situated in rural and suburban areas. A previous study indicated that in rural and suburban areas, traditional rice farming and intensive pig breeding provide an ideal environment for both mosquito development and the transmission of JEV among humans. Our findings align with those of previous studies [48-50]. In other words, there are always traces of the JEV in Taiwan that present a threat to public health and increase the clinical medical burden. At this juncture, it is imperative to address these significant challenges head-on, employing strategies to prevent and control them. It is imperative that the Taiwanese government's health department proactively propose local prevention strategies and implement effective surveillance operations to control the epidemic and reduce the number of cases, thereby eliminating the health threat.

JEV is spread by mosquitoes within an enzootic cycle that includes reservoirs among pigs and birds, with humans serving as unintentional dead-end hosts. In the past decade, both serological and genetic evidence from previous studies suggest the presence of JEV in the local fauna [51]. As in the previous study, workplaces of cases in Taiwan are commonly found in pig farms, pigeon farms, and paddy fields [52,53]. This study recommends that Taiwan's government health departments should increase publicity about measures to prevent JE by raising pigs and pigeons in rural areas.

A previous study demonstrated that while the impacts of the COVID-19 lockdown on various sectors, including the economy,

XSL•FO
RenderX

research, travel, education, and sports, were clear, the effects on the occurrence of vector-borne diseases (VBDs) were less apparent [54]. The number of cases reported was compared with the number of cases predicted for each year from 2015 to 2020 for the VBDs (JE) under study to infer whether the COVID-19 lockdown had any impact on their prevalence in Taiwan. The actual number of cases was 21, both before and after the lockdown (2019 and 2020). The predicted number of cases for 2020, based on the trend of the previous 5 years (2015 - 2019), was also close to the actual number of cases. This suggests that the JEV has been consistently present in the local area of Taiwan for a long time. During the COVID-19 period, the number of JE cases did not fluctuate significantly. The analysis suggests that the lockdown did not have a significant impact on the incidence of VBDs such as JE.

In the majority of cases, patients infected with JE do not exhibit any clinical symptoms [55,56]. Consequently, the clinical symptoms observed in JE cases in Taiwan are consistent with those previously documented in other studies [57,58]. It is challenging to differentiate between JEV and other pathogens that can cause encephalitis and meningitis, including enterovirus, mumps virus, mycoplasma, herpes virus, and other viruses. This necessitates the use of viral culture or serological examination [59]. Therefore, this study proposes that the current understanding of JE serum complex flavivirus cross-reactivity, which may result in variable clinical outcomes, could inform future preventive and therapeutic interventions.

During the investigation period of this study, Taiwan exhibited a consistently low incidence rate of JE and its associated mortality from 2008 to 2020. The average incidence and mortality rates were 1.03 cases per 1,000,000 people and 0.09 cases per 1,000,000 people, respectively. A total of 10 cases of JE were reported in children. Consequently, the study indicates that the JE vaccination offered a moderate level of protection among children in JE-endemic Taiwan, a finding consistent with previous studies [60]. The implementation of the JE immunization program has been a pivotal factor in controlling the spread of JE. It is necessary to sustain a high vaccination coverage rate for JE and reinforce the disease surveillance system to ensure the efficacious control and eventual eradication of JE [61].

The prospective effects of climate change on public health are a rapidly expanding field of investigation. This encompasses the immediate consequences of more intense heat waves and declining food security, in addition to the indirect influences on the prevalence of infectious diseases [62-66]. Research has demonstrated that the proximity of collection sites to human dwellings (adjusted odds ratio [AOR] 2.02, *P*=.009) and a relative humidity exceeding 80% (AOR 2.40, *P*=.001) are significant independent risk factors for the transmission of JE viruses [67]. This finding is consistent with the present study. The results of this study indicate a positive correlation between relative humidity and JE cases (β=0.380, *P*=.02) through multiple linear regression analysis. The findings of this study indicate that climate change is a significant factor influencing the incidence of infectious diseases. It is therefore recommended that public health and epidemic prevention experts give this issue their attention, with the implementation of early planning measures to reduce the risk of disease.

Air pollution represents one of the most significant challenges of our era, affecting not only climate change but also public and individual health through increased morbidity and mortality [68]. A multitude of pollutants are significant contributors to human disease [69]. A previous study indicated that a comparison of the combined anthropogenic and environmental risk factors of major mosquito-borne diseases, specifically JE in Thailand, revealed that higher SO_2 surface concentrations were negatively associated with disease case counts [70]. However, the results of the linear regression analysis indicated that there was a positive association between SO_2 concentration and JE (β=2.184, *P*=.02). The concentration of O_3 was found to be negatively correlated with JE (β=−0.157, *P*=.01). To the best of our knowledge, this is the first study to demonstrate that the number of cases of JE increased as the concentration of the air pollutant SO_2 increased, and that there was a positive correlation. The study suggests that the high concentration of SO_2 in summer, which is a season prone to JE, may be a contributing factor. It was hypothesized that there was a correlation between these air pollutants and the disease. Consequently, this study proposes that Taiwan's official policy should implement a rigorous monitoring program to assess local climate factors and air pollution concentration changes. In the event of minor fluctuations or significant fluctuations in environmental data, it is imperative that the media and the public be promptly informed. This will enable individuals to take prompt action to mitigate environmental pressures and the threat to public health.

Limitations

One limitation of the present study is that the data provided by the Taiwan National Infectious Disease Statistics System lack information about the genotypes or strains of the JEV isolated. Consequently, the type of JEV strain that spread to Taiwan and the affinity between virus strains in Taiwan and other countries were not analyzed in this study. Nevertheless, the advantage of this study was the access to the diverse data provided by Taiwan's public sector on its web-based platform (including the initial version of the platform) and the evaluation of the impact of the COVID-19 pandemic on the epidemiological features of typhoid and paratyphoid. The information from existing public network platforms in Taiwan is both timely and accurate.

Conclusions

This study is the first to report on the epidemiological characteristics, clinical symptomatology, climatic factors, and air pollutants associated with JE cases in Taiwan between 2008 and 2020. In 2018, the incidence rate of JE was the highest (1.57 per million people). In recent years, the risk of local cases has increased at a rapid pace, resulting in a significant burden of disease, public health challenges, and epidemic prevention. The distribution of JE in Taiwan is regional in nature. The risk of JE among patients residing in different areas increases with age. The clinical manifestations of JE are variable, and the primary areas of endemicity are pig farms, pigeon farms, and paddy

XSL•FO
RenderX

fields. The JE vaccine coverage rate in Taiwan is high, with an average of 94.9%. The concentration of certain environmental factors (SO_2, O_3) may influence the increase or decrease in JE cases. The identification of crucial data will facilitate future surveillance and research activities in Taiwan.

Conflicts of Interest

None declared.

References

1. Key facts of japanese encephalitis. WHO. 2022. URL: https://www.who.int/news-room/fact-sheets/detail/japanese-encephalitis [accessed 2025-01-10]
2. Japanese encephalitis: symptoms, diagnosis, and treatment. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/japanese-encephalitis/symptoms-diagnosis-treatment/?CDC_AAref_Val=https://www.cdc.gov/japaneseencephalitis/symptoms/index.html [accessed 2025-01-10]
3. Wilder-Smith A, Halstead SB. Japanese encephalitis: update on vaccines and vaccine recommendations. Curr Opin Infect Dis 2010 Oct;23(5):426-431. [doi: 10.1097/QCO.0b013e32833c1d01] [Medline: 20581670]
4. Leinikki P. Japanese encephalitis. In: Zuckerman AJ, Banatvala JE, Pattison JR, editors. Principles and Practice of Clinical Virology, 5th edition: Wiley; 2004:459-466.
5. Lindenbach B, Rice C. Flaviviridae: the viruses and their replication. In: Fields BN, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, et al, editors. Fields Virology, 4th edition: Lippincott Williams & Wilkins; 2001:991-1041.
6. Ghosh D, Basu A. Japanese encephalitis-a pathological and clinical perspective. PLoS Negl Trop Dis 2009 Sep 29;3(9):e437. [doi: 10.1371/journal.pntd.0000437] [Medline: 19787040]
7. Hong C, Zhang Q, Zhang Y, et al. Impacts of climate change on future air quality and human health in China. Proc Natl Acad Sci USA 2019 Aug 27;116(35):17193-17200. [doi: 10.1073/pnas.1812881116] [Medline: 31405979]
8. Zhang Y, Wang H, Du J, et al. Population genetic structure of Culex tritaeniorhynchus in different types of climatic zones in China. BMC Genomics 2024;25(1):673. [doi: 10.1186/s12864-024-10589-4]
9. Centers for Disease Control and Prevention (CDC). Japanese encephalitis surveillance and immunization—Asia and Western pacific, 2012. MMWR Morb Mortal Wkly Rep 2013;62:658-662.
10. Heffelfinger JD, Li X, Batmunkh N, et al. Japanese encephalitis surveillance and immunization—Asia and Western Pacific Regions, 2016. Wkly Epidemiol Rep 2017;92:321-332. [doi: 10.15585/mmwr.mm6622a3]
11. Le Flohic G, Porphyre V, Barbazan P, Gonzalez JP. Review of climate, landscape, and viral genetics as drivers of the Japanese encephalitis virus ecology. PLoS Negl Trop Dis 2013;7(9):e2208. [doi: 10.1371/journal.pntd.0002208] [Medline: 24069463]
12. Fischer M, Hills S, Staples E, Johnson B, Yaich M, Solomon T. Japanese encephalitis prevention and control: advances, challenges, and new initiatives. In: Scheld WM, Hammer SM, Hughes JM, editors. Emerging Infections 8: ASM Press; 2008:93-124.
13. Chang YK, Chang HL, Wu HS, Chen KT. Epidemiological features of Japanese encephalitis in Taiwan from 2000 to 2014. Am J Trop Med Hyg 2017 Feb 8;96(2):382-388. [doi: 10.4269/ajtmh.16-0330] [Medline: 27821699]
14. Average GDP per person (USD) [Article in Mandarin]. National Statistics, ROC (Taiwan). 2022. URL: https://www.stat.gov.tw/point.asp?index=1 [accessed 2025-1-10]
15. Su CL, Yang CF, Teng HJ, et al. Molecular epidemiology of Japanese encephalitis virus in mosquitoes in Taiwan during 2005-2012. PLoS Negl Trop Dis 2014 Oct;8(10):e3122. [doi: 10.1371/journal.pntd.0003122] [Medline: 25275652]
16. Wu YH, Chen JM, Su CY. Occurrence of old age Japanese encephalitis: current situation in Taiwan. J Formos Med Assoc 2017 Dec;116(12):915-916. [doi: 10.1016/j.jfma.2017.04.010] [Medline: 28528139]
17. Hsu JY, Hung CC, Tsou TP, Chen WC. Epidemiology and risk factors of Japanese encephalitis in Taiwan, 2010-2022. PLoS Negl Trop Dis 2023 Oct;17(10):e0011421. [doi: 10.1371/journal.pntd.0011421] [Medline: 37782654]
18. Lin CL, Chang HL, Lin CY, Chen KT. Seasonal patterns of Japanese encephalitis and associated meteorological factors in Taiwan. Int J Environ Res Public Health 2017 Oct 29;14(11):1317. [doi: 10.3390/ijerph14111317] [Medline: 29109371]
19. Ho YC, Su BH, Su HJ, et al. The association between the incidence of mumps and meteorological parameters in Taiwan. Hum Vaccin Immunother 2015;11(6):1406-1412. [doi: 10.1080/21645515.2015.1029687] [Medline: 25891825]
20. Taiwan National Infectious Disease Statistics System [Website in Mandarin]. URL: http://nidss.cdc.gov.tw [accessed 2025-01-10]
21. Solomon T, Thao TT, Lewthwaite P, et al. A cohort study to assess the new WHO Japanese encephalitis surveillance standards. Bull World Health Organ 2008 Mar;86(3):178-186. [doi: 10.2471/blt.07.043307] [Medline: 18368204]
22. Hsu LC, Chen YJ, Hsu FK, et al. The incidence of Japanese encephalitis in Taiwan--a population-based study. PLoS Negl Trop Dis 2014 Jul;8(7):e3030. [doi: 10.1371/journal.pntd.0003030] [Medline: 25058573]

XSL•FO
RenderX

23. Ravi V, Robinson JS, Russell BJ, et al. Evaluation of IgM antibody capture enzyme-linked immunosorbent assay kits for detection of IgM against Japanese encephalitis virus in cerebrospinal fluid samples. Am J Trop Med Hyg 2009 Dec;81(6):1144-1150. [doi: 10.4269/ajtmh.2009.09-0144] [Medline: 19996450]
24. Clark CH, Casals J. Techniques for hemagglutination inhibition with arthropod viruses. Am J Trop Med Hyg 1958;7:561-573. [doi: 10.4269/ajtmh.1958.7.561]
25. Manual for the Laboratory Diagnosis of Japanese Encephalitis Virus Infection: World Health Organization; 2007.
26. Johnson BW, Goodman CH, Jee Y, Featherstone DA. Differential diagnosis of Japanese encephalitis virus infections with the Inbios JE Detect™ and DEN Detect™ MAC-ELISA Kits. Am Soc Trop Med Hyg 2016;94(4):820-828. [doi: 10.4269/ajtmh.15-0631]
27. Sapkal GN, Wairagkar NS, Ayachit VM, Bondre VP, Gore MM. Detection and isolation of Japanese encephalitis virus from blood clots collected during the acute phase of infection. Am J Trop Med Hyg 2007 Dec;77(6):1139-1145. [Medline: 18165537]
28. Mathur A, Kumar R, Sharma S, Kulshreshtha R, Kumar A, Chaturvedi UC. Rapid diagnosis of Japanese encephalitis by immunofluorescent examination of cerebrospinal fluid. Indian J Med Res 1990 Jan;91:1-4. [Medline: 2160913]
29. Environmental statistics inquiry network [Article in Mandarin]. Ministry of Environment. URL: https://statis.moenv.gov.tw/epanet [accessed 2025-01-13]
30. Monthly weather [Article in Mandarin]. Central Weather Administration. URL: https://www.cwa.gov.tw/V8/C/C/Statistics/monthlydata.html [accessed 2025-01-13]
31. Holland KM, Jones C, Vivolo-Kantor AM, et al. Trends in US Emergency Department visits for mental health, overdose, and violence outcomes before and during the COVID-19 pandemic. JAMA Psychiatry 2021 Apr 1;78(4):372-379. [doi: 10.1001/jamapsychiatry.2020.4402] [Medline: 33533876]
32. Pan CY, Liu WL, Su MP, et al. Epidemiological analysis of the Kaohsiung city strategy for dengue fever quarantine and epidemic prevention. BMC Infect Dis 2020 May 15;20(1):347. [doi: 10.1186/s12879-020-4942-y] [Medline: 32414340]
33. Communicable Disease Control Act. Laws & Regulations Database of The Republic of China (Taiwan). 2023. URL: https://law.moj.gov.tw/ENG/LawClass/LawAll.aspx?pcode=L0050001 [accessed 2025-01-10]
34. Lopalco PL, Biasio LR. Japanese encephalitis can be devastating. Ann Ig 2024;36(3):370-375. [doi: 10.7416/ai.2024.2616] [Medline: 38436079]
35. Gordis L. The dynamics of disease transmission. In: Gordis L, editor. Epidemiology, 3rd edition 2004:1934-2015.
36. Stanley SL Jr. Amoebiasis. Lancet 2003 Mar 22;361(9362):1025-1034. [doi: 10.1016/S0140-6736(03)12830-9] [Medline: 12660071]
37. Srivastava N, Deval H, Mittal M, et al. Extent of disability among paediatric Japanese encephalitis survivors and predictors of poor outcome: a retrospective cohort study in North India. BMJ Open 2022 Oct 31;12(10):e060795. [doi: 10.1136/bmjopen-2022-060795] [Medline: 36316071]
38. Pommier JD, Gorman C, Crabol Y, et al. Childhood encephalitis in the Greater Mekong region (the SouthEast Asia Encephalitis Project): a multicentre prospective study. Lancet Glob Health 2022 Jul;10(7):e989-e1002. [doi: 10.1016/S2214-109X(22)00174-7] [Medline: 35714649]
39. Udeze AO, Odebisi-Omokanye MB. Sero-evidence of silent Japanese encephalitis virus infection among inhabitants of Ilorin, North-central Nigeria: a call for active surveillance. J Immunoassay Immunochem 2022 May 4;43(3):250-258. [doi: 10.1080/15321819.2021.1993897] [Medline: 34809529]
40. Wu D, Chen X, Liu W, et al. Emergence of Japanese encephalitis among adults 40 years of age or older in northern China: epidemiological and clinical characteristics. Transbound Emerg Dis 2021 Nov;68(6):3415-3423. [doi: 10.1111/tbed.13945] [Medline: 33283432]
41. Kwak BO, Hong YJ, Kim DH. Changes in age-specific seroprevalence of Japanese encephalitis virus and impact of Japanese encephalitis vaccine in Korea. Clin Exp Pediatr 2022 Mar;65(3):108-114. [doi: 10.3345/cep.2020.01984] [Medline: 34592804]
42. Lin FH, Chen BC, Chou YC, et al. The epidemiology of Entamoeba histolytica infection and its associated risk factors among domestic and imported patients in Taiwan during the 2011-2020 period. Medicina (Kaunas) 2022 Jun 17;58(6):820. [doi: 10.3390/medicina58060820] [Medline: 35744083]
43. Lee PI, Huang YC, Hwang KP, et al. Recommendations for the use of Japanese encephalitis vaccines. Pediatr Neonatol 2020 Feb;61(1):3-8. [doi: 10.1016/j.pedneo.2019.11.009] [Medline: 31870559]
44. Griffith MM, Fukusumi M, Kobayashi Y, et al. Epidemiology of vaccine-preventable diseases in Japan: considerations for pre-travel advice for the 2019 Rugby World Cup and 2020 Summer Olympic and Paralympic Games. Western Pac Surveill Response J 2018;9(2):26-33. [doi: 10.5365/wpsar.2017.8.4.002] [Medline: 30057855]
45. Folly AJ, Dorey-Robinson D, Hernández-Triana LM, et al. Temperate conditions restrict Japanese encephalitis virus infection to the mid-gut and prevents systemic dissemination in Culex pipiens mosquitoes. Sci Rep 2021 Mar 17;11(1):6133. [doi: 10.1038/s41598-021-85411-2] [Medline: 33731761]
46. Lo SH, Tang HJ, Lee SSJ, et al. Determining the clinical characteristics and prognostic factors for the outcomes of Japanese encephalitis in adults: a multicenter study from southern Taiwan. J Microbiol Immunol Infect 2019 Dec;52(6):893-901. [doi: 10.1016/j.jmii.2019.08.010] [Medline: 31628089]

XSL•FO
RenderX

47. Monthly average climate [Article in Mandarin]. Central Weather Administration. URL: https://www.cwa.gov.tw/V8/C/C/Statistics/monthlymean.html [accessed 2025-01-10]
48. Tian HY, Bi P, Cazelles B, et al. How environmental conditions impact mosquito ecology and Japanese encephalitis: an eco-epidemiological approach. Environ Int 2015 Jun;79:17-24. [doi: 10.1016/j.envint.2015.03.002] [Medline: 25771078]
49. Di Francesco J, Choeung R, Peng B, et al. Comparison of the dynamics of Japanese encephalitis virus circulation in sentinel pigs between a rural and a peri-urban setting in Cambodia. PLoS Negl Trop Dis 2018 Aug;12(8):e0006644. [doi: 10.1371/journal.pntd.0006644] [Medline: 30138381]
50. Lindahl JF, Ståhl K, Chirico J, Boqvist S, Thu HTV, Magnusson U. Circulation of Japanese encephalitis virus in pigs and mosquito vectors within Can Tho city, Vietnam. PLoS Negl Trop Dis 2013;7(4):e2153. [doi: 10.1371/journal.pntd.0002153] [Medline: 23593520]
51. Lim MJ, Loh ZY, Yeo HL, et al. Isolation and genetic characterization of Japanese encephalitis virus two decades after its elimination in Singapore. Viruses 2022 Nov 28;14(12):2662. [doi: 10.3390/v14122662] [Medline: 36560666]
52. Williams CR, Webb CE, Higgs S, van den Hurk AF. Japanese encephalitis virus emergence in Australia: public health importance and implications for future surveillance. Vector Borne Zoonotic Dis 2022 Nov;22(11):529-534. [doi: 10.1089/vbz.2022.0037] [Medline: 36354964]
53. Walsh MG, Pattanaik A, Vyas N, et al. High-risk landscapes of Japanese encephalitis virus outbreaks in India converge on wetlands, rain-fed agriculture, wild Ardeidae, and domestic pigs and chickens. Int J Epidemiol 2022 Oct 13;51(5):1408-1418. [doi: 10.1093/ije/dyac050] [Medline: 35355081]
54. Mayilsamy M, Vijayakumar A, Veeramanoharan R, Rajaiah P, Balakrishnan V, Kumar A. Impact of COVID-19 lockdown during 2020 on the occurrence of vector-borne diseases in India. J Vector Borne Dis 2023;60(2):207. [doi: 10.4103/0972-9062.364762]
55. Sahu RC, Suthar T, Pathak A, Jain K. Interventions for the prevention and treatment of Japanese encephalitis. Curr Infect Dis Rep 2022 Nov;24(11):189-204. [doi: 10.1007/s11908-022-00786-1]
56. Turtle L, Solomon T. Japanese encephalitis - the prospects for new treatments. Nat Rev Neurol 2018 Apr 26;14(5):298-313. [doi: 10.1038/nrneurol.2018.30] [Medline: 29697099]
57. Pham D, Howard-Jones AR, Hueston L, et al. Emergence of Japanese encephalitis in Australia: a diagnostic perspective. Pathology (Phila) 2022 Oct;54(6):669-677. [doi: 10.1016/j.pathol.2022.07.001] [Medline: 35995617]
58. Guo H, Sun L, Shen X, Hu W. A retrospective study of the clinical characteristics of Japanese encephalitis in adults. J Integr Neurosci 2022 Jul 18;21(5):125. [doi: 10.31083/j.jin2105125] [Medline: 36137966]
59. Wang LP, Yuan Y, Liu YL, et al. Etiological and epidemiological features of acute meningitis or encephalitis in China: a nationwide active surveillance study. Lancet Reg Health West Pac 2022 Mar;20:100361. [doi: 10.1016/j.lanwpc.2021.100361] [Medline: 35036977]
60. Khare B, Kuhn RJ. The Japanese encephalitis antigenic complex viruses: from structure to immunity. Viruses 2022 Oct 8;14(10):2213. [doi: 10.3390/v14102213] [Medline: 36298768]
61. Tandale BV, Khude PM, Deshmukh PS, et al. Effectiveness of Japanese encephalitis vaccination among children in central India. J Med Virol 2023 Jan;95(1):e28399. [doi: 10.1002/jmv.28399] [Medline: 36512338]
62. Zhang WX, Zhao S, Pan C, et al. Mass immunisation to eradicate Japanese encephalitis: real-world evidence from Guizhou Province in 2005-2021. J Virus Erad 2024 Mar;10(1):100366. [doi: 10.1016/j.jve.2024.100366] [Medline: 38586471]
63. Mora C, Dousset B, Caldwell IR, et al. Global risk of deadly heat. Nat Clim Chang 2017 Jul;7(7):501-506. [doi: 10.1038/nclimate3322]
64. Mazdiyasni O, AghaKouchak A, Davis SJ, et al. Increasing probability of mortality during Indian heat waves. Sci Adv 2017 Jun;3(6):e1700066. [doi: 10.1126/sciadv.1700066] [Medline: 28630921]
65. Schmidhuber J, Tubiello FN. Global food security under climate change. Proc Natl Acad Sci USA 2007 Dec 11;104(50):19703-19708. [doi: 10.1073/pnas.0701976104] [Medline: 18077404]
66. Springmann M, Mason-D'Croz D, Robinson S, et al. Global and regional health effects of future food production under climate change: a modelling study. Lancet 2016 May 7;387(10031):1937-1946. [doi: 10.1016/S0140-6736(15)01156-3] [Medline: 26947322]
67. Wu X, Lu Y, Zhou S, Chen L, Xu B. Impact of climate change on human infectious diseases: empirical evidence and human adaptation. Environ Int 2016 Jan;86:14-23. [doi: 10.1016/j.envint.2015.09.007] [Medline: 26479830]
68. Diptyanusa A, Herini ES, Indarjulianto S, Satoto TBT. Estimation of Japanese encephalitis virus infection prevalence in mosquitoes and bats through nationwide sentinel surveillance in Indonesia. PLoS One 2022;17(10):e0275647. [doi: 10.1371/journal.pone.0275647] [Medline: 36223381]
69. Manisalidis I, Stavropoulou E, Stavropoulos A, Bezirtzoglou E. Environmental and health impacts of air pollution: a review. Front Public Health 2020;8:14. [doi: 10.3389/fpubh.2020.00014] [Medline: 32154200]
70. Tewari P, Ma P, Gan G, et al. Non-linear associations between meteorological factors, ambient air pollutants and major mosquito-borne diseases in Thailand. PLoS Negl Trop Dis 2023 Dec;17(12):e0011763. [doi: 10.1371/journal.pntd.0011763] [Medline: 38150471]

XSL•FO
RenderX

Abbreviations

AOR: adjusted odds ratio
CSF: cerebrospinal fluid
DF: dengue fever
HI: hemagglutination inhibition
JE: Japanese encephalitis
JEV: Japanese encephalitis virus
PM: particulate matter
TCDC: Taiwan Centers for Diseases Control
VBD: vector-borne disease

Edited by C Argyropoulos; submitted 11.06.24; peer-reviewed by D Valvi, RJPS Guimaraes, YL Cheong; revised version received 21.11.24; accepted 23.11.24; published 28.01.25.

Please cite as:
Lin FH, Chou YC, Hsieh CJ, Yu CP
Epidemiological Features, Clinical Symptoms, and Environmental Risk Factors for Notifiable Japanese Encephalitis in Taiwan From 2008 to 2020: Retrospective Study
JMIR Public Health Surveill 2025;11:e63053
URL: https://publichealth.jmir.org/2025/1/e63053
doi:10.2196/63053

© Fu-Huang Lin, Yu-Ching Chou, Chi-Jeng Hsieh, Chia-Peng Yu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Effectiveness of a Health Education Program to Reduce Recurrence of Stroke by Controlling Modifiable Risk Factors in a Specialized Hospital in Bangladesh: Randomized Controlled Trial

Mahabuba Afrin[1], RN, MPH, MSS; K A T M Ehsanul Huq[1], MBBS, DTM, MMS, PhD; Sharif Uddin Khan[2], MBBS, MD; Subir Chandra Das[2], MBBS, FCPS; Mohammad Shah Jahirul Hoque Chowdhury[2], MBBS, MD; Yasuko Fukuoka[3], RN, MSN, PhD; Yasuko Fukushima[1], RN, MSN; Michiko Moriyama[1], RN, MSN, PhD

[1]Graduate School of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan

[2]National Institute of Neurosciences and Hospital, Dhaka, Bangladesh

[3]Department of Nursing, Ube Frontier University, Yamaguchi, Japan

Corresponding Author:
Michiko Moriyama, RN, MSN, PhD
Graduate School of Biomedical and Health Sciences
Hiroshima University
Kasumi 1-2-3 Minami-ku
Hiroshima, 734-8553
Japan
Phone: 81 81 82 257 5365
Fax: 81 81 82 257 5369
Email: morimich@hiroshima-u.ac.jp

Abstract

Background: Health education could be an effective way to increase knowledge regarding behavioral changes to prevent the recurrence of stroke; however, the evidence is ambiguous. A lack of both knowledge and compliance with treatment to control modifiable risk factors and unhealthy lifestyles increases the risk of stroke recurrence.

Objective: This study aimed to evaluate the effectiveness of a health education program among patients with stroke postdischarge and their family caregivers in a tertiary specialized hospital in Bangladesh to reduce stroke recurrence by controlling modifiable risk factors.

Methods: A parallel (1:1), open-label, prospective randomized controlled trial was conducted in Bangladesh. A total of 432 patients with first-time stroke, aged ≥18 years and a modified Rankin Scale (mRS) score of 0-4, were randomly enrolled at the National Institute of Neuroscience & Hospital. We stratified the patients by age and type of stroke and randomly allocated to an intervention group (IG) and a control group (CG). We collected sociodemographic and clinical data by using a structured questionnaire. The IG received (self) management education, including monitoring blood pressure (BP), medication, diet, and exercise for 12 months, and the CG received usual care. The outcomes were (1) recurrence after 28 days of stroke and (2) all adverse events related to stroke after 12-month follow-up.

Results: Of 432 patients (n=216, 50%, in each group), stroke recurrence was observed 14 (6.5%) patients in the IG and 8 (3.7%) patients in the CG, and the difference was not significant (*P*=.19). Death was lower in the IG (n=39, 18.1%, vs n=56, 25.9%) compared to the CG. In Cox regression analysis, the unadjusted model showed death was significant (hazard ratio [HR] 1.531, 95% CI 1.017-2.304; *P*=.04); however, the difference was not significant after adjusting the mRS score at baseline (HR 0.818, 95% CI 0.540-1.238; *P*=.34). Patients' medication adherence significantly improved at 6-month (*P*<.001) and 12-month (*P*=.002) follow-up after the intervention.

Conclusions: This study failed to demonstrate the effectiveness of health education in reducing recurrence, death, and stroke-related adverse events. However, health education enhanced medication adherence. Some causes of death could not be diagnosed due to inadequate health care systems. Further research with diagnostic facilities and a long observation period is necessary to explore the underlying cause of recurrence. The results suggest the importance of structuring acute care management for patients with stroke into the health care system of Bangladesh.

Trial Registration: ClinicalTrials.gov NCT05520034; https://clinicaltrials.gov/ct2/show/NCT05520034

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e72233) doi:10.2196/72233

KEYWORDS

stroke; health education; recurrence; modifiable risk factors

Introduction

The stroke burden is a significant and growing public health concern. Stroke was the second-most fatal and third-most leading cause of disability worldwide in 2022. Its prevalence is rising rapidly in low- and middle-income countries (LMICs) [1]. Every year, more than 13.7 million strokes occur worldwide, which leads to 5.5 million deaths. Although strokes are more likely to affect older adults, there is an increased frequency of strokes among younger people [2,3]. In 2019, the foremost risk factors for stroke were elevated systolic blood pressure (BP), a high BMI, elevated fasting plasma glucose concentrations, and smoking [4]. Studies have indicated that the risk of recurrent stroke is the highest within the first few months after the initial event, and up to 25% of survivors experience a recurrent stroke within 5 years [5]. Recurrent strokes are associated with fatal outcomes, increased disability, and higher mortality rates compared to first-time strokes. It is essential to manage modifiable risk factors for reducing stroke recurrence and improving long-term outcomes for survivors of stroke [5].

In Bangladesh, stroke was the top cause of death per 100,000 population for both males and females and all ages in 2019 [6]. In 2022, Bangladesh had a stroke rate of 79.9%, 15.7%, and 4.6%, which was caused by strokes that involved ischemia, hemorrhage, and subarachnoid hemorrhage, respectively, and affected 1,139,000 people [7]. One study conducted in a tertiary medical university hospital in Bangladesh evaluated the rate of recurrence of stroke and found that the cumulative recurrence rate was 14.7% at 3 months, 15.3% at 6 months, 17.3% at 9 months, and 20% at 1 year. The most frequent age group was >75 years, which represented 44.4% of recurrent stroke cases [8].

Despite the advancement of thrombolytic therapy and acute stroke management, preventing the recurrence of stroke remains a significant challenge. Moreover, the ability of patients to manage complications and avoid underlying diseases, such as hypertension, diabetes, and dyslipidemia, is impacted by substandard primary care and a lack of referral mechanisms. Patients in underdeveloped health care systems are not consistently diagnosed or treated because of a lack of information, infrastructure, resources, shortage of equipment, a lack of diagnostic accuracy, a lack of health education, and budgetary constraints, which prevents them from undergoing periodical medical examinations [9,10].

One systematic review concluded that health education may improve BP targets but not improve other risk factors or reduce recurrent events, and patient education alone might not prevent the risk of recurrent stroke [11]. A study in Bangladesh found that a small sample size of 150 is insufficient for identifying the stroke recurrence frequency. A larger sample size and health education efforts are recommended to reduce recurrence [8]. Another study involving a sample size of 321 in Japan provided self-management health education for 6- and 24-month follow-up. This study reduced the recurrence rate by half after the intervention; however, the difference was not statistically significant for the prevention of stroke recurrence [12].

Since there is no established patient follow-up system and health education offered across hospitals in many high-income countries, this intervention study was conducted to incorporate patient health education following discharge from acute care hospitals. Even though hospitals provide health education at the time of discharge to improve patients' self-motivation at home, patients with stroke are found to lack the knowledge to follow medical advice [13]. As most of the patients are taken care of by family caregivers, their health education can significantly reduce risk factors and complications related to stroke [14].

In Bangladesh, health education for patients with stroke regarding self-management skills at home from the hospital level is limited. The reasons could be sociocultural, economic, and political, as well as poor planning and improper implementation of health policies and programs within the health system. Patients with stroke and their family caregivers should know the importance of controlling modifiable risk factors [8]. Therefore, it is crucial to establish a well-organized patient–disease management education system.

Based on these findings, we hypothesized that a health education program among patients with first-time stroke and their family caregivers could reduce stroke recurrence by controlling modifiable risk factors. We tried to evaluate the effectiveness of a health education program in improving risk factors and reducing the recurrence of stroke and other stroke-related adverse events.

Methods

Study Design and Participants

This was a parallel (1:1), open-label, prospective randomized controlled trial and adhered to the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Multimedia Appendix 1) [15]. This study was conducted at the National Institute of Neuroscience & Hospital (NINS&H) in Dhaka, Bangladesh, between October 2022 and March 2024. The NINS&H is the only state-run referral tertiary care hospital in Bangladesh focused on neurological disorders. Patients with neurological disorders, including strokes, and from different socioreligious groups visit this hospital from all over the country. We enrolled participants who had their first stroke and were in the inpatient department or the emergency department before discharge or who were visiting the outpatient department of the NINS&H. The participants were aged 18 years and above, both men and women, and were willing to participate in this study. We included patients with ischemic (large artery atherosclerosis, cardioembolism, small vessel occlusion [SVO], lacunar subtype, and transient ischemic attack [TIA]) strokes diagnosed by

XSL•FO
RenderX

physicians. In addition, as a sufficient number of patients with hemorrhagic stroke visited the hospital, we included those patients too in order to obtain the required sample size. Enrolling both types of patients in this study helped us understand their clinical presentations, complications, and recovery patterns and also provided comprehensive data for better stroke management strategies and resource allocation in resource-limited settings. We included only patients with a modified Rankin Scale (mRS) score of 0-4. As most of the patients with stroke with mRS score=0-4 in Bangladesh have family caregivers to provide health care support, family caregivers, living with the patients and providing support to them, were also included [16]. We excluded patients who had mRS scores of >4, had mental illness or cognitive impairment (diagnosed cases), had participated in other clinical trials, and had multiorgan failure or terminal disease.

Randomization

Participants were randomly assigned to an intervention group (IG) and a control group (CG) in a 1:1 ratio using a pregenerated allocation list based on block randomization with variable block sizes (2-4). The list was stratified by age group (18-64 vs ≥65 years) and stroke type (1: large artery atherosclerosis; 2: cardioembolism; 3: SVO/lacunar/TIA; and 4: hemorrhage). Statistics for the age of patients with recurrent stroke and the types of strokes were not available at the study site, so we used the study findings as references [17,18]. We used this stratified block randomization technique as recurrence rates varied depending on the age and type of the stroke [19]. This randomization procedure was conducted by an experienced investigator who was not involved in the study intervention or analysis. This design minimized bias by controlling key confounders at this stage.

Sampling Technique and Study Procedures

A convenient sampling technique was used to enroll participants. Neurologists at the hospital and research assistant (RA) nurses identified patients with stroke, then checked the patients' eligibility criteria by referring to their medical records. We used sealed envelopes to keep the serial number of the allocated randomization group. These envelopes were prepared and opened sequentially only after participant enrollment to allocate patients to the respective groups. The RA nurses obtained each participant's sociodemographic information (age, gender, marital status, education, occupation, living status) and clinical data (medical history, comorbidities, treatments, prescribed medications, self-reported medication adherence, types of stroke, mRS score, and risk factors of stroke). Additionally, laboratory data, including glycated hemoglobin (HbA1c), lipid profile, and BP, were obtained by interviewing the participants and from hospital medical records. The medication adherence questionnaire included participants' responses to whether they followed the doctor's prescription for taking medications.

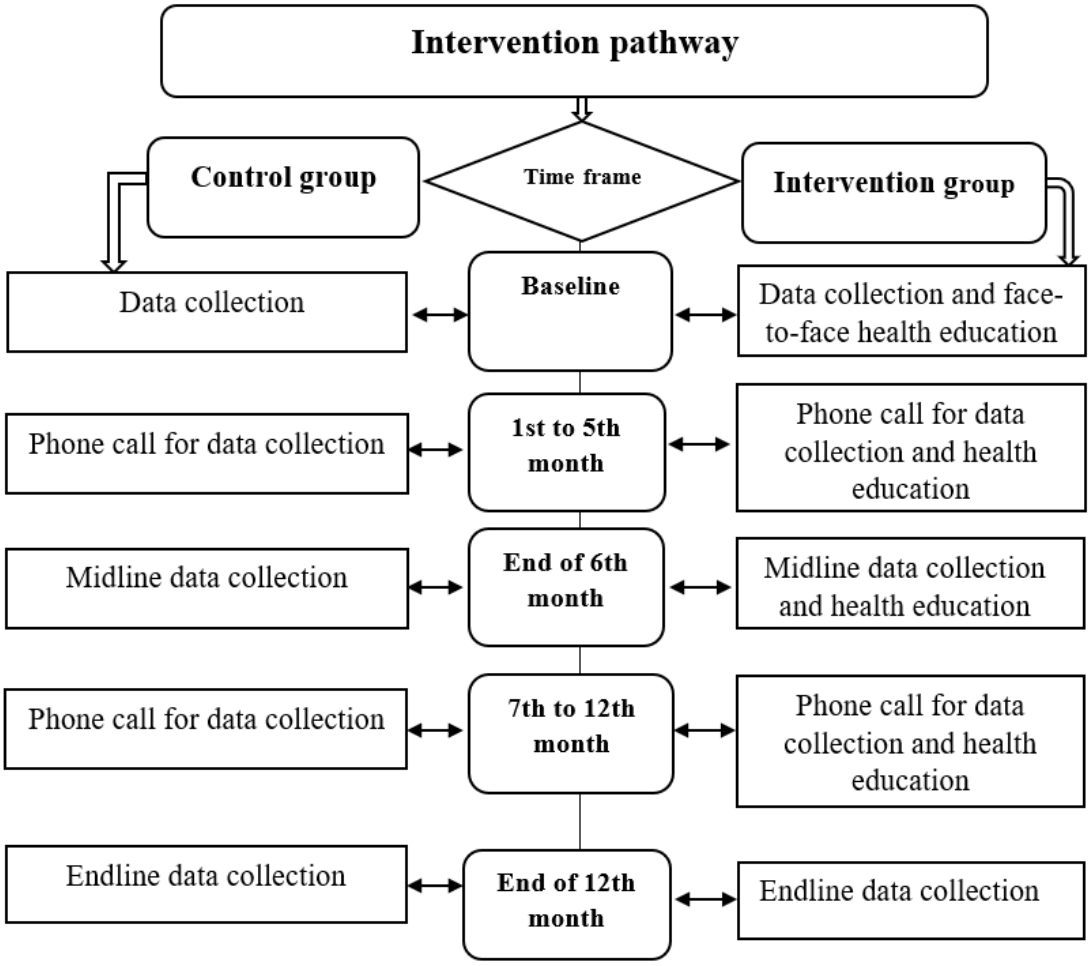
Intervention Group

After enrollment at the hospital, RA nurses conducted a face-to-face health education and video presentation for the patients and their family caregivers for 45 minutes. The RA nurses followed up with the participants through phone calls every month (months 1-3, twice a month; months 4-6, once a month) and sent SMS text messages to remind them to record their daily monitoring in a health education book. At midline after 6 months and at endline after 12 months, the RA nurses called the participants to the hospital for follow-up visits, provided the same health education as baseline, collected data, and conducted laboratory tests. Figure 1 shows the intervention pathway.

If a patient and their family caregiver were unable to visit the hospital for any reason, then the RA nurses provided health education over the telephone and advised them to undergo the laboratory test at a nearby hospital. The RA nurses also visited the participants' homes to provide health education and for sample collection, if needed. If an RA nurse failed to communicate with a patient/caregiver over a phone call for 3 consecutive months, the patient/caregiver was considered a dropout. For participants, 6 months of follow-up visits at the hospital were found to be low due to economic, social, and transportation difficulties, particularly in rural areas. The RA nurses arranged home visits at 12-month follow-up. This significantly increased participant involvement at 12 months compared to 6-month follow-up visits, emphasizing the importance of flexible research designs in LMICs, such as Bangladesh. This study has been described elsewhere [20].

XSL•FO
RenderX

Figure 1. Study activities of the IG (n=216) and the CG (n=216). CG: control group; IG: intervention group.



Health Education Procedures

Trained RA nurses provided health education sessions to participants using a researcher-developed booklet, which was provided to the patients for self-education at home. We delivered 2 health education sessions using a multimedia approach, which included visual aids and interactive discussions, alongside the booklet material. For patients who could not attend intervention sessions, health education was provided at their bedside using the same educational booklet to maintain consistency and ensure accessibility of information.

Components of the Health Education Intervention

Lecture-Based Education

A booklet-based health education session was conducted for about 20 minutes. Each participant received 2 sessions of health education: once at enrollment and again during the 6-month follow-up hospital visit. The booklet guided patients on how to interpret food labels to manage stroke-related risk factors and included a reminder section for upcoming clinic appointments. Additionally, it featured a health-monitoring notebook for following daily BP, physical activity, and medication adherence by patients or caregivers.

Video-Based Education

Subsequently, participants watched a 15-minute educational video covering stroke risk factors, their management, and their prevention strategies. The educational video included simulated demonstrations of essential tasks, such as measuring the BP, organizing medications using boxes, and using salt-measuring spoons.

Following this, a physiotherapist conducted a 10-minute orientation session to teach participants about appropriate home-based daily physical exercises. To reinforce learning and support retention, RA nurses and physicians repeated the educational content during the 6-month follow-up visit according to the protocol.

Monthly motivational SMS reminders were sent to patients and caregivers to encourage adherence to daily BP monitoring, reduced salt intake, healthier eating habits, regular physical activity, and consistent medication use.

Control Group

Participants received standard medical care as per the hospital's protocols. As part of routine poststroke care, they received medical treatment, including prescribed medications and follow-up based on the attending physician's recommendations. Importantly, no additional health education related to stroke recurrence prevention was provided during the intervention

XSL•FO
RenderX

period. This ensured that only the IG was exposed to the educational component of the study.

Sample Size

We considered the stroke recurrence rate to be 20% [8]. The incidence rate of recurrence was 1.9/100 and 3.8/100 person-years for the IG and the CG, respectively, with 50% reduction [12] and 80% power. The expected proportion of outcomes from the IG was 10%. Here, "n" was the sample size, according to assumptions P_1 was the proportion of outcomes from the IG (10%), P_2 was the proportion of outcomes from the CG (20%), α was the level of significance (5%), $1 - \beta$ was the power of the test (80%), $Z_{1-\alpha/2}$=Z α value level of significance (1.96), and $Z_{1-\beta}$=Z β value level of the corresponding level of power (0.84). The estimated sample size was 196 for each arm (total 196 × 2=392 for both arms). Considering a 10% dropout rate, the final sample size was 432 (n=216, 50%, for each arm).

Operational Definitions

Recurrence of Stroke

Recurrence was defined as a new stroke event occurring at least 28 days after the incident event [17] confirmed by a physician who treated and diagnosed the patients in this study. We considered this definition of recurrence as >28 days to see the effects of the intervention in reducing stroke recurrence.

Adverse Events

Adverse events, including all causes of death and cardiovascular events, referred to any case that required a diagnosis. For instance, the physician treating and diagnosing the participants confirmed whether any recurrence of stroke, new onset of acute myocardial infarction and unstable angina, heart failure, and peripheral arterial disease required hospitalization.

Outcomes

The primary outcome of this study was to reduce the recurrence rate of stroke by providing health education after 12 months of follow-up. The secondary outcomes were to (1) reduce all adverse events (stroke recurrence, all causes of death, and cardiovascular events) and (2) increase medication adherence by providing health education after a 12-month follow-up period.

The effectiveness of health education was evaluated using 2 structured tools: a knowledge assessment questionnaire focused on stroke risk factors, symptoms, and prevention strategies and a behavior change questionnaire designed to measure lifestyle modifications, such as diet, physical activity, medication adherence, and BP monitoring. All tools were developed by the research team. These results will be published in a separate paper.

Data Monitoring and Quality Assurance

We conducted pretesting to observe the RA nurses' consent-taking, data collection, and physical examination procedures using 5% of the estimated samples of the total 432 participants and changed the procedures where necessary. These participants were not included in the main study. In our study, investigators (medical professionals) had direct involvement in patient care. They held regular biweekly meetings with the RA nurses to discuss the progress of the study and resolve any confusion and concerns. Furthermore, as part of routine quality control checking, the investigators performed field testing by independently verifying 5% of the data collected from participants on the same day. The RA nurses consulted with the investigators for any queries: questionnaires were completed by the RA nurses, who then input the information into a password-protected computer to maintain confidentiality. The investigators examined the information and answered the questions.

Statistical Analysis

An intention-to-treat (ITT) analysis was conducted. Descriptive statistics were performed to assess baseline data. Categorical variables were reported as frequencies and percentages and compared between the IG and the CG using the chi-square (χ^2) test or the Fisher exact test, where applicable. Continuous variables were reported as the mean (SD), and after checking the normality test, we performed the Mann-Whitney U test as our variables were not normally distributed. To estimate the time to events during the 12-month follow-up period from enrollment to the occurrence of events, study completion, dropout, and death, we used Kaplan-Meier estimates and compared them between the IG and the CG using the log-rank test. To compare the primary and secondary endpoints between the IG and the CG, we used Cox proportional hazards regression, and results were presented as hazard ratios (HRs) with 95% CIs. Because statistical significance was observed in mRS scores at baseline, we adjusted for the mRS score as a covariate in the Cox regression model to mitigate its potential impact on the primary and secondary outcomes. We also conducted the χ^2 test for medication adherence to compare the 2 groups. The significance level was set as P<.05. Data were analyzed using SPSS version 27.0 (IBM Corp).

Ethical Considerations

This study was approved by the Ethical Review Committee of the NINS&H (approval number: IRB/NINS/2022/151), Dhaka, Bangladesh. All participants provided written informed consent. In case the patients could not take care of themselves and needed family caregivers' support for their daily activities, we obtained written informed consent from the family caregivers instead, considering them as participants. Data were deidentified and assigned unique identification numbers for analysis, and no personal information was used in any publication or presentation resulting from this research. Allocation concealment was securely maintained by the principal investigator (PI) in a password-protected personal computer. Participants were compensated when they performed laboratory tests by themselves, and we reimbursed the same amount based on the receipt. To maintain ethical standards, health education materials were distributed to all CG participants after the intervention period was completed.

Results

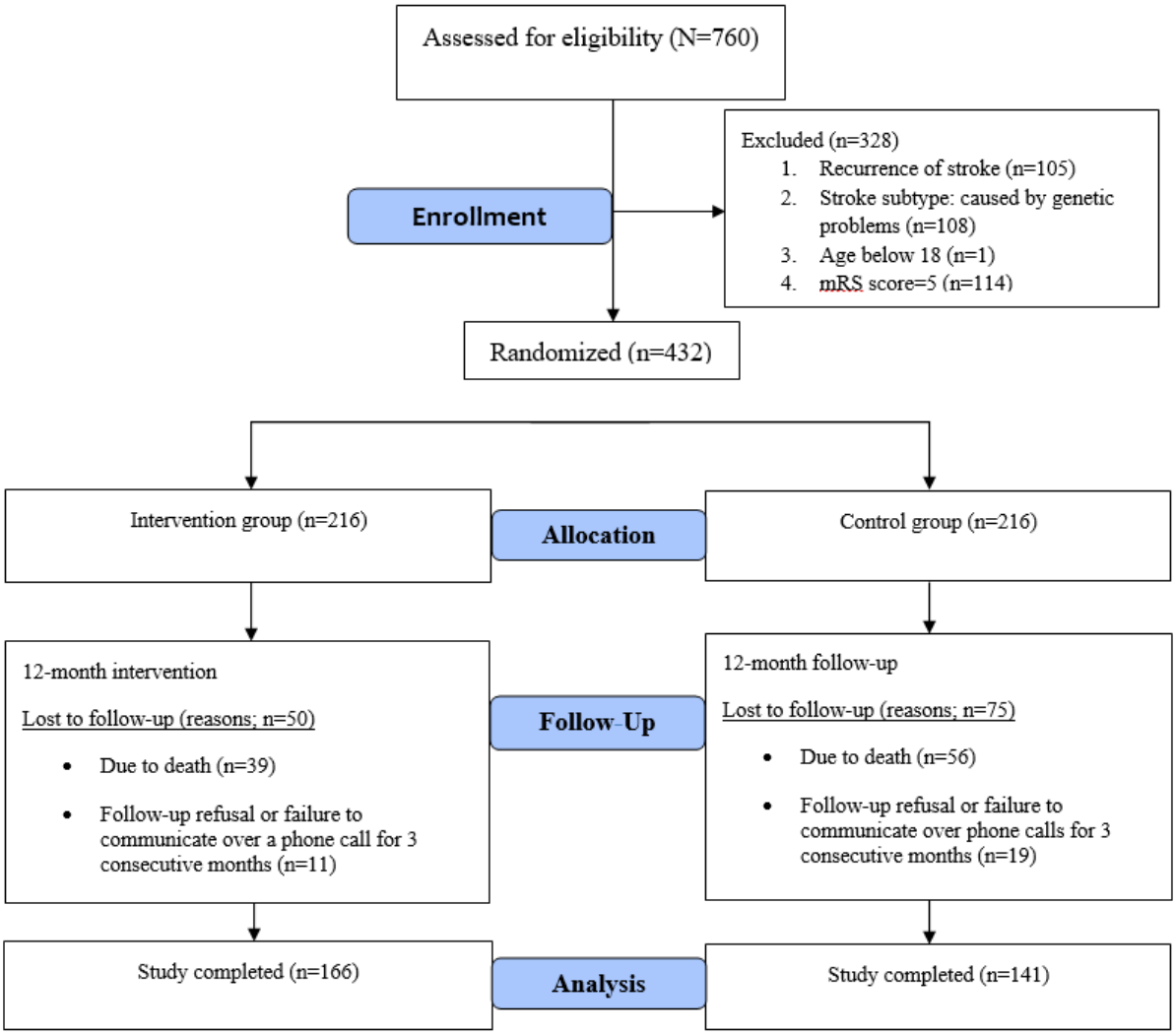
Participant Details

Of the 760 participants, we excluded 328 (43.2%) participants, and 432 (56.8%) met the inclusion criteria and consented to

XSL•FO
RenderX

participate in the study. The study flowchart is presented in Figure 2. Subsequently, participants were equally allocated to the IG (n=216, 50%) and the CG (n=216, 50%). After 12 months of follow-up, 307 (71.1%) participants completed the study (IG: 166/216, 76.9 %; CG: 141/216, 65.3%) and 95 (22%) died (IG: 39/216, 18.1%; CG: 56/216, 25.9%). Furthermore, 30 (6.9%) participants dropped out (IG: 11/216, 5.1%; CG: 19/216, 8.8%).

Figure 2. Study flowchart based on the CONSORT statement.



Sociodemographic Characteristics and Clinical Profiles of Participants

The average age of the participants was 55.2 years (SD 13.5), with 273 (63.2%) males and 263 (60.9%) participants from rural areas. The mean duration of onset of stroke to enrollment (*P*<.001), dyslipidemia (*P*=.03), taking hyperlipidemic medication (*P*=.05), and mRS score=1-3 (*P*<.001) were significantly higher in the IG compared to the CG. In contrast, the CG had significantly more rural residents (*P*=.02), a history of myocardial infarction (*P*=.01), and mRS score=4 (*P*<.001). Age (*P*=.19) and stroke types (*P*=.96) were well allocated (Table 1).

XSL•FO
RenderX

Table 1. Comparison of baseline characteristics and clinical profiles between the IG[a] and the CG[b].

Characteristics	Total participants (N=432)	IG (n=216)	CG (n=216)	*P* value
Age (years), mean (SD)	55.2 (13.5)	54.4 (13.7)	56.1 (13.2)	.19[c]
Males, n (%)	273 (63.2)	137 (63.4)	136 (63.0)	.92[d]
Married, n (%)	427 (98.8)	215 (99.5)	212 (98.1)	.18[d]
Living with family, n (%)	425 (98.4)	212 (98.1)	213 (98.6)	.70[d]
Education: literate, n (%)	309 (71.5)	160 (74.1)	149 (68.9)	.24[d]
Employed, n (%)	222 (51.4)	116 (53.7)	106 (49.1)	.34[d]
Residence: rural, n (%)	263 (60.9)	120 (55.6)	143 (66.2)	.02[d]
Monthly family income (US $)	184.39 (169.72)	189.45 (141.07)	174.05 (194.22)	.07[c]
Alcohol habit, n (%)	16 (3.7)	10 (4.6)	6 (1.4)	.31[e]
Smoking, n (%)	136 (31.5)	69 (31.9)	67 (31.0)	.84[d]
Clinical history of the patient				
Time from onset of stroke to enrollment (days), mean (SD)	7.75 (6.213)	8.85 (6.67)	6.64 (5.51)	<.001[f,g]
History of hypertension, n (%)	315 (72.9)	156 (72.2)	159 (73.6)	.75[d]
History of diabetes mellitus, n (%)	91 (21.1)	43 (19.9)	48 (22.2)	.56[d]
History of dyslipidemia, n (%)	295 (68.3)	158 (73.1)	137 (63.4)	.03[f]
History of arrhythmia, n (%)	30 (6.9)	15 (6.9)	15 (6.9)	.99[d]
History of myocardial infarction, n (%)	95 (22.0)	36 (16.7)	59 (27.3)	.01[d,f]
Stroke type, n (%)				
Cardioembolism	69 (16.0)	32 (14.8)	37 (17.1)	.96[e]
Atherothrombosis	131 (30.3)	65 (30.1)	66 (30.6)	.96[e]
Lacunar subtype	7 (1.6)	4 (1.9)	3 (1.4)	.96[e]
Hemorrhage	223 (51.6)	114 (52.3)	109 (50.5)	.96[e]
TIA[h]	2 (0.5)	1 (0.5)	1 (0.5)	.96[e]
Medications, n (%)				
Antihypertensives, n (%)	317 (73.4)	158 (73.1)	159 (73.6)	.91[d]
Antidiabetics, n (%)	92 (21.3)	43 (19.9)	49 (22.7)	.48[d]
Antihyperlipidemics	295 (68.3)	157 (72.7)	138 (63.9)	.05[d,f]
Antiarrhythmics	31 (7.2)	15 (6.9)	16 (7.4)	.85[d]
Cardiac drugs	95 (22.0)	36 (16.7)	59 (27.3)	.01[d,f]
mRS[i] score, n (%)				
0	1 (0.2)	0	1 (0.5)	<.001[e,f]
1	57 (13.2)	48 (22.2)	9 (4.2)	<.001[e,f]
2	34 (7.9)	21 (9.7)	13 (6.0)	<.001[e,f]
3	152 (35.2)	84 (38.9)	68 (31.5)	<.001[e,f]
4	188 (43.5)	63 (29.2)	125 (57.9)	<.001[e,f]

XSL•FO
RenderX

[a]IG: intervention group.
[b]CG: control group.
[c]Mann-Whitney U test.
[d]Chi-square test.
[e]Fisher exact test.
[f]*P*<.05.
[g]*t* Test.
[h]TIA: transient ischemic attack.
[i]mRS: modified Rankin Scale.

Primary Endpoint

The recurrence rate after 28 days was 5.1% (22/432) in the total participants, 6.5% (14/216) in the IG, and 3.7% (8/216) in the CG, and the difference was not statistically significant (*P*=.19) between the groups (Table 2).

Table 2. Comparison of stroke recurrence between the IGa and the CGb during 12-month follow-up.

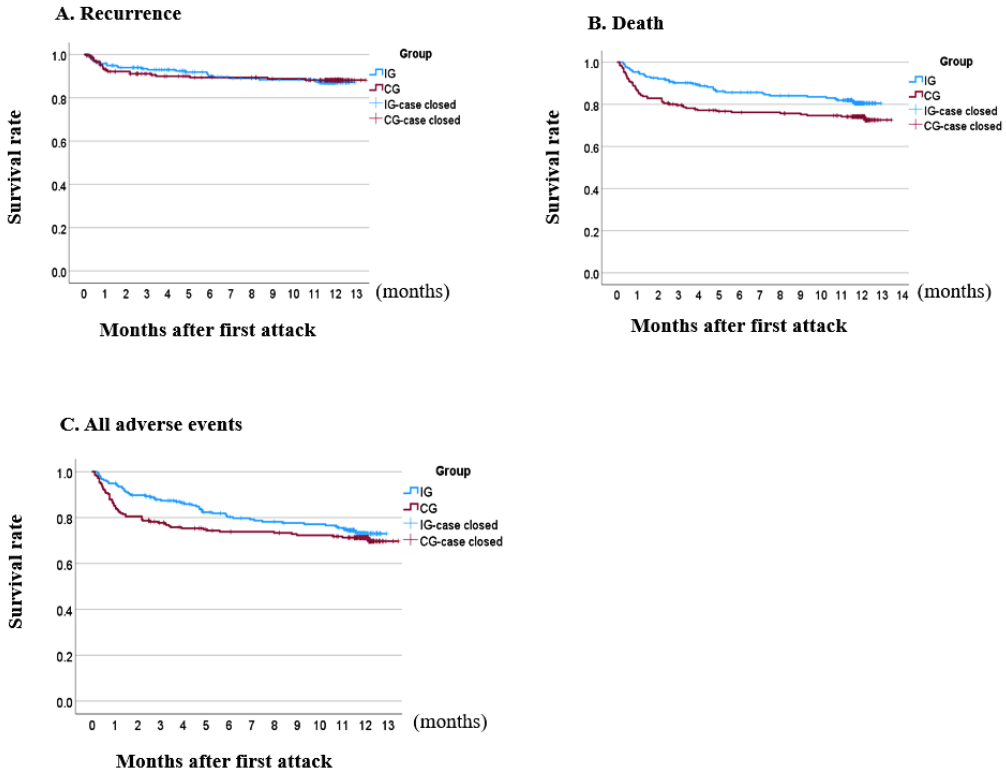
Recurrence onset	Total participants (N=432), n (%)	IG (n=216), n (%)	CG (n=216), n (%)	χ^2 (*df*)	*P* value
Within 28 days	26 (6.0)	11 (5.1)	15 (6.9)	0.66 (1)	.42
After 28 days	22 (5.1)	14 (6.5)	8 (3.7)	1.73 (1)	.19
Total	48 (11.1)	25 (11.5)	23 (10.6)	0.09 (1)	.76

[a]IG: intervention group.
[b]CG: control group.

Secondary Endpoints

The Kaplan-Meier survival curves in Figure 3A-3C show the estimated survival days for recurrence, death, and all adverse events at 12-month follow-up. These graphs illustrate time-to-event differences between the IG and the CG over the 12-month follow-up period.

Figure 3. Kaplan-Meier survival curve estimates the survival rate (days) for 12-month follow-up from the onset of the first attack: (A) recurrence, (B) death, and (C) all adverse events (recurrence, death, and cardiac events). CG: control group; IG: intervention group.



XSL•FO
RenderX

Stroke-Related Events

The HR for recurrence was 0.978 (95% CI 0.555-1.722, *P*=.94), and the total number of deaths was 95 (22%), with 39 (18.1%) deaths in the IG and 56 (25.9%) in the CG, and the difference was significant (HR 1.531, 95% CI 1.017-2.304; *P*=.04). The observed recurrence rate was significantly lower than expected, as mentioned before at 6.5% (14/216) in the IG and 3.7% (8/216) in the CG. The total number of all stroke-related adverse events, including recurrence, death, and cardiac events, in both groups was 116 (26.9%), with 54 (25%) adverse events in the IG and 62 (28.7%) in the CG, and the difference was not significant (HR 1.216, 95% CI 0.844-1.752; *P*=.29). To account for the significant baseline difference in mRS scores between the 2 groups, deaths were found to be not statistically significant after adjusting for the mRS score as a covariate. The adjusted HR was 1.234 (95% CI 0.693-2.197, *P*=.48) for recurrence, 0.818 (95% CI 0.540-1.238, *P*=.34) for death, and 1.016 (95% CI 0.701-1.473, *P*=.93) for all adverse events. The total number of cardiac events was 8 (1.8%), with 5 (2.3%) cardiac events in the IG and 3 (1.4%) in the CG. Therefore, the intervention effect was not statistically significant (Table 3).

In our Cox regression models, the proportional hazards assumption was tested using Schoenfeld residuals. For the model predicting evaluating the effect of the group on recurrence (with group as the sole covariate), the test statistic was χ^2_1=1.11 (*P*=.29). In the model for death, the group test statistic was χ^2_1=8.24 (*P*<.05). In the model for all adverse events, the group test statistic was χ^2_1=12.87 (*P*<.05).

Table 3. Cox regression analysis of patients with stroke (N=432[a]) for endpoints between the IG[b] and the CG[c].

Variables	Total participants (N=432), n (%)	IG (n=216), n (%)	CG (n=216), n (%)	Nonadjusted variables		Variables adjusted by baseline data (mRS[d])	
				HR[e] (95% CI)	*P* value	HR (95% CI)	*P* value
Recurrence (day 1 to 12 months)	48 (11.1)	25 (11.5)	23 (10.6)	0.978 (0.555-1.722)	.94	1.234 (0.693-2.197)	.48
Secondary endpoints							
Death	95 (22.0)	39 (18.1)	56 (25.9)	1.531 (1.017-2.304)	.04	0.818 (0.540-1.238)	.34
Cardiac events	8 (1.8)	5 (2.3)	3 (1.4)	—[f]	—	—	—
All adverse events[g]	116 (26.9)	54 (25.0)	62 (28.7)	1.216 (0.844-1.752)	.29	1.016 (0.701-1.473)	.93

[a]Note that of the 95 (22%) patients who died, only 2 (2.1%) were diagnosed with cardiac events and 33 (34.7%) with recurrence as the cause of death.

[b]IG: intervention group.

[c]CG: control group.

[d]mRS: modified Rankin Scale (mRS score was the covariate).

[e]HR: hazard ratio.

[f]Not applicable.

[g]Calculated by adding total recurrence, death, and cardiac events.

The monthly distribution of all causes of recurrence and death in 12 months was calculated and is shown in Multimedia Appendix 2. Nearly half of the cases of recurrence (27/48, 56.3%) and deaths (47/95, 49.5%) occurred within the first month. We also analyzed the distribution of death across different stroke types and mRS groups (Multimedia Appendix 3) and found that recurrence and death were the highest in the hemorrhagic subtype (n=23, 47.9%, and n=48, 50.5%, respectively) and in patients with an mRS score of 4 (n=31, 64.6%, and n=59, 62.1%, respectively).

Changes in Medication Adherence Over Time

Medication adherence strongly affects the recurrence of stroke. In this study, we compared the medication intake adherence between the IG and the CG at baseline, midline, and endline. We found significant improvement in the IG at midline and endline (*P*<.001 and *P*=.002, respectively). Taking medication “as doctor’s prescription for 6-7 days per week” significantly increased (Table 4).

XSL•FO
RenderX

Table 4. Comparison using the Fisher exact test of medication adherence between the IG[a] and the CG[b] at each time point.

Time points and groups	Not at all, n (%)	Occasionally, n (%)	Once/week, n (%)	2-3 days/week, n (%)	4-5 days/week, n (%)	6-7days/week, n (%)
Baseline (*P*=.78)						
IG (n=216)	67 (31.0)	54 (25.0)	1 (0.5)	6 (2.8)	23 (10.6)	65 (30.1)
CG (n=216)	57 (26.4)	71 (32.9)	0 (0.0)	2 (0.9)	25 (11.6)	61 (28.2)
Midline[c] (*P*<.001)						
IG (n=149)	12 (8.1)	2 (1.3)	0 (0.0)	10 (6.7)	4 (2.7)	121 (81.2)
CG (n=130)	26 (20.0)	1 (0.8)	0 (0.0)	0 (0.0)	0 (0.0)	103 (79.2)
Endline (*P*=.002)						
IG (n=166)	2 (1.2)	0 (0.0)	7 (4.2)	5 (3.0)	11 (6.6)	141 (84.9)
CG (n=141)	11 (7.8)	0 (0.0)	18 (12.8)	5 (3.5)	6 (4.3)	101 (71.6)

[a]IG: intervention group.

[b]CG: control group.

[c]Patient attendance was low at midline.

Discussion

Principal Findings

This is the first intervention study that followed patients with stroke for 12 months to evaluate the risk factors for recurrence and other adverse events related to stroke in Bangladesh. We provided health education for controlling modifiable risk factors to prevent the recurrence of stroke in patients who had a stroke attack.

We compared patients who received the health education intervention with those who received usual care, and no significant difference was found in reducing the recurrence of stroke in the IG after 12 months of follow-up. A similar finding was observed in another study: recurrence did not seem to be reduced by only self-management education with treatment coordination [12]. We included disproportionately more patients with severe stroke (mRS score=4) in the CG; therefore, more patients died in the CG. Among all deaths, 61.1% were undiagnosed: 27.4% in the IG and 33.7% in the CG. This caused our intervention to be ineffective.

Patients with stroke need to undergo laboratory investigations, such as computerized tomography (CT) scans and magnetic resonance imaging (MRI) for their diagnosis. However, most of the patients in this study could not undergo these tests for a subsequent attack, recurrence, and other adverse events, including the cause of death, due to the unavailability of investigation facilities and the patients' financial constraints. Therefore, it was not possible to diagnose recurrence and the cause of death in patients with stroke. In addition, many of our patients could not undergo these tests as they died at home before visiting the hospital for their diagnosis; those deaths were considered of having unknown causes. Among patients with recurrent stroke, less than three-quarters (33/48; n=13 in the IG vs n=20 in the CG) died. Therefore, recurrence was considered a main risk factor for death [21].

We included patients with ischemic and hemorrhagic stroke. The risk of recurrence and death for patients with hemorrhagic stroke is higher compared to those with ischemic stroke [22]. This caused an overall 22% rate of death among the study participants, which might have had a significant impact on the results. Moreover, we did not consider the differences in risk factors, admission severity, and discharge status between the 2 groups, which was also one reason for the insignificant results of this study. It was also difficult to evaluate the effects of the intervention within this short time. Many patients were first diagnosed with new comorbidities, such as hypertension, diabetes, and dyslipidemia, after they visited the hospital and started treatment. Therefore, the risk of recurrence after a stroke remained and made the patients more vulnerable to other adverse events. It also increased the risk of stroke attack after discharge from the hospital due to inadequate monitoring of the risk factors at home. Although the patients were on antihyperlipidemic medication for dyslipidemia, there was still a high risk of developing stroke recurrence [23]. These might have acted as confounding factors that could not reduce the recurrence of stroke significantly through the health education intervention.

The risk of recurrence in our study was higher within the first 6 months compared to the subsequent 6 months. Another study also found that the first 3 months are crucial for the recurrence of stroke [8]. Therefore, intensive intervention is necessary to control the risk factors within this time.

Even though we failed to show the effectiveness of health education in this study, we still cannot ignore the importance of health education. Many previous studies have shown the effectiveness of health education on hypertension, diabetes, and ischemic heart diseases [24-27]. In Bangladesh, postdischarge health education is not paid as much attention in hospitals and primary care settings. Therefore, stroke and other noncommunicable diseases are highly prevalent as a result of this reactive strategy. Nationwide health education programs, routine health checkups, and regular follow-up visits could be preventive measures for the early detection and treatment of the risk factors of stroke. This can also lower the health care costs and enhance long-term health results. Regular health education and health check-up campaigns can increase knowledge about

XSL•FO
RenderX

the risk factors of stroke and motivate people to change their behavior. In Bangladesh, patients with stroke attacks usually attend local health care facilities. As the primary and secondary health care systems are weak, patients are not diagnosed properly and treated (lack of qualified physicians and diagnostic facilities) and miss an opportunity to receive acute care management. In our study, patients visited the hospital 7.75 days after the onset of stroke on average, and the IG delayed visiting the hospital more compared to the CG. This higher delay made the IG more vulnerable to developing recurrence and other stroke-related adverse events compared to the CG.

The NINS&H is the only specialized referral hospital in Bangladesh and receives more patients compared to its limited beds. Therefore, only patients with severe stroke (mRS score=3-4) visit the hospital and are admitted, which pushed us to enroll participants from the outpatient department to obtain mRS scores of 0-2.

The sociodemographic attributes of the participants may also be a factor in determining the results. The average age of our participants was 55 years, with a minimum of 18 years to a maximum of 92 years, and males were predominant. One meta-analysis conducted in Japan reported that the average age of the patients was 67.2 years [28]. Considering their age, older adults feel inconvenienced when attending health care facilities due to the long distance from their homes. Despite the desire and necessity of many participants to visit hospitals, financial constraints prevent them from covering treatment-related associated expenses as there is no national health insurance system in Bangladesh [29,30].

It is alarming that young people are also at risk of stroke in this population. Some of them were vulnerable to stroke as they had modifiable risk factors, such as smoking [31,32]. Three-quarters of them had hypertension, and less than three-quarters had dyslipidemia, which are the major risk factors for the recurrence of stroke [23,33]. This study found that medication adherence significantly improved in the IG during monthly follow-ups. This could be a positive effect of health education. Regular medication adherence is crucial for controlling underlying diseases.

About half of our participants were unemployed, and some of them had financial constraints. Their average monthly income was far below (US $ 184.39) the national average income of US $ 277 [34]. Therefore, they might have limited access to hospitals or any health care facilities for medical checkups.

Limitations

The study has a few limitations. We followed a convenience sampling method using a specialized tertiary hospital that receives patients from diverse geographic regions across the country. Although this approach may have introduced selection bias and limited the representativeness of the sample, we believe the findings provide valuable insights and may still reflect broader patterns relevant to the nationwide setting.

During enrollment, we did not allocate participants to the 2 groups based on the severity of stroke (mRS score); usually, the stroke subtype and the mRS score are correlated [16]. Therefore, there was an uneven distribution of participants, which influenced the findings of our study.

A huge number of participants died at their homes without any investigation for the diagnosis of the cause of death. The limited diagnostic facilities, a lack of health care resources, and patients' economic difficulties influenced our study findings. The overall sample size was small, and the follow-up time was insufficient, which might also have a significant impact on the results.

Moreover, the number of recurrence cases and the number of events in this study were small, which may have limited the statistical power and widened the CIs, and effects were cautiously interpreted. However, the small number of events and visual inspection suggested time-related potential instability. Causal interpretation should be avoided, and further studies with larger samples and longer follow-ups are needed.

As this study was conducted in the context of Bangladesh, generalizability may be limited due to differences in health care infrastructure, socioeconomic factors, and stroke management practices compared to other countries. We acknowledge that some results, particularly those related to lifestyle behaviors, medication adherence, and self-monitoring, were based on self-reports from patients and caregivers. Although clear guidance was provided to the RA team to ensure accuracy, self-reported data are characteristically topic to social attractiveness and recall biases, which may affect the dependability of the answers. Multicenter studies are recommended to validate the effect of health education for patients poststroke.

This study used a single-blind design, which may have introduced observer bias during outcome assessment, particularly in a resource-limited setting, such as Bangladesh, where familiarity between participants and assessors could influence data collection. To reduce the possibility of assessment bias, outcome evaluations, such as recurrence and medication adherence evaluations, were carried out by an experienced investigator who was not involved in the study.

Conclusion

We provided health education to patients with stroke and found it had no effect in reducing recurrence and stroke-related adverse events, including death. This study emphasizes the importance of health education in enhancing medication adherence among patients. In both groups, patient with hemorrhagic stroke and an mRS score of 4 had more stroke recurrences and deaths overall. Therefore, the acute care system needs to pay more attention to prevent recurrences and deaths. Patients should visit the hospital without any delay to get treatment in their acute phase. We could not identify the cause of death and recurrence among patients who had not attended health care facilities for their diagnosis and treatment. Introducing regular health checkup systems, postdischarge health education, and follow-up should be incorporated into the national health policy**.** Further interventional studies with large sample sizes to address all limitations, including diagnostic facilities for stroke, will determine the underlying factors for patients with stroke developing recurrence and even dying.

XSL•FO
RenderX

Acknowledgments

We would like to express our sincere gratitude to the study participants. We extend our thanks to the National Institute of Neuroscience & Hospital authorities for their support and the research assistant nurses and laboratory staff for their unwavering assistance. The study was supported by a restricted research grant from Hiroshima University, Japan

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization, investigation, methodology, resources, software, validation, and visualization were handled by MA and MM; data curation by MA; formal analysis by MA, YF (Yasuko Fukuoka), and YF (Yasuko Fukushima); funding acquisition by MM; project administration by MA, SUK, SCD, MSJHC, and MM; supervision by SUK, SCD, MSJHC, and MM; writing—original draft by MA, KATMEH; and writing—review and editing by KATMEH and MM.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT 2010 checklist.
[PDF File (Adobe PDF File), 126 KB - publichealth_v11i1e72233_app1.pdf]

Multimedia Appendix 2
Comparison of monthly recurrence and death that occurred during the 6- and 12-month follow-up.
[DOCX File , 25 KB - publichealth_v11i1e72233_app2.docx]

Multimedia Appendix 3
Distribution of recurrence and death due to stroke by type and mRS (modified Rankin Scale) score within the 12-month follow-up period.
[DOCX File , 23 KB - publichealth_v11i1e72233_app3.docx]

References

1. Owolabi M, Thrift A, Mahal A, Ishida M, Martins S, Johnson WD, Stroke Experts Collaboration Group. Primary stroke prevention worldwide: translating evidence into action. Lancet Public Health 2022 Jan;7(1):e74-e85 [FREE Full text] [doi: 10.1016/S2468-2667(21)00230-9] [Medline: 34756176]
2. Campbell B, De Silva DA, Macleod M, Coutts SB, Schwamm LH, Davis SM, et al. Ischaemic stroke. Nat Rev Dis Primers 2019 Oct 10;5(1):70 [FREE Full text] [doi: 10.1038/s41572-019-0118-8] [Medline: 31601801]
3. Ekker M, Boot E, Singhal A, Tan K, Debette S, Tuladhar A, et al. Epidemiology, aetiology, and management of ischaemic stroke in young adults. Lancet Neurol 2018 Sep;17(9):790-801 [FREE Full text] [doi: 10.1016/s1474-4422(18)30233-3]
4. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820 [FREE Full text] [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
5. Global health estimates: Leading causes of death. World Health Organization. URL: https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death [accessed 2025-05-16]
6. Mondal M, Hasan A, Khan N. Prevalence and risk factors of stroke in Bangladesh: a nationwide population-based survey. medRxiv Preprint posted online 2021. [doi: 10.1101/2021.06.17.21259097]
7. Recurrent Strokes: everything that you need to know. Echelon Health. URL: https://www.echelon.health/recurrent-strokes-everything-that-you-need-to-know/ [accessed 2025-05-16]
8. Kanti Dey S, Bakshi L, Ahmed A, Md Shahidullah , Habib A, Chowdhury A, et al. Recurrence of ischemic stroke patients with common risk factors. Bangla J Neurosci 2019 Jan 31;35(1):33-38. [doi: 10.3329/bjn.v35i1.57623]
9. Peters D, Garg A, Bloom G, Walker DG, Brieger WR, Rahman MH. Poverty and access to health care in developing countries. Ann N Y Acad Sci 2008;1136:161-171 [FREE Full text] [doi: 10.1196/annals.1425.011] [Medline: 17954679]
10. United Nations. Dangers of poor quality health care revealed ‘in all countries’: WHO report. UN News. 2018. URL: https://news.un.org/en/story/2018/07/1013942 [accessed 2024-09-30]

XSL•FO
RenderX

11. Bridgwood B, Lager KE, Mistri AK, Khunti K, Wilson AD, Modi P. Interventions for improving modifiable risk factor control in the secondary prevention of stroke. Cochrane Database Syst Rev 2018 May 07;5(5):CD009103 [FREE Full text] [doi: 10.1002/14651858.CD009103.pub3] [Medline: 29734470]
12. Fukuoka Y, Hosomi N, Hyakuta T, Omori T, Ito Y, Uemura J, et al. Effects of a disease management program for preventing recurrent ischemic stroke. Stroke 2019 Mar;50(3):705-712 [FREE Full text] [doi: 10.1161/strokeaha.118.020888]
13. Trivedi S, Corderman S, Berlinberg E. Assessment of patient education delivered at time of hospital discharge. JAMA Intern Med 2023;183(5):417-423 [FREE Full text] [doi: 10.1001/jamainternmed.2023.0070]
14. Shoshi K, Sheikh S. Awareness regarding stroke among the caregivers of stroke survivors at a tertiary level hospital in Bangladesh. Open Access Library J 2023;10:e10285 [FREE Full text] [doi: 10.4236/oalib.1110285]
15. Grant S, Mayo-Wilson E, Montgomery P, Macdonald G, Michie S, Hopewell S, on behalf of the CONSORT-SPI Group. CONSORT-SPI 2018 explanation and elaboration: guidance for reporting social and psychological intervention trials. Trials 2018 Jul 31;19(1):406 [FREE Full text] [doi: 10.1186/s13063-018-2735-z] [Medline: 30060763]
16. Banks J, Marotta C. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. Stroke 2007;38(3):1091-1096 [FREE Full text] [doi: 10.1161/01.str.0000258355.23810.c6]
17. Han J, Mao W, Ni J, Wu Y, Liu J, Bai L, et al. Rate and determinants of recurrence at 1 year and 5 years after stroke in a low-income population in rural China. Front Neurol 2020;11:2 [FREE Full text] [doi: 10.3389/fneur.2020.00002] [Medline: 32038470]
18. Zhang J, Zhu P, Liu B, Yao Q, Yan K, Zheng Q, et al. Time to recurrence after first-ever ischaemic stroke within 3 years and its risk factors in Chinese population: a prospective cohort study. BMJ Open 2019 Dec 18;9(12):e032087 [FREE Full text] [doi: 10.1136/bmjopen-2019-032087] [Medline: 31857306]
19. Katan M, Luft A. Global burden of stroke. Semin Neurol 2018 Apr;38(2):208-211 [FREE Full text] [doi: 10.1055/s-0038-1649503] [Medline: 29791947]
20. Afrin M, Khan S, Das S, Huq KATME, Moriyama M. Effectiveness of a health education program for patients who had a stroke and their caregivers by controlling modifiable risk factors to reduce stroke recurrence in a tertiary hospital in Bangladesh: protocol for a randomized controlled trial. JMIR Res Protoc 2023 Dec 15;12:e51178 [FREE Full text] [doi: 10.2196/51178] [Medline: 38100172]
21. Chen Y, Wright N, Guo Y, Turnbull I, Kartsonaki C, Yang L, China Kadoorie Biobank Collaborative Group. Mortality and recurrent vascular events after first incident stroke: a 9-year community-based study of 0·5 million Chinese adults. Lancet Glob Health 2020 Apr;8(4):e580-e590 [FREE Full text] [doi: 10.1016/S2214-109X(20)30069-3] [Medline: 32199124]
22. Andersen KK, Olsen TS, Dehlendorff C, Kammersgaard LP. Hemorrhagic and ischemic strokes compared: stroke severity, mortality, and risk factors. Stroke 2009 Jun;40(6):2068-2072. [doi: 10.1161/strokeaha.108.540112]
23. Coull A, Rothwell P. Underestimation of the early risk of recurrent stroke: evidence of the need for a standard definition. Stroke 2004 Aug;35(8):1925-1929 [FREE Full text] [doi: 10.1161/01.str.0000133129.58126.67]
24. Lukitasari M, Nugroho D, Rohman M, Mardhotillah H, Natasya D, Fitriyawat F, et al. An intervention study for impact assessment of health education by empowered community health workers in improving treatment and diet adherence in hypertension. Indian J Community Med 2021;46(4):618 [FREE Full text] [doi: 10.4103/ijcm.ijcm_895_20]
25. Nazar C, Bojerenu M, Safdar M, Marwat J. Effectiveness of diabetes education and awareness of diabetes mellitus in combating diabetes in the United Kigdom; a literature review. J Nephropharmacol 2016;5(2):110-115 [FREE Full text] [Medline: 28197516]
26. Oliveira A, Cavalcante A, Carneiro C, Santos VB, Moorhead S, Lopes J, et al. Health education: the effectiveness of interventions in patients with heart failure. Rev Bras Enferm 2020;73(2):e20180782 [FREE Full text] [doi: 10.1590/0034-7167-2018-0782] [Medline: 32159695]
27. Zhu L, Ho S, Wong T. Effectiveness of health education programs on exercise behavior among patients with heart disease: a systematic review and meta-analysis. J Evid Based Med 2013 Nov;6(4):265-301 [FREE Full text] [doi: 10.1111/jebm.12063] [Medline: 24325420]
28. Kitagawa K, Yamamoto Y, Arima H, Maeda T, Sunami N, Kanzawa T, Recurrent Stroke Prevention Clinical Outcome (RESPECT) Study Group. Effect of standard vs intensive blood pressure control on the risk of recurrent stroke: a randomized clinical trial and meta-analysis. JAMA Neurol 2019 Nov 01;76(11):1309-1318 [FREE Full text] [doi: 10.1001/jamaneurol.2019.2167] [Medline: 31355878]
29. Population and housing census 2022. Bangladesh Bureau of Statistics. 2023. URL: http://www.bbs.gov.bd/site/page/47856ad0-7e1c-4aab-bd78-892733bc06eb/Population-and-Housing-Census [accessed 2025-05-16]
30. Mohammad K, Kawser A, Rebeka S. Study report on health insurance for public servants in Bangladesh: a study on its social security implications and challenges. Cabinet Division, Government of Bangladesh. 2022 Sep. URL: https://tinyurl.com/3puda9tw [accessed 2025-05-16]
31. Shah RS, Cole JW. Smoking and stroke: the more you smoke the more you stroke. Expert Rev Cardiovasc Ther 2010 Jul 10;8(7):917-932 [FREE Full text] [doi: 10.1586/erc.10.56] [Medline: 20602553]
32. Markidan J, Cole J, Cronin C, Merino J, Phipps M, Wozniak M, et al. Smoking and risk of ischemic stroke in young men. Stroke 2018 May;49(5):1276-1278 [FREE Full text] [doi: 10.1161/strokeaha.117.018859]

XSL•FO
RenderX

33. Wang Y, Xu J, Zhao X, Wang D, Wang C, Liu L, et al. Association of hypertension with stroke recurrence depends on ischemic stroke subtype. Stroke 2013 May;44(5):1232-1237 [FREE Full text] [doi: 10.1161/strokeaha.111.000302]
34. Bangladesh Bureau of Statistics, Statistics and Informatics Division,Ministry of Planning. Household income and expenditure survey. HIES 2022. Bangladesh Bureau of Statistics. 2023 Apr 12. URL: https://bbs.portal.gov.bd/sites/default/files/files/bbs.portal.gov.bd/page/57def76a_aa3c_46e3_9f80_53732eb94a83/2023-04-13-09-35-ee41d2a35dcc47a94a595c88328458f4.pdf [accessed 2025-05-16]

Abbreviations

BP: blood pressure
CG: control group
CONSORT: Consolidated Standards of Reporting Trials
HR: hazard ratio
IG: intervention group
LMIC: low- and middle-income country
mRS: modified Rankin Scale
NINS&H: National Institute of Neuroscience & Hospital
RA: research assistant
SVO: small vessel occlusion
TIA: transient ischemic attack

Edited by A Mavragani; submitted 06.02.25; peer-reviewed by T Hyakuta, A Protopapas; comments to author 11.04.25; revised version received 30.04.25; accepted 07.05.25; published 27.05.25.

Please cite as:
Afrin M, Huq KATME, Khan SU, Das SC, Chowdhury MSJH, Fukuoka Y, Fukushima Y, Moriyama M
Effectiveness of a Health Education Program to Reduce Recurrence of Stroke by Controlling Modifiable Risk Factors in a Specialized Hospital in Bangladesh: Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e72233
URL: https://publichealth.jmir.org/2025/1/e72233
doi:10.2196/72233
PMID:

©Mahabuba Afrin, K A T M Ehsanul Huq, Sharif Uddin Khan, Subir Chandra Das, Mohammad Shah Jahirul Hoque Chowdhury, Yasuko Fukuoka, Yasuko Fukushima, Michiko Moriyama. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Impact of Commercial Food Environments on Local Type 2 Diabetes Burden: Cross-Sectional and Ecological Multimodeling Study

Kurubaran Ganasegeran[1,2,3,4*], MBBS, MSc; Mohd Rizal Abdul Manaf[1*], MPH, PhD; Lance A Waller[5*], MS, PhD; Nazarudin Safian[1], MPH, PhD; Muhammad Faid Mohd Rizal[6], BSc, MSc; Wye Lee Chiew[7], MPH, DrPH; Feisul Mustapha[8], MPH; Khairul Nizam Abdul Maulud[6,9], MSc, PhD

[1]Department of Public Health Medicine, Faculty of Medicine, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

[2]Clinical Research Centre, Seberang Jaya Hospital, Ministry of Health Malaysia, Seberang Perai, Penang, Malaysia

[3]Occupational Safety and Health Unit, Seberang Jaya Hospital, Ministry of Health Malaysia, Seberang Perai, Penang, Malaysia

[4]Public Health Unit, Seberang Jaya Hospital, Ministry of Health Malaysia, Seberang Perai, Penang, Malaysia

[5]Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[6]Earth Observation Centre (EOC), Institute of Climate Change, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

[7]Penang State Health Department, Ministry of Health Malaysia, Georgetown, Penang, Malaysia

[8]Perak State Health Department, Ministry of Health Malaysia, Ipoh, Perak, Malaysia

[9]Department of Civil Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

[*]these authors contributed equally

Corresponding Author:
Mohd Rizal Abdul Manaf, MPH, PhD
Department of Public Health Medicine
Faculty of Medicine
Universiti Kebangsaan Malaysia
Jalan Yaacob Latif
Kuala Lumpur, 56000
Malaysia
Phone: 60 193208925
Email: mrizal@hctm.ukm.edu.my

Abstract

Background: Neighborhoods resulting from rapid urbanization processes are often saturated with eateries for local communities, potentially increasing exposure to unhealthy foods and creating diabetogenic residential habitats.

Objective: We examined the association between proximity of commercial food outlets to local neighborhood residences and type 2 diabetes (T2D) cases to explore how local T2D rates vary by location and provide policy-driven metrics to monitor food outlet density as a potential control for high local T2D rates.

Methods: This cross-sectional ecological study included 11,354 patients with active T2D aged ≥20 years geocoded using approximate neighborhood residence aggregated to area-level rates and counts by subdistricts (*mukims*) in Penang, northern Malaysia. We used the National Diabetes Registry complemented with data from medical records across 29 primary care clinics throughout the state. Food establishment data were retrieved from the Open Data Portal sourced through the Penang GeoHub, and urbanization indicators were retrieved from MyCensus 2020. We executed point-level proximity- and density-based area-level analysis through multimodel aspatial and spatial regression methods.

Results: Our final hierarchical linear regression revealed that the distance to food complexes, hawker markets, *kopitiams* (a type of coffee shop), 24-7 convenience stores, fast food outlets, and public markets showed statistically significant associations (P<.05) with the age and BMI of patients with T2D. In the multiscale geographically weighted regression model, the adjusted R^2 values ranged from 0.15 to 0.62, with lower values observed across the mainland. The multiscale geographically weighted regression model yielded average β coefficients for densities of *kopitiams* (β=0.256), fast food outlets (β=−0.061), 24-7 convenience stores (β=0.028), supermarkets (β=0.122), public markets (β=0.067), and *nasi kandar* (a type of rice dish) restaurants (β=−0.064), urban growth rate (β=0.189), and population density (β=−0.080; $t_{65.835}$≥1.96 in all cases). We established population-attributable

XSL•FO
RenderX

fractions suggesting that, if local neighborhoods underwent township restructuring to remove food complexes, hawker markets, or *kopitiams*, an estimated reduction of 0.21%, 0.27%, and 0.09%, respectively, in the risk of T2D cases in Penang would be anticipated. However, if local neighborhoods underwent township restructuring to add hawker complexes, *nasi kandar* restaurants, fast food outlets, 24-7 convenience stores, public markets, or supermarkets, an estimated reduction of between 0.07% and 0.64% in the number of residents with risk of T2D was estimated.

Conclusions: The reported variations provide insights into the associations between high neighborhood T2D rates and the density of a range of food outlets. We observed that these associations varied by place, providing insight into potential monitoring and policy considerations. This work provides evidence for interpretation at the individual and aggregate levels, shifting public health interventions from a generic to a targeted approach and prompting township planners to restructure food outlet accessibility or availability in local neighborhoods and to develop health behavior interventions for local communities for healthy food purchase and consumption.

(JMIR Public Health Surveill 2025;11:e70045) doi:10.2196/70045

KEYWORDS

local eateries; neighborhood; health; diabetes; policies; attributable fraction; township; urbanization; spatial

Introduction

Background

Local changes to land use and neighborhoods are structured within ongoing worldwide urbanization changes, impacting access to food, health, and human livelihood across towns, cities, modern villages, and even distant remote areas [1,2]. As societies consolidate in newer environments further removed from food sources, food production, distribution, and accessibility evolve away from local options to more processed, imported, and commercially driven choices [3]. Due to this, and coupled with increasing options for consumption of food with high fat, caloric, or sugar content, societies are changing and assuming dietary transition roles that coincide with lifestyle and behavior changes often in a more sedentary direction and that potentially increase the risk of noncommunicable diseases because of dietary implications [4]. Urban transformations collectively restructure townships to accommodate more connected human habitats, but such transfigurations also accommodate socioeconomic, infrastructural, and socioecological changes supporting the lifestyle dynamics of urban space development coupled with capitalist industrialization processes aimed to catalyze local economic advancement.

Contemporary neighborhoods resulting from ongoing urbanization processes are often saturated with commercial eateries for access by local communities [4]. Urbanization forcibly transforms food systems according to geographical patterns mediated by food supply, food demand, food delivery, and local consumer choices, driving changing market landscapes and interconnectedness between commercial food system actors and consumers [3-5]. Some local neighborhoods are afflicted by social inequities or inequalities and can be even more susceptible to local food commercialization activities that tend to influence peoples' preferences regarding food purchasing behaviors according to the availability and accessibility of different food outlets [4,5].

As place matters, we posit a local nutritional transition theory from a geographical epistemology perspective collectively nested within neighborhoods' modernization and urbanization processes, diving across the commercial and sociospatial complexities shifting access to available food choices for local neighborhood consumers [3]. Under these local neighborhood circumstances, it becomes increasingly difficult for local public health organizations and governments to achieve or even maintain sustainable food consumption in line with United Nations Sustainable Development Goal 2; despite food being available and accessible to people to end hunger, it is challenging for community planners to provide access to healthy and nutritious food within these changing foodscapes [5].

Neighborhood food environments are collectively contextualized with respect to how individuals from residential homes or settlements encounter foods and engage in decision-making as to what foods to purchase and consume [6]. The linkages between neighborhood food environments and local burdens of noncommunicable diseases have been widely established in the United States [7-10], South Asia [11], and Australia [12], but to date, none have been established in Malaysia. As measures of availability and accessibility to food outlets in previous work vary significantly according to geographical settings, culture, or resources available [13], this study explored the linkage between food outlet density and proximity within neighborhoods and type 2 diabetes (T2D) rates in the Malaysian setting.

Studies examining correlations involving food environment availability and accessibility across neighborhoods have accumulated over time. Although earlier work has used different metrics (ie, density- or proximity-based measures) for the associations between neighborhood food environments and the local burden of noncommunicable diseases across different settings based on available resources [13], none have explored how the efforts and metrics used to establish connections could cumulatively be used as evidence to motivate and inform targeted interventions and policies related to the restructuring of local food environments across neighborhoods. Despite the implementation of numerous policies and interventions by the Malaysian government to drive healthy food consumption among citizens (eg, the Malaysia sugar tax, Healthier Choice Logo, and Malaysia Dietary Guidelines), the incidence of obesity and T2D has steadily increased over the past decade [14]. These policies were primarily focused on individual-level behavior regarding adopting healthy eating habits but

XSL•FO
RenderX

substantially failed to address the changing landscapes of the distal determinants of health, the environment in which we live, and the exposure to food commercialization activities that can nudge individuals to access or purchase ultraprocessed or cheaper food choices offered by commercial actors within reach, in addition to using poor implementation strategies to tackle or control the changing food landscapes [15].

Objectives

Such circumstances were not due to failure of the strategies implemented but, instead, highlight how those interventions could not be targeted based on priorities due to limited evidence and metrics used to assess and monitor the impacts of the local food environments on residents' health. To expand this work and establish baseline evidence for local public health and nutrition epidemiologists to use in implementing targeted place-based interventions for T2D control in Penang, our study aimed to aid in data-based crafting of local policies for healthy township restructuring through the following objectives:

1. Determine the proximity of food outlets to the locus of analysis (ie, neighborhood residence) while concurrently yielding associations between proximity to food outlets and age and BMI of patients with T2D by different distance classifications from the locus of analysis
2. Understand how local T2D rates vary by location across subdistricts (*mukims*) in Penang based on urbanization indicators and available food outlet density
3. Establish a policy-driven metric based on population-attributable fractions (PAFs) through an exponentiated algorithmic regression function within *mukims* to estimate how much a particular food outlet density in a town needs to be restructured to control local T2D rates within *mukims*

Methods

Study Design and Setting

We designed and implemented a population-based cross-sectional (ie, to determine proximity to food outlets at the individual level to address objective 1) and ecological (ie, to determine T2D rates that vary by location at the aggregate level, as highlighted in objectives 2 and 3) study design that applied aspatial and spatial methodological approaches to determine the relationships among the built environment (specifically, the food environment), urbanization indicators, individual-level demographics, and T2D burden through proximity- and density-based analytics using small-area methods of investigation [16] by leveraging spatial data analysis.

This study was conducted in the northern region of Malaysia within the state of Penang, a United Nations Educational, Scientific, and Cultural Organization heritage site that hosts high numbers of local commercial food enterprises and many tourism activities. The state consists of 83 *mukims* housing approximately 1.77 million people as of 2020 [17].

Study Population and Geolocation

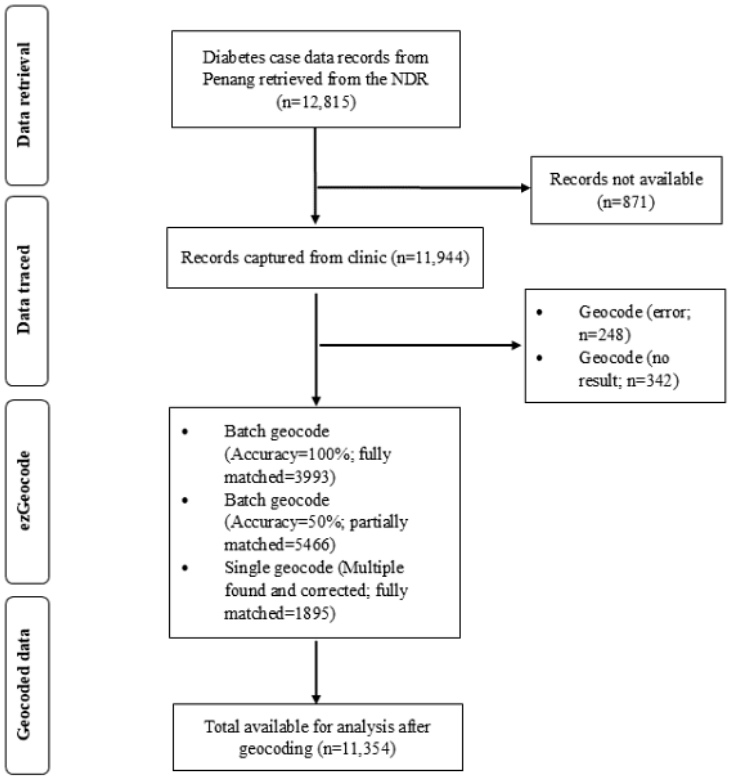
Data on 12,815 adults with active T2D aged ≥20 years between 2016 and 2020 from 29 primary care clinics (*klinik kesihatan*) throughout Penang by administrative district were retrieved from the state-level National Diabetes Registry (NDR), a subset of the countrywide NDR [18]. These 12,815 active T2D cases were traced individually from each clinic in Penang using patient records to obtain residential address information, *mukim* information, and individual-level covariate data for point pattern analyses.

All T2D cases' residential address information at the time of diagnosis was geocoded into a coordinate system (longitude and latitude) upon retrieval from the clinic registry. This enhanced temporal accuracy (ie, allowing neighborhood exposures to be more accurately correlated with T2D cases at the time of diagnosis) as neighborhood landscapes and environment change over time in line with urbanization processes and township development. The valid anonymized residential addresses of patients with T2D (in the standard Malaysian format of unit or block, street, neighborhood or village name, and postal code) in Penang underwent geocoding processes to convert locations to the World Geodetic System 84 coordinate system (longitude and latitude). The geocoding process was conducted using the ezGeocode app (ez34) [19], a tool that allows users to geocode addresses in batches to obtain latitude, longitude, and geographical details for each address within a Google spreadsheet.

The process was accomplished by purchasing a quota of 13,000 geocoding locations via the ezGeocode app. The app provided precision information on the batch geocoding results, allowing users to assess the accuracy of the data, with extraction of Plus Codes to provide an additional level of precision via a single geocode [19]. The final sample included 11,354 T2D cases, each of which had a valid residential address successfully geocoded to be compatible within the geographic information system environment for spatial overlay to execute spatial regression analyses. The geocoding process flow and the matched rates of geocoded addresses are available in Figure 1.

XSL•FO
RenderX

Figure 1. Flowchart of the geocoding process. Geocode (error) means coding returned an error due to an incomplete address or address computed incorrectly. Geocode (no results) means coding returned no results due to limited address information for location detection or no address information. Batch geocoding that returned multiple results was repeated using a single geocode after manual checking for misspelled addresses or street names and plotting on Google Maps via Plus Codes, thus ensuring 100% accuracy of locations via corrected features. NDR: National Diabetes Register.



Data Sources

We obtained the official state shapefile (ie, the base map of administrative districts and *mukims*) alongside a digital copyright license for permitted use from the Department of Survey and Mapping Penang (code J72499B; issue date: September 14, 2022) to facilitate spatial overlay for the development of cartographic maps and spatial join of attribute data from various sources.

At the state-wide level, geolocated coordinates of neighborhood food outlets were retrieved from the Penang Open Data Portal sourced through the geographic information system–based Penang GeoHub [20] administrated by the local state government of Penang. A total of 1042 geolocated food outlets were divided into 9 different categories based on outlet type (ie, hawker complexes, food complexes, hawker markets, *kopitiams* (a type of coffee shop), fast food outlets, *nasi kandar* (a type of rice dish) restaurants, 24-7 convenience stores, supermarkets, and public markets); all of these food outlets were classified according to the commercial and economic activities stipulated by the Malaysian Standard Industrial Classification coding scheme for food, beverages, household goods, and grocery services [21] (see the complete classifications and harmonization of definitions for the local setting in Penang in Multimedia Appendix 1). Outlets excluded from the analysis included cafeterias and canteens for childcare and educational institutions, bars, taverns, hotels, caterers, delivery-only outlets, dessert shops, theaters, cinemas, hospitals, bistros, clubs, motels, lodges, campgrounds, suppliers of alcohol, and refreshment rooms as these services were not commonly present within the local neighborhoods [22].

Urban population growth rates, defined as urban population shifts between the years of 2010 and 2020 by *mukim* in Penang, were based on MyCensus 2020 data [23]. Population density, defined as the number of inhabitants living in an area per kilometer squared by *mukim*, was based on MyCensus 2020 data [23].

Individual-level data on T2D cases were obtained from the NDR [18]. The NDR is the largest diabetes register in Malaysia, with patients with active T2D captured across 966 primary care clinics at the local level across all states in Malaysia. The NDR is a web-based application systematically capturing patients with active T2D diagnosed by a clinician along with data entered into the system by trained clinic nurses or medical assistants from patient medical records. The registry captures sociodemographic, diagnosis date, complications, and comorbidities of the disease [24]. We acknowledge that data from the NDR rely heavily on documentation by health care providers in the patient case report notes. When data were classified as “unknown” by the registry due to inability to capture them from the clinical documentation [24], they were excluded from the analysis to establish better data accuracy and analysis quality. However, BMI (ie, body weight divided by the square of the height, expressed as kg/m^2) of patients with T2D and address information were not readily available from the registry. To obtain this information, we referred to the source documentation from the medical records of the diagnosed T2D cases in the clinics. We trained research assistants to visit each

XSL•FO
RenderX

clinic after approval from the Penang State Health Department, and with the assistance of a local clinic nurse or medical assistant in charge of managing the NDR in each of the 29 primary care clinics, these research assistants manually traced patient medical records that matched the registration number of T2D cases as captured in the registry. The required individual data (eg, BMI and address information) were extracted manually and tabulated into the master dataset.

Measures

For point-level food outlet proximity analyses, we used T2D case counts by *mukim* in Penang. For density-based analysis, our main outcome measure was *mukim*-wise T2D crude rates computed from aggregated cases reported on the NDR divided by the total adult population aged ≥20 years by *mukim* in Penang. We log10 transformed T2D crude rates to meet the assumption of normal errors and provide a more symmetrical distribution of the outcome and applied empirical Bayes smoothing to allow for fit of a local spatial regression model that predicts T2D burden varying by space [25]. Food outlet coverage was computed as density per 1000 *mukim* residents.

Statistical Analyses

Proximity-Based Measures

The proximity between each T2D case residential address (ie, the locus of analysis) and food outlets was computed by using the "distance to the nearest hub" feature in the QGIS Bialowieza processing toolbox (version 3.22; QGIS Development Team), an algorithm that computes the distance between point features taken as origin and their closest destination line or polygon feature [13,26,27]. Descriptive statistics, alongside their associated measures of dispersion with minimum and maximum values of the distance, were reported.

Next, we executed locally estimated scatterplot smoothing (LOESS) regression. The LOESS regression is a nonparametric regression technique that uses locally weighted regression to fit a smooth curve through points in a scatterplot (ie, associations between proximity to food outlets in kilometers and age and BMI of T2D cases). If the trend line shows a near smooth line of associations within the weighted scatterplot, the associations could be modeled through conventional multivariate linear regression. The proximity to food environments was analyzed as both a continuous and categorical covariate. The following cutoffs for categories were determined: <0.5 km, 0.5 to <1.0 km, and ≥1.0 km. A minimum cutoff of 0.5 km was first set based on previous evidence that this distance is highly walkable within neighborhoods [28], followed by increments of 500 m for subsequent categories up to ≥1.0 km.

Conventional hierarchical multivariable linear regression analyses using the Enter technique [29] were subsequently executed if the LOESS regression showed a relatively smooth linear pattern of distribution in the correlation between 2 variables (ie, proximity to food outlets and age and BMI of T2D cases). The linear regression analysis took the form of the following equation: $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + ... + \beta_pX_p$, where Y is the predicted or expected value of the dependent variable (ie, BMI of T2D cases), X_1 to X_p are p distinct covariates (ie, age and different level of proximity to food outlets), β_0 is the value of Y when all covariates (X_1 to X_p) are equal to 0, and β_1 to β_p are estimated regression coefficients for each of the p parameters. Each regression coefficient represents the change in Y relative to a unit change in the corresponding covariate when all other independent variables are held constant (ie, when the remaining covariates are fixed). The model performance at each step was evaluated by calculating the model base and change statistic parameters. The analysis was conducted using the SPSS software (version 23.0; IBM Corp).

When determining variable selection and fit assessment for models that used conventional regression approaches, covariates were selected based on theoretical and plausible importance. Variables of food outlet proximity were included as these variables have plausible influences on township restructuring for healthy urbanism in preventing T2D. Individual-level factors such as age were included as it was biologically plausible for them to influence T2D. Model fit for linear regression analysis was assessed using R^2 values, and potential multicollinearity between covariates was assessed based on the values of the variance inflation factor exceeding 5. We acknowledge the difficulty of model building for real-life interpretations and acknowledge the dictum of building models with a purpose based on theoretical knowledge that sets the direction of the findings to be adopted for public health policy change [30,31].

Density-Based Measures Across Mukims Using Spatial Weights

We applied multiscale geographically weighted regression (MGWR) to explore spatial heterogeneity in the observed associations among neighborhood food outlet density covariates, urbanization indicators, and logged T2D crude rates across space (hereafter simply referred to as T2D rates). The MGWR model yields a set of locally varying regression coefficients that are mapped and geovisualized to document varying spatial associations between neighborhood covariates and T2D rates across *mukims* [32,33]. We initially fit using a log-linear regression (ie, the ordinary least squares model of log-rates) and, subsequently, using the MGWR (which is the log-linear regression), but in this case, we allowed for a smooth variation of the associations (β values) over space.

MGWR is an extension of geographically weighted regression, which is defined by the following equation:

In this equation, y_i is the value of the outcome variable (ie, log10 of T2D rates by *mukim*) for location i, β_{i0} is the intercept coefficient specific to location i, X_{ik} is the value of the kth explanatory variable specific to location i, β_{ik} denotes the local regression coefficients for the kth explanatory variable specific to location i, and ε_i is the random error terms at location i assumed to follow an independent normal distribution of zero mean and homogeneous variance [34].

The MGWR allows the scale of all covariates of associations with the main outcome variable to be specific to each covariate.

XSL•FO
RenderX

At different spatial locations of the study setting (ie, *mukims* across Penang), these associations may exhibit different impacts on local log-T2D rates. MGWR facilitates spatial variation in estimated associations via a weighting scheme (ie, kernel bandwidth) for each spatial feature using the covariate and main outcome variable of the focal feature areas and their neighbors, with neighboring regions closer to the focal feature area assigned higher weights and larger influence in the local regression model. The geographic scale of correlation is determined via distance-based kernel weights (with either a location-adaptive kernel bandwidth or a simpler fixed bandwidth over space [32]). The kernel weights, defined based on distance between regions (d_{ij}) and the bandwidth h, are given by the following equation [34]:

Bandwidths define the effective number of parameters (ENP) that exhibit patterns for interpretations of spatial heterogeneity, where larger ENP values suggest greater spatial heterogeneity. Specifically, ENP values nearing 1 suggest little spatial variation in estimates of associations in local regressions. Covariate-specific adjusted t values were obtained [35].

This study applied MGWR modeling to food outlet density features and urbanization indicators at the *mukim* level within the state of Penang, allowing estimated covariate associations to change over time and to operate at different spatial scales or features. As the density of food environment features changes over time at different spatial scales (perhaps because underlying township restructuring and urbanization processes change over time and location as well), understanding local variations in coefficient estimates yielded through MGWR can provide more useful interpretations and insights for local public health interventions and policy planning [36]. In this study, we executed MGWR using adaptive kernels via the MGWR software (version 2.2; Spatial Analysis Research Centre, Arizona State University).

Density-Based Measures Within Mukims

We then modeled count outcomes directly by applying a Poisson regression that adjusted for population size by region (ie, *mukims*). A typical Poisson regression is defined by the following equation: $E(Y_i/n_i) = \exp(\beta_0 + \beta_1 \times X_{1i} + ... \beta_p \times X_{pi})$.

On the basis of this equation, the law of probability yields the following final Poisson regression equation: $E(Y_i) = \exp[\ln(n_i) + \beta_0 + \beta_1 \times X_{1i} + ... \beta_p \times X_{pi}]$, where β for the log-population value ($\ln[n_i]$) is fixed and equal to 1, defined as the *offset* variable.

The associations between T2D case counts with an offset variable of log-population for *mukims* in Penang and the density of food outlets was explored through Poisson regression modeling executed through generalized linear models using the GENLIN command in the SPSS software (version 23.0).

The Poisson regression analysis yields β coefficients that, when exponentiated, can be interpreted as approximations to the relative risks (ie, rate ratios [RRs]) associated with the outcome. Next, we used model-predicted counts to define the PAFs associated with each covariate to enhance real-life interpretations for proposing public health policy changes. More specifically, the PAF is defined as the proportional reduction of a population being attributed to a disease if a particular risk factor is removed or eliminated. The PAF was calculated after determining the strength of associations between risk exposures and T2D rates through yielded estimates of adjusted RRs (ARRs) and 95% CIs in the Poisson regression analysis using the formula by Miettinen as follows:

In this formula, *RRadj* is the estimate of the ARR, and *Pc* is the prevalence computed from exposure among the cases [37-39].

Ethical Considerations

Ethics approval was granted by the ethics committee of the National University of Malaysia; the Ministry of Higher Education Malaysia (JEP-2022-445); and the Medical Research Ethics Committee, Ministry of Health Malaysia (ID-22-01264-EE7; investigator initiated research). The registry data used in this study were deidentified and anonymized to protect patients' confidentiality, privacy, and anonymity. No information is presented in this paper that would allow for the direct or indirect identification of individuals. The Medical Research Ethics Committee approval covered permissions for secondary data analysis without additional consent. No financial compensation was provided as this study did not involve human participants in primary data collection.

Results

Proximity of Local Commercial Food Outlets to the Locus of Analysis

A total of 11,354 people with T2D had access to approximately 1042 food outlets across neighborhoods in Penang. The mean distance to any food outlet accessible to patients with T2D was 0.42 (SD 0.515) km, and the distance ranged from 0 to 5.11 km. The median distance was 0.25 (IQR 0.339) km. Being the highest-distributed food outlet type in Penang, hawker markets (n=175) had the shortest distance of accessibility by local neighborhood residents with a mean of 0.69 (SD 0.642) km, and the distance ranged from 0 to 5.11 km. The median distance to the nearest hawker market was 0.50 (IQR 0.593) km. In contrast, food complexes and supermarkets had the lowest distribution throughout the state (n=43 each), and these outlets were the furthest from local neighborhood residents with a mean distance of 6.58 (SD 7.434) and 2.11 (SD 2.240) km, respectively (Table 1).

XSL•FO
RenderX

Table 1. "Distance to the nearest hub" analysis of proximity of food outlets to the neighborhood residences of type 2 diabetes cases (N=11,354).

Outlet type	Outlets, n	Distance (km), mean (SD)	Distance (km), median (IQR)	Range (km)
Any food outlet	1042	0.42 (0.515)	0.25 (0.339)	0-5.11
Hawker complex	116	1.17 (1.168)	0.72 (1.187)	0-7.69
Food complex	43	6.58 (7.434)	3.31 (11.074)	0-33.31
Hawker market	175	0.69 (0.642)	0.50 (0.593)	0-5.11
Kopitiam	106	1.87 (1.906)	1.11 (1.580)	0-12.30
Nasi kandar restaurant	169	1.03 (1.286)	0.59 (0.828)	0-11.40
Fast food outlet	174	1.63 (1.742)	0.90 (1.473)	0-11.30
24-7 convenience store	159	1.07 (1.341)	0.60 (0.851)	0-10.20
Public market	57	1.32 (1.303)	0.83 (1.396)	0.12-8.80
Supermarket	43	2.11 (2.240)	1.17 (2.155)	0.02-13.06

Multivariable Linear Regression Analysis on the Associations Between Proximity to Food Outlets and Age and BMI of T2D Cases

The LOESS regression interpolation associations between BMI and food outlet type showed a relatively smooth fit line indicative of linear correlations, and we focus on linear regressions in the following paragraphs (Multimedia Appendices 2-4).

Table 2 shows the results of multivariable linear regression analysis on the associations between proximity to food outlets and age and BMI of T2D cases. A total of 21 covariates accounted for 6.5% of the total variance for the BMI values of T2D cases within neighborhoods in Penang. The base covariates in step 1 accounted for 1.1% of the variance in the BMI of T2D cases within neighborhoods ($F_{6,11,046}$ change=21.061), and among those variables, proximity to food complexes of <0.5 km and 0.5 to <1 km (β=−1.130 and −0.756, respectively; *P*<.001) and proximity to hawker markets of <0.5 km and 0.5 to <1 km (β=−0.325 and −0.498, respectively; *P*=.02 and *P*<.001) showed statistically significant negative associations with the BMI of T2D cases. The associations between BMI and proximity to hawker complexes were not statistically significant (*P*>.05; Table 2).

The total variance explained increased slightly by 0.3% after inclusion of *kopitiam*, *nasi kandar* restaurant, and fast food outlet proximity in the second step ($F_{6,11,040}$ change=5.007), and the statistically significant variables associated with BMI included proximity to food complexes of <0.5 km and 0.5 to <1 km (β=−1.073 and −0.685, respectively; *P*<.001), proximity to hawker markets of <0.5 km (β=−0.308; *P*=.03) and 0.5 to <1 km (β=−0.481; *P*=.002), proximity to *kopitiams* of <0.5 km (β=0.303; *P*=.049) and 0.5 to <1 km (β=0.502; *P*=.001), and proximity to fast food outlets of <0.5 km and 0.5 to <1 km (β=−0.694 and −0.660, respectively; *P*<.001). Associations between BMI and proximity to hawker complexes and *nasi kandar* restaurants were not statistically significant (*P*>.05; Table 2).

Adding covariates of 24-7 convenience stores, public markets, and supermarkets in the third step contributed minimally to an additional 0.1% of the explained variance in the BMI of T2D cases ($F_{6,11,034}$ change=1.382). The statistically significant covariates associated with the BMI of T2D cases were proximity to food complexes of <0.5 km and 0.5 to <1 km (β=−1.041 and −0.649, respectively; *P*<.001), proximity to hawker markets of 0.5 to <1 km (β=−0.444; *P*=.006), proximity to *kopitiams* of <0.5 km (β=0.348; *P*=.04) and 0.5 to <1 km (β=0.560; *P*=.001), proximity to fast food outlets of <0.5 km (β=−0.605; *P*<.001) and 0.5 to <1 km (β=−0.576; *P*=.001), and proximity to 24-7 convenience stores of <0.5 km (β=−0.376; *P*=.03) and 0.5 to <1 km (β=−0.438; *P*=.01). None of the associations between BMI and proximity to hawker complexes; proximity to hawker markets of <0.5 km; and proximity to *nasi kandar* restaurants, public markets, or supermarkets were statistically significant (*P*>.05; Table 2).

Finally, adding age covariates explained an additional 5% of the variance in the BMI of T2D cases ($F_{3,11,031}$ change=198.191). The final step in the model found that proximity to food complexes of <0.5 km (β=−0.855; *P*<.001) and 0.5 to <1 km (β=−0.449; *P*=.004); proximity to hawker markets of <0.5 km (β=−0.308; *P*<.04) and 0.5 to <1 km (β=−0.474; *P*=.002); proximity to *kopitiams* of 0.5 to <1 km (β=0.551; *P*=.001); proximity to fast food outlets of <0.5 km (β=−0.457; *P*=.005) and 0.5 to <1 km (β=−0.536; *P*=.001); proximity to 24-7 convenience stores of 0.5 to <1 km (β=−0.348; *P*=.04); proximity to public markets of <0.5 km (β=0.344; *P*=.01); and ages of 20 to 34 years, 35 to 49 years, and 50 to 64 years (β=4.180, 3.221, and 1.386, respectively; *P*<.001) associated with BMI change were statistically significant. Associations between BMI and proximity to hawker complexes, proximity to *kopitiams* of <0.5 km, proximity to *nasi kandar* restaurants, proximity to public markets of 0.5 to <1 km, and proximity to supermarkets were not statistically significant (*P*>.05; Table 2).

XSL•FO
RenderX

Table 2. Multivariable linear regression analysis on the associations between proximity to food outlets and age and BMI of type 2 diabetes cases (N=11,047)[a].

Characteristic	Step 1[b]			Step 2[c]			Step 3[d]			Step 4[e]		
	β (SE)	*P* value	VIF[f]	β (SE)	*P* value	VIF	β (SE)	*P* value	VIF	β (SE)	*P* value	VIF
Hawker complex at <0.5 km	−0.028 (0.131)	.83	1.641	0.115 (0.143)	.42	1.955	0.109 (0.144)	.45	1.974	0.173 (0.140)	.22	1.975
Hawker complex at 0.5 to <1 km	−0.247 (0.147)	.09	1.534	−0.062 (0.156)	.69	1.725	−0.055 (0.157)	.73	1.750	0.015 (0.153)	.92	1.751
Food complex at <0.5 km	−1.130 (0.139)	<.001[g]	1.337	−1.073 (0.140)	<.001[g]	1.375	−1.041 (0.151)	<.001[g]	1.582	−0.855 (0.147)	<.001[g]	1.587
Food complex at 0.5 to <1 km	−0.756 (0.148)	<.001[g]	1.303	−0.685 (0.151)	<.001[g]	1.363	−0.649 (0.158)	<.001[g]	1.492	−0.449 (0.154)	.004[g]	1.496
Hawker market at <0.5 km	−0.325 (0.142)	.02[g]	1.985	−0.308 (0.145)	.03[g]	2.096	−0.280 (0.150)	.06	2.236	−0.308 (0.146)	.04[g]	2.237
Hawker market at 0.5 to <1 km	−0.498 (0.154)	.001[g]	1.842	−0.481 (0.156)	.002[g]	1.894	−0.444 (0.161)	.006[g]	2.001	−0.474 (0.156)	.002[g]	2.001
Kopitiam at <0.5 km	—[h]	—	—	0.303 (0.154)	.049[g]	1.710	0.348 (0.166)	.04[g]	1.981	0.264 (0.162)	.10	1.982
Kopitiam at 0.5 to <1 km	—	—	—	0.502 (0.154)	.001[g]	1.619	0.560 (0.167)	.001[g]	1.911	0.551 (0.163)	.001[g]	1.911
Nasi kandar restaurant at <0.5 km	—	—	—	0.002 (0.153)	.99	2.324	0.068 (0.163)	.68	2.642	−0.026 (0.159)	.87	2.644
Nasi kandar restaurant at 0.5 to <1 km	—	—	—	−0.162 (0.160)	.31	1.937	−0.069 (0.171)	.69	2.220	−0.134 (0.167)	.42	2.220
Fast food outlet at <0.5 km	—	—	—	−0.694 (0.154)	<.001[g]	1.900	−0.605 (0.167)	<.001[g]	2.237	−0.457 (0.163)	.005[g]	2.241
Fast food outlet at 0.5 to <1 km	—	—	—	−0.660 (0.154)	<.001[g]	1.803	−0.576 (0.166)	.001[g]	2.094	−0.536 (0.162)	.001[g]	2.095
24-7 convenience store at <0.5 km	—	—	—	—	—	—	−0.376 (0.172)	.03[g]	2.924	−0.289 (0.168)	.09	2.926
24-7 convenience store at 0.5 to <1 km	—	—	—	—	—	—	−0.438 (0.174)	.01[g]	2.387	−0.348 (0.170)	.04[g]	2.389
Public market at <0.5 km	—	—	—	—	—	—	0.208 (0.138)	.13	1.718	0.344 (0.135)	.01[g]	1.722
Public market at 0.5 to <1 km	—	—	—	—	—	—	0.117 (0.148)	.43	1.548	0.187 (0.145)	.20	1.549
Supermarket at <0.5 km	—	—	—	—	—	—	−0.067 (0.168)	.69	1.910	−0.053 (0.164)	.75	1.911
Supermarket at 0.5 to <1 km	—	—	—	—	—	—	−0.031 (0.167)	.85	1.892	0.005 (0.162)	.97	1.893
Age of 20 to 34 y	—	—	—	—	—	—	—	—	—	4.180 (0.293)	<.001[g]	1.098
Age of 35 to 49 y	—	—	—	—	—	—	—	—	—	3.221 (0.144)	<.001[g]	1.493
Age of 50 to 64 y	—	—	—	—	—	—	—	—	—	1.386 (0.120)	<.001[g]	1.512

[a]The reference for the age variable was those aged ≥65 years; the reference for proximity to all food outlets was ≥1 km.

[b]$F_{6,11,046}$ change=21.061; R^2=0.011; R^2 change=0.011.

[c]$F_{6,11,040}$ change=5.007; R^2=0.014; R^2 change=0.003.

[d]$F_{6,11,034}$ change=1.382; R^2=0.015; R^2 change=0.001.

[e]$F_{3,11,031}$ change=198.191; R^2=0.065; R^2 change=0.050.

[f]VIF: variance inflation factor.

XSL•FO
RenderX

[g]Statistical significance at $P<.05$.

[h]Not applicable.

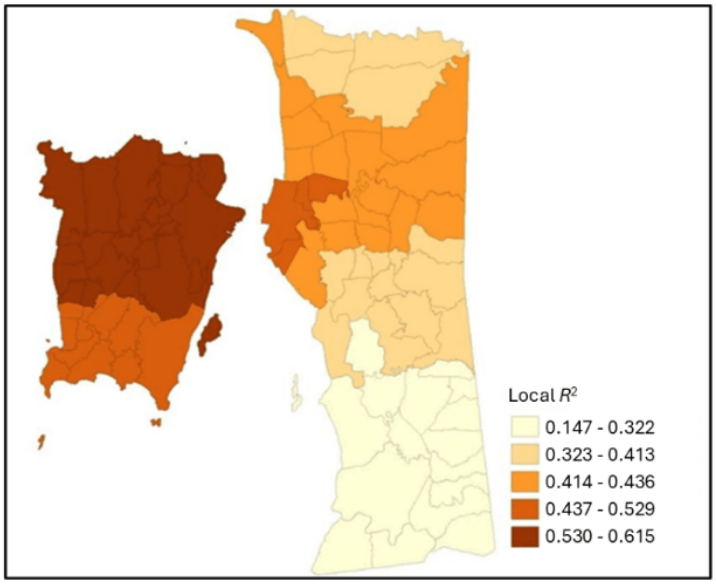
Local Associations Among Neighborhood Food Environments, Urban Indicators, and T2D Rates by *Mukim* in Penang

The MGWR model concluded that variations in local logged T2D rates were positively associated with the density of *kopitiams* per 1000 residents (β=0.256), density of 24-7 convenience stores per 1000 residents (β=0.028), density of supermarkets per 1000 residents (β=0.122), density of public markets per 1000 residents (β=0.067), urban growth rate (β=0.189), and population density (β=0.080). In contrast, the MGWR model found that variations in local logged T2D rates were negatively associated with the density of fast food outlets per 1000 residents (β=−0.061) and the density of *nasi kandar* restaurants per 1000 residents (β=−0.064; $t_{65.835}$≥1.96; Multimedia Appendix 5).

Variation of Pooled Factors by Location

The local R^2 values from the MGWR model are mapped in Figure 2. The local R^2 values indicate how well the modeled covariates (ie, neighborhood food outlet density and urban indicators) explain the variation in T2D rates in different *mukims* in Penang. Higher R^2 values in certain *mukims* suggest that the model performed well for those locations (ie, the covariates significantly influenced T2D rates in those areas), whereas lower R^2 values suggest that the model explained less of the variance, suggesting that there are other unaccounted for factors that may be influencing T2D rates in those areas. The largest values, ranging from 0.53 to 0.62, were observed across *mukims* in Timur Laut and half of all *mukims* in the Barat Daya district, suggesting that the model has significant predictive power in those areas. Overall, local R^2 values ranged from 0.15 to 0.62 across *mukims* in Penang, but lower values were predominantly observed in mainland Penang, particularly across *mukims* in Seberang Perai Selatan and parts of *mukims* in Seberang Perai Tengah and Seberang Perai Utara.

Figure 2. Local R^2 values of pooled covariates.



Variation of Each Factor That Influenced T2D Rates by Location

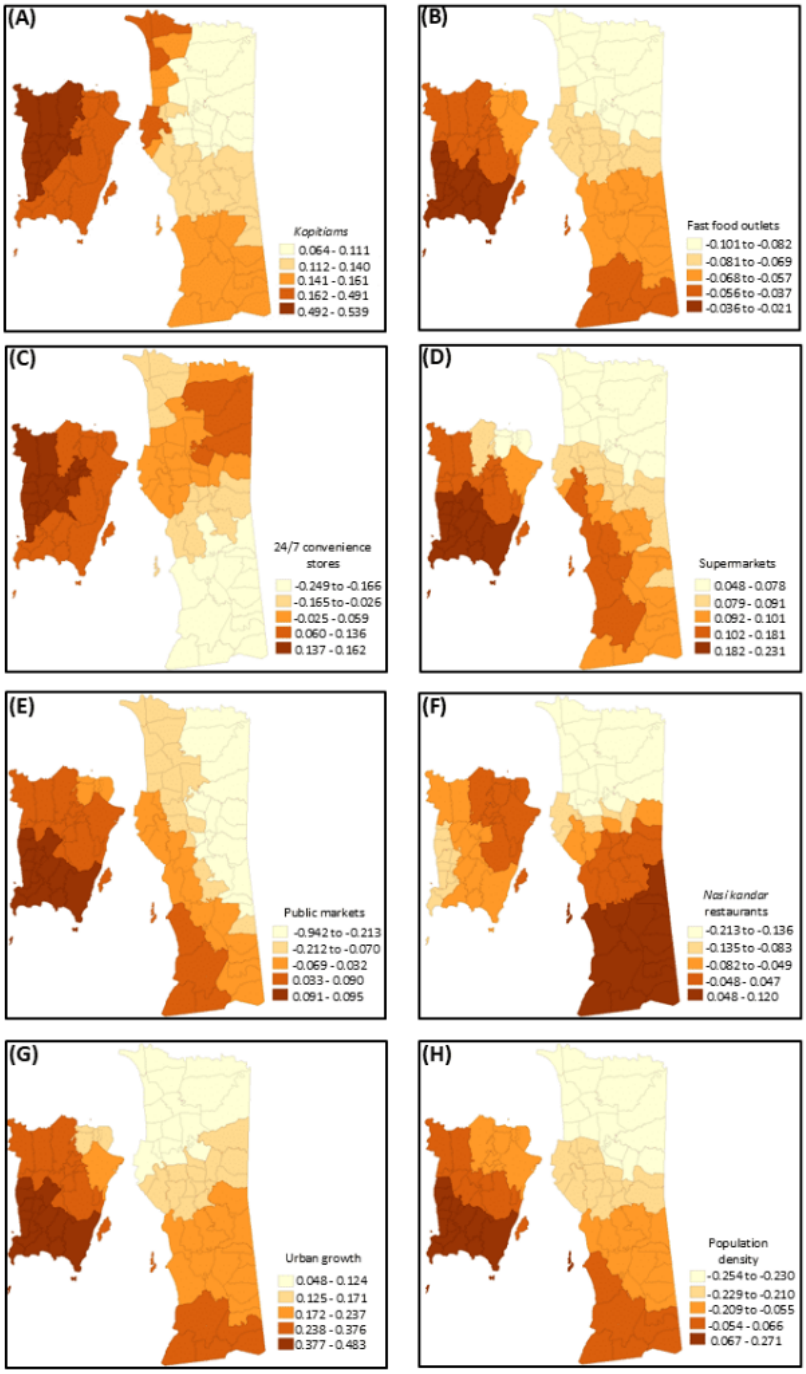
The β values for the locally adjusted covariates are visualized via quantile maps in Figures 3A-3H. The local MGWR model showed that the density of *kopitiams* per 1000 residents was positively associated with logged T2D rates across *mukims* in almost half of the Barat Daya district (β range 0.492-0.539) followed by *mukims* in the entire Timur Laut district and part of the southern Barat Daya district and some *mukims*, particularly *mukims* 1, 2, 4, and 15, in the Seberang Perai Utara district (β range 0.162-0.491; Figure 3A). In contrast, the density of fast food outlets per 1000 residents was associated with lower logged T2D rates across all *mukims* in Penang, and this negative association was the strongest in almost half of all *mukims* in the Barat Daya district (β range −0.036 to −0.021; Figure 3B).

The density of 24-7 convenience stores per 1000 residents showed the highest associations of logged T2D rates across *mukims* in the west of the Barat Daya district (β range 0.137-0.162) followed by *mukims* in the entire Timur Laut district and the southern Barat Daya district (β range 0.060-0.136). Overall, coefficient estimates ranged from −0.249 to 0.162 with relatively lower logged T2D rates observed across *mukims* within the Seberang Perai Selatan district (Figure 3C). Supermarket density per 1000 residents showed positive relationships with logged T2D rates across all *mukims* in Penang (β range 0.048-0.231); the highest associations with logged T2D rates were found in almost half of the Barat Daya district (β range 0.182-0.231), whereas the lowest associations with logged

XSL•FO
RenderX

T2D rates were distributed across *mukims* in the Seberang Perai Utara district (β range 0.048-0.078; Figure 3D).

Figure 3. β quantile maps for (A) density of kopitiams per 1000 residents, (B) density of fast food outlets per 1000 residents, (C) density of 24/7 convenience stores per 1000 residents, (D) density of supermarkets per 1000 residents, (E) density of public markets per 1000 residents, (F) density of nasi kandar restaurants per 1000 residents, (G) urban growth rate, and (H) population density.



The density of public markets per 1000 residents was associated with logged T2D rates across *mukims* in almost half of the Barat Daya district (β range 0.091-0.095) followed by *mukims* in the northwest of the Barat Daya district, most *mukims* in the Timur Laut district, and *mukims* in part of the Seberang Perai Selatan district (β range 0.033-0.090); however, the presence of public markets had lower strength of association with logged T2D rates across most *mukims* in the Seberang Perai Utara and Seberang Perai Tengah districts (Figure 3E). Except for *mukims* in the Seberang Perai Selatan district that showed positive relationships between logged T2D rates and the density of *nasi kandar* restaurants per 1000 residents (β range 0.048-0.120), the rest of the *mukims* in Penang showed negative relationships between logged T2D rates and the density of *nasi kandar* restaurants per 1000 residents (Figure 3F).

With respect to urbanization indicators, urban growth was positively associated with logged T2D rates across all *mukims* in Penang, with the highest associations with logged T2D rates being observed across *mukims* in the south of the Barat Daya district (β range 0.377-0.483; Figure 3G). A similar positive association was observed between population density and logged T2D rates, with the highest association with logged T2D rates

XSL•FO
RenderX

observed in the south of the Barat Daya district (β range 0.067-0.271), although the rest of the *mukims* across Penang showed negative associations (Figure 3H).

Estimated Associations Between Density of Food Outlets and Area-Level T2D Cases Within *Mukims* Through Univariate Poisson Regression

Table 3 shows the associations between density of different types of food outlets per 1000 population within *mukims* and area-level T2D cases through univariate Poisson regression. The density of food complexes per 1000 residents (RR=1.63, 95% CI 1.55-1.71; *P*<.001), hawker markets per 1000 residents (RR=1.84, 95% CI 1.68-2.01; *P*<.001), and fast food outlets per 1000 residents (RR=1.14, 95% CI 1.05-1.23; *P*=.002) increased the risk of T2D as compared to the absence of such food outlets. The density of *nasi kandar* restaurants per 1000 residents (RR=0.64, 95% CI 0.61-0.67; *P*<.001), 24-7 convenience stores per 1000 residents (RR=0.80, 95% CI 0.75-0.87; *P*<.001), and supermarkets per 1000 residents (RR=0.81, 95% CI 0.78-0.84; *P*<.001) decreased the risk of T2D as compared to the absence of such food outlets (Table 3).

Table 3. The association between density of food outlets and area-level type 2 diabetes cases within mukims through univariate Poisson regression[a].

Characteristic	Exponentiated Poisson regression coefficient (β crude; 95% CI)	*P* value
Density of hawker complexes (per 1000 residents)		.45
At least one outlet	0.98 (0.94-1.03)	
No outlet[b]	1	
Density of food complexes (per 1000 residents)		<.001[c]
At least one outlet	1.63 (1.55-1.71)	
No outlet	1	
Density of hawker markets (per 1000 residents)		<.001[c]
At least one outlet	1.84 (1.68-2.01)	
No outlet	1	
Density of *kopitiams* (per 1000 residents)		.06
At least one outlet	1.06 (0.99-1.11)	
No outlet	1	
Density of *nasi kandar* restaurants (per 1000 residents)		<.001[c]
At least one outlet	0.64 (0.61-0.67)	
No outlet	1	
Density of fast food outlets (per 1000 residents)		.002[c]
At least one outlet	1.14 (1.05-1.23)	
No outlet	1	
Density of 24-7 convenience stores (per 1000 residents)		<.001[c]
At least one outlet	0.80 (0.75-0.87)	
No outlet	1	
Density of public markets (per 1000 residents)		.15
At least one outlet	0.97 (0.93-1.01)	
No outlet	1	
Density of supermarkets (per 1000 residents)		<.001[c]
At least one outlet	0.81 (0.78-0.84)	
No outlet	1	

[a]β provides the crude rate ratio.

[b]"No outlet" represents the reference category and does not have a 95% CI.

[c]Statistical significance at *P*<.05.

Estimated Associations Between Density of Food Outlets and Area-Level T2D Cases Within *Mukims*

XSL•FO
RenderX

Through Multivariate Poisson Regression Complemented With PAFs

Table 4 shows the association between the density of different types of food outlets per 1000 residents within *mukims* and area-level T2D cases through multivariate Poisson regression complemented with PAFs. For each of the covariates with statistically significant associations, exponentiating the estimated coefficient provided the multiplicative change in the number of T2D cases in the presence of hawker complexes (ARR=0.94, 95% CI 0.89-0.99; *P*=.03), food complexes (ARR=1.24, 95% CI 1.20-1.29; *P*<.001), hawker markets (ARR=1.32, 95% CI 1.22-1.43; *P*<.001), *kopitiams* (ARR=1.09, 95% CI 0.99-1.20; *P*=.08), *nasi kandar* restaurants (ARR=0.63, 95% CI 0.58-0.68; *P*<.001), fast food outlets (ARR=0.92, 95% CI 0.88-0.95; *P*<.001), 24-7 convenience stores (ARR=0.72, 95% CI 0.69-0.75; *P*<.001), public markets (ARR=0.92, 95% CI 0.86-0.97; *P*=.005), and supermarkets (ARR=0.72, 95% CI 0.66-0.78; *P*<.001) per 1000 residents.

Table 4. Estimated associations between density of food outlets and area-level type 2 diabetes cases within mukims through multivariate Poisson regression complemented with population-attributable fractions (PAFs)[a].

Characteristic	Exponentiated Poisson regression coefficient (β adjusted; 95% CI)	*P* value	PAF (%)
Density of hawker complexes (per 1000 residents)		.03[c]	−0.07
At least one outlet	0.94 (0.89-0.99)		
No outlet	1[b]		
Density of food complexes (per 1000 residents)		<.001[c]	0.21
At least one outlet	1.24 (1.20-1.29)		
No outlet	1		
Density of hawker markets (per 1000 residents)		<.001[c]	0.27
At least one outlet	1.32 (1.22-1.43)		
No outlet	1		
Density of *kopitiams* (per 1000 residents)		.08	0.09
At least one outlet	1.09 (0.99-1.20)		
No outlet	1		
Density of *nasi kandar* restaurants (per 1000 residents)		<.001[c]	−0.64
At least one outlet	0.63 (0.58-0.68)		
No outlet	1		
Density of fast food outlets (per 1000 residents)		<.001[c]	−0.10
At least one outlet	0.92 (0.88-0.95)		
No outlet	1		
Density of 24-7 convenience stores (per 1000 residents)		<.001[c]	−0.43
At least one outlet	0.72 (0.69-0.75)		
No outlet	1		
Density of public markets (per 1000 residents)		.005[c]	−0.10
At least one outlet	0.92 (0.86-0.97)		
No outlet	1		
Density of supermarkets (per 1000 residents)		<.001[c]	−0.43
At least one outlet	0.72 (0.66-0.78)		
No outlet	1		

[a]β provides the adjusted rate ratio.

[b]"No outlet" represents the reference category and does not have a 95% CI.

[c]Statistical significance at *P*<.05.

With regard to parameter interpretations of PAFs, it was found that, if local neighborhoods (within *mukims*) underwent township restructuring to remove food complexes, hawker markets, or *kopitiams*, the result would be an estimated reduction of 0.21%, 0.27%, and 0.09% in the number of *mukim*-level T2D cases, respectively. In contrast, if the local neighborhoods

XSL•FO
RenderX

(within *mukims*) underwent township restructuring to add hawker complexes, *nasi kandar* restaurants, fast food outlets, 24-7 convenience stores, public markets or supermarkets, the result would be an estimated reduction of 0.07% to 0.64% in the number of residents within those *mukims* having T2D (Table 4).

Discussion

Principal Findings

Our aim was to first assess how close commercial food establishments are to the residences of individuals with T2D within neighborhoods; subsequently, explore how rates of T2D differ by location; and, finally, evaluate the extent to which the density of specific food outlets in a neighborhood needs to be modified to manage local T2D rates. Our analysis revealed variations in the associations between proximity to food outlets and the BMI of T2D cases when different food outlets were populated in the hierarchical model adjusted for age, with model performance ranging from 1.1% to 6.5%. The MGWR β coefficients showed how different food outlet density varied by space, indicating areas that should be prioritized for local public health interventions. Finally, the PAFs suggested that, if local neighborhoods underwent township restructuring to remove food complexes, hawker markets, or *kopitiams*, an estimated reduction of 0.21%, 0.27%, and 0.09%, respectively, in the risk of T2D cases in Penang would be anticipated but, if local neighborhoods underwent township restructuring to add hawker complexes, *nasi kandar* restaurants, fast food outlets, 24-7 convenience stores, public markets or supermarkets, an estimated reduction of between 0.07% and 0.64% in the number of residents with risk of T2D was estimated. These findings provide a better overview of a local neighborhood landscape regarding where interventions should be focused to tackle local T2D burden.

Comparison With Prior Work

Our linear regression approach assisted us in understanding how much variation persists and how food environments change with neighborhood development and urbanization processes. In the first step, if the neighborhood was only structured with access to hawker food outlets, the likely change in the BMI of T2D cases would account for approximately 1.1% of the variance in the model, whereas in the subsequent step, if the town was further structured or developed to facilitate accessibility to *kopitiams*, *nasi kandar* restaurants, and fast food outlets, the change in BMI among T2D cases was likely to account for approximately 1.4% of the variance. If the town became more populated with convenience stores, public markets, and supermarkets for access by local communities, the likely change in BMI would account for approximately 1.5% of the variance. Finally, to make it more sensible as towns are inhabited by individuals of different demographics, the final model that adjusted for age yielded a huge increase in the BMI change among T2D cases, accounting for 6.5% of the overall variance. These findings highlight that proximity to food outlets was associated with BMI concurrently mediated by age [40-43], perhaps based on different food preferences and ease of access to neighborhood food outlets as perceived by different age groups.

In the final model, proximity to food complexes and hawker markets within a kilometer showed negative associations with BMI, whereas proximity to *kopitiams* within 1 km showed positive associations with BMI. These local eateries offer different food choices and are common venues for local communities to access for breakfast, lunch, or dinner [44]. Similarly, while proximity to *kopitiams* was positively associated with BMI change in T2D cases, the direction of the association was negative for food complexes and hawker markets, likely attributed to the local food classifications or subject to the effect of the spatial scale applied for this investigation [45].

Proximity to fast food outlets showed negative associations with BMI, consistent with the findings reported by Jeffery et al [46] but inconsistent with those reported by Kusuma et al [11] and van Erpecum et al [28]. The negative association between fast food outlet proximity and BMI mirrors some earlier results [47]; however, it is important to consider the ecologic nature of our estimates and the small local sample sizes involved in estimating local impacts of fast food access [48]. Supermarket proximity showed no statistically significant relationship with BMI change, inconsistent with the work by Garg et al [49]. Proximity to 24-7 convenience stores and BMI change showed negative associations, with the literature showing variations across different geographical areas [50,51]. Proximity to public markets unexpectedly showed positive associations with BMI change; however, such a finding was also anticipated in a previous study that suggested that low numbers of public markets in different areas could be associated with a nonprotective effect of public market accessibility on BMI change among individuals with T2D [52].

The local bandwidths of the MGWR model showed that associations among local T2D prevalence rates; the density of fast food outlets, supermarkets, public markets, and *nasi kandar* restaurants; and urban growth occurred at a global scale, whereas the association with the density of *kopitiams* and 24-7 convenience stores occurred at a regional scale and, finally, associations between T2D prevalence and population density occurred at the local scale. The adjusted R^2 quantile maps and β weight quantile maps from the MGWR regression could provide clues for priority setting [35,53]. This study found higher local R^2 values from the pooled covariates in the MGWR model that ranged from 0.53 to 0.62 covering all *mukims* in the Timur Laut district and half of all *mukims* in the Barat Daya district. These areas are urbanized towns with higher population density nested within neighborhoods with high densities of food establishments offering food deemed to be *unhealthy*. The overall adjusted R^2 values ranged from 0.15 to 0.62, with lower values observed across *mukims* in mainland Penang, particularly in Seberang Perai Selatan and parts of *mukims* in Seberang Perai Tengah and Seberang Perai Utara. These areas show lower urban growth and population density, with food establishments operating according to community needs or demands. The β weight coefficient maps obtained from the MGWR results support the trend of highly concentrated *kopitiams* and 24-7

XSL•FO
RenderX

convenience stores increasing T2D rates in almost half of the *mukims* in the Barat Daya district, whereas high concentrations of supermarkets and public markets in parts of *mukims* across the Barat Daya and Timur Laut districts showed positive correlations with T2D rates alongside increased urban growth and population density as compared to the more rural areas in the mainland. Examining relationships across diverse geographical settings is a complex endeavor, as in this study, where areal urbanization processes vary across the most urbanized towns to suburban or small towns and, finally, rural settings, where patterns of food establishments and their correlation with T2D could vary substantially [54].

Previous works have identified supermarkets as having more diverse and healthy food options as compared to convenience stores; however, the availability of the latter in rural settings or areas of lower urban growth as the main source of food purchases may increase local communities' risk of T2D [55,56]. Nevertheless, fast food outlet concentrations in those areas showed negative correlations with T2D rates, and the plausibility of such associations could be attributed to classifications of definitions from the available sources [7]. Except for *mukims* in the Seberang Perai Selatan district that showed positive relationships between T2D rates and the density of *nasi kandar* restaurants, the rest of the areas in Penang showed negative relationships. The results of the MGWR analysis were in general alignment with the global model; however, the local insights of the MGWR estimates pose questions as to where significant associations occur, the local strength of such associations, and how these could be prioritized for local interventions. A similar argument was put forward in a previous study that used MGWR [35].

Our work illustrates how the neighborhood food environment influences T2D rates through local patterning. Traditionally, health care sector prevention efforts have frequently directed interventions toward diet, weight loss, or lifestyle [12,57], but these interventions are difficult and a costly enterprise to upscale to population-level interventions [58]. Despite these challenges, policy advocates and decision makers are aware of the need for coordinated multisectoral prevention action to promote health within neighborhoods throughout the human life course. Unifying multiple stakeholders or decision makers requires a single convincing metric to initiate preventive actions in what our results show is a complicated system with multiple moving parts. One metric considered in this work was PAFs [45,59], allowing for the epidemiological interpretation that if local neighborhoods underwent township restructuring to remove food complexes, hawker markets, or *kopitiams*, an estimated reduction of 0.21%, 0.27%, and 0.09%, respectively, in the number of T2D cases would be anticipated. Similarly, if local neighborhoods underwent township restructuring to add hawker complexes, *nasi kandar* restaurants, fast food outlets, 24-7 convenience stores, public markets or supermarkets, an estimated reduction of 0.07% to 0.64% in the number of residents with T2D was estimated. While the former was convincing, the latter could be debatable because of the local classification of those food establishments [7] and the spatial scale that was used [45].

Policy Implications

Our 3 focused objectives yielded 3 crucial metrics that set the foundation for public health policymaking in the quest to reduce local T2D burden in Penang. First, our hierarchical multivariable linear regression showed how the distance of different food outlets from the locus of analysis was associated with change in the BMI of T2D cases in Penang. We documented the strength of association and model performance when a particular food outlet was added in a hierarchical way to the model adjusted for age—keeping in mind how neighborhoods undergo restructuring in line with urbanization processes—thus allowing for interpretations regarding dietary choices and accessibility among different age groups. Proximity measures may influence zoning decisions by local authorities to facilitate food outlet licensing within neighborhoods while creating opportunities to advocate for healthier food options.

Second, across neighborhoods, the area-level local MGWR β weights and their residuals and bandwidths revealed how the factors pooled together were associated with the concentration of T2D rates by areal *mukims* and, if interventions or policies were planned to be executed, whether they should be targeted at one specific focal area or be applied to the entire population in that area. For example, if the density of certain food outlets were found to be associated significantly with area-level T2D rates across *mukims*, local city councils, local assemblypersons, township neighborhood planners, or local community boards could advocate for the relocation or removal of those food outlets, limit license approval for newer establishments of similar food outlets that could have detrimental effects on health, or even plan community health education programs on healthy food purchase. This approach could minimize the harmful effects of commercial determinants of health such as unhealthy food promotion or availability to neighborhood residents.

The MGWR bandwidth metrics convey important messages for public health responses, providing evidence regarding at what scale and for what targets local programs or interventions should be implemented based on their predicted effect magnitude, as well as providing insight into how interventions should be planned, implemented, and evaluated. We recommend an interplay between health promotion (ie, how local health departments, community organizations, and policy makers could foster the overall well-being of communities by addressing the commercial determinants of food environments, such as restructuring the town) and health education (ie, how community health leaders, health care providers, or government agencies could impart knowledge or skills to promote healthy behaviors, such as community programs on healthy food purchase or implementation of local subsidized voucher programs for purchasing healthy food options) interventions [60]. The former would be more appropriate to tackle associations at a global or regional scale, whereas the latter would be more appropriate for a local scale.

Alongside the synthesis of MGWR β weights and bandwidths, the local R^2 values could assist public health policy makers in understanding where certain interventions (eg, food outlet relocation, township restructuring, or targeted community-based behavioral interventions) need to be prioritized. When the

XSL•FO
RenderX

model's predictive power is higher across specific regions or areas, intervening in local factors with associated high associations and in areas exhibiting high local *R*-squared values, it would be wise for additional investigations to be conducted first to uncover other potential risk factors before interventions are implemented for optimal outcomes that provide value for money.

While the RRs estimated changes in the local T2D burden in conjunction with the local measures of food outlet density within a neighborhood, the PAFs provide another important interpretable metric for policy makers for food environment interventions; that is, if local neighborhoods (within *mukims*) undergo township restructuring to remove certain food premises, what would be the estimated reduction in T2D cases within those *mukims*. These metrics complement the findings from the MGWR (ie, measures across *mukims* or their neighbors), but in this case, PAFs are interpretable within *mukims* (ie, without their neighbors), which motivates a more holistic approach useful for policy planning in developing or restructuring towns to include healthy urbanism neighborhood features.

Limitations

We note that our interpretations are not causal but associative (ie, using ecological and cross-sectional designs); however, these designs are relatively suitable to frame our regional findings for policy implications. We opted for an ecological study design that can be subject to ecological fallacy (ie, interpretation of risk factors observed at the group level that cannot be individualized). While this study acknowledged this limitation by convention, it should be considered from a geographical point of view by interpreting where and why T2D occurs as the aim of any spatial epidemiological study is to provide evidence by place for targeted interventions. We frame our results regionally as many public health policies are usually made at the aggregate level involving populations or communities within neighborhoods.

We note that MGWR is a regression model that assumes that the residuals follow a normal distribution. To see whether we had independent observations that violated Tobler's First Law of Geography, we fit a standard linear ordinary least squares regression with the log-rates as outcome variables using independent, normally distributed error terms and compared model fit to guard against this potential limitation. As an alternative, we propose to fit a Poisson regression with spatially correlated random effects or the integrated nested Laplace approximation for similar data analysis in the future.

Next, we acknowledge that, while addresses provide point locations for residences and food outlets, we are required to preserve patient confidentiality and anonymity. One limitation of our analysis is that it currently does not allow each individual's mobility and preferences regarding food purchase to result in dining patterns covering a broader area than the closest food outlets. Although our work focused on small-area summaries of both our outcome and the exposure to nearby food outlets, our results can motivate future research to explore finer-scale associations between point patterns of residence and mobility zones using point process summaries such as kernel density estimates [25].

In addition, although there is evidence suggesting that ethnicity plays a critical role in determining associations with T2D at an individual level, this scenario may not result in the same associations with risk measured at an area level as different proportions of ethnic groups (ie, either ethnic minority or majority groups) may reside across different settings. Under these circumstances, we did not include ethnic proportions by areas as a potential variable to be adjusted for in any models as this is not a local feature that can be changed via policies. Future work can adopt a more nuanced approach by exploring interactions between local policy impacts and local analysis interpretations stratified by ethnicity.

Finally, while we established the correlation between T2D rates and the density of food outlets by administrative levels, we caution that our approach and analyses may be subject to the modifiable areal unit problem (ie, a geographic manifestation of the ecological fallacy where estimated associations may differ when translated to smaller areal units, such as local authority areas, or larger areal units, such as administrative districts) when it is replicated within different administrative boundaries [13]. We also note that, based on the establishment type and the food choices being offered, the estimates for PAFs were relatively low for hawker complexes and fast food outlets. Under modern commercialization landscapes, it would be difficult to eliminate such establishments completely; however, we could first reduce their density of availability, which could result in reduced health impacts on local neighborhoods. Another relevant query is to find what spatial extent of the densities of food outlets is relevant to dietary behaviors that, in turn, could accelerate the risk of T2D, a scenario known as the uncertain geographic context problem [61] that is difficult to adjust for in spatial studies.

Conclusions

We derived 3 important indicators to engage local epidemiologists, township planners, and public health policy advocates in setting priorities for T2D control in Penang. Our results showed how BMI and age change with different proximities to food outlets among residents with T2D, followed by how the density of different food outlets in different spaces was correlated at different levels with T2D rates across space in Penang. We next explored how much the town's neighborhood food environment should be restructured to reduce the incidence of T2D in Penang. Our findings suggest that areas with high T2D rates and different food outlet densities vary by *mukim* and coupled with the relative ease of access adjusting for age should be considered for targeted interventions.

Using spatial analytics in epidemiological investigations and studies has offered better alternatives for researchers and public health advocates to explore real-time local disease burden within communities across areas. This approach can explore distal determinants of health as compared to proximal ones to identify local area-level risks for targeted intervention. We note that T2D is a common problem, but despite efforts to control and interventions to limit the risk of the disease, T2D seems to be increasing across the population. Although interventions have been targeted to increase awareness of the risk or promote healthy lifestyle behaviors at the individual level, public health systems continue to struggle in tackling the T2D burden and its

XSL•FO
RenderX

complications. Our work offers insights into distal determinants of health within real-time neighborhood features across the urbanization process and a paradigm shift toward crafting strategies at the area level.

Acknowledgments

The authors thank the various government agencies and ministries that supported this work. They acknowledge the use of population census data (MyCensus 2020) sourced from the Department of Statistics Malaysia under the Ministry of Economy Malaysia. They acknowledge the support from the Ministry of Higher Education Malaysia and the National University of Malaysia in funding this work (Fundamental Research Grant Scheme; FRGS/1/2022/SKK04/UKM/01/1) and for the resources provided to conduct this study. They also thank the Non-Communicable Disease Section, Public Health Division of the Ministry of Health Malaysia, and the Office of the Director General of Public Health for providing the National Diabetes Registry data and for the support of this study to be in line with Malaysia's public health priority target. The authors are grateful to all lead family medicine specialists, medical assistants, and nurses in charge of managing the National Diabetes Registry from the district health offices and primary care clinics throughout the state of Penang for their assistance with data collection. They thank the Department of Survey and Mapping under the jurisdiction of the Ministry of Natural Resources and Environmental Sustainability for providing the authors with an official digital copyright license for permitted use of shapefiles (code J72499B; issue date September 14, 2022), the Department of Town and Country Planning under the Ministry of Housing and Local Government for providing land use data for the state of Penang, the Penang GeoHub, the Penang State Health Department, and the Penang State government for access to relevant built environment facility (food outlets) data. The authors thank the director general of health for the support and permission for the publication of their study findings.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

KG and MRAM designed the study. NS, MFMR, FM, and WLC had access to and collected the data. NS, MFMR, and WLC managed the data. KG and LAW conducted the data analysis. KG, MRAM, NS, LAW, KNAM, and FM contributed to the interpretation of the data. KG drafted the first version of the manuscript. MRAM, LAW, and KNAM critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Malaysian Standard Industrial Classification codes for food, beverages, household goods, and grocery services.
[DOCX File , 17 KB - publichealth_v11i1e70045_app1.docx]

Multimedia Appendix 2
Locally estimated scatterplot smoothing regression interpolation on influence of proximity of hawker complexes, food complexes and hawker markets on the BMI of type 2 diabetes cases.
[DOCX File , 94 KB - publichealth_v11i1e70045_app2.docx]

Multimedia Appendix 3
Locally estimated scatterplot smoothing regression interpolation on influence of proximity of *nasi kandar*, *kopitiam* and fast-food outlets on the BMI of type 2 diabetes cases.
[DOCX File , 102 KB - publichealth_v11i1e70045_app3.docx]

Multimedia Appendix 4
Locally estimated scatterplot smoothing regression interpolation on influence of proximity of 24-7 convenience stores, public markets and supermarkets on the BMI of type 2 diabetes cases.
[DOCX File , 99 KB - publichealth_v11i1e70045_app4.docx]

Multimedia Appendix 5
Multiscale geographically weighted regression results and model performance.
[DOCX File , 19 KB - publichealth_v11i1e70045_app5.docx]

XSL•FO
RenderX

References

1. Nhema AG, Zinyama T. Modernization, dependency and structural adjustment development theories and Africa: a critical appraisal. Int J Soc Sci Res 2016;4(1):151-166 [FREE Full text] [doi: 10.5296/ijssr.v4i1.9040]
2. Steyn NP, McHiza ZJ. Obesity and the nutrition transition in Sub-Saharan Africa. Ann N Y Acad Sci 2014 Apr 11;1311(1):88-101. [doi: 10.1111/nyas.12433] [Medline: 24725148]
3. Popkin BM. The nutrition transition and its health implications in lower-income countries. Public Health Nutr 1998 Mar 01;1(1):5-21. [doi: 10.1079/phn19980004] [Medline: 10555527]
4. Mackay H. A feminist geographic analysis of perceptions of food and health in Ugandan cities. Gend Place Cult 2019 Apr 23;26(11):1519-1543 [FREE Full text] [doi: 10.1080/0966369x.2018.1555148]
5. Summerhayes L, Baker D, Vella K. Food diversity and accessibility enabled urban environments for sustainable food consumption: a case study of Brisbane, Australia. Humanit Soc Sci Commun 2024 Sep 16;11(1):97 [FREE Full text] [doi: 10.1057/s41599-024-03724-9]
6. Toure D, Herforth A, Pelto GH, Neufeld LM, Mbuya MN. An emergent framework of the market food environment in low- and middle-income countries. Curr Dev Nutr 2021 Apr;5(4):nzab023 [FREE Full text] [doi: 10.1093/cdn/nzab023] [Medline: 33948531]
7. AlHasan DM, Eberth JM. An ecological analysis of food outlet density and prevalence of type II diabetes in South Carolina counties. BMC Public Health 2016 Jan 05;16(1):10 [FREE Full text] [doi: 10.1186/s12889-015-2681-6] [Medline: 26728042]
8. Babey SH, Diamant AL, Hastert TA, Harvey S. Designed for disease: the link between local food environments and obesity and diabetes. UCLA Center for Health Policy Research. URL: https://escholarship.org/uc/item/7sf9t5wx [accessed 2024-10-09]
9. Mazidi M, Speakman JR. Higher densities of fast-food and full-service restaurants are not associated with obesity prevalence. Am J Clin Nutr 2017 Aug 31;106(2):603-613 [FREE Full text] [doi: 10.3945/ajcn.116.151407] [Medline: 28566310]
10. Zick CD, Curtis DS, Meeks H, Smith KR, Brown BB, Kole K, et al. The changing food environment and neighborhood prevalence of type 2 diabetes. SSM Popul Health 2023 Mar;21:101338 [FREE Full text] [doi: 10.1016/j.ssmph.2023.101338] [Medline: 36691490]
11. Kusuma D, Atanasova P, Pineda E, Anjana RM, De Silva L, Hanif AA, et al. Food environment and diabetes mellitus in South Asia: a geospatial analysis of health outcome data. PLoS Med 2022 Apr 26;19(4):e1003970 [FREE Full text] [doi: 10.1371/journal.pmed.1003970] [Medline: 35472059]
12. Feng X, Astell-Burt T, Badland H, Mavoa S, Giles-Corti B. Modest ratios of fast food outlets to supermarkets and green grocers are associated with higher body mass index: Longitudinal analysis of a sample of 15,229 Australians aged 45 years and older in the Australian National Liveability Study. Health Place 2018 Jan;49:101-110. [doi: 10.1016/j.healthplace.2017.10.004] [Medline: 29253808]
13. Ganasegeran K, Abdul Manaf MR, Safian N, Waller L, Abdul Maulud KN, Mustapha. GIS-based assessments of neighborhood food environments and chronic conditions: an overview of methodologies. Annu Rev Public Health 2024 May;45(1):109-132. [doi: 10.1146/annurev-publhealth-101322-031206] [Medline: 38061019]
14. National Health and Morbidity Survey (NHMS) 2019: non-communicable diseases, healthcare demand and health literacy: key findings. Institute for Public Health. 2019. URL: https://iku.gov.my/nhms-2019 [accessed 2024-06-03]
15. Ng S, Swinburn B, Kelly B, Vandevijvere S, Yeatman H, Ismail MN, et al. Extent of implementation of food environment policies by the Malaysian Government: gaps and priority recommendations. Public Health Nutr 2018 Dec 02;21(18):3395-3406 [FREE Full text] [doi: 10.1017/S1368980018002379] [Medline: 30277185]
16. Piel FB, Fecht D, Hodgson S, Blangiardo M, Toledano M, Hansell AL, et al. Small-area methods for investigation of environment and health. Int J Epidemiol 2020 Apr 01;49(2):686-699 [FREE Full text] [doi: 10.1093/ije/dyaa006] [Medline: 32182344]
17. Population indicators: Pulau Pinang. Department of Statistics Malaysia. URL: https://www.dosm.gov.my/v1/index.php?r=column/cone&menu_id=SEFobmo1N212cXc5TFlLVTVxWUFXZz09 [accessed 2024-05-03]
18. National diabetes registry system 2022. Ministry of Health Malaysia. URL: http://ndr.moh.gov.my/account/login?return=/dashboard/home [accessed 2022-02-05]
19. Giron S. Geocode. ez34. 2017. URL: https://ez34.net/ [accessed 2023-06-06]
20. Open data portal. Penang GeoHub. URL: https://pegis.penang.gov.my/ [accessed 2024-02-10]
21. Malaysia standard industrial classification. Department of Statistics Malaysia. 2008. URL: https://msic.stats.gov.my/bi/index.php [accessed 2024-02-15]
22. Wiki J, Kingham S, Campbell M. A geospatial analysis of type 2 diabetes mellitus and the food environment in urban New Zealand. Soc Sci Med 2021 Nov;288:113231. [doi: 10.1016/j.socscimed.2020.113231] [Medline: 32741687]
23. MyCensus 2020 population indicators. Department of Statistics Malaysia. URL: https://newss.statistics.gov.my/newss-portalx/ep/epLogin.seam [accessed 2023-02-10]
24. National diabetes registry report 2023. Ministry of Health Malaysia. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Laporan/Umum/NDR_Report_2023_Final.pdf [accessed 2025-05-29]
25. Waller LA, Gotway CA. Applied Spatial Statistics for Public Health Data. Hoboken, NJ: John Wiley & Sons; 2004.

XSL•FO
RenderX

26. Charreire H, Casey R, Salze P, Simon C, Chaix B, Banos A, et al. Measuring the food environment using geographical information systems: a methodological review. Public Health Nutr 2010 Nov 21;13(11):1773-1785. [doi: 10.1017/S1368980010000753] [Medline: 20409354]
27. Amici L. Vector operations: proximity analysis QGIS tutorials. Lorenzo Amici Revision. URL: https://qgis3-10-geoanalysis-un.readthedocs.io/en/latest/vector/proximity.html [accessed 2024-01-15]
28. van Erpecum CL, van Zon SK, Bültmann U, Smidt N. The association between fast-food outlet proximity and density and body mass index: findings from 147,027 lifelines cohort study participants. Prev Med 2022 Feb;155:106915 [FREE Full text] [doi: 10.1016/j.ypmed.2021.106915] [Medline: 34922992]
29. Ross A, Willson VL. Hierarchical multiple regression analysis using at least two sets of variables (in two blocks). In: Ross A, Willson VL, editors. Basic and Advanced Statistical Tests: Writing Results Sections and Creating Tables and Figures. Cham, Switzerland: Springer; 2017:61-74.
30. Henderson HV, Velleman PF. Building multiple regression models interactively. Biometrics 1981 Jun;37(2):391 [FREE Full text] [doi: 10.2307/2530428]
31. Harrell Jr FE. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Cham, Switzerland: Springer; 2015.
32. Fotheringham A, Brunsdon C, Charlton M. Geographically weighted summary statistics — a framework for localised exploratory data analysis. Comput Environ Urban Syst 2002 Nov;26(6):501-524 [FREE Full text] [doi: 10.1016/S0198-9715(01)00009-6]
33. Fotheringham SA, Brunsdon C, Charlton M. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. Hoboken, NJ: Wiley–Blackwell; 2003.
34. Gebreab SY. Statistical methods in spatial epidemiology. In: Duncan DT, Kawachi I, editors. Neighborhoods and Health. New York, NY: Oxford University Press; 2018.
35. Oshan TM, Li Z, Kang W, Wolf LJ, Fotheringham AS. mgwr: a Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. ISPRS Int J Geoinf 2019 Jun 08;8(6):269 [FREE Full text] [doi: 10.3390/ijgi8060269]
36. How Multiscale Geographically Weighted Regression (MGWR) works? ESRI ArGIS Pro Tool Reference. URL: https://tinyurl.com/56byykb2 [accessed 2024-01-15]
37. Mansournia MA, Altman DG. Population attributable fraction. BMJ 2018 Feb 22;360:k757. [doi: 10.1136/bmj.k757] [Medline: 29472187]
38. Northridge ME. Public health methods--attributable risk as a link between causality and public health action. Am J Public Health 1995 Sep;85(9):1202-1204. [doi: 10.2105/ajph.85.9.1202] [Medline: 7661224]
39. Rockhill B, Newman B, Weinberg C. Use and misuse of population attributable fractions. Am J Public Health 1998 Jan;88(1):15-19. [doi: 10.2105/ajph.88.1.15] [Medline: 9584027]
40. Li L, Wang D. Do neighborhood food environments matter for eating through online-to-offline food delivery services? Appl Geogr 2022 Jan;138:102620 [FREE Full text] [doi: 10.1016/j.apgeog.2021.102620]
41. Block J, Seward M, James P. Food environment and health. In: Duncan DT, Kawachi I, editors. Neighborhoods and Health. New York, NY: Oxford University Press; 2018:247-277.
42. Zhang YT, Laraia BA, Mujahid MS, Blanchard SD, Warton EM, Moffet HH, et al. Is a reduction in distance to nearest supermarket associated with BMI change among type 2 diabetes patients? Health Place 2016 Jul;40:15-20 [FREE Full text] [doi: 10.1016/j.healthplace.2016.04.008] [Medline: 27160530]
43. Sachdeva B, Puri S, Aeri BT. Availability and accessibility of healthy and unhealthy foods in neighborhood and their association with noncommunicable diseases: a scoping review. Indian J Public Health 2024 Jan 01;68(1):95-105 [FREE Full text] [doi: 10.4103/ijph.ijph_436_23] [Medline: 38847640]
44. Naidoo N, van Dam RM, Ng S, Tan CS, Chen S, Lim JY, et al. Determinants of eating at local and western fast-food venues in an urban Asian population: a mixed methods approach. Int J Behav Nutr Phys Act 2017 May 25;14(1):69 [FREE Full text] [doi: 10.1186/s12966-017-0515-x] [Medline: 28545458]
45. Laxer RE, Janssen I. The proportion of excessive fast-food consumption attributable to the neighbourhood food environment among youth living within 1 km of their school. Appl Physiol Nutr Metab 2014 Apr;39(4):480-486 [FREE Full text] [doi: 10.1139/apnm-2013-0208] [Medline: 24669990]
46. Jeffery RW, Baxter J, McGuire M, Linde J. Are fast food restaurants an environmental risk factor for obesity? Int J Behav Nutr Phys Act 2006 Jan 25;3(1):2 [FREE Full text] [doi: 10.1186/1479-5868-3-2] [Medline: 16436207]
47. Sarkar C, Webster C, Gallacher J. Are exposures to ready-to-eat food environments associated with type 2 diabetes? A cross-sectional study of 347 551 UK Biobank adult participants. Lancet Planet Health 2018 Oct;2(10):e438-e450 [FREE Full text] [doi: 10.1016/S2542-5196(18)30208-0] [Medline: 30318101]
48. Ntarladima AM, Karssenberg D, Poelman M, Grobbee DE, Lu M, Schmitz O, et al. Associations between the fast-food environment and diabetes prevalence in the Netherlands: a cross-sectional study. Lancet Planet Health 2022 Jan;6(1):e29-e39 [FREE Full text] [doi: 10.1016/S2542-5196(21)00298-9] [Medline: 34998457]

XSL•FO
RenderX

49. Garg G, Tedla YG, Ghosh AS, Mohottige D, Kolak M, Wolf M, et al. Supermarket proximity and risk of hypertension, diabetes, and CKD: a retrospective cohort study. Am J Kidney Dis 2023 Feb;81(2):168-178 [FREE Full text] [doi: 10.1053/j.ajkd.2022.07.008] [Medline: 36058428]
50. Ghosh-Dastidar M, Hunter G, Collins RL, Zenk SN, Cummins S, Beckman R, et al. Does opening a supermarket in a food desert change the food environment? Health Place 2017 Jul;46:249-256 [FREE Full text] [doi: 10.1016/j.healthplace.2017.06.002] [Medline: 28648926]
51. Lind PL, Jensen PV, Glümer C, Toft U. The association between accessibility of local convenience stores and unhealthy diet. Eur J Public Health 2016 Aug 06;26(4):634-639. [doi: 10.1093/eurpub/ckv242] [Medline: 26851817]
52. Gbenro Jr M, Student M, Brace A, Matthews T. The relationship between food deserts, farmers' markets, Nutrition Benefits, and health in Delaware census tracts; 2017. Dela J Public Health 2019 Dec;5(5):16-23 [FREE Full text] [doi: 10.32481/djph.2019.12.005] [Medline: 34467066]
53. Kauhl B, Schweikart J, Krafft T, Keste A, Moskwyn M. Do the risk factors for type 2 diabetes mellitus vary by location? A spatial analysis of health insurance claims in Northeastern Germany using kernel density estimation and geographically weighted regression. Int J Health Geogr 2016 Nov 03;15(1):38 [FREE Full text] [doi: 10.1186/s12942-016-0068-2] [Medline: 27809861]
54. Thorpe LE, Adhikari S, Lopez P, Kanchi R, McClure LA, Hirsch AG, et al. Neighborhood socioeconomic environment and risk of type 2 diabetes: associations and mediation through food environment pathways in three independent study samples. Diabetes Care 2022 Apr 01;45(4):798-810 [FREE Full text] [doi: 10.2337/dc21-1693] [Medline: 35104336]
55. Cannuscio CC, Tappe K, Hillier A, Buttenheim A, Karpyn A, Glanz K. Urban food environments and residents' shopping behaviors. Am J Prev Med 2013 Nov;45(5):606-614. [doi: 10.1016/j.amepre.2013.06.021] [Medline: 24139774]
56. Chrisinger BW, Kallan MJ, Whiteman ED, Hillier A. Where do U.S. households purchase healthy foods? An analysis of food-at-home purchases across different types of retailers in a nationally representative dataset. Prev Med 2018 Jul;112:15-22 [FREE Full text] [doi: 10.1016/j.ypmed.2018.03.015] [Medline: 29555187]
57. Uusitupa M, Khan TA, Viguiliouk E, Kahleova H, Rivellese AA, Hermansen K, et al. Prevention of type 2 diabetes by lifestyle changes: a systematic review and meta-analysis. Nutrients 2019 Nov 01;11(11):2611 [FREE Full text] [doi: 10.3390/nu11112611] [Medline: 31683759]
58. Leeder SR, Downs SM. Diabetes and the human condition. Med J Aust 2014 Aug 18;201(4):185-186. [doi: 10.5694/mja14.00943] [Medline: 25164832]
59. Congdon P. Spatial variation in attributable risks. Spat Spatiotemporal Epidemiol 2015 Jan;12:39-52. [doi: 10.1016/j.sste.2015.02.002] [Medline: 25779908]
60. de Vries H, Kremers S, Lippke S. Health education and health promotion: key concepts and exemplary evidence to support them. In: Fisher EB, Cameron LD, Christensen AJ, Ehlert U, Guo Y, Oldenburg B, et al, editors. Principles and Concepts of Behavioral Medicine: A Global Handbook. Cham, Switzerland: Springer; 2018:489-432.
61. Chen X, Kwan MP. Contextual uncertainties, human mobility, and perceived food environment: the uncertain geographic context problem in food access research. Am J Public Health 2015 Sep;105(9):1734-1737. [doi: 10.2105/AJPH.2015.302792] [Medline: 26180982]

Abbreviations

ARR: adjusted rate ratio
ENP: effective number of parameters
km: kilometers
LOESS: locally estimated scatterplot smoothing
MGWR: multiscale geographically weighted regression
NDR: National Diabetes Registry
PAF: population attributable fraction
RR: rate ratio
T2D: type 2 diabetes

Edited by A Mavragani; submitted 13.12.24; peer-reviewed by S Heydari, K Musa, SSSA Rahim; comments to author 01.05.25; revised version received 05.07.25; accepted 24.07.25; published 08.09.25.

Please cite as:
Ganasegeran K, Abdul Manaf MR, Waller LA, Safian N, Mohd Rizal MF, Chiew WL, Mustapha F, Abdul Maulud KN
Impact of Commercial Food Environments on Local Type 2 Diabetes Burden: Cross-Sectional and Ecological Multimodeling Study
JMIR Public Health Surveill 2025;11:e70045
URL: https://publichealth.jmir.org/2025/1/e70045
doi:10.2196/70045
PMID:40921070

XSL•FO
RenderX

©Kurubaran Ganasegeran, Mohd Rizal Abdul Manaf, Lance A Waller, Nazarudin Safian, Muhammad Faid Mohd Rizal, Wye Lee Chiew, Feisul Mustapha, Khairul Nizam Abdul Maulud. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 08.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Behavioral Profiles and Sociodemographic Predictors of Planetary Health Diet Engagement Among Health Care Professionals to Inform Public Health Promotion: Cross-Sectional Study

Florence Carrouel[1], Prof Dr; Rita Nugem[1], PhD; Laurie Fraticelli[1], PhD; Yohan Fayet[2], PhD; Lama Basbous[3], MScD; Corélie Salque[1], MScD; Claude Dussart[1,3], Prof Dr; Marie-Thérèse Charreyre[4], PhD; Romain Lan[1,5], PhD; Audrey Murat-ringot[1,3], PhD

[1]Health Systemic Process Laboratory (P2S), UR4129, University Claude Bernard Lyon 1, University of Lyon, Lyon, France

[2]Department of Geography, Université Clermont Auvergne, AgroParisTech, INRAE, Clermont-Ferrand, France

[3]Hospices Civils de Lyon, Lyon, France

[4]Universite Claude Bernard Lyon 1, INSA Lyon, Université Jean Monnet, Centre National de la Recherche Scientifique (CNRS) UMR 5223, Ingénierie des Matériaux Polymères, Villeurbanne, France

[5]Laboratory Anthropologie Bio-Culturelle, Droit, Éthique et Santé (ADES), Aix Marseille University, Centre National de la Recherche Scientifique (CNRS), Établissement Français du Sang (EFS), Marseille, France

Corresponding Author:
Florence Carrouel, Prof Dr
Health Systemic Process Laboratory (P2S), UR4129
University Claude Bernard Lyon 1
University of Lyon
7 Rue Guillaume Paradin
Lyon, 69008
France
Phone: 33 0478785745
Email: florence.carrouel@univ-lyon1.fr

Abstract

Background: The planetary health diet (PHD) promotes dietary habits that are beneficial to human health and environmental sustainability, two closely related goals of modern public health. Health care professionals are expected to lead by example and use their position to encourage their patients and communities to adopt healthier and more sustainable behaviors. However, little is known about health care professionals' knowledge, attitudes, and behaviors (KAB) related to the PHD, which limits the ability to design targeted interventions for this key population.

Objective: This study aimed to analyze KAB profiles related to the PHD among French health care professionals and assess associations with individual characteristics.

Methods: This cross-sectional study was conducted from April 2024 to June 2024 among health care professionals. Participants were recruited using a nonprobabilistic convenience sampling method. Inclusion criteria were to be aged ≥18 years, fluent in French, and working as a health care professional at Hospices Civils de Lyon. Data were collected through an online questionnaire including sociodemographic and geographic variables, KAB items related to the PHD, and dietary intake. Adherence to the PHD was calculated using a validated scoring system. KAB items were analyzed using k-means clustering to identify distinct profiles. Associations with sociodemographic and geographic variables were explored using chi-square tests and ANOVA.

Results: Among 1104 respondents (n=927, 83.97% women and n=882, 79.89% aged between 30 and 59 years), 3 KAB clusters were identified: cluster 1 ("knowledge-attitude gap"; n=481, 43.57%) included professionals with good knowledge but moderate attitudes and partially consistent behaviors; cluster 2 ("high intent, low action"; n=318, 28.8 included those with strong knowledge and positive attitudes but limited behavioral change, and cluster 3 ("behavior-driven alignment"; n=305, 27.63%) included participants who reported pro-PHD behaviors despite lower knowledge and less favorable attitudes. PHD adherence was significantly associated with being female, having a vegetarian or flexitarian diet, and reporting environmental concerns (P<.01 in all cases).

XSL•FO
RenderX

Conclusions: Distinct KAB profiles among health care professionals suggest differing levels of engagement regarding the PHD. Findings suggest that tailored interventions addressing knowledge, attitudes, or behaviors could improve health care professionals' alignment with planetary health goals within similar institutional contexts.

(JMIR Public Health Surveill 2025;11:e67633) doi:10.2196/67633

KEYWORDS

planetary health diet; sustainable nutrition; food systems; dietary transition; One Health

Introduction

The planetary health diet (PHD), introduced by the EAT-Lancet Commission, represents a transformative dietary paradigm that seeks to simultaneously promote human health and environmental sustainability [1]. Since the publication of the EAT-Lancet report, growing empirical evidence has confirmed the dual benefit of planetary health–aligned diets for both human and environmental outcomes. Recent global analyses show that higher adherence to plant-forward dietary patterns reduces the incidence of noncommunicable diseases (NCDs) and mortality while simultaneously decreasing greenhouse gas emissions and land use [2-4]. In addition, moderate shifts toward sustainable diets could prevent millions of premature deaths annually and contribute significantly to climate mitigation targets [5,6]. Beyond individuals, health care systems themselves are increasingly recognized as strategic environments for implementing and modeling sustainable food choices [7,8]. Health care professionals are also increasingly viewed as key actors in the transition toward climate-resilient and sustainable health systems. Recent reports emphasize their dual role as credible messengers and institutional change agents in promoting planetary health principles within hospitals and clinical settings [9-11].

Poor dietary patterns are among the leading modifiable risk factors of NCDs, which account for approximately 74% of global deaths [12]. Diets high in ultraprocessed foods, red and processed meats, added sugars, and saturated fats have been strongly linked to cardiovascular disease, type 2 diabetes, obesity, certain cancers, and chronic respiratory diseases [13,14]. Moreover, many countries—particularly in rapidly urbanizing regions—face a double burden of malnutrition, where undernutrition and overweight or obesity coexist within the same communities or even individuals [15]. Nutrition is recognized as a key pathway through which social determinants of health exert their effects. Socioeconomic factors, education, and environmental contexts shape dietary behaviors and contribute to health disparities [16]. Consequently, transforming food systems is vital not only for curbing NCDs but also for securing long-term planetary health.

Responding to these interlinked crises, the field of planetary health has emerged to integrate human health with the sustainability of the Earth's ecosystems [17]. The PHD provides a concrete application of this vision, aligning nutritional science with environmental goals. However, despite its potential, empirical studies suggest a persistent "knowledge-attitude gap": although awareness of healthy and sustainable diets is growing, actual behavioral adoption remains limited [18]. This discrepancy highlights the complex interplay of psychological, social, and structural factors that shape dietary behavior.

To investigate this complexity, the knowledge, attitudes, and behaviors (KAB) model offers a useful lens. Rooted in health psychology, it posits that knowledge influences attitudes, which subsequently guide behaviors [19,20]. However, real-world deviations from this linear sequence suggest the need to explore individual variability in how people relate to new dietary models such as the PHD.

Cluster analysis has emerged as a powerful tool to address such variability by identifying subgroups within populations based on shared patterns of KAB [21]. Identifying such patterns is especially important among health care professionals, who are expected to champion evidence-based health behaviors and influence broader public adoption.

This study aimed to identify distinct profiles of health care professionals based on their KAB regarding the PHD and investigate which sociodemographic factors are associated with these behavioral profiles. By answering this research question, this study sought to contribute to the development of more tailored and effective strategies for promoting sustainable dietary transitions in line with the goals of human and planetary health.

Methods

Study Design

This study was a descriptive, cross-sectional, and quantitative approach to examine KAB related to the PHD adhering to the Checklist for Reporting of Survey Studies (Multimedia Appendix 1). It constitutes the behavioral component of the BIOQUALIM project [22], a transdisciplinary One Health research program assessing the role of the PHD in the primary and tertiary prevention of NCDs through 4 complementary studies (clinical, chemical, behavioral, and psychosocial).

Study Setting

This study was conducted in France at the Hospices Civils de Lyon (HCL) between April 15, 2024, and June 15, 2024. HCL is the second-largest public university hospital network in France, encompassing 13 hospitals and approximately 24,000 employees across the Lyon metropolitan area [23]. It covers a wide range of clinical specialties, research units, and training departments. The workforce includes physicians, nurses, midwives, pharmacists, technicians, allied health staff, and administrative personnel, constituting a highly diverse and representative health care environment.

XSL•FO
RenderX

Participant Eligibility

The participants included (1) were aged ≥18 years, (2) spoke fluent French, and (3) were health care professionals at HCL.

Recruitment and Data Collection Procedures

Sampling Strategy

A nonprobabilistic convenience sampling method was used. A systematic dissemination plan was implemented across all HCL hospital sites to ensure broad and consistent recruitment coverage. The survey was distributed in both digital and paper formats. Digital invitations were circulated via institutional mailing lists targeting all professional categories within the HCL system. A pop-up link was displayed on the hospital intranet and on shared computers in staff rooms and administrative offices. Flyers with QR codes were also distributed on meal trays and posted in common areas. This multichannel strategy ensured that both clinical and nonclinical staff could access the survey regardless of their work schedule or digital access constraints.

Questionnaire Development and Validation

The survey was created to assess KAB regarding the PHD and was based on a literature review and the World Health Organization's (WHO) methodology for designing KAB surveys [24]. Initial items were inspired by established tools, notably the Food Choice Questionnaire [25]. After generating a preliminary version of 136 closed questions, expert review was undertaken to validate both face and content aspects. An expert panel—including nutritionists, dietitians, and general practitioners—reviewed and refined the items. Cognitive testing was conducted through qualitative cognitive interviews with 20 volunteer respondents using think-aloud and probing techniques to evaluate comprehension and clarity of the items in line with best practices for questionnaire pretesting [26]. The final validated version comprised 88 scored items across knowledge (n=36), attitudes (n=21), and behavior (n=31) domains, plus 1 open-ended question [22].

Scoring System

To quantify responses, a scoring system was developed through a Delphi consensus process involving a panel of experts. Knowledge items were assessed dichotomously (+1 for correct and −1 for incorrect answers) as they referred to factual information with a single correct answer.

In contrast, attitude and behavior items were initially designed using Likert-type scales (eg, frequency or level of agreement). To ensure comparability and interpretive coherence across the 3 KAB dimensions, these Likert-type responses were subsequently transformed into normalized values (+1, 0, and −1) by the expert panel. This transformation reflected the degree of alignment between each response and the principles of the PHD as determined through consensus among Delphi participants [24,27,28].

This standardized conversion was chosen to maintain the semantic richness of the original Likert items while allowing for a unified scoring framework across domains. It also minimized subjective interpretation and facilitated statistical aggregation of KAB scores within cluster analysis. Such normalization is consistent with WHO knowledge, attitude, and practice methodologies, which favor simplified, direction-oriented coding to describe behavioral tendencies at the population level [24,29].

Scores for each KAB dimension were normalized on a scale from 0 to 40, with total KAB scores ranging from 0 to 120. Categories were set as follows: inadequate (0-70), marginal (71-88), and adequate (89-120).

Survey Platform

Participants completed the questionnaire using LimeSurvey (version 6.5.7; LimeSurvey GmbH). Measures were taken to ensure secure, anonymous participation: CAPTCHA verification, cookies to block duplicate submissions, and unique tokens for log-in. Respondents could pause and resume the survey as needed.

Study Outcomes

The primary outcome was to identify distinct KAB profiles regarding the PHD among health care professionals. The secondary outcome was to investigate how sociodemographic characteristics relate to these profiles in this specific population.

Statistical Analysis

Sample Size Calculation

On the basis of an estimated population of 8000 HCL staff members regularly using the hospital catering service (collective kitchen facilities), a minimum of 367 responses was calculated to ensure a 95% confidence level and a 5% margin of error for proportion estimates (finite population correction applied). This target was established to guarantee statistical precision for proportion estimates in the reference population.

Geospatial and Demographic Variables

Municipality-level variables were incorporated to explore contextual disparities, including population density; average annual income; and the Geographical Classification for Health Studies [30], which evaluates socioeconomic and environmental conditions.

Data Analysis

Descriptive statistics were used to analyze the sociodemographic and geographic characteristics of the participants and KAB scores.

To identify distinct profiles based on KAB dimensions, a hierarchical cluster analysis was first conducted using the Ward method and squared Euclidean distance. The optimal number of clusters was determined through visual inspection of the dendrogram. This structure was then used to guide the final k-means clustering procedure. A principal-component analysis was applied to visualize the separation of clusters.

To assess the association between individual characteristics and cluster membership, a multinomial logistic regression model was conducted. Independent variables included gender, age group, profession, dietary pattern, household size, population density, income quintile, Geographical Classification for Health Studies, and weekly food budget. Odds ratios (ORs) with 95% CIs and *P* values were reported.

XSL•FO
RenderX

In addition, a mediation analysis was conducted to examine whether attitudes mediated the relationship between knowledge and behavior. The analysis was based on the ordinary least squares regression method, and indirect effects were estimated via bootstrapping with 5000 samples. Unstandardized coefficients (*b*), SEs, 2-tailed *t* values, *P* values, and 95% CIs were reported for all effects. Statistical significance was set at $P<.05$.

Ethical Considerations

This study was approved by the ethics committee of HCL (approval IRB00014232 2024 06 13_12, issued on June 19, 2024) and complied with the Declaration of Helsinki and the French data protection authority MR-004 norms. Written informed consent was obtained from all participants.

Results

Characteristics of the Participants

Figure 1 shows the flowchart of the study. Of the 8000 persons assessed for eligibility, 1229 (15.36%) were included, and 1104 (13.8%) fully completed the questionnaire. This sample size, well above the calculated minimum, ensured high statistical precision (–2.7% to +2.7% margin of error), whereas the voluntary recruitment process may limit full population representativeness.

Figure 1. Flowchart of the study. Recruitment and inclusion process for a cross-sectional knowledge, attitudes, and behaviors survey on adherence to the planetary health diet among Hospices Civils de Lyon health care professionals (France; April 2024-June 2024).

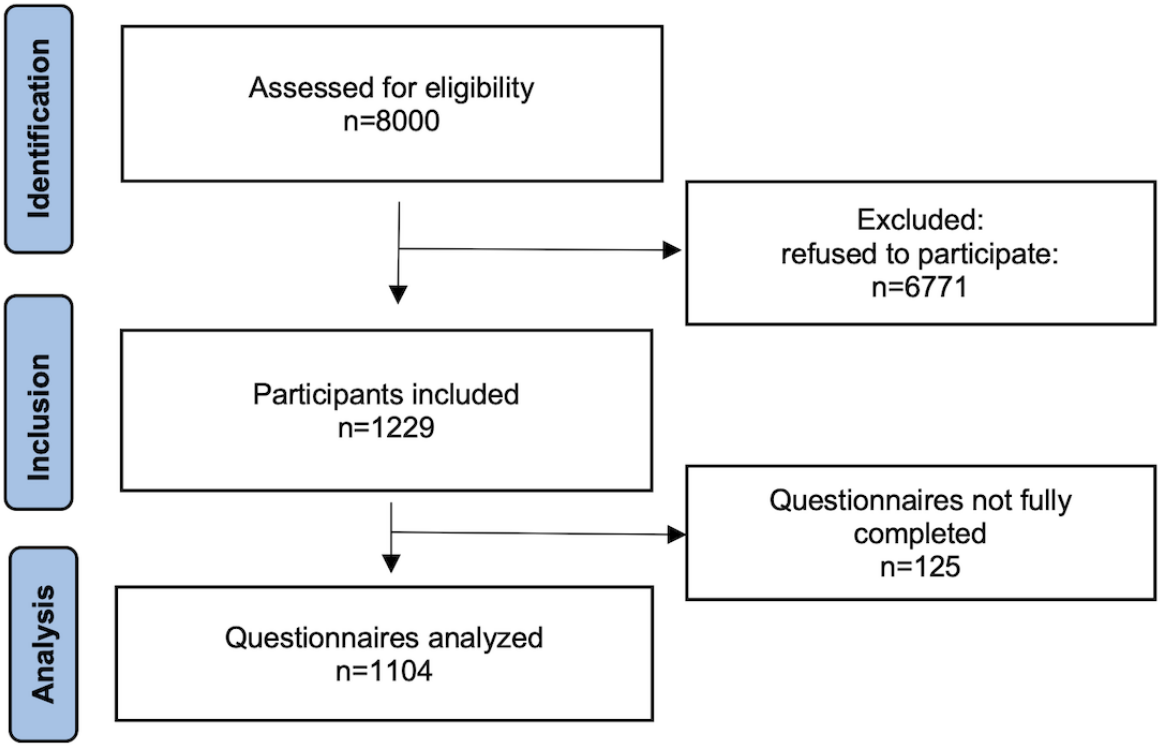


Table 1 shows the sociodemographic and geographic characteristics of the participants. Most participants were women (927/1104, 83.97%) aged between 30 and 59 years (882/1104, 79.89%), lived in small households (990/1104, 89.67%), and were in intermediate professions (720/1104, 65.22%). Omnivorous (702/1104, 63.59%) and flexitarian (331/1104, 29.98%) diets were predominant. Half of the participants (598/1104, 54.17%) lived in urbanized areas (quintile 4-5), and 69.02% (762/1104) resided in wealthy metropolitan areas. A high proportion (876/1104, 79.3%) belonged to municipalities in the top 2 income quintiles. Weekly food expenses varied, with 46.01% (508/1104) spending less than €100 (US $116.56) and 49.55% (547/1104) spending more.

XSL•FO
RenderX

Table 1. Sociodemographic and geographic characteristics of the study population and cluster membership (N=1104)[a].

	All, n (%)	Cluster 1: knowledge-attitude gap (n=481), n (%)	Cluster 2: high intent, low action (n=318), n (%)	Cluster 3: behavior-driven alignment (n=305), n (%)
Gender				
Women	927 (83.97)	389 (80.9)	270 (84.9)	268 (87.9)
Men	173 (15.67)	90 (18.7)	46 (14.5)	37 (12.1)
Other	4 (0.36)	2 (0.4)	2 (0.6)	0 (0)
Age group (y)				
18-29	146 (13.22)	68 (14.1)	30 (9.4)	48 (15.7)
30-39	263 (23.82)	112 (23.3)	86 (27)	65 (21.3)
40-49	295 (26.72)	138 (28.7)	74 (23.3)	83 (27.2)
50-59	324 (29.35)	143 (29.7)	95 (29.9)	86 (28.2)
≥60	76 (6.89)	138 (28.7)	134 (42.1)	65 (21.3)
Household size				
1-4 people	990 (89.67)	436 (90.6)	281 (88.4)	273 (89.5)
5-9 people	114 (10.33)	45 (9.4)	37 (11.6)	32 (10.5)
Profession				
Senior executive or intellectual employee	337 (30.53)	138 (28.7)	134 (42.1)	65 (21.3)
Intermediate profession	720 (65.22)	332 (69.0)	177 (55.7)	211 (69.2)
Other categories	47 (4.25)	11 (2.3)	7 (2.2)	29 (9.5)
Student	10 (0.91)	2 (0.4)	0 (0.0)	8 (2.6)
Laborer or worker	7 (0.63)	2 (0.4)	1 (0.3)	4 (1.3)
Other	30 (2.71)	7 (1.5)	6 (1.9)	17 (5.6)
Main diet				
Flexitarian	331 (29.98)	105 (21.8)	159 (50.0)	67 (22)
Omnivorous	702 (63.59)	362 (75.3)	118 (37.1)	222 (72.8)
Other diets	71 (6.43)	14 (2.9)	41 (12.9)	16 (5.2)
Pescatarian	13 (1.18)	2 (0.4)	8 (2.5)	3 (1.0)
Vegan	6 (0.54)	0 (0)	5 (1.6)	1 (0.3)
Vegetarian	30 (2.72)	4 (0.8)	22 (6.9)	4 (1.3)
Other	22 (1.99)	8 (1.7)	6 (1.9)	8 (2.6)
Population density categorized into quintiles (people per km^2)				
Quintile 1	59 (5.34)	22 (4.6)	14 (4.4)	23 (7.5)
Quintile 2	171 (15.49)	68 (14.1)	46 (14.5)	57 (18.7)
Quintile 3	276 (25.00)	113 (23.5)	88 (27.7)	75 (24.6)
Quintile 4	205 (18.57)	90 (18.7)	64 (20.1)	51 (16.7)
Quintile 5	393 (35.60)	188 (39.1)	106 (33.3)	99 (32.5)
Income per consumption unit in quintiles per year				
Quintile 1 (€15,015-€18,072 [US $17,500.70-$21,063.80])	34 (3.08)	15 (3.1)	8 (2.5)	11 (3.6)
Quintile 2 (€18,110-€19,320 [US $21,108.10-$22,518.40])	99 (8.97)	31 (6.4)	37 (11.6)	31 (10.2)
Quintile 3 (€19,480-€20,732 [US $22,704.90-$24,164.20])	95 (8.60)	27 (5.6)	38 (11.9)	30 (9.8)

XSL•FO
RenderX

	All, n (%)	Cluster 1: knowledge-attitude gap (n=481), n (%)	Cluster 2: high intent, low action (n=318), n (%)	Cluster 3: behavior-driven alignment (n=305), n (%)
Quintile 4 (€20,774-€22,974 [US $24,213.10-$26,777.30])	430 (38.95)	208 (43.2)	114 (35.8)	108 (35.4)
Quintile 5 (€23,021-€36,562 [US $26,832.10-$42,614.80])	446 (40.40)	200 (41.6)	121 (38.1)	125 (41.0)
Geographical Classification for Health Studies				
Wealthy metropolitan areas	762 (69.02)	367 (76.3)	206 (64.8)	189 (62.0)
Precarious population districts	218 (19.75)	72 (15.0)	83 (26.1)	63 (20.7)
Periurban and rural areas	124 (11.23)	42 (8.7)	29 (9.1)	53 (17.4)
Weekly food budget				
€1-€50 (US $1.17-$58.28)	134 (12.14)	52 (10.8)	31 (9.7)	51 (16.7)
€51-€100 (US $59.44-$116.56)	374 (33.88)	172 (35.8)	109 (34.3)	93 (30.5)
€101-€150 (US $117.72-$174.83)	295 (26.72)	118 (24.5)	91 (28.6)	86 (28.2)
>€150 (US $174.83)	252 (22.83)	116 (24.1)	74 (23.3)	62 (20.3)
Did not know	49 (4.43)	23 (4.8)	13 (4.1)	13 (4.3)

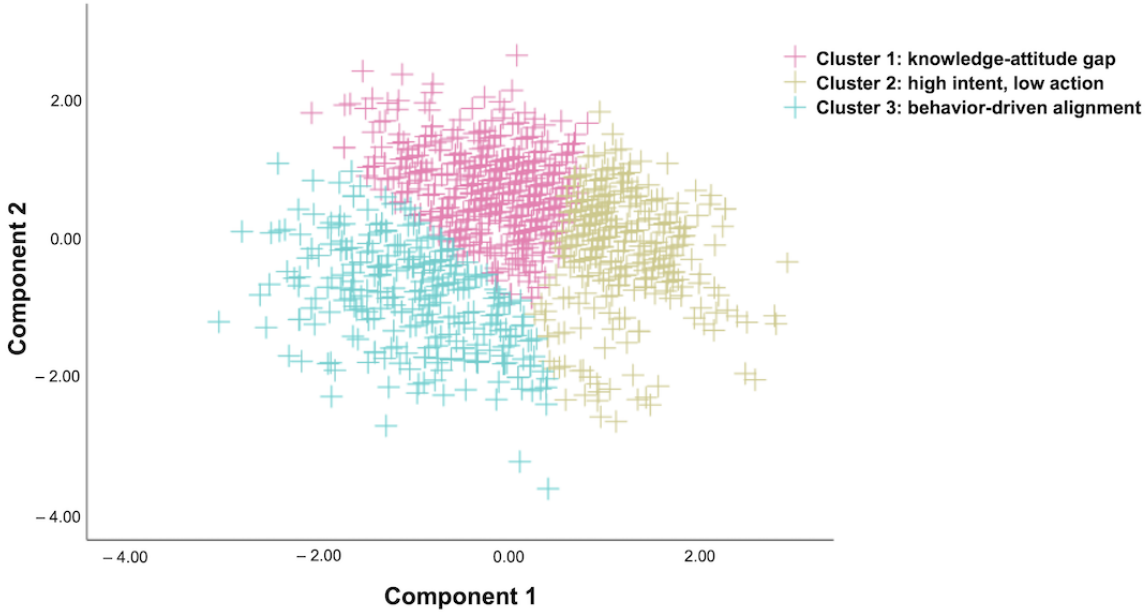
[a]Data from a cross-sectional survey on planetary health diet engagement conducted among health care professionals working at Hospices Civils de Lyon (France) between April 2024 and June 2024. Cluster allocation based on knowledge, attitudes, and behaviors profiles were derived from k-means analysis.

Clustering of Health Care Professionals Based on KAB

To better understand variations in KAB toward the PHD, a cluster analysis was conducted using standardized KAB scores. The internal consistency of the solution was supported by a total within-cluster sum of squares of 103,797. This clustering allocation (Table 1) and analysis (Figure 2 and Table 2) revealed 3 distinct profiles in health care professionals' engagement with the PHD.

Cluster 1, named "knowledge-attitude gap" (481/1104, 43.57%), exhibited a relatively high knowledge score (32.8) but moderate scores in both attitude (13.8) and behavior (14.6), resulting in a total KAB score of 61.2. Cluster 2, named "high intent, low action" (318/1104, 28.8%), had the highest knowledge (33.1) and attitude (23.3) scores but the lowest behavior score (10.2), yielding a total KAB score of 76.6. Cluster 3, named "behavior-driven alignment" (305/1104, 27.63%), had moderate knowledge (21.4) and attitude (11.7) scores and the highest behavior score (15.7), resulting in a total KAB score of 48.8.

Figure 2. Principal-component analysis (PCA) of standardized knowledge, attitudes, and behaviors (KAB) scores related to the planetary health diet among health care professionals (N=1104). Each point represents an individual, colored by cluster membership from k-means analysis. The PCA revealed distinct KAB profiles.



XSL•FO
RenderX

Table 2. Descriptive analysis of knowledge, attitudes, and behaviors (KAB) scores by cluster[a].

	Knowledge section score (0-40), mean (SD)	Attitude section score (0-40), mean (SD)	Behavior section score (0-40), mean (SD)	KAB score (0-120), mean (SD)
Cluster 1: knowledge-attitude gap (n=481)	32.8 (3.1)	13.8 (5.6)	14.6 (3.1)	61.2 (6.7)
Cluster 2: high intent, low action (n=318)	33.1 (4.4)	23.3 (4.8)	10.2 (3.4)	76.6 (6.0)
Cluster 3: behavior-driven alignment (n=305)	21.4 (3.6)	11.7 (6.3)	15.7 (3.7)	48.8 (8.2)

[a]Results from a cross-sectional survey on planetary health diet engagement conducted among health care professionals at Hospices Civils de Lyon, France (April 2024-June 2024). Clusters were identified through k-means clustering on standardized KAB scores to profile engagement patterns with the PHD.

Sociodemographic, Geographic, and Clinical Characteristics of the Clusters

Cluster 1 ("knowledge-attitude gap"; 481/1104, 43.57% of the participants) was predominantly composed of middle-aged women in intermediate professions with a largely omnivorous diet and living in wealthy metropolitan areas (Table 1). Cluster 2 ("high intent, low action"; 318/1104, 28.8% of the participants) included more older participants and had the highest proportion of flexitarians. This group showed more diversity in occupational status and place of residence, with more frequent location in low- and middle-income municipalities. Cluster 3 ("behavior-driven alignment"; 305/1104, 27.63% of the participants) was the youngest group, also predominantly female, with mixed diets and a more frequent location in rural areas. Weekly food expenses were similar across clusters, mostly ranging from €51 to €150 (US $59.44-$174.83).

Identification of Sociodemographic, Geographic, and Clinical Determinants of Low Engagement and Cluster Alignment Toward the PHD

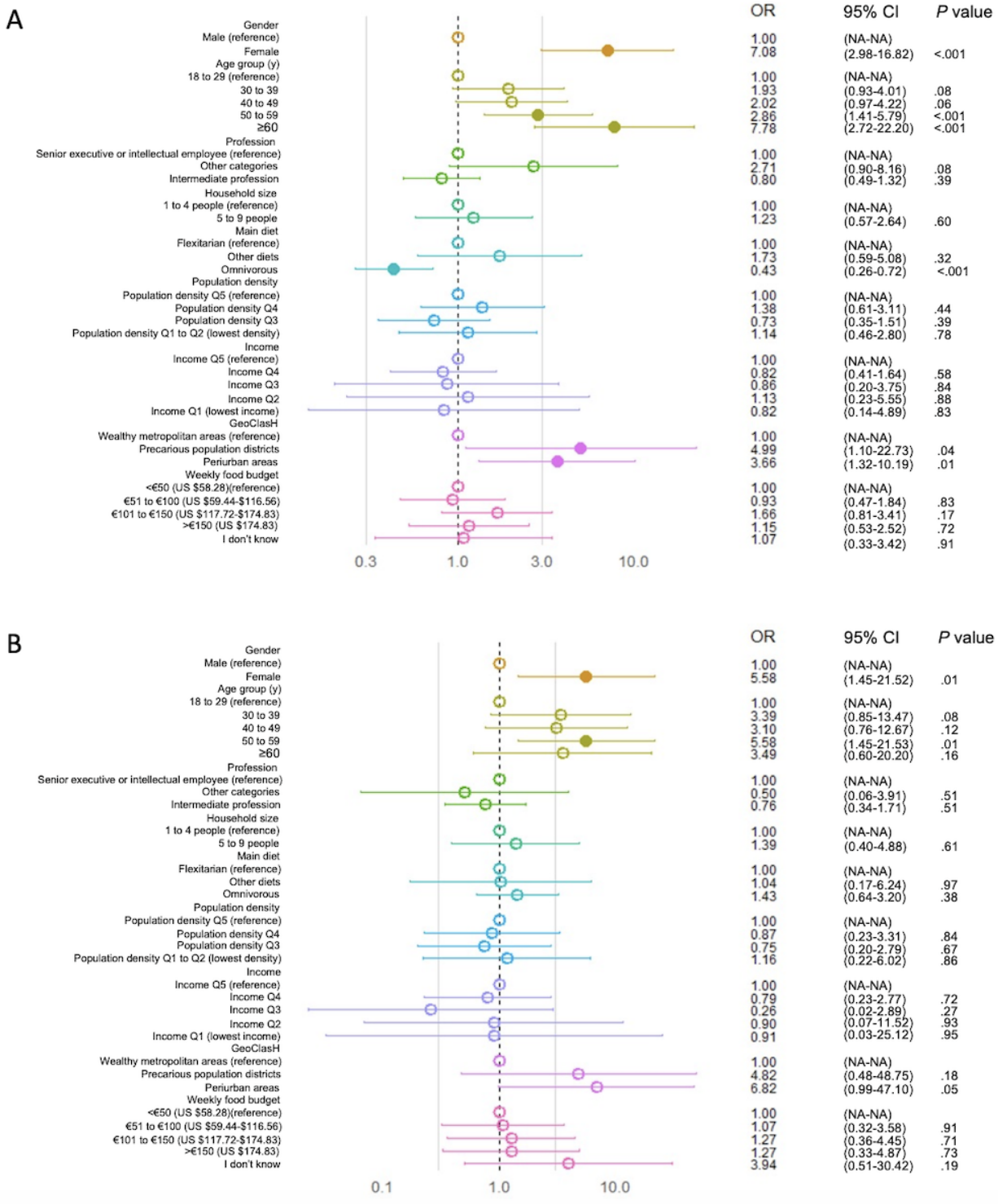
A multinomial logistic regression evaluated associations between individual characteristics and cluster membership considering the "behavior-driven alignment" cluster (cluster 3) as the reference category (Figure 3).

The comparison of cluster 1, "knowledge-attitude gap," with cluster 3, "behavior-driven alignment" (Figure 3A), demonstrated that cluster 1 was significantly associated with being female (OR 7.08, 95% CI 2.98-16.8; *P*<.001) and of older age, especially ≥60 years (OR 7.78, 95% CI 2.72-22.20; *P*<.001) and 50 to 59 years (OR 2.86, 95% CI 1.41-5.79; *P*=.004). Omnivorous individuals were less likely than flexitarians to belong to cluster 1 (OR 0.43, 95% CI 0.26-0.72; *P*=.001). Living in periurban areas (OR 3.66, 95% CI 1.32-10.19; *P*=.01) or precarious population districts (OR 4.99, 95% CI 1.10-22.73; *P*=.04) also increased the probability of alignment with this cluster. Other variables showed no significant associations.

The comparison of cluster 2, "high intent, low action," with cluster 3, "behavior-driven alignment" (Figure 3B), demonstrated that the cluster 2 profile was more likely associated with women (OR 5.58, 95% CI 1.45-21.52; *P*=.01) and individuals aged 50 to 59 years (OR 5.58, 95% CI 1.45-21.53; *P*=.01). A marginal trend of alignment with this cluster was noted for those in periurban areas (OR 6.82, 95% CI 0.99-47.10; *P*=.05). No other factors were significantly associated.

XSL•FO
RenderX

Figure 3. Sociodemographic and geographic determinants of low engagement and alignment clusters with the planetary health diet. (A) Comparison of cluster 1, "knowledge-attitude gap," with cluster 3, "behavior-driven alignment." (B) Comparison of cluster 2, "high intent, low action," with cluster 3, "behavior-driven alignment." An empty circle indicates a *P* value greater than .05, whereas a filled circle indicates a *P* value of .05 or less. GeoClasH: Geographical Classification for Health Studies; NA: not applicable; OR: odds ratio; Q: quintile.



Exploration of Mediating KAB Relationships

The results of the mediation analysis (Table 3) revealed that knowledge significantly predicted attitudes (*b*=0.2893 (0.2230-0.3556); SE 0.0338; t_{1102}=8.56; *P*<.001). In turn, attitudes significantly predicted behaviors (*b*=0.1798, 95% CI 0.1472-0.2124; SE 0.0166; t_{1102}=10.83; *P*<.001). However, the direct effect of knowledge on behaviors controlling for attitudes was not statistically significant (*b*=0.0072, 95% CI −0.0306 to 0.0449; SE 0.0192; t_{1102}=0.37; *P*=.71).

The indirect effect of knowledge on behaviors via attitudes was significant (*b*=0.0520), with a 95% CI (0.0379-0.0682) that did not include 0. This indicates a full mediation effect: knowledge influences behaviors only indirectly through its effect on attitudes.

XSL•FO
RenderX

Table 3. Mediation analysis examining the indirect effect of attitudes (A) in the relationship between knowledge (K) and behavior (B) (K→A→B)[a].

Path	Coefficient (*b*; SE; 95% CI)	*t* test	*P* value
Effect of knowledge on attitude (path a)	0.2893 (0.0338; 0.2230 to 0.3556)	8.56	<.001
Effect of attitude on behavior (path b)	0.1798 (0.0166; 0.1472 to 0.2124)	10.83	<.001
Direct effect of knowledge on behavior controlling for attitude	0.0072 (0.0192; −0.0306 to 0.0449)	0.37	.71
a × b (indirect effect)[b]	0.0520 (0.0077; 0.0379 to 0.0682)	—[c]	—

[a]Data from a cross-sectional knowledge, attitudes, and behaviors survey on planetary health diet adherence among health care professionals working at Hospices Civils de Lyon (France; April 2024-June 2024).

[b]Estimated via bootstrapping (5000 samples).

[c]Empty cells indicate that t or *P* values were not computed for the indirect effect.

Discussion

Principal Findings and Comparison With Prior Work

In a French context and among HCL health care professionals, this study addressed global health by identifying key behavioral typologies regarding the PHD. The PHD represents a transformative framework for addressing interlinked crises of human health and environmental sustainability. However, our findings highlight persistent gaps in its adoption even among health-competent populations. Cluster analysis of KAB identified three distinct profiles: (1) "knowledge-attitude gap"; (2) "high intent, low action"; and (3) "behavior-driven alignment." These profiles underline both opportunities and barriers shaping dietary transitions within health care systems.

The largest cluster, "knowledge-attitude gap" (481/1104, 43.57% of the participants) demonstrated high knowledge about the PHD but failed to adopt corresponding dietary behaviors. This knowledge-behavior disconnect has been noted across public health domains, particularly among health care workers [31,32], who have dual responsibilities: modeling health-promoting behaviors and guiding patients toward sustainable lifestyles. Similar gaps have been identified in tobacco cessation [33] and preventive oncology [34], suggesting that professional knowledge does not ensure personal action. Our analysis is consistent with the KAB framework, indicating that, within this sample, knowledge appeared to influence behaviors indirectly through attitudes. Thus, affective and motivational factors may be as critical as informational ones. Environmental factors, institutional norms, and perceived behavioral control likely moderate this pathway, highlighting the importance of structural interventions alongside education [19,35]. Our mediation analysis suggests that targeted interventions should emphasize strengthening attitudes and motivations rather than simply providing information. Approaches such as motivational interviewing, normative messaging, and role-modeling within health care teams might significantly reduce the knowledge-attitude gap [36,37].

The "high intent, low action" cluster (318/1104, 28.8% of the participants) demonstrated high knowledge and positive attitudes but low PHD-aligned behaviors. This profile illustrates a different form of behavioral inertia: intention without implementation. Contributing factors may include workplace culture, lack of time, workplace food options, or insufficient social support. Previous studies show that, without standardized and accessible sustainable consumption options, even health-conscious individuals struggle to act on their intentions [38]. This profile was more prevalent among older and flexitarian participants, suggesting generational or cultural barriers to adopt the PHD. These individuals may benefit from interventions promoting action planning and institutional support, such as providing plant-based meals in hospitals or developing peer-led sustainability programs.

Conversely, the "behavior-driven alignment" cluster (305/1104, 27.63% of the participants), characterized by the highest behavioral alignment, had only moderate knowledge and attitude scores. This challenges the assumption that knowledge is a prerequisite for sustainable dietary practice. It underscores the role of social and contextual determinants on food choices. Behavioral science research suggests that daily routines, cultural norms, and environmental prompts can foster health-promoting behaviors even without strong cognitive drivers [39,40]. This cluster, predominantly young and periurban, may reflect evolving dietary behaviors driven by cultural trends, peer influence, or institutional exposure rather than formal health education [41,42]. Exploiting these archetypes could inform peer-led interventions within health systems.

Collectively, these profiles reveal that no group simultaneously achieved high knowledge and high behavioral adherence to the PHD. This overall pattern underscores the persistence of a structural knowledge-behavior gap across all segments. The absence of a cluster combining both high knowledge and high adoption of PHD practices reflects the well-established knowledge-behavior gap in health psychology and sustainability research. Numerous studies have shown that knowledge or intention alone seldom predicts actual behavior as the translation into action depends on psychological, social, and contextual moderators [43,44]. Self-efficacy and perceived behavioral control play pivotal roles in bridging this gap, influencing whether individuals feel capable of acting upon what they know [45]. Despite the proliferation of evidence-based guidelines, barriers to adherence persist (eg, guideline complexity, environmental context, and resource constraints)—as documented in recent implementation research [46,47]. Frameworks highlight multistage processes in which motivation, environmental triggers, and sustained self-regulation determine whether intentions lead to consistent behavioral adoption [48]. Similar intention-action discrepancies have been observed in sustainable food consumption, where positive attitudes coexist with low behavioral uptake due to contextual and social

XSL•FO
RenderX

constraints [49]. Altogether, these findings suggest that promoting sustainable dietary practices among health care staff requires multilevel interventions that combine education with environmental and organizational enablers.

Beyond knowledge and attitudes, professional identity and perceived institutional norms significantly influence behavior. When sustainability is not embedded in the clinical culture, a disconnect can arise between personal values and visible action. Positioning the PHD as a personal choice and a professional standard could reinforce adoption and enable culture change [50]. Interventions integrating social role-modeling, normative messaging, and organizational alignment with PHD principles may be more effective than education alone.

The coexistence of these 3 profiles within a single health system reveals that health care professionals, despite having received scientific training and facing NCDs, do not uniformly model sustainable dietary behaviors. Lifestyle behaviors among health care professionals reflect inconsistencies found in the general population [51]. Given their direct involvement in managing NCDs and their role as trusted communicators, health care professionals' dietary practices influence public health far beyond personal well-being [52]. By adopting and advocating for sustainable diets aligned with frameworks such as the PHD, which integrates human and environmental health goals, they can enhance clinical credibility and support institutional shifts toward healthier food policies [53,54].

Both the WHO [18] and the EAT-Lancet Commission [1] emphasize the PHD potential to reduce premature mortality, alleviate environmental degradation, and contribute to the United Nations Sustainable Development Goals [55]. Therefore, aligning health care professional behavior with these recommendations is a systemic opportunity. A PHD-adherent workforce is not only healthier but also more credible in promoting planetary health.

Health care institutions must recognize that individual knowledge is not sufficient to drive change. Behavioral interventions need embedding within broader institutional commitments to sustainability, including food supply policies, cafeteria reforms, cooking staff education, and food environment restructuring. Differentiating between structural barriers (eg, food access) and psychological barriers (eg, habit and motivation) should enable targeted actions on both fronts. Systems combining policy shifts with environmental engagements such as menu redesign, sustainable meal defaults, or workplace programs are more likely to succeed [56]. By integrating PHD principles into daily operations and wellness programs, health systems can act as exemplars and catalysts for the planetary health agenda, contributing directly to climate resilience, disease prevention, and institutional leadership [57].

Limitations

This study has several limitations. First, it was conducted in a single hospital network and among health care professionals only using a convenience sampling strategy. These features may limit external validity and prevent direct extrapolation to the general population or other institutional contexts. In addition, analyses were descriptive by design, intended to identify behavioral profiles rather than develop predictive models. Second, self-reported data may have introduced social desirability bias, particularly among health care professionals who may overreport sustainable behaviors. Third, the cross-sectional design limits the ability to infer causality or examine the progression from knowledge to behavior over time, highlighting the need for longitudinal studies. Fourth, the sample was disproportionately female, reflecting HCL staffing demographics (75% women) but potentially influencing the results and limiting representativeness. Fifth, the study did not distinguish between clinical staff directly involved in patient care and administrative personnel. In addition, variables such as seniority or cultural or ethnic background were not collected due to institutional and regulatory restrictions, which may have constrained the exploration of factors influencing KAB. Sixth, although structural and institutional influences were conceptually discussed, they were not directly measured, limiting empirical assessment of their moderating effects. Finally, while the PHD offers a globally relevant framework, its applicability varies across cultural and socioeconomic contexts, requiring local adaptation. Future research should explore institutional culture, perceived norms, and structural facilitators to better understand their specific roles in shaping dietary adoption.

Perspectives

In French hospital settings (like the HCL), the 3 KAB profiles can guide targeted action: improve motivation and perceived control for the knowledge-attitude gap; remove practical barriers and adjust the food environment for high intent, low action; and maintain momentum for behavior-driven alignment. Indeed, given its scale and structure, HCL provides a relevant pilot environment to test and refine such approaches before broader implementation across other health care systems. As one of the largest public hospital networks in France [23], it offers a realistic institutional platform to evaluate feasibility and inform national deployment strategies.

Short-term institutional priorities are to align cafeteria offerings with the PHD (default options, placement, and pricing), deliver brief training with clear normative cues, and track change over time using the current KAB scoring framework (light-touch monitoring).

Medium-term priorities are to replicate the approach in other French health care contexts and link results to routine indicators (diet adherence, purchasing, and waste) to evaluate health and sustainability impacts. If routine standardized measurement is needed, a separate program could undertake formal psychometric validation; this is not required to implement the aforementioned actions.

Conclusions

Our study reveals heterogeneity in how health care professionals relate to the PHD. While knowledge remains important, attitudes and contextual factors strongly shape behavior. The identification of a behaviorally aligned cluster with only moderate knowledge underlines the power of cultural and environmental levers. To advance planetary health, health care must move beyond education toward institutional alignment.

XSL•FO
RenderX

Within hospital systems, health care professionals may play an important role in promoting alignment with planetary health principles, potentially contributing to both personal and institutional sustainability goals.

Practical steps such as policy reform, sustainable food guidelines, staff training, and integrating sustainability into curricula are essential to embed these values within health care systems and support meaningful, long-term dietary transformation.

Acknowledgments

The authors would like to thank all survey respondents (professionals of the Hospices Civils de Lyon), the Delphi method experts (Dr Frédéric Valla, nutritionist; Dr Loïc Blanchet-Mazuel, general practitioner; Anne Lastmann, sustainable development consultant; and Cécile Legendre, dietician), and Laurence Mayaud. This work was supported by a grant from the French National Research Agency as part of the Investissements d'Avenir ExcellencES program from France 2030 (SHAPE-Med@Lyon; ANR-22-EXES-0012).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Checklist for Reporting of Survey Studies (CROSS).
[DOCX File , 26 KB - publichealth_v11i1e67633_app1.docx]

References

1. Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, et al. Food in the anthropocene: the EAT-Lancet commission on healthy diets from sustainable food systems. Lancet 2019 Mar 02;393(10170):447-492. [doi: 10.1016/S0140-6736(18)31788-4] [Medline: 30660336]
2. Ye YX, Geng TT, Zhou YF, He P, Zhang JJ, Liu G, et al. Adherence to a planetary health diet, environmental impacts, and mortality in Chinese adults. JAMA Netw Open 2023 Oct 02;6(10):e2339468 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.39468] [Medline: 37874563]
3. Chen H, Wang X, Ji JS, Huang L, Qi Y, Wu Y, et al. Plant-based and planetary-health diets, environmental burden, and risk of mortality: a prospective cohort study of middle-aged and older adults in China. Lancet Planet Health 2024 Aug;8(8):e545-e553 [FREE Full text] [doi: 10.1016/S2542-5196(24)00143-8] [Medline: 39122323]
4. Gu X, Bui LP, Wang F, Wang DD, Springmann M, Willett WC. Global adherence to a healthy and sustainable diet and potential reduction in premature death. Proc Natl Acad Sci U S A 2024 Dec 10;121(50):e2319008121 [FREE Full text] [doi: 10.1073/pnas.2319008121] [Medline: 39621925]
5. Hilary S, Safi S, Sabir R, Numan AB, Zidan S, Platat C. Development and validation of a tool to assess knowledge, attitudes, and practices toward diet sustainability. Front Sustain Food Syst 2024 Aug 16;8:33. [doi: 10.3389/fsufs.2024.1432057]
6. Fresán U, Vidal-Carou MC, Ramos-Truchero G, de Pipaon MS, Moreno LA, Salas-Salvadó J. Knowledge, attitude, and patient advice on sustainable diets among Spanish health professionals. Front Nutr 2023 Jul 17;10:1182226 [FREE Full text] [doi: 10.3389/fnut.2023.1182226] [Medline: 37528999]
7. Spoyalo K, Viduka N, Dixon S, MacNeill AJ, Zhao J. Using salience and availability to promote sustainable and healthy food choices in hospital cafeterias. Sci Rep 2024 Nov 01;14(1):26265 [FREE Full text] [doi: 10.1038/s41598-024-76579-4] [Medline: 39487229]
8. Chatzipavlou M, Karayiannis D, Chaloulakou S, Georgakopoulou E, Poulia KA. Implementation of sustainable food service systems in hospitals to achieve current sustainability goals: a scoping review. Clin Nutr ESPEN 2024 Jun;61:237-252. [doi: 10.1016/j.clnesp.2024.03.030] [Medline: 38777440]
9. Ihsan FR, Bloomfield JG, Monrouxe LV. Triple planetary crisis: why healthcare professionals should care. Front Med (Lausanne) 2024;11:1465662 [FREE Full text] [doi: 10.3389/fmed.2024.1465662] [Medline: 39359929]
10. Frumkin H. Health professionals and the climate crisis. JAMA Netw Open 2024 Mar 04;7(3):e241410 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.1410] [Medline: 38517443]
11. Kotcher J, Maibach E, Miller J, Campbell E, Alqodmani L, Maiero M, et al. Views of health professionals on climate change and health: a multinational survey study. Lancet Planet Health 2021 May;5(5):e316-e323 [FREE Full text] [doi: 10.1016/S2542-5196(21)00053-X] [Medline: 33838130]
12. Non communicable diseases. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2025-04-29]

XSL•FO
RenderX

13. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
14. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2019 May 11;393(10184):1958-1972 [FREE Full text] [doi: 10.1016/S0140-6736(19)30041-8] [Medline: 30954305]
15. Popkin BM, Corvalan C, Grummer-Strawn LM. Dynamics of the double burden of malnutrition and the changing nutrition reality. Lancet 2020 Jan 04;395(10217):65-74 [FREE Full text] [doi: 10.1016/S0140-6736(19)32497-3] [Medline: 31852602]
16. Marmot M. Capabilities, human flourishing, and the health gap. J Hum Dev Capab 2017 Oct 05;18(3):370-383. [doi: 10.1080/19452829.2017.1342362]
17. Myers SS. Planetary health: protecting human health on a rapidly changing planet. Lancet 2017 Dec 23;390(10114):2860-2868. [doi: 10.1016/S0140-6736(17)32846-5] [Medline: 29146123]
18. Vermeir I, Weijters B, De Houwer J, Geuens M, Slabbinck H, Spruyt A, et al. Environmentally sustainable food consumption: a review and research agenda from a goal-directed perspective. Front Psychol 2020;11:1603 [FREE Full text] [doi: 10.3389/fpsyg.2020.01603] [Medline: 32754095]
19. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
20. Nigg CR, Geller KS, Motl RW, Horwath CC, Wertin KK, Dishman RK. A research agenda to examine the efficacy and relevance of the transtheoretical model for physical activity behavior. Psychol Sport Exerc 2011 Jan;12(1):7-12 [FREE Full text] [doi: 10.1016/j.psychsport.2010.04.004] [Medline: 21113323]
21. Mackert M, Walker LO. Cluster analysis identifies subpopulations for health promotion campaign design. Public Health Nurs 2011 Apr 11;28(5):451-457. [doi: 10.1111/j.1525-1446.2011.00948.x] [Medline: 22092430]
22. Murat-Ringot A, Lan R, Fraticelli L, Fayet Y, Bourgeois D, Nugem R, et al. An innovative one health approach: BIOQUALIM, a transdisciplinary research action protocol-from cultivated biodiversity to human health prevention. Nutrients 2024 Oct 15;16(20):3495 [FREE Full text] [doi: 10.3390/nu16203495] [Medline: 39458490]
23. 13 hôpitaux publics d'excellence. Hospices Civils de Lyon. URL: https://www.chu-lyon.fr/hospices-civils-de-lyon [accessed 2025-10-16]
24. Advocacy, communication and social mobilization for TB control: a guide to developing knowledge, attitude and practice surveys. World Health Organization. 2008. URL: https://www.who.int/publications/i/item/9789241596176 [accessed 2025-07-13]
25. Steptoe A, Pollard TM, Wardle J. Development of a measure of the motives underlying the selection of food: the Food Choice Questionnaire. Appetite 1995 Dec;25(3):267-284. [doi: 10.1006/appe.1995.0061] [Medline: 8746966]
26. Lenzer T, Hadler P, Neuert C. Cognitive pretesting. GESIS - Leibniz Institute for the Social Sciences. 2016. URL: https://search.gesis.org/instruments_tools/sg-049-e [accessed 2025-05-29]
27. Yusoff MS. ABC of content validation and content validity index calculation. Educ Med J 2019 Jun 28;11(2):49-54. [doi: 10.21315/eimj2019.11.2.6]
28. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Res Nurs Health 2007 Aug;30(4):459-467. [doi: 10.1002/nur.20199] [Medline: 17654487]
29. Launiala A. How much can a KAP survey tell us about people's knowledge, attitudes and practices? Some observations from medical anthropology research on malaria in pregnancy in Malawi. Anthropol Matters 1970 Jan 01;11(1):e35. [doi: 10.22582/am.v11i1.31]
30. Fayet Y, Praud D, Fervers B, Ray-Coquard I, Blay J, Ducimetiere F, et al. Beyond the map: evidencing the spatial dimension of health inequalities. Int J Health Geogr 2020 Nov 09;19(1):46 [FREE Full text] [doi: 10.1186/s12942-020-00242-0] [Medline: 33298076]
31. Kelly MP, Barker M. Why is changing health-related behaviour so difficult? Public Health 2016 Jul;136:109-116. [doi: 10.1016/j.puhe.2016.03.030] [Medline: 27184821]
32. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
33. Chan MF, Alsaidi Y, Al-Sumri S, Al-Maskari B, Al-Hamrashdi H. Knowledge, attitudes, practice and barriers of physicians to provide tobacco dependence treatment: a cluster analysis. Pan Afr Med J 2021;38:193 [FREE Full text] [doi: 10.11604/pamj.2021.38.193.27047] [Medline: 33995799]
34. Khalis M, Elbadisy I, Bouaddi O, Luo A, Bendriouich A, Addahri B, et al. Cluster analysis of cancer knowledge, attitudes and behaviors in the Moroccan population. BMC Cancer 2024 Jun 01;24(1):669 [FREE Full text] [doi: 10.1186/s12885-024-12226-5] [Medline: 38824496]
35. Mauro F, Greenman P, Di Trani M. Normativity and health promotion across biopsychosocial ecosystems. Front Psychol 2025 Apr 25;16:1570813 [FREE Full text] [doi: 10.3389/fpsyg.2025.1570813] [Medline: 40400744]
36. Roel Lutkenhaus RL, Courtney McLarnon-Silk CM, Frances Walker FW. Norms-shifting on social media. Rev Commun Res 2023 May 14;11:127-149. [doi: 10.52152/rcr.v11.5]

XSL•FO
RenderX

37. Timm A, Kragelund Nielsen K, Alvesson HM, Jensen DM, Maindal HT. Motivation for behavior change among women with recent gestational diabetes and their partners-a qualitative investigation among participants in the face-it intervention. Nutrients 2023 Sep 07;15(18):3906 [FREE Full text] [doi: 10.3390/nu15183906] [Medline: 37764690]
38. Syed S, Acquaye A, Khalfan MM, Obuobisa-Darko T, Yamoah FA. Decoding sustainable consumption behavior: a systematic review of theories and models and provision of a guidance framework. Resour Conserv Recycl Adv 2024 Nov;23:200232. [doi: 10.1016/j.rcradv.2024.200232]
39. Bardenheier BH, Lindley MC, Ball SW, de Perio MA, Laney S, Gravenstein S. Cluster analysis: vaccination attitudes and beliefs of healthcare personnel. Am J Health Behav 2020 May 01;44(3):302-312 [FREE Full text] [doi: 10.5993/AJHB.44.3.3] [Medline: 32295678]
40. Gao CX, Dwyer D, Zhu Y, Smith CL, Du L, Filia KM, et al. An overview of clustering methods with guidelines for application in mental health research. Psychiatry Res 2023 Sep;327:115265 [FREE Full text] [doi: 10.1016/j.psychres.2023.115265] [Medline: 37348404]
41. Rahmania T. Exploring school environmental psychology in children and adolescents: the influence of environmental and psychosocial factors on sustainable behavior in Indonesia. Heliyon 2024 Sep 30;10(18):e37881 [FREE Full text] [doi: 10.1016/j.heliyon.2024.e37881] [Medline: 39315168]
42. Brownlee MT, Powell RB, Hallo JC. A review of the foundational processes that influence beliefs in climate change: opportunities for environmental education research. Environ Educ Res 2013 Feb;19(1):1-20. [doi: 10.1080/13504622.2012.683389]
43. Sheeran P, Webb TL. The intention–behavior gap. Soc Pers Psychol Compass 2016 Sep 05;10(9):503-518. [doi: 10.1111/spc3.12265]
44. Conner M, Norman P. Understanding the intention-behavior gap: the role of intention strength. Front Psychol 2022;13:923464 [FREE Full text] [doi: 10.3389/fpsyg.2022.923464] [Medline: 35992469]
45. Rimal RN. Closing the knowledge-behavior gap in health promotion: the mediating role of self-efficacy. Health Commun 2000;12(3):219-237. [doi: 10.1207/S15327027HC1203_01] [Medline: 10938914]
46. Qumseya B, Goddard A, Qumseya A, Estores D, Draganov PV, Forsmark C. Barriers to clinical practice guideline implementation among physicians: a physician survey. Int J Gen Med 2021;14:7591-7598 [FREE Full text] [doi: 10.2147/IJGM.S333501] [Medline: 34754231]
47. Salari H, Najm F, Yazdankhahfard M, Esfandiari A. Challenges, barriers and solutions for implementing clinical practice guidelines: a qualitative study in southern Iran. BMJ Open Qual 2024 Jul 23;13(3):e002595 [FREE Full text] [doi: 10.1136/bmjoq-2023-002595] [Medline: 39043580]
48. Kok G, Gottlieb NH, Peters GY, Mullen PD, Parcel GS, Ruiter RA, et al. A taxonomy of behaviour change methods: an intervention mapping approach. Health Psychol Rev 2016 Sep;10(3):297-312 [FREE Full text] [doi: 10.1080/17437199.2015.1077155] [Medline: 26262912]
49. Borges-Tiago MT, Almeida A, Tiago FG, Avelar SM. Bridging the innovative attitude–behavior gap: a dual-level analysis. J. Innov Knowl 2024 Oct;9(4):100561. [doi: 10.1016/j.jik.2024.100561]
50. Godin G, Bélanger-Gravel A, Eccles M, Grimshaw J. Healthcare professionals' intentions and behaviours: a systematic review of studies based on social cognitive theories. Implement Sci 2008 Jul 16;3(1):36 [FREE Full text] [doi: 10.1186/1748-5908-3-36] [Medline: 18631386]
51. Khan PB. The role of lifestyle interventions in preventing and managing non-communicable diseases. Non-Commun Dis Care 2025;2(1):60.
52. Hammoudi Halat D, Kassem II, Osman M, Manageiro V. Editorial: world antimicrobial awareness week. Front Public Health 2024;12:1543642 [FREE Full text] [doi: 10.3389/fpubh.2024.1543642] [Medline: 39839430]
53. Carino S, Malekpour S, Porter J, Collins J. The drivers of environmentally sustainable hospital foodservices. Front Nutr 2021;8:740376 [FREE Full text] [doi: 10.3389/fnut.2021.740376] [Medline: 34722609]
54. Braun LT, Dunbar SB. Preventive Cardiovascular Nursing : Resilience across the Lifespan for Optimal Cardiovascular Wellness. London, UK: Browns Books; 2023.
55. The 17 GOALS | sustainable development. United Nations. URL: https://sdgs.un.org/goals [accessed 2025-05-08]
56. Thaler RR, Sunstein CR. Nudge: The Final Edition. New York, NY: Penguin Books; 2021.
57. Health care climate footprint report. Health Care Without Harm. 2024. URL: https://global.noharm.org/resources/health-care-climate-footprint-report [accessed 2025-05-08]

Abbreviations

HCL: Hospices Civils de Lyon
KAB: knowledge, attitudes, and behaviors
NCD: noncommunicable disease
OR: odds ratio
PHD: planetary health diet
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 28.07.25; peer-reviewed by F Denis, C Neuzillet; comments to author 13.10.25; revised version received 16.10.25; accepted 20.10.25; published 12.11.25.

Please cite as:
Carrouel F, Nugem R, Fraticelli L, Fayet Y, Basbous L, Salque C, Dussart C, Charreyre MT, Lan R, Murat-ringot A
Behavioral Profiles and Sociodemographic Predictors of Planetary Health Diet Engagement Among Health Care Professionals to Inform Public Health Promotion: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e67633
URL: https://publichealth.jmir.org/2025/1/e67633
doi:10.2196/67633
PMID:

©Florence Carrouel, Rita Nugem, Laurie Fraticelli, Yohan Fayet, Lama Basbous, Corélie Salque, Claude Dussart, Marie-Thérèse Charreyre, Romain Lan, Audrey Murat-ringot. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing COVID-19 Mortality in Serbia's Capital: Model-Based Analysis of Excess Deaths

Dane Cvijanovic[1*], MD; Nikola Grubor[2*], MD; Nina Rajovic[2], MD; Mira Vucevic[3], MD; Svetlana Miltenovic[3], MD; Marija Laban[4], MD; Tatjana Mostic[5], MD, PhD; Radica Tasic[6], MD, PhD; Bojana Matejic[7*], MD, PhD; Natasa Milic[2,8*], MD, PhD

[1]Department of Cardiovascular Diseases, University Clinical Center Zvezdara, Belgrade, Serbia

[2]Institute of Medical Statistics and Informatics, Faculty of Medicine, University of Belgrade, Dr Subotica 15, Belgrade, Serbia

[3]The City Institute for Public Health, Belgrade, Serbia

[4]Clinic of Pulmonology, University Clinical Center of Serbia, Belgrade, Serbia

[5]Department of Anesthesiology, University Clinical Center of Serbia, Belgrade, Serbia

[6]The College of Health Sciences, Academy of Applied Studies Belgrade, Belgrade, Serbia

[7]Institute of Social Medicine, Faculty of Medicine, University of Belgrade, Belgrade, Serbia

[8]Division of Nephrology and Hypertension, Department of Internal Medicine, Mayo Clinic, Rochester, MN, United States

[*]these authors contributed equally

Corresponding Author:
Natasa Milic, MD, PhD
Institute of Medical Statistics and Informatics, Faculty of Medicine, University of Belgrade, Dr Subotica 15, Belgrade, Serbia

Abstract

Background: Concerns have been raised about discrepancies in COVID-19 mortality data, particularly between preliminary and final datasets of vital statistics in Serbia. In the original preliminary dataset, released daily during the ongoing pandemic, there was an underestimation of deaths in contrast to those reported in the subsequently released yearly dataset of vital statistics.

Objective: This study aimed to assess the accuracy of the final mortality dataset and justify its use in further analyses. In addition, we quantified the relative impact of COVID-19 on the death rate in the Serbian capital's population. In the process, we aimed to explore whether any evidence of cause-of-death misattribution existed in the final published datasets.

Methods: Data were sourced from the electronic databases of the Statistical Office of the Republic of Serbia. The dataset included yearly recorded deaths and the causes of death of all citizens currently living in the territory of Belgrade, the capital of the Republic of Serbia, from 2015 to 2021. Standardization and modeling techniques were utilized to quantify the direct impact of COVID-19 and to estimate excess deaths. To account for year-to-year trends, we used a mixed-effects hierarchical Poisson generalized linear regression model to predict mortality for 2020 and 2021. The model was fitted to the mortality data observed from 2015 to 2019 and used to generate mortality predictions for 2020 and 2021. Actual death rates were then compared to the obtained predictions and used to generate excess mortality estimates.

Results: The total number of excess deaths, calculated from model estimates, was 3175 deaths (99% CI 1715-4094) for 2020 and 8321 deaths (99% CI 6975-9197) for 2021. The ratio of estimated excess deaths to reported COVID-19 deaths was 1.07. The estimated increase in mortality during 2020 and 2021 was 12.93% (99% CI 15.74%-17.33%) and 39.32% (99% CI 35.91%-39.32%) from the expected values, respectively. Those aged 0 - 19 years experienced an average decrease in mortality of 22.43% and 23.71% during 2020 and 2021, respectively. For those aged up to 39 years, there was a slight increase in mortality (4.72%) during 2020. However, in 2021, even those aged 20 - 39 years had an estimated increase in mortality of 32.95%. For people aged 60 - 79 years, there was an estimated increase in mortality of 16.95% and 38.50% in 2020 and 2021, respectively. For those aged >80 years, the increase was estimated at 11.50% and 34.14% in 2020 and 2021, respectively. The model-predicted deaths matched the non-COVID-19 deaths recorded in the territory of Belgrade. This concordance between the predicted and recorded non-COVID-19 deaths provides evidence that the cause-of-death misattribution did not occur in the territory of Belgrade.

Conclusions: The finalized mortality dataset for Belgrade can be safely used in COVID-19 impact analysis. Belgrade experienced a significant increase in mortality during 2020 and 2021, with most of the excess mortality attributable to SARS-CoV-2. Concerns about increased mortality from causes other than COVID-19 in Belgrade seem misplaced as their impact appears negligible.

(JMIR Public Health Surveill 2025;11:e56877) doi:10.2196/56877

XSL•FO
RenderX

KEYWORDS

COVID-19; COVID-19 impact; SARS-Cov-2; coronavirus; respiratory; infectious disease; pulmonary; pandemic; excess mortality; death rate; death toll; centralized health care; urban; Serbia; dense population; public health; surveillance

Introduction

Discrepancies in COVID-19 mortality data for Serbia, particularly between the preliminary and final datasets of vital statistics, have raised concerns. In the original preliminary dataset, released daily during the ongoing pandemic, there was an underestimation of deaths in contrast to those reported in the subsequently released yearly vital statistics [1,2]. This finding led to the recommendation to use only the final datasets from Serbia for further research. Therefore, in this study, we aimed to validate the final mortality dataset for Belgrade, the capital of the Republic of Serbia, by collecting data on COVID-19 and non-COVID-19 causes of death and analyzing it for any evidence of cause-of-death misattribution. We utilized standardization and modeling techniques to quantify the direct impact of COVID-19 on specific age groups and to estimate subgroup-specific excess mortality. Excess mortality, defined as the increase in overall mortality rates beyond what can be anticipated by historical trends, has traditionally served as a valuable metric for estimating the mortality impact of pandemics and other extraordinary occurrences [3]. Monitoring all-cause excess deaths has been lauded as a good metric to determine the COVID-19 impact [4]. It is unaffected by the underlying clinical diagnosis and between-country differences in age and comorbidity because it facilitates comparisons between the same population [5]. While excess mortality does not equate precisely to COVID-19 fatalities, it is widely regarded as the most objective indicator available in numerous countries [3,6]. Death rate standardization and model-derived rates are explicitly compared to highlight the strengths and weaknesses of both approaches.

The literature has shown differences in the case definitions, leading to heterogeneity in COVID-19 mortality reporting and its associated metrics. Faced with the unknown threat of a growing pandemic, several approaches have emerged, ranging from counting only autopsy-confirmed cases, counting clinical diagnoses with or without laboratory confirmation, to counting all suspected cases [5,7]. The chosen approach leads to undercounting if the criteria are too stringent and overcounting if the definition is too broad. Test availability can influence the reported mortality rate if laboratory confirmation is required. In many regions, deaths attributable to COVID-19 were likely undercounted, as official reports often required a positive test result to confirm the virus as the cause of death. During the early stages of the pandemic, limited access to testing, particularly in long-term care facilities, meant that many deaths went unattributed to COVID-19. Furthermore, the systems for registering deaths vary widely between countries, affecting the consistency and comprehensiveness of reported mortality data [4,8,9]. Numerous studies have since attempted to estimate the true mortality impact of the pandemic, with the World Health Organization (WHO) estimating 3 million excess deaths globally for 2020 alone [10]. However, more detailed estimates by location remain unavailable [4,11]. It is unclear whether excess death calculations should be attempted in countries with unreliable data. Hence, focusing on settings with high-quality data may be more productive, such as cohorts with intensive data collection and accurate death ascertainment [12]. The pandemic has highlighted that death registration systems are not sufficiently robust to accommodate such large-scale events.

However, case definitions and data manipulation do not account for all differences. In addition to deaths directly caused by the SARS-CoV-2 infection, pandemic-related restrictions, such as social distancing mandates and lockdowns, likely influenced mortality rates from other causes. For example, some diseases and injuries, such as road accidents, may have decreased during the pandemic. At the same time, deaths from chronic and acute conditions might have increased due to delayed care in overwhelmed health care systems [13,14]. These shifts in mortality patterns complicate the calculation of excess deaths, as it becomes difficult to separate the impact of SARS-CoV-2 from the broader societal, economic, and behavioral changes triggered by the pandemic [8]. This challenge is particularly pronounced in countries needing more detailed cause-of-death data. While distinguishing between the myriads of influences on excess mortality would be illuminating, estimating excess mortality is a more attainable and equally important goal. Seemingly similar urban environments differed in their resilience to the COVID-19 pandemic. Certain aspects of these urban environments, such as response efficiency, optimized testing and detection strategies, and practical immunization efforts, make them more or less resilient to the spread of the SARS-CoV-2 [15]. Analyzing the impact of COVID-19 on urban environments can shed light on the epidemiological burden and help quantify the effectiveness of government policies [16]. Different geographical layouts and heterogeneous urban design have demonstrated the potential to lead to a highly nonuniform epidemic distribution [11,17]. Taking this into account, the aim of this study was to validate the final mortality dataset and assess the impact of COVID-19 on excess mortality in Belgrade, the capital of the Republic of Serbia.

Methods

Data Collection

The study focused on the recorded deaths and the causes of death of all citizens currently living in the territory of Belgrade for the period from 2015 to 2021. The age-specific and sex-stratified midyear population estimates for the territory of Belgrade were collected from the electronic databases of the Statistical Office of the Republic of Serbia for the given period [18]. These midyear population estimates were calculated considering the natural population change and net migration rates. Overall mortality data, as well as data on the mortality from COVID-19 as the main cause of death, were collected. The mortality data for COVID-19 as the primary cause of death were coded per the WHO guidelines [19]. These guidelines define COVID-19 as a cause of death if clinical death occurs,

XSL•FO
RenderX

which is compatible with COVID-19 disease, in a case with confirmed or probable COVID-19 disease, and no alternative, more probable causes exist. For this analysis, we used the Codes for special purposes (U00-U89), Chapter XXII, *ICD-10*, where only deaths with the main cause of death–COVID-19 (U07.1, U07.2) were exclusively attributed for 2020 and 2021. This coding scheme was used to derive crude COVID-19 and non-COVID-19 mortality estimates. The data were also sourced from the electronic databases of the Statistical Office of the Republic of Serbia [18].

Excess Death Estimation

We first calculated the standardized direct death rates for the populations from 2015 to 2019 to estimate the excess deaths [20]. The 2020 and 2021 population estimates used this standardized death rate to generate age-specific and sex-stratified mortality. This total was then compared to the actual crude death rate. Direct standardization is not without its limits. It can severely underestimate excess deaths due to COVID-19. The 5-year average estimate does not capture seasonal variation and year-to-year mortality. To account for these year-to-year trends, we use a mixed-effects hierarchical Poisson generalized linear regression model (GLM) to predict mortality for 2020 and 2021 using a yearly time series of sex, age group membership, and the logarithm of absolute population count in the region. The model was fit to mortality data observed from 2015 to 2019. To fit the Bayesian model, we estimated the parameter posterior distributions using Hamiltonian Markov Chain Monte Carlo, as implemented in Stan (version 2.26.1). The posterior distributions quantify which parameter values are most consistent with the data and are subsequently used for generating predictions. We assessed convergence by the inspection of trace plots and R-hat values, which should be below 1.01 [21], and the effective sample size, which should be greater than 1000 [22]. These quality assurance measures indicate that the sampler has achieved robust convergence and that the tails of the parameter distributions are reliable when estimating confidence intervals of excess mortality. Model out-of-sample predictive performance was assessed using leave-one-out cross-validation approximated via Pareto smoothed importance sampling [23]. Leave-one-out cross-validation metrics estimate model performance on unseen data and are useful measures of out-of-sample predictive performance. The leave-one-out cross-validation indicated that the predictors contain valuable predictive information. Priors were chosen through prior predictive simulation so that predata predictions span the range of scientifically plausible outcomes. Using weakly informative priors ensures that the model is partially regularized and protects against implausible parameter values. We constructed posterior predictive distributions to assess the model fit adequacy to evaluate predictions at the observed data's mean (Multimedia Appendix 1) [24]. Posterior predictive distributions evaluated the models' predictions for consistency against the observed data before any inference could be made. Furthermore, to assess model calibration across the whole observed distribution, we compared the probability integral transformed values of the predictive distribution to our data's empirical (observed) cumulative distribution function (Multimedia Appendix 2) [25]. The model convergence diagnostic is shown in Multimedia Appendix 3. After successful model fitting, we generated predictions for each age group for the next two pandemic years and compared the expected and the observed mortalities.

The raw number of excess deaths was not comparable across countries due to significant differences in the population. To enable easier comparison across countries, we calculated the P-score. The P-Score measures the percentage difference between the reported and projected deaths.

All calculations were performed using R software (version 4.3.1; R Foundation for Statistical Computing).

Ethical Considerations

This study was the secondary analysis of publicly available data collected by the Statistical Office of the Republic of Serbia. Data collection adhered to strict ethical guidelines ensuring the confidentiality and privacy of all participants. All collected data were fully anonymized to prevent the identification of individuals.

Results

Population Characteristics

From 2015 to 2019, the territory of Belgrade had an average population of 1,688,185, of which 797,080 (47.22%) citizens were male and 891,105 (52.78%) were female. The largest number of people were included in the 40 - 59 year age group (459,345), closely followed by the 20 - 39 year age group (459,032). The number of people aged >80 years was the lowest (72,757). These demographics remained relatively stable throughout 2021 (Multimedia Appendix 4). The raw population characteristics are provided in Multimedia Appendix 5.

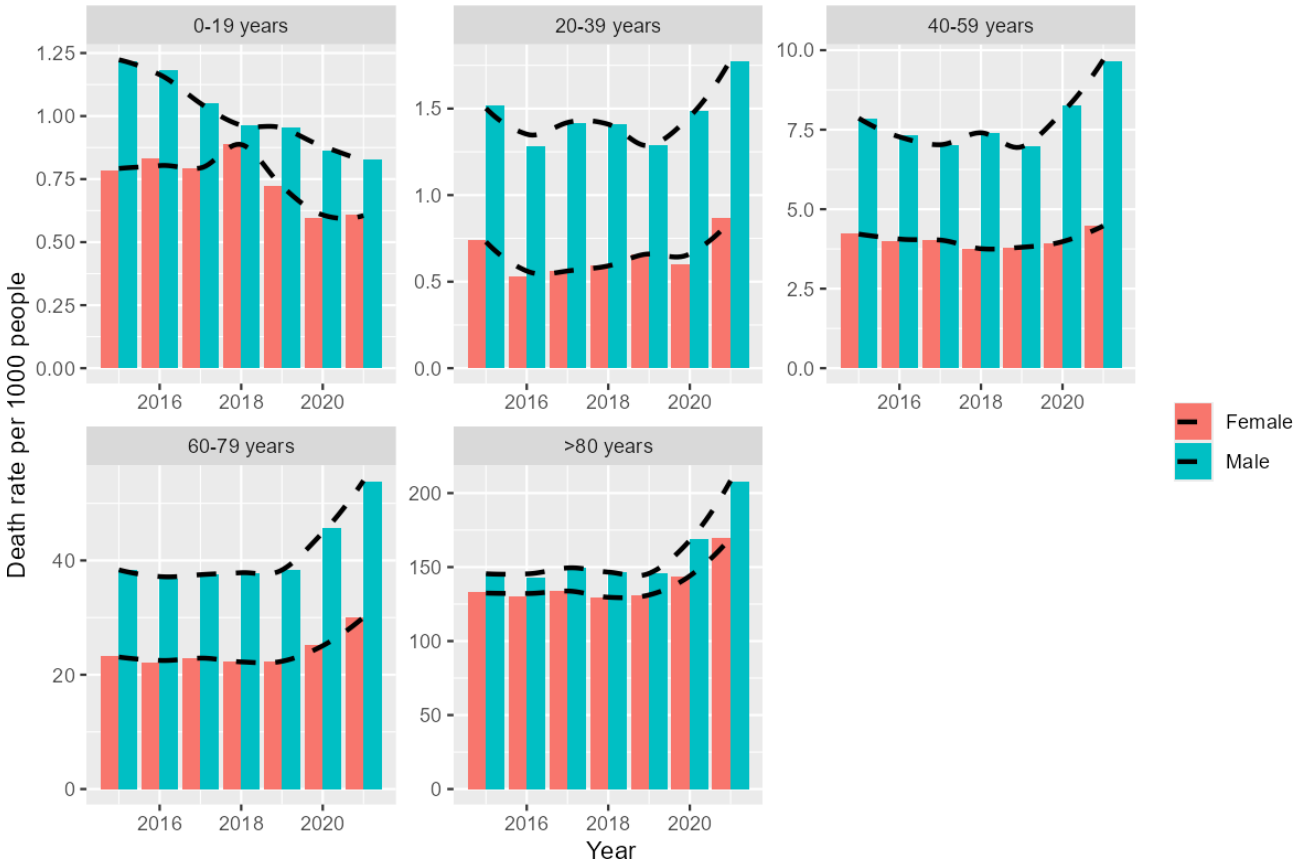
Death Rates

In the territory of Belgrade, there were 27,722 and 32,593 deaths in 2020 and 2021, respectively. The average number of deaths in the previous 5-year period was 23,792. In 2020 and 2021, there were 3028 (12%) and 7669 (24%) deaths when COVID-19 was coded as the primary cause.

The crude death rate increased significantly in 2020 and 2021 from its average of 14.1 per 1000 people in the previous 5-year period to an average of 17.85 per 1000 people. Those in their seventh and eighth decade were the most severely impacted age groups. Age-specific crude death rates for citizens in their seventh decade of life were 33.90 per 1000 people and 40.08 per 1000 people in 2020 and 2021, respectively. Those in their eight-decade had crude death rates of 153.24 per 1000 people and 183.91 per 1000 people for 2020 and 2021, respectively. Raw death rate trends are shown in Figure 1.

XSL•FO
RenderX

Figure 1. Trends of crude death rates per 1000 people stratified by sex and age group for 2015 - 2021. A decrease in death rate among children and young people is evident, while all other age groups showed an increase in death rate during the pandemic (2020 and 2021) relative to the prepandemic period.

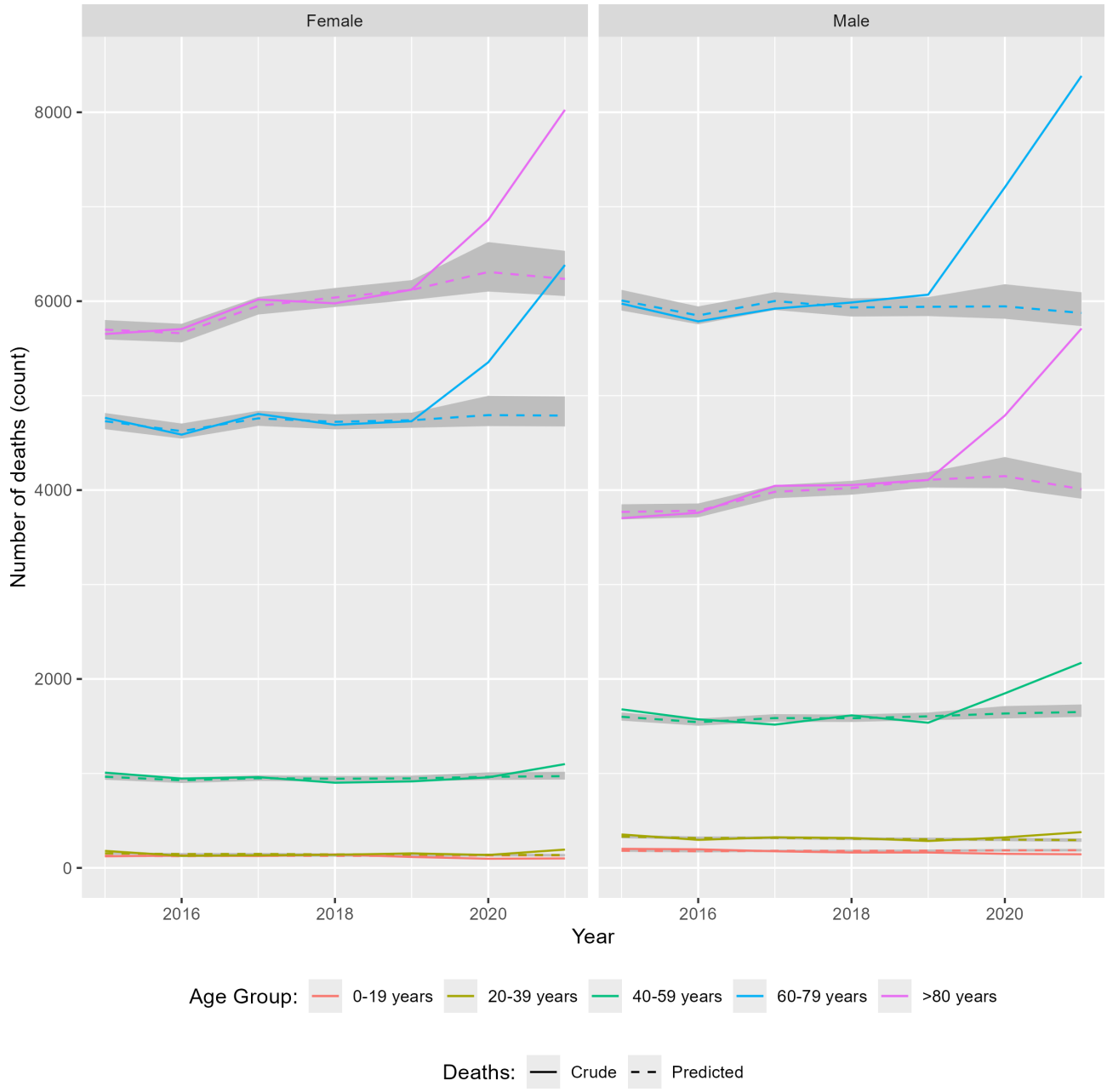


Excess Mortality

The total number of excess deaths, calculated from model estimates, was 3175 deaths (99% CI 1715-4094) for 2020 and 8321 deaths (99% CI 6975-9197) for 2021. The ratio of estimated excess deaths to reported COVID-19 deaths was 1.07. Excess mortality, stratified by age group and sex, is shown in Figures 2-4. The estimated increase in mortality during 2020 was 12.93% (99% CI 15.74-17.33%). During 2021, mortality increased by 39.32% (99% CI 35.91-39.32%) from the expected value. The average increase in mortality during 2020 and 2021 was 23.55%, compared to the expected mortality extrapolated from the previous 5-year period. There seemed to be fewer deaths than expected based on the predicted trends in children and young people, which aligned with the raw observed data. To be precise, between 40 and 106 and between 44 and 110 fewer deaths occurred among children and young people in 2020 and 2021, respectively; this amounted to an average decrease in mortality of 22.43% and 23.71% for this age group. The most impacted groups were people in their seventh and eighth decades. People aged 60 - 79 years experienced an estimated 1279 to 2143 and 3594 to 4423 excess deaths in 2020 and 2021, respectively. People aged >80 years experienced an estimated 504 to 1625 and 2895 to 3858 excess deaths in 2020 and 2021, respectively. In people aged 60 to 79 years, there was an estimated increase in mortality of 16.95% and 38.50% in 2020 and 2021, respectively. For those aged >80 years, the increase was estimated at 11.50% and 34.14% in 2020 and 2021, respectively. For those aged up to 39 years old, there was little increase in mortality (4.72%) during 2020. However, in 2021, even those aged 20 - 39 had an estimated increase in mortality of 32.95%. The crude, standardized, and model-predicted death rates are directly compared in Figure 3 and Tables 1 and 2. The model-predicted deaths matched the number of non-COVID deaths recorded in the territory of Belgrade (Figure 3). This concordance between the predicted and recorded non-COVID deaths provided evidence that the cause-of-death misattribution did not occur in the territory of Belgrade. The error whiskers represent 99% credible intervals.

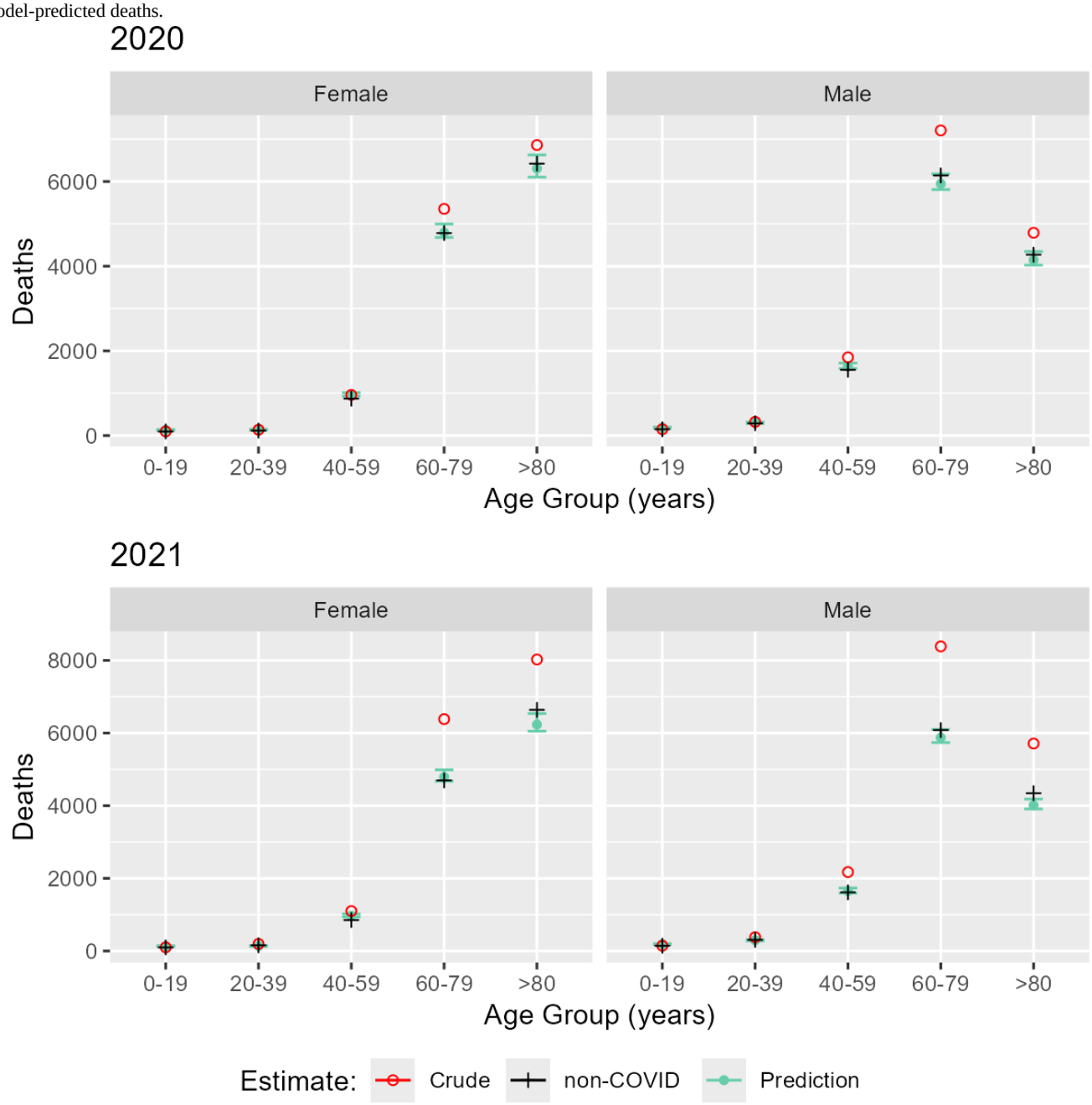
XSL•FO
RenderX

Figure 2. Raw death counts compared with the model-predicted death counts for the study period (2015 - 2021). The raw number of deaths per year, stratified by sex and age group, is shown. Dashed lines represent model-predicted deaths. The grey-shaded area around the dashed lines represents the 99% confidence interval. An excess number of deaths is evident as a rise in the observed deaths (solid lines) compared to the prepandemic model predictions (dashed lines). A larger discrepancy between the lines indicates a larger impact of COVID-19 on those subgroups.



XSL•FO
RenderX

Figure 3. Crude, non-COVID, and model-predicted death comparisons. The model-predicted deaths fall within the confidence intervals of non-COVID deaths recorded in the territory of Belgrade. Agreement between the predicted and recorded non-COVID deaths provides evidence that cause-of-death misattribution did not occur in the territory of Belgrade for the period 2020 - 2021. The error whiskers represent 99% confidence intervals of model-predicted deaths.



XSL•FO
RenderX

Figure 4. Excess deaths stratified by age group and sex during 2020 and 2021. The heights of each bar represent accumulated excess deaths for their respective year with associated 99% confidence intervals (black error whiskers). Children and young people had fewer deaths than expected, while all other age groups showed greater excess mortality. Older age groups are more severely impacted as evident by the relative height of the bars.

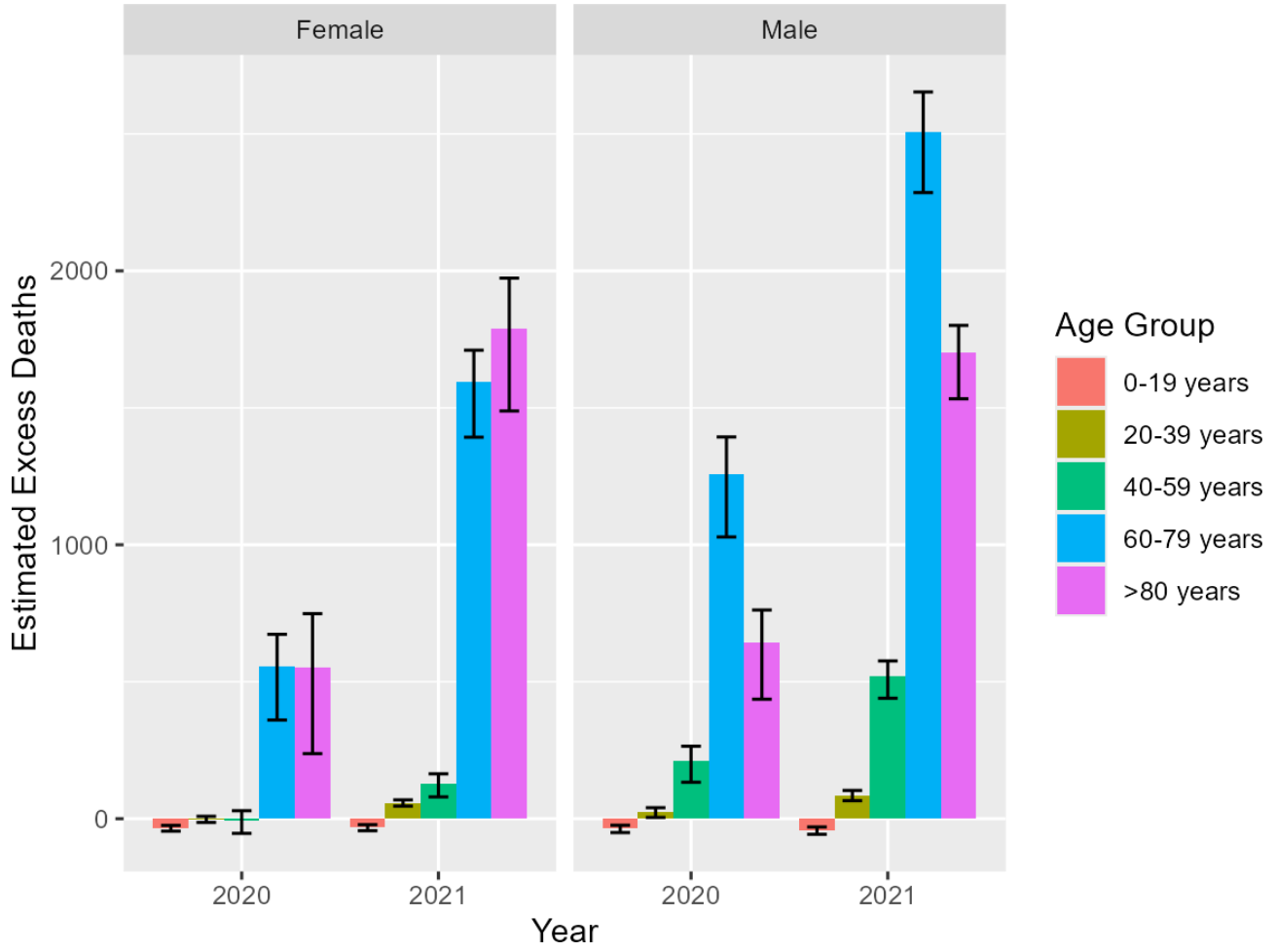


Table . Female-specific death rates per 1000 people.

Year	Crude death rate	Standardized death rate	Predicted death rate (99% CI)
2015	13.23	13.75	12.82 - 13.54
2016	12.94	13.23	12.59 - 13.27
2017	13.53	13.62	13.08 - 13.74
2018	13.28	13.20	13.08 - 13.78
2019	13.48	13.31	13.18 - 13.86
2020	14.99	14.61	13.24 - 14.66
2021	17.71	17.32	13.23 - 14.56

Table . Male-specific death rates per 1000 people.

Year	Crude death rate	Standardized death rate	Predicted death rate (99% CI)
2015	15.01	15.41	14.57 - 15.43
2016	14.60	14.77	14.30 - 15.07
2017	15.04	15.05	14.77 - 15.54
2018	15.21	15.10	14.68 - 15.46
2019	15.20	15.01	14.78 - 15.58
2020	17.90	17.66	14.71 - 16.14
2021	21.08	21.02	14.58 - 15.86

XSL•FO
RenderX

Discussion

Study Findings and Comparison With Previous Works

When revised death statistics were released in Serbia, some authors noted significant discrepancies between the preliminary and finalized COVID-19 dataset [1,2]. Specifically, considerable mortality underreporting occurred in the initial dataset compared to the finalized COVID-19 mortality dataset. We collected data on the number of deaths in Belgrade, coding the SARS-CoV-2 as the primary cause, to validate the accuracy of this finalized dataset. We opted for a modeling-based approach because classical standardization does not address the hypothesized problem of death misattribution, which involves coding another primary cause of death as a COVID-19 death [20]. We fitted a Poisson GLM model to the acquired data for 2015 - 2019, and then used the model to generate mortality predictions for 2020 and 2021. Given the expected age and comorbidity structure within the population over the studied time frame, the model's mortality estimates reflect the expected deaths in the absence of the pandemic. When we removed COVID-19 as a cause of death from the dataset, the model's predictions matched the reported deaths, indicating that there were no significant cause-of-death misattributions in the final published dataset for Belgrade. Furthermore, this establishes a firm link between excess deaths and COVID-19 in Belgrade.

Possible mechanisms for this undercount discrepancy have yet to be fully investigated. Differences in the case definition were a common explanation for reporting heterogeneity between countries [5,26,27]. However, according to official reports, Serbia adopted the case definition proposed by the WHO early on during the pandemic, which remained unchanged throughout [1]. The ratio of excess mortality to the officially reported COVID-19 death count (undercount ratio) for the amended dataset was 1.07, indicating slight undercounting. This is a substantial improvement compared to the ratio of greater than 3.00 reported in the preliminary dataset [1]. The undercount ratio was larger in 2020 than in 2021, with a slight decrease in divergence. This improvement is consistent with a worldwide improvement in undercount ratios throughout the pandemic, attributed to improvements in accuracy, case definition revisions, and better reporting [3]. However, substantial undercounting remained in the preliminary dataset. Our analysis indicates that the revised datasets accurately reflect actual COVID-19 mortality, at least in Belgrade. This suggests that accurate death certificate data were collected, and that the problem was localized to the reporting system. No prior reporting system of this scale had been used in Serbia; thus, a learning curve could have existed initially, causing delays. However, the divergence between the actual and reported deaths continues into 2021, making any attribution to a learning curve unlikely.

Despite intense government measures to curb the spread of the SARS-CoV-2, the territory of Belgrade had the most considerable death rates in the country [2]. There is a complex interplay of contributing causes to this increase in mortality. Belgrade is home to one of the world's largest medical health care centres. The University Clinical Centre of Serbia possesses a hospital floor capacity of 3150 beds [28]. The health care system in Serbia is heavily centralized, with most medical services provided by the state. Overreliance on the centralized health care system may have contributed to delays in reporting death, particularly given worldwide reports of barriers to access in rural areas [29-31]. Health care disparities between communities of differing socioeconomic status and mobility are a well-recognized moderator of pandemic impact [31-34]. The government's response included the incorporation of many hospitals into a 'COVID-19 Response System,' which had limited capacity to treat non-COVID-19 diseases. Thus, treatment priorities could have contributed to limited health care access and might have increased morbidity and mortality due to other causes. However, as our model's death predictions account for the non-COVID deaths almost perfectly, especially in the lower age subgroups, deaths from other causes could not have been acute—or at least, could not have occurred during 2020 and 2021. Increased mortality and morbidity due to the delayed diagnosis of oncological and other chronic diseases in the following years can be an expected consequence, which could manifest as a delayed increases in disease burden. Further studies in this area would shed light on the delayed impact of COVID-19.

On assessing the stringency index, a composite measure of 9 government response metrics, we found that Serbia quickly reached the maximal (strictest response) by March 22, 2020. Both stringency as well as containment and health indexes (an extension of the stringency index) settled into a 40 - 70 range during 2020 - 2021, briefly dropping below 40 for the period leading up to and including the June 21, 2020 elections [35]. Increased deaths led to more stringent government responses, while every reduction led to waning compliance [36]. Over time, people become less risk-averse, and government policies lose effectiveness, especially after initiating the vaccination campaign. The relative percentage of people who received all doses prescribed by the initial vaccination protocol was 46.87% by January 1, 2022, and showed significant plateauing [37]. The relative effect of nonpharmaceutical interventions in general declined by 12% from May 2021 due to less stringent methods and the introduction of vaccination strategies [38].

Mortality from COVID-19 is more significant among those with advanced age, mimicking overall mortality [39-41]. The territory of Belgrade observed a reduction in the mortality rate of children and young people (aged 0 - 19 years) during 2020 and 2021. COVID-19 deaths in this age group were very low (1 case in 2020 for males and 0 for females). Most of the mortality risk in children and young people comes from injuries and infectious diseases [42]. Government intervention and the mobility restrictions during the pandemic probably contributed to lower mortality in children and young people by indirectly lowering their overall risk. For example, traffic accident fatalities decreased in the Western Balkans during lockdowns [43]. The evidence suggests a decrease in pediatric intensive care unit admissions between 2020 and 2021. Additionally, research has found no increase in adolescent self-harm or suicide [44]. These effects contributed to a decrease in the death rate among the child and adolescent population. However, the changes in health care functioning during the same period could have had

XSL•FO
RenderX

long-term consequences in this age group, which have yet to manifest.

Limitations

We acknowledge that the estimates obtained are Bayesian and are posterior distributions of expected deaths. As some of the raw non-COVID-19 deaths are on the edges of our estimates, where the predicted response parameter is least likely (given our data), there is room to argue that deaths from other causes might have increased for specific age groups. In contrast, other age groups showed fewer deaths than expected. Additionally, the resolution of our data limited our modeling choice, as the most accurate dataset only includes measurements from yearly reports.

Conclusion

The finalized mortality dataset for Belgrade can be safely used in COVID-19 impact analysis. In 2020 and 2021, Belgrade experienced a significant increase in mortality, with the majority of the excess mortality attributable to the SARS-CoV-2. Concerns about increased mortality from causes other than COVID-19 seem misplaced, as their impact appears negligible. However, we do caution that the delayed effects of the SARS-CoV-2, barriers to health care, and the psychological impact of the COVID-19 pandemic on Belgrade territory remain unexplored.

Data Availability

The datasets analyzed during this study are publicly available from the electronic databases of the Statistical Office of the Republic of Serbia [45,46].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Model posterior predictive check. Observed data (black line) compared with 100 draws from the posterior predictive distribution (light blue). Bayesian models are generative and can simulate data after successful fitting. The posterior predictive check can ensure the model has successfully learned from the data and indicates that the fit accurately predicts the mean number of expected deaths given the predictors.
[PNG File, 25 KB - publichealth_v11i1e56877_app1.png]

Multimedia Appendix 2
Model probability integral transform (PIT) values plotted against the empirical cumulative distribution function (ECDF). The curves fall across the diagonal and stay within the confidence bands, indicating that the model's predictive distribution is well-calibrated. This ensures that the predictions are reliable across the whole range of the distribution and that the model is not under- or overestimating deaths for certain subgroups.
[PNG File, 35 KB - publichealth_v11i1e56877_app2.png]

Multimedia Appendix 3
Model diagnostic Markov Chain Monte Carlo (MCMC) chain trace plots show adequate chain mixing for estimated parameters. MCMC chains are used to sample parameter values from the model and their adequate mixing ensures that the predictions are stable and consistent with the observed data. The chains ran for 4000 iterations, with an adaptive warm-up period of 1000 iterations.
[PNG File, 188 KB - publichealth_v11i1e56877_app3.png]

Multimedia Appendix 4
Demographic pyramid comparison between the years 2015 and 2021. Relative age and sex subgroups remained stable over the study period ensuring that the baseline populations are comparable.
[PNG File, 32 KB - publichealth_v11i1e56877_app4.png]

Multimedia Appendix 5
Midyear serial baseline population structure in Belgrade, Serbia's capital stratified by age and sex. Midyear population estimates were calculated considering natural population change and net migration rates.
[XLSX File, 11 KB - publichealth_v11i1e56877_app5.xlsx]

References

1. Galjak M, Marinković I. Discrepancies between preliminary and final COVID-19 mortality data-the case of Serbia. Ann Epidemiol 2023 Aug;84:41-47. [doi: 10.1016/j.annepidem.2023.05.006] [Medline: 37196849]

XSL•FO
RenderX

2. Marinkovic I, Galjak M. Excess mortality in Europe and Serbia during the COVID-19 pandemic in 2020. Stanovnishtvo 2021;59(1):61-73. [doi: 10.2298/STNV2101061M]
3. Karlinsky A, Kobak D. Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset. Elife 2021 Jun 30;10:e69336. [doi: 10.7554/eLife.69336] [Medline: 34190045]
4. COVID-19 Excess Mortality Collaborators. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. Lancet 2022 Apr 16;399(10334):1513-1536. [doi: 10.1016/S0140-6736(21)02796-3] [Medline: 35279232]
5. Corrao G, Rea F, Blangiardo GC. Lessons from COVID-19 mortality data across countries. J Hypertens 2021 May 1;39(5):856-860. [doi: 10.1097/HJH.0000000000002833] [Medline: 33657587]
6. Beaney T, Clarke JM, Jain V, et al. Excess mortality: the gold standard in measuring the impact of COVID-19 worldwide? J R Soc Med 2020 Sep;113(9):329-334. [doi: 10.1177/0141076820956802] [Medline: 32910871]
7. Feyissa GT, Tolu LB, Ezeh A. COVID-19 death reporting inconsistencies and working lessons for low- and middle-income countries: opinion. Front Med (Lausanne) 2021;8:595787. [doi: 10.3389/fmed.2021.595787] [Medline: 33681241]
8. Lau H, Khosrawipour T, Kocbach P, Ichii H, Bania J, Khosrawipour V. Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters. Pulmonol 2021;27(2):110-115. [doi: 10.1016/j.pulmoe.2020.05.015] [Medline: 32540223]
9. Kobak D. Excess mortality reveals Covid's true toll in Russia. Signif (Oxf) 2021 Feb;18(1):16-19. [doi: 10.1111/1740-9713.01486] [Medline: 33821160]
10. World Health Organization. The true death toll of COVID-19: estimating global excess mortality. URL: https://www.who.int/data/stories/the-true-death-toll-of-covid-19-estimating-global-excess-mortality [accessed 2024-10-01]
11. GBD 2021 Demographics Collaborators. Global age-sex-specific mortality, life expectancy, and population estimates in 204 countries and territories and 811 subnational locations, 1950-2021, and the impact of the COVID-19 pandemic: a comprehensive demographic analysis for the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):1989-2056. [doi: 10.1016/S0140-6736(24)00476-8] [Medline: 38484753]
12. Ioannidis JPA, Zonta F, Levitt M. Flaws and uncertainties in pandemic global excess death calculations. Eur J Clin Invest 2023 Aug;53(8):e14008. [doi: 10.1111/eci.14008] [Medline: 37067255]
13. Mungmunpuntipantip R, Wiwanitkit V. The COVID-19 pandemic and traffic accidents. S Afr Med J 2021 Mar 2;111(3):13227. [doi: 10.7196/SAMJ.2021.v111i3.15531] [Medline: 33944735]
14. Salottolo K, Caiafa R, Mueller J, et al. Multicenter study of US trauma centers examining the effect of the COVID-19 pandemic on injury causes, diagnoses and procedures. Trauma Surg Acute Care Open 2021;6(1):e000655. [doi: 10.1136/tsaco-2020-000655] [Medline: 33884307]
15. Truszkowska A, Fayed M, Wei S, et al. Urban determinants of COVID-19 spread: a comparative study across three cities in New York state. J Urban Health 2022 Oct;99(5):909-921. [doi: 10.1007/s11524-022-00623-9] [Medline: 35668138]
16. Salama AM. Coronavirus questions that will not go away: interrogating urban and socio-spatial implications of COVID-19 measures. Emerald Open Res 2020;2:14. [doi: 10.35241/emeraldopenres.13561.1]
17. Brizuela NG, García-Chan N, Gutiérrez Pulido H, Chowell G. Understanding the role of urban design in disease spreading. Proc R Soc A 2021 Jan;477(2245):20200524. [doi: 10.1098/rspa.2020.0524]
18. Republika srbija [Article in Serbian]. Republički zavod za statistiku. URL: https://www.stat.gov.rs [accessed 2024-12-23]
19. World Health Organization. International Guidelines for Certification and Classification (Coding) of COVID-19 as Cause of Death. URL: https://www.who.int/publications/m/item/international-guidelines-for-certification-and-classification-(coding)-of-covid-19-as-cause-of-death [accessed 2023-10-19]
20. Naing NN. Easy way to learn standardization: direct and indirect methods. Malays J Med Sci 2000 Jan;7(1):10-15. [Medline: 22844209]
21. Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner PC. Rank-normalization, folding, and localization: an improved Rˆ for assessing convergence of MCMC (with discussion). Bayesian Anal 2021 Jun 1;16(2). [doi: 10.1214/20-BA1221]
22. Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. J Stat Soft 2017;80(1):1-28. [doi: 10.18637/jss.v080.i01]
23. Vehtari A, Simpson D, Gelman A, Yao Y, Gabry J. Pareto smoothed importance sampling. arXiv. Preprint posted online on Aug 4, 2022. [doi: 10.48550/arXiv.1507.02646]
24. Gabry J, Simpson D, Vehtari A, Betancourt M, Gelman A. Visualization in Bayesian workflow. J R Stat Soc Ser A Stat Soc 2019 Feb 1;182(2):389-402. [doi: 10.1111/rssa.12378]
25. Säilynoja T, Bürkner PC, Vehtari A. Graphical test for discrete uniformity and its applications in goodness-of-fit evaluation and multiple sample comparison. Stat Comput 2022 Apr 15;32(2). [doi: 10.1007/s11222-022-10090-6]
26. Marinković I, Tramošljanin A, Galjak M. Assessing the availability and quality of COVID-19 mortality data in Europe: a comparative analysis. Eur J Public Health 2023 Oct 10;33(5):944-946. [doi: 10.1093/eurpub/ckad088] [Medline: 37263603]
27. Seghezzo G, Allen H, Griffiths C, et al. Comparison of two COVID-19 mortality measures used during the pandemic response in England. Int J Epidemiol 2024 Feb 1;53(1):dyad116. [doi: 10.1093/ije/dyad116] [Medline: 37615969]
28. Uredba o planu mreže zdravstvenih ustanova [Article in Serbian]. URL: https://www.paragraf.rs/propisi/uredba_o_planu_mreze_zdravstvenih_ustanova.html [accessed 2023-10-25]

XSL•FO
RenderX

29. Azar KMJ, Shen Z, Romanelli RJ, et al. Disparities in outcomes among COVID-19 patients in a large health care system in California. Health Aff (Millwood) 2020 Jul;39(7):1253-1262. [doi: 10.1377/hlthaff.2020.00598] [Medline: 32437224]
30. Núñez A, Sreeganga SD, Ramaprasad A. Access to healthcare during COVID-19. Int J Environ Res Public Health 2021 Mar 14;18(6):2980. [doi: 10.3390/ijerph18062980] [Medline: 33799417]
31. Abedi V, Olulana O, Avula V, et al. Racial, economic, and health inequality and COVID-19 infection in the United States. J Racial Ethn Health Disparities 2021 Jun;8(3):732-742. [doi: 10.1007/s40615-020-00833-4] [Medline: 32875535]
32. Bambra C, Riordan R, Ford J, Matthews F. The COVID-19 pandemic and health inequalities. J Epidemiol Community Health 2020 Nov;74(11):964-968. [doi: 10.1136/jech-2020-214401] [Medline: 32535550]
33. Solís Arce JS, Warren SS, Meriggi NF, et al. COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries. N Med 2021 Aug;27(8):1385-1394. [doi: 10.1038/s41591-021-01454-y] [Medline: 34272499]
34. Dror AA, Eisenbach N, Taiber S, et al. Vaccine hesitancy: the next challenge in the fight against COVID-19. Eur J Epidemiol 2020 Aug;35(8):775-779. [doi: 10.1007/s10654-020-00671-y] [Medline: 32785815]
35. Hale T, Angrist N, Goldszmidt R, et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). Nat Hum Behav 2021 Apr;5(4):529-538. [doi: 10.1038/s41562-021-01079-8] [Medline: 33686204]
36. Agyapon-Ntra K, McSharry PE. A global analysis of the effectiveness of policy responses to COVID-19. Sci Rep 2023 Apr 6;13(1):5629. [doi: 10.1038/s41598-023-31709-2] [Medline: 37024541]
37. Mathieu E, Ritchie H, Rodés-Guirao L, et al. Coronavirus Pandemic (COVID-19): Our World in Data; 2020.
38. Ge Y, Zhang WB, Wu X, et al. Untangling the changing impact of non-pharmaceutical interventions and vaccination on European COVID-19 trajectories. Nat Commun 2022 Jun 3;13(1):3106. [doi: 10.1038/s41467-022-30897-1] [Medline: 35661759]
39. Goldstein JR, Lee RD. Demographic perspectives on the mortality of COVID-19 and other epidemics. Proc Natl Acad Sci U S A 2020 Sep 8;117(36):22035-22041. [doi: 10.1073/pnas.2006392117] [Medline: 32820077]
40. Shang W, Wang Y, Yuan J, Guo Z, Liu J, Liu M. Global excess mortality during COVID-19 pandemic: a systematic review and meta-analysis. Vaccines 2022 Oct 12;10(10):1702. [doi: 10.3390/vaccines10101702] [Medline: 36298567]
41. Safavi-Naini SAA, Farsi Y, Alali WQ, Solhpour A, Pourhoseingholi MA. Excess all-cause mortality and COVID-19 reported fatality in Iran (April 2013-September 2021): age and sex disaggregated time series analysis. BMC Res Notes 2022 Apr 5;15(1):130. [doi: 10.1186/s13104-022-06018-y] [Medline: 35382865]
42. Flaxman S, Whittaker C, Semenova E, et al. Assessment of COVID-19 as the underlying cause of death among children and young people aged 0 to 19 years in the US. JAMA Netw Open 2023 Jan 3;6(1):e2253590. [doi: 10.1001/jamanetworkopen.2022.53590] [Medline: 36716029]
43. Transport Community. Fatalities for 2020 annual statistics for Western Balkans. : Transport Community Permanent Secretariat; 2021 URL: https://www.transport-community.org/publications/fatalities-for-2020-annual-statistics-for-western-balkans-2/ [accessed 2024-12-11]
44. McCluskey CK, Black TR, Zee-Cheng J, et al. Suicide and self-harm in adolescents during the COVID-19 pandemic: A U.S. virtual pediatric systems, LLC, database study of PICU admissions, 2016-2021. Pediatr Crit Care Med 2024 Feb 1;25(2):e73-e81. [doi: 10.1097/PCC.0000000000003381] [Medline: 37812055]
45. Statistical Office of the Republic of Serbia. Deaths by causes of death, sex and age. URL: https://data.stat.gov.rs/Home/Result/18030302?languageCode=sr-Latn [accessed 2025-04-17]
46. Statistical Office of the Republic of Serbia. Estimates of population by age and sex (beginning, middle and end of year). URL: https://data.stat.gov.rs/Home/Result/18010403?languageCode=sr-Latn [accessed 2025-04-17]

Abbreviations

GLM: generalized linear regression model
WHO: World Health Organization

Edited by A Mavragani; submitted 05.02.24; peer-reviewed by SA Ahmad Safavi-Naini, S Ouyang; revised version received 01.10.24; accepted 02.10.24; published 17.04.25.

Please cite as:
Cvijanovic D, Grubor N, Rajovic N, Vucevic M, Miltenovic S, Laban M, Mostic T, Tasic R, Matejic B, Milic N
Assessing COVID-19 Mortality in Serbia's Capital: Model-Based Analysis of Excess Deaths
JMIR Public Health Surveill 2025;11:e56877
URL: https://publichealth.jmir.org/2025/1/e56877
doi:10.2196/56877

XSL•FO
RenderX

© Dane Cvijanovic, Nikola Grubor, Nina Rajovic, Mira Vucevic, Svetlana Miltenovic, Marija Laban, Tatjana Mostic, Radica Tasic, Bojana Matejic, Natasa Milic. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Survey–Based Tracing of COVID-19 Over the Early Pandemic: Comprehensive Geospatial and Symptomatic Analysis in Lebanon

Youssef Bassim[1*], MD; Abir Abdelrahman[2*], PhD; Amal Iaaly[3*], PhD; Gavin M Douglas[4], PhD; Patrick Daou[3], MSc; Mayda Finianos[2], MSc; Rim Hassan[2], MSc; Ibrahim Nassif[2], MSc; Layal Greige[2], MSc; Liza Dib[2], MSc; Jean Marc Mardirossian[2,5], MSc; Mira El Chaar[2], PhD

[1]Faculty of Medicine, University of Balamand, Koura, Lebanon

[2]Faculty of Health Sciences, University of Balamand, Rond Point Saloumeh-Dekouaneh, P.O. Box: 55251 Sin El Fil, Beirut, Lebanon

[3]Faculty of Engineering, University of Balamand, Koura, Lebanon

[4]Department of Biology, University of New Brunswick, Fredericton, NB, Canada

[5]Department of Nephrology and Hypertension, Mayo Clinic, Jacksonville, FL, United States

[*]these authors contributed equally

Corresponding Author:
Mira El Chaar, PhD
Faculty of Health Sciences, University of Balamand, Rond Point Saloumeh-Dekouaneh, P.O. Box: 55251 Sin El Fil, Beirut, Lebanon

Abstract

Background: In response to the early spread of COVID-19 in Lebanon, the University of Balamand developed the HAYATI app, a community-focused, geographic information system (GIS)–based digital health platform aimed at enhancing public health surveillance. At the time, while the Lebanese Ministry of Public Health utilized centralized dashboards to report confirmed cases and monitor national trends, no interactive tool existed to engage the public directly in real-time risk assessment and surveillance, especially in underserved regions.

Objective: The aim of this study was to design, implement, and evaluate the effectiveness of the HAYATI app as a GIS-integrated digital surveillance tool to identify high-risk individuals and support targeted testing and contact tracing during the early stages of the COVID-19 pandemic in Lebanon.

Methods: The HAYATI app was launched in March 2020 using ArcGIS Survey123 and real-time dashboards, incorporating a risk scoring algorithm based on 21 clinical and behavioral criteria. Between April 2020 and March 2021, self-reported data were collected from 10,235 individuals across Lebanon. Participants identified as high or major risk through the automated scoring algorithm were referred for free polymerase chain reaction testing at the University of Balamand. Test results were securely communicated to local municipalities and the Ministry of Public Health. Data were analyzed for associations between symptoms and positivity rates, as well as geographic and demographic trends using spatial analysis tools.

Results: Of the 10,235 individuals who submitted data, 1782 were classified as high or major risk and referred for polymerase chain reaction testing. Among them, 394 (22.1%) tested positive for SARS-CoV-2. Loss of smell and taste was strongly associated with positive test results (P<.001). The highest positivity rates were observed among individuals aged 18 - 29 years and in the North Governorate. GIS mapping enabled real-time visualization of case clusters, which informed localized containment responses.

Conclusions: The HAYATI app effectively filled a critical surveillance gap during the early pandemic phase in Lebanon. By integrating GIS technology, automated risk stratification, and community-level engagement, it provided a scalable model for public health surveillance in resource-limited settings. This approach has potential for broader applications in managing future outbreaks and endemic diseases through decentralized, real-time digital health strategies.

(JMIR Public Health Surveill 2025;11:e80331) doi:10.2196/80331

KEYWORDS

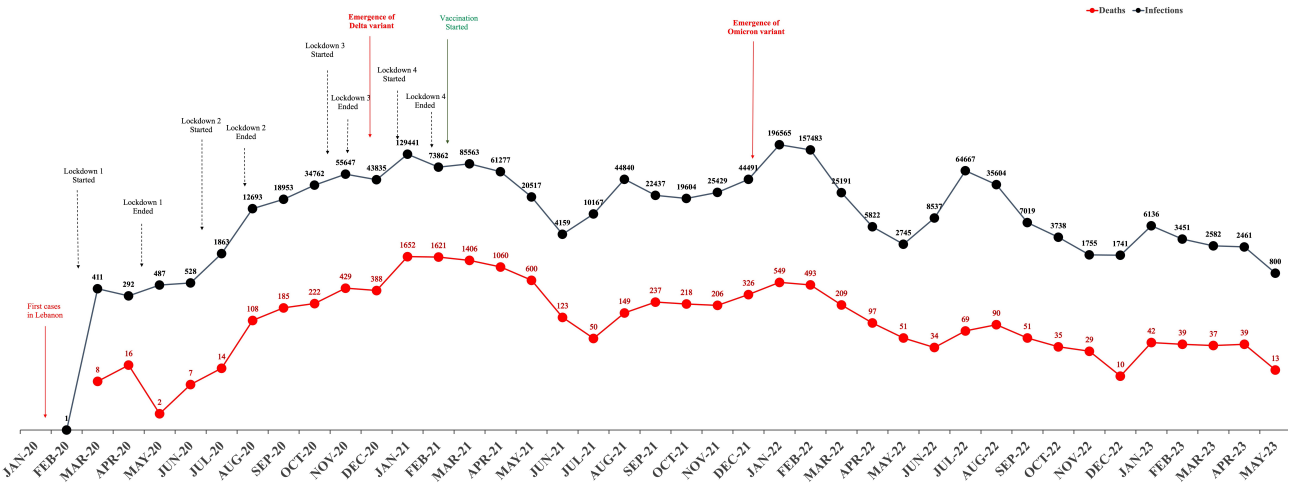
COVID-19; community engagement; digital health; GIS; Lebanon; risk stratification; surveillance; geographic information system

XSL•FO
RenderX

Introduction

The emergence of the novel coronavirus SARS-CoV-2 in Wuhan, China, in December 2019, set off a chain of events that would lead to a global health emergency unprecedented in recent history [1-3]. The first confirmed case of COVID-19 in Lebanon, reported on February 21, 2020, marked the beginning of the nation's encounter with the pandemic, leading to a series of public health responses aimed at curbing the virus's spread [4]. Lebanon's COVID-19 epidemiological trajectory was characterized by fluctuating infection and mortality rates, a pattern mirrored by public health responses that ranged from stringent lockdowns to comprehensive vaccination campaigns (Figure 1). The challenges faced by the government of Lebanon in response to the emerging pandemic were enormous due to very limited resources as the outbreak came at a time when Lebanon was going through the worst economic crisis in its history and during an unstable and conflicting political situation. The coronavirus outbreak also stretched the health sector that was already strained due to the lack of finance caused by the banking crisis that imposed severe restrictions on foreign currency fund transfers from the Lebanese Central Bank. This has caused a setback in Lebanon's effort in its fight against COVID-19 on top of all other economic disasters. Despite all the efforts taken by the Lebanese government to mobilize resources to equip public hospitals, the unmet needs were immense and the hospitals remained underequipped. At the start of the pandemic, COVID-19 testing was limited to a single designated facility at Rafic Hariri Public Hospital in Beirut, with a restricted supply of testing kits that could not meet the high market demand. Parallel to that, there were concerns that the coronavirus outbreak would hit vulnerable areas (poor and overpopulated) in Lebanon, such as the North Governorate, where people could not afford polymerase chain reaction (PCR) testing, and government intervention was limited in combating the spread of the pandemic. Given the limited testing capacity and resources, it became necessary to develop a targeted testing strategy to rationalize and prioritize the use of PCR kits. This included the smart identification of high-risk individuals, the establishment of a contact tracing mechanism, and the implementation of a robust surveillance system. In response to these challenges, the GIS Center at the Faculty of Engineering at the University of Balamand harnessed the power of geographic information systems (GISs) to enable rapid screening and tracing of COVID-19 cases, while also supporting the targeted use of PCR testing. By mid-March 2020, just 2 weeks after the first confirmed case in Lebanon, the GIS Center launched the HAYATI app, built using the Esri Geospatial Cloud. During this time, the strategic application of GIS emerged as an essential tool in navigating the complexities introduced by the pandemic [5]. GIS played a pivotal role in tracking and tracing the spread of COVID-19 by enabling spatial visualization, real-time data integration, and informed decision-making. Through interactive dashboards, GIS facilitated the monitoring of infection rates, mortality, and recovery data across global, national, and local scales. Health authorities used GIS tools to identify hotspots, track the movement of infected individuals, and model the spread of the virus, which was critical in guiding quarantine efforts and allocating medical resources. Additionally, GIS supported contact tracing initiatives by integrating spatial and temporal data, helping to map the networks of potential exposure. By linking demographic, mobility, and health data, GIS allowed for targeted interventions and effective communication with the public, underscoring its value as an essential tool in public health crisis management.

Figure 1. Monthly trends of infection and death cases with key pandemic events and interventions (January 2020-May 2023). The black line indicates active cases, while the red line denotes deaths, with each point on the lines corresponding to the number of cases reported monthly. Key events are marked, including the start and end of 4 lockdown periods, the emergence of the Delta and Omicron variants, and the beginning of vaccination efforts. Following May 2023, the data for both active cases and deaths are considered incomplete due to the discontinuation of regular reporting by centers, which no longer publish positivity rate statistics.



In this study, we highlight how the HAYATI app provided valuable insights into the interplay between clinical symptoms, demographics, and public health interventions, demonstrating the effectiveness of GIS in characterizing the pandemic. We also discuss the continued relevance of digital health tools like the HAYATI project in managing the endemic phase of COVID-19 and emphasize their critical role in addressing current and future public health challenges.

XSL•FO
RenderX

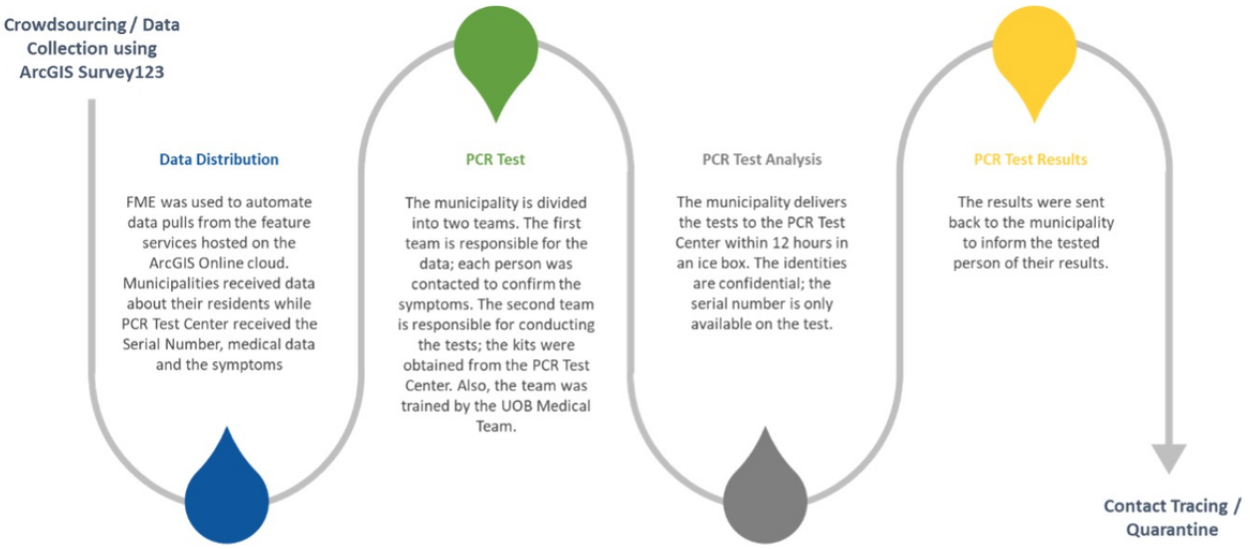
Methods

The HAYATI App

By mid-March 2020, just 2 weeks after the first COVID-19 case was reported in Lebanon, the GIS Center at the Faculty of Engineering at the University of Balamand launched the HAYATI app, a digital public health solution built using the Esri Geospatial Cloud. The app was developed as part of the University's social responsibility efforts to support national containment strategies, particularly in underserved areas of North Lebanon. It was made freely accessible to municipalities across the country to support the rapid identification and management of COVID-19 cases.

The HAYATI app integrated core Esri technologies, including ArcGIS Survey123 (version 3.9.149) for smart data collection and ArcGIS Dashboards for real-time data visualization and sharing. To streamline data processing, the Feature Manipulation Engine (version 2020.2.0.0; Safe Software Inc.) was used to extract and automate data transfers from ArcGIS Online feature services. The Esri ArcGIS Portal Feature Service was further embedded into the Feature Manipulation Engine pipeline to ensure timely dissemination of collected data to both municipal stakeholders and the University of Balamand's PCR Testing Center (Figure 2). This fully coordinated system created a closed-loop workflow involving municipalities, the PCR test center, and the University of Balamand medical team. Figure 2 illustrates the end-to-end process, from data crowdsourcing and distribution to PCR testing, laboratory analysis, result reporting, and follow-up actions such as quarantine and contact tracing. This workflow enabled efficient identification, risk-based triage, and timely intervention for individuals likely to have contracted COVID-19.

Figure 2. Operational workflow of the HAYATI app for COVID-19 risk assessment and targeted testing. This figure outlines the end-to-end workflow implemented through the HAYATI app, developed by the University of Balamand to support targeted polymerase chain reaction (PCR) testing and contact tracing during the COVID-19 pandemic. The process begins with data crowdsourcing using ArcGIS Survey123, where respondents provide information related to symptoms, demographics, and risk factors. Feature Manipulation Engine (FME) automates the extraction of these data from ArcGIS Online and distributes the data to municipalities and the PCR testing center. Municipalities divide their teams: one validates symptom data and contacts individuals, while the second handles logistics for PCR sample collection using kits provided by the University of Balamand (UOB) medical team. Samples are transported to the PCR testing center within 12 hours for analysis, with patient identities protected by serial number coding. Test results are reported back to municipalities, which then notify individuals and initiate contact tracing or quarantine measures based on the outcomes.



At the core of the app was a COVID-19 risk calculator developed by the University of Balamand, in line with recommendations from the World Health Organization, the Centers for Disease Control and Prevention, and the Lebanese Society of Infectious Diseases and Clinical Microbiology. The calculator incorporated 21 criteria grouped into 3 categories. Group 1 (critical exposures and symptoms) included 5 criteria: recent travel, contact with a traveler, contact with a suspected case, and sudden loss of smell or taste. The presence of any single Group 1 criterion automatically classified the respondent as high risk and mandated referral for PCR testing. Group 2 (key clinical symptoms) included 4 criteria: dry cough, sore throat, fever, and shortness of breath. The presence of any 2 Group 2 criteria also triggered PCR referral. Group 3 (additional epidemiological and clinical risk factors) included 12 criteria: public transportation use, exposure to crowded areas, hospital visits, rhinorrhea, productive cough, common cold, diarrhea, myalgia, headache, fatigue, abdominal pain, and cyanosis. Each of these criteria was assigned a weight ranging from 1 to 3 points. When neither Group 1 nor Group 2 conditions were met, the cumulative weight of Group 3 criteria determined the overall risk category.

Based on the total score, individuals were stratified into 4 categories: minor (<10% of total weight; score 0 - 2), moderate (10% - 30%; score 2 - 6), major (30% - 60%; score 6 - 12), and high (>60%; score 12 - 20) (Table 1). Only those categorized *as* major *or* high *risk* were referred for free PCR testing at the University's testing center. This targeted testing strategy was crucial during the country's economic crisis, as it enabled rational allocation of limited testing kits and resources. The full list of variables and their assigned weights is provided in Table S1 in Multimedia Appendix 1, while Table 1

XSL•FO
RenderX

summarizes the scoring thresholds and corresponding testing recommendations.

Table . COVID-19 risk score calculation and polymerase chain reaction (PCR) testing strategy.

Risk level	Minimum risk score	Maximum risk score	PCR test
Minor	0	2	No PCR test
Moderate	2	6	No PCR test
Major	6	12	PCR recommended
High	12	20	PCR needed

To complement the risk stratification system, an operational dashboard was developed using ArcGIS to display risk scores and testing outcomes in real time (Figure 3, Figure S1 in Multimedia Appendix 2). The dashboard provided geographic visualization of COVID-19 risk levels by plotting respondent data on an interactive map. Users could monitor daily changes in case distribution and view demographic breakdowns (by age, gender, and district) of high-and major-risk respondents. It also displayed epidemiological indicators such as recent travel history and hallmark COVID-19 symptoms (eg, fever, cough, loss of smell), alongside aggregated PCR test results at the municipal level.

Figure 3. Visualization of COVID-19 risk distribution via the HAYATI dashboard. This figure presents a comprehensive screenshot of the HAYATI dashboard developed by the GIS Center at the University of Balamand, showcasing the geographic and demographic distribution of COVID-19 risk across Lebanon. The central map displays the locations of respondents, with color-coded dots representing different risk levels: red for high risk, orange for major risk, yellow for moderate risk, and green for minor risk. At the top of the dashboard, the current number of high, major, moderate, and minor risk cases is displayed, with a comparison to the previous day's data. On the left sidebar, the distribution of cases by governorate highlights the regions most affected. To the right of the map, bar charts illustrate the breakdown of high and major risk cases by gender, age group, and caza (district). Additional charts below include data on respondents' travel history and the presence of loss of smell, a key symptom of COVID-19. At the bottom of the dashboard, summary statistics show the total number of reported respondents (n=21,175), high-risk cases (n=12,650), and major-risk cases (n=908).

Beyond its real-time monitoring capabilities, the HAYATI app was uniquely designed to integrate risk prediction, testing triage, and surveillance within a single platform. Unlike Go.Data, a digital tool developed by the WHO that focuses primarily on case investigation and contact follow-up, the HAYATI app combined algorithm-based risk scoring with geospatial visualization and diagnostic referral. This integration enabled immediate classification and response, ensuring that high-risk individuals received timely diagnostic testing and public health follow-up.

Participant Enrollment and App Accessibility

The app was primarily developed to support surveillance efforts in North Lebanon; however, it was made publicly accessible to all individuals across Lebanon, without geographic restriction. Any person residing in Lebanon could access the app, complete the survey, and be referred for testing, regardless of their governorate of residence. Our survey app was deployed between April 2020 and March 2021, during the first and second waves of the COVID-19 pandemic in Lebanon.

XSL•FO
RenderX

Nucleic Acid Extraction and SARS-CoV-2 Testing

Sample collection began with a nasopharyngeal swab from each participant. The swab was immediately placed into viral transport medium and transported to the Microbiology and Molecular Biology Laboratory at the University of Balamand, a facility certified for COVID-19 testing and operating in coordination with the Lebanese Ministry of Public Health. Upon arrival, SARS-CoV-2 RNA extraction was performed using the AccuPrep Viral RNA Extraction Kit. Extracted RNA was then subjected to real-time reverse transcription polymerase chain reaction testing using the GeneFinder COVID-19 Plus RealAmp Kit, which targets the nucleocapsid, envelope, and RNA-dependent RNA polymerase genes of the virus. RNase P was used as an internal control to ensure sample integrity and extraction efficiency.

Data Processing

Although no age restrictions were applied during participant recruitment, we performed several quality-control filters on the survey data to ensure data consistency and reliability for analysis. First, we removed 13 entries that shared the same participant serial number, retaining only 1 representative response per duplicate set. We then excluded 60 participants under 5 years old and 3 participants over 90 years old. These extreme age values were removed due to concerns about data accuracy and the ability of very young participants to self-report symptoms reliably. After filtering, a total of 10,954 participant entries were retained for downstream analyses.

Statistical Analysis

Statistical analyses were performed using the R programming language (v4.2.2; R Core Team) [6]. Enriched demographic factors and symptoms in SARS-CoV-2 positive versus negative individuals were identified based on Fisher exact tests, implemented via the exact2x2 R package (v1.6.9) [7]. In cases where frequency tables larger than 2×2 were analyzed, Fisher exact tests were performed using Monte Carlo simulations with the fisher.test function from the stats package. All symptoms were initially tested individually, and statistical significance was determined using Benjamini-Hochberg-adjusted *P* values, with a threshold of <.05. For the multisymptom analysis, we focused on symptom combinations that were observed in at least 10 participants to ensure sufficient statistical power. This more lenient threshold was adopted due to decreased statistical power and was considered justified as the analysis aimed to recapitulate well-known symptomatic signatures based on digital survey data rather than to robustly identify novel associations.

We conducted a PERMANOVA (permutational multivariate analysis of variance) using the adonis2 function from the vegan R package (v2.6.4) [8], based on pairwise Jaccard distances computed from the symptom presence or absence matrix. Only individuals reporting at least 2 symptoms were included in this analysis.

All data visualizations were also generated in R. Heatmaps were produced using the ComplexHeatmap package (v2.14.0) [9], and UpSet plots were generated using ComplexUpset (v1.3.3) [10]. Shapefiles representing Lebanese administrative boundaries were downloaded from the Humanitarian Data Exchange [11], maintained by the Office for the Coordination of Humanitarian Affairs. All additional figures were created using the ggplot2 package (v3.4.0) [9]. The complete analysis code is publicly available on GitHub [12].

Ethical Considerations

A pivotal aspect of the Hayati project's integrity and commitment to ethical standards was its approval by the institutional review board at the University of Balamand (IRB-REC/o/023-07/1123). This approval highlights the project's strict adherence to ethical research principles, ensuring that all data collection and participant engagement were conducted responsibly and with full respect for individual privacy and welfare.

Participation in the Hayati study was open to the public and entirely voluntary. Individuals who suspected exposure to COVID-19, experienced COVID-19-like symptoms, or were simply concerned about their health status were encouraged to participate. There were no specific inclusion or exclusion criteria, and no age restrictions were imposed. Any person willing to undergo screening and complete the survey was eligible. In cases where the participant was a minor, the application was filled out by a parent or legal guardian, and children were accompanied by a guardian during testing and enrollment. Recruitment occurred through community outreach in coordination with municipalities, health care centers, and public awareness campaigns.

All participants provided informed consent through the application, which clearly described the purpose of the study, procedures involved, potential risks and benefits, and strict confidentiality measures. Data collected via the app were securely stored in the ArcGIS online cloud platform, with full compliance with Lebanese data protection laws. Data access and processing were handled solely by the University of Balamand, with offline backup and secure archiving maintained at the GIS Center within the Faculty of Engineering.

Results

Demographic Distribution by Region, Gender, and Age

Our filtered dataset includes results from 10,954 individuals tested for COVID-19. Table 2 summarizes the demographic statistics of these individuals, categorized by governorate, gender, and age group. The largest proportion of individuals resided in the North governorate, ~70% (n=7635) of the population, followed by Mount Lebanon with ~25% (n=2781). The gender distribution was balanced, with males (n=5640, 51.5%) slightly outnumbering females (n=5116, 46.7%). Age information was available for the majority of participants; however, 42 individuals did not provide their age and were therefore excluded from age-specific analyses but retained in other parts of the dataset. Among those with available age data, the largest group was aged 18-29 years, indicating a youthful demographic, while those aged above 90 constituted the smallest group. The mean age of participants with known age was 38 years (SD=17) (Figure 4).

XSL•FO
RenderX

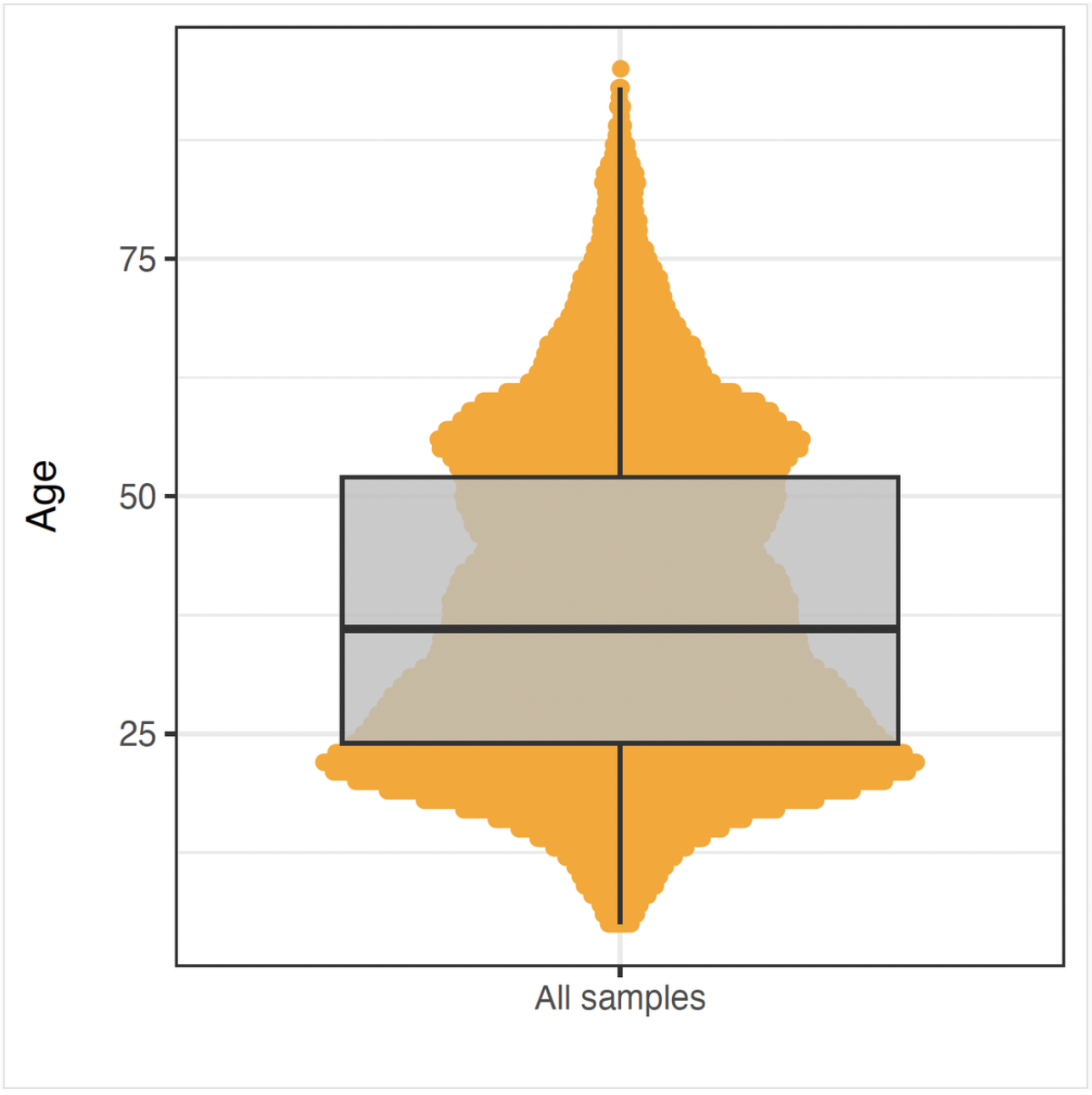
Table . Demographic and clinical characteristics of polymerase chain reaction (PCR)–tested individuals across governorates in Lebanon.

	Total, n (%)	Positive for SARS-CoV-2, n (%)
Governate		
Akkar	411 (3.8)	37 (9)
Baalbek-Hermel	2 (0.0)	0 (0)
Beirut	103 (0.9)	6 (5.8)
Beqaa	15 (0.1)	1 (6.7)
Mount Lebanon	2781 (25.4)	319 (11.5)
El Nabatieh	1 (0.0)	0 (0)
North	7635 (69.7)	907 (11.9)
South	6 (0.1)	0 (0)
Missing	0 (0)	0 (0)
Gender		
Female	5116 (46.7)	605 (11.8)
Male	5640 (51.5)	639 (11.3)
Missing	198 (1.8)	26 (13.1)
Blood group		
A+	3278 (29.9)	465 (14.2)
B+	784 (7.2)	96 (12.2)
AB+	299 (2.7)	40 (13.4)
O+	2908 (26.5)	360 (12.4)
A–	263 (2.4)	40 (15.2)
B–	72 (0.7)	11 (15.3)
AB–	44 (0.4)	10 (22.7)
O–	325 (3)	37 (11.4)
Missing	2981 (27.2)	211 (7)
Social distancing		
Yes	4370 (39.9)	427 (9.8)
No	6525 (59.6)	840 (12.9)
Missing	59 (0.5)	3 (5.1)
Public transportation		
Yes	1530 (14)	124 (8.1)
No	9420 (86)	1146 (12.2)
Missing	4 (0.0)	0 (0)
Crowded area		
Yes	3599 (32.9)	317 (8.8)
No	7351 (67.1)	953 (13.0)
Missing	4 (0.0)	0 (0)
Recent travel		
Yes	1141 (10.4)	25 (2.2)
No	9809 (89.5)	1245 (12.7)
Missing	4 (0.0)	0 (0)
Contact with travelers		
Yes	653 (6)	28 (4.3)

XSL•FO
RenderX

	Total, n (%)	Positive for SARS-CoV-2, n (%)
No	10,297 (94)	1242 (12.1)
Missing	4 (0.0)	0 (0)
Contact with COVID positive patients		
Yes	4133 (37.7)	595 (14.4)
No	6819 (62.3)	675 (9.9)
Missing	2 (0.0)	0 (0)
Hospital visit		
Yes	941 (8.6)	84 (8.9)
No	10,009 (91.4)	1186 (11.8)
Missing	4 (0.0)	0 (0)

Figure 4. Age distribution of all samples tested for COVID-19. The violin plot illustrates the density of different ages, showing both the range and the frequency of ages within the tested population. The boxplot within the violin plot provides a summary of the central tendency and dispersion of ages, with the median age indicated by the thick black line.



XSL•FO
RenderX

Associations Between Demographic and Behavioral Factors With SARS-CoV-2 Positivity Rates

Our dataset provides a detailed breakdown of SARS-CoV-2 positivity rates across various demographic and behavioral factors, including gender, age group, geographic region, quarantine practices, and blood group, as illustrated in Table 2. Overall, the COVID-19 positivity rate was approximately 11.5% (1270/10,954). Rates were comparable between females (605/5116, 11.8%) and males (639/5640, 11.3%), indicating no significant difference by gender.

While age groups showed relatively consistent positivity rates, slightly higher rates were observed among participants aged 30‐39 years. In contrast, more pronounced differences appeared across blood group types, with the lowest positivity rate among individuals with blood type O– (37/325, 11.4%) and the highest among those with AB– (10/44, 22.7%). These trends prompted further statistical analysis of the association between blood group and SARS-CoV-2 positivity, presented below.

To assess the relationship between blood group and COVID-19 test positivity, we performed Fisher exact tests, comparing the distribution of positive and negative cases for each blood group against all other groups combined. A statistically significant association was identified: individuals with blood type O+ were significantly less likely to test positive (OR 0.716, 95% CI 0.72-0.93; BH-corrected P=.01).

We also tested for associations between SARS-CoV-2 positivity and 7 self-reported behavioral factors, all of which showed statistically significant associations (BH-corrected P<.006, Fisher exact tests). Two behaviors followed expected patterns: individuals who reported adhering to social distancing had a lower positivity rate (427/4370, 9.8%) compared to those who did not (840/6525, 12.9%), and those who reported contact with confirmed COVID-19 cases had a higher positivity rate (595/4133, 14.4%) compared to those without known contact (675/6819, 10.9%).

In contrast, additional behaviors, which would be expected to be associated with higher positivity rates (were all possible confounders excluded), were significantly associated with lower positivity rates. In particular, the use of public transportation was associated with an 8.1% (124/1530) positivity rate, whereas those not using public transportation had a higher rate of 12.2% (1146/9420). Being in crowded areas was associated with a lower positivity rate of 8.8% (317/3599) compared to a 13.0% (953/7351) positivity rate for those who avoided crowded areas. Similarly, recent travel was negatively related to positivity rates, with recent travelers showing a rate of 2.2% (25/1141) against 12.7% (1245/9809) for those who had not traveled. In addition, individuals who had contact with travelers showed a lower COVID-19 positivity rate (28/653, 4.3%) compared to those who did not have such contact (1242/10,297, 12.1%). Separately, participants who reported visiting a hospital recently had a positivity rate of 8.9% (84/941), whereas those who had not visited a hospital had a higher rate of 11.8% (1186/10,009).

Age-Specific Analysis of COVID-19 Positivity Rates

In the analysis of COVID-19 positivity rates across different age groups, we observed notable variation in positivity percentages (Multimedia Appendix 3). The study included participants aged 5 years and above, following quality-control filtering that excluded individuals under the age of 5 years. Age groups ranged from 5 to over 90 years, with data analyzed accordingly.

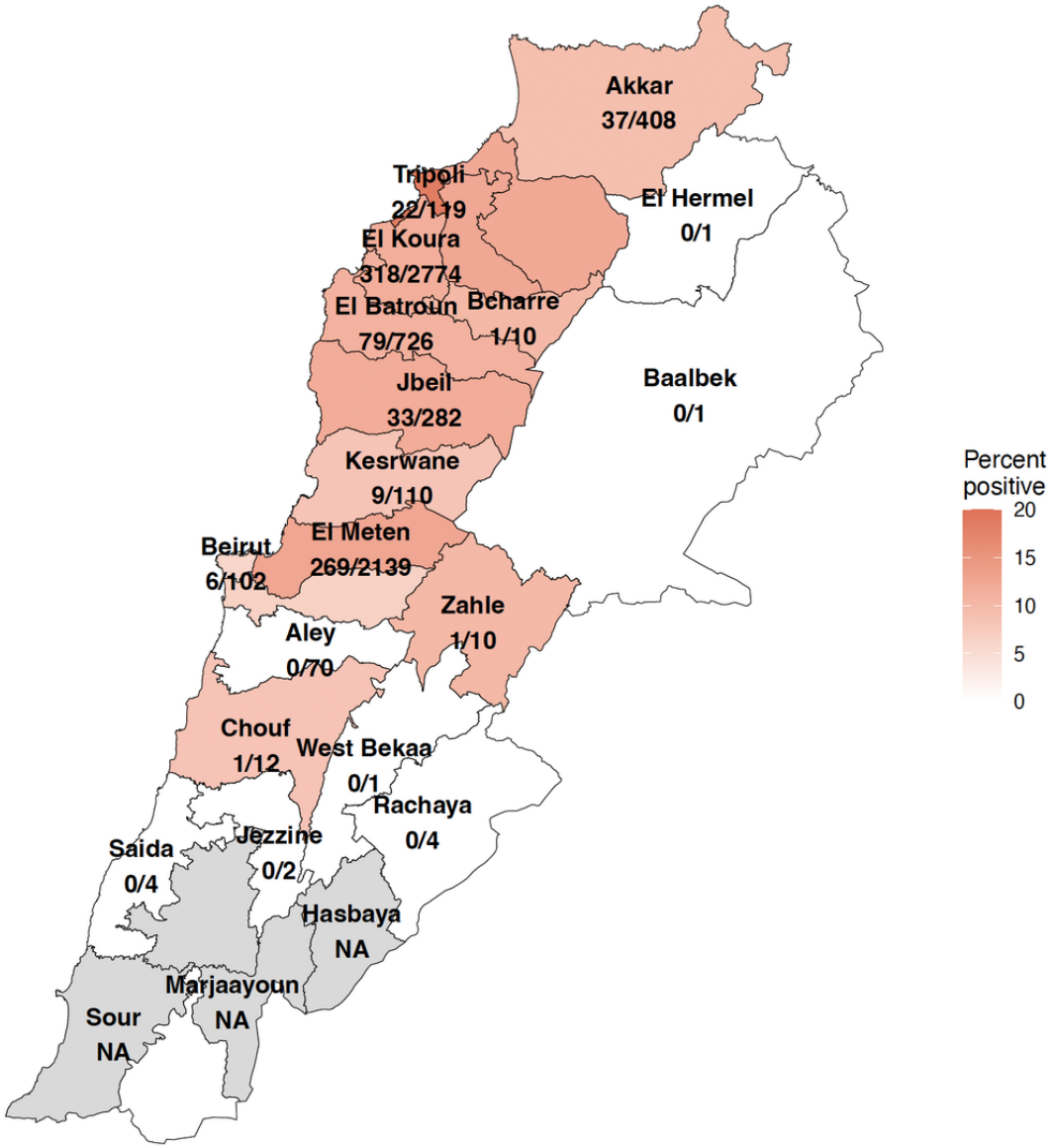
A total of 1270 PCR-positive cases were identified from the tested individuals. The age-specific positivity rates highlight variation in infection distribution across different age groups (Multimedia Appendix 3). Notably, higher positivity rates were observed among individuals aged 11‐17 years and those over 80 years old, compared to other age cohorts. For example, the 11‐17 age group exhibited a higher positivity rate than younger children. In contrast, younger adults and middle-aged groups showed relatively lower positivity rates. These patterns were recorded during the study period, which coincided with phases of school closures, lockdowns, and movement restrictions in Lebanon. While individual-level data were collected, household-level clustering (eg, siblings or cohabitants) was not tracked.

COVID-19 Positivity Rates by Caza

We analyzed COVID-19 positivity rates based on participants' reported caza of current residence, as shown in Figure 5. This variability likely reflects differences in viral spread and testing efficacy across regions. Notably, El Koura and El Meten reported the highest positivity rates among cazas with sufficient testing, at 11.4% (318/2774) and 12.5% (269/2139), respectively, suggesting they are focal points of infection. Tripoli, meanwhile, exhibited the highest positivity rate at 18.5% (22/119). Akkar and El Batroun also showed considerable positivity rates of 9.1% (37/408) and 10.9% (79/726), indicating moderate transmission levels.

XSL•FO
RenderX

Figure 5. Geographical distribution of COVID-19 positivity rates across different cazas over any time during the survey period. Shaded areas indicate the positivity percentage, with darker shades denoting higher rates. Numerical overlays show the fraction of positive cases compared to total tests administered regionally.

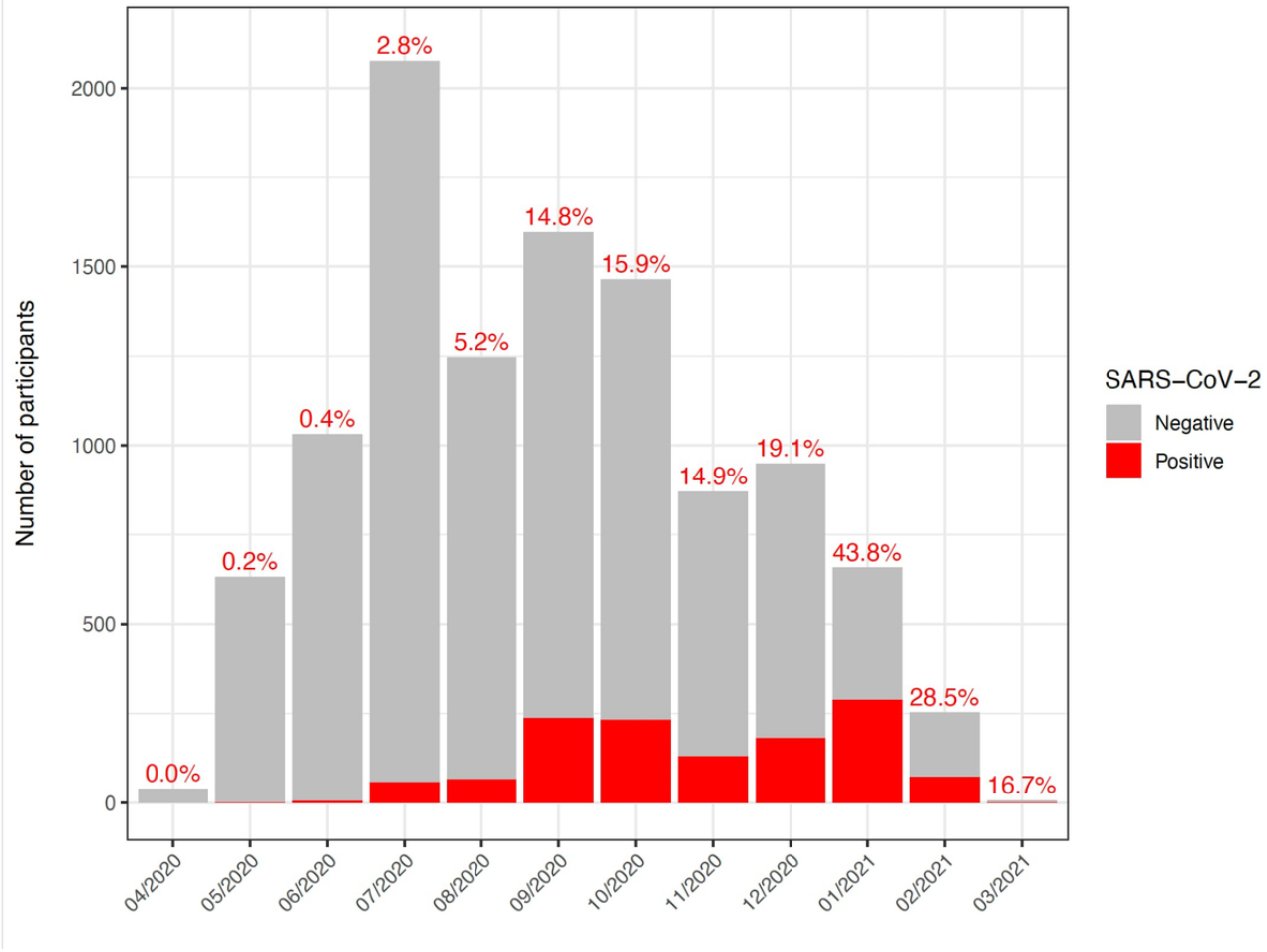


Trends in COVID-19 Positivity Rates: An Analysis From April 2020 to March 2021

Between April 2020 and March 2021, COVID-19 positivity rates, derived from PCR-confirmed cases following submissions made through the HAYATI app, varied significantly, as shown in Figure 6. The rates started at negligible levels, gradually increased during the spring and early summer months, and showed a marked surge in the fall, peaking in October. After a slight decline in November, a sharp spike was recorded in January 2021, coinciding with the emergence of the Delta variant and representing the highest positivity rate observed during the study period. This was followed by a notable decline through March 2021. A similar temporal trend was observed in the official COVID-19 case data published by the Ministry of Public Health (MoPH).

XSL•FO
RenderX

Figure 6. Monthly COVID-19 positivity trends. Stacked bars represent the total number of participants tested each month, divided into negative (gray) and positive (red) test results. The percentages atop each bar indicate the positivity rate for that month.

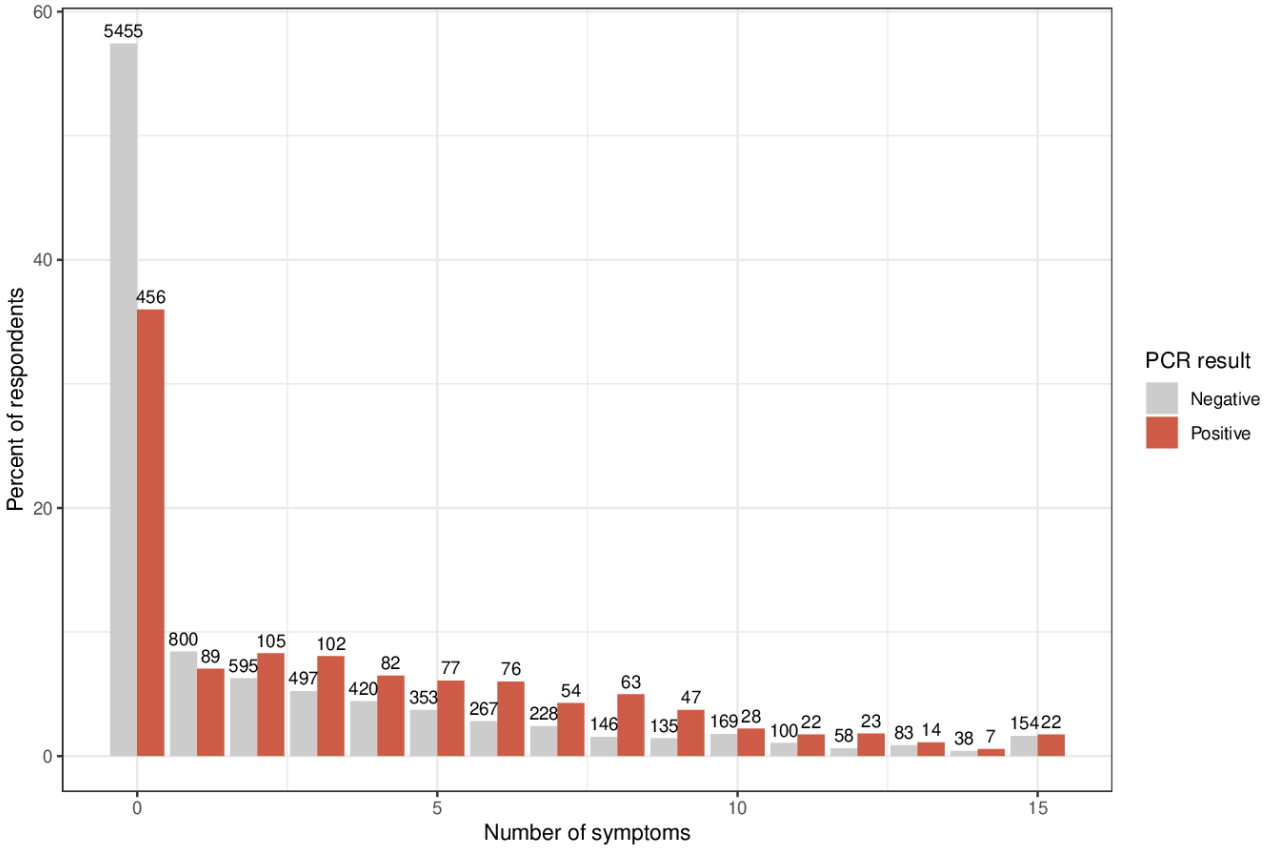


Symptom Distribution Among Individuals Screened for COVID-19

The analysis of symptoms was based on self-reported data collected through the HAYATI app and linked to PCR test results. Individuals presenting at the testing center exhibited a range of symptoms, leading to assessments for COVID-19. Notably, symptom presence was not a requirement for participation, allowing for the inclusion of asymptomatic individuals in the study. Among those who reported symptoms via the app, fever was recorded in 1273 individuals, with a corresponding positivity rate of 23%. Headaches and sore throats, reported in 3058 and 1844 cases, respectively, each had a positivity rate of 17%. Additionally, the loss of smell and taste was reported by 1065 and 914 individuals, respectively, correlated with a higher positivity rate of 25%. Figure 7. Our analysis demonstrates that loss of smell and taste was more strongly associated with positive COVID-19 test results than other common symptoms such as fever, myalgia, and cold. This correlation is illustrated in Figures 7 and 8, which visualize the distribution and frequency of self-reported symptoms within the tested cohort. Figure 8 reveals distinct symptom clusters that are more commonly observed in PCR-positive cases.

XSL•FO
RenderX

Figure 7. Symptom frequency in SARS-CoV-2 polymerase chain reaction (PCR) results. Bar plot of the number of symptoms reported by individuals with positive and negative SARS-CoV-2 PCR results.



XSL•FO
RenderX

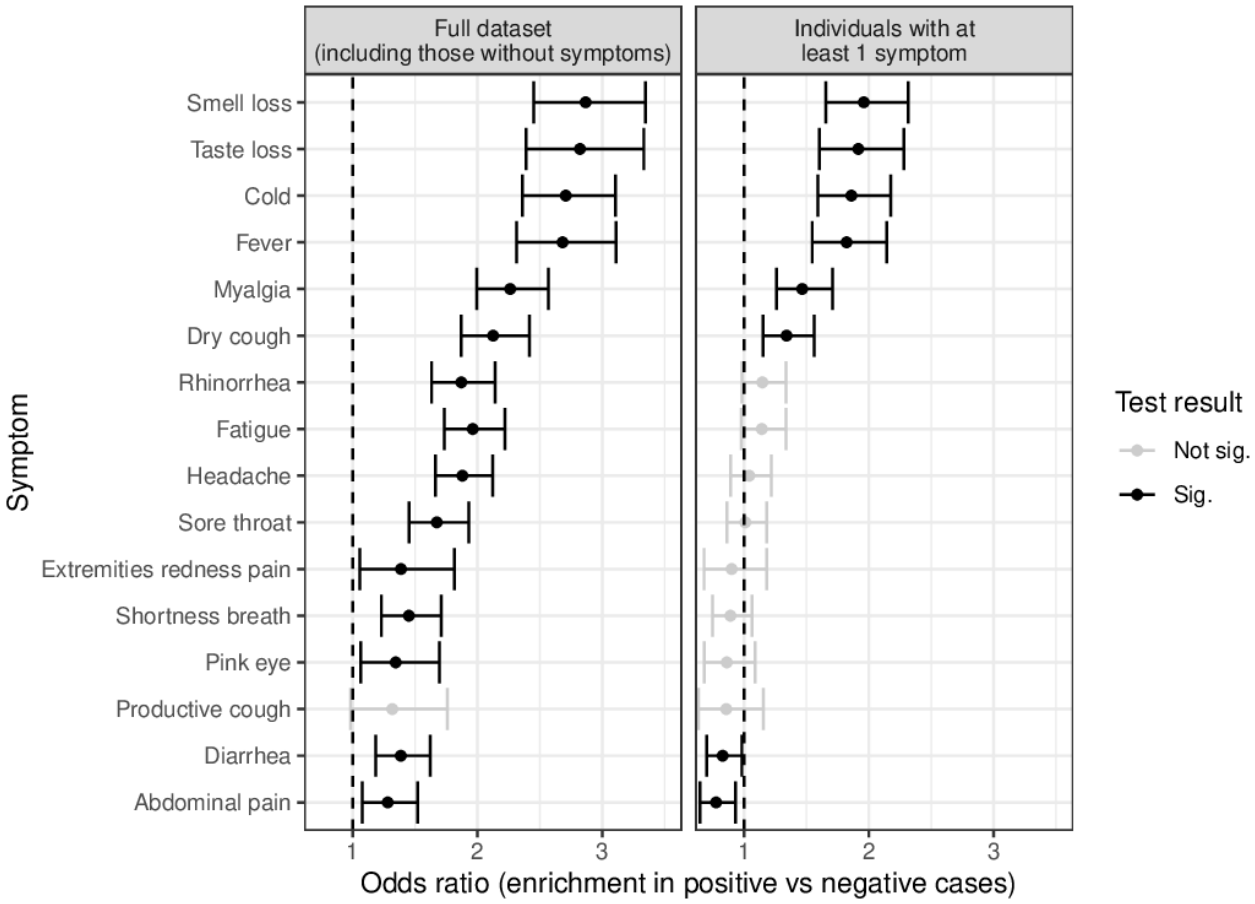
Figure 8. Overview of the co-occurrence of surveyed symptoms with SARS-CoV-2 PCR test results. Each row represents a different symptom, and the columns are individual participants, from any time during the survey period. Blue indicates that the participants reported the symptom, while grey indicates they did not report having the symptom. The overall percent prevalence ('Prev.') of each symptom is indicated in orange on the right side.

Further exploration of the HAYATI app data examined the relationship between the number of reported symptoms and test positivity (Figure 7). A trend of increasing positivity rates was observed as the number of symptoms increased, particularly from 3 symptoms onward.

To evaluate the consistency of survey-based symptom reporting with known clinical patterns, we assessed symptom enrichment among positive cases (Figure 9). Significant symptom sets, including fever, myalgia, and notably, smell and taste loss, were markedly enriched in PCR-positive individuals. Analysis of specific symptom combinations further revealed elevated odds ratios for patterns involving respiratory and systemic symptoms such as cold, dry cough, and myalgia (Figure 10).

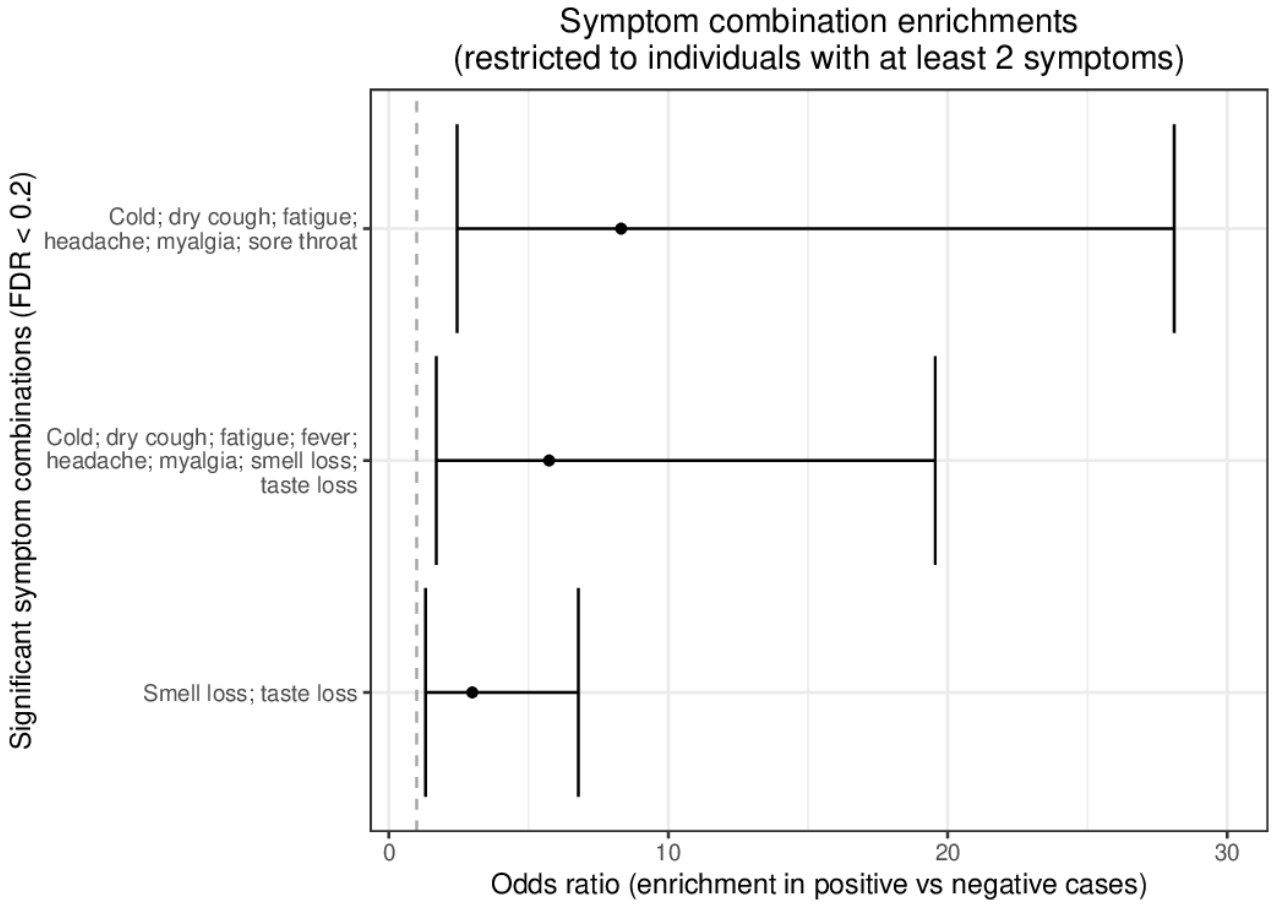
XSL•FO
RenderX

Figure 9. Intersection analysis of symptoms in positive and negative cases. Summary of reported symptoms enriched in participants with positive versus negative test results. Summary of reported symptoms enriched in participants with positive versus negative test results. The odds ratios based on a Fisher exact test are indicated for each individual symptom (bars indicate 95% CIs). Odds ratios >1 are enriched in positive cases. The left panel shows the odds ratio for the full dataset, including asymptomatic individuals, who are much more common among participants who tested negative. The right panel shows the same analysis restricted to participants with at least one symptom. Black points signify statistically significant differences (Benjamini-Hochberg corrected *P* values <.05) in symptom prevalence between positive and negative cases. The dotted line indicates an odds ratio of 1. Sig: significant.



XSL•FO
RenderX

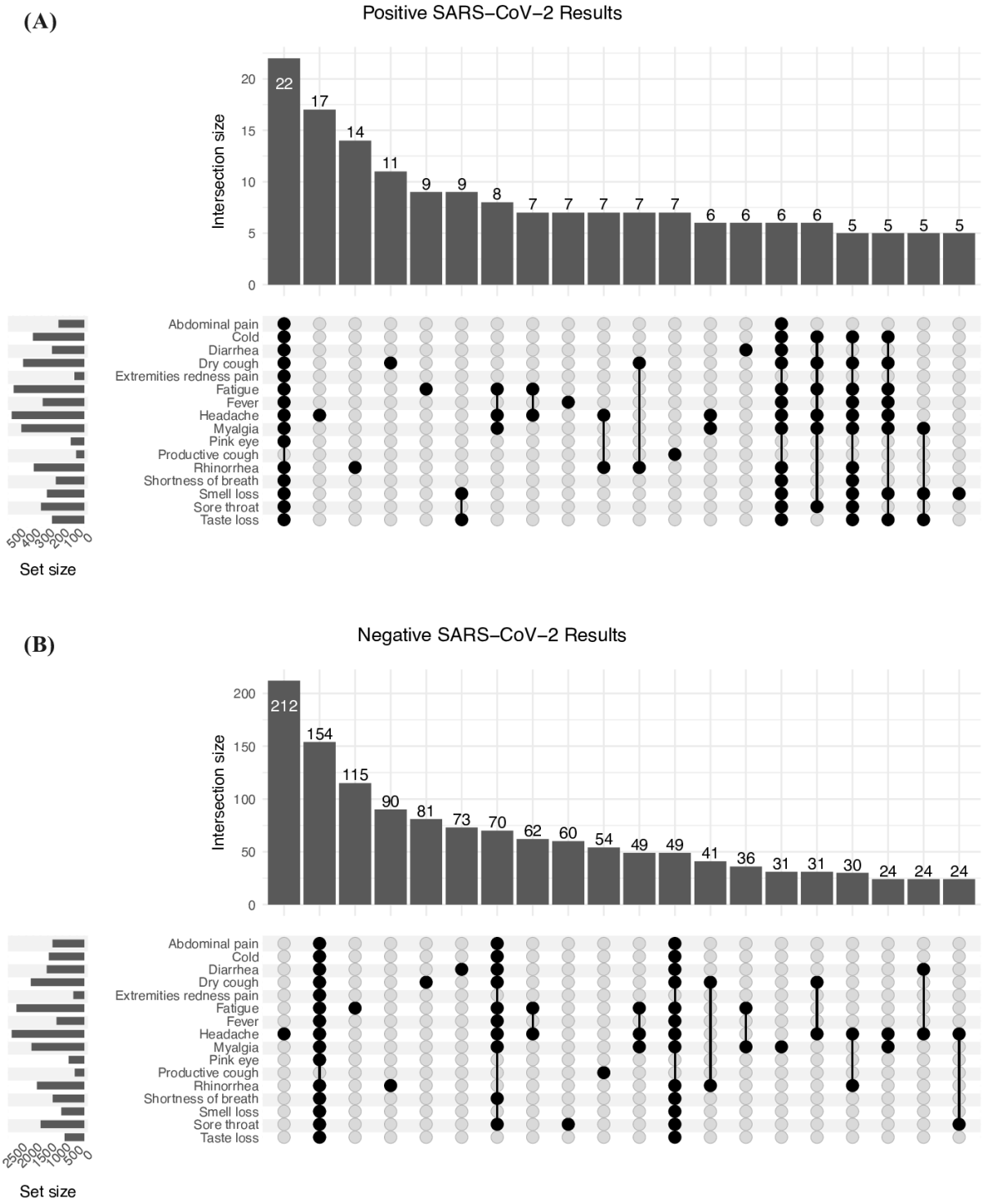
Figure 10. Enriched symptom combinations in COVID-19 positive cases. Significantly enriched symptom combinations in individuals with COVID-19 compared to individuals who tested negative (based on Fisher exact tests). Only individuals with at least 2 symptoms and symptom combinations that occurred at least 10 times were considered for this analysis. The plot illustrates the odds ratios for symptom combinations that are significantly more common in positive cases, with a Benjamini-Hochberg corrected *P* value <.2. The vertical bars represent the 95% confidence intervals for these ratios. The dotted line indicates an odds ratio of 1.



Additionally, the UpSet plots (Figure 11) derived from HAYATI app data illustrate the frequency and co-occurrence of symptoms among PCR-positive and PCR-negative individuals, highlighting the predominance of smell and taste loss among those who tested positive. These patterns were further investigated through a principal coordinates analysis (Figure 12), which graphically represents individual cases based on symptom similarity. Although a statistically significant difference was detected (PERMANOVA: R^2=0.0046, P=.001), the effect size was minimal, suggesting that while differences in symptom profiles between positive and negative cases exist, they explain only a small portion of the overall variation.

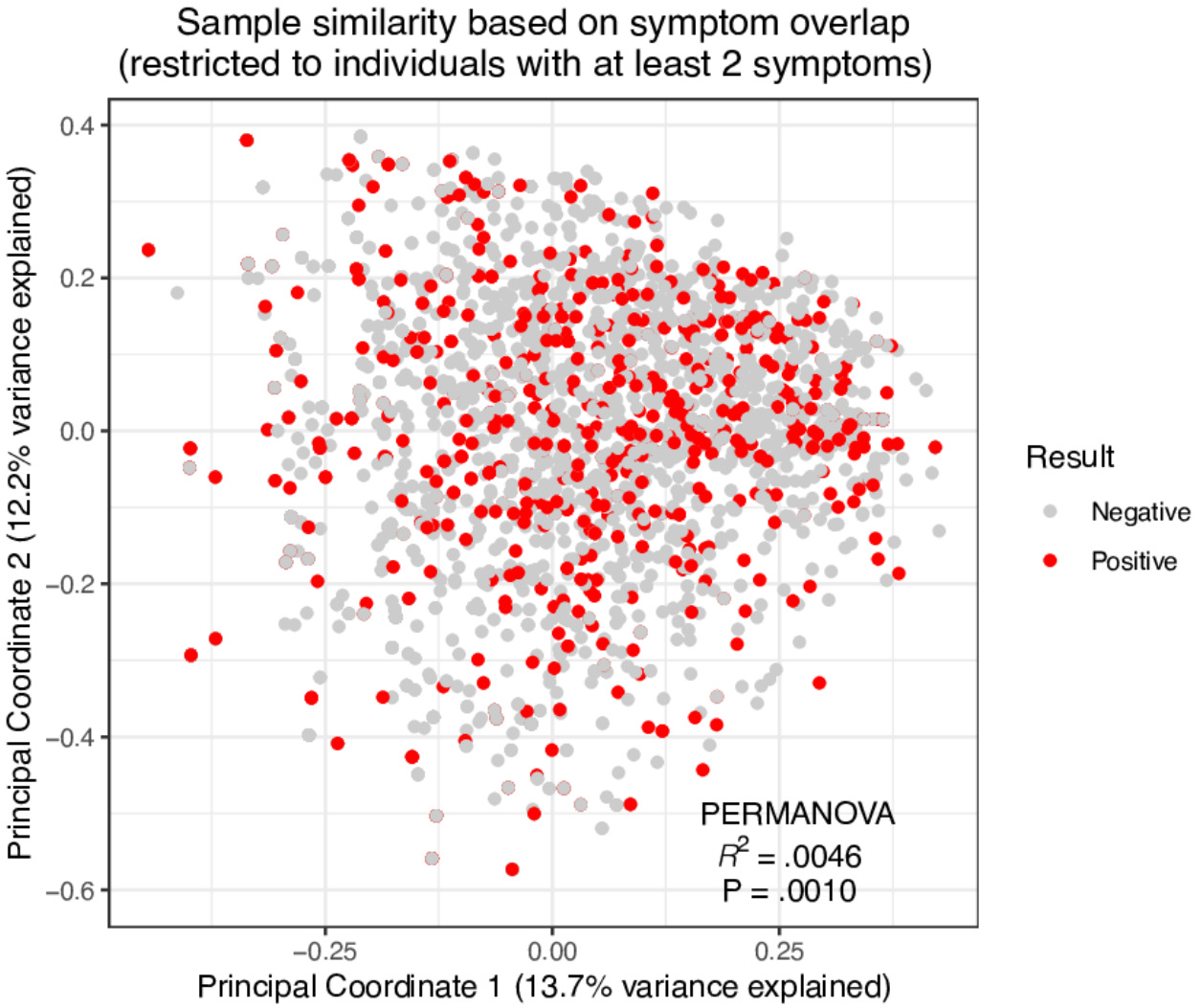
XSL•FO
RenderX

Figure 11. Intersection analysis of COVID-19 symptoms in positive and negative cases. The figure juxtaposes the symptom profiles of individuals with positive versus negative SARS-CoV-2 results using an UpSet plot format. Panel (A) focuses on the symptom intersections within the positive cases, while panel (B) does the same for the negative cases. The vertical bars represent the number of individuals reporting each intersection of symptoms, with the connected dots below delineating the specific symptoms within each intersection. This visualization elucidates the common symptom combinations in both cohorts (restricted to the top 20 most common symptoms in each case). The number of participants with each individual symptom is indicated by the horizontal bars on the left.



XSL•FO
RenderX

Figure 12. Symptom overlap and SARS-CoV-2 test results of principal coordinate analysis (PCoA). PCoA plot illustrating the similarity in symptom reporting between individuals with at least 2 symptoms, segmented by their SARS-CoV-2 polymerase chain reaction (PCR) test results. Each point represents a participant, with red indicating a positive test result and gray indicating a negative result. The proximity of points suggests the degree of similarity in symptoms reported. The effect size and *P* value of a PERMANOVA (permutational multivariate analysis of variance) statistical test of whether the PCR test result explains variation across samples is indicated.



Discussion

Transforming Infectious Disease Management: The Hayati Project's Role

The Hayati project was developed as a geospatial public health tool during a period of economic hardship, aiming to fill critical gaps in national surveillance by capturing data from individuals in underserved regions of North Lebanon. While national trends in COVID-19 cases were broadly reported by the Lebanese Ministry of Public Health, the HAYATI app provided several distinct advantages. First, the app enabled early, community-level surveillance by collecting self-reported symptoms and exposure data in near real-time, even before PCR confirmation. This type of participatory surveillance extended beyond the scope of MoPH systems, which primarily focused on confirmed case reporting. Hayati thus enabled proactive identification of potential hotspots before official case numbers increased. Second, and critically, the app was technically integrated with ArcGIS Pro, a professional-grade desktop GIS software. This allowed for automated spatial analysis, enabling researchers and public health officials to map the distribution of reported symptoms and confirmed cases, identify emerging clusters, and visualize epidemiologic trends in real time. This feature enhanced the utility of the data for rapid response and spatial decision-making, especially at the district (caza) level, where local municipalities could respond with targeted testing, outreach, or awareness efforts. Third, the initiative highlighted the role of a private academic institution in supporting national efforts through community-based, digital health tools. By offering free testing and building a regional digital infrastructure for data collection and spatial analysis, the University of Balamand demonstrated how academia can contribute actionable insights and public health resources during crises, particularly in resource-limited settings. Although some trends observed in the Hayati dataset aligned with national-level epidemiologic patterns, the value of the app lies in how the data were collected (community-driven), analyzed (spatially, via ArcGIS Pro), and acted upon (locally), providing operational support to public health authorities in a decentralized, real-time manner. A potential limitation of this study lies in the variability of

XSL•FO
RenderX

self-reported data obtained through the app, particularly during the early phase of implementation when user compliance and reporting accuracy could not be consistently verified. This limitation is common in digital surveillance tools launched during public health emergencies and should be considered when interpreting risk profiles and exposure data.

Alignment With Global COVID-19 Patterns

The patterns observed in the Hayati project are consistent with globally recognized characteristics of the COVID-19 pandemic, including variation in infection rates by age, symptom profile, and exposure-related behaviors. While these trends were broadly documented in national MoPH reports and international literature [13-17], the HAYATI app provided a community-driven, real-time dataset that mirrored these findings through self-reported inputs. For example, the strong association between loss of taste and smell and positive PCR results in our dataset reinforces the global understanding of these symptoms as key indicators of infection.

Although the Hayati project did not generate new epidemiological insights beyond those already known, it demonstrated that digital self-reporting platforms can effectively capture meaningful and reliable public health data, even in decentralized or resource-limited settings. The app's ability to collect, visualize, and analyze spatially referenced data at the community level, particularly in underserved regions, offers a complementary tool to traditional surveillance systems.

Moreover, the implementation of the Hayati project during Lebanon's economic and public health crisis serves as a case study for the feasibility and utility of digital health interventions in emergency contexts [18-20]. This model aligns with international efforts that highlight the value of early digital surveillance, public engagement, and spatial mapping to enhance pandemic preparedness and response [21-26].

Demographic and Behavioral Associations With COVID-19 Positivity Rates

Our study reveals important insights into the associations between demographic factors, behaviors, and SARS-CoV-2 positivity rates. We observed a slight variance in positivity rates between genders, with females showing marginally higher rates than males. Contrary to findings from previous studies, significant disparities in positivity rates were observed across different blood types. Specifically, individuals with the O+ blood type were significantly less likely to test positive for COVID-19 compared to individuals with other blood types [27,28].

Our study reveals that behavioral factors, particularly adherence to social distancing and contact with confirmed cases, demonstrated a strong correlation with infection rates. Specifically, individuals who adhered to social distancing had a lower positivity rate of 9.9% compared to 13.0% for those who did not. Additionally, contact with COVID-19 positive patients was associated with a markedly higher positivity rate of 14.6%, in contrast to 10.0% for those without such contact.

Interestingly, several behaviors that we would expect to be associated with higher positivity rates were instead linked to lower positivity rates. For example, the use of public transportation was associated with a lower positivity rate of 8.2%, compared to 12.3% for those not using public transportation. Similarly, being in crowded areas was associated with a lower positivity rate of 8.9%, compared to 13.1% for those who avoided crowded areas. Recent travel was also negatively related to positivity rates, with recent travelers showing a significantly lower positivity rate of 2.2%, compared to 12.9% for those who had not traveled. Moreover, the rate of positivity was higher in people who did not have contact with travelers (12.2%) versus those who did (4.3%). Hospital visits were associated with a positivity rate of 9.0%, whereas those who had not recently visited a hospital had a higher rate of 12.0%. These counterintuitive findings may be explained by the context in which the data were collected, specifically during intermittent lockdowns and movement restrictions in Lebanon between April 2020 and February 2021. During these periods, individuals who traveled, used public transportation, or accessed public spaces were often essential workers or lower-risk individuals who adhered strictly to preventive measures such as mask-wearing, hand hygiene, and distancing. Conversely, people remaining at home may have been at increased risk of household transmission, which was a dominant mode of infection during lockdowns. Additionally, the possibility of uncontrolled confounding, such as differing levels of risk awareness or health status, must be considered. For instance, healthier individuals may have been more likely to engage in public activities while taking effective precautions. Finally, we acknowledge that all behavioral data were self-reported, which may introduce information and recall biases that could affect the accuracy of the observed associations.

Our data collection was not based on a random sample of the population, which introduces another layer of potential bias. Younger individuals, possibly more concerned about contracting COVID-19, may have been more likely to participate. This bias in participant demographics could affect the generalizability of our findings.

By recognizing these limitations, we can better understand the complexities of COVID-19's transmission dynamics and further validate the critical role of public health guidelines.

The detailed analysis of how specific behaviors and demographic factors influence infection rates provides valuable insights for policymakers and public health officials in crafting targeted interventions to control the spread of the virus.

Age-Specific Analysis and the Impact on Different Demographics

The age-specific analysis of COVID-19 positivity rates in our study reveals notable variation across age groups, with the highest number of positive cases observed among individuals aged 18-29 years. While this may reflect greater social activity or increased likelihood of seeking testing within this age group, it is essential to interpret these findings with caution. The study was based on a self-selected, nonrandom sample, which introduces potential selection bias. Younger individuals may have been more inclined to participate due to greater digital access, heightened health awareness, or perceived risk, which could influence the observed age distribution.

XSL•FO
RenderX

Moreover, participation in the study may have been shaped by socioeconomic and behavioral factors, such as access to mobile technology, willingness to engage with digital platforms, or the ability to reach testing centers. These potential confounders limit the generalizability of our findings to the broader population.

Given these limitations, the results should be viewed as indicative rather than representative, offering insights into symptom trends and positivity rates within a specific, engaged segment of the population. Future studies using randomized or more representative sampling strategies are needed to more accurately characterize COVID-19 transmission dynamics across different age groups.

Conclusion

The study's results highlight the impact of the Hayati project in deepening our understanding of COVID-19's epidemiological characteristics within Lebanon. While the detailed analysis of demographic, behavioral, and clinical factors associated with SARS-CoV-2 positivity rates was conducted post-pandemic, it validates known global patterns and provides localized insights crucial for informed public health decision-making. This retrospective analysis underscores the potential benefits of real-time data analysis. Had these data been analyzed as they were collected, it could have significantly enhanced the responsiveness of health interventions during the pandemic. The successful application of GIS technology and self-reported data in tracking pandemic trends demonstrates the robustness and reliability of digital health tools. These tools hold promise for augmenting traditional epidemiological methods and offer valuable lessons for managing future health crises, particularly by leveraging real-time analysis to enable more timely and effective responses.

Acknowledgments

We are very grateful to Theresa Maarawi, Hayat Zaatiti, Samer Bazzi, Hortensia Klaymeh, Christina Sayhoun, Lea Nemr, and Jean Marc Mardorissian for their contribution to COVID-19 testing and literature research. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

Conceptualization: AA, AI, MEC, YB
Formal analysis: AI, GMD, MEC
Data curation: AI, GMD, MEC, PD, YB
Investigation: AI, IN, LD, LG, MEC, MF, RH, TH, YB
Methodology: AA, AI, MEC, YB
Project administration: YB
Software: AI, PD
Supervision: YB
Visualization: AI, GMD, MEC, YB
Writing – original draft: AI, GMD, MEC
Writing – review & editing: AA, AI, GMD, JM, IN, LD, LG, MEC, MF, PD, RH, TH, YB

Conflicts of Interest

None declared.

Multimedia Appendix 1
COVID-19 questionnaire variables and assigned weights in the HAYATI app.
[DOCX File, 18 KB - publichealth_v11i1e80331_app1.docx]

Multimedia Appendix 2
HAYATI COVID-19 survey questionnaire by the University of Balamand GIS Center. This figure shows a part of the HAYATI COVID-19 survey questionnaire interface, developed by the GIS Center at the University of Balamand, as it appears while being filled out.
[PNG File, 356 KB - publichealth_v11i1e80331_app2.png]

Multimedia Appendix 3

XSL•FO
RenderX

COVID-19 positivity rates by age group. The bar graph quantifies the percentage of positive cases within each age category (over all surveyed time periods). The absolute count of positive cases is indicated above each bar.
[PNG File, 142 KB - publichealth_v11i1e80331_app3.png]

References

1. Hui DS, I Azhar E, Madani TA, et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health—the latest 2019 novel coronavirus outbreak in Wuhan, China. Int J Infect Dis 2020 Feb;91:264-266. [doi: 10.1016/j.ijid.2020.01.009] [Medline: 31953166]
2. Wang W, Tang J, Wei F. Updated understanding of the outbreak of 2019 novel coronavirus (2019-nCoV) in Wuhan, China. J Med Virol 2020 Apr;92(4):441-447. [doi: 10.1002/jmv.25689] [Medline: 31994742]
3. Zhu H, Wei L, Niu P. The novel coronavirus outbreak in Wuhan, China. Glob Health Res Policy 2020;5(1):6. [doi: 10.1186/s41256-020-00135-6] [Medline: 32226823]
4. Epidemiological surveillance program of COVID-19. Republic of Lebanon Ministry of Public Health. 2020. URL: https://www.moph.gov.lb/en/Pages/2/24870/novel-coronavirus-2019 [accessed 2025-11-01]
5. Charlotte D, Smith JM. Incorporating geographic information science and technology in response to the COVID-19 pandemic. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/pcd/issues/2020/20_0246.htm [accessed 2025-11-01]
6. R: a language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.R-project.org [accessed 2025-11-01]
7. Fay MP, Lumbard K. Confidence intervals for difference in proportions for matched pairs compatible with exact McNemar's or sign tests. Stat Med 2021 Feb 28;40(5):1147-1159. [doi: 10.1002/sim.8829] [Medline: 33263202]
8. Oksanen JSG, Blanchet F, Kindt R, Legendre P, et al. Vegan: community ecology package. The Comprehensive R Archive Network. 2022. URL: https://CRAN.R-project.org/package=vegan [accessed 2025-11-01]
9. Wickham H, editor. Ggplot2: Elegant Graphics for Data Analysis: Springer; 2016. URL: https://ggplot2-book.org/ [accessed 2025-11-02]
10. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: visualization of intersecting sets. IEEE Trans Vis Comput Graph 2014 Dec;20(12):1983-1992. [doi: 10.1109/TVCG.2014.2346248] [Medline: 26356912]
11. Lebanon humanitarian data. Humanitarian Data Exchange. URL: https://data.humdata.org/group/lbn [accessed 2025-11-02]
12. Douglas GM. Lebanon_COVID. GitHub. 2024. URL: https://github.com/gavinmdouglas/Lebanon_COVID [accessed 2025-11-01]
13. Vaira LA, Salzano G, Deiana G, De Riu G. Anosmia and Ageusia: common findings in COVID-19 patients. Laryngoscope 2020 Jul;130(7). [doi: 10.1002/lary.28692] [Medline: 32237238]
14. Moein ST, Hashemian SM, Mansourafshar B, Khorram-Tousi A, Tabarsi P, Doty RL. Smell dysfunction: a biomarker for COVID-19. Int Forum Allergy Rhinol 2020 Aug;10(8):944-950. [doi: 10.1002/alr.22587] [Medline: 32301284]
15. Lechien JR, Chiesa-Estomba CM, De Siati DR, et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. Eur Arch Otorhinolaryngol 2020 Aug;277(8):2251-2261. [doi: 10.1007/s00405-020-05965-1] [Medline: 32253535]
16. Spinato G, Fabbris C, Polesel J, et al. Alterations in smell or taste in mildly symptomatic outpatients with SARS-CoV-2 infection. JAMA 2020 May 26;323(20):2089-2090. [doi: 10.1001/jama.2020.6771] [Medline: 32320008]
17. Parma V, Ohla K, Veldhuizen MG, et al. More than smell—COVID-19 is associated with severe impairment of smell, taste, and chemesthesis. Chem Senses 2020 Oct 9;45(7):609-622. [doi: 10.1093/chemse/bjaa041] [Medline: 32564071]
18. Peeri NC, Shrestha N, Rahman MS, et al. The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? Int J Epidemiol 2020 Jun 1;49(3):717-726. [doi: 10.1093/ije/dyaa033] [Medline: 32086938]
19. Ting DSW, Carin L, Dzau V, Wong TY. Digital technology and COVID-19. Nat Med 2020 Apr;26(4):459-461. [doi: 10.1038/s41591-020-0824-5] [Medline: 32284618]
20. Lalmuanawma S, Hussain J, Chhakchhuak L. Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: a review. Chaos Solitons Fractals 2020 Oct;139:110059. [doi: 10.1016/j.chaos.2020.110059] [Medline: 32834612]
21. Min-Allah N, Alahmed BA, Albreek EM, et al. A survey of COVID-19 contact-tracing apps. Comput Biol Med 2021 Oct;137:104787. [doi: 10.1016/j.compbiomed.2021.104787] [Medline: 34482197]
22. Schmeelk S, Davis A, Li Q, et al. Monitoring symptoms of COVID-19: review of mobile apps. JMIR Mhealth Uhealth 2022 Jun 1;10(6):e36065. [doi: 10.2196/36065] [Medline: 35609313]
23. Holl F, Flemisch F, Swoboda W, Schobel J. Apps for Covid-19 in Germany: assessment using the German Mobile App Rating Scale. JAMIA Open 2022 Dec;5(4):ooac082. [doi: 10.1093/jamiaopen/ooac082] [Medline: 36225895]
24. Park J, Han J, Kim Y, Rho MJ. Development, acceptance, and concerns surrounding app-based services to overcome the COVID-19 outbreak in South Korea: web-based survey study. JMIR Med Inform 2021 Jul 30;9(7):e29315. [doi: 10.2196/29315] [Medline: 34137726]

XSL•FO
RenderX

25. Wymant C, Ferretti L, Tsallis D, et al. The epidemiological impact of the NHS COVID-19 app. Nature New Biol 2021 Jun;594(7863):408-412. [doi: 10.1038/s41586-021-03606-z] [Medline: 33979832]
26. Samuel G, Roberts SL, Fiske A, et al. COVID-19 contact tracing apps: UK public perceptions. Crit Public Health 2022 Jan 1;32(1):31-43. [doi: 10.1080/09581596.2021.1909707] [Medline: 35221546]
27. Gopal SV, Sanker V, Pandian S, et al. Association between blood groups and COVID-19 CT severity: a retrospective analysis from a tertiary care center. Cureus 2023 Oct;15(10):e46506. [doi: 10.7759/cureus.46506] [Medline: 37927694]
28. Rubin R. Why blood type seems to be linked with COVID-19 risk. JAMA 2023 Sep 5;330(9):795. [doi: 10.1001/jama.2023.15996]

Abbreviations

GIS: geographic information system
MoPH: Ministry of Public Health
PCR: polymerase chain reaction
PERMANOVA: permutational multivariate analysis of variance

Edited by H Taha; submitted 09.07.25; peer-reviewed by C Heumann, O Lawal; revised version received 23.09.25; accepted 26.09.25; published 20.11.25.

Please cite as:
Bassim Y, Abdelrahman A, Iaaly A, Douglas GM, Daou P, Finianos M, Hassan R, Nassif I, Greige L, Dib L, Mardirossian JM, El Chaar M
Digital Survey–Based Tracing of COVID-19 Over the Early Pandemic: Comprehensive Geospatial and Symptomatic Analysis in Lebanon
JMIR Public Health Surveill 2025;11:e80331
URL: https://publichealth.jmir.org/2025/1/e80331
doi:10.2196/80331

© Youssef Bassim, Abir Abdelrahman, Amal Iaaly, Gavin M Douglas, Patrick Daou, Mayda Finianos, Rim Hassan, Ibrahim Nassif, Layal Greige, Liza Dib, Jean Marc Mardirossian, Mira El Chaar. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Development of a Conceptual Framework of Health Misinformation During the COVID-19 Pandemic: Systematic Review of Reviews

Javier Alvarez-Galvez[1], PhD; Jesus Carretero-Bravo[1], PhD; Carolina Lagares-Franco[2], PhD; Begoña Ramos-Fiol[1], MSc; Esther Ortega-Martin[1], MSc

[1]Computational Social Science DataLab (CS2 DataLab), University Research Institute for Sustainable Social Development (INDESS), Universidad de Cádiz, Jerez de la Frontera, Andalusia, Spain

[2]Department of Statistics and Operative Research, Universidad de Cádiz, Cádiz, Andalusia, Spain

Corresponding Author:
Javier Alvarez-Galvez, PhD
Computational Social Science DataLab (CS2 DataLab)
University Research Institute for Sustainable Social Development (INDESS)
Universidad de Cádiz
Jerez de la Frontera, Andalusia
Spain
Phone: 34 956167217
Email: javier.alvarezgalvez@uca.es

Abstract

Background: Despite the wide variety of studies that have focused on the recent COVID-19 infodemic, defining health mis- or disinformation remains a challenge due to the dynamic nature of the social media ecosystem and, in particular, the different terminologies from different fields of knowledge.

Objective: In this work, we aim to develop a conceptual framework of health misinformation during pandemic contexts that will enable the establishment of an interoperable definition of this concept and consequently a better management of these problems in the future.

Methods: We conducted a systematic review of reviews to develop a conceptual framework for health misinformation during the pandemic context as a case study.

Results: This review comprises 51 reviews from which we developed a conceptual framework that integrates 6 key domains—sources, drivers, content, dissemination channels, target audiences, and health-related effects of mis- or disinformation—offering a structured approach to analyze and categorize health misinformation. These 6 domains collectively form the basis of our proposed conceptual framework.

Conclusions: Our results highlight the complexity and multifaceted nature of health disinformation and underscore the need for a common language across disciplines addressing this global problem in order to use interoperable definitions and advance this evolving field of study. By offering a structured conceptual framework, we also provide a valuable foundation for interventions aimed at surveillance, public communication, and digital content moderation in future health emergencies.

(JMIR Public Health Surveill 2025;11:e62693) doi:10.2196/62693

KEYWORDS

misinformation; disinformation; infodemic; social media; social networks; COVID-19; review; SARS-CoV-2; coronavirus; respiratory; infectious; scoping review; conceptual framework; management; internet; technology; technologies; disease surveillance; public health; digital health; health informatics

Introduction

The rapid expansion of the internet and, more recently, the widespread adoption of social media platforms have led to an unprecedented acceleration in the production and dissemination of information. While these digital ecosystems have facilitated access to knowledge, they have also fostered the proliferation of misinformation. This phenomenon has increasingly drawn the attention of researchers, particularly within the social and behavioral sciences [1,2]. Over time, interest in the effects of misinformation on human behavior and social interaction has extended beyond these disciplines to fields such as health

XSL•FO
RenderX

sciences, computer science, physics, and mathematics, where scholars seek to understand the cascading effects of false or misleading information [2-5]. The COVID-19 pandemic further amplified this issue, as the global health crisis created a fertile ground for the spread of misinformation. Uncertainty, urgency, and the high demand for scientific evidence contributed to the rapid circulation of misleading claims, ranging from misconceptions about disease prevention to false treatments and conspiracy theories [6,7]. Moreover, both political figures and scientists, whether intentionally or unintentionally, played a role in amplifying misinformation, sometimes by disseminating preliminary findings without sufficient scientific scrutiny [8-10]. The increasing presence of automated agents, such as social bots and artificial intelligence (AI) tools based on deep learning and large language models, has further complicated the landscape by enabling the mass production and dissemination of misleading content [11,12].

Studies on health misinformation have emerged as a specific subfield that focuses on social and health-related topics such as vaccines, epidemics and pandemics, drugs and new tobacco products, noncommunicable diseases, eating disorders and diets, and new health treatments [2,13]. However, although the most frequent topics of debate are relatively well identified, in scientific practice, we still encounter many difficulties in delimiting the concept of health misinformation, which is ultimately affecting the advancement of research in this realm of knowledge [14]. This underscores the need for a comprehensive conceptual framework capable of integrating the multifaceted elements of misinformation—its origins, pathways, and impacts—to provide both clarity and actionable insight. In fact, we can find all kinds of definitions depending on the field and focus of the study, the analytical approach, or the type of materials analyzed (eg, internet texts, tweets, posts, and online videos, among others) [15]. On the one hand, the highly dynamic nature of social media ecosystems makes it difficult to create a stable definition that characterizes all the agents involved, the topics of conversation, the channels of misinformation, and the respective health effects [2,16]. On the other hand, there is a differential use of terminology from different fields of knowledge that makes it difficult to configure a unitary language that allows us to identify the current evidence in this field of study [17].

While the concept of health misinformation had been around since the early days of the first social media platforms [18], the COVID-19 pandemic would take this term to a new dimension never known before. During this period, health misinformation has had a clear impact on the management of the health crisis in the different countries and population groups [19,20], but it has also significantly affected the scientific field and, particularly, the way scientific evidence is generated. Indeed, the need to generate rapid evidence that would take us out of the context of the health crisis has contributed to an increase in the rate of scientific production, in some cases reducing the quality of studies [21]. With the pandemic, studies based on opportunistic samples, rapid reviews, and studies led by scientists with no previous experience in the field of public health proliferated, which, despite having an altruistic purpose, may have indirectly contributed to the increase in misinformation on health issues fundamental to the proper management of the health crisis [22-24].

Consequently, although research on health misinformation and its impact on health systems has become more prevalent, clarifying the concept continues to be difficult given the constantly changing landscape of social media and the wide array of health-related issues it involves, especially within the COVID-19 context [2]. To navigate this complexity, we adopt an inclusive term of health misinformation in this study, considering any health-related claim that is based on anecdotal evidence, lacks scientific validation, or is misleading due to the absence of empirical support [11], regardless of whether it was disseminated intentionally or unintentionally. Accordingly, our definition distinguishes between 2 primary forms: misinformation, which consists of false information shared without the intent to cause harm [25], and disinformation or malinformation, which involves the deliberate creation or manipulation of false or partially true information with the purpose of misleading or harming specific individuals, social groups, institutions, or nations [26,27].

Although efforts have been made in the existing literature to construct both conceptual frameworks and taxonomies on misinformation, for example, linked to specific topics such as vaccines [28,29] or preventive measures to combat misinformation [30-32], to the best of our knowledge, there is still no conceptual model that allows for an adequate operationalization of the concept of health misinformation. To address this research gap, the present study aims to develop a comprehensive conceptual framework for health misinformation in the context of pandemics. This tool seeks to provide an interoperable definition that integrates the key dimensions of misinformation—sources, drivers, message content, dissemination channels, audiences, and health-related effects—while addressing the interdisciplinary nature of the problem. By offering a structured approach to understanding and categorizing misinformation, this study contributes to the advancement of research in this field and provides a foundation for the development of strategies to manage and mitigate misinformation in future health crises. In particular, we selected the case of COVID-19 for several reasons: (1) the wide range of agents that have been involved in the pandemic (health professionals, academics, politicians, influencers, etc); (2) the broad impact that health misinformation has had as global health phenomenon; and (3) the usefulness of a conceptual framework that can serve as a basis for the semantic interoperability of the term in interdisciplinary scientific fields and for the management of future health emergencies. Thus, the final purpose is to lay the foundations of a taxonomy of pandemic misinformation that can be extended and adjusted to other health topics, but specifically in the context of health crises.

To carry out the present objective, a systematic review of reviews was conducted (including systematic reviews, qualitative synthesis, scoping reviews, among others), from which the fundamental dimensions that would compose our conceptual framework of health misinformation would be extracted.

XSL•FO
RenderX

Methods

Study Design

The present systematic review of reviews was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [33].

Search Strategy

We searched in the following databases: PubMed, Scopus, Web of Science, Ovid, and Cochrane in January 2023, based on the peak of COVID-19 misinformation literature. The search was limited to papers published before January 10, 2023. The search strategy consisted of 2 parts: 1 focused on COVID-19 and the other on misinformation using Boolean terms. The following keywords to develop the search strategy were used: (1) for COVID-19: "covid 19", "covid-19", "sars-cov-2 infection", "2019 novel coronavirus disease", "2019 novel coronavirus infection", "2019-ncov disease", "2019 ncov disease", "2019-ncov diseases", "covid-19 virus infection," "covid 19 virus infection", "covid-19 virus infections", "coronavirus disease 2019", "coronavirus disease-19", "coronavirus disease 19", "severe acute respiratory syndrome coronavirus 2 infection", "sars coronavirus 2 infection", "covid-19 virus disease", "covid 19 virus disease", "covid-19 virus diseases", "disease, covid-19 virus", "2019-ncov infection", "2019 ncov infection", "2019-ncov infections", "covid19", "covid-19 pandemic", "covid 19 pandemic", "covid-19 pandemics"; (2) for Misinformation: "inaccurate information", "misleading information", "seeking information", "rumour", "rumor", "gossip", "hoax", "urban legend", "urban legends", "myth", "fallacy", "fallacies", "conspiracy theories", "conspiracy theory", "malinformation", "disinformation", "misinformation".

The search strategy was adapted to each database, and the key terms used in each database are summarized in Multimedia Appendix 1.

Eligibility Criteria

The inclusion criteria encompassed studies addressing misinformation related to health within the pandemic context, covering topics such as COVID-19, vaccination, virus-related treatments, policies, and humorous content such as jokes and memes. Reviews published in peer-reviewed journals and written in English, Spanish, and French were considered eligible. Accepted types of literature reviews included narrative, qualitative, systematic, meta-analysis, scoping, and umbrella reviews that provided insights into information sources, messages, channels, audiences, or health outcomes.

Exclusion criteria involved nonresearch papers in scientific journals (abstracts, commentaries, opinions, editorials, letters, and books) and nonscientific content reviews, including those from social media, blogs, and newspapers. Additionally, publications that did not satisfy the inclusion criteria were excluded.

Data Extraction

The papers found were exported to Microsoft Excel 365 (Microsoft Corp) for duplicate removal and screening. Titles and abstracts were first examined by 2 reviewers (EO-M and JC-B) independently. These reviewers selected studies based on the inclusion criteria if the title or abstract contained pandemic, health, or misinformation-related terms. Papers without an abstract went directly to full-text screening. After this screening, the results were blinded so that researchers could discuss contradictions until agreement was reached. Subsequently, the first 30% of the full-text studies were reviewed twice to establish a uniform method. In case of disagreement, this was resolved by discussion and consensus among the reviewers. In case of persistent disagreement between reviewers (EO-M and JC-B), an additional protocol involving an external reviewer was applied (JA-G). Once the method to be followed was clear and defined, 4 reviewers (BR-F, CL-F, EO-M, and JC-B) screened the full text according to the inclusion and exclusion criteria. The references of the included papers were also screened to find new studies that met the selection criteria.

Data Synthesis

Descriptive information and information on misinformation during the COVID-19 pandemic were extracted from the selected papers in 2 tables. The first table included the title, authors, topics covered, year of publication, objectives, and a summary of the authors' conclusions. For studies that did not include the conclusions in the abstract, a summary of the conclusions was made, taking into account the most relevant aspects. On the other hand, the second table collected information on misinformation during the COVID-19 pandemic: title, source, message, channel, audience, and result of the misinformation (Multimedia Appendix 2). To perform the analysis of the misinformation data in this second table, each of its elements was classified. For this classification, some categories have been combined because of their similarity. The mass media category, for example, covers newspapers, radio, and television. Within communication channels, certain social media platforms have also been grouped, Twitter (subsequently rebranded X; X Corp) includes Twitter, Gab (Gab AI, Inc), Weibo (Weibo Corporation), and Parler; Instagram (Meta) includes Snapchat (Snap Inc) and Pinterest; Facebook (Meta) combines Facebook and Mondo Sporco; while WhatsApp (Meta) combines WhatsApp, Telegram (Telegram Messenger Inc), and WeChat. Similarly, in terms of the consequences or impact of reduced trust, a group was formed to reduce trust in government, scientists, and medical professionals.

The development of the conceptual framework was guided by a thematic synthesis approach, involving iterative coding and integration of topics across included reviews (ie, papers mentioning misinformation sources, specific message content, drivers, channels, effects, etc). Through consensus discussions among the research team, recurring patterns were grouped into 6 dimensions that represent the underlying mechanisms and manifestations of health misinformation.

Quality Assessment of Included Reviews

To evaluate the methodological quality of the studies included in this review, we used the QUEST (Quality Evaluation of Scientific Texts) Checklist [34], a comprehensive appraisal tool specifically designed to assess the rigor and transparency of empirical research across diverse methodological designs,

XSL•FO
RenderX

including quantitative, qualitative, and mixed-methods studies. QUEST comprises 20 items grouped into 6 core domains: theoretical framework, research questions and objectives, methods, results, discussion, and additional information. Each item is scored as "yes," "no," or "not applicable," enabling a structured assessment of key quality indicators such as theoretical grounding, appropriateness of study design, clarity of analytical techniques, reporting of results, discussion of limitations, and ethical considerations. This instrument was chosen for its adaptability to interdisciplinary research designs.

Conceptual Framework

Based on the references of this review, we developed a conceptual framework. With our results, we identified the main dimensions and subdimensions addressing misinformation during the COVID-19 health crisis. All subdimensions were then ranked (assigning a corresponding score) according to the relevance within their dimension and their potential impact on the level of misinformation. We determined the relevance by considering the percentage of results within their respective dimensions. Additionally, the research team assessed the potential impact on the level of misinformation of the population, considering the benchmarks of our review. Each of the items in the conceptual framework included a score that would be agreed upon by the research team that intervened in the development of the systematic review.

The conceptual framework was developed by categorizing misinformation sources based on their profile (human expert vs nonexpert and declared bot vs nondeclared) to capture their distinct roles in the dissemination process. Within bot-generated content, a key distinction was made between declared and nondeclared bots, as previous studies have shown that declared bots maintained a neutral tone during the pandemic, whereas nondeclared bots were more likely to spread negative narratives, criticize public health measures, and promote misinformation [11,24]. Additionally, the audience dimension incorporates health literacy levels, recognizing the higher susceptibility of vulnerable groups to misinformation. This integration is supported by existing research, which has consistently demonstrated that lower levels of health literacy are strongly associated with greater exposure to and belief in health-related misinformation. By structuring the framework around these factors, this model provides a comprehensive approach to understanding the pathways through which misinformation spreads and impacts public health decision-making [35,36].

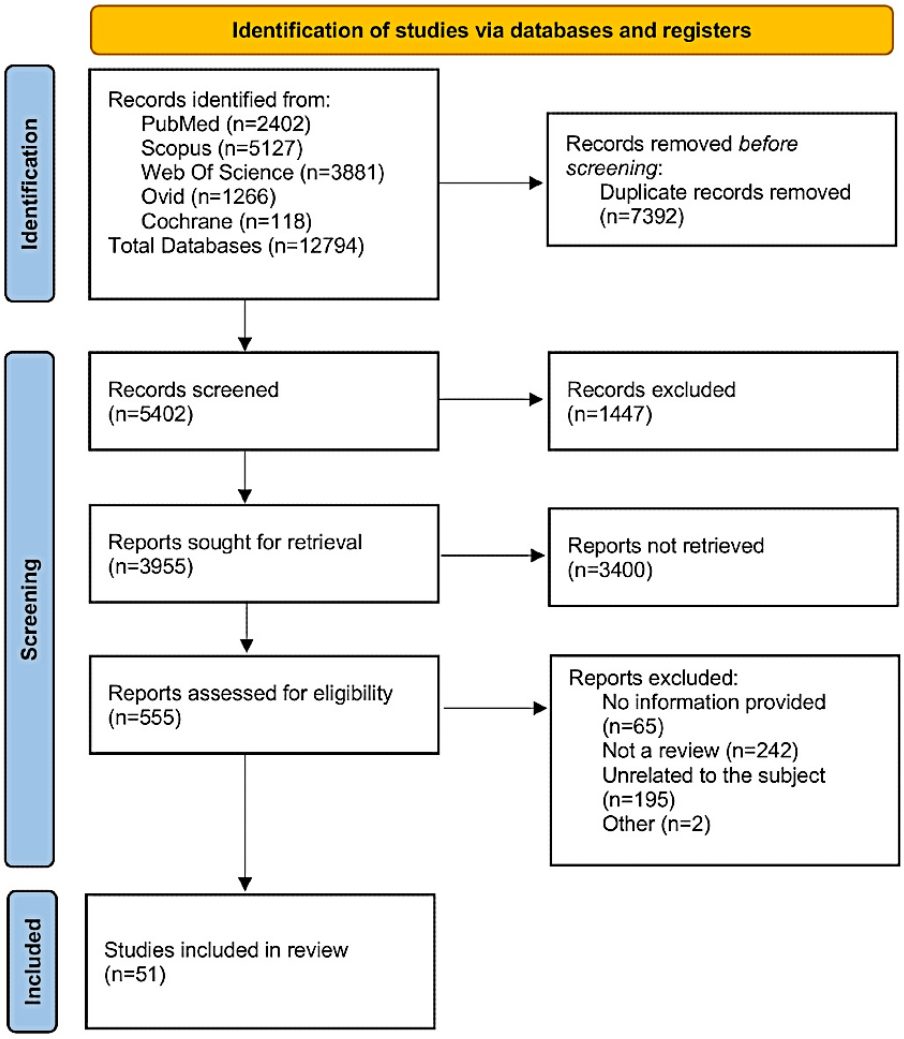
Results

Selection of Studies

We identified a total of 12,794 results, from which 5402 remained after removing duplicates (Figure 1). Then, we screened titles and abstracts, selecting 556 papers for a full review. Finally, after a full-text examination, 51 papers were included in our review. Applying the QUEST instrument to our sample of included reviews, we found that most studies demonstrated high methodological quality. Of the evaluated papers, the mean number of criteria satisfied ("yes") was 16.4 of 20, with only 2 "no" responses and 1.8 "not applicable" on average per study. Most studies provided adequate conceptual frameworks, clearly stated research questions and aims, and appropriately described their methods and analyses. However, some limitations were observed, particularly in the justification of sample sizes and the explicit reporting of data availability for replication. Detailed quality assessment results for each study are provided in Multimedia Appendix 3.

XSL•FO
RenderX

Figure 1. PRISMA flowchart. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

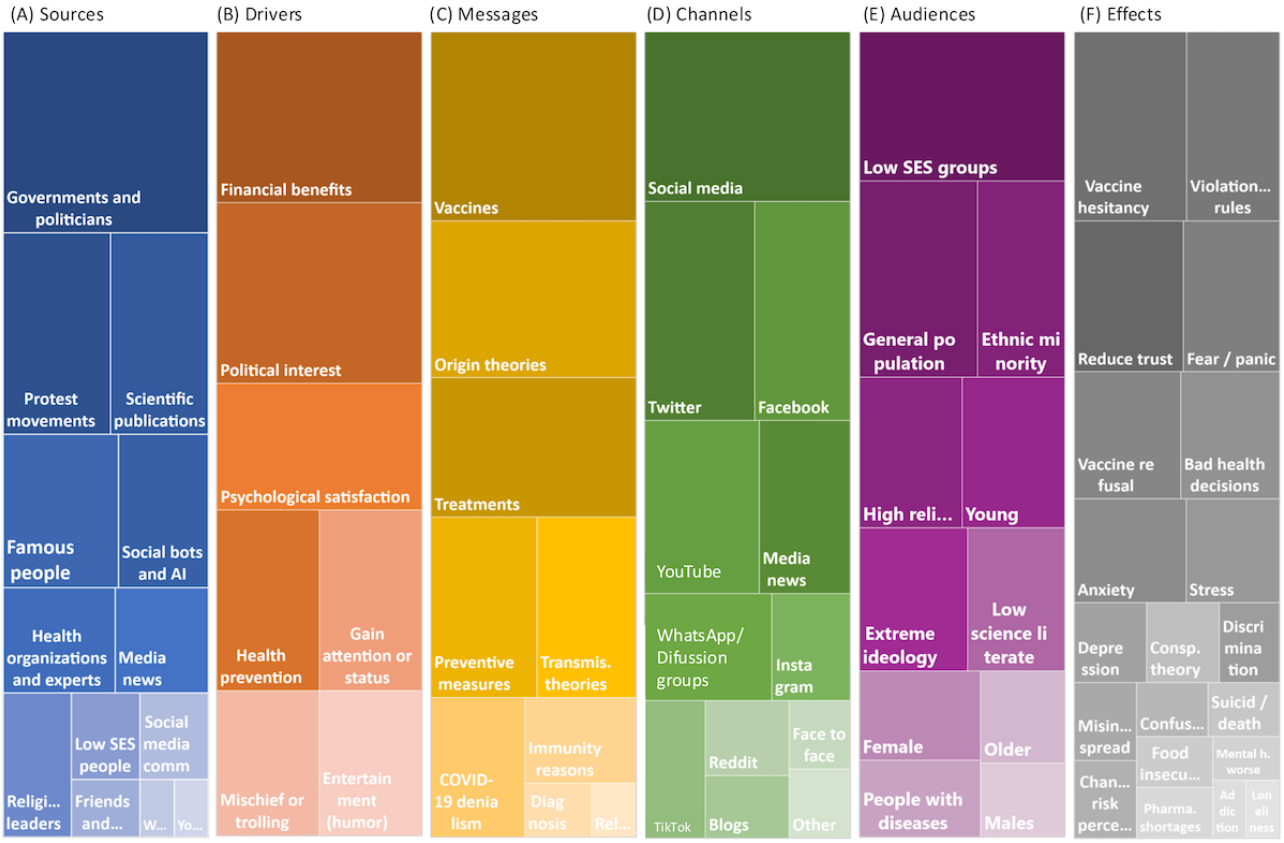


This instrument demonstrated high reliability, with reported intraclass correlation coefficients exceeding 0.98 and a Cronbach α of 0.99, indicating excellent interrater consistency and internal coherence.

Figure 2 provides a visual depiction of the different categories that make up the 6 domains (or dimensions) that we identified as those characterizing the concept of health misinformation during the pandemic context. This group of elements is composed of (1) the different sources of misinformation, (2) the drivers, (3) the specific messages, (4) the propagation channels, (5) the various audiences, and (6) the respective social and health effects. Around this set composed of 6 dimensions, we detected a wide variability of health misinformation traits that, in this specific context, are indistinctly mentioned and combined among the 51 papers selected.

XSL•FO
RenderX

Figure 2. Domains and components of health misinformation during the pandemic. AI: artificial intelligence; Pharma: pharmacy; SES: socioeconomic status.



Sources of Health Misinformation

One of the most critical aspects of health misinformation is identifying its origin. During the COVID-19 pandemic, the sources of health misinformation that were most frequently mentioned in the studies located in the literature were those related to governmental and political actors (26%) [38-59]. Among them, Donald Trump was mentioned in several papers as a figure in the spread of misinformation [44,45,49,55]. It was followed by individuals linked to protest movements (13%) [40,44-46,49,57,59-63], composed mainly of antivaxxers, and scientists (12%) [38,40,46,50,53,54,58,64-66] (including the Nobel Prize winner Luc Montagnier [67]). However, other sources were also mentioned, such as famous people (11%) [41,44,47,49,51,58,68-70], (declared and nondeclared) social bots or new AI tools (8%) [47,52,71-74], health organization experts (7%) [42,43,54,56,62,67], mass media opinion leaders (6%) [55,60,61,68,75], religious leaders (6%) [41,49-51,61], people of low socioeconomic status (4%) [41,49-51,61], and social media community members (4%) [47,55,76].

Drivers of Health Misinformation

The studies highlight several underlying motivations behind the dissemination of health misinformation. The drivers of health misinformation during the COVID-19 pandemic were fundamentally divided between the achievement of political interest (22%) [39,41,46,49,59,63,68,71,73,76], economic interest (22%) [39,40,46,54,55,58,75-78] (eg, saying that the CDC is exaggerating the pandemic to damage Donald Trump's reputation or selling fraudulent products, respectively [40,46,71]), and psychological satisfaction (16%) [39,47,52,58,76,79,80] or social status (11%) [41,49,68,79,80] (eg, this is the case of influencers or social media users that tried to gain attention in these new platforms [41]), which, to a certain extent, corresponded to the typology of topics mentioned as sources of misinformation in the previous section. However, misinformation was also observed as a driver of health prevention from both (scientific or health) experts, policymakers, and health organizations (11%) [39,54,61,76,81], that is, misleading, false, and erroneous information that could possibly be produced unintentionally [54,61,81]. In this sense, we found a clear division between misinformation for selfish purposes (ie, for personal or institutional gain in a context of crisis) and misinformation for altruistic purposes (ie, for helping others). Among the reasons for misinformation was also mentioned the need of troll users to damage the reputation, create chaos, directly annoy others (9%) [54,55,58,72], or entertain (9%) [54,76,79,80].

In short, while informative drivers showed a lower risk of misinformation, others, such as economic, political (which could have a greater impact on the opinions of the population), or those that were directly aimed at generating chaos and damaging the image of individuals, groups, or institutions, presented a greater risk.

Content of Health Misinformation Messages

The health misinformation circulating during the COVID-19 pandemic covered a wide range of health-related topics. In the contents of the messages, misinformation was observed in relation to vaccines (24%)

XSL•FO
RenderX

[40,41,43-47,49,51,52,54,57-63,69-72,74,76-78,80-87], where several papers related vaccines to the insertion of microchips with 5G [49,52,58,60,74,84] or the possible side effects. Furthermore, among the prevalent misinformation were various theories about the origin of COVID-19 (19%) [16,39,41,44,47,49,51,52,54,55,58-61,63,64,67,68,70,71,77-79,81,84,86], including unfounded claims suggesting that COVID-19 is a biological weapon [30,42,51,59,68,74] or is produced by the 5G network [63,68]. Additionally, we found health misinformation about certain treatments and home remedies (17%) [39-42, 44, 48, 50, 52, 54-56, 58, 61, 65, 71, 73, 75, 76, 78, 79, 81, 86-88] such as the use of herbal medicines [40,61,71,81] or even the use of cow urine [41,77]. Among these home remedies, there were also references to the use of sodium hypochlorite (ie, bleach) and alcohols to prevent COVID-19 infection [44,55,61,69,88] or even to eliminate its effects after being infected [61,77,86]. Other relevant topics were the possible preventive measures (11%) [39,41-43,49,50,54-56,58,61,69,74,76-78], theories about the transmission of the new disease (11%) [39,41-43,49,50,54-56,58,61,69,74,76-78], denialist messages about COVID-19 (8%) [44,48,49,51,52,56,59,69,73,78,87], theories about immunity (6%) [40,41,49-51,61,62,79], including the arguments about the supposed natural immunity [40,50,62] or ethnic immunity [41,49,51,61] of certain groups (ie, vegan immunity [41]), about diagnoses (2%) [41,64,76] and even with religious content (1%) [59,84], specifically about religious conspiracies, such as the false claim that the COVID-19 vaccine contained gelatin derived from pigs, leading to concerns among Muslim communities [84].

Among the “less harmful” messages, reference was made to humorous topics (satire or parody), which, although not generally intended to cause harm [54,76,79,80], had a certain potential to mislead, so that humor could also be used to intentionally spread rumors and conspiracies in a simple way. Other studies referred to false connections (when, as in the case of clickbait, headlines do not support the content), misleading content to frame information (selective selection of images, quotes, or statistics) or false contexts (when genuine content is shared with false contextual information) [39,46,58,59], practices that, as some studies point out, seem to undermine trust in the media and promote polarization. At a higher level of misinformation, studies were also located that mention of impostor content (the impersonation of genuine sources, for example, through phishing and smishing techniques), direct manipulation of content (deliberate modification of information or images with the aim of deceiving) or, in the most extreme cases, content fabrication, which would involve the creation of content that is totally false and that is created with the purpose of deceiving and causing harm (this, for example, would be the case of the so-called deepfakes) [39,58].

Health Misinformation Dissemination Channels

A key factor in understanding health misinformation is analyzing how it spreads. Among the main channels of health misinformation, social media were frequently mentioned, which, as a whole, were identified by 80% of the studies analyzed. Many studies referred to social media platforms as the main health misinformation dissemination channel (22%) [35,38-47,49,50,52,56-58,61-63,65-68,70,71,73,74,77-79,81,86,87,89]; however, this percentage varies depending on specific platforms such as Twitter (15%) [38, 45, 47, 48, 52, 54, 55, 57, 58, 60, 65, 71-77, 79-81, 86], Facebook (13%) [40, 44, 45, 47, 50, 54-58, 60, 65, 69, 71, 74-76, 80, 81, 86], YouTube (Google LLC; 12%) [39, 40, 45, 48, 55-58, 60, 66, 69, 71, 74, 76, 78, 80, 84, 86, 89], WhatsApp and Telegram (8%) [38, 40, 45, 49, 57, 58, 60, 61, 69, 74, 76, 80, 86], Instagram (5%) [45, 56, 57, 60, 68, 74, 76, 81], TikTok (5%) [41, 45, 56, 57, 69, 74-76] and Reddit (Reddit, Inc; 4%) [38, 47, 57, 58, 74, 76]. Reference was also made to mass media news (9%) [38, 40, 42, 46, 51, 54, 57, 58, 65, 68, 70, 71, 74, 80, 86], misinformation spread through blogs (3%) [45, 54, 57, 58, 60], direct contacts in face-to-face relationships (2%) [45, 50, 57, 84], and other sources (2%) [40, 46, 55, 65], among which scientific media or web-based platforms such as LinkedIn were also mentioned [65].

In the case of channels, it was observed that much of the misinformation fell within the framework of social networks and different online platforms, while it was less common among official sources such as health organizations and the media.

Target Audiences

Although exposure to health misinformation is transversal to different social groups, as also indicated by different studies (13% of the studies pointed to the general population [38,39,42-45,50,51,54,55,61,63,68,79,80]), a greater predisposition to misinformation was also identified among socially vulnerable population groups. Thus, for example, 20% of the studies pointed to groups with low socioeconomic status [19,39,41,43,45,46,49,51,52,55,57,59,61,62,70,73,75,76,81,82,84], ethnic minorities (10%) [39,41,50,52,61,62,73,79,81,82,84], high religiosity groups with low scientific literacy (9%) [40,43,46,49-52,62,75,84], young people because of their greater exposure to online media (9%) [39,46,49,52,55,61,75,76,80,84], extreme ideological groups (ie, commonly partisan individuals who are ideologically positioned and less critical of their own groups; 9%) [46,52,59,66,68,70,71,76,78,84], people with low scientific and health literacy (8%) [41,46,47,51,52,59,67,70,75], women (6%) [45,51,52,57,80,82,84], people with illness or disability (6%) [41,51,55,56,60,61], older adults (5%) [41,52,61,75,86], while others pointed to the greater exposure of men (4%) [52,75,76,84]. In this sense, despite the generalized susceptibility of the population to health misinformation, a higher risk was observed in groups with lower socioeconomic status (including vulnerable groups such as ethnic minorities) and religious positions (9%), as well as in those individuals who were more exposed to extreme political and lower scientific literacy (8%).

Among the various types of audiences, health literacy was observed as a fundamental articulating variable when processing and evaluating the various health contents. For this reason, to facilitate the classification in the conceptual framework, we would opt for this indicator divided into 4 categories, ranging from 0 health literacy (0 points) to high health literacy (4 points).

Social and Health Effects of Misinformation

The impact of health misinformation extends beyond simple misconceptions, influencing both social trust and individual

XSL•FO
RenderX

behaviors. The effects of health misinformation identified in the specialized literature were varied. Among the most relevant impacts were identified the increase in doubts about vaccination (13%) [40,46,47,49,51,52,57,59,62,66,70,73,76,78,80-87,89], as well as the violation of the norms established by political agents (11%) [38,41,47,50-54,63,66,68,73-76,78,80,84,85], the generalized reduction of trust in government, politicians, health institutions, and scientists (10%) [38,41,43-46,50,52,53,58,60,66,68,75,78,80,81], and the increase in vaccine refusal behavior (8%) [40,43-45,54,59,60,62,63,70,73,78,83,84]. However, as a whole, the significant effects that health misinformation may have had on the mental health of the population were also noted. The different effects included an increase in fear and panic among the population (9%) [38, 41, 42, 45, 46, 51, 52, 54, 56, 66, 71, 75, 81, 86, 89], anxiety (7%) [39,41,42,52,54,57,66,71,74,75,77,86], incorrect health decision-making [38,39,46,47,50,53,55,56,58,63,79,86,87] (8%, eg, an increase in advanced lung cancer cases among patients due to hesitancy to approach health care facilities is seen [87]), stress (6%) [42,46,52,55,56,66,71,78,81], depression (3%) [42,46,52,66,77,86], changes in risk perception (3%) [39,54,56,70,71], confusion (2%) [52,78,81,86], worsening of general mental health (2%) [66,78,86], increase in suicides (2%) [41,42,77,87], loneliness (1%) [66,81], and addictions (1%) [41,46], problems related to mental health which together accounted for 44% of the impacts identified. However, the impact of health misinformation on discrimination (3%) [41,42,46,54,71], propagation of erroneous or poor quality health information (3%) [39,55,66,68,75], conspiracy theories about the COVID-19 pandemic (3%) [47,52,58,60,68,80], impulsive buying of food and consumables (2%) [41,52,71,78], or shortage of medicines (2%) [50,55,66,78] were also mentioned.

Conceptual Framework

Figure 3 presents a conceptual framework on COVID-19 misinformation that comprehensively systematizes the 6 identified dimensions characterizing the various types of health misinformation that have emerged during the recent pandemic: (1) sources; (2) drivers; (3) types; (4) channels; (5) audiences; and (6) health misinformation effects on COVID-19. For the development of these dimensions, the results of this review were considered, as well as some references for the source and audience. First, the sources are divided into human sources [45,54,56,57,67,81] and artificial sources (social bots or AI) [46,47,52,71-73], both of which may or may not be experts. For example, in the case of artificial sources, we would distinguish between declared and nondeclared artificial sources, that is, social bots or AIs that could be declared (self-identifying as bots, usually for informational purposes) and undeclared (those acting for some covert purpose). This basic distinction would allow us to differentiate between expert and nonexpert sources, while also differentiating the positive work of social bots (eg, at the level of providing information) versus the negative effects that bots or covert AIs might generate.

XSL•FO
RenderX

Figure 3. HMCF6: the conceptual framework illustrates the flow of misinformation from reception to impact. From left to right, the process begins with the reception of information, which is then classified by source type (eg, human or nonhuman and expert or nonexpert), followed by categorization of the message content (eg, misleading, fabricated, or false context). The information is then disseminated through specific channels such as social media, mass media, or health apps, and reaches different audience groups based on their level of health literacy (low, medium, or high). Finally, the process results in potential effects—cognitive, affective, or behavioral—or may produce no effect at all. This sequential model provides a structured understanding of how misinformation travels and influences public health outcomes. AI: artificial intelligence; HMCF6: health misinformation conceptual framework with six dimensions.

In the case of drivers, we would move on a continuum ranging from offering health information with the goal of helping [39,54,61,76,81] to the extreme case of trolling [54,55,58,72] whose objective would be to create chaos (ie, damaging reputation or annoying others, sometimes without even a clear interest beyond personal satisfaction). The humorous messages [54,76,79,80] would be defined as the least harmful, assuming that audiences would be able to grasp the double meaning of the information (although logically it is not free of misunderstandings), while at the other extreme, we would have content fabrication. The channel would be divided along a gradient from information from health expert communication channels to open use of information by non–health expert groups, covering the following categories: social media [40,41,53,65,73,89], health websites, health apps, health forums, mass media [38,40,58,65], and health authorities, as the least likely to misinform. Taking into account the wide diversity of profiles of the misinformation target groups, the types of audiences would be delimited by the level of health literacy, and the effects would be as follows: no effects, cognitive effects, affective effects, behavioral effects, and, finally, direct impact on physical and mental health [38,47,63,66,71,73,78]. Table 1 provides a full operational description of the 6 dimensions.

XSL•FO
RenderX

Table 1. Operationalization of conceptual framework dimensions.

Dimension, application, and category	Example
Source	
Source profiling and classification to detect recurrent misinformation actors	
Undeclared bot or AI[a]	Automated Twitter account spreading antivaccine narratives
Declared bot or AI	Verified health bot disseminating WHO[b] updates
Nonexpert human	Celebrity sharing unverified home remedy
Expert human	Physician issuing early advice without evidence
Drivers	
Intent detection via motive-based message tagging	
Create chaos or trolling	Troll account spreading fake death rates to incite panic
Political	Post undermining public health policy during election
Economic	Promoting false cure to sell supplements
Psychological	Seeking attention via viral misinformation
Informative	Health worker unintentionally sharing outdated guidelines
Message	
Risk-level classification of misinformation types based on verifiability and fabrication	
Fabricated content	Claim that vaccines contain surveillance chips
Manipulated content	Edited video misrepresenting a public health official
Imposter content	Fake website mimicking the CDC[c] logo
False context	Old photo presented as a current event
Misleading content	Selective use of statistics to discredit vaccines
False connection	Clickbait headline about vaccine danger
Satire or parody	Humorous meme misinterpreted as fact
Channel	
Channel mapping and platform-specific risk alerting	
Social media	Facebook groups spreading antimask claims
Health websites	Alternative health blog promoting nonevidence-based remedies
Health apps	Unregulated app suggesting treatments based on symptoms
Health forums	Discussion thread recommending dangerous practices
Mass media	Sensationalist television report linking vaccines to infertility
Health authorities	Official health bulletin (less likely to misinform)
Audience	
Segmentation of target populations by health literacy or vulnerability	
Not health literate or fanatic	Users in conspiracy groups believing 5G causes COVID-19
Lowly health literate	Older adult individuals forwarding chain messages on WhatsApp
Mid health literate	General public with moderate science education
Highly health literate	Medical professionals engaging in fact-checking
Outcome	
Monitoring behavioral, emotional, or cognitive impacts postexposure	
Health outcome	Delayed cancer diagnosis due to fear of hospitals
Behavioral effect	Refusal to wear masks or get vaccinated
Affective effect	Anxiety and panic from repeated exposure to false claims

XSL•FO
RenderX

Dimension, application, and category	Example
Cognitive effect	Confusion about the safety of vaccines
No effect	Message ignored or dismissed due to lack of credibility

[a]AI: artificial intelligence.

[b]WHO: World Health Organization.

[c]CDC: Centers for Disease Control and Prevention.

According to the different dimensions and categories, the resulting conceptual framework assigns numerical scores to different elements to represent their relative impact on the spread and consequences of health misinformation. These scores range from 0 to 20, where 0 indicates no impact and 20 represents the highest level of health misinformation risk. The source of misinformation is assigned a baseline score (0), acknowledging that different actors—whether human or artificial—can promote misinformation depending on their nonexpertise or undeclared character (1). Drivers of misinformation go from unintentional misinformative motives (0) to specific motives for deceiving or harming others (3). Message types are ranked on a scale from 0 to 6, with satirical or parodic content receiving the lowest score (0) due to its lower likelihood of deception, while fabricated content ranks highest (6) as it represents fully false and intentionally misleading information. Dissemination channels are evaluated between 0 and 3, with health authorities at the lower end of the spectrum (formal channels) and social media at the upper end (informal channels), highlighting their dominant role in the rapid and widespread diffusion of misinformation. The audience's health literacy level also affects susceptibility, with individuals with high health literacy scoring the lowest (0) in comparison with fanatic audiences (3) in terms of vulnerability to misinformation. Finally, the outcomes of misinformation are structured into 4 levels, ranging from no effect (0) to severe health outcomes (4), which include cognitive, affective, and behavioral consequences that can ultimately lead to detrimental health decisions. As a rule, when a misinformation message falls into multiple categories within a given domain (eg, both mass media and social media as sources), the highest applicable score within the framework should be used. This approach ensures that the scoring system provides a consistent and quantifiable assessment of misinformation dynamics, thereby facilitating the identification of the most critical intervention points for mitigating its impact.

Table 2 presents concrete examples of health misinformation messages and their classification according to the proposed conceptual framework. It also includes an estimated risk score based on the scale defined in Figure 3, indicating their relative harmfulness across dimensions.

Based on the application of the conceptual framework to the examples in the table, the severity and potential impact of different types of health misinformation can be assessed. First, messages that combine high fabrication in content, dissemination through high-reach channels (eg, social media), and targeting of vulnerable audiences (eg, low health literacy or fanatic groups) consistently result in higher total risk scores. These examples, such as vaccine microchip conspiracies or bleach remedies, not only pose direct health risks but also undermine public trust and amplify uncertainty. Second, the source and driver dimensions are critical in distinguishing between unintentional misinformation (eg, experts presenting misinformative messages) and strategically crafted disinformation, with politically, economically, or trolling motivated cases scoring significantly higher (eg, misinformative content). Third, the framework presents content validity in differentiating between harmful and low-impact content—evidenced by the much lower score of a satirical meme, which carries minimal risk when its humorous intent is recognizable. Overall, the framework enables a nuanced evaluation that integrates both structural (source or channel) and psychosocial (audience or effect) factors, providing a robust foundation for prioritizing intervention strategies.

XSL•FO
RenderX

Table 2. Examples of health misinformation, framework mapping, and MRAS[a].

Misinformation example and domain and category	Examples	Risk score
1. "COVID-19 vaccines contain microchips for surveillance" (social bot, source: Telegram, in a low health literate group)		
Source	Undeclared bots	1
Driver	Political distrust	2
Content	Fabricated conspiracy	6
Channel	Telegram and Facebook	3
Audience	Low health literacy	2
Effect	Vaccine hesitancy	3
MRAS	Total	17
2. "Bill Gates, pharmaceutical companies, 5G, vaccines, chips... it's all connected. They want to control us." (singer, source: Instagram, Facebook, and Twitter)		
Source	Famous person (nonexpert)	1
Driver	Psychological	1
Content	Fabricated conspiracy	6
Channel	Instagram, Facebook, and Twitter	3
Audience	General public (mid health literates)	1
Effect	Vaccine hesitancy	3
MRAS	Total	15
3. "Hydroxychloroquine—I don't know, it's looking like it's having some good results. That would be a phenomenal thing." (politician, source: Twitter)		
Source	Political figures (nonexpert)	1
Driver	Political agenda	2
Content	Conspiracy theory	6
Channel	Mass or social media	3
Audience	Polarized groups (fanatic)	3
Effect	Polarization and distrust	3
MRAS	Total	18
4. "COVID-19 is turning people into zombies" (humorist, source: television comedy)		
Source	Social media users (nonexpert)	1
Driver	Informative	0
Content	Satirical misinformation	0
Channel	Instagram and Twitter	3
Audience	General public (mid health literates)	1
Effect	Misinterpretation	1
MRAS	Total	6

[a]MRAS: Misinformation Risk Assessment Scores.

Discussion

Principal Findings

Our systematic review of reviews of health misinformation during the COVID-19 pandemic provides a comprehensive characterization of the sources, drivers, message content, channels, audiences, and health and social outcomes associated with the spread of misinformation. The findings presented in this paper highlight the complexity and multifaceted nature of health misinformation, emphasizing the need for the use of a common language among the various disciplines addressing this global problem to use interoperable definitions.

The prominence of government and political actors, especially the mention of Donald Trump, as major sources of misinformation underscores the polarizing capacity of political figures on public perception [90]. Additionally, the influence of protest movements, scientific publications, social bots and AI, celebrities, health organizations, and mass media evidences

XSL•FO
RenderX

the great diversity of sources that can contribute to the spread of misinformation [47,59,91-94]. This highlights the need to recognize these varied origins to develop interventions aimed at different social groups (eg, by limiting access to underage groups, through fact-checking of content, or the establishment of rules against the deliberate creation and manipulation of content [37]). However, it is striking to note the confusion between misinformation sources and misinformation channels. Health studies refer to social networks as "misinformation sources," whereas, from the perspective of communication sciences and social sciences, they could be considered "misinformation channels." In this sense, the need for a clear conceptualization becomes evident to address the problem adequately [15], which points to the need to incorporate training programs aimed at the adequate dissemination of health information by professionals [95].

In the case of drivers, the identification of political, economic, and psychological motivations as determinants of health misinformation in the context of the recent pandemic reaffirms the interconnectedness of misinformation with broader societal issues (national elections, political debates, new markets and products, viral marketing strategies, and new ways of working and relating to others on social networking platforms). Similarly, it is worth highlighting the relevance of the unintentional production of misinformation for disease prevention and health promotion (including from the perspective of health experts). Although the ultimate goal of these forms of misinformation is usually altruistic (assuming that the aim of professionals is to help), it can also have a significant impact on opinions, attitudes, and behaviors that, directly or indirectly, can affect the health of the population when the evidence is erroneous [21,22].

The wide variety of misinformation, covering topics such as vaccines, treatments, the origin of COVID-19, preventive measures, transmission theories, denialist messages, immunity, diagnoses, and religious content, underscores the adaptability of misinformation to different dimensions of the pandemic [96]. This wide range of misinformation highlights the multifaceted nature of the infodemic [6], where false or misleading information infiltrates various aspects of public health discourse. The variety of issues uniquely linked to the pandemic demonstrates the difficulty of developing a single definition of health misinformation [15], a fact that again underscores the need for a shared conceptual framework upon which to advance the development of concrete and tailored definitions of different health issues susceptible to misinformation (vaccines, eating disorders, treatments, among others). These definitions should not only recognize the complexity of the sources and drivers of misinformation but also provide an appropriate framework for developing effective interventions to mitigate the impact of misinformation on health [2].

The significant role played by social media platforms, including Twitter, Facebook, YouTube, WhatsApp, and Telegram, in the spread of misinformation underscores the pressing need for collaborative efforts between these platforms and regulatory authorities to effectively mitigate the spread of false information. The dynamic social media ecosystem demands a nuanced understanding of the intricate relationship between traditional mass media and emerging social networking platforms [94,97]. While these platforms offer unprecedented reach and immediacy, they also present challenges related to the rapid circulation of information, the amplification of certain narratives, and the potential for the viral spread of misinformation [75]. In navigating this complex landscape, it is crucial to recognize that social media platforms operate as influential channels for information consumption, shaping public perceptions and influencing attitudes on a global scale [88]. As observed in this study, efforts to address misinformation should not only focus on these platforms in isolation but also consider the broader media ecosystem, where mass media news plays a role alongside social media, and direct interpersonal communication. In this context, a holistic approach to addressing health misinformation involves collaboration between social media platforms, traditional media outlets, health authorities, and other stakeholders [94,98]. This collaboration is essential for developing comprehensive strategies that encompass content moderation, fact-checking mechanisms, and educational campaigns aimed at promoting media literacy [94]. Understanding the synergies and interplay between mass media and social media is fundamental for creating effective interventions that can withstand the challenges posed by the rapid dissemination of information in the digital age [99]. Therefore, an integral and collaborative approach is necessary to build resilience against the detrimental effects of misinformation on public perception and, ultimately, public health.

Health misinformation affects the entire population, but it is crucial to recognize specific vulnerable groups, such as those with low incomes, ethnic minorities, and people with low scientific literacy. These groups face unique challenges due to socioeconomic disparities, cultural differences, and unequal access to quality education and scientific knowledge. Therefore, addressing misinformation in these groups requires adapted strategies considering the intersectionality of their social profiles [35]. For example, people with low socioeconomic status may encounter barriers to accessing accurate information, ethnic minorities may face cultural misunderstandings, or even more educated people may sometimes assume certain information arrogantly and uncritically. Thus, understanding the differential susceptibility among these population groups is critical to designing targeted interventions and, ultimately, building their resilience against misinformation.

Additionally, the diverse and wide-ranging effects of health misinformation on trust in institutions, vaccination rates, adherence to (sometimes misguided) treatments, adherence to health norms, and mental health highlight the need to address this problem from a global perspective. The substantial effects on mental health outcomes, including fear, stress, anxiety, and depression (especially in young groups), emphasize the urgent need for mental health support services in future infodemics.

Given the growing risk of future infodemics, our proposed conceptual framework offers not only a systematized understanding of health misinformation but also a foundation for developing actionable tools. It can inform the design of real-time surveillance systems that monitor emerging misinformation by categorizing it according to its source, content, or dissemination channel. Similarly, it provides valuable

XSL•FO
RenderX

guidance for crafting targeted public health messaging aimed at vulnerable audiences identified as being at greater risk. The framework also has applications in enhancing content moderation systems by enabling the detection of high-impact misinformation patterns—for instance, fabricated claims spread by undeclared bots directed at individuals with low health literacy. By classifying misinformation across its key dimensions, more precise and effective countermeasures can be implemented to limit its reach and mitigate its consequences. These practical applications underscore the framework's relevance for navigating the complex and evolving landscape of digital misinformation. Despite increased scholarly interest, few studies have contextualized the problem with such specificity. Existing models tend to focus on isolated aspects or topics, whereas our framework introduces an integrated and operational structure applicable to both academic analysis and intervention strategies [99-101]. Therefore, our framework enables systematic classification across all 6 domains, facilitating tailored responses such as educational initiatives, content moderation, or targeted fact-checking.

Our study contributes to the current debate on misinformation by offering a specific health-oriented conceptual framework, contrasting with broader approaches focused on science-related misinformation [100,101]. While the academic consensus has identified the multiple misinformation sources, dissemination mechanisms, and impacts of misinformation on science, our work refines this understanding by specifically addressing the health crisis context and, in particular, the impact of health misinformation on attitudinal and health outcomes at the individual level. While previous studies have developed taxonomies and conceptual frameworks on misinformation, often focusing on general classifications, media dissemination, or the risks associated with medical communication and social media, this work proposes a more integrative approach by providing a structured framework specifically tailored to health misinformation, which underscores the necessity of a common language and semantic interoperability. Therefore, this conceptual framework has the potential to enhance the management of current and future health emergencies (ie, epidemics, pandemics, natural disasters, and humanitarian crises) by facilitating the rapid identification and classification of health misinformation dimensions within the evolving digital communication landscape.

Beyond identifying the 6 individual dimensions, our findings suggest important interrelations among these elements that shape the dynamics of health misinformation. For instance, certain sources (such as political figures or nondeclared bots) often align with specific drivers such as political or economic agendas, thereby amplifying misleading content through high-reach channels such as social media. Similarly, vulnerable audiences characterized by low health literacy may be disproportionately exposed to complex misinformation narratives, intensifying affective and behavioral health impacts. These interconnected pathways underscore that the spread and impact of health misinformation are not merely additive but synergistic, with drivers and channels interacting to reinforce certain message types and ultimately influence trust, perceptions, and health behaviors. Recognizing these interdependencies provides a more holistic understanding of how misinformation ecosystems function, which is essential for designing targeted, multidimensional interventions.

Limitations and Strengths

This study has several limitations. First, given the rapid advancement of technology, recent publications on bots, AI, or large language models may have emerged after the completion of our review. Second, while misinformation studies have historical roots dating back to the early 20th century, their interaction with evolving technologies and digital health communication is a more recent phenomenon. The expansion of online platforms and social media has significantly altered the drivers and mechanisms of misinformation spread, adding layers of complexity to its study. Additionally, a language bias is present, as only publications in English, Spanish, and French were included. This may have limited the global scope of our findings, potentially excluding valuable perspectives from studies published in other languages, which could provide deeper insights into the spread of health misinformation across diverse cultural and geopolitical contexts. Moreover, for future studies, there is a need to work on the validation of the conceptual framework that has been developed from evidence in other cultural and linguistic contexts.

Despite these limitations, this study also presents several notable strengths. While technological advancements and linguistic constraints may have influenced the scope of our review, our work lays the groundwork for a taxonomy that, although initially developed in the context of the COVID-19 pandemic, can be adapted for future pandemic preparedness. Moreover, the synthesis of findings from this review underscores the urgent need for interoperable definitions of health misinformation for improved measurement of this diffused phenomenon. Given the complexity and evolving nature of misinformation, definitions should be flexible yet aligned with a common framework, facilitating comparisons across different dimensions, including sources, drivers, content, channels, audiences, and impacts. This study highlights that addressing health misinformation and its negative effects requires a multilateral approach involving researchers, governments, social media platforms, health organizations, and communities, as well as the necessity for targeted interventions and educational campaigns tailored to specific demographic groups.

Conclusions

This study highlights the complexity and multifaceted nature of health misinformation during the COVID-19 pandemic, emphasizing the need for an interoperable conceptual framework to facilitate the identification, measurement, and management of this socially harmful phenomenon. By characterizing the main sources, drivers, message types, dissemination channels, audiences, and health-related effects of misinformation, our findings underscore the importance of a shared language across disciplines to improve understanding and intervention strategies. The development of a structured approach to health misinformation contributes not only to advancing research in this field but also to enhancing public health responses in future crises. Furthermore, the impact of misinformation on vulnerable groups, trust in (health) institutions, and overall public health

XSL•FO
RenderX

outcomes underscores the necessity of targeted interventions that integrate media literacy, fact-checking mechanisms, and collaborative efforts between researchers, policymakers, health organizations, and digital platforms. As the digital information ecosystem continues to evolve, addressing misinformation requires a dynamic and interdisciplinary approach that accounts for both technological advancements and societal behavioral patterns. Our proposed framework serves as a foundational step toward mitigating the spread of health misinformation and fostering a more resilient and informed society in the face of future health crises.

Acknowledgments

We would like to acknowledge the support of the University Research Institute for Sustainable Social Development and the University of Cádiz. The publication is part of the project NETDYNAMIC (CNS2022-135907), funded by MCIN/AEI/10.13039/501100011033 and by the European Union "Next Generation EU"/PRTR. This study has also been supported by the project DCODES (PID2020-118589RB-I00), granted by the Spanish Ministry of Science and Innovation and financed by MCIN/AEI/10.13039/501100011033.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 15 KB - publichealth_v11i1e62693_app1.docx]

Multimedia Appendix 2
Data collection.
[XLSX File (Microsoft Excel File), 187 KB - publichealth_v11i1e62693_app2.xlsx]

Multimedia Appendix 3
PRISMA checklist.
[PDF File (Adobe PDF File), 116 KB - publichealth_v11i1e62693_app3.pdf]

References

1. Allcott H, Gentzkow M, Yu C. Trends in the diffusion of misinformation on social media. Res Polit Internet 2019;6(2):1-15 [FREE Full text] [doi: 10.3386/w25500]
2. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021;23(1):e17187 [FREE Full text] [doi: 10.2196/17187] [Medline: 33470931]
3. Törnberg P. Echo chambers and viral misinformation: modeling fake news as complex contagion. PLoS One 2018;13(9):e0203958 [FREE Full text] [doi: 10.1371/journal.pone.0203958] [Medline: 30235239]
4. Xian J, Yang D, Pan L, Wang W, Wang Z. Misinformation spreading on correlated multiplex networks. Chaos 2019;29(11):113123. [doi: 10.1063/1.5121394] [Medline: 31779364]
5. Fernandez-Luque L, Imran M. Humanitarian health computing using artificial intelligence and social media: a narrative literature review. Int J Med Inform 2018;114:136-142. [doi: 10.1016/j.ijmedinf.2018.01.015] [Medline: 29395987]
6. Alvarez-Galvez J, Suarez-Lledo V, Rojas-Garcia A. Determinants of infodemics during disease outbreaks: a systematic review. Front Public Health 2021;9:603603 [FREE Full text] [doi: 10.3389/fpubh.2021.603603] [Medline: 33855006]
7. Roozenbeek J, Schneider C, Dryhurst S, Kerr J, Freeman A, Recchia G, et al. Susceptibility to misinformation about COVID-19 around the world. R Soc Open Sci 2020;7(10):201199 [FREE Full text] [doi: 10.1098/rsos.201199] [Medline: 33204475]
8. Ho K, Chan J, Chiu D. Fake news and misinformation during the pandemic: what we know and what we do not know. IT Prof 2022;24(2):19-24. [doi: 10.1109/mitp.2022.3142814]
9. Lwin M, Lee S, Panchapakesan C, Tandoc E. Mainstream news media's role in public health communication during crises: assessment of coverage and correction of COVID-19 misinformation. Health Commun 2023;38(1):160-168. [doi: 10.1080/10410236.2021.1937842] [Medline: 34157919]
10. Parker L, Byrne J, Goldwater M, Enfield N. Misinformation: an empirical study with scientists and communicators during the COVID-19 pandemic. BMJ Open Sci 2021;5(1):e100188 [FREE Full text] [doi: 10.1136/bmjos-2021-100188] [Medline: 35047703]
11. Suarez-Lledo V, Alvarez-Galvez J. Assessing the role of social bots during the COVID-19 pandemic: infodemic, disagreement, and criticism. J Med Internet Res 2022;24(8):e36085 [FREE Full text] [doi: 10.2196/36085] [Medline: 35839385]

XSL•FO
RenderX

12. Meyrowitsch D, Jensen A, Sørensen JB, Varga TV. AI chatbots and (mis)information in public health: impact on vulnerable communities. Front Public Health 2023;11:1226776 [FREE Full text] [doi: 10.3389/fpubh.2023.1226776] [Medline: 38026315]
13. Krishna A, Thompson T. Misinformation about health: a review of health communication and misinformation scholarship. Am Behav Sci 2019;65(2):316-332. [doi: 10.1177/0002764219878223]
14. Vraga E, Bode L. Defining misinformation and understanding its bounded nature: using expertise and evidence for describing misinformation. Polit Commun 2020;37(1):136-144 [FREE Full text] [doi: 10.1080/10584609.2020.1716500]
15. El Mikati IK, Hoteit R, Harb T, El Zein O, Piggott T, Melki J, et al. Defining misinformation and related terms in health-related literature: scoping review. J Med Internet Res 2023;25:e45731 [FREE Full text] [doi: 10.2196/45731] [Medline: 37556184]
16. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
17. Camargo CQ, Simon FM. Mis-and disinformation studies are too big to fail: six suggestions for the field's future. Misinformation Review. Cambridge: Harvard Kennedy School; 2022. URL: https://misinforeview.hks.harvard.edu/article/mis-and-disinformation-studies-are-too-big-to-fail-six-suggestions-for-the-fields-future/ [accessed 2025-11-04]
18. Larson H. The biggest pandemic risk? Viral misinformation. Nature 2018;562(7727):309. [doi: 10.1038/d41586-018-07034-4] [Medline: 30327527]
19. Pickles K, Cvejic E, Nickel B, Copp T, Bonner C, Leask J, et al. COVID-19 Misinformation Trends in Australia: prospective Longitudinal National Survey. J Med Internet Res 2021;23(1):e23805 [FREE Full text] [doi: 10.2196/23805] [Medline: 33302250]
20. Kim S, Capasso A, Ali S, Headley T, DiClemente R, Tozan Y. What predicts people's belief in COVID-19 misinformation? A retrospective study using a nationwide online survey among adults residing in the United States. BMC Public Health 2022;22(1):2114 [FREE Full text] [doi: 10.1186/s12889-022-14431-y] [Medline: 36401186]
21. O'Grady C. Poor trials of health steps are worse than none, scientists say. Science 2021;374(6572):1180-1181. [doi: 10.1126/science.acx9739] [Medline: 34855506]
22. Bagdasarian N, Cross G, Fisher D. Rapid publications risk the integrity of science in the era of COVID-19. BMC Med 2020;18(1):192 [FREE Full text] [doi: 10.1186/s12916-020-01650-6] [Medline: 32586327]
23. Schonhaut L, Costa-Roldan I, Oppenheimer I, Pizarro V, Han D, Díaz F. Scientific publication speed and retractions of COVID-19 pandemic original articles. Rev Panam Salud Publica 2022;46:e25 [FREE Full text] [doi: 10.26633/RPSP.2022.25] [Medline: 35432503]
24. Giachanou A, Zhang X, Barrón-Cedeño A, Koltsova O, Rosso P. Online information disorder: fake news, bots and trolls. Int J Data Sci Anal 2022;13(4):265-269 [FREE Full text] [doi: 10.1007/s41060-022-00325-0] [Medline: 35574264]
25. Fetzer JH. Disinformation: the use of false information. Minds Mach 2004;14:231-240 [FREE Full text]
26. Hernon P. Disinformation and misinformation through the internet: findings of an exploratory study. Gov Inf Q 1995;12(2):133-139.
27. Baptista JP, Gradim A. Who believes in fake news? Identification of political (a)symmetries. Soc Sci 2022;11(10):460 [FREE Full text] [doi: 10.3390/socsci11100460]
28. Stureborg R, Nichols J, Dhingra B, Yang J, Orenstein W, Bednarczyk R, et al. Development and validation of vaxConcerns: a taxonomy of vaccine concerns and misinformation with crowdsource-viability. Vaccine 2024;42(10):2672-2679. [doi: 10.1016/j.vaccine.2024.02.081] [Medline: 38521676]
29. Fasce A, Schmid P, Holford D, Bates L, Gurevych I, Lewandowsky S. A taxonomy of anti-vaccination arguments from a systematic literature review and text modelling. Nat Hum Behav 2023;7(9):1462-1480. [doi: 10.1038/s41562-023-01644-3] [Medline: 37460761]
30. van der Linden S. Misinformation: susceptibility, spread, and interventions to immunize the public. Nat Med 2022;28(3):460-467. [doi: 10.1038/s41591-022-01713-6] [Medline: 35273402]
31. Bastani P, Hakimzadeh SM, Bahrami MA. Designing a conceptual framework for misinformation on social media: a qualitative study on COVID-19. BMC Res Notes 2021;14(1):408 [FREE Full text] [doi: 10.1186/s13104-021-05822-2] [Medline: 34727969]
32. Ishizumi A, Kolis J, Abad N, Prybylski D, Brookmeyer K, Voegeli C, et al. Beyond misinformation: developing a public health prevention framework for managing information ecosystems. Lancet Public Health 2024;9(6):e397-e406 [FREE Full text] [doi: 10.1016/S2468-2667(24)00031-8] [Medline: 38648815]
33. Page M, McKenzie J, Bossuyt P, Boutron I, Hoffmann T, Mulrow C, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
34. Serrano-Macías M, Álvarez-Gálvez J, Camacho-García M. QUEST Checklist Internet. Zenodo. 2025. URL: https://doi.org/10.5281/zenodo.15697534 [accessed 2025-10-18]
35. Häfliger C, Diviani N, Rubinelli S. Communication inequalities and health disparities among vulnerable groups during the COVID-19 pandemic - a scoping review of qualitative and quantitative evidence. BMC Public Health 2023;23(1):428 [FREE Full text] [doi: 10.1186/s12889-023-15295-6] [Medline: 36879229]

XSL•FO
RenderX

36. Marques MJ, Gama A, Mendonça J, Fernandes A, Osborne R, Dias S. Assessing health literacy among migrants and associated socioeconomic factors. Eur J Public Health Internet 2021;31(Supplement_3):ckab165.410. [doi: 10.1093/eurpub/ckab165.410]
37. Roozenbeek J, Culloty E, Suiter J. Countering misinformation. Eur Psychol 2023;28(3):189-205 [FREE Full text] [doi: 10.1027/1016-9040/a000492]
38. Casino G. [Communication in times of pandemic: information, disinformation, and provisional lessons from the coronavirus crisis]. Gac Sanit 2022;36 Suppl 1:S97-S104 [FREE Full text] [doi: 10.1016/j.gaceta.2022.01.003] [Medline: 35781157]
39. Ahmad T, Aliaga-Lazarte EAA, Mirjalili S. A systematic literature review on fake news in the COVID-19 pandemic: can AI propose a solution? Appl Sci 2022;12(24):12727 [FREE Full text] [doi: 10.3390/app122412727]
40. Chavda VP, Sonak SS, Munshi NK, Dhamade PN. Pseudoscience and fraudulent products for COVID-19 management. Environ Sci Pollut Res Int 2022;29(42):62887-62912 [FREE Full text] [doi: 10.1007/s11356-022-21967-4] [Medline: 35836045]
41. Mukhtar S. Psychology and politics of COVID-19 misinfodemics: why and how do people believe in misinfodemics? Int Sociol 2020;36(1):111-123 [FREE Full text] [doi: 10.1177/0268580920948807]
42. Ying W, Cheng C. Public emotional and coping responses to the COVID-19 infodemic: a review and recommendations. Front Psychiatry 2021;12:755938 [FREE Full text] [doi: 10.3389/fpsyt.2021.755938] [Medline: 34970164]
43. Sarkar MA, Ozair A, Singh KK, Subash N, Bardhan M, Khulbe Y. SARS-CoV-2 vaccination in India: considerations of hesitancy and bioethics in global health. Ann Glob Health 2021;87(1):124 [FREE Full text] [doi: 10.5334/aogh.3530] [Medline: 34963880]
44. Hakim MS. SARS-CoV-2, COVID-19, and the debunking of conspiracy theories. Rev Med Virol 2021;31(6):e2222 [FREE Full text] [doi: 10.1002/rmv.2222] [Medline: 33586302]
45. Cascini F, Pantovic A, Al-Ajlouni Y, Failla G, Puleo V, Melnyk A, et al. Social media and attitudes towards a COVID-19 vaccination: a systematic review of the literature. eClinicalMedicine 2022;48:101454 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101454] [Medline: 35611343]
46. Ries M. The COVID-19 infodemic: mechanism, impact, and counter-measures—a review of reviews. Sustainability 2022;14(5):2605 [FREE Full text] [doi: 10.3390/su14052605]
47. Dow BJ, Johnson AL, Wang CS, Whitson J, Menon T. The COVID-19 pandemic and the search for structure: social media and conspiracy theories. Soc Personal Psychol Compass 2021;15(9):e12636 [FREE Full text] [doi: 10.1111/spc3.12636] [Medline: 34512798]
48. Giordani RCF, Donasolo JPG, Ames VDB, Giordani R. The science between the infodemic and other post-truth narratives: challenges during the pandemic. Cien Saude Colet 2021;26(7):2863-2872 [FREE Full text] [doi: 10.1590/1413-81232021267.05892021] [Medline: 34231699]
49. Perveen S, Akram M, Nasar A, Arshad-Ayaz A, Naseem A. Vaccination-hesitancy and vaccination-inequality as challenges in Pakistan's COVID-19 response. J Community Psychol 2022;50(2):666-683 [FREE Full text] [doi: 10.1002/jcop.22652] [Medline: 34217150]
50. Shobowale O. A systematic review of the spread of information during pandemics: a case of the 2020 COVID-19 virus. J African Media Stud 2021;13:221-234 [FREE Full text] [doi: 10.1386/jams_00045_1]
51. Tharwani ZH, Kumar P, Marfani WB, Shaeen S, Adnan A, Mohanan P, et al. What has been learned about COVID-19 vaccine hesitancy in Pakistan: insights from a narrative review. Health Sci Rep 2022;5(6):e940 [FREE Full text] [doi: 10.1002/hsr2.940] [Medline: 36425896]
52. van Mulukom V, Pummerer L, Alper S, Bai H, Čavojová V, Farias J, et al. Antecedents and consequences of COVID-19 conspiracy beliefs: a systematic review. Soc Sci Med 2022;301:114912 [FREE Full text] [doi: 10.1016/j.socscimed.2022.114912] [Medline: 35354105]
53. La Bella E, Allen C, Lirussi F. Communication vs evidence: what hinders the outreach of science during an infodemic? A narrative review. Integr Med Res 2021;10(4):100731 [FREE Full text] [doi: 10.1016/j.imr.2021.100731] [Medline: 34141575]
54. Hansson S, Orru K, Torpan S, Bäck A, Kazemekaityte A, Meyer S, et al. COVID-19 information disorder: six types of harmful information during the pandemic in Europe. Journal of Risk Research 2021;24(3-4):380-393 [FREE Full text] [doi: 10.1080/13669877.2020.1871058]
55. Shan Y, Ji M, Xie W, Zhang X, Ng Chok H, Li R, et al. COVID-19-related health inequalities induced by the use of social media: systematic review. JMIR Infodemiology 2022;2(2):e38453 [FREE Full text] [doi: 10.2196/38453] [Medline: 36420437]
56. Gabarron E, Oyeyemi S, Wynn R. COVID-19-related misinformation on social media: a systematic review. Bull World Health Organ 2021;99(6):455-463A [FREE Full text] [doi: 10.2471/BLT.20.276782] [Medline: 34108756]
57. Riaz W, Lee YS. Examining the association between social media misinformation and COVID-19 vaccine hesitancy: a systematic review. KPAQ 2021;33(2):437-457 [FREE Full text] [doi: 10.21888/kpaq.2021.6.33.2.437]
58. Clemente-Suárez VJ, Navarro-Jiménez E, Simón-Sanjurjo JA, Beltran-Velasco A, Laborde-Cárdenas CC, Benitez-Agudelo J, et al. Mis-dis information in COVID-19 health crisis: a narrative review. Int J Environ Res Public Health 2022;19(9):5321 [FREE Full text] [doi: 10.3390/ijerph19095321] [Medline: 35564714]

XSL•FO
RenderX

59. Ullah I, Khan KS, Tahir MJ, Ahmed A, Harapan H. Myths and conspiracy theories on vaccines and COVID-19: potential effect on global vaccine refusals. Vacunas 2021;22(2):93-97 [FREE Full text] [doi: 10.1016/j.vacun.2021.01.001] [Medline: 33727904]
60. Kużelewska E, Tomaszuk M. Rise of conspiracy theories in the pandemic times. Int J Semiot Law 2022;35(6):2373-2389 [FREE Full text] [doi: 10.1007/s11196-022-09910-9] [Medline: 35910405]
61. Kemei J, Alaazi D, Tulli M, Kennedy M, Tunde-Byass M, Bailey P, et al. A scoping review of COVID-19 online mis/disinformation in Black communities. J Glob Health 2022;12:05026. [doi: 10.7189/jogh.12.05026] [Medline: 35866205]
62. Biswas MR, Alzubaidi MS, Shah U, Abd-Alrazaq A, Shah Z. A scoping review to find out worldwide COVID-19 vaccine hesitancy and its underlying determinants. Vaccines (Basel) 2021;9(11):1243 [FREE Full text] [doi: 10.3390/vaccines9111243] [Medline: 34835174]
63. Lewandowsky S, Holford D, Schmid P. Public policy and conspiracies: the case of mandates. Curr Opin Psychol 2022;47:101427 [FREE Full text] [doi: 10.1016/j.copsyc.2022.101427] [Medline: 36029701]
64. Nouri SN, Cohen Y, Madhavan MV, Slomka PJ, Iskandrian AE, Einstein AJ. Preprint manuscripts and servers in the era of coronavirus disease 2019. J Eval Clin Pract 2021;27(1):16-21. [doi: 10.1111/jep.13498] [Medline: 33094906]
65. Vlasschaert C, Topf JM, Hiremath S. Proliferation of papers and preprints during the coronavirus disease 2019 pandemic: progress or problems with peer review? Adv Chronic Kidney Dis 2020;27(5):418-426 [FREE Full text] [doi: 10.1053/j.ackd.2020.08.003] [Medline: 33308508]
66. Pian W, Chi J, Ma F. The causes, impacts and countermeasures of COVID-19 "Infodemic": a systematic review using narrative synthesis. Inf Process Manag 2021;58(6):102713 [FREE Full text] [doi: 10.1016/j.ipm.2021.102713] [Medline: 34720340]
67. de Rosa AS, Mannarini T. The “Invisible Other”: social representations of COVID-19 pandemic in media and institutional discourse. Pap Soc Represent Internet 2020;29(2):5.1-5.35 [FREE Full text]
68. Meese J, Frith J, Wilken R. COVID-19, 5G conspiracies and infrastructural futures. Media Int Aust 2020;177(1):30-46. [doi: 10.1177/1329878x20952165]
69. Abidin C, Lee J, Barbetta T, Miao W. Influencers and COVID-19: reviewing key issues in press coverage across Australia, China, Japan, and South Korea. Media Int Aust 2020;178(1):114-135 [FREE Full text] [doi: 10.1177/1329878x20959838]
70. Romate J, Rajkumar E, Gopi A, Abraham J, Rages J, Lakshmi R, et al. What contributes to COVID-19 vaccine hesitancy? A systematic review of the psychological factors associated with COVID-19 vaccine hesitancy. Vaccines (Basel) 2022;10(11):1777 [FREE Full text] [doi: 10.3390/vaccines10111777] [Medline: 36366286]
71. Ali S. Hu Arenas 2022;5(2):337-352. [doi: 10.1007/s42087-020-00139-1]
72. Aiyer I, Shaik L, Kashyap R, Surani S. COVID-19 misinformation: a potent co-factor in the COVID-19 pandemic. Cureus 2022;14(10):e30026 [FREE Full text] [doi: 10.7759/cureus.30026] [Medline: 36348909]
73. Arigbede OM, Aladeniyi OB, Buxbaum SG, Arigbede O. The use of five public health themes in understanding the roles of misinformation and education toward disparities in racial and ethnic distribution of COVID-19. Cureus 2022;14(10):e30008 [FREE Full text] [doi: 10.7759/cureus.30008] [Medline: 36225248]
74. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt Z. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021;3(3):e175-e194 [FREE Full text] [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
75. Balakrishnan V, Rahman LHA, Tan JK, Lee Y. COVID-19 fake news among the general population: motives, sociodemographic, attitude/behavior and impacts – a systematic review. Online Inf Rev 2023;47:944-973 [FREE Full text] [doi: 10.1108/oir-02-2022-0082]
76. Balakrishnan V, Ng WZ, Soo MC, Han G, Lee C. Infodemic and fake news - a comprehensive overview of its global magnitude during the COVID-19 pandemic in 2021: a scoping review. Int J Disaster Risk Reduct 2022;78:103144 [FREE Full text] [doi: 10.1016/j.ijdrr.2022.103144] [Medline: 35791376]
77. Sharma L, Joshi K, Acharya T, Dwivedi M, Sethy G. Infodemics during era of COVID-19 pandemic: a review of literature. J Fam Med Prim Care 2022;11(8):4236-4239 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_2446_21] [Medline: 36352968]
78. Czudy Z, Matuszczak M, Donderska M, Haczy?ski J. The role of (Dis)information in society during the COVID-19 pandemic. Probl Zarzadzania 2021;19(92):49-63 [FREE Full text] [doi: 10.7172/1644-9584.92.3]
79. Ali SM, Hashmi A, Hussain T. Causes and treatment of COVID-19: myths vs facts. Pak J Pharm Sci 2020;33(4):1731-1734. [Medline: 33583809]
80. Joseph AM, Fernandez V, Kritzman S, Eaddy I, Cook O, Lambros S, et al. COVID-19 misinformation on social media: a scoping review. Cureus 2022;14(4):e24601 [FREE Full text] [doi: 10.7759/cureus.24601] [Medline: 35664409]
81. Goldsmith LP, Rowland-Pomp M, Hanson K, Deal A, Crawshaw A, Hayward S, et al. Use of social media platforms by migrant and ethnic minority populations during the COVID-19 pandemic: a systematic review. BMJ Open 2022;12(11):e061896 [FREE Full text] [doi: 10.1136/bmjopen-2022-061896] [Medline: 36396309]
82. Hussain B, Latif A, Timmons S, Nkhoma K, Nellums L. Overcoming COVID-19 vaccine hesitancy among ethnic minorities: a systematic review of UK studies. Vaccine 2022;40(25):3413-3432 [FREE Full text] [doi: 10.1016/j.vaccine.2022.04.030] [Medline: 35534309]

XSL•FO
RenderX

83. Roy D, Biswas M, Islam E, Azam M. Potential factors influencing COVID-19 vaccine acceptance and hesitancy: a systematic review. PLOS One 2022;17(3):e0265496 [FREE Full text] [doi: 10.1371/journal.pone.0265496] [Medline: 35320309]
84. Tsamakis K, Tsiptsios D, Stubbs B, Ma R, Romano E, Mueller C, et al. Summarising data and factors associated with COVID-19 related conspiracy theories in the first year of the pandemic: a systematic review and narrative synthesis. BMC Psychol 2022;10(1):244 [FREE Full text] [doi: 10.1186/s40359-022-00959-6] [Medline: 36320071]
85. Ripp T, Röer JP. Systematic review on the association of COVID-19-related conspiracy belief with infection-preventive behavior and vaccination willingness. BMC Psychol 2022;10(1):66 [FREE Full text] [doi: 10.1186/s40359-022-00771-2] [Medline: 35292110]
86. Rocha YM, de Moura GA, Desidério GA, de Oliveira CH, Lourenço FD, de Figueiredo Nicolete LD. The impact of fake news on social media and its influence on health during the COVID-19 pandemic: a systematic review. J Public Health 2023;31:1007-1016 [FREE Full text] [doi: 10.1007/s10389-021-01658-z] [Medline: 34660175]
87. Caceres MMF, Sosa JP, Lawrence JA, Sestacovschi C, Tidd-Johnson A, Rasool MHU, et al. The impact of misinformation on the COVID-19 pandemic. AIMS Public Health 2022;9(2):262-277 [FREE Full text] [doi: 10.3934/publichealth.2022018] [Medline: 35634019]
88. do Nascimento IJB, Pizarro AB, Almeida JM, Azzopardi-Muscat N, Gonçalves MA, Björklund M, et al. Infodemics and health misinformation: a systematic review of reviews. Bull World Health Organ 2022;100(9):544-561 [FREE Full text] [doi: 10.2471/BLT.21.287654] [Medline: 36062247]
89. de Albuquerque Veloso Machado M, Roberts B, Wong BLH, van Kessel R, Mossialos E. The relationship between the COVID-19 pandemic and vaccine hesitancy: a scoping review of literature until August 2021. Front Public Health 2021;9:747787 [FREE Full text] [doi: 10.3389/fpubh.2021.747787] [Medline: 34650953]
90. Jungkunz S. Political polarization during the COVID-19 pandemic. Front Political Sci 2021;3:622512. [doi: 10.3389/fpos.2021.622512]
91. Paakkari L, Okan O. COVID-19: health literacy is an underestimated problem. Lancet Public Health 2020;5(5):e249-e250 [FREE Full text] [doi: 10.1016/S2468-2667(20)30086-4] [Medline: 32302535]
92. Himelein-Wachowiak M, Giorgi S, Devoto A, Rahman M, Ungar L, Schwartz H, et al. Bots and misinformation spread on social media: implications for COVID-19. J Med Internet Res 2021;23(5):e26933 [FREE Full text] [doi: 10.2196/26933] [Medline: 33882014]
93. Sule S, DaCosta MC, DeCou E, Gilson C, Wallace K, Goff S. Communication of COVID-19 misinformation on social media by physicians in the US. JAMA Netw Open 2023;6(8):e2328928 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.28928] [Medline: 37581886]
94. Aïmeur E, Amri S, Brassard G. Fake news, disinformation and misinformation in social media: a review. Soc Netw Anal Min 2023;13(1):30 [FREE Full text] [doi: 10.1007/s13278-023-01028-5] [Medline: 36789378]
95. Kaper MS, de Winter AF, Bevilacqua R, Giammarchi C, McCusker A, Sixsmith J, et al. Positive outcomes of a comprehensive health literacy communication training for health professionals in three european countries: a multi-centre pre-post intervention study. Int J Environ Res Public Health 2019;16(20):3923 [FREE Full text] [doi: 10.3390/ijerph16203923] [Medline: 31619010]
96. Enders AM, Uscinski JE, Klofstad C, Stoler J. The different forms of COVID-19 misinformation and their consequences. Harvard Kennedy Sch Misinformation Rev. 2022. URL: https://misinforeview.hks.harvard.edu/article/the-different-forms-of-covid-19-misinformation-and-their-consequences/ [accessed 2025-10-18]
97. Caled D, Silva MJ. Digital media and misinformation: an outlook on multidisciplinary strategies against manipulation. J Comput Soc Sci 2022;5(1):123-159. [doi: 10.1007/s42001-021-00118-8] [Medline: 34075349]
98. Chou WYS, Oh A, Klein WMP. Addressing health-related misinformation on social media. JAMA 2018;320(23):2417-2418. [doi: 10.1001/jama.2018.16865] [Medline: 30428002]
99. Sohn D. Spiral of silence in the social media era: a simulation approach to the interplay between social networks and mass media. Commun Res 2019;49(1):139-166 [FREE Full text] [doi: 10.1177/0093650219856510]
100. Southwell BG, Brennen JSB, Paquin R, Boudewyns V, Zeng J. Defining and measuring scientific misinformation. Ann Am Acad Political Soc Sci 2022;700(1):98-111 [FREE Full text] [doi: 10.1177/00027162221084709]
101. National Academies of Sciences Engineering and Medicine. In: Viswanath K, Taylor TE, Rhodes HG, editors. Understanding and Addressing Misinformation About Science. Washington, D.C: National Academies Press; 2024.

Abbreviations

AI: artificial intelligence
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QUEST: Quality Evaluation of Scientific Texts

XSL•FO
RenderX

Edited by AS Bhagavathula; submitted 29.05.24; peer-reviewed by Y Zhang, N Ismail, B Southwell; comments to author 04.01.25; revised version received 26.02.25; accepted 14.07.25; published 21.11.25.

Please cite as:
Alvarez-Galvez J, Carretero-Bravo J, Lagares-Franco C, Ramos-Fiol B, Ortega-Martin E
Development of a Conceptual Framework of Health Misinformation During the COVID-19 Pandemic: Systematic Review of Reviews
JMIR Public Health Surveill 2025;11:e62693
URL: https://publichealth.jmir.org/2025/1/e62693
doi:10.2196/62693
PMID:41269748

©Javier Alvarez-Galvez, Jesus Carretero-Bravo, Carolina Lagares-Franco, Begoña Ramos-Fiol, Esther Ortega-Martin. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Social Media Use, Influencer Status, and Outdoor Risk-Taking in Australian Adults: Cross-Sectional Survey

Samuel Cornell[1,2], BSc, MSc; Amy Peden[1,2], BA, MPP, PhD

[1]UNSW Beach Safety Research Group, University of New South Wales Sydney, Kensington, Sydney, Australia

[2]School of Population Health, Faculty of Medicine and Health, University of New South Wales Sydney, Sydney, Australia

Corresponding Author:
Samuel Cornell, BSc, MSc
UNSW Beach Safety Research Group, University of New South Wales Sydney, Kensington, Sydney, Australia

Abstract

Background: There is growing awareness of the broader health-related harms of social media; yet, research on social media–related injury mortality and morbidity remains limited. Emerging evidence suggests links between excessive social media use and increased risks of self-harm, cyberbullying-related distress, and dangerous viral challenges, but there has been limited research on the link between time spent on social media and environmental risk-taking, such as risky selfies. However, comprehensive epidemiological studies and policy-driven interventions remain scarce, highlighting the need for further investigation into the public health implications of digital engagement.

Objective: This research aimed to examine the relationship among self-reported time spent on social media, influencer status, and risk-taking behaviors among Australians, considering implications for injury prevention.

Methods: A cross-sectional survey of Australian social media users (N=509) was conducted using stratified quotas to approximate national distributions by age, sex, and geographical location. Participants reported their average daily time spent on social media, whether they identified as a social media influencer, and whether they had ever engaged in risk-taking behavior to create social media content. Associations between categorical variables (eg, influencer status and risk-taking) were examined using Pearson chi-square tests and supplemented with odds ratios (ORs) and 95% CIs. Independent samples 2-tailed *t* tests were used to compare mean time spent on social media between risk-takers and non–risk-takers.

Results: Among participants, 48 (9.4%) self-reported engaging in risk-taking behavior in the outdoors. Influencers were significantly more likely to report risk-taking (28/58, 48.3%) compared to noninfluencers (20/451, 4.4%; χ^2_1=110.57, *P*<.001). Risk-takers (n=48) also spent significantly more time on social media (mean=2.05, SD 1.04) compared to non-risk-takers (n=461; mean 1.37, SD 1.04; $t_{57.22}$=4.31, *P*<.001). In multivariate analyses, influencers (OR 20.11), males (OR 2.00), and younger age groups (eg, OR 33.06 for 18 - 24 vs 55 - 64 years) had significantly higher odds of reporting risk-taking.

Conclusions: Outdoor risk-taking for content creation is associated with influencer status and greater time spent on social media. These findings suggest that policy makers should prioritize regulations addressing risky social media behaviors and hold platforms accountable for promoting harmful content. Social media platforms should implement real-time alerts, pop-up warnings, and geolocated safety information to discourage risky behaviors. Public health practitioners should engage influencers to promote safer content norms and develop targeted injury prevention strategies.

(JMIR Public Health Surveill 2025;11:e73089) doi:10.2196/73089

KEYWORDS

social media; Australia; selfies; outdoors; leisure; tourism; safety

Introduction

The rise in social media use has impacted human behavior, including outdoor activities. One notable change driven by smartphone and social media use is the increasing tendency to take photographs and selfies in risky outdoor settings [1]. While most selfies and photographs are harmless, the pursuit of visually engaging content for social media has led some individuals to take significant risks, often in hazardous locations such as cliffs, waterfalls, and other remote areas [2]. This phenomenon, fueled by social media platforms' algorithmically driven reward systems (likes, comments, and increased visibility), has emerged as a public health concern, resulting in injuries and deaths [3,4].

Social media platforms are intentionally designed to maximize user engagement through algorithmic content curation, variable reward schedules, and metrics like likes, shares, and follower counts. These mechanisms reinforce high-engagement content, which often includes dramatic or risky imagery, indirectly incentivizing risk-taking behaviors [5,6], which we define in

XSL•FO
RenderX

this study as self-reported engagement in actions in outdoor settings that could endanger personal safety, undertaken specifically for the purpose of creating content for social media platforms.

Theoretical frameworks help explain why individuals may engage in these behaviors. The uses and gratifications theory suggests that people actively choose media and content to satisfy specific psychological and social needs, such as entertainment, social interaction, and self-expression [7]. In the context of social media, this theory explains why users—especially influencers—may take risks to create engaging content that meets these needs and attracts audience attention [8].

Previous research has highlighted the link between social media-driven risk-taking and injury, particularly in the context of extreme selfies and dangerous outdoor activities. Studies suggest that social media engagement reinforces risky behaviors through social validation mechanisms [9], yet the specific role of time spent on social media and influencer status in driving these behaviors has remained underexplored.

Negative impacts of social media use on mental health and well-being are increasingly recognized. Excessive social media use has been linked to higher rates of anxiety, depression, self-harm, and exposure to cyberbullying [8,10,11]. These mental health stressors can increase sensation-seeking behavior and risk-taking, particularly among young users seeking validation or coping with digital pressures [12].

Furthermore, previous studies have shown that poor mental health outcomes such as anxiety and depression can contribute to impulsivity and reduced risk appraisal, particularly among adolescents and young adults [12]. These psychological factors may increase the likelihood of engaging in risky behaviors to seek social validation or to cope with negative affect, further compounding the harms associated with excessive social media use.

Influencers, who derive social and economic capital from online visibility, often push content boundaries to maintain engagement [11,13-15]. Their behaviors are highly visible and frequently imitated, contributing to the normalization of risk-taking in pursuit of content creation [16].

While previous studies have documented incidents of injury and death associated with risky selfies and outdoor stunts [1], none have quantitatively examined behavioral correlates such as time spent on social media or self-identified influencer status. These variables are central to understanding the extent to which social media engagement drives behavioral risk.

This study is the first to specifically investigate the relationship between time spent on social media, influencer status, and self-reported risk-taking behavior in outdoor environments. Such behaviors not only pose risks to individuals but can also burden emergency services [3], complicate park and land management, and influence the behavior of broader online audiences. Understanding these relationships is critical for developing targeted interventions that address the underlying social and psychological drivers of risky behaviors for content creation.

Methods

Ethical Considerations

This study involved human participants and was approved by the University of New South Wales Sydney Human Research Ethics Committee (project number HC230479). All participants provided informed consent before commencing the survey. Data were collected anonymously through an established third-party panel provider (Dynata) and were fully deidentified before analysis. No personally identifying information was accessible to the research team at any stage. Participants received standard, non–research-specific compensation in the form of points via Dynata's rewards system. No identifiable images or personal visual data are included in this manuscript or the supplementary materials.

Study Design and Setting

This study used a cross-sectional, self-report survey design to investigate social media use and risk-taking behaviors, which we define in this study as self-reported engagement in actions in outdoor settings that could endanger personal safety. The survey was administered online using the Dynata panel survey platform, which provided a sample that was demographically balanced through stratified quotas based on age, sex, and geographical location. Data collection occurred from October 26 to November 2, 2023.

Participants and Recruitment

Participants were eligible if aged 18 years or older, resided in Australia, and used a social media platform. Recruitment was managed by Dynata, an established international online panel provider. Dynata uses multisource recruitment, including partnerships with websites, email lists, and mobile apps, to reach a broad and diverse participant pool. Stratified quota sampling was used to approximate the Australian population distribution by age, sex, and geographical location, including urban versus regional residence based on postal code. Geographical stratification was applied based on state or territory population distribution and urban versus regional residence based on the residential postal code (zip code) of the respondent. Participants received compensation through Dynata's points-based reward system. The research team was not involved in recruitment or incentives. Although quotas were used to improve demographic balance, the sample was drawn from a nonprobability online panel, and no post-stratification weighting was applied. As such, results should not be interpreted as nationally representative but rather as reflecting patterns among a diverse sample of Australian social media users.

Survey Measures

The wider survey included a combination of closed- and open-ended questions. This study focuses on 3 key items, which were asked using the following exact wording:

1. Time spent on social media: *"On average, how much time do you spend on social media each day?"* Response options: Less than 30 minutes, 30 minutes to 1 hour, 1 to 2 hours, 2 to 3 hours, more than 3 hours.

XSL•FO
RenderX

2. Influencer status: *"Do you consider yourself a social media influencer?"* Response options: Yes/No.
3. Risk-taking behavior: *"Have you ever taken risks in outdoor or recreational settings specifically to create content for social media?"* Response options: Yes/No.

Demographic variables, including age, gender, and geographical location, were also collected.

Analysis

Categorical responses for time spent on social media (eg, "Less than 30 min") were converted into approximate numerical values (eg, 0.5 h for "Less than 30 min"). This allowed the creation of a continuous variable to calculate means and standard deviations, which supports clearer comparisons between groups. This approach is commonly used in social science and health research where ordinal categories represent underlying continuous behaviors, and it allows for the application of parametric tests [17]. Analyses were conducted in R (version 2024.09.0+375; R Core Team) using dplyr for data manipulation, tidyr for data cleaning, and ggplot2 for visualization.

Summary statistics (means, SDs, and frequencies) were calculated for all key variables. Chi-square tests were used to assess associations between categorical variables, such as influencer status and risk-taking behavior. Independent samples 2-tailed *t* tests were used to compare mean time spent on social media between self-reported risk-takers and non–risk-takers. Odds ratios (ORs) and 95% CIs were also calculated from 2×2 contingency tables to quantify the strength of associations between risk-taking and key predictors, including influencer status, gender, and time spent on social media.

Results

The survey had a total of 509 respondents (age: mean 45.9, SD 16.8 years). In total, 256 (50.3%) identified as female, and 11.4% (n=58) considered themselves social media influencers. Participants reported spending an average of 1.43 (SD 1.06) hours per day on social media. Among the sample, 9.43% (n=48) of them reported having taken risks to create social media content.

A significant association was also found between age and risk-taking behavior (χ^2_5=62.61, *P*<.001), with younger participants being more likely to report engaging in risk-taking. Compared to adults aged 55-64 years (reference group), the odds of reporting risk-taking were significantly higher among younger age groups: 18 - 24 years (OR 33.06, 95% CI 6.57 - 806.26; *P*<.001), 25 - 34 years (OR 12.91, 95% CI 2.49 - 318.16; *P*=.0007), and 35 - 44 years (OR 7.31, 95% CI 1.31 - 185.01; *P*=.0223). The proportion of participants engaging in risk-taking was highest among those aged 18 - 24 years (32.3%) and progressively declined with age, with only 1.2% - 1.3% of individuals aged 55 and older engaging in such behavior. Gender differences in risk-taking behavior were also statistically significant (χ^2_1=4.30, *P*=.038). Males (12.5%) were significantly more likely to report engaging in risk-taking behavior compared to females (6.6%). A chi-square test revealed a significant association between identifying as a social media influencer and reporting risk-taking behavior (χ^2_1=110.57, *P*<.001). Cross-tabulation showed that 48.3% (n=28) of influencers reported risk-taking, compared to only 4.4% (n=20) of noninfluencers (Table 1). An independent samples 2-tailed *t* test demonstrated that participants who reported risk-taking behavior spent significantly more time on social media (mean 2.05, SD 1.04) compared to those who did not report risk-taking (mean 1.37, SD 1.04; $t_{57.22}$=4.31, *P*<.001).

XSL•FO
RenderX

Table . Chi-square and *t* test analysis of risk-taking behavior by age, gender, influencer status, and social media usage.

Variable and category	Risk-taking, No, n (%)	Risk-taking, Yes, n (%)	Chi-square (*df*; *P* value)
Age group (years)			62.607 (5; <.001)
18 - 24	44 (67.7)	21 (32.3)	
25 - 34	76 (84.4)	14 (15.6)	
35 - 44	87 (90.6)	9 (9.4)	
45 - 54	90 (97.8)	2 (2.2)	
55 - 64	80 (98.7)	1 (1.3)	
≥65	84 (98.8)	1 (1.2)	
Gender			4.300 (1; .038)
Male	218 (87.5)	31 (12.5)	
Female	239 (93.4)	17 (6.6)	
Other or prefer not to say	4 (100.0)	0 (0.0)	
Influencer status			110.574 (1; <.001)
Yes (influencer)	30 (51.7)	28 (48.3)	
No (noninfluencer)	431 (95.6)	20 (4.4)	
Average daily time spent on social media			23.201 (4; <.001)
<30 min	99 (99.0)	1 (1.0)	
30 min-1 h	135 (92.5)	11 (7.5)	
1 - 2 h	118 (91.5)	11 (8.5)	
2 - 3 h	53 (79.1)	14 (20.9)	
>3 h	56 (83.5)	11 (16.4)	

In bivariate analyses, influencer status, male gender, younger age, and higher daily social media use were all significantly associated with increased odds of reporting risk-taking behavior (Table 2). Influencers had 20.11 times the odds of reporting risk-taking compared to noninfluencers (95% CI 10.16 - 39.81, *P*<.001). Males had 2 times the odds of reporting risk-taking compared to females (95% CI 1.04 - 3.85, *P*=.033). Compared to adults aged 55-64 years (reference group), the odds of reporting risk-taking were significantly higher among younger age groups: 18 - 24 years (OR 33.06, 95% CI 6.57 - 806.26; *P*<.001), 25 - 34 years (OR 12.91, 95% CI 2.49 - 318.16; *P*=.0007), and 35-44 years (OR 7.31, 95% CI 1.31 - 185.01; *P*=.0223). No significant differences were observed for the 45- to 54-year or 65- to 99-year age groups. Individuals spending more than 1 hour per day on social media had 3.09 times the odds of reporting risk-taking compared to those spending 1 hour or less per day (95% CI 1.57 - 6.09, *P*=.001).

Table . Odds ratios (ORs), 95% CIs, and *P* values for reporting risk-taking behavior by influencer status, gender, age, and daily social media use. Reference categories: non-influencer, female, 55 - to 64-year age group, and ≤1 hour/day social media use. “Ref” denotes the reference group.

Predictor and comparison	OR	95% CI	*P* value
Influencer status: influencer vs noninfluencer (ref)	20.11	10.16 - 39.81	<.001
Gender: male vs female (ref)	2.00	1.04 - 3.85	.033
Daily social media use: >1 hour/day vs ≤1 hour/day (ref)	3.09	1.57 - 6.09	.001
Age group (years)			
18 - 24 vs 55 - 64 (ref)	33.06	6.57 - 806.26	<.001
25 - 34 vs 55 - 64 (ref)	12.91	2.49 - 318.16	.0007
35 - 44 vs 55 - 64 (ref)	7.31	1.31 - 185.01	.0223
45 - 54 vs 55 - 64 (ref)	1.67	0.13 - 53.05	1.000
65 - 99 vs 55 - 64 (ref)	0.95	0.02 - 37.58	1.000

XSL•FO
RenderX

Discussion

Principal Findings

This study provides important insights into the associations between social media use, influencer status, and risk-taking behaviors, with implications for injury prevention strategies in outdoor settings. The strongest statistical finding shows that risk-taking behavior for social media is significantly more prevalent among younger individuals, particularly those aged 18 - 24, who had over 30 times the odds of reporting risk-taking compared to adults aged 55 - 64. Adults aged 25-34 years and those aged 35-44 years also had significantly increased odds of risk-taking. These findings align with existing literature on developmental psychology, where younger adults exhibit greater sensation-seeking tendencies and risk-taking behavior [12,18]. Public health interventions should therefore focus on younger demographics, where risk perception and decision-making processes are still developing.

Social media influencers and individuals who self-reported greater average daily time on social media were also significantly more likely to engage in risk-taking behaviors for content creation. Such behaviors, which can include climbing cliffs [1], entering unsafe waters, or ignoring safety warnings, may elevate the risk of injury or death when in the outdoors [3].

Gender differences in risk-taking were also observed. Males had twice the odds of reporting risk-taking compared to females. This finding is consistent with previous research indicating that men are more inclined toward risk-taking, particularly in outdoor and physically demanding contexts [19]. Future research should further explore the interaction between gender and social media use, including whether platform dynamics amplify existing gender-based differences in risk propensity.

Social media has been characterized as addictive, with research suggesting that its design leverages psychological mechanisms such as variable reward schedules, social validation, and engagement-driven algorithms to encourage prolonged use [20] as well as platform business models strategically designed to amplify user engagement, thereby encouraging prolonged and frequent use [21]. The link between time spent on social media and changes in behavior has been well researched in the domains of well-being, mental health, and anxiety [9,10,22]. Studies often find no strong relationship and inconclusive results [10].

On the contrary, our findings indicate a strong association between social media use and risk-taking behaviors in outdoor settings. The association raises the possibility that increased social media use may influence users' willingness to take risks, although causality cannot be determined from cross-sectional data.

Our study is novel in that it directly asks social media users to self-assess whether they engage in risk-taking behaviors in outdoor settings. The perception of one's behavior as risky may be influenced by societal norms and the extent to which individuals believe others perceive their actions as hazardous [23]. In contrast, previous studies on social media impacts have primarily focused on subjective well-being, a construct that may be less dependent on external perceptions and more intrinsically determined [24,25].

Influencer status was also strongly associated with increased risk-taking behavior. In this study, individuals who identified as influencers had over 20 times the odds of reporting risk-taking compared to noninfluencers. Influencers often set trends that prioritize visually striking or extreme content, which may prompt followers to imitate risky behaviors in pursuit of attention, validation, or perceived status [4].

These trends have significant implications for injury prevention efforts. Social media plays a critical role in shaping peer norms and pressures, particularly among young adults and impressionable users [26,27]. Addressing this issue requires interventions that directly target the influence of social media on behavioral norms; perhaps also the influence of the influencer.

Given the strong association between time spent on social media and risk-taking behavior, public health interventions should consider targeting excessive social media use as a potential risk factor but also advocate for the removal of content that glorifies risk-taking behaviors. Addressing compulsive social media engagement through media literacy initiatives may contribute to reducing outdoor risk-taking. Future research is needed to establish whether social media exposure directly increases risk-taking or if individuals predisposed to risk-seeking behaviors are more likely to engage with social media. Platforms have a responsibility to collaborate with safety organizations and influencers to encourage the creation and dissemination of responsible content. Public health practitioners should seek to collaborate with influencers who have sway over a section of the population that they are trying to reach. These efforts are crucial for breaking the cycle of risky behaviors perpetuated by social media and preventing avoidable injuries.

These findings align with established public health frameworks that emphasize modifying environmental, social, and individual-level risk factors. For example, the Haddon Matrix [28], widely used in injury prevention, highlights the importance of addressing pre-event, event, and postevent factors, which could inform digital safety interventions. Similarly, behavior change theories such as the theory of planned behavior [29] suggest that modifying perceived norms and enhancing self-efficacy are key strategies to reduce risky behaviors online. Applying these frameworks can help guide multifaceted interventions involving policy changes, platform design modifications, and targeted education campaigns.

Several potential confounding factors should be considered when interpreting these associations. Individual traits such as sensation-seeking [12], impulsivity, and susceptibility to peer influence [16] may predispose individuals both to greater social media engagement and to risk-taking behaviors. Other factors, including socioeconomic status [18], education level, and previous experiences with outdoor activities, could also influence both social media use patterns and risk-taking propensity. These confounders should be addressed in future longitudinal or multivariate studies.

XSL•FO
RenderX

While this study highlights potential interventions—such as platform-based warnings, real-time alerts, and collaboration with influencers—it does not offer a detailed implementation or evaluation framework. This reflects the exploratory nature of the work, which aimed to establish baseline associations rather than test interventions. Further research is needed to co-design, implement, and evaluate specific harm-reduction strategies with stakeholders, including social media platforms, public health agencies, and tourism authorities. Doing so will be critical for translating these findings into actionable outcomes. The findings, and other work being conducted in this area, suggest that a public health approach is necessary to consider these types of behaviors and resulting incidents as a widespread issue that is having repercussions on a population-level basis.

Potential ethical considerations must also be acknowledged when designing interventions aimed at reducing risk-taking behaviors [30]. Efforts to remove or limit harmful online content could raise concerns about censorship, freedom of expression, and the overreach of platform policies. Balancing safety with the rights of users to share content is a complex challenge. There may need to be more discussion between ethicists, legal scholars, public health academics, and the public to determine what people are willing to lose for the sake of increased online safety.

These findings suggest the need for policy makers to consider regulations or guidelines addressing risky online behaviors, including the promotion of hazardous content. Social media platforms should be encouraged or required to implement design changes such as pop-up warnings, geolocation-based risk alerts, and clearer community guidelines discouraging risk-taking behaviors.

Limitations

A key limitation of this study is its cross-sectional design, which restricts the ability to draw causal inferences. While significant associations were found between time spent on social media, influencer status, and self-reported risk-taking behaviors, the direction of these relationships cannot be determined. It remains unclear whether increased social media use contributes to greater risk-taking in outdoor settings or whether individuals predisposed to risk-taking are more likely to engage with social media platforms and content. Longitudinal or experimental research designs are needed to establish causal pathways and better understand the mechanisms driving these behaviors. Nevertheless, our study reveals that risk-takers spend significantly more time on social media. Although quota sampling was used to improve demographic balance, the sample was drawn from a nonprobability online panel and no statistical corrections were applied to adjust for sampling bias. These factors limit the generalizability of the findings, which should be interpreted with caution.

This study did not examine broader contextual or psychosocial factors that may influence risk-taking behavior in the social media environment, such as peer influence, personality traits, or platform-specific dynamics. Future research should consider these variables to better understand the mechanisms driving risky content creation. Finally, while ORs provided insights into demographic and behavioral predictors, some confidence intervals were wide, reflecting small numbers of risk-takers in certain subgroups. This should be considered when interpreting effect sizes.

Conclusions

This study offers a brief overview of the role of time spent on social media in shaping risk-taking in the outdoors and its broader implications for public health. The study determined a significant link between time spent on social media, influencer status, and risk-taking. Time spent on social media may need to be considered by public health practitioners and researchers when designing or testing interventions to improve safety in the outdoors at places frequented by social media users. Future research is critical to explore causal pathways between social media use and injury risks, as well as platform-specific trends that influence behavior. Policy makers should prioritize regulations addressing risky online behaviors and encourage platform accountability. Social media companies should also be engaged to co-design technological solutions aimed at discouraging risky content creation and promoting safer digital engagement.

Acknowledgments

The authors thank the participants who took part in this research. We also acknowledge the support of the UNSW Beach Safety Research Group and the School of Population Health at UNSW Sydney. This research was supported by Royal Life Saving Society–Australia to aid in the prevention of drowning. Research at Royal Life Saving Society–Australia is supported by the Australian Government. Samuel Cornell is funded by an Australian Government Research Training Program (RTP) Scholarship. AP is supported by a National Health and Medical Research Council (NHMRC) Emerging Leadership Fellowship (grant ID: APP2009306). The authors also acknowledge the support of Surf Life Saving Australia. SC is funded by an Australian Government University Postgraduate Award Scholarship via the University of New South Wales. The research is also supported by Surf Life Saving Australia.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Cornell S, Brander R, Peden A. Selfie-related incidents: narrative review and media content analysis. J Med Internet Res 2023 Sep 27;25(1):e47202. [doi: 10.2196/47202] [Medline: 37756044]
2. Cornell S, Peden AE. Visual "scrollytelling": mapping aquatic selfie-related incidents in Australia. Interact J Med Res 2024 May 23;13(1):e53067. [doi: 10.2196/53067] [Medline: 38781002]
3. Cornell S, Brander RW, Peden AE. Preventing selfie-related incidents: taking a public health approach to reduce unnecessary burden on emergency medicine services. Emerg Med Australas 2023 Aug;35(4):691-693. [doi: 10.1111/1742-6723.14219] [Medline: 37454361]
4. Cornell S, Brander RW, Peden AE. Likes, comments, and emergency responses: exploring the burden of social media tourism on land managers and the need for risk communication. Health Literacy Commun Open 2024 Dec 31;2(1):2313846. [doi: 10.1080/28355245.2024.2313846]
5. Montag C, Lachmann B, Herrlich M, Zweig K. Addictive features of social media/messenger platforms and freemium games against the background of psychological and economic theories. Int J Environ Res Public Health 2019 Jul 23;16(14):2612. [doi: 10.3390/ijerph16142612] [Medline: 31340426]
6. Zuboff S. Surveillance capitalism and the challenge of collective action. New Labor Forum 2019 Jan;28(1):10-29. [doi: 10.1177/1095796018819461]
7. Katz E, Blumler JG, Gurevitch M. Uses and gratifications research. Public Opin Q 1973;37(4):509. [doi: 10.1086/268109]
8. Plackett R, Blyth A, Schartau P. The impact of social media use interventions on mental well-being: systematic review. J Med Internet Res 2023 Aug 11;25:e44922. [doi: 10.2196/44922] [Medline: 37565693]
8. Plackett R, Blyth A, Schartau P. The impact of social media use interventions on mental well-being: systematic review. J Med Internet Res 2023 Aug 11;25(1):e44922. [doi: 10.2196/44922] [Medline: 37565693]
10. Jones CN, Rudaizky D, Mahalingham T, Clarke PJF. Investigating the links between objective social media use, attentional control, and psychological distress. Soc Sci Med 2024 Nov;361:117400. [doi: 10.1016/j.socscimed.2024.117400] [Medline: 39388755]
11. Barry CT, Sidoti CL, Briggs SM, Reiter SR, Lindsey RA. Adolescent social media use and mental health from adolescent and parent perspectives. J Adolesc 2017 Dec;61(1):1-11. [doi: 10.1016/j.adolescence.2017.08.005] [Medline: 28886571]
12. Khurana A, Bleakley A, Ellithorpe ME, Hennessy M, Jamieson PE, Weitz I. Sensation seeking and impulsivity can increase exposure to risky media and moderate its effects on adolescent risk behaviors. Prev Sci 2019 Jul;20(5):776-787. [doi: 10.1007/s11121-019-0984-z] [Medline: 30659453]
13. Abidin C. "Aren't these just young, rich women doing vain things online?": Influencer selfies as subversive frivolity. Social Media Soc 2016 Apr 1;2(2):2056305116641342. [doi: 10.1177/2056305116641342]
14. Arriagada A, Bishop S. Between commerciality and authenticity: the imaginary of social media influencers in the platform economy. Commun Culture Critique 2021 Nov 26;14(4):568-586. [doi: 10.1093/ccc/tcab050]
15. Djafarova E, Trofimenko O. 'Instafamous' – credibility and self-presentation of micro-celebrities on social media. Inf Commun Soc 2019 Aug 24;22(10):1432-1446. [doi: 10.1080/1369118X.2018.1438491]
16. Reniers R, Beavan A, Keogan L, Furneaux A, Mayhew S, Wood SJ. Is it all in the reward? Peers influence risk-taking behaviour in young adulthood. Br J Psychol 2017 May;108(2):276-295. [doi: 10.1111/bjop.12195] [Medline: 26990740]
17. Bland JM, Altman DG. Statistics notes: the use of transformation when comparing two means. Br Med J 1996;312. [doi: 10.1136/bmj.312.7039.1153] [Medline: 8620137]
18. Steinberg L, Icenogle G, Shulman EP, et al. Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation. Dev Sci 2018 Mar;21(2):e12532. [doi: 10.1111/desc.12532] [Medline: 28150391]
19. Frick B. Gender differences in risk-taking and sensation-seeking behavior: empirical evidence from "ExtremeSports". Economy 2021 Feb;169(1):5-20. [doi: 10.1007/s10645-020-09373-y]
20. Andreassen CS. Online social network site addiction: a comprehensive review. Curr Addict Rep 2015;2:175-184 [FREE Full text] [doi: 10.1007/s40429-015-0056-9]
21. Claussen J, Kretschmer T, Mayrhofer P. The effects of rewarding user engagement: the case of Facebook apps. Inf Syst Res 2013 Mar;24(1):186-200. [doi: 10.1287/isre.1120.0467]
22. Orben A. Teenagers, screens and social media: a narrative review of reviews and key studies. Soc Psychiatry Psychiatr Epidemiol 2020 Apr;55(4):407-414. [doi: 10.1007/s00127-019-01825-4] [Medline: 31925481]
23. Slovic P. Perception of risk. Science 1987 Apr 17;236(4799):280-285. [doi: 10.1126/science.3563507] [Medline: 3563507]
24. Wirtz D, Tucker A, Briggs C, Schoemann AM. How and why social media affect subjective well-being: multi-site use and social comparison as predictors of change across time. J Happiness Stud 2021 Apr;22(4):1673-1691. [doi: 10.1007/s10902-020-00291-z]
25. Webster D, Dunne L, Hunter R. Association between social networks and subjective well-being in adolescents: a systematic review. Youth Soc 2021 Mar;53(2):175-210. [doi: 10.1177/0044118X20919589]
26. Keles B, McCrae N, Grealish A. A systematic review: the influence of social media on depression, anxiety and psychological distress in adolescents. Int J Adolesc Youth 2020 Dec 31;25(1):79-93. [doi: 10.1080/02673843.2019.1590851]

XSL•FO
RenderX

27. Cornell S, Piatkowski T, Taba M, Peden AE. Meeting people where they are: leveraging influencers' social capital and trust to promote safer behaviours. Health Lit Commun Open 2024 Dec 31;2(1):2350155. [doi: 10.1080/28355245.2024.2350155]
28. Haddon W. Advances in the epidemiology of injuries as a basis for public policy. Public Health Rep 1980;95(5):411-421. [Medline: 7422807]
29. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
30. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. Pharm Ther 2014 Jul;39(7):491-520 [FREE Full text]

Abbreviations

OR: odds ratio

Edited by A Mavragani; submitted 24.02.25; peer-reviewed by K Elechi, Y Sarikhani; revised version received 28.05.25; accepted 28.05.25; published 21.08.25.

Please cite as:
Cornell S, Peden A
Social Media Use, Influencer Status, and Outdoor Risk-Taking in Australian Adults: Cross-Sectional Survey
JMIR Public Health Surveill 2025;11:e73089
URL: https://publichealth.jmir.org/2025/1/e73089
doi:10.2196/73089

© Samuel Cornell, Amy Peden. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Dissemination of Federal Risk Communication by News Media Outlets During Enteric Illness Outbreaks: Canadian Content Analysis

Hisba Shereefdeen[1,2], MSc; Lauren Elizabeth Grant[1], PhD; Vayshali Patel[1,2], BSc; Melissa MacKay[1], PhD; Andrew Papadopoulos[1], PhD; Leslie Cheng[2], MPH; Melissa Phypers[2], MPA, MSc; Jennifer Elizabeth McWhirter[1], PhD

[1]Department of Population Medicine, Ontario Veterinary College, University of Guelph, Population Medicine Building, Room 201, 50 Stone Rd E, Guelph, ON, Canada

[2]Outbreak Management Division, Centre for Foodborne, Environmental & Zoonotic Infectious Diseases, Public Health Agency of Canada, Guelph, ON, Canada

Corresponding Author:
Lauren Elizabeth Grant, PhD
Department of Population Medicine, Ontario Veterinary College, University of Guelph, Population Medicine Building, Room 201, 50 Stone Rd E, Guelph, ON, Canada

Abstract

Background: Effective dissemination of federal risk communication by news media during multijurisdictional enteric illness outbreaks can increase message reach to rapidly contain outbreaks, limit adverse outcomes, and promote informed decision-making by the public. However, dissemination of risk communication from the federal government by mass media has not been evaluated.

Objective: This study aimed to describe and assess the dissemination of federal risk communication by news media outlets during multijurisdictional enteric illness outbreaks in Canada.

Methods: A comprehensive systematic search of 2 databases, Canadian Newsstream and Canadian Business & Current Affairs, was run using search terms related to the source of enteric illnesses, general outbreak characteristics, and relevant enteric pathogen names to retrieve news media articles issued between 2014 and 2023, corresponding to 46 public health notices (PHNs) communicating information about multijurisdictional enteric illness outbreaks during the same period. A codebook comprised of 3 sections—general characteristics of the article, consistency and accuracy of information presented between PHNs and news media articles, and presence of health belief model constructs—was developed and applied to the dataset. Data were tabulated and visualized using RStudio (Posit).

Results: News media communicated about almost all PHNs (44/46, 96%). News media commonly developed their own articles (320/528, 60.6%) to notify the public about an outbreak and its associated product recall (121/320, 37.8%), but rarely communicated about the conclusion of an outbreak (12/320, 3.8%). News media communicated most outbreak characteristics, such as the number of cases (237/319, 74.3%), but the number of deaths was communicated less than half the time (114/260, 43.8%). Benefit and barrier constructs of the health belief model were infrequently present (50/243, 20.6% and 15/243, 6.2%, respectively).

Conclusions: Canadian news media disseminated information about most multijurisdictional enteric illness outbreaks. However, differences in coverage of multijurisdictional enteric illness outbreaks by news media were evident. Federal organizations can improve future risk communication of multijurisdictional enteric illness outbreaks by news media by maintaining and strengthening interorganizational connections and ensuring the information quality of PHNs as a key information source for news media.

(JMIR Public Health Surveill 2025;11:e68724) doi:10.2196/68724

KEYWORDS

risk communication; health communication; enteric illness; foodborne illness; zoonotic disease; media; content analysis; health belief model; public health; Canada

Introduction

Risk communication is a fundamental component of risk analysis that involves the exchange of information about risk and risk-related factors using a range of appropriate tools and strategies [1]. Effective risk communication is instrumental in promoting awareness [2], increasing behavior uptake [3], and mitigating adverse health events [4,5]. Food safety risk communication has 2 aims depending on the audience. For those in public and private sectors, it ensures a clear and transparent decision-making process to assess and manage food safety risks. For the public, it enhances awareness of relevant food safety risks and promotes health-protective behaviors and corrective actions for ill persons [6,7]. In our study, risk communication

XSL•FO
RenderX

is defined as conveying information to increase awareness or promote behavior change to prevent exposure or manage illness during a multijurisdictional enteric illness outbreak.

Enteric illnesses are an important public health concern that impacts an estimated 4 million Canadians each year [8,9]. Enteric pathogens are transmitted via the fecal-oral route through consumption of contaminated food or water, animal contact, or person-to-person transmission [10-12]. Enteric illnesses are usually acute and self-limited but can cause serious illness or death, particularly in vulnerable populations [13-16]. In addition, enteric illnesses can result in chronic sequelae and morbidity following acute infection [17-20]. As a result, rapid and effective control of enteric illness outbreaks that limit potential widespread exposure to enteric pathogens and resultant adverse health outcomes and health care utilization is essential. One of the broadest and fastest tools that can be deployed once an outbreak is detected is risk communication to the public, and as such, effective risk communication is essential for controlling outbreaks and safeguarding human health [21-23].

In Canada, the response to multijurisdictional enteric illness outbreaks is guided by the Foodborne Illness Outbreak Response Protocol [24]. This protocol outlines the coordinated approach of cross-sectoral organizations, including the Public Health Agency of Canada (PHAC), the Canadian Food Inspection Agency (CFIA), Health Canada, and provincial or territorial organizations. When an enteric illness outbreak occurs in 2 or more provinces or territories or between Canada and another country (multijurisdictional), PHAC manages the outbreak response to ensure rapid and coordinated action across multiple jurisdictions [24]. If deemed necessary, PHAC will issue web-based public health notices (PHNs) and social media posts to communicate outbreak information and behavioral recommendations [24]. This information is intended to be used by health organizations, news media, and the public as a central source of factual, timely information.

News media are not required to convey outbreak information; however, their ability to relay information and influence health decision-making make them an important communication pathway during an outbreak [25-27]. News media use multiple communication channels to reach their target audiences including traditional (eg, television and print media) and online formats (eg, web, social media, and podcasts). These platforms are a popular source of health information among Canadians, with 78% relying on the internet to access health information during the COVID-19 pandemic [28]. Given that 95% of Canadians are internet users, online formats have the broadest potential reach. On average, Canadian adults spend over 1 hour reading press media online each day [29]. Despite the digitalization of print publications, print newspapers remain the most trusted news source for Canadians [30]. Both print and digital newspapers maintain high readership across generations, with 83% of Canadian millennials, 83% of Gen X adults, and 78% of boomers engaging with these formats [31]. As intermediaries between public health agencies and the public, news media have the potential to effectively shape how risk information is framed and received. However, the coverage of information may be influenced by factors beyond public health priorities, such as agenda-setting priorities, where news media may highlight or downplay certain information to promote or limit the public's understanding [32-34].

News media are valuable information sources with broad public reach that can disseminate PHAC's public communication during multijurisdictional enteric illness outbreaks to ensure timely delivery of information to at-risk populations. Given their potential reach and impact, it is imperative to optimize this communication pathway to maximize the number of Canadians receiving information during an outbreak. However, a formal assessment of how and the extent to which news media disseminate federal risk communication in this context has not been conducted. This information is needed to inform how public communication during enteric illness outbreaks can be improved to increase access to information and support adoption of protective behaviors, access to appropriate treatment options, and ultimately limit adverse population health outcomes. In this study, dissemination is assessed by the frequency of communication, determined by the number of news media articles, and the inclusion of information from PHNs in news media articles. The aim of this study is to assess the communication pathway between PHAC and news media by describing how, and the extent to which, Canadian news media disseminate federal risk communication during multijurisdictional enteric illness outbreaks.

Methods

Data Collection

In consultation with the research team and a Research Librarian from the University of Guelph, search strings in English and French were developed using a combination of general enteric illness outbreak terms and specific pathogens that have caused multijurisdictional enteric illness outbreaks in Canada:

1. (food OR waterborne OR zoono* OR case? OR illness* OR contamina* OR outbreak? OR recall* OR toxic* OR advisor* OR alert* OR sick* OR disease*) AND (norovirus OR salmonella OR cyclospora OR hepatitis OR hep OR listeria OR listeriosis OR vibrio OR Ecoli OR "E.coli" OR "E coli" OR "E. coli" OR "Escherichia coli" OR Pseudomonas OR botulism).
2. (aliment OR "maladies d'origine hydrique" OR "zoonotiques" OR cas OR maladie, OR contaminent OR éclosion OR rappel OR toxique OR avis OR alerte OR malade) AND (norovirus OR salmonella OR Cyclospora OR hépatite OR listeria OR listériose OR vibrio OR Ecoli OR "E.coli" OR "E coli" OR "E. coli" OR "Escherichia coli" OR pseudomonas OR botulisme).

Search strings were applied to the Canadian Newsstream and Canadian Business & Current Affairs databases, accessible through the University of Guelph, on November 12, 2023. Data were excluded if they did not meet the eligibility criteria, which were as follows: the news article must (1) communicate about a multijurisdictional enteric illness outbreak (eg, original outbreak notice, updates about outbreaks, or food recall notice) in Canada; (2) be published between January 1, 2014, and November 12, 2023, referencing a multijurisdictional enteric illness outbreak for which there is an available PHN on the

XSL•FO
RenderX

Government of Canada website [35]; (3) be published by at least 1 Canadian news media outlet; (4) be written in English or French; and (5) have been originally communicated in written format. A start date of 2014 was selected as, at the time of writing, PHNs published before 2014 were archived and not publicly accessible. These eligibility criteria and search terms were also applied to a targeted Google search of randomly selected news media outlets (Multimedia Appendix 1). Retrieved results were exported and uploaded to Covidence (Veritas Health Innovation) [36], where articles were deduplicated and underwent title and abstract and full-text screening by 2 researchers (HS and VP).

At the time of data collection, a unique ID, a numerical identifier corresponding to the relevant PHN (Multimedia Appendix 2), and the geographical scope of the news media outlet were assigned. Local and regional news media were distinguished by their coverage areas, with local outlets providing information to specific neighborhoods or towns, while regional outlets encompassed broader regions, such as provinces or territories.

Content Analysis

The content analysis used the health belief model (HBM) [37] as a theoretical framework and followed a previously described established methodology for qualitative content analysis [38]. The codebook (Multimedia Appendix 3) was comprised of 3 sections. Section 1 collected information about the general characteristics of the news article, including the language of communication, type of article, and intention of the article.

Section 2 assessed the inclusion of outbreak information in news media articles to evaluate the accuracy and consistency of information communicated between PHNs and news media articles (recall of item, location of the outbreak, morbidity and mortality, health protective behaviors, and symptoms). In our study, consistency refers to the regular reporting of outbreak characteristics. This includes always reporting the number of illnesses, or explicitly stating the source when this information is available. Accuracy, which is conceptualized as the reporting of the same information indicated in PHAC's PHNs, includes indicating the correct number of hospitalizations. Consistency and accuracy were determined by tracing the posting date of news media articles and comparing the information with that presented in PHNs of the same period.

Section 3 included constructs from the HBM, a theoretical framework that posits an individual's motivation to implement a behavior is influenced by their perception of the severity and susceptibility of a given outcome [37,39]. These constructs include (1) cues to action, which describe any stimuli that promotes the implementation of a recommended behavior, (2) self-efficacy, which refers to the level of confidence an individual has in performing the recommended behavior, (3) perceived susceptibility, which refers to an individuals' subjective perception of the risk of acquiring an illness or disease, (4) perceived severity, which refers to an individuals' perception of the seriousness of the illness or disease, (5) perceived benefits, which refer to an individual's perception of the effectiveness of the behaviors intended to reduce the threat of an illness or disease, and (6) perceived barriers, which refer to the individual's perception of the obstacles to performing the recommended behaviors. In this study, self-efficacy was conceptualized as communicating about, or encouraging the public, to perform the recommended behaviors outlined in a news media article.

Data Analysis

A 10% (77/763) sample of news media articles was independently coded by 2 researchers (HS and VP), and the Cohen κ statistic was calculated using the *irr* package in R Studio v.4.2.3 (Posit) to assess interrater reliability. A value of 10% was chosen by the research teams to ensure that a sufficient subsample was analyzed. The statistical analysis indicated almost perfect agreement between the 2 coders (κ=0.9) [40]. The 2 researchers resolved disagreements in coding, and no significant changes were made to the codebook. The remaining articles were evenly split and coded by HS (n=343) and VP (n=343). All coding was performed using Microsoft Excel.

Total and stratified descriptive statistics were calculated. We examined differences in general characteristics of unique original news media articles by geographic scope (national [n=97], regional [n=123], and local [n=100]) and by source of outbreak (foodborne [n=286], zoonotic [n=7], and unknown [n=27]). In addition, we examined differences in the inclusion of HBM constructs in original news media articles providing at least 1 behavioral recommendation (n=243) by geographic scope (national [n=68], regional [n=94], and local [n=81]) and by source of outbreak (foodborne [n=228], zoonotic [n=4], and unknown [n=11]). Chi-square tests were used to test for differences in communication variables. A significance level of .05, a conventionally accepted threshold, was used to determine statistical significance. All analyses were conducted and visualized using RStudio base packages, *tidyr* [41] and *ggplot2* [42].

Ethical Considerations

Ethics approval was not required. This study used publicly available data from news media articles indexed in the Canadian Newsstream and Canadian Business & Current Affairs databases.

Results

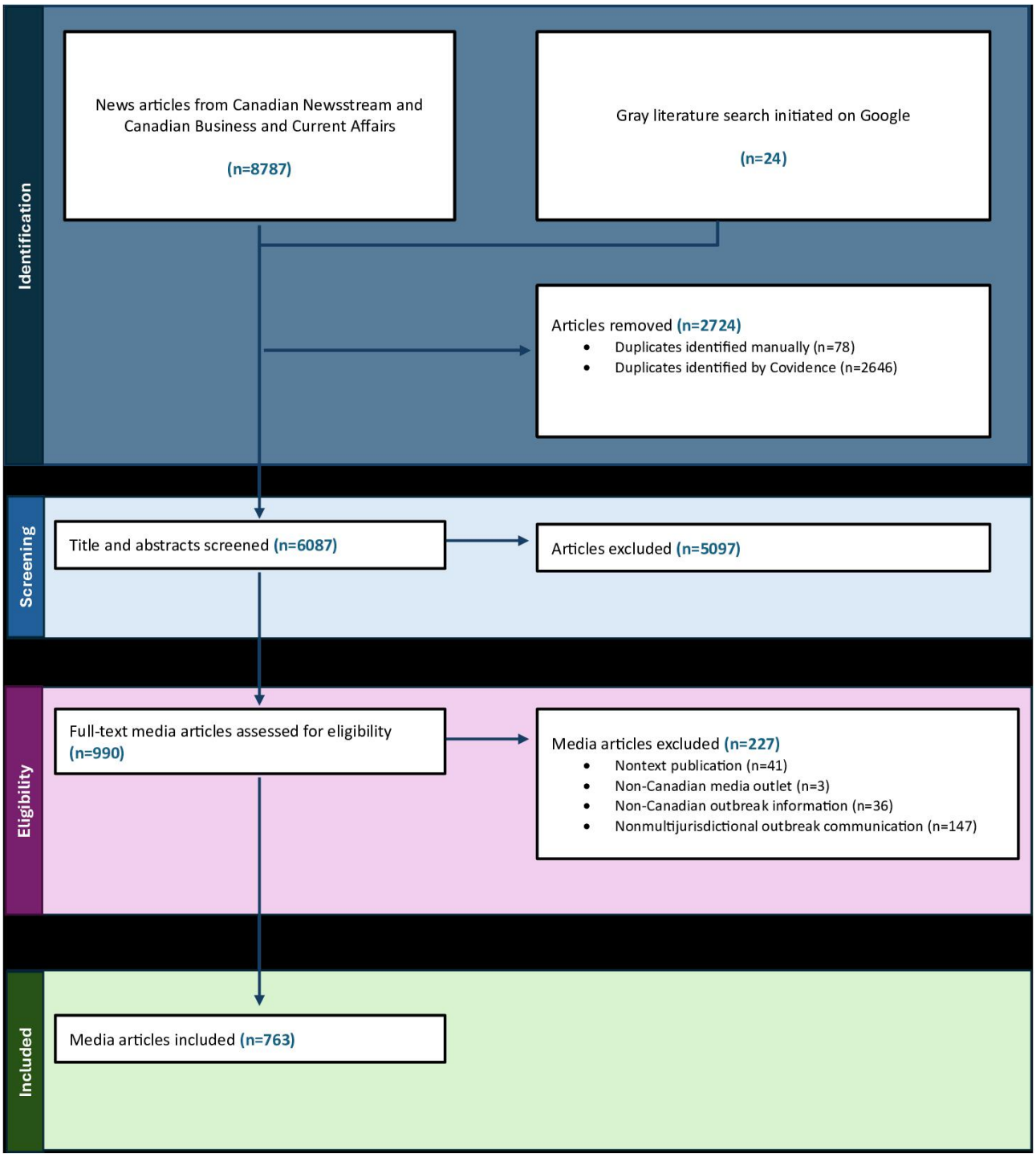
Dissemination of PHN Information by News Media

Dataset Creation

Citation searches identified 8787 news media articles from news databases (Figure 1). After removal of duplicate citations (n=2724), 6063 database articles and 24 gray literature articles underwent title and abstract screening. After screening, 5097 news media articles were excluded. Full-text news media articles (n=990) were then assessed for their eligibility, and 227 articles were excluded because they did not communicate about a multijurisdictional enteric illness outbreak (n=147), were nontext publications (n=41), did not communicate about Canadian outbreaks (n=36), or were issued by a non-Canadian news media outlet (n=3). This process yielded 763 unique articles related to 46 PHNs.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) diagram depicting the systematic retrieval of news media articles from databases and a targeted Google search.

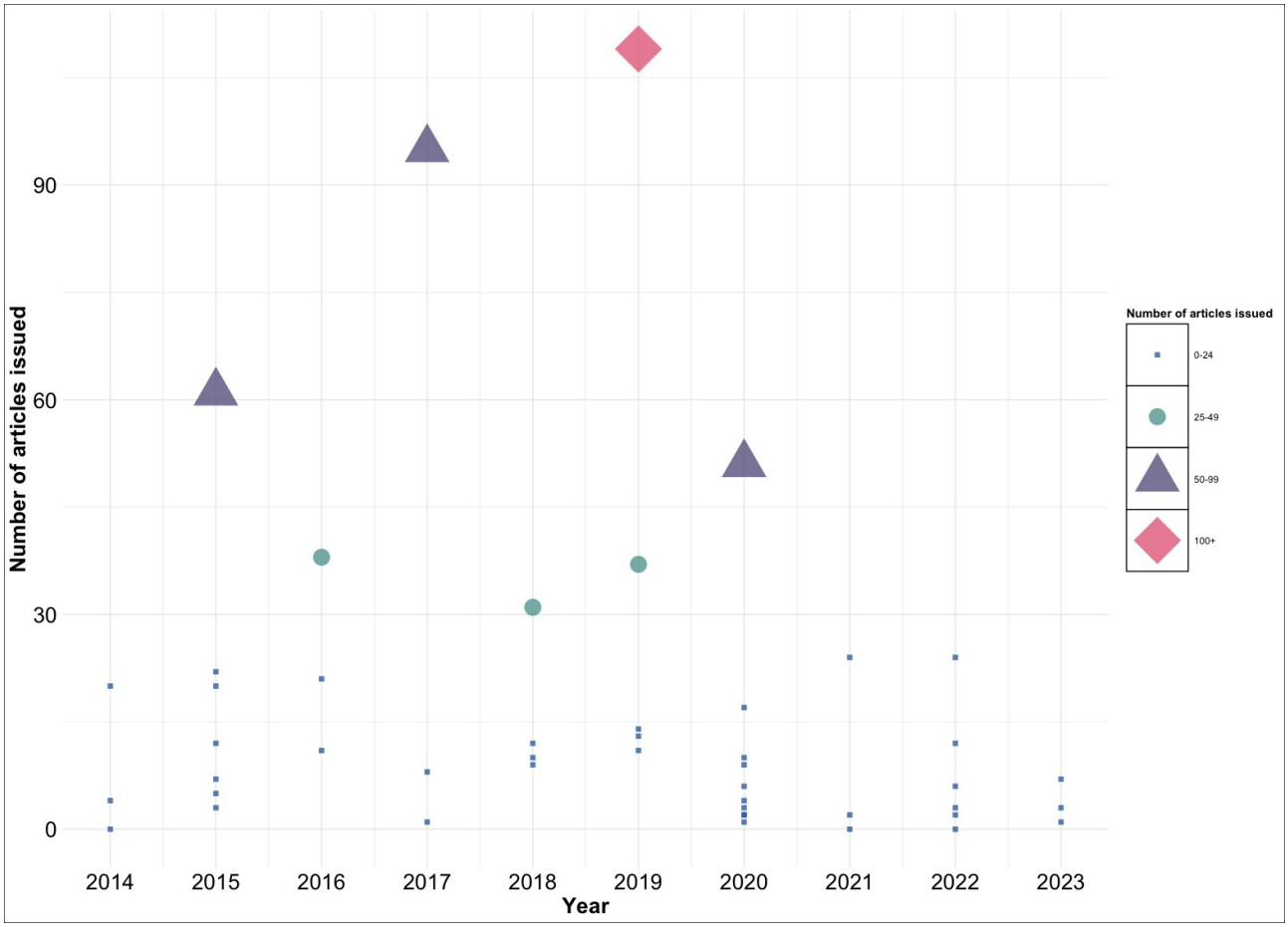


Extent of Dissemination per PHN

News media communicated about almost all PHNs (44/46, 96%) by issuing at least 1 article per PHN (Figure 2). On average, news media issued 16.6 (SD 22.7) articles (0 - 109 articles) per PHN. The most disseminated PHNs by news media were the 2017 - 2019 *Salmonella* outbreaks involving frozen raw breaded chicken products (109/763, 14.3%) and the 2017 *Escherichia coli* (*E. coli*) O121 PHN involving flour (95/763, 12.5%). The remaining PHNs were disseminated 0 - 65 times each, with the majority of PHNs being disseminated up to 10 times each (Multimedia Appendix 4).

XSL•FO
RenderX

Figure 2. Bubble chart visualizing the number of news media articles issued per public health notice. Each bubble represents 1 public health notice. The size and color of the bubble correlates to the number of posts issued for the public health notice.



Content Analysis

Dataset Creation

Content analysis of original news media articles further excluded articles that shared official government and industry statements verbatim (n=208; PHNs [n=127]; food recall notices [n=73]; industry statements [n=6]; other Government of Canada statements [n=2]), yielding 555 articles. In addition, duplicates were removed (n=235), yielding 320 unique original news media articles for content analysis. Duplicates included the same news media article published by multiple outlets or in both English and French. Where relevant, the denominator representing the number of articles was further adjusted to support valid interpretations and is indicated in the text.

Communication Approaches

Original news media articles were most often issued by regional news media (123/320, 38.4%), followed by local news media (100/320, 31.3%), and national news media (97/320, 30.3%) (Table 1). Considering both original news media articles (n=320) and resharing of verbatim statements (n=208), 60.6% (320/528) of articles were original news media articles and 24.1% (127/528) shared PHNs verbatim. For articles published about a PHN with an associated food recall (n=378), 19.3% (73/378) shared food recall notices verbatim. To assess language distribution, we considered both original news media articles and duplicates to avoid excluding French communications (n=555). English was the predominant language used (533/555, 96%) compared with French (22/555, 4%). The language of news media articles significantly differed by the scope of news media coverage (Multimedia Appendix 5): local news media produced only 1 article (1/239, 0.4%) in French and regional news media only communicated in English (188/188, 100%; $P<.001$).

Most original news media articles were regarding foodborne outbreaks (286/320, 89.3%) (Table 2). Other news media articles communicated about zoonotic outbreaks (7/320, 2.2%) and outbreaks where the source was unknown (27/320, 8.4%). Original news media articles often notified the public of an ongoing food recall (121/320, 37.8%). Others were intended to provide an initial notice about an outbreak (107/320, 33.4%). News media articles notifying the public about the end of an outbreak were infrequent (12/320, 3.8%). For original news media articles published regarding an outbreak with an associated recall (n=204), news media consistently mentioned the ongoing recall of products (194/204, 95.1%).

XSL•FO
RenderX

Table . Descriptive statistics of variables assessing the communication methods by news media.

Variable and variable level		Value, n (%)
Geographic scope of news media article (n=320)		
	Regional	123 (38.4)
	Local	100 (31.3)
	National	97 (30.3)
Type of news media article (n=528; 320+208 verbatim articles)		
	Original article by news media outlet	320 (60.6)
	Shares PHN[a] verbatim	127 (24.1)
	Shares CFIA's[b] food recall notices verbatim (n=378 articles for PHNs with an associated recall)	73 (19.3)
	Shares official industry communication verbatim (eg, Safety Notice by Producer)	6[c]
	Shares other official communication verbatim (eg, statement by the Chief Public Health Officer obf Canada)	2[c]
Language of news media article (n=555; 320+235 duplicates)		
	English	533 (96)
	French	22 (4)

[a]PHN: public health notice.

[b]CFIA: Canadian Food Inspection Agency.

[c]Percentage value (%) not applicable.

Table . Descriptive statistics of variables assessing the general characteristics of news media articles.

Variable and variable level		Value, n (%)
Distribution of news media articles by source type (n=320)		
	Foodborne	286 (89.3)
	Unknown	27 (8.4)
	Zoonotic	7 (2.2)
Intention of news media article (n=320)		
	Food recall notice	121 (37.8)
	Notice of outbreak or advisory	107 (33.4)
	Update of ongoing outbreak	30 (9.4)
	Unclear	30 (9.4)
	Reminder or best practice	19 (5.9)
	End of outbreak or advisory	12 (3.8)
	Other	1 (0.3)
Indication of item recall (n=204; 378 articles for PHNs[a] with an associated recall – 159 shared articles – 15 articles where recall issued after article was published)		
	Yes	194 (95.1)
	No	10 (4.9)

[a]PHN: public health notice.

XSL•FO
RenderX

Inclusion of Outbreak Characteristics and Preventive Behavioral Recommendations

Considering news media articles where the associated PHN had relevant information regarding outbreak characteristics and behavioral recommendations, news media articles commonly included the number of cases (237/319, 74.3%), and to a lesser extent, the location of the outbreak (181/316, 57.3%), number of hospitalizations (144/277, 52%), and number of deaths (114/260, 43.8%) (Multimedia Appendix 5). Local news media less frequently communicated about several outbreak characteristics, including the location of the outbreak (48/100, 48%; *P*=.04), number of hospitalizations (33/94, 35%; *P*<.001), and number of deaths (24/77, 31%; *P*=.03) compared with regional and national news media.

News media articles often included preventive behavior information (243/320, 75.9%) and symptoms of illness (201/320, 62.8%) (Table 3). For news media articles that included at least 1 behavioral recommendation (n=243), behavioral recommendations if an individual becomes ill (eg, visiting the doctor if you become ill) were not commonly communicated (62/243, 25.5%). Local news media more frequently provided behavioral recommendations for what to do if an individual becomes ill (29/81, 36%; *P*=.01) (Multimedia Appendix 5).

Table . Descriptive statistics of variables assessing how news media promotes behavioral uptake.

Variable and variable level	Value, n (%)
Indication of symptoms (n=320)	
Yes	201 (62.8)
No	119 (37.2)
Provision of behavioral recommendations to prevent illness (n=320)	
Yes	243 (75.9)
No	77 (24.1)
Provision of behavioral recommendations if individual becomes ill (n=243)	
Yes	62 (25.5)
No	181 (74.6)

Inclusion of HBM Constructs in News Media Articles With Behavioral Recommendations

Considering news media articles with at least 1 behavioral recommendation provided (n=243), cues to action (202/243, 83.1%) and self-efficacy constructs (191/243, 78.6%) were often present. National news media articles less frequently included the self-efficacy construct (44/68, 65%) compared with local (68/81, 84%) and regional news media (79/94, 84%; *P*=.004). Likewise, news media articles about zoonotic outbreaks never contained the self-efficacy construct, whereas articles about foodborne outbreaks frequently included this construct (187/228, 82%; *P*<.001). News media articles often mentioned information relevant to the susceptibility (190/243, 78.2%) and severity of illness (161/243, 66.3%). However, they less commonly included information pertaining to the benefits of implementing the recommended behaviors (50/243, 20.6%), and seldom addressed the potential barriers an individual may encounter (15/243, 6.2%).

Discussion

Principal Findings

Effective communication during enteric illness outbreaks is important for disease control. Coordination of communication efforts and repetition of information across sources are important for increasing the public's receptiveness to the message. To describe the dissemination of federal risk communication during multijurisdictional enteric illness outbreaks by news media, we performed a content analysis of news media articles issued by local, regional, and national Canadian outlets over a 9-year period. The media plays an important role in risk management as an arms-length source of information about government decisions and responses to a health event by contributing to the public's awareness of hazards and influencing their behavior [43]. The ability to effectively communicate timely, accurate, and clear information to the media is a key responsibility of public health, especially during an emergency [44]. An evaluation of news media communication during multijurisdictional enteric illness outbreaks provides a snapshot of how news media disseminate federal risk communication during an outbreak. It also identifies key areas of opportunity to improve this communication pathway.

Differences in News Media Dissemination of PHNs and the Importance of Coordinated Risk Communication

Effective dissemination of critical health information by news media can enhance message reach and visibility to the public and promote adherence to recommended behaviors. During the COVID-19 pandemic, Canadians frequently sought information from online sources that were not government or public health web pages [45]. They also perceived news media to be equally credible to national government and public health websites [45], indicating that news media are trusted sources that can effectively promote the uptake of protective behaviors.

Our results indicate that local, regional, and national news media communicated about almost all PHNs at least once, primarily in English. While the scope of media sources in our study was predominantly regional, outbreak communication by national sources included dissemination by wire services, highlighting

XSL•FO
RenderX

the wide reach of information to diverse audiences across the country. As expected, most news media articles were about foodborne outbreaks. However, dissemination per PHN varied, with 2 PHNs being disseminated to a greater extent than others. The PHN with the highest media coverage, with 109 news media articles, occurred between 2017 and 2019 and was associated with raw chicken and frozen raw breaded chicken products [46]. The second most disseminated PHN, with 95 news media articles, was the 2017 *E. coli* O121 PHN associated with flour [47]. This outbreak prompted a significant nationwide recall directed at consumers, restaurants, and retailers. The coverage of these high-risk outbreaks is expected and in alignment with other studies showing that news media coverage increases during significant public health events, such as large food recalls [48]. In addition, the 2017 - 2019 PHN associated with frozen raw breaded chicken products covered 18 individual outbreaks; however, they were treated as a single multijurisdictional outbreak due to the issuance of only 1 PHN. This may have overlooked distinctions in the extent of dissemination between individual outbreaks.

Ensuring that all multijurisdictional enteric outbreaks are disseminated is important for increasing the reach and visibility of protective information. A qualitative evaluation of the Centers for Disease Control and Prevention's risk communication methods during multistate foodborne outbreaks revealed that consumers want to be notified about all outbreaks, regardless of their direct relevance [49]. Journalists have identified information dissemination as their primary role during food incidents [50], underscoring the importance of close collaboration between public health officials and news media to ensure regular and consistent communication of information to the public.

Importance of Consistent, Accurate, and Timely Communication in Dissemination of PHNs by News Media

When faced with uncertainty, individuals seek information from various sources, including news media [51], highlighting the need for accurate and consistent information. A key component of effective risk communication is the communication of accurate and consistent information by stakeholders [52], as this enables the public to understand the risks and implications of the outbreak, facilitating informed decision-making and increasing the uptake of protective behaviors [53].

Although traditional media sources generally overreport death tolls [54,55], our study found a significant difference in reporting across local, regional, and national news media, with a general trend toward underreporting case fatality numbers. While sensationalized reporting often contributes to public fear and erosion of trust [56,57], Zhang et al [58] found an increased implementation of protective actions during the H1N1 flu outbreaks when individuals were exposed to media coverage that incited fear. Nevertheless, reporting of death tolls is important for providing the public with information necessary to make informed health decisions. Inaccurate reporting can heighten public fear [51] and undermine trust in risk communication [59]. In some cases, downplaying health risks resulted in reduced public urgency to implement health protective behaviors [60].

News media often communicated about ongoing food recalls associated with an enteric illness outbreak, with some news media articles directly sharing food recall notices shared by CFIA. Often, CFIA's food recall notices directed readers to visit PHNs to learn more about the outbreak. This enables the public to avoid contaminated food and protect themselves against illness, an important goal during outbreaks. This study noted frequent communication at the onset of multijurisdictional enteric illness outbreaks; however, communication about the conclusion of an outbreak was less common. This may impact the public's access to timely information. Outbreaks often evolve rapidly, posing potential challenges for news media to access up-to-date information. However, timely risk communication is important for ensuring that the public is aware of how outbreaks evolve over time [61].

Integration of HBM Constructs When Communicating About Multijurisdictional Enteric Illness Outbreaks

News media can importantly influence the public's perception of the outbreak [43]. Risk perception is influenced by threat severity; that is, how susceptible an individual feels to the threat or hazard, and the perceived efficacy of protecting themselves [53]. Models, such as the HBM, provide guidance for health message development that can influence risk perception and encourage the uptake of risk-protective behaviors [62]. To assess how news media disseminate information that helps the public to safeguard their health, we examined the inclusion of HBM constructs. These constructs provide an assessment of whether the public can act on the information being communicated, aligning with the core objective of food safety risk communication, which promotes informed decision-making and uptake of protective behaviors [7].

In this study, news media consistently incorporated the cues to action construct when disseminating PHAC's public communication during multijurisdictional enteric illness outbreaks. Cues to action can increase behavioral acceptance [63], and news media regularly provided recommendations during an outbreak and after its conclusion. These included advising the public to check their shelves for the presence of recalled items. In addition, news media frequently mentioned symptoms associated with the illness. While news media regularly provided cues to action such as behavioral recommendations for individuals to prevent the onset of illness, our study noted the infrequent provision of behavioral recommendations for ill persons (eg, encouraging ill individuals to visit the doctor) by national news media and when communicating about zoonotic outbreaks. Cues to action are important for individuals to develop readiness to act in response to a perceived threat [64] and should be provided to enhance risk perception.

Self-efficacy is particularly crucial in food safety communication, as this construct promotes safe food handling intentions [65]. In this study, self-efficacy was conceptualized as communication about or encouraging the public to perform the recommended behaviors using the health protective behaviors outlined in news media articles. For instance, when

XSL•FO
RenderX

directing consumers to wash lettuce to prevent *E. coli* contamination, news media often provided specific recommendations, including the length of time and how each leaf should be washed. While the self-efficacy construct was generally present in news media articles, our study noted that articles by national news media conveyed this construct less frequently compared to local and regional sources. In addition, reporting of this construct was absent in articles regarding zoonotic outbreaks. It is important that news media articles integrate the self-efficacy construct, as this construct influences the public's confidence in implementing the recommended behaviors [66].

Susceptibility and severity constructs were also frequently present in news media articles. These constructs directly influence risk perception and are crucial factors driving the uptake of behavioral recommendations [67]. News media often communicated about the susceptibility of high-risk groups, such as pregnant women, for contracting the illness. Similarly, news media provided information about the fatality of some enteric pathogens, encouraging individuals to consider their vulnerability to the illness and the seriousness and consequences of the outbreak.

Benefits and barrier constructs, observed less frequently in news media articles, are thought to be the strongest predictors of behavioral change [68]. News media frequently reported the susceptibility and severity of illness, while underreporting the benefits and barriers an individual may face when performing the recommended behaviors. News media is more inclined to sensationalize risk information by overemphasizing risks and negative consequences [69], impeding risk perception [55,70]. In food safety, it is vital to communicate about the risks, benefits, and barriers as this facilitates informed food choices [7]. Although this study did not assess federal risk communication materials, it is important to note that HBM constructs may not be present in official federal communication products such as PHNs. It is reasonable to expect that news media will reflect the information presented in PHNs. In addition, other factors, such as agenda-setting priorities, may influence what and how information is presented to the audience. For instance, news media may emphasize the severity of an outbreak as opposed to the benefits of implementing a protective behavior, as this can lead to increased audience engagement [71]. However, federal partners and news media may benefit from integrating all HBM constructs in future risk communication efforts by highlighting the ease of implementing recommended behaviours, emphasizing their positive impact on health, and addressing common concerns that may act as barriers to promoting health behavior.

When constructing risk communication messages, using the HBM constructs in combination may be effective in promoting behavior change [62,68,72,73]. For instance, risk communication could include stimuli to cue behavior change, instill confidence to adopt behavioral recommendations, and provide information about the susceptibility and severity of illness, and benefits and barriers an individual may encounter when performing the recommended behavior.

Limitations

First, the data collection process involved running search strings on 2 databases accessed through the University of Guelph. While these databases cover a range of news outlets, not all Canadian news media may be indexed. Similarly, the gray literature search was limited to a subset of Canadian media outlets; this may have not captured all news media articles on multijurisdictional enteric illness outbreaks. Therefore, dissemination may be higher than observed in this study. In addition, the search strings used in this study, while comprehensive, may not have captured all news media articles related to multijurisdictional enteric illness outbreaks.

Second, this study assessed the presence of HBM constructs in articles issued by news media. However, this information may not be reflected in PHAC's communication products and, as a result, may be absent from news media articles. Comparing these constructs between the 2 sources was outside this study's scope. Future research should build on our findings and seek to better understand the inclusion of HBM constructs in PHNs and subsequent news media articles.

Conclusions

This study explored the dissemination of federal risk communication by news media during multijurisdictional enteric illness outbreaks. Our findings suggest news media disseminated information on almost all multijurisdictional enteric illness outbreaks; however, there are opportunities for improving how federal risk communication is disseminated by news media. Federal organizations can improve future risk communication of multijurisdictional enteric illness outbreaks by news media by maintaining and strengthening interorganizational connections and ensuring the information quality of PHNs as a key information source for news media. Optimized communication can help to ensure the public has the information necessary to implement protective behaviors.

Acknowledgments

Funding was provided by the Public Health Agency of Canada, in part by the Federal Student Work Experience Program.

Authors' Contributions

HS contributed to methodology, investigation, formal analysis, and writing – original draft. LEG contributed to conceptualization, methodology, writing – review & editing, supervision, and project administration. VP contributed to formal analysis and writing – review & editing. MM contributed to methodology and writing – review & editing. AP contributed to supervision and writing – review & editing. LC contributed to conceptualization, methodology, and writing – review & editing. MP contributed to

XSL•FO
RenderX

conceptualization, methodology, writing – review & editing, and funding acquisition. JEM contributed to conceptualization, methodology, writing – review & editing, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results of gray literature searches of 4 Canadian news media outlets.
[DOCX File, 36 KB - publichealth_v11i1e68724_app1.docx]

Multimedia Appendix 2
Multijurisdictional enteric illness outbreaks in Canada from 2014 to 2023.
[DOCX File, 31 KB - publichealth_v11i1e68724_app2.docx]

Multimedia Appendix 3
Codebook used for media content analysis.
[DOCX File, 48 KB - publichealth_v11i1e68724_app3.docx]

Multimedia Appendix 4
Distribution of multijurisdictional enteric illness outbreaks by number of news media articles issued.
[DOCX File, 24 KB - publichealth_v11i1e68724_app4.docx]

Multimedia Appendix 5
Descriptive statistics of communication variables stratified by scope of news media coverage or outbreak source.
[DOCX File, 53 KB - publichealth_v11i1e68724_app5.docx]

References

1. Weaver JB, Weaver SS, DiClemente RJ. Risk communication. In: Heggenhougen HKK, editor. International Encyclopedia of Public Health: Academic Press; 2008. [doi: 10.1016/B978-012373960-5.00582-7]
2. Winters M, Jalloh MF, Sengeh P, et al. Risk communication and Ebola-specific knowledge and behavior during 2014-2015 outbreak, Sierra Leone. Emerg Infect Dis 2018 Feb;24(2):336-344. [doi: 10.3201/eid2402.171028] [Medline: 29350151]
3. Wang X, Lin L, Xuan Z, Xu J, Wan Y, Zhou X. Risk communication on behavioral responses during COVID-19 among general population in China: a rapid national study. J Infect 2020 Dec;81(6):911-922. [doi: 10.1016/j.jinf.2020.10.031] [Medline: 33144190]
4. Infanti J, Sixsmith J, Barry M, Núñez-Córdoba J, Oroviogoicoechea-Ortega C, Guillén-Grima F. A literature review on effective risk communication for the prevention and control of communicable diseases in Europe. : European Centre for Disease Prevention and Control (ECDC); 2013 URL: https://www.ecdc.europa.eu/sites/default/files/media/en/publications/Publications/risk-communication-literary-review-jan-2013.pdf [accessed 2025-04-03]
5. Xiang D, Kontos C, Veloudaki A, Baka A, Karnaki P, Linos A. Risk communication in times of an epidemic or pandemic. : Zadig S.r.l; 2017 URL: https://www.asset-scienceinsociety.eu/sites/default/files/paper_series_5_-_risk_communication_in_time_of_an_epidemic_or_pandemic.pdf [accessed 2025-04-03]
6. When food is cooking up a storm: proven recipes for risk communications. Publications Office of the European Union. 2017. URL: https://op.europa.eu/en/publication-detail/-/publication/654b67b4-57bf-11e7-a5ca-01aa75ed71a1/ [accessed 2025-04-03]
7. Risk Communication Applied to Food Safety: Handbook: World Health Organization and Food and Agriculture Organization of the United Nations; 2016. URL: https://iris.who.int/handle/10665/250083 [accessed 2025-04-03]
8. Thomas MK, Murray R, Flockhart L, et al. Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. Foodborne Pathog Dis 2013 Jul;10(7):639-648. [doi: 10.1089/fpd.2012.1389] [Medline: 23659355]
9. Thomas MK, Murray R, Nesbitt A, Pollari F. The incidence of acute gastrointestinal illness in Canada, Foodbook Survey 2014-2015. Can J Infect Dis Med Microbiol 2017;2017:5956148. [doi: 10.1155/2017/5956148] [Medline: 29410684]
10. Butler AJ, Thomas MK, Pintar KDM. Expert elicitation as a means to attribute 28 enteric pathogens to foodborne, waterborne, animal contact, and person-to-person transmission routes in Canada. Foodborne Pathog Dis 2015 Apr;12(4):335-344. [doi: 10.1089/fpd.2014.1856] [Medline: 25835810]
11. Butler AJ, Pintar KDM, Thomas MK. Estimating the relative role of various subcategories of food, water, and animal contact transmission of 28 enteric diseases in Canada. Foodborne Pathog Dis 2016 Feb;13(2):57-64. [doi: 10.1089/fpd.2015.1957] [Medline: 26863428]

XSL•FO
RenderX

12. Hurst M, Nesbitt A, Kadykalo S, Dougherty B, Arango-Sabogal JC, Ravel A. Attributing salmonellosis cases to foodborne, animal contact and waterborne routes using the microbial subtyping approach and exposure weights. Food Control 2023 Jun;148:109636. [doi: 10.1016/j.foodcont.2023.109636]
13. Delahoy MJ, Shah HJ, Weller DL, et al. Preliminary incidence and trends of infections caused by pathogens transmitted commonly through food - Foodborne Diseases Active Surveillance Network, 10 U.S. Sites, 2022. MMWR Morb Mortal Wkly Rep 2023 Jun 30;72(26):701-706. [doi: 10.15585/mmwr.mm7226a1] [Medline: 37384552]
14. Ahn JW, Scallan Walter E, White AE, McQueen RB, Hoffmann S. Identifying sepsis from foodborne hospitalization: incidence and hospitalization cost by pathogen. Clin Infect Dis 2022 Sep 14;75(5):857-866. [doi: 10.1093/cid/ciab1045] [Medline: 34950950]
15. Thomas MK, Murray R, Flockhart L, et al. Estimates of foodborne illness-related hospitalizations and deaths in Canada for 30 specified pathogens and unspecified agents. Foodborne Pathog Dis 2015 Oct;12(10):820-827. [doi: 10.1089/fpd.2015.1966] [Medline: 26259128]
16. Morton VK, Thomas MK, McEwen SA. Estimated hospitalizations attributed to norovirus and rotavirus infection in Canada, 2006-2010. Epidemiol Infect 2015 Dec;143(16):3528-3537. [doi: 10.1017/S0950268815000734] [Medline: 25991407]
17. Majowicz SE, Panagiotoglou D, Taylor M, et al. Determining the long-term health burden and risk of sequelae for 14 foodborne infections in British Columbia, Canada: protocol for a retrospective population-based cohort study. BMJ Open 2020 Aug 31;10(8):e036560. [doi: 10.1136/bmjopen-2019-036560] [Medline: 32868357]
18. Gohari MR, Taylor M, MacKinnon MC, et al. Patterns of enteric infections in a population-wide cohort study of sequelae, British Columbia, Canada. Epidemiol Infect 2022 Dec 14;151:e7. [doi: 10.1017/S0950268822001911] [Medline: 36515015]
19. Scallan Walter EJ, Crim SM, Bruce BB, Griffin PM. Postinfectious irritable bowel syndrome after Campylobacter infection. Am J Gastroenterol 2019 Oct;114(10):1649-1656. [doi: 10.14309/ajg.0000000000000408] [Medline: 31567167]
20. Scallan Walter EJ, Crim SM, Bruce BB, Griffin PM. Incidence of Campylobacter-associated Guillain-Barré syndrome estimated from Health Insurance Data. Foodborne Pathog Dis 2020 Jan;17(1):23-28. [doi: 10.1089/fpd.2019.2652] [Medline: 31509036]
21. Patel V, Grant LE, Shereefdeen H, et al. Evaluating multi-jurisdictional enteric illness outbreak messaging in Canada: a content analysis of Public Health Notices. Health Commun 2024 Aug 22:1-15. [doi: 10.1080/10410236.2024.2391207] [Medline: 39169881]
22. Patel V, Grant LE, Shereefdeen H, et al. A content analysis of government-issued social media posts during multi-jurisdictional enteric illness outbreaks in Canada. Front Commun 2024;9. [doi: 10.3389/fcomm.2024.1512014]
23. Mucević H, McWhirter JE, Macleod J, et al. Best practices in public risk communication during enteric illness outbreak investigations: A scoping review. Front Commun 2025;9. [doi: 10.3389/fcomm.2024.1509940]
24. Canada's Foodborne Illness Outbreak Response Protocol (FIORP): a guide to multi-jurisdictional enteric outbreak response. : Public Health Agency of Canada; 2017 URL: https://www.canada.ca/content/dam/phac-aspc/documents/services/publications/health-risks-safety/64-02-17-1879-FIORP-2015-EN-04.pdf [accessed 2025-04-03]
25. Wakefield MA, Loken B, Hornik RC. Use of mass media campaigns to change health behaviour. Lancet 2010 Oct 9;376(9748):1261-1271. [doi: 10.1016/S0140-6736(10)60809-4] [Medline: 20933263]
26. Pieri E. Media framing and the threat of global Pandemics: the Ebola crisis in UK media and policy response. Sociol Res Online 2019 Mar;24(1):73-92. [doi: 10.1177/1360780418811966] [Medline: 34285639]
27. Pinto R, Lacerda J, Silva L, et al. Text mining analysis to understand the impact of online news on public health response: case of syphilis epidemic in Brazil. Front Public Health 2023;11:1248121. [doi: 10.3389/fpubh.2023.1248121] [Medline: 38026344]
28. How have Canadians been using the internet during the COVID-19 pandemic. Statistics Canada. URL: https://www.statcan.gc.ca/o1/en/plus/198-how-have-canadians-been-using-internet-during-covid-19-pandemic [accessed 2025-04-03]
29. Guttmann A. Daily time spent with selected media among adults in Canada as of 3rd quarter 2023 (in hours.minutes). Statista. URL: https://www.statista.com/statistics/237478/daily-time-spent-with-media-among-adults-in-canada/ [accessed 2025-03-04]
30. Watson A. Trust in news sources in Canada as of November 2023, by platform. Statista. URL: https://www.statista.com/statistics/1426915/trust-news-source-platform-canada/ [accessed 2025-04-03]
31. Watson A. Weekly print and digital newspaper reach in Canada from March 2019 to November 2023, by generation. Statista. URL: https://www.statista.com/statistics/261801/newspaper-reach-in-canada-by-age/ [accessed 2025-04-03]
32. Happer C, Philo G. The role of the media in the construction of public belief and social change. J Soc Polit Psych 2013;1(1):321-336. [doi: 10.5964/jspp.v1i1.96]
33. McCombs M, Valenzuela S. Setting the Agenda: Mass Media and Public Opinion, 3rd edition: Polity Press; 2020.
34. Wang Y, Wu R, Zeng J, Xue P. Research on the public opinion guidance mechanism of major public health incidents. Front Psychol 2022;13:872464. [doi: 10.3389/fpsyg.2022.872464] [Medline: 35846618]
35. Public health notices. Government of Canada. URL: https://www.canada.ca/en/public-health/services/public-health-notices.html [accessed 2025-04-03]
36. Covidence. URL: www.covidence.org [accessed 2025-04-03]

XSL•FO
RenderX

37. Rosenstock IM. Historical origins of the health belief model. Health Educ Monogr 1974 Dec;2(4):328-335. [doi: 10.1177/109019817400200403]
38. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
39. Janz NK, Becker MH. The health belief model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
40. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282. [Medline: 23092060]
41. tidyr: tidy messy data. The Comprehensive R Archive Network. 2024. URL: https://CRAN.R-project.org/package=tidyr [accessed 2025-04-03]
42. ggplot2. 2016. URL: https://ggplot2.tidyverse.org [accessed 2025-04-03]
43. Mheidly N, Fares J. Leveraging media and health communication strategies to overcome the COVID-19 infodemic. J Public Health Policy 2020 Dec;41(4):410-420. [doi: 10.1057/s41271-020-00247-w] [Medline: 32826935]
44. Hyer RN, Covello VT. Effective media communication during public health emergencies. Institutional Repository for Information Sharing (IRIS). 2005. URL: https://tinyurl.com/4ca3bc77 [accessed 2025-04-03]
45. Parsons Leigh J, Fiest K, Brundin-Mather R, et al. A national cross-sectional survey of public perceptions of the COVID-19 pandemic: self-reported beliefs, knowledge, and behaviors. PLoS ONE 2020;15(10):e0241259. [doi: 10.1371/journal.pone.0241259] [Medline: 33095836]
46. Outbreaks of Salmonella infectious linked to raw chicken, including frozen raw breaded chicken products. Government of Canada. URL: https://www.canada.ca/en/public-health/services/public-health-notices/2018/outbreaks-salmonella-infections-linked-raw-chicken-including-frozen-raw-breaded-chicken-products.html [accessed 2025-04-03]
47. Public health notice – outbreak of E.coli infections linked to various flours and flour products. Government of Canada. URL: https://www.canada.ca/en/public-health/services/public-health-notices/2017/public-health-notice-outbreak-e-coli-infections-linked-various-flours-flour-products.html [accessed 2025-04-03]
48. Tiozzo B, Pinto A, Neresini F, et al. Food risk communication: analysis of the media coverage of food risk on Italian online daily newspapers. Qual Quant 2019 Nov;53(6):2843-2866. [doi: 10.1007/s11135-019-00897-3]
49. Ablan M, McFadden K, Jhung M, et al. A qualitative evaluation of the Centers for Disease Control and Prevention risk communication methods during multistate foodborne outbreaks. Food Prot Trends 2021 Nov;41(6):547-554. [Medline: 37200708]
50. Wilson AM, Henderson J, Coveney J, et al. Media actors' perceptions of their roles in reporting food incidents. BMC Public Health 2014 Dec 18;14:1305. [doi: 10.1186/1471-2458-14-1305] [Medline: 25524217]
51. Vermeer S, Kruikemeier S, Trilling D, de Vreese C. Using panel data to study political interest, news media trust, and news media use in the early stages of the COVID-19 pandemic. Journal Stud 2022 Apr 26;23(5-6):740-760. [doi: 10.1080/1461670X.2021.2017790]
52. Coombs WT. Conceptualizing crisis communication. In: Heath RL, O'Hair D, editors. Handbook of Risk and Crisis Communication: Routledge; 2010, Vol. 1.
53. Narimani S, Mozaffari N, Habibi Soola A, et al. Risk communication and risk perception along with its influencing factors in Covid-19 disease: focusing on the Extended Parallel Process Model. Int J Disaster Risk Reduct 2023 Feb 15;86:103547. [doi: 10.1016/j.ijdrr.2023.103547] [Medline: 36713632]
54. Goodall C, Sabo J, Cline R, Egbert N. Threat, efficacy, and uncertainty in the first 5 months of national print and electronic news coverage of the H1N1 virus. J Health Commun 2012;17(3):338-355. [doi: 10.1080/10810730.2011.626499] [Medline: 22188164]
55. Généreux M, David MD, O'Sullivan T, et al. Communication strategies and media discourses in the age of COVID-19: an urgent need for action. Health Promot Int 2021 Aug 30;36(4):1178-1185. [doi: 10.1093/heapro/daaa136] [Medline: 33294917]
56. Hsu YC, Chen YL, Wei HN, Yang YW, Chen YH. Risk and outbreak communication: lessons from Taiwan's experiences in the post-SARS era. Health Secur 2017;15(2):165-169. [doi: 10.1089/hs.2016.0111] [Medline: 28418746]
57. Van Scoy LJ, Snyder B, Miller EL, et al. Public anxiety and distrust due to perceived politicization and media sensationalism during early COVID-19 media messaging. J Commun Healthc 2021 Jul 3;14(3):193-205. [doi: 10.1080/17538068.2021.1953934]
58. Zhang L, Kong Y, Chang H. Media use and health behavior in H1N1 flu crisis: the mediating role of perceived knowledge and fear. Atl J Commun 2015 Mar 15;23(2):67-80. [doi: 10.1080/15456870.2015.1013101]
59. Nahum A, Drekonja DM, Alpern JD. The erosion of public trust and SARS-CoV-2 vaccines- more action is needed. Open Forum Infect Dis 2021 Feb;8(2):ofaa657. [doi: 10.1093/ofid/ofaa657] [Medline: 34141815]
60. Sandell T, Sebar B, Harris N. Framing risk: communication messages in the Australian and Swedish print media surrounding the 2009 H1N1 pandemic. Scand J Public Health 2013 Dec;41(8):860-865. [doi: 10.1177/1403494813498158] [Medline: 23873631]
61. Lundgren RE, McMakin AH. Risk Communication: A Handbook for Communicating Environmental, Safety, and Health Risks, 6th edition: John Wiley & Sons; 2018.

XSL•FO
RenderX

62. Jones CL, Jensen JD, Scherr CL, Brown NR, Christy K, Weaver J. The health belief model as an explanatory framework in communication research: exploring parallel, serial, and moderated mediation. Health Commun 2015;30(6):566-576. [doi: 10.1080/10410236.2013.873363] [Medline: 25010519]
63. Wong MCS, Wong ELY, Huang J, et al. Acceptance of the COVID-19 vaccine based on the health belief model: a population-based survey in Hong Kong. Vaccine (Auckl) 2021 Feb 12;39(7):1148-1156. [doi: 10.1016/j.vaccine.2020.12.083] [Medline: 33461834]
64. Hochbaum GM, United States Public Health Service Division of Special Health Services. Public Participation in Medical Screening Programs; a Socio-Psychological Study: U.S. Dept. of Health, Education, and Welfare, Public Health Service, Bureau of State Services, Division of Special Health Services, Tuberculosis Program; 1958.
65. Chow S, Mullan B. Predicting food hygiene. An investigation of social factors and past behaviour in an extended model of the health action process approach. Appetite 2010 Feb;54(1):126-133. [doi: 10.1016/j.appet.2009.09.018] [Medline: 19815040]
66. Reid AE, Aiken LS. Integration of five health behaviour models: common strengths and unique contributions to understanding condom use. Psychol Health 2011 Nov;26(11):1499-1520. [doi: 10.1080/08870446.2011.572259] [Medline: 21678166]
67. DeDonno MA, Longo J, Levy X, Morris JD. Perceived susceptibility and severity of COVID-19 on prevention practices, early in the pandemic in the State of Florida. J Community Health 2022 Aug;47(4):627-634. [doi: 10.1007/s10900-022-01090-8] [Medline: 35451692]
68. Carpenter CJ. A meta-analysis of the effectiveness of health belief model variables in predicting behavior. Health Commun 2010 Dec;25(8):661-669. [doi: 10.1080/10410236.2010.521906] [Medline: 21153982]
69. van der Meer T, Kroon AC, Vliegenthart R. Do news media kill? How a biased news reality can overshadow real societal risks, the case of aviation and road traffic accidents. Social Forces 2022 Jul 20;101(1):506-530. [doi: 10.1093/sf/soab114]
70. Henderson J, Wilson A, Meyer SB, et al. The role of the media in construction and presentation of food risks. Health Risk Soc 2014 Nov 17;16(7-8):615-630. [doi: 10.1080/13698575.2014.966806]
71. Kim HS, Suh YJ, Kim EM, et al. Fact-checking and audience engagement: a study of content analysis and audience behavioral data of fact-checking coverage from news media. Digital Journalism 2022 May 28;10(5):781-800. [doi: 10.1080/21670811.2021.2006073]
72. Orji R, Vassileva J, Mandryk R. Towards an effective health interventions design: an extension of the health belief model. Online J Public Health Inform 2012;4(3):3. [doi: 10.5210/ojphi.v4i3.4321] [Medline: 23569653]
73. Li ZT, Yang SS, Zhang XX, Fisher EB, Tian BC, Sun XY. Complex relation among health belief model components in TB prevention and care. Public Health (Fairfax) 2015 Jul;129(7):907-915. [doi: 10.1016/j.puhe.2015.04.008] [Medline: 26021833]

Abbreviations

CFIA: Canadian Food Inspection Agency
HBM: health belief model
PHAC: Public Health Agency of Canada
PHN: public health notice

Edited by A Mavragani; submitted 12.11.24; peer-reviewed by A King, R Rosemary; revised version received 27.02.25; accepted 04.03.25; published 10.04.25.

Please cite as:
Shereefdeen H, Grant LE, Patel V, MacKay M, Papadopoulos A, Cheng L, Phypers M, McWhirter JE
Assessing the Dissemination of Federal Risk Communication by News Media Outlets During Enteric Illness Outbreaks: Canadian Content Analysis
JMIR Public Health Surveill 2025;11:e68724
URL: https://publichealth.jmir.org/2025/1/e68724
doi:10.2196/68724

© Hisba Shereefdeen, Lauren Elizabeth Grant, Vayshali Patel, Melissa MacKay, Andrew Papadopoulos, Leslie Cheng, Melissa Phypers, Jennifer Elizabeth McWhirter. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mindfulness Intervention for Health Information Avoidance in Older Adults: Mixed Methods Study

Chenyu Gu[1], PhD; Liquan Qian[2], PhD; Xiaojie Zhuo[2], PhD

[1]School of Journalism and Communication, Minjiang University, Fuzhou, China

[2]School of Arts and Media, Wuhan College, Wuhan, China

Corresponding Author:
Xiaojie Zhuo, PhD
School of Arts and Media
Wuhan College
No. 333 Huangjiahu Avenue
Jiangxia District, Hubei Province
Wuhan, 430212
China
Phone: 86 180 5922 1673
Email: 3074@mju.edu.cn

Abstract

Background: The global aging population and rapid development of digital technology have made health management among older adults an urgent public health issue. The complexity of online health information often leads to psychological challenges, such as cyberchondria, exacerbating health information avoidance behaviors. These behaviors hinder effective health management; yet, little research examines their mechanisms or intervention strategies.

Objective: This study investigates the mechanisms influencing health information avoidance among older adults, emphasizing the mediating role of cyberchondria. In addition, it evaluates the effectiveness of mindfulness meditation as an intervention strategy to mitigate these behaviors.

Methods: A mixed methods approach was used, combining quantitative and qualitative methodologies. Substudy 1 developed a theoretical model based on self-determination theory to explore internal (positive metacognition and health self-efficacy) and external (subjective norms and health information similarity) factors influencing health information avoidance, with cyberchondria as a mediator. A cross-sectional survey (N=236) was conducted to test the proposed model. Substudy 2 involved a 4-week mindfulness meditation intervention (N=94) to assess its impact on reducing health information avoidance behaviors.

Results: Study 1 showed that positive metacognition (β=.26, *P*=.002), health self-efficacy (β=.25, *P*<.001), and health information similarity (β=.29, *P*<.001) significantly predicted health information avoidance among older adults. Cyberchondria mediated these effects: positive metacognition (effect=0.106, 95% CI 0.035-0.189), health self-efficacy (effect=0.103, 95% CI 0.043-0.185), and health information similarity (effect=0.120, 95% CI 0.063-0.191). Subjective norms did not significantly predict health information avoidance (β= .11, *P*=.13), and cyberchondria did not mediate this relationship (effect= 0.045, 95% CI 0.102 to 0.016). Study 2 found that after the 4-week mindfulness intervention, the intervention group (group 1: n=46) exhibited significantly higher mindfulness levels than the control group (group 2: n=48; M_{group1}=4.122, M_{group2}=3.606, *P*<.001) and higher levels compared with preintervention (M_{t2}=4.122, M_{t1}=3.502, *P*<.001, where t1=preintervention and t2=postintervention). However, cyberchondria levels did not change significantly (M_{t1}=2.848, M_{t2}=2.685, *P*=.18). Nevertheless, the results revealed a significant interaction effect between mindfulness and cyberchondria on health information avoidance (effect= 0.357, *P*=.002, 95% CI 0.580 to 0.131), suggesting that mindfulness intervention effectively inhibited the transformation of cyberchondria into health information avoidance behavior.

Conclusions: This study reveals the role of cyberchondria in health information avoidance and validates mindfulness meditation as an effective intervention for mitigating such behaviors. Findings offer practical recommendations for improving digital health information delivery and health management strategies for older adults.

(JMIR Public Health Surveill 2025;11:e69554) doi:10.2196/69554

XSL•FO
RenderX

KEYWORDS

health information avoidance; cyberchondria; self-determination theory; mindfulness; elderly

Introduction

Background

With the global aging population and the rise of digital health information, the health information behavior of older adults has become a critical focus [1]. The adoption of internet and age-friendly technologies has enabled older adults to seek health information online. However, the overwhelming volume and inconsistent quality of information, combined with lower information literacy among older adults, often hinder efficient access to reliable health information, posing challenges to their health management.

Health information avoidance is a growing concern [2]. It involves individuals deliberately avoiding health-related information, often driven by anxiety and uncertainty, which can lead to severe health consequences [3]. Given their greater need for health information, avoidance behaviors in older adults can hinder effective health management and delay timely diagnosis or treatment [4]. However, research on its causes and interventions remains limited. Cyberchondria, characterized by excessive health information searching, is another critical issue. It disrupts daily activities and exacerbates health anxiety, despite being initially aimed at reducing uncertainty [5]. While it has been linked to problems like internet addiction and sleep disorders, its connection to health information avoidance remains underexplored [6,7]. This study posits that cyberchondria amplifies anxiety, which in turn motivates information avoidance as a self-protection mechanism [8]. Viewing cyberchondria as a psychological driver of health information avoidance offers valuable insights into this behavior.

This study introduces mindfulness meditation as a potential intervention. Rooted in Buddhist traditions, mindfulness has been shown to reduce anxiety and improve emotional regulation and cognitive flexibility [9,10]. By applying mindfulness meditation with older adults, this study evaluates its effects on cyberchondria and health information avoidance, providing a cost-effective health management strategy.

In summary, this study aims to explore the psychological mechanisms underlying health information avoidance among older adults, with a particular focus on cyberchondria as a mediating factor. Drawing from self-determination theory, we propose a model that incorporates both internal (positive metacognition and health self-efficacy) and external (subjective norms and health information similarity) factors to explain how they contribute to health information avoidance. Furthermore, we aim to evaluate the effectiveness of mindfulness meditation as a practical intervention to mitigate the impact of cyberchondria on information avoidance behaviors. This study seeks to address the following two research questions:

1. What are the formation pathways of health information avoidance among older adults?
2. Can mindfulness meditation effectively intervene in cyberchondria and reduce health information avoidance behaviors among older adults?

By addressing these questions, this research intends to contribute theoretical and practical insights to enhance older adults' health management capabilities in the digital era.

Theory Background and Hypothesis Development

Health Information Avoidance Among Older Adults

In the context of global digitalization and an aging population, the health information behavior of older adults has become a critical topic in health management [11]. Health information avoidance, the deliberate decision to avoid health-related information, is particularly prevalent among older adults due to lower information literacy, anxiety about health issues, and fear of disease severity [12]. This behavior poses significant risks, such as missed medical advice or delayed disease prevention and treatment, as older adults often experience poorer health than younger individuals [1].

Although avoidance may temporarily reduce stress, it often exacerbates anxiety by failing to address underlying health concerns and increasing uncertainty [13,14]. It also undermines proactive health management, reducing engagement in health-promoting behaviors and increasing the risk of chronic diseases [15]. Despite growing research, key gaps remain, particularly in understanding how external environments interact with psychological factors and in developing targeted interventions for older adults. Cyberchondria, defined as repeated health searches that heighten anxiety, offers a valuable framework for understanding avoidance behaviors [16]. It highlights the uncertainty and negative emotions triggered by health information, which align with the drivers of avoidance. Incorporating this framework can help explain the anxiety-avoidance cycle in older adults and support the development of targeted interventions to break this cycle.

Cyberchondria

Cyberchondria refers to compulsive and excessive searches for health-related information, often intended for self-reassurance. However, these searches usually provide only temporary relief, with anxiety often worsening during and after the process. Despite its negative consequences, such as harm to mental health, overmedicalization, and susceptibility to medical scams, cyberchondria has become a significant public health concern with the rise of mobile internet use [17,18].

Research on cyberchondria highlights 2 main areas. First, contributing factors include low self-esteem, anxiety sensitivity, low tolerance for uncertainty, and high neuroticism [19-21]. Second, its negative impacts include reduced trust in doctors, increased self-treatment behaviors, medical errors, lower quality of life, and wasted health care resources [22,23]. It has also been linked to problematic internet use [24]. Despite these findings, little research has examined the link between cyberchondria and health information avoidance. This study

XSL•FO
RenderX

aims to fill this gap by exploring how cyberchondria contributes to avoidance behaviors, providing new insights into health behaviors [25,26]. Understanding this connection is crucial for designing interventions to help vulnerable populations better manage health information and reduce the negative effects of cyberchondria.

Self-Determination Theory

Self-determination theory (SDT), proposed by Deci and Ryan [27], explains human motivation and behavior through the fulfillment of 3 psychological needs: competence, relatedness, and autonomy. Behavior is influenced by internal motivation, driven by psychological cognition, and external motivation, shaped by external pressures or rewards [28]. This dual perspective has been widely applied in health psychology, effectively explaining behaviors such as the adoption of personal health record systems driven by autonomous traits (internal factors) and compulsive social media use influenced by social interaction (external factors) [29-31].

Although extensively applied, SDT has not been used to explain cyberchondria. By integrating internal motivations and external environments, SDT provides a valuable framework for understanding its development and guiding effective health interventions [32]. Responding to calls for further exploration of cyberchondria's antecedents and consequences [33], this study uses SDT to analyze the factors influencing cyberchondria and its connection to health information avoidance.

Cyberchondria and Health Information Avoidance

Cyberchondria is characterized by compulsive and frequent internet-based searching for health-related information, which intensifies individuals' anxiety and concerns about their health. This behavior often leads to cognitive and emotional responses, such as excessive worry about illness and overinterpretation of symptoms, ultimately impacting health behaviors [34]. While cyberchondria has been linked to behaviors like internet addiction, evidence suggests that internet addiction may eventually result in information avoidance. Individuals with cyberchondria often experience heightened anxiety and negative emotional responses when exposed to health-related information [35]. These negative emotions reinforce their fear of health issues, leading to a hypersensitivity toward disease-related information. Based on protection motivation theory, individuals may perceive excessive health information as a source of emotional distress and adopt avoidance behaviors to reduce anxiety and stress [36]. Furthermore, individuals with cyberchondria often interpret ambiguous or uncertain information as a potential threat. This heightened threat perception can result in strong psychological resistance, causing them to proactively reduce exposure to health information to avoid discomfort. Over time, this behavior manifests as health information avoidance [37]. In light of these mechanisms, the following hypotheses are proposed in this study:

H1: Cyberchondria positively influences the health information avoidance.

Internal Factors: Positive Metacognition and Health Self-Efficacy

Next, we explore the factors that lead to cyberchondria in older adults. According to the SDT mentioned earlier, an individual's behavioral intention is influenced by internal cognitive factors. This study selects positive metacognition as one of the key internal factors influencing cyberchondria. Metacognition is commonly defined as "thinking about thinking" and encompasses the beliefs, strategies, and methods individuals use to regulate their internal cognitive processes [38]. Metacognitive beliefs can be categorized as either positive or negative. Positive metacognition refers to the belief that focusing on health issues and extensively thinking about them enhances one's sense of safety and control. This construct is widely used to explain health anxiety [39].

Research has shown that positive metacognition effectively predicts attentional biases toward health-related information. Individuals with strong positive metacognitive tendencies are more vigilant about their health and more likely to seek health-related information [40]. Such individuals often believe that worrying about their health and gaining more knowledge about illnesses provide a sense of security. This biased thinking pattern drives compulsive health information-seeking behaviors, which are hallmarks of cyberchondria [41]. However, excessive information-seeking often leads to heightened anxiety, cognitive overload, and feelings of inadequacy in managing one's health, which can result in health information avoidance as a coping mechanism. Therefore, this study hypothesizes that positive metacognition positively influences an individual's likelihood of developing cyberchondria, which, in turn, leads to health information avoidance. Based on this logic, the following research hypotheses are proposed:

H2a: Positive metacognition positively influences cyberchondria.

H2b: Cyberchondria mediates the effect of positive metacognition on health information avoidance.

The concept of self-efficacy was first introduced by Bandura [42] in 1977, referring to an individual's confidence and belief in their ability to achieve specific behavioral goals. Ajzen [43] later incorporated it into the theory of planned behavior, which suggests that self-efficacy is a significant predictor of an individual's behavioral intention. Numerous studies have since confirmed that self-efficacy is an important component of the internal cognitive system, significantly affecting personal behavior and decision-making [44]. Self-efficacy has been widely studied in the field of health behaviors [45], where it is often referred to as health self-efficacy. According to Lee et al [46], health self-efficacy refers to an individual's confidence in their ability to manage their health. Individuals with higher levels of health self-efficacy are more likely to feel capable of addressing health issues independently and have a stronger inclination toward self-healing. This self-healing tendency reduces communication with health care professionals and encourages individuals to search for health and treatment-related information for self-treatment [47]. Furthermore, the tendency toward self-healing has been shown to be significantly associated with cyberchondria [22]. Therefore, this study

XSL•FO
RenderX

speculates that individuals with higher levels of health self-efficacy are more inclined to gather health-related information through internet-based channels, which may, in turn, lead to a tendency toward cyberchondria, causing these users to become trapped in a health information bubble. Based on this reasoning, the following research hypotheses are proposed:

H3a: Health self-efficacy positively influences cyberchondria.

H3b: Cyberchondria mediates the effect of health self-efficacy on health information avoidance.

External Factors: Health Information Similarity and Subjective Norms

Health information similarity refers to the consistency or resemblance in the content and structure of health information encountered on digital platforms, often driven by personalized recommendation algorithms tailored to users' preferences and behaviors [48,49]. These algorithms frequently expose users to overlapping health content, which can create cognitive confusion and negative emotional responses. For instance, similar symptom descriptions across diseases make it difficult to differentiate conditions, increasing uncertainty and anxiety [50]. This anxiety often triggers more frequent searches, leading to information overload and reinforcing a cycle of heightened anxiety.

While previous studies link information similarity to health anxiety, its impact on cyberchondria remains underexplored. Repeated exposure to overlapping health content can cause individuals to misinterpret ordinary symptoms such as severe illnesses, further amplifying uncertainty and intensifying cyberchondria [51]. This exacerbates excessive information searching and, in turn, contributes to health information avoidance. Based on this reasoning, the following hypotheses are proposed:

H4a: Health information similarity positively influences cyberchondria.

H4b: Cyberchondria mediates the effect of health information similarity on health information avoidance.

Subjective norms refer to the perceived social pressure or expectations from others, reflecting the social acceptance of a particular behavior. Both the risk information seeking and processing model and the theory of planned behavior emphasize that social factors (subjective norms) significantly influence an individual's intention to seek health information [52]. In this study, subjective norms specifically refer to the expectations from important others, such as family and friends, regarding an individual's health information-seeking behavior. When individuals perceive strong interest or encouragement from their social circle, subjective norms can create pressure to continue searching for health information. This social pressure often leads individuals to feel a heightened sense of responsibility, causing them to persist in health information searches despite experiencing anxiety and discomfort. Such pressure can exacerbate anxiety, trapping individuals in a cycle of frequent searches and information overload [53]. Furthermore, subjective norms can make it difficult for individuals to exercise self-restraint, increasing doubts about their health and ultimately contributing to cyberchondria.

Subjective norms may also worsen cyberchondria by influencing self-perception. When individuals believe that those around them place great importance on health information searching, they may see not engaging in such behavior as irresponsible, further intensifying their psychological burden and health anxiety [54]. To alleviate this emotional distress, individuals may choose to avoid health information altogether, attempting to reduce the psychological burden caused by social pressure. Based on this reasoning, the following hypotheses are proposed:

H5a: Subjective norms positively influence cyberchondria.

H5b: Cyberchondria mediates the effect of subjective norms on health information avoidance.

Mindfulness as a Proactive Psychological Intervention Method

The resource model of self-control and ego depletion theory posits that self-control tasks, such as compulsive health information searches, deplete cognitive resources ("ego depletion"), impairing subsequent performance and leading to avoidance of related activities like processing additional health information [55,56]. However, self-control can be improved through consistent training. Mindfulness, rooted in Buddhist traditions, is defined as present-focused attention and nonjudgmental acceptance [57]. It includes 4 elements: being present, attention, awareness, and acceptance [58]. Mindfulness training enhances self-control, reduces cognitive depletion, and improves emotional regulation, making it effective for addressing anxiety, depression, and stress [59,60].

In the context of cyberchondria, individuals experience intense anxiety due to excessive focus on health information, leading to repeated searches that amplify uncertainty and anxiety. This cycle depletes cognitive resources, increasing health information avoidance. Mindfulness training helps break this cycle by enhancing awareness and acceptance, thereby reducing anxiety and distress [61]. Improved emotional regulation allows individuals to manage anxiety without excessive searching, disrupting the "anxiety–information searching–greater anxiety–information avoidance" loop. In addition, mindfulness promotes cognitive flexibility, helping individuals identify and resist cognitive biases common in cyberchondria, such as misinterpreting ordinary symptoms as severe illnesses [62]. By encouraging observation over reaction, mindfulness reduces health anxiety and prevents avoidance behaviors stemming from excessive concern about health information.

H6a: Mindfulness training significantly reduces individuals' levels of cyberchondria.

H6b: Mindfulness training negatively moderates the positive effect of cyberchondria on health information avoidance.

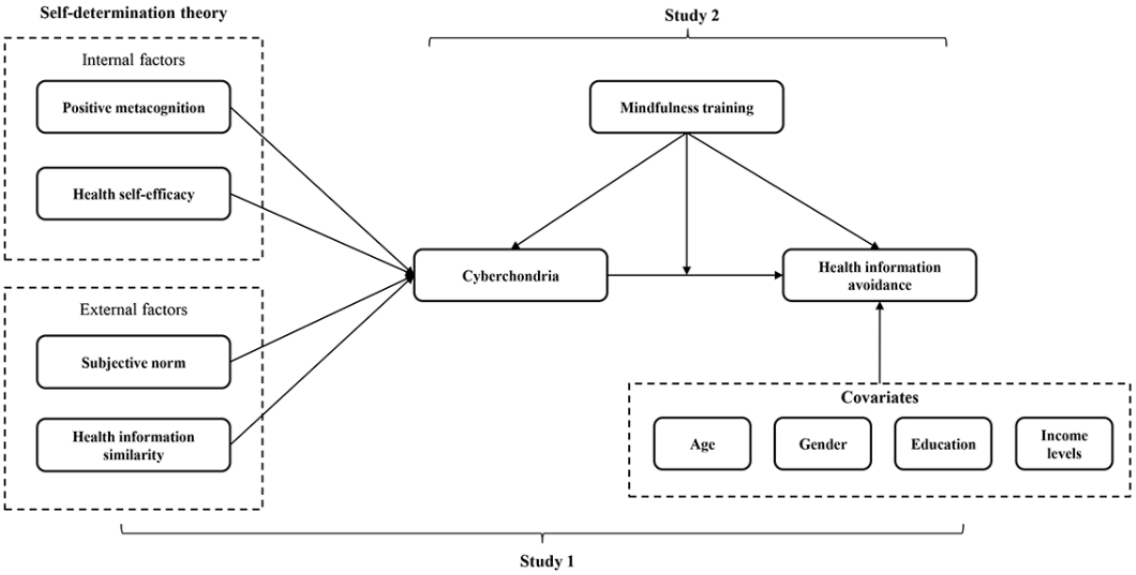
Conceptual Model

In summary, this study consists of 2 substudies. Study 1 aims to explore the causes of cyberchondria among older adults in the digital age and the impact model of health information avoidance behavior among older adults. Study 2 aims to validate

XSL•FO
RenderX

the intervention effects of mindfulness training on cyberchondria and health information avoidance behavior in older adults. The specific research models are shown in Figure 1.

Figure 1. Conceptual model of health information avoidance among older adults.



Methods

Study 1

Participants and Procedure

Participants for study 1 were primarily recruited through health communities and social media platforms commonly used by older adults, such as WeChat Moments (Tencent), Douyin (the Chinese version of TikTok, ByteDance), and Jinri Toutiao (Today's Headlines). The target population included individuals aged 60 years and older who were capable of independently completing internet-based questionnaires. To enhance sample diversity and ensure an adequate sample size, additional outreach was conducted through local community networks, such as senior activity centers.

Data for study 1 was collected through internet-based survey platform. Before the survey began, the researchers clearly informed all participants about the purpose of the study and assured them that all responses would be strictly confidential and used solely for academic purposes. The survey was conducted anonymously, with no personal information being disclosed. Participants were also informed of the importance of providing truthful responses and were reminded of their right to withdraw from the survey at any time without any conditions. Informed consent was obtained from all participants.

Initially, 328 questionnaires were collected, with 92 invalid responses excluded, resulting in 236 valid questionnaires (mean age 63.93 years; 91, 38.56% males; 145, 61.44% females). Demographic details are provided in the Multimedia Appendix 1. The response rate of valid questionnaires was 71.95%. The criteria for selecting valid questionnaires were correctly answering screening questions, a total response time of more than 1 minute, and not selecting the same response option for 9 consecutive items.

Measurements

Study 1 measured the following latent variables: health information avoidance [63], cyberchondria [64], positive metacognition [65], health self-efficacy [66], subjective norms, and health information similarity [67,68]. The detailed measurement scales can be found in Multimedia Appendix 1. To ensure validity, all measurement tools were adapted from established research and adjusted to fit the context of health information dissemination. The average score of the items in each scale was used to represent each latent variable, and all questionnaires were measured using a 5-point Likert scale. A pretest was also conducted with 10 volunteers to ensure that all items were clearly stated and accurately understood.

Measurement of Model

Since this study is exploratory in nature, the research model, although built based on theoretical and logical reasoning, does not originate from an existing model. Therefore, it is appropriate to use "Smart PLS" for partial least squares (PLS) path analysis () to test the research model in this context [69].

Reliability and validity testing. As shown in Table 1, the factor loadings for all measurement items ranged from 0.728 to 0.950, indicating that all items met the retention criteria. In addition, the Cronbach α values for each latent variable ranged from 0.870 to 0.915, and the composite reliability (CR) values for all latent variables were greater than the acceptable threshold of 0.7, indicating that the reliability of the scales was satisfactory. Furthermore, the average variance extracted (AVE) values for all variables exceeded the standard acceptance value of 0.5, demonstrating adequate convergent validity. The variance inflation factor (VIF) values for all factors were below 10, indicating that there were no multicollinearity issues in the scales used in this study [70,71].

XSL•FO
RenderX

Table 1. Reliability and convergent validity results of latent variables.

Latent variable and items	Factor loadings	VIF[a]	Cronbach α values	CR[b]	AVE[c]
Health information avoidance			.912	.925	.739
HIA[d] 1	.841	2.608			
HIA 2	.903	3.805			
HIA 3	.803	2.056			
HIA 4	.854	2.461			
HIA 5	.894	2.929			
Cyberchondria			.915	.916	.796
CC[e] 1	.904	3.421			
CC 2	.912	3.329			
CC 3	.898	3.127			
CC 4	.854	2.202			
Positive metacognition			.898	.884	.710
PM[f] 1	.847	2.528			
PM 2	.854	2.408			
PM 3	.797	1.955			
PM 4	.869	2.744			
PM 5	.844	2.373			
Health self-efficacy			.870	.903	.718
HS[g] 1	.873	2.145			
HS 2	.823	2.053			
HS 3	.851	2.377			
HS 4	.843	2.100			
Subjective norm			.912	.936	.850
SN[h] 1	.944	3.727			
SN 2	.882	2.521			
SN 3	.940	3.908			
Health information similarity			.880	.927	.731
IS[i] 1	.728	2.246			
IS 2	.877	2.815			
IS 3	.850	2.239			
IS 4	.950	4.355			

[a]VIF: variance inflation factor.
[b]CR: composite reliability.
[c]AVE: average variance extracted.
[d]HIA: health information avoidance.
[e]CC: cyberchondria.
[f]PM: positive metacognition.
[g]HS: health self-efficacy.
[h]SN: subjective norm.
[i]IS: health information similarity.

Next, discriminant validity among the variables was verified, and the specific results are shown in Table 2. The square root of the AVE for all variables (on the diagonal) was greater than the Pearson correlation coefficients between the variables,

XSL•FO
RenderX

indicating that the scales had satisfactory discriminant validity. A common method bias test was also conducted using the single-factor test. The variance explained by the first unrotated factor was 31.999%, which is below the critical threshold of 40%, indicating that common method bias is not an issue in this study [72].

Table 2. Discriminant validity analysis of constructs in study 1.

Constructs	HIA[a]	CC[b]	HS[c]	PM[d]	SN[e]	IS[f]
HIA	.860[g]	—[h]	—	—	—	—
CC	.410	.892[g]	—	—	—	—
HS	.346	.382	.848[g]	—	—	—
PM	.212	.343	.184	.843[g]	—	—
SN	.158	.137	.019	.651	.922[g]	—
IS	.339	.435	.291	.371	.246	.855[g]

[a]HIA: health information avoidance.

[b]CC: cyberchondria.

[c]HS: Health self-efficacy.

[d]PM: Positive metacognition.

[e]SN: subjective norm.

[f]IS: health information similarity.

[g]square roots of the average variance extracted.

[h]Not applicable.

Model Testing

The PLS algorithm was used to calculate the R^2 values for each variable, all of which were greater than the standard acceptance value of 0.1, indicating good predictive accuracy of the model [73]. Blindfolding was then conducted, and the results showed that the Stone-Geisser Q^2 values for all variables were greater than 0, indicating that the model effectively predicts the relationships between variables [74]. In addition, the SRMR value was 0.054, which is below the standard threshold of 0.08, the NFI value was 0.931, and the RMS θ value was 0.103. These results demonstrate that the model has good fit [75].

Study 2

Participants and Procedure

The required sample size for the intervention study was estimated using GPower, with an effect size greater than 0.8 considered appropriate for the experimental study. Setting α=.05, 1-β=.8, and effect size=0.4, the minimum required sample size was calculated to be 44 [76].

The mindfulness training required participants to invest a significant amount of time, making it challenging to recruit a sufficient number of participants. To address this difficulty, our research team collaborated with students from 3 courses at 3 universities, encouraging them to invite their older family members (older than 60 years old and capable of independently using social media) to participate in the mindfulness training program. To motivate the students, the researchers offered course credits and a reward of approximately US $20. To improve the quality of the mindfulness training and the responses to the survey, the researchers informed participants before the experiment began that their responses would be kept strictly confidential and used only for academic purposes. Participants were also informed of the importance of honest responses and were made aware of their right to withdraw from the study at any time without conditions. Informed consent was obtained from all participants. The research was conducted in accordance with the standards of the Declaration of Helsinki.

We initially recruited 101 participants and divided them into 2 groups: group 1 (mindfulness training group, n=51) and group 2 (control group, n=50). The research process was conducted in 3 stages.

Stage 1 (t1)

At the start of the study, a survey measured participants' levels of cyberchondria, health information avoidance, state mindfulness (Brown and Ryan [77]), and demographic variables (gender, age, income, and education). An independent samples *t* test showed no significant differences between the 2 groups for cyberchondria (M_{group1}=2.848, M_{group2}=3.110, P=.12, t_{99}=–1.557), health information avoidance (M_{group1}=3.063, M_{group2}=3.384, P=.11, t_{99}=–1.631), or state mindfulness (M_{group1}=3.502, M_{group2}=3.524, P=.77, t_{99}=–.297).

Stage 2

Mindfulness training started 1 week after the initial survey to avoid bias from survey content. Group 1 underwent a short-term intervention based on the mindfulness-based stress reduction (MBSR) program (Kabat-Zinn [78]), while group 2 received no intervention. The training included five 10-minute audio sessions focusing on breathing, bodily sensations, thoughts, and emotions. Participants were instructed to complete a daily session at noon with randomly assigned recordings and record their practice (refer to Multimedia Appendix 1 for audio excerpts).

XSL•FO
RenderX

Stage 3 (t2)

Following 4 weeks of training, consistent with previous findings on mindfulness duration (Rooks et al [79]), 5 participants from group 1 and 2 from group 2 withdrew or failed to complete the study, leaving a final valid sample of 46 in group 1 and 48 in group 2. Cyberchondria, health information avoidance, and state mindfulness levels were reassessed. Reanalysis of initial data showed no significant differences between retained participants, ensuring data integrity.

Ethical Considerations

This study strictly adhered to the ethical principles outlined in the Declaration of Helsinki and was approved by the Ethics Committee of the School of Journalism and Communication at Minjiang University (Ref: MJUCER20240107). Both substudies were conducted in strict accordance with relevant ethical guidelines to protect participants' rights and ensure data confidentiality.

All participants provided informed consent before participating in the study. They were fully informed about the study's objectives, procedures, data usage policies, and their right to withdraw from the study at any time without penalty. As the data were collected through questionnaires, the consent forms were obtained electronically.

To protect privacy and data security, all participant data were anonymized and deidentified during both the data collection and analysis processes. No personally identifiable information was recorded during the study, and all data were securely stored in an encrypted system to prevent unauthorized access.

Participants received appropriate compensation based on their involvement in the study. In substudy 1, a cross-sectional survey, participants were rewarded US $2 upon completing the questionnaire. In substudy 2, a 4-week longitudinal survey, participants received a gift valued at approximately US $20 after completing the study. These compensations were intended to acknowledge participants' time and effort while avoiding undue influence on their decision to participate. This manuscript and its supplementary materials do not include any images or materials that could identify participants.

Results

Study 1

There were no missing values in the data. A bootstrapping test with a sample size of 5000 was conducted on the collected data to explore the path coefficients and their significance within the model. The final test results are presented in Tables 3 and 4.

The results of the data analysis indicate that cyberchondria exacerbates health information avoidance among older adults (β=.41, *P*<.001), supporting H1. From the perspective of internal factors, both positive metacognition and health self-efficacy have a positive influence on cyberchondria (β=.26, *P*=.002; β=.25, *P*<.001), which in turn further increases health information avoidance behavior (95% Boot CI 0.035-0.189; 95% Boot CI 0.043-0.185). Thus, H2a, H2b, H3a, and H3b are all supported. Regarding external environmental factors, the similarity of health information also significantly positively affects cyberchondria (β=.29, *P*<.001) and leads to health information avoidance through cyberchondria (95% Boot CI 0.063-0.191), supporting H4a and H4b. However, subjective norms did not show a significant direct relationship with cyberchondria (β=–.11, *P*=.13), and cyberchondria did not mediate the relationship between subjective norms and health information avoidance (95% Boot CI –0.102 to 0.016), leading to the rejection of H5a and H5b.

The results of study 1 indicate that cyberchondria significantly exacerbates health information avoidance among older adults, influenced by both internal factors (positive metacognition and health self-efficacy) and external factors (health information similarity). However, subjective norms did not have a significant effect on cyberchondria, suggesting that external social pressure is not a major driver of health information avoidance behavior in older adults. Based on these findings, the next study will focus on validating whether mindfulness training can effectively intervene in cyberchondria and the resulting health information avoidance behavior. The specific findings of study 1 are shown in Figure 2.

XSL•FO
RenderX

Table 3. Hypothesis testing results of study 1 (direct effect).

Direct effect	β values	*t* test (*df*)	*P* values
H1: CC[a]→HIA[b]	.410	7.016 (226)	<.001[c]
H2a: PM[d]→CC	.259	3.035 (226)	.002[e]
H3a: HS[f]→CC	.251	3.453 (227)	.001[e]
H4a: IS[g]→CC	.293	4.344 (227)	<.001[c]
H5a: SN[h]→CC	-.109	–1.507 (228)	.13

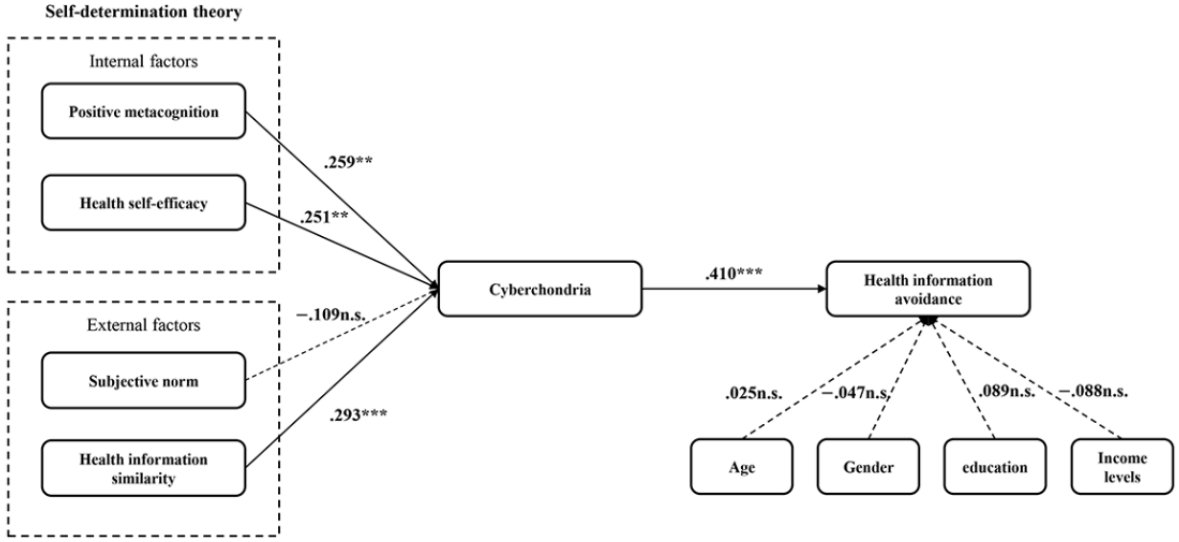
[a]CC: cyberchondria.
[b]HIA: health information avoidance.
[c]*P*<.001.
[d]PM: positive metacognition.
[e]*P*<.01.
[f]HS: health self-efficacy.
[g]IS: health information similarity.
[h]SN: subjective norm.

Table 4. Hypothesis testing results of study 1 (indirect effect).

Indirect effect	Effect value	*t* values (*df*)	95% CI	*P* values
H2b: PM[a]→CC[b]→HIA[c]	0.106	2.697 (220)	0.035 to 0.189	.007[d]
H3b: HS[e]→CC→HIA	0.103	2.821 (221)	0.043 to 0.185	.005[d]
H4b: IS[f]→CC→HIA	0.120	3.691 (221)	0.063 to 0.191	<.001[g]
H5b: SN[h]→CC→HIA	–0.045	1.459 (222)	–0.102 to 0.016	.15

[a]PM: positive metacognition.
[b]CC: cyberchondria.
[c]HIA: health information avoidance.
[d]*P*<.01.
[e]HS: health self-efficacy.
[f]IS: health information similarity.
[g]*P*<.001.
[h]SN: subjective norm.

Figure 2. Result model of study 1: path analysis.***P*<.01, ****P*<.001, n.s. *P*>.05.



XSL•FO
RenderX

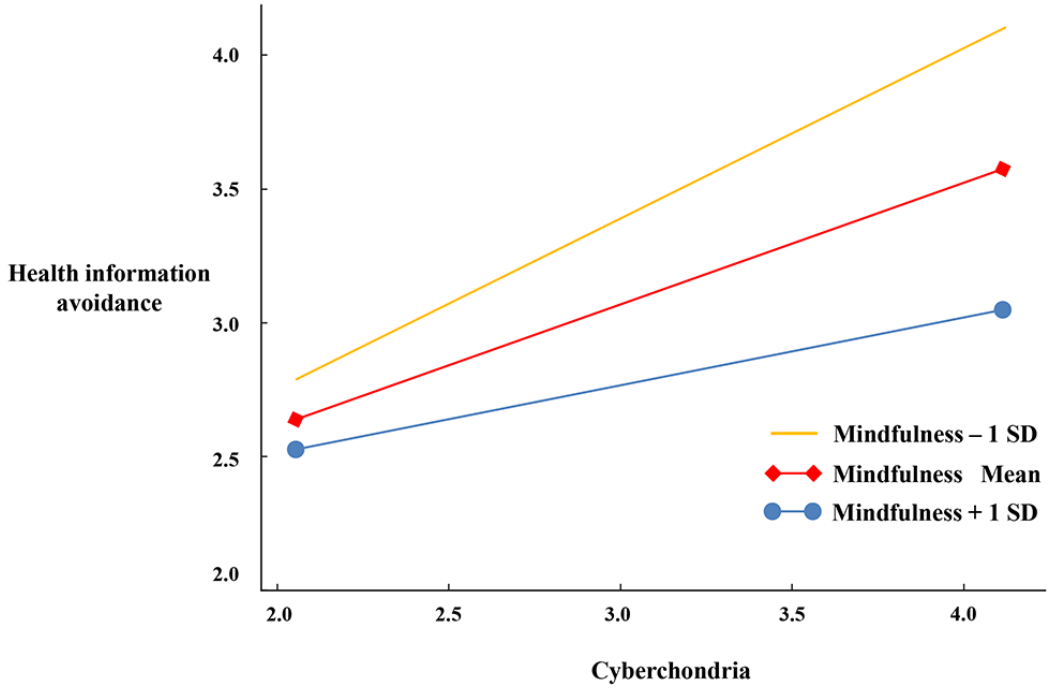
Study 2

Before hypothesis testing, the effectiveness of the mindfulness intervention was first examined. To control for the effects of covariates such as age, gender, and education level, an ANCOVA (analysis of covariance) was conducted to test the manipulation of state mindfulness. The analysis showed that the state mindfulness score of the intervention group was significantly higher than that of the control group (M_{group1}=4.122, M_{group2}=3.606, $F_{1,93}$=16.866, P<.001). Subsequently, the within-group differences in t1 and t2 for both the intervention and control groups were compared. The results showed that the state mindfulness level of the intervention group at t2 was significantly higher than at t1 (M_{t2}=4.122, M_{t1}=3.502, $F_{1,96}$=85.024, P<.001), while there was no significant difference between t1 and t2 in the control group (M_{t2}=3.524, M_{t1}=3.583, $F_{1,97}$=.222, P=.64). In summary, the manipulation of state mindfulness was successful.

Next, ANCOVA was conducted to test H6a, with age, gender, education level, and income level as covariates. The results showed that there was no significant difference in the level of cyberchondria between t1 and t2 in the intervention group (M_{t1}=2.848, M_{t2}=2.685, $F_{1,96}$=1.835, P=.18), indicating that H6a was not supported, meaning that mindfulness training did not improve the level of cyberchondria. Subsequently, H6b was tested using Hayes' suggested Bootstrap method with SPSS (IBM Corp) process model 1 to examine the interaction effect, with 5000 bootstrap samples and a 95% CI. The data analysis showed that the interaction effect of mindfulness level and cyberchondria on health information avoidance was significant (effect= 0.357, P=.002, 95% CI 0.580 to 0.131). Thus, H6b was supported.

In conclusion, the results of study 2 supported H6b but rejected H6a. Although mindfulness training did not directly reduce the level of cyberchondria, it effectively inhibited the transformation of cyberchondria into health information avoidance behavior (Figure 3).

Figure 3. The Interaction effect of mindfulness and cyberchondria on health information avoidance.



Discussion

Principal Findings

Cyberchondria and health information avoidance have become increasingly pressing public health issues in the digital age, posing significant threats to the health of older adults. Therefore, exploring their underlying mechanisms and potential intervention strategies is of great theoretical and practical significance. Based on the self-determination theory, this study conducted 2 substudies to explore the causes of cyberchondria and health information avoidance among older adults, as well as the intervention effects of mindfulness training.

Internal and External Factors Leading to Health Information Avoidance: Cyberchondria as the Mediator

The results reveal that internal factors (positive metacognition and health self-efficacy) and external factors (health information similarity) significantly predict cyberchondria among older adults, leading to health information avoidance.Positive metacognition fosters overthinking and health-related anxiety, as excessive focus on health concerns amplifies uncertainty. Previous research links positive metacognition to problematic behaviors [80], and this study confirms its role in exacerbating cyberchondria and health information avoidance. Health self-efficacy also influences cyberchondria. Although high self-efficacy promotes confidence in managing health, ambiguous health information undermines this confidence,

XSL•FO
RenderX

driving excessive searches and heightening anxiety. This aligns with self-efficacy theory, where unresolved issues lead to repetitive checking and stress [81].

Health information similarity indirectly contributes to avoidance by increasing anxiety and cognitive overload. Repeated exposure to similar information, often driven by algorithms, makes it difficult to differentiate between diseases, heightening uncertainty and driving avoidance behaviors through cyberchondria [82]. However, subjective norms did not significantly affect cyberchondria, suggesting that older adults rely more on personal judgment and experience rather than external social influences [83,84]. This finding underscores the need for interventions targeting individual-level factors rather than external pressures.

Mindfulness Training as an Effective Intervention for Cyberchondria and Health Information Avoidance

The results also indicate that while mindfulness training did not directly reduce the level of cyberchondria, it significantly mitigated the transition from cyberchondria to health information avoidance behavior. Mindfulness meditation, by enhancing emotional regulation and cognitive flexibility, helps older adults avoid being dominated by anxiety when faced with health information, even though it may not directly prevent the onset of cyberchondria. This effectively reduces their tendency to avoid health information. These findings are consistent with previous studies on the effects of mindfulness, supporting the conclusion that mindfulness training can help individuals better accept current emotions and reduce fear of future uncertainties [85]. Furthermore, our findings emphasize the role of mindfulness training in regulating health information behavior, demonstrating that it not only improves the emotional state of older adults but also increases their acceptance of health information, thereby reducing avoidance behaviors. This highlights the potential of mindfulness training to effectively break the vicious cycle of "anxiety–information search–greater anxiety–information avoidance." Mindfulness can enable older adults to maintain a more balanced perspective when encountering health-related information, which prevents anxiety from escalating into avoidant behavior.

Implications and Future Directions

This Study Makes Three Theoretical Contributions

First, it makes significant contributions to SDT by deepening the understanding of how internal and external factors influence health behaviors through psychological mechanisms. It validates the mediating role of cyberchondria in health information avoidance among older adults and demonstrates how positive metacognition, health self-efficacy, and health information similarity interact within this process. This fills a critical gap in previous research by uncovering the mediating mechanisms of health information avoidance and provides new insights into the cognitive-emotional processes of older adults, an information-vulnerable population. Furthermore, this study extends the applicability of SDT by highlighting the dynamic balance between intrinsic motivation and external support in older adults navigating complex digital health environments. In addition, this study systematically explores the role of mindfulness interventions in managing health information behaviors. While mindfulness did not directly reduce cyberchondria, it significantly inhibited its progression to information avoidance. By enhancing older adults' capacity to cope with uncertainty and anxiety, mindfulness fosters a more open attitude toward health information. This expands mindfulness beyond emotional regulation to include its impact on complex health decision-making, providing theoretical support for its application in public health interventions. Finally, this study challenges the universal applicability of subjective norms in shaping health behaviors, as proposed in the theory of planned behavior [86]. It finds that older adults rely more on autonomous judgment than societal expectations due to their life experience and independence, emphasizing the importance of considering demographic differences and group heterogeneity in health behavior research.

Practically, This Study Offers Three Recommendations

First, for internal factors (eg, positive metacognition and excessive health self-efficacy), psychological education and cognitive behavioral therapy can help older adults reduce overfocus on health risks and avoid unnecessary health information searches [87]. Second, optimizing recommendation algorithms on health information platforms to reduce repetitive content and providing personalized information tailored to older adults' health conditions can alleviate anxiety caused by information overload and similarity. Finally, enhancing digital health literacy through health seminars and digital skills training, combined with mindfulness programs, can offer a comprehensive intervention strategy to address health information avoidance. These measures collectively aim to reduce health anxiety, improve engagement with health information, and enable older adults to achieve proactive health management.

Limitations

This study provides valuable insights into the causes of cyberchondria and health information avoidance among older adults and evaluates mindfulness meditation as an intervention. However, several limitations should be noted. First, the 4-week mindfulness intervention, while effective in the short term, does not address the sustainability of its effects. Future research should explore longer interventions and conduct follow-ups to assess lasting impacts. Combining mindfulness with approaches such as health education or social support may yield stronger and more enduring outcomes. In addition, while positive metacognition and health self-efficacy emerged as critical internal factors, specific intervention strategies for addressing these were not explored. Future research should investigate techniques such as cognitive behavioral therapy and health coaching to mitigate their negative impacts and further reduce health information avoidance. Expanding on these areas will help refine interventions and improve health management strategies for older adults.

Conclusion

This study used 2 substudies to explore the causes of cyberchondria and health information avoidance among older adults, as well as the intervention effects of mindfulness

XSL•FO
RenderX

meditation. Study 1 constructed an influence model based on self-determination theory, using a cross-sectional survey to verify that internal factors such as positive metacognition (β=.26, *P*=.002) and health self-efficacy (β=.25, *P*<.001), as well as external factors such as health information similarity (β=.29, *P*<.001), were significant predictors of cyberchondria among older adults, which in turn led to health information avoidance behavior (β=.41, *P*<.001). However, subjective norms did not show a significant association with cyberchondria among older adults (β=–.11, *P*=.13). Study 2 conducted a 4-week mindfulness intervention, demonstrating the effectiveness of mindfulness in inhibiting the transition from cyberchondria to health information avoidance (effect=–0.357, *P*=.002, 95% CI –0.580 to –0.131]). However, it was also found that mindfulness did not directly reduce the occurrence of cyberchondria (*P*=.18).

Acknowledgments

The authors thank all the participants in this study. This research was funded by Minjiang University Research Start-up Funding (grant 324-32404314). The authors confirm that no generative AI tools, including ChatGPT or similar models, were used in the writing, editing, or preparation of this manuscript. All content was developed independently by the authors, based on original research and analysis.

Data Availability

The dataset of this study is accessible through science data bank website. In compliance with ethical guidelines and to protect participant privacy, all personally identifiable information has been removed from the dataset. Researchers are welcome to access the data through the provided link. For additional inquiries or specific questions about the dataset, please contact the corresponding author.

Authors' Contributions

CG contributed to conceptualization, methodology, software, formal analysis, data curation, writing-original draft preparation, writing-review and editing, visualization, and project administration. XZ handled methodology, validation, formal analysis, and resources. LQ managed validation, investigation, resources, visualization, and project administration.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A: Demographic information B: Measures C: Study 2 Stimulus materials D: Checklist for Reporting Results of Internet E-Surveys (CHERRIES).
[DOCX File , 26 KB - publichealth_v11i1e69554_app1.docx]

References

1. Zhao YC, Zhao M, Song S. Online health information seeking behaviors among older adults: systematic scoping review. J Med Internet Res 2022;24(2):e34790 [FREE Full text] [doi: 10.2196/34790] [Medline: 35171099]
2. Cheng X, Huang X, Yang B, Xiao Y. Unveiling the paradox of technostress: Impacts of technology-driven stressors on the elderly's avoidance behaviors. Inf. Manag 2023;60(8):103875. [doi: 10.1016/j.im.2023.103875]
3. Barbour JB, Rintamaki LS, Ramsey JA, Brashers DE. Avoiding health information. J Health Commun 2012;17(2):212-229. [doi: 10.1080/10810730.2011.585691] [Medline: 22004015]
4. Sun Y, Feng Y, Shen X, Guo X. Fear appeal, coping appeal and mobile health technology persuasion: a two-stage scenario-based survey of the elderly. Inf. Technol. People 2022;36(1):362-386. [doi: 10.1108/itp-07-2021-0519]
5. Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use. Expert Rev Neurother 2013;13(2):205-213. [doi: 10.1586/ern.12.162] [Medline: 23368807]
6. Vismara M, Varinelli A, Pellegrini L, Enara A, Fineberg NA. New challenges in facing cyberchondria during the coronavirus disease pandemic. Curr Opin Behav Sci 2022;46:101156 [FREE Full text] [doi: 10.1016/j.cobeha.2022.101156] [Medline: 35581995]
7. Zheng H, Chen X, Jiang S, Sun L. How does health information seeking from different online sources trigger cyberchondria? The roles of online information overload and information trust. Inf. Process. Manag 2023;60(4):103364. [doi: 10.1016/j.ipm.2023.103364]
8. Chae J. Who avoids cancer information? Examining a psychological process leading to cancer information avoidance. J Health Commun 2016;21(7):837-844. [doi: 10.1080/10810730.2016.1177144] [Medline: 27337343]
9. Linardon J. Rates of attrition and engagement in randomized controlled trials of mindfulness apps: Systematic review and meta-analysis. Behav Res Ther 2023;170:104421 [FREE Full text] [doi: 10.1016/j.brat.2023.104421] [Medline: 37862854]

XSL•FO
RenderX

10. Creswell JD. Mindfulness Interventions. Annu Rev Psychol 2017;68:491-516. [doi: 10.1146/annurev-psych-042716-051139] [Medline: 27687118]
11. Davodi SR, Zendehtalab H, Zare M, Behnam Vashani H. Effect of health promotion interventions in active aging in the elderly: a randomized controlled trial. Int J Community Based Nurs Midwifery 2023;11(1):34-43 [FREE Full text] [doi: 10.30476/IJCBNM.2022.96246.2117] [Medline: 36650846]
12. Howell JL, Shepperd JA. Reducing health-information avoidance through contemplation. Psychol Sci 2013;24(9):1696-1703. [doi: 10.1177/0956797613478616] [Medline: 23842956]
13. Leyva B, Taber JM, Trivedi AN. Medical care avoidance among older adults. J Appl Gerontol 2020;39(1):74-85. [doi: 10.1177/0733464817747415] [Medline: 29249189]
14. Case D, Andrews JE, Johnson JD, Allard SL. Avoiding versus seeking: the relationship of information seeking to avoidance, blunting, coping, dissonance, and related concepts. J Med Libr Assoc 2005;93(3):353-362 [FREE Full text] [Medline: 16059425]
15. Howell JL, Lipsey NP, Shepperd JA. Health information avoidance. The Wiley Encyclopedia of health psychology 2020:279-286. [doi: 10.1002/9781119057840.ch77]
16. Schenkel SK, Jungmann SM, Gropalis M, Witthöft M. Conceptualizations of cyberchondria and relations to the anxiety spectrum: systematic review and meta-analysis. J Med Internet Res 2021;23(11):e27835 [FREE Full text] [doi: 10.2196/27835] [Medline: 34792473]
17. Vismara M, Caricasole V, Starcevic V, Cinosi E, Dell'Osso B, Martinotti G, et al. Is cyberchondria a new transdiagnostic digital compulsive syndrome? A systematic review of the evidence. Compr Psychiatry 2020;99:152167 [FREE Full text] [doi: 10.1016/j.comppsych.2020.152167] [Medline: 32146315]
18. Tyrer P, Cooper S, Tyrer H, Wang D, Bassett P. Increase in the prevalence of health anxiety in medical clinics: possible cyberchondria. Int J Soc Psychiatry 2019;65(7-8):566-569 [FREE Full text] [doi: 10.1177/0020764019866231] [Medline: 31379243]
19. Bajcar B, Babiak J. Self-esteem and cyberchondria: The mediation effects of health anxiety and obsessive–compulsive symptoms in a community sample. Curr Psychol 2019;40(6):2820-2831. [doi: 10.1007/s12144-019-00216-x]
20. Fergus TA. Anxiety sensitivity and intolerance of uncertainty as potential risk factors for cyberchondria: a replication and extension examining dimensions of each construct. J Affect Disord 2015;184:305-309. [doi: 10.1016/j.jad.2015.06.017] [Medline: 26141807]
21. Bajcar B, Babiak J. Neuroticism and cyberchondria: The mediating role of intolerance of uncertainty and defensive pessimism. Pers. Individ. Differ 2020;162:110006. [doi: 10.1016/j.paid.2020.110006]
22. Khan AW, Pandey J. Dark side consequences of cyberchondria: an empirical investigation. Aslib J. Inf. Manag 2022;74(5):801-817. [doi: 10.1108/ajim-08-2021-0222]
23. Mathes BM, Norr AM, Allan NP, Albanese BJ, Schmidt NB. Cyberchondria: Overlap with health anxiety and unique relations with impairment, quality of life, and service utilization. Psychiatry Res 2018;261:204-211. [doi: 10.1016/j.psychres.2018.01.002] [Medline: 29324396]
24. Fergus TA, Spada MM. Cyberchondria: Examining relations with problematic Internet use and metacognitive beliefs. Clin Psychol Psychother 2017;24(6):1322-1330. [doi: 10.1002/cpp.2102] [Medline: 28621035]
25. Brown RJ, Skelly N, Chew-Graham CA. Online health research and health anxiety: A systematic review and conceptual integration. Clin. Psychol.: Sci. Pract 2020;27(2):e12299. [doi: 10.1111/cpsp.12299]
26. Starcevic V. Keeping Dr. Google under control: how to prevent and manage cyberchondria. World Psychiatry 2023;22(2):233-234 [FREE Full text] [doi: 10.1002/wps.21076] [Medline: 37159368]
27. Deci EL, Ryan RM. Self-determination theory. Handbook of theories of social psychology 2012;1(20):416-436. [doi: 10.4135/9781446249215.n21]
28. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000;55(1):68-78. [doi: 10.1037//0003-066x.55.1.68] [Medline: 11392867]
29. Gillison FB, Rouse P, Standage M, Sebire SJ, Ryan RM. A meta-analysis of techniques to promote motivation for health behaviour change from a self-determination theory perspective. Health Psychol Rev 2019;13(1):110-130. [doi: 10.1080/17437199.2018.1534071] [Medline: 30295176]
30. Assadi V, Hassanein K. Consumer adoption of personal health record systems: a self-determination theory perspective. J Med Internet Res 2017;19(7):e270 [FREE Full text] [doi: 10.2196/jmir.7721] [Medline: 28751301]
31. Ali F, Zhang Q, Tauni MZ, Butt AH, Ahsan T. Contingent self-esteem, social interaction fears, and compulsive WeChat usage. Behav. Inf. Technol 2022;42(6):1-12. [doi: 10.1080/0144929x.2022.2039768]
32. Li Y, Wang X. Seeking health information on social media: a perspective of trust, self-determination, and social support. j organ end user com 2018;30(1):1-22. [doi: 10.4018/JOEUC.2018010101]
33. Laato S, Islam AKMN, Islam MN, Whelan E. What drives unverified information sharing and cyberchondria during the COVID-19 pandemic? Eur. J. Inf. Syst 2020;29(3):288-305. [doi: 10.1080/0960085x.2020.1770632]
34. Yalçın İ, Boysan M, Eşkisu M, Çam Z. Health anxiety model of cyberchondria, fears, obsessions, sleep quality, and negative affect during COVID-19. Curr Psychol 2024;43(9):8502-8519 [FREE Full text] [doi: 10.1007/s12144-022-02987-2]

XSL•FO
RenderX

35. Zhu X, Zheng T, Ding L, Zhang X. Exploring associations between eHealth literacy, cyberchondria, online health information seeking and sleep quality among university students: A cross-section study. Heliyon 2023;9(6):e17521 [FREE Full text] [doi: 10.1016/j.heliyon.2023.e17521] [Medline: 37408886]
36. Sweeny K, Melnyk D, Miller W, Shepperd JA. Information avoidance: who, what, when, and why. Rev. Gen. Psychol 2010;14(4):340-353. [doi: 10.1037/a0021288]
37. McMullan RD, Berle D, Arnáez S, Starcevic V. The relationships between health anxiety, online health information seeking, and cyberchondria: Systematic review and meta-analysis. J Affect Disord 2019;245:270-278. [doi: 10.1016/j.jad.2018.11.037] [Medline: 30419526]
38. Spada MM, Nikcević AV, Moneta GB, Wells A. Metacognition as a mediator of the relationship between emotion and smoking dependence. Addict Behav 2007;32(10):2120-2129. [doi: 10.1016/j.addbeh.2007.01.012] [Medline: 17307299]
39. Nadeem F, Malik N, Atta M, Ullah I, Martinotti G, Pettorruso M, et al. Relationship between health-anxiety and cyberchondria: role of metacognitive beliefs. J Clin Med 2022;11(9):2590 [FREE Full text] [doi: 10.3390/jcm11092590] [Medline: 35566713]
40. Kaur A, Butow PN, Thewes B. Do metacognitions predict attentional bias in health anxiety? Cogn Ther Res 2011;35(6):575-580. [doi: 10.1007/s10608-011-9387-6]
41. Starcevic V, Baggio S, Berle D, Khazaal Y, Viswasam K. Cyberchondria and its relationships with related constructs: a network analysis. Psychiatr Q 2019;90(3):491-505 [FREE Full text] [doi: 10.1007/s11126-019-09640-5] [Medline: 31098922]
42. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
43. Ajzen I. The theory of planned behavior. Organ. Behav. Hum. Decis. Process 1991;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
44. Aminuddin HB, Jiao N, Jiang Y, Hong J, Wang W. Effectiveness of smartphone-based self-management interventions on self-efficacy, self-care activities, health-related quality of life and clinical outcomes in patients with type 2 diabetes: a systematic review and meta-analysis. Int J Nurs Stud 2021;116:103286. [doi: 10.1016/j.ijnurstu.2019.02.003] [Medline: 30827741]
45. Simonetti V, Durante A, Ambrosca R, Arcadi P, Graziano G, Pucciarelli G, et al. Anxiety, sleep disorders and self-efficacy among nurses during COVID-19 pandemic: a large cross-sectional study. J Clin Nurs 2021;30(9-10):1360-1371 [FREE Full text] [doi: 10.1111/jocn.15685] [Medline: 33534934]
46. Lee SY, Hwang H, Hawkins R, Pingree S. Interplay of negative emotion and health self-efficacy on the use of health information and its outcomes. Commun. Res 2008;35(3):358-381. [doi: 10.1177/0093650208315962]
47. Essery R, Geraghty AWA, Kirby S, Yardley L. Predictors of adherence to home-based physical therapies: a systematic review. Disabil Rehabil 2017;39(6):519-534. [doi: 10.3109/09638288.2016.1153160] [Medline: 27097761]
48. Liang H, Fu K. Information overload, similarity, and redundancy: unsubscribing information sources on Twitter. J Comput-Mediat Comm 2016;22(1):1-17. [doi: 10.1111/jcc4.12178]
49. Cinelli M, De Francisci Morales G, Galeazzi A, Quattrociocchi W, Starnini M. The echo chamber effect on social media. Proc Natl Acad Sci U S A 2021;118(9):e2023301118 [FREE Full text] [doi: 10.1073/pnas.2023301118] [Medline: 33622786]
50. Malloch YZ, Zhang J. Seeing others receive support online: effects of self-disclosure and similarity on perceived similarity and health behavior intention. J Health Commun 2019;24(3):217-225. [doi: 10.1080/10810730.2019.1595226] [Medline: 30912708]
51. Deng Z, Deng Z, Fan G, Wang B, Fan W, Liu S. More is better? Understanding the effects of online interactions on patients health anxiety. Asso for Info Science & Tech 2023;74(11):1243-1264. [doi: 10.1002/asi.24822]
52. Griffin RJ, Dunwoody S, Neuwirth K. Proposed model of the relationship of risk information seeking and processing to the development of preventive behaviors. Environ Res 1999;80(2 Pt 2):S230-S245. [doi: 10.1006/enrs.1998.3940] [Medline: 10092438]
53. Zheng H, Kyung Kim H, Joanna Sin S, Theng Y. A theoretical model of cyberchondria development: Antecedents and intermediate processes. Telemat Inform 2021;63:101659. [doi: 10.1016/j.tele.2021.101659]
54. Honora A, Wang K, Chih W. How does information overload about COVID-19 vaccines influence individuals' vaccination intentions? The roles of cyberchondria, perceived risk, and vaccine skepticism. Comput Human Behav 2022;130:107176 [FREE Full text] [doi: 10.1016/j.chb.2021.107176] [Medline: 35013641]
55. Inzlicht M, Schmeichel BJ. What is ego depletion? Toward a mechanistic revision of the resource model of self-control. Perspect Psychol Sci 2012;7(5):450-463. [doi: 10.1177/1745691612454134] [Medline: 26168503]
56. Reinecke L, Hartmann T, Eden A. The guilty couch potato: the role of ego depletion in reducing recovery through media use. J Commun 2014;64(4):569-589. [doi: 10.1111/jcom.12107]
57. Ma J, Zhao D, Xu N, Yang J. The effectiveness of immersive virtual reality (VR) based mindfulness training on improvement mental-health in adults: a narrative systematic review. Explore (NY) 2023;19(3):310-318 [FREE Full text] [doi: 10.1016/j.explore.2022.08.001] [Medline: 36002363]
58. Shapiro SL, Carlson LE, Astin JA, Freedman B. Mechanisms of mindfulness. J Clin Psychol 2006;62(3):373-386. [doi: 10.1002/jclp.20237] [Medline: 16385481]

XSL•FO
RenderX

59. Cullen M. Mindfulness-based interventions: an emerging phenomenon. Mindfulness 2011;2(3):186-193. [doi: 10.1007/s12671-011-0058-1]
60. Hathaisaard C, Wannarit K, Pattanaseri K. Mindfulness-based interventions reducing and preventing stress and burnout in medical students: a systematic review and meta-analysis. Asian J Psychiatr 2022;69:102997. [doi: 10.1016/j.ajp.2021.102997] [Medline: 34995839]
61. Toivonen KI, Zernicke K, Carlson LE. Web-based mindfulness interventions for people with physical health conditions: systematic review. J Med Internet Res 2017;19(8):e303 [FREE Full text] [doi: 10.2196/jmir.7487] [Medline: 28860106]
62. Yogeswaran V, El Morr C. Effectiveness of online mindfulness interventions on medical students' mental health: a systematic review. BMC Public Health 2021;21(1):2293 [FREE Full text] [doi: 10.1186/s12889-021-12341-z] [Medline: 34920715]
63. Howell JL, Shepperd JA. Establishing an information avoidance scale. Psychol Assess 2016;28(12):1695-1708. [doi: 10.1037/pas0000315] [Medline: 27099980]
64. Jokić-Begić N, Mikac U, Čuržik D, Sangster Jokić C. The development and validation of the short Cyberchondria Scale (SCS). J Psychopathol Behav Assess 2019;41(4):662-676. [doi: 10.1007/s10862-019-09744-z]
65. Bailey R, Wells A. Development and initial validation of a measure of metacognitive beliefs in health anxiety: The MCQ-HA. Psychiatry Res 2015;230(3):871-877. [doi: 10.1016/j.psychres.2015.11.035] [Medline: 26626951]
66. Sheer VC. A meta-synthesis of health-related self-efficacy instrumentation: problems and suggestions. J Nurs Meas 2014;22(1):77-93. [doi: 10.1891/1061-3749.22.1.77] [Medline: 24851665]
67. Ajzen I. Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. J Applied Social Pyschol 2006;32(4):665-683. [doi: 10.1111/j.1559-1816.2002.tb00236.x]
68. Walsh G, Mitchell V. Consumer vulnerability to perceived product similarity problems: scale development and identification. J. Macromarketing 2005;25(2):140-152. [doi: 10.1177/0276146705280613]
69. Hair JF, Sarstedt M, Ringle CM, Mena JA. An assessment of the use of partial least squares structural equation modeling in marketing research. J. of the Acad. Mark. Sci 2011;40(3):414-433. [doi: 10.1007/s11747-011-0261-6]
70. Hair JF, Ringle CM, Gudergan SP, Fischer A, Nitzl C, Menictas C. Partial least squares structural equation modeling-based discrete choice modeling: an illustration in modeling retailer choice. Bus Res 2018;12(1):115-142. [doi: 10.1007/s40685-018-0072-4]
71. Fornell C, Larcker DF. Structural equation models with unobservable variables and measurement error: algebra and statistics. J. Mark. Res 1981;18(3):382-388. [doi: 10.2307/3150980]
72. Podsakoff PM, MacKenzie SB, Lee J, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
73. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: Indeed a silver bullet. J. Mark. Theory Pract 2014;19(2):139-152. [doi: 10.2753/MTP1069-6679190202]
74. Dijkstra TK, Henseler J. Consistent and asymptotically normal PLS estimators for linear structural equations. Comput Stat Data Anal 2015;81:10-23. [doi: 10.1016/j.csda.2014.07.008]
75. Hair JF, Risher JR, Sarstedt M, Ringle CM. When to use and how to report the results of PLS-SEM. Eur. Bus. Rev 2019;31(1):2-24. [doi: 10.1108/ebr-11-2018-0203]
76. Cohen J. Statistical power analysis for the behavioral science. Cambridge: Academic Press; 2013.
77. Brown KW, Ryan RM. The benefits of being present: mindfulness and its role in psychological well-being. J Pers Soc Psychol 2003;84(4):822-848. [doi: 10.1037/0022-3514.84.4.822] [Medline: 12703651]
78. Kabat-Zinn J. Mindfulness-based interventions in context: past, present, and future. Clinical Psychology: science and practice 2003;10(2):144-156. [doi: 10.1093/clipsy.bpg016]
79. Rooks JD, Morrison AB, Goolsarran M, Rogers SL, Jha AP. "We Are Talking About Practice": the influence of mindfulness vs. relaxation training on athletes' attention and well-being over high-demand intervals. J Cogn Enhanc 2017;1(2):141-153. [doi: 10.1007/s41465-017-0016-5]
80. Casale S, Rugai L, Fioravanti G. Exploring the role of positive metacognitions in explaining the association between the fear of missing out and social media addiction. Addict Behav 2018;85:83-87. [doi: 10.1016/j.addbeh.2018.05.020] [Medline: 29864680]
81. Muris P. Relationships between self-efficacy and symptoms of anxiety disorders and depression in a normal adolescent sample. Pers. Individ. Differ 2002;32(2):337-348. [doi: 10.1016/S0191-8869(01)00027-7]
82. Ahmed A, Aziz S, Toro CT, Alzubaidi M, Irshaidat S, Serhan HA, et al. Machine learning models to detect anxiety and depression through social media: a scoping review. Comput Methods Programs Biomed Update 2022;2:100066 [FREE Full text] [doi: 10.1016/j.cmpbup.2022.100066] [Medline: 36105318]
83. Ahn J, Kahlor LA. No regrets when it comes to your health: anticipated regret, subjective norms, information insufficiency and intent to seek health information from multiple sources. Health Commun 2020;35(10):1295-1302. [doi: 10.1080/10410236.2019.1626535] [Medline: 31177829]
84. Heid AR, Zarit SH, Fingerman KL. "My Parent is so stubborn!"-perceptions of aging parents' persistence, insistence, and resistance. J Gerontol B Psychol Sci Soc Sci 2016;71(4):602-612 [FREE Full text] [doi: 10.1093/geronb/gbu177] [Medline: 25573152]

XSL•FO
RenderX

85. Tudor K, Maloney S, Raja A, Baer R, Blakemore S, Byford S, MYRIAD Team, et al. Universal mindfulness training in schools for adolescents: a scoping review and conceptual model of moderators, mediators, and implementation factors. Prev Sci 2022;23(6):934-953 [FREE Full text] [doi: 10.1007/s11121-022-01361-9] [Medline: 35267177]
86. Korn L, Haynie DL, Luk JW, Sita K, Simons-Morton BG. Attitudes, subjective norms, and perceived behavioral control associated with age of first use of cannabis among adolescents. J Sch Health 2021;91(1):50-58. [doi: 10.1111/josh.12977] [Medline: 33216363]
87. Podina IR, Cosmoiu A, Rusu P, Chivu AC. Positive thinking is not adaptive thinking: a cognitive-behavioral take on interpretation bias modification for social anxiety. J Rat-Emo Cognitive-Behav Ther 2020;38(3):424-444. [doi: 10.1007/s10942-020-00344-5]

Abbreviations

ANCOVA: analysis of covariance
AVE: average variance extracted
CR: composite reliability
MBSR: mindfulness-based stress reduction
PLS: partial least squares
SDT: self-determination theory
VIF: variance inflation factor

Edited by A Mavragani; submitted 02.12.24; peer-reviewed by S You, Q Duan; comments to author 25.12.24; revised version received 29.12.24; accepted 30.12.24; published 28.01.25.

Please cite as:
Gu C, Qian L, Zhuo X
Mindfulness Intervention for Health Information Avoidance in Older Adults: Mixed Methods Study
JMIR Public Health Surveill 2025;11:e69554
URL: https://publichealth.jmir.org/2025/1/e69554
doi:10.2196/69554
PMID:

©Chenyu Gu, Liquan Qian, Xiaojie Zhuo. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Alternative Presentations of Overall and Statistical Uncertainty for Adults' Understanding of the Results of a Randomized Trial of a Public Health Intervention: Parallel Web-Based Randomized Trials

Christine Holst[1,2], PhD; Steven Woloshin[2,3], MD, MS; Andrew D Oxman[1,2], MD; Christopher Rose[1,4], PhD; Sarah Rosenbaum[1,2], PhD; Heather Menzies Munthe-Kaas[1,2], PhD

[1]Centre for Epidemic Interventions Research, Norwegian Institute of Public Health, Oslo, Norway

[2]Lisa Schwartz Foundation for Truth in Medicine, Norwich, VT, United States

[3]Dartmouth Institute for Health Policy and Clinical Practice, Geisel Medical School, West Lebanon, NH, United States

[4]Cluster for Reviews and Health Technology Assessments, Division for Health Services, Norwegian Institute of Public Health, Oslo, Norway

Corresponding Author:
Christine Holst, PhD
Centre for Epidemic Interventions Research
Norwegian Institute of Public Health
Myrens Verksted 3L
Oslo, 0213
Norway
Phone: 47 48234044
Email: Christine.Holst@fhi.no

Abstract

Background: Well-designed public health messages can help people make informed choices, while poorly designed messages or persuasive messages can confuse, lead to poorly informed decisions, and diminish trust in health authorities and research. Communicating uncertainties to the public about the results of health research is challenging, necessitating research on effective ways to disseminate this important aspect of randomized trials.

Objective: This study aimed to evaluate people's understanding of overall and statistical uncertainty when presented with alternative ways of expressing randomized trial results.

Methods: Two parallel, web-based, individually randomized trials (3×2 factorial designs) were conducted in the United States and Norway. Participants were randomized to 1 of 6 versions of a text (summary) communicating results from a study examining the effects of wearing glasses to prevent COVID-19 infection. The summaries varied in how *overall uncertainty* ("Grading of Recommendations Assessment, Development and Evaluation [GRADE] language," "plain language," or "no explicit language") and *statistical uncertainty* (whether a margin of error was shown or not) were presented. Participants completed a web-based questionnaire exploring 4 coprimary outcomes: 3 to measure understanding of overall uncertainty (benefits, harms, and sufficiency of evidence), and one to measure statistical uncertainty. Participants were adults who do not wear glasses recruited from web-based research panels in the United States and Norway. Results of the trials were analyzed separately and combined in a meta-analysis.

Results: In the US and Norwegian trials, 730 and 497 individuals were randomized, respectively; data for 543 (74.4%) and 452 (90.9%) were analyzed. More participants had a correct understanding of uncertainty when presented with plain language (United States: 37/99, 37% and Norway: 40/76, 53%) than no explicit language (United States: 18/86, 21% and Norway: 34/80, 42%). Similar positive effect was seen for the GRADE language in the United States (26/79, 33%) but not in Norway (30/71, 42%). There were only small differences between groups for understanding the uncertainty of harms. Plain language improved correct understanding of evidence sufficiency (odds ratio 2.05, 95% CI 1.17-3.57), compared to no explicit language. The effect of GRADE language was inconclusive (odds ratio 1.34, 95% CI 0.79-2.28). The understanding of statistical uncertainty was improved when the participants were shown the margin of error compared to not being shown: Norway: 16/75, 21% to 24/71, 34% vs 1/71, 1% to 2/76, 3% and the United States: 21/101, 21% to 32/90, 36% vs 0/86, 0% to 3/79, 4%).

XSL•FO
RenderX

Conclusions: Plain language, but not GRADE language, was better than no explicit language in helping people understand overall uncertainty of benefits and harms. Reporting margin of error improved understanding of statistical uncertainty around the effect of wearing glasses, but only for a minority of participants.

Trial Registration: ClinicalTrials.gov NCT05642754; https://tinyurl.com/4mhjsm7s

(JMIR Public Health Surveill 2025;11:e62828) doi:10.2196/62828

KEYWORDS

communication; Grading of Recommendations Assessment, Development, and Evaluation language; GRADE language; statistical uncertainty; overall uncertainty; randomized trial

Introduction

Background

Public health messaging matters—it shapes how people understand important health risks and what can be done to mitigate them. Ideally, such messages are based on findings from robust research, but even important messages based on solid research can fail to properly reach the target audience when poorly communicated, as was seen during the COVID-19 pandemic [1]. Decision scientists have articulated basic principles for effective health communication, such as using simple and familiar wording, using clear visual design, presenting structured comparisons of alternatives, and careful testing in the target audiences [2-5].

A recent randomized trial assessed the effect of these principles on communication effectiveness in the context of COVID-19 home test kit instructions using a real example [6]. The trial showed that individuals randomized to instructions that did not follow best decision science principles (ie, the actual instructions authorized by the Food and Drug Administration), compared to those that did (ie, carefully pretested intervention instructions), were more likely to fail to quarantine when quarantine was the right choice (33% vs 14% failed; 95% CI for the 19% difference being 6%-31%).

Evidence from randomized trials documents the importance of both the format and content of health messaging. Formatting examples include how the use of percentages (eg, 10%) versus frequency formats (eg, 10 in 100) can improve comprehension [7] and how absolute versus relative risk measures for communicating treatment effects are better understood [3,8] and help people make decisions more consistent with their values [9].

Content examples include the importance of presenting both benefits and harms when describing interventions [10] and highlighting study limitations, such as how simple nondirective explanations about surrogate outcomes and newly approved drugs enhance evidence-based decision-making about prescription drugs [11,12].

Furthermore, there is evidence supporting the importance of communication about the uncertainty of research findings in both research summaries aimed at professionals [13-16] and plain language summaries for the public [17]. This includes both *statistical uncertainty* (ie, imprecision) [16] and the *overall uncertainty*, due to the risk of bias, inconsistency, indirectness, and publication bias [13,15]. However, there are still open questions about the best formats and language for presenting both kinds of uncertainty, and how people understand and react to such information.

A recent pair of trials [18] found that including “quality cues” in communications tempered the publics’ tendency to assume that the quality of evidence presented without such cues is high (when it was not), and reporting that evidence quality was low decreased trust, perception of intervention efficacy, and the likelihood of adopting it.

The certainty of the evidence can affect the decisions that people make. If the purpose of a message is to inform people rather than to persuade them [19], it is necessary to include information about the degree and source of uncertainty related to the effect estimate. Not doing so can be misleading.

This study is the first of several planned studies to evaluate strategies to improve communication of research findings (Message Lab). The goal of Message Lab is to develop and promote best practices in message development (ie, attending to the foregoing communication principles and evidence); facilitate user testing; and conduct randomized trials assessing the effects of public health messages on the public’s understanding of the messages, beliefs, decisions, and behaviors. Therefore, this study is also designed as a proof-of-concept exercise to develop and test a method to efficiently and effectively evaluate communication strategies intended to summarize the results of randomized trials using web-based trial platforms.

Objectives

This study aimed to evaluate the effect of alternative formats for communicating overall and statistical uncertainty on the public’s understanding of uncertainty and the sufficiency of evidence. The objectives of the study were (1) to compare the effects of 3 ways of communicating the overall uncertainty of the effect of wearing glasses on reducing the chance of acquiring COVID-19, and (2) to compare the effects of including the margin of error (MOE; also called CI) compared to not including it.

Methods

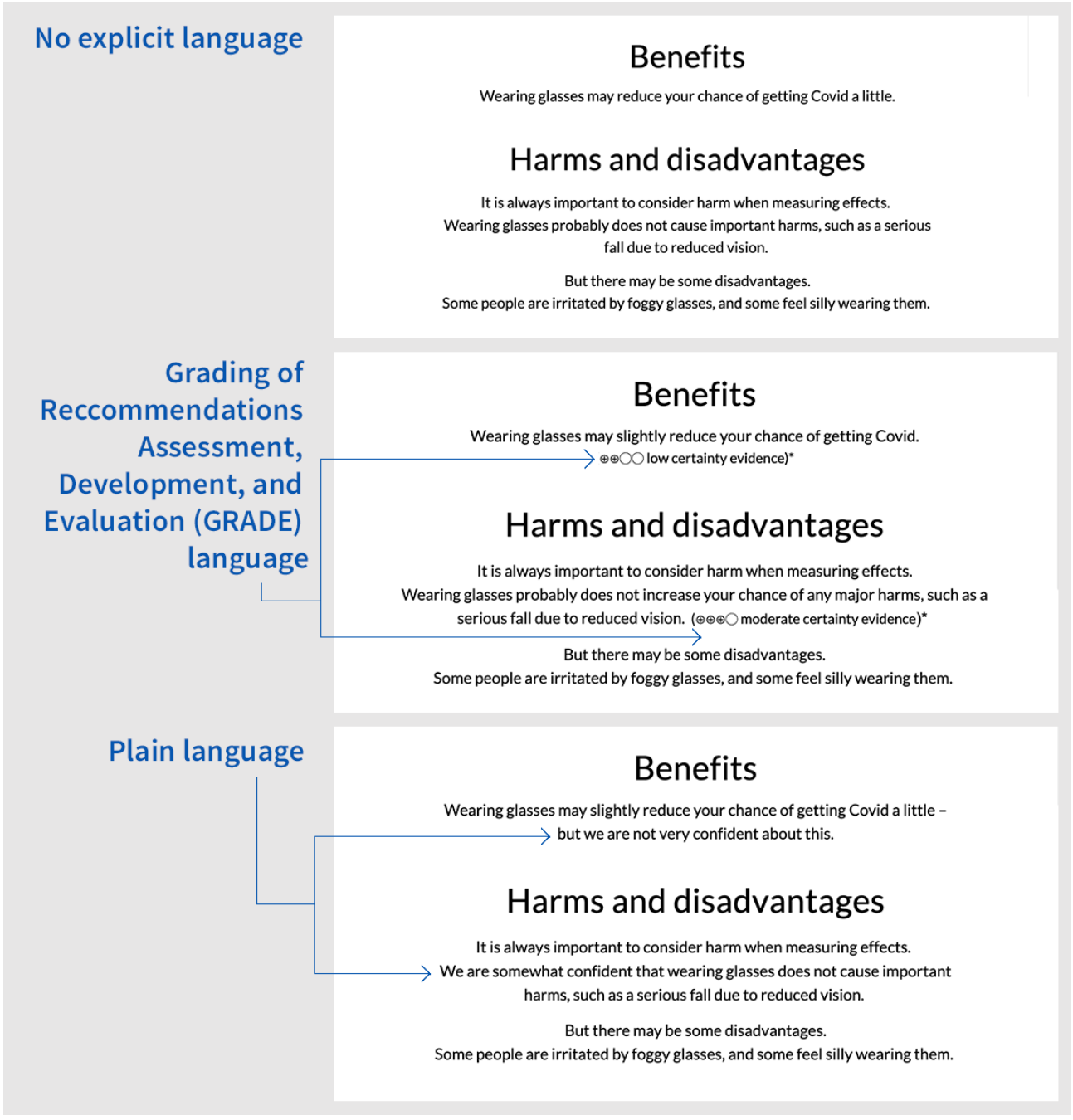
Design

We designed a web-based, parallel-group, individually randomized, pragmatic trial to compare the effects of different ways of communicating uncertainty when reporting the results of a randomized trial to the public. The trial was prespecified,

XSL•FO
RenderX

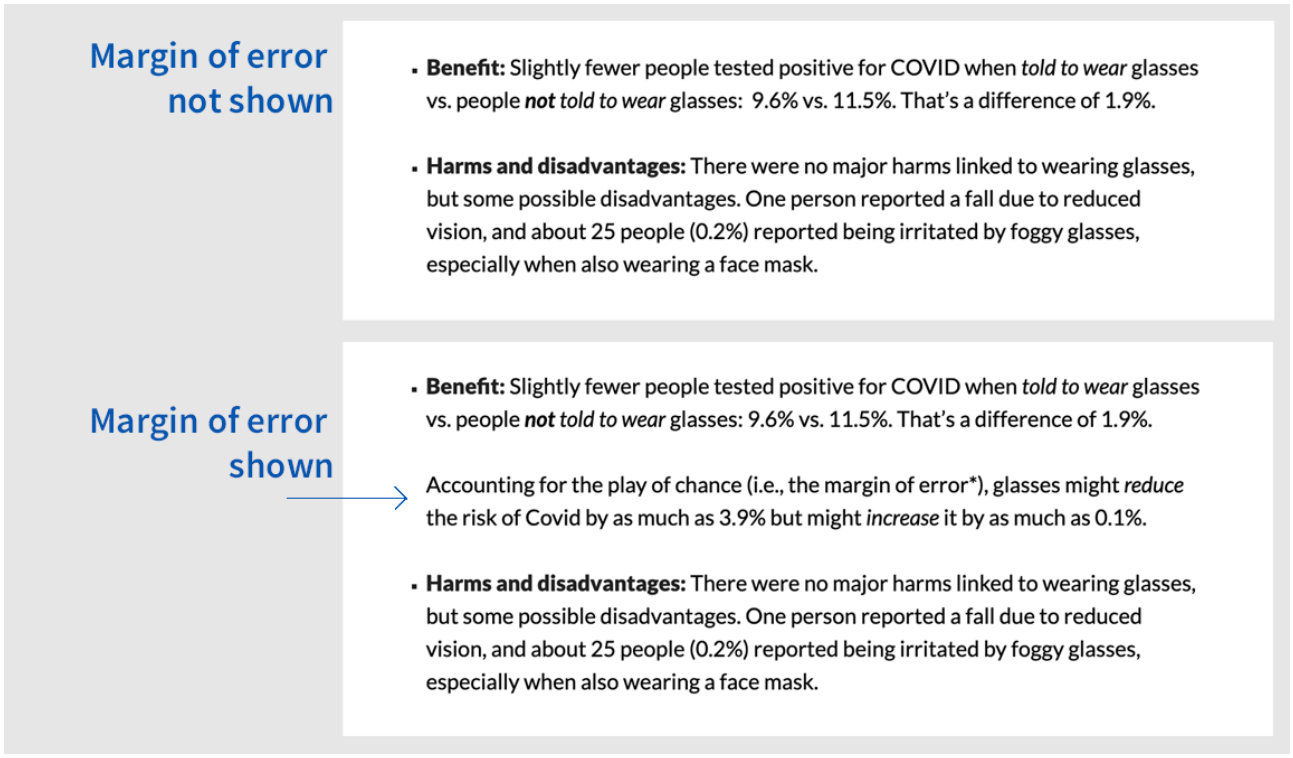
registered at ClinicalTrials.gov (NCT05642754) Zenodo (7428981), and, except as noted, conducted following a published protocol [20]. We used a published trial assessing the effect of wearing glasses on the risk of being infected with COVID-19 as our example [21]. We used a 3×2 factorial trial design because we were interested in the effects of presenting overall uncertainty in each of 3 ways (ie, Grading of Recommendations Assessment, Development and Evaluation [GRADE] language, plain language, or no explicit language) combined with the effects of presenting or not presenting statistical uncertainty (in the form of an MOE). This resulted in 6 (3×2) groups differing in how the COVID-19 study was summarized (Figures 1 and 2).

Figure 1. Three alternative versions presented to participants in the US trial communicating overall uncertainty about the possible benefits and harms of wearing glasses to reduce the chances of acquiring COVID-19.



XSL•FO
RenderX

Figure 2. Two alternative versions presented to participants in the US trial communicating statistical uncertainty (margin of error).



Recruitment of Participants

We implemented the design in 2 nearly identical trials in Norway and the United States in April 2023 and May 2023. For each trial, we used quota sampling from a web-based platform (Prolific; Prolific Academic Ltd) of volunteer research participants in the US trial [22] and from an independent commercial research agency (Opinion) [23] with a panel of 120,000 people living in Norway. All participants were thus part of established research panels, which provide recruitment and management of participants for online research. Participants were eligible if they were aged ≥18 years and said that they did not regularly wear glasses (based on a prescreen feature in Prolific and a screener question in Opinion). In Norway, 2667 persons received an invitation from Opinion to participate, while the study was open to 34,242 eligible persons in the Prolific platform (see the Results section for more details). Participants were literate in English (used in the US version) or Norwegian (used in the Norwegian version). Both platforms applied processes to prevent bots and fraudulent participants [24], and to further increase data quality, we built in attention and comprehension checks [25].

Randomization

Eligible participants were randomized in a 1:1:1 ratio to 1 of the 6 comparison groups (described above). In the United States, participants were randomized via the web-based questionnaire after clicking on the link to the summary. In Norway, the participants were randomized to each summary by the research agency.

Interventions and Control

Overview

The alternative versions of the presentation of overall and statistical uncertainty included in the 6 summaries are provided in Figures 1 and 2, and the full summaries are provided in Multimedia Appendix 1. Each version included one of the alternatives for communicating overall uncertainty about the benefits and harms (ie, GRADE, plain, or no explicit language) and either included or did not include a presentation of the MOE:

- Version 1: no explicit language and MOE not shown (reference or control version)
- Version 2: no explicit language and MOE shown
- Version 3: plain language and MOE not shown
- Version 4: plain language and MOE shown
- Version 5: GRADE language and MOE not shown
- Version 6: GRADE language and MOE shown

Development and User Testing of the Interventions

The summaries were drafted in English by 4 researchers (SW, AO, HMK, SR) and translated to Norwegian by 2 researchers (IHE, CH). Note that the summaries underwent multiple rounds of user testing and modifications as needed before the trial. We used human-centered design methods to develop the summaries to present to the participants [26]. Initially, we invited 4 native English speakers and 3 native Norwegian speakers to unmoderated user testing using Loop11, a digital user experience platform [27]. In total, 7 user test participants (n=5, 71% female and n=2, 29% male; n=2, 29% with secondary school education and n=5, 71% with postsecondary education; n=2, 29% living in Canada and n=5, 71% living in Norway; and n=3, 43% aged >55 years and n=4, 57% aged between 30 and 45 years) were introduced to the project and asked to imagine themselves in the following scenario: “Imagine you hear that glasses may

XSL•FO
RenderX

reduce your chance of becoming infected with COVID. You go online to find out more information and find a website that says... [insert COVID trial summary version 1, 2 or 3]."

They were then asked to read 5 pieces of text: 3 versions of the summaries with progressively more information (eg, the first version was the shortest summary with no text related to uncertainty or the MOE, the second version had text related to uncertainty but not the MOE, and the third version had text related to uncertainty and the MOE); the trial questions; and the text that would be available via hyperlinks and hover text in the summaries. For each piece of text, the user test participants w asked for their first set of impressions and then a series of follow-up questions related to content, font, format, language, and anything else.

We revised the summaries according to feedback. The translations were reviewed by a third researcher. The English summaries were presented to 1 more participant in English (moderated, in-person user experience interview) using the same format as the first 4 interviews. The summaries were subsequently revised, and changes were also made to the Norwegian summaries.

We gathered feedback on the Norwegian summaries from colleagues and conducted 2 in-person user experience interviews with 2 native Norwegian speakers using the same question guide as the English user experience interviews. Revisions were made according to the feedback, and where appropriate, these revisions were back translated into the English versions of the COVID-19 trial summaries.

Furthermore, we gathered feedback from the user experience participants on other materials related to the process of participating in the randomized trial (eg, from the invitation to participate to the text sent after participants completed the questionnaire).

Outcomes

We defined 3 binary coprimary outcomes to measure understanding of overall uncertainty and 1 coprimary outcome to measure understanding of statistical uncertainty (precision of the effect estimate for the benefit). Each of the 3 outcomes for overall uncertainty were measured by comparing participants' answers to questions (Table 1) about uncertainty to expected (correct) answers based on the size of the effects and the certainty of the evidence (the full questionnaire is provided in the Multimedia Appendix 2).

The 3 *coprimary outcomes* were as follows (all used 4-point ordinal response sets):

- Understanding of the uncertainty of the benefit (question 9)
- Understanding of the sufficiency of the evidence (question 15)
- Understanding of the uncertainty of important harms (question 12)

An additional coprimary outcome was included for the 3 versions reporting the MOE, assessing the understanding of statistical uncertainty (ie, the precision of the effect estimates: choose which among the 4 statements was most consistent with the information provided, eg, question 10: "wearing glasses may *reduce* the chance of COVID a little, but might *reduce it a lot*").

Secondary outcomes included questions (Table 1) regarding the following:

- The perceived benefit and harms of wearing glasses to reduce the chance of acquiring COVID-19 infections (questions 8 and 11)
- Intended behavior (whether participants would wear glasses to reduce the chance of acquiring COVID-19 in areas with high and low COVID-19 infection rates; questions 3 and 7)
- Perceptions of the information; trustworthiness (question 13), sufficiency (question 14), clarity (questions 17 and 18), and helpfulness (question 19); and the likelihood of sharing it with others (question 20)
- Decisional conflict (questions 4 and 5) [28]

XSL•FO
RenderX

Table 1. Outcomes and the corresponding questionnaire items.

Outcome	Questionnaire
Primary outcomes	
Understanding uncertainty	• Question 9. How sure are you about the effect of wearing glasses on your chance of getting COVID?
Understanding uncertainty of important harms	• Question 12. How sure are you about whether wearing glasses to reduce COVID can cause important harms?
Understanding sufficiency	• Question 15. Not enough is known to be sure about the effects of wearing glasses to reduce the chance of getting COVID.
Understanding statistical uncertainty	• Question 10. Which of the following statements is most consistent with the information provided?
Secondary outcomes	
Perceived benefit and harm	• Question 8. What is the possible effect of wearing glasses on your chance of getting COVID? • Question 11. How likely do you think it is that wearing glasses to reduce the chance of getting COVID can cause any important harms?
Intended behavior	• Question 3. If there were a surge of COVID-19 cases in your area, how likely would you be to wear glasses or recommend wearing glasses to reduce the chance of getting COVID? • Question 7. If there were very few COVID cases in your area, how likely would you be to wear glasses or recommend wearing glasses to reduce the chance of getting COVID?
Perceptions of the information	• Question 13. This information seems like a trustworthy summary of what is known about the effects of wearing glasses to reduce the chance of getting COVID. • Question 14. The summary gives me enough information to understand what is known about the effects of wearing glasses to reduce the chance of getting COVID. • Question 17. I think the information about whether wearing glasses affects the chance of getting COVID was... • Question 18. I think the information about whether wearing glasses to prevent COVID has important harms was... • Question 19. If you were making a decision about wearing glasses to prevent COVID, would you find the information we showed you helpful? • Question 20. Say you knew someone who heard that wearing glasses might affect your chance of getting COVID. How likely would you be to share the information you just saw with them?
Decisional conflict	• Question 4. The answer about wearing glasses if there were a surge of COVID was hard for me to give. • Question 5. The information in the summary helped me make an informed decision about wearing glasses if there were a surge of COVID.

Data Collection

In the United States, we collected data using SurveyMonkey (SurveyMonkey Inc). Participants were directed to the platform after agreeing to participate in the study. They were asked to enter their personal identifying number (as a panel member on the relevant recruiting platform). In Norway, Opinion collected the data in their own panel via an email invitation and a link to the survey on their own platform. After reading initial information, the participants were randomized to read one of the 6 summaries; they were asked to answer 18 questions about the summary; 4 questions about themselves (eg, whether they wore glasses, what was their highest level of education, and what was their concern regarding COVID-19) to assess saliency of the scenario; and 3 questions to assess their numeracy (Multimedia Appendix 2). They submitted their responses electronically, using the “submit” button (refer to Statistical Analysis section).

Statistical Analysis

The US and Norwegian trials were analyzed in the same way, except as noted. We excluded participants who completed the survey in <3 minutes, reported regularly wearing glasses, or failed the attention checks. The attention checks (questions 6 and 21) included 2 true or false questions to verify the consistency of participants’ responses, ensuring that they were paying attention and not just randomly answering. Data were duplicated for some US participants who appeared to have submitted the same responses multiple times in a short period. We assumed that these participants clicked the submit button several times or refreshed their browsers, so we analyzed only the first data submitted by these participants. All analyses were performed before unblinding as prespecified according to the intention-to-treat principle—all randomized participants meeting the inclusion criteria were included and analyzed in the arms to which they were randomized [20].

All outcomes were binomial. We used logistic regression to estimate odds ratios (ORs) for the treatments and their

XSL•FO
RenderX

interactions. Model fit was assessed using the Hosmer-Lemeshow test. To aid in interpretation, we reexpressed ORs as risk differences, accounting for statistical uncertainty on baseline odds, main effects, and interactions. No data were missing for any of the participants meeting the inclusion criteria.

We used fixed effects meta-analysis to pool estimates across the trials, obtain overall estimates of effect, and assess the country as a potential effect modifier. We performed prespecified subgroup analyses for both trials to explore differences in treatment effect with respect to numeracy and, in the US trial, saliency. Numeracy was defined as scoring 3 (vs <3) on a validated instrument [29]. Saliency was defined as being very or extremely worried about acquiring COVID-19 and considering it very or extremely important to take action to reduce the chance of acquiring COVID-19. It was not possible to perform the analysis for saliency for the Norwegian trial because only 20 participants met the saliency criteria.

Furthermore, we performed nonprespecified analyses for potential effect modification. In the US trial, we explored effect modification by response time (responding within 7 min vs >7 min). The choice of 7 minutes was data driven, chosen to be the whole number of minutes closest to the median response time. We also explored effect modification by education level (having graduated vs not having graduated from a university).

We presented 2-sided 95% CIs and *P* values, wherever applicable, throughout the study. Meta- and subgroup analyses were presented using forest plots, with *P* values testing null hypotheses of homogeneity (there was no difference between the estimates for the 2 trials or no effect modification). All analyses were performed using Stata 18 (StataCorp LLC).

Ethical Considerations

This study was considered for ethics approval by the Regional Committee for Medical and Health Research Ethics in Norway and was found not to require ethics approval because it falls outside the committee's mandate under the Health Research Act (reference 557972).

It was also deemed exempt from further review by the Committee for the Protection of Human Subjects (STUDY00032615) and the Dartmouth College Institutional Review Board.

By clicking on the study link in Prolific and Opinions platforms, participants implicitly consented to take part in the study. Everyone invited to participate were given information about who was conducting the research (Dartmouth University was emphasized in the US trial, while Norwegian Institute of Public Health was emphasized in the Norwegian trial), the aim of the research, what they would be asked to do, and how long it would take (Multimedia Appendix 3). It was clarified that participation was voluntary and that participants could discontinue participation at any time. We did not collect personal data from the participants (as per policy of the recruitment agencies), all collected data were anonymous, and we could not trace any data back to the participants. We did not inform the participants how the data would be stored or for how long; however, we did provide contact information to participants to ask any questions.

However, the final process of obtaining informed consent did not completely align with the procedures described in the study protocol [20]. During user testing, we received feedback that the introductory text was too long and cumbersome. Thus, we attempted to shorten it. Given that the study was completely anonymous (no personal or identifying information was gathered from the participants), posed minimal risk, and involved consenting members of the web-based panel platforms we used, we decided that we did not need to provide information regarding how data would be stored. In addition, we could remove a specific sentence asking for informed consent, as implied consent via clicking on the link to participate in the study was deemed sufficient. However, considering emerging best practice guidelines for web-based trials, we will provide more information and seek explicit consent in future trials.

We paid participants 100 Norwegian Kroner (US $8.40) and US $4.30 (50 Norwegian Kroner) in the Norwegian and US trials, respectively.

Patient and Public Involvement

We involved members of the public in user testing during the development of the survey and COVID-19 trial summaries.

Results

Overview

We invited 2667 Norwegians to participate in the trial, of which 1782 (66.82%) did not respond, 388 (14.55%) were not eligible, and 497 (18.63%) were randomized to one of the 6 intervention groups (Figure 3).

Of the 497 participants, 45 (9%) were excluded from the analysis for the reasons provided in Figure 3.

The study was open for 34,242 people in the United States to participate in the trial, of which 33,512 (97.88%) did not respond in time and 730 (2.13%; required sample size) were randomized to one of the 6 intervention groups (Figure 4).

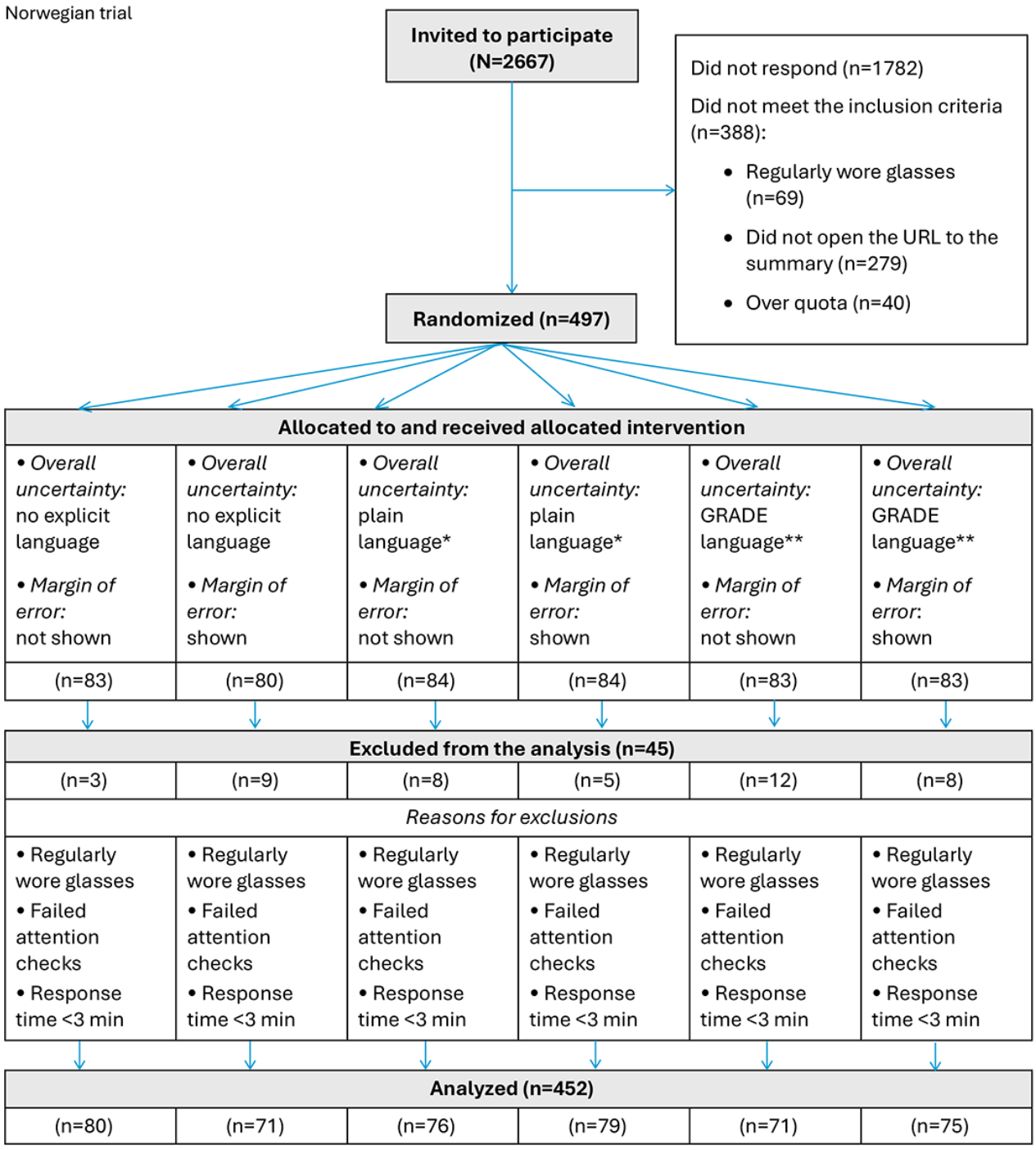
Of the 730 participants, 187 (25.6%) were later excluded from the analysis for the reasons provided in Figure 4. The average age of participants was 45 (SD 16.8) years in the Norwegian trial and 38 (SD 13.2) years in the US trial (Table 2).

In total, 50% (226/452) of the Norwegians and 48.3% (262/543) of the US participants were female. The educational level was higher in the US participants (284/543, 52.3% had at least a college degree) than in the Norwegian participants (197/452, 43.6% had at least a college degree). Altogether, approximately half (517/995, 52%) of participants in both trials failed to answer all 3 numeracy questions correctly (Table 2).

There were only minor differences among the comparison groups in both the Norwegian and US trials (Tables S1 and S2 in Multimedia Appendix 4).

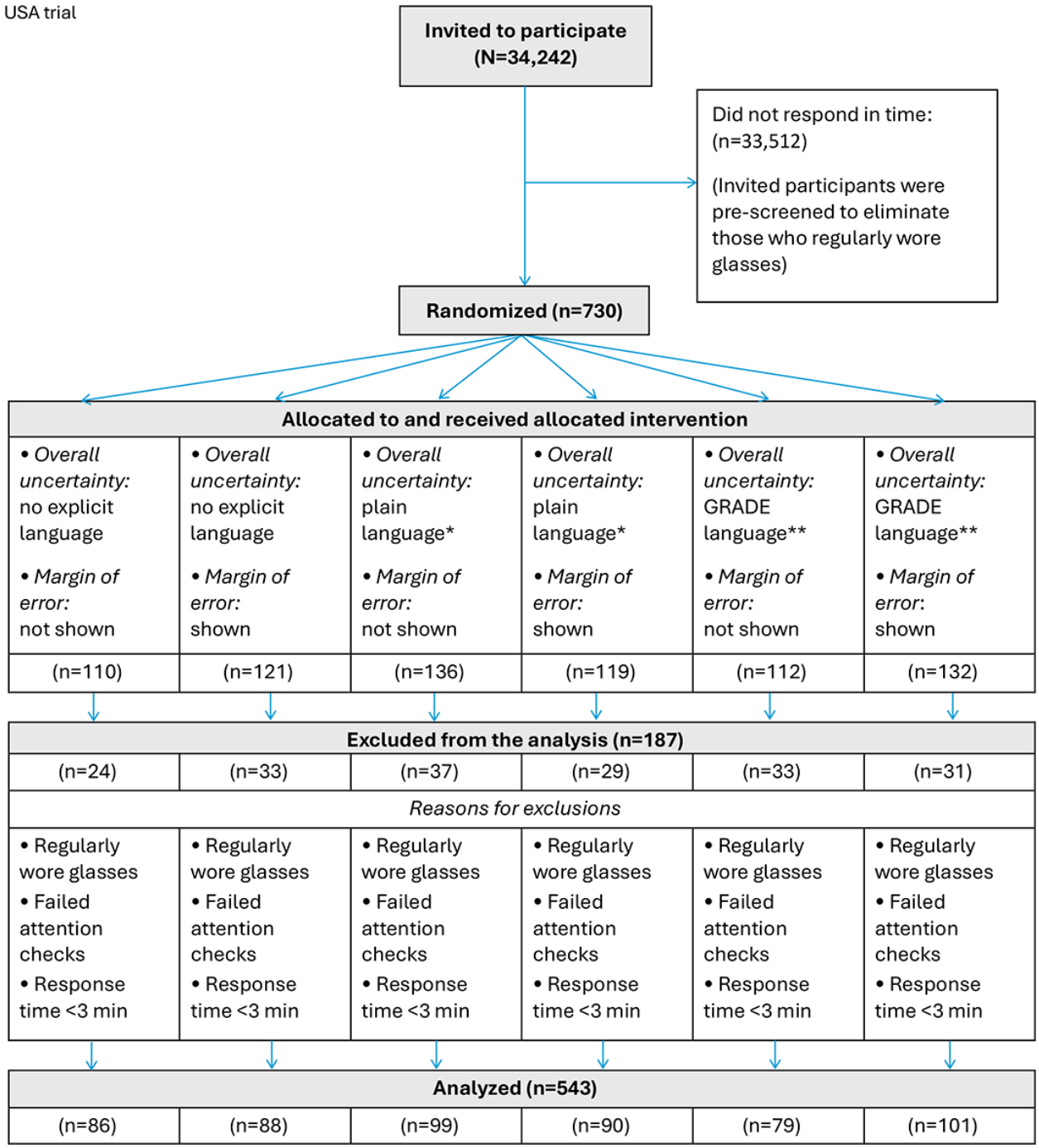
XSL•FO
RenderX

Figure 3. Modified CONSORT (Consolidated Standards of Reporting Trials) flow diagram of the participant inclusion process in April 2023. The flow diagram shows the eligible, recruited, and allocated participants from the Opinion panel in the Norwegian randomized controlled trial. *A less formal expression of the overall uncertainty language used in ordinary or familiar conversation, corresponding to the same Grading of Recommendations Assessment, Development and Evaluation (GRADE) assessment of the certainty of the evidence. **On the basis of the Cochrane Effective Practice and Organization of Care Group's guidance for communicating the certainty of evidence based on the GRADE approach to assessing the certainty of evidence [30].



XSL•FO
RenderX

Figure 4. Modified CONSORT (Consolidated Standards of Reporting Trials) flow diagram of the participant inclusion process in April 2023. The flow diagram shows the eligible, recruited, and allocated participants from the Prolific panel in the US randomized controlled trial. *A less formal expression of the overall uncertainty language used in ordinary or familiar conversation, corresponding to the same Grading of Recommendations Assessment, Development and Evaluation (GRADE) assessment of the certainty of the evidence. **On the basis of the Cochrane Effective Practice and Organization of Care Group's guidance for communicating the certainty of evidence based on the GRADE approach to assessing the certainty of evidence [30].



XSL•FO
RenderX

Table 2. Characteristics of the Norwegian and American participants from the web-based research panels taking part in the randomized controlled trial on evaluating the people's understanding of overall and statistical uncertainty when presented with alternative expressions in April 2023.

	Norwegian trial (n=452)	US trial (n=543)	Total (N=995)
Age (y), mean (SD)	45 (16.8)	38 (13.2)	41 (15.4)
Age group (y), n (%)			
18 to 29	109 (24.1)	155 (28.5)	264 (26.5)
30 to 39	74 (16.4)	195 (35.9)	269 (27)
40 to 49	81 (17.9)	98 (18)	179 (18)
50 to 59	88 (19.5)	27 (5)	115 (11.6)
60 to 69	59 (13.1)	58 (10.7)	117 (11.8)
≥70	41 (9.1)	10 (1.8)	51 (5.1)
Sex, n (%)			
Female	226 (50)	262 (48.3)	488 (49)
Male	226 (50)	281 (51.7)	507 (51)
Employment status, n (%)			
Full time	209 (46.2)	213 (39.2)	422 (42.4)
Part time	34 (7.5)	65 (12)	99 (9.9)
Unpaid[a]	—[b]	55 (10.1)	55 (5.5)
Unemployed or seeking work	—	51 (9.4)	51 (5.1)
Welfare[c]	114 (25.2)	—	114 (11.5)
Self-employed	22 (4.9)	—	22 (2.2)
Missing	17 (3.8)	129 (23.8)	146 (14.7)
Student	49 (10.8)	—	49 (4.9)
Other	7 (1.5)	30 (5.5)	37 (3.7)
Highest education, n (%)			
<High school	33 (7.3)	25 (4.6)	58 (5.8)
High school degree	131 (29)	92 (16.9)	223 (22.4)
Some college	91 (20.1)	142 (26.2)	233 (23.4)
College degree	124 (27.4)	204 (37.6)	328 (33)
Graduate or professional school	73 (16.2)	80 (14.7)	153 (15.4)
Numeracy, n (%)			
<3	235 (52)	282 (51.9)	517 (52)
3	217 (48)	261 (48.1)	478 (48)

[a]Includes participants who are homemakers, retired, or disabled.

[b]Not applicable.

[c]Includes participants in any welfare program, maternity leave, pensioner, or unemployment benefits.

Understanding of the Uncertainty of the Benefit

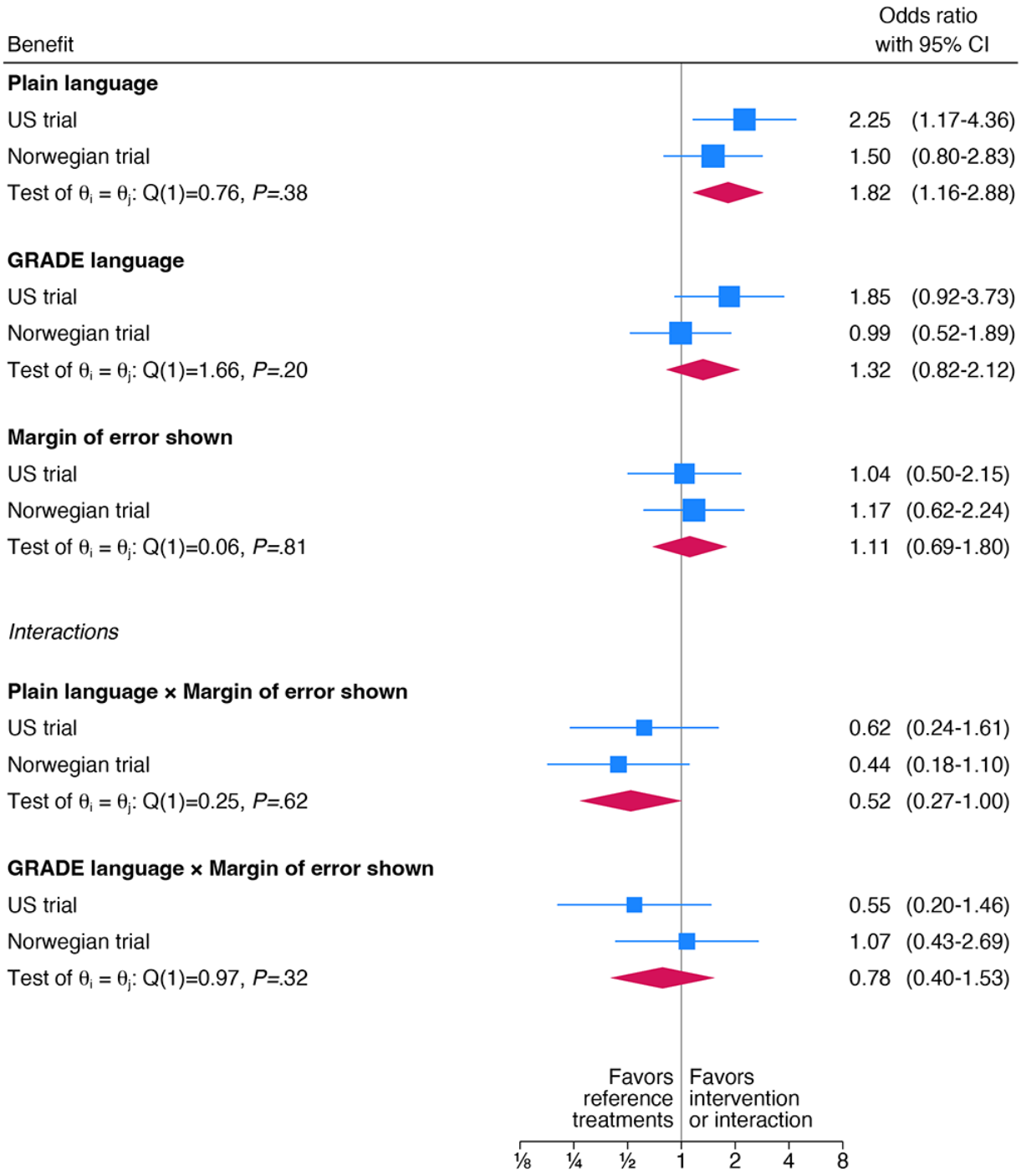
Overall

Overall, plain (but not GRADE) language, compared to no explicit language, improved correct understanding of the overall uncertainty of the effect of wearing glasses on the chance of acquiring COVID-19 (ie, "mixed but more unsure than sure;" OR 1.82, 95% CI 1.16-2.88; Figure 5).

When the MOE was shown together with plain language, the effect on understanding of the overall uncertainty was substantially reduced (OR 0.52, 95% CI 0.27-1.00; Figure 5): participants were more likely to perceive the evidence as "very unsure" rather than "mixed but more unsure than sure" (the correct response) (Table S3 and S4 in Multimedia Appendix 4).

XSL•FO
RenderX

Figure 5. Meta-analysis of the understanding of the uncertainty of the benefit of wearing glasses. Odds ratios for answering "Mixed but more unsure than sure" to the question "How sure are you about the effect of wearing glasses on your chance of getting COVID?" in the Norwegian and US randomized controlled trials on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions. GRADE: Grading of Recommendations Assessment, Development and Evaluation.



Individual Trials

Overall, on average, across all 6 comparison groups, there were 25.2% (114/452) of Norwegian participants who responded that the evidence was very unsure compared to US participants, of whom 16.4% (89/543) reported the same (Table S3 and S4 in Multimedia Appendix 4).

The risk difference for the Norwegian trial was less certain (risk difference of 10.1%, 95% CI –5.5 to 25.7; Table 3) than in the US trial.

In the US trial, using plain language increased correct understanding of how sure we can be about the effect of wearing glasses on the chance of acquiring COVID-19, from 21% (18/86) to 37% (37/99; risk difference 16.4%; 95% CI 3.6-29.3; Table 4).

XSL•FO
RenderX

Table 3. Risk differences for the understanding of the overall uncertainty of the benefit of wearing glasses in the Norwegian randomized controlled trial on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions for the question "How sure are you about the effect of wearing glasses on your chance of getting COVID?" with responses including very sure, mixed but more sure than unsure, mixed but more unsure than sure (correct), and very unsure.

Uncertainty	Margin of error	Participants[a], n (%)	Odds ratio[b] (95% CI)	Risk difference[c] (%), (95% CI)	*P* value
No explicit language (n=80)	Not shown	34 (42.5)	1	0	—[d]
No explicit language (n=71)	Shown	33 (46.5)	1.17 (0.62 to 2.24)	4.0 (–11.89 to 19.85)	.62
Plain language (n=76)	Not shown	40 (52.6)	1.50 (0.80 to 2.83)	10.1 (–5.47 to 25.73)	.21
Plain language (n=79)	Shown	29 (36.7)	0.78 (0.42 to 1.48)	–5.8 (–20.97 to 9.39)	.46
GRADE[e] language (n=71)	Not shown	30 (42.3)	0.99 (0.52 to 1.89)	–0.3 (–16.04 to 15.54)	.98
GRADE language (n=75)	Shown	36 (48)	1.25 (0.66 to 2.35)	5.5 (–10.16 to 21.16)	.49

[a]Participants randomized to the intervention answering correctly or as anticipated.

[b]Odds ratios include the main and interaction effects.

[c]Risk differences account for uncertainty on the baseline odds.

[d]Not applicable.

[e]GRADE: Grading of Recommendations Assessment, Development and Evaluation.

Table 4. Risk differences for the understanding of the overall uncertainty of the benefit of wearing glasses in the US randomized controlled trial on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions for the question "How sure are you about the effect of wearing glasses on your chance of getting COVID?" with responses including very sure, mixed but more sure than unsure, mixed but more unsure than sure (correct), and very unsure.

Overall uncertainty	Margin of error	Participants[a], n (%)	Odds ratio (95% CI)[b]	Risk difference[c] (%), (95% CI)	*P* value
No explicit language (n=86)	Not shown	18 (20.9)	1	0	—[d]
No explicit language (n=88)	Shown	19 (21.6)	1.04 (0.50 to 2.15)	0.7 (–11.5 to 12.8]	.92
Plain language (n=99)	Not shown	37 (37.4)	2.25 (1.17 to 4.36)	16.4 (3.6 to 29.3)	.02
Plain language (n=90)	Shown	25 (27.8)	1.45 (0.73 to 2.91)	6.8 (–5.8 to 19.5)	.29
GRADE[e] language (n=79)	Not shown	26 (32.9)	1.85 (0.92 to 3.73)	12.0 (–1.5 to 25.4)	.08
GRADE language (n=101)	Shown	22 (21.8)	1.05 (0.52 to 2.12)	0.9 (–10.9 to 12.6)	.89

[a]Participants randomized to the intervention answering correctly or as anticipated.

[b]Odds ratios include the main and interaction effects.

[c]Risk differences account for uncertainty on the baseline odds.

[d]Not applicable.

[e]GRADE: Grading of Recommendations Assessment, Development and Evaluation.

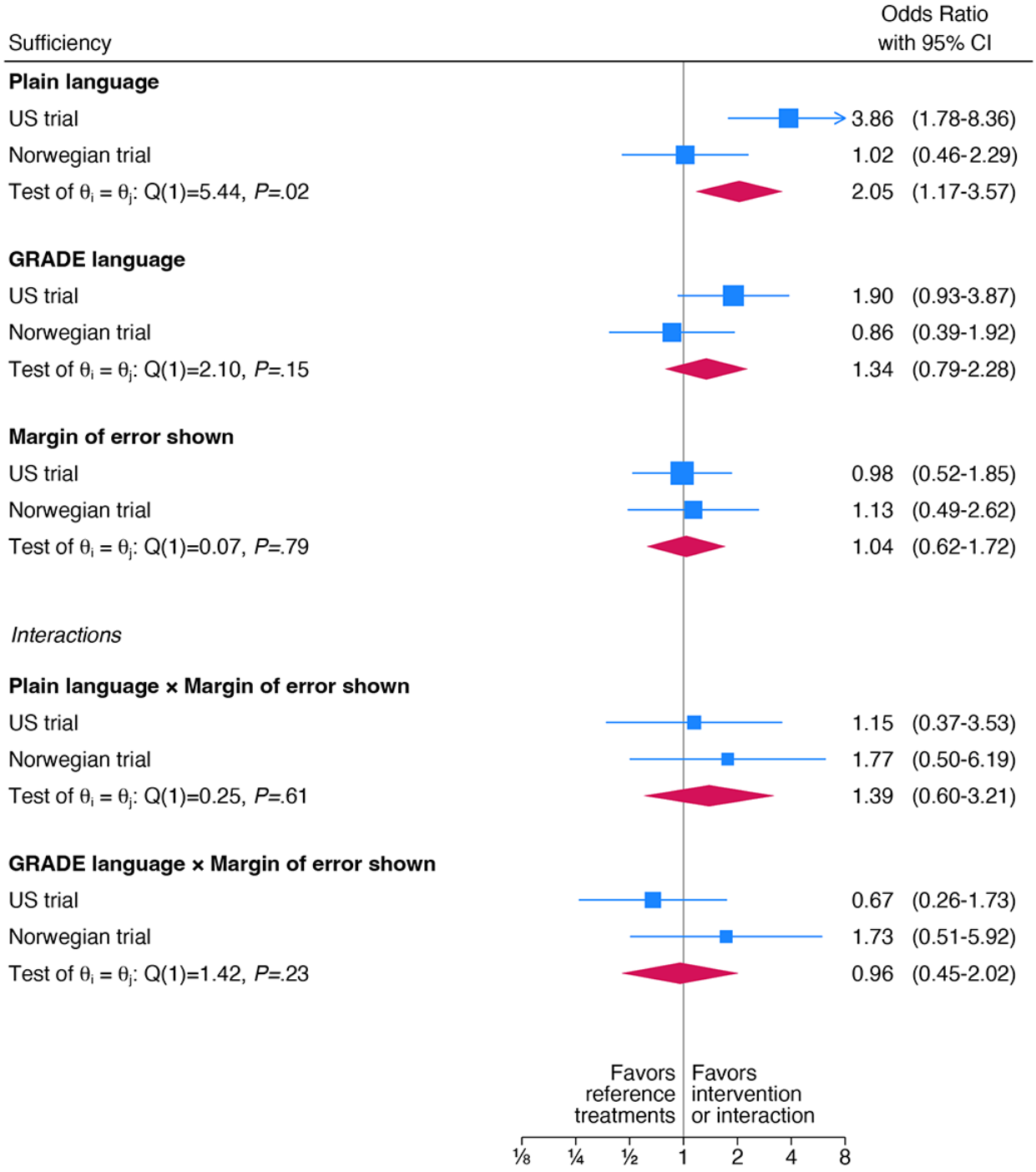
Understanding of the Sufficiency of the Evidence

Overall

Overall, plain, compared to no explicit language, improved correct understanding of the sufficiency of the evidence (OR 2.05, 95% CI 1.17-3.57; Figure 6). However, the results of the 2 trials were heterogeneous (Cochran Q_1=5.44; *P*=.02). The OR for the Norwegian trial was 1.02 (95% CI 0.46-2.29), whereas the OR for the US trial was 3.86 (95% CI 1.78-8.36). The effect of GRADE language, compared to no explicit language, was inconclusive on understanding of the sufficiency of the evidence (OR 1.34, 95% CI 0.79-2.28; Figure 6).

XSL•FO
RenderX

Figure 6. Meta-analysis presenting understanding the sufficiency of the evidence. Odds ratios for agreeing or strongly agreeing with "Not enough is known to be sure about the effects of wearing glasses to reduce the chance of getting COVID" in the Norwegian and US randomized controlled trials on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions. GRADE: Grading of Recommendations Assessment, Development and Evaluation.



Individual Trials

The results of the Norwegian trial for this outcome likely reflected a ceiling effect. In total, 81% (65/80) of the participants, who were shown no explicit language for overall uncertainty and no MOE (Table 5), agreed that not enough was known to be sure about the effects of wearing glasses to reduce the chance of acquiring COVID-19.

On the basis of the results of the US trial, using plain language increased the proportion of participants who agreed or strongly agreed that not enough was known to be sure about the effects of wearing glasses to reduce the chance of getting COVID-19, from 67% (58/86) to 89% (88/99); risk difference 21.4%, 95% CI 9.8-33.1 (Table 6). The effect of using the GRADE language was less certain.

Participants randomized to the intervention answering correctly or as anticipated.

XSL•FO
RenderX

Table 5. Risk differences for sufficiency of the evidence in the Norwegian randomized controlled trial on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions for the question "not enough is known to be sure about the effects of wearing glasses to reduce the chance of getting COVID," with responses including strongly agree, agree (correct), disagree, and strongly disagree.

Overall uncertainty	Margin of error	Participants[a], n (%)	Odds ratio (95% CI)[b]	Risk difference[c] (%; 95% CI)	*P* value
No explicit language (n=80)	Not shown	65 (81.2)	1	0	—[d]
No explicit language (n=71)	Shown	59 (83.1)	1.13 (0.49 to 2.62)	1.9 (–10.36 to 14.06)	.77
Plain language (n=76)	Not shown	62 (81.6)	1.02 (0.46 to 2.29)	0.3 (–11.88 to 12.54)	.96
Plain language (n=79)	Shown	71 (89.9)	2.05 (0.81 to 5.15)	8.6 (–2.21 to 19.46)	.13
GRADE[e] language (n=71)	Not shown	56 (78.9)	0.86 (0.39 to 1.92)	–2.4 (–15.16 to 10.40)	.72
GRADE language (n=75)	Shown	66 (88)	1.69 (0.69 to 4.14)	6.8 (–4.53 to 18.03)	.25

[a]Participants randomized to the intervention answering correctly or as anticipated.

[b]Odds ratios include the main and interaction effects.

[c]Risk differences account for uncertainty on the baseline odds.

[d]Not applicable.

[e]GRADE: Grading of Recommendations Assessment, Development and Evaluation.

Table 6. Risk differences for sufficiency of the evidence in the US randomized controlled trial on evaluating people's understanding of overall and statistical uncertainty when presented with alternative expressions for the question "not enough is known to be sure about the effects of wearing glasses to reduce the chance of getting COVID," with responses including strongly agree, agree (correct), disagree, and strongly disagree.

Overall uncertainty	Margin of error	Participants[a], n (%)	Odds ratio (95% CI)[b]	Risk difference[c] (%; 95% CI)	*P* value
No explicit language (n=86)	Not shown	58 (67.4)	1	0	—[d]
No explicit language (n=88)	Shown	59 (67)	0.98 (0.52 to 1.85)	–0.4 (–14.34 to 13.55)	.96
Plain language (n=99)	Not shown	88 (88.9)	3.86 (1.78 to 8.36)	21.5 (9.77 to 33.13)	<.001
Plain language (n=90)	Shown	81 (90)	4.34 (1.91 to 9.90)	22.6 (10.87 to 34.24)	<.001
GRADE[e] language (n=79)	Not shown	63 (79.7)	1.90 (0.93 to 3.87)	12.3 (–0.98 to 25.59)	.08
GRADE language (n=101)	Shown	73 (72.3)	1.26 (0.67 to 2.36)	4.8 (–8.37 to 18.04)	.47

[b]Odds ratios include the main and interaction effects.

[c]Risk differences account for uncertainty on the baseline odds.

[d]Not applicable.

[e]GRADE: Grading of Recommendations Assessment, Development and Evaluation.

Understanding of the Uncertainty of Important Harms

Overall

There was no difference between plain and GRADE languages, compared to no explicit language, on the ability to correctly understand the uncertainty of important harms (OR 0.98, 95% CI 0.68-1.63 and OR 1.23, 95% CI 0.85-1.78; see Figure S1 in Multimedia Appendix 4).

Individual Trials

Understanding of the uncertainty of important harms varied across the trials. In total, 14.2% (64/452) of the participants in the Norwegian trial and 31.1% (169/543) in the US trial correctly understood what we aimed to communicate about the certainty of the evidence for important harms (Tables S5 and S6 in Multimedia Appendix 4). In both trials, there were only small differences between the comparison groups that could have occurred by chance alone.

Understanding Statistical Uncertainty (MOE)

Overall

In both countries, showing, compared to not showing, the MOE increased the proportion of people who chose the answer most consistent with the information provided (may reduce the chance of COVID a little but might increase it a little).

Individual Trials

In the Norwegian trial, 27.1% (61/225 and between 21%-34% in each of the 3 groups) of the participants answered correctly when the MOE was shown. Only 1.8% (4/227 and between 1%-3% in each of the 3 groups) answered correctly when MOE was not shown. In the US trial, 27.6% (77/279 and between 21%-36% in each of the 3 groups)) answered correctly when MOE was shown, while 1.9% (5/264 and between 0%-4% in each of the 3 groups) answered correctly when the MOE was not shown (Tables S7 and S8 in Multimedia Appendix 4).

The majority of the total number of participants shown the MOE (366/504, 72.6%) from both trials (and between 64% and 79% in each of the 3 groups) *failed* to correctly understand the MOE.

XSL•FO
RenderX

As expected, very few people correctly guessed the MOE when it was not shown. A correct answer when not shown the MOE would have been "Don't know." The proportion of participants who responded "Don't know" when the MOE was shown was 4.8% (24/504) compared to 4.7% (23/491) when MOE was not shown.

Secondary Outcomes

Interest in Wearing Glasses to Reduce COVID-19 Risk During a Surge in Cases

Interest in wearing glasses to reduce COVID-19 risk during a surge was consistently lower in the Norwegian versus US trials (Figure S2 in Multimedia Appendix 4). In the Norwegian trial, plain language with or without the MOE and the GRADE language without the MOE reduced the proportion of participants who responded that they would be likely or very likely to wear or recommend wearing glasses to reduce the chance of acquiring COVID-19 if there were a surge of COVID-19 cases (risk difference: –13.6%, 95% CI –24.7 to –2.5; –18.6%, 95% CI –28.7 to –8.4; and −15.5%, 95% CI –26.4 to –4.5, respectively; Table S9 in Multimedia Appendix 4).

In the US trial, the combination of using either plain or GRADE languages to communicate the overall uncertainty of the benefit of wearing glasses and showing the MOE reduced the proportion of participants who responded that they would be likely or very likely to wear or recommend wearing glasses to reduce the chance of acquiring COVID-19 if there were a surge of COVID-19 cases (risk difference: –20.5%, 95% CI –33.1 to –7.9 and 21.2%, 95% CI –33.5 to –8.9, respectively; Table S10 in Multimedia Appendix 4).

The difference between the 2 trials is in part due to the lower proportion of participants in the Norwegian reference group (18/80, 23%) compared to the US reference group (31/86, 36%), who responded that they would be likely or very likely to wear or recommend wearing glasses to reduce the chance of acquiring COVID-19 if there were a surge of COVID-19 cases. This might be due to the translation from English to Norwegian.

Interest in Wearing Glasses to Reduce COVID-19 Risk if No Surge in Cases

We did not find an effect of plain or GRADE languages, compared to no explicit language, on the likelihood of wearing glasses or recommending wearing glasses to reduce the chance of acquiring COVID-19 if there were very few COVID-19 cases, the perceived benefit of wearing glasses, or the perceived chance of important harms (Figure S3 in Multimedia Appendix 4).

Perceptions of the Information Provided

In both trials, perceptions of the information provided differed little across plain and GRADE languages compared to no explicit language, with few exceptions; in both trials (US trial more so than Norway), perceptions of *helpfulness* were somewhat lower with plain language, with or without the MOE (Tables S11 and S12 in Multimedia Appendix 4).

In the Norwegian trial, with plain or GRADE languages, reporting the MOE reduced perceptions of the *trustworthiness* of the information (Table S13 and S14 in Multimedia Appendix 4).

In the Norwegian trial, the GRADE language without the MOE reduced the perception that information about whether wearing glasses affects the chance of acquiring COVID-19 was *sufficient* (Table S15 and S16 in Multimedia Appendix 4).

In the Norwegian trial, plain language without the MOE and the GRADE language with it reduced perceptions that the information about the benefit of glasses was *clear.* A similar effect was seen in the US trial for plain language with the MOE (Tables S17 and S18 in Multimedia Appendix 4).

In the Norwegian trial, plain language with the MOE reduced perceptions that information about whether glasses had important harms was *clear*. Similarly, this effect was seen in the US trial when the MOE was shown with no explicit language or plain language (Tables S19 and S20 in Multimedia Appendix 4).

In the Norwegian trial, plain language with the MOE reduced the proportion of participants who responded that they definitely or probably would *share the information* with someone who heard that wearing glasses might affect the chance of acquiring COVID-19 (Table S21 and S22 in Multimedia Appendix 4).

Decisional Conflict

Plain language without the MOE increased the feeling that the decision about wearing glasses if there were a surge of COVID-19 cases was hard to make (Table S23 and S24 in Multimedia Appendix 4). Most participants felt that they made an informed decision about wearing glasses if there were a surge of COVID-19 (264/452, 58.4%; between 52.6% and 67.6% in the Norwegian trial and 406/543, 74.8%; between 70.9% and 83% in the US trial; Table S25 and S26 in Multimedia Appendix 4).

Potential Modifying Factors

We did not find credible evidence of effect modification for numeracy, education, salience, or the time taken to complete the questionnaire.

Discussion

Principal Findings

The US and Norwegian trials comparing different ways of communicating the overall and statistical uncertainty of research results to the public generated mixed results. Plain language improved readers' understanding of the overall uncertainty of the benefit but only to a modest extent. The effect of the GRADE language was uncertain but, at best, had a modest effect. Furthermore, reporting the MOE reduced understanding by making the evidence seem more uncertain than it actually was. Reporting the MOE did improve understanding of statistical uncertainty around the effect of glasses but only for a minority group of people. A more detailed discussion of the most important findings follows in the subsequent sections.

XSL•FO
RenderX

Communicating Overall Uncertainty of the Benefit

Plain language is probably more accessible than the GRADE language (eg, the phrase "not very confident" seems easier to understand than "moderate uncertainty") for communicating the overall uncertainty of benefit. Given the modest findings, more work is needed on the language; other ways of communicating uncertainty, such as visualizations; and easily accessible explanations.

It should be noted that we displayed the GRADE symbol together with the GRADE language. The GRADE symbols for uncertainty are similar to widely used symbols used for ranking the quality of, for example, hotels, restaurants, and consumer products [30]. Nonetheless, they may be unfamiliar in this context and not easily understood without further explanation.

In another small trial [31], the GRADE symbols were compared to letters to convey the quality of evidence, and both letters and symbols were well understood.

Both the plain language and the GRADE language summaries included text explaining why we were "not very confident" (it was "uncertain" that wearing glasses may slightly reduce the chance of acquiring COVID-19). These explanations were available under a tab labeled "Keep in mind." Neither the number of participants who read this text nor the effect of the text on their understanding of the uncertainty of the evidence is known. More Norwegians responded that the evidence was very unsure compared to US participants. One reason for the differences between the Norwegian and US results may be that "may reduce" was translated to "kan muligens redusere" ("can possibly reduce") in Norwegian (based on the approved translation used by the Norwegian Institute of Public Health) [32]. The word "kan" in Norwegian can mean either "may" or "can," and these 2 terms may differ with regards to their emphasis on the degree of uncertainty, which is why the moderating adverb "muligens" (directly translating to "possibly") was included in the Norwegian version. Although we user-tested the summaries before the trials, further exploration of the extent to which people find this text helpful and how it impacts their understanding of the uncertainty of the evidence is warranted. Furthermore, cultural understandings and comfort with terms related to or describing the concept of "uncertainty" could be further explored to communicate uncertainty more accurately in languages other than English.

Communicating Overall Uncertainty of Important Harm

Neither plain language nor the GRADE language improved the understanding of the overall uncertainty of important harm. This could be because we did not explain why we said that there was moderate certainty evidence and prima facie, it seemed implausible to participants that wearing glasses could cause important harm. In fact, about one-third of the Norwegian participants were very sure, and this might be because they considered it implausible that wearing glasses could cause important harm. On the other hand, the other participants (about half) might have been somewhat or very unsure because there was so little information. The difference between the Norwegian and US trials might have been due to translation. The English version for the control group (No explicit language) stated that "wearing glasses probably does not cause important harms, such as serious fall due to reduced vision" while the Norwegian version, if back translated stated "Using glasses probably doesn't cause serious injury, for example after falling" but does not refer to "reduced vision." Roughly twice as many participants in the US trial, compared to participants in the Norwegian trial, correctly understood the overall uncertainty of the evidence for important harm. The most likely reason for this is the difference between the English summaries and the Norwegian translations. In this case, the explanation for why important harms (serious injury after falling is plausible because of reduced vision) was not mentioned in the Norwegian translation. This highlights the need for more extensive user testing and the use of back translation or other means of ensuring that translations are correctly understood.

Effect of MOE on Overall Uncertainty

Showing the MOE did help people understand statistical uncertainty (ie, glasses may reduce the chance of acquiring COVID-19 a little *but might* increase it a little). While this is encouraging, the modest effect suggests more work is needed to ensure that people understand what MOE implies.

In the group of participants who were shown plain language, showing the MOE (for benefit) decreased the proportion of people who answered this question correctly. This probably happened because the effect of glasses was small, and the MOE really highlighted that adding uncertainty to uncertainty in the context of a small effect, to begin with, made people feel the effect was less certain than it was. This finding underscores the need for more work to help people calibrate their sense of uncertainty. It would be interesting to see what would happen in another example where the intervention effect was bigger than the effect of glasses. Another reason for this may be that participants who were shown the MOE were also shown 2 reasons for rating the evidence as low certainty (wide MOE *and* important study limitations), whereas those who were not shown the MOE were only presented with 1 source of uncertainty (important study limitations). Findings from a 2020 study [33] looking at the effect of communicating uncertainty found that participants who were shown 3 sources of uncertainty were more likely to report a weaker perception of the effectiveness of the intervention (drug) than those who were presented with only 2 sources of uncertainty. More research needs to be conducted to explore this hypothesis.

Effect of MOE on Intended Behavior

Reporting overall uncertainty using plain language or the GRADE language or reporting the MOE decreased the likelihood that participants would wear glasses to protect themselves against COVID-19 if there was a surge in cases. This finding is consistent with other research findings that suggest reporting uncertainty decreased the likelihood that people would use eye protection to reduce their chance of acquiring COVID-19 [18]. It is also logical that the less certain one is about the benefits of doing something, the less likely it is to be done. These results highlight the importance of effectively communicating uncertainty if the intention is to inform people rather than to persuade them [19].

XSL•FO
RenderX

Effect of MOE on Perceptions of Information

In Norwegian trial, showing the MOE in combination with plain language reduced the perceived trustworthiness of the information (the results were similar, although smaller, and not statistically significant for the other versions). This is consistent with the findings of Schneider et al [18], which found that including a clue that evidence of the effect of eye protection was low quality or certainty reduced the perceived trustworthiness of an infographic. In contrast, in our US trial, the reduction in perceived trustworthiness when showing the MOE was substantially smaller and not statistically significant in any version. This suggests that we cannot conclude that communicating uncertainty necessarily reduces the perceived trustworthiness of information about the effects of interventions. Indeed, more candid communication over time might make changes in recommendations seem less arbitrary and help preserve people's trust in health authorities [14].

There is some evidence that plain language with or without the MOE and the GRADE language with the MOE reduced the perception that the summary was sufficient. This may be due to the participants confusing the sufficiency of the evidence (ie, a small, uncertain effect) with the sufficiency of the summary, which was meant to communicate the effect. While the question we asked specifically aimed at the latter (ie, whether the information was a sufficient summary of what is known about the effects of wearing glasses to reduce the chance of acquiring COVID-19), future qualitative work should be done to help distinguish these 2 kinds of sufficiency. Similarly, showing the MOE also reduced the perception that the summaries were clear, perhaps for the same reasons. In the Norwegian trial, the likely ceiling effect may reflect the mentioned translation of "may" to "can possibly" in the Norwegian summary.

Nevertheless, most participants found the decision about wearing glasses to reduce the chance of acquiring COVID-19 hard to make, with or without the MOE, regardless of which summary they were shown.

Limitations

Study limitations include the weak, uncertain effect that we were trying to summarize (from the Glasses trial), which magnified the communication challenges, and several language translation issues, which may limit the cross-country comparisons. Furthermore, our results may have been influenced by a lack of saliency, that is, COVID-19 infection rates were relatively low when we conducted the study, and participants might have responded differently in a more realistic or pressing scenario. Furthermore, we need to explore the effect of paying participants in web-based trials on the quality of the responses [24].

As this study was the research team's first attempt at conducting a web-based trial, we encountered and reflected on a number of challenges and opportunities. First, while we were satisfied with the web-based platforms we eventually chose, it was time-consuming and challenging to identify a proper platform to facilitate the conduct of the trial, and we eventually needed to incorporate 2 platforms to meet our specific needs. Another challenge was related to obtaining informed consent—it was unclear what elements of the consent were covered by the web-based platform or our invitation text. Since the conduct of our trial, more guidance has been provided on the platforms and other academic websites offering guidance on informed consent in web-based trials.

In the next Message Lab trial for assessing different ways of communicating evidence, we would like to set up Google Analytics (Google LLC), if using Google Sites to present the health messages to the participants. This will allow us to investigate further on how much time people spend on the different sections and how they navigate the texts presented.

Conclusions

Our study has several strengths, including the randomized factorial design, which let us explore the interaction between uncertainty language (GRADE, plain, or no explicit) and the MOE on understanding the benefits, harms, and the corresponding uncertainties and allowed replication in 2 distinct populations. Our study shows that explicitly reporting uncertainty affects peoples' understanding, perceptions, and intended actions. We found that plain language was better than no explicit language in helping people understand the overall uncertainty of the evidence, although most participants still did not correctly understand how sure they could be. Reporting MOE reduced understanding of the overall uncertainty by making people feel that the evidence was even less certain. Reporting MOE improved the interpretation of statistical uncertainty around the effect of glasses, but only for a minority group of participants.

This study underscores how much more work needs to be done to develop effective ways to communicate overall uncertainty and just how unclear the numbers are (statistical uncertainty). If this communication is done poorly, it may simply add to confusion and lead to poor decisions. If done properly, effective communication around uncertainty can help people to make the best decisions they can, given the evidence that is known.

Acknowledgments

This research is funded by the Norwegian Institute of Public Health. The authors would like to thank Ingeborg Hess Elgersma and the Centre for Epidemic Interventions Research team for their valuable feedback on the translations and questionnaire development.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
All summaries in both English and Norwegian.
[PDF File (Adobe PDF File), 4235 KB - publichealth_v11i1e62828_app1.pdf]

Multimedia Appendix 2
Questionnaires in English and Norwegian.
[PDF File (Adobe PDF File), 432 KB - publichealth_v11i1e62828_app2.pdf]

Multimedia Appendix 3
Information provided to the participants.
[PDF File (Adobe PDF File), 561 KB - publichealth_v11i1e62828_app3.pdf]

Multimedia Appendix 4
Additional result tables and figures.
[PDF File (Adobe PDF File), 893 KB - publichealth_v11i1e62828_app4.pdf]

Multimedia Appendix 5
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1253 KB - publichealth_v11i1e62828_app5.pdf]

References

1. Fischhoff B. The COVID communication breakdown: how to fix public health messaging. Foreign Affairs. 2021 Oct 4. URL: https://www.foreignaffairs.com/articles/united-states/2021-10-04/covid-communication-breakdown [accessed 2025-03-10]
2. Fischhoff B. The sciences of science communication. Proc Natl Acad Sci U S A 2013 Aug 20;110 Suppl 3(Suppl 3):14033-14039 [FREE Full text] [doi: 10.1073/pnas.1213273110] [Medline: 23942125]
3. Fischhoff B. Risk perception and communication. In: Detels R, Karim QA, Baum F, Li L, Leyland AH, editors. Oxford Textbook of Global Public Health, Seventh Edition. Oxford, UK: Oxford University Press; 2021.
4. Oxman AD, Glenton C, Flottorp S, Lewin S, Rosenbaum S, Fretheim A. Development of a checklist for people communicating evidence-based information about the effects of healthcare interventions: a mixed methods study. BMJ Open 2020 Jul 21;10(7):e036348 [FREE Full text] [doi: 10.1136/bmjopen-2019-036348] [Medline: 32699132]
5. Woloshin S, Yang Y, Fischhoff B. Communicating health information with visual displays. Nat Med 2023 May 08;29(5):1085-1091. [doi: 10.1038/s41591-023-02328-1] [Medline: 37156935]
6. Woloshin S, Dewitt B, Krishnamurti T, Fischhoff B. Assessing how consumers interpret and act on results from at-home COVID-19 self-test kits: a randomized clinical trial. JAMA Intern Med 2022 Mar 01;182(3):332-341 [FREE Full text] [doi: 10.1001/jamainternmed.2021.8075] [Medline: 35099501]
7. Woloshin S, Schwartz LM. Communicating data about the benefits and harms of treatment: a randomized trial. Ann Intern Med 2011 Jul 19;155(2):87-96 [FREE Full text] [doi: 10.7326/0003-4819-155-2-201107190-00004] [Medline: 21768582]
8. Akl EA, Oxman AD, Herrin J, Vist GE, Terrenato I, Sperati F, et al. Using alternative statistical formats for presenting risks and risk reductions. Cochrane Database Syst Rev 2011 Mar 16;2011(3):CD006776 [FREE Full text] [doi: 10.1002/14651858.CD006776.pub2] [Medline: 21412897]
9. Carling CL, Kristoffersen DT, Montori VM, Herrin J, Schünemann HJ, Treweek S, et al. The effect of alternative summary statistics for communicating risk reduction on decisions about taking statins: a randomized trial. PLoS Med 2009 Aug 25;6(8):e1000134 [FREE Full text] [doi: 10.1371/journal.pmed.1000134] [Medline: 19707575]
10. Schwartz LM, Woloshin S. The drug facts box: improving the communication of prescription drug information. Proc Natl Acad Sci U S A 2013 Aug 20;110 Suppl 3(Suppl 3):14069-14074 [FREE Full text] [doi: 10.1073/pnas.1214646110] [Medline: 23942130]
11. Blalock SJ, Sage A, Bitonti M, Patel P, Dickinson R, Knapp P. Communicating information concerning potential medication harms and benefits: what gist do numbers convey? Patient Educ Couns 2016 Dec;99(12):1964-1970 [FREE Full text] [doi: 10.1016/j.pec.2016.07.022] [Medline: 27444232]
12. Schwartz LM, Woloshin S. Communicating uncertainties about prescription drugs to the public: a national randomized trial. Arch Intern Med 2011 Sep 12;171(16):1463-1468. [doi: 10.1001/archinternmed.2011.396] [Medline: 21911629]
13. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011 Apr;64(4):401-406. [doi: 10.1016/j.jclinepi.2010.07.015] [Medline: 21208779]

XSL•FO
RenderX

14. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol 2011 Dec;64(12):1283-1293. [doi: 10.1016/j.jclinepi.2011.01.012] [Medline: 21839614]
15. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is "quality of evidence" and why is it important to clinicians? BMJ 2008 May 01;336(7651):995-998. [doi: 10.1136/bmj.39490.551019.be]
16. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines: 4. Rating the quality of evidence--study limitations (risk of bias). J Clin Epidemiol 2011 Apr;64(4):407-415. [doi: 10.1016/j.jclinepi.2010.07.017] [Medline: 21247734]
17. Santesso N, Rader T, Nilsen ES, Glenton C, Rosenbaum S, Ciapponi A, et al. A summary to communicate evidence from systematic reviews to the public improved understanding and accessibility of information: a randomized controlled trial. J Clin Epidemiol 2015 Feb;68(2):182-190 [FREE Full text] [doi: 10.1016/j.jclinepi.2014.04.009] [Medline: 25034199]
18. Schneider CR, Freeman AL, Spiegelhalter D, van der Linden S. The effects of quality of evidence communication on perception of public health information about COVID-19: two randomised controlled trials. PLoS One 2021 Nov 17;16(11):e0259048 [FREE Full text] [doi: 10.1371/journal.pone.0259048] [Medline: 34788299]
19. Oxman AD, Fretheim A, Lewin S, Flottorp S, Glenton C, Helleve A, et al. Health communication in and out of public health emergencies: to persuade or to inform? Health Res Policy Syst 2022 Mar 05;20(1):28 [FREE Full text] [doi: 10.1186/s12961-022-00828-z] [Medline: 35248064]
20. Woloshin S, Oxman AD, Rose CJ, Rosenbaum SE, Holst C, Diaz-Hemard L, et al. Communicating uncertainty about the effects of wearing glasses to reduce the chance of getting COVID: protocol for a randomized trial. Zenodo. 2022 Nov 18. URL: https://zenodo.org/records/7428981 [accessed 2025-03-10]
21. Fretheim A, Elgersma IH, Helleve A, Elstrøm P, Kacelnik O, Hemkens LG. Effect of wearing glasses on risk of infection with SARS-CoV-2 in the community: a randomized clinical trial. JAMA Netw Open 2022 Dec 01;5(12):e2244495 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.44495] [Medline: 36454571]
22. Prolific homepage. Prolific. URL: https://www.prolific.com/ [accessed 2024-03-15]
23. Vi måler nordmenns meninger, og leverer innsikt som bringer deg videre. Opinion. URL: https://www.opinion.no/ [accessed 2024-03-15]
24. Bradley P. Bots and data quality on crowdsourcing platforms. Prolific. 2018 Aug 10. URL: https://www.prolific.com/resources/bots-and-data-quality-on-crowdsourcing-platforms [accessed 2024-02-20]
25. Douglas BD, Ewell PJ, Brauer M. Data quality in online human-subjects research: comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. PLoS One 2023 Mar 14;18(3):e0279720 [FREE Full text] [doi: 10.1371/journal.pone.0279720] [Medline: 36917576]
26. Ergonomics of human-system interaction part 210: human-centred design for interactive systems. International Organization for Standardization. 2019. URL: https://www.iso.org/standard/77520.html [accessed 2025-03-10]
27. Loop11 homepage. Loop11. URL: https://www.loop11.com/ [accessed 2024-03-15]
28. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 1995 Feb 01;15(1):25-30. [doi: 10.1177/0272989x9501500105]
29. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med 1997 Dec 01;127(11):966-972. [doi: 10.7326/0003-4819-127-11-199712010-00003] [Medline: 9412301]
30. Schünemann HJ, Best D, Vist G, Oxman AD, GRADE Working Group. Letters, numbers, symbols and words: how to communicate grades of evidence and recommendations. CMAJ 2003 Sep 30;169(7):677-680 [FREE Full text] [Medline: 14517128]
31. Akl EA, Maroun N, Guyatt G, Oxman AD, Alonso-Coello P, Vist GE, et al. Symbols were superior to numbers for presenting strength of recommendations to health care consumers: a randomized trial. J Clin Epidemiol 2007 Dec;60(12):1298-1305. [doi: 10.1016/j.jclinepi.2007.03.011] [Medline: 17998085]
32. Formuleringer i GRADE for å uttrykke resultater fra kunnskapsoppsummeringer om effekt. Folkehelseinstituttet (Norwegian Institute of Public Health). URL: https://www.fhi.no/ku/kunnskaps-og-beslutningsstotte/metodeboka/framgangsmate/hente-ut-data-sammenfatte-og-gradere/?term= [accessed 2024-03-15]
33. Büchter RB, Betsch C, Ehrlich M, Fechtelpeter D, Grouven U, Keller S, et al. Communicating uncertainty in written consumer health information to the public: parallel-group, web-based randomized controlled trial. J Med Internet Res 2020 Aug 10;22(8):e15899 [FREE Full text] [doi: 10.2196/15899] [Medline: 32773375]

Abbreviations

GRADE: Grading of Recommendations Assessment, Development and Evaluation
MOE: margin of error
OR: odds ratio

XSL•FO
RenderX

Edited by A Mavragani; submitted 04.06.24; peer-reviewed by R Büchter, H Pilabré; comments to author 11.12.24; revised version received 30.01.25; accepted 06.02.25; published 18.03.25.

Please cite as:
Holst C, Woloshin S, Oxman AD, Rose C, Rosenbaum S, Munthe-Kaas HM
Alternative Presentations of Overall and Statistical Uncertainty for Adults' Understanding of the Results of a Randomized Trial of a Public Health Intervention: Parallel Web-Based Randomized Trials
JMIR Public Health Surveill 2025;11:e62828
URL: https://publichealth.jmir.org/2025/1/e62828
doi:10.2196/62828
PMID:

©Christine Holst, Steven Woloshin, Andrew D Oxman, Christopher Rose, Sarah Rosenbaum, Heather Menzies Munthe-Kaas. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of a Digital Media Campaign to Promote Knowledge and Awareness of the GPFirst Program for Nonurgent Conditions: Repeated Survey Study

Rebecca Ong Hui Shan[1], MPH; Hong Choon Oh[1,2], PhD; Priscilla Goh Sook Kheng[3], AdvDipLaw; Lyndia Lee Sze Hui[3], BSIT; Mas Riza Bte Mohd Razali[3], BCom; Edris Atikah Ahmad[3], MScAppGer; Jagadesan Raghuram[4], MB BCh BAO; Choon How How[4,5], MBBS, MMed; Steven Lim Hoon Chin[6], MBBS, MRCSEd

[1]Health Services Research, Changi General Hospital, SingHealth, Singapore, Singapore

[2]Centre for Population Health Research and Implementation, SingHealth, Singapore, Singapore

[3]Primary Care Integration, Changi General Hospital, SingHealth, Singapore, Singapore

[4]Chairman Medical Board's Office, Changi General Hospital, SingHealth, Singapore, Singapore

[5]Family Medicine, Academic Clinical Programme, Duke-NUS, Singapore, Singapore

[6]Department of Emergency Medicine, Changi General Hospital, SingHealth, Singapore, Singapore

Corresponding Author:
Rebecca Ong Hui Shan, MPH
Health Services Research
Changi General Hospital
SingHealth
2 Simei Street 3
Singapore, 529889
Singapore
Phone: 65 64267479
Fax: 65 67826049
Email: rebecca.ong.h.s@singhealth.com.sg

Abstract

Background: GPFirst is a primary care partnership program designed to encourage patients with nonurgent conditions to seek care at participating general practitioner clinics instead of visiting the emergency department. In 2019, a digital media campaign (DMC) was launched to raise awareness and knowledge about GPFirst among residents in eastern Singapore.

Objective: This study aims to assess the DMC's impact on awareness and knowledge of GPFirst across different age groups, and the acceptability and satisfaction of GPFirst.

Methods: The DMC, comprising Facebook posts and a website designed using the Andersen behavioral model, was evaluated through 2 repeated cross-sectional surveys. The first cross-sectional survey (CS1) was conducted with eastern Singapore residents aged 21 years and older, 2 1 year before the campaign's launch, and the second survey (CS2) 4 months after. Satisfaction was measured on a 5-point Likert scale (very poor to excellent) about GPFirst experiences. Acceptability was assessed with 3 yes or no questions on decisions to visit or recommend GPFirst clinics. Analyses used tests of proportions, adjusted multiregression models, and age-stratified secondary analyses.

Results: The Facebook posts generated 38,404 engagements within 5 months, with "#ThankYourGP" posts being the most viewed (n=24,602) and engaged (n=2618). Overall, 1191 and 1161 participants completed CS1 and CS2 respectively. Compared to CS1, CS2 participants were more aware (odds ratio [OR] 2.64, 95% CI 2.11-3.31; *P*<.001) and knowledgeable of GPFirst (OR 4.20, 95% CI 2.62-6.73; *P*<.001). Awareness was higher among married individuals (OR 1.31, 95% CI 1.04-1.66; *P*=.03), those without a regular primary care physician (OR 1.79, 95% CI 1.44-2.22; *P*<.001), and with higher education levels. Similarly, knowledge was greater among individuals with secondary (OR 2.88, 95% CI 1.35-6.17; *P*=.006) and preuniversity education (OR 2.56, 95% CI 1.14-5.70; *P*=.02), and those without a regular primary care physician (OR 1.54, 95% CI 1.02-2.34; *P*=.04). For acceptability, among participants who visited a GPFirst clinic, 98.2% (163/166) reported they would continue to visit a GPFirst clinic before the emergency department in the future, 95.2% (158/166) would recommend the clinic, 60.2% (100/166) cited the clinic's participation in GPFirst as a factor in their provider's choice and 87.3% (145/166) were satisfied with GPFirst.

XSL•FO
RenderX

Among those unaware of GPFirst, 88.3% (1680/1903) would consider visiting a GPFirst clinic before the emergency department in the future.

Conclusions: The DMC improved awareness and knowledge of GPFirst, with high satisfaction and acceptability among participants. Age-dependent strategies may improve GPFirst participation. The "#ThankYourGP" campaign demonstrated the potential of user-generated content to boost social media engagement, a strategy that international health systems could adopt.

(JMIR Public Health Surveill 2025;11:e66062) doi:10.2196/66062

KEYWORDS

digital media campaign; public awareness campaign; primary care partnership; social media; nonurgent emergency department visits; Andersen model

Introduction

Overcrowding in emergency departments (EDs) is a global concern, exacerbated by nonurgent patient visits [1-3], which could be managed in primary care settings. In the United States, studies consistently report nonurgent ED visit rates exceeding 30% [3], leading to adverse outcomes such as medication errors and increased mortality [4,5]. Redirecting these patients to primary care providers could potentially mitigate health care costs [6], improve patient safety and optimize ED resource allocation for urgent cases. Singapore faces a parallel challenge, with nonurgent visits, defined as mild-to-moderate conditions that could be managed in primary care, contributing to over 50% of cases in public hospitals as of 2013 [7]. In response, the GPFirst primary care partnership program was launched in 2014 at Changi General Hospital's (CGH) ED department [8]. GPFirst is an educational program that integrates financial incentives to encourage patients with nonurgent conditions to seek initial care at participating GPFirst clinics. The program educates the public on recognizing these nonurgent conditions. Additionally, GPFirst fosters collaboration with general practitioners to enhance the management of nonurgent conditions in primary care. To reinforce these educational efforts, the program offers a financial incentive in the form of an SG $50 (US $36.50) discount on ED fees for patients who first seek care at participating GPFirst clinics and subsequently require a referral to CGH's ED. Research indicated that this amount adequately covers most primary care expenses [9].

Recognizing the need to enhance awareness and understanding of GPFirst among the 1 million residents within its catchment area [10], CGH initiated a digital media campaign (DMC) in 2019. The campaign aimed at promoting visits to participating GPFirst clinics over ED visits for nonurgent health issues. This was done through the dissemination of educational materials via social media and a website, which provided information on nonurgent conditions manageable by general practitioners, and the location of participating GPFirst clinics. The GPFirst team selected digital media (DM) for the campaign due to its proven ability to reach diverse demographic groups and promote health-related behavioral changes through integrated health promotion content, as highlighted by existing literature [11-19].

In Singapore, individuals aged 39 years and younger account for over 50% of nonurgent visits, with the 40-59 years age group contributing 25%, and those aged 60 years and older constituting 15.2% [20]. A 2019 national survey highlighted that 88% of adults aged 25 years and older use the internet for the prevalent activity of social networking [21]. Particularly, Facebook is popular among individuals aged 25-44 years, with approximately 57% actively using the platform in 2019 [22]. This offered an opportunity to leverage digital platforms like Facebook to disseminate health educational material and catalyze health behavior changes [23,24].

While numerous studies have explored the use of DM in promoting healthier lifestyle practices [15,25], gaps remain in understanding its impact in promoting specific health program awareness and knowledge and how this impact may vary across diverse age groups. Addressing these gaps is also crucial for advancing health equity [26,27], as awareness and understanding of health programs can vary significantly among different demographic segments, which could lead to unequal access to its benefits. Younger individuals or those with higher education levels might be more adept at finding and understanding health care information, whereas older adults or those with lower health literacy may be less aware of the program's benefits [28].

This study primarily aims to assess the impact of a new DMC on the awareness and knowledge of GPFirst among residents who reside in eastern Singapore. The study also has a secondary aim of assessing the overall acceptability and satisfaction of GPFirst within this community since its introduction in 2014.

Methods

Overview

The roll-out of DMC entailed two key phases, namely (1) campaign development and (2) impact assessment. Details of these 2 phases are described in the following sections.

DMC Development

To capitalize on the popularity of Facebook and the internet in the general local population, DMC leveraged information-seeking models [29] that delineate 2 primary mechanisms for acquiring digital health information: purposive information-seeking and information encountering. While purposive seeking involves intentional efforts to obtain specific information, information encountering occurs incidentally during routine digital activities [29]. Research shows that individuals frequently encounter health-related information incidentally more than they actively seek it out [30,31]. The study team postulated that placement of GPFirst-related information on frequently visited DM platforms increases the likelihood that target residents would encounter messages promoting GPFirst during their routine digital activities. This approach could

XSL•FO
RenderX

potentially enable the DMC to reach a broad audience, including those who may not actively seek out health information but could benefit from the program's messages.

The development of the DMC entailed a collaborative effort among a multidisciplinary team from CGH, including ED and community care practitioners, researchers, data managers, corporate communications personnel, GPFirst program executives, and a communication agency (GOLIN/HARRIS International Pte Ltd). This agency was commissioned to create the educational materials featured in this study, with actors and models specifically hired for their production. All images, videos, and content are owned and protected under CGH's intellectual property rights. Furthermore, the Primary Care Advisory Council, comprising health care administrators and clinicians dedicated to advancing primary health care delivery models and initiatives in eastern Singapore, provided advisory support on the campaign's content and branding.

The objective of this stage was to generate content aimed at fostering community understanding and recognition of the legitimacy of ED overcrowding, while also delivering educational material essential for behavioral change. This approach is rooted in established behavior change models such as the Andersen behavior model [32-34]. According to the Andersen model, the use of health services is influenced by 3 key components: predisposing factors (personal characteristics and beliefs affecting service use likelihood), enabling factors (resources that facilitate or impede health care access), and need factors (individual perception of health needs driving care-seeking decisions) [32-34]. In the context of individuals who have nonurgent conditions, we postulated that they are less likely to visit the ED if (1) they perceive their general practitioners as capable of managing their health conditions and understand that EDs are reserved for urgent cases (predisposing factors), (2) they are educated on identifying nonurgent conditions and have access to participating GPFirst clinics (enabling factors), and (3) they are able to comprehend the perceived severity of their health condition (need factors). As such, the campaign content was developed so that it was embedded adequately with the aforementioned predisposing, enabling, and need factors. The development process involved regular weekly meetings to craft campaign messaging, educational content, and design elements. This collaborative endeavor led to the creation of 2 digital platforms aimed at promoting GPFirst, detailed in the subsequent sections.

GPFirst Facebook Page

The GPFirst Facebook page [35] adopted a targeted strategy to engage its social media audience by sharing 2-3 posts weekly from end-July 2019 to December 2020. The campaign included informative content (Multimedia Appendix 1), video reels, education on various nonurgent conditions, festival greetings, regular quizzes, and special giveaway events like the "8 Weeks of Gifting," where participants could win prizes by correctly answering questions posted on the GPFirst Facebook page. These activities are aimed at engaging and educating the audience, addressing the Andersen predisposing factors by shaping beliefs and attitudes toward using general practitioner services over EDs. An event titled "#ThankYourGP" encouraged individuals to share personal stories and show appreciation for their general practitioners (Multimedia Appendix 1), with a chance to win a nonmonetary prize. This initiative was intended to build trust and positive perceptions of general practitioners, further addressing predisposing factors.

GPFirst leveraged the Facebook platform to share educational videos about the appropriate use of ED services. Each video was crafted to highlight the nonurgent conditions manageable by general practitioners, emphasizing the Andersen enabling factors by informing the audience about available resources and appropriate health care access points. One notable concept was a 3-part series featuring live actors (Figure 1). This series followed the story of a youth named Zac, who, having a fever, opted to visit the ED accompanied by his sister (Cassy). Zac and Cassy observed individuals in the ED with conditions that did not warrant emergency care. Through Cassy's guidance, Zac learned that general practitioners are a viable alternative for such conditions and discovered the benefits of GPFirst.

Figure 1. Zac and Cassy 3-part video series. The individuals in the images are paid models or actors hired by Golin Pte Ltd. The intellectual property rights for all designs and deliverables (images/videos) produced by the vendor (Golin Pte Ltd) belong to Changi General Hospital, and are used here with permission.

The narrative of Zac's story informed users of the inappropriate use of the ED for minor ailments like mild fevers and colds, addressing the Andersen need factors by helping individuals understand the severity of their health conditions and the appropriate care settings. The series concluded with Zac and Cassy returning home to share their newfound knowledge about GPFirst with their parents and neighbors, thereby spreading awareness within their community and reinforcing the Andersen enabling factors by promoting access to GPFirst services.

XSL•FO
RenderX

GPFirst Website

GPFirst website [36] was designed to address Andersen needs and enabling factors. The website offers both an overview and a detailed explanation of GPFirst, emphasizing its purpose and benefits. The smartphone-friendly design ensured a seamless browsing experience on smartphones and tablets. The homepage introduces visitors to GPFirst Facebook videos and provides links to numerous other educational materials and resources. A dedicated section titled "Conditions your GP can treat" provides an extensive list of various medical conditions, indicating whether they necessitate a visit to the general practitioner, ED, or managed at home. This helps visitors determine the appropriate course of action based on the presented symptoms, aligning with the need factors of the Andersen model. Enabling factors are addressed through an "Info for GPs" page, which outlines the enrollment process for general practitioner clinics into GPFirst, ensuring that health care providers can participate and contribute to the initiative. Additionally, a dedicated web page helps visitors locate the nearest participating GPFirst clinic, facilitating easy access to services.

GPFirst Awareness

To assess the campaign's impact on community knowledge and awareness of GPFirst, we conducted 2 repeated cross-sectional surveys. The first cross-sectional survey (CS1) took place from July 2018 to August 2018, a year before the campaign started. The second survey (CS2) was conducted from November 2019 to December 2019, 4 months after the campaign began. Data collection was conducted by EA Research and Consulting Pte Ltd (EARC), a market research organization that was commissioned and compensated through the study grant. We measured awareness of GPFirst by asking participants in both CS1 and CS2 if they had heard of the program before participating in the surveys.

GPFirst Knowledge

Knowledge about GPFirst was assessed through 3 specific questions (Multimedia Appendix 2) and was limited to those who reported being aware of GPFirst. Participants who answered all 3 questions correctly were categorized as "knowledgeable," while those who did not were categorized as "not as knowledgeable."

Acceptability and Satisfaction of GPFirst

Acceptability was evaluated using responses from both CS1 and CS2, targeting 2 groups of participants: those who were aware of GPFirst and had visited a participating GPFirst clinic within the past 3 months (group A), and those who were unaware of GPFirst (group B). For group A, acceptability was assessed through three dichotomous yes or no questions: (1) whether they would revisit the participating GPFirst clinic, (2) whether they would recommend the clinic to others, and (3) whether their decision to visit the clinic was influenced by its participation in GPFirst (Multimedia Appendix 2). For group B, after receiving a brief overview of the program, acceptability was assessed by asking whether they would consider visiting participating GPFirst clinics before going to the ED in the future (Multimedia Appendix 1). Satisfaction with GPFirst was measured based on group A's ratings of the overall experience with the program, using a 5-point Likert scale ranging from "very poor" to "excellent" (Multimedia Appendix 2).

Recruitment

Quota sampling was used to recruit participants from three groups: (1) the ED group comprising of individuals who had visited CGH ED within the last 3 months, (2) the general practitioner group consisting of individuals who had visited general practitioners in eastern Singapore within the last 3 months, and (3) the public group composed of individuals who had not sought care at CGH ED or general practitioners in eastern Singapore within the last 3 months. If a participant had visited both general practitioner and CGH ED over the last 3 months, the most recent visit was used to determine the participant's classification. Recruitment was conducted at public locations in eastern Singapore (Bedok, Pasir Ris, Tampines, and Simei), including train and bus stations, and food centers. A market research organization EARC was commissioned to help recruit participants and administer the surveys. EARC surveyors received training from the study team before being deployed to the designated areas between 11 AM and 9 PM, on both weekdays and weekends, to approach the public. To be eligible, participants were required to (1) be Singaporean citizens or permanent residents, (2) be aged 21 years and older, (3) be proficient in English, (4) not be employed by CGH, (5) have a residential address within eastern Singapore, and (6) belong to one of the three study groups.

Sample Size Estimation

The sample size was determined using the Yamane formula with an error tolerance set at 0.05 [37], resulting in an estimated sample of 400 per study group, 1200 participants per survey, and a total of 2400 participants across both surveys.

Statistical Analysis

Data analysis was performed using SPSS (version 27; IBM Corp), with a significance level of .05. Standard descriptive statistics including means, SDs, and frequencies were reported as appropriate. Chi-square, and where appropriate, Fisher exact tests were used to compare changes in outcomes between the CS1 and CS2 surveys. Analyses were also performed to examine the impact of the campaign across different age groups: (1) 21-39 years, (2) 40-59 years, and (3) 60 years and older [20], and multilogistic regression models adjusted for gender and ethnicity were performed to examine factors associated with GPFirst awareness. To address the potential inflation of type I error rates due to multiple comparisons, the false discovery rate (FDR) method was applied. The FDR method was chosen as it offers a more balanced approach, controlling false positives without being as overly conservative as the Bonferroni correction. While Bonferroni adjustment reduces the likelihood of type I errors, it also increases the risk of type II errors by being overly stringent [38]. The FDR method, on the other hand, maintains greater statistical power by allowing for a controlled proportion of false positives, making it more suitable for exploratory studies and public health research involving multiple hypothesis testing. Marital status was recoded into 2 groups: "married" and "single," education level was recoded into 4 groups: "primary and below," "secondary," "preuniversity,"

XSL•FO
RenderX

and “university and higher,” and residential type was recorded into 3 groups: “1-3 room public flats,” “4-5 room public flats,” and “private and others.” Participants who were part of the military service were excluded from the analysis. This exclusion was justified by studies suggesting that out-of-pocket costs can influence ED attendance [39]. Given that existing financing policies waive public hospital ED fees for military service personnel, their health-seeking behaviors may differ from those of other participants. This study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist (Multimedia Appendix 3).

Ethical Considerations

This study was approved by the SingHealth Centralized Institutional Review Board (reference 2018/2196). Informed consent was obtained from all participants, and data were deidentified for analysis. Participants who completed the survey received an SGD $5 (US $3.73) supermarket voucher as a token of appreciation. Models present in images provided written informed consent to allow their image to be published.

Results

Overview

Within 5 months of the DMC rollout, the GPFirst Facebook page garnered 3162 likes, 3199 followers, and 38,404 engagements. An engagement is defined as a reaction, comment, share, or click on the GPFirst Facebook page [40]. The “#ThankYourGP” post achieved the highest number of views (n=24,602) and had the highest engagement (n=2618). Due to data limitations, the profiles of engaged individuals, or the platforms used (eg, smartphones, laptops) to access the Facebook page could not be analyzed in this study.

Data from 1191 participants for CS1 and 1161 participants for CS2 were analyzed (Table 1). Compared to CS1 participants, those in CS2 had higher proportions of women (54% vs 52%; *P*=.04), Chinese (65.5% vs 58.8%; *P*=.005), married individuals (66.3% vs 62%; *P*=.008), and those with higher education levels (*P*=.008). Additionally, CS2 participants had lower proportions of individuals with regular primary care physicians (*P*=.04). Participants’ characteristics stratified by age are detailed in Multimedia Appendix 4.

XSL•FO
RenderX

Table 1. Characteristics of participants from the precampaign (CS1[a]) and postcampaign (CS2[b]) surveys.

Variable	CS1 (n=1191)	CS2 (n=1161)
Age groups (years), n (%)		
21-39	414 (34.8)	389 (33.5)
40-59	467 (39.2)	471 (40.6)
60 and older	310 (26)	301 (25.9)
Gender, n (%)		
Women	572 (48)	653 (56.2)
Men	619 (52)	508 (43.8)
Ethnicity, n (%)		
Chinese	700 (58.8)	762 (65.6)
Indian	78 (7)	74 (7)
Malay	380 (31.9)	297 (25.6)
Others	33 (3)	28 (2)
Current marital status, n (%)		
Single	453 (38)	391 (33.7)
Married	738 (62)	770 (66.3)
Residential type, n (%)		
1-3 room public flats	316 (26.5)	283 (24.4)
4-5 room public flats	730 (61.3)	724 (62.4)
Private and others	145 (12.2)	154 (13.3)
Highest education level, n (%)		
Primary and lower	239 (20.1)	173 (14.9)
Secondary	445 (37.4)	440 (37.9)
Preuniversity	297 (24.9)	319 (27.5)
University and higher	210 (17.6)	229 (19.7)
Regular primary care physician, n (%)		
Yes	674 (56.4)	608 (52.4)
No	517 (43.6)	553 (47.6)

[a]CS1: cross-sectional survey 1 (baseline).

[b]CS2: cross-sectional survey 2.

GPFirst Awareness

The proportion of participants who were aware of GPFirst in CS1 and CS2 was 11.5% (137/1191 participants) and 26.9% (312/1161 participants), respectively (Multimedia Appendix 5). This more than 2-fold increase in GPFirst awareness was also observed across all age groups. After adjustment (Table 2), the odds of being aware of GPFirst in CS2 were 2.7 times higher than in CS1 (odds ratio [OR] 2.69, 95% CI 2.15-3.37; *P*<.001). In each of the 3 age groups (ie, 21-39 years, 40-59 years, and 60 years and older), the odds of being aware GPFirst in CS2 were also 2.3 to 3.2 times higher than in CS1 (Table 3). Overall, married individuals had 1.3 times higher odds (95% CI 1.03-1.65, *P*=.03) of being aware of GPFirst compared to individuals who were single (Table 3). Higher education levels were also associated with GPFirst awareness, particularly among those aged 40-59 years. Additionally, individuals without a regular primary care physician had 1.8 times higher odds (95% CI 1.44-2.22; *P*<.001) of being aware of GPFirst compared to those with a regular primary care physician.

XSL•FO
RenderX

Table 2. Adjusted multilogistic regression analyses of factors associated with GPFirst awareness among participants from the precampaign (CS1[a]) and postcampaign (CS2[b]) surveys (N=2352).

Variable	Overall	
	Adjusted OR[c] (95% CI)	*P* value
Time		
CS1	1.00	—[d]
CS2	2.64 (2.11-3.31)	<.001
Marital status		
Single	1.00	—
Married	1.31 (1.04-1.66)	.02
Gender		
Men	1.00	—
Women	1.16 (0.93-1.44)	.19
Ethnicity		0.27[e]
Chinese	1.00	—
Indian	0.89 (0.57-1.39)	.61
Malay	0.94 (0.74-1.21)	.65
Others	0.43 (0.18-1.01)	.05
Education level		< .001[e]
Primary and lower	1.00	—
Secondary	2.12 (1.47-3.05)	<.001
Preuniversity	2.17 (1.47-3.20)	<.001
University and higher	1.95 (1.28-2.97)	.002
Residential type		.21[e]
1 to 3 room public flats	1.00	—
4 to 5 room public flats	1.18 (0.90-1.56)	.23
Private and others	1.41 (0.96-2.08)	.08
Regular primary care physician		
Yes	1.00	—
No	1.79 (1.44-2.22)	<.001

[a]CS1: cross-sectional survey 1 (baseline).

[b]CS2: cross-sectional survey 2.

[c]OR: odds ratio.

[d]Not applicable.

[e]Omnibus test.

XSL•FO
RenderX

Table 3. Age-stratified adjusted multilogistic regression analysis of factors associated with GPFirst awareness among participants from the precampaign (CS1[a]) and postcampaign (CS2[b]) surveys (N=2352).

Variable	21-39 age group		40-59 age group		60 and older age group	
	Adjusted OR[c] (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Time						
CS1	1.00	—[d]	1.00	—	1.00	—
CS2	3.17 (2.16-4.67)	<.001	2.33 (1.67-3.26)	<.001	2.92 (1.75-4.88)	<.001
Marital status						
Single	1.00	—	1.00	—	1.00	—
Married	1.33 (0.90-1.95)	.15	1.16 (0.76-1.80)	.49	1.73 (0.90-3.30)	.10
Gender						
Men	1.00		1.00	—	1.00	—
Women	1.20 (0.83-1.73)	.35	1.14 (0.82-1.59)	.44	1.24 (0.76-2.01)	.39
Ethnicity		.36[e]		.21[e]		.34[e]
Chinese	1.00	—	1.00	—	1.00	—
Indian	0.54 (0.25-1.18)	.12	1.34 (0.69-2.60)	.39	0.86 (0.31-2.41)	.77
Malay	0.84 (0.55-1.28)	.41	1.18 (0.80-1.72)	.51	0.53 (0.27-1.05)	.07
Others	0.50 (0.11-2.29)	.37	0.17 (0.02-1.26)	.08	0.94 (0.26-3.46)	.93
Education level		.36[e]		.05[e]		.15[e]
Primary and lower	1.00	—	1.00	—	1.00	—
Secondary	3.95 (0.49-31.95)	.20	2.18 (1.25-3.79)	.006	1.56 (0.88-2.76)	.13
Preuniversity	3.75 (0.46-30.45)	.22	1.97 (1.07-3.63)	.03	2.03 (0.97-4.24)	.06
University and higher	2.87 (0.35-23.58)	.33	1.96 (1.00-3.82)	.05	2.71 (0.94-7.86)	.07
Residential type		.19[e]		.31[e]		.66[e]
1 to 3 room public flats	1.00	—	1.00	—	1.00	—
4 to 5 room public flats	1.03 (0.64-1.65)	.91	1.36 (0.87-2.12)	.18	1.21 (0.68-2.16)	.51
Private and others	1.71 (0.87-3.35)	.12	1.55 (0.86-2.81)	.15	0.92 (0.39-2.16)	.85
Regular primary care physician						
Yes	1.00	—	1.00	—	1.00	—
No	1.85 (1.27-2.70)	.001	1.65 (1.19-2.28)	.003	2.05 (1.28-3.30)	.003

[a]CS1: cross-sectional survey 1 (baseline).

[b]CS2: cross-sectional survey 2.

[c]OR: odds ratio.

[d]Not applicable.

[e]Omnibus test.

GPFirst Knowledge

Among the 449 participants who self-reported being aware of GPFirst, significant differences were observed in the knowledge levels of GPFirst between CS2 and CS1 participants (Multimedia Appendix 6). The proportion of knowledgeable individuals was significantly higher in CS2 compared to CS1, particularly in the 21-39 age group (CS1: 19.6%; CS2: 66.7%; *P*<.001) and the 40-59 age group (CS1: 25.4%; CS2: 57.8%; *P*<.001). No significant difference in proportions was observed for the 60 years and older age group. A significantly higher proportion of CS2 participants correctly answered questions B (CS1: 38.0% vs CS2: 69.6%; *P*<.001) and C (CS1: 70.1% vs CS2: 87.2%; *P*<.001). No significant difference was observed for question A. After adjustment (Table 4), the odds of being knowledgeable about GPFirst were 4.2 times higher in CS2 than in CS1 (OR 4.20, 95% CI 2.62-6.73; *P*<.001). Compared to individuals with primary education or lower, those with secondary education and preuniversity education had 2.9 times (95% CI 1.35-6.17; *P*=.006) and 2.6 times (95% CI 1.14-5.70; *P*=.02) higher odds of being knowledgeable about GPFirst respectively. Additionally, individuals without a regular primary care physician had 1.5 times higher odds (95% CI 1.02-2.34;

XSL•FO
RenderX

P=.04) of being knowledgeable about GPFirst compared to those with a regular primary care physician.

Table 4. Adjusted multilogistic regression analysis of factors associated with GPFirst knowledge among participants aware of GPFirst from the precampaign (CS1[a]) and postcampaign (CS2[b]) surveys (N=449).

Variable[c]	Adjusted OR[d] (95% CI)	*P* value
Time		
CS1	1.00	—[e]
CS2	4.20 (2.62-6.73)	<.001
Marital status		
Single	1.00	—
Married	1.02 (0.65-1.59)	.94
Gender		
Men	1.00	—
Women	0.82 (0.55-1.24)	.35
Ethnicity		.14[f]
Chinese	1.00	—
Indian	0.50 (0.22-1.15)	.10
Malay	0.82 (0.51-1.32)	.41
Others	5.82 (0.57-59.73)	.14
Education level		.03[f]
Primary and lower	1.00	—
Secondary	2.88 (1.35-6.17)	.006
Preuniversity	2.56 (1.14-5.70)	.02
University and higher	1.90 (0.80-4.51)	.14
Residential type		.25[f]
1 to 3 room public flats	1.00	—
4 to 5 room public flats	0.76 (0.44-1.31)	0.32
Private and others	1.21 (0.58-2.51)	0.61
Regular primary care physician		
Yes	1.00	—
No	1.54 (1.02-2.34)	0.04

[a]CS1: cross-sectional survey 1 (baseline).

[b]CS2: cross-sectional survey 2.

[c]Age was not included for adjustment due to multicollinearity with marital status.

[d]OR: odds ratio.

[e]Not applicable.

[f]Omnibus test.

Acceptability and Satisfaction of GPFirst

Among the 166 participants who were aware of GPFirst and had visited participating GPFirst clinics in the past 3 months in both CS1 and CS2 (Multimedia Appendix 7), 60.2% (100/166 participants) chose these clinics specifically because they were part of GPFirst. Overall, 98.2% (163/166 participants) expressed their willingness to first visit participating GPFirst clinics in the future, and 95.2% (158/166 participants) said they would recommend participating GPFirst clinics to family and friends as a first stop before the ED. Regarding their overall experience with the GPFirst program, 87.3% (145/166 participants) rated it as “Good” or “Excellent.” Adjusted logistic regression analyses indicated that the 40-59 age group was more likely to choose participating GPFirst clinics because of the clinic’s involvement in the program compared to those in the younger age group (OR 2.33, 95% CI 1.13-4.82; *P*=.02; Multimedia Appendix 7).

XSL•FO
RenderX

Among the 1903 participants in both surveys who were unaware of GPFirst, 88.3% (1680/1903 participants) expressed their willingness to first visit a participating GPFirst clinic in the future. There was a significant association between age groups and the choice of visiting a participating GPFirst clinic (P<.001). Adjusted logistic regression (Multimedia Appendix 7) indicated that individuals aged 60 years and older were less likely to consider visiting a participating GPFirst clinic in the future compared to those aged 21-39 (OR 0.58, 95% CI 0.41-0.82) years (P=.002).

Discussion

Principal Findings

Our study found that awareness of GPFirst increased significantly from CS1 to CS2. Married individuals, those with higher education, and participants without a regular primary care physician were more likely to be aware of GPFirst. Knowledge also improved in CS2, particularly among younger age groups (21-39 years and 40-59 years). Among participants aware of GPFirst, 60.2% (100/166 participants) chose to visit their clinics because they were part of the program, 98.2% (163/166 participants) expressed willingness to return, and 95.2% (158/166 participants) would recommend these clinics. Satisfaction was high, with 87.3% (145/166 participants) rating their experience as “Good” or “Excellent.” However, those aged 40-59 years were less likely to choose clinics specifically due to GPFirst participation. Even among those unaware of GPFirst, 88.3% (1680/1903 participants) were willing to visit a participating clinic instead of the ED in the future. However, older participants (60 years and older) were less likely to consider GPFirst clinics compared to younger groups.

It is evident from our study that the DMC was an effective strategy in engaging the general community and increasing awareness and knowledge levels of GPFirst. Notably, the awareness of GPFirst among CS2 participants increased to 27%, which was 2.7 times higher than that of CS1 after adjusting for potential confounding differences between these 2 groups of participants. While 27% might seem modest compared to other Singapore national health campaigns such as the War on Diabetes, which achieved 61% exposure rates [41,42], it is in line with other awareness campaigns, like those for age-related macular degeneration, which achieved a 28% awareness rate after a 5-year campaign [43]. By leveraging digital platforms, particularly social media, which are widely accessed by diverse age cohorts in Singapore [22], the GPFirst campaign has succeeded in capturing the attention of a broad audience. Interestingly, even among participants in the 60 years and older age group, which typically exhibits lower DM use [22], the odds of being GPFirst aware in the CS2 group was 2.92 compared to the CS1 group. Our findings also revealed an interesting association between not having a regular primary care physician and a higher likelihood of being aware of GPFirst across diverse age groups. The reason for this was unclear in this study and warrants further investigation in future studies. Within our study cohort, a substantial proportion (47.2%) do not have a regular primary care physician. This suggests the potential need for future GPFirst publicity campaigns to include educational elements emphasizing the importance of having a regular primary care physician. Encouraging individuals to have a regular primary care physician can lead to well-documented benefits, such as decreased health care use [44,45]. Additionally, this move will align GPFirst with Singapore’s Healthier SG initiative, a national effort by the Ministry of Health to promote preventive care and empower patients to take steps toward better health [46]. A key aspect of this initiative is to encourage all residents to enroll with a regular family doctor who serves as the first point of contact for health management.

Our results also offered evidence supporting the effectiveness of the campaign in enhancing knowledge levels about GPFirst among the target population. This suggests that DM-driven health promotion campaigns can deliver reliable and trustworthy information to the general public helping to combat misinformation [47-50]. Despite the improvement in knowledge from CS1 to CS2 regarding general practitioners participating in the GPFirst program, 30% (41/137 respondents) still answered the criteria about subsidy eligibility incorrectly. This highlights the need for targeted content to address misconceptions about this aspect in future campaign designs.

The increased proportion of GPFirst knowledgeable respondents among the 21-39 years and 40-59 years age groups suggests that the campaign delivered relevant health promotion and educational materials that resonated with this demographic, potentially leading to a better understanding of GPFirst among young to middle-aged adults. This success might be partially due to our content design and production, which were guided by evidence-based models such as the Andersen model. The 60 years and older age group did not show a statistically significant increase in GPFirst knowledgeable respondents between CS1 and CS2. It is possible that the campaign content was more aligned with the preferences and digital behavior of younger participants, resulting in a smaller impact on older adults. However, the 12% increase in knowledgeable respondents in this age group still reflects some success in engagement, which is important given that this age group contributed to about 15% of nonurgent ED visits [20]. These findings underscore the importance of tailoring health campaigns to effectively target different age groups, ensuring that accurate health information is accessible and well-received by diverse populations [51]. This approach is crucial for the success of public health initiatives in reaching and educating all segments of the community. Similar to our findings on GPFirst awareness, our results also revealed an unexpected association between not having a regular primary care physician and a higher likelihood of being knowledgeable about GPFirst. This should also be addressed in future investigations.

Upon being introduced to the GPFirst program during the survey to participants who were previously unaware of GPFirst, 88% of them expressed their willingness to consider first visiting a participating GPFirst clinic instead of the ED in the future. Significant associations were observed for age groups and choice of health care providers. Older adults who were unaware of GPFirst still have an inclination preference toward EDs even after being introduced to GPFirst. This preference could be due to barriers in accessing primary care, which could include challenges related to transportation and mobility [52], and an

XSL•FO
RenderX

expectation for more timely and specialized care [53,54]. Clearly, this suggests the need to have age-dependent strategies to further improve the uptake of participating GPFirst clinics, particularly among older residents.

Among individuals who visited a participating GPFirst clinic within the last 3 months during the study periods, high levels of acceptability of GPFirst were observed across all age groups. Nearly all respondents expressed their willingness to visit the participating GPFirst clinic again. About 95% also indicated their intention to recommend their family or friends to seek care at the participating GPFirst clinic before the ED in the future. The high percentage of participants who rated their GPFirst experience as "good" or "excellent" also suggests a high satisfaction level of their respective GPFirst experience. Among survey participants who were aware of GPFirst and who visited participating GPFirst clinics in the past 3 months of the study periods, only approximately 60% of participants indicated that they chose to visit their general practitioner clinics because they were participating in GPFirst clinics, with those aged 40-59 years more likely to cite GPFirst participation as a reason for their clinic choice. Evidently, these results highlighted 2 important points. First, the fact that about 40% of individuals who were GPFirst aware did not visit the participating GPFirst clinics because of their affiliation to GPFirst suggests that benefits (eg, financial subsidy and higher priority at ED) offered by GPFirst may not be their key considerations when they decided to visit these participating clinics. Insights from previous qualitative studies may provide possible explanations for this finding [55,56]. While financial subsidies are an important factor, other factors such as positive and trusting general practitioner–patient relationships, clinical accessibility, and interpersonal influences, such as advice from family and friends were also observed to play important roles in shaping health-seeking decisions. Second, the findings also highlighted again the need to have age-dependent strategies to further improve the uptake of participating GPFirst clinics among residents while accounting for other relevant determinants that may drive health-seeking behavior among residents with nonurgent medical conditions [56].

Public Health Implications

The success of the "#ThankYourGP" campaign, which achieved the highest engagement among all post categories, demonstrated the public health potential of user-generated content in fostering interaction on social media, and as a tool for health communication and advocacy. Digital platforms offer a participatory model of engagement, where individuals share personal experiences and narratives, fostering engagement, community-driven health promotion, and peer-to-peer knowledge exchanges [13,57,58]. Beyond increasing engagement and content dissemination, DMCs have demonstrated their ability to drive behavioral change [13,57]. Given the growing reliance on DM for health information, policy makers and health agencies can leverage DMCs as part of broader health promotion strategies, extending their use beyond awareness-building initiatives to influence longer-term health behaviors and service use.

The effectiveness of a DM-driven campaign in raising awareness and knowledge about GPFirst highlights the growing role of digital platforms as essential tools for public health education. Social media and digital platforms allow rapid, scalable, and cost-effective dissemination of information [58]. These platforms can improve accessibility to health information, especially in underserved communities and health systems where traditional media may have limited reach and younger populations who primarily consume health-related content digitally [15,16,24,58]. As digital engagement continues to shape public health communication, integrating such content into national health initiatives could improve health literacy, service use, and preventive health behaviors. Furthermore, algorithm-driven content distribution enables the delivery of tailored health messages to specific demographic groups based on their digital behavior, interests, and search patterns. Health policy makers should consider expanding DMCs beyond awareness and knowledge-building to facilitate behavior change. For instance, linking social media content with telehealth services, appointment scheduling, or digital screening tools to enable seamless access with a single click [58].

The findings of acceptability and satisfaction are particularly relevant to Singapore, given its distinctive health care financing structure and emphasis on public-private partnerships. Singapore's financing system, characterized by a combination of government subsidies, mandatory savings schemes like MediSave, and private insurance, creates a health care landscape that differs from many other systems [59]. These structural elements shape how individuals perceive and engage with health care programs, including initiatives like GPFirst. Additionally, Singapore's emphasis on public-private partnerships facilitates collaboration between government health care providers and private providers like general practitioners, this may serve as a model for other similar health care systems, where alignment and partnerships between public and private providers could improve appropriate health-seeking behaviors [60]. However, scalability beyond Singapore requires consideration of local health system structures [46,55,56]. For health care systems with strong hospital-community partnerships, DMCs for similar programs like GPFirst may be more effective when integrated into primary care outreach programs or coordinated with community public health workers to maximize impact. In settings with fragmented health care systems or high out-of-pocket costs, awareness-building DMCs alone may have limited effectiveness due to existing barriers to care access. In such contexts, additional policy interventions may be necessary to encourage changes in health-seeking behaviors.

Limitations

There are several limitations in our study that should be acknowledged. First, the COVID-19 pandemic significantly impacted ED visits, making it difficult to reliably obtain rates of nonurgent ED visits after the rolling out of the DM-driven GPFirst campaign. The confirmation of the first COVID-19 case in Singapore on January 23, 2020, followed by the development of multiple local transmission clusters and the implementation of a lockdown in April 2020, prevented us from conducting a meaningful comparison of nonurgent ED visit rates before and after the roll-out of DM-driven GPFirst

XSL•FO
RenderX

campaign. Second, the sample of individuals with a recent encounter with a participating GPFirst clinic was relatively small. Therefore, their perceptions of acceptability and satisfaction with GPFirst may not be fully representative of the past GPFirst patients' views and experiences. Third, the inclusion of only English-literate participants may affect the generalizability of results to Singapore's multicultural society. Finally, it is important to note that this study used a convenience sample at 2 different time points, with recruitment conducted at public locations in eastern Singapore. This approach may have introduced selection bias, as individuals who frequent public places might differ systematically from those who do not, such as housebound individuals or those who rely on private transportation. Additionally, variations in participant characteristics across the 2 time points, as well as potential self-selection bias (where those more interested or available chose to participate), could have influenced the findings. To address this, a more rigorous longitudinal study design using household surveys would be preferable to track changes in knowledge, awareness, and behavior over time.

Conclusions

This study has offered evidence that a carefully crafted DM-driven campaign can have a positive impact on the awareness and knowledge of GPFirst among the target population across all age groups. Our surveys show high population-wide acceptance of GPFirst and high satisfaction with GPFirst among the program participants. The findings also highlighted the need to have age-dependent strategies to further improve participation among residents. Overall, our experience does suggest that a publicity campaign that uses a mix of social media platforms (Facebook) and digital methods may be adopted for other public health campaigns with similar objectives of modifying the population's health-seeking behaviors through public health education and promotion efforts.

Acknowledgments

The authors would like to thank all past and present GPFirst program team members, such as TK Udairam, Mohan Tiruchittampalam, and Yap Mei Foon. The authors gratefully acknowledge the support and contributions of the Regional Health System GP Engagement team, CGH's departments and teams, including ED, data management and informatics, corporate communications, finance, procurement, and the vendors engaged for the campaign implementation and data collection. The authors would like to thank Mon Hnin Tun for her guidance in addressing the type 1 error correction in the data analysis. The authors would also like to thank all of the study participants for their time. This study was supported by the Ministry of Health (reference MH006:013), and the publication fee of this paper was funded by the NMRC through the SingHealth PULSES II Centre Grant (CG21APR1013). The funders did not have any role in the study design, data collection, analysis, interpretation, and preparation of the manuscript.

Authors' Contributions

ROHS wrote the manuscript, conducted all statistical analyses, curated data, developed methodology, and provided data for all tables. HCO, PGSK, LLSH, MRBMR, and EAA secured funding, conceptualized the study, supervised, managed administration, and contributed to data curation, visualization, and manuscript review. JR, CHH, and SLHC secured funding and contributed to the study conceptualization, supervision, and methodology, with SLHC also supporting visualization. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
GPFirst Facebook Posts.
[DOCX File , 278 KB - publichealth_v11i1e66062_app1.docx]

Multimedia Appendix 2
GPFirst survey.
[DOCX File , 26 KB - publichealth_v11i1e66062_app2.docx]

Multimedia Appendix 3
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.
[DOCX File , 32 KB - publichealth_v11i1e66062_app3.docx]

Multimedia Appendix 4
Survey participants' characteristics stratified by age groups.
[DOCX File , 17 KB - publichealth_v11i1e66062_app4.docx]

Multimedia Appendix 5

XSL•FO
RenderX

Comparison of GPFirst awareness levels.
[DOCX File , 16 KB - publichealth_v11i1e66062_app5.docx]

Multimedia Appendix 6
Comparison of correct response rates for each knowledge question.
[DOCX File , 18 KB - publichealth_v11i1e66062_app6.docx]

Multimedia Appendix 7
Acceptability and satisfaction of GPFirst.
[DOCX File , 23 KB - publichealth_v11i1e66062_app7.docx]

References

1. Lee A, Lau FL, Hazlett CB, Kam CW, Wong P, Wong TW, et al. Factors associated with non-urgent utilization of accident and emergency services: a case-control study in Hong Kong. Soc Sci Med 2000;51(7):1075-1085. [doi: 10.1016/s0277-9536(00)00039-3] [Medline: 11005394]
2. Trzeciak S, Rivers EP. Emergency department overcrowding in the United States: an emerging threat to patient safety and public health. Emerg Med J 2003;20(5):402-405 [FREE Full text] [doi: 10.1136/emj.20.5.402] [Medline: 12954674]
3. Uscher-Pines L, Pines J, Kellermann A, Gillen E, Mehrotra A. Emergency department visits for nonurgent conditions: systematic literature review. Am J Manag Care 2013;19(1):47-59 [FREE Full text] [Medline: 23379744]
4. Richardson DB. Increase in patient mortality at 10 days associated with emergency department overcrowding. Med J Aust 2006;184(5):213-216. [doi: 10.5694/j.1326-5377.2006.tb00204.x] [Medline: 16515430]
5. Kulstad EB, Sikka R, Sweis RT, Kelley KM, Rzechula KH. ED overcrowding is associated with an increased frequency of medication errors. Am J Emerg Med 2010;28(3):304-309. [doi: 10.1016/j.ajem.2008.12.014] [Medline: 20223387]
6. Xin H, Kilgore ML, Sen BP, Blackburn J. Can nonurgent emergency department care costs be reduced? Empirical evidence from a U.S. nationally representative sample. J Emerg Med 2015;49(3):347-354. [doi: 10.1016/j.jemermed.2015.01.034] [Medline: 25846051]
7. Pang M. The Straits Times. 2013. URL: https://web.archive.org/web/20241124230511/http://www.straitstimes.com:80/singapore/ae-units-flooded-with-non-emergency-cases [accessed 2025-03-06]
8. Oh HC, Sridharan S, Yap MF, Goh PSK, Lee LSH, Venkataraman N, et al. Impact of a primary care partnership programme on accident and emergency attendances at a regional hospital in Singapore: a pilot study. Singapore Med J 2023;64(8):534-537 [FREE Full text] [doi: 10.11622/smedj.2021157] [Medline: 34628785]
9. Tay A, Choo KW, Koh GCH. Singapore GP fee survey 2013: a comparison with past surveys. Singapore Fam Physician 2017;43(1):42-51. [doi: 10.33591/sfp.44.1.o1]
10. Hospital overview. Changi General Hospital. URL: https://www.cgh.com.sg/about-us/Pages/hospital-overview.aspx [accessed 2024-07-03]
11. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40. [doi: 10.2196/jmir.2952] [Medline: 24550083]
12. Powell R, Rosenthal J, August EM, Frey M, Garcia L, Sidibe T, et al. Ante La Duda, Pregunta: a social marketing campaign to improve contraceptive access during a public health emergency. Health Commun 2022;37(2):177-184. [doi: 10.1080/10410236.2020.1828534] [Medline: 33016136]
13. Grantham JL, Verishagen CL, Whiting SJ, Henry CJ, Lieffers JRL. Evaluation of a social media campaign in Saskatchewan to promote healthy eating during the COVID-19 pandemic: social media analysis and qualitative interview study. J Med Internet Res 2021;23(7):e27448. [doi: 10.2196/27448] [Medline: 34133314]
14. Hui A, Wong PWC, Fu KW. Evaluation of an online campaign for promoting help-seeking attitudes for depression using a Facebook advertisement: an online randomized controlled experiment. JMIR Ment Health 2015;2(1):e5 [FREE Full text] [doi: 10.2196/mental.3649] [Medline: 26543911]
15. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
16. Woolley P, Peterson M. Efficacy of a health-related Facebook social network site on health-seeking behaviors. Soc Mark Q 2012;18(1):29-39. [doi: 10.1177/1524500411435481]
17. Booth RG, Allen BN, Bray Jenkyn KM, Li L, Shariff SZ. Youth mental health services utilization rates after a large-scale social media campaign: population-based interrupted time-series analysis. JMIR Ment Health 2018;5(2):e27. [doi: 10.2196/mental.8808] [Medline: 29625954]
18. Roberts-Lewis S, Baxter H, Mein G, Quirke-McFarlane S, Leggat FJ, Garner H, et al. Examining the effectiveness of social media for the dissemination of research evidence for health and social care practitioners: systematic review and meta-analysis. J Med Internet Res 2024;26:e51418 [FREE Full text] [doi: 10.2196/51418] [Medline: 38838330]

XSL•FO
RenderX

19. Zhang X, Chen X, Kourkoumelis N, Gao R, Li G, Zhu C. A social media-promoted educational community of joint replacement patients using the WeChat app: survey study. JMIR Mhealth Uhealth 2021;9(3):e18763 [FREE Full text] [doi: 10.2196/18763] [Medline: 33734094]
20. Oh HC, Chow WL, Gao Y, Tiah L, Goh SH, Mohan T. Factors associated with inappropriate attendances at the emergency department of a tertiary hospital in Singapore. Singapore Med J 2020;61(2):75-80 [FREE Full text] [doi: 10.11622/smedj.2019041] [Medline: 31044259]
21. Annual survey on infocomm usage in households and by individuals for 2019. Infocomm Media Development Authority. URL: https://www.imda.gov.sg/-/media/imda/files/industry-development/fact-and-figures/infocomm-usage-business/infocomm-usage-survey-public-report-2019.pdf [accessed 2024-12-09]
22. Meltwater WAS. Digital 2019 Singapore. DataReportal. 2019. URL: https://datareportal.com/reports/digital-2019-singapore?rq=singapore [accessed 2024-07-03]
23. Evans-Lacko S, Henderson C, Thornicroft G, McCrone P. Economic evaluation of the anti-stigma social marketing campaign in England 2009-2011. Br J Psychiatry Suppl 2013;55:s95-101. [doi: 10.1192/bjp.bp.112.113746] [Medline: 23553701]
24. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
25. Lau PWC, Wang JJ, Ransdell LL, Shi L. The effectiveness of Facebook as a social network intervention to increase physical activity in Chinese young adults. Front Public Health 2022;10:912327 [FREE Full text] [doi: 10.3389/fpubh.2022.912327] [Medline: 35937270]
26. Braveman P. Health disparities and health equity: concepts and measurement. Annu Rev Public Health 2006;27:167-194. [doi: 10.1146/annurev.publhealth.27.021405.102103] [Medline: 16533114]
27. Braveman P, Arkin E, Orleans T, Proctor D, Acker J, Plough A. What is health equity? Behav Sci Policy 2018;4(1):1-14. [doi: 10.1353/bsp.2018.0000]
28. Hahn RA, Truman BI. Education improves public health and promotes health equity. Int J Health Serv 2015;45(4):657-678 [FREE Full text] [doi: 10.1177/0020731415585986] [Medline: 25995305]
29. Wilson T. Models in information behaviour research. J Doc 1999;55:249-270. [doi: 10.1108/EUM0000000007145] [Medline: 25995305]
30. Pálsdóttir A. The connection between purposive information seeking and information encountering. J Doc 2010;66(2):224-244. [doi: 10.1108/00220411011023634]
31. Niederdeppe J, Hornik RC, Kelly BJ, Frosch DL, Romantan A, Stevens RS, et al. Examining the dimensions of cancer-related information seeking and scanning behavior. Health Commun 2007;22(2):153-167. [doi: 10.1080/10410230701454189] [Medline: 17668995]
32. Andersen RM. Revisiting the behavioral model and access to medical care: Does it matter? J Health Soc Behav 1995;36(1):1-10. [Medline: 7738325]
33. Andersen RM. National health surveys and the behavioral model of health services use. Med Care 2008;46(7):647-653. [doi: 10.1097/MLR.0b013e31817a835d] [Medline: 18580382]
34. Babitsch B, Gohl D, von Lengerke T. Re-revisiting andersen's behavioral model of health services use: a systematic review of studies from 1998-2011. Psychosoc Med 2012;9:Doc11 [FREE Full text] [doi: 10.3205/psm000089] [Medline: 23133505]
35. GPFirst Facebook. Facebook. URL: https://www.facebook.com/GPFirst.sg/ [accessed 2024-09-02]
36. GPFirst: Your family doctor, your first stop. SingHealth. URL: https://www.singhealth.com.sg/rhs/get-well/gpfirst [accessed 2024-12-05]
37. Taro Y. Statistics: An Introductory Analysis. 2nd edition. New York, NY: Harper and Row; 1967.
38. Narum SR. Beyond Bonferroni: less conservative analyses for conservation genetics. Conserv Genet 2006;7(5):783-787. [doi: 10.1007/s10592-005-9056-y]
39. Cunningham P, May J. Insured Americans drive surge in emergency department visits. Issue Brief Cent Stud Health Syst Change 2003(70):1-6. [Medline: 14577417]
40. Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. Health Promot Pract 2012;13(2):159-164. [doi: 10.1177/1524839911433467] [Medline: 22382491]
41. Khow YZ, Lim TLY, Ng JSP, Wu J, Tan CS, Chia KS, et al. Behavioral impact of national health campaigns on healthy lifestyle practices among young adults in Singapore: a cross-sectional study. BMC Public Health 2021;21(1):1601. [doi: 10.1186/s12889-021-11628-5] [Medline: 34461867]
42. Bee YM, Tai ES, Wong TY. Singapore's "War on Diabetes". Lancet Diabetes Endocrinol 2022;10(6):391-392. [doi: 10.1016/S2213-8587(22)00133-4] [Medline: 35483379]
43. Sanjay S, Chin YC, Teo HT, Ong SX, Toh SHF, Khong MH, et al. A follow-up survey on the knowledge of age-related macular degeneration and its risk factors among Singapore residents after 5 years of nation-wide awareness campaigns. Ophthalmic Epidemiol 2014;21(4):230-236. [doi: 10.3109/09286586.2014.929708] [Medline: 24983763]
44. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. Milbank Q 2005;83(3):457-502 [FREE Full text] [doi: 10.1111/j.1468-0009.2005.00409.x] [Medline: 16202000]

XSL•FO
RenderX

45. van Walraven C, Oake N, Jennings A, Forster AJ. The association between continuity of care and outcomes: a systematic and critical review. J Eval Clin Pract 2010;16(5):947-956. [doi: 10.1111/j.1365-2753.2009.01235.x] [Medline: 20553366]
46. Promoting overall healthier living while targeting specific sub-populations. Ministry of Health Singapore. URL: https://www.moh.gov.sg/news-highlights/details/promoting-overall-healthier-living-while-targeting-specific-sub-populations [accessed 2024-07-03]
47. Mheidly N, Fares J. Leveraging media and health communication strategies to overcome the COVID-19 infodemic. J Public Health Policy 2020;41(4):410-420 [FREE Full text] [doi: 10.1057/s41271-020-00247-w] [Medline: 32826935]
48. Ortiz RR, Smith A, Coyne-Beasley T. A systematic literature review to examine the potential for social media to impact HPV vaccine uptake and awareness, knowledge, and attitudes about HPV and HPV vaccination. Hum Vaccin Immunother 2019;15(7-8):1465-1475 [FREE Full text] [doi: 10.1080/21645515.2019.1581543] [Medline: 30779682]
49. Hauer MK, Sood S. Using social media to communicate sustainable preventive measures and curtail misinformation. Front Psychol 2020;11 [FREE Full text] [doi: 10.3389/fpsyg.2020.568324] [Medline: 33178073]
50. Farsi D, Martinez-Menchaca HR, Ahmed M, Farsi N. Social media and health care (part II): narrative review of social media use by patients. J Med Internet Res 2022;24(1):e30379 [FREE Full text] [doi: 10.2196/30379] [Medline: 34994706]
51. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85. [doi: 10.2196/jmir.1933] [Medline: 23615206]
52. Colorafi K, Thomas A, Wilson M, Corbett CF. Perspectives of aging adults who frequently seek emergency department care. Pain Manag Nurs 2021;22(2):184-190. [doi: 10.1016/j.pmn.2020.10.003] [Medline: 33317936]
53. Lowthian JA, Smith C, Stoelwinder JU, Smit DV, McNeil JJ, Cameron PA. Why older patients of lower clinical urgency choose to attend the emergency department. Intern Med J 2013;43(1):59-65. [doi: 10.1111/j.1445-5994.2012.02842.x] [Medline: 22646852]
54. Goodridge D, Stempien J. Understanding why older adults choose to seek non-urgent care in the emergency department: the patient's perspective. CJEM 2019;21(2):243-248. [doi: 10.1017/cem.2018.378] [Medline: 29843840]
55. Goh WX, Ong RHS, Rebello SA, Goh PSK, Ahmad EA, How CH, et al. Patients' perspectives on a programme to reduce emergency department use for low acuity conditions. Proc Singapore Healthc 2024;33. [doi: 10.1177/20101058241268494]
56. Koh VTJ, Ong RHS, Chow WL, Tiah L, Oh HC, Yow WQ, et al. Understanding patients' health-seeking behaviour for non-emergency conditions: a qualitative study. Singapore Med J 2023 [FREE Full text] [doi: 10.4103/singaporemedj.SMJ-2020-494] [Medline: 37675670]
57. Heldman AB, Schindelar J, Weaver JB. Social media engagement and public health communication: implications for public health organizations being truly "social". Public Health Rev 2013;35(1):13. [doi: 10.1007/BF03391698]
58. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021;23(5):e17917 [FREE Full text] [doi: 10.2196/17917] [Medline: 33978589]
59. Tan CC, Lam CSP, Matchar DB, Zee YK, Wong JEL. Singapore's health-care system: key features, challenges, and shifts. Lancet 2021;398(10305):1091-1104. [doi: 10.1016/S0140-6736(21)00252-X] [Medline: 34481560]
60. Torchia M, Calabrò A, Morner M. Public–private partnerships in the health care sector: a systematic review of the literature. Public Manag Rev 2013;17(2):236-261. [doi: 10.1080/14719037.2013.792380]

Abbreviations

CGH: Changi General Hospital
CS1: cross-sectional survey 1
CS2: cross-sectional survey 2
DM: digital media
DMC: digital media campaign
EARC: EA Research and Consulting Pte Ltd
ED: emergency department
FDR: false discovery rate
OR: odds ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 03.09.24; peer-reviewed by CR Wells, LV Ivanitskaya; comments to author 09.11.24; revised version received 29.12.24; accepted 21.02.25; published 14.04.25.

Please cite as:
Ong Hui Shan R, Oh HC, Goh Sook Kheng P, Lee Sze Hui L, Riza Bte Mohd Razali M, Ahmad EA, Raghuram J, How CH, Lim Hoon Chin S
Evaluation of a Digital Media Campaign to Promote Knowledge and Awareness of the GPFirst Program for Nonurgent Conditions: Repeated Survey Study
JMIR Public Health Surveill 2025;11:e66062
URL: https://publichealth.jmir.org/2025/1/e66062
doi:10.2196/66062
PMID:40228291

©Rebecca Ong Hui Shan, Hong Choon Oh, Priscilla Goh Sook Kheng, Lyndia Lee Sze Hui, Mas Riza Bte Mohd Razali, Edris Atikah Ahmad, Jagadesan Raghuram, Choon How How, Steven Lim Hoon Chin. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The American Transformative HIV Study: Protocol for a US National Cohort of Sexual and Gender Minority Individuals With HIV

Christian Grov[1], MPH, PhD; Alexa B D'Angelo[1], MPH, PhD; Chloe Mirzayi[1], MPH, PhD; Michelle Dearolf[1], MPH; Elena Hoeppner[1], MPH; Yan Guo[1], MS; Nicole Richards[1], MPH; Rifa Ehsan[1], MPH; Sarah Kulkarni[1], MPH; Denis Nash[1], MPH, PhD; Viraj V Patel[2], MD; Dustin T Duncan[3], ScD; Meredith Ray[4], PhD; Tyler Bartholomew[5], PhD; Jennifer Manuzak[6], PhD; Jennifer Manuel[7], PhD; Kathryn McCollister[5], PhD; Drew Westmoreland[8], PhD; Adam W Carrico[9], PhD

[1]City University of New York Graduate School of Public Health and Health Policy, 55 West 125 Street, New York, NY, United States

[2]Albert Einstein College of Medicine, Montefiore Health System, New York, NY, United States

[3]Columbia University, New York, NY, United States

[4]The University of Memphis, Memphis, TN, United States

[5]University of Miami, Coral Gables, FL, United States

[6]Tulane University, New Orleans, LA, United States

[7]University of California San Francisco, San Francisco, CA, United States

[8]University of Florida, Gainesville, FL, United States

[9]Florida International University, Miami, FL, United States

Corresponding Author:
Christian Grov, MPH, PhD
City University of New York Graduate School of Public Health and Health Policy, 55 West 125 Street, New York, NY, United States

Abstract

Background: Sexual and gender minority (SGM) individuals represent 2% - 5% of the US population, yet continue to account for more than two-thirds of new HIV infections annually.

Objective: This study seeks to identify multilevel (ie, structural, psychological, and social) and biobehavioral (ie, rectal cytokines or chemokines) determinants of amplified HIV seroconversion risk for SGM individuals, including those who use methamphetamine.

Methods: The American Transformative HIV Study is an ongoing web-based cohort study of 5364 SGM individuals from all 50 US states and Puerto Rico, enrolled in 2022 and 2023, and will be followed through 2027. We oversampled persons who use methamphetamine (2846/5364, 53.1%). We used established web-based strategies to enroll individuals aged 16 - 49 years at high risk of HIV acquisition via sexual networking apps. To be eligible, participants had to report meeting objective criteria for HIV pre-exposure prophylaxis (PrEP) care, but not be taking PrEP. Participants complete annual web-based surveys (baseline, 12, 24, and 36 months) and are asked to provide self-collected oral fluid samples for HIV testing and 2 rectal swabs (the Aptima Multitest Swab and the Zymo DNA/RNA Shield swab) following each assessment. Oral fluid samples are analyzed immediately, while rectal swabs are banked for a future nested case-cohort analysis to assess changes in inflammatory markers following a new infection.

Results: Nearly all participants (4542/5364, 84.7%) were cisgender men, 3.7% (201/5364) were transgender women, and 1.1% (61/5364) were transgender men who have sex with men. There were also 560 (10.4%) individuals who self-identified outside of the gender binary—all reported being assigned male sex at birth. In total, 56.5% (3031/5364) were persons of color, and 31.8% (1714/5365) were aged 16 to 29 years. In total, 4054 baseline HIV test kits were returned, including 371 HIV reactive samples—3.3% (69/2210) were HIV-positive among those who did not report methamphetamine use, and 15.5% (302/1944) were HIV-positive among those reporting methamphetamine use. Based on participant's HIV results as well as self-reporting when their most recent prior HIV-negative test was, we estimated that the incidence rate in this cohort in the 12-month period leading up to study enrollment was 10.06 (95% CI 8.65 - 11.64) per 100 person-years among those reporting methamphetamine use compared with 2.04 (95% CI 1.49 - 2.73) among those not reporting methamphetamine use per 100 person-years.

XSL•FO
RenderX

Conclusions: A large, US national, and racially diverse web-based cohort of SGM individuals at high risk for HIV has been successfully enrolled and will be followed through 2027. Persons who use methamphetamine have been oversampled and demonstrated an exceptionally greater risk for HIV. Our study will offer insight into the development and implementation of new interventions, which aim to have a meaningful impact on HIV transmission.

(JMIR Public Health Surveill 2025;11:e66921) doi:10.2196/66921

KEYWORDS

HIV testing; rectal cytokines or chemokines; cohort research; nested case-cohort designs; American Transformative HIV Study; AMETHST; United States; sexual and gender minority; methamphetamine; web-based survey; oral fluid; rectal swab

Introduction

Background

Sexual and gender minority (SGM) individuals represent 2% - 5% of the US population, yet continue to account for more than two-thirds of new HIV infections annually [1-4]. The Centers for Disease Control and Prevention has estimated that 1 in 6 sexual minority men will acquire HIV in their lifetime, including half of Black sexual minority men and one-quarter of Latino sexual minority men [2,5]. Meanwhile, gender minority individuals (transgender men and transgender women) accounted for 2% of new HIV infections in 2019 (while representing ~1% of the US population) [6-8], with transwomen of color most impacted [8]. A Centers for Disease Control and Prevention survey of transgender women across 7 US cities found that 42% of respondents were living with HIV, with 65% of American Indian or Alaska Natives, 62% of Black transwomen, and 35% of Latinx transwomen living with HIV [9]. Our ability to avert new infections may be further challenged by racial disparities in pre-exposure prophylaxis (PrEP) uptake [10].

Methamphetamine use is an enduring and resurgent epidemic in SGM communities [11-15]. Rates of methamphetamine use are high among SGM individuals, ranging from 7.4% to 30%. Although methamphetamine use may have declined somewhat after significant public health attention in the early- and mid-2000s [16,17], it is again on the rise [16,18-22]. Findings from our team and others [23,24] demonstrate that methamphetamine use is disproportionately impacting SGM individuals who are from racial and ethnic minority groups and accounts for one-in-three new HIV infections in SGM individuals [25]. There are behavioral, structural, psychosocial, and biological pathways through which methamphetamine increases the risk for HIV transmission.

Mechanisms That Increase HIV Vulnerability

Behavioral

Methamphetamine increases sexual libido while decreasing inhibitions [26-28]. Because it is a stimulant, sexual minority men using methamphetamine will often engage in condomless anal sex (CAS) with multiple partners for prolonged periods (1 - 3 days) [28-30], with less regard to HIV prevention modalities.

Structural

High rates of methamphetamine use and its social acceptability in sexual minority men communities potentiate exposure risk through social or sexual networks [31-34]. Many sexual minority men are also highly likely to reside in concentrated neighborhoods within urban centers where the public health impact of methamphetamine use is amplified. Meanwhile, institutionalized and cultural stigma, homophobia, and discrimination affect the health of SGM individuals [35-41], which can keep them from seeking HIV prevention services and substance use disorder treatment [42]. These can include individualized minority stress exposures (eg, being subject to a homophobic slur) and structural determinants (eg, living in an area with unfavorable SGM policies or laws).

Psychosocial

There are known psychosocial determinants that increase the risk for methamphetamine use and HIV. Syndemic theory posits that multiple co-occurring epidemics work synergistically to increase—for instance—HIV vulnerability in SGM individuals [43-48]. In syndemic theory, “epidemics” refers both to epidemics of disease [49] (eg, mental health issues, methamphetamine use, and HIV) and social epidemics (eg, poverty, racism, and homophobia). Among SGM individuals—especially SGM persons of color—simultaneous and synergistic epidemics increase HIV vulnerability [43,47,48,50,51]. For example, studies of Black and Latino sexual minority men draw associations among methamphetamine use, mental health outcomes, and condomless sex, all of which could amplify HIV vulnerability [51].

Biological

Methamphetamine potentiates rectal immune dysregulation in SGM individuals, which could amplify the biological risk for HIV. Approximately 70% of HIV infections in sexual minority men occur during receptive CAS [52-54], but little is known about the correlates and consequences of rectal immune dysregulation relevant to HIV [55]. Methamphetamine directly decreases parasympathetic tone [56], increasing intestinal permeability and altering rectal immune responses by changing the gut microbiome [57]. Methamphetamine also directly damages gut barrier integrity, which could amplify rectal immune dysregulation [58].

Study Objective and Hypothesis

The objective of this study is to identify multilevel (ie, structural, psychological, and social) and biological (ie, rectal cytokines or chemokines) determinants of amplified HIV seroconversion risk for SGM individuals who use methamphetamine to guide targeted, combination interventions to reduce HIV incidence. We hypothesize that those reporting methamphetamine use

XSL•FO
RenderX

would have higher HIV prevalence as well as higher recent HIV incidence than those not reporting methamphetamine use.

Methods

Overview

We sought to recruit, via sexual networking apps, a racially and geographically diverse sample of HIV-negative SGM individuals who are not on PrEP, to better inform the design, implementation, scale-up, and evaluation of HIV prevention programs. Moreover, to understand the multilevel biobehavioral pathways through which methamphetamine amplifies HIV risk, we sought to overenroll persons who use methamphetamine. We describe the protocol and baseline participant characteristics for the American Transformative HIV Study (AMETHST).

Target Population

The AMETHST cohort study used established [59,60], web-based strategies to enroll a large sample of HIV-negative men, transmen, and transwomen who have sex with men, aged 16 to 49 years, and who are at high risk of HIV acquisition—defined as meeting objective criteria for PrEP care but *not* being on PrEP at the time of enrollment. The cohort will be followed prospectively for 36 months to examine outcomes including PrEP uptake or discontinuation, HIV acquisition, changing patterns of methamphetamine use, and rectal immune dysregulation. We aimed to enroll a cohort of participants at high risk for HIV that was geographically diverse (ie, with representation from every US state), racially and ethnically diverse (n=3031, 56.5% are participants of color), and young (n=1704, 31.8% aged 16 - 29 years). We overenrolled persons who reported recent (<3 months) methamphetamine use (n=2518). All study procedures were approved by the City University of New York institutional review board, and we obtained a waiver of parental consent for those aged 16 and 17 years. Those aged 16 or 17 years were presented with additional multiple-choice questions about their comprehension of the study consent form (that they were required to answer correctly) as part of the assent process. We include those aged 16 and 17 years because data suggest that this is about the age when many men who have sex with men begin to become sexually active with other men [61], and thus vulnerable to HIV. We obtained a waiver of parental consent, given the minimal risk posed by the study and that requiring parental consent could potentially endanger participants who may not be "out" to their parents [62].

Cohort Eligibility and Recruitment

Enrollment for AMETHST began in August 2022 and concluded in August 2023. Participants were recruited via advertisements on men-for-men geosocial sexual networking mobile phone apps. Some advertisements were more generic in nature, while others used imagery (eg, smoke exhaled from methamphetamine use) and textual cues that would suggest the study was seeking to enroll those who use methamphetamine (see Multimedia Appendix 1 for example advertisements). Of note, app developers forbid both its user base as well as advertisers from mentioning drug use directly. To subvert this, users have developed evolving coded language or emojis to communicate interest in finding meth-using partners without setting off automatic filters or censors [63,64]. Our advertisements were geotargeted to individuals using apps inside the United States. Eligibility criteria are shown in Table 1; however, once we reached saturation of individuals who do *not* use methamphetamine, we limited enrollment to just those who met these criteria and reported recent methamphetamine use (defined as at least 1 instance of use in the past 3 months). Apps themselves require users to *indicate* that they are older than 18 years of age. Individuals younger than 18 years of age are known to bypass this by saying that they are older than 18 years when they join. It is for this reason we allowed for individuals aged 16 and 17 years of age to enroll in our study (n=123).

Potential participants were directed to a secure screening survey in their device's web browser and presented with a screen describing study participation and eliciting informed consent. The screening survey took approximately 5 - 10 minutes to complete and included core eligibility criteria for the study. As the mpox outbreak was peaking in the United States at the time recruitment launched, the survey also included questions on mpox, the results of which have been published elsewhere [65]. There was no incentive for completing the screening survey. However, those who were eligible were sent a baseline survey, for which they received an incentive. Upon completion of the baseline survey, participants were mailed at-home collection kits for HIV testing (ie, oral fluid sample) and rectal swabs—all to be returned via mail to a laboratory for testing or storage. Upon receipt at the laboratory, participants received an incentive for their HIV test kit and a second incentive for returning the rectal swabs (described below).

XSL•FO
RenderX

Table . Eligibility criteria for the American Transformative HIV Study.

Eligibility criteria	Participants who used methamphetamine in the last 3 months (n=2846), n (%)	Participants who had not used methamphetamine in the last 3 months (n=2518), n (%)	Chi-square (*df*)	*P* value
Core eligibility criteria (all participants must meet these criteria)				—[a]
Aged 16 to 49 years	2846 (100)	2518 (100)	—	
At least 2 men sex partners in the past 90 days	2846 (100)	2518 (100)	—	
Not currently participating in a clinical trial for an HIV vaccine or pre-exposure prophylaxis	2846 (100)	2518 (100)	—	
Not currently on pre-exposure prophylaxis	2846 (100)	2518 (100)	—	
Never diagnosed with HIV (self-report)	2846 (100)	2518 (100)	—	
Currently residing in the United States or territories	2846 (100)	2518 (100)	—	
Not cisgender woman	2846 (100)	2518 (100)	—	
Additional eligibility criteria (participants must meet at least 1)				
≥1 receptive condomless anal sex acts with a men partner in the last 6 months	2384 (83.8)	2102 (83.5)	0.1 (1)	.78
≥3 insertive condomless anal sex acts with a men partner in the last 6 months	1913 (67.2)	1277 (50.7)	150.9 (1)	<.001
Used methamphetamine in the last 3 months	2846 (100)	0 (0)	—	—
Syphilis diagnosis in the last 6 months	149 (5.2)	55 (2.2)	34.0 (1)	<.001
Shared injection drug needles in the last 12 months	236 (8.3)	3 (0.1)	209.6 (1)	<.001

[a]Not applicable.

Preventing Fraudulent Participation

We enacted multiple measures to minimize repeat participation and fraudulent participants, including recording IP addresses and using cookies to block repeated attempts. Our databases flagged duplicate contact information. Mailing addresses were validated with the US postal service, and participants were required to provide a unique valid phone number (ie, text verification). Multiple entries were identified by email addresses, mailing addresses, or phone numbers. In addition, the data manager manually checked for duplicate entries during baseline data collection. We also assessed the time to completion of our web-based surveys and checked for variability in response sets.

Baseline Web-Based Survey

Participants were emailed and texted a unique link to a baseline survey. The web-based survey assessed psychosocial, demographic, social network, and behavioral characteristics (behaviors including sexual behavior and substance use). The full measures for the screening survey and baseline survey are included in Multimedia Appendix 2. The survey took approximately 45 minutes to complete, and participants were compensated US $40 for completing the survey.

Biomarkers

Upon completing the baseline survey, participants were sent at-home specimen collection kits. This kit included the OraSure HIV-1 specimen collection device and 2 rectal swabs (the Aptima Multitest Swab and the Zymo DNA/RNA Shield swab). Participants were also provided access to a study video along with printed instructions on collecting their specimens. Upon receipt at the laboratory, oral fluid samples were used for HIV antibody testing using the Avioq HIV-1 enzyme immunoassay [66]. Any samples testing positive with this assay were subsequently tested with the Food and Drug Administration–cleared BioRad Geenius HIV-1/2 as a supplemental antibody test, which has been shown to reliably detect HIV-1 antibodies in oral fluid specimens with mean enzyme immunoassay signal-to-cutoff greater than or equal to 3.00 [66]. The laboratory returned all results to our team, and we in turn informed the participants. Those with HIV-negative results were informed by email, and those with HIV-reactive

XSL•FO
RenderX

results were informed via phone following a published protocol [67].

Rectal swabs received by the laboratory are being banked or frozen until all follow-up assessments are complete, at which point we will be conducting a nested case-cohort study [68] to examine biological vulnerability to HIV. This is described in greater detail below. Participants were compensated US $25 for returning their HIV test kit and US $25 for returning rectal swabs.

Participants With Unknown Baseline HIV Status

Participants who screened eligible and completed a baseline questionnaire, but did not return an HIV test kit (baseline serostatus not confirmed with a biomarker), will continue to be followed and asked at regular intervals to submit an oral fluid sample for HIV testing using the at-home sampling kit. For the purposes of prospectively estimating HIV incidence in the AMETHST cohort, these individuals will be excluded. However, we will conduct post hoc sensitivity analyses that make assumptions about having similar, lower, or higher HIV risk profiles than the AMETHST participants for whom baseline HIV status was determined. Key questions answered by including these participants will include identifying psychosocial and demographic factors associated with a timeline to sample return (for those who return a sample at a future wave) as well as how these factors differentiate those who do not return a sample going forward.

HIV Incidence in the 12 Months Before Cohort Enrollment

We estimated pre-enrollment HIV incidence using baseline data on HIV serostatus and HIV testing history prior to joining the study. Using these data, we reconstructed a cohort of individuals who could all be classified as HIV-negative as of 12 months before cohort enrollment. For those HIV-positive at enrollment, we estimated the time of seroconversion using self-reported data on the timing of their last HIV-negative test. Specifically, those self-reporting a negative HIV test within 6 months of AMETHST enrollment were classified as being HIV-negative as of 6 months before enrollment, with seroconversion timing assumed to be distributed evenly during the 6 months leading up to study enrollment. Similarly, participants self-reporting a negative HIV test 7 to 12 months before AMETHST enrollment were classified as being HIV-negative as of 12 months before enrollment, with seroconversion timing assumed to be distributed evenly during the 12 months leading up to study enrollment. We also estimated HIV incidence for the 6-month period leading up to cohort enrollment. For the purposes of identifying boundaries around minimum and maximum incidence rates, we conducted sensitivity analyses representing assumptions for the timing of seroconversion at both extremes.

All participants are provided information on where they can access free and low-cost HIV testing outside of the study as well as information on PrEP and where they can get access to PrEP for no or low cost. Participants are also welcome to reach out to our team if they need assistance getting connected to HIV prevention resources or HIV care.

Statistical Power

This study is sufficiently powered for all future analyses assessing the main aims of the cohort (ie, impact of methamphetamine on HIV seroconversion and PrEP uptake) considering expected HIV seroconversion rates, PrEP uptake rates, methamphetamine use, and response rates based on our prior work [69].

Ethical Considerations

All participants provided informed consent. The AMETHST study protocol was approved by the Institutional Review Board of the City University of New York (approval 2022-0378-PHHP). Participant contact information is stored separately from their data or responses and linked only by a unique ID. At enrollment, participants were compensated US $40 for completing the survey, US $25 for returning their HIV test kit, and US $25 for returning rectal swabs.

Results

Overview

In total, 71,205 individuals began our screening survey, and 42,132 (59%) completed it. Of the noncompleters, most closed their browser window on the informed consent page (ie, immediately). Of the 42,132 who completed the screening survey, 9115 appeared to be eligible and were sent a baseline computer-assisted self-interview (CASI) survey, 88.3% (n=8049) of whom completed that survey. However, among those completing this baseline CASI (n=8049), 2716 were removed from further consideration (1675/2716, 61.7% of these were explicitly removed because the participant did not verify a working phone number with the study team; a measure in place to prevent fraudulent or duplicate participants).

Of the participants who completed the baseline CASI and verified a phone number, 5364 were eligible to receive and were mailed HIV and rectal test kits. We consider these 5364 individuals to be enrolled in the cohort. Participants resided in all 50 states and Puerto Rico (Figures 1 and 2).

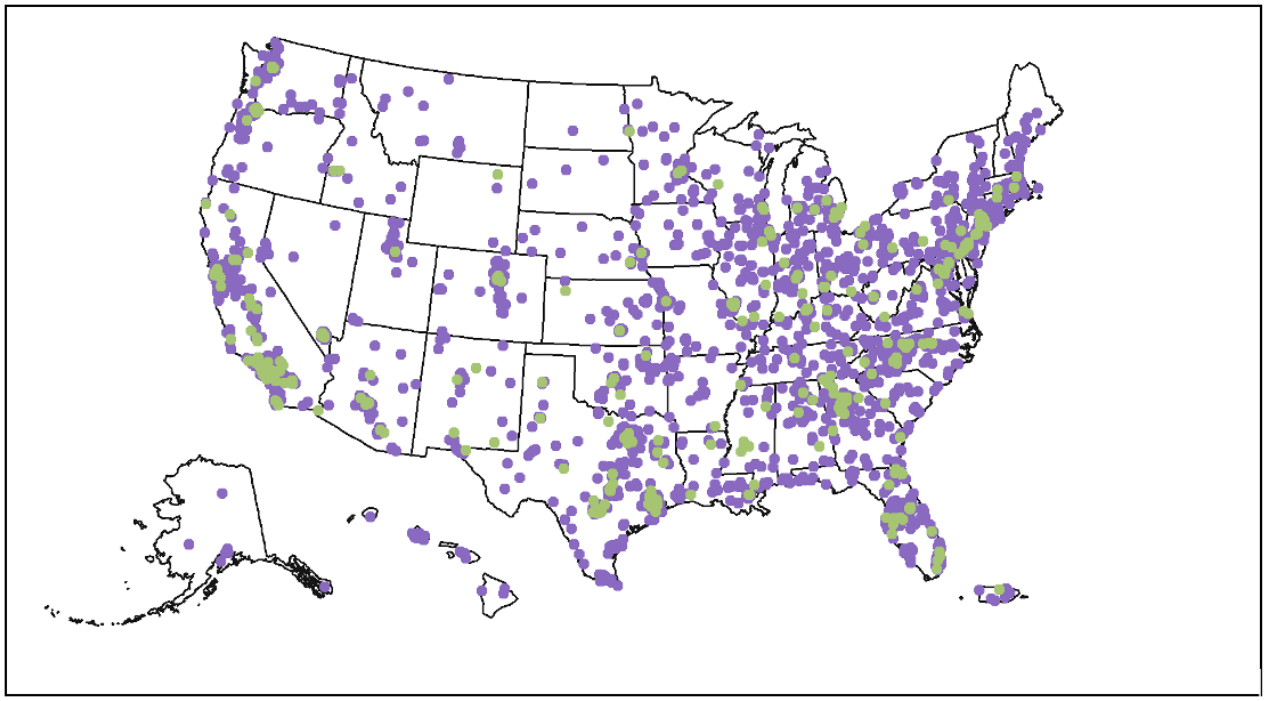
Of those 5364 who were sent a kit, 78.2% returned at least 1 of the specimens (be it an HIV test kit or at least 1 rectal sample). The majority of those sent kits returned their HIV test kit (n=4054, 76%). Throughout the course of enrollment, we had to resend packages to participants a total of 1720 times (this includes 920 participants who had to have a package resent once and 800 participants who had to have a package resent more than once). Of these resent packages, 469 were partial kits (eg, an HIV test kit or a rectal swab). Reasons for resent kits varied and included participants reporting a package not having arrived to them (n=505), packages not received by the laboratory (n=71), some issue with the kit itself requiring the participant to resample (eg, insufficient sample and sample damaged in transit; n=524), the participant misplacing the kit (n=224), the participant moving or asking to send the kit to a different address (n=275), or other factors or a combination of the aforementioned (n=196). Of the participants who returned kits (n=4191), 4085 (97.5%) returned their HIV test kit and at least 1 rectal sample, 73 (1.7%) participants returned just their HIV test kits, and 33

XSL•FO
RenderX

(0.8%) participants returned at least 1 rectal sample but no HIV test kit.

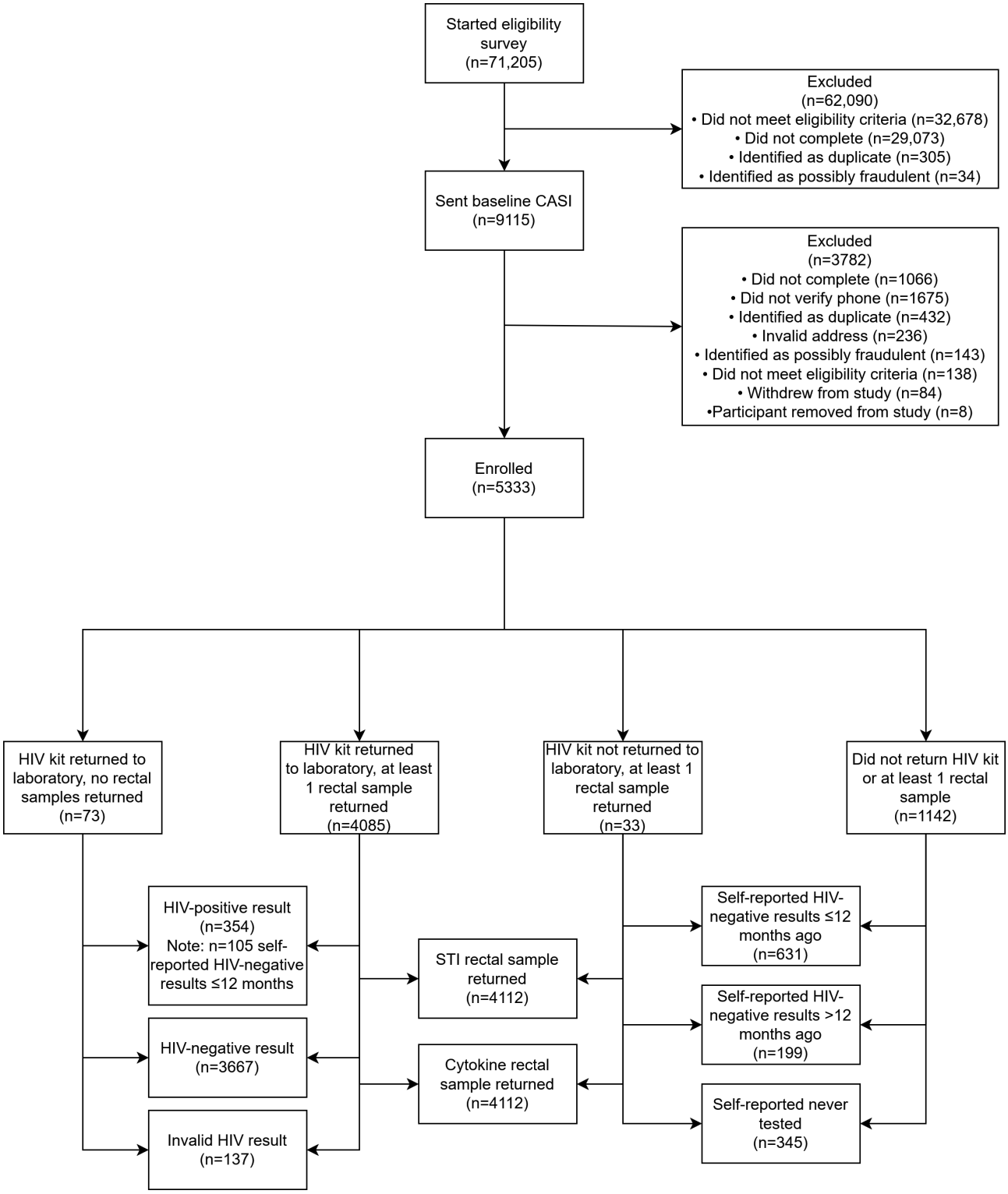
As shown in Table 1, those reporting methamphetamine use were significantly more likely to report nearly all the "additional eligibility" criteria, including 3 or more insertive CAS acts with a man in the last 3 months (67.2% vs 50.7%), a syphilis diagnosis in the last 6 months (5.2% vs 2.2%), and sharing injection drug needles in the last 12 months (8.3% vs 0.1%).

Figure 1. Geographic distribution of participants. Those in green had HIV-positive test results.



XSL•FO
RenderX

Figure 2. Participant flow through study procedures at enrollment. CASI: computer-assisted self-interview; STI: sexually transmitted infection.



Baseline Characteristics of the AMETHST Cohort

Nearly all (4542/5364, 84.7%) were cisgender men, 3.7% (201/5364) were transgender women, and 1.1% (61/5364) were transgender men who have sex with men (Table 2). There were also 560 (10.4%) individuals who self-identified outside of the gender binary—all reported being assigned male sex at birth. In total, 56.5% (3031/5364) were persons of color, and 31.8% (1714/5364) were younger individuals aged 16 to 29 years.

As shown in Table 2, those reporting methamphetamine use were more likely to be White, transgender women or nonbinary, nongay identified, older, and to have never heard of PrEP. They were less likely to have health insurance, have a primary care provider, and were significantly more likely to have tested for HIV more than a year ago or had never been tested for HIV in their lives.

XSL•FO
RenderX

Table . Demographic characteristics for the American Transformative HIV Study cohort study.

	Participants who used methamphetamine in the last 3 months (n=2846), n (%)	Participants who had not used methamphetamine in the last 3 months (n=2518), n (%)	Chi-square (*df*)	*P* value
Race or ethnicity			242.8 (4)	<.001
Asian or Pacific Islander	37 (1.3)	153 (6.1)		
Hispanic or Latinx	649 (22.8)	829 (32.9)		
Non-Hispanic Black	350 (12.3)	416 (16.5)		
Non-Hispanic White	1452 (51)	881 (35)		
Multiracial or other	358 (12.6)	239 (9.5)		
Gender			40.5 (3)	<.001
Cisgender male	2342 (82.3)	2200 (87.4)		
Transgender woman	137 (4.8)	64 (2.5)		
Transgender man	25 (0.9)	36 (1.4)		
Nonbinary or other	342 (12)	218 (8.7)		
Sexual identity			123.8 (2)	<.001
Gay	1760 (61.9)	1912 (75.9)		
Bisexual	664 (23.3)	355 (14.1)		
Other	422 (14.8)	251 (10)		
Age (years)			240.3 (6)	<.001
16 - 19	31 (1.1)	92 (3.7)		
20 - 24	179 (6.3)	369 (14.7)		
25 - 29	467 (16.4)	566 (22.5)		
30 - 34	731 (25.7)	613 (24.3)		
35 - 39	630 (22.1)	438 (17.4)		
40 - 44	505 (17.7)	273 (10.8)		
45 - 49	303 (10.6)	167 (6.6)		
Has health insurance	1947 (68.4)	1989 (79)	76.5 (1)	<.001
Has a primary care provider	1174 (41.3)	1505 (59.8)	183.3 (1)	<.001
Never heard of PrEP[a]	83 (2.9)	41 (1.6)	9.8 (1)	.001
HIV testing history			175.9 (3)	<.001
In the last 6 months	882 (31)	1017 (40.4)		
In the last 7 - 12 months	862 (30.3)	940 (37.3)		
More than a year ago	882 (31)	417 (16.6)		
Never been tested for HIV	220 (7.7)	144 (5.7)		

[a]PrEP: pre-exposure prophylaxis.

Completion of Baseline HIV Testing and Returning Rectal Swabs

All 5364 participants enrolled in the cohort were mailed specimen collection kits (oral fluid for HIV testing and 2 rectal swabs) via United States Postal Service, of which 74.5% (3998/5364) returned both their HIV test kit and at least 1 rectal sample. An additional 3.6% (194/5364) partially returned their samples (either an HIV test kit or at least 1 rectal sample). The remaining did not return their kits, and failure to return kits was higher among those who reported methamphetamine use (28.6% vs 14.3%) despite being sent multiple reminders to do so. Among those having returned an HIV test kit, 3.3% were HIV-positive among those who did not report methamphetamine use, and 15.5% were HIV-positive among those reporting methamphetamine use. Among those with HIV-negative results (n=3683), those reporting methamphetamine use were significantly more likely to report their last HIV test was either more than a year ago (27% vs 15.5%) or having either never tested for HIV or not knowing the results of their last test (6.5%

XSL•FO
RenderX

vs 3.9%). In contrast, among those with HIV-positive results (n=371), there were no significant differences in how long ago they previously tested HIV-negative (Table 3).

Table . Biomarkers and HIV results for the American Transformative HIV Study.

	Participants who used methamphetamine in the last 3 months (n=2846), n (%)	Participants who had not used methamphetamine in the last 3 months (n=2518), n (%)	Chi-square (*df*)	*P* value
Returned sampling kits to the laboratory			185.2 (2)	<.001
Did not return either kit	813 (28.6)	359 (14.3)		
Partial kit return (HIV or rectal)	128 (4.5)	66 (2.6)		
Returned all kits (HIV and at least 1 rectal)	1905 (66.9)	2093 (83.1)		
HIV test kit result (only those with valid HIV test kit results, n=4054)			183.1 (1)	<.001
HIV-negative	1642 (84.5)	2041 (96.7)		
HIV-positive	302 (15.5)	69 (3.3)		
Date of most recent HIV test (only those with valid HIV-negative results, n=3683)			95.4 (3)	<.001
Self-reported HIV-negative test <6 months ago	585 (35.6)	843 (41.3)		
Self-reported HIV-negative test 7 - 12 months ago	507 (30.9)	801 (39.2)		
Self-reported HIV-negative test >12 months ago	443 (27)	317 (15.5)		
No previous HIV test, indeterminate result, or do not know or do not remember	107 (6.5)	80 (3.9)		
Date of most recent HIV test (only those with valid HIV-positive results, n=371)			4.7 (3)	.19
Self-reported HIV-negative test<6 months ago	54 (17.9)	13 (18.8)		
Self-reported HIV-negative test 7 - 12 months ago	86 (28.5)	23 (33.3)		
Self-reported HIV-negative test >12 months ago	137 (45.4)	23 (33.3)		
No previous HIV test, indeterminate result, or do not know or do not remember	25 (8.3)	10 (14.5)		

HIV Status and Estimating the Cohort HIV Incidence Rate Before AMETHST Enrollment

Based on participants' HIV results as well as self-reporting when their most recent HIV-negative test was, we estimated that the incidence rate in this cohort in the 12-month period leading up to study enrollment was 10.06 (95% CI 8.65 - 11.64) per 100 person-years among those reporting methamphetamine use compared with 2.04 (95% CI 1.49 - 2.73) among those not reporting methamphetamine use per 100 person-years. The estimate for the 6-month period leading up to enrollment was 8.85 (95% CI 7.02 - 11.01) per 100 person-years among those reporting methamphetamine use compared with 1.85 (95% CI 1.14 - 2.83) among those not reporting methamphetamine use per 100 person-years (Table 4).

XSL•FO
RenderX

Table . HIV incidence among American Transformative HIV Study participants in the 12- and 6-month periods before study enrollment.

	Participants who used methamphetamine in the last 3 months (n=2820)	Participants who had not used methamphetamine in the last 3 months (n=2584)	Incidence rate ratio
12-month period before enrollment			
Presumed recent seroconversions, n[a]	174	42	—[b]
Person-years at risk among people who seroconverted[c], n	87	21	—
Person-years at risk among persons with HIV-negative test results, n	1642	2041	—
Incidence rate per 100 person-years (95% CI)[d]	10.06 (8.65-11.64)	2.04 (1.49-2.73)	4.93 (3.55-6.99)
6-month period before enrollment			
Presumed recent seroconversions, n[e]	76	19	—
Person-years at risk among people who seroconverted[c], n	38	9.5	—
Person-years at risk among persons with HIV-negative test results, n	821	1020.5	—
Incidence rate per 100 person-years (95% CI)[d]	8.85 (7.02-11.01)	1.85 (1.14-2.83)	4.78 (2.94-8.12)

[a]To accurately approximate the number of recent seroconversions for the 12-month incidence calculation, we first totaled the number of HIV-positive results for testing categories ≤6 months and 7 - 12 months (Table 3). Then, we added 25% of the adjacent category of those who reported last testing >12 months ago.

[b]Not applicable.

[c]Person-years calculated as (1) mid-year estimates for those presumed to have seroconverted in the last year (12-month incidence rate), (2) full-year estimates for people who did not seroconvert (12-month incidence rate), (3) quarter-year estimates for those presumed to have seroconverted in the last 6 months (6-month incidence rate), and (4) half year estimates for those who did not seroconvert (6-month incidence rate).

[d]Mid-P exact test 95% CIs calculated using OpenEpi software.

[e]To accurately approximate the number of recent seroconversions for the 6-month incidence calculation, we used the number of HIV-positive results for the ≤6 months category. Then, we added 25% of those who reported the last testing 7-12 months ago.

Longitudinal Follow-Up and Measurement

Prospective closed follow-up of AMETHST participants includes completion of web-based surveys beginning 12 months after the baseline survey, again at 24 months, and finally at 36 months, as well as annual self-administered at-home HIV test and rectal swabs. However, at follow-up, participants who self-report being on PrEP (ie, PrEP adoption outside of the study) are asked to provide proof in the form of a photo of the medication bottle with their prescription information or other medical documentation of their prescription. Likewise, those who report having been diagnosed with HIV between study assessments are asked to provide proof in the form of a picture of a document or medication bottle with their prescription information. Those known to be HIV-positive (either via testing HIV-positive during a prior wave of assessments or documented proof of HIV seroconversion between study assessments) will not be asked to test for HIV within the study again. Our reason for not asking those on PrEP to test with us is because our test is a third-generation HIV test, and those on PrEP should be receiving a fourth-generation HIV test (blood) from their medical provider, and we do not want participants to inadvertently use our test in lieu of their providers. Should participants report discontinuing PrEP at a follow-up assessment, they will be asked to test with us again.

A comprehensive list of key study measurements by study wave is included in Multimedia Appendix 3. We will follow the cohort of 5364 individuals to characterize methamphetamine use or uptake or discontinuation and its role in the following: the rate of PrEP uptake or discontinuation, individual or network or contextual-level determinants of PrEP uptake or discontinuation, patterns of PrEP use (eg, daily or on-demand), the rate of new HIV seroconversions and other missed HIV prevention opportunities (ie, sexually transmitted infections while not on PrEP), individual or network or contextual-level determinants of HIV seroconversion and missed HIV prevention opportunities, racial or ethnic disparities in HIV incidence and their trends over time, and the influence of PrEP uptake on racial or ethnic disparities in HIV incidence.

Given the nature of this web-based US nationwide cohort, we elected to longitudinally follow all participants who mailed an

XSL•FO
RenderX

HIV test kit at baseline rather than only those who completed HIV testing. Participants who did not complete HIV testing will have the opportunity to provide those data or samples at future assessments, and this will enable us to learn more about differential rates of participation and attrition in the cohort moving forward.

Biological Vulnerability to HIV—Nested Case-Cohort Study

At the completion of all follow-up assessments, we will conduct a nested case-cohort study [68] to examine biological vulnerability to HIV. We anticipate that upward of 450 individuals will have seroconverted to HIV-positive, and we will match their samples against 450 HIV-negative participants (similar in age and race). The rectal swabs for these 900 participants will be retrieved from the biospecimen repository. For those who are HIV seroconvert, swabs preceding seroconversion will be selected. One swab (the Aptima Multitest Swab) will be tested for rectal gonorrhea and chlamydia. The second swab (the Zymo DNA/RNA Shield swab) will be examined using the LEGENDPlex assays to quantify rectal cytokine or chemokine levels. We will use the Human Inflammation Panel, which allows for the measurement of 13 cytokines or chemokines [70]. This will allow to determine changes in cytokine or chemokine profiles before and after HIV seroconversion—matched against those of HIV-negative participants—and to determine if methamphetamine use is associated with any immune changes in the rectal mucosa among those who acquire HIV as well as in the comparison group.

Discussion

Principal Findings

The AMETHST cohort study enrolled an entirely web-based US national cohort of cisgender men, transgender men, and transgender men who have sex with men and are at very high risk for HIV, into longitudinal follow-up. Greater than half of the participants were persons of color, and 53.1% had recently used methamphetamine. Importantly, none of the AMETHST participants were on PrEP at enrollment, but all met the clinical criteria for PrEP use. This will allow our study to fill critical knowledge gaps that can improve our understanding of the role that methamphetamine plays in (barriers to) PrEP uptake or discontinuation among those most in need of HIV prevention interventions—and at a time in which PrEP has been Food and Drug Administration–approved for more than a decade.

Although all participants were similar in their having met clinical criteria for PrEP, we observed several characteristics that distinguished people who used methamphetamine at baseline from those who did not in ways that put them at greater risk for HIV. People who used methamphetamine were more likely to report greater than 3 recent insertive CAS acts with men, to have been diagnosed recently with syphilis, to report sharing injection drug needles in the past year, trended toward being older, having never tested for HIV, and to not have health insurance, nor a primary care provider, or have heard of PrEP. These facets all compound to elevate HIV risk. This is perhaps evidenced best through the staggering number of HIV infections (baseline prevalence) identified among those who use methamphetamine (n=302, 15.5% vs n=69, 3.3% among those not using). This high number was despite lower HIV kit return among those who use methamphetamine. Our baseline data highlight the challenges people who use methamphetamine face as well as the challenges we as researchers may have in engaging and retaining this population longitudinally. Our estimates of HIV incidence in the 12-month period leading up to enrollment (2.04% per year among those who do not use methamphetamine and 10.06% per year among those who use methamphetamine) suggest that, without significant PrEP uptake as the cohort moves forward, the risk of HIV acquisition that we will observe prospectively will also be quite high. As noted, per enrollment criteria, no participants were taking PrEP at the time of enrollment; however, we provide participants information and resources on how to access PrEP and will be monitoring that outcome going forward.

Strengths

Major strengths of the AMETHST cohort include its large sample size, geographic diversity, self-reported data related to substance use, HIV risk and PrEP, biomarkers for direct measurement of HIV status and seroconversion and inflammation, recruitment of participants independent of their access to or engagement with the health care system, and inclusion of large numbers of racial or ethnic minority SGM individuals as well as those younger than the age of 29 years. Having overrepresented persons who use methamphetamine uniquely positions this cohort to answer key causal questions about the role that methamphetamine plays in disrupting HIV prevention. As well, our future nested case-cohort study will enable our team to answer key questions regarding the biological role that methamphetamine plays in HIV seroconversion.

Importantly, the AMETHST cohort study design, with its infrequent (ie, annual) surveys, reduces the potential for participation and questionnaire-response bias introduced by the Hawthorne effect. Studies involving frequent contact can cause participants to adopt behaviors that make them less representative of the vulnerable populations from which they were drawn. For example, McCambridge et al [71] discussed how research participant effects—an expansion on the Hawthorne effect [72]—can introduce bias, including demand characteristics [73]. Studies involving high levels of staff contact with participants may induce behavior change by repeatedly engaging participants outside of their natural context, artificially impacting results and reducing generalizability [74,75]. That being said, we recognize that any self-reporting of a behavior introduces some potential for a participant to reflect upon, and thus modify, their behavior.

Limitations

Recruiting participants and gathering data on the web are a less controlled research environment than face-to-face studies. Web-based recruitment increases vulnerability to repeat participation and fraudulent manipulation of HIV testing procedures. We took several measures to avoid this, including requiring all participants to have a unique email address, a unique and valid phone number (with text verification), and a valid mailing address (verified with the US postal service). We

XSL•FO
RenderX

tracked IP addresses and blocked responses from outside the United States on our screening survey. Our advertising materials did not indicate full eligibility criteria (eg, age 16 to 49 years, cannot be HIV-positive, and meeting criteria for PrEP care but not being on PrEP), and the screening survey included additional items not related to study eligibility such as cannabis use, mpox infection or vaccination, health insurance status, and doxycycline postexposure prophylaxis. However, no method is entirely infallible. At future waves of follow-up, data will be compared to responses at baseline for consistency (eg, age and race).

Next, participants were recruited via sexual networking apps and are thus not representative of all SGM individuals highly vulnerable to HIV. These apps have emerged as the modal means through which SGM individuals meet sex partners and have become an essential tool for researchers and providers to reach this population.

Finally, HIV seroconversions and PrEP uptake may be incompletely ascertained, and the timing of HIV seroconversions must be estimated using a midpoint approach with broad intervals, potentially introducing bias in our baseline HIV incidence estimates. For example, our study eligibility or inclusion criterion of never diagnosed with HIV (self-report) effectively excludes frequent HIV testers who seroconverted and were diagnosed in the period immediately before the study launch.

Conclusions

The AMETHST cohort is a large racially and geographically diverse group of SGM individuals who report high vulnerability to HIV. Its focus to identify multilevel (ie, structural, psychological, and social) and biobehavioral (ie, rectal cytokines or chemokines) determinants of amplified HIV seroconversion risk for SGM individuals, including those who use methamphetamine, will serve to guide targeted, combination interventions to reduce HIV incidence and increase PrEP uptake. Its added nested case-cohort study [68] will determine changes in cytokine or chemokine profiles before and after HIV seroconversion as well as the role that methamphetamine plays in that process. Our team welcomes new collaborations both in secondary analysis of the data generated as well as proposals for new measures or data to be gathered from participants in future waves of follow-up. Instructions and a concept proposal form are available by emailing the principal investigators (CG and AWC). Submitted concept proposals will be reviewed by the principal investigators (CG and AWC) and a core group of AMETHST investigators with rapid turnaround.

Acknowledgments

American Transformative HIV Study was funded by the National Institutes of Health (NIH; UG3 AI169652—CG or AWC). Other forms of support include the City University of New York (CUNY) Institute for Implementation Science in Population Health, the Einstein, Rockefeller, and CUNY Center for AIDS Research (P30 AI124414). DW was supported in part by K01 AA029047. JM is supported in part by the Tulane National Primate Research Center NIH Core Grant (P51OD011104). The authors especially thank Matthew Stief as well as NIH program staff: Gerald Sharp, Lori Zimand, Richard Jenkins, Michael Stirratt, and Sonia Lee. Although NIH funded this research, they do not necessarily endorse these findings.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample recruitment advertisements.
[DOCX File, 1887 KB - publichealth_v11i1e66921_app1.docx]

Multimedia Appendix 2
Baseline survey.
[DOCX File, 150 KB - publichealth_v11i1e66921_app2.docx]

Multimedia Appendix 3
Screening survey.
[DOCX File, 47 KB - publichealth_v11i1e66921_app3.docx]

References

1. HIV in the United States: at a glance. Centers for Disease Control and Prevention. 2015. URL: http://www.cdc.gov/hiv/statistics/basics/ataglance.html [accessed 2025-07-22]
2. HIV among gay and bisexual men. Centers for Disease Control and Prevention. 2016. URL: http://www.cdc.gov/hiv/group/msm [accessed 2016-08-22]
3. Diagnoses of HIV infection among adolescents and young adults in the United States and 6 dependant areas 2010-2014. HIV surveillance supplemental report. : Centers for Disease Control and Prevention; 2016.

XSL•FO
RenderX

4. HIV and gay and bisexual men. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/hiv/group/msm/index.html [accessed 2018-10-17]
5. Half of Black gay men and a quarter of Latino gay men projected to be diagnosed within their lifetime. Centers for Disease Control and Prevention. 2016. URL: http://www.cdc.gov/nchhstp/newsroom/2016/croi-press-release-risk.html [accessed 2016-04-26]
6. What percentage of the US population is transgender? USA Facts. 2023. URL: https://usafacts.org/articles/what-percentage-of-the-us-population-is-transgender [accessed 2025-04-29]
7. How many adults and youth identify as transgender in the United States? The Williams Institute. 2022. URL: https://williamsinstitute.law.ucla.edu/publications/trans-adults-united-states [accessed 2025-04-29]
8. HIV surveillance report 2019. Centers for Disease Control and Prevention. 2021. URL: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html [accessed 2022-08-04]
9. HIV infection, risk, prevention, and testing behaviors among transgender women. : Centers for Disease Control and Prevention; 2021.
10. Bush S, Magnuson D, Rawlings MK, Hawkins T, McCallister S, Mera Giler R. Racial characteristics of FTC/TDF for pre-exposure prophylaxis (PrEP) users in the US #2651. Presented at: American Society for Microbiology: ICAAC; Jun 16-20, 2016; Boston, MA, United States.
11. Carrico AW, Zepf R, Meanley S, Batchelder A, Stall R. Critical review: when the party is over: a systematic review of behavioral interventions for substance-using men who have sex with men. J Acquir Immune Defic Syndr 2016 Nov 1;73(3):299-306. [doi: 10.1097/QAI.0000000000001102] [Medline: 27258233]
12. Thiede H, Valleroy LA, MacKellar DA, et al. Regional patterns and correlates of substance use among young men who have sex with men in 7 US urban areas. Am J Public Health 2003 Nov;93(11):1915-1921. [doi: 10.2105/ajph.93.11.1915] [Medline: 14600066]
13. Das M, Santos D, Matheson T, et al. Feasibility and acceptability of a phase II randomized pharmacologic intervention for methamphetamine dependence in high-risk men who have sex with men. AIDS 2010 Apr 24;24(7):991-1000. [doi: 10.1097/qad.0b013e328336e98b] [Medline: 20397286]
14. Colfax G, Shoptaw S. The methamphetamine epidemic: implications for HIV prevention and treatment. Curr HIV/AIDS Rep 2005 Nov;2(4):194-199. [doi: 10.1007/s11904-005-0016-4] [Medline: 16343378]
15. Compton WM, Jones CM. Substance use among men who have sex with men. N Engl J Med 2021 Jul 22;385(4):352-356. [doi: 10.1056/NEJMra2033007] [Medline: 34289278]
16. Pantalone DW, Bimbi DS, Holder CA, Golub SA, Parsons JT. Consistency and change in club drug use by sexual minority men in New York City, 2002 to 2007. Am J Public Health 2010 Oct;100(10):1892-1895. [doi: 10.2105/AJPH.2009.175232] [Medline: 20724693]
17. Grov C, Parsons JT, Bimbi DS, Sex and Love v3.0 Research Team. In the shadows of a prevention campaign: sexual risk behavior in the absence of crystal methamphetamine. AIDS Educ Prev 2008 Feb;20(1):42-55. [doi: 10.1521/aeap.2008.20.1.42] [Medline: 18312066]
18. Reback CJ, Shoptaw S, Grella CE. Methamphetamine use trends among street-recruited gay and bisexual males, from 1999 to 2007. J Urban Health 2008 Nov;85(6):874-879. [doi: 10.1007/s11524-008-9326-3] [Medline: 18843536]
19. Reback CJ, Fletcher JB, Shoptaw S, Grella CE. Methamphetamine and other substance use trends among street-recruited men who have sex with men, from 2008 to 2011. Drug Alcohol Depend 2013 Nov 1;133(1):262-265. [doi: 10.1016/j.drugalcdep.2013.06.007] [Medline: 23890490]
20. Kuo I, Patrick R, Opoku J, Rawls A, Magnus M, Peterson J, editors. Changing pattern of crystal meth use in Black & White MSM, Washington, DC, 2008–2014. Presented at: Conference on Retroviruses and Opportunistic Infections; Feb 13-16, 2017; Seattle, WA, United States.
21. Hood JE, Buskin SE, Golden MR, Glick SN, Banta-Green C, Dombrowski JC. The changing burden of HIV attributable to methamphetamine among men who have sex with men in King County, Washington. AIDS Patient Care STDS 2018 Jun;32(6):223-233. [doi: 10.1089/apc.2017.0306] [Medline: 29851502]
22. HIV infection risk, prevention, and testing behaviors among men who have sex with men—national HIV behavioral surveillance, 20 U.S. cities 2014. : Centers for Disease Control and Prevention; 2016.
23. Hoenigl M, Chaillon A, Moore DJ, Morris SR, Smith DM, Little SJ. Clear links between starting methamphetamine and increasing sexual risk behavior: a cohort study among men who have sex with men. J Acquir Immune Defic Syndr 2016 Apr 15;71(5):551-557. [doi: 10.1097/QAI.0000000000000888] [Medline: 26536321]
24. Rivera AV, Harriman G, Carrillo SA, Braunstein SL. Trends in methamphetamine use among men who have sex with men in New York City, 2004-2017. AIDS Behav 2021 Apr;25(4):1210-1218. [doi: 10.1007/s10461-020-03097-5] [Medline: 33185774]
25. Grov C, Westmoreland D, Morrison C, Carrico AW, Nash D. The crisis we are not talking about: one-in-three annual HIV seroconversions among sexual and gender minorities were persistent methamphetamine users. J Acquir Immune Defic Syndr 2020 Nov 1;85(3):272-279. [doi: 10.1097/QAI.0000000000002461] [Medline: 32740370]
26. Halkitis PN, Parsons JT, Stirratt MJ. A double epidemic: crystal methamphetamine drug use in relation to HIV transmission among gay men. J Homosex 2001;41(2):17-35. [doi: 10.1300/J082v41n02_02] [Medline: 11482426]

XSL•FO
RenderX

27. Isaiah Green A, Halkitis PN. Crystal methamphetamine and sexual sociality in an urban gay subculture: an elective affinity. Cult Health Sex 2006 Jul 1;8(4):317-333. [doi: 10.1080/13691050600783320]
28. Baskin-Sommers A, Sommers I. The co-occurrence of substance use and high-risk behaviors. J Adolesc Health 2006 May;38(5):609-611. [doi: 10.1016/j.jadohealth.2005.07.010] [Medline: 16635777]
29. Semple SJ, Zians J, Strathdee SA, Patterson TL. Sexual marathons and methamphetamine use among HIV-positive men who have sex with men. Arch Sex Behav 2009 Aug;38(4):583-590. [doi: 10.1007/s10508-007-9292-y] [Medline: 18185990]
30. Semple SJ, Patterson TL, Grant I. Binge use of methamphetamine among HIV-positive men who have sex with men: pilot data and HIV prevention implications. AIDS Educ Prev 2003 Apr;15(2):133-147. [doi: 10.1521/aeap.15.3.133.23835] [Medline: 12739790]
31. Halkitis PN, Fischgrund BN, Parsons JT. Explanations for methamphetamine use among gay and bisexual men in New York City. Subst Use Misuse 2005;40(9-10):1331-1345. [doi: 10.1081/JA-200066900] [Medline: 16048820]
32. Braine N, Acker CJ, van Sluytman L, Friedman S, Des Jarlais DC. Drug use, community action, and public health: gay men and crystal meth in NYC. Subst Use Misuse 2011;46(4):368-380. [doi: 10.3109/10826081003720899] [Medline: 20394522]
33. Treloar C, Hopwood M, Drysdale K, et al. Stigma as understood by key informants: a social ecological approach to gay and bisexual men's use of crystal methamphetamine for sex. Int J Drug Policy 2021 Aug;94:103229. [doi: 10.1016/j.drugpo.2021.103229] [Medline: 33774423]
34. Aguilar JP, Sen S. The culture of methamphetamine: reframing gay men's methamphetamine use. J Hum Behav Soc Environ 2013 Apr;23(3):370-382. [doi: 10.1080/10911359.2013.764204]
35. Bostwick WB, Boyd CJ, Hughes TL, West BT, McCabe SE. Discrimination and mental health among lesbian, gay, and bisexual adults in the United States. Am J Orthopsychiatry 2014;84(1):35-45. [doi: 10.1037/h0098851] [Medline: 24826824]
36. Balaji AB, Bowles KE, Hess KL, Smith JC, Paz-Bailey G, Taussig J. Association between enacted stigma and HIV-related risk behavior among MSM, National HIV Behavioral Surveillance System, 2011. AIDS Behav 2017 Jan;21(1):227-237. [doi: 10.1007/s10461-016-1599-z] [Medline: 27830344]
37. Lee JH, Gamarel KE, Bryant KJ, Zaller ND, Operario D. Discrimination, mental health, and substance use disorders among sexual minority populations. LGBT Health 2016 Aug;3(4):258-265. [doi: 10.1089/lgbt.2015.0135] [Medline: 27383512]
38. Mays VM, Cochran SD. Mental health correlates of perceived discrimination among lesbian, gay, and bisexual adults in the United States. Am J Public Health 2001 Nov;91(11):1869-1876. [doi: 10.2105/ajph.91.11.1869] [Medline: 11684618]
39. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychol Bull 2003 Sep;129(5):674-697. [doi: 10.1037/0033-2909.129.5.674] [Medline: 12956539]
40. Meyer IH, Schwartz S, Frost DM. Social patterning of stress and coping: does disadvantaged social statuses confer more stress and fewer coping resources? Soc Sci Med 2008 Aug;67(3):368-379. [doi: 10.1016/j.socscimed.2008.03.012] [Medline: 18433961]
41. Meyer IH. Minority stress and mental health in gay men. J Health Soc Behav 1995 Mar;36(1):38-56. [Medline: 7738327]
42. HIV and gay and bisexual men. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/hiv/group/msm/index.html [accessed 2021-07-14]
43. Parsons JT, Millar BM, Moody RL, Starks TJ, Rendina HJ, Grov C. Syndemic conditions and HIV transmission risk behavior among HIV-negative gay and bisexual men in a U.S. national sample. Health Psychol 2017 Jul;36(7):695-703. [doi: 10.1037/hea0000509] [Medline: 28541070]
44. Stall R, Mills TC, Williamson J, et al. Association of co-occurring psychosocial health problems and increased vulnerability to HIV/AIDS among urban men who have sex with men. Am J Public Health 2003 Jun;93(6):939-942. [doi: 10.2105/ajph.93.6.939] [Medline: 12773359]
45. O'Leary A, Jemmott JB 3rd, Stevens R, Rutledge SE, Icard LD. Optimism and education buffer the effects of syndemic conditions on HIV status among African American men who have sex with men. AIDS Behav 2014 Nov;18(11):2080-2088. [doi: 10.1007/s10461-014-0708-0] [Medline: 24705710]
46. Brennan J, Kuhns LM, Johnson AK, et al. Syndemic theory and HIV-related risk among young transgender women: the role of multiple, co-occurring health problems and social marginalization. Am J Public Health 2012 Sep;102(9):1751-1757. [doi: 10.2105/AJPH.2011.300433] [Medline: 22873480]
47. van den Berg JJ, Isabel Fernández M, Fava JL, et al. Using syndemics theory to investigate risk and protective factors associated with condomless sex among youth living with HIV in 17 U.S. cities. AIDS Behav 2017 Mar;21(3):833-844. [doi: 10.1007/s10461-016-1550-3] [Medline: 27624727]
48. Wim VB, Christiana N, Marie L. Syndemic and other risk factors for unprotected anal intercourse among an online sample of Belgian HIV negative men who have sex with men. AIDS Behav 2014 Jan;18(1):50-58. [doi: 10.1007/s10461-013-0516-y] [Medline: 23681697]
49. Singer M. AIDS and the health crisis of the U.S. urban poor; the perspective of critical medical anthropology. Soc Sci Med 1994 Oct;39(7):931-948. [doi: 10.1016/0277-9536(94)90205-4] [Medline: 7992126]
50. Halkitis PN, Moeller RW, Siconolfi DE, Storholm ED, Solomon TM, Bub KL. Measurement model exploring a syndemic in emerging adult gay and bisexual men. AIDS Behav 2013 Feb;17(2):662-673. [doi: 10.1007/s10461-012-0273-3] [Medline: 22843250]

XSL•FO
RenderX

51. Wilson PA, Nanin J, Amesty S, Wallace S, Cherenack EM, Fullilove R. Using syndemic theory to understand vulnerability to HIV infection among Black and Latino men in New York City. J Urban Health 2014 Oct;91(5):983-998. [doi: 10.1007/s11524-014-9895-2] [Medline: 25155096]
52. Kelley CF, Kraft CS, de Man TJ, et al. The rectal mucosa and condomless receptive anal intercourse in HIV-negative MSM: implications for HIV transmission and prevention. Mucosal Immunol 2017 Jul;10(4):996-1007. [doi: 10.1038/mi.2016.97] [Medline: 27848950]
53. Hladik F, McElrath MJ. Setting the stage: host invasion by HIV. Nat Rev Immunol 2008 Jun;8(6):447-457. [doi: 10.1038/nri2302] [Medline: 18469831]
54. Sullivan PS, Salazar L, Buchbinder S, Sanchez TH. Estimating the proportion of HIV transmissions from main sex partners among men who have sex with men in five US cities. AIDS 2009 Jun 1;23(9):1153-1162. [doi: 10.1097/QAD.0b013e32832baa34] [Medline: 19417579]
55. Gonzalez SM, Aguilar-Jimenez W, Su RC, Rugeles MT. Mucosa: key interactions determining sexual transmission of the HIV infection. Front Immunol 2019;10(144):144. [doi: 10.3389/fimmu.2019.00144] [Medline: 30787929]
56. Henry BL, Minassian A, Perry W. Effect of methamphetamine dependence on heart rate variability. Addict Biol 2012 May;17(3):648-658. [doi: 10.1111/j.1369-1600.2010.00270.x] [Medline: 21182570]
57. Cook RR, Fulcher JA, Tobin NH, et al. Alterations to the gastrointestinal microbiome associated with methamphetamine use among young men who have sex with men. Sci Rep 2019 Oct 16;9(1):14840. [doi: 10.1038/s41598-019-51142-8] [Medline: 31619731]
58. Persons AL, Bradaric BD, Dodiya HB, et al. Colon dysregulation in methamphetamine self-administering HIV-1 transgenic rats. PLoS ONE 2018;13(1):e0190078. [doi: 10.1371/journal.pone.0190078] [Medline: 29293553]
59. Nash D, Stief M, MacCrate C, et al. A web-based study of HIV prevention in the era of pre-exposure prophylaxis among vulnerable HIV-negative gay and bisexual men, transmen, and transwomen who have sex with men: protocol for an observational cohort study. JMIR Res Protoc 2019 Sep 17;8(9):e13715. [doi: 10.2196/13715] [Medline: 31538945]
60. Grov C, Westmoreland DA, Carneiro PB, et al. Recruiting vulnerable populations to participate in HIV prevention research: findings from the Together 5000 cohort study. Ann Epidemiol 2019 Jul;35:4-11. [doi: 10.1016/j.annepidem.2019.05.003] [Medline: 31182379]
61. Grov C, Bimbi DS, Nanin JE, Parsons JT. Race, ethnicity, gender, and generational factors associated with the coming-out process among lesbian, and bisexual individuals. J Sex Res 2006 May;43(2):115-121. [doi: 10.1080/00224490609552306] [Medline: 16817058]
62. Mustanski B. Ethical and regulatory issues with conducting sexuality research with LGBT adolescents: a call to action for a scientifically informed approach. Arch Sex Behav 2011 Aug;40(4):673-686. [doi: 10.1007/s10508-011-9745-1] [Medline: 21528402]
63. Lang N. There's a meth problem on Grindr and Scruff—and no one is doing much about it. MIC. 2016. URL: https://www.mic.com/articles/138644/there-s-a-meth-problem-on-grindr-and-no-one-is-doing-much-about-it [accessed 2019-10-01]
64. Ziegler C. HosTing Now: methamphetamine crisis in the gay community. KTSW 899. URL: https://ktswblog.net/2019/04/25/hosting-now-methamphetamine-crisis-in-the-gay-community [accessed 2025-04-29]
65. Grov C, Zohra F, Mirzayi C, et al. Sexual and gender minorities' vaccine uptake and behavioral change in response to the mpox outbreak in the United States: August 2022 through November 2022. Clin Infect Dis :ciad793. [doi: 10.1093/cid/ciad793]
66. Elma M, Krusel K, Crumpler M, et al. Use of the Bio-Rad Geenius HIV-1/2 supplemental assay for the testing of oral fluids for the presence of HIV antibody. J Clin Virol 2020 Jul;128:104422. [doi: 10.1016/j.jcv.2020.104422] [Medline: 32464307]
67. D'Angelo AB, Morrison CA, Lopez-Rios J, et al. Experiences receiving HIV-positive results by phone: acceptability and implications for clinical and behavioral research. AIDS Behav 2021 Mar;25(3):709-720. [doi: 10.1007/s10461-020-03027-5] [Medline: 32915328]
68. Katz DA, Dombrowski JC, Bell TR, Kerani RP, Golden MR. HIV incidence among men who have sex with men after diagnosis with sexually transmitted infections. Sex Transm Dis 2016 Apr;43(4):249-254. [doi: 10.1097/OLQ.0000000000000423] [Medline: 26967302]
69. Grov C, Guo Y, Westmoreland DA, et al. Factors associated with PrEP-era HIV seroconversion in a 4-year U.S. national cohort of n = 6059 sexual and gender minority individuals who have sex with men, 2017-2022. J Int AIDS Soc 2024 Jun;27(6):e26312. [doi: 10.1002/jia2.26312] [Medline: 38924359]
70. McManus K, McGaugh A, Miller C, Manuzak J, Broedlow C. Bio-behavioral correlates of rectal inflammation in HIV-negative sexual minority men. Presented at: Second Annual Mucosal Systems Meeting; May 21-22, 2019; Chicago, IL, United States.
71. McCambridge J, Kypri K, Elbourne D. Research participation effects: a skeleton in the methodological cupboard. J Clin Epidemiol 2014 Aug;67(8):845-849. [doi: 10.1016/j.jclinepi.2014.03.002] [Medline: 24766858]
72. Adair JG. The Hawthorne effect: a reconsideration of the methodological artifact. J Appl Psychol 1984;69(2):334-345. [doi: 10.1037/0021-9010.69.2.334]

XSL•FO
RenderX

73. McCambridge J, de Bruin M, Witton J. The effects of demand characteristics on research participant behaviours in non-laboratory settings: a systematic review. PLoS ONE 2012;7(6):e39116. [doi: 10.1371/journal.pone.0039116] [Medline: 22723942]
74. Mbulaiteye SM, Mahe C, Ruberantwari A, Whitworth JAG. Generalizability of population-based studies on AIDS: a comparison of newly and continuously surveyed villages in rural southwest Uganda. Int J Epidemiol 2002 Oct;31(5):961-967. [doi: 10.1093/ije/31.5.961] [Medline: 12435768]
75. Bouchet C, Guillemin F, Briançon S. Nonspecific effects in longitudinal studies: impact on quality of life measures. J Clin Epidemiol 1996 Jan;49(1):15-20. [doi: 10.1016/0895-4356(95)00540-4] [Medline: 8598506]

Abbreviations

AMETHST: American Transformative HIV Study
CAS: condomless anal sex
CASI: computer-assisted self-interview
PrEP: pre-exposure prophylaxis
SGM: sexual and gender minority

Edited by DI Yaras; submitted 26.09.24; peer-reviewed by J Jones, R Schnall; revised version received 19.03.25; accepted 02.04.25; published 22.05.25.

Please cite as:
Grov C, D'Angelo AB, Mirzayi C, Dearolf M, Hoeppner E, Guo Y, Richards N, Ehsan R, Kulkarni S, Nash D, Patel VV, Duncan DT, Ray M, Bartholomew T, Manuzak J, Manuel J, McCollister K, Westmoreland D, Carrico AW
The American Transformative HIV Study: Protocol for a US National Cohort of Sexual and Gender Minority Individuals With HIV
JMIR Public Health Surveill 2025;11:e66921
URL: https://publichealth.jmir.org/2025/1/e66921
doi:10.2196/66921

© Christian Grov, Alexa B D'Angelo, Chloe Mirzayi, Michelle Dearolf, Elena Hoeppner, Yan Guo, Nicole Richards, Rifa Ehsan, Sarah Kulkarni, Denis Nash, Viraj V Patel, Dustin T Duncan, Meredith Ray, Tyler Bartholomew, Jennifer Manuzak, Jennifer Manuel, Kathryn McCollister, Drew Westmoreland, Adam W Carrico. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 22.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Validation of the Perceived Barriers to Antiretroviral Therapy Adherence (PEDIA) Scale Among Gay, Bisexual, and Other Men Who Have Sex With Men and Transgender and Nonbinary Persons: Cross-Sectional Study

Victor C Matos[1], MSc; Dyego Carlos Araújo[2], PhD; Thiago S Torres[3], PhD; Paula M Luz[3], PhD

[1]Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, Rua Leopoldo Bulhões, 1480, Rio de Janeiro, Brazil

[2]Departamento de Ciências Farmacêuticas, Universidade Federal do Espírito Santo, Vitória, Brazil

[3]Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Corresponding Author:
Victor C Matos, MSc
Escola Nacional de Saúde Pública Sergio Arouca, Fundação Oswaldo Cruz, Rua Leopoldo Bulhões, 1480, Rio de Janeiro, Brazil

Abstract

Background: Approximately 30% of people living with HIV in Brazil have suboptimal adherence to antiretroviral therapy (ART). Theoretical models of health behaviors and objective evidence support the role of perceived barriers in hindering health behaviors.

Objective: We aimed to evaluate the psychometric properties of the perceived barriers to ART adherence (PEDIA) scale among gay, bisexual, and other men who have sex with men (MSM) and transgender and nonbinary (TGNB) persons in Brazil.

Methods: Using a cross-sectional study design, we evaluated the factor structure, construct validity, convergent and discriminant factorial validity, and internal reliability of the PEDIA scale using 2 convenience samples of participants aged 18 years or older living in Brazil. Sample 2020 was recruited through an internet-based survey between February and March 2020 and Sample 2021 between May 2021 and January 2022. The study instrument included sociodemographic information, ART use, and the 3 measures of ART adherence. Using Sample 2020, we performed exploratory factor analysis (EFA) with parallel analysis to evaluate PEDIA's factor structure. Based on the goodness of fit measures and theoretical relevance, we developed a reduced version of PEDIA (PEDIAr). Using Sample 2021, we performed a psychometric evaluation of PEDIAr, which included confirmatory factor analysis (CFA), examination of modification indexes and item-item and item-total correlations, and an assessment of its correlation with different measures of adherence to ART. Goodness of fit was evaluated based on multiple indices.

Results: EFA conducted on Sample 2020 (n=1692) revealed a 2-factor structure with 3 factor loadings <0.4 (excluded). Using Sample 2021 (n=4893), modification indices from the CFA and item-item and item-total correlations along with item relevance analysis suggested the exclusion of 5 additional items. CFA fit indices for PEDIAr were adequate (root mean square error of approximation=0.07, comparative fit index=0.95, Tucker-Lewis Index=0.94, standardized root mean square residual=0.05). Construct validity was supported by factor loadings above 0.6 and negative correlations between PEDIAr scores and 3 measures of adherence to ART. McDonald omega was 0.795 and 0.859 for factors 1 and 2, respectively. The square root of the average variance extracted (AVE) was 0.704 and 0.711 for factors 1 and 2, respectively, and the difference between AVE and the square of the factor correlations with other items was small (0.001 and 0.009, respectively) and not statistically significant for both factors (*P*=.94 and *P*=.55, respectively).

Conclusions: PEDIAr, the 10-item reduced version of PEDIA, proved to be valid among gay, bisexual, and other MSM and TGNB persons in Brazil. This shorter instrument was able to capture 2 distinguished dimensions of the perceived barriers to adherence to ART (practical aspects and psychological aspects). By proactively identifying individuals struggling with adherence to ART, PEDIAr can facilitate timely interventions and improve personalized care.

(JMIR Public Health Surveill 2025;11:e67005) doi:10.2196/67005

KEYWORDS

antiretroviral therapy; adherence; perceived barriers; psychometric analysis; human immunodeficiency virus; Brazil; HIV

XSL•FO
RenderX

Introduction

Worldwide estimates point to 30.7 million out of 39.9 million (76.9%) people living with HIV with access to antiretroviral therapy (ART) at the end of 2023 [1]. However, ART adherence is a challenge for people living with HIV across multiple settings and populations [2]. Brazil's HIV epidemic is characterized as a concentrated epidemic, with specific populations bearing a disproportionate burden of the disease. Among these groups, gay, bisexual, other men who have sex with men (MSM), and transgender and nonbinary (TGNB) persons are one of the most severely affected population groups. This population experiences significantly higher HIV prevalence [3] and HIV incidence rates [4]. In Brazil, nationally representative data on the cascade of HIV care is still scarce, with a 2016 study conducted in the city of Rio de Janeiro among 793 MSM and 37 transwomen indicating that 72.5% of those diagnosed were linked to care, 68.7% retained in care, and 61.1% were taking ART [5]. In terms of ART adherence in Brazil, a systematic review with meta-analysis that included 24 studies using various adherence measures published between 2005 and 2016 estimated that only 64% of people living with HIV achieved optimal ART adherence, frequently defined as ≥95% though it varied by study [6].

Numerous obstacles to achieve and sustain ART adherence have been investigated, notably at the individual level age, education, knowledge about ART, mental and physical health, and use of alcohol and illicit drugs [7-9]. Nonetheless, interpersonal factors have gained considerable attention, such as HIV-related stigma, lack of family and community support, and patient–health care provider mistrust [10,11]. In consonance, perceived barriers of people living with HIV, which denote an individual's assessment of social, personal, and environmental hurdles to fulfill a goal [12], have risen in importance [13]. Evidence for other chronic diseases, such as cancer and heart failure, show that perceived barriers are strong predictors of health behaviors [14,15]. In the context of HIV care, perceived barriers of treatment refer to the struggles people living with HIV face to initiate and sustain ART adherence [16]. Daily ART adherence is the best predictor of viral suppression [17], and it hinders HIV transmission, disease progression, and the development of drug-resistant viral strains [18].

Person-reported outcome measures provide a unique perspective on health-related outcomes that are centered on the person as opposed to what is perceived by health professionals [19]. Existing person-reported instruments related to ART adherence focus on constructs such as HIV-related stigma [20], depression [21], and alcohol use [22]. These instruments, while important, do not necessarily capture how individuals perceive these factors as barriers to achieving adherence. Furthermore, even instruments specifically targeting perceived barriers, as the Perceived Social Support (PSS-HIV) [23] and the Perceived HIV Self-Management Scale (PHIVSMS) [24], address isolated barriers rather than the complex, interacting set of challenges that hinder ART adherence. As such, the development and availability of multidimensional instruments to assess perceived barriers to achieving ART adherence is warranted [25].

The Perceived Barriers to Antiretroviral Therapy Adherence (in Portuguese, Percepção de Dificuldades com o Tratamento Antirretroviral—PEDIA) scale is an instrument in Brazilian Portuguese comprising 18 statements that describe challenges of living with HIV and barriers to taking ART medication. This instrument was developed through an initial qualitative analysis of 3 open-ended questions on the aspects that facilitate and hinder daily ART adherence which were answered by 598 people living with HIV across all Brazilian regions. Subsequently, 47 statements were elaborated and a content validity was performed, eliminating 7 items [16]. The final version of the scale was proposed after its evaluation in a sample of people living with HIV receiving health care in Belo Horizonte, Brazil [26].

As hypothesized, the PEDIA scale may be able to predict ART nonadherence, and thus, provide an opportunity to intervene before the therapy fails. In addition, it can aid our understanding of the specific barriers faced by MSM and TGNB persons living with HIV, however an assessment of the instrument's validity in this population is lacking. In this study, we sought to evaluate the psychometric properties of the PEDIA scale among MSM and TGNB persons in Brazil.

Methods

Study Design

Using a cross-sectional study design, we evaluated the factor structure, construct validity, and internal reliability of the PEDIA scale using 2 convenience samples recruited through internet-based surveys in Brazil.

Study Population

A total of 2 independent samples of Brazilian MSM and TGNB persons were recruited through location-based dating apps and social media. Sample 2020 was recruited between February 20, 2020, and March 29, 2020 through Hornet [Queer Networks Inc.], where users received an inbox invite to participate in the research [27]. Sample 2021 was recruited between May 14, 2021, and January 29, 2022 using 3 different location-based dating apps (Hornet, Grindr [Grindr LLC], and Scruff [Perry Street Software]) and social media (Facebook and Instagram [Meta Platforms]) [28]. Recruitment strategies varied by app: direct inbox messaging was used for Hornet, whereas banners were used on Grindr and Scruff. Boosted posts were used on Facebook and Instagram. Individuals were not compensated for their participation as per Brazilian research regulations.

Adult (age ≥18 y) participants who provided electronic informed consent were directed to the study's questionnaires. Participants who (1) identified themselves as cisgender women, (2) reported living abroad, (3) incorrectly answered any of the attention questions (only applicable to Sample 2020) [29], or (4) abandoned the questionnaire before reaching the last question were excluded. Furthermore, in line with the study's objective to evaluate an instrument addressing barriers to sustaining ART adherence, participants who reported their HIV status as negative or unknown were excluded. Duplicate responses were identified and removed based on the IP address, with only the first instance

XSL•FO
RenderX

retained for analysis. To protect respondent privacy, a masked IP was generated, used for duplicate removal and then deleted.

Recruitment Apps

Grindr, Hornet, and Scruff are geosocial networking (GSN) apps designed for dating and sexual encounters mostly used by MSM. According to the Google play store, as of July 2024, Grindr and Hornet had each surpassed 10 million downloads and Scruff, 5 millions in Brazil alone. These apps are quite popular among MSM and recruitment based on them has been previously used in other cross-sectional studies in Brazil [27,28,30-35]. Instagram and Facebook are social media apps widely used in Brazil. Although these figures vary over time, as of June 2024, 29.6% of Brazilians actively use Instagram, whereas 33% use Facebook [36].

Study Instruments

Questionnaires were administered via Alchemer and consisted of about 100 questions, though the actual number of questions was dependent on the participant's responses given conditional logic. Participants were able to change and review any of their answers and always had the possibility to choose a nonresponse option, usually labeled as "Prefer not to answer." Usability and technical functionality of questionnaires were reviewed in different devices and operating systems before launching the study.

The questionnaires had 3 main sections. Section 1 included sociodemographic characteristics of the participants (age, gender, sexual orientation, race, education, income, and state of residency). Section 2 addressed the use of ART, and if in use, the level of ART adherence. Section 3 focused on possible barriers to achieve ART adherence, such as HIV stigma, people-reported barriers, alcohol use, drug use, among others.

The WebAd-Q is an instrument developed and validated in Brazil to assess ART adherence. The validity and internal consistency were evaluated in a sample of 74 adults taking ART for at least 3 months. Construct validity was supported by high correlations between responses to the WebAd-Q items and 5 different measures of ART adherence. Test-retest measures indicated a Kappa coefficient of 0.77 [37]. The WebAd-Q instrument consists of 3 questions inquiring on not following scheduled hours for ART intake, missing ART doses, and incorrectly taking more or less pills than prescribed. For all questions, the recall period was 7 days and the response options included "Yes," "No," or "I don't remember." As per instrument's instructions, a participant was deemed adherent if they answered "No" to all 3 questions. Following the three WebAd-Q questions, the participants were prompted with a visual analogue scale (VAS) item: "please mark below the value that corresponds to how much of your antiretroviral medication you took in the past 30 days" that varied from 0% to 100%. A dichotomous variable, denominated "adherence to ART based on VAS dummy," was created by categorizing <100% as nonadherence and 100% as adherence [38,39]. In sum, the 3 measures of ART adherence were the WebAd-Q instrument, the 0% - 100% VAS, and the "Yes/No" adherence based on VAS.

The PEDIA scale is an instrument developed in Brazil to assess various barriers to achieve ART adherence as reported by people living with HIV. Initial validity of the PEDIA scale was assessed in a study of 415 adults accompanied in 3 health care facilities in Brazil, with results showing good reliability (McDonald's omega of 0.97) and construct validity (one point increase in PEDIA's score increases the odds of nonadherence to ART in 12%) [26]. PEDIA's 18 statements should be answered on a 3-point Likert scale of agreement: "disagree" (coded as 1), "neither agree nor disagree" (coded as 2), and "agree" (coded as 3). Items 5 and 10 were reverse coded so that all statements reflect barriers and challenges of living with HIV and of daily medication. An overall score of the scale is given by the sum of each item's response, with higher scores indicative of more perceived barriers (score range 18 - 54). PEDIA captures 3 different dimensions, labelled as "patient's fears and feelings," "cognitive and routine problems," and "medication and health concerns" (see Table 1). For each dimension, the corresponding score is also calculated by summing the respective items, the score range of each dimension is as follows: "patient's fears and feelings," 8 items, score range 8 - 24, "cognitive and routine problems," 4 items, score range 4 - 12, and "medication and health concerns," 6 items, score range 6 - 18.

XSL•FO
RenderX

Table . The 18 items of the perceived barriers to antiretroviral therapy adherence (PEDIA) scale by dimension (adapted from the original paper [26]).

Item number	Item content
Patient's fears and feelings	
1	The main problem of living with HIV is the stigma around it.
2	I am afraid to be identified as living with HIV when I go to the health care facility to get my HIV meds refill.
3	It frustrates me to think that I need to take the HIV meds in order to be alive.
7	I do not like to take my HIV meds around others.
11	It is tiresome to take my HIV meds every day.
14	It is difficult to tell people that I am living with HIV.
15	I am worried about the HIV meds stopping to work in the future.
17	It bothers me that I have to get my HIV meds refill in the health care facility's pharmacy.
Cognitive and routine problems	
6	It is difficult to take my HIV meds at home.
8	Sometimes I forget to take my HIV meds because I get distracted.
9	It is difficult to take my HIV meds at work.
18	It is harder to keep track of my HIV meds on weekends.
Medication and health concerns	
4	Sometimes I skip taking my HIV meds because I want to avoid side effects.
5	Despite my HIV status, I live a normal life.
10	I believe that my HIV meds make me healthy.
12	I find it difficult to swallow the pills.
13	When I feel depressed I do not want to take my HIV meds.
16	It is hard to get used to the side effects.

Statistical Analysis

We used a split-sample approach, using separate samples for exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). All EFAs were conducted using Sample 2020, and all CFAs were conducted using Sample 2021.

Sample 2020: Exploratory Factor Analysis

First, we calculated the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and performed the Bartlett test of sphericity to explore the presence of latent factors and the suitability of the data for EFA [40]. The KMO measure was deemed acceptable if greater than 0.6 and the significance level for the Bartlett test of sphericity was 0.05.

Then, we proceeded to implement an EFA (EFA 1) with polychoric correlation matrix and robust diagonally weighted least squares as the extraction method, given the 3-points Likert scale response format [41]. We first considered the originally proposed 3-factor structure, using parallel analysis (PA) with minimum rank factor analysis to determine the optimal number of factors [42] and robust promin rotation [43]. Based on the PA result, we reran our EFA (EFA 2) with the same specifications aforementioned and the optimal number of factors given by the PA. Items with standardized factor loadings below 0.4 were removed from subsequent analyses, aiming a more meaningful and robust factor structure [44,45]. While theoretical relevance was also considered during item removal, the primary criterion was the 0.4 factor loading threshold. Low loadings indicate that the item does not strongly represent the factor, and retaining such items can add noise and reduce the clarity of the factor structure [46].

Sample 2021: Confirmatory Factor Analysis

Akin to the initial steps undertaken for Sample 2020, we first calculated the KMO measure of sampling adequacy and performed Bartlett's test of sphericity to assess the presence of latent factors and the suitability of the data for CFA. We used the optimal number of factors with the remaining items from our second EFA to run this CFA. Our CFA was performed using a polychoric correlation matrix and weighted least squares estimator with a diagonal weight matrix and mean- and variance-adjusted χ^2 test statistic (WLSMV), robust standard errors and delta parametrization [41]. Similarly to our EFA results, we removed items with standardized loadings below 0.4 [44,45].

In addition, modification indices were used to identify changes that could improve model fit such as items that were loading on more than 1 factor, or items heavily correlated with other

XSL•FO
RenderX

items, among others. Theoretically plausible modifications were discussed by the study team. Finally, a second CFA was performed on the final version of the scale using a polychoric correlation matrix and WLSMV estimator, robust standard errors, and delta parametrization [41].

Since the optimal number of factors given by the PA was different from that of the original PEDIA scale, we performed a CFA on the original PEDIA scale with all the original items grouped into 3 factors, as described previously. We sought to show the difference in model fit and factor loadings for this version (see Table S1 in Multimedia Appendix 1).

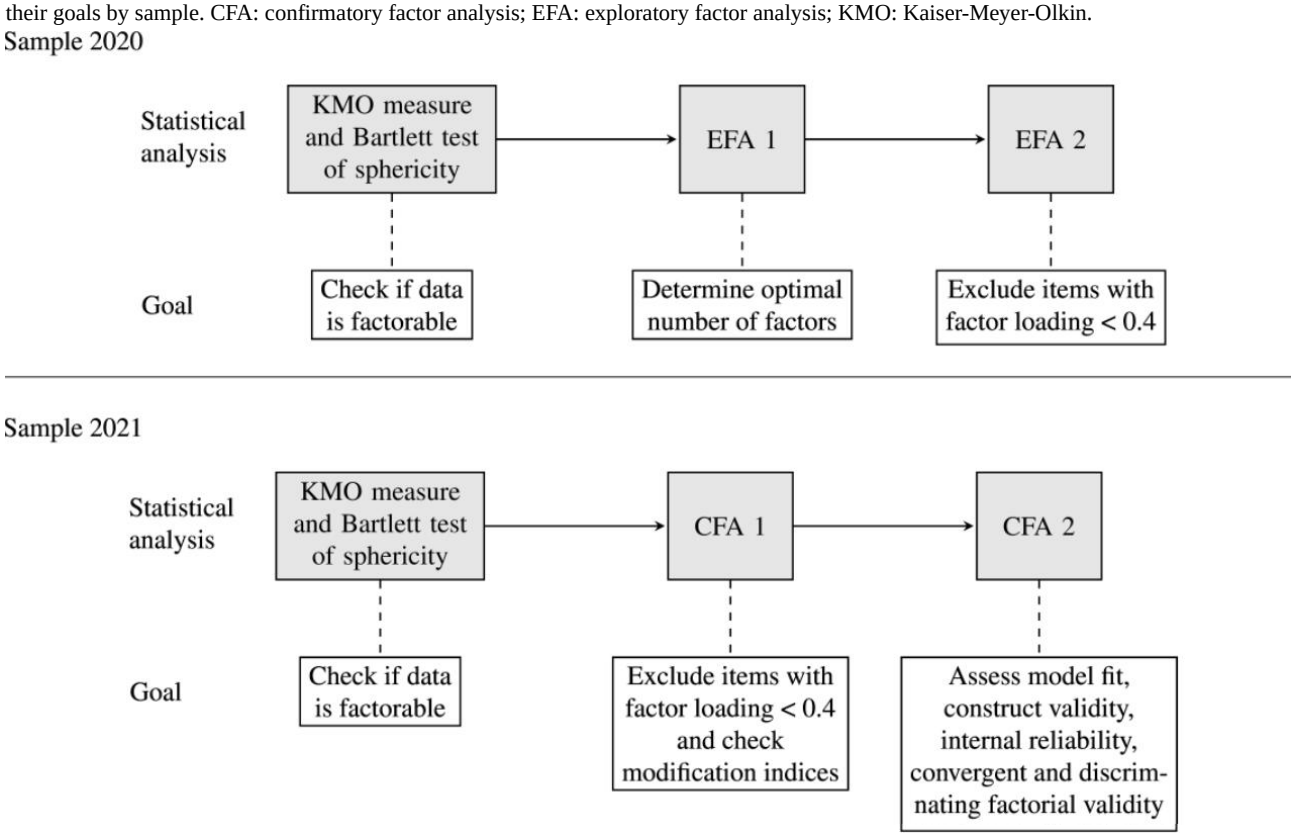
Model fit was assessed using root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean square residual (SRMR). The χ^2 statistic was not emphasized in our results due to its sensitivity to sample size and rejection of well-fitting models [47]. Model fitting was deemed acceptable if CFI≥0.95, TLI≥0.95, RMSEA≤0.06, and SRMR≤0.08 [48].

Construct validity was verified by standardized factor loadings above 0.5 and correlation assessment of the final version of the PEDIA scale with the 3 previously described measures of ART adherence. A negative correlation between perceived barriers as measured by PEDIA and ART adherence was taken as evidence of construct validity [49]. Spearman correlation coefficient was used for continuous variables, whereas point-biserial correlation was employed for categorical variables.

Internal consistency (reliability) was assessed using McDonald's omega, which is considered an appropriate measure when handling ordinal variables (polychoric correlation matrices) and non-normally distributed variables [46]. Convergent factorial validity was evaluated based on the average variance extracted (AVE), which indicates convergence when $\sqrt{AVE} \geq 0.7$ [50]. Finally, discriminating factorial validity was assessed by testing the statistical difference between the AVE of each factor and the square of its correlations with others [50]. If the difference was statistically not significant at 5%, then discriminating factorial validity was endorsed [51].

Figure 1 shows the statistical analyses and their goals as performed in each sample. EFAs were conducted in Factor 12.04.05 Lorenzo-Seva and Ferrando [52] and CFAs in MPLUS 8.8 Muthén and Muthén [53]. Correlations were calculated in R (version 4.2.3; R Foundation for Statistical Computing) [54].

Figure 1. Flowchart of all statistical analyses performed to define the final version of perceived barriers to antiretroviral therapy adherence (PEDIA) and their goals by sample. CFA: confirmatory factor analysis; EFA: exploratory factor analysis; KMO: Kaiser-Meyer-Olkin.



Ethical Considerations

This study received approval from the human subject's ethics committee at Instituto Nacional de Infectologia Evandro Chagas of Fundação Oswaldo Cruz. The study received the following ethics approvals: #2.531.022, issued on March 7, 2018, and #4.300.120, issued on September 25, 2020. All study participants provided electronic informed consent before survey initiation. To prevent duplicate responses, IP addresses were collected and subsequently discarded after data cleaning. No other personally identifiable information was collected. Individuals were not compensated for their participation as per Brazilian research regulations.

XSL•FO
RenderX

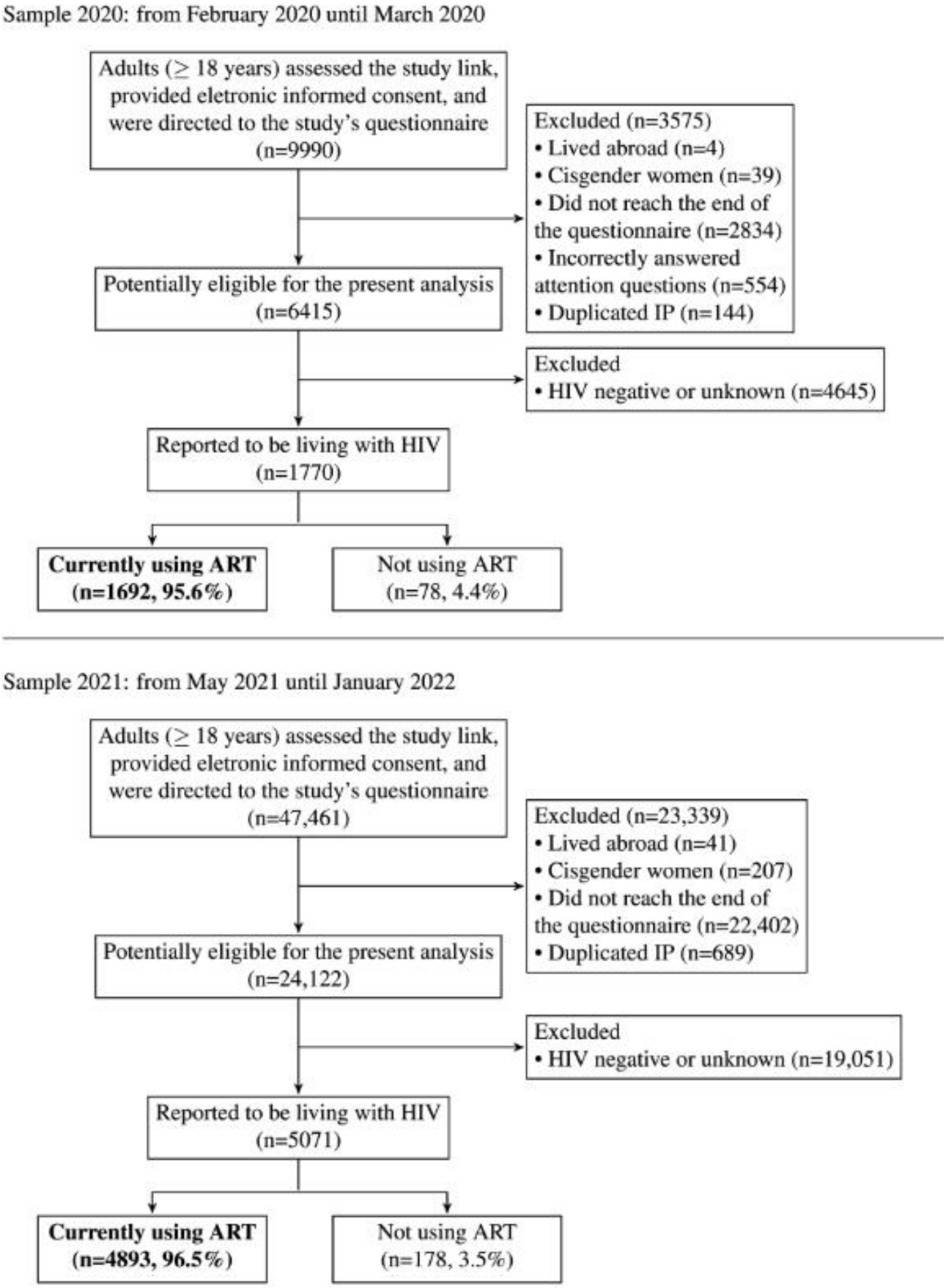
Results

Participant Characteristics

Recruitment for Sample 2020 reached 9990 adults, out of which 3575 were excluded (see Figure 2 for reasons). Amidst those living with HIV (n=1770), 4.4% (n=78) reported not having initiated ART and 95.6% (n=1692) reported having initiated ART. Recruitment for Sample 2021 reached 47,641 adults, out of which 23,339 were excluded (see Figure 2 for reasons). Amidst those living with HIV (n=5071), 3.5% (n=178) reported not having initiated ART or chose not to answer this question whereas 96.5% (n=4893) reported having initiated ART. In total, Sample 2020 and Sample 2021 were composed, respectively, of 1692 and 4893 participants who were living with HIV and had initiated ART (Figure 2).

XSL•FO
RenderX

Figure 2. Flowchart of the study population and exclusion criteria by sample. ART: antiretroviral therapy.



In Sample 2020 (see Table 2), median age was 36 years (IQR 30 - 44 y), most participants were cisgender men (n=1665, 98.4%), self-identified as gay (n=1555, 91.9%) and White (n=1024, 60.5%), reported having a college degree or higher (n=1158, 68.4%), and were living in the Southeast region of Brazil (n=1386, 81.9%). A total of 971 (57.4%) were deemed

XSL•FO
RenderX

adherent to ART based on the WebAd-Q. According to the VAS, 69.6% (n=1778) participants indicated 100% adherence to ART. Median PEDIA total score was 36 (IQR 32 - 40), and breaking down by dimension, median scores were 17 (IQR 14 - 19) for patient's fears and feelings, 6 (IQR 4 - 8) for cognitive and routine problems, and 11 (IQR 10 - 12) for medication and health concerns. Sample 2021 exhibited a more diversified regional distribution (region of residency) when compared with Sample 2020 and some small *P* values emerged in demographic comparisons. All measures of ART adherence and PEDIAr scores were similar across samples (see Table 2).

XSL•FO
RenderX

Table . Characteristics of the 2 study samples.

Characteristics	Sample 2020 (n=1692)	Sample 2021 (n=4893)	*P* value
Age (years), median (IQR)	36 (30-44)	38 (31-45)	<.001
Age group (years), n (%)			.003
18 - 24	90 (5.3)	241 (4.9)	
25 - 29	283 (16.7)	651 (13.3)	
30 - 39	667 (39.4)	1900 (38.8)	
40 - 49	392 (23.2)	1317 (26.9)	
50 - 59	223 (13.2)	679 (13.9)	
60+	37 (2.2)	105 (2.1)	
Gender, n (%)			.68
Cisgender man	1665 (98.4)	4800 (98.1)	
Trans man	2 (0.1)	5 (0.1)	
Trans woman	2 (0.1)	9 (0.2)	
Travesti	23 (1.4)	74 (1.5)	
Nonbinary	0 (0)	5 (0.1)	
Sexual orientation, n (%)			.003
Gay	1555 (91.9)	4344 (88.8)	
Bisexual	118 (7.0)	451 (9.2)	
Heterosexual	7 (0.4)	41 (0.8)	
Other[a]	12 (0.7)	57 (1.2)	
Race, n (%)			.003
Asian	20 (1.2)	43 (0.9)	
Black	173 (10.2)	577 (11.8)	
Pardo/mixed	419 (24.8)	1426 (29.1)	
Indigenous	18 (1.1)	52 (1.1)	
White	1024 (60.5)	2767 (56.6)	
Missing	38 (2.2)	28 (0.6)	
Schooling, n (%)			.08
High school or less	514 (30.4)	1646 (33.6)	
Undergraduate degree	650 (38.4)	1819 (37.2)	
Graduate degree	508 (30.0)	1416 (28.9)	
Missing	20 (1.2)	12 (0.2)	
Household income[b], n (%)			.13
Low	457 (27)	1428 (29.2)	
Middle	776 (45.9)	2116 (43.2)	
High	459 (27.1)	1297 (26.5)	
Missing	0 (0)	52 (1.1)	
Region of residency, n (%)			<.001
North	10 (0.6)	120 (2.5)	
Northeast	38 (2.2)	431 (8.8)	
Central-west	55 (3.3)	330 (6.7)	
Southeast	1386 (81.9)	3317 (67.8)	

XSL•FO
RenderX

Characteristics	Sample 2020 (n=1692)	Sample 2021 (n=4893)	*P* value
South	203 (12.0)	695 (14.2)	
Adherence to ART[c]-based on WebAd-Q[d] , n (%)			>.99
No	721 (42.6)	2087 (42.7)	
Yes	971 (57.4)	2806 (57.3)	
Adherence to ART based on VAS[e], median (IQR)	100 (98-100)	100 (98-100)	.24
Adherence to ART based on VAS dummy[f], n (%)			.17
No	514 (30.4)	1398 (28.6)	
Yes	1178 (69.6)	3495 (71.4)	
PEDIA[g] scale			
Total score, median (IQR)	36 (32-40)	36 (32-40)	.03
Fear and feelings, median (IQR)	17 (14-19)	17 (14-19)	.11
Cognitive and routines problems, median (IQR)	6 (4-8)	6 (4-8)	.50
Medication and health concerns, median (IQR)	11 (10-12)	11 (10-12)	.01

[a]"Other" includes responses of asexual, pansexual, and other.

[b]Household income was based on the ratio of household income relative to the number of minimum wages monthly: low ≤2, middle >2 and ≤ 6, and high >6.

[c]ART: antiretroviral therapy.

[d]Answered "no" to all 3 questions.

[e]VAS: visual analog scale.

[f]Dummy equals 1 if answered "100" on VAS, and 0 otherwise.

[g]PEDIA: perceived barriers to antiretroviral therapy adherence.

Exploratory Factor Analysis

The KMO measure of sampling adequacy was 0.870 and the Bartlett test of sphericity *P* value was <.001, indicating that the data was factorable and that the EFA could be conducted. The first EFA (EFA 1) was performed using the original number of factors, that is, 3. The parallel analysis based on minimum rank factor analysis indicated the presence of 2 factors. Accordingly, the EFA was repeated (EFA 2) using 2 factors (see Table 3). Results showed that most items had standardized factor loadings ≥0.4 except for items 5, 10, and 16 and these items were thus removed from subsequent analyses (see Table 3).

XSL•FO
RenderX

Table . Standardized factor loadings for the perceived barriers to antiretroviral therapy adherence (PEDIA) scale by type of factory analysis used.[a]

		EFA 2		CFA 1		CFA 2	
Item number	Item content	F1	F2	F1	F2	F1	F2
1	The main problem of living with HIV is the stigma around it.	0.518	−0.232	0.363[b]	—[c]	—	—
2	I am afraid to be identified as living with HIV when I go to the health care facility to get my HIV meds refill.	0.918	−0.276	0.646[b]	—	—	—
3	It frustrates me to think that I need to take the HIV meds in order to be alive.	0.472	0.200	0.633	—	0.673	—
4	Sometimes I skip taking my HIV meds because I want to avoid side effects.	−0.067	0.793	—	0.647	—	0.648
5	Despite my HIV status, I live a normal life.	−0.229[d]	−0.366[d]	—	—	—	—
6	It is difficult to take my HIV meds at home.	−0.186	0.890	—	0.756	—	0.778
7	I do not like to take my HIV meds around others.	0.430	0.261	0.700	—	0.682	—
8	Sometimes I forget to take my HIV meds because I get distracted.	−0.265	0.781	—	0.600	—	0.638
9	It is difficult to take my HIV meds at work.	−0.096	0.850	—	0.818	—	0.829
10	I believe that my HIV meds make me healthy.	−0.117[d]	−0.312[d]	—	—	—	—
11	It is tiresome to take my HIV meds every day.	0.390	0.434	0.775	—	0.835	—
12	I find it difficult to swallow the pills.	0.136	0.459	—	0.597[b]	—	—
13	When I feel depressed I do not want to take my HIV meds.	0.046	0.740	—	0.705	—	0.680
14	It is difficult to tell people that I am living with HIV.	0.869	−0.233	0.646[b]	—	—	—

XSL•FO
RenderX

		EFA 2		CFA 1		CFA 2	
Item number	Item content	F1	F2	F1	F2	F1	F2
15	I am worried about the HIV meds stopping to work in the future.	0.518	0.049	0.491[b]	—	—	—
16	It is hard to get used to the side effects.	0.161[d]	0.397[d]	—	—	—	—
17	It bothers me that I have to get my HIV meds refill in the health care facility's pharmacy.	0.739	−0.039	0.699	—	0.606	—
18	It is harder to keep track of my HIV meds on weekends.	0.022	0.601	—	0.685	—	0.675

[a]Two types of analysis were conducted: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). All analyses used a 2-factor structure. EFA 2 was performed using Sample 2020. CFAs 1 and 2 were performed using Sample 2021.

[b]Item removal in CFA 1 was based on modification indices and theoretical relevance.

[c]Not applicable.

[d]Item removal in EFA 2 was primarily determined by standardized factor loadings below 0.4, supplemented by theoretical considerations.

Confirmatory Factor Analysis

Fit indices for the 2-factor model of the PEDIA scale with 15 items were poor (RMSEA=0.08, 90% CI 0.081 - 0.086; CFI=0.89; TLI=0.87; and SRMR=0.08). Items 1 ("The main problem of living with HIV is the stigma around it") and 14 ("It is difficult to tell people that I am living with HIV") do not specifically relate to the challenges of treatment. Items 12 ("I find it difficult to swallow the pills") and 15 ("I am worried about the HIV meds stopping to work in the future") tackle an issue that is not as relevant nowadays given the large advancements in the ART medication. In addition, item 1 did not load above 0.4, and items 12 and 15 were the lowest factor loadings in their corresponding factors. The team decided that these four items should be excluded (see Table 3).

The modification indices were then explored and discussed by the study team. Results showed that item 2 ("I am afraid to be identified as living with HIV when I go to the health care facility to get my HIV meds refill") was strongly correlated with items 11, 14 and especially 17, suggesting that its content was overlapping with the aforementioned items. The idea proposed in item 2 ("I am afraid to be identified as living with HIV when I go to the health care facility to get my HIV meds refill") also seemed already encompassed in item 17 ("It bothers me that I have to get my HIV meds refill in the health care facility's pharmacy"). The team then decided to exclude item 2 (see Table 3).

Subsequent to the exclusion of items, the reduced version of the PEDIA scale, henceforth referred to as PEDIAr, retained 10 items with a 2-factor structure. Factor 1, labelled "psychological aspects," encompassed items 3, 7, 11, and 17, whereas factor 2, labelled "practical aspects," grouped items 4, 6, 8, 9, 13, and 18 (see Table 4). PEDIAr showed an acceptable fit (RMSEA=0.07, 90% CI 0.067 - 0.075; CFI=0.95; TLI=0.94; SRMR=0.05). Construct validity was also supported by factor loadings well above 0.5 (see Figure 3). Internal consistency was verified by McDonald omega: 0.795 for factor 1, and 0.859 for factor 2. Both factors exhibited convergent factorial validity, as evidenced by their √AVE values of 0.704 and 0.711 for factor 1 and 2, respectively. In addition, discriminating factorial validity was supported, with no significant differences between the AVE of each factor and the square of its correlations with others: 0.001 (P=.94) for factor 1 and 0.009 (P=.55) for factor 2.

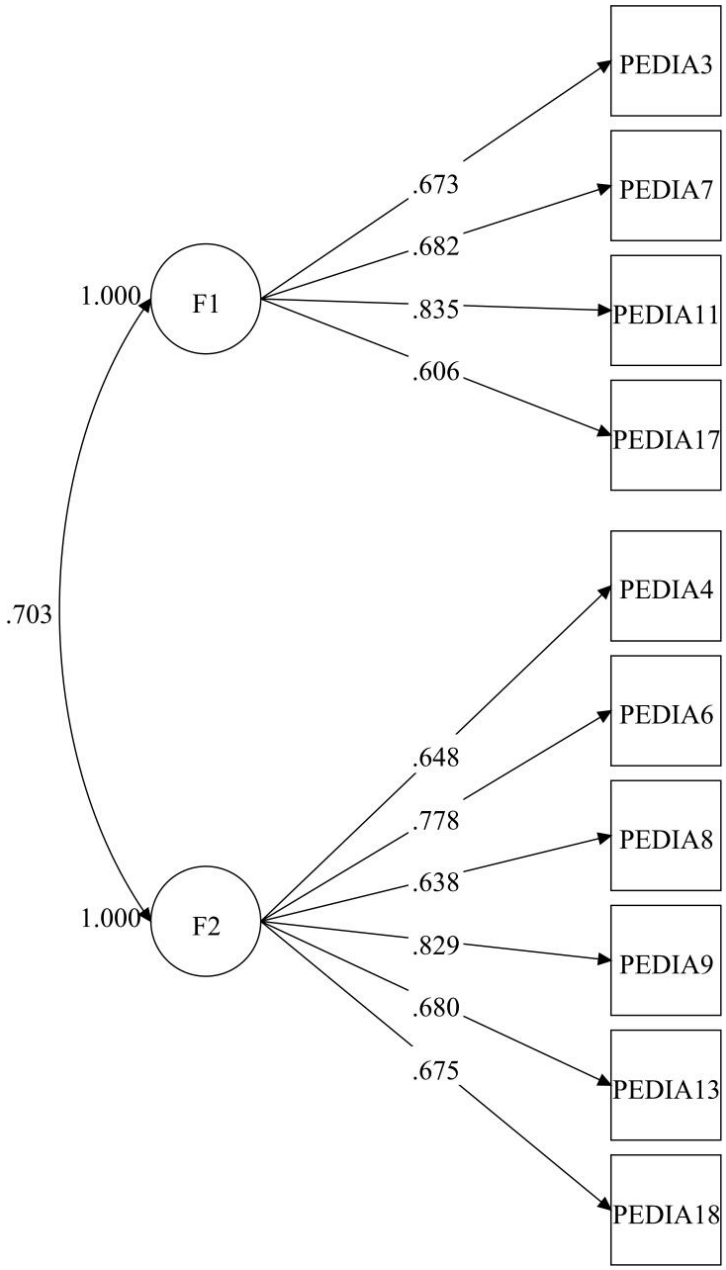
XSL•FO
RenderX

Table . The 10 items of the reduced version of the perceived barriers to antiretroviral therapy adherence scale (PEDIAr) by factor.

Item number	Item content
Psychological aspects	
3	It frustrates me to think that I need to take the HIV meds in order to be alive.
7	I do not like to take my HIV meds around others.
11	It is tiresome to take my HIV meds every day.
17	It bothers me that I have to get my HIV meds refill in the health care facility's pharmacy.
Practical aspects	
4	Sometimes I skip taking my HIV meds because I want to avoid side effects.
6	It is difficult to take my HIV meds at home.
8	Sometimes I forget to take my HIV meds because I get distracted.
9	It is difficult to take my HIV meds at work.
13	When I feel depressed I do not want to take my HIV meds.
18	It is harder to keep track of my HIV meds on weekends.

XSL•FO
RenderX

Figure 3. Standardized factor loadings for the reduced version of the perceived barriers to antiretroviral therapy adherence (PEDIAr) scale by factor and the correlation between the factors. F1: factor 1, labelled as "psychological aspects"; F2: factor 2, labelled as "practical aspects"; PEDIA: perceived barriers to antiretroviral therapy adherence.



Complementarily to the factor loadings, the correlations between the 10-item PEDIAr scale and all measures of ART adherence were negative and with a relevant magnitude, especially between factor 2 (6 items, labeled “practical aspects”) and ART adherence measured by WebAd-Q and between the total score of the PEDIAr and ART adherence measured by WebAd-Q in both samples (see Table 5). It is also noteworthy that the correlations were quite similar in the 2 samples.

XSL•FO
RenderX

Table . Correlations of the reduced version of the perceived barriers to antiretroviral therapy adherence scale (PEDIAr)—by factor and total score—and antiretroviral therapy (ART) adherence measures.

	Type	Sample 2020 (95% CI)	Sample 2021 (95% CI)
PEDIAr Factor 1			
ART adherence WedAd-Q	Biserial	−0.187 (−0.232 to −0.140)	−0.183 (−0.210 to −0.156)
ART adherence VAS[a]	Spearman	−0.175 (−0.221 to −0.129)	−0.147 (−0.175 to −0.120)
ART adherence VAS dummy[b]	Biserial	−0.165 (−0.211 to −0.118)	−0.136 (−0.164 to −0.109)
PEDIAr Factor 2			
ART adherence WedAd-Q	Biserial	−0.517 (−0.551 to −0.482)	−0.496 (−0.517 to −0.475)
ART adherence VAS	Spearman	−0.442 (−0.461 to −0.382)	−0.395 (−0.418 to −0.371)
ART adherence VAS dummy	Biserial	−0.407 (−0.446 to −0.366)	−0.376 (−0.340 to −0.352)
PEDIAr (10-items)			
ART adherence WedAd-Q	Biserial	−0.432 (−0.470 to −0.392)	−0.408 (−0.431 to −0.384)
ART adherence VAS	Spearman	−0.355 (−0.396 to −0.313)	−0.315 (−0.340 to −0.289)
ART adherence VAS dummy	Biserial	−0.349 (−0.391 to −0.307)	−0.308 (−0.333 to −0.282)

[a]VAS: visual analog scale.

[b]ART adherence VAS dummy=1, if the participant marked 100 on VAS, 0 otherwise.

Discussion

Principal Findings

As described previously, the PEDIA scale was first assessed among 415 people living with HIV (n=336, 81% men and n=79, 19% women; mean age 35, SD 10.45 years) using ART from 3 health care facilities in the city of Belo Horizonte, Brazil, who were recruited between September 2015 and October 2017. In this study, the original 3-factor structure of the PEDIA scale was not supported by the parallel analysis carried out in our EFA 1 and was further disconfirmed by its CFA fit indices that did not meet any of the conventionally used cutoffs (see Table S1 in Multimedia Appendix 1). Therefore, for our population that was predominantly composed of cisgender MSM (n=4795, 98%, mean age 39, SD 9.84 years), we proposed a reduced version, the PEDIAr scale, for which we have provided evidence of construct validity, convergent and discriminating factorial validity, and a well-fitted 2-factor structure. As recommended [46,51], this proposed refinement of the PEDIA scale was achieved based on both item-item and item-total correlations, along with theoretical relevance. Accordingly, we believe the PEDIAr scale retains PEDIA's original aim to assess perceived barriers to adherence from the perspective of the person, that is, to assess which factors hinder daily adherence to ART.

In comparison with the original structure, some items initially assigned to the dimensions representing "cognitive and routine problems" (4 items) and "medication and health concerns" (6 items) were merged into a single dimension "practical aspects" with 6 items (items 5, 10, 12, and 16 of the original scale were excluded), whereas the dimension "patient's fears and feelings" remained as one of the factors, renamed "psychological aspects," with 4 items (items 1, 2, 14, and 15 were excluded). PEDIAr presented good fitness measures (TLI, CFI, and SRMR), good internal consistency provided by the McDonald's omega, construct validity through high factor loadings, and convergent and discriminating factorial validity assessed by the AVE. Our results suggest a restructuring of the original 3-factor model into 2: one focusing on subjective barriers to daily medication adherence, including thoughts and feelings (denominated "psychological aspects"), and another focusing on the objective challenges to achieving daily adherence, including habit formation and normal disruptions to daily routines, hence the denomination "practical aspects."

Some items of PEDIAr tangentially refer to internalized HIV-related stigma and depressive symptoms, which are factors that intertwine with perceived barriers to sustain ART adherence. Item 9 ("It is difficult to take my HIV meds at work") touches on the stigma faced by people living with HIV, that is, taking antiretroviral drugs in public or alongside peers poses a threat to the person's well-being and sense of worth [55]. Item 13 ("When I feel depressed I do not want to take my HIV meds") alludes to the intricate relationship between depression and ART adherence. A meta-analysis across 29 studies found that the treatment of depression and psychological distress improved ART adherence [56]. Recent longitudinal evidence from Uganda (n=143) suggests a bidirectional relationship between depression and ART adherence [57]. Furthermore, other intersecting stigmas, such as internalized homonegativity, might play a role deterring ART adherence through the aforementioned factors. There is a large body of literature showing that internalized homonegativity is positively associated with depressive symptoms [58-60]. Similarly, a meta-analysis including 176 studies published between 2000 and 2021 found a positive association between internalized homonegativity and various forms of HIV-related stigma, including internalized HIV stigma [61]. When people living with HIV's ability to adhere to ART is compromised, PEDIAr overall scores would likely increase. However, further research is needed to specifically examine the relationship between internalized homonegativity and PEDIAr scores.

XSL•FO
RenderX

Our results provide further evidence on PEDIAr's construct validity based on the correlations found between the scale and ART adherence [49]. Not only did the correlations show a relevant magnitude and expected direction of effect (ie, higher PEDIAr scores correlate with poor adherence), but they were consistent through varied measures of ART adherence. Furthermore, the results were similar in both samples, which corroborates to the reliability of the scale [51]. In this regard, it is important to highlight that some of the items excluded in our analysis did not tackle issues directly related to ART adherence such as items 1 ("The main problem of living with HIV is the stigma around it"), 5 ("Despite my HIV status, I live a normal life"), and 14 ("It is difficult to tell people that I am living with HIV"), while others were excluded because they referred to issues that were more prevalent with older, less efficient and more toxic medications, such as items 4 ("Sometimes I skip taking my HIV meds because I want to avoid side effects"), 12 ("I find it difficult to swallow the pills"), 15 ("I am worried about the HIV meds stopping to work in the future"), and 16 ("It is hard to get used to the side effects"), which are likely less relevant nowadays [62]. Taken together, these findings suggest that PEDIAr may be more specific in its assessment of current barriers to ART adherence.

In comparison with other existing instruments, most assess specific constructs derived from qualitative research with people living with HIV that have been shown to correlate with ART adherence [25]. Other instruments which assess perceived barriers per se, meaning the perception of factors as barriers to adherence, do it as unidimensional scales and this can hinder relevant information about the interplay among them [63]. In this instance, a major advantage of the original PEDIA scale is its multidimensionality, which was retained in PEDIAr. Furthermore, PEDIAr was consistently correlated with ART adherence, an essential attribute to be used as an insightful measure of a person's perception of barriers to sustain ART adherence [26]. In terms of clinical practice, PEDIAr offers the advantage of brevity, comprising only 10 items, and may be self-administered by people living with HIV during routine visits, reducing time burden for both people seeking care and health care providers. While further research may be needed to explore its application within other populations, our study offers compelling evidence of PEDIAr's applicability, especially among MSM, suggesting its value as a tool for routine clinical assessments.

Our study has some limitations that should be acknowledged. First, our sampling strategy relied on GSN apps and social media and thus results may not be generalizable to all Brazilian gay, bisexual and other MSM or, to an even greater extent, to other TGNB persons. Adding to the sampling limitations, both samples were primarily composed of individuals residing in the Southeast region of Brazil, the most populous and wealthiest region, potentially limiting the generalizability of our findings to other regions of the country. Nonetheless, given the challenges of reaching MSM and TGNB populations, especially in a nationwide context, the HIV prevalence in both samples was consistent with national studies targeting similar population groups, suggesting a degree of representativeness within the inherent limitations of internet-based sampling. In Sample 2020, HIV prevalence was estimated at 26.4% (n=6415), slightly higher than the average reported for Brazil in a respondent-driven sampling carried out in 12 cities in 2016 (n=3952, 18.4%) [64]. However, Sample 2020 was largely composed of participants from the southeast region, which exhibited a higher prevalence in the aforementioned study, especially in Rio de Janeiro (n=256, 15.3%) and São Paulo (n=341, 24.8%). In contrast, Sample 2021 was more diversified, capturing participants from other regions of Brazil. As a consequence, HIV prevalence found in Sample 2021 (n=24,122, 21.0%) was similar to the Brazilian average. Another limitation of our study is the possibility that participants from Sample 2020 were also included in Sample 2021 (meaning that they answered both surveys), which would violate the assumption of independence of our samples. However, we assessed the extent of this overlap based on the IP address and found it to be minimal: only 12 overlapping responses were identified, representing less than 1% of the total number of participants. Given the small magnitude of this overlap and that our split-sample approach does not rely on strict independence between the two samples, we did not remove these duplicates from the analysis. In addition, both of our instruments (PEDIA scale and ART adherence) were self-reported, which could be subject to recall and social desirability bias. However, existing evidence indicates that these self-reported measures are reliable [26,37].

Conclusions

Perceived barriers to adherence to ART can help address an important challenge faced by ~30% to 40% of people living with HIV across different settings: nonadherence to therapy. This study evaluated the original PEDIA scale in 2 large samples of MSM and TGNB persons living with HIV and validated a reduced version of the instrument, the PEDIAr scale. PEDIAr offers a 2-factor structure that distinguishes between psychological and practical aspects influencing adherence. This distinction provides valuable insights into designing interventions to tackle specific challenges faced by people living with HIV. Our results also stress PEDIAr's applicability, particularly among cisgender MSM, and its brevity and ease of administration make it accessible to diverse populations. Furthermore, shorter instruments not only reduce the participant's burden, but may also increase retention. By identifying individuals struggling with adherence, PEDIAr can pave the way for personalized care and interventions, ultimately improving health-related outcomes. Future studies should explore PEDIAr's performance in other populations and the relationship between barriers to adherence and ART adherence across different durations of ART use.

XSL•FO
RenderX

Acknowledgments

Initial results were presented as an oral presentation at the 13th Congresso Brasileiro de Saúde Coletiva in November 2022 [65]. PML was supported by Programa Inova FIOCRUZ, Edital Geração do Conhecimento/VPPCB, the National Council of Technological and Scientific Development (CNPq; #316401/2021-8) and Carlos Chagas Filho Foundation for Research Support in the State of Rio de Janeiro (FAPERJ; #E-26/201.133/2021). TST was supported by CNPq (#402916/2021-2 and #311871/2021-6) and FAPERJ (#E-26/211.577/2021 and #E-26/201.270/2022). VCM was supported by FAPERJ (#E-03/202.304/2024). Funders did not have any role in the study design, data collection, data analysis, interpretation, or writing of this report.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the sensitive nature of the information but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Standardized factor loadings of the original 3-factor perceived barriers to antiretroviral therapy adherence (PEDIA) scale using Sample 2021 data, and the reduced version of the PEDIA scale's final 10 items in Brazilian Portuguese, grouped by factor.
[DOCX File, 17 KB - publichealth_v11i1e67005_app1.docx]

References

1. UNAIDS. Latest global and regional statistics on the status of the AIDS epidemic. The United Nations Programme on HIV/AIDS. 2023. URL: https://www.humanitarianlibrary.org/sites/default/files/2023/10/UNAIDS_FactSheet_en.pdf [accessed 2024-04-18]
2. Achappa B, Madi D, Bhaskaran U, Ramapuram JT, Rao S, Mahalingam S. Adherence to antiretroviral therapy among people living with HIV. N Am J Med Sci 2013 Mar;5(3):220-223. [doi: 10.4103/1947-2714.109196] [Medline: 23626959]
3. Jalil CM, Jalil EM, Hoagland B, et al. The rising tide of HIV among young men who have sex with men in Brazil: insights from the Conectad@s study. Lancet Reg Health Am 2024 Aug;36:100798. [doi: 10.1016/j.lana.2024.100798] [Medline: 38978784]
4. Torres TS, Teixeira SLM, Hoagland B, et al. Recent HIV infection and annualized HIV incidence rates among sexual and gender minorities in Brazil and Peru (ImPrEP seroincidence study): a cross-sectional, multicenter study. Lancet Reg Health Am 2023 Dec;28:100642. [doi: 10.1016/j.lana.2023.100642] [Medline: 38076411]
5. Castro R, Ribeiro-Alves M, Corrêa RG, et al. The men who have sex with men HIV care cascade in Rio de Janeiro, Brazil. PLoS ONE 2016;11(6):e0157309. [doi: 10.1371/journal.pone.0157309] [Medline: 27299524]
6. Costa JDM, Torres TS, Coelho LE, Luz PM. Adherence to antiretroviral therapy for HIV/AIDS in Latin America and the Caribbean: systematic review and meta-analysis. J Int AIDS Soc 2018 Jan;21(1):e25066. [doi: 10.1002/jia2.25066] [Medline: 29356390]
7. Haas AD, Lienhard R, Didden C, et al. Mental health, ART adherence, and viral suppression among adolescents and adults living with HIV in South Africa: a cohort study. AIDS Behav 2023 Jun;27(6):1849-1861. [doi: 10.1007/s10461-022-03916-x] [Medline: 36592251]
8. Meireles MV, Pascom ARP, Duarte EC. Factors associated with early virological response in HIV-infected individuals starting antiretroviral therapy in Brazil (2014–2015): results from a large HIV surveillance cohort. J Acquir Immune Defic Syndr 2018 Aug 1;78(4):e19-e27. [doi: 10.1097/QAI.0000000000001684] [Medline: 29557856]
9. Jones D, Cook R, Rodriguez A, Waldrop-Valverde D. Personal HIV knowledge, appointment adherence and HIV outcomes. AIDS Behav 2013 Jan;17(1):242-249. [doi: 10.1007/s10461-012-0367-y] [Medline: 23143751]
10. Buh A, Deonandan R, Gomes J, Krentel A, Oladimeji O, Yaya S. Barriers and facilitators for interventions to improve ART adherence in Sub-Saharan African countries: a systematic review and meta-analysis. PLoS ONE 2023;18(11):e0295046. [doi: 10.1371/journal.pone.0295046] [Medline: 38032918]
11. Campbell L, Masquillier C, Thunnissen E, et al. Social and structural determinants of household support for art adherence in low- and middle-income countries: a systematic review. Int J Environ Res Public Health 2020 May 27;17(11):3808. [doi: 10.3390/ijerph17113808] [Medline: 32471153]
12. Glasgow RE. Perceived barriers to self-management and preventive behaviors. United States National Cancer Institute (NIH). 2020. URL: https://cancercontrol.cancer.gov/sites/default/files/2020-06/barriers.pdf [accessed 2024-04-18]
13. Kall M, Marcellin F, Harding R, Lazarus JV, Carrieri P. Patient-reported outcomes to enhance person-centred HIV care. Lancet HIV 2020 Jan;7(1):e59-e68. [doi: 10.1016/S2352-3018(19)30345-5] [Medline: 31776101]

XSL•FO
RenderX

14. Tripp MK, Diamond PM, Vernon SW, Swank PR, Dolan Mullen P, Gritz ER. Measures of parents' self-efficacy and perceived barriers to children's sun protection: construct validity and reliability in melanoma survivors. Health Educ Res 2013 Oct;28(5):828-842. [doi: 10.1093/her/cys114] [Medline: 23204537]
15. de Vries AE, van der Wal MHL, Nieuwenhuis MMW, et al. Perceived barriers of heart failure nurses and cardiologists in using clinical decision support systems in the treatment of heart failure patients. BMC Med Inform Decis Mak 2013 Apr 26;13(1):54. [doi: 10.1186/1472-6947-13-54] [Medline: 23622342]
16. Almeida-Brasil CC, Nascimento E, Oliveira Costa J, Silveira MR, Fátima Bonolo P, Braga Ceccato MG. Desenvolvimento e validação do conteúdo da escala de percepções de dificuldades com o tratamento antirretroviral [Article in Portuguese]. Rev Med Minas Gerais 2016;26(5):S56-S64.
17. Maggiolo F, Airoldi M, Kleinloog HD, et al. Effect of adherence to HAART on virologic outcome and on the selection of resistance-conferring mutations in NNRTI- or PI-treated patients. HIV Clin Trials 2007;8(5):282-292. [doi: 10.1310/hct0805-282] [Medline: 17956829]
18. Anglemyer A, Rutherford GW, Horvath T, et al. Antiretroviral therapy for prevention of HIV transmission in HIV-discordant couples. Cochrane Database of Systematic Reviews 2013;2013(4). [doi: 10.1002/14651858.CD009153.pub3]
19. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 2006 Oct 11;4(1):79. [doi: 10.1186/1477-7525-4-79] [Medline: 17034633]
20. Berger BE, Ferrans CE, Lashley FR. Measuring stigma in people with HIV: psychometric assessment of the HIV stigma scale. Res Nurs Health 2001 Dec;24(6):518-529. [doi: 10.1002/nur.10011]
21. Monahan PO, Shacham E, Reece M, et al. Validity/reliability of PHQ-9 and PHQ-2 depression scales among adults living with HIV/AIDS in western Kenya. J GEN INTERN MED 2009 Feb;24(2):189-197. [doi: 10.1007/s11606-008-0846-z] [Medline: 19031037]
22. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the alcohol use disorders identification test (audit): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun;88(6):791-804. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
23. Cortes A, Hunt N, McHale S. Development of the scale of perceived social support in HIV (PSS-HIV). AIDS Behav 2014 Dec;18(12):2274-2284. [doi: 10.1007/s10461-014-0902-0] [Medline: 25245475]
24. Wallston KA, Osborn CY, Wagner LJ, Hilker KA. The perceived medical condition self-management scale applied to persons with HIV/AIDS. J Health Psychol 2011 Jan;16(1):109-115. [doi: 10.1177/1359105310367832] [Medline: 20656769]
25. Mills EJ, Nachega JB, Bangsberg DR, et al. Adherence to HAART: a systematic review of developed and developing nation patient-reported barriers and facilitators. PLoS Med 2006 Nov;3(11):e438. [doi: 10.1371/journal.pmed.0030438] [Medline: 17121449]
26. Almeida-Brasil CC, Nascimento ED, Silveira MR, Bonolo PDF, Ceccato MDG. New patient-reported outcome measure to assess perceived barriers to antiretroviral therapy adherence: the PEDIA scale. Cad Saude Publica 2019;35(5):e00184218. [doi: 10.1590/0102-311x00184218] [Medline: 31166419]
27. Matos VC, Torres TS, Luz PM. Adherence to antiretroviral therapy among cisgender gay, bisexual and other men who have sex with men in Brazil: Evaluating the role of HIV-related stigma dimensions. In: Demant D, editor. PLoS ONE 2024;19(8):e0308443. [doi: 10.1371/journal.pone.0308443] [Medline: 39116156]
28. Silva KRO, Ferreira RC, Coelho LE, et al. Knowledge of HIV transmission, prevention strategies and U = U among adult sexual and gender minorities in Brazil. J Int AIDS Soc 2024 Feb;27(2):e26220. [doi: 10.1002/jia2.26220] [Medline: 38379186]
29. Meade AW, Craig SB. Identifying careless responses in survey data. Psychol Methods 2012 Sep;17(3):437-455. [doi: 10.1037/a0028085] [Medline: 22506584]
30. Blair KJ, Torres TS, Hoagland B, et al. Pre-exposure prophylaxis use, HIV knowledge, and internalized homonegativity among men who have sex with men in Brazil: a cross-sectional study. Lancet Reg Health Am 2022 Feb;6:100152. [doi: 10.1016/j.lana.2021.100152] [Medline: 35291206]
31. Ferreira RC, Torres TS, Marins LMS, Ceccato MDG, Bezerra DRB, Luz PM. HIV knowledge and its correlation with the Undetectable = Untransmittable slogan in Brazil. Rev Saude Publica 2022;56:87. [doi: 10.11606/s1518-8787.2022056004168] [Medline: 36228232]
32. Ferreira RC, Torres TS, Ceccato M, et al. Development and evaluation of short-form measures of the HIV/AIDS knowledge assessment tool among sexual and gender minorities in Brazil: cross-sectional study. JMIR Public Health Surveill 2022 Mar 29;8(3):e30676. [doi: 10.2196/30676] [Medline: 35348470]
33. Torres TS, De Boni RB, de Vasconcellos MT, et al. Awareness of prevention strategies and willingness to use preexposure prophylaxis in Brazilian men who have sex with men using apps for sexual encounters: online cross-sectional study. JMIR Public Health Surveill 2018 Jan 22;4(1):e11. [doi: 10.2196/publichealth.8997] [Medline: 29358160]

XSL•FO
RenderX

34. Torres TS, Luz PM, De Boni RB, et al. Factors associated with PrEP awareness according to age and willingness to use HIV prevention technologies: the 2017 online survey among MSM in Brazil. AIDS Care 2019 Oct;31(10):1193-1202. [doi: 10.1080/09540121.2019.1619665] [Medline: 31122033]
35. Torres TS, Konda KA, Vega-Ramirez EH, et al. Factors associated with willingness to use pre-exposure prophylaxis in Brazil, Mexico, and Peru: web-based survey among men who have sex with men. JMIR Public Health Surveill 2019 Jun 17;5(2):e13771. [doi: 10.2196/13771] [Medline: 31210143]
36. Social media stats Brazil Jun 2023 – Jun 2024. Statcounter Global Stats. URL: https://gs.statcounter.com/social-media-stats/all/brazil [accessed 2024-07-11]
37. Vale FC, Santa-Helena ED, Santos MA, et al. Development and validation of the WebAd-Q Questionnaire to monitor adherence to HIV therapy. Rev Saude Publica 2018;52:62. [doi: 10.11606/s1518-8787.2018052000337] [Medline: 29846437]
38. Torres TS, Harrison LJ, La Rosa AM, et al. Poor quality of life and incomplete self-reported adherence predict second-line ART virological failure in resource-limited settings. AIDS Care 2021 Oct;33(10):1340-1349. [doi: 10.1080/09540121.2021.1874275] [Medline: 33487029]
39. Parker RA, Rabideau DJ, Sax PE, et al. Impact of medication adherence on virologic failure in A5202: a randomized, partially blinded, phase 3B study. Clin Infect Dis 2017 Jun 1;64(11):1612-1614. [doi: 10.1093/cid/cix176] [Medline: 28329243]
40. Howard MC. A review of exploratory factor analysis decisions and overview of current practices: what we are doing and how can we improve? Int J Hum Comput Interact 2016 Jan 2;32(1):51-62. [doi: 10.1080/10447318.2015.1087664]
41. Li CH. Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. Behav Res 2016 Sep;48(3):936-949. [doi: 10.3758/s13428-015-0619-7]
42. Timmerman ME, Lorenzo-Seva U. Dimensionality assessment of ordered polytomous items with parallel analysis. Psychol Methods 2011 Jun;16(2):209-220. [doi: 10.1037/a0023353] [Medline: 21500916]
43. Lorenzo-Seva U, Ferrando PJ, Department of Psychology, Universitat Rovira i Virgili, Spain, Department of Psychology, Universitat Rovira i Virgili, Spain. Robust Promin: a method for diagonally weighted factor rotation. Liberabit 2019 Jun 27;25(1):99-106. [doi: 10.24265/liberabit.2019.v25n1.08]
44. Ford JK, MacCALLUM RC, Tait M. The application of exploratory factor analysis in applied psychology: a critical review and analysis. Pers Psychol 1986 Jun;39(2):291-314. [doi: 10.1111/j.1744-6570.1986.tb00583.x]
45. Hinkin TR. A review of scale development practices in the study of organizations. J Manage 1995;21(5):967-988. [doi: 10.1016/0149-2063(95)90050-0]
46. Raykov T. Introduction to Psychometric Theory: Routledge; 2011. [doi: 10.4324/9780203841624]
47. Reise SP, Widaman KF, Pugh RH. Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. Psychol Bull 1993 Nov;114(3):552-566. [doi: 10.1037/0033-2909.114.3.552] [Medline: 8272470]
48. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
49. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. Am J Health Syst Pharm 2008 Dec 1;65(23):2276-2284. [doi: 10.2146/ajhp070364] [Medline: 19020196]
50. Hair JF. Multivariate Data Analysis, 7th edition: Pearson Prentice Hall; 2010.
51. Reichenheim M, Bastos JL. O quê, para quê e como? Desenvolvendo instrumentos de aferiçãO em epidemiologia [Article in Portuguese]. Rev saúde pública 2021 Aug 9;55:40. [doi: 10.11606/s1518-8787.2021055002813]
52. Lorenzo-Seva U, Ferrando PJ. FACTOR: A computer program to fit the exploratory FACTOR analysis model. Behav Res Methods 2006 Feb;38(1):88-91. [doi: 10.3758/bf03192753] [Medline: 16817517]
53. Muthén LK, Muthén BO. Mplus user's guide. Mplus. 2017. URL: https://www.statmodel.com/download/usersguide/MplusUserGuideVer_8.pdf [accessed 2025-06-11]
54. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2023. URL: https://www.R-project.org/ [accessed 2025-06-11]
55. Wagener MN, van Opstal SEM, Miedema HS, van Gorp ECM, Roelofs P. Work-related stigma and disclosure: a daily challenge for people living with HIV. WOR 2017 Dec 13;58(4):537-548. [doi: 10.3233/WOR-172650]
56. Sin NL, DiMatteo MR. Depression treatment enhances adherence to antiretroviral therapy: a meta-analysis. Ann Behav Med 2014 Jun;47(3):259-269. [doi: 10.1007/s12160-013-9559-6] [Medline: 24234601]
57. Wagner GJ, Ghosh-Dastidar B, Mukasa B, Linnemayr S. Changes in ART adherence relate to changes in depression as well! evidence for the bi-directional longitudinal relationship between depression and ART adherence from a prospective study of HIV clients in Uganda. AIDS Behav 2020 Jun;24(6):1816-1824. [doi: 10.1007/s10461-019-02754-8] [Medline: 31813077]
58. Igartua KJ, Gill K, Montoro R. Internalized homophobia: a factor in depression, anxiety, and suicide in the gay and lesbian population. Can J Commun Ment Health 2003;22(2):15-30. [doi: 10.7870/cjcmh-2003-0011] [Medline: 15868835]
59. Lee H, Operario D, Yi H, Choo S, Kim SS. Internalized homophobia, depressive symptoms, and suicidal ideation among lesbian, gay, and bisexual adults in South Korea: an age-stratified analysis. LGBT Health 2019;6(8):393-399. [doi: 10.1089/lgbt.2019.0108] [Medline: 31746660]

XSL•FO
RenderX

60. Yolaç E, Meriç M. Internalized homophobia and depression levels in LGBT individuals. Perspect Psychiatr Care 2021 Jan;57(1):304-310. [doi: 10.1111/ppc.12564] [Medline: 32557669]
61. van der Kooij YL, den Daas C, Bos AER, Willems RA, Stutterheim SE. Correlates of internalized HIV stigma: a comprehensive systematic review. AIDS Educ Prev 2023 Apr;35(2):158-172. [doi: 10.1521/aeap.2023.35.2.158] [Medline: 37129595]
62. Menéndez-Arias L, Delgado R. Update and latest advances in antiretroviral therapy. Trends Pharmacol Sci 2022 Jan;43(1):16-29. [doi: 10.1016/j.tips.2021.10.004] [Medline: 34742581]
63. Gerend MA, Shepherd MA, Shepherd JE. The multidimensional nature of perceived barriers: global versus practical barriers to HPV vaccination. Health Psychol 2013 Apr;32(4):361-369. [doi: 10.1037/a0026248] [Medline: 22059622]
64. Kerr L, Kendall C, Guimarães MDC, et al. HIV prevalence among men who have sex with men in Brazil: results of the 2nd national survey using respondent-driven sampling. Medicine (Baltimore) 2018 May;97(1S Suppl 1):S9-S15. [doi: 10.1097/MD.0000000000010573] [Medline: 29794604]
65. Matos VC, Torres TS, Luz PM. Análise fatorial confirmatória da escala PEDIA de avaliação de barreiras percebidas para a adesão à terapia antirretroviral (PEDIA) [Web page in Portuguese]. Proceedings.science. 2022. URL: https://proceedings.science/abrascao-2022/trabalhos/analise-fatorial-confirmatoria-da-escala-pedia-de-avaliacao-de-barreiras-percebi [accessed 2024-04-18]

Abbreviations

ART: antiretroviral therapy
AVE: average variance extracted
CFA: confirmatory factor analysis
CFI: comparative fit index
EFA: exploratory factor analysis
GSN: geosocial networking
KMO: Kaiser-Meyer-Olkin
MSM: men who have sex with men
PA: parallel analysis
PEDIA: perceived barriers to antiretroviral therapy adherence scale
PEDIAr: reduced version of the perceived barriers to antiretroviral therapy adherence scale
PHIVSMS: Perceived HIV Self-Management Scale
PSS-HIV: Perceived Social Support in HIV
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TGNB: transgender and nonbinary
TLI: Tucker-Lewis Index
VAS: visual analog scale
WLSMV: weighted least squares estimator with a diagonal weight matrix and mean- and variance-adjusted $\chi 2$ test statistic

Edited by GJB Sousa; submitted 29.09.24; peer-reviewed by M Alagi, MS Torres Silva, Z Yang; revised version received 23.03.25; accepted 20.04.25; published 27.06.25.

Please cite as:
Matos VC, Araújo DC, Torres TS, Luz PM
Validation of the Perceived Barriers to Antiretroviral Therapy Adherence (PEDIA) Scale Among Gay, Bisexual, and Other Men Who Have Sex With Men and Transgender and Nonbinary Persons: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e67005
URL: https://publichealth.jmir.org/2025/1/e67005
doi:10.2196/67005

© Victor C Matos, Dyego Carlos Araújo, Thiago S Torres, Paula M Luz. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Intention to Use Pre-Exposure Prophylaxis Among Age Groups of Brazilian Men Who Have Sex With Men: National Cross-Sectional Study

Alvaro Francisco Lopes de Sousa[1,2], PhD; Caique Jordan Nunes Ribeiro[3], PhD; Guilherme Reis de Santana Santos[3], BSc; Emerson Lucas Silva Camargo[4], MSc; Herica Emilia Félix de Carvalho[5], PhD; Guilherme Schneider[4], MSc; Leticia Genova Vieira[4], MSc; Liliane Moretti Carneiro[6], PhD; Ana Paula Morais Fernandes[1,4], PhD; Talita Morais Fernandes[7], PhD; Márcio Bezerra-Santos[8], PhD; Rita de Cassia Dias Nascimento[9], RN; Lucas Almeida Andrade[10], MSc; Anderson Reis de Sousa[9], PhD; Inês Fronteira[1], PhD; Isabel Amélia Costa Mendes[4], PhD

[1]Public Health Research Centre, Comprehensive Health Research Center, CHRC, REAL, CCAL, NOVA National School of Public Health, Universidade Nova de Lisboa, Avenida Padre Cruz, Lisbon, Portugal

[10]Graduate Program in Health Sciences, Universidade Federal de Sergipe, Aracaju, Brazil

[2]Instituto de Ensino e Pesquisa, Hospital Sírio-Libanês, São Paulo, Brazil

[3]››Graduate Program in Nursing, Universidade Federal de Sergipe, São Cristovão, Brazil

[4]Ribeirão Preto College of Nursing, Universidade de São Paulo, RIbeirão Preto, Brazil

[5]Universidade Estadual do Maranhão, Campus Coroatá, Maranhão, Brazil

[6]Program in Health and Development in the Central-West Region, Universidade Federal de Mato Grosso do Sul, Campo Grande, Brazil

[7]Ribeirão Preto Faculty of Medicine, Universidade de São Paulo, Ribeirão Preto, Brazil

[8]Medical and Nursing Science Complex, Universidade Federal de Alagoas, Arapiraca, Brazil

[9]Nursing College, Universidade Federal da Bahia, Salvador, Brazil

Corresponding Author:
Alvaro Francisco Lopes de Sousa, PhD
Public Health Research Centre, Comprehensive Health Research Center, CHRC, REAL, CCAL, NOVA National School of Public Health, Universidade Nova de Lisboa, Avenida Padre Cruz, Lisbon, Portugal

Abstract

Background: In Brazil, men who have sex with men (MSM) are disproportionately affected by the HIV epidemic, mirroring global trends. Despite advancements in HIV prevention, such as pre-exposure prophylaxis (PrEP), uptake remains uneven among different MSM age groups, influenced by various sexual behaviors, risk perceptions, and social stigmas.

Objective: This study aimed to investigate factors associated with PrEP intentions among Brazilian MSM across different age groups.

Methods: A cross-sectional electronic survey was conducted with Brazilian MSM aged 18 years and above.

Results: PrEP intentions were expressed by 55.1% (2390/4341) of participants. Among those under 25 years old, frequent barebacking was associated with a 25.4-fold higher prevalence of PrEP intentions (95% CI 18.18 - 35.48). Conversely, having one or more casual sexual partners in the last 30 days (2164/3838) was associated with a 22% lower frequency of PrEP intentions (95% CI 0.72 - 0.85). The practice of chemsex was also associated with a lower prevalence of PrEP intentions (adjusted prevalence ratio [aPR] 0.96; 95% CI 0.94 - 0.98). Among individuals aged 25 - 49 years, the practice of double penetration was associated with a higher prevalence of PrEP intentions (aPR 1.22; 95% CI .13 - 1.32), as was being a receptive or versatile partner (aPR 1.27; 95% CI 1.04 - 1.55 and aPR 1.23; 95% CI 1.01 - 1.49). In addition, within this age range, a higher prevalence of PrEP intentions (over 10% [258/429]) was found among those with specific characteristics, such as a previous diagnosis of syphilis (aPR 1.12; 95% CI 1.03 - 1.22). However, being single (aPR 0.90; 95% CI 0.83 - 0.97) and having a partner who used PrEP (aPR 0.82; 95% CI 0.72 - 0.93) were associated with a lower prevalence of PrEP intentions. For MSM aged ≥50 years, a higher prevalence of PrEP intentions was associated with the practice of double penetration (aPR 1.31; 95% CI 1.20 - 1.43), as well as being a receptive (aPR 1.27; 95% CI 1.04 - 1.55) or versatile partner (aPR 1.23; 95% CI 1.01 - 1.50). The practice of group sex was independently associated with a lower prevalence of PrEP intentions exclusively among MSM aged ≥50 years (aPR 0.82; 95% CI 0.75 - 0.89).

Conclusions: The study highlights significant generational differences in factors influencing PrEP intentions among Brazilian MSM. It underscores the need for tailored HIV prevention strategies that consider the unique behaviors and perceptions of different

XSL•FO
RenderX

age groups. By addressing these nuances, public health initiatives can more effectively promote PrEP use, catering to the diverse needs of the MSM community and contributing to the reduction of HIV infection rates.

(JMIR Public Health Surveill 2025;11:e58405) doi:10.2196/58405

KEYWORDS

age groups; pre-exposure prophylaxis; HIV; men's health; sexual and gender minorities; PrEP; men who have sex with men; MSM; sexual behaviors; preventive practices; Brazil; HIV prevention; epidemic; global trends; social stigma; cross-sectional study; electronic survey; prevalence; syphilis; infections; sexually transmitted disease

Introduction

Men who have sex with men (MSM) experience a disproportionately high burden of HIV infection compared to the general population and even other groups vulnerable to the virus [1-4]. Since the beginning of the HIV epidemic in the 1980s, MSM have been recognized as one of the most affected populations due to factors such as social stigma, discrimination, barriers to accessing health care services, and higher-risk sexual behaviors. These longstanding challenges have made HIV a significant public health concern among MSM, despite advances in infection prevention and treatment [3,5].

According to the World Health Organization (WHO), in 2022, approximately 39 million people were living with HIV worldwide, with 29.8 million receiving antiretroviral therapy, contributing to a 62% reduction in the AIDS mortality rate since 2000 [6]. However, treatment does not fully meet the needs of the MSM population, highlighting the need to target Antiretroviral Therapy across different age groups to reduce new HIV infections [7].

The number of individuals infected with HIV varies based on age and sexual behavior. In Brazil, approximately 1 million people live with HIV, with a growing proportion of infections among those aged 13 years to 39 years, particularly among MSM (526/1.000, 52.6%). However, the prevalence within this group is not uniform and varies by age group. According to official Brazilian reports, the HIV prevalence among younger MSM (aged 13 - 24 y) is 24% (240/1.000), among young adults (aged 25 - 34 y) it is 40% (400/1.000), and among those over 35 years old, it stands at 17% (170/1.000). These rates are significantly higher than those observed in the general Brazilian population within the corresponding age groups, at 0.3%, 0.7%, and 0.8%, respectively [8].

The differences in HIV/AIDS prevalence among MSM of various age groups can be attributed to several factors. Understanding how sexual behaviors vary with age is a critical aspect of epidemiology in this population, as it significantly impacts HIV exposure risk [9-13].

Older MSM are more likely to exhibit risk factors for HIV infection, which may include, in addition to cumulative exposure over time, increased stigma and discrimination that can impede access to healthcare services and HIV-related information, including testing, antiretroviral treatment, and condom use [14-17]. This population continues to represent a substantial portion of those living with HIV. According to WHO data, in 2022, 22% of people living with HIV worldwide were MSM, with 17% of them aged 35 years or older [6]. In Brazil, the proportion of HIV diagnoses among individuals aged 50 years and older increased from 9.3% to 12.7% between 2009 and 2019 and continues to rise each year [16].

On the other hand, young MSM are more prone to engaging in risk behaviors due to new forms of relationships and sexual practices that increase their vulnerability to HIV acquisition and other sexually transmitted infections (STIs). One example is chemsex, the practice of using drugs to enhance sexual experiences, which has gained popularity among younger MSM. Chemsex is associated with prolonged sexual encounters, reduced inhibition, and less consistent condom use, all of which contribute to higher risks of HIV transmission [18-20]. In addition, younger MSM are more likely to engage in casual relationships, often facilitated by dating apps, leading to multiple sexual partners and further increasing their likelihood of HIV exposure [21,22].

HIV prevention efforts, such as pre-exposure prophylaxis (PrEP), often focus predominantly on young MSM, while there is limited literature on the subject regarding older MSM [15,23]. PrEP has been shown to be effective in preventing HIV among MSM of all ages. For instance, a study conducted in the United States demonstrated that PrEP is equally effective among young MSM (aged 18 - 25 y) and older MSM (aged 26 - 50 y) [24]. Consequently, HIV prevention efforts, including PrEP, should be expanded to encompass older age groups as well.

The objective of this study was to investigate factors associated with PrEP intention among Brazilian MSM of different age groups. We hypothesized that there are differences in the factors associated with PrEP intention among MSM of different age groups residing in Brazil.

Methods

Study Design

This is a cross-sectional online survey study that includes MSM throughout Brazil, aged 18 years and older, from January 2020 to May 2021.

Population and Sample

A sample size calculation was performed using G*Power software (version 3.1.9.7; Heinrich Heine University Düsseldorf) to determine the sample size needed for the study. The calculation considered the male population in Brazil, a tolerable SE of 3% , a confidence level of 95%, and a presumed outcome prevalence of 50%.

It is noteworthy that when it is not possible to determine the true prevalence of the phenomenon of interest in local or

XSL•FO
RenderX

national literature, a presumed prevalence of 50% is used, as this value maximizes the sample and its use does not pose a problem [25,26].

Participants who identified as men (cis or trans), had sexual relations with another man in the last 12 months, and were 18 years or older were included.

Individuals diagnosed with HIV infection and those already using PrEP were excluded from the study.

Data Collection Procedures

To recruit survey participants, we used the snowball sampling technique adapted for the virtual environment, a method previously validated for the MSM population [9,27]. In this approach, participants help recruit other individuals with similar characteristics through their social networks and contacts by sending them an invitation to participate. Given that our study was conducted in a country of continental dimensions, we intentionally selected the first 30 participants with diverse social and demographic characteristics—including age group, region or district of residence, race or skin color (White or non-White), income, and education—to enhance the generalizability of our data [9,27,28].

To reach these initial 30 participants without requiring researchers to travel across all regions of Brazil, 2 of our cisgender male researchers, who identify as MSM and were thoroughly trained and calibrated with one another, created public profiles on 2 geolocation-based dating apps, Grindr and Hornet, to identify the "seeds."

The decision to use Grindr and Hornet was based on their widespread recognition as popular platforms within the MSM community, offering a broad user base. In addition, their geolocation features facilitated a more targeted and efficient approach to identifying participants in various regions of Brazil. Another significant factor was the familiarity and comfort many MSM community members already have with these platforms, as they are commonly used for social interaction and relationship-seeking.

Participants were offered the option to maintain anonymity until they felt comfortable disclosing their identity to the researchers, fostering a safe and welcoming environment for study participation.

This innovative strategy not only optimized the recruitment process but also demonstrated our commitment to inclusion and cultural sensitivity, which are crucial for the success of the research.

Within these apps, researchers approached the first available online individuals who met the study's inclusion criteria and satisfied the criteria for seed generalization, as recommended by previous studies [29,30]. After engaging with each seed, we provided a link to participate in the study and instructed them to invite other MSM from their social networks until reaching the required sample size through a dissemination strategy.

Data Collection Instruments

The survey form was hosted on the SurveyMonkey platform for data collection. This platform incorporates security features that permit only one response per IP address. To ensure the validity of the survey form, a content and face validation process was conducted with 10 expert judges in the field. In addition, a pretest involving 10 participants was carried out to assess the clarity and effectiveness of the form.

The study's dependent variable was the intention to use PrEP, with outcomes categorized as yes or no. The explanatory variables included sociodemographic data such as age, education level, place of residence, gender identity, and relationship status. Participants were also queried on their practices related to disclosing HIV status on dating apps, their knowledge of PrEP, and the "undetectable=untransmittable" concept, as well as their previous diagnoses of STIs and history of HIV testing, either in their lifetime or within the past 12 months.

Subsequent sections of the instrument examined variables associated with sexual practices, including preventive measures taken during sex to avoid HIV transmission, sexual activity with a partner living with HIV, involvement in group sex (defined as sexual encounters involving 3 or more people), self-assessment of perceived HIV infection risk, preferred sexual positions, involvement in receptive anal sex without a condom, recent engagement in unprotected sex, drug use during sexual encounters or sex under the influence of drugs within the last 6 months, sexual activity in public places, and specific practices such as bareback sex, double penetration, or fisting. The frequency with which participants sought health care services was also assessed.

We provided clear and accessible definitions for uncommon or potentially unfamiliar terms such as "Cruising" [31], "Double Penetration" (DP) [9,19], "Chemsex" [29], and "Gouinage" [15,32] to ensure participants' understanding. For the term "Chemsex," participants were asked if they had consumed illicit drugs immediately before and during sexual intercourse within the past 12 months. Those who responded affirmatively were presented with a multiple-choice list to indicate specific drugs consumed [29].

The practices were defined based on previous studies [33-35], that are (1) DP: simultaneous sexual penetration by two or more penises. (2) Fisting or footing: anal penetration using the fist or foot. (3) Cruising: free, consensual, and anonymous sex practiced between men in public spaces such as parks, forests, beaches, or parking lots. (4) Gouinage: sexual activity without penetration.

Data Analysis

Statistical analyses were performed using IBM SPSS 27.0 software (SPSS Inc). An initial exploratory analysis described the distribution of predictor variables across the study's age groups of interest (<25 y, 25 - 49 y, and ≥50 y). Data were presented in absolute frequencies, and percentage frequencies were calculated based on column totals, with denominators representing the subsamples from each age group. Associations between predictor variables and age groups were assessed using Pearson χ^2 test with the Monte Carlo permutation method (999 simulations and a 95% confidence level), with a significance level of 5% (P value <.05).

XSL•FO
RenderX

Bivariate analyses were then conducted for each age group. Results were expressed in absolute frequencies, with percentage frequencies calculated based on row totals of the contingency table, using denominators representing the fraction of exposed and nonexposed categories for each predictor variable. Variables for inclusion in the multivariate model were selected using Pearson χ^2 test with the Monte Carlo permutation method (999 simulations and a 95% confidence level), applying a statistical criterion of *P* value <.20 at this stage. Crude prevalence ratios (PRs) with their respective 95% CIs were calculated to measure the strength of association between the outcome (intention to use PrEP among MSM aged <25 y, 25 - 49 y, and ≥50 y) and predictor variables, including social characteristics, sexual partnerships, forms of HIV prevention adopted, reasons for not using condoms in recent relationships, sexual practices, STI history, and willingness to use PrEP, given that the prevalence of PrEP intention exceeded 10%.

The final stage involved multivariate modeling to identify factors independently associated with the intention to use PrEP within each age group. We used the generalized linear model with a Poisson distribution and a log-linear link function. Adherence to the Poisson distribution was tested using the Kolmogorov-Smirnov test (*P* value >.05). In additionally, the equidispersion assumption was evaluated by comparing the variance and the mean of each outcome.

To compute adjusted prevalence ratios (aPRs) and their corresponding 95% CI, we used a hybrid parameter estimation method with a robust variance estimator and Type III analysis to test model effects. The omnibus test was used to evaluate whether the final multivariate model provided a better explanation of the prevalence of PrEP intention compared to a model that only included the intercept, with a significance level of 5% (*P* value <.05). The Akaike Information Criterion, deviance, and log-likelihood parameters served as criteria for selecting the best-fitting model, with lower values indicating better fit. The significance of aPRs for variables in the final model was determined using the Wald χ^2 test, with variables having a *P* value <.05 considered significant.

Ethical Considerations

The project obtained approval from the Ethics Committee of the Ribeirão Preto School of Nursing, Brazil, under opinion number 4,163,084. An online Informed Consent Form (was provided and the platform used to host the collected data allowed for anonymization of all participants. All participants provided informed consent electronically before initiating the survey.

The online data collection platform ensured the anonymity of participants through mechanisms that prevented the identification of IP addresses and other traceable information. No identifying personal data were collected. Responses were de-identified at the source, and all data were stored on secure, password-protected servers with restricted access.

Following the completion of the study, participants received HIV prevention information and were directed to institutional websites for additional details on HIV/AIDS prevention. For individuals interested in PrEP usage who provided their contact email, a list of centers offering PrEP consultations in their state/region was supplied.

No monetary or material compensation was offered to participants. Participation was entirely voluntary and unpaid.

Results

We recorded 8620 accesses to the survey form, of which 4341 responses or participants met the study's inclusion criteria. Table 1 presents the baseline characteristics of the participants, who were predominantly cisgender (4216/4341, 97.1%), with ≥12 years of education (3298/4341, 76.0%), and native to Brazil (3686/4341; 84.9%). The majority identified as homosexual (3553/4341; 81.8%), were single (3000/4341; 69.1%), and reported having both fixed and casual partners (4010/4341; 92.4%). In addition, most had at least one casual sexual partner in the last 30 days (3838/4341; 88.4%) and reported meeting 1 to 3 casual sexual partners through apps in the past 6 months (2018/4341; 46.5%).

XSL•FO
RenderX

Table . Bivariate analysis of social, demographic and behavioral factors associated with intention to use PrEP among MSM by age group, a cross-sectional study, 2020 - 2021, Brazil.

Variables	Age group											
	<25 years (n=1702)				25 - 49 years (n=2238)				≥50 years (n=401)			
	Intention to use PrEP											
	Yes (n=925)	No (n=777)	PR (95% CI)	*P* value	Yes (n=1233)	No (n=1005)	PR (95% CI)	*P* value	Yes (n=232), n (%)	No (n=169), n (%)	PR (95% CI)	*P* value
	n (%)	n (%)			n (%)	n (%)			n (%)	n (%)		
Social characteristics												
Gender identity				.07[a]				.15[a]				.66
Cisgender [ref]	911 (54.3)	766 (45.7)	—		1186 (55.2)	961 (44.8)	—		227 (57.9)	165 (42.1)	—	
Transgender	9 (81.8)	2 (18.2)	1.51 (1.14 - 2.00)		25 (62.5)	15 (37.5)	1.13 (0.89 - 1.44)		4 (66.7)	2 (33.3)	1.15 (0.65 - 2.04)	
Other	5 (35.7)	9 (64.3)	0.66 (0.33 - 1.33)		22 (43.1)	29 (56.9)	0.78 (0.57 - 1.07)		1 (33.3)	2 (66.7)	0.58 (0.12 - 2.86)	
Schooling				.39				.04[a]				.33
< 12 years [ref]	278 (56.0)	218 (44.0)	0.96 (0.87 - 1.05)		250 (59.7)	169 (40.3)	0.91 (0.83 - 0.99)		79 (61.7)	49 (38.3)	0.91 (0.76 - 1.08)	
≥ 12 years	647 (53.6)	559 (46.4)			983 (54.0)	836 (46.0)			153 (56.0)	120 (44.0)		
Immigrant				.20				.008[a]				.13
Yes [ref]	178 (51.3)	169 (48.7)	1.07 (0.96 - 1.20)		44 (42.3)	60 (57.7)	1.32 (1.05 - 1.65)		126 (61.8)	78 (38.2)	0.87 (0.74 - 1.03)	
No	747 (55.1)	608 (44.9)			1189 (55.7)	945 (44.3)			106 (53.8)	91 (46.2)		
Partnerships												
Are you attracted to women?				.25				.96				.42
Yes	152 (51.4)	144 (48.6)	0.93 (0.83 - 1.05)		235 (55.3)	190 (44.7)	1.00 (0.91 - 1.09)		42 (62.7)	25 (37.3)	1.10 (0.90 - 1.35)	
No [ref]	773 (55.0)	633 (45.0)			998 (55.0)	815 (45.0)			190 (56.9)	144 (43.1)		
Type of relationship				.17[a]				.02[a]				.32
Single	614 (53.0)	544 (47.0)	—		834 (53.3)	732 (46.7)	0.89 (0.82 - 0.96)		153 (55.4)	123 (44.6)	0.87 (0.73 - 1.03)	
Steady [ref]	278 (57.9)	202 (42.1)	0.91 (0.83 - 1.01)		352 (60.2)	233 (39.8)	—		71 (64.0)	40 (36.0)	—	
Polygamous	33 (51.6)	31 (48.4)	0.89 (0.69 - 1.14)		47 (54.0)	40 (46.0)	0.90 (0.73 - 1.10)		8 (57.1)	6 (42.9)	0.89 (0.56 - 1.44)	
Type of sexual partnership				<.001[a]				.02[a]				>.99

XSL•FO
RenderX

Variables	Age group															
	<25 years (n=1702)						25 - 49 years (n=2238)						≥50 years (n=401)			
	Intention to use PrEP															
	Yes (n=925)		No (n=777)		PR (95% CI)	*P* value	Yes (n=1233)		No (n=1005)		PR (95% CI)	*P* value	Yes (n=232), n (%)		No (n=169), n (%)		PR (95% CI)	*P* value
	n (%)		n (%)				n (%)		n (%)				n (%)		n (%)			
Steady [ref]	96	67.6	46	32.4	0.79 (0.69 - 0.89)		104	64.2	58	35.8	0.85 (0.75 - 0.96)		16	59.3	11	40.7	0.97 (0.70 - 1.35)	
Casual	829	53.1	731	46.9			1129	54.4	947	45.6			216	57.8	158	42.2		
Preferred sexual position						.03[a]						.01[a]						.17[a]
Only oral [ref]	38	43.7	49	56.3	-		59	42.8	79	57.2	-		10	40.0	15	60.0	-	
Insertive (top)	90	49.7	91	50.3	1.14 (0.86 - 1.51)		132	53.7	114	46.3	1.25 (1.00 - 1.57)		20	50.0	20	50.0	1.25 (0.71 - 2.21)	
Receptive (bottom)	348	58.0	252	42.0	1.33 (1.04 - 1.70)		447	57.4	332	42.6	1.34 (1.10 - 1.64)		71	59.2	49	40.8	1.48 (0.89 - 2.45)	
Versatile	449	53.8	385	46.2	1.23 (0.96 - 1.58)		595	55.3	480	44.7	1.29 (1.06 - 1.58)		131	60.6	85	39.4	1.52 (0.93 - 2.48)	
Declare HIV sorology status in the apps						.01[a]						.22						.18[a]
Yes	77	45.6	92	54.4	0.82 (0.65 - 1.04)		162	58.7	114	41.3	1.07 (0.97 - 1.20)		75	63.0	44	37.0	1.13 (0.95 - 1.35)	
No [ref]	848	55.3	685	44.7			1071	54.6	891	55.4			157	55.7	125	44.3		
Number of partners in the last 30 days						.01[a]						.56						.17
None [ref]	145	62.8	86	37.2	-		140	58.3	100	41.7	-		23	71.9	9	28.1	-	
1	391	54.0	333	46.0	0.86 (0.76 - 0.97)		541	54.5	451	45.5	0.93 (0.83 - 1.05)		65	53.7	56	46.3	0.75 (0.57 - 0.98)	
≥2	389	52.1	358	47.9	0.83 (0.73 - 0.94)		552	54.9	454	45.1	0.94 (0.83 - 1.06)		144	58.1	104	41.9	0.81 (0.63 - 1.03)	
Number of partners from apps in the last 6 mo																		
None [ref]	162	56.8	123	43.2	-	.46	181	54.5	151	45.5	-	.83	37	56.1	29	43.9	-	.57
1 - 3	453	54.8	374	55.2	0.96 (0.86 - 1.09)		568	54.6	473	45.4	1.00 (0.89 - 1.12)		92	61.3	58	38.7	1.09 (0.85 - 1.40)	
≥4	310	52.5	280	47.5	0.92 (0.81 - 1.05)		484	56.0	381	44.0	1.03 (0.91 - 1.15)		103	55.7	82	44.3	0.99 (0.77 - 1.27)	
Forms of HIV prevention adopted.																		
Sexual abstinence																		

XSL•FO
RenderX

Variables	Age group																
	<25 years (n=1702)						25 - 49 years (n=2238)						≥50 years (n=401)				
	Intention to use PrEP																
	Yes (n=925)		No (n=777)		PR (95% CI)	*P* value	Yes (n=1233)		No (n=1005)		PR (95% CI)	*P* value	Yes (n=232), n (%)		No (n=169), n (%)		PR (95% CI)	*P* value
	n (%)		n (%)				n (%)		n (%)				n (%)		n (%)			
Yes [ref]	78	67.2	38	32.8	0.79 (0.69 - 0.91)	.01[a]	109	64.1	61	35.9	0.85 (0.75 - 0.95)	.02[a]	16	66.7	8	33.3	0.86 (0.64 - 1.15)	.40
No	847	53.4	739	46.6			1124	54.4	944	45.6			216	57.3	161	42.7		
To cum outside																		
Yes [ref]	305	63.1	178	36.9	0.81 (0.74 - 0.88)	<.001[a]	394	59.8	265	40.2	0.89 (0.82 - 0.96)	.004[a]	77	63.6	44	36.4	0.87 (0.73 - 1.03)	.15
No	620	50.9	599	49.1			839	53.1	740	46.9			155	55.4	125	44.6		
Being top																		
Yes [ref]	23	45.1	28	54.9	1.21 (0.89 - 1.65)	.18[a]	33	55.0	27	45.0	1.00 (0.79 - 1.26)	>.99	6	66.7	3	33.3	0.87 (0.54 - 1.38)	.74
No	902	54.6	749	45.4			1200	55.1	978	44.9			226	57.7	166	42.3		
Gouinage																		
Yes [ref]	8	100.0	-	-	0.54 (0.52 - 0.57)	.01[a]	6	100.0	-	-	0.55 (0.53 - 0.57)	.04[a]	2	66.7	1	33.3	0.87 (0.39 - 1.94)	>.99
No	917	54.1	777	45.9			1227	55.0	1,005	45.0			230	57.8	168	42.2		
Condom use																		
Yes [ref]	83	59.3	57	40.7	0.91 (0.79 - 1.05)	.22	116	54.2	98	45.8	1.02 (0.89 - 1.16)	.83	14	51.9	13	48.1	1.12 (0.77 - 1.63)	.55
No	842	53.9	720	46.1			1117	55.2	907	44.8			218	58.3	156	41.7		
Justification for not using condoms in recent relationships																		
Partner uses PrEP																		
Yes [ref]	78	67.2	38	32.8	0.79 (0.69 - 0.91)	.01[a]	109	64.1	61	35.9	0.85 (0.75 - 0.95)	.02[a]	16	66.7	8	33.3	0.86 (0.64 - 1.15)	.40
No	847	53.4	739	46.6			1124	54.4	944	45.6			21	57.3	161	42.7		
Known/repeat partner																		
Yes [ref]	620	50.9	599	49.1	1.24 (1.08 - 1.42)	<.001[a]	839	53.1	740	46.9	1.13 (1.04 - 1.22)	.004[a]	155	55.4	125	44.6	1.15 (0.97 - 1.36)	.15
No	305	63.1	178	36.9			394	59.8	265	40.2			77	63.6	44	36.4		
Good physical appearance																		
Yes [ref]	23	45.1	28	54.9	1.21 (0.89 - 1.65)	.18[a]	33	55.0	27	45.0	1.00 (0.79 - 1.26)	>.99	6	66.7	3	33.3	0.87 (0.54 - 1.38)	.74
No	902	54.6	749	45.4			1200	55.1	978	44.9			226	57.7	166	42.3		
Partner claims not to have an STI																		
Yes [ref]	106	60.6	69	39.4	0.89 (0.72 - 1.08)	.09[a]	125	59.0	87	41.0	0.93 (0.82 - 1.05)	.25	23	62.2	14	37.8	0.92 (0.71 - 1.21)	.61
No	819	53.6	708	46.4			1108	54.7	918	45.3			209	57.4	155	42.6		
Partner reports having recently been tested																		
Yes [ref]	830	53.9	711	46.1	1.10 (0.89 - 1.35)	.21	1113	54.8	919	45.2	1.06 (0.94 - 1.20)	.34	210	56.9	159	43.1	1.21 (0.94 - 1.55)	.26
No	95	59.0	66	41.0			120	58.3	86	41.7			22	68.8	10	31.3		
Sex intensity																		

XSL•FO
RenderX

Variables	Age group																
	<25 years (n=1702)						25 - 49 years (n=2238)						≥50 years (n=401)				
	Intention to use PrEP																
	Yes (n=925)		No (n=777)		PR (95% CI)	*P* value	Yes (n=1233)		No (n=1005)		PR (95% CI)	*P* value	Yes (n=232), n (%)		No (n=169), n (%)		PR (95% CI)	*P* value
	n (%)		n (%)				n (%)		n (%)				n (%)		n (%)			
Yes	222	64.0	125	36.0	1.23 (1.06 - 1.43)	<.001[a]	309	60.9	198	39.1	1.14 (1.05 - 1.24)	.003[a]	61	64.9	33	35.1	1.17 (0.97 - 1.39)	.12
No [ref]	703	51.9	652	48.1			924	53.4	807	46.6			171	55.7	136	44.3		
Sexually gifted partner																		
Yes	56	56.6	43	43.4	1.04 (0.80 - 1.37)	.65	100	54.1	85	45.9	0.98 (0.85 - 1.13)	.82	14	58.3	10	41.7	1.01 (0.71 - 1.43)	>.99
No [ref]	869	54.2	734	45.8			1133	55.2	920	44.8			218	57.8	159	42.2		
Sexual practices																		
The last person you had sex with was HIV+																		
Yes	25	62.5	15	37.5	1.15 (0.90 - 1.47)	.29	26	60.5	17	39.5	1.10 (0.86 - 1.40)	.54	7	87.5	1	12.5	1.53 (1.16 - 2.01)	.15
No [ref]	900	54.2	762	45.8			1207	55.0	988	45.0			225	57.3	168	42.7		
Group sex																		
Yes	327	42.3	446	57.7	0.66 (0.60 - 0.72)	<.001[a]	458	48.5	486	51.5	0.81 (0.75 - 0.88)	<.001[a]	81	46.8	92	53.2	0.71 (0.59 - 0.85)	<.001
No [ref]	598	64.4	331	35.6			775	59.9	519	40.1			151	66.2	77	33.8		
Frequent barebacking																		
Yes	893	99.9	1	0.1	25.22 (17.73 - 35.89)	<.001[a]	1233	100.0	-	-	-	<.001[a]	232	98.3	4	1.7	-	<.001
No [ref]	32	4.0	776	96.0			-	-	1,005	100.0			-	-	165	100.0		
Double penetration																		
Yes	262	71.2	106	28.8	1.43 (1.24 - 1.65)	<.001[a]	320	66.1	164	33.9	1.27 (1.17 - 1.37)	<.001	57	73.1	21	26.9	1.35 (1.14 - 1.59)	.003
No [ref]	663	49.7	671	50.3			913	52.1	841	47.9			175	54.2	148	45.8		
Chemsex practice																		
Yes	270	58.3	193	41.7	1.10 (0.96 - 1.28)	.05[a]	333	55.0	273	45.0	1.00 (0.92 - 1.08)	.96	59	66.3	30	33.7	1.20 (1.00 - 1.43)	.07
No [ref]	655	52.9	584	47.1			900	55.1	732	44.9			173	55.4	139	44.6		
Fisting/footing																		
Yes	85	57.8	62	42.2	1.07 (0.86 - 1.34)	.38	100	52.4	91	47.6	0.95 (0.82 - 1.09)	.45	19	59.4	13	40.6	1.03 (0.76 - 1.39)	>.99
No [ref]	840	54.0	715	46.0			1133	55.3	914	44.7			213	57.7	156	42.3		
Cruising practice																		
Yes	18	64.3	10	35.7	1.19 (0.74 - 1.89)	.28	20	62.5	12	37.5	1.14 (0.87 - 1.49)	.48	-	-	-	-	-	-
No [ref]	907	54.2	767	45.8			1213	55.0	993	45.0			232	57.9	169	42.1		
STIs History																		
Had clamydia diagnosis																		

XSL•FO
RenderX

Variables	Age group																	
	<25 years (n=1702)						25 - 49 years (n=2238)						≥50 years (n=401)					
	Intention to use PrEP																	
	Yes (n=925)		No (n=777)		PR (95% CI)	*P* value	Yes (n=1233)		No (n=1005)		PR (95% CI)	*P* value	Yes (n=232), n (%)		No (n=169), n (%)		PR (95% CI)	*P* value
	n (%)		n (%)				n (%)		n (%)				n (%)		n (%)			
Yes	89	53.3	78	46.7	0.98 (0.79 - 1.22)	.77	220	59.8	148	40.2	1.10 (1.00 - 1.21)	.05[a]	25	52.1	23	47.9	0.89 (0.67 - 1.18)	.44
No [ref]	836	54.5	699	45.5			1013	54.2	857	45.8			207	58.6	146	41.4		
Had gonorrhea diagnosis																		
Yes	103	54.2	87	45.8	1.00 (0.81 - 1.22)	.96	223	59.8	150	40.2	1.10 (1.00 - 1.21)	.05[a]	28	50.9	27	49.1	0.86 (0.66 - 1.14)	.30
No [ref]	822	54.4	690	45.6			1010	54.2	855	45.8			204	59.0	142	41.0		
Had syphilis diagnosis																		
Yes	125	51.0	120	49.0	0.93 (0.77 - 1.12)	.25	258	60.1	171	39.9	1.12 (1.02 - 1.22)	.02[a]	38	62.3	23	37.7	1.09 (0.88 - 1.35)	.48
No [ref]	800	54.9	657	45.1			975	53.9	834	46.1			194	57.1	146	42.9		

[a]variables that met the statistical criterion of a *P* value <.20.

In terms of sexual practices, nearly half of the sample identified as versatile (2125/4341; 49.0%) and reported engaging in group sex (1890/4341; 43.5%). Furthermore, the majority indicated frequent bareback sex (2363/4341; 54.4%). Higher-risk HIV exposure practices, such as DP (930/4341; 21.4%) and chemsex (1,158/4341; 26.7%), were reported by approximately one-quarter of the respondents.

Regarding HIV/AIDS prevention methods, most MSM reported widespread use of less effective practices such as interrupted intercourse (3078/4341; 70.9%), insertive sexual practices (4324/4341; 99.6%), gouinage (4324/4341; 99.6%), and sexual abstinence (4031/4341; 92.9%). Condom use during at least one of the most recent sexual encounters was reported by 91.2% (3960/4341) of participants. Among the reasons for not using condoms during sexual encounters, participants cited not using them with known sexual partners (3078/4341; 70.9%) and with partners who had recently tested for HIV (3,942/4341; 90.8%).

The majority of participants reported no previous history of chlamydia (3758/4341; 86.6%), gonorrhea (3723/4341; 85.8%), or syphilis (3606/4341; 83.1%), and frequently attended health care services (3120/4341; 71.9%).

PrEP intention was expressed by more than half of the participants (2390/4341; 55.1%). Bivariate analysis was conducted for each of the 3 age groups, and variables meeting the statistical criteria for inclusion in the multivariate modeling stage were selected accordingly (Table 1).

Table 2 presents the independent factors associated with the intention to use PrEP in each age group, along with their respective measures of association. A total of 5 variables were independently associated with the outcome of interest in the models for the age subgroups <25 and ≥50 years, while 9 variables were included in the final model for adults aged 25 - 49 years.

XSL•FO
RenderX

Table . Multivariate analysis of factors associated with intention to use PrEP among MSM by age group, a cross-sectional study, 2020 - 2021, Brazil.

Variables	β	Adjusted prevalence ratio	95% CI		*P* value
			Lower	Superior	
<25 years[a]					
Intercept	−2.99	0.05	0.03	0.07	<.001
Frequent bareback	3.23	25.40	18.18	35.48	<.01
Transgender identity	.03	1.03	1.01	1.05	.01
Casual sexual partnership	−.02	0.98	0.97	0.99	.01
Chemsex practice	−.04	0.96	0.94	0.98	<.01
≥1 sexual partner in the last 30 days	−.25	0.78	0.72	0.85	<.01
25 - 49 years[b]					
Intercept	−.74	0.48	0.39	0.59	<.01
Adopt a receptive position	.24	1.27	1.04	1.55	.02
Being an immigrant	.23	1.26	1.02	1.56	.03
Be versatile	.21	1.23	1.01	1.49	.04
Double penetration practice	.20	1.22	1.13	1.32	<.01
Inconsistent condom use with casual partners	.13	1.14	1.05	1.24	.01
Belief that condoms can be abandoned due to age of sex	.12	1.13	1.04	1.23	.04
Previous history of syphilis	.11	1.12	1.03	1.22	.01
Being single	−.11	0.90	0.83	0.97	.07
Belief that condoms can be abandoned if the partner is on PrEP[c]	−.20	0.82	0.72	0.93	.01
Gouinage practice	−.67	0.51	0.46	0.57	<.01
≥50 years[d]					
Intercept	−.73	0.48	0.40	0.58	<.01
Double penetration practice	.27	1.31	1.20	1.43	<.01
Adopt a receptive position	.24	1.27	1.04	1.55	.02
Being an immigrant	.24	1.27	1.03	1.58	.03
Be versatile	.21	1.23	1.01	1.50	.04
Chemsex practice	−.18	0.83	0.76	0.92	<.01
Group sex practice	−.20	0.82	0.75	0.89	<.01

[a]Deviance: 201.32; Akaike Information Criterion: 2,067.32; BIC: 2,110.84; Omnibus test: $\chi^2(7)=926.74$; *P* value <.001.

[b]Deviance: 1,433.85; Akaike Information Criterion: 3,925.85; BIC: 4,000.12; Omnibus test: $\chi^2(12)=36.21$; *P* value <.001.

[c]PrEP: pre-exposure prophylaxis.

[d]Deviance: 1,435.809; Akaike Information Criterion: 3,917.81; BIC: 3,963.52; Omnibus test: $\chi^2(7)=34.25$; *P* value <.001).

XSL•FO
RenderX

The practice of chemsex was associated with a lower prevalence of the intention to use PrEP among MSM aged <25 years (aPR: 0.96; 95% CI 0.94 - 0.98) and those aged ≥50 years (aPR 0.83; 95% CI 0.76 - 0.92). Certain variables were shared among the final models of different subgroups. The practice of DP was associated with a higher prevalence of the intention to use PrEP among MSM aged 25 - 49 years (aPR: 1.22; 95% CI 1.13 - 1.32) and those aged ≥50 years (aPR: 1.31; 95% CI 1.20 - 1.43). Being bottom or receptive (aPR: 1.27; 95% CI 1.04 - 1.55) or versatile (aPR 1.23; 95% CI 1.01 - 1.50) was also associated with a higher prevalence of PrEP intention among MSM aged 25 - 49 years and ≥50 years.

Frequent bareback sex was associated with a prevalence 25.4 times higher for the intention to use PrEP among MSM aged <25 years. Conversely, having one or more casual sexual partners in the last 30 days was associated with a 22% lower frequency of PrEP intention.

Among MSM aged 25 - 49 years, the prevalence of intention to use PrEP was over 10% (258/429) higher among those with specific characteristics, such as a previous diagnosis of syphilis (aPR 1.12; 95% CI 1.03 - 1.22) and nonuse of condoms due to the intensity of sex (aPR 1.13; 95% CI 1.04 - 1.23) or when sexual partners were known (aPR: 1.14; 95% CI 1.05 - 1.24). However, being single (aPR: 0.90; 95% CI 0.83 - 0.97), having a sexual partner who used PrEP (aPR: 0.82; 95% CI 0.72 - 0.93), and using gouinage as an HIV/AIDS prevention practice were exclusively associated with lower prevalences of PrEP intention.

The practice of group sex was independently associated with a lower prevalence of PrEP intention exclusively among MSM aged ≥50 years (aPR: 0.82; 95% CI 0.75 - 0.89).

Discussion

Principal Findings

This study, which explores factors associated with the intention to use PrEP among MSM across various age groups, found that more than half of the participants expressed an intention to adopt PrEP. However, this intention varies significantly between age groups, emphasizing distinct behaviors and perceptions that may influence the decision to use PrEP. Gaining an understanding of these age-specific nuances can aid in the development and implementation of tailored strategies to enhance PrEP awareness and adherence within this population.

In examining the factors that influence the intention to use PrEP among different age groups, our findings reveal that this intention is significantly shaped by sexual practices, risk perception, and sociodemographic characteristics. While some factors are shared across age groups, recognizing the unique characteristics and needs of each age category provides a more comprehensive understanding of the topic and highlights opportunities for targeted interventions.

Intention to Use PrEP Among Young Individuals (Under 25 Years)

In the age group under 25 years, the practice of unprotected sex or frequent barebacking was associated with a substantial increase in the intention to use PrEP, with a prevalence 25.4 times higher. This finding is crucial for discussions on the need for targeted interventions aimed at young individuals engaging in higher-risk behaviors for HIV infection. A 2021 study analyzing the influence of consuming bareback-type sexually explicit media among Brazilian MSM found that a preference for condomless scenes, casual sexual partnerships, and multiple partners was linked to an increased likelihood of engaging in condomless anal sex, particularly among younger MSM populations [36].

Another study conducted in Brazil demonstrated that MSM aged 15 - 19 who engage in barebacking had a higher likelihood of initiating PrEP use [37]. However, data also show that in Brazil, MSM aged 18 - 24 years are twice as likely to discontinue PrEP treatment after the first month compared to older MSM. Risky behaviors, transactional sex (in exchange for gifts or money), a lower perception of HIV risk, and illicit drug use further contribute to a higher prevalence of HIV infection within this group [38].

A study involving adolescent MSM in 3 Brazilian capitals reported a 79.9% prevalence of barebacking within the past 6 months. Another noteworthy finding indicated that being transgender was associated with a higher intention to use PrEP, potentially due to the unique risk factors this population faces throughout their lives [39].

Similarly, in our study, the practice of chemsex among individuals under 25 years was associated with a lower prevalence of PrEP intention, as was having multiple casual sexual partners in the past 30 days. This suggests complex behavioral nuances within this subgroup. One possible interpretation is that young individuals involved in chemsex may significantly underestimate their HIV exposure risk [40]. This underestimation could stem from the effects of substances consumed during chemsex, which may impair judgment, alter risk perception, and provide a temporary but false sense of invulnerability [41].

The prevalence of chemsex among young individuals may reflect sociocultural, moral, and health care dynamics linked to the pursuit of pleasure characteristic of this age group. This behavior often involves the discovery of the body and intensified emotional connections during social events or parties [42-44]. Such a focus on immediate gratification can overshadow long-term health considerations, making HIV prevention efforts seem less pressing or relevant [45,46]. Early exposure to social networks and the early initiation of sexual activity further increase the risks of engaging in chemsex and partially adhering to PrEP [47].

In addition, substance use associated with chemsex may have psychological effects, such as depression or anxiety, which influence decision-making. These mental health conditions can decrease motivation to seek and adhere to HIV prevention strategies like PrEP [48].

Thus, sociocultural factors, including limited awareness of PrEP, internalized stigma due to negative attitudes toward homosexuality and sexual activity, and the stigmatization of

XSL•FO
RenderX

PrEP users, emerge as critical barriers to adopting this preventive measure.

Intention to Use PrEP Among Adults (Intermediate Age Range of 25 to 49 Years)

For the intermediate adult age group, 9 variables showed significant associations. In this group, "personal experiences with STIs" appear to serve as a "wake-up call" for the risk of contracting HIV, increasing awareness of the risk and encouraging the pursuit of combined or additional prevention measures, such as PrEP. Unlike what was observed in younger MSM, past experiences in this group directly influence risk perception and the decision to adopt preventive measures, highlighting how accumulated experiences over time can be a critical factor for awareness and adherence to PrEP [49]. This result aligns with existing literature, which emphasizes that older age can be significantly associated with increased PrEP use among MSM [10,16,49].

PrEP intention is positively impacted by the role of sexual position, with individuals assuming insertive, versatile, or DP roles being more likely to adopt PrEP, while those engaging in nonpenetrative practices (such as gouinage) are less inclined to do so [50,51]. This finding may reflect an awareness that certain sexual positions carry higher HIV transmission risks due to increased exposure to the virus [52]. However, it is essential to interpret these results cautiously, as any unprotected sexual exposure carries inherent risks. Furthermore, practices such as DP elevate transmission risk due to the challenging nature of the practice, which can increase the likelihood of microlesions and, consequently, vulnerability to infections [34,53].

Situational factors, such as the intensity of sexual practices and trust in partners who claim to be using PrEP, emerged as significant contributors to the likelihood of PrEP adherence. These factors underscore the multifactorial nature of motivations for PrEP use among MSM in this age range, highlighting the importance of clinical consultations that adopt a holistic approach, seeing MSM comprehensively. It is crucial for these consultations to consider the complexity of individual experiences and risk perceptions to understand and effectively implement PrEP within the context of MSM's lives.

An interesting finding in this group pertains to the context of migration, where immigrant MSM were more likely to adhere to PrEP. MSM immigrants may face challenges such as limited health care access, stigma, and conditions that heighten their vulnerability, making them more inclined to seek additional protective measures like PrEP as a form of individual health control [20,54].

This underscores the need for prevention and education campaigns about PrEP to encompass a broad spectrum of life situations and risk perceptions. These variations highlight the urgent need for personalized and adaptable prevention strategies that consider the unique characteristics and needs of each age subgroup [55,56]. This context emphasizes the complexity of decisions regarding PrEP usage and reinforces the significance of culturally sensitive public health approaches informed by the diverse realities experienced within the MSM community.

Intention to Use PrEP Among Older MSM (Age Range Above 50 Years)

Limited attention has been given to analyzing the factors related to PrEP use among older MSM, and the findings from our study imply a range of implications and underlying dynamics that require careful examination [10,16,57]. Social stigma surrounding HIV/AIDS, combined with the marginalization of non-conventional sexual behaviors, can lead to reluctance or hesitation in seeking information and prevention services, including HIV testing and counseling. Older individuals may feel uncomfortable discussing their sexual practices [10,16].

As MSM age, changes in sexual practices, social networks, and risk perception often occur, which can impact their decision to initiate PrEP [16]. In this age group, there is an increasing convergence of factors with younger adults, particularly regarding the adoption of sexual positions (passive and versatile individuals are more likely to adopt PrEP) and engagement in high-risk sexual activities. This finding is supported by a multicenter study conducted with older Portuguese-speaking MSM, which showed that those who adopted a receptive or versatile sexual position were more likely to use PrEP [16]. The study emphasizes that HIV exposure among older MSM is influenced not only by age or isolated behavioral factors but also by an accumulated perception of risk and willingness to adopt protective measures.

In this age group, it is noteworthy that chemsex and group sex practices were associated with a lower intention to use PrEP, a finding that contrasts with trends observed in younger counterparts [19,58]. One possible explanation is that among older MSM, chemsex tends to occur less frequently and in more situational contexts, often being perceived as an isolated event rather than a routine sexual behavior. This occasional nature of their practices may lead them to underestimate their risk, viewing PrEP as an unnecessary or excessive measure for situations they perceive as exceptional. Unlike younger individuals, who often see chemsex as part of a continuous high-risk practice associated with parties and sexual encounters, and thus opt for PrEP to mitigate frequent exposure, older MSM may not perceive themselves as being in a constant risk context that would justify regular preventive use [16,19,59].

Additionally, risk perception among older MSM is often shaped by traditional or conservative views on sexuality and self-image [17], leading them to perceive PrEP as a strategy targeted at younger or more sexually "promiscuous" populations. This internalized stigma, which associates PrEP use with behaviors deemed morally transgressive, can discourage interest in the prophylaxis. For some, adherence to PrEP may seem unnecessary or even damaging to their identity, revealing a significant barrier related to self-image and the influence of social norms on the use of preventive strategies [16,17].

Another relevant factor is what can be described as "prevention fatigue" among older MSM. After decades of exposure to the HIV/AIDS epidemic and adherence to prevention campaigns primarily focused on condom use and reducing risky practices, many may feel exhausted by the constant vigilance required for prevention. This fatigue can lead to apathy or resignation toward new preventive interventions, such as PrEP, resulting in passive

XSL•FO
RenderX

resistance to its adoption. This fatigue is particularly complex, as it reflects not only practical exhaustion but also emotional strain from prolonged exposure to stigma and fear surrounding HIV.

Additionally, it is important to consider that older MSM may experience social isolation, lower health literacy, and limited access to support networks or groups that discuss sexual health. This can hinder their access to updated information on PrEP and other preventive interventions. Such isolation, often exacerbated by age-related stigma and barriers within health care settings, contributes to a lack of awareness and misinformation about PrEP, reinforcing perceptions that it is unnecessary or unsuitable for this age group [16,17].

These findings underscore the need for personalized prevention approaches that consider the specific contexts and values of MSM over the age of 50. Interventions targeting this population should adopt inclusive language that is sensitive to age-related issues, emphasizing the benefits of PrEP without associating it exclusively with a specific age group or lifestyle.

Implications for Public Health and Surveillance

Our findings carry critical and urgent implications for public health and surveillance. Public health strategies must consider generational differences in attitudes, knowledge, and behaviors related to PrEP and HIV. Educational campaigns and awareness programs should be tailored to address the unique and specific needs of each age group, using the most effective language and communication channels for each, rather than using a single, generic approach.

Furthermore, public health initiatives should work to expand access to PrEP by addressing specific barriers faced by different generations. This includes reducing stigma surrounding PrEP use, ensuring financial and physical accessibility, and providing targeted education to healthcare providers on how to effectively address issues of sexuality and HIV prevention with MSM of all ages.

Our conclusions also underscore the importance of integrating sexual and mental health services, particularly to address issues such as stigma, substance use, and social isolation. Integrated services can better address the complex health needs of MSM, promoting a holistic and effective approach to HIV prevention.

Moreover, public health must continue to monitor trends in PrEP use and HIV infection rates across different generational groups of MSM. This surveillance can help identify gaps in prevention and treatment efforts, informing the development of targeted interventions. In addition, research should investigate the underlying reasons for generational differences in PrEP intention to inform more effective intervention strategies.

It is also important to consider the variability of socialization patterns and affective and sexual interactions, sociocultural transformations, and the influence of sexually explicit media content. The appeal of engaging in high-risk sexual practices, such as chemsex and DP—practices highlighted in our study—can lead to greater involvement in group sexual activities, increasing the risk of sexually transmitted infections, especially when there is limited knowledge of individual serological status and preventive measures taken by each participant [9,60,61].

In addition, the study of health behaviors among adult and elderly men, who are often overlooked in health priority agendas, must not be neglected. Research has highlighted the consequences of this oversight, indicating that older men often have insufficient information about HIV/AIDS compared to younger men, a gap that is reflected in their perceptions of individual health, sexual activity, and HIV testing behaviors.

Limitations

Our study presents several limitations that warrant discussion. First, the data collection method may introduce significant information biases. While the use of an online survey and a “snowball” sampling strategy adapted to the virtual environment is both innovative and practical, it may not capture a fully representative sample of the Brazilian MSM population. This approach tends to include individuals with greater internet access and technological familiarity, potentially excluding those from less urbanized regions or with limited digital access—a challenge that is especially pronounced in a vast country like Brazil. Moreover, our initial recruitment through 2 dating apps may have skewed the sample toward individuals who are more sexually active or seeking casual relationships, potentially not reflecting the broader practices and intentions regarding PrEP use among all MSM.

Additionally, although the study addresses key sociodemographic and behavioral variables, it may not fully encompass the complex sociostructural contexts that influence PrEP utilization intentions. Critical factors such as social stigma, discrimination, social support, mental health, and access to health care services play significant roles in individual health decisions and may vary widely among different MSM subgroups. Moreover, cultural and media narratives, particularly concerning chemsex and other sexual practices, may have distinct impacts across different age groups, which were not fully explored in this study.

Conclusions

The adoption of HIV preventive measures, such as PrEP, among Brazilian MSM reveals complex and varied nuances across different age groups. Our findings indicate that more than half of the participants expressed an intention to adopt PrEP; however, this intention varies significantly among different age groups, reflecting distinct behaviors and perceptions that influence PrEP use decisions.

For individuals under 25 years of age, high-risk practices such as frequent bareback sex were strongly associated with a greater intention to use PrEP, potentially reflecting a heightened perception of HIV risk or a response to their vulnerability. In contrast, having multiple casual sexual partners in the last 30 days was associated with a lower likelihood of intending to use PrEP, suggesting potential challenges in risk perception or PrEP accessibility for these individuals.

Among those in the intermediate age range (25 to 49 years), factors such as engaging in DP, being passive or versatile, and having a previous diagnosis of syphilis were associated with a

XSL•FO
RenderX

higher intention to use PrEP. For MSM aged 50 years or older, DP and receptive or versatile sexual positions were linked to a greater intention to use PrEP, while group sex was negatively associated. These differences underscore the importance of personalized PrEP promotion approaches that consider the diverse behaviors, attitudes, and challenges faced by MSM across different ages. Understanding these nuances is essential for developing more effective public health strategies aimed at increasing PrEP awareness and adherence, ultimately contributing to a reduction in HIV infection rates among Brazilian MSM.

Acknowledgments

This research was funded by Fundação para a Ciência e Tecnologia (Portugal) through national funds to the Associated Laboratory in Translation and Innovation Towards Global Health REAL (LA/P/0117/2020) and was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance Code 001

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization was carried out by AFLS, CJNR, IACM, and IF. Data curation was performed by AFLS, CJNR, and GRSS. Formal analysis was conducted by CJNR and GRSS. Funding acquisition was managed by AFLS, CJNR, APMF, IACM, and IF. Investigation was undertaken by AFLS, CJNR, GRSS, ELSC, HEFC, GS, LGV, LMC, APMF, TMF, MBS, RCDN, LAA, ARS, IACM, and IF. Methodology was developed by AFLS, CJNR, GRSS, ELSC, HEFC, GS, LGV, LMC, APMF, TMF, MBS, RCDN, LAA, ARS, IACM, and IF. Project administration was handled by AFLS, CJNR, IACM, and IF. Visualization was prepared by AFLS, CJNR, GRSS, ELSC, HEFC, GS, LGV, LMC, APMF, TMF, MBS, RCDN, LAA, ARS, IACM, and IF. Writing – review and editing was done by all authors.

Conflicts of Interest

None declared.

References

1. Zhang H, Sanchez T, Zlotorzynska M, Baral S, Hecht J. Experiences of stigma and HIV status associated with awareness and usage of safer sex functions in dating apps among MSM, 2019 American men's internet survey (AMIS) study. AIDS Behav 2023 Nov;27(11):3603-3611. [doi: 10.1007/s10461-023-04074-4] [Medline: 37421546]
2. Ribeiro Banze Á, Muleia R, Nuvunga S, Boothe M, Semá Baltazar C. Trends in HIV prevalence and risk factors among men who have sex with men in Mozambique: implications for targeted interventions and public health strategies. BMC Public Health 2024 Apr 27;24(1):1185. [doi: 10.1186/s12889-024-18661-0] [Medline: 38678173]
3. Mayer KH, Nelson L, Hightow-Weidman L, et al. The persistent and evolving HIV epidemic in American men who have sex with men. Lancet 2021 Mar 20;397(10279):1116-1126. [doi: 10.1016/S0140-6736(21)00321-4] [Medline: 33617771]
4. Sullivan PS, Satcher Johnson A, Pembleton ES, et al. Epidemiology of HIV in the USA: epidemic burden, inequities, contexts, and responses. The Lancet 2021 Mar;397(10279):1095-1106. [doi: 10.1016/S0140-6736(21)00395-0]
5. Cruz Neto J, Maia AC, de Siqueira Holanda VM, et al. From HIV to Mpox: socio-discursive analysis of communicable diseases, stigma, and pathologization in sexual minorities in Brazil. Arch Sex Behav 2024 Nov;53(10):3809-3824. [doi: 10.1007/s10508-024-02968-5] [Medline: 39227499]
6. UNAIDS. Estatísticas Globais do HIV. 2023. URL: https://unaids.org.br/estatisticas [accessed 2025-06-24]
7. Geneva: Joint United Nations Programme on HIV/AIDS. The path that ends aids 2023 unaids global aids update. 2023. URL: http://www.wipo.int/amc/en/mediation/rules [accessed 2024-03-13]
8. Brasil. Boletim Epidemiológico - HIV e Aids 2023. 2023. URL: https://www.gov.br/aids/pt-br/central-de-conteudo/boletins-epidemiologicos/2023/hiv-aids/boletim-epidemiologico-hiv-e-aids-2023.pdf/view [accessed 2025-06-24]
9. Teixeira JRB, Lima S, Sousa AD, et al. Determinants of sexual exposure to HIV in Portuguese and Brazilian adolescents: a path analysis. Rev Lat Am Enfermagem 2022;30(spe):e3715. [doi: 10.1590/1518-8345.6222.3715] [Medline: 36197395]
10. Hulstein SH, Zimmermann HML, de la Court F, et al. Factors associated with the intention to use HIV preexposure prophylaxis for young and older men who have sex with men. Sex Transm Dis 2022 May 1;49(5):343-352. [doi: 10.1097/OLQ.0000000000001599] [Medline: 35001016]
11. Szwarcwald CL, Souza Júnior PD, Pascom ARP, et al. HIV incidence estimates by sex and age group in the population aged 15 years or over, Brazil, 1986-2018. Rev Soc Bras Med Trop 2022;55(suppl 1):e0231. [doi: 10.1590/0037-8682-0231-2021] [Medline: 35107522]

XSL•FO
RenderX

12. Mitsch A, Singh S, Li J, Balaji A, Linley L, Selik R. Age-associated trends in diagnosis and prevalence of infection with HIV among men who have sex with men - United States, 2008-2016. MMWR Morb Mortal Wkly Rep 2018 Sep 21;67(37):1025-1031. [doi: 10.15585/mmwr.mm6737a2] [Medline: 30235184]
13. Risher KA, Cori A, Reniers G, et al. Age patterns of HIV incidence in eastern and southern Africa: a modelling analysis of observational population-based cohort studies. Lancet HIV 2021 Jul;8(7):e429-e439. [doi: 10.1016/S2352-3018(21)00069-2] [Medline: 34197773]
14. Roomaney RA, van Wyk B, Pillay-van Wyk V. Aging with HIV: increased risk of HIV comorbidities in older adults. Int J Environ Res Public Health 2022 Feb 18;19(4):2359. [doi: 10.3390/ijerph19042359] [Medline: 35206544]
15. Sousa AFL, Queiroz A, Fronteira I, Lapão L, Mendes IAC, Brignol S. HIV testing among middle-aged and older men who have sex with men (MSM): a blind spot? Am J Mens Health 2019;13(4):1557988319863542. [doi: 10.1177/1557988319863542] [Medline: 31288596]
16. Nunes RV, de Araújo TME, de Oliveira LB, et al. Sexual practices and predisposition to PrEP use among men ages 50 years and older who have sex with men: a cross-sectional study. J Assoc Nurses AIDS Care 2023;34(6):548-565. [doi: 10.1097/JNC.0000000000000433] [Medline: 37815844]
17. Queiroz A, Sousa AFL, Sousa AR. Sexuality “in apps”? older men who have sex with men and their use of dating apps in brazil. In: Ageing, Men and Social Relations 2023. [doi: 10.46692/9781447363088]
18. Johnston LG, Soe P, Widihastuti AS, et al. Alarmingly high HIV prevalence among adolescent and young men who have sex with men (MSM) in urban Indonesia. AIDS Behav 2021 Nov;25(11):3687-3694. [doi: 10.1007/s10461-021-03347-0] [Medline: 34143341]
19. de Sousa Á, Nunes Ribeiro CJ, Santos GDS, et al. Intention to use PrEP among men who have sex with men and engage in chemsex: an international descriptive study. Ther Adv Infect Dis 2023;10:20499361231206918. [doi: 10.1177/20499361231206918] [Medline: 37900982]
20. Sousa Á, Lima S, Ribeiro CJN, et al. Pre-exposure prophylaxis among Brazilian men who have sex with men: a comparative study between migrants and non-migrants. Front Public Health 2023;11:1198339. [doi: 10.3389/fpubh.2023.1198339] [Medline: 37663850]
21. Wang H, Zhang J, Chu Z, et al. Risk-taking behaviors and adherence to HIV pre-exposure prophylaxis in users of geosocial networking apps: real-world, multicenter study. J Med Internet Res 2020 Oct 14;22(10):e22388. [doi: 10.2196/22388] [Medline: 33052132]
22. Hentges B, Knauth DR, Vigo A, et al. Inconsistent condom use with casual partners among men who have sex with men in Brazil: a cross-sectional study. Rev Bras Epidemiol 2023;26:e230019. [doi: 10.1590/1980-549720230019] [Medline: 36995830]
23. Farley SM, Wang C, Bray RM, et al. Progress towards the UNAIDS 90-90-90 targets among persons aged 50 and older living with HIV in 13 African countries. J Int AIDS Soc 2022 Sep;25 Suppl 4(Suppl 4):e26005. [doi: 10.1002/jia2.26005] [Medline: 36176030]
24. Weiss KM, Prasad P, Sanchez T, Goodreau SM, Jenness SM. Association between HIV PrEP indications and use in a national sexual network study of US men who have sex with men. J Int AIDS Soc 2021 Oct;24(10):e25826. [doi: 10.1002/jia2.25826] [Medline: 34605174]
25. Luiz RR, Magnanini MMF, Medronho AR, Carvalho DM, Bloch BV. O tamanho da amostra em investigações epidemiológicas [Book in Spanish]. In: Epidemiologia: São Paulo: Atheneu; 2011:415-427.
26. Hulley S, Browner WS, Cummings SR, Grady D, Newman TB. Delineando a Pesquisa Clinica: Uma Abordagem Epidemiológica [Book in Spanish]: Porto Alegre: ArtmedArtmed; 2015.
27. Camargo ELS, de Oliveira BIA, Siffoni IF, et al. Low psychological well-being in men who have sex with men (MSM) during the shelter-in-place orders to prevent the COVID-19 spread: results from a nationwide study. Sex Res Social Policy 2022;19(1):391-400. [doi: 10.1007/s13178-021-00550-5] [Medline: 33643482]
28. Sousa LRM, Elias HC, Fernandes NM, Gir E, Reis RK. Knowledge of PEP and PrEP among people living with HIV/aids in Brazil. BMC Public Health 2021 Jan 7;21(1):64. [doi: 10.1186/s12889-020-10135-3] [Medline: 33413241]
29. Chone JS, Lima S, Fronteira I, et al. Factors associated with chemsex in Portugal during the COVID-19 pandemic. Rev Lat Am Enfermagem 2021;29:e3474. [doi: 10.1590/1518-8345.4975.3474] [Medline: 34468628]
30. Sousa ÁD, Queiroz A, Lima S, et al. Chemsex practice among men who have sex with men (MSM) during social isolation from COVID-19: multicentric online survey. Cad Saude Publica 2020;36(12):e00202420. [doi: 10.1590/0102-311X00202420] [Medline: 33237252]
31. Dias S, Gama A, Pingarilho M, Simões D, Mendão L. Health services use and HIV prevalence among migrant and national female sex workers in Portugal: are we providing the services needed? AIDS Behav 2017 Aug;21(8):2316-2321. [doi: 10.1007/s10461-016-1511-x] [Medline: 27475942]
32. Sousa AD, Camargo ELS, Mendes IAC. Chemsex and its repercussions on the health of men who have sex with men (MSM): a global health perspective. Rev Bras Enferm 2023;76(3):3-7. [doi: 10.1590/0034-7167-2023-0004]
33. Santos GDS, Ribeiro CJN, Lima S, et al. Chemsex among men who have sex with men during the Mpox health crisis in Brazil: A nationwide web survey. Public Health Nurs 2024;41(3):589-601. [doi: 10.1111/phn.13308] [Medline: 38528749]

XSL•FO
RenderX

34. Sousa AD, Lima S, Ribeiro CJN, et al. Adherence to pre-exposure prophylaxis (PrEP) among men who have sex with men (MSM) in Portuguese-speaking countries. Int J Environ Res Public Health 2023 Mar 10;20(6):4881. [doi: 10.3390/ijerph20064881] [Medline: 36981792]
35. Felipe da Cruz AX, Berté R, Oliveira ADBL, et al. Barriers and facilitators to HIV/AIDS testing among Latin immigrant men who have sex with men (MSM): a systematic review of the literature. TOAIDJ 2023 Aug 7;17(1):1. [doi: 10.2174/18746136-v17-230720-2023-12]
36. Martins ADA, Queiroz AAFLN, Frota OP, et al. Consumption of sexually explicit media and unprotected anal sex in men who have sex with men. Cien Saude Colet 2021 Nov;26(11):5841-5849. [doi: 10.1590/1413-812320212611.30532020] [Medline: 34852113]
37. Soares F, Magno L, Eustorgio Filho M, et al. Important steps for PrEP uptake among adolescent men who have sex with men and transgender women in Brazil. PLoS ONE 2023;18(4):e0281654. [doi: 10.1371/journal.pone.0281654] [Medline: 37014830]
38. Echeverría-Guevara A, Coelho LE, Veloso VG, et al. Travestis, transgender women and young MSM are at high risk for PrEP early loss to follow-up in Rio de Janeiro, Brazil. Braz J Infect Dis 2023;27(1):102733. [doi: 10.1016/j.bjid.2022.102733] [Medline: 36586721]
39. Rosário R, Dourado I, Pereira M, et al. Factors associated with condomless anal sex among adolescent men who have sex with men and transgender women in three Brazilian state capitals: a PrEP1519 study. Rev Saude Publica 2024;58(suppl 1):8s. [doi: 10.11606/s1518-8787.2020054005462] [Medline: 39417518]
40. Jalil EM, Torres TS, de A Pereira CC, et al. High rates of sexualized drug use or chemsex among Brazilian transgender women and young sexual and gender minorities. Int J Environ Res Public Health 2022 Feb 2;19(3):1704. [doi: 10.3390/ijerph19031704] [Medline: 35162728]
41. Irfan SD, Sarwar G, Emran J, Khan SI. An uncharted territory of sexualized drug use: exploring the dynamics of chemsex among young and adolescent MSM including self-identified gay men in urban Dhaka, Bangladesh. Front Psychol 2023;14:1124971. [doi: 10.3389/fpsyg.2023.1124971] [Medline: 37425169]
42. Rungmaitree S, Werarak P, Pumpradit W, et al. A pilot program of HIV pre-exposure prophylaxis in Thai youth. PLoS ONE 2024;19(2):e0298914. [doi: 10.1371/journal.pone.0298914] [Medline: 38386680]
43. Tanner MR, Miele P, Carter W, et al. Preexposure prophylaxis for prevention of HIV acquisition among adolescents: clinical considerations, 2020. MMWR Recomm Rep 2020 Apr 24;69(3):1-12. [doi: 10.15585/mmwr.rr6903a1] [Medline: 32324724]
44. Ramautarsing RA, Meksena R, Sungsing T, et al. Evaluation of a pre-exposure prophylaxis programme for men who have sex with men and transgender women in Thailand: learning through the HIV prevention cascade lens. J Int AIDS Soc 2020 Jun;23 Suppl 3(Suppl 3):e25540. [doi: 10.1002/jia2.25540] [Medline: 32602660]
45. Hovaguimian F, Kouyos RD, Kusejko K, et al. Incidence of sexually transmitted infections and association with behavioural factors: time-to-event analysis of a large pre-exposure prophylaxis (PrEP) cohort. HIV Med 2024 Jan;25(1):117-128. [doi: 10.1111/hiv.13543] [Medline: 37771207]
46. Hammond R, Cambiano V, Lampe FC, et al. Predictors of starting and stopping chemsex in men who have sex with men in England: findings from the AURAH2 prospective study. Sex Transm Infect 2023 Nov;99(7):474-481. [doi: 10.1136/sextrans-2023-055774] [Medline: 37321843]
47. Tan RKJ, O'Hara CA, Koh WL, et al. Social capital and chemsex initiation in young gay, bisexual, and other men who have sex with men: the pink carpet Y cohort study. Subst Abuse Treat Prev Policy 2021 Feb 19;16(1):18. [doi: 10.1186/s13011-021-00353-2] [Medline: 33608005]
48. Tomkins A, George R, Kliner M. Sexualised drug taking among men who have sex with men: a systematic review. Perspect Public Health 2019 Jan;139(1):23-33. [doi: 10.1177/1757913918778872] [Medline: 29846139]
49. Morgan E, Moran K, Ryan DT, Mustanski B, Newcomb ME. Threefold increase in PrEP uptake over time with high adherence among young men who have sex with men in Chicago. AIDS Behav 2018 Nov;22(11):3637-3644. [doi: 10.1007/s10461-018-2122-5] [Medline: 29728949]
50. Goedel WC, Schneider JA, Hambrick HR, et al. Are anal sex roles associated with preferences for pre-exposure prophylaxis administration modalities among men who have sex with men? Arch Sex Behav 2018 Oct;47(7):2123-2133. [doi: 10.1007/s10508-017-1083-5] [Medline: 29192368]
51. Hudrudchai S, Suwanwong C, Prasittichok P, Mohan KP, Janeaim N. Pre-exposure prophylaxis adherence among men who have sex with men: a systematic review and meta-analysis. J Prev Med Public Health 2024 Jan;57(1):8-17. [doi: 10.3961/jpmph.23.345] [Medline: 38147821]
52. Mabire X, Puppo C, Morel S, et al. Pleasure and PrEP: pleasure-seeking plays a role in prevention choices and could lead to PrEP initiation. Am J Mens Health 2019;13(1):1557988319827396. [doi: 10.1177/1557988319827396] [Medline: 30819060]
53. de Araújo TME, Costa Chaves FR, de Oliveira M, et al. Sexual practices and HIV risk perception among men who have sex with men in Brazil. J Int Assoc Provid AIDS Care 2024;23:23259582241283196. [doi: 10.1177/23259582241283196] [Medline: 39314093]

XSL•FO
RenderX

54. Sun S, Yang C, Zaller N, Zhang Z, Zhang H, Operario D. PrEP willingness and adherence self-efficacy among men who have sex with men with recent condomless anal sex in urban China. AIDS Behav 2021 Nov;25(11):3482-3493. [doi: 10.1007/s10461-021-03274-0] [Medline: 33932186]
55. Sun Z, Gu Q, Dai Y, et al. Increasing awareness of HIV pre-exposure prophylaxis (PrEP) and willingness to use HIV PrEP among men who have sex with men: a systematic review and meta-analysis of global data. J Int AIDS Soc 2022 Mar;25(3):e25883. [doi: 10.1002/jia2.25883] [Medline: 35255193]
56. Sudarto B, Chow EPF, Medland N, et al. "How PrEPared are you?": Knowledge of and attitudes toward PrEP among overseas-born and newly arrived gay, bisexual, and other men who have sex with men in Australia. Front Public Health 2022;10:946771. [doi: 10.3389/fpubh.2022.946771] [Medline: 36062118]
57. Klein H, Washington T. Older versus younger men who have sex with men: awareness of and potential barriers to the use of pre-exposure prophylaxis (PrEP) medication to prevent the transmission of HIV. J AIDS HIV Treat 2021;2(2):42-50. [doi: 10.33696/AIDS.2.006]
58. Flores Anato JL, Panagiotoglou D, Greenwald ZR, et al. Chemsex practices and pre-exposure prophylaxis (PrEP) trajectories among individuals consulting for PrEP at a large sexual health clinic in Montréal, Canada (2013-2020). Drug Alcohol Depend 2021 Sep 1;226:108875. [doi: 10.1016/j.drugalcdep.2021.108875] [Medline: 34218004]
59. Amundsen E, Muller AE, Reierth E, Skogen V, Berg RC. Chemsex among men who have sex with men: a systematic scoping review of research methods. J Homosex 2024 May 11;71(6):1392-1418. [doi: 10.1080/00918369.2023.2170757] [Medline: 36939142]
60. Oliveira JD, Sousa AD, Araújo IFM. Sexually transmitted infections in men in the prison system: integrative review. Rev baiana enferm 2022;36. [doi: 10.18471/rbe.v36.38071]
61. Chaves ADC, Araújo TD, Sousa E, Almeida PD, Nunes RV, Lulendo EM. Sexually explicit media consumption and HIV/AIDS risk practices in northeastern Brazil. Rev Enferm UFPI 2023;12(1):1-12. [doi: 10.26694/reufpi.v12i1.3933]

Abbreviations

aPR: adjusted prevalence ratio
DP: double penetration
PR: prevalence ratio
PrEP: Pre-Exposure Prophylaxis
STIs: sexually transmitted infections
WHO: World Health Organization

Edited by A Mavragani; submitted 14.03.24; peer-reviewed by Q Adeyemo, T Finlayson; revised version received 07.11.24; accepted 20.11.24; published 08.07.25.

Please cite as:
Sousa AFLD, Ribeiro CJN, Santos GRDS, Camargo ELS, Carvalho HEFD, Schneider G, Vieira LG, Carneiro LM, Fernandes APM, Fernandes TM, Bezerra-Santos M, Nascimento RDCD, Andrade LA, Sousa ARD, Fronteira I, Mendes IAC
Intention to Use Pre-Exposure Prophylaxis Among Age Groups of Brazilian Men Who Have Sex With Men: National Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e58405
URL: https://publichealth.jmir.org/2025/1/e58405
doi:10.2196/58405

© Alvaro Francisco Lopes de Sousa, Caique Jordan Nunes Ribeiro, Guilherme Reis de Santana Santos, Emerson Lucas Silva Camargo, Herica Emilia Félix de Carvalho, Guilherme Schneider, Leticia Genova Vieira, Liliane Moretti Carneiro, Ana Paula Morais Fernandes, Talita Morais Fernandes, Márcio Bezerra-Santos, Rita de Cassia Dias Nascimento, Lucas Almeida Andrade, Anderson Reis de Sousa, Inês Fronteira, Isabel Amélia Costa Mendes. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Knowledge Access on Risky Sexual Behaviors Among Chinese Youths to Improve HIV Prevention: Cross-Sectional Study

Yun Zhang[1], MPH; Jie Jin[2], MPH, PhD; Feng Cheng[3,4], MPH, PhD; Xingliang Zhang[2*], MPH; Junfang Xu[1*], MPH, PhD

[1]The Second Affiliated Hospital, School of Public Health, Zhejiang University School of Medicine, Number 866, Yuhangtang Road, Hangzhou, China

[2]Hangzhou Center for Disease Control and Prevention (Hangzhou Health Supervision Institution), Hangzhou, China

[3]Vanke School of Public Health, Tsinghua University, Beijing, China

[4]Institute for Healthy China, Tsinghua University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Junfang Xu, MPH, PhD
The Second Affiliated Hospital, School of Public Health, Zhejiang University School of Medicine, Number 866, Yuhangtang Road, Hangzhou, China

Abstract

Background: HIV infection significantly challenges Chinese youth, with most infections resulting from risky sexual behaviors. Health communication is critical in preventing risky behaviors. Different HIV-related information sources exhibit unique characteristics, affecting young adults' perception of HIV and sexual behaviors.

Objective: This study aimed to examine the relationship between knowledge accesses and risky behaviors, analyzing the mediating effects of HIV-related knowledge to provide evidence for more effective HIV knowledge dissemination and the reduction of risky behaviors among youth.

Methods: A questionnaire on HIV-related knowledge and behavior of young students was used for the online survey. Knowledge accesses included school education, mass communication, and interpersonal communication. HIV-related knowledge was categorized into fundamental and behavioral guidance knowledge. Risky behavior was defined as multiple sexual partners or unprotected sex. Binary logistic regression was used to analyze the relationship between knowledge accesses, HIV-related knowledge, and risky behaviors. Multiple mediation models constructed by the Process macro (version 4.1) were conducted to determine whether the knowledge level mediated the relationship between knowledge accesses and risky behaviors.

Results: Totally 20,602 respondents participated, with 9541 (46.31%) males and 11,061 (53.69%) females, averaging 20.14 (SD 1.91) years. Mass communication reached 19,030 (92.37%) students, school education reached 17,949 (87.12%), and 11,274 (54.72%) accessed interpersonal communication. Among the 2423 sexually active students, 363 (14.98%) had multiple sexual partners, and 830 (34.26%) engaged in unprotected sex. School education (odds ratio [OR] 0.60, 95% CI 0.46 to 0.78), mass communication (OR 0.64, 95% CI 0.47 to 0.87) and behavioral guidance knowledge (OR 0.79, 95% CI 0.69 to 0.90) were negatively associated with multiple sexual partners, while interpersonal communication (OR 1.41, 95% CI 1.13 to 1.75) was positively associated. School education (OR 0.75, 95% CI 0.62 to 0.91) and behavioral guidance knowledge (OR 0.81, 95% CI 0.73 to 0.89) were negatively correlated with unprotected sexual behaviors, while interpersonal communication (OR 1.27, 95% CI 1.09 to 1.46) and fundamental knowledge (OR 1.17, 95% CI 1.08 to 1.28) were significantly positively correlated. Knowledge accesses indirectly influenced risky behaviors through fundamental and behavioral guidance knowledge. School education had opposing indirect effects of −0.02 and 0.02 through behavioral guidance and fundamental knowledge, mass communication showed effects of −0.04 and 0.01, while interpersonal communication demonstrated indirect effects of 0.01 and 0.01 for both pathways.

Conclusions: Students acquiring HIV knowledge from school education and mass media are less likely to engage in risky behaviors than those learning through interpersonal communication. We should fully leverage the advantages of school education and mass communication, and an emphasis should be placed on behavioral guidance knowledge to promote changes in risky behavior. Knowledge access primarily influences risky behaviors by shaping cultural values and behavioral norms, rather than transmitting information.

(JMIR Public Health Surveill 2025;11:e68339) doi:10.2196/68339

KEYWORDS

Chinese young student; knowledge access; risky sexual behavior; HIV knowledge; health communication; cross-sectional study

XSL•FO
RenderX

Introduction

HIV infection presents a significant public health challenge due to its high fatality rate and widespread transmission worldwide. Until 2023, a cumulative 39.9 million HIV positive cases, including 1.3 million new infections, were estimated according to the United Nations Program on HIV/AIDS [1]. While China's concerted efforts have maintained overall HIV prevalence at a low level (approximately <0.1%), there is an alarming rise in infections among young students [2]. Worryingly, nearly 3000 new cases of HIV infection are reported among young adults aged 15 to 24 years every year, increasing at an annual rate of 30% - 50% [3,4]. Therefore, college students in China have become a high-risk group for HIV infection. Sexual transmission is the main mode of HIV spread among young students. Being sexually active and open to casual sex, without correct sexual behavior guidance, young students are prone to have risky behaviors (such as having multiple sexual partners and not using condoms), thereby increasing the risk of HIV infection. Given that, transforming risky behaviors is considered a preventive and control measure for restricting the spread of the HIV pandemic.

Health communication is found to be critical in behavioral intervention [5]. To reduce HIV infection among young students effectively, China conducted HIV health education for students in its 6 major projects aimed at curbing AIDS in 2019 [6]. The public can acquire HIV-related information through various knowledge access points, which can be commonly categorized into organizational communication, mass communication, interpersonal communication, and other types of communication studies [7-9]. Information access exposure exerts an unstoppable and subtle influence on individuals' cognition and behaviors. Media Richness Theory, proposed by Daft and Lengel [10], suggests that differences in information accesses result in variations in the ability to transmit information. Different HIV-related information sources assign different labels to the nature of the information that affects young adults' understanding and perception of HIV, ultimately shaping their sexual behaviors. Therefore, the association between different knowledge accesses, knowledge levels, and behaviors deserves attention, as it is essential for identifying effective methods to disseminate HIV-related knowledge and reduce risky behaviors.

However, research on sources of HIV information primarily focuses on the correlation between information sources and HIV-related knowledge [11-13], which fails to fully capture the process of transforming into sexual behavior. Students adopt safe sexual behavior and reject risky behavior based on their accumulation of knowledge, which is considered the ultimate goal of HIV knowledge transmission. Only then can the risk of HIV infection among students be truly reduced [14]. Few studies have explored the relationship between communication channels and sexual behaviors, but they have often neglected to organize these channels [15,16]. Given the diversity of information sources, examining each one individually risks offering incomplete perspectives. Instead of focusing on each specific communication channel, such as distributing leaflets, watching television, or chatting with classmates, it is essential to classify these different channels, which enables the identification of commonalities and differences among them, clarifying their respective strengths and limitations in the dissemination of information.

With the aim of seeking effective information accesses, providing evidence for disseminating knowledge, and reducing risky behavior among young students, this study investigated the knowledge access, knowledge level, and risky behaviors of students, focusing on whether knowledge accesses are associated with sexual behaviors, determining if the HIV-related knowledge mediates the relationship between knowledge accesses and risky behaviors.

Methods

Participants

A nonprobability convenience sampling approach was used to recruit young students from 5 colleges located in Zhejiang province, covering comprehensive education, science and technology, medicine, and vocational training, which could represent different types of higher education institutions in Zhejiang province. The sampling strategy aimed to capture the diversity of college students, ensuring a robust representation of this study's population. The sample size was estimated with the following sample size formula.

$$N=Z2\times p\times(1-p)E2$$

Z=1.96 (95% confidence level); *P*=.5 (maximum heterogeneity assumption), d=0.05 (5% sampling error), and a response rate of 75% was considered. Given the inherent selection bias of convenience sampling, the minimum calculated sample size was inflated by 30%. As a result, 669 students were required. Finally, 20,602 was investigated and incorporated in the analysis with a strong statistical power.

Recruitment

The survey was conducted on SoJump (Changsha Ranxing Information Technology Co, Ltd), a widely used Chinese online platform with a user-friendly interface and broad applicability in academic research [17]. The SoJump platform uses cookies to prevent duplicate submissions, and we set a mandatory requirement that participants complete all survey items before submission to ensure data integrity. Participants were recruited voluntarily through QR codes distributed by school counselors via WeChat (Tencent Holdings Limited) and QQ (Tencent Holdings Limited) groups, ensuring accessibility across diverse student populations. Data collection took place from November 2022 to June 2023. The inclusion criteria of the college students were (1) registered students of higher education institutions during the survey period and (2) willing to participate in the survey.

Variable Measurement

A questionnaire on HIV/AIDS knowledge and behavior among young college students was used to conduct the online survey. The content of the questionnaire included demographic characteristics (age, gender, major, etc), knowledge accesses, HIV-related knowledge, and sexual behaviors. Knowledge accesses in our study included school education, mass communication, and interpersonal communication, which are

XSL•FO
RenderX

always discussed in educational communication [18]. As an educational organization, a school can disseminate HIV-related knowledge by organizing classes and giving lectures, which play an important role in the education of young students. Mass communication means the way to transmit information to the public via mass media, such as the internet, radio, television, books, newspapers, and so on. Interpersonal communication refers to direct interaction between 2 or more people to exchange information, such as acquiring HIV-related knowledge from parents, classmates, or friends.

A commonly used scale, "8 questions for youths," was used to measure HIV-related knowledge [19,20], which refers to the 8 items of HIV/AIDS core knowledge proposed by the Chinese Center for Disease Control and Prevention in 2016 [21]. To better understand the roles of different types of knowledge in cognitive construction and behavioral intervention, HIV-related knowledge was classified into 2 major categories: fundamental knowledge and behavioral guidance knowledge. The classification aligns with the framework of the "Core Information on HIV/AIDS Prevention Education for Young Students (2021)" issued by the China Center for Disease Control and Prevention [22], which originally divides HIV knowledge into 4 modules: awareness of harm, legal regulations, preventive measures, and testing and treatment. However, to ensure measurement reliability and stability, dimension grouping must consider item correlations and internal consistency. With only 8 items, subdividing into 4 modules risks low reliability and modeling challenges. Therefore, the first 2 modules (awareness of harm, legal regulations) were merged into fundamental knowledge to establish a basic understanding of HIV/AIDS and its societal impacts, such as "HIV/AIDS is a serious, incurable infectious disease" and "national laws protect the rights of people living with HIV, including marriage, employment, and education." The latter 2 modules (preventive measures, testing, and treatment) were combined into behavioral guidance knowledge, focusing on specific preventive actions and behavioral recommendations, such as "daily life and learning contact do not transmit HIV" and "seeking HIV testing and counseling after risky behaviors is essential."

Risky behavior in our study was defined as having multiple sexual partners or engaging in unprotected sexual behaviors. Having more than 2 sexual partners was considered as having multiple sexual partners. Not using a condom every time during sex was considered unprotected sexual behavior, while consistent condom use was regarded as never engaging in unprotected sex.

Statistical Analysis

Data analysis was conducted using SPSS (version 19.0; IBM Corp). Chi-square test was used to analyze the differences in knowledge acquisition methods among people with different characteristics. Taking multiple sexual partners (no multiple sexual partners=0 and had multiple sexual partners=1) and unprotected sexual behavior (no unprotected sexual behavior=0 and had unprotected sexual behaviors=1) as dependent variables. The 3 forms of knowledge accesses, including school education, mass communication, and interpersonal communication, and HIV-related knowledge (containing fundamental knowledge and behavioral guidance knowledge of HIV), were used as the core independent variables. The effects of knowledge access and HIV-related knowledge on sexual behaviors were analyzed using binary logistic regression. Specifically, confounding factors such as gender, age, major, and grade were incorporated as covariates in the logistic regression models to adjust for their effects.

Process macro (version 4.1, model 4) [23] program compiled by Hayes [24] was used to construct multiple mediation models of "knowledge accesses-knowledge-risky behaviors" to explore how different knowledge accesses (school education, mass communication, and interpersonal communication) directly affect risky behaviors and how they indirectly affect risky behaviors through fundamental knowledge and behavioral guidance knowledge. Multiple mediation (model 4) analysis was based on bootstrap, with 5000 repeated sampling times. The confounding factors, such as gender, age, major, and grade, were controlled for analysis. The significance test was 2-sided, and the level of significance was set at $P<.05$.

Ethical Considerations

This study's protocol and informed consent procedures were approved by the Ethics Review Committee of the School of Public Health, Zhejiang University (ZGL202306-9). Before filling out the questionnaire, participants were informed about the purpose, process, potential risks, and benefits of the study, and it was voluntary to withdraw or end the investigation. No personally identifiable information was collected. Participants were provided with nonmonetary incentives, including free access to HIV/AIDS risk assessment tools and health consultation services.

Results

Sample Characteristics

The study population comprised 20,602 students. There were 9541 (46.31%) males and 11,061 (53.69%) females. The average age was 20.14 (SD 1.91) years. The number of freshmen, sophomores, juniors, and seniors was 10,339 (50.18%), 4679 (22.71%), 2532 (12.29%), and 974 (4.73%), respectively. A total of 2078 (10.09%) were graduate students. In terms of major, 7835 (38.03%) were majoring in medicine, 6729 (32.66%) were enrolled in science and technology programs, 4947 (24.01%) were studying literature, arts, or management, and 1091 (5.30%) were in vocational education. A total of 2423 students had sex. Among them, 363 (14.98%) had multiple sexual partners, and 830 (34.26%) engaged in unprotected sexual behaviors.

HIV-Related Knowledge and Knowledge Accesses

The average HIV-related knowledge score of 20,602 college students was 6.95 (SD 1.32). A total of 8485 (41.19%) students answered all questions correctly. Further, 125 (0.61%) students scored 0. The average score of fundamental knowledge was 3.26 (SD 0.88), and the average score of behavioral guidance knowledge was 3.69 (SD 0.69; Table 1).

XSL•FO
RenderX

Table . Distribution of responses to HIV-related knowledge among college students in China, a cross-sectional online survey in 2022 - 2023.

HIV-related knowledge	Correct count	Accuracy (%)
Fundamental knowledge		
AIDS is a serious and incurable infectious disease.	15,623	75.83
The main transmission mode of AIDS among young students in China is male homosexual, followed by heterosexual behavior.	15,860	76.98
A person infected with HIV can be identified by their appearance.	17,918	86.97
The rights of people living with HIV, including marriage, employment, and education, are protected by national law.	17,669	85.76
Behavioral guidance knowledge		
Can daily life and learning contact infect you with HIV?	18,499	89.79
Consistent and correct condom use reduces the risk of acquiring and transmitting HIV.	19,671	95.48
The use of new psychoactive substances increases the risk of HIV infection.	17,985	87.30
Seeking HIV testing and counseling after engaging in risky behaviors is essential.	19,894	96.56

The audience for mass communication included 19,030 (92.37%) students; 17,949 (87.12%) students received school education to acquire HIV-related knowledge; 11,274 (54.72%) obtained HIV-related knowledge via interpersonal communication, accounting for a relatively lower proportion. Further, 9645 (46.82%) students used all 3 forms of information channels to obtain knowledge, whereas only 2596 (12.60%) students relied on a single source for obtaining knowledge (Multimedia Appendix 1).

Distribution Characteristics of Knowledge Accesses of Young Students

Students with different genders, grades, ages, majors, and sexual behaviors exhibited varying preferences for information sources. Most of those relying on interpersonal communication were male (χ^2_1=4.19, P=.04). In terms of age, students under 20 primarily acquired knowledge through school education (χ^2_1=176.04, P<.01) and interpersonal communication (χ^2_1=82.67, P<.01), and the audience of mass communication (χ^2_1=25.48, P<.01) was mainly composed of students aged 20 years and older. Medical students were the main audience for both school education (χ^2_3=160.17, P<.01) and mass communication (χ^2_3=54.22, P<.01), whereas vocational students mostly preferred interpersonal communication (χ^2_3=58.72, P<.01). Students who never had sexual behavior mainly chose school education (χ^2_1=80.53, P<.01) and mass communication (χ^2_1=4.19, P=.04). In terms of risky sexual behaviors, those with multiple sexual partners tended to obtain HIV-related knowledge through interpersonal communication (χ^2_1=4.69, P=.03; Table 2).

XSL•FO
RenderX

Table . Chi-square test of HIV-related knowledge acquisition approaches among young students in China, a cross-sectional online survey in 2022 - 2023.

Items	Number	School education	Mass communication	Interpersonal communication
Gender				
Male	9541	8121 (85.12)[a]	8555 (89.67)[a]	5294 (55.49)[a]
Female	11,061	9828 (88.85)[a]	10,475 (94.70)[a]	5980 (54.06)[a]
χ^2 *(df)*		63.72 (1)	184.36 (1)	4.19 (1)
P		<.01	<.01	.04
Age (years)				
≤20	13,835	12,353 (89.29)[a]	12,689 (91.72)[a]	7876 (56.93)[a]
>20	6767	5596 (82.70)[a]	6341 (93.70)[a]	3398 (50.21)[a]
χ^2 *(df)*		176.04 (1)	25.48 (1)	82.67 (1)
P		<.01	<.01	<.01
Major				
Medicine	7835	7098 (90.59)[a]	7301 (93.18)[a]	4521 (57.70)[a]
Science and technology	6729	5650 (83.96)[a]	6193 (92.03)[a]	3541 (52.62)[a]
Literature, arts, or management	4947	4228 (85.47)[a]	4587 (92.72)[a]	2578 (52.11)[a]
Vocational education	1091	973 (89.18)[a]	949 (86.98)[a]	634 (58.11)[a]
χ^2 *(df)*		160.17 (3)	54.22 (3)	58.72 (3)
P		<.01	<.01	<.01
Grade				
Freshman	10,339	9289 (89.84)[a]	9421 (91.12)[a]	5945 (57.5)[a]
Sophomore	4679	4152 (88.74)[a]	4344 (92.84)[a]	2625 (56.1)[a]
Junior	2532	2107 (83.21)[a]	2395 (94.59)[a]	1309 (51.70)[a]
Senior	974	807 (82.85)[a]	907 (93.12)[a]	499 (51.23)[a]
Graduate student	2078	1594 (76.71)[a]	1963 (94.47)[a]	896 (43.12)[a]
χ^2 *(df)*		330.30 (4)	55.78 (4)	162.87 (4)
P		<.01	<.01	<.01
Having sexual intercourse?				
Yes	2423	1972 (81.39)[a]	2213 (91.33)[a]	1358 (56.05)[a]
No	18,179	15,977 (87.89)[a]	16,817 (92.51)[a]	9916 (54.55)[a]
χ^2 *(df)*		80.53 (1)	4.19 (1)	1.94 (1)
P		<.01	.04	.16
Having multiple sexual partners?				
Yes	363	281 (77.41)[a]	313 (86.23)[a]	219 (60.33)[a]
No	20,239	17,668 (87.3)[a]	18,717 (92.48)[a]	11,055 (54.62)[a]
χ^2 *(df)*		31.07 (1)	19.79 (1)	4.69 (1)
P		<.01	<.01	.03
Having unprotected sex?				
Yes	830	673 (81.08)[a]	747 (90.00)[a]	478 (57.59)[a]

XSL•FO
RenderX

Items	Number	School education	Mass communication	Interpersonal communication
No	19,772	17,276 (87.38)[a]	18,283 (92.47)[a]	10,796 (54.60)[a]
χ^2 *(df)*		28.11 (1)	6.89 (1)	2.87 (1)
P		<.01	<.01	.09

[a]n (%).

Association Between Knowledge Accesses, HIV-Related Knowledge, and Multiple Sexual Partner Behaviors

According to multivariate logistic regression analysis, compared with students who did not acquire knowledge through school education and mass media, those who acquired knowledge through school education (odds ratio [OR] 0.60, 95% CI 0.46 to 0.78) and mass media (OR 0.64, 95% CI 0.47 to 0.87) were less likely to have multiple sexual partners. Students who acquired knowledge through interpersonal communication (OR 1.41, 95% CI 1.13 to 1.75) were more likely to have multiple sexual partners than those who did not. A higher level of behavioral guidance knowledge (OR 0.79, 95% CI 0.69 to 0.90) was associated with a lower likelihood of having multiple sexual partners; the level of fundamental knowledge had no significant effect on the behavior of having multiple sexual partners (Table 3).

Table . Association between knowledge accesses, HIV-related knowledge, and multiple sexual partner behaviors among Chinese college students[a].

Variables	B	SE	Wald	*P* value	OR (95% Cl)
Knowledge accesses					
School education	−0.51	0.13	14.94	<.01	0.60 (0.46, 0.78)
Mass communication	−0.45	0.16	8.05	<.01	0.64 (0.47, 0.87)
Interpersonal communication	0.34	0.11	9.33	<.01	1.41 (1.13, 1.75)
HIV-related knowledge					
Fundamental knowledge	0.11	0.06	2.69	.10	1.11 (0.98, 1.26)
Behavioral guidance knowledge	−0.24	0.07	12.53	<.01	0.79 (0.69, 0.90)

[a]Logistic regression analysis was conducted to examine the relationship between having multiple sexual partners (independent variable) and knowledge accesses and levels (dependent variables), controlling for gender, age, major, and grade. Hosmer-Lemeshow test *P*>.5.

Association Between Knowledge Accesses, HIV-Related Knowledge, and Unprotected Sexual Behaviors

The regression results showed that students who acquired knowledge through school education (OR 0.75, 95% CI 0.62 to 0.91) were less likely to engage in unprotected sexual behaviors than those who did not; students who obtained knowledge through interpersonal communication (OR 1.27, 95% CI 1.09 to 1.46) were more likely to engage in unprotected sexual behaviors. There was no significant association between mass media exposure and unprotected sex behaviors. Fundamental knowledge (OR 1.17, 95% CI 1.08 to 1.28) was associated with an increased likelihood of engaging in unprotected sexual behaviors. Behavioral guidance knowledge (OR 0.81, 95% CI 0.73 to 0.89) was negatively correlated with unprotected sexual behaviors (Table 4).

XSL•FO
RenderX

Table . Association of knowledge accesses, HIV-related knowledge, and unprotected sexual behaviors among Chinese college students.[a]

Variables	B	SE	Wald	*P* value	OR (95% Cl)
Knowledge accesses					
School education	−0.29	0.1	8.95	<.01	0.75 (0.62, 0.91)
Mass communication	−0.19	0.12	2.44	.12	0.83 (0.65, 1.05)
Interpersonal communication	0.24	0.07	9.96	<.01	1.27 (1.09, 1.46)
HIV-related knowledge					
Fundamental knowledge	0.16	0.05	12.73	<.01	1.17 (1.08, 1.28)
Behavioral guidance knowledge	−0.21	0.05	18.47	<.01	0.81 (0.73, 0.89)

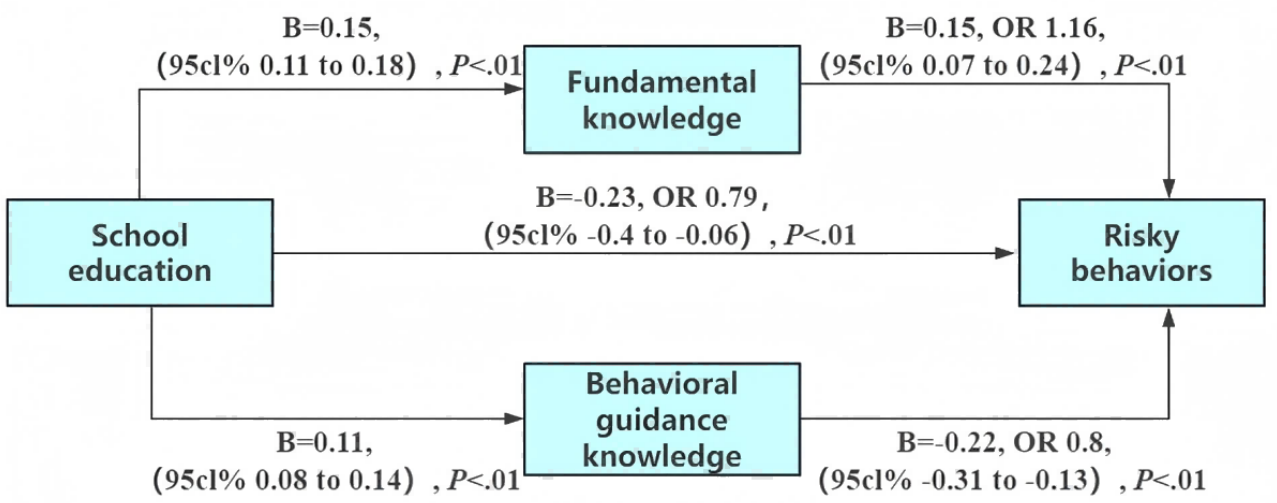
[a]Logistic regression analysis was conducted to examine the relationship between having unprotected sex (independent variable) and knowledge accesses and levels (dependent variables), controlling for gender, age, major, and grade. Hosmer-Lemeshow test *P*>.5.

Mediating Effect of HIV-Related Knowledge on the Correlation Between Knowledge Accesses and Risky Behaviors

The results of multiple mediation models showed that both fundamental knowledge and behavioral guidance knowledge of HIV acted as joint mediators in the relationship between various knowledge accesses (school education, mass communication, and interpersonal communication) and risky behaviors.

The regression results indicated that school education was positively associated with behavioral guidance knowledge (B=0.11, *P*<.01) and fundamental knowledge (B=0.15, *P*<.01). When simultaneously predicting risky behaviors with school education, behavioral guidance knowledge, and fundamental knowledge, both school education (B=−0.23, OR =0.79, 95% CI −0.40 to −0.06, *P*<.01) and behavioral guidance knowledge (B=−0.22, OR 0.8, 95% CI −0.31 to −0.13, *P*<.01) were negatively associated with risky behaviors, while fundamental knowledge was positively associated with risky behaviors (B=0.15, OR 1.16, 95% CI 0.07 to 0.24, *P*<.01; Figure 1). The indirect effects of behavioral guidance knowledge and fundamental knowledge were −0.02 and 0.02, respectively. The total indirect effect was nonsignificant, suggesting that the opposing effects of the 2 mediating pathways may offset each other (Table S3 in Multimedia Appendix 2).

Figure 1. Mediation model illustrating the impact of school education on risky behaviors among Chinese students via fundamental knowledge and behavioral guidance knowledge. Path diagram illustrating the dual mediation mechanism through fundamental knowledge (**M1**) and behavioral guidance knowledge (**M2**) between school education (**X**) and risky behaviors (**Y**). Values on the arrows represent unstandardized regression coefficients (**B**). ORs and their 95% CIs are also reported. The model controlled for gender, age, major, and grade. OR: odds ratio.

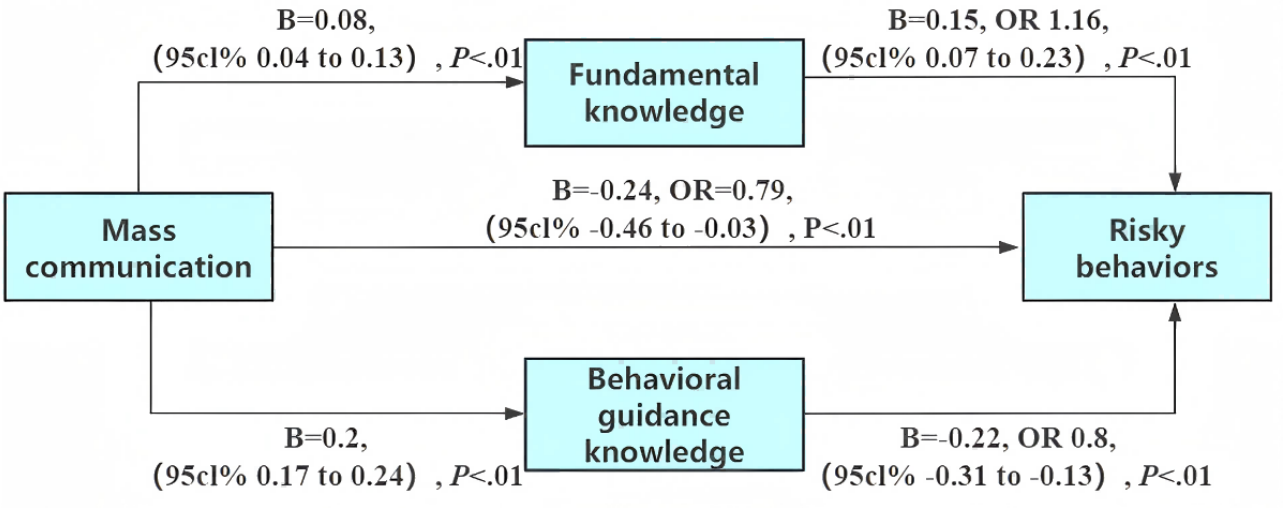


Mass communication was positively associated with behavioral guidance knowledge (B=0.2, *P*<.01) and fundamental knowledge (B=0.08, *P*<.01). When predicting risky behaviors with mass communication, behavioral guidance knowledge, and fundamental knowledge, both mass communication (B=−0.24, OR 0.79, 95% CI −0.46 to −0.03, *P*<.01) and behavioral guidance knowledge (B=−0.22, OR 0.8, 95% CI −0.31 to −0.13, *P*<.01) were negatively associated with risky behaviors, while fundamental knowledge was positively associated (B=0.15, OR 1.16, 95% CI 0.07 to 0.23, *P*<.01; Figure 2). The indirect effects of behavioral guidance knowledge and fundamental knowledge

XSL•FO
RenderX

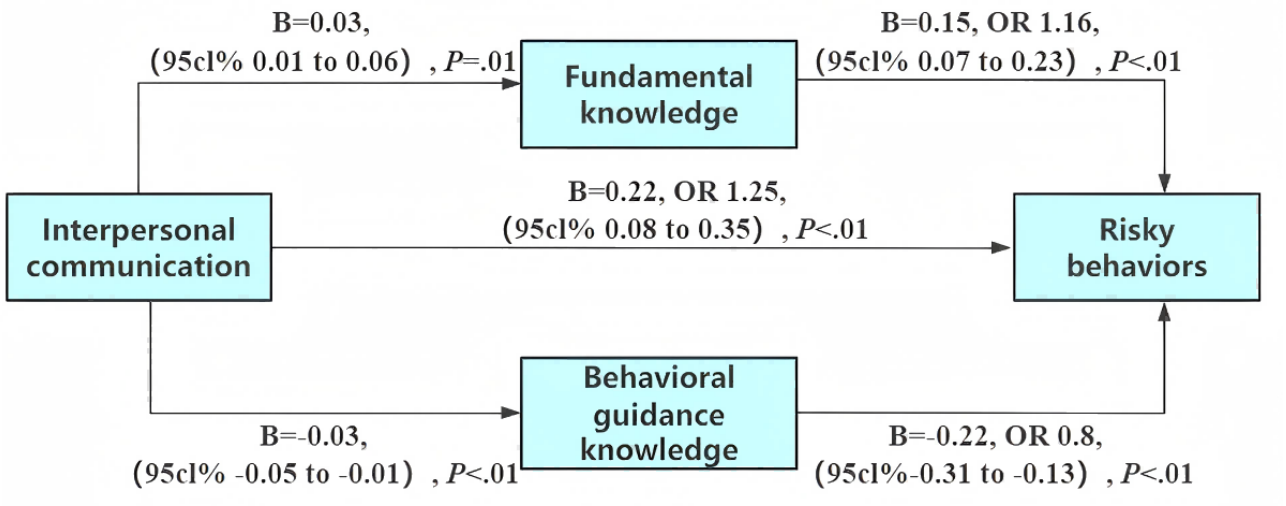
were −0.04 and 0.01, respectively (Table S5 in Multimedia Appendix 2).

Figure 2. Mediation model illustrating the impact of mass communication on risky behaviors among Chinese students via fundamental knowledge and behavioral guidance knowledge. Path diagram illustrating the dual mediation mechanism through fundamental knowledge (**M1**) and behavioral guidance knowledge (**M2**) between mass communication (**X**) and risky behaviors (**Y**). Values on the arrows represent unstandardized regression coefficients (**B**). ORs and their 95% CIs are also reported. The model controlled for gender, age, major, and grade. OR: odds ratio.



Interpersonal communication was negatively associated with behavioral guidance knowledge (B=−0.03, *P*<.01) and positively associated with fundamental knowledge (B=0.03, *P*<.01). Interpersonal communication showed a direct positive association with risky behaviors (B=0.22, OR 1.25, 95% CI 0.08 to 0.35, *P*<.01), while both behavioral guidance knowledge (B=−0.22, OR 0.8, 95% CI −0.31 to −0.13, *P*<.01) and fundamental knowledge (B=0.15, OR 1.16, 95% CI 0.07 to 0.23, *P*<.01) demonstrated significant effects on risky behaviors. Interpersonal communication directly had a positive impact on risky behaviors and increased risky behaviors by improving fundamental knowledge and reducing behavioral guidance knowledge (Figure 3). The indirect effects of behavioral guidance knowledge and fundamental knowledge were 0.01 and 0.01, respectively (Table S7 in Multimedia Appendix 2).

Figure 3. Mediation model illustrating the impact of interpersonal communication on risky behaviors among Chinese students via fundamental knowledge and behavioral guidance knowledge. Path diagram illustrating the dual mediation mechanism through fundamental knowledge (**M1**) and behavioral guidance knowledge (**M2**) between interpersonal communication (**X**) and risky behaviors (**Y**). Values on the arrows represent unstandardized regression coefficients (**B**). ORs and their 95% CIs are also reported. The model controlled for gender, age, major, and grade. OR: odds ratio.



Discussion

Principal Findings

Based on the investigation of 20,602 Chinese university students, this cross-sectional study revealed distinct associations between HIV knowledge, access channels, and risky sexual behaviors (unprotected sex and multiple partners). While school education and mass media access were associated with reduced risky sexual behaviors, interpersonal communication correlated with increased risk. What is more, higher HIV behavioral guidance knowledge was linked to safer behaviors, whereas fundamental knowledge paradoxically was associated with more risky behaviors. Mediation analysis indicated that the direct

XSL•FO
RenderX

influence of knowledge accesses, likely via shaping norms, was more dominant than their indirect effects through knowledge transmission.

A score of 6 or more is considered to be up to the standard for HIV health education knowledge [22,25]. In this survey, 89.80% (18,501/20,602) of the students answered more than 6 questions correctly. There is still a long way to go to meet the goal of over 95% of young students being aware of HIV, as outlined in the Implementation Plan for Curbing the Spread of HIV/AIDS (2019‐2022). College students mainly obtain HIV-related information through mass media and school education, which aligns with previous studies conducted in China [26-28]. This may be attributed to the rapid development of the internet, allowing young adults to easily obtain various information about HIV/AIDS through search engines, social media, and other platforms [29,30]. School, being central to students' daily lives, offers a structured environment conducive to HIV educational campaigns, which has the advantage of centralized publicity. Therefore, mass communication and school education are always favored by students. However, fewer students reported acquiring HIV-related knowledge through interpersonal communication, which may be affected by Chinese traditional culture. Sexual content is often regarded as an obscure and sensitive topic, and parents or friends will deliberately avoid such topics. Similar patterns, rooted in shared East Asian cultural influences, which are culturally conservative and emphasize collectivism, have been observed in studies on Korean youths, who also primarily obtain HIV knowledge from mass media and school, rather than interpersonal channels [31]. However, a substantial body of Western literature documents the importance of parent-child communication and peer interaction in preventing risky sexual behaviors [32-34]. Cultural values can influence the sources of HIV information sought by young people, with collectivist and conservative cultures potentially prioritizing indirect or institutional channels over direct, potentially confrontational, and interpersonal dialogues.

Access to HIV-related knowledge through school education and mass media may have a positive association with a reduction in behaviors such as having multiple sexual partners and engaging in unprotected sex, as indicated by our regression analysis. As a vital form of organizational communication, school education holds a high level of authority and credibility, which is embodied in the strict censorship system of knowledge dissemination, so young students are inclined to trust the information they receive in school. Additionally, organizational communication may be mandatory. For instance, schools can introduce compulsory HIV education courses or make it mandatory for students to attend relevant lectures, which can break through the personal "information cocoons" to some extent, where students might otherwise only seek information aligned with their existing interests or biases. As a result, students are exposed to more comprehensive and accurate HIV-related knowledge than they might seek out on their own. The mass media, including newspapers, radio, television, and the internet, also have unique advantages. Mass media is characterized by its speed of transmission and the breadth of its content, allowing it to rapidly achieve widespread information coverage. The ability of mass media to disseminate knowledge in engaging and entertaining formats helps overcome the discomfort associated with "face-to-face" discussions on sensitive topics, making it easier for students to accept and retain the information.

The association between school education and risky behaviors is stronger than that of mass media, which is consistent with the research of Chinese scholars [35]. Unlike online platforms, where information can be overwhelming and fragmented, school curricula are designed to be comprehensive and progressive, allowing students to build their understanding step by step. Additionally, the accuracy of information from mass media is often difficult to guarantee. While the internet has become the primary medium through which young students access information, its less regulated nature requires students to exercise caution. The sheer volume of content, along with the prevalence of unreliable or false information, underscores the importance of seeking official and authoritative channels to acquire accurate HIV knowledge. This preference for structured and authoritative sources, along with their effectiveness, is deeply rooted in China's collectivist culture [36], particularly its emphasis on institutional trust, hierarchical order, and the role of established authority in knowledge dissemination. As state-sanctioned institutions, schools provide a reliable and systematic avenue for addressing sensitive topics such as sexual health. Culturally, students are conditioned to view knowledge imparted by teachers and formal curricula as vetted and trustworthy.

Mass communication did not have a significant effect on unprotected sex, but did have a significant effect on the reduction of multiple sexual partners. When disseminating information, mass media always tailor content and format to align with the public's interests and concerns. Compared to topics such as "correct condom use during sex," the issue of multiple sexual partners is much more provocative and sensitive, which can arouse the curiosity of college students. It is suggested that mass media should not only consider the attractiveness of the content, but also shoulder the social responsibility of providing more comprehensive coverage of HIV/AIDS prevention.

By regression analysis, it was found that students who acquired knowledge through interpersonal communication were more likely to have multiple sexual partners and engage in unprotected sex. This may be due to the lack of clarity or comprehensiveness in the information disseminated through personal networks, which undermines effective risk perception and the adoption of protective behaviors. In the mediation analysis, we found that interpersonal communication negatively predicted behavioral guidance knowledge, which may be able to support this view. First, HIV-related knowledge acquired from family members or friends is relatively superficial and unsystematic [29]. Knowledge transmitted may be based on personal understanding, hearsay, or limited media information, making it difficult to guarantee the accuracy and comprehensiveness. The widespread cultural constraints in East Asian societies have further limited the acquisition of knowledge through interpersonal communication, where sexuality is often considered a sensitive and taboo topic [34,37]. The cultural stigma makes open and direct discussion about sexual health challenging within family

XSL•FO
RenderX

and peer circles. Even when discussions occur, the focus tends to be on the moral consequences or general harms of the disease, rather than on practical prevention methods, largely due to discomfort with explicit sexual discourse. Sensitive topics, such as condom use or specific sexual practices, might be perceived as impolite and embarrassing. Family members or friends, constrained by these cultural norms, might default to abstract moral warnings or general admonitions rather than concrete and actionable prevention strategies. In Latino families that emphasize conservative culture, studies have also pointed out that Latinos are less willing to provide children with information about sex and AIDS, telling them to protect themselves, but not giving them specific information on how to do it [38]. While the intention may be to guide young students toward safe sexual behavior, the cultural barriers to open dialogue create a critical gap between the desired outcome and the actual knowledge and skills acquired.

Moreover, the efficacy of knowledge acquisition through interpersonal communication is often undermined by personal filters. Emotional factors frequently mediate the transmission of information, leading to its simplification or distortion. When receiving HIV knowledge from a close friend, the emotional bond might lead to a downplaying of risks. Casual language or humor used to address the topic can create a false sense of security and diminish the perceived seriousness of HIV's harms. The personal filtering effect is amplified in cultural contexts that prioritize social harmony and discomfort avoidance [39]. Out of a desire to protect the listener from anxiety or maintain normalcy, friends or family members may inadvertently or intentionally simplify or downplay the risks of HIV. Furthermore, within Chinese culture, the concept of filial piety may discourage younger individuals from questioning or challenging potentially inaccurate information from parents or older adults, even if they intuitively perceive discrepancies. It is recommended that young students, especially the main audience group of interpersonal communication, consult HIV curricula in schools or official websites to avoid being misled by false information when discussing HIV with parents, relatives, or friends. When acting as communicators, young students must ensure the accuracy of the information, prioritize the transmission of behavioral guidance knowledge, and treat the topic with appropriate seriousness.

However, this study could not determine whether the acquisition of HIV-related knowledge through interpersonal communication occurs in succession with risky behaviors, so it could not clarify whether interpersonal communication promotes the occurrence of risky sexual behaviors. As mentioned above, sexual topics are inherently highly sensitive and difficult to discuss openly in China. Individuals who proactively engage in interpersonal communication about HIV information are likely to be those who have already engaged in high-risk behaviors, and similar views can be found [40]. Their exchanges are more often based on personal experiences or seeking social support, rather than systematically acquiring scientific prevention knowledge. This tendency to seek social support after engaging in risky behaviors may explain why acquiring knowledge through interpersonal communication is positively correlated with risky sexual behaviors.

According to the "Knowledge, Attitude, and Practice" model, knowledge is the foundation of behavior change, but a high level of HIV knowledge does not always lead to safe behaviors, reflecting the phenomenon of "separation." This study found that a higher level of fundamental knowledge of HIV may be related to more risky sexual behaviors, while behavioral guidance knowledge plays a positive role in reducing such behaviors. The findings suggest that different types of knowledge have varying effects on risky behaviors. Fundamental knowledge provides a broad and foundational understanding of HIV, but may not effectively translate into safe sexual practices. In fact, a strong grasp of fundamental knowledge may foster a sense of overconfidence or "risk underestimation," causing students to lower their vigilance during sexual behaviors. Additionally, students may misunderstand the legal protection for HIV-positive individuals, regarding such policies as shelters for risky sexual behaviors. Consequently, fundamental knowledge may be insufficient to reduce risky behaviors alone. Behavioral guidance knowledge is relatively more specific and practical, directly guiding individuals to adopt HIV prevention measures, such as correctly using condoms or knowing when and where to get tested after potential exposure. Unlike fundamental knowledge, which may raise awareness but leave gaps in practical application, behavioral guidance knowledge equips students with concrete tools to lower their risk and facilitates the translation of knowledge into safe behaviors. Therefore, the focus should be on strengthening the dissemination of HIV behavioral guidance knowledge. However, it does not imply that we should reduce the publicity of HIV fundamental knowledge, which serves as the foundation for building a comprehensive understanding of HIV. HIV health education should emphasize behavioral guidance knowledge while constantly reinforcing fundamental knowledge [41], which enables college students to gain a more systematic and comprehensive grasp of HIV information. By combining a solid foundation of fundamental knowledge with a stronger focus on prevention, young students will be better positioned to apply information in real-life situations, significantly lowering their likelihood of engaging in risky sexual behaviors.

Based on the results of the mediation analysis, it was found that school education, mass communication, and interpersonal communication all influence risky behaviors through both behavioral guidance knowledge and fundamental knowledge. However, the overall mediation effects were relatively small, and the influence paths varied across these channels. It suggests that the impact of these knowledge accesses on behavior may not primarily operate through knowledge acquisition, but is instead driven more by its direct influence. Media systems serve not only as tools for knowledge transmission but also as sites for the production and reproduction of cultural capital, influencing individuals by shaping cultural values and behavioral norms. School education's influence on risky behaviors may arise less from knowledge acquisition and more from the structured and organized environment it provides, where students may feel external pressures or benefit from behavioral norms enforced by the school setting. Interpersonal communication easily fosters an open atmosphere for young students, making students more likely to have a fluke or curiosity, and ultimately leading to a higher likelihood of

XSL•FO
RenderX

engaging in risky behaviors. The indirect effect of mass communication on reducing risky behaviors by improving the knowledge level of individuals is relatively strong, especially through the enhancement of behavioral guidance knowledge. It may be due to the voluntary nature of information seeking in mass communication that highlights the importance of knowledge acquisition initiatives. This personalized, needs-driven communication approach is more effective than passive teaching methods in influencing behavior by increasing the level of knowledge.

Limitations

The study is subject to some limitations. First, while the cross-sectional design revealed associations between knowledge access channels and risky behaviors, our study could not explore the causality of knowledge accesses and risky behaviors. Future research should use longitudinal or quasi-experimental designs to clarify causality. Second, the generalizability of the findings in this study is limited by its scope. The level of economic development and the open environment in Zhejiang province may have led to an overestimation of students' HIV knowledge and their openness to discussing the topic. The results of this study may also be influenced by cultural context; while Zhejiang province is somewhat representative of broader Chinese values, caution is required when directly applying these findings to all university student populations or to distinctly different cultural contexts, given the inherent geographical and cultural differences. This also underscores the need for future cross-cultural research to further analyze and explore effective approaches to HIV education across diverse cultural backgrounds. Third, by combining the 8 questions for measuring HIV-related knowledge into 2 dimensions, we may not have fully explored the intrinsic connections among the questions. Future research should develop more refined knowledge classification systems and measurement tools, ensuring a sufficient number of items to capture the nuanced effects of different knowledge types on behavior. Moreover, considering the sensitivity of sexual behaviors, the young students may provide social desirability responses, which may influence the results. Finally, for statistical analysis, more demographic variables deserve to be included, such as sexual orientation and family economic status. Interaction effects of variables can provide valuable insights into the conditional relationships between key factors, which future research could focus on.

Conclusions and Policy Implications

This study highlights that HIV prevention education should strategically combine the strengths of different transmission pathways. Specifically, it should leverage the authority and structured nature of school-based programs while capitalizing on the broad reach of mass media. Guidance should be provided for schools to reform educational content and formats, emphasizing the encouragement of voluntary student participation rather than mandatory intervention. Online public discourse should be monitored, and the government must intervene to correct misinformation that misleads students. To address the potential risks of informal interpersonal communication, it is crucial to enhance its quality. A knowledge certification system for peer educators should be established, and student leaders should be trained to disseminate accurate prevention information. Additionally, interpersonal communication quality should be included in HIV/AIDS prevention assessment metrics. Regarding knowledge structure, we advocate prioritizing actionable, practical HIV/AIDS behavioral guidance knowledge, focusing on risk-reduction skills to bridge the knowledge-behavior gap. Transmission pathways not only convey information but also shape behavioral norms. Schools can use rituals such as anti-AIDS pledge ceremonies to strengthen students' responsibility, while mass media can reshape perceptions and promote supportive environments for behavior change.

Acknowledgments

We are grateful to all study participants for their valuable participation.

Data Availability

All of the main data generated or analyzed during this study are included in this published paper in the Results section. Additional materials with details generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

All authors were responsible for the structure of this paper. YZ conducted the data analysis and drafted this paper. JX contributed to this study conception and design, data interpretation, and critical revision of this paper. JJ and FC contributed to this study's design and provided additional support. JX and XZ served as co-corresponding authors. Contact details for co-corresponding author XZ: xingliangyufang@163.com; Hangzhou Center for Disease Control and Prevention (Hangzhou Health Supervision Institution), No. 568, Mingshi Road, Jianqiao Street, Shangcheng District, Hangzhou, Zhejiang, China, 310020. All authors approved the final version of this paper for submission.

The research was supported by the Fundamental Research Funds for the Central Universities and Zhejiang Province Research Project on youths and youth work (ZQ2025020).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Distribution of AIDS knowledge acquisition approaches among college students in China, a cross-sectional online survey in 2022-2023.
[PNG File, 199 KB - publichealth_v11i1e68339_app1.png]

Multimedia Appendix 2
Comprehensive overview of sample characteristics and mediation analysis results.
[DOCX File, 34 KB - publichealth_v11i1e68339_app2.docx]

References

1. 2024 global AIDS report — the urgency of now: AIDS at a crossroads. : UNAIDS; 2024 URL: https://www.unaids.org/sites/default/files/media_asset/2024-unaids-global-aids-update_en.pdf [accessed 2025-08-02]
2. Han MJ. Analysis on HIV/AIDS epidemic situation and prospect of prevention and control in China. Chin J AIDS STD 2023;29(3):247-250. [doi: 10.13419/j.cnki.aids.2023.03.01]
3. Li GQ, Jiang Y, Zhang LQ. HIV upsurge in China's students. Science 2019 May 24;364(6442):711-711. [doi: 10.1126/science.aay0799]
4. Wu ZY. Characteristics of HIV sexually transmission and challenges for controlling the epidemic in China. Chin J Epidemiol 2018;39(6):707-709. [doi: 10.3760/cma.j.issn.0254-6450.2018.06.002]
5. Fineberg HV. Education to prevent AIDS: prospects and obstacles. Science 1988 Feb 5;239(4840):592-596. [doi: 10.1126/science.3340845] [Medline: 3340845]
6. China NHC. Implementation plan to curb the spread of AIDS. 2019 URL: https://www.gov.cn/xinwen/2019-10/13/content_5439036.htm [accessed 2025-08-07]
7. Palupi R, Purwanto B, Wuryanto T. Role-playing method for students' interpersonal communication improvement. Khatulistiwa: JPendidikan SosHumaniora 2023;3:259-270. [doi: 10.55606/khatulistiwa.v3i2.2085]
8. Yu D, Wang LL. Analysis on the communication channels and effects of major events in China in 2009- based on a survey of urban residents in Beijing, Shanghai, and Guangzhou. Mod Commun (J Commun Univ China) 2010(9):35-41. [doi: 10.19997/j.cnki.xdcb.2010.09.008]
9. Wang D, Shi ZP. The influence of communication channels exposure on the public's infectious disease protection behavior: mediated by health beliefs. Journalism Commun 2023;30(8):63-79.
10. Daft RL, Lengel RH. Organizational information requirements, media richness and structural design. Manage Sci 1986 May;32(5):554-571. [doi: 10.1287/mnsc.32.5.554]
11. Tung WC, Cook DM, Lu M, Ding K. A comparison of HIV knowledge, attitudes, and sources of STI information between female and male college students in Taiwan. Health Care Women Int 2015;36(8):870-882. [doi: 10.1080/07399332.2014.962136] [Medline: 25271399]
12. Abiona T, Balogun J, Yohannes E, et al. HIV/AIDS knowledge, perception of knowledge and sources of information among university students in USA, Turkey, South Africa and Nigeria. Health Educ J 2014 Nov;73(6):755-767. [doi: 10.1177/0017896914550943]
13. Nwakasi C, Esiaka D, Staab T, Philip AA, Nweke C. HIV knowledge and information access among women cancer survivors in Nigeria. J Cancer Policy 2024 Mar;39:100456. [doi: 10.1016/j.jcpo.2023.100456] [Medline: 37989454]
14. Niu WY. Health Communication. Chin J Health Educ 2004;04:43-46.
15. Nwagwu W. Use of HIV/AIDS information sources and sexual behaviour of adolescents in a rural community in Southeastern Nigeria. Malays J Library Inf Sci 2017;17(1):1-19 [FREE Full text]
16. Tung WC, Cook DM, Lu M, Yang W. HIV knowledge and behavior among Chinese college students in China and the United States. West J Nurs Res 2013 Oct;35(9):1171-1183. [doi: 10.1177/0193945913486201] [Medline: 23618822]
17. Del Ponte A, Li L, Ang L, Lim N, Seow WJ. Evaluating SoJump.com as a tool for online behavioral research in China. J Behav Exp Finance 2024 Mar;41:100905. [doi: 10.1016/j.jbef.2024.100905]
18. Purohit N. Health communication and behavior change. In: Gupta SD, editor. Healthcare System Management: Methods and Techniques: Singapore, Springer Nature Singapore; 2022:295-312. [doi: 10.1007/978-981-19-3076-8_12]
19. Zhang L, Yu H, Luo H, et al. HIV/AIDS-related knowledge and attitudes among Chinese college students and associated factors: a cross-sectional study. Front Public Health 2021;9:804626. [doi: 10.3389/fpubh.2021.804626] [Medline: 35096751]
20. Lai XH, et al. Status of knowledge and behavior on AIDS among college students in Zhongshan. Chin J Sch Health 2017;38(10):1558-1559. [doi: 10.4103/0253-7184.196888]
21. China CDC. Notice on the issuance of the questionnaire on the core knowledge of AIDS publicity and education and the awareness rate of AIDS knowledge. 2016. URL: https://ncaids.chinacdc.cn/zlk/cllb/xcy/xcdhbzy/202112/P020211221500983637737.pdf [accessed 2025-08-02]

XSL•FO
RenderX

22. China CDC. Core information on HIV/AIDS prevention education for young students. 2021. URL: https://view.officeapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.chinaaids.cn%2Fzxzx%2Fzxdteff%2F202112%2FW020211207416031984780.docx&wdOrigin=BROWSELINK [accessed 2025-08-02]
23. The PROCESS macro for SPSS, SAS, and R. URL: https://www.processmacro.org/index.html [accessed 2025-08-07]
24. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach, 3rd edition: Guilford Press; 2022.
25. Ma Z, Guo L, Zhou M, Zuo H. HIV/AIDS-related knowledge and attitudes towards HIV rapid testing among Chinese college students: findings from a cross-sectional survey. Prev Med Rep 2023 Dec;36:102409. [doi: 10.1016/j.pmedr.2023.102409] [Medline: 37719792]
26. Liu H, et al. Analysis on AIDS awareness and ways to acquire AIDS knowledge in middle school students in Zhangbei county. Mod Prev Med 2014;41(1):37-40 [FREE Full text]
27. Zhang YP. Investigation and evaluation on AIDS knowledge, attitude and practice of students in a university of Shenyang City, Liaoning Province. Chin J Health Educ 2011;27(8):596-599. [doi: 10.16168/j.cnki.issn.1002-9982.2011.08.003]
28. Tan Y, Yang J, Cui Y, et al. Knowledge, attitudes and practices related to AIDS amongst transgender women in Chongqing, China: a cross-sectional study. BMC Public Health 2024 Sep 16;24(1):2522. [doi: 10.1186/s12889-024-19901-z] [Medline: 39285344]
29. Wang XX. Changes of HIV knowledge and attitude among students in Beijing from 2018 to 2022. Chin J Sch Health 2024;45(1):51-54. [doi: 10.16835/j.cnki.1000-9817.2024023]
30. Obeagu EI, Obeagu GU, Ede MO, Odo EO, Buhari HA. Translation of HIV/AIDS knowledge into behavior change among secondary school adolescents in Uganda: a review. Medicine (Baltimore) 2023 Dec 8;102(49):e36599. [doi: 10.1097/MD.0000000000036599] [Medline: 38065920]
31. Yoo H, Lee SH, Kwon BE, Chung S, Kim S. HIV/AIDS knowledge, attitudes, related behaviors, and sources of information among Korean adolescents. J Sch Health 2005 Dec;75(10):393-399. [doi: 10.1111/j.1746-1561.2005.00054.x] [Medline: 16313510]
32. Fasula AM, Miller KS. African-American and Hispanic adolescents' intentions to delay first intercourse: parental communication as a buffer for sexually active peers. J Adolesc Health 2006 Mar;38(3):193-200. [doi: 10.1016/j.jadohealth.2004.12.009] [Medline: 16488815]
33. Paikoff RL, Parfenoff SH, Williams SA, McCormick A, Greenwood GL, Holmbeck GN. Parenting, parent–child relationships, and sexual possibility situations among urban African American preadolescents: preliminary findings and implications for HIV prevention. J Fam Psychol 1997;11(1):11-22. [doi: 10.1037/0893-3200.11.1.11]
34. Tinsley BJ, Lees NB, Sumartojo E. Child and adolescent HIV risk: familial and cultural perspectives. J Fam Psychol 2004 Mar;18(1):208-224. [doi: 10.1037/0893-3200.18.1.208] [Medline: 14992622]
35. Hong X, et al. Investigation on health education demand and knowledge about AIDS in medical college students. Chin J Public Health 2007;09:1036-1037. [doi: 10.11847/zgggws2007-23-09-06]
36. Meng M, Li X, Zhao J, Hao Y. When western concept meets eastern culture: exploring the impact of Confucianism on shared decision-making in China. Asia Pac J Oncol Nurs 2024 Nov;11(11):100586. [doi: 10.1016/j.apjon.2024.100586] [Medline: 39444586]
37. Gao E, Zuo X, Wang L, Lou C, Cheng Y, Zabin LS. How does traditional Confucian culture influence adolescents' sexual behavior in three Asian cities? J Adolesc Health 2012 Mar;50(3 Suppl):S12-S17. [doi: 10.1016/j.jadohealth.2011.12.002] [Medline: 22340851]
38. Lescano CM, Brown LK, Raffaelli M, Lima LA. Cultural factors and family-based HIV prevention intervention for Latino youth. J Pediatr Psychol 2009;34(10):1041-1052. [doi: 10.1093/jpepsy/jsn146] [Medline: 19181820]
39. Badanta B, González-Cano-Caballero M, Suárez-Reina P, Lucchetti G, de Diego-Cordero R. How does Confucianism influence health behaviors, health outcomes and medical decisions? A scoping review. J Relig Health 2022 Aug;61(4):2679-2725. [doi: 10.1007/s10943-022-01506-8] [Medline: 35141796]
40. Duong HT, Nguyen LTV, McFarlane SJ, Nguyen HT, Nguyen KT. Preventing the COVID-19 outbreak in Vietnam: social media campaign exposure and the role of interpersonal communication. Health Commun 2023 Feb;38(2):394-401. [doi: 10.1080/10410236.2021.1953729] [Medline: 34278892]
41. Wang JJ, Chen QF, Han MJ. Survey on AIDS knowledge awareness of male college students in 7 universities in China. Chin J AIDS STD 2020;26(5):493-495. [doi: 10.13419/j.cnki.aids.2020.05.10]

Abbreviations

OR: odds ratio

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 04.11.24; peer-reviewed by A Onzia, M Beltrão; revised version received 04.07.25; accepted 06.07.25; published 29.08.25.

Please cite as:
Zhang Y, Jin J, Cheng F, Zhang X, Xu J
Impact of Knowledge Access on Risky Sexual Behaviors Among Chinese Youths to Improve HIV Prevention: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e68339
URL: https://publichealth.jmir.org/2025/1/e68339
doi:10.2196/68339

© Yun Zhang, Jie Jin, Feng Cheng, Xingliang Zhang, Junfang Xu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

HIV Incidence and Associated Risk Factors Among Young Men Who Have Sex With Men in Tianjin, China: Retrospective Cohort Study

Yan Guo[1*], BMed; Hui Gong[1*], MSc; Xiangyu Yan[2], PhD; Xinying Zhang[3], BMed; Tielin Ning[1], BMed; Zhongquan Liu[1], BMed; Long Li[1], MSc; Jie Yang[4]; Liu Li[1], BMed; Changxu Ma[1], BMed; Zhuang Cui[5], PhD; Maohe Yu[1], MSc

[1]Tianjin Centers for Disease Control and Prevention, 6 Hua Yue Road, Hedong District, Tianjin, China

[2]School of Disaster and Emergency Medicine, Tianjin University, Tianjin, China

[3]Tianjin Hongqiao Center for Disease Control and Prevention, Tianjin, China

[4]Shenlan Public Health Consulting Service Center, Tianjin, China

[5]Department of Epidemiology and Biostatistics, Tianjin Medical University, Tianjin, China

[*]these authors contributed equally

Corresponding Author:
Maohe Yu, MSc
Tianjin Centers for Disease Control and Prevention, 6 Hua Yue Road, Hedong District, Tianjin, China

Abstract

Background: Young men who have sex with men (YMSM) have a higher risk of HIV infection. However, evidence of HIV incidence from large cohort studies is limited in this key population, particularly among Chinese YMSM.

Objective: This study aimed to investigate the HIV incidence and associated risk factors among YMSM aged 16 - 24 years in Tianjin, China.

Methods: We conducted a retrospective cohort study among men who have sex with men aged 16 - 24 years from October 2017 to December 2022 through the largest local nongovernmental organization serving men who have sex with men in Tianjin. Participants who responded to the investigations at least twice during the study period were included. They completed questionnaires on demographic information, sexual behaviors, psychoactive substance use, and sexually transmitted infection status. Simultaneously, their blood samples were collected for HIV testing. HIV incidence was calculated by dividing the sum of observed HIV seroconversions by the observed person-years (PYs). A Cox proportional hazards regression model was used to identify risk factors associated with HIV incidence.

Results: A total of 1367 HIV-negative YMSM were included in the cohort, among whom 62 seroconversions occurred, contributing 2384.2 observed PYs; the total incidence was 2.6 (95% CI 2.0 - 3.2) per 100 PYs. The segmented HIV incidence rates were 3 (95% CI 1.5 - 4.5), 2.4 (95% CI 1.5 - 3.3), and 2.7 (95% CI 1.5 - 3.9) per 100 PYs for 2017 - 2018, 2019 - 2020, and 2021 - 2022, respectively. Compared to YMSM who had been followed up fewer than three times, those followed up three or more times had a relatively lower risk of HIV infection (Adjusted hazard ratio [AHR] 0.553, 95% CI 0.325 - 0.941). YMSM who preferred finding sexual partners offline had a higher risk of HIV infection compared to those who preferred finding sexual partners online (AHR 2.207, 95% CI 1.198 - 4.066). Compared to YMSM without syphilis, those infected with syphilis had an increased risk of HIV infection (AHR 2.234, 95% CI 1.137 - 4.391). Additionally, YMSM who used psychoactive substances had a higher risk of HIV infection compared to those who did not use such substances (AHR 2.467, 95% CI 1.408 - 4.321).

Conclusions: Our study complements data on HIV incidence among YMSM in large cities across China. Syphilis infection and the use of psychoactive substances were risk factors associated with HIV occurrence, demonstrating an urgent need for tailored prevention and control interventions for this key population.

(JMIR Public Health Surveill 2025;11:e66487) doi:10.2196/66487

KEYWORDS

young men who have sex with men; YMSM; incidence; HIV; cohort study; risk factor

Introduction

The Global AIDS report released by the Joint United Nations Programme on HIV/AIDS (UNAIDS) [1] estimated that there were 350,000 new HIV infections among people aged 15 - 24 years in 2022. Young people accounted for 26% of HIV infections in the Asia-Pacific region in 2020; among them, 52% were men who have sex with men (MSM), making them a key

XSL•FO
RenderX

population that requires focused attention for HIV prevention. Meanwhile, the prevalence of HIV among young men who have sex with men (YMSM) has more than doubled in Indonesia (ie, from 6% in 2011 to 13% in 2019) and nearly tripled in Malaysia (ie, 6% in 2012 to 15% in 2022), and Vietnam (ie, 3% in 2011 to 11% in 2020) [2]. Previous studies suggest that young people from key populations experience high vulnerability and challenges in staying healthy and safe in situations marked by stigma, discrimination and harassment, punitive laws, and social taboos [2]. Moreover, young people living with HIV have the lowest rate of awareness about their infection status [3].

Meanwhile, in recent years, the HIV epidemic in China has risen rapidly among young people. The number of new HIV cases reported yearly among people aged 15‐24 years increased from 9373 in 2010 to 15,790 in 2019; a total of more than 140,000 HIV cases have been reported in this age group [4]. HIV has become a major infectious disease that endangers the health of young people in China, particularly in some regions where the proportions of reported cases have increased [4]. The spread of HIV among young people is concerning, posing a serious threat to their lives and health. Due to them being in a stage of rapid physical and mental development, lack of sexual health knowledge, and the increase in high-risk behaviors, they are more vulnerable to HIV.

Given that YMSM are susceptible to HIV infection [5,6], studies have shown that HIV infection rates among YMSM are higher than those in older populations [5,6,7,8]. YMSM under 25 years of age represent a demographic with increasing numbers of new HIV infections even as HIV incidence is declining globally [9]. Although there were some data on HIV prevalence among YMSM [9,10], the epidemic remains poorly defined among in this population [11], particularly the data of incidence is scarcely reported. Understanding HIV incidence—especially among subgroups of key populations—is essential for effectively targeting strategies to prevent new HIV infections [12].

Tianjin, located in the Bohai Gulf, is the second most populous city in northern China and one of the four major municipalities, with 16 districts and a resident population of 13.73 million as of 2021 [13,14]. The majority of young HIV cases in Tianjin were attributed to homosexual transmission, and the prevalence of HIV and syphilis among YMSM in the city is high [15]. However, the incidence of HIV among YMSM in Tianjin remains unknown.

This study aimed to (1) identify the HIV-related behavioral characteristics of YMSM, (2) understand the trends of HIV incidence among YMSM, and (3) investigate the association between related factors and HIV seroconversion among YMSM. The findings of this study will provide policymakers with evidence to inform prevention strategies aimed at controlling the HIV epidemic among YMSM.

Methods

Study Design

A retrospective cohort study on HIV incidence among YMSM was conducted in Tianjin from 2017 to 2022. First, to enhance geographical coverage and participant accessibility, seven voluntary HIV counselling and testing sites were set up for on-site surveys. Second, the on-site surveys were conducted at Shenlan, the largest local nongovernmental organization (NGO) serving MSM. The surveys were implemented by professional staff at the NGO, who also comprised MSM. To ensure the quality of the surveys, all professional staff at the 7 voluntary HIV counselling and testing sites received systematic and standard training; furthermore, the surveys were carried out through a unified method. Third, to guarantee the representativeness of the participants, two recruitment approaches were adopted. One approach involved professional staff recruiting YMSM through various strategies, such as offline hotspots (eg, gay baths, gay bars, public toilets) and online platforms (eg, WeChat, QQ, Blued). The second approach involved mobilizing the YMSM recruited through the first approach to introduce new YMSM. Fourth, all participants were required to participate in a baseline survey, followed by a follow-up survey every 6 months.

During the study, the fingerprints of the participants' right index fingers were collected, and a unique fingerprint code was assigned. A verification code was sent to the participants' cell phones to verify their phone numbers. After successful verification, training investigators conducted face-to-face questionnaire interviews. After the completion of the questionnaires, professional blood collectors obtained venous blood samples from the participants. The blood was sent to a professional laboratory for HIV testing. HIV test and questionnaire responses were linked using the participant's fingerprint code, which was unique and served as a proxy identifier to match the HIV testing results to each participant. Participants' fingerprint codes were recorded into an unidentifiable ID number to make the dataset anonymous. The Tianjin Centers for Disease Control and Prevention questionnaire was developed for the target population and presurvey.

Participant's Eligibility Criteria

The target population of this study were YMSM who participated in the survey and resided in Tianjin during the study period. Eligible criteria for study enrollment included: (1) born a male biologically, (2) aged 16‐24 years old, (3) who engaged in penetrative oral or anal sexual intercourse with males in the past year, (4) voluntarily consented to questionnaires, fingerprint registration, blood sample collection, and related testing, and (5) self-reported negative HIV tests or were unaware of their HIV infection status before participation in the survey.

Questionnaire Interview

The questionnaire included 4 parts. The first part collected information on sociodemographic characteristics such as age, residential district, occupation, education, marital status, and duration of residence. The second part collected information on sexual orientation and sexual behaviors in the last 6 months, including heterosexual behaviors, homosexual behaviors, commercial homosexual partners, primary venue to find sexual partners, and condom use. The third part collected information on the use of psychoactive substances (eg, Rush and Capsule Zero), injection drug use in the past year, syphilis infection status, and peer education received in the past year. As smoking,

XSL•FO
RenderX

alcohol consumption, and psychoactive substance use are addictive in nature, and previous research by Guo et al [16] showed a positive relationship between smoking and psychoactive substance use among YMSM, this study further explored these relationships. In this study, drinking was defined as consuming alcohol at least twice within the past 30 days, while smoking was defined as continuous or cumulative smoking for at least 6 months before the survey and smoking within the past 30 days.

HIV Laboratory Test Procedure

Venous blood samples were collected from the study participants. First, rapid HIV test reagents were used; those who tested reactive were retested using ELISA (enzyme-linked immunosorbent assay). If reactive, they were subjected to an HIV supplemental test. The supplemental test was first confirmed using the Western Blot method. If the Western Blot yielded indeterminate or negative results in the confirmatory test, the participants underwent a quantitative nucleic acid test using viral load detection reagents. Participants with nonreactive findings in the initial rapid test or the enzyme immunoassay were further tested using pooled nucleic acid tests.

Statistical Analysis

Based on the baseline data, frequency analyses stratified by HIV seroconversion status were performed. χ^2 tests were used to evaluate the differences in characteristics between participants with and without HIV seroconversion. For HIV seroconversions, the midpoint between the last negative and the first positive test dates was assumed to be the estimated onset date of HIV infection. For participants who did not seroconvert (ie, test positive for HIV), the date of their last survey was used as the endpoint for our analysis. Individual observation time was calculated as the interval between the participant's baseline survey and their latest HIV test during the study period. HIV incidence was calculated by dividing the number of observed HIV seroconversions by the total observed person-time. The 95% CI for HIV incidence were calculated using person-years (PYs) over the observed time as the denominator. The proportional hazards assumptions for the Cox models were tested by examining the interactions between variables included in the model and the logarithm of the follow-up time; no violations were observed. Therefore, Cox proportional hazards models were used to estimate the associations between various factors and HIV incidence. A multivariate analysis was conducted based on variables identified in univariate analysis; marginally significant variables ($P \leq .10$) in univariate analysis and factors that may be associated with new HIV infections were included in a multivariate Cox proportional hazards regression model to identify factors associated with HIV incidence. All analyses were performed using SPSS software (version 24.0; IBM Corp). A two-tailed $P<.05$ was considered statistically significant.

Ethical Considerations

This study was approved by the Ethics Committee of Tianjin Centers for Disease Control and Prevention (approval number: TJCDC-R-2023‐020) and was carried out in accordance with the Declaration of Helsinki. Ethical guidelines were followed during all the stages of this study. Comprehensive details about the study were provided to the participants during their first survey, and participation in the study was entirely voluntary and all participants provided written informed consent. Participants had the right to opt out of the study at any time. When accessing the dataset of this study, it was entirely deidentifed, and there was no effective way to link the data to the participants; further, the data were completely anonymized during data processing and statistical analyses. We did not pay any monetary compensation to the participants; however, we provided free HIV counseling and testing services.

Results

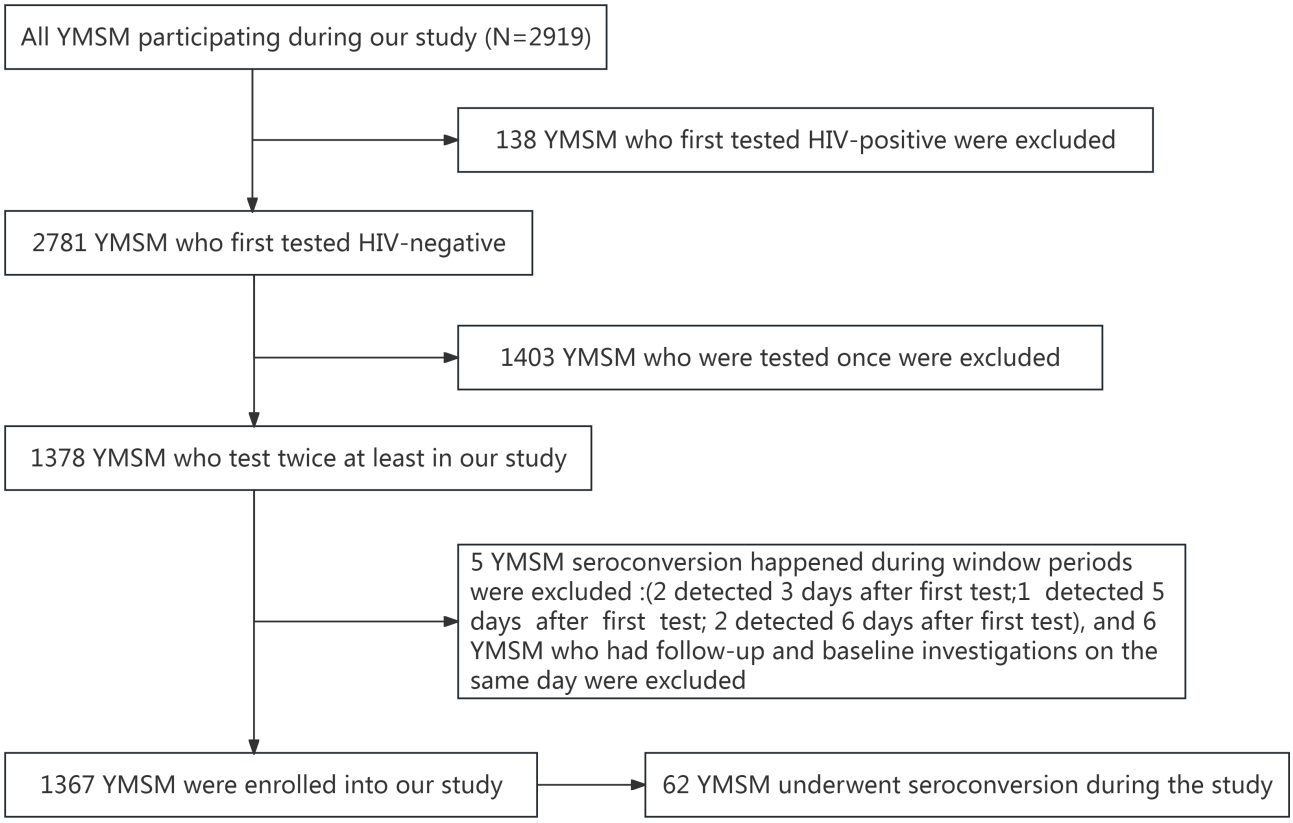
The Cohort

A total of 2919 YMSM completed the baseline survey in this study, of which 138 YMSM yielded positive results for HIV and were excluded from the initial investigation, with 2781 HIV-negative YMSM remaining after the first test. An additional 1403 YMSM who completed only the baseline survey were further excluded from the cohort. A total of 1378 YMSM then participated in two or more surveys during the study. Among these, 6 YMSM who underwent follow-up and baseline investigations on the same day were excluded. Additionally, 5 YMSM were excluded due to positive seroconversion within the window period (7d) after the baseline investigation. Lastly, a total of 1367 HIV-negative YMSM aged 16‐24 years, each with at least two HIV test records were identified and included in our cohort. Among the 1367 HIV-negative YMSM who were enrolled in our retrospective cohort study, 62 YMSM were newly infected with HIV during the study period (Figure 1). All participants in the cohort were followed up and tested simultaneously.

A total of 1403 YMSM were lost to follow-up in our study. Compared to those retained in the cohort YMSM, lost to follow-up YMSM had a higher proportion of out-of-school individuals, a higher frequency of anal intercourse in the last week, lower condom use during their last anal sex experience, and were less likely to use psychoactive substances (Table 1).

XSL•FO
RenderX

Figure 1. Retrospective cohort study screening process among YMSM in Tianjin, China,2017 - 2022. YMSM: young men who have sex with men.



XSL•FO
RenderX

Table . Comparisons of characteristics between the cohort YMSM[a] and the lost to follow-up YMSM in Tianjin, China, 2017 - 2022.

Variables	The cohort YMSM (N=1367), n (%)	The lost to follow-up YMSM (n=1403), n (%)	*P* value
Education level			.67
High school or below	266 (19.5)	282 (20.1)	
College and above	1101 (80.5)	1121 (79.9)	
Duration of residence in Tianjin (years)			.10
<2	304 (22.2)	349 (24.9)	
≧2	1063 (77.8)	1054 (75.1)	
Occupation			.03
Out-of-school YMSM	742 (54.3)	818 (58.3)	
Students YMSM	625 (45.7)	585 (41.7)	
Sexual orientation			.57
Homosexual	1283 (93.9)	1324 (94.4)	
Bisexual/heterosexual	84 (6.1)	79 (5.6)	
Primary venue to find sexual partners			.37
Online	1193 (87.3)	1208 (86.1)	
Offline	174 (12.7)	195 (13.9)	
Number of anal intercourse last week			.02
<2	1127 (82.4)	1109 (79)	
≧2	240 (17.6)	294 (21)	
Used condoms for anal sex last time			.002
No	451 (33)	541 (38.6)	
Yes	916 (67)	862 (61.4)	
Psychoactive substance use			.001
No	699 (51.1)	802 (57.2)	
Yes	668 (48.9)	601 (42.8)	

[a]YMSM: young men who have sex with men.

Demographic Characteristics

The majority (1101/1367, 80.5%) of participants completed college or university education. Most participants (n=1063, 77.8%) lived in Tianjin for over two years. A total of 45.7% (n=625) participants were students. The cohort participants' education level and duration of residence in Tianjin were statistical significant between different HIV seroconversion statuses (P<.05), whereas occupation and number of follow-up visits did not significantly differ by different statuses of HIV seroconversion (P>.05) (Table 2).

XSL•FO
RenderX

Table . Demographic characteristics of the cohort YMSM[a] in Tianjin,China,2017 - 2022.

Variables	Participants (N=1367), n (%)	HIV seroconversion		*P* value
		Yes, n (%)	No, n (%)	
Education level				.02
High school or below	266 (19.5)	19 (30.6)	247 (18.9)	
College and above	1101 (80.5)	43 (69.4)	1058 (81.1)	
Duration of residence in Tianjin (years)				.03
<2	304 (22.2)	7 (11.3)	297 (22.8)	
≧2	1063 (77.8)	55 (88.7)	1008 (77.2)	
Occupation				.26
Out-of-school YMSM	742 (54.3)	38 (61.3)	704 (53.9)	
Student YMSM	625 (45.7)	24 (38.7)	601 (46.1)	
Times of follow-up in the study				.47
<3	875 (64)	37 (59.7)	838 (64.2)	
≥3	492 (36)	25 (40.3)	467 (35.8)	

[a]YMSM: young men who have sex with men.

Sexual Behaviors, Syphilis Infection, and Psychoactive Substance Use

Among participants, 93.9% (1283/1367) identified themselves as homosexual, and 87.3% (n=1193) mainly found sexual partners online. In the past 6 months, 97.4% (n=1331) had anal intercourse, 17.6% (n=240) had more than two instances of anal intercourse in the last week, and 33.0%(n=451) did not use condoms during their last anal sex experience. Most (n=999, 73.1%) participants had their first homosexual intercourse between the ages of 16 - 19 years. Additionally, 7.3% (n=100) were male sex workers and 7.8% (n=107) participants were infected with syphilis. A total of 48.9% (n=668) used psychoactive substances, 25.2% (n=345) were smokers, and 41.6% (n=569) consumed alcohol (Table 3). The rate of psychoactive drug use was higher among smokers (209/345, 60.6%) than nonsmokers (459/1022, 44.9%*; P*<.001). Similarly, the rate of psychoactive drug use was more common among alcohol drinkers (310/569, 54.5%) than nondrinkers at (358/798, 44.9%; *P*<.001).

XSL•FO
RenderX

Table . Sexual behaviors, syphilis infection, and psychoactive substances use among the cohort YMSM[a] in Tianjin, China, 2017 - 2022.

Variables	Participants (N=1367), n (%)	HIV seroconversion		*P* value
		Yes, n (%)	No, n (%)	
Sexual orientation				.48
Homosexual	1283 (93.9)	60 (96.8)	1223 (93.7)	
Bisexual/heterosexual	84 (6.1)	2 (3.2)	82 (6.3)	
Primary venue to find sexual partners				.02
Online	1193 (87.3)	48 (77.4)	1145 (87.7)	
Offline	174 (12.7)	14 (22.6)	160 (12.3)	
Number of anal intercourse last week				.32
<2	1127 (82.4)	54 (87.1)	1073 (82.2)	
≧2	240 (17.6)	8 (12.9)	232 (17.8)	
Use a condom for anal sex last time				.50
No	451 (33.0)	18 (29)	433 (33.2)	
Yes	916 (67.0)	44 (71)	872 (66.8)	
Age of first homosexual intercourse (years)				.02
<16	70 (5.1)	8 (12.9)	62 (4.8)	
16 - 19	999 (73.1)	42 (67.7)	957 (73.3)	
>19	298 (21.8)	12 (19.4)	286 (21.9)	
Syphilis infection				.006
No	1260 (92.2)	51 (82.3)	1209 (92.6)	
Yes	107 (7.8)	11 (17.7)	96 (7.4)	
Male sex worker				.13
No	1267 (92.7)	61 (98.4)	1206 (92.4)	
Yes	100 (7.3)	1 (1.6)	99 (7.6)	
Psychoactive substance use				.001
No	699 (51.1)	19 (30.6)	680 (52.1)	
Yes	668 (48.9)	43 (69.4)	625 (47.9)	
Smoking				.69
No	1022 (74.8)	45 (72.6)	977 (74.9)	
Yes	345 (25.2)	17 (27.4)	328 (25.1)	
Alcohol consumption				.17
No	798 (58.4)	31 (50)	767 (58.8)	
Yes	569 (41.6)	31 (50)	538 (41.2)	
Know about PrEP[bc]				.03
No	513 (63.3)	26 (81.3)	487 (62.6)	
Yes	297 (36.7)	6 (18.7)	291 (37.4)	

[a]YMSM: young men who have sex with other men.

[b]There is missing data.

[c]PrEP: pre-exposure prophylaxis.

The primary venue for finding sexual partners, age of first homosexual intercourse, syphilis infection, and psychoactive substance use were statistically significant between different HIV seroconversion statuses (all *Ps*<.05). However, the number of instances of anal intercourse in the last week, condom use during last anal sex, male sex worker status, smoking status, and alcohol consumption status were not significantly associated with HIV seroconversion (all *Ps*>.05) (Table 3).

XSL•FO
RenderX

A total of 95.8% (n=599) student MSM considered themselves homosexual men, which was higher than the proportion among out-of-school YMSM; additionally, 93.0%(n=581) of student YMSM mainly found sexual partners online, which was higher than among out-of-school YMSM. Student YMSM also had lower rates of using condoms for anal sex, syphilis infection, psychoactive substance use, smoking, and alcohol consumption compared to out-of-school YMSM (Table 4).

Table . Comparisons of characteristics between the out-of-school YMSM[a] and the student YMSM among the YMSM cohort in Tianjin, China, 2017 - 2022.

Variables	Out-of-school YMSM, n (%)	Students YMSM, n (%)	*P* value
Sexual orientation			.005
Homosexual	684 (92.2)	599 (95.8)	
Bisexual/heterosexual	58 (7.8)	26 (4.2)	
Primary venue to find sexual partners			<.001
Online	612 (82.5)	581 (93.0)	
Offline	130 (17.5)	44 (7.0)	
Frequency of anal intercourse last week			<.001
<2	575 (77.5)	552 (88.3)	
≧2	167 (22.5)	73 (11.7)	
Use of a condom for anal sex during last sexual encounter			<.001
No	189 (25.5)	262 (41.9)	
Yes	553 (74.5)	363 (58.1)	
Age at first homosexual intercourse (years)			<.001
<16	42 (5.6)	28 (4.5)	
16 - 19	497 (67.0)	502 (80.3)	
>19	203 (27.4)	95 (15.2)	
Syphilis infection			<.001
No	657 (88.5)	603 (96.5)	
Yes	85 (11.5)	22 (3.5)	
Psychoactive substance use			<.001
No	321 (43.3)	378 (60.5)	
Yes	421 (56.7)	247 (39.5)	
Smoking			<.001
No	485 (65.4)	537 (85.9)	
Yes	257 (34.6)	88 (14.1)	
Alcohol consumption			<.001
No	383 (51.6)	415 (66.4)	
Yes	359 (48.4)	210 (33.6)	

[a]YMSM: young men who have sex with men.

HIV Incidence and the Associated Factors

Follow-up encounters ranged from 1 to 20, with the follow-up intervals ranging from 0.4 to 60.3 months with a median value of 16.5 (IQR 8.3 - 31.5) months among participants. The total person-time observed was 2384.2 PYs. The overall HIV incidence was 2.6 per 100 PYs (95% CI 2.0 - 3.2). To reduce the impact of random errors on the incidence rate and provide a more comprehensive perspective on observing incidence trends, we used a 2-year incidence rate. The segmented HIV incidence rates were 3.0 (95% CI 1.5 - 4.5), 2.4 (95% CI 1.5 - 3.3), and 2.7 (95% CI 1.5 - 3.9) per 100 PYs in 2017 - 2018, 2019 - 2020, and 2021 - 2022, respectively (Figure 2).

XSL•FO
RenderX

Figure 2. Trends in HIV incidence and 95% CI among YMSM in Tianjin, China, 2017 - 2022. PY: person-year. YMSM: young men who have sex with men.

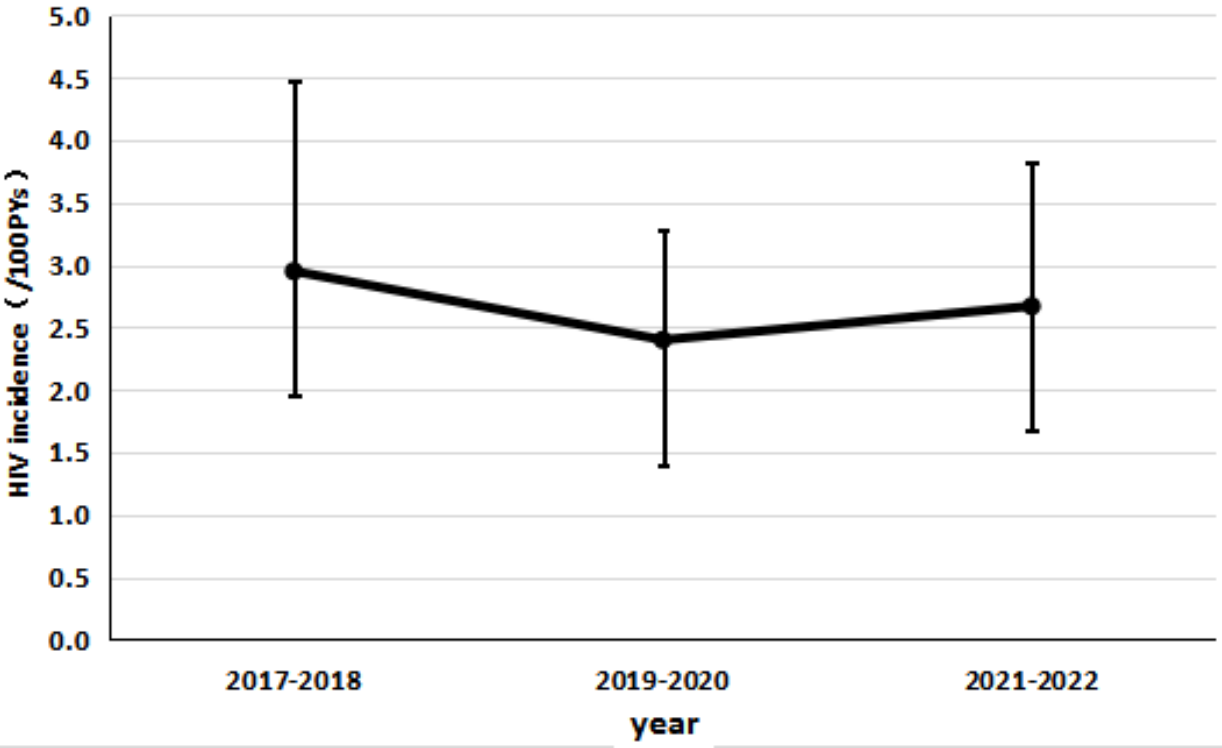


Table 5 shows that YMSM who were followed-up at least three times in the study had an incidence of 2.0 per 100 PYs. and were at a lower risk of HIV infection compared to those followed-up fewer than three times during the study (Adjusted hazard ratio [AHR] 0.553, 95% CI 0.325-0.941). The members who found sexual partners offline had an HIV incidence of 4.7 per 100 PYs. These individuals were at more than twice the risk of HIV infection compared to individuals who find sexual partners online (AHR 2.207, 95% CI 1.198-4.066). People infected with syphilis had an HIV incidence of 5.6 per 100 PYs and had a 2.234-fold higher risk of HIV infection compared to those without infected syphilis (AHR 2.234, 95% CI 1.137-4.391). The YMSM who used psychoactive substances had an HIV incidence of 3.7 per 100 PYs and a 2.467-fold greater risk of HIV infection compared to those who did not use psychoactive substances (AHR 2.467, 95% CI 1.408-4.321).

XSL•FO
RenderX

Table . Factors associated with HIV incidence identified using a proportional hazards regression model among YMSM[a] in Tianjin, China, 2017 - 2022.

Variables	Incidence (per 100 PYs)[b]	Univariate analysis		Multivariate analysis	
		HR[c] value (95% CI)	*P* value	HR value (95% CI)	*P* value
Education level			.08		.08
High school or below	3.7	1.000		1.000	
College and above	2.3	0.616 (0.359-1.057)		0.593 (0.328-1.073)	
Duration of residence in Tianjin (years)			.07		.13
<2	1.4	1.000		1.000	
≧2	2.9	2.061 (0.938-4.527)		1.887 (0.837-4.252)	
Occupation			.63		.53
Nonstudents	2.7	1.000		1.000	
Students	2.4	0.882 (0.528-1.472)		1.201 (0.678-2.126)	
Times of follow-up in the study			.11		.03
<3	3.2	1.000		1.000	
≧3	2.0	0.658 (0.392-1.104)		0.553 (0.325-0.941)	
Sexual orientation			.31		.65
Homosexual	2.7	1.000		1.000	
Bisexual/heterosexual	1.3	0.484 (0.118-1.978)		0.716 (0.170-3.010)	
Primary venue to find sexual partners			.02		.01
Online	2.3	1.000		1.000	
Offline	4.7	2.054 (1.132-3.725)		2.207 (1.198-4.066)	
Number of anal intercourse last week			.48		.55
<2	2.7	1.000		1.000	
≧2	2.1	0.764 (0.364-1.606)		0.791 (0.369-1.695)	
Use a condom for anal sex last time			.65		.88
No	2.9	1.000		1.000	
Yes	2.5	0.881 (0.508-1.530)		0.956 (0.542-1.687)	
Age of first homosexual intercourse (years)					
<16	6.2	1.000	—[d]	1.000	—[d]
16 - 19	2.4	0.383 (0.180-0.815)	.01	0.490 (0.226-1.059)	.07
>19	2.3	0.376 (0.153-0.919)	.03	0.547 (0.217-1.380)	.20
Syphilis Infection			.007		.02
No	2.3	1.000		1.000	
Yes	5.6	2.440 (1.272-4.683)		2.234 (1.137-4.391)	
Male sex worker			.12		.13
No	2.8	1.000		1.000	
Yes	0.8	0.207 (0.029-1.496)		0.209 (0.027-1.614)	
Psychoactive substance use			.002		.002
No	1.6	1.000		1.000	
Yes	3.7	2.369 (1.380-4.065)		2.467 (1.408-4.321)	

[a]YMSM: young men who have sex with men.

[b]PY: person-year.

[c]HR: hazard ratio.

XSL•FO
RenderX

[d]not applicable.

Discussion

Principal Findings

Our study is the first long-term cohort study to assess the incidence of HIV and its associated factors among YMSM in Tianjin. The HIV incidence identified in this study was similar to the incidence among MSM under 24 years old (2.63/100PYs) in Shanghai, published in 2019 [17]. This study found that the incidence of HIV among YMSM in Tianjin showed a stable trend from 2017 to 2022 and was maintained at an elevated level. This steady trend in HIV incidence suggests that risk factors for HIV infection continue to exist; additionally, as a group whose capacity for self-regulation has not fully matured [18], YMSM are more likely to engage in risky behaviors [19]. In our study, approximately 80% of the cohort members had their first homosexual encounter before the age of 20. Individuals who had a sexual debut at an earlier age are more likely to be cognitively immature and involved in high-risk sexual behaviors. Earlier sexual debut is associated with an increased HIV transmission risk [20]. Furthermore, a previous study identified YMSM as potentially more vulnerable to HIV infection than older MSM, reinforcing that YMSM in Tianjin are still at high risk of HIV infection [21].

Our study found a decreased risk of HIV incidence among YMSM who found sexual partners online. In contrast, for depth analysis, among those who found sexual partners offline, the primary venue (135/174, 77.6%) was gay bathhouses, where YMSM are more likely to engage in risky sexual behaviors [22]. Additionally, a study demonstrated that MSM who find partners online have more sexual encounters and multiple partners; however, this behavior was not associated with a greater risk of HIV infection [23]. This may be due to increased distrust and suspicion of partners met online.

As reported in a previous study [24], syphilis infection is a known risk factor associated with HIV incidence, as both HIV and syphilis share the same transmission route [24]. Therefore, we should conduct routine screening and treatment for syphilis infection among YMSM as a measure to prevent HIV infection [25].

Consistent with our findings, previous studies have shown that YMSM who participate in follow-up visits more frequently have a lower HIV incidence [26]. This could be attributed to these individuals being more concerned about HIV infection, which could reduce their sexual risk behaviors, such as decreasing the likelihood of condomless anal sex with a man whose HIV status was unknown. Although data on knowledge of HIV pre-exposure prophylaxis (PrEP) was partially missing in our study, YMSM who experienced HIV seroconversion had a lower awareness rate of HIV PrEP than those who did not. Therefore, we should take effective prevention measures such as PrEP, in addition to routine interventions, to decrease new HIV infections among YMSM.

The use of psychoactive substances was identified as a risk factor for HIV incidence in our study. Additionally, 48.9% of YMSM reported using psychoactive substances, which is higher than that among all MSM in the same region [27]. Psychoactive substances are mostly sold online, and 90% of YMSM in our study mainly found sexual partners online, a higher proportion than among the general MSM population [27]. Furthermore, with the development of network technology and improved internet access, more young people seek health services, particularly HIV-related services, through the internet [28]. Previous studies revealed that MSM who use gay social networking apps reported a higher number of sexual partners, suggesting that we should explore effective intervention strategies using the internet [29,30].

Another finding in our study was that YMSM who smoked or consumed alcohol had a higher prevalence of psychoactive substance use, which is a risk factor for HIV infection. This suggests that strengthening health education initiatives for reducing smoking and alcohol consumption among YMSM may help lower the incidence of HIV.

YMSM in Tianjin remain a vulnerable subpopulation for HIV infection and should be given priority for HIV intervention programs. Ongoing HIV control efforts among young students in schools in Tianjin, such as lectures on HIV, large-scale publicity and education, knowledge contests on HIV prevention, and debate competitions. Given the ongoing HIV epidemic among young students in Tianjin, we designed posters for student YMSM and displayed them in all universities, which yielded a good effect. However, performing effective work is challenging for the out-of-school YMSM due to the absence of a fixed setting, such as a school for intervention. In addition, our study found that, compared to the student YMSM, out-of-school YMSM were more inclined to engage in high-risk behaviors and had a higher proportion of bisexual or heterosexual orientation. Consequently, out-of-school YMSM are more likely to serve as a bridge population for HIV transmission from high-risk MSM to the general population. Therefore, additional prevention options and efforts should be made available to this group.

There are several limitations to this study. First, it was impossible to follow up on the HIV infection status of the individuals who did not enter the cohort; we aim to conduct the follow-up study in the future. Second, as we conducted our study through an NGO, the target population who did not seek HIV testing through the NGO were excluded, potentially leading to selection bias. Finally, given the sensitive nature of the topics, such as sexual behavior, reporting bias may exist.

Conclusions

HIV incidence among YMSM in Tianjin remains high and has not changed over time. Syphilis infection and the use of psychoactive substances continue to be risk factors associated with HIV incidence. Additionally, risky behaviors such as the use of psychoactive substances are widely present among YMSM; therefore, tailored interventions for this population should be strengthened. Our study supplements the data and evidence for research on new HIV infections among YMSM, both in and out-of-school in China's megacities and has significant practical implications.

XSL•FO
RenderX

Acknowledgments

This study was supported by the National Natural Science Foundation of China (72374153), Tianjin Health Research Project (grant No: TJWJ2023QN092), and Tianjin Medical Key Discipline Project (grant number: TJYXZDXK-50A). We would like to express our gratitude to all the staff and study participants.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy restrictions but are available from the corresponding author for reasonable request.

Authors' Contributions

Conceptualization: HG, MY, YG
Supervision: HG, MY, YG
Writing – original draft – HG, YG
Writing – review & editing – CM, JL, LL[1], LL[2],TN, XY, XZ, ZC, ZL.

Conflicts of Interest

None declared.

References

1. The path that ends AIDS: UNAIDS global AIDS update 2023. UNAIDS. 2023. URL: https://idpc.net/publications/2023/07/the-path-that-ends-aids-2023-unads-global-aids-update [accessed 2024-07-13]
2. UNAIDS. Putting young key populations first: HIV and young people from key populations in the asia and pacific region 2022. Geneva: Joint United Nations Programme on HIV/AIDS. 2022. URL: https://www.unaids.org/en/resources/documents/2022/2022-HIV-young-people-key-populations-asia-pacific [accessed 2023-07-05]
3. Baggaley R, Armstrong A, Dodd Z, Ngoksin E, Krug A. Young key populations and HIV: a special emphasis and consideration in the new WHO Consolidated Guidelines on HIV prevention, diagnosis, treatment and care for key populations. J Int AIDS Soc 2015 Feb;18(2S1). [doi: 10.7448/IAS.18.2.19438] [Medline: 25724509]
4. Zhao H, Liu H, Wang L, et al. Epidemiological characteristics of newly- reported HIV cases among youth aged 15-24 years-China, 2010 to 2019. China CDC Weekly 2020;2(48):913. [doi: 10.46234/ccdcw2020.249]
5. Garofalo R, Hotton AL, Kuhns LM, Gratzer B, Mustanski B. Incidence of HIV infection and sexually transmitted infections and related risk factors among very young men who have sex with men. J Acquir Immune Defic Syndr 2016 May 1;72(1):79-86. [doi: 10.1097/QAI.0000000000000933] [Medline: 26745827]
6. Dong Y, Wang L, Burgner DP, et al. Infectious diseases in children and adolescents in China: analysis of national surveillance data from 2008 to 2017. BMJ 2020 Apr 2;369:m1043. [doi: 10.1136/bmj.m1043] [Medline: 32241761]
7. UNAIDS. Young people and HIV. 2021. URL: https://www.unaids.org/en/resources/documents/2021/young-people-and-hiv [accessed 2021-05-07]
8. Mao X, Wang Z, Hu Q, et al. HIV incidence is rapidly increasing with age among young men who have sex with men in China: a multicentre cross-sectional survey. HIV Med 2018 Jun 19;19(8):513-522. [doi: 10.1111/hiv.12623] [Medline: 29923304]
9. Zhang H, Chen CZ, Lin SY, Xu HH, Xue DC, He ZG, et al. HIV/AIDS surveillance in men who have sex with men aged 15-24 based on internet in Fuzhou, 2016-2021. Chin J Epidemiol 2022;43(11):1761-1767. [doi: 10.3760/cma.j.cn112338-20220430-00365]
10. Coelho LE, Torres TS, Veloso VG, et al. The prevalence of HIV among men who have sex with men (MSM) and young MSM in Latin America and the Caribbean: a systematic review. AIDS Behav 2021 Oct;25(10):3223-3237. [doi: 10.1007/s10461-021-03180-5] [Medline: 33587242]
11. Yan G, Fangfang C, Hui G, et al. Trends in Incidence of HIV and Syphilis Among Men Who Have Sex with Men — Tianjin Municipality, China, 2013–2022. China CDC Wkly 2024;6(27):651-657. [doi: 10.46234/ccdcw2024.121]
12. Mwaniki SW, Mugo PM, Phillips TP. Project BESPOKE (Integrated bio-behavioral assessment of HIV and STI among young tertiary student men who have sex with men in Nairobi, Kenya): a respondent-driven sampling survey protocol. Front Public Health 2021 Oct 11;9:619694. [doi: 10.3389/fpubh.2021.619694] [Medline: 34708012]
13. State Council Seventh National Census Leading Group Office.: China Statistics Press; 2020. URL: https://www.stats.gov.cn/sj/ndsj/2020/indexch.htm [accessed 2024-07-13]
14. Tianjin bureau of statistics. Tianjin Statistical Yearbook. 2022. URL: https://stats.tj.gov.cn/nianjian/2022nj/zk/indexch.htm [accessed 2024-07-13]

XSL•FO
RenderX

15. Guo Y, Jie Y, Jianyun B. Study on the new incidence of HIV infection among young men who have sex with men using psychoactive substances in Tianjin. Chin J Viral Dis 2025;15(1):42-48. [doi: 10.16505/j.2095-0136.2024.0062]
16. Guo Y, Yang J, Bai JY, et al. Analysis on psychoactive substance use in young men who have sex with men in Tianjin[J]. Chin J Epidemiol 2024;45(5):687-693. [doi: 10.3760/cma.j.cn112338-20211104-00851]
17. Yin FL, Zhong PS, Zhang Y, Wang JX, Lu M, et al. HIV incidence,retention rate and related factors in a cohort study among men who have sex with men in Shanghai suburb. Chin J AIDS STD 2019 Sep 19;25(9):947-950. [doi: 10.13419/j.cnki.aids.2019.09.19]
18. Mi GD, Zhang BB, Yu F, et al. HIV incidence among men who have sex with men using geosocial networking smartphone application in Beijing, China: an open cohort study. Infect Dis Poverty 2021 Apr 2;10(1):27. [doi: 10.1186/s40249-021-00814-7] [Medline: 33795011]
19. Steinberg L. Risk taking in adolescence: what changes, and why? Ann N Y Acad Sci 2004 Jun;1021(1):51-58. [doi: 10.1196/annals.1308.005] [Medline: 15251873]
20. Balaji AB, An Q, Smith JC, et al. High human immunodeficiency virus incidence and prevalence and associated factors among adolescent sexual minority males-3 cities, 2015. Clin Infect Dis 2018 Mar 5;66(6):936-944. [doi: 10.1093/cid/cix902] [Medline: 29069298]
21. Wei L, Tian J, Guo M, et al. Trajectories of sexual risk behaviors and the associated factors among young men who have sex with men in China. Front Public Health 2022;10:854616. [doi: 10.3389/fpubh.2022.854616] [Medline: 35387185]
22. Wu Z, Xu J, Liu E, et al. HIV and syphilis prevalence among men who have sex with men: a cross-sectional survey of 61 cities in China. Clin Infect Dis 2013 Jul;57(2):298-309. [doi: 10.1093/cid/cit210] [Medline: 23580732]
23. Hong H, Xu J, McGoogan J, Dong H, Xu G, Wu Z. Relationship between the use of gay mobile phone applications and HIV infection among men who have sex with men in Ningbo, China: a cross-sectional study. Int J STD AIDS 2018 Apr;29(5):491-497. [doi: 10.1177/0956462417738468] [Medline: 29099328]
24. Macapagal K, Moskowitz DA, Li DH, et al. Hookup app use, sexual behavior, and sexual health among adolescent men who have sex with men in the United States. J Adolesc Health 2018 Jun;62(6):708-715. [doi: 10.1016/j.jadohealth.2018.01.001] [Medline: 29784114]
25. US Preventive Services Task Force, Mangione CM, Barry MJ, et al. Screening for syphilis infection in nonpregnant adolescents and adults: US Preventive Services Task Force Reaffirmation Recommendation Statement. JAMA 2022 Sep 27;328(12):1243-1249. [doi: 10.1001/jama.2022.15322] [Medline: 36166020]
26. Hui S, Chen F, Li Y, et al. Factors associated with newly HIV infection and transmitted drug resistance among men who have sex with men in Harbin, P.R. China. Front Public Health 2022;10:860171. [doi: 10.3389/fpubh.2022.860171] [Medline: 35719611]
27. Guo Y, Yang J, Bai JY, et al. Analysis on psychoactive substances use in men who have sex with men in Tianjin. Chin J Epidemiol 2022;43(5):686-691. [doi: 10.3760/cma.j.cn112338-20211104-00851] [Medline: 33041319]
28. You X, Gilmour S, Cao W, et al. HIV incidence and sexual behavioral correlates among 4578 men who have sex with men (MSM) in Chengdu, China: a retrospective cohort study. BMC Public Health 2021 Apr 26;21(1):802. [doi: 10.1186/s12889-021-10835-4] [Medline: 33902499]
29. Buhi ER, Daley EM, Oberne A, Smith SA, Schneider T, Fuhrmann HJ. Quality and accuracy of sexual health information web sites visited by young people. J Adolesc Health 2010 Aug;47(2):206-208. [doi: 10.1016/j.jadohealth.2010.01.002] [Medline: 20638014]
30. Zou H, Fan S. Characteristics of men who have sex with men who use smartphone geosocial networking applications and implications for HIV interventions: a systematic review and meta-analysis. Arch Sex Behav 2017 May;46(4):885-894. [doi: 10.1007/s10508-016-0709-3] [Medline: 27043837]

Abbreviations

AHR: adjusted hazard ratio
MSM: men who have sex with men
NGO: nongovernmental organization
PrEP: pre-exposure prophylaxis
PY: person-years
YMSM: young men who have sex with men

XSL•FO
RenderX

Edited by A Mavragani; submitted 20.09.24; peer-reviewed by J Li, K Jiao; revised version received 05.02.25; accepted 06.02.25; published 19.03.25.

Please cite as:
Guo Y, Gong H, Yan X, Zhang X, Ning T, Liu Z, Li L, Yang J, Li L, Ma C, Cui Z, Yu M
HIV Incidence and Associated Risk Factors Among Young Men Who Have Sex With Men in Tianjin, China: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e66487
URL: https://publichealth.jmir.org/2025/1/e66487
doi:10.2196/66487

© Yan Guo, Hui Gong, Xiangyu Yan, Xinying Zhang, Tielin Ning, Zhongquan Liu, Long Li, Jie Yang, Liu Li, Changxu Ma, Zhuang Cui, Maohe Yu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Disclosing to Patients the Use of Antiretroviral Resistance Testing Results for Molecular HIV Surveillance: A Randomized Experiment in 2 National Surveys

Jeremy Sugarman[1,2*], MD, MPH, MA; Juli Bollinger[1*], MS; Jose Agostini[3], BSW; Kevin Weinfurt[4], PhD; Gail Geller[1,2], MHS, ScD; Sheethal Jose[1,5], PhD; Marissa Hannah[3], MPH; O. Winslow Edwards[3], MPH; Leslie Meltzer Henry[1,6], JD, PhD; Travis Sanchez[3], DVM, MPH

[1]Berman Institute of Bioethics, Johns Hopkins University, 1809 Ashland Avenue, Baltimore, MD, United States

[2]School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[3]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[4]Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, United States

[5]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[6]Carey School of Law, University of Maryland, Baltimore, MD, United States

[*]these authors contributed equally

Corresponding Author:
Jeremy Sugarman, MD, MPH, MA
Berman Institute of Bioethics, Johns Hopkins University, 1809 Ashland Avenue, Baltimore, MD, United States

Abstract

Background: Molecular HIV surveillance (MHS) can be used to help identify and respond to emerging clusters of rapidly spreading HIV transmissions, a practice known as cluster detection and response (CDR). In the United States, MHS relies on HIV gene sequences obtained from routine clinical antiretroviral resistance testing (ARVRT). By law, ARVRT results are reported to public health agencies for MHS and individuals are not asked for their specific consent to do so. This practice has raised ethical concerns, including the lack of consent for, and transparency surrounding, public health uses of these clinical data. Such concerns have spurred debate and could have a chilling effect on the willingness of people living with HIV to agree to ARVRT when recommended clinically and jeopardize the utility of MHS-informed HIV prevention efforts. In response to the lack of routine disclosure of use of ARVRT results for MHS, in 2022, the Presidential Advisory Council on HIV/AIDS (PACHA) issued a resolution calling on the US Centers for Disease Control to "require that providers explain MHS/CDR and the laboratory test results that are collected and used in these surveillance activities to their patients."

Objective: This study aimed to examine the effect of clinician disclosure of the public health uses of ARVRT results for MHS versus clinician nondisclosure on patient willingness to undergo recommended ARVRT.

Methods: We conducted a randomized survey experiment examining the effect of clinician disclosure of the public health uses of ARVRT results for MHS versus clinician nondisclosure (the current standard of care) and subsequent discovery of such uses through a "trusted media source" on patient willingness to undergo recommended ARVRT. Study participants were respondents to 1 of 2 national web-based surveys conducted annually in the United States: the American Men's Internet Survey (AMIS) and the Transgender Women's Internet Survey and Testing (TWIST).

Results: Overall, 4348 AMIS participants (n=2151 disclosure; n=2197 nondisclosure) and 3314 TWIST participants (n=1670 disclosure; n=1644 nondisclosure) completed survey items regarding the randomly assigned vignettes. The majority were willing to undergo ARVRT regardless of which vignette they saw (1670/2151, 82.7% [AMIS] and 1326/1670, 80.8% [TWIST] in the disclosure group; and 1399/2197, 68% [AMIS] and 1101/1674, 68.45% [TWIST] in the nondisclosure group) after later learning about public health uses of ARVRT results.

Conclusions: The majority of respondents expressed willingness to undergo ARVRT even with disclosure of public health uses of these data, but willingness markedly decreased when learning about these uses after the fact, highlighting the importance of transparency in MHS programs. Accordingly, in line with the ethical principle of respect for autonomy and the likelihood that the potential public health benefits of MHS programs will not be compromised, consideration should be given to encouraging clinicians to disclose public health uses of ARVRT at the time ARVRT is recommended.

(JMIR Public Health Surveill 2025;11:e64663) doi:10.2196/64663

XSL•FO
RenderX

KEYWORDS

HIV; disclosure; consent; antiretroviral; HIV gene; ARVRT; molecular HIV surveillance; antiretroviral resistance; antiretroviral resistance testing; molecular HIV; HIV surveillance; public health; people living with HIV; web-based survey; United States; AMIS; TWIST; prevention and care; American Men's Internet Survey; Transgender Women's Internet Survey and Testing

Introduction

Molecular HIV surveillance (MHS) is used in public health to help identify and respond to clusters of rapidly spreading HIV transmissions, a practice known as cluster detection and response (CDR). In the United States, MHS relies on partial HIV gene sequences obtained from antiretroviral resistance testing (ARVRT) performed during routine clinical care of people living with HIV [1]. As opposed to time-space clusters that rely on the identification of an increase in new HIV diagnoses in a geographic area, MHS/CDR uses ARVRT data to identify emerging clusters of highly similar HIV sequences, which is an indicator of rapid HIV transmission [2]. Research has demonstrated that rapidly growing clusters contribute disproportionally to new HIV infections [3]. The use of MHS/CDR allows health departments to direct and prioritize resources toward groups experiencing rapid HIV transmission [2].

In the interests of public health and as is the case with other reportable conditions, ARVRT results are reported to public health agencies for MHS and individuals are not asked for their specific consent to do so. In 2018, the US Centers for Disease Control and Prevention (CDC) mandated that health departments include MHS in their CDR procedures; and in 2019, the use of MHS for CDR was identified as a pillar of the US Ending the HIV Epidemic initiative [4].

While difficult to quantify, the CDC has cited beneficial results of using MHS for CDR, including reduction in new HIV diagnoses, improved HIV testing and diagnoses, increased pre-exposure prophylaxis uptake, expanded syringe services program use, and improved linkage to care [5-7]. Nevertheless, current practices regarding such uses have raised ethical and social concerns regarding disclosure, consent, privacy, stigma, and the potential for HIV-related criminal prosecution [8-11]. Critics have argued that the identification of individuals in transmission clusters for intervention and prevention services, and the publicity surrounding those networks, could potentially further increase stigma and legal risk for marginalized individuals and groups including people of color, sex workers, and those already subject to enhanced policing and surveillance [12,13]. Several advocacy organizations have opposed using MHS for CDR [12,13], with some calling upon the CDC to halt its use altogether [14], citing concerns that the personal and medical information of people living with HIV are being used for surveillance purposes without meaningful engagement of people living with HIV in program planning, individuals' knowledge about how their ARVRT results are used in surveillance activities, and lack of informed consent. In response to these concerns, the Presidential Advisory Council on HIV/AIDS (PACHA) passed a nonbinding resolution in 2022 calling for substantial action by public health agencies and others, including the need to meaningfully engage with communities regarding MHS to minimize harms associated with its implementation, to gather evidence concerning its use, to obtain consent for it, to review criminal and public health laws that disparately treat people living with HIV, and to address punitive laws related to HIV criminalization [15]. To deal with the lack of transparency surrounding the use of clinical ARVRT results for MHS/CDR, the resolution called upon CDC to "require that providers explain MHS/CDR and the laboratory test results that are collected and used in these surveillance activities to their patients" [15].

While use of MHS continues to spur vigorous debate, to date there has not been a systematic assessment of how the risks and benefits of MHS are perceived among people living with HIV and people without HIV at increased risk of acquiring it. If left unidentified and unaddressed, these perceptions and risks could conceivably have a chilling effect on the willingness of people living without HIV at increased risk to pursue HIV testing and people living with HIV to agree to ARVRT when recommended clinically as well as decrease trust in both the health care system and public health.

To this end, our multidisciplinary team is conducting a mixed-methods, multiphase research project titled, "Study of Ethics and Stakeholder Attitudes towards Molecular Epidemiology (SESAME)" to collect data from a wide range of stakeholders to inform developing policies and practices.

In this report, we describe the results of a randomized survey experiment examining the effect of clinician disclosure of the public health uses of ARVRT results for MHS versus clinician nondisclosure and subsequent discovery of such uses through a "trusted media source" on patient willingness to undergo recommended ARVRT.

Methods

Study Population and Procedures

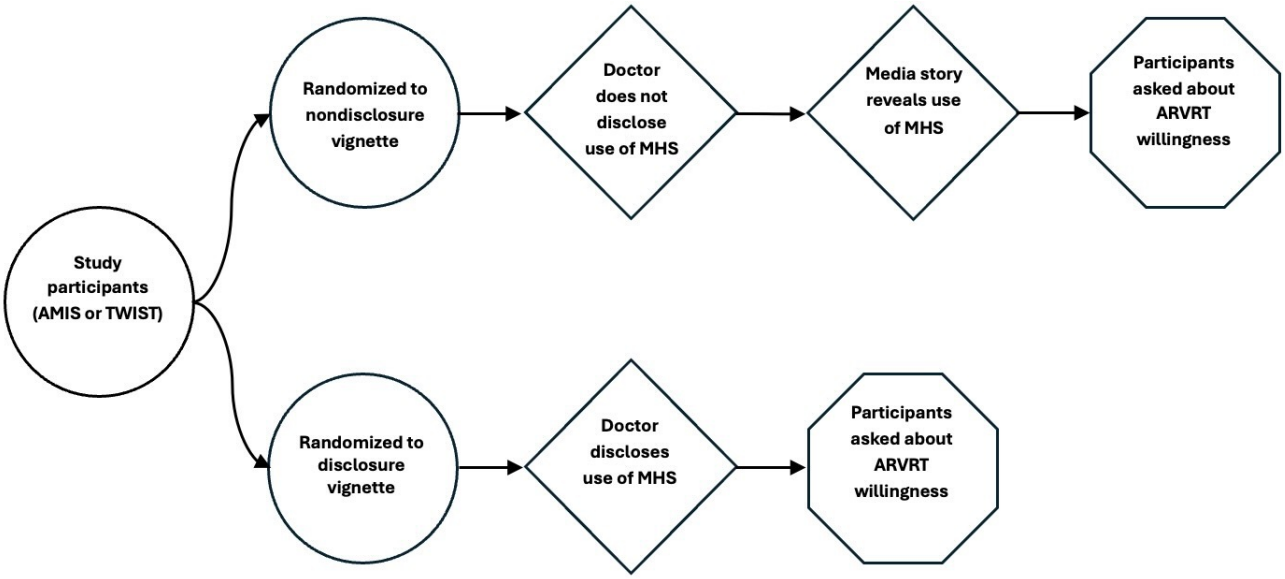
Participants in this study were respondents to 1 of 2 national web-based surveys fielded annually in the United States: The American Men's Internet Survey (AMIS) conducted from October 2022 to October 2023 and the Transgender Women's Internet Survey and Testing (TWIST) conducted from June 2022 to October 2023. AMIS is a cross-sectional study that seeks participants from eligible cisgender gay, bisexual, same gender loving, and other men who have sex with men (GBMSM) to provide information useful for monitoring trends in sexual health [16]. Briefly, participants were recruited through online advertisements on social media sites and apps among the target population of GBMSM. All participants who clicked on advertisements links were taken to the survey website for screening. Respondents were eligible for AMIS if they were at least 15 years of age, reported male sex at birth and male gender identity, resided in the United States, and reported ever having oral or anal sex with a male partner. TWIST involves transgender women and transfeminine nonbinary people

XSL•FO
RenderX

assigned male at birth [17]. Recruitment for TWIST was similar to AMIS and was conducted with advertisements on social media platforms and apps popular among the target population. Respondents were eligible for TWIST if they were at least 15 years of age, reported male sex at birth and feminine (woman or transgender woman) gender, resided in the United States, and reported ever having oral, anal, or vaginal sex.

In each study, participants were randomly assigned to 1 of 2 hypothetical vignettes that both set the scenario as them being someone living with HIV seeing a clinician who asked them about their willingness to undergo standard ARVRT. This scenario was the same regardless of the participant's self-reported HIV status in the survey (Figure 1.).

Figure 1. Visual depiction of the survey experiment. Respondents were randomly assigned to view 1 of 2 hypothetical vignettes that both set the scenario as them being someone living with HIV seeing a clinician who asked them about their willingness to undergo standard antiretroviral resistance testing (ARVRT). Survey arms differed by whether the treating clinician disclosed public health use of ARVRT results for molecular HIV surveillance or the respondent learns about this later through a media story. AMIS: American Men's Internet Survey; MHS: molecular HIV surveillance; TWIST: Transgender Women's Internet Survey and Testing.



In the "disclosure" vignette, the clinician initially discloses the public health reporting and MHS associated with ARVRT, and then participants are asked about their willingness to have the test performed. In the "nondisclosure" vignette, which generally reflects the current standard of care, the public health reporting and MHS practices are not initially shared with the participant, and they are asked about their willingness to undergo ARVRT. Then, they are told about a "media story" that indicates ARVRT results are reported to public health agencies for MHS. Willingness to have ARVRT was then asked again. (The text of the 2 vignettes and questions regarding them are available in the Multimedia Appendices 1 and 2)

Before fielding, the vignette descriptions and survey questions were iteratively tested and revised through a series of 12 cognitive interviews with MSM to test for clarity and comprehension. Potential interviewees were recruited using a convenience sample of previous AMIS respondents who provided their email addresses so they could be contacted regarding participation in future studies. Cognitive interviews were conducted virtually over Zoom (Zoom Communications) by trained interviewers using a cognitive interview guide (Multimedia Appendix 3) developed by the research team.

Measures

The dependent measure in this study is willingness to undergo ARVRT. For the disclosure vignette, this was asked only once. For the nondisclosure vignette, the measure of interest for this study was the willingness to undergo ARVRT both before and after the participant learned of the public health uses through the media story. The response scale was "very unwilling, unwilling, neither willing nor unwilling, willing, and very willing." For this analysis, responses were dichotomized to combine "very willing" and "willing" responses into "willing" and all other responses as "unwilling."

The primary independent measure was the randomization to either the disclosure or nondisclosure vignette. Other independent measures were participants' race and ethnicity, self-reported HIV status, and anticipated health care stigma. Race was captured with a question asking participants to select all races that apply from a list of possible options and a separate yes/no question about Hispanic or Latinx ethnicity. For this analysis, responses were categorized into Hispanic or Latinx, non-Hispanic Black, non-Hispanic White, and other or multiple race [16]. HIV status was self-reported based on the result of respondents' most recent HIV test. Anticipated health care stigma is a variable created with responses to a pair of questions about whether they ever "avoided" or "felt afraid" getting health care services because of how they would be treated [18]. AMIS participants were asked about stigma related to having sex with men. TWIST participants were asked about anticipated stigma related to who they have sex with or their gender identity. Participants who answered yes to either of these questions were recorded as having anticipated health care stigma, and those who reported no fear or avoidance of health care services were labeled as not having anticipated stigma.

XSL•FO
RenderX

Analyses

Analyses were conducted with unduplicated successful survey responses from participants who answered the ARVRT willingness questions. The procedures for deduplication and successful completion metrics have been previously described [16,17]. Comparisons in dichotomous willingness to have ARVRT by vignette were calculated as prevalence ratios with 95% CI. Multivariable modeling was not deemed necessary for this randomized study with no failure of randomization detected in participant characteristics between arms. Prevalence ratios greater than the null (1) would indicate that those in the disclosure vignette were more willing to have ARVRT than those in the nondisclosure vignette. To explore heterogeneity in the effects, these comparisons were also stratified by participant race and ethnicity, self-reported HIV status, and anticipated health care stigma. Analyses were conducted in SAS version 9.4 (SAS Institute) and significance level was α=.05.

Ethical Considerations

Cognitive Interviews

Cognitive interview participants were sent the oral consent script by email in advance of the virtual interview and provided oral consent before beginning the interview. Each participant was provided US $20 for completing the virtual interview. All data related to the cognitive interviews were stored in a secure analytic environment at Johns Hopkins University in a project folder accessible only to study team members who conducted the interviews using password-protected computers. The cognitive interview protocol was reviewed and approved as an expedited study by the Johns Hopkins Medicine Institutional Review Board (IRB00331966).

American Men's Internet Survey and Transgender Women's Internet Survey

AMIS and TWIST participants provided online informed consent before beginning the online self-administered survey. AMIS participants were not compensated; TWIST participants were compensated US $10 for survey completion. The AMIS and TWIST surveys are administered on a secure online survey platform (Alchemer) and include questions on demographic characteristics, sexual and substance use behaviors, HIV testing history, stigma and discrimination, and mental health. Personal identifiers other than email address and IP address are not collected. Original study data are stored on secure network drives with access restricted to authorized study staff. Datasets for analyses are deidentified, and no identifiable information is presented in this manuscript or supplemental files. The AMIS and TWIST studies were reviewed and approved by Emory University's human subjects research review panel (protocol 47676 and 108784, respectively).

All project data are protected by a US federal Certificate of Confidentiality that prevents court-ordered disclosure of research data.

Results

American Men's Internet Survey

Overall, 4348 AMIS participants completed the vignette items (n=2151 disclosure vignette and n=2197 nondisclosure vignette) (Table 1). Most were aged 30 years or older and non-Hispanic White. People living with HIV comprised 12.2% (529/4348) of AMIS participants. Among those who answered the anticipated health care stigma question (n=2260), 33.9% (767/2260) reported ever experiencing this stigma. The majority were willing to undergo ARVRT regardless of which vignette they saw (1670/2151, 82.7% in the disclosure group; and 2186/2197, 92.4% in the nondisclosure group before the media story and 1399/2197, 68% after the media story), but those in the disclosure group were significantly more likely to be willing to have ARVRT compared with the nondisclosure group after the media story disclosed MHS uses of ARVRT information (prevalence ratio=1.22, 95% CI 1.17-1.26). In stratifications by age, race and ethnicity, HIV status, and anticipated health care stigma, the effects were all in the same direction and of similar magnitude to the overall finding.

XSL•FO
RenderX

Table . Willingness to have HIV antiretroviral resistance testing under randomized scenarios of clinician disclosure of HIV molecular epidemiology by key participant characteristics, United States Cisgender Men Who Have Sex with Men and Transfeminine Persons, 2022 - 2023.

Characteristic	Overall, n	Randomization group		Prevalence ratio (95% CI)
		Disclosure, n/N (%)	Nondisclosure[a] (ref), n/N (%)	
All men who have sex with men (AMIS[b] participants)	4348	1670/2151 (82.7)	1399/2197 (68)	1.22 (1.17-1.26)
Age (years)				
15 - 24	335	121/169 (77.6)	106/166 (69.3)	1.12 (0.98-1.28)
25 - 29	407	173/215 (85.2)	135/192 (74.6)	1.14 (1.03-1.27)
30 - 39	1014	397/496 (85.4)	346/518 (70.8)	1.21 (1.13-1.29)
40+	2592	979/1271 (81.9)	812/1321 (65.8)	1.24 (1.19-1.31)
Race and ethnicity				
Non-Hispanic Black	473	161/221 (79.7)	163/252 (69.4)	1.15 (1.03-1.28)
Hispanic	441	162/222 (78.7)	143/219 (69.8)	1.13 (1.01-1.27)
Non-Hispanic White	3015	1195/1501 (84)	962/1514 (67.7)	1.24 (1.19-1.3)
Other	383	146/194 (80.2)	123/189 (70.3)	1.14 (1.01-1.29)
HIV status				
Positive	529	202/258 (82.1)	172/271 (67.5)	1.22 (1.1-1.35)
Negative or unknown	3819	1468/1893 (82.8)	1227/1926 (68.1)	1.22 (1.17-1.26)
Anticipated health care stigma				
Yes	767	277/374 (79.4)	234/393 (63.9)	1.24 (1.13-1.36)
No	1493	572/732 (83.9)	487/761 (68.7)	1.22 (1.15-1.3)
All transfeminine persons (TWIST[c] participants)	3314	1326/1670 (80.8)	1101/1644 (68.4)	1.18 (1.13-1.23)
Age (years)				
15 - 24	1126	475/568 (85.1)	375/558 (69.1)	1.22 (1.16-1.32)
25 - 29	719	283/370 (77.5)	248/349 (71.9)	1.08 (0.99-1.18)
30 - 39	900	343/442 (79)	314/458 (69.8)	1.13 (1.05-1.22)
40+	569	225/290 (79.2)	164/279 (60.3)	1.31 (1.17-1.47)
Race and ethnicity				
Non-Hispanic Black	280	97/140 (71.3)	71/140 (53.8)	1.33 (1.1-1.61)
Hispanic	299	116/150 (78.4)	101/149 (71.1)	1.1 (0.96-1.26)
Non-Hispanic White	2566	1052/1297 (82.5)	877/1269 (70.1)	1.18 (1.13-1.23)
Other	135	52/68 (77.6)	44/67 (63.1)	1.23 (0.98-1.54)
HIV status				
Positive	49	22/26 (88)	10/23 (45.5)	1.94 (1.2-3.13)
Negative or Unknown	3265	1304/1644 (80.7)	1091/1621 (68.7)	1.18 (1.13-1.22)

XSL•FO
RenderX

Characteristic	Overall, n	Randomization group		Prevalence ratio (95% CI)
		Disclosure, n/N (%)	Nondisclosure[a] (ref), n/N (%)	
Anticipated health care stigma				
Yes	1111	435/552 (80.3)	363/559 (66)	1.22 (1.13-1.31)
No	531	207/257 (81.8)	201/274 (74.7)	1.1 (1-1.2)

[a]After media story.

[b]AMIS: American Men's Internet Survey.

[c]TWIST: Transgender Women's Internet Survey and Testing.

Transgender Women's Internet Survey

Overall, 3314 TWIST participants completed the vignette items (n=1670 disclosure vignette and n=1644 nondisclosure vignette). Most were aged 30 years or younger and non-Hispanic White (Table 1). People living with HIV comprised 1.5% (49/3314) of TWIST participants. Among those who answered the anticipated health care stigma question (n=1642), 68.4% (n=1111) reported ever experiencing this stigma. The majority were willing to undergo ARVRT regardless of which vignette they saw (1326/1670, 80.8% in disclosure and 1446/1674, 89.5% in the nondisclosure group before the media story and 1101/1674, 68.4% after the media story) but those in the disclosure group were significantly more likely to be willing to have ARVRT compared with the nondisclosure group after the media story disclosed MHS uses of ARVRT information (prevalence ratio=1.18; 95% CI=1.13-1.23). In stratifications by age and anticipated health care stigma, the changes in willingness were all in the same direction and of similar magnitude to the overall finding. The association between disclosure vignette and willingness to have ARVRT was even more pronounced for TWIST participants living with HIV (prevalence ratio=1.94; 95% CI=1.2-3.13) than those who reported being HIV negative or of unknown HIV status.

Discussion

In this study with GBMSM and transgender women or transfeminine persons from across the United States, majorities of both study populations were willing to undergo ARVRT recommended by an HIV care clinician regardless of the requirement for public health reporting of those test results and the manner in which they learned about the reporting requirement. Despite these majorities, however, our results show that disclosing the public health reporting requirement to patients at the time ARVRT is ordered resulted in significantly higher willingness to pursue ARVRT than subsequently finding out about this reporting requirement from a "media story." The disclosure-related disparities in willingness to pursue ARVRT were most evident among transfeminine persons who were Black or who were living with HIV infection.

Importantly, our data suggest that informing patients upfront that ARVRT results are shared with public health authorities has only a modest effect on ARVRT uptake. Higher levels of ARVRT uptake were reported among nondisclosure vignette respondents when asked before learning about the public health reporting requirement through the "media story" and, thus, when they were "unaware" of the reporting requirement. However, when those in the nondisclosure group later learned of the public health reporting requirement from a media story, their subsequent willingness to pursue testing dropped significantly, far below the level reported by those who were informed by the clinician at the time of testing (disclosure group). This steep drop in willingness to pursue recommended ARVRT among those in the nondisclosure group is not surprising and is likely attributed to respondents' learning that their ARVRT results were shared with public health authorities without their knowledge.

Critics of MHS have advocated for incorporating informed consent into sharing ARVRT data with public health authorities for MHS, citing a lack of transparency about the process and concerns about individual privacy and autonomy. In their 2022 resolution PACHA urged "CDC to work with states and jurisdictions, alongside community members, providers, tribes, tribal epidemiology centers, members of networks of people living with HIV, and partner agencies within HHS to create a stronger system of informed consent around the uses of molecular HIV data" and called upon CDC to require states to "provide plain language notifications to individuals living with HIV on the types of surveillance and analyses being conducted and opt-out options for the use of MHS patient data for CDR activities" [15]. Nevertheless, while CDC's recently updated HIV CDR Guidance for Health Departments includes calls for public health practitioners to be more transparent about the use of MHS, it does not endorse waivers or opting out of it [19].

In the absence of informed consent or an opt-out option for sharing ARVRT results with public health agencies for MHS, our data suggest the potential value of simply disclosing the public health use of ARVRT results to people living with HIV, preferably at the time testing is recommended. While disclosure is not equivalent to obtaining informed consent, upfront disclosure about public health reporting may resolve some of the ethical and community concerns regarding the lack for transparency about how personal data are used in MHS/CDR without jeopardizing the potential public health utility of MHS-informed HIV prevention efforts.

These findings also align with results from an earlier survey we conducted as part of our SESAME project where we examined 2 methods for explaining MHS to MSM living with HIV and those at-risk without HIV in the 2021 AMIS survey. Specifically, participants were randomly assigned to view a

XSL•FO
RenderX

brief video or text explaining MHS. In the video group, respondents with high video engagement were less likely to be concerned about MHS; and in the text group, discomfort with MHS decreased as awareness of different public health activities increased [20].

The results of our study should be interpreted with some limitations in mind. First, while responses to hypothetical vignettes can be informative, it is unknown how well the findings translate to real-world situations. For example, we do not know why some participants in both arms of the study declined ARVRT testing at the time of their visit with the clinician. Clearly there were other factors at play besides the disclosure of public health use of these test results. Ideally, future demonstration projects in actual situations could assess the impact of disclosure on willingness to undergo ARVRT. In addition, unmeasured contextual factors may play a role in real-world decision making, but it is unlikely such factors would bias the results given randomization. Although the vignettes were simple, there may have been differential understanding of them based upon the vignette type. Of note, to overcome this possibility, we conducted extensive cognitive testing of the scenarios (as well as the response options) before fielding them. Furthermore, AMIS and TWIST are online convenience sampling surveys and results are not generalizable to all people living with HIV and those without HIV at increased risk of acquiring it. The convenience sampling approach, even though conducted with multiple types of websites and social networking applications intended to increase sample diversity, increases the potential for selection biases and in ways that cannot be directly measured. Racial and ethnic minority persons in both AMIS and TWIST and younger persons in AMIS were also substantially underrepresented, potentially reducing our ability to discern key differences among these groups.

Conclusions

The majority of respondents expressed willingness to undergo ARVRT even with disclosure of public health uses of these data, but willingness markedly decreased when people learn about these uses after the fact. Accordingly, in line with the ethical principle of respect to autonomy, and the likelihood that the potential public health benefits of MHS programs will not be compromised, consideration should be given to encouraging clinicians to disclose public health uses of ARVRT at the time testing is suggested.

Acknowledgments

The authors thank the American Men’s Internet Survey (AMIS) and Transgender Women’s Internet Survey and Testing (TWIST) respondents who participated in this study. Work on this project was supported by the National Institute of Mental Health, “Risks, Benefits, and Stakeholder Perspectives of Molecular Epidemiology for HIV Prevention (HIV-ME),” (5R01MH124590).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to the sensitive nature of the information but are available from the corresponding author on reasonable request.

Authors' Contributions

JS contributed to conceptualization, methodology, writing – reviewing & editing, supervision, and funding acquisition. JB managed conceptualization, methodology, writing – original draft, and project administration. JA handled formal analysis, writing – original draft, and visualization. KW managed methodology and writing – review & editing. GG handled conceptualization, methodology, writing – review & editing, and funding acquisition. SJ contributed to conceptualization, methodology, and writing – review & editing. MH managed software, data curation, and writing – review & editing. OWE supervised resources, investigation, writing – review & editing, and project administration. LMH managed conceptualization and writing – reviewing & editing. TS handled conceptualization, methodology, formal analysis, writing – review & editing, and supervision.

Conflicts of Interest

JS is a consultant to Merck KGaA, IQVIA, and Merck; and a member of Aspen Neurosciences' Clinical Advisory Panel. TS is Editor-in-Chief of *JMIR Public Health and Surveillance*. He had no editorial role in the review of this paper.

Multimedia Appendix 1
American Men’s Internet Survey-Transgender Women’s Internet Survey and Testing vignette-disclosure.
[DOCX File, 32 KB - publichealth_v11i1e64663_app1.docx]

Multimedia Appendix 2
American Men’s Internet Survey-Transgender Women’s Internet Survey and Testing vignette-no disclosure.
[DOCX File, 34 KB - publichealth_v11i1e64663_app2.docx]

Multimedia Appendix 3
Cognitive interview guide for vignette survey.

XSL•FO
RenderX

[DOCX File, 30 KB - publichealth_v11i1e64663_app3.docx]

References

1. Oster AM, France AM, Panneer N, et al. Identifying clusters of recent and rapid HIV transmission through analysis of molecular surveillance data. J Acquir Immune Defic Syndr 2018 Dec 15;79(5):543-550. [doi: 10.1097/QAI.0000000000001856] [Medline: 30222659]
2. Oster AM, Lyss SB, McClung RP, et al. HIV cluster and outbreak detection and response: the science and experience. Am J Prev Med 2021 Nov;61(5 Suppl 1):S130-S142. [doi: 10.1016/j.amepre.2021.05.029] [Medline: 34686282]
3. Dennis AM, Hué S, Billock R, et al. Human Immunodeficiency Virus type 1 phylodynamics to detect and characterize active transmission clusters in North Carolina. J Infect Dis 2020 Mar 28;221(8):1321-1330. [doi: 10.1093/infdis/jiz176] [Medline: 31028702]
4. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. JAMA 2019 Mar 5;321(9):844-845. [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
5. Science brief: evidence that HIV cluster and outbreak detection and response strengthens HIV prevention and care services. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/hiv/programresources/guidance/cluster-outbreak/science-brief.html#print [accessed 2024-4-21]
6. Saldana C, Philpott DC, Mauck DE, et al. Public health response to clusters of rapid HIV transmission among Hispanic or Latino gay, bisexual, and other men who have sex with men — Metropolitan Atlanta, Georgia, 2021–2022. MMWR Morb Mortal Wkly Rep 2023 Mar 10;72(10):261-264. [doi: 10.15585/mmwr.mm7210a3] [Medline: 36893048]
7. McClung RP, Atkins AD, Kilkenny M, et al. Response to a large HIV outbreak, Cabell County, West Virginia, 2018-2019. Am J Prev Med 2021 Nov;61(5 Suppl 1):S143-S150. [doi: 10.1016/j.amepre.2021.05.039] [Medline: 34686283]
8. Crowley JS, Bland SE. Big idea: policy action can increase community support for HIV cluster detection. O'Neill Institute for National and Global Health Law. 2019. URL: https://oneill.law.georgetown.edu/wp-content/uploads/2021/06/Policy-Can-Increase-Community-Support-For-HIV-Cluster-Detection.pdf [accessed 2024-04-21]
9. Molldrem S, Smith AKJ, Subrahmanyam V. Toward consent in molecular HIV surveillance?: Perspectives of critical stakeholders. AJOB Empir Bioeth 2024;15(1):66-79. [doi: 10.1080/23294515.2023.2262967] [Medline: 37768111]
10. Shook AG, Buskin SE, Golden M, et al. Community and provider perspectives on molecular HIV surveillance and cluster detection and response for HIV prevention: qualitative findings from King County, Washington. J Assoc Nurses AIDS Care 2022;33(3):270-282. [doi: 10.1097/JNC.0000000000000308] [Medline: 35500058]
11. PROMPT Bioethics Study Team, Dennis A, Day S, Rennie S, Sullivan K. Prompt study: revitalizing community engagement in the public health use of molecular HIV epidemiology. Dennis Research Group. 2023. URL: https://dennis.unc.edu/wp-content/uploads/sites/3288/2023/04/MHE_WhitePaper_2023-04-28.pdf [accessed 2025-03-26]
12. Molecular health surveillance. The Center for HIV Law and Policy (CHLP). 2024. URL: https://www.hivlawandpolicy.org/projects/molecular-hiv-surveillance#:~:text=People%20living%20with%20HIV%2C%20advocates,from%20HIV%20stigma%20and%20criminalization [accessed 2025-04-11]
13. Trejo A, McClelland A. Molecular HIV surveillance: a global review of human rights implications. HIV Justice Worldwide. URL: https://www.hivjusticeworldwide.org/wp-content/uploads/2021/10/HJWW-MHS-Paper-English-Final.pdf [accessed 2025-03-26]
14. Open letter to the CDC: networks of people living with HIV demand a moratorium on molecular HIV surveillance. US PLHIV Caucus. 2024. URL: https://actionnetwork.org/forms/open-letter-to-the-cdc-we-demand-a-moratorium-on-molecular-hiv-surveillance/ [accessed 2025-03-26]
15. Molecular HIV surveillance and cluster detection and response resolution. Presidential Advisory Council on HIV/AIDS (PACHA). 2022. URL: https://files.hiv.gov/s3fs-public/2022-11/PACHA-resolution-on-Molecular-HIV-Surveillance-Cluster-Detection-and-Response-10-17-22.pdf [accessed 2025-04-01]
16. Sanchez TH, Sineath RC, Kahle EM, Tregear SJ, Sullivan PS. The annual American Men's internet survey of behaviors of men who have sex with men in the United States: protocol and key indicators report 2013. JMIR Public Health Surveill 2015;1(1):e3. [doi: 10.2196/publichealth.4314] [Medline: 27227126]
17. Zlotorzynska M, Sanchez TH, Scheim AI, et al. Transgender women's internet survey and testing: protocol and key indicators report. Transgend Health 2021 Oct;6(5):256-266. [doi: 10.1089/trgh.2020.0071] [Medline: 34993298]
18. Augustinavicius JL, Baral SD, Murray SM, et al. Characterizing cross-culturally relevant metrics of stigma among men who have sex with men across 8 sub-Saharan African countries and the United States. Am J Epidemiol 2020 Jul 1;189(7):690-697. [doi: 10.1093/aje/kwz270] [Medline: 31942619]
19. HIV CDR guidance for health departments. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/hiv/programresources/guidance/hiv-cluster-detection-and-response-guidance/index.html [accessed 2024-04-24]
20. Valenzuela Lara M, Sanchez TH, Bollinger JM, Geller G, Little SJ, Sugarman J. Perceptions of molecular HIV surveillance among men who have sex with men in the United States. AIDS Behav 2025 Mar;29(3):760-768. [doi: 10.1007/s10461-024-04557-y] [Medline: 39636553]

XSL•FO
RenderX

Abbreviations

AMIS: American Men's Internet Survey
ARVRT: antiretroviral resistance testing
CDC: US Centers for Disease Control and Prevention
CDR: cluster detection and response
GBMSM: gay, bisexual, same gender loving, and other men who have sex with men
MHS: molecular HIV surveillance
PACHA: Presidential Advisory Council on HIV/AIDS
TWIST: Transgender Women's Internet Survey and Testing

Edited by A Mavragani; submitted 22.08.24; peer-reviewed by L Dawson, M Zlotorzynska; revised version received 05.03.25; accepted 08.03.25; published 11.04.25.

Please cite as:
Sugarman J, Bollinger J, Agostini J, Weinfurt K, Geller G, Jose S, Hannah M, Edwards OW, Henry LM, Sanchez T
Impact of Disclosing to Patients the Use of Antiretroviral Resistance Testing Results for Molecular HIV Surveillance: A Randomized Experiment in 2 National Surveys
JMIR Public Health Surveill 2025;11:e64663
URL: https://publichealth.jmir.org/2025/1/e64663
doi:10.2196/64663

© Jeremy Sugarman, Juli Bollinger, Jose Agostini, Kevin Weinfurt, Gail Geller, Sheethal Jose, Marissa Hannah, O Winslow Edwards, Leslie Meltzer Henry, Travis Sanchez. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 11.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effect of a Brief Video-Based Intervention to Improve AIDS Prevention in Older Men: Randomized Controlled Trial

Tian Huaju[1*], BD; Xie Rendie[2*], BD; Xiao Lu[1], MD; Li Mei[3], MD; Luo Yue[3], MD; Zhang Daiying[1], BD; Chen Yanhua[4], DD; Ren Jianlan[1*], MD

[1]Department of the Operating Room, the Affiliated Hospital of Southwest Medical University, 25 Taiping Street, Luzhou, Sichuan Province, China

[2]Department of Anesthesiology, the Affiliated Hospital of Southwest Medical University, Luzhou, China

[3]School of Nursing, Southwest Medical University, Luzhou, China

[4]Department of Nursing, the Affiliated Hospital of Southwest Medical University, Luzhou, China

[*]these authors contributed equally

Corresponding Author:
Ren Jianlan, MD
Department of the Operating Room, the Affiliated Hospital of Southwest Medical University, 25 Taiping Street, Luzhou, Sichuan Province, China

Abstract

Background: The AIDS epidemic among older people is becoming more serious. Evidence-based, acceptable, and effective preventive interventions are urgently needed. Video-based interventions have become an innovative way to change behaviors, and we have developed a brief video-based intervention named Sunset Without AIDS.

Objective: In this study, we tested the effectiveness of a brief video-based intervention targeting older men's understanding of AIDS prevention.

Methods: A randomized controlled trial was conducted from June 20 to July 3, 2023. In total, 100 older men were randomly divided into the intervention group (n=50) and the control group (n=50) using the envelope extraction method. The intervention group was shown the Sunset Without AIDS video; the control group viewed a standard AIDS education video. A questionnaire was used to measure the effect of Sunset Without AIDS after 2 interventions. AIDS-related high-risk behaviors were followed up 1 and 3 months after the intervention. The difference was statistically significant at $P \leq .05$.

Results: After 2 interventions, the total awareness rates (%) of AIDS-related knowledge in the intervention and control groups were 84% (42/50) and 66% (33/50), respectively ($P=.04$). The mean stigma attitude scores of the 2 groups were 2.53 (SD 0.45) and 2.58 (SD 0.49), respectively ($P=.55$), but there was a statistically significant difference in the first dimension (fear of infection) between the 2 groups ($P<.001$). The mean positive scores of attitudes of AIDS-related high-risk behaviors of the 2 groups were 83.33 (SD 21.56) and 75.67 (SD 26.77), respectively ($P=.58$). In addition, 82% reported that they were satisfied with the educational content within the Sunset Without AIDS video. At 1- and 3-month follow-ups conducted after the intervention, participants in the 2 groups did not report AIDS-related high-risk behaviors. After watching the 2 videos, more people accepted and were satisfied with Sunset Without AIDS.

Conclusions: Sunset Without AIDS could improve the ability of older men in China to follow best practices for AIDS prevention and provide a certain basis for the innovation of AIDS education in the older adult population.

Trial Registration: Chinese Clinical Trial Registry, ChiCTR2100045708; https://www.chictr.org.cn/showproj.html?proj=125478

(JMIR Public Health Surveill 2025;11:e65674) doi:10.2196/65674

KEYWORDS

older men; AIDS prevention; video; video-based intervention; HIV; AIDS

Introduction

With the acceleration of global population aging, AIDS has become increasingly prevalent among older adults, and the age composition of people living with HIV has gradually tilted toward older people. According to the Joint United Nations Programme on HIV/AIDS data [1], there were approximately 38 million people living with HIV worldwide and about 8.1 million of them were aged ≥50 years old. According to data released by the Chinese Center for Disease Control and Prevention (CDC) AIDS, the number of HIV-positive men over 60 years of age rose from 8391 in 2012 to 19,815 in 2017 [2] and 24,465 in 2018 [3].

On November 30, 2020, the Chinese CDC released core information on AIDS prevention. The core information has

XSL•FO
RenderX

shown that in recent years, more than 95% of newly diagnosed people living with HIV in China contracted HIV through sexual means, and heterosexual transmission accounts for about 70% of these cases; unsafe sex is the main cause of sexual transmission of HIV [4]. The main route of HIV infection in older people in China is sexual transmission, and the imbalance of sexual demand is one of the causes of HIV infection in older people [5]. People remain sexually active after the age of 50, and high-risk sexual behaviors such as having multiple sexual partners and having unprotected sex with casual partners are common in the older population [6]. Due to insufficient AIDS-related knowledge or some personal factors, such as divorce, death of a spouse, or a spouse with no sexual needs, some older people, especially older men, may take part in unsafe sex under the influence of their friends and environment [5,7]. The survey found that up to 80% of older men with HIV admitted to having visited prostitutes, while older women were more likely to be infected within a marriage [7]. A survey showed that 94.7% of older men infected with HIV never used condoms when having sex with a commercial partner [8]. This shows that older men have a low degree of understanding of safe sex and do not take protective measures when having sex with strangers, so they are at risk of infection with and transmission of HIV [9].

Therefore, it is necessary to take effective measures to control the prevalence of HIV/AIDS among older men. Although there is no HIV vaccine available, other biomedical options, such as preexposure prophylaxis (PrEP) and postexposure prophylaxis, have had multiple challenges in terms of the rollout and implementation in China. In China, the key vulnerable population has insufficient awareness of these prevention methods and a low utilization rate. For example, only 20% of men who have sex with men were aware of PrEP and postexposure prophylaxis, and only <1% initiated PrEP [10,11].

Health education is an important tool in curbing the HIV/AIDS epidemic [12]. At present, AIDS health education targets mainly focus on adolescents [13-15] and men who have sex with men [16-18], with less focus on the older adult population. Furthermore, the existing AIDS awareness and education work is not accurate enough, and there is a lack of specific and strong popular science materials for older people [19]. At present, HIV/AIDS health education resources and methods for older adults are obsolete, and now awareness pamphlets, centralized lectures, publicity columns, and banners are more often used [20]. Older adults generally have poor eyesight, low levels of education, few ways of obtaining knowledge, and a lower ability to accept new knowledge. This makes it difficult for them to obtain correct HIV/AIDS prevention and control knowledge, so the effectiveness of education efforts is often not ideal [21,22]. In addition, there is limited discussion of the study's global relevance beyond China [23,24]. Therefore, there is an urgent need to develop innovative models of AIDS education for older people and improve the educational effect, which can be feasibly implemented in overburdened health systems.

Video-based interventions are a promising but underused way to deliver AIDS education to older people. As a kind of teaching media, video interventions can deliver standardized information. At the same time, some useful information can be woven into engaging storylines, piloted to ensure cultural relevance, and delivered at critical teachable moments [25]. In addition to the aesthetic and entertainment value of such videos, they also contain great value in terms of health education.

In recent years, some randomized controlled trials were conducted in the United States and Africa targeting people living with HIV and high-risk populations. The trials found that video-based interventions were highly cost-effective [26] and demonstrated efficacy in improving health knowledge, supporting partner disclosure, increasing medication adherence, promoting HIV screening, and reducing sexual partners, as well as fostering attitude and behavior change [27-33]. However, due to regional and cultural differences, these video-based interventions do not conform to the national characteristics of HIV infection among older people in China. Meanwhile, these videos lack systematization, specialization, and guidance in the knowledge, attitude, and skills related to HIV infection among older people, so their promotion and application are limited. In China, there are few evidence-based studies on HIV/AIDS educational video interventions, especially randomized controlled trials to evaluate the effects of a video-based intervention on older men.

Considering the current HIV epidemic among older men in China, while recognizing that evidence-based, scalable, and accessible interventions to improve the ability of older men to prevent HIV are critical, we developed a brief video-based intervention (Sunset Without AIDS), mainly targeting education to increase AIDS-related knowledge, improve AIDS prevention motivation, and strengthen AIDS prevention behaviors among older men. The video is suitable for China's national conditions and addresses the primary sexual issues related to HIV transmission among older people and can be promoted and applied in the older adult population.

In this study, we tested whether the Sunset Without AIDS video was effective in increasing AIDS-related knowledge, reducing AIDS-related stigma, and reducing AIDS-related high-risk behaviors among older men in Luzhou City, Southwest China. Here we report the results from this study.

Methods

Video Design

The AIDS educational video used in this study is named Sunset Without AIDS and is 25 minutes long. In conceptualizing the video, the Information-Motivation-Behavioral (IMB) Skills Model was selected. The IMB Skills Model involves three aspects: AIDS-related information, motivation, and behavior skills [34]. It states that knowledge alone is insufficient for behavior change it must be linked to both motivation and behavioral skills to overcome barriers critical to achieving adherence [27]. Sunset Without AIDS incorporates all components of the IMB Skills Model and integrates other evidence-based video techniques to facilitate behavior change, such as video modeling and gain-framed information [35]. The gain-framed messaging emphasizes the advantages of behavior change, rather than the disadvantages of not changing behavior. It was selected to promote behavior change due to its

XSL•FO
RenderX

demonstrated superiority over loss-framed messaging [36]. The script of Sunset Without AIDS is based on the true stories of older people living with HIV in China. The development of video content was through a community participatory approach, involving in-depth research and interviews with older men and people living with HIV to ensure that the educational content was acceptable and compelling. The script underwent iterative review and editing by a scriptwriter, until the expert advisory group felt the script addressed these issues in a compelling fashion. The characters in the video are all experienced actors and they speak Chinese. Sunset Without AIDS consists of two parts; one is about 18 minutes of story content, while the other is approximately 7 minutes of expert guidance [35]. In Sunset Without AIDS, the video provides information about the main route of transmission in the older population, the importance of HIV prevention, and education on HIV testing and treatment. Sunset Without AIDS is divided into multiple independent scenes in which topics including AIDS-related knowledge, attitude, and personal protection skills are interwoven into the dialogues and behavior attitudes of characters in specific plots. At the same time, the end of the story is supplemented with expert guidance, through expert authority, to increase awareness regarding HIV prevention and reduce high-risk behaviors related to HIV. The video covers three different health education themes: (1) facing sexual needs and guiding safe sex; (2) protecting yourself from HIV; (3) facing the disease bravely and preventing AIDS scientifically [25]. Screenshots of Sunset Without AIDS are shown in Figure 1.

Figure 1. Screenshots of the Sunset Without AIDS intervention video.

Study Design

A parallel design randomized controlled trial was conducted. The sample size calculation was based on the primary outcome, that is, AIDS-related knowledge. Sufficient statistical power to detect changes in AIDS-related knowledge after the intervention was considered in the sample size calculation, and 36 participants were required in each group to achieve 80% statistical power. The study took place in Kuangchang Town, Luzhou City, which is a Chinese administrative division containing 11 communities. To facilitate the implementation of the trial, a simple random sampling method was used to select 3 communities in Kuangchang Town. Based on the size of the older population in these 3 communities, a total of 100 older men who met the requirements of the trial were recruited in the 3 communities using quota sampling through community advertising and on-site recruitment. Research assistants delivered talks to all eligible participants describing the trial. When the participants agreed to join the study and completed the follow-up measurement, an informed consent form needed to be signed for confirmation. Once participants' eligibility was determined, they were scheduled to complete a baseline survey. The baseline survey included demographic information, AIDS-related knowledge, stigma attitude, AIDS-related attitudes to high-risk behaviors, and AIDS-related high-risk behaviors.

To reduce contamination, the trial took place at 2 different community centers. For the study, 100 participants were randomly assigned to 1 of the 2 intervention groups on a 1:1 basis. Opaque envelopes were prepared for randomization. Each of the 100 envelopes contained a note with the words "Sunset Without AIDS intervention group" or "Normal video group" (50 envelopes per group). Participants (n=100) were randomly assigned a computer number from 1 to 100, and they randomly

XSL•FO
RenderX

selected an envelope in numbered order. For each envelope selected by participants, the total number of envelopes was reduced by 1 until the end. Participants who selected the envelopes in turn opened them in front of a third-party staff member and were assigned to different groups based on the information on the envelope. Another staff member recorded the allocation information of the participants.

After 2 interventions, all participants were evaluated for their AIDS-related knowledge, stigma attitude, and AIDS-related high-risk behaviors attitude. Meanwhile, face-to-face interviews were used to evaluate the feelings of all participants in the intervention group toward Sunset Without AIDS. The two groups' AIDS-related high-risk behaviors were followed up 1 and 3 months after the intervention. Three months after the intervention, the convenience sampling method was used to select 10 older men from the intervention group. After watching the videos of the intervention group and the control group, the 10 older men were interviewed face-to-face to evaluate their satisfaction with and acceptance of the intervention.

We had the following hypotheses: (1) Sunset Without AIDS can increase the awareness of AIDS-related knowledge, reduce AIDS-related stigma, and reduce HIV-related high-risk behaviors among older men; and (2) in a comparison between the Sunset Without AIDS intervention group and the normal video group, the Sunset Without AIDS intervention group would integrate various knowledge and behavioral skills into the story. So, the Sunset Without AIDS intervention could have better acceptability and feasibility for older men's education.

Participants

We recruited 100 older men in three communities in Kuangchang Town, Luzhou City in Southwest China through community advertising and on-site recruitment. The eligibility criteria for participation were as follows: (1) male aged over 60 years old; (2) clear consciousness, no mental illness and serious visual impairment, has a certain level of reading ability and language expression ability; and (3) plans to stay in the trial catchment area for over 3 months. The exclusion criteria were as follows: (1) ongoing or already participating in an AIDS education intervention; and (2) cognitive and mental impairment, inability to communicate properly, or blurred vision or blindness.

Intervention

The intervention group and control group received the corresponding intervention from June 20 to July 3, 2023. The intervention group received 25 minutes of the Sunset Without AIDS video once per week for 2 weeks. During the viewing, a member of the study team was responsible for answering questions. To ensure the compliance of participants, each participant could receive recognition and a reward of 50 RMB (US $6.91) in cash after completing each video viewing and follow-up assessment.

Participants in the normal video group received an AIDS education video intervention downloaded from the website of the Chinese CDC, which was composed of 5 short video clips, about 25 minutes in total. Intervention times and duration were the same as the Sunset Without AIDS group. The video contained similar educational content as the Sunset Without AIDS, including AIDS-related knowledge and attitude and behavior education. The only difference is that it is a popular science informational video, which lacks a storyline and is not specifically aimed at older people. As in the intervention group, a member of the study team was responsible for answering questions from the participants during the viewing.

Outcomes and Measurements

The primary outcomes comprised change in score, that is, increased AIDS-related knowledge, decreased AIDS-related stigma attitude, and improved AIDS-related high-risk behaviors attitude. The secondary outcomes were the feelings of the intervention group after 2 interventions and their AIDS-related high-risk behaviors 1 and 3 months after the intervention. At the same time, the satisfaction and acceptance of both interventions were also evaluated after 3 months of intervention.

AIDS-related knowledge was assessed through 8 core knowledge questions on AIDS commonly used by the Chinese AIDS Prevention and Control Committee [19]. One point was given for a correct answer. The higher the score, the more knowledge participants have obtained. The total awareness rate of AIDS knowledge refers to the proportion of the respondents who correctly answered 6 or more of the 8 basic AIDS knowledge items simultaneously.

AIDS-related stigma attitude was assessed through the Chinese version of Zelaya et al's [37] HIV/AIDS Stigma Scale. The scale was translated into Chinese by Xing et al in 2014 [38]. All items are divided into four dimensions: fear of infection, stigma prejudice, personal discrimination, and social discrimination, with 6 items for each dimension, for a total of 24 items. The scale contains 6 positive items and 18 negative items, and the answers are strongly agree, agree, uncertain, oppose, and strongly disagree. Items are scored on a 5-point scale (1 to 5) and negative items are reverse scored. The higher the participants' discrimination score, the more severe the AIDS discrimination [39].

AIDS-related high-risk behaviors attitude was assessed through a 6-item questionnaire. It included attitude toward AIDS-related high-risk behaviors such as seeking HIV testing and counseling, seeking sexual partners through commercial/social/internet methods, having one-night stands, having multiple sexual partners, and having unprotected sex. The higher the questionnaire score, the lower the participant's risk of engaging in AIDS-related high-risk behaviors. The content validity of the questionnaire was evaluated by 5 experts; the item-level content validity index and the scale-level convent validity index were both 1.00.

AIDS-related high-risk behaviors were assessed through a self-made questionnaire with 3 items [22]. The questionnaire included whether someone sought sexual partners through business, social, or online means, and whether they had sex with strangers. The fewer AIDS-related high-risk behaviors the participants had, the better their ability to prevent HIV/AIDS.

The outline of the user experience interview mainly included: (1) whether you accepted the education of AIDS related-knowledge in the form of watching videos; (2) were you

XSL•FO
RenderX

satisfied with the educational content of the Sunset Without AIDS video; (3) whether the story and theme of the video could attract you and arouse your attention; (4) after watching the video, whether you had a clear understanding of the main route of HIV infection in older people, and emphasized the use of condoms; and (5) did you know how and where to get tested for HIV after watching the video?

Data Analysis

SPSS (version 21.0; IBM Corp) statistical software was used for data analysis. Percentage was used for statistical descriptions of classified data, and χ^2 test was used for qualitative data to compare each indicator; the statistical description of the quantitative data was given using mean (SD), while *t* test was used to compare each indicator between groups, and $P \leq .05$ was considered a statistically significant difference. The data analysts were blinded to the allocation.

Ethical Considerations

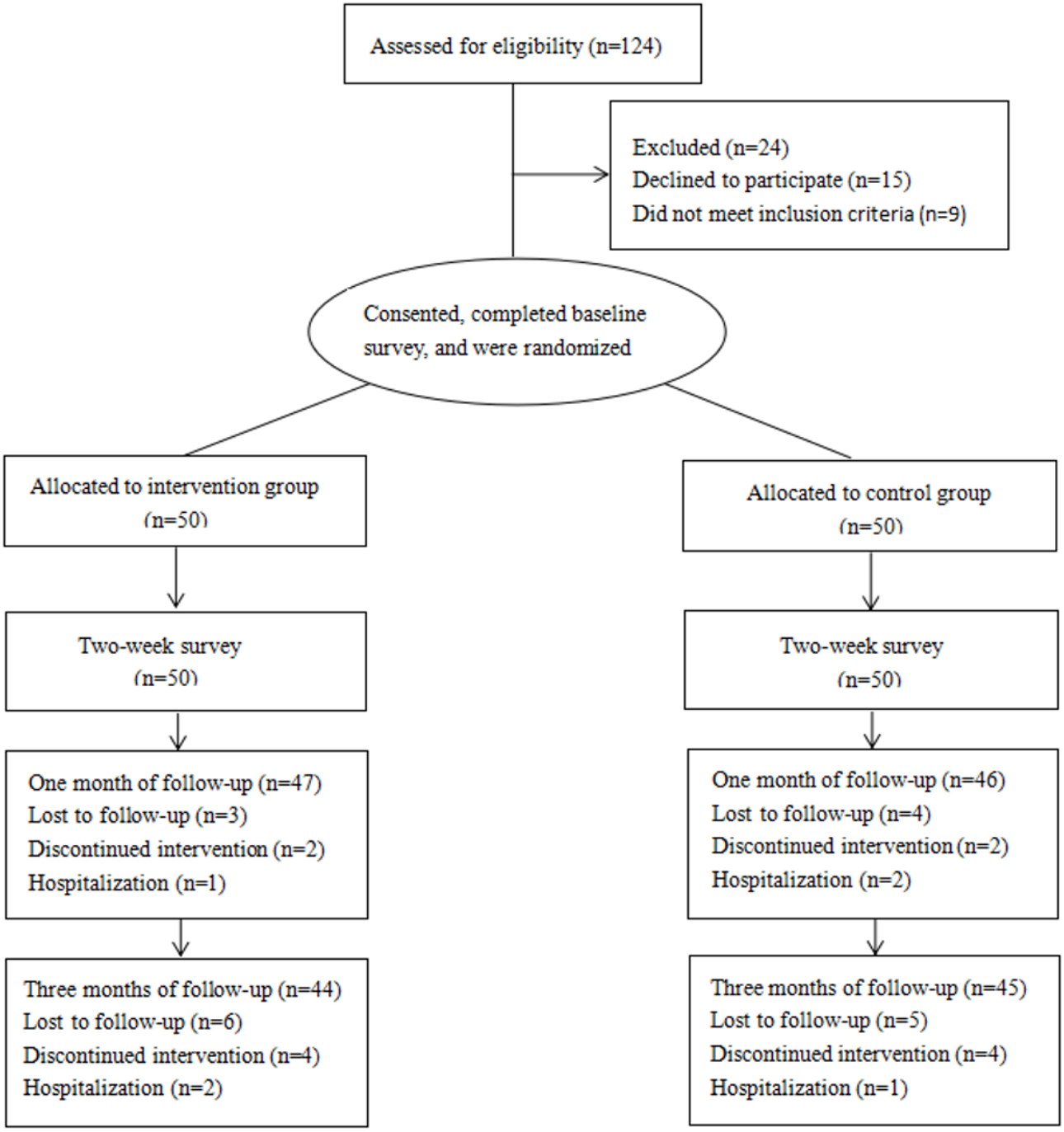
Study approval was gained from the ethics review committees of the Affiliated Hospital of Southwest Medical University (number KY2021071). This study was registered at the Chinese Clinical Trial Registry (ChiCTR2100045708). All participants who participated in the survey signed an informed consent form. Each participant could receive 50 RMB (US $6.91) in cash after completing each video viewing and follow-up assessment. The informed consent form also included a statement that study data were anonymous or would be deidentified and could be used in a secondary analysis of data without additional consent.

Results

Figure 2 depicts participant flow throughout the study from screening to follow-up.

XSL•FO
RenderX

Figure 2. The participant flow throughout the study from screening to follow-up.



Description of Study Sample

The study sample included 100 participants (intervention group: n=50, control group: n=50) aged between 60 and 78 years old, all Han nationality. Urban older people accounted for 72% (n=72) of the sample, and rural older people accounted for 28% (n=28). Married, divorced, and widowed people accounted for 77% (n=77), 8% (n=8), and 15% (n=15), respectively. Illiteracy, primary school, junior high school, and academic education accounted for 8% (n=8), 33% (n=33), 34% (n=34), and 25% (n=25), respectively. Participants were mainly from enterprise institutions (n=30, 30%) and farming (n=27, 27%). The 2 groups of participants acquired AIDS-related knowledge mainly through television and hearing it from others. There were no significant demographic differences between the 2 groups, with comparability (Table 1).

XSL•FO
RenderX

Table . Participant demographics (n=100).

Characteristics		Intervention group (n=50)	Control group (n=50)	*t* test/chi-square (*df=98*)	*P* value
Age (years), mean (SD)		68.62 (4.72)	68.48 (4.99)	0.129	.90
Marriage, n (%)				0.684	.71
	Married	37 (74)	40 (80)		
	Divorced	5 (10)	3 (6)		
	Widowed	8 (16)	7 (14)		
Education, n (%)				1.25	.74
	Illiterate	3 (6)	5 (10)		
	Primary school	15 (30)	18 (36)		
	Junior high school	18 (36)	16 (32)		
	Academic	14 (28)	11 (22)		
Occupation, n (%)				0.477	.98
	Enterprise institution	14 (28)	16 (32)		
	Business	3 (6)	2 (4)		
	Farming	14 (28)	13 (26)		
	No job	9 (18)	8 (16)		
	Other	10 (20)	11 (22)		
Inhabiting information, n (%)				0.88	.83
	Solitary	10 (20)	9 (18)		
	Couple (partner) living together	10 (20)	14 (28)		
	Live with children	18 (36)	16 (32)		
	Three or more generations live together	12 (24)	11 (22)		
Access to AIDS information (multiple choices available), n (%)					
	TV	44 (88)	42 (84)	0.332	.56
	Newspaper	14 (28)	8 (16)	2.098	.15
	Internet	19 (38)	12 (24)	2.291	.13
	Broadcast	12 (24)	15 (30)	0.457	.50
	Community and hospital publicity	13 (26)	10 (20)	0.508	.48
	Hearing	35 (70)	40 (80)	1.333	.25
	Other	2 (4)	5 (10)	1.382	.24

Comparison of Indicators Between the Two Groups Before the Intervention

Two people in the intervention group reported seeking sexual partners through business, social networking, or the internet and having sex with strangers. Both reported not using condoms when having sex with strangers. Three people in the control group reported seeking sexual partners through business, social networking, or the internet and having sex with strangers. Among them, 2 reported not using condoms when having sex with strangers, and 1 reported sometimes using condoms when having sex with strangers. Before the intervention, there was no statistically significant difference in the AIDS-related knowledge, stigma attitude, AIDS-related high-risk behaviors attitude, and AIDS-related high-risk behaviors between the 2 groups (Table 2).

XSL•FO
RenderX

Table . Comparison of indicators between the two groups before intervention (n=100).

Items	Intervention group (n=50)	Control group (n=50)	*t* test/chi-square (*df=98*)	*P* value
Total awareness rate of AIDS-related knowledge, correct rate, n (%)	15 (30)	20 (40)	1.099	.30
1. Is it possible to get AIDS if you share syringes with a person with HIV?	37 (74)	32 (64)	1.169	.28
2. Can HIV infection be seen on the outside?	32 (64)	33 (66)	0.044	.83
3. Are children born to women with HIV likely to develop AIDS?	26 (52)	31 (62)	1.02	.31
4. Can proper condom use reduce HIV transmission?	36 (72)	40 (80)	0.877	.35
5. Can mosquito bites spread HIV?	8 (16)	13 (26)	1.507	.22
6. Can you get AIDS from eating with people with HIV or sick people?	27 (54)	29 (58)	0.162	.69
7. Does having sex with only one partner reduce HIV transmission risk?	35 (70)	32 (64)	0.407	.52
8. Can you get AIDS from HIV-infected blood?	36 (72)	37 (74)	0.051	.82
Stigma score, mean (SD)	2.91 (0.38)	2.99 (0.52)	0.897	.37
Dimension 1: fear of infection	3.16 (0.57)	3.22 (0.74)	0.482	.63
Dimension 2: stigma prejudice	3.12 (0.52)	3.12 (0.70)	0.05	.96
Dimension 3: personal discrimination	2.48 (0.47)	2.46 (0.53)	0.13	.90
Dimension 4: social discrimination	2.89 (0.77)	3.16 (0.69)	1.898	.06
Positive attitude rate of AIDS-related high-risk behaviors, %, mean (SD)	58.33 (32.34)	59.33 (32.67)	0.15	.88
1. If you have high-risk behavior, will you take the initiative to seek HIV testing and counseling?	54.00 (50.35)	56.00 (50.14)	0.216	.83
2. Do you agree with the behavior of seeking sexual partners through business, social networking, or the internet?	62.00 (49.03)	66.00 (47.85)	0.389	.70
3. Do you approve of having a one-night stand with a stranger?	72.00 (45.36)	70.00 (46.29)	0.24	.81
4. Do you approve of having multiple sexual partners?	72.00 (45.36)	70.00 (46.29)	0.227	.82
5. Do you agree to use condoms if you have sex with strangers?	46.00 (50.35)	48.00 (50.47)	0.24	.81

XSL•FO
RenderX

Items	Intervention group (n=50)	Control group (n=50)	*t* test/chi-square (*df=98*)	*P* value
6. Do you approve of taking Viagra and other drugs during sex?	44.00 (50.14)	46.00 (50.35)	0.184	.86

Comparison of Indicators Between the Two Groups Before and After the Intervention

After 2 interventions, the total awareness rate of AIDS-related knowledge in the intervention group was higher than that in the control group ($P<.05$). The difference in the stigma score was not statistically significant after the intervention ($P=.55$). However, there was a statistical difference in the first dimension (fear of infection) between the 2 groups ($P<.001$). The difference in the positive attitude rate of AIDS-related high-risk behaviors was also not statistically significant after the intervention ($P=.12$). Before and after the intervention, the total awareness rate of AIDS-related knowledge and positive attitude rate of high-risk AIDS-related behaviors in both groups were higher than before the intervention, and the stigma score was lower than before the intervention; the difference was statistically significant ($P<.01$; Table 3).

XSL•FO
RenderX

Table . Comparison of indicators between the 2 groups before and after 2 interventions.

Items	Intervention group (n=50)				Control group (n=50)				Difference between the 2 groups after the intervention	
	Before intervention	After intervention	*t* test/chi-square (*df=98*)	*P* value	Before intervention	After intervention	*t* test/chi-square (*df=98*)	*P* value	*t* test/chi-square (*df=98*)	*P* value
Total awareness rate of AIDS-related knowledge, correct rate, n (%)	15 (30)	42 (84)	29.743	<.001	20 (40)	33 (66)	6.784	.01	4.32	.04
1. Is it possible to get AIDS if you share syringes with a person with HIV?	37 (74)	47 (94)	7.44	.006	32 (64)	43 (86)	6.453	.01	1.778	.18
2. Can HIV infection be seen on the outside?	32 (64)	40 (80)	3.175	.08	33 (66)	32 (64)	0.044	.83	3.175	.08
3. Are children born to women with HIV likely to develop AIDS?	26 (52)	46 (92)	19.841	<.001	31 (62)	43 (86)	7.484	.006	0.919	.34
4. Can proper condom use reduce HIV transmission?	36 (72)	46 (92)	6.775	.009	40 (80)	41 (82)	0.065	.78	2.21	.14
5. Can mosquito bites spread HIV?	8 (16)	29 (58)	18.919	<.001	13 (26)	17 (34)	0.762	.38	5.797	.02
6. Can you get AIDS from eating with people with HIV or sick people?	27 (54)	40 (80)	7.644	.006	29 (58)	35 (70)	1.563	.21	1.333	.25
7. Does having sex with only one partner reduce HIV transmission risk?	35 (70)	43 (86)	3.73	.053	32 (64)	41 (82)	4.11	.04	0.298	.58

XSL•FO
RenderX

Items	Intervention group (n=50)				Control group (n=50)				Difference between the 2 groups after the intervention	
	Before intervention	After intervention	*t* test/chi-square (*df=98*)	*P* value	Before intervention	After intervention	*t* test/chi-square (*df=98*)	*P* value	*t* test/chi-square (*df=98*)	*P* value
8. Can you get AIDS from HIV-infected blood?	36 (72)	47 (94)	8.575	.003	37 (74)	46 (92)	5.741	.02	0.154	.70
Stigma score, mean (SD)	2.91 (0.38)	2.53 (0.45)	6.20	<.001	2.99 (0.52)	2.58 (0.49)	9.531	<.001	0.604	.55
Dimension 1: fear of infection	3.16 (0.57)	2.69 (0.67)	5.86	<.001	3.22 (0.74)	2.79 (0.76)	5.488	<.001	3.921	<.001
Dimension 2: stigma prejudice	3.12 (0.52)	2.80 (0.61)	3.662	.001	3.12 (0.70)	2.71 (0.69)	6.779	<.001	0.69	.49
Dimension 3: personal discrimination	2.48 (0.47)	2.14 (0.45)	4.975	<.001	2.46 (0.53)	2.27 (0.59)	3.023	.004	1.375	.18
Dimension 4: social discrimination	2.89 (0.77)	2.50 (0.64)	4.18	<.001	3.16 (0.69)	2.57 (0.52)	7.382	<.001	0.628	.53
Positive attitude rate of AIDS-related high-risk behaviors, %, mean (SD)	58.33 (32.34)	83.33 (21.56)	5.935	<.001	59.33 (32.67)	75.67 (26.77)	4.085	<.001	1.592	.12
1. If you have high-risk behaviors, will you take the initiative to seek HIV testing and counseling?	54.00 (50.35)	88.00 (32.83)	4.629	<.001	56.00 (50.14)	82.00 (38.81)	3.487	.001	0.829	.41
2. Do you agree with the behavior of seeking sexual partners through business, social networking, or the internet?	62.00 (49.03)	92.00 (27.41)	4.20	<.001	66.00 (47.85)	86.00 (35.05)	3.13	.003	1.00	.32

XSL•FO
RenderX

Items	Intervention group (n=50)				Control group (n=50)				Difference between the 2 groups after the intervention	
	Before intervention	After intervention	*t* test/chi-square (*df=98*)	*P* value	Before intervention	After intervention	*t* test/chi-square (*df=98*)	*P* value	*t* test/chi-square (*df=98*)	*P* value
3. Do you approve of having a one-night stand with a stranger?	72.00 (45.36)	92.00 (27.41)	3.13	.003	70.00 (46.29)	92.00 (27.41)	3.348	.002	0.00	>.99
4. Do you approve of having multiple sexual partners?	72.00 (45.36)	88.00 (32.83)	2.419	.02	70.00 (46.29)	86.00 (35.05)	2.682	.01	0.299	.77
5. Do you agree to use condoms if you have sex with strangers?	46.00 (50.35)	74.00 (44.31)	4.365	<.001	48.00 (50.47)	60.00 (49.49)	1.52	.14	1.549	.13
6. Do you approve of taking Viagra and other drugs during sex?	44.00 (50.14)	66.00 (47.85)	2.852	.006	46.00 (50.35)	48.00 (50.47)	0.216	.83	1.843	.07

Evaluation of User Experience in the Intervention Group

After 2 interventions, the user feelings of the intervention group after watching the Sunset Without AIDS video were evaluated by face-to-face interview. Overall, 92% (46/50) reported that they were willing to accept AIDS-related knowledge education in the form of watching videos and 82% (41/50) reported that they were satisfied with the educational content of the Sunset Without AIDS video. In addition, 94% (47/50) reported that the story and theme of the video could attract them and arouse their attention and 90% (45/50) reported that after watching the video, they had a clear understanding of the main route of HIV infection in older people and emphasized the use of condoms. Finally, 88% (44/50) reported that they knew how and where to conduct HIV testing after watching the video. The specific feelings of the intervention group after watching the video mainly focused on the 3 aspects displayed in Textbox 1.

XSL•FO
RenderX

Textbox 1. The specific feelings of the intervention group.

AIDS awareness

1. Some understanding about HIV/AIDS.
2. Know how HIV is spread.
3. AIDS had been thought to be a disease of the young, but it turned out to be a disease of older people as well.

Personal and social behavior

1. Be responsible for our family, be responsible for our children.
2. We should not discriminate against people with AIDS.
3. The sexual needs of older people should be faced squarely.
4. Society should strengthen the propaganda and guidance of AIDS prevention.
5. Pay attention to safe sex and use condoms.

Evaluation of the Sunset Without AIDS intervention

1. The video was very meaningful it felt like the people and things in the video were happening around us.
2. It was closer to our life for older people.
3. The story of the video was very good and could attract me very much.
4. This way was very suitable for older adults.

Follow-Up Evaluation of AIDS-Related High-Risk Behaviors

After 1 and 3 months of the intervention, the 2 groups were followed up to evaluate the changes to AIDS-related high-risk behaviors. The follow-up rates of the intervention group and the control group were 94% (47/50) and 92% (46/50) 1 month after the intervention. Three months after the intervention, the follow-up rates of the intervention group and control group were 88% (44/50) and 90% (45/50), respectively. At 2 follow-up visits, participants in both groups reported that they did not seek sexual partners through business, social networking, or the internet and did not have sexual activities with strangers after the intervention. The main characteristics of sexual behavior were compared between those who were lost to follow-up and those who had participated in the follow-up evaluation. The results showed that before the intervention, none of those lost to follow-up reported seeking sexual partners through business, social networking, or the internet and engaging in sexual activities with strangers. Compared with before the intervention, AIDS-related high-risk behaviors decreased in the 2 groups, but there was no statistical significance ($P>.05$).

Evaluation of Acceptance and Satisfaction

After 3 months of follow-up, 10 older men in the intervention group were interviewed after watching the videos of the intervention group and the control group. The results showed that, compared with traditional forms of health education such as AIDS book publicity and lectures, 7 reported that they were more willing to watch videos, 2 reported that they were more willing to use traditional health education, and 1 reported that there was no difference. After watching the Sunset Without AIDS video and normal AIDS education video, 9 accepted and were satisfied with the Sunset Without AIDS video.

Discussion

Principal Findings

In this study, a randomized controlled trial was conducted to evaluate the effectiveness of Sunset Without AIDS on preventing HIV among older men in China. The results showed that before the intervention, 100 participants had a very low awareness rate of AIDS-related knowledge and a serious discriminatory attitude toward AIDS, which was basically consistent with the relevant survey results [22,40-43]. Among AIDS-related knowledge, for the question “Can mosquito bites spread HIV?” only 21% were correct. It showed that there was a certain misunderstanding of the method of HIV transmission and a lack of understanding of AIDS-related knowledge among the older men. This suggested that it is urgent to carry out effective AIDS health education for older people.

The study found that after the intervention, the total awareness rate and item 5 awareness rate of AIDS-related knowledge in the intervention group were higher than those in the control group; the difference was statistically significant ($P<.05$). It was suggested that the education effect of the Sunset Without AIDS video was better than that of the normal AIDS propaganda video in improving the knowledge level of AIDS in older men. At present, HIV/AIDS prevention and treatment focused on scientific and accurate propaganda content, but ignored the relevance of the propaganda content for the target population [44]. Sunset Without AIDS considered and told the story from the perspective of older people, which was not only more targeted, but also in line with the daily life of older people. In the form of a film, Sunset Without AIDS vividly showed the experience of HIV infection among older people, which conveyed educational information, and greatly attracted users’ interest in learning and enhanced their learning motivation. At the same time, participants had access to targeted expert

XSL•FO
RenderX

guidance to strengthen the relevant knowledge to improve the effect of education.

Studies have shown that all film and television materials have the potential to profoundly affect emotions and knowledge [45]; this was consistent with the results of this study. After the intervention, the total awareness rate of AIDS-related knowledge in both groups was higher than that before the intervention (P<.01). This result was consistent with the application effect of video-based intervention in college students [45]. After the application of the Sunset Without AIDS intervention, the total awareness rate of AIDS-related knowledge in the intervention group and the awareness rate of AIDS-related knowledge in the other items except item 2 and item 7 were all higher than before the intervention and the differences were statistically significant. Among them, the awareness rate of item 3 and item 5 increased most obviously (P<.001). This showed that Sunset Without AIDS had a good effect on the propaganda and popularization of nonsexual transmission of AIDS under the educational background that focused on sexual transmission of AIDS.

Influenced by traditional concepts in China, older adults generally avoid talking about sexual issues, and this group had greater discrimination and prejudice against HIV/AIDS than other age groups [19]. Sunset Without AIDS showed some effect in reducing AIDS-related stigma among older men; it has the potential to become an effective educational tool to reduce AIDS-related stigma among older men. The study found that although there was no statistically significant difference in the AIDS-related stigma attitude score between the 2 groups after the intervention, there was a statistical difference in the first dimension (fear of infection*;* P<.001). This might be related to the positive emotion emphasized in Sunset Without AIDS, and the portrayal of optimistic and strong images presented by the media. On the one hand, it uses sympathy and empathy to make the viewer think and consider themselves in that position, helping them understand the plight of people living with HIV to reduce discrimination against them. On the other hand, the intervention depicts the brave face of people living with HIV while HIV/AIDS prevention volunteers answer the older adults' questions, underscoring that normal daily life communication will not lead to the spread of HIV/AIDS so as to ease public concern.

Attitude has a very important effect on the occurrence of behavior. At present, interventions targeted at AIDS-related stigma reduction are mainly in the form of peer education, health lectures, and face-to-face counseling. After intervention, a high positive rate of AIDS-related stigma attitudes can be maintained in the short term, but the long-term effect is not good or needs to be studied [46-48]. This suggests that in changing AIDS-related stigma in older adults, we need to adopt some better, more sustainable and effective intervention tools and intervention methods.

The results showed that after 2 interventions, the total positive attitude rate of AIDS-related high-risk behaviors and the positive rate of each item in the intervention group were increased; the improvement degree was statistically significant (P<.05). Among them, items 1, 2, and 5 showed the most significant improvement (P<.001). This shows that Sunset Without AIDS could improve the positive attitude rate of AIDS-related high-risk behaviors in older men. This might be because the video introduced the core knowledge of AIDS in an easy-to-understand way while adding some warning educational content. Through watching this content, the older men could realize the severe situation of the AIDS epidemic and enhance their awareness of self-protection [49], thus promoting a positive change in AIDS-related high-risk behaviors attitude.

The user experience evaluation results of Sunset Without AIDS suggested that the users had a high overall evaluation of this video, especially in the evaluation of "whether the story and theme of the video could attract them and arouse their attention"—as high as 94% of the older men said yes. In terms of the evaluation of users' specific feelings, they were particularly satisfied with the story and immersion of the film. It showed that in terms of design, the film with story combined the scientific aspects of AIDS education with the entertainment value of a film, which could fully attract and mobilize the interest of users and satisfy their curiosity and thirst for knowledge. It was worth noting that in the interviews with users, some older men put forward that: "The sexual needs of older people should be faced squarely." Compared with previous studies, Sunset Without AIDS did not blindly emphasize abstinence, but directly faced the primary sexual issue involved in the spread of AIDS among older people.

Follow-up data found that, although there was a reduction in AIDS-related high-risk behaviors attitude between the two groups, the difference was not statistically significant compared with before the intervention. The reasons might be as follows: (1) the sample size of this study was small, which might lead to insignificant differences; and (2) the intervention was only carried out twice and the follow-up survey of AIDS-related high-risk behaviors was only 1 and 3 months after the intervention. If a long-term intervention could be conducted and the follow-up time could be extended to evaluate the long-term effects, significant differences might be obtained.

The study found that there was no significant difference in some educational effects between the Sunset Without AIDS intervention and the normal video intervention, but compared with the normal video intervention, having one with a storyline that integrated various knowledge and behavioral skills led to better acceptance and satisfaction and it was more favored by the older men. Therefore, as a potential, targeted, and innovative AIDS health education tool, Sunset Without AIDS is worth sharing and promoting among older men.

Limitations

This study is a pilot randomized controlled trial and the total sample size was small (100 cases). In addition, the intervention time was short—only 2 weeks. The participants of this study were only recruited in Kuangchang Town, Luzhou City. It did not include research participants from different cities and regions, so the representativeness of the research participants might be insufficient. Among the samples, the urban older adults accounted for a high proportion, and most of them were between 60 and 70 years old, which might result in the overrepresentation of this group. In the next study, we will conduct a large-scale,

XSL•FO
RenderX

continuous, multicenter, gender-specific study to improve the scientific nature of the study.

HIV/AIDS and sex are both relatively private and sensitive topics; therefore, in Sunset Without AIDS, there may be some educational content that the older men cannot accept. In the future, we will improve the content of the video, integrate AIDS education into topics that the older men are interested in, and make the educational content more simple and easier to understand to have a greater effect on AIDS-related knowledge, stigma, and sexual behavior.

In addition, the next step will also be to evaluate the long-term effect of the Sunset Without AIDS video and study the mechanism of action, laying the foundation for the large-scale promotion and application of this video.

Conclusion

Sunset Without AIDS could improve the ability of older men in China to prevent AIDS and provide a certain basis for the innovation of AIDS education in the older adult population. At the same time, it could also have a broader impact on public health policy formulation and practice. It could effectively increase AIDS-related knowledge among older men, but there was no observed effect in reducing AIDS-related stigma, improving attitudes toward high-risk AIDS behaviors, and reducing AIDS-related high-risk behaviors. Large-scale and long-term effects studies are needed to assess the efficacy of the Sunset Without AIDS intervention.

Acknowledgments

We wish to thank the members of the research team for their support and Roger Watson for carefully revising the paper. This project is funded by National Natural Science Foundation of China (82304256), Luzhou Social Science Applied Research Quality Improvement Program (LZSK-GH029), and Luzhou Municipal People's Government-Southwest Medical University Science and Technology Strategic Cooperation Program (2024LZXNYDR009). The funding body had no role in the design of the study; collection, analysis, and interpretation of data; and writing the manuscript.

Data Availability

The datasets used and/or analyzed as well as data collection forms and model consent forms of this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH V1.6.
[PDF File, 94 KB - publichealth_v11i1e65674_app1.pdf]

References

1. AIDSinfo: people living with HIV. UNAIDS. 2023 Nov 8. URL: https://aidsinfo.unaids.org/ [accessed 2025-02-28]
2. National Health Commission of the People's Republic of China. 2018 Nov 23. URL: http://www.nhc.gov.cn/wjw/xwdt/201811/b0ca3817224e44899a3248a0f6e48948.shtml [accessed 2025-02-28]
3. World AIDS Day Focus: these changes have emerged among domestic infected people! [Website in Chinese]. Chinanews. 2020 Dec 1. URL: http://www.chinanews.com/sh/2020/12-01/9351177.shtml [accessed 2025-02-28]
4. The latest core information on AIDS prevention in 2020 is released! [Website in Chinese]. Chinese Center for Disease Control and Prevention. 2023 Nov 8. URL: https://www.chinacdc.cn/ [accessed 2025-02-28]
5. Gao Y, Lu HZ. Research progress on influencing factors of AIDS-related high-risk sexual behavior in the elderly. Chin J AIDS STD 2020;26(4):452-454. [doi: 10.13419/j.cnki.aids.2020.04.31]
6. Liu DD, Cao XB. Research progress on risk factors of HIV/AIDS infection in male ≥50 years old in China. Chin J AIDS 2020;26(3). [doi: 10.13419/j.cnki.aids.2020.03.28]
7. He B LP, Nong QX, et al. Analysis of the middle-aged and old people at high risk of AIDS and the influencing factors of their high-risk behavior. Chin J Dermatovenereol 2017(6):641-644. [doi: 10.13735/j.cjdv.1001-7089.201609015]
8. Jie L. Investigation and analysis of HIV infection pathway and related influencing factors in the elderly in western Hunan. CNKI. 2010. URL: https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD2011&filename=2010254641.nh [accessed 2025-02-28]
9. Lovejoy TI, Heckman TG. Depression moderates treatment efficacy of an HIV secondary-prevention intervention for HIV-positive late middle-age and older adults. Behav Med 2014;40(3):124-133. [doi: 10.1080/08964289.2014.893982] [Medline: 25090365]
10. Wang H, Tang W, Shang H. Expansion of PrEP and PEP services in China. Lancet HIV 2022 Jul;9(7):e455-e457. [doi: 10.1016/S2352-3018(22)00138-2] [Medline: 35688167]

XSL•FO
RenderX

11. Zhang Y, Wang M, Tan SM, et al. The progress and promotion of HIV pre-exposure prophylaxis in MSM. Chin J AIDS STD 2022;28(7):872-877. [doi: 10.13419/j.cnki.aids.2022.07.28]
12. Wang L, Yu XM. Evaluation of the effect of health education on AIDS prevention for adolescents outside school. Chin J AIDS STD 2016;22(7):532-542. [doi: 10.13419/j.cnki.aids.2016.07.13]
13. Yang M, Yang S, Cao CY. Current situation and thinking of AIDS health education for college students. China Higher Medical (8):20-21.
14. Kim YK, Small E, Okumu M. School-based HIV/AIDS education, risky sexual behaviors, and HIV testing among high school students in the United States. Soc Work Health Care 2019 Mar;58(3):258-273. [doi: 10.1080/00981389.2018.1558163] [Medline: 30556490]
15. Sarma H, Islam MA, Khan JR, Chowdhury KIA, Gazi R. Impact of teachers training on HIV/AIDS education program among secondary school students in Bangladesh: a cross-sectional survey. PLoS One 2017;12(7):e0181627. [doi: 10.1371/journal.pone.0181627] [Medline: 28742103]
16. Chan ASW, Tang PMK. Application of novel psychoactive substances: chemsex and HIV/AIDS policies among men who have sex with men in Hong Kong. Front Psychiatry 2021;12:680252. [doi: 10.3389/fpsyt.2021.680252] [Medline: 34335329]
17. Chan ASW, Tang PMK, Yan E. Chemsex and its risk factors associated with human immunodeficiency virus among men who have sex with men in Hong Kong. World J Virol 2022 Jul 25;11(4):208-211. [doi: 10.5501/wjv.v11.i4.208] [Medline: 36159612]
18. Chan ASW, Leung LM, Li JSF, et al. Impacts of psychological wellbeing with HIV/AIDS and cancer among sexual and gender minorities: a systematic review and meta-analysis. Front Public Health 2022;10:912980. [doi: 10.3389/fpubh.2022.912980] [Medline: 36523575]
19. Ren JL, Li M, Xie RD, Chen YH. Intervention effect of microfilm in AIDS health education of elderly population. Chin J AIDS STD 2022;28(3):340. [doi: 10.13419/j.cnki.aids.2022.03.18]
20. He WY, Chen JW, Deng XY. Explore the extension and connotation of health education in rural digital film health popularization promotion project. Presented at: Proceedings of the 10th Guangzhou Health Education Academic Exchange Conference; Nov 29, 2016; Guangzhou p. 3. [doi: 10.26914/c.cnkihy.2016.007622]
21. Jiang F, Wu WQ, Liang AH. The effect of group discussion on AIDS education among older people in the community. Chin J AIDS STD 2018;24(1):97-98. [doi: 10.13419/j.cnki.aids.2018.01.27]
22. Ren JL, Luo Y, Li M, et al. Investigation on AIDS related knowledge, attitude and behavior among older people in Luzhou city. Chin J AIDS STD 2021;27(2):146-149. [doi: 10.13419/j.cnki.aids.2021.02.08]
23. Kalula SZ, Blouws T. Older persons' knowledge of HIV and AIDS prevention in a province of South Africa. Afr J Prim Health Care Fam Med 2024 Aug 9;16(1):e1-e10. [doi: 10.4102/phcfm.v16i1.4264] [Medline: 39221733]
24. Kalula SZ, Blouws T, Ramathebane M, Sayed AR. HIV and AIDS prevention: knowledge, attitudes, practices and health literacy of older persons in the Western Cape and KwaZulu-Natal Provinces, South Africa and in Lesotho. BMC Geriatr 2023 May 8;23(1):279. [doi: 10.1186/s12877-023-04009-7] [Medline: 37158837]
25. O'Donnell L, San Doval A, Duran R, O'Donnell CR. The effectiveness of video-based interventions in promoting condom acquisition among STD clinic patients. Sex Transm Dis 1995;22(2):97-103. [doi: 10.1097/00007435-199503000-00004] [Medline: 7624818]
26. Sweat M, O'Donnell C, O'Donnell L. Cost-effectiveness of a brief video-based HIV intervention for African American and Latino sexually transmitted disease clinic clients. AIDS 2001 Apr 13;15(6):781-787. [doi: 10.1097/00002030-200104130-00014] [Medline: 11371693]
27. Kim MH, Ahmed S, Tembo T, et al. VITAL Start: video-based intervention to inspire treatment adherence for life-pilot of a novel video-based approach to HIV counseling for pregnant women living with HIV. AIDS Behav 2019 Nov;23(11):3140-3151. [doi: 10.1007/s10461-019-02634-1] [Medline: 31410618]
28. Tuong W, Larsen ER, Armstrong AW. Videos to influence: a systematic review of effectiveness of video-based education in modifying health behaviors. J Behav Med 2014 Apr;37(2):218-233. [doi: 10.1007/s10865-012-9480-7] [Medline: 23188480]
29. Neumann MS, Plant A, Margolis AD, et al. Effects of a brief video intervention on treatment initiation and adherence among patients attending human immunodeficiency virus treatment clinics. PLoS One 2018;13(10):e0204599. [doi: 10.1371/journal.pone.0204599] [Medline: 30289884]
30. Neumann MS, Plant A, Margolis AD, Flores SA. Observed reactions among patients attending HIV treatment facilities to a brief video intervention on treatment initiation and adherence. AIDS Care 2020 May;32(5):656-665. [doi: 10.1080/09540121.2019.1695729] [Medline: 31766857]
31. Chiasson MA, Shaw FS, Humberstone M, Hirshfield S, Hartel D. Increased HIV disclosure three months after an online video intervention for men who have sex with men (MSM). AIDS Care 2009 Sep;21(9):1081-1089. [doi: 10.1080/09540120902730013] [Medline: 20024766]
32. Ingersoll KS, Farrell-Carnahan L, Cohen-Filipic J, et al. A pilot randomized clinical trial of two medication adherence and drug use interventions for HIV+ crack cocaine users. Drug Alcohol Depend 2011 Jul 1;116(1-3):177-187. [doi: 10.1016/j.drugalcdep.2010.12.016] [Medline: 21306837]

XSL•FO
RenderX

33. Essien EJ, Mgbere O, Monjok E, Ekong E, Holstad MM, Kalichman SC. Effectiveness of a video-based motivational skills-building HIV risk-reduction intervention for female military personnel. Soc Sci Med 2011 Jan;72(1):63-71. [doi: 10.1016/j.socscimed.2010.10.012] [Medline: 21106284]
34. Fisher JD, Fisher WA. Changing AIDS-risk behavior. Psychol Bull 1992 May;111(3):455-474. [doi: 10.1037/0033-2909.111.3.455] [Medline: 1594721]
35. Ren J, Li M, Luo Y, et al. Sunset Without AIDS: protocol for a randomized controlled trial of a brief video-based intervention to improve the ability of AIDS prevention in elderly men. Trials 2022 Feb 14;23(1):146. [doi: 10.1186/s13063-022-06069-3] [Medline: 35164846]
36. O'Keefe DJ, Jensen JD. The relative persuasiveness of gain-framed and loss-framed messages for encouraging disease prevention behaviors: a meta-analytic review. J Health Commun 2007;12(7):623-644. [doi: 10.1080/10810730701615198] [Medline: 17934940]
37. Zelaya CE, Sivaram S, Johnson SC, Srikrishnan AK, Solomon S, Celentano DD. HIV/AIDS stigma: reliability and validity of a new measurement instrument in Chennai, India. AIDS Behav 2008 Sep;12(5):781-788. [doi: 10.1007/s10461-007-9331-7] [Medline: 18030613]
38. Xing HY, Wen DH, Chen SM, et al. Reliability and validity of Chinese version Zelaya's HIV/AIDS stigma scale. Chin J nurs 2012;47(12):1098. [doi: 10.3761/j.issn.0254-1769.2012.12.016]
39. Tang J, Zheng Y, Zhang D, et al. Evaluation of an AIDS educational mobile game (AIDS Fighter · Health Defense) for young students to improve AIDS-related knowledge, stigma, and attitude linked to high-risk behaviors in China: randomized controlled trial. JMIR Serious Games 2022 Jan 24;10(1):e32400. [doi: 10.2196/32400] [Medline: 34870603]
40. Wang S, Huang Q, Jiang L. The status quo and influencing factors of AIDS-related knowledge and discrimination among community elderly in Sichuan Province. J Occup Health 2021;37(10):1360-1362. [doi: 10.13329/j.cnki.zyyjk.2021.0338]
41. Zhang XG, Yin P, Lu GH, et al. Investigation on knowledge and attitude of community residents about HIV/AIDS in Yunnan Province. Appl Prev Med 2015;21(4):1673-1758. [doi: 10.3969/j.issn.1673-758X.2015.04.006]
42. Peng XB, Tan Y, Ye M, et al. Analysis of AIDS surveillance results of 60 year old male in Kaixian city, Chongqing in 2015. Chin J AIDS STD 2017;23(7):669-670. [doi: 10.13419/j.cnki.aids.2017.07.26]
43. Liu QX, Yang Y, Li Y, et al. Analysis of HIV/AIDS discrimination and its influencing factors in rural middle-aged and elderly men based on random forest algorithm. Mod Prev Med 2023;50(2):354-384. [doi: 10.20043/j.cnki.MPM.202206159]
44. Wu ZY. Advances in AIDS prevention technology and prevention strategies. Chin J Prev Med 2018;52(12):1204-1209. [doi: 10.3760/cma.j.issn.0253-9624.2018.12.002]
45. Liang ZJ, Wei T, Zhou ZR. Effects of micro-film model on AIDS health education for freshmen in a university. Chin J School Health 2017;38(6):828-831. [doi: 10.16835/j.cnki.1000-9817.2017.06.010]
46. Social organizations participate in the prevention of AIDS awareness action and exploration. China social organization. China Foundation for Prevention of Sexually Transmitted Diseases and AIDS. 2018. URL: https://kns.cnki.net/kcms2/article/
[accessed 2025-02-28]
47. Han X. Evaluation on effect of peer education intervention on AIDS prevention in a secondary vocational school.: Zhengzhou University; 2018. URL: https://kns.cnki.net/kcms2/article/
[accessed 2025-02-28]
48. Wang DH. Study on Intervention of AIDS High-Risk Behaviors Among Rural Elderly in Guangxi Based on IMB Model: Guilin Medical College; 2023. [doi: 10.27806/d.cnki.gglyx.2023.000218]
49. Liu DD, Li HL, Lu XC, et al. AIDS high-risk sexual behaviors and its influencing factors among 50 years old men in Hangzhou city. Chin J AIDS STD 2019;25(5):459-463. [doi: 10.13419/j.cnki.aids.2019.05.06]

Abbreviations

CDC: Center for Disease Control and Prevention
IMB: Information-Motivation-Behavioral
PrEP: preexposure prophylaxis

XSL•FO
RenderX

Edited by A Mavragani; submitted 22.08.24; peer-reviewed by J Ren, K Bond; revised version received 12.02.25; accepted 13.02.25; published 18.04.25.

Please cite as:
Huaju T, Rendie X, Lu X, Mei L, Yue L, Daiying Z, Yanhua C, Jianlan R
The Effect of a Brief Video-Based Intervention to Improve AIDS Prevention in Older Men: Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e65674
URL: https://publichealth.jmir.org/2025/1/e65674
doi:10.2196/65674

© Tian Huaju, Xie Rendie, Xiao Lu, Li Mei, Luo Yue, Zhang Daiying, Chen Yanhua, Ren Jianlan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Awareness of Sexual Partner's HIV Status Among Men Who Have Sex With Men in China: Cross-Sec. tional Survey Study

Qian Ma[1], MA; Tingting Jiang[2], MA; Wanjun Chen[2], MA; Shaoqiang Jiang[3], BA; Jinlei Zheng[2], MA; Hui Wang[2], MA; Lin He[2], MA

[1]Department of Rheumatology Immunology & Allergy, Zhejiang University School of Medicine, Hangzhou, China

[2]Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China

[3]Coastal Service Center, Qiantang District, Hangzhou, Zhejiang, China

Corresponding Author:
Lin He, MA
Zhejiang Provincial Center for Disease Control and Prevention, Hangzhou, China

Abstract

Background: Men who have sex with men (MSM) constitute a significant proportion of individuals living with human immunodeficiency virus. Over the past few years, China has implemented various strategies aimed at increasing the rate of HIV testing and reducing HIV transmission among MSM. Among these, the disclosure of HIV serostatus is an effective prevention strategy.

Objective: This study aimed to assess HIV serostatus disclosure and identify factors associated with awareness of sexual partners' HIV status among MSM to provide a scientific basis for promoting HIV testing and reducing HIV transmission.

Methods: A cross-sectional study based on a large-scale web-based survey was conducted among MSM in Zhejiang province, China, between July and December 2023. MSM who were HIV-negative or had an unknown HIV status were recruited from the Sunshine Test, a web-based platform that uses location-based services to provide HIV prevention services. Participants were required to complete a questionnaire on demographic characteristics, sexual behavior, rush popper use, awareness of sexual partners' HIV status, and knowledge of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP). A multinomial regression model was used to identify the factors associated with awareness of sexual partners' HIV status.

Results: A total of 7629 MSM participated in the study, with 45.2% (n=3451) being aware, 35.4% (n=2701) being partially aware, and 19.4% (n=1477) being unaware of their sexual partner's HIV status. The multinomial logistic regression analysis revealed the following results. Compared to those who were unaware of their sexual partner's HIV status, participants who were students (adjusted odds ratio [aOR] 1.43, 95% CI 1.09 - 1.86), had a monthly income of more than US $1400 (aOR 1.36, 95% CI 1.03 - 1.80), had insertive anal sex (aOR 1.35, 95% CI 1.12 - 1.63), had only male sexual partners (aOR 1.53, 95% CI 1.28 - 1.82), had 1 sexual partner in the past 3 months (aOR 2.36, 95% CI 2.01 - 2.77), had used condoms for the past 3 months (aOR 1.72, 95% CI 1.33 - 2.22), had frequently used rush poppers in the past 3 months (aOR 2.27, 95% CI 1.81 - 2.86), were aware of HIV PrEP (aOR 2.04, 95% CI 1.68 - 2.48), were aware of HIV PEP (aOR 1.69, 95% CI 1.39 - 2.06), used mail reagent self-testing (aOR 1.19, 95% CI 1.04 - 1.36), and had previously undergone HIV testing (aOR 1.40, 95% CI 1.16 - 1.69) were associated with increased odds of being aware of their sexual partner's HIV status.

Conclusions: Overall, 45.2% of MSM who were HIV-negative or had an unknown status were aware of their sexual partners' HIV status in China. We suggest strengthening targeted interventions through web-based platforms and gay apps to promote the disclosure of HIV serostatus and reduce HIV transmission among MSM.

(JMIR Public Health Surveill 2025;11:e66139) doi:10.2196/66139

KEYWORDS

human immunodeficiency virus; HIV; serostatus disclosure; pre-exposure prophylaxis; PrEP; men who have sex with men; MSM; web-based survey; HIV awareness

Introduction

Men who have sex with men (MSM) constitute a significant proportion of individuals living with human immunodeficiency virus. The overall HIV prevalence among MSM in China has remained at 8% for the past several years [1]. The annual number of newly diagnosed HIV infections is >100,000, with 25.7% of new cases resulting from homosexual transmission in 2023 [2]. The Joint United Nations Program on HIV/AIDS (UNAIDS) proposed a 95-95-95 target by 2030 [3]. The target was to achieve 95% of patients infected with HIV diagnosed, 95% of diagnosed patients treated, and 95% of treated patients

XSL•FO
RenderX

achieving viral suppression. In China, the second and third 95% targets (95.1% and 97.3% in 2023, respectively) were achieved with the implementation of a free treatment policy and lifelong follow-up, but only 84.3% of patients have been diagnosed [2]. There are challenges to achieving the first 95% target in China. Scaling up HIV screening for MSM is an important means to achieving the first 95% target, but there are still major challenges. Data from 2023 showed that only 88% of MSM had ever been tested for HIV [4]. Various testing strategies, including regular testing, self-testing, web-based mail testing, and serostatus disclosure, have been promoted to expand HIV testing among MSM [5].

The serostatus disclosure strategy is an effective measure for reducing and controlling HIV transmission. Obermeyer et al [6] have defined HIV serostatus disclosure as "the process of revealing a person's HIV status, whether positive or negative." Serostatus disclosure requires a person to ask or know their partner's HIV status before having sex and to only have sex if they have the same HIV status or use protective measures. Self-disclosure is the most common form of disclosure, which requires a person to share information about his or her HIV status directly with another person. In recent years, serostatus disclosure strategies have been the focus of intervention efforts among people with a high risk of HIV infection and have played an important role in HIV transmission risk reduction. One study showed that disclosure was associated with a 45% reduction in HIV transmission [7]. Empirical studies have shown that disclosure of HIV serostatus by people living with HIV can reduce the risk of HIV transmission by 17.9% - 40.6% compared to no disclosure [8]. Disclosure to sexual partners could also increase the HIV testing rate and reduce transmission risk behaviors [9]. A benefit of disclosing HIV status is that it empowers sexual partners to make shared decisions about how to protect their health while engaging in an active sexual life [10]. Prevalence of HIV disclosure among MSM living with HIV ranges from 12% to 53% [11]. Accordingly, most HIV infections occur when there is a lack of knowledge of the HIV-positive status of the other partner, coupled with a low uptake of preventive measures and inconsistent condom use [12]. Common barriers to disclosure among MSM were a lower perceived risk of HIV infection, history of sexually transmitted infections, engagement in receptive sex, and having sex with casual partners [13-15]. Patients with HIV viral suppression have a lower disclosure rate [16], as it is now established that there is no risk of HIV transmission when the viral load is undetectable [17,18]. Other studies have shown that being unmarried, living in rural formal areas [15], being aged 50 years and over [19], lacking family support [20], poverty, having multiple sexual partners, and stigma [21,22] lead to a lower rate of HIV serostatus disclosure. In addition, substance use prior to sexual encounters is less likely in lower-risk locations compared to highest-risk locations (ie, bathhouse and public sex environment) [23].

In China, individuals infected with HIV are required by law to disclose an HIV infection to their spouses. Furthermore, more than 95% of patients' spouses know their HIV status, and more than 92% of spouses have been tested for HIV yearly [24]. Therefore, China's HIV serostatus disclosure strategy is predominantly promoted among high-risk populations susceptible to HIV infections, including MSM. This is achieved by implementing an HIV serostatus disclosure strategy with the aim of expanding HIV testing coverage, enabling the early detection of individuals infected with HIV, and subsequently reducing HIV transmission among MSM. However, there are barriers to implementing this strategy, including the prevalence of multiple sexual partners in the MSM population, fear of being rejected by sexual partners, and the personal privacy of one's infection status. In addition, the implementation of the "undetectable equals untransmittable" strategy means that many people living with HIV either do not disclose or do not ask about the HIV status of their sexual partners. Therefore, this could be a factor in a partner's decision to use a condom during sexual intercourse [10]. When implementing the serostatus disclosure strategy, MSM who are HIV-negative or unaware of their infection status are advised to ask their partners about their HIV infection status before sex, to self-test for HIV, and to take appropriate safety measures according to their HIV status. It has also been shown that HIV-negative MSM become more aware of their partners' serostatuses over time [25].

With the development of the digital economy, an increasing number of MSM are looking for sexual partners through the internet and dating apps. This study shows that the level of HIV knowledge among MSM can be improved and HIV testing can be expanded by relying on the internet. Consequently, web-based interventions and HIV testing among MSM populations, including web-based counseling and mailing of HIV self-test kits, are becoming increasingly common. A previous meta-analysis showed that using mobile health technology could increase the linkage to care among MSM using HIV self-testing [26].

Zhejiang province is economically developed, with an active digital economy. Approximately 66,000 HIV tests were performed among MSM in 2023, accounting for 10% of the national proportion [2], and over 70% of HIV testing among MSM in Zhejiang province relied on the internet. The promotion of pre-exposure prophylaxis (PrEP) and postexposure prophylaxis (PEP) prevention services in China began in 2020. To work toward achieving the target of 95% of HIV infections diagnosed globally and in China, a large-scale, web-based cross-sectional study was conducted. This study aimed to assess the awareness of a sexual partner's HIV status before engaging in sexual intercourse among MSM, especially for those who were not infected with HIV or did not know their infection status, which could be causing an information gap among MSM in China [13]. In the context of PrEP and PEP prevention services and the implementation of the "undetectable equals untransmittable" strategy, this study aimed to assess HIV serostatus disclosure and identify factors associated with awareness of sexual partners' HIV status among MSM to provide a scientific basis for promoting HIV testing and reducing HIV transmission.

XSL•FO
RenderX

Methods

Sampling

A cross-sectional survey was conducted between July and December 2023. The survey was based on a large web-based survey of MSM. To determine the required sample size, we used the following formula:

$$n=Z\alpha/2\ P(1-P)$$

The α level was .05, *P* was estimated as 20% - 30%, and δ was 0.05 resulting in a minimum sample size of 1329.

Study Participants

The following inclusion criteria were used to recruit MSM participants: (1) at least 16 years of age, (2) having had sex with men in the past year, (3) HIV-negative or unknown HIV status, and (4) living in Zhejiang province. Participants were excluded if they had previously received 2 or more HIV tests.

Participant Recruitment and Data Collection

Sunshine Coast Public Welfare is a social service organization that uses the internet to provide HIV outreach and testing services to the MSM population. Through the use of location-based services, Sunshine Coastal Public Welfare established a digital HIV prevention service. MSM in Zhejiang province could apply for free HIV-testing services by visiting the official account of the Sunshine Test on WeChat (a popular communication software in China). There were two ways for MSM to access HIV testing services: they could choose to have the test reagent sent by courier, or they could choose to self-test after receiving the reagent. Alternatively, they could choose offline services, where volunteers in their neighborhood provided HIV testing services, or they could visit a nearby voluntary counseling and testing site. MSM who applied for HIV testing services through the Sunshine Test were required to complete a questionnaire that covered demographic characteristics such as age, marital status, and education level, as well as sexual behavior characteristics such as sexual roles, number of sexual partners, use of rush poppers, sexual history, PrEP use, PEP use, and HIV testing history. All the questionnaires were completed using the Sunshine Test [4].

Statistical Analysis

For descriptive analyses, categorical variables were presented as frequencies and proportions, whereas continuous variables were presented as medians and IQRs or means and SDs. Differences in general demographic characteristics of the study population were compared using a χ^2 test. The main outcome was awareness (participants knowing all of their sexual partners' HIV statuses) and partial awareness (participants only knowing some of their partners' HIV statuses) of their sexual partners' HIV status, with those who were unaware of their sexual partner's HIV status being used as the reference group. Multinomial logistic regression was used to identify the factors associated with the awareness and partial awareness of a partners HIV status compared to the unaware group. Factors from the univariate analysis were included in the multinomial regression models. The α level was .05 and β was .1. All statistical analyses were performed using SPSS version 19.0 (IBM Corp).

Ethical Considerations

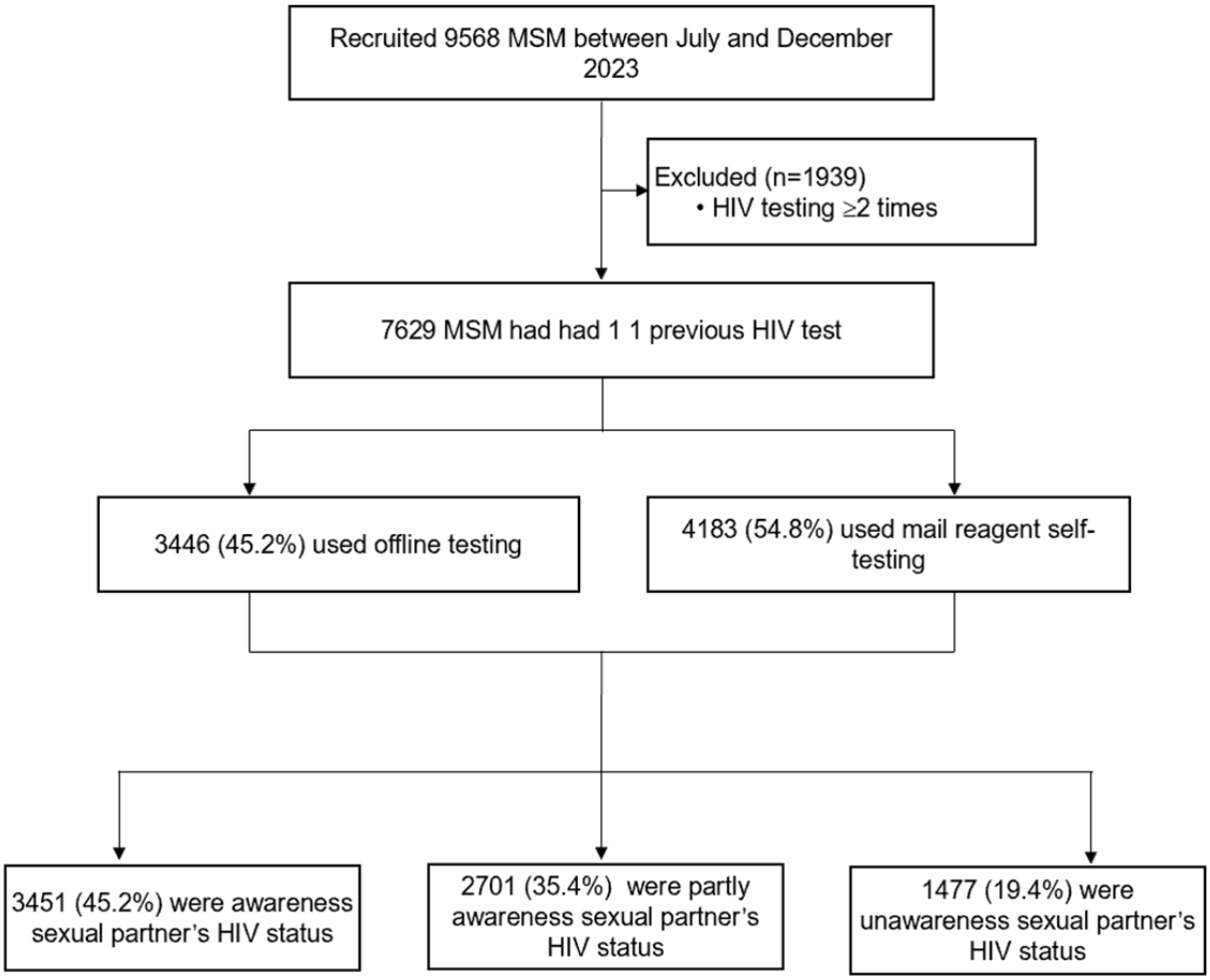
This study was approved by the Zhejiang Provincial Center for Disease Control and Prevention (approval number 2022-011-01). Written informed consent was obtained from all participants prior to survey completion. Participants received free HIV testing and health counseling, and all study procedures were conducted in accordance with approved guidelines and regulations. Participants could dop out during any stage of the survey. All the data was anonymized no amount of compensation was provided.

Results

A total of 7629 MSM participated in the study (Figure 1), and 64.9% (4956/7629) were aged less than 30 years. Among them, 50.7% (3869/7629) had a bachelor's degree or higher, 16.4% (1254/7629) were students, and 39.1% (2985/7629) had a monthly income of less than US $700. The majority (78.2%, 5964/7629) had only male sexual partners, while 33.3% (2543/7629) had more than 2 sexual partners in the past 3 months. Additionally, 31.2% (2380/7629) reported using rush poppers during sexual activity. The awareness rates of PrEP and PEP were 76.2% (5816/7629) and 78.1% (5960/7629), respectively. Furthermore, 54.8% (4183/7629) had HIV testing using a mail reagent self-test, and 12.9% (982/7629) were undergoing HIV testing for the first time (Table 1).

XSL•FO
RenderX

Figure 1. Study design.



XSL•FO
RenderX

Table . Social demographic and behavioral characteristics among men who have sex with men.

Variables	Total (n=7629), n (%)	Awareness of sexual partner's HIV status			Chi-square (*df*)	*P* value
		Yes (n=3451)	Partially (n=2701)	No (n=1477)		
Age[a] (years), n (%)					33.458 (4)	<.001
<20	430 (5.6)	185 (43)	136 (31.6)	109 (25.3)		
20 - 29	4526 (59.3)	2143 (47.3)	1528 (33.8)	855 (18.9)		
≥30	2673 (35)	1123 (42)	1037 (38.8)	513 (19.2)		
Marital status, n (%)					35.719 (4)	<.001
Single	5997 (78.6)	2804 (46.8)	2032 (33.9)	1161 (19.4)		
Married	1388 (18.2)	558 (40.2)	571 (41.1)	259 (18.7)		
Divorced/separated	244 (3.2)	89 (36.5)	98 (40.2)	57 (23.4)		
Education, n (%)					77.958 (6)	<.001
High school and below	1851 (24.3)	683 (36.9)	736 (39.8)	432 (23.3)		
College	1909 (25)	898 (47)	639 (33.5)	372 (19.5)		
Bachelor's degree	3145 (41.2)	1511 (48)	1099 (34.9)	535 (17)		
Master's degree or above	724 (9.5)	359 (49.6)	227 (31.4)	138 (19.1)		
Occupation, n (%)					54.237 (6)	<.001
Student	1254 (16.4)	608 (48.5)	375 (29.9)	271 (21.6)		
Company employee	2843 (37.3)	1370 (48.2)	1009 (35.5)	464 (16.3)		
Freelance	1022 (13.4)	426 (41.7)	383 (37.5)	213 (20.8)		
Others	2510 (32.9)	1047 (41.7)	934 (37.2)	529 (21.1)		
Monthly income (US$), n (%)					38.718 (6)	<.001
<$350	1320 (17.3)	577 (43.7)	419 (31.7)	324 (24.5)		
$350-$699	1665 (21.8)	721 (43.3)	633 (38)	311 (18.7)		
$700-$1399	3259 (42.7)	1477 (45.3)	1178 (36.1)	604 (18.5)		
≥$1400	1385 (18.2)	676 (48.8)	471 (34)	238 (17.2)		
Sex roles, n (%)					131.721 (4)	<.001
Receptive anal sex	1750 (22.9)	796 (45.5)	675 (38.6)	279 (15.9)		
Insertive anal sex	2652 (34.8)	1319 (49.7)	949 (35.8)	384 (14.5)		
Both	3227 (42.3)	1336 (41.4)	1077 (33.4)	814 (25.2)		
Sex of partner, n (%)					284.131 (2)	<.001
Male only	5964 (78.2)	2852 (47.8)	2197 (36.8)	915 (15.3)		
Male and female	1665 (21.8)	599 (36)	504 (30.3)	562 (33.8)		
Number of sexual partners in the past 3 months, n (%)					748.589 (4)	<.001
0	1277 (16.7)	475 (37.2)	296 (23.2)	506 (39.6)		
1	3809 (49.9)	2124 (55.8)	1145 (30.1)	540 (14.2)		

XSL•FO
RenderX

Variables	Total (n=7629), n (%)	Awareness of sexual partner's HIV status			Chi-square (*df*)	*P* value
		Yes (n=3451)	Partially (n=2701)	No (n=1477)		
≥2	2543 (33.3)	852 (33.5)	1260 (49.5)	431 (16.9)		
Condom use in the past 3 months, n (%)					681.686 (6)	<.001
Every time	4266 (55.9)	2171 (50.9)	1445 (33.9)	650 (15.2)		
Sometimes	2005 (26.3)	751 (37.5)	958 (47.8)	296 (14.8)		
No sex	874 (11.5)	288 (33)	168 (19.2)	418 (47.8)		
Never	484 (6.3)	241 (49.8)	130 (26.9)	113 (23.3)		
Ever used rush poppers during sexual behavior, n (%)					169.468 (4)	<.001
Never	5249 (68.8)	2355 (44.9)	1704 (32.5)	1190 (22.7)		
Occasionally	1312 (17.2)	536 (40.9)	601 (45.8)	175 (13.3)		
Often	1068 (14)	560 (52.4)	396 (37.1)	112 (10.5)		
History of STIs[b], n (%)					7.553 (2)	.02
No	7428 (97.4)	3376 (45.4)	2612 (35.2)	1440 (19.4)		
Yes	201 (2.6)	75 (37.3)	89 (44.3)	37 (18.4)		
Awareness of HIV PrEP[c], n (%)					336.193 (2)	<.001
No	1813 (23.8)	551 (30.4)	669 (36.9)	593 (32.7)		
Yes	5816 (76.2)	2900 (49.9)	2032 (34.9)	884 (15.2)		
Knowledge of HIV PEP[d], n (%)					313.701 (2)	<.001
No	1669 (21.9)	528 (31.6)	579 (34.7)	562 (33.7)		
Yes	5960 (78.1)	2923 (49)	2122 (35.6)	915 (15.4)		
HIV testing pathway, n (%)					26.881 (2)	<.001
Offline testing	3446 (45.2)	1526 (44.3)	1165 (33.8)	755 (21.9)		
Mail reagent self-testing	4183 (54.8)	1925 (46)	1536 (36.7)	722 (17.3)		
HIV status, n (%)					7.114 (2)	.03
Positive	82 (1.1)	27 (32.9)	31 (37.8)	24 (29.3)		
Negative	7547 (98.9)	3424 (45.4)	2670 (35.4)	1453 (19.3)		
First time undergoing HIV testing, n (%)					59.378 (2)	<.001
No	6647 (87.1)	3024 (45.5)	2421 (36.4)	1202 (18.1)		
Yes	982 (12.9)	427 (43.5)	280 (28.5)	275 (28)		

[a]The median age was 27 (IQR 23-33).

[b]STI: sexually transmitted infection.

[c]PrEP: pre-exposure prophylaxis.

[d]PEP: postexposure prophylaxis.

Of the 7629 participants, 45.2% (3451/7629) were aware, 35.4% (2701/7629) were partially aware, and 19.4% (1477/7629) were unaware of their sexual partner's HIV status. Factors associated with high awareness of their sexual partner's HIV status

XSL•FO
RenderX

included being 20 - 29 years old (47.3%, 2143/4526), being single (46.8%, 2804/5997), having a high education level (49.6%, 359/724), having a high monthly income (48.8%, 676/1385), performing insertive anal sex (49.7%, 1319/2652), having one sexual partner in the past 3 months (55.8%, 2124/3809), using condoms in the past 3 months (50.9%, 2171/4266), frequently using rush poppers (52.4%, 560/1068), being aware of HIV PrEP (49.9%, 2900/5816) and PEP (49.0%, 2923/5960), and being HIV-negative (45.4%, 3424/7547). The χ^2 test showed that all factors were associated with awareness of sexual partners' HIV status (Table 1).

As presented in Table 2, multinomial logistic regression analysis showed that, compared to those who were unaware of their sexual partner's HIV status, participants who were students (adjusted odds ratio [aOR] 1.43, 95% CI 1.09 - 1.86), had a monthly income of more than US $1400 (aOR 1.36, 95% CI 1.03 - 1.80), had insertive anal sex (aOR 1.35, 95% CI 1.12 - 1.63), had only male sexual partners (aOR 1.53, 95% CI 1.28 - 1.82), had 1 sexual partner in the past 3 months (aOR 2.36, 95% CI 2.01 - 2.77), used condoms in the past 3 months (aOR 1.72, 95% CI 1.33 - 2.22), frequently used rush poppers in the past 3 months (aOR 2.27, 95% CI 1.81 - 2.86), were aware of HIV PrEP (aOR 2.04, 95% CI 1.68 - 2.48), were aware of HIV PEP (aOR 1.69, 95% CI 1.39 - 2.06), used mail reagent self-testing (aOR 1.19, 95% CI 1.04 - 1.36), and were undergoing HIV testing for the first time (aOR 1.40, 95% CI 1.16 - 1.69) were associated with increased odds of awareness of their sexual partner's HIV status.

XSL•FO
RenderX

Table . Factors associated with awareness of sexual partner's HIV status among men who have sex with men.

Variables	Partial awareness		Awareness	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value
Occupation				
Student	1.05 (0.80 - 1.39)	.71	1.43 (1.09 - 1.86)	.009
Company employee	1.14 (0.96 - 1.36)	.14	1.17 (0.99 - 1.38)	.07
Freelance	1.01 (0.82 - 1.25)	.93	1.01 (0.82 - 1.25)	.91
Others	1.00 (reference)	—[b]	1.00 (reference)	—
Monthly income (US$)				
<$350	1.00 (reference)	—	1.00 (reference)	—
$350-$699	1.30 (1.01 - 1.67)	.045	1.36 (1.03 - 1.80)	.03
$700-$1399	1.11 (0.86 - 1.41)	.43	1.36 (1.07 - 1.73)	.01
≥$1400	0.97 (0.73 - 1.30)	.85	1.36 (1.03 - 1.80)	.03
Sex roles				
Receptive anal sex	1.00 (reference)	—	1.00 (reference)	—
Insertive anal sex	1.08 (0.89 - 1.31)	.35	1.35 (1.12 - 1.63)	.002
Both	0.92 (0.76 - 1.10)	.45	1.02 (0.85 - 1.23)	.82
Sex of partner				
Male only	1.39 (1.16 - 1.66)	<.001	1.53 (1.28 - 1.82)	<.001
Male and female	1.00 (reference)	—	1.00 (reference)	—
Number of sexual partners in the past 3 months				
0	0.84 (0.61 - 1.14)	.27	1.73 (1.27 - 2.32)	<.001
1	0.91 (0.78 - 1.07)	.25	2.36 (2.01 - 2.77)	<.001
≥2	1.00 (reference)	—	1.00 (reference)	—
Condom use in the past 3 moths				
Every time	1.91 (1.45 - 2.52)	<.001	1.72 (1.33 - 2.22)	<.001
Sometimes	2.57 (1.92 - 3.45)	<.001	1.42 (1.08 - 1.88)	.01
No sex	0.64 (0.42 - 0.97)	.04	0.65 (0.44 - 0.95)	.03
Never	1.00 (reference)	—	1.00 (reference)	—
Ever used rush poppers during sex behavior				
Never	1.00 (reference)	—	1.00 (reference)	—
Occasionally	1.63 (1.34 - 1.98)	<.001	1.35 (1.10 - 1.64)	.003
Often	1.77 (1.40 - 2.24)	<.001	2.27 (1.81 - 2.86)	<.001
Awareness of HIV PrEP[c]				
No	1.00 (reference)	—	1.00 (reference)	—
Yes	1.22 (1.00 - 1.49)	.048	2.04 (1.68 - 2.48)	<.001
Awareness of HIV PEP[d]				
No	1.00 (reference)	—	1.00 (reference)	—
Yes	1.56 (1.28 - 1.91)	<.001	1.69 (1.39 - 2.06)	<.001
HIV testing pathway				
Offline testing	1.00 (reference)	—	1.00 (reference)	—

XSL•FO
RenderX

Variables	Partial awareness		Awareness	
	aOR[a] (95% CI)	*P* value	aOR (95% CI)	*P* value
Mail reagent self-testing	1.20 (1.04 - 1.38)	.01	1.19 (1.04 - 1.36)	.01
First time undergoing HIV testing				
No	1.59 (1.30 - 1.93)	<.001	1.40 (1.16 - 1.69)	<.001
Yes	1.00 (reference)	—	1.00 (reference)	—

[a]aOR: adjusted odds ratio.

[b]Statistical analysis was not done for the reference group.

[c]PrEP: pre-exposure prophylaxis.

[d]PEP: postexposure prophylaxis.

Multinomial logistic regression analysis showed that, compared to those who were unaware of their sexual partner's HIV status, participants who had a monthly income of US $350-$699 (aOR 1.30, 95% CI 1.01 - 1.67), only had male sexual partners (aOR 1.39, 95% CI 1.16 - 1.66), used condoms in the past 3 months (aOR 1.91, 95% CI 1.45 - 2.52), frequently used rush poppers in the past 3 months (aOR 1.77, 95% CI 1.40 - 2.24), were aware of HIV PrEP (aOR 1.22, 95% CI 1.00 - 1.49), were aware of HIV PEP (aOR 1.56, 95% CI 1.28 - 1.91), used mail reagent self-testing (aOR 1.20, 95% CI 1.04 - 1.38), and were undergoing HIV testing for the first time (aOR 1.59, 95% CI 1.30 - 1.93) were associated with increased odds of partial awareness of their sexual partner's HIV status (Table 2).

Discussion

Principal Findings

This study described the awareness of sexual partners' HIV status and its associated factors among MSM with an HIV-negative or unknown status in Zhejiang province, China. Based on data from 7629 participants, we found 45.2% (n=3451) of participants were aware of their sexual partner's HIV status and 19.4% (1477) were unaware of their partner's HIV status. Our results show higher percentages compared with studies from Guangzhou, China, which revealed that 36% of MSM were aware of the HIV serostatus of all their sexual partners [27], 37.8% of MSM had received their partner's HIV testing report before engaging in sex [28], and 30.4% of MSM asked all partners about their HIV status [13]. A previous study from the United States found that 55% of participants consistently disclosed their HIV status to their sexual partners, 30% inconsistently disclosed, and 15% did not disclose [29], and these rates were lower than the disclosure rates of 61.2% in Korea [30] and 64.3% in Thailand [31]. The HIV serostatus disclosure strategy has proven effective in preventing and reducing HIV transmission. HIV prevalence among the MSM population has remained at 8% and only 84.3% of HIV-positive people have been diagnosed. We recommend the continued promotion and implementation of serostatus disclosure strategies among MSM, especially proactively asking about their sexual partners' HIV status, taking appropriate protective measures based on HIV status, reducing HIV transmission, and undergoing early diagnosis testing for HIV.

Consistent with previous studies [28,32], we found that MSM who had undergone HIV testing were more likely to obtain their sexual partners' HIV status. MSM who engage in HIV testing and seek their partners' HIV status may reflect a heightened awareness and proactive approach toward HIV management. Individuals who have already undergone HIV testing may be better informed about the importance of knowing their partners' HIV status, thereby reducing HIV transmission risks and promoting safer sexual practices. We found that MSM who used mail reagent self-testing had higher awareness of their partner's HIV status. This suggests that HIV testing through web-based platforms or gay apps could increase HIV serostatus disclosure among MSM [33]. These findings suggest that interventions aimed at increasing HIV testing rates among MSM could indirectly enhance partner disclosure and engagement in HIV-prevention strategies. It supports the need for continued efforts to promote regular HIV testing and web-based self-testing to increase the HIV serostatus disclosure rate. Our study provides evidence that understanding and engaging in HIV testing behaviors may contribute to HIV serostatus disclosure among MSM.

MSM who were aware of HIV PrEP and PEP were more likely to obtain their sexual partners' HIV status. A previous study showed that awareness of and access to PrEP might reduce the need and motivation to discuss HIV status with partners because they may assume that it is the partner's responsibility to use PrEP to protect themselves from HIV infection [13]. PrEP users expressed significantly greater openness to serodifferent partnering than participants who had never used PrEP [34], and MSM were more likely to have sex with a person who disclosed being on PrEP [35]. Of those who had not disclosed their status to their partners, 71% stated PrEP would encourage them to have sex [36]. However, another study showed that PrEP awareness was not associated with HIV serostatus disclosure among Chinese MSM [13]. Knowledge of PrEP and PEP may drive more thorough discussions on HIV status, aligning with a proactive approach to reducing transmission risks. Against the background of a 78% awareness of PrEP and PEP in China, we recommend targeted interventions to increase awareness of PrEP and PEP, thereby increasing the rate of partners' HIV testing and disclosure.

Consistent with a previous study [37], we found that MSM who used condoms and had fewer sexual partners were more likely

XSL•FO
RenderX

to obtain the HIV status from their partners. MSM who adhere to consistent condom use and have fewer sexual partners often exhibit a heightened awareness of HIV risks and a commitment to preventing transmission. Findings showed that HIV status disclosure was significantly associated with lower rates of condomless sex compared to nondisclosure [33,38,39]. The probability of HIV exposure through condomless anal intercourse was substantially lower after serostatus disclosure than after nondisclosure [22]. One study showed that MSM engaging in receptive sex were less likely to ask their partners about their HIV status [13]. MSM with an HIV-negative or unknown status have a lower proportion of sexual events with a partner with an unknown HIV status [25]. Studies have demonstrated that MSM who disclose their HIV status are more likely to practice safer sexual behavior with HIV-infected partners than with partners with an HIV-negative or unknown status [23]. These findings highlight the importance of promoting consistent condom use and supporting MSM in maintaining low partner numbers to enhance the disclosure of HIV serostatus.

Moreover, we found that MSM who often used rush poppers had high awareness of their partners' HIV status. Similar to other studies, drug use was independently associated with disclosure [40]. The possible reasons were as follows. First, MSM who used rush poppers had a higher rates of previous HIV testing, which was associated with greater awareness of their partner's HIV status [4]. Second, due to the high use of rush poppers among MSM, rush popper use increases the risk of unprotected sexual behavior [39] and HIV transmission [41]. Targeted interventions were conducted for rush users among MSM [42], including reducing rush popper use, condom promotion, HIV serostatus disclosure, and other measures, such as knowing the HIV status of sexual partners before sex. We suggest that targeted interventions should be continued among MSM who use rush poppers, and this approach could help further improve the proportion of HIV serostatus disclosure.

Limitations

This study has some limitations. First, the study used routine surveillance questionnaires, which may not have captured all factors associated with HIV serostatus disclosure in previous studies. For example, the questionnaire only investigated the awareness of PrEP but did not investigate whether they had used it, and previous studies showed that the use of PrEP could also improve the rates of HIV serostatus disclosure [35]. Second, the study focused on participants actively asking about their sexual partners' HIV status and did not investigate whether the participants disclosed their HIV status to their sexual partners. Finally, the study recruited MSM from web-based applications, 64.9% of whom were under 30 years old. It did not include a few MSM who were over 50 years of age. Previous studies have shown that this population has lower HIV serostatus disclosure [19]. Therefore, the results of this study apply only to MSM who rely on the internet.

Conclusion

The rate of awareness of sexual partners' HIV status is low among MSM with an HIV-negative or unknown status in China. We suggest strengthening targeted interventions through web-based platforms and dating apps to promote HIV serostatus disclosure among MSM. Future studies could explore strategies to actively ask for a sexual partner's HIV status among underrepresented demographics, such as older MSM and those who used PrEP, ultimately helping to improve the HIV disclosure rate and reduce HIV transmission.

Acknowledgments

We express our gratitude to the participants and Sunshine Coast Public Welfare for their contributions to the study. The study was funded by a grant from the Health Science and Technology Project in Zhejiang Province Social (2023KY632 and 2021RC048).

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LH conceived the study design, coordinated the conduct of this research in the field, and performed the statistical analysis. QM drafted the paper. TJ, WC, SJ, JZ, and HW designed the research and played a major role in the survey. All the authors read and approved the final paper.

Conflicts of Interest

None declared.

References

1. Dong MJ, Peng B, Liu ZF, et al. The prevalence of HIV among MSM in China: a large-scale systematic analysis. BMC Infect Dis 2019 Nov 27;19(1):1000. [doi: 10.1186/s12879-019-4559-1] [Medline: 31775654]
2. National center for AIDS STD control and prevention CCFD: HIV prevention and control progress report. : China Center for Disease Control and Prevention; 2023 URL: https://www.chinaaids.cn/ [accessed 2025-01-06]

XSL•FO
RenderX

3. Understanding Fast-Track. UNAIDS. 2015 Jul 7. URL: https://www.unaids.org/en/resources/documents/2015/201506_JC2743_Understanding_FastTrack [accessed 2025-01-06]
4. He L, Jiang T, Chen W, et al. Examining HIV testing coverage and factors influencing first-time testing among men who have sex with men in Zhejiang province, China: cross-sectional study based on a large internet survey. JMIR Public Health Surveill 2024 Jun 14;10:e56906. [doi: 10.2196/56906] [Medline: 38875001]
5. Chen S, Zhang Q, Chan CK, et al. Evaluating an innovative HIV self-testing service with web-based, real-time counseling provided by an artificial intelligence chatbot (HIVST-Chatbot) in increasing HIV self-testing use among chinese men who have sex with men: protocol for a noninferiority randomized controlled trial. JMIR Res Protoc 2023 Jun 30;12:e48447. [doi: 10.2196/48447] [Medline: 37389935]
6. Obermeyer CM, Baijal P, Pegurri E. Facilitating HIV disclosure across diverse settings: a review. Am J Public Health 2011 Jun;101(6):1011-1023. [doi: 10.2105/AJPH.2010.300102] [Medline: 21493947]
7. O'Connell AA, Reed SJ, Serovich JA. The efficacy of serostatus disclosure for HIV transmission risk reduction. AIDS Behav 2015 Feb;19(2):283-290. [doi: 10.1007/s10461-014-0848-2] [Medline: 25164375]
8. Pinkerton SD, Galletly CL. Reducing HIV transmission risk by increasing serostatus disclosure: a mathematical modeling analysis. AIDS Behav 2007 Sep;11(5):698-705. [doi: 10.1007/s10461-006-9187-2] [Medline: 17082982]
9. Simoni JM, Pantalone DW. Secrets and safety in the age of AIDS: does HIV disclosure lead to safer sex? Top HIV Med 2004;12(4):109-118. [Medline: 15516708]
10. Irungu E, Chersich MF, Sanon C, et al. Changes in sexual behaviour among HIV-infected women in West and East Africa in the first 24 months after delivery. AIDS 2012 May 15;26(8):997-1007. [doi: 10.1097/QAD.0b013e3283524ca1] [Medline: 22343965]
11. Wei C, Lim SH, Guadamuz TE, Koe S. HIV disclosure and sexual transmission behaviors among an Internet sample of HIV-positive men who have sex with men in Asia: implications for prevention with positives. AIDS Behav 2012 Oct;16(7):1970-1978. [doi: 10.1007/s10461-011-0105-x] [Medline: 22198313]
12. Eyawo O, de Walque D, Ford N, Gakii G, Lester RT, Mills EJ. HIV status in discordant couples in sub-Saharan Africa: a systematic review and meta-analysis. Lancet Infect Dis 2010 Nov;10(11):770-777. [doi: 10.1016/S1473-3099(10)70189-4] [Medline: 20926347]
13. Cao W, Li J, Sun S, et al. HIV serostatus disclosure among men who have sex with men in China in the era of U=U and PrEP. AIDS Behav 2022 May;26(5):1477-1488. [doi: 10.1007/s10461-021-03502-7] [Medline: 34697704]
14. Murphy DA, de Wit JBF, Donohoe S, Adam PCG. The need to know: HIV status disclosure expectations and practices among non-HIV-positive gay and bisexual men in Australia. AIDS Care 2015;27(Sup1):90-98. [doi: 10.1080/09540121.2015.1062077] [Medline: 26616130]
15. Simbayi LC, Zungu N, Evans M, et al. HIV serostatus disclosure to sexual partners among sexually active people living with HIV in South Africa: results from the 2012 national population-based household survey. AIDS Behav 2017 Jan;21(1):82-92. [doi: 10.1007/s10461-015-1278-5] [Medline: 26767538]
16. Daskalopoulou M, Rodger AJ, Phillips AN, et al. Attitudes to disclosure of HIV-serostatus to new sexual partners and sexual behaviours among HIV-diagnosed gay, bisexual and other men who have sex with men in the UK. AIDS Care 2020 Oct;32(10):1323-1332. [doi: 10.1080/09540121.2020.1728218] [Medline: 32114800]
17. Eisinger RW, Dieffenbach CW, Fauci AS. HIV viral load and transmissibility of HIV infection: undetectable equals untransmittable. JAMA 2019 Feb 5;321(5):451-452. [doi: 10.1001/jama.2018.21167] [Medline: 30629090]
18. Rodger AJ, Cambiano V, Bruun T, et al. Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive PARTNER taking suppressive antiretroviral therapy (PARTNER): final results of a multicentre, prospective, observational study. Lancet 2019 Jun 15;393(10189):2428-2438. [doi: 10.1016/S0140-6736(19)30418-0] [Medline: 31056293]
19. Lian J, Su XY, Chen XY, et al. Receiving human immunodeficiency virus serostatus disclosure from male sexual partners and related factors among men who have sex with men aged 50 and above. Zhong Yi Xue Ke Xue Yuan Xue Bao 2022 Apr;44(2):221-226. [doi: 10.3881/j.issn.1000-503X.14171] [Medline: 35538756]
20. Mutambara J, Vandirayi S, January J, Zirima H. Factors that influence HIV status disclosure among the elderly: perspectives from Shurugwi, Zimbabwe. AIDS Care 2022 Jan;34(1):112-117. [doi: 10.1080/09540121.2021.1930997] [Medline: 34039191]
21. Smith R, Rossetto K, Peterson BL. A meta-analysis of disclosure of one's HIV-positive status, stigma and social support. AIDS Care 2008 Nov;20(10):1266-1275. [doi: 10.1080/09540120801926977] [Medline: 18608080]
22. Kavuma D, Kirwana VB, Taani M. Factors associated with HIV positive serostatus disclosure to sexual partners among sexually active young people on anti-retroviral therapy in Central Uganda. HIV AIDS (Auckl) 2023;15:293-311. [doi: 10.2147/HIV.S407535] [Medline: 37312814]
23. Serovich JM, Reed SJ, O'Connell AA, Laschober TC. Relationship between serostatus disclosure and categories of HIV transmission risk in men who have sex with men living with HIV. Int J STD AIDS 2018 Jul;29(8):744-750. [doi: 10.1177/0956462417752267] [Medline: 29431024]
24. Han M. Analysis of the epidemic situation of AIDS in China and prospects for its prevention and control. Chin J AIDS STD 2023;29(3):247-250. [doi: 10.13419/j.cnki.aids.2023.03.01]

XSL•FO
RenderX

25. Sang JM, Wang L, Moore DM, et al. Undetectable or unknown? A longitudinal event-level analysis of disclosure of HIV serostatus and undetectability among gay, bisexual, and other men who have sex with men (gbMSM) in metro Vancouver. AIDS Behav 2021 Aug;25(8):2630-2643. [doi: 10.1007/s10461-021-03224-w] [Medline: 33751314]
26. Chen S, Fang Y, Chan PSF, Kawuki J, Mo P, Wang Z. Counseling supporting HIV self-testing and linkage to care among men who have sex with men: systematic review and meta-analysis. JMIR Public Health Surveill 2024 Jan 24;10:e45647. [doi: 10.2196/45647] [Medline: 38265866]
27. Zeng W, Luo Y, Liu F, et al. Impact of awareness of sexual partners' HIV serostatus on the HIV acquisition among men who have sex with men in Guangzhou, China: a nested case-control study. Arch Sex Behav 2024 Oct;53(9):3655-3662. [doi: 10.1007/s10508-024-02971-w] [Medline: 39152320]
28. Sun HT, Fan XR, Gu YZ, et al. WeChat-based HIV e-report, a new approach for HIV serostatus requests and disclosures among men who have sex with men: prospective subgroup analysis of a randomized controlled trial. JMIR Mhealth Uhealth 2023 May 8;11:e44513. [doi: 10.2196/44513] [Medline: 37155223]
29. Cook CL, Staras SAS, Zhou Z, Chichetto N, Cook RL. Disclosure of HIV serostatus and condomless sex among men living with HIV/AIDS in Florida. PLoS ONE 2018;13(12):e0207838. [doi: 10.1371/journal.pone.0207838] [Medline: 30557362]
30. Kim K, Woo J. Predictors associated with HIV status non-disclosure in Korea. J Korean Med Sci 2022 Apr 25;37(16):e125. [doi: 10.3346/jkms.2022.37.e125] [Medline: 35470599]
31. Morkphrom E, Ratanasuwan W, Sittironnarit G, Rattanaumpawan P. Non-disclosure of HIV serostatus to sexual partners: prevalence, risk factors and clinical impact in patients with HIV. HIV Med 2021 Mar;22(3):194-200. [doi: 10.1111/hiv.13005] [Medline: 33145968]
32. Tang W, Liu C, Cao B, et al. Receiving HIV serostatus disclosure from partners before sex: results from an online survey of Chinese men who have sex with men. AIDS Behav 2018 Dec;22(12):3826-3835. [doi: 10.1007/s10461-018-2062-0] [Medline: 29470809]
33. Konkor I, Antabe R, Mkandawire P, et al. Knowledge of sexual partner's HIV serostatus and the practice of safer sex among heterosexual men of African descent in London, Ontario. Ethn Health 2022 Feb;27(2):375-387. [doi: 10.1080/13557858.2020.1734776] [Medline: 32115995]
34. Kalwicz DA, Rao S, Modrakovic DX, et al. The implications of PrEP use, condom use, and partner viral load status for openness to serodifferent partnering among US sexual minority men (SMM). AIDS Behav 2024 Feb;28(2):524-534. [doi: 10.1007/s10461-023-04241-7] [Medline: 38329558]
35. Medina MM, Crowley C, Montgomery MC, et al. Disclosure of HIV serostatus and pre-exposure prophylaxis use on internet hookup sites among men who have sex with men. AIDS Behav 2019 Jul;23(7):1681-1688. [doi: 10.1007/s10461-018-2286-z] [Medline: 30267365]
36. Bracho-Sanchez E, Stonbraker S, Halpern M, Wood S, Lowenthal E. HIV pre-exposure prophylaxis: assessment of acceptability among members of HIV serodiscordant/serostatus unknown couples. Int J STD AIDS 2022 Jun;33(7):687-693. [doi: 10.1177/09564624221097753] [Medline: 35488451]
37. Sullivan MC, Cruess DG, Huedo-Medina TB, Kalichman SC. Substance use, HIV serostatus disclosure, and sexual risk behavior in people living with HIV: an event-level analysis. Arch Sex Behav 2020 Aug;49(6):2005-2018. [doi: 10.1007/s10508-019-01531-x] [Medline: 31863314]
38. Okafor CN, Li MJ, Hucks-Ortiz C, Mayer KH, Shoptaw S. Disclosure of HIV status and HIV sexual transmission behaviors among HIV-positive Black men who have sex with men in the BROTHERS (HPTN 061) study. J Urban Health 2020 Oct;97(5):692-703. [doi: 10.1007/s11524-020-00419-9] [Medline: 32020466]
39. Fisher MP, Ramchand R, Bana S, Iguchi MY. Risk behaviors among HIV-positive gay and bisexual men at party-oriented vacations. J Stud Alcohol Drugs 2013 Jan;74(1):158-167. [doi: 10.15288/jsad.2013.74.158] [Medline: 23200162]
40. Daskalopoulou M, Rodger A, Phillips AN, et al. Recreational drug use, polydrug use, and sexual behaviour in HIV-diagnosed men who have sex with men in the UK: results from the cross-sectional ASTRA study. Lancet HIV 2014 Oct;1(1):e22-e31. [doi: 10.1016/S2352-3018(14)70001-3] [Medline: 26423813]
41. He L, Pan X, Wang N, et al. New types of drug use and risks of drug use among men who have sex with men: a cross-sectional study in Hangzhou, China. BMC Infect Dis 2018 Apr 17;18(1):182. [doi: 10.1186/s12879-018-3091-z] [Medline: 29665785]
42. Huang PX, Duan Q, Liao MZ, et al. Analysis of rush poppers use and related factors in men who have sex with men in Jinan and Qingdao, 2016. Zhonghua Yu Fang Yi Xue Za Zhi 2020 Aug 6;54(8):861-866. [doi: 10.3760/cma.j.cn112150-20190815-00659] [Medline: 32842316]

Abbreviations

MSM: men who have sex with men
STI: sexually transmitted infections
PEP: postexposure prophylaxis
PrEP: pre-exposure prophylaxis
aOR: adjusted odds ratio
UNAIDS: Joint United Nations Program on HIV/AIDS (UNAIDS)

XSL•FO
RenderX

Edited by A Mavragani; submitted 05.09.24; peer-reviewed by J Jiang, S Chen; revised version received 18.11.24; accepted 19.11.24; published 13.01.25.

Please cite as:
Ma Q, Jiang T, Chen W, Jiang S, Zheng J, Wang H, He L
Awareness of Sexual Partner's HIV Status Among Men Who Have Sex With Men in China: Cross-Sec. tional Survey Study
JMIR Public Health Surveill 2025;11:e66139
URL: https://publichealth.jmir.org/2025/1/e66139
doi:10.2196/66139

© Qian Ma, Tingting Jiang, Wanjun Chen, Shaoqiang Jiang, Jinlei Zheng, Hui Wang, Lin He. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effects of MyChoices and LYNX Mobile Apps on HIV Testing and Pre-Exposure Prophylaxis Use by Young US Sexual Minority Men: Results From a National Randomized Controlled Trial

Katie B Biello[1,2], MPH, PhD; Kenneth H Mayer[2,3], MD; Hyman Scott[4,5], MD; Pablo K Valente[6], MPH, PhD; Jonathan Hill-Rorie[2], MPH; Susan Buchbinder[4,5], MD; Lucinda Ackah-Toffey[7], MPH; Patrick S Sullivan[7], DVM, PhD; Lisa Hightow-Weidman[8], MPH, PhD; Albert Y Liu[4,5], MPH, MD

[1]Department of Epidemiology, School of Public Health, Brown University, Providence, RI, United States

[2]The Fenway Institute, Fenway Health, Boston, MA, United States

[3]Department of Medicine, Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, MA, United States

[4]Bridge HIV, San Francisco Department of Public Health, San Francisco, CA, United States

[5]Department of Medicine, University of California, San Francisco, San Francisco, CA, United States

[6]University of Connecticut, Waterbury, CT, United States

[7]Rollins School of Global Public Health, Emory University, Atlanta, GA, United States

[8]Institute on Digital Health and Innovation, College of Nursing, Florida State University, Tallahassee, FL, United States

Corresponding Author:
Katie B Biello, MPH, PhD
Department of Epidemiology
School of Public Health
Brown University
121 South Main Street
Box G-S121-2
Providence, RI, 02912
United States
Phone: 1 4018633082
Email: katie_biello@brown.edu

Abstract

Background: Young sexual minority men have among the highest rates of HIV in the United States; yet, the use of evidence-based prevention strategies, including routine HIV testing and pre-exposure prophylaxis (PrEP), remains low. Mobile apps have enormous potential to increase HIV testing and PrEP use among young sexual minority men.

Objective: This study aims to assess the efficacy of 2 theory- and community-informed mobile apps—LYNX (APT Mobility) and MyChoices (Keymind)—to improve HIV testing and PrEP initiation among young sexual minority men.

Methods: Between October 2019 and May 2022, we implemented a 3-arm, parallel randomized controlled trial in 9 US cities to test the efficacy of the LYNX and MyChoices apps against standard of care (SOC) among young sexual minority men (aged 15-29 years) reporting anal sex with cisgender male or transgender female in the last 12 months. Randomization was 1:1:1 and was stratified by site and participant age; there was no masking. The co-primary outcomes were self-reported HIV testing and PrEP initiation over 6 months of follow-up.

Results: A total of 381 young sexual minority men were randomized. The mean age was 22 (SD 3.2) years. Nearly one-fifth were Black, non-Hispanic (n=67, 18%), Hispanic or Latino men (n=67, 18%), and 60% identified as gay (n=228). In total, 200 (53%) participants resided in the Southern United States. At baseline, participants self-reported the following: 29% (n=110) had never had an HIV test and 85% (n=324) had never used PrEP. Sociodemographic and behavioral characteristics did not differ by study arm. Compared to SOC (n=72, 59%), participants randomized to MyChoices (n=87, 74%; *P*=.01) were more likely to have received at least 1 HIV test over 6 months of follow-up; those randomized to LYNX also had a higher proportion of testing (n=80, 70%) but it did not reach the a priori threshold for statistical significance (*P*=.08). Participants in both MyChoices (n=23, 21%) and LYNX (n=21, 20%) arms had higher rates of starting PrEP compared to SOC (n=19, 16%), yet these differences were not statistically significant (*P*=.52).

XSL•FO
RenderX

Conclusions: In addition to facilitating earlier treatment among those who become aware of their HIV status, given the ubiquity of mobile apps and modest resources required to scale this intervention, a 25% relative increase in HIV testing among young sexual minority men, as seen in this study, could meaningfully reduce HIV incidence in the United States.

Trial Registration: ClinicalTrials.gov NCT03965221; https://clinicaltrials.gov/study/NCT03965221

(JMIR Public Health Surveill 2025;11:e63428) doi:10.2196/63428

KEYWORDS

HIV testing; adolescents; sexual minority men; mobile apps; pre-exposure prophylaxis; youths; randomized controlled trial; sexual minority; United States; efficacy; LYNX; MyChoices; sociodemographic; behavioral health; HIV prevention; HIV infection; HIV care; transmission; public health; mHealth; mobile phones

Introduction

Sexual minority men (eg, gay, bisexual, and other men who have sex with men) account for nearly two-thirds of all people living with HIV in the United States while comprising less than 10% of the population. Sexually active young sexual minority men, aged 13-25 and 25-34 years, have among the highest rates of HIV [1,2]. Young sexual minority men are also the least likely to be aware of their HIV status, further enhancing HIV transmission risks and delaying HIV care [2].

The Centers for Disease Control and Prevention (CDC) recommends all sexually active sexual minority men get tested for HIV at least annually, and consider testing more often (ie, every 3 to 6 months) for sexual minority men who are sexually active [3]. HIV testing is essential for early detection, early care engagement, and early treatment—resulting in a lower likelihood of onward transmission and improved quality of life [3]. For those who test negative, it is also an opportunity to engage in pre-exposure prophylaxis (PrEP) care. The US Preventive Services Task Force expanded the recommendation for PrEP in August 2023 to include adolescents, in addition to adults, who report anal or vaginal sexual activity and disclose one of the following criteria: sexual partner with HIV, a recent bacterial sexually transmitted infection (STI), or inconsistent or no condom use with sexual partners [4]. Despite the efficacy of and need for these evidence-based prevention strategies—routine HIV testing and PrEP—use remains suboptimal among young sexual minority men [5-7].

Reasons for a lack of routine HIV testing and limited PrEP use among young sexual minority men are multilevel and complex [8]. For example, at the individual level, young sexual minority men tend to underestimate their personal risk of acquiring HIV [9,10]; at the interpersonal level, young sexual minority men may be concerned that their partner will feel threatened if they test for HIV or initiate PrEP [11,12]; at the clinical level, young sexual minority men may not be comfortable disclosing their sexual activity to providers [13-15]; at the structural level, young sexual minority men may not know where to access HIV testing or PrEP, or have the autonomy or ability to access a testing site or PrEP clinic on their own [11,16]. Any intervention or program to increase routine HIV testing and PrEP use will need to address a plethora of multilevel challenges [8].

Mobile apps—accessed via smartphones which have nearly complete coverage in the United States, particularly among youths—have enormous potential to increase routine HIV testing and PrEP use among young sexual minority men [17-19]. Ninety-five percent of youth report having a smartphone [20], and youth commonly use their mobile devices for a range of activities, including downloading health-related apps and accessing health information [21]. Implemented through mobile apps, mobile health (mHealth) interventions may be a promising tool for promoting behaviors such as HIV testing and PrEP use among young sexual minority men who have been poorly reached through standard clinic services [17-19].

In prior work, 2 distinct apps, MyChoices (Keymind) and LYNX (APT Mobility), each designed to increase HIV testing and PrEP Initiation among young sexual minority men, were developed by separate research teams using different theoretical models to effect behavior change and different approaches for engagement (Multimedia Appendix 1). The process for development and pilot testing of these apps has been described previously [22-25]. In brief, both apps were developed with extensive formative research with young sexual minority men, including interviews and focus groups, beta testing, and pilot randomized controlled trials (RCTs). In parallel pilot RCTs conducted as part of the National Institutes of Health's Adolescent Medicine Trials Network (ATN) for HIV Interventions, both MyChoices and LYNX were found to be feasible and acceptable, meeting predetermined go or no-go criteria for this current full-scale efficacy trial.

This study presents the results of a 3-arm, RCT conducted in 9 US cities in order to assess the efficacy of the LYNX and MyChoices apps to improve HIV testing and PrEP initiation among young sexual minority men.

Methods

Trial Design

This was a 3-arm, RCT with a 1:1:1 allocation ratio conducted as part of the ATN and the University of North Carolina/Emory Center for Innovative Technology (iTech), and implemented between October 2019 and May 2022 [26,27]. Consented participants completed a baseline assessment and follow-up assessments at 3- and 6-months post baseline. Some participants also completed an additional follow-up assessment between 6 and 12 months post baseline (see Adjustments in Trial Design in Response to the COVID-19 Pandemic section for more details).

XSL•FO
RenderX

Participants

To be eligible participants had to (1) identify as cisgender men, (2) be aged 15-29 years, (3) self-report being HIV serostatus negative or HIV serostatus unknown at screening, (4) not have had an HIV test in the past 3 months, (5) not be currently taking PrEP, (6) self-report at least 1 episode of anal intercourse with a male or transgender female partner during the last 12 months, (7) had to own or have regular and ongoing access to an iOS or Android smartphone, (8) be willing and able to download the MyChoices or LYNX app, and (9) be able to understand, read, and speak English.

Participants were recruited through social media, flyers, and direct outreach at local venues and clinical sites across 9 sites: Atlanta, Georgia (study site: PRISM Health); Boston, Massachusetts (study site: Fenway Health); Bronx, New York City (study site: the Adolescent AIDS Program at Montefiore); Chapel Hill, North Carolina (study site: University of North Carolina, Chapel Hill); Charlotte, North Carolina (study site: RAIN); Chicago, Illinois (study site: AYAR at CORE Center); Houston, Texas (study site: Texas Children's Hospital); Philadelphia, Pennsylvania (study site: Children's Hospital of Philadelphia); and Tampa, Florida (study site: University of South Florida).

Ethical Considerations

The study procedures were reviewed and approved by the University of North Carolina institutional review board (IRB) as a single IRB-of-record (study 19-0260). IRB authorization agreements with all participating research entities were enacted. A waiver of parental consent was obtained for individuals aged 15-17 years. All enrolled participants completed a site-specific consent or assent, which contained the essential elements, including a description of the purpose, study procedures, and risks and benefits. No social harms or adverse events were reported. Compensation for completing the survey was set by each site in accordance with norms, cost of living, and input from their IRBs. Participants were not compensated for using the app. The protocol for ATN 143: COMPARE (Comparison of Men's Prevention Apps to Research Efficacy) is registered at ClinicalTrials.gov (NCT03965221).

Randomization and Masking

Randomization was computer generated and was equally distributed 1:1:1 for study arms. Randomization was stratified by site and participant age (15-18 and 19-29 years). This trial did not use masking. Following randomization, study staff supported participants randomized to the experimental arms through downloading the app, and then provided a brief, structured demonstration of the main components of the app.

Interventions

Standard of Care

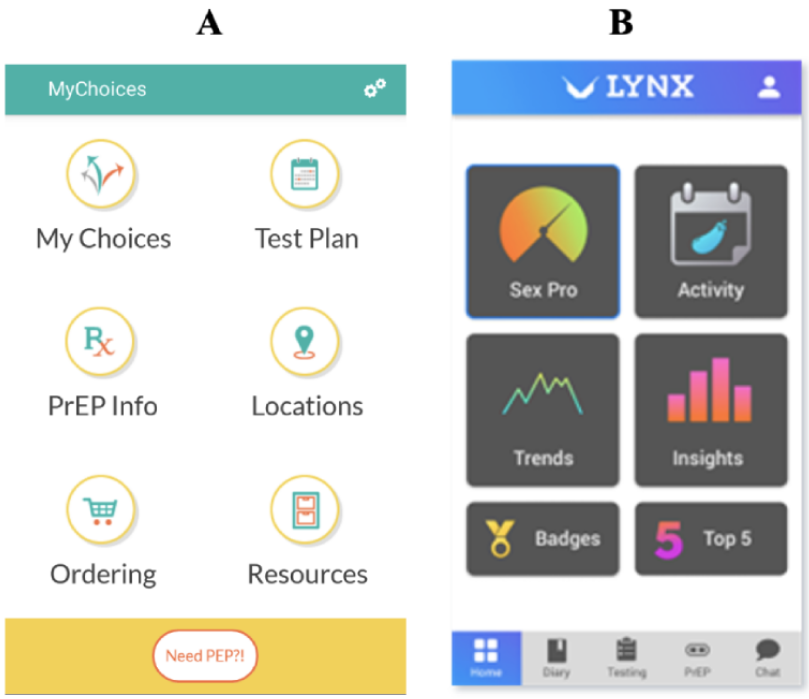
Participants in all study conditions received CDC fact sheets on HIV testing and PrEP Initiation and had access to any services available as standard of care (SOC) at each study site [28,29].

Experimental Condition: MyChoices Mobile App

As described and detailed previously [23,24], MyChoices was adapted, through extensive formative research and pilot testing, from an app previously demonstrated to increase HIV testing and PrEP Initiation in adult sexual minority men [22-24,30-32]. The MyChoices app is informed by the social cognitive theory [33] and includes the following primary components: sexual health information using a variety of formats (eg, infographics, videos, weblinks to educational resources); access to free home HIV and STI testing kits as well as condoms and lubricants; individualized, tailored HIV prevention recommendations and HIV testing plans based on monthly check-ins; reminders for HIV testing with geolocation-based notifications when near HIV testing sites and GPS-enabled maps with local HIV and STI testing locations and PrEP providers (Figure 1A).

XSL•FO
RenderX

Figure 1. (A) MyChoices and (B) LYNX app landing pages.



Experimental Condition: LYNX Mobile App

As described previously [25], LYNX was developed through iterative formative research and pilot testing with young sexual minority men. The LYNX app is guided by the information-motivation-behavioral skills (IMB) model [34] and includes the following primary components: Sexual Health Promotion (Sex Pro) score that provides a personalized HIV risk score; sexual diary to help participants track sexual partners and encounters; reminders for HIV or STI testing; ability to order free home HIV and STI testing kits as well as condoms and lubricant; a geolocation map of HIV or STI testing sites and PrEP clinics; a PrEP information page including a roadmap on how to get on PrEP; a "trends and insights" page including infographics summarizing partner types and sexual activities; and a 2-way chat feature to answer questions and provide real-time assistance with PrEP navigation (Figure 1B).

Privacy

To enhance confidentiality, in order to access either app, users initially entered their email address and a registration key (supplied following randomization) and set up a user-generated password. On subsequent visits to the app, users were required to supply their username and password or create a PIN. The apps timed out every time users left the app or when the phone went into sleep mode.

Adjustments in Trial Design in Response to the COVID-19 Pandemic

Enrollment for the COMPARE study began in October 2019. At this point, all enrollment visits were conducted in person at the enrolling sites (n=83). All sites paused enrollment due to COVID-19 from March 2020 to June 2020. In June 2020, sites began to resume enrollment, with all enrollment visits occurring via videoconferencing. This continued until November 2021 when enrollment closed. Given the entirely remote procedures, in October 2021 and November 2021, enrollment was expanded to the entire United States. Participants were assigned to the study site that was closest to their zip code of residence. See Multimedia Appendix 2 for the enrollment timeline. In addition to changes in enrollment procedures, the COVID-19 pandemic significantly slowed down enrollment. In response, in collaboration with the ATN-wide Study Monitoring Committee, the decision was made to extend the enrollment period and shorten the full follow-up period for participants enrolled between June 2021 and November 2022 (follow-up ranging from 6 months to 11 months; n=40 participants were only eligible for a 6-month visit because they were enrolled in November 2022; the remaining 341 participants were eligible to complete an additional follow-up survey between 7 and 12 months post baseline). As such, our statistical analysis plan (also approved by the Study Monitoring Committee and IRB) was revised to shift the primary endpoint to be measured at the 6-month follow-up visit (rather than the 12-month follow-up visit). Importantly, based on eligibility criteria, all participants would be recommended to have an HIV test every 3 to 6 months according to CDC guidelines; as such, the outcome time frame remained clinically meaningful.

Assessments

Baseline assessments were conducted at the enrolling study site or via videoconferencing and self-administered using a secure web-based data entry system. All follow-up assessments were self-administered and conducted off-site through the same secure web-based system. Follow-up assessments occurred at 6-month postbaseline for all participants, and an additional follow-up assessment occurred between 7- and 12-month post baseline for all participants enrolled prior to November 2022.

XSL•FO
RenderX

Primary Efficacy Outcome Measures

HIV testing and PrEP initiation over study follow-up were assessed by self-report at each follow-up visit. Specifically, participants were asked "How many times have you had an HIV test in the past 3 months?" with response options: "I have not had an HIV test in the past 3 months, 1 time, 2 times, 3 times, and 4 or more times." For PrEP, participants were asked "In the past 3 months, have you taken PrEP (pre-exposure prophylaxis) to reduce the chance of getting HIV?" with response options: "Yes, I am on PrEP right now," "Yes, but I'm not on PrEP anymore," and "No."

Secondary Outcome Measures

Time-to-event outcomes. As secondary outcomes, we examined the time to first HIV test and time to PrEP initiation from the baseline assessment using all available follow-ups. At each assessment, participants were asked the month and year of all their HIV tests and PrEP Initiation (if applicable) since the previous assessment. Since the specific dates when HIV tests and PrEP initiation took place were not assessed, we considered these events to happen on the last day of the month in which they reported the events.

HIV Testing Frequency

Participants were asked the number of times they were tested for HIV in the past 3 months (at 3- and 6-month visits) and 6 months (at 12-month visits). Responses were summed to obtain the total number of times participants were tested for HIV through the 12 months of follow-up, which ranged from 0 to 4 or more times.

PrEP IMB Scale

We created a 5-item scale to assess the 3 domains of the IMB model related to PrEP uptake: information (1 item), motivation (1 item), and behavioral skills (3 items). Sample items included, "I understand whether PrEP would be a good fit for me," "I feel motivated to get on PrEP," and "I know how to get PrEP". Responses on a 5-point scale were averaged to reach a total IMB score. The Cronbach α=0.83 indicated strong internal consistency.

PrEP Self-Efficacy

We adopted the PrEP Adherence Self-Efficacy Scale [35] to assess PrEP self-efficacy in the past month with 10 items asking how confident participants had been that they can or could "integrate PrEP into your daily routine" or "take a PrEP pill every day even when you aren't feeling well." Response options ranged from 0="could not do at all" to 10="completely certain could do." We obtained a total PrEP self-efficacy score based on the average of all 10 responses. The Cronbach α=0.96 indicated strong internal consistency.

PrEP Stigma

We assessed PrEP stigma with 12 items drafted by the study team, including "I can tell my friends that I am using PrEP," "I would not want a sexual partner to see my PrEP pills in my medicine cabinet," and "People would think I am sexually risky if they find out that I take PrEP." Responses on a 5-point scale were reverse-coded as needed and averaged across 12 items such that higher scores indicated higher PrEP stigma. The Cronbach α=0.76 indicated moderate internal consistency.

Discussing PrEP With Health Care Provider

In order to measure the first steps in the PrEP care continuum, we assessed whether participants talked to a health care provider about PrEP (yes vs no).

PrEP Interest

Participants who had not initiated PrEP were asked about their interest in taking PrEP with a single item with response options ranging from 1="not at all interested" to 5="extremely interested."

App Use Measures

We assessed the overall number of log-ins in each of the apps, the number of calendar days with app logins, and the total amount of time spent in the apps (minutes).

Sociodemographic Measure

To characterize the sample and examine potential moderating effects, we collected participants' age, race and Hispanic or Latinx ethnicity, current relationship status (ie, single vs not single), and zip code of residence (which was aggregated to census regions: South, Northeast, Midwest, and West). We also assessed recent history of condomless anal sex by asking participants a series of questions about past-3-month sex partners, type of sex with those partners, and condom use during sex [36,37].

Power and Sample Size

Using data from our pilot study, we powered the study to see a significant difference between each experimental condition compared to the standard of care assuming a baseline past-year HIV testing rate of 40%, HIV testing rate in the interventions arm at 77% (risk ratio [RR] 1.56), any app use in intervention arms of 60%, α=.05, and power=.80. With 70% retention at 12 months follow-up and equal allocation, we estimated a sample size of 450 (n=150 per arm, 33%). Notably, in June 2021, at the request of the ATN-wide Study Monitoring Committee, we reran our sample size calculation with all of the same assumptions but based on the higher-than-expected retention demonstrated at that point in the study (approximately 90% at each visit). The new targeted sample size was 351 (n=117 per arm, 33%).

Statistical Analysis

Means for continuous variables and frequencies for categorical variables were calculated to describe the participants at baseline, overall, and stratified by intervention condition to assess balance in baseline characteristics.

An intent-to-treat approach was used, analyzing participants in the study arm to which they were assigned. As missing data was minimal (less than 10%), we used complete case analysis for each study aim. The primary efficacy analyses compared HIV testing and PrEP Initiation between the study arms over the study follow-up (ie, 6 months) using chi-square tests. Risk ratios and their corresponding 95% CIs were also calculated using log-binomial regression. Moreover, we examined

XSL•FO
RenderX

differences between the study arms in the hypothesized mediators, including information, motivation, and behavioral skills related to PrEP Initiation, PrEP self-efficacy, PrEP stigma, discussing PrEP with a health care provider, and PrEP interest.

Secondary outcomes included time-to-event analyses for HIV testing and PrEP Initiation through the 12 months of follow-up using Kaplan-Meier survival curves. We examined overall differences across study arms using log-rank tests, and Cox proportional hazard models to obtain hazard ratios by study arm. We also examined the frequency of HIV testing over 12 months using Poisson regression models with robust standard errors.

We also assessed whether the intervention effect on the outcomes differed by baseline characteristics (ie, age, race or ethnicity, region, condomless anal sex, relationship status, depression and anxiety, and prepandemic enrollment). These moderator analyses were chosen a priori (with the exception of the pandemic measure) and were evaluated with interaction terms; we probed potential moderation in post hoc stratified analyses if the *P* value of interaction terms was less than .20.

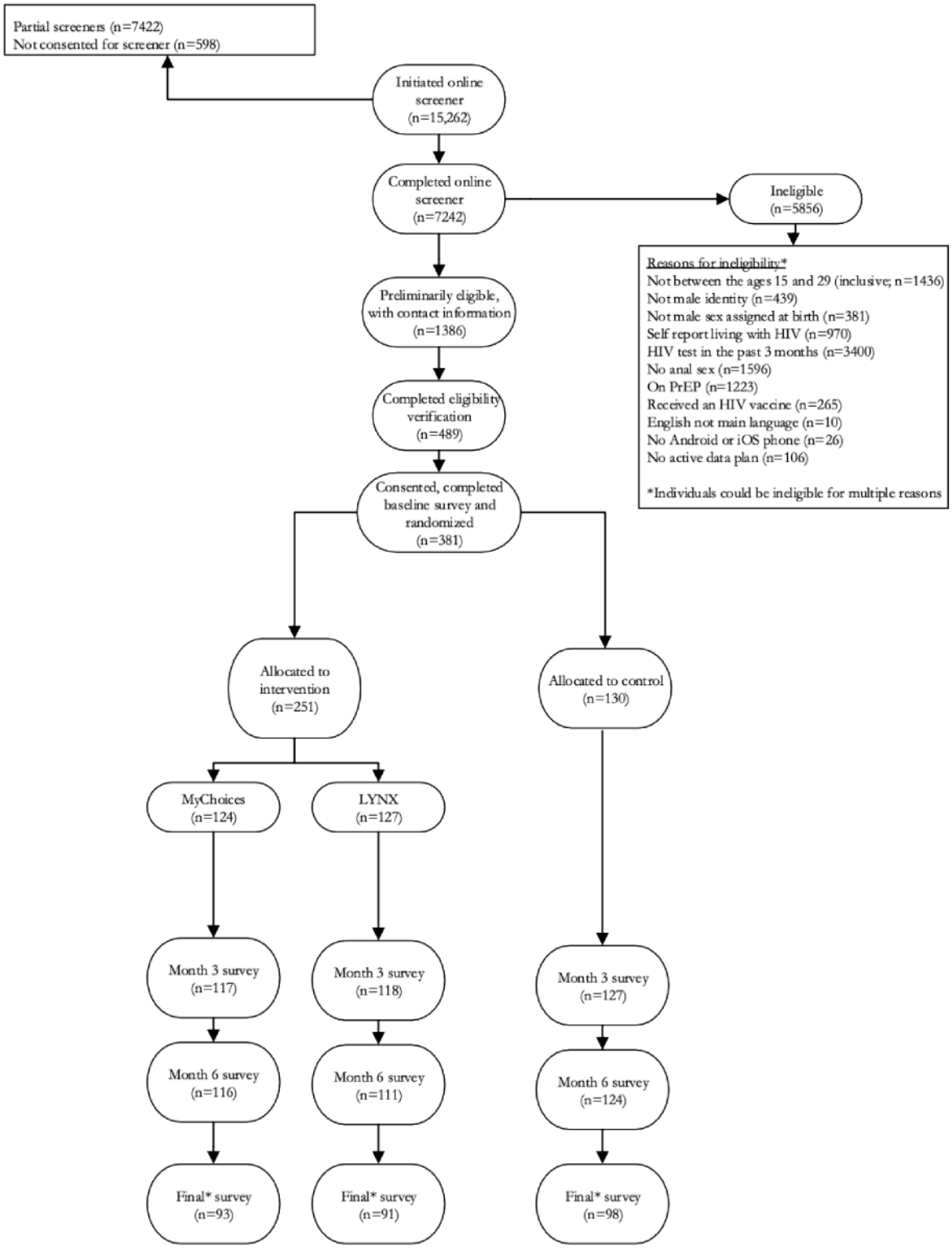
In secondary analyses, we explored dose-response relationships by examining the effect of measures of usage of each app on HIV testing and PrEP Initiation at 6 months using log-binomial regression. In all dose-response analyses, we only considered app usage until the date of the assessment when outcomes of interest (ie, HIV testing and PrEP Initiation) were first reported. Since app usage measures were right-skewed, we winsorized these variables at the 98th percentile in all regression analyses [38].

Results

Online eligibility screeners were completed by 7,242 individuals. Of those, 1,386 screened initially eligible and provided contact information. Common reasons for ineligibility included having a recent HIV test and not reporting recent anal sex. Four hundred and eighty-nine individuals were able to be contacted and attended the initial visit. Among these, 381 participants consented, completed a baseline assessment and were randomized to either the MyChoices app (n=124), the LYNX app (n=127) or the comparison condition (n=130). Among randomized participants, 95% completed a 3-month follow up assessment and 92% completed a 6-month follow up assessment. Figure 2 depicts participant flow throughout the study from screening to follow-up.

XSL•FO
RenderX

Figure 2. COMPARE (Comparison of Men's Prevention Apps to Research Efficacy) study CONSORT (Consolidated Standards of Reporting Trials) diagram, 2019-2022. PrEP: pre-exposure prophylaxis. *Final follow-up period ranged from 7-12 months postbaseline. Of the 381 enrolled, 341 were eligible to complete this final follow-up (n=111 in MyChoice; n=115 in LYNX; n=115 in standard of care).



Sample characteristics at baseline are described in Table 1. The mean age was 22.4 (SD 3.2); 51% were White, non-Hispanic (n=194), 18% were Black, non-Hispanic (n=67), and 18% were Latino (n=67). 60% (n=228) identified as gay. 53% (n=200) resided in the Southern United States. At baseline, participants self-reported the following: 29% (n=110) had never had an HIV test, and 67% (n=253) had not had an HIV test in the prior 3 months. 85% (n=324) had never used PrEP, and an additional 13% (n=51) had not used PrEP in the prior 3 months. Sociodemographic and behavioral characteristics did not differ significantly by study arm.

XSL•FO
RenderX

Table 1. Baseline characteristics for the COMPARE[a] study, overall and by study condition, 2019-2022.

	Total (N=381)	SOC[b] (n=130)	LYNX (n=127)	MyChoices (n=124)
Age (years), mean (SD)	22.4 (3.17)	22.8 (3.34)	22.3 (3.20)	22.2 (2.96)
Race and ethnicity, n (%)				
Asian, non-Hispanic	34 (9)	10 (7.8)	12 (9.5)	12 (9.8)
Black, non-Hispanic	67 (17.7)	25 (19.4)	26 (20.5)	16 (13)
Latino or Hispanic	67 (17.7)	29 (22.5)	23 (18.1)	15 (12.2)
White, non-Hispanic	194 (51.2)	58 (45)	59 (46.5)	77 (62.6)
Other or Mixed	17 (4.5)	7 (5.4)	7 (5.5)	3 (2.4)
Census region of current residence, n (%)				
Midwest	56 (14.7)	19 (14.6)	18 (14.2)	19 (15.3)
Northeast	112 (29.4)	36 (27.7)	39 (30.7)	37 (29.8)
South	200 (52.5)	68 (52.3)	67 (52.8)	65 (52.4)
West	13 (3.4)	7 (5.4)	3 (2.4)	3 (2.4)
Insurance status, n (%)				
Uninsured	43 (11.3)	12 (9.2)	12 (9.5)	19 (15.5)
Own insurance	136 (35.8)	46 (35.4)	47 (37)	43 (35)
Parent’s insurance	196 (51.6)	70 (53.9)	67 (52.8)	59 (48)
Do not know	5 (1.3)	2 (1.5)	1 (0.8)	2 (1.6)
Sexual orientation, n (%)				
Bisexual	69 (18.1)	23 (17.7)	23 (18.1)	23 (18.6)
Gay or homosexual	228 (59.8)	78 (60)	78 (61.4)	72 (58.1)
Queer or same sex or other	77 (20.2)	26 (20)	22 (17.3)	29 (23.4)
Heterosexual or straight	7 (1.8)	3 (2.3)	4 (3.2)	0 (0)
Relationship status, n (%)				
Single	182 (47.9)	62 (47.7)	61 (48.0)	59 (48)
Has boyfriend, girlfriend, partner, or lover	119 (31.3)	39 (30)	39 (30.7)	41 (33.3)
Casually dating	73 (19.2)	27 (20.8)	27 (21.3)	19 (15.5)
Married, civil union, or domestic partnership	6 (1.6)	2 (1.5)	0 (0)	4 (3.3)
Employment status, n (%)				
Unemployed	121 (31.9)	37 (28.7)	48 (37.8)	36 (29.3)
Employed part-time	126 (33.3)	41 (31.8)	40 (31.5)	45 (36.6)
Employed full-time	132 (34.8)	51 (39.5)	39 (30.7)	42 (34.2)
Student status, n (%)				
Currently enrolled	235 (61.8)	79 (60.8)	77 (60.6)	79 (64.2)
Not currently enrolled	145 (38.2)	51 (39.2)	50 (39.4)	44 (35.8)
Prior HIV test, n (%)				
Never	110 (29.1)	40 (30.8)	35 (27.9)	35 (28.7)
More than 3 months	253 (66.9)	87 (66.9)	85 (67.5)	81 (66.4)
Within past 3 months	15 (4.0)	3 (2.3)	6 (4.8)	6 (4.9)
Prior PrEP[c] use, n (%)				
Never	324 (85.3)	112 (86.2)	110 (86.6)	102 (82.9)
More than 3 months	51 (13.4)	15 (11.5)	17 (13.4)	19 (15.5)

XSL•FO
RenderX

	Total (N=381)	SOC[b] (n=130)	LYNX (n=127)	MyChoices (n=124)
Within past 3 months	5 (1.3)	3 (2.3)	0 (0)	2 (1.6)
Condomless anal or vaginal sex (past 3 months), n (%)				
Yes	150 (39.4)	56 (43.1)	52 (40.9)	42 (33.9)
No	231 (60.6)	74 (56.9)	75 (59.1)	82 (66.1)
Enrolled Pre–COVID-19 shutdown, n (%)				
Yes	82 (21.5)	27 (20.8)	30 (23.6)	25 (20.2)
No	299 (78.5)	103 (79.2)	96 (76.4)	98 (79.8)

[a]COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b]SOC: standard of care.

[c]PrEP: pre-exposure prophylaxis.

HIV Testing

Compared to SOC (n=72, 59%), participants randomized to MyChoices were significantly more likely to have received at least 1 HIV test over 6 months of follow-up (n=87, 74%; *P*=.010; Table 2). Participants randomized to LYNX also had a higher proportion of testing but it did not meet the a priori threshold for statistical significance (n=80, 70%; *P*=.08). Individuals randomized to the MyChoices condition also reported more frequent HIV testing over follow-up compared to SOC (incidence rate ratio 1.25, 95% CI 1.01-1.54; *P*=.04), while this relationship did not reach significance for the LYNX condition compared to SOC (incidence rate ratio 1.13, 0.91-1.40; *P*=.26; Table 3). In time-to-event analyses (Table 4 and Multimedia Appendix 3), using all data (up to 12 months), those randomized to the MyChoices condition had a 45% higher likelihood (hazard ratio 1.45, 95% CI 1.07-1.95; *P*=.02) and those in the LYNX condition had a 34% higher likelihood (hazard ratio 1.34, 95% CI 1.00-1.81; *P*=.05) of receiving an HIV test compared to those in the SOC condition (log-rank *P*=.04).

Table 2. Estimated uptake of HIV testing and PrEP[a] in the COMPARE[b] study, 2019-2022.

	Any HIV test over 6 months			Any PrEP initiation over 6 months		
	n (%)	Risk ratio (95% CI)	*P* value	n (%)	Risk ratio (95% CI)	*P* value
SOC[c] (reference)	72 (58.5)	1.0	—[d]	19 (15.6)	1.0	—
MyChoices	87 (74.4)	1.27 (1.06-1.53)	.01	23 (21.3)	1.37 (0.79-2.37)	.26
LYNX	80 (70)	1.19 (0.98-1.44)	.08	21 (19.6)	1.26 (0.72-2.21)	.42

[a]PrEP: pre-exposure prophylaxis.

[b]COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[c]SOC: standard of care.

[d]Not applicable.

Table 3. Frequency of HIV testing in the COMPARE[a] study, 2019-2022.

	Incidence rate ratio (95% CI)	*P* value
SOC[b] (reference)	—[c]	—
MyChoices	1.25 (1.01-1.54)	.04
LYNX	1.13 (0.91-1.40)	.26

[a]COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b]SOC: standard of care.

[c]Not applicable.

XSL•FO
RenderX

Table 4. Time to first HIV test and PrEP[a] initiation in the COMPARE[b] study, 2019-2022.

	Time to first HIV test over 12 months		Time to PrEP initiation over 12 months	
	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
SOC[c] (reference)	—[d]	—	—	—
MyChoices	1.45 (1.07-1.95)	.015	0.90 (0.47-1.71)	.75
LYNX	1.34 (1.00-1.81)	.052	1.17 (0.64-2.13)	.61

[a]PrEP: pre-exposure prophylaxis.

[b]COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[c]SOC: standard of care.

[d]Not applicable.

According to the omnibus tests, the baseline characteristics did not significantly modify the effect of either app on HIV testing compared to SOC. However, in post hoc analyses (Table 5), both apps were significantly associated with HIV testing compared to SOC among White, non-Hispanic participants (MyChoices: RR 1.6, 95% CI 1.2-2.1; LYNX: RR 1.4, 95% CI 1.0-2.0); however, the conditions did not differ among non-White participants (MyChoices: RR 1.06, 95% CI 0.8-1.4; LYNX: RR 1.04, 95% CI 0.8-1.3). Among participants who reported engaging in condomless sex with a male or transgender female partner at baseline (Table 6), participants randomized to the MyChoices condition were nearly 6 times as likely to get an HIV test compared to SOC (RR 5.6, 95% CI 1.7-18.1; *P*=.004); the likelihood did not significantly differ among those not reporting recent condomless sex (RR 1.5, 95% CI 0.8-3.0; *P*=.21). Finally, participants who enrolled after the COVID-19 shutdown and were randomized to the MyChoices and LYNX conditions were significantly more likely to obtain an HIV test compared to the SOC condition (*P*=.005 and *P*=.02, respectively); the conditions did not differ for participants randomized prior to the COVID-19 shutdown (Table 7).

Table 5. Post hoc stratified analysis to assess moderation for HIV testing at 6 months by race or ethnicity in the COMPARE[a] study, 2019-2022.

	White, non-Hispanic			Non-White		
	n (%)	RR[b] (95% CI)	*P* value	n (%)	RR (95% CI)	*P* value
SOC[c] (reference)	28 (48.3)	1.0	—[d]	43 (67.2)	1.0	—
MyChoices	56 (75.7)	1.57 (1.17-2.11)	.003	30 (71.4)	1.06 (0.82-1.37)	.64
LYNX	35 (68.6)	1.42 (0.67-3.00)	.03	45 (70.3)	1.05 (0.83-1.32)	.70

[a]COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b]RR: risk ratio.

[c]SOC: standard of care.

[d]Not applicable.

Table 6. Post hoc stratified analyses to assess moderation for HIV testing at 6 months by condomless anal or vaginal sex (past 3 months) in the COMPARE[a] study, 2019-2022.

	Yes			No		
	n (%)	RR[b] (95% CI)	*P* value	n (%)	RR (95% CI)	*P* value
SOC[c] (reference)	33 (62.3)	1.0	—[d]	39 (55.7)	1.0	—
MyChoices	37 (90.2)	5.61 (1.74-18.09)	.004	50 (65.8)	1.53 (0.78-2.98)	.21
LYNX	36 (76.6)	1.98 (0.83-4.75)	.13	44 (64.7)	1.46 (0.73-2.89)	.28

[a]COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b]RR: risk ratio.

[c]SOC: standard of care.

[d]Not applicable.

XSL•FO
RenderX

Table 7. Post hoc stratified analyses to assess moderation for HIV testing at 6 months by pre–COVID-19 enrollment in the COMPARE[a] study, 2019-2022.

	Yes			No		
	n (%)	RR[b] (95% CI)	*P* value	n (%)	RR (95% CI)	*P* value
SOC[c] (reference)	13 (68.4)	1.0	—[d]	59 (56.7)	1.0	—
MyChoices	13 (59.1)	1.04 (0.68-1.59)	.86	74 (77.9)	1.37 (1.13-1.68)	.005
LYNX	12 (54.6)	0.90 (0.57-1.40)	.63	68 (73.1)	1.29 (1.04-1.59)	.02

[a]COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[b]RR: risk ratio.

[c]SOC: standard of care.

[d]Not applicable.

PrEP Initiation

While participants in both MyChoices (21%) and LYNX (20%) arms had higher rates of initiating PrEP compared to SOC (16%), these differences were not statistically significant (*P*=.52; Table 2). The time-to-event analysis also did not demonstrate a significant difference in PrEP Initiation between study conditions (log-rank *P*=.72; Table 4 and Multimedia Appendix 3). Omnibus tests did not demonstrate an effect modification of PrEP Initiation by baseline characteristics or time of enrollment (all *P*>.10).

Given the null findings for PrEP Initiation, we explored, in post hoc analyses, whether the apps were efficacious in improving the hypothesized PrEP-related mediators. Participants randomized to the MyChoices condition had significantly higher scores for the PrEP IMB scale compared to the SOC condition (β=.24; 95% CI 0.03-0.44; *P*=.02). Among participants who had not initiated PrEP, participants randomized to the LYNX arm had a lower interest in PrEP compared to SOC (β=–.43; 95% CI –0.78 to –0.08; *P*=.02); PrEP interest among those randomized to the MyChoices arm who did not initiate PrEP was not significantly different from SOC (β=–.17, 95% CI –0.53 to 0.18; *P*=.33). The arms did not differ for PrEP self-efficacy, PrEP stigma scores, nor having spoken to a provider about PrEP.

App Use

MyChoices and LYNX app use data are described in Table 4. In brief, for MyChoices, the mean number of log-ins was 8.55 (SD 17.0; range 1-186), and the mean total minutes on the app was 31.00 (SD 26.16; range 2.95-148.83). For LYNX, the mean number of log-ins was 12.43 (SD 13.10; range 1-74), and the mean total minutes on the app was 34.52 (SD 33.57; range 0.00-209.17). For either app, among those randomized to the experimental conditions, the number of app logins and time on the app were not significantly associated with HIV testing or PrEP Initiation (Table 8).

Table 8. Use and dose-response effects of MyChoices and LYNX apps on HIV testing and PrEP[a] uptake at 6 months in the COMPARE[b] study, 2019-2022.

	Range	Mean (SD)	HIV testing at 6 months		PrEP Initiation at 6 months	
			RR[c] (95% CI)	*P* value	RR (95% CI)	*P* value
LYNX (n=119)						
Number of log-ins	1-74	12.43 (13.10)	1.00 (0.98-1.02)	.97	0.97 (0.92-1.02)	.21
Number of days with log-in to app	1-49	10.0 (9.78)	1.00 (0.98-1.02)	.98	0.94 (0.87-1.01)	.09
Total minutes spent on app	<0.01-209.17	34.52 (33.57)	1.00 (0.99-1.01)	.72	1.00 (1.00-1.00)	.34
Minutes per login	<0.01-13.39	3.35 (2.38)	0.97 (0.92-1.03)	.39	0.99 (0.98-1.01)	.37
MyChoices (n=124)						
Number of log-ins	1-186	8.55 (16.96)	1.01 (0.99-1.04)	.29	0.94 (0.85-1.04)	.23
Number of days with app log-in	1-153	7.58 (14.03)	1.02 (0.99-1.06)	.24	0.95 (0.85-1.06)	.35
Total minutes spent on app	2.95-148.83	31.00 (26.16)	1.00 (1.00-1.00)	.44	0.99 (0.97-1.01)	.23
Minutes per log-in	0.51-16.58	4.86 (3.04)	1.01 (0.97-1.04)	.72	0.97 (0.85-1.12)	.68

[a]PrEP: pre-exposure prophylaxis.

[b]COMPARE: Comparison of Men's Prevention Apps to Research Efficacy.

[c]RR: risk ratio

XSL•FO
RenderX

Discussion

HIV remains a significant public health problem in the United States, with particular subgroups at increased risk, including young sexual minority men. mHealth interventions, delivered through mobile apps, have the potential to reach this population given their patterns of use. Routine HIV testing and PrEP use are 2 strategies that have been shown to be efficient and cost-effective at reducing HIV transmission, particularly among groups at higher risk for HIV [39,40]. However, barriers to widespread acceptance of these interventions among young sexual minority men have reduced their population-level effects. Our 2 community-driven and theory-informed apps, MyChoices and LYNX, increased the likelihood that a large, diverse sample of at-risk young sexual minority men across the United States received at least 1 HIV test over 6 months of follow-up compared to SOC; however, the apps did not statistically significantly improve the PrEP initiation. While these results were mixed, they are illuminating.

Both MyChoices and LYNX had a meaningful impact on HIV testing compared to SOC; though the effect size for MyChoices was slightly higher and was statistically significant. An approximately 25% relative increase in HIV testing among young sexual minority men, as seen in this study, is similar to those seen in the evidence-based interventions promoted by the CDC for HIV testing [41]. It is estimated that people living with HIV who are unaware of their serostatus account for up to 40% of new sexually transmitted HIV cases, and knowing one's status reduces HIV sexual transmission by 71% [42]. Given the potential reach and successful implementation, these apps could have a substantial impact on the epidemic among young sexual minority men in the United States.

Over the course of the study follow-up, the amount of time that youths engaged with the apps was relatively modest. Future work should explore the role of different types of reminders for repeat testing and assess whether additional information (eg, about having more fulfilling sex, signs and symptoms of STI, other sexual health topics of interest) might lead to more sustained use, and greater impact in establishing regular testing patterns and PrEP use [43]. That said, notably, among young sexual minority men randomized to MyChoices and LYNX, the number of log-ins and the time spent on the app were not associated with each outcome, suggesting that it may not necessarily be the dose but the content and functionalities that matter the most. Future studies should examine the role of app engagement to examine the optimal balance between quantity versus quality [44].

Despite omnibus tests for interactions not showing evidence for effect modification by baseline characteristics, stratified analyses demonstrated that both apps were significantly associated with HIV testing compared to SOC among White, non-Hispanic participants; however, the conditions did not differ among non-White participants. The null finding among non-White participants may have been due to the lack of statistical power due to small sample sizes. However, given that young sexual minority men of color are at disproportionate risk of acquiring HIV, it is essential to further explore how to support HIV testing uptake in this population [45], and future iterations of these apps should attempt to address any gaps. On the other hand, the MyChoices app was substantially more effective in improving HIV testing among participants who had engaged in recent condomless sex with a male or transgender female partner, suggesting that this app is appropriately routing young sexual minority men who are at higher risk of HIV acquisition toward HIV testing.

These mHealth interventions and the format of the efficacy trial facilitated the continuation of this study during the COVID-19 pandemic. Even more, post hoc analyses suggested that individuals enrolled after the COVID-19 shutdown received the greatest benefit of both MyChoices and LYNX. It is possible that the availability of at-home HIV tests, condoms, and lubes through the apps was even more essential to isolate young sexual minority men during this time frame. Given the prediction that additional pandemics are increasingly more likely due to climate change and ease of global travel, this is an important finding. Additionally, it suggests that young sexual minority men who may be isolated for reasons other than pandemics (eg, living in rural areas) might further benefit from these apps and the resources that they provide [19].

MyChoices and LYNX demonstrated a small and nonsignificant impact on PrEP Initiation among young sexual minority men in this study, including among subgroups. Modeling studies have suggested that even a modest increase in PrEP Initiation could have meaningful impacts on HIV incidence [46]. Reasons for the small effect sizes (particularly in comparison to the larger impact on HIV testing) are not fully understood; however, barriers to PrEP Initiation are greater than those for HIV testing—it is less widely available, requires clinical appointments (vs HIV testing within community-based organizations or even at home), has more cost implications and is a medicine that must be taken regularly. In this study, for those who did not initiate PrEP by 6 months, the most common reasons reported include not having a perceived need (47.3%), costs (31.2%), worry about side effects (32.2%), and concern about long-term effects (30.1%). Future app enhancements should further address these barriers, including highlighting the potential benefits of PrEP, linkage to a PrEP navigator and telehealth services for PrEP, more diverse content on side effects, and linkage to programs that help defray the costs. Additionally, since these apps were developed, several different PrEP modalities have been approved, including event-driven oral medication and injectable PrEP that can be administered every 2 months. Future iterations of these apps will need to accommodate the opportunities and challenges for youths now that a menu of PrEP choices is available.

This large, randomized efficacy trial has a number of strengths, including methodological rigor, community engagement throughout the app development and study planning, and the diversity of participants in terms of race or ethnicity, geography, employment, and insurance status. However, there are also limitations to consider. First, due to the timing of this study, the enrollment of participants was slowed due to the COVID-19 pandemic. As a result, we had to extend our enrollment dates and shorten our follow-up for a subset of participants to 6 months instead of the preplanned 12 months. However, given

XSL•FO
RenderX

the enrollment criteria limited the sample to young sexual minority men at higher risk of HIV and who had not had a recent HIV test, according to the CDC, each participant would be recommended to an HIV test at least every 6 months, and thus this shorter follow-up time frame remains meaningful. Additionally, our time-to-event analysis, where all available data was used, demonstrated a similar pattern of results. This study also relied on self-reported HIV testing and PrEP Initiation, which increases the likelihood of social desirability bias and overestimation of the impact of the apps on these practices. This potential bias was reduced by the use of self-administered, web-based follow-up assessments. Additionally, the alternative would have required participants to show proof of HIV testing and PrEP Initiation, and we anticipated that the increased chances of missing data would have led to increased selection bias.

In spite of these limitations, the findings of this study are compelling and merit future exploration. In summary, future studies should enhance the functionalities related to PrEP to increase the apps' impact on this important outcome, as well as add content and functionalities related to emerging prevention technologies as they arise, including doxycycline prophylaxis for STI prevention. Additionally, given the 25% relative increase in HIV testing among young sexual minority men who used the MyChoices app and the potential for a large population-level effect, implementation studies that test and determine best practices for dissemination of this app to young sexual minority men across the United States are warranted [47].

Acknowledgments

The authors thank the study participants and community partners for their time and commitment. The authors would also like to acknowledge the work and contributions of Mr. Kenneth Coleman. This study was funded by the ATN for HIV Interventions, with primary funding by the Eunice Kennedy Shriver National Institute of Child Health and Human Development and supplemental funding from the National Institute of Mental Health, the National Institute on Drug Abuse, and the National Institute on Minority Health and Health Disparities (grant U01HD068040). This study also received support from the Emory Center for AIDS Research (P30AI050409), The Providence/Boston Center for AIDS Research (P30AI042853), and the MAC AIDS Fund. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or other funders.

Data Availability

The datasets generated during and/or analyzed during this study are available in the National Institute of Child Health and Human Development's Data and Specimen Hub data repository.

Authors' Contributions

KBB was responsible for conceptualization, methodology, formal analysis and investigation, original draft preparation, review and editing of the manuscript, funding acquisition, and supervision. KHM was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. HS was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. PKV was responsible for the formal analysis and review and editing of the manuscript. JHR was responsible for the review and editing of the manuscript and supervision. SB was responsible for the conceptualization, methodology, and review and editing of the manuscript. LAT was responsible for the formal analysis and review and editing of the manuscript. PS was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. LHW was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. AL was responsible for conceptualization, methodology, review and editing of the manuscript, and funding acquisition. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Comparison of theoretical models, key components and user experiences for the LYNX and MyChoices apps.
[PNG File , 153 KB - publichealth_v11i1e63428_app1.png]

Multimedia Appendix 2
Enrollment by month for the COMPARE (Comparison of Men's Prevention Apps to Research Efficacy) study, 2019-2022. Key: Red solid line=COVID-19 Shutdown (March 15, 2020); Green dashed line=re-opening of initial sites for enrollment (July 2020); Green solid line=all sites re-opened for enrollment (November 2020); Gray dashed line=expanded to national enrollment (October 2021).
[PNG File , 26 KB - publichealth_v11i1e63428_app2.png]

XSL•FO
RenderX

Multimedia Appendix 3
Time to first HIV test (a) and PrEP Initiation (b) in the COMPARE study, 2019-2022.
[PNG File , 74 KB - publichealth_v11i1e63428_app3.png]

Multimedia Appendix 4
CONSORT eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1134 KB - publichealth_v11i1e63428_app4.pdf]

References

1. Agwu A. Sexuality, sexual health, and sexually transmitted infections in adolescents and young adults. Top Antivir Med 2020;28(2):459-462 [FREE Full text] [Medline: 32886466]
2. HIV Surveillance Supplemental Report: Estimated HIV incidence and prevalence in the United States, 2017?2021. Centers for Disease Control and Prevention. 2023. URL: https://stacks.cdc.gov/view/cdc/149080 [accessed 2024-09-12]
3. DiNenno EA, Prejean J, Irwin K, Delaney KP, Bowles K, Martin T, et al. Recommendations for HIV screening of gay, bisexual, and other men who have sex with men—United States, 2017. MMWR Morb Mortal Wkly Rep 2017;66(31):830-832 [FREE Full text] [doi: 10.15585/mmwr.mm6631a3] [Medline: 28796758]
4. US Preventive Services Task Force, Barry MJ, Nicholson WK, Silverstein M, Chelmow D, Coker TR, et al. Preexposure prophylaxis to prevent acquisition of HIV: US Preventive Services Task Force recommendation statement. JAMA 2023;330(8):736-745. [doi: 10.1001/jama.2023.14461] [Medline: 37606666]
5. Gleason N, Serrano PA, Muñoz A, French A, Hosek S. Limited interaction targeted epidemiology of HIV in sexual and gender minority American adolescents and adults: feasibility of the keeping it LITE study. JMIR Form Res 2021 Nov 26;5(11):e30761 [FREE Full text] [doi: 10.2196/30761] [Medline: 34346403]
6. Whitfield THF, Parsons JT, Rendina HJ. Rates of pre-exposure prophylaxis use and discontinuation among a large U.S. national sample of sexual minority men and adolescents. Arch Sex Behav 2020;49(1):103-112 [FREE Full text] [doi: 10.1007/s10508-019-01602-z] [Medline: 31845148]
7. Kamitani E, Johnson WD, Wichser ME, Adegbite AH, Mullins MM, Sipe TA. Growth in proportion and disparities of HIV PrEP use among key populations identified in the United States national goals: systematic review and meta-analysis of published surveys. J Acquir Immune Defic Syndr 2020;84(4):379-386 [FREE Full text] [doi: 10.1097/QAI.0000000000002345] [Medline: 32205721]
8. Wang Y, Mitchell JW, Zhang C, Liu Y. Evidence and implication of interventions across various socioecological levels to address pre-exposure prophylaxis uptake and adherence among men who have sex with men in the United States: a systematic review. AIDS Res Ther 2022;19(1):28 [FREE Full text] [doi: 10.1186/s12981-022-00456-1] [Medline: 35754038]
9. Estimated HIV incidence and prevalence in the United States, 2015-2019. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-supplemental-report-vol-26-1.pdf [accessed 2024-12-03]
10. Moskowitz D, Moran K, Matson M, Alvarado-Avila A, Mustanski B. The PrEP cascade in a national cohort of adolescent men who have sex with men. J Acquir Immune Defic Syndr 2021;86(5):536-543. [doi: 10.1097/QAI.0000000000002613] [Medline: 33399311]
11. Estem KS, Catania J, Klausner JD. HIV self-testing: a review of current implementation and fidelity. Curr HIV/AIDS Rep 2016;13(2):107-115. [doi: 10.1007/s11904-016-0307-y] [Medline: 26879653]
12. Dellucci TV, Jones SS, Adhemar C, Feldstein Ewing SW, Lovejoy TI, Starks TJ. Correlates of HIV Testing across the lifespan—adolescence through later adulthood—among sexual minority men in the US who are not on PrEP. J Behav Med 2022;45(6):975-982 [FREE Full text] [doi: 10.1007/s10865-022-00341-2] [Medline: 35916966]
13. Fisher CB, Fried AL, Macapagal K, Mustanski B. Patient-provider communication barriers and facilitators to HIV and STI preventive services for adolescent MSM. AIDS Behav 2018;22(10):3417-3428 [FREE Full text] [doi: 10.1007/s10461-018-2081-x] [Medline: 29546468]
14. Fuzzell L, Fedesco HN, Alexander SC, Fortenberry JD, Shields CG. "I just think that doctors need to ask more questions": Sexual minority and majority adolescents' experiences talking about sexuality with healthcare providers. Patient Educ Couns 2016;99(9):1467-1472. [doi: 10.1016/j.pec.2016.06.004] [Medline: 27345252]
15. Babel RA, Wang P, Alessi EJ, Raymond HF, Wei C. Stigma, HIV risk, and access to HIV prevention and treatment services among men who have sex with men (MSM) in the United States: a scoping review. AIDS Behav 2021;25(11):3574-3604 [FREE Full text] [doi: 10.1007/s10461-021-03262-4] [Medline: 33866444]
16. Newcomb ME, Moran K, Li DH, Mustanski B. Demographic, regional, and political influences on the sexual health care experiences of adolescent sexual minority men. LGBT Health 2020;7(1):28-36 [FREE Full text] [doi: 10.1089/lgbt.2019.0122] [Medline: 31750760]
17. LaBelle M, Strong C, Tseng YC. mHealth strategies to promote uptake and adherence to PrEP: a systematic review. 2020 Presented at: 12th International Conference on Cross-Cultural Design (CCD 2020); July 19–24, 2020; Copenhagen, Denmark. [doi: 10.1007/978-3-030-49913-6_9]

XSL•FO
RenderX

18. Mulawa MI, Rosengren AL, Amico KR, Hightow-Weidman LB, Muessig KE. mHealth to reduce HIV-related stigma among youth in the United States: a scoping review. Mhealth 2021;7:35 [FREE Full text] [doi: 10.21037/mhealth-20-68] [Medline: 33898604]
19. Sullivan PS, Hightow-Weidman L. Mobile apps for HIV prevention: How do they contribute to our epidemic response for adolescents and young adults? Mhealth 2021;7:36 [FREE Full text] [doi: 10.21037/mhealth-20-71] [Medline: 33898605]
20. Teens and internet, device access fact sheet. Pew Research Center. 2024. URL: https://www.pewresearch.org/internet/fact-sheet/teens-and-internet-device-access-fact-sheet/ [accessed 2024-09-12]
21. Reyes E, Silvis J, Gandhi M, Shi Y, Greene M. Telehealth access and experiences of older adults with HIV during the COVID-19 pandemic: lessons for the future. J Am Geriatr Soc 2024;72(9):2816-2824. [doi: 10.1111/jgs.18970] [Medline: 38819631]
22. Biello KB, Hill-Rorie J, Valente PK, Futterman D, Sullivan PS, Hightow-Weidman L, et al. Development and evaluation of a mobile app designed to increase HIV testing and pre-exposure prophylaxis use among young men who have sex with men in the United States: open pilot trial. J Med Internet Res 2021;23(3):e25107 [FREE Full text] [doi: 10.2196/25107] [Medline: 33759792]
23. Biello KB, Daddario SR, Hill-Rorie J, Futterman D, Sullivan PS, Hightow-Weidman L, et al. Uptake and acceptability of MyChoices: results of a pilot RCT of a mobile app designed to increase HIV testing and PrEP initiation among young American MSM. AIDS Behav 2022;26(12):3981-3990. [doi: 10.1007/s10461-022-03724-3] [Medline: 35829971]
24. Biello KB, Marrow E, Mimiaga MJ, Sullivan P, Hightow-Weidman L, Mayer KH. A mobile-based app (MyChoices) to increase uptake of HIV testing and pre-exposure prophylaxis by young men who have sex with men: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2019;8(1):e10694 [FREE Full text] [doi: 10.2196/10694] [Medline: 30617042]
25. Liu A, Coleman K, Bojan K, Serrano PA, Oyedele T, Garcia A, et al. Developing a mobile app (LYNX) to support linkage to HIV/sexually transmitted infection testing and pre-exposure prophylaxis for young men who have sex with men: protocol for a randomized controlled trial. JMIR Res Protoc 2019;8(1):e10659 [FREE Full text] [doi: 10.2196/10659] [Medline: 30681964]
26. Lee S, Kapogiannis BG, Allison S. Improving the youth HIV prevention and care continuums: the Adolescent Medicine Trials Network for HIV/AIDS Interventions. JMIR Res Protoc 2019;8(3):e12050 [FREE Full text] [doi: 10.2196/12050] [Medline: 30912750]
27. Hightow-Weidman LB, Muessig K, Rosenberg E, Sanchez T, LeGrand S, Gravens L, et al. University of North Carolina/Emory Center for Innovative Technology (iTech) for addressing the HIV epidemic among adolescents and young adults in the United States: protocol and rationale for center development. JMIR Res Protoc 2018;7(8):e10365 [FREE Full text] [doi: 10.2196/10365] [Medline: 30076126]
28. PrEP brochure. Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/hiv/media/pdfs/2024/04/cdc-hiv-stsh-prep-brochure-english.pdf [accessed 2023-12-11]
29. HIV testing 101. Centers for Disease Control and Prevention. 2018. URL: https://www.cdc.gov/hiv/resourcelibrary/infosheets/cdc-hiv-consumer-info-sheet-testing-2023-508.pdf [accessed 2024-09-18]
30. Sullivan PS, Stephenson R, Hirshfield S, Mehta CC, Zahn R, Bauermeister JA, et al. Behavioral efficacy of a sexual health mobile app for men who have sex with men: randomized controlled trial of mobile messaging for men. J Med Internet Res 2022;24(2):e34574 [FREE Full text] [doi: 10.2196/34574] [Medline: 35025755]
31. Biello KB, Horvitz C, Mullin S, Mayer KH, Scott H, Coleman K, et al. HIV self-testing and STI self-collection via mobile apps: experiences from two pilot randomized controlled trials of young men who have sex with men. Mhealth 2021;7:26 [FREE Full text] [doi: 10.21037/mhealth-20-70] [Medline: 33898595]
32. Sullivan PS, Driggers R, Stekler JD, Siegler A, Goldenberg T, McDougal SJ, et al. Usability and acceptability of a mobile comprehensive HIV prevention app for men who have sex with men: a pilot study. JMIR Mhealth Uhealth 2017;5(3):e26 [FREE Full text] [doi: 10.2196/mhealth.7199] [Medline: 28279949]
33. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
34. Fisher JD, Fisher WA, Shuper PA. The information-motivation-behavioral skills model of HIV preventive behavior. In: DiClemente RJ, Crosby RA, Kegler MC, editors. Emerging theories in health promotion practice and research (2nd ed.) 2009:21-63.
35. Golub SA, Starbuck L, Fikslin R, Gamarel KE. Psychometric evaluation and predictive validity of an adapted Adherence Self-Efficacy Scale for PrEP. AIDS Behav 2023 Jan;27(1):218-230. [doi: 10.1007/s10461-022-03758-7] [Medline: 35809144]
36. Donenberg GR, Emerson E, Bryant FB, Wilson H, Weber-Shifrin E. Understanding AIDS-risk behavior among adolescents in psychiatric care: links to psychopathology and peer relationships. J Am Acad Child Adolesc Psychiatry 2001;40(6):642-653 [FREE Full text] [doi: 10.1097/00004583-200106000-00008] [Medline: 11392341]
37. Kurth AE, Martin DP, Golden MR, Weiss NS, Heagerty PJ, Spielberg F, et al. A comparison between audio computer-assisted self-interviews and clinician interviews for obtaining the sexual history. Sex Transm Dis 2004;31(12):719-726. [doi: 10.1097/01.olq.0000145855.36181.13] [Medline: 15608586]
38. Barnett V, Lewis T. Outliers in Statistical Data. 3rd edition. Chichester, United Kingdom: John Wiley & Sons; 1994.

XSL•FO
RenderX

39. Juusola JL, Brandeau ML, Owens DK, Bendavid E. The cost-effectiveness of preexposure prophylaxis for HIV prevention in the United States in men who have sex with men. Ann Intern Med 2012;156(8):541-550 [FREE Full text] [doi: 10.7326/0003-4819-156-8-201204170-00001] [Medline: 22508731]
40. Sanders GD, Bayoumi AM, Sundaram V, Bilir SP, Neukermans CP, Rydzak CE, et al. Cost-effectiveness of screening for HIV in the era of highly active antiretroviral therapy. N Engl J Med 2005;352(6):570-585. [doi: 10.1056/NEJMsa042657] [Medline: 15703422]
41. HIV compendium of best practices. Centers for Disease Control and Prevention. 2023. URL: https://www.cdc.gov/hivpartners/php/compendium/index.html [accessed 2023-06-12]
42. Marks G, Crepaz N, Janssen RS. Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. AIDS 2006;20(10):1447-1450. [doi: 10.1097/01.aids.0000233579.79714.8d] [Medline: 16791020]
43. Hightow-Weidman LB, Bauermeister JA. Engagement in mHealth behavioral interventions for HIV prevention and care: making sense of the metrics. Mhealth 2020;6:7 [FREE Full text] [doi: 10.21037/mhealth.2019.10.01] [Medline: 32190618]
44. Iribarren SJ, Akande TO, Kamp KJ, Barry D, Kader YG, Suelzer E. Effectiveness of mobile apps to promote health and manage disease: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2021;9(1):e21563 [FREE Full text] [doi: 10.2196/21563] [Medline: 33427672]
45. Nittas V, Daniore P, Chavez SJ, Wray TB. Challenges in implementing cultural adaptations of digital health interventions. Commun Med 2024;4(1):7 [FREE Full text] [doi: 10.1038/s43856-023-00426-2] [Medline: 38182750]
46. Sewak A, Lodi S, Li X, Shu D, Wen L, Mayer KH, et al. Causal effects of stochastic PrEP interventions on HIV incidence among men who have sex with men. Am J Epidemiol 2024;193(1):6-16 [FREE Full text] [doi: 10.1093/aje/kwad097] [Medline: 37073419]
47. Li DH, Macapagal K, Mongrella M, Saber R, Mustanski B. "Your package could not be delivered": The state of digital HIV intervention implementation in the US. Curr HIV/AIDS Rep 2024;21(3):152-167. [doi: 10.1007/s11904-024-00693-1] [Medline: 38502421]

Abbreviations

ATN: Adolescent Medicine Trials Network
CDC: Centers for Disease Control and Prevention
COMPARE: Comparison of Men's Prevention Apps to Research Efficacy
IMB: information-motivation-behavioral skills
IRB: institutional review board
mHealth: mobile health
PrEP: pre-exposure prophylaxis
RCT: randomized controlled trial
SOC: standard of care
STI: sexually transmitted infection

Edited by A Mavragani; submitted 19.06.24; peer-reviewed by S John, J Zucker, R Amico, AAS Sawitri; comments to author 13.08.24; revised version received 19.11.24; accepted 29.11.24; published 05.02.25.

Please cite as:
Biello KB, Mayer KH, Scott H, Valente PK, Hill-Rorie J, Buchbinder S, Ackah-Toffey L, Sullivan PS, Hightow-Weidman L, Liu AY The Effects of MyChoices and LYNX Mobile Apps on HIV Testing and Pre-Exposure Prophylaxis Use by Young US Sexual Minority Men: Results From a National Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e63428
URL: https://publichealth.jmir.org/2025/1/e63428
doi:10.2196/63428
PMID:

©Katie B Biello, Kenneth H Mayer, Hyman Scott, Pablo K Valente, Jonathan Hill-Rorie, Susan Buchbinder, Lucinda Ackah-Toffey, Patrick S Sullivan, Lisa Hightow-Weidman, Albert Y Liu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 05.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Use of Behavior Change Techniques in Digital HIV Prevention Programs for Adolescents and Young People: Systematic Review

Phoenix Kit-han Mo[1], PhD; Luyao Xie[1], PhD; Tsz Ching Lee[1], MPH; Angela Yuen Chun Li[1], BSCS

Center for Health Behaviours Research, JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Phoenix Kit-han Mo, PhD
Center for Health Behaviours Research
JC School of Public Health and Primary Care
Chinese University of Hong Kong
No. 30-32 Ngan Shing Street
Hong Kong
China (Hong Kong)
Phone: 852 22528765
Email: phoenix.mo@cuhk.edu.hk

Abstract

Background: HIV infections have caused severe public health and economic burdens to the world. Adolescents and young people continue to constitute a large proportion of newly diagnosed HIV cases. Digital health interventions have been increasingly used to prevent the rising HIV epidemic. Behavior change techniques (BCTs) are intervention components designed to modify the underlying processes that regulate behavior. The BCT taxonomy offers a systematic approach to identifying, extracting, and coding these components, providing valuable insights into effective intervention strategies. However, few reviews have comprehensively identified the use of BCTs in digital HIV interventions among adolescents and young people.

Objective: This study aimed to synthesize existing evidence on the commonly used BCTs in effective digital HIV prevention programs targeting adolescents and young people.

Methods: In total, 4 databases (PubMed, Embase, Cochrane Library, and APA PsycINFO) were searched, and studies from January 2008 to November 2024 were screened. Reference lists of relevant review studies were reviewed to identify any additional sources. Eligible randomized controlled trials with 1 of 3 HIV prevention outcomes (ie, HIV knowledge, condom-use self-efficacy, and condom use) were included. Basic study characteristics, intervention strategies, and study results were extracted and compared for data analysis. For the included interventions, BCTs were identified according to the BCT taxonomy proposed by Abraham and Michie in 2008, and the frequencies of BCTs used in these interventions were counted.

Results: Searches yielded 383 studies after duplicates were removed, with 34 (8.9%) publications finally included in this review. The most frequently used BCTs included *prompting intention formation* (34/34, 100%), *providing information about behavior-health link* (33/34, 97%), *providing information on consequences* (33/34, 97%), and *providing instruction* (33/34, 97%). Interventions with significant improvements in HIV knowledge (11/34, 32%) more frequently used BCTs with a provision nature, such as providing information about behavior-health link (11/11, 100%), information on consequences (11/11, 100%), encouragement (10/11, 91%), and instruction (10/11, 91%). Those with significant increases in condom-use self-efficacy (7/34, 20%) used BCTs toward initiating actions, such as prompts for intention formation (7/7, 100%), barrier identification (7/7, 100%), and practice (5/7, 71%). In addition, studies showing significant improvements in condom use (14/34, 41%) included BCTs focused not only on provision and initiation but also on behavioral management and maintenance, such as use follow-up prompts (5/14, 36%), relapse prevention (4/14, 29%), prompt self-monitoring of behavior (3/14, 21%), and prompt review of behavioral goals (3/14, 21%).

Conclusions: This is the first systematic review that examined the use of BCTs in digital HIV prevention interventions for adolescents and young adults. The identified BCTs offer important reference for developing more effective digital interventions, with implications for enhancing their HIV knowledge, condom-use self-efficacy, and condom use in youth.

(JMIR Public Health Surveill 2025;11:e59519) doi:10.2196/59519

XSL•FO
RenderX

KEYWORDS

HIV; prevention; adolescent; young adult; behavior change techniques

Introduction

HIV Infections Among Adolescents and Young People

HIV continues to be a global public health concern, with around 1.3 million new HIV infections occurring in 2022 worldwide [1]. Until 2022, there were around 39 million people living with HIV worldwide. HIV imposes a tremendous health and economic burden on the world. Since the epidemic began, more than 40 million lives have been lost due to HIV-related illnesses [2]. Lifelong antiretroviral treatments require long-term health care services and medications [3]. A study by Schackman et al [4] estimated that the discounted cost treating one patient with HIV, infected at the age of 35 years, for a lifetime in the United States is approximately US $326,500. By preventing 1 HIV infection, US $229,800 could be saved [4].

Despite a decreasing trend [5], adolescents and young people still occupy a significant proportion of new HIV infection. In 2020, adolescents and young people aged 13 years to 24 years accounted for one-fifth of newly diagnosed cases in the United States [6]. The number of new HIV infections among adolescents and young people from 2010 to 2019 were 170,000 and 460,000, respectively worldwide [5]. Meanwhile, the number of HIV-positive cases of young men who have sex with men (MSM) between the age of 13 years and 24 years has significantly increased from 2015 to 2019, indicating an increasing trend in HIV infection among young MSM in the United States [7]. Significant efforts remain to be made to achieve the global target of ending the HIV epidemic by 2030 [8].

Behavioral Interventions for HIV Prevention

The disproportionate burden of HIV among adolescents and young people is demonstrated to be primarily driven by HIV risky behaviors, such as unprotected sexual intercourse [9]. It was reported that the percentage of condomless sex among MSM in the United States increased from 46% in 2012 to 70.5% in 2017 [10]. A study suggested that only half (54.3%) of high school students in the United States reported having used condoms during their last sexual intercourse [11]. Hence, improving their condom-use behaviors is important. Previous literature has suggested that HIV knowledge and condom-use self-efficacy were 2 essential cognitive factors influencing condom use [12-15]. Some studies also showed that condom-use self-efficacy was a significant mediator between HIV knowledge and adolescents' condom use [16]. Nevertheless, the HIV knowledge and condom-use self-efficacy of youth remains suboptimal. It was found that less than half of adolescents possess comprehensive knowledge about HIV prevention [17]. Another study also documented that more than half of university students reported low levels of condom-use self-efficacy [18]. Therefore, to control and prevent HIV among adolescents and young people, interventions to promote their HIV-related knowledge, condom-use self-efficacy, and condom-use behaviors would be particularly important.

The Internet as a Useful Platform for Health Promotion

In recent decades, increasing use of the internet and smartphones to deliver health interventions has been observed in the literature [19,20]. Numerous advantages of web-based interventions have been identified, such as easy accessibility, availability, and assimilation to participants' everyday lives. The internet has become the most common platform for young people to search for health-related information [21].

Information on sexual health and sexually transmitted diseases is commonly sought through the internet, especially by ethnic minority groups and minority groups with sexual orientation differences [21,22]. Young MSM are more likely to look for HIV-related information digital compared to the general population [23]. The growing HIV epidemic and the proliferation of the internet use warrants the application of innovative strategies and IT for behavior change interventions. The internet provides an alternative platform for delivering health interventions to the minority population. Accumulated evidence has shown that digital interventions are effective in promoting HIV knowledge and condom use, with varying effect size [24-28].

Behavior Change Techniques in Health Interventions

One factor contributing to the varying effectiveness of digital HIV prevention can be attributed to the difference in the use of behavior change techniques (BCTs). BCTs are the observable, replicable, and irreducible intervention components that are designed to modify the processes that regulate behavior. In 2008, Abraham and Michie [29] developed a taxonomy of BCTs and defined 26 techniques with respective theoretical support, providing a standardized method of classifying intervention content [29,30]. Theoretical frameworks most frequently identified in the use of BCTs include the information-motivation-behavioral skills model, theory of planned behavior, theory of reasoned action, social cognitive theory, control theory, and operant conditioning [29,31]. Furthermore, as suggested by the Transtheoretical Model, individuals experience stages of precontemplation, contemplation, preparation, action, and maintenance to achieve and sustain a behavioral change [32]. People situated at different stages of change have various needs in driving them onto the next stage, including recognizing the pros and cons of behavior and acquiring the skills to act as well as managing the action to sustain changes. The 26 BCTs categorized in the taxonomy can be classified into various hierarchical levels fulfilling the diverse needs a person requires to produce a behavioral change under the Transtheoretical Model.

Evidence suggests that specific BCTs would be effective in improving a range of health behaviors [33-39]. For example, a systematic review of 19 studies on Type 2 diabetes prevention interventions published in 2020 [37] showed that effective interventions used in general 3.7 to 5.6 times more BCTs than noneffective interventions, and BCTs of social support, goal setting, feedback on behavior, and self-monitoring of outcomes

XSL•FO
RenderX

of behavior were identified in >90% of effective interventions. A review and meta-regression published in 2022 [39] identified BCTs used in smartphone app interventions, which found that BCTs of action planning and graded tasks had medium-positive associations with increasing physical activity in people with cardiovascular disease. Another review of interventions that promoted walking and cycling [36] found that prompting self-monitoring of behavior and prompting intention formation were the most frequently coded BCTs in the included studies. Other review studies on interventions in promoting health behavior change also support that those interventions that incorporated more BCTs tended to have larger effects compared to interventions that incorporated fewer techniques [35]. One review specifically targeting mobile health interventions [40] found that techniques of personalization, feedback and monitoring, and associations were most commonly used in mobile health interventions, and prompts and cues were the most common BCTs used in effective trials.

The BCT taxonomy provides a standardized and reliable way to extract and code the information from behavior change interventions and describe their components. It also offers a nomenclature system of intervention techniques that could be applied to health promotion programs aiming to drive behavioral changes [29]. The use of well-defined BCTs allows investigators to evaluate the effectiveness of an intervention more systematically and to conduct accurate replications of efficient interventions. However, past reviews on digital HIV prevention tended to focus on the interventions' effectiveness, rather than their techniques in leading behavioral changes [25,26,41-43]. There are limited reviews on the BCTs used in digital HIV interventions targeting adolescents and young people [44]. As digital HIV interventions may vary in their content and effectiveness, identifying the most effective BCTs used in different HIV prevention studies is important. This can help to delineate potential mechanisms between components and outcomes and inform the development of more targeted and efficient interventions in the future.

This Study

This study aimed to identify the commonly used BCTs in effective digital HIV prevention programs targeting adolescent and young people using a systematic review.

Methods

Search for Studies

A literature search was conducted in 4 databases, including PubMed, Embase, Cochrane Library, and APA PsycINFO, to identify publications written in English and published from January 2008 (the beginning of the popularization of smartphones) to November 2024 [45]. Search keywords were developed in the following 5 domains: setting, intervention, outcome, population, and study design, using Boolean logic. The search terms in the 4 databases are presented in Multimedia Appendix 1. In addition, bibliographies of existing systematic reviews were also reviewed for a comprehensive data search. No protocol or registration of this review was prepared. This review was performed following the PRISMA (Preferred Reporting Items for Systematic Reviews And Meta-Analyses) statement [46]. The PRISMA checklist is provided in Multimedia Appendix 2.

Eligibility Criteria

The eligible criteria for this systematic review, developed based on the population, intervention, comparison, outcomes, and study design framework, are provided in Multimedia Appendix 3. Briefly, randomized controlled trials (RCTs) of HIV prevention in adolescents and young people, implementing digital health interventions aiming to improve key outcomes, such as HIV knowledge, condom-use self-efficacy, and condom use, were considered for inclusion.

Study Selection

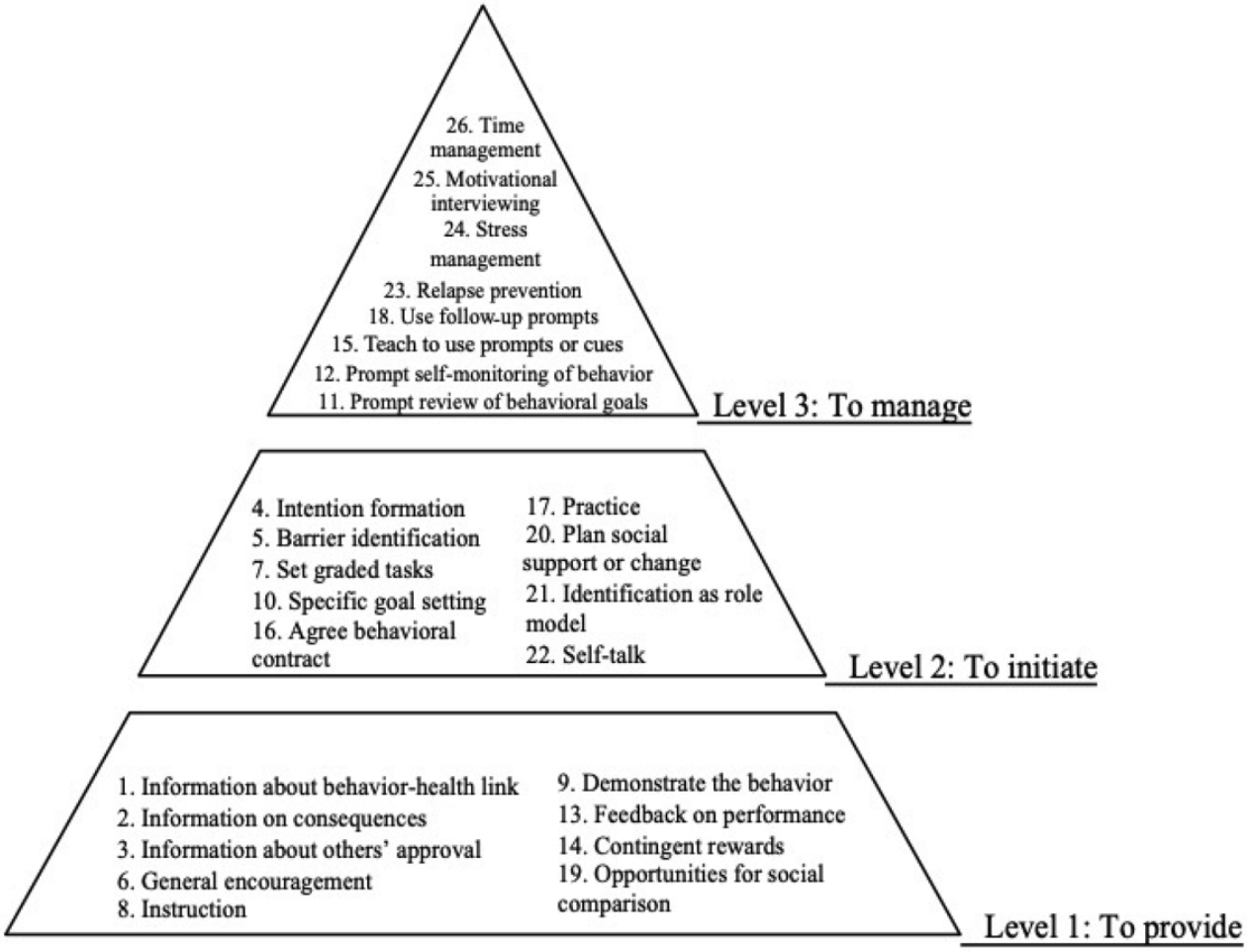
Articles were reviewed after a comprehensive search and screening based on the inclusion and exclusion criteria. The screening and selection process was conducted by 2 independent reviewers (TCL and LX) to identify eligible studies. Any disagreements were resolved through discussion and were referred to by a third reviewer (PK-hM) if necessary.

Data Extraction and Quality Assessment

For the included studies, data extraction was performed by 2 independent reviewers to collect the following information: studies' bibliographic information, participant demographics, intervention and control group, models or theoretical frameworks use, intervention intensity, duration, settings, completion rate, and results. The intervention strategies used in each study were reviewed and the BCTs used were identified based on the taxonomy of 26 BCTs developed by Abraham and Michie [29]. The BCTs were further classified into 3 hierarchical levels for individuals with different stages under the transtheoretical model (Figure 1 [29]). In particular, the level 1 BCTs have a provision nature and are used *to provide* information and resources to participants, mainly from a top-down approach. Examples include providing information, providing instruction, and providing general encouragement. Level 2 BCTs aim *to initiate* actions by prompting preparation and planning. Examples include prompt intention formation, barrier identification, specific goal setting, and practice. Level 3 BCTs are used in strategies *to manage* and hence maintain desired health behaviors for a sustainable change during the late action and maintenance stages. Examples are prompt self-monitoring of behaviors, relapse prevention, and time management.

XSL•FO
RenderX

Figure 1. Three hierarchical levels of 26 behavior change techniques developed by Abraham and Michie [29].



The methodological quality and risks of bias were assessed by the 2 reviewers, using the Joanna Briggs Institute critical appraisal tool for assessment of risk of bias for RCTs [47]. In total, 11 categories, including randomization, concealed allocation, blinding of participants and deliverers, and outcome measurements, were evaluated. An overall percentage score (0% to 100%) was calculated for each study representing the overall study quality. The 2 reviewers scored them independently, and discussion was made with a third investigator in case of any disagreements.

Data Analysis and Results Interpretation

Basic study characteristics of all RCTs included were summarized in table form. Characteristics included bibliographic information, settings, sample sizes, participant demographics, inclusion criteria, measurements used, and study duration. Results of RCTs were summarized and compared while significant improvements on predefined outcome indicators, as well as intervention completion rate were recorded.

Identified BCTs from intervention of selected RCTs were analyzed and compared based on study outcomes. Application of health models and theories in those studies were noted and matched with relative BCTs. Interventions that were developed or derived from previously published interventions were remarked for potential further investigations. The overall frequencies of BCTs used among all studies collected for this systematic review were counted and summarized. In addition, the frequencies of BCTs used in every significant outcome were counted to explore the commonly used BCTs in effective digital HIV prevention programs targeting adolescent and young audience.

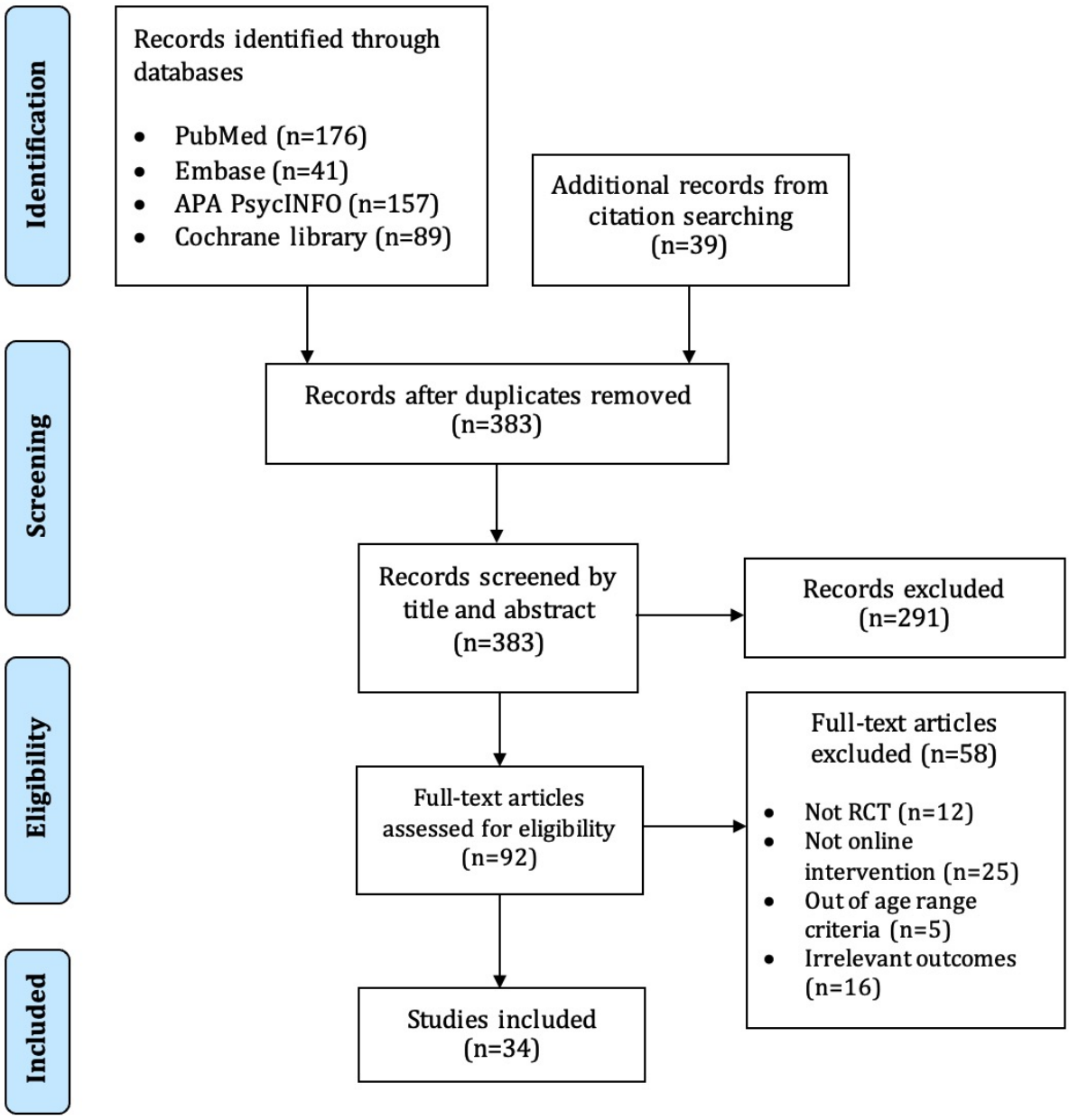
Results

Included Studies

On the basis of the literature search, 463 records were identified from 4 databases; an additional 39 records were identified through searching for the bibliographies of previous systematic reviews [43,48-50] on digital HIV interventions. Of these records, 119 (25.7%) duplicated records were removed, and the remaining 383 records were screened for eligibility. Following this, an additional 291 (63%) studies were removed after screening by title and abstract, as they did not meet the inclusion criteria. A further full-text review was conducted on the remaining 92 records, and 51 (11%) of them were excluded for the following reasons: non RCT design (n=12, 24%), irrelevant outcomes (n=16, 31%), not digital interventions (eg, in-person interventions, text messaging interventions, or telephone interventions; n=25, 49%), and out of age range (n=5, 10%). Finally, 34 studies were included in this systematic review. The selection process is shown in the PRISMA flowchart in Figure 2 [51].

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) Flowchart of the study screening process for this review. RCT: randomized controlled trial.



Study Characteristics

Table 1 provides an overview of the characteristics of the included studies. Among the 34 studies, 24 (71%) were conducted in the United States, 4 (12%) in African countries (Nigeria [52] and Uganda [53-55]), 5 (15%) in Asian countries (Hong Kong [56], China [57,58], Thailand [59], and Cambodia [60]), and 1 (3%) in Europe (Germany) [61]. Of the 34 studies, 14 (41%) included participants around early or middle adolescence (aged <18 years), and the other 20 (59%) included late adolescents and young people (average age: ≥18 years). In addition, 20 studies (20/34, 59%) targeted at those high-risk populations for HIV infection, involving subgroups such as MSM (n=12, 35%) [56,59,62-71], sexual minority men (N=1) [57], transgender women (n=1, 3%) [59], Black people (n=3, 9%) [67,68,70], low-income [72,73] or homeless individuals [74], refugees (n=1, 3%) [55], female entertainment workers (n=1, 3%) [60] and people with substance use [75,76]. The sample size in the 34 included studies ranged from 40 [62] to 1578 [77].

XSL•FO
RenderX

Table 1. Sample characteristics of included studies (N=34).

Author (year)	Country	Sample size, n	Recruitment	Participants			
				Targeted population (age range, if mentioned) and their mean age	Race and ethnicity	Male (%)	If were considered a high-risk population
Anand et al [59], (2020)	Thailand	76 (IG[a]:37, CG[b]:39)	Websites and testing clinics	Late adolescents and young people (aged ≥18 years), median 28 years	Thai nationals	75 (98.7; MSM[c] and 1 transgender women)	Yes (MSM or transgender women)
Bauermeister et al [70], (2019)	United States	180 (IG:120, CG:60)	Websites and social media	Late adolescents and young people (aged 18 to24 years), mean 21.7 years	White, Black, Asian, Middle Eastern, Native American, and mixed race	100	Yes (Black and MSM)
Brody et al [60], (2022)	Cambodia	1118 (IG: 435; CG: 683)	Entertainment venues	Female entertainment workers (aged 18 to 30 years), mean 24.6 years	Cambodian	0	Yes (female entertainment workers)
Bull et al [78], (2009)	United States	1565 (IG:799, CG: 766)	Websites	Late adolescents and young adulthood (aged 18-24 years), mean22 years	White, Asian, Black, Hispanic, multiracial, and others	42	No
Bull et al [77], (2012)	United States	1578 (IG:942, CG:636)	Websites and newspapers	Late adolescents and young adulthood (aged 15 to 25 years), mean 20 years	Hispanic, African American, Native American, Asian, Hawaiian, White, and others	Unknown	No
Christensen et al [69], (2013)	United States	935	Websites	Late adolescents and young adulthood (aged 18 to 24 years), mean: 21 years	Black, Latino, and White	100	Yes (MSM)
Cordova et al [79], (2020)	United States	50 (IG:25, CG:25)	Health clinics	Middle and late adolescents (aged 13 to 21 years), mean 19 years	White, Black, Native American, and others	18	No
Ezegbe et al [52], (2018)	Nigeria	80 (IG:40, CG: 40)	Schools	Middle adolescents (junior secondary school students), mean 14.78 years	Not mentioned	52.5	No
Fiellin et al [80], (2017)	United States	333 (IG: 166, CG: 167)	Schools	Early adolescents (aged 11 to 14 years), mean 12.9 years	White, Black, and others	53	No
Hightow-Weidman et al [68], (2012)	United States	50 (IG: 25, CG: 25)	Websites and community flyers	Late adolescents and young people (aged 18 to 30 years), mean 23.7 years	Black	100	Yes (Black and MSM)
Hightow-Weidman et al [67], (2019)	United States	474 (IG: 238, CG:236)	Websites and local flyers	Late adolescents and young people (aged 18 to 30 years), mean 24.3 years	Black	100	Yes (Black and MSM)
Jones et al [81], (2013)	United States	295 (IG:149, CG:146)	Testing clinics and community centers	Late adolescents and young people (aged 18to 29 years) mean 22 years	African American, Latino, and Caribbean	0	No
Klein and Card [82], (2011)	United States	178 (IG: 91, CG:87)	Websites and community flyers	Middle and late adolescents (aged 14 to 19 years), mean 15.8 years	Black	0	No
Lau et al [56], (2016)	Hong Kong SAR	402 (IG1: 133, IG2:133, CG:136)	Websites and gay venues	Late adolescents and young people (aged ≥18 years), mean or median not mentioned	Asian	100	Yes (MSM)

XSL•FO
RenderX

Author (year)	Country	Sample size, n	Recruitment	Participants			
				Targeted population (age range, if mentioned) and their mean age	Race and ethnicity	Male (%)	If were considered a high-risk population
Levy et al [61], (2021)	Germany	149 (IG:72, CG:77)	Universities	Late adolescents and young people (university students, aged 18 to 30 years), mean 21.48 years	Not mentioned	15.4	No
Logie et al [55], (2023)	Uganda	450 (IG1: 152; IG2: 157; CG: 141)	Refugee settlements	Youth (aged 16 to 24 years), mean 20 years	Not mentioned	50.7	Yes (refugee)
Marsch et al [75], (2011)	United States	56 (IG:28, CG:28)	Clinics	Early and middle adolescents (aged 12 to 18), mean: 16 years	African American, White, and others	70	Yes (substance abuse)
Marsch et al [76], (2015)	United States	141 (IG:69, CG:72)	Clinics	Early and middle adolescents (aged 12 to 18 years), mean: 16 years	Black, White, multiracial, and others	77	Yes (substance abuse)
McCrimmon et al [83], (2024)	United States	457 (IG: 233, CG: 224)	High School	9th and 11th graders Adolescents (no age restriction reported), mean 15.1 years	White, Black, Hispanic, and others	40	No
Mustanski et al [66], (2013)	United States	102 (IG:50, CG:52)	Testing clinics	Late adolescents and young adulthood (aged 18 to 24 years), mean 21 years	Latino, White, African American, and others	100	Yes (MSM)
Mustanski et al [65], (2018)	United States	901 (IG 445, CG:456)	Testing clinics	Late adolescents and young people (aged 18 to 29 years)	White, Black, Latino, and others	100	Yes (MSM)
Nelson et al [64], (2022)	United States	154 (IG:77, CG:77)	Websites and social media	Middle adolescents (age 14-17), mean: 16	White, Black, Latino, mixed race, and others	100	Yes (MSM)
Newcomb et al [71], (2023)	United States	400 (IG: 200, CG: 200)	Social media and dating or hookup apps	Young men (aged 18 to 29 years) mean 28.3 years	White, Hispanic or Latinx, Black, Asian, Native Hawaiian or Pacific Islander, Native American or Alaskan Native, multiracial, and others	100	Yes (MSM)
Peskin et al [73], (2015)	United States	1374 (IG:768, CG:606)	Schools	Middle adolescents (Eighth grade students), mean 14.3 years	African American, Hispanic, and others	41	Yes (Most students were economically disadvantaged)
Santa Maria et al [74], (2021)	United States	97 (IG: 48, CG:49)	Community centers	Late adolescents and young adulthood (aged 18 to 25 years), mean: 21.2 years	White, Black, Latino, others, and multiracial	57	Yes (homeless)
Schnall et al [63], (2022)	United States	763 (IG: 382, CG:381)	Websites and community centers	Middle adolescents (aged 13 to 18 years), mean 16.2 years	Black, Latino, and White	100	Yes (same-sex attracted adolescent males)

XSL•FO
RenderX

Author (year)	Country	Sample size, n	Recruitment	Participants			
				Targeted population (age range, if mentioned) and their mean age	Race and ethnicity	Male (%)	If were considered a high-risk population
Swendeman et al [84], (2024)	United States	1482 (IG1: 266, IG2: 279, IG3: 295, CG: 642)	Community centers and clinics	Adolescents (aged 12 to 24 years), mean: 21.1 years	Black or African American, Latinx or Hispanic, White (non-Hispanic), Asian or Pacific Islander, and Other or mixed background	93	No
Widman et al [72], (2018)	United States	222 (IG:107, CG:115)	Schools	Middle adolescents (10th grade students), mean: 15.2 years	White, black, and Hispanic	0	Yes (low-income)
Widman et al [85], (2020)	United States	226 (IG: 113, CG:113)	Schools	Middle adolescents (10th and 11th grade students), mean 16.3 years	White, Black, Hispanic, and others	42	No
Wray et al [62], (2019)	United States	40 (IG:20, CG: 20)	Gay-oriented smartphone dating apps	Late adolescents and young people (aged 18to 30 years), mean 28 years	White, black, Asian, and multiracial	100	Yes (MSM with heavy drinking problems)
Ybarra et al [53], (2013)	Uganda	366 (IG: 183, CG: 183)	Schools	Middle adolescents (Secondary students, aged >12 years), mean 16.1 years	Not mentioned	84	No
Ybarra et al [54], (2015)	Uganda	366 (IG: 183, CG: 183)	Schools	Middle adolescents (Secondary students, aged >12 years), mean: 16.1 years	Not mentioned	84	No
Yi et al [57], (2024)	China	120 (IG: 60, CG: 60)	Social media for sexual minority men and community centers	Young men (aged 16 to 30 years), mean 23.2 years	Han, Hui, Miao, and Tujia	100	Yes (Sexual minority men)
Zhang et al [58], (2024)	China	247 (IG: 125, CG: 122)	Factories	Men (>18 years), mean 36.3 years	Han, minority group	100	No

[a]IG: intervention group.

[b]CG: control group.

[c]MSM: men who have sex with men.

Regarding recruitment methods of the included RCTs, 8 (23%) recruited participants from educational institutions (n=7, 12% from secondary schools [52-54,72,73,80,83,85] and n=1, 3% from a university [61]), 7 (12%) recruited through a combination of digital advertisements and community-based organizations [56,57,59,63,67,68,77,82], 7 (12%) recruited from community centers or treatment clinics [65,66,74-76,79,81,84], 1 (3%) recruited from factories [58], 1 (3%) recruited from refugee settlements [64], 1 (3%) recruited from entertainment venues [60], and 5 (15%) recruited entirely digital via websites, social media, or dating apps [62,64,69-71,78]. Most of the studies focused on participants who were HIV-negative (Table 1).

Intervention Characteristics

An overview of the characteristics of interventions included is listed in Tables 2 and 3. Most of the studies (29/34, 85%) were 2-arm intervention design with 1 experimental group and 1 control group. In total, 4 studies [53-56] had 2 experimental groups together with 1 control group. Only 1 study [84] had 3 experiential groups together with 1 control group. In addition, 6 studies [56,58,62,72,83,85] were one-off interventions; the remaining 28 studies had multiple sessions, and their intervention duration ranged from 2 weeks [76] to 24 months [84].

XSL•FO
RenderX

Table 2. Intervention conditions and effectiveness of included studies (N=34).

Author (year)	Modality type	On-site session	Time of follow-up (from baseline)	Completion rate after intervention (%)
Anand et al [59], (2020)	Website	No	6 mo and 12 mo	88.2
Bauerm-eister et al [70], (2019)	Mobile app	No	1 mo, 2 mo, and 3 mo	79.4
Brody et al [60], (2022)	SMS text messaging	No	6 mo and 12 mo	34.7
Bull et al [78], (2009)	Website	No	2 mo (internet sample) or 3 mo (clinic sample)	Not mentioned
Bull et al [77], (2012)	Social media	No	After intervention (8 wk from baseline), 6 mo	69
Christensen et al [69], (2013)	Video game	No	After intervention and 3-mo after intervention	98.5
Cordova et al [79], (2020)	Mobile app	Yes	After intervention, 30 d	98
Ezegbe et al [52], (2018)	Social media	No	8 wk and follow-up after intervention	100
Fiellin et al [80], (2017)	Video game	Yes	6 wk, 3 mo, 6 mo, and 12 mo	83
Hightow-Weidman et al [68], (2012)	Website	No	1 and 3 mo	90
Hightow-Weidman et al [67], (2019)	Website	No	3, 6, and 12 mo	85.2
Jones et al [81], (2013)	Website	No	3 and 6 mo	80.7
Klein and Card [82], (2011)	Multimedia software	Yes	3 mo	91
Lau et al [56], (2016)	Website	No	After intervention, 1 mo, and 3 mo	77.4
Levy et al [61], (2021)	Website	No	1 mo	69.1
Logie et al [55], (2023)	SMS	No	8 mo, 12 mo, and 16 mo	71.1
Marsch et al [75], (2011)	Multimedia software	Yes	After intervention, 1-mo after test, and 3-mo after test	93
Marsch et al [76], (2015)	Multimedia software	Yes	After intervention (within 2 wk)	100
McCrimmon et al [83], (2024)	Website	Yes	After intervention	100
Mustanski et al [66], (2013)	Website	No	After intervention, 6, and 12 wk	88.2
Mustanski et al [65], (2018)	Website	No	3, 6, and 12 mo	81.2
Nelson et al [64], (2022)	Website	No	3 wk and 15 wk	89
Newcomb et al [71], (2023)	Videoconference	No	3 mo, 6 mo, 9 mo, and 12 mo	58.3
Peskin et al [73], (2015)	Video game	Yes	1 y	89.2
Santa Maria et al [74], (2021)	Mobile app	No	6 wk	Not mentioned
Schnall et al [63], (2022)	Mobile app	No	3 mo, 6, and 9 mo	81.7
Swendeman et al [84], (2024)	SMS, social media, and videoconference	No	4 mo, 8 mo, 12 mo, 16 mo, 20 mo, and 24 mo	42.2
Widman et al [72], (2018)	Website	No	After intervention, 4 mo	99.5
Widman et al [85], (2020)	Website	No	After intervention	91.6
Wray et al [62], (2019)	Mobile app	Yes	1, 2, and 3 mo	97.5
Ybarra et al [53], (2013)	Website	Yes	3 and 6 mo	94.8
Ybarra et al [54], (2015)	Website	Yes	3 and 6 mo	94.8
Yi et al [57], (2024)	Website	No	4 mo and 8 mo	95
Zhang et al [58], (2024)	Videos	No	6 mo and 12 mo	89.1

XSL•FO
RenderX

Table 3. Conditions and outcomes of included interventions.

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
Anand et al [59], (2020)	2	An digital video-based discussion and interaction platform, Vialogues, with monthly HIV education sessions. The website was integrated with counseling and testing support services, as well as video-uploading and messaging functions for counselors and participants to discuss.	N/A[a]	Private clinic-based HIV counseling and testing.	HIV knowledge, percentage of condom use for anal intercourse, condom-use self-efficacy	Unchanged HIV knowledge level at baseline and after intervention in both arms; significant increase in percentage of condom use in the intervention group; no significant changes in condom-use self-efficacy among intervention group participants.
Bauerm-eister et al [70], (2019)	2	A mobile app, myDEX, equipped with 6 sessions on sexual health information and skills tailored for YMSM[b]. Each session consisted of a core message, a deep discussion regarding the topics, as well as an interactive activity. Features like a diary, role-playing, quizzes, and practice opportunities were included.	N/A	An Information-only static site with HIV prevention contents split into 6 sessions.	Condomless anal sex events and condom-use self-efficacy	Intervention participants were less likely to engage in condomless anal sex than participants from the control group; control participants were more likely to forego condoms than participants from the intervention group
Brody et al [60], (2022)	2	Participants could choose to receive text or voices messages. 180 messages comprising of 10 health themes will be sent, including gender-based violence, cervical cancer, vaginal health, contraception, hygiene, pregnancy termination, general health information, alcohol use, HIV and STI[c] transmission and prevention, pregnancy and miscarriage. A message was sent 2 times for 10 wk, and the message from each topic area was repeated every 10 wk for 60 wk. An option to be contacted by outreach workers for free was provided after each message.	N/A	Existing standard care.	HIV test, STI testing when symptoms arise, contraceptive usage, and condom use with nonpaying and paying partner	No significant differences in HIV-preventive behaviors, including HIV or STI testing, contraceptive usage and condom use with nonpaying and paying partner.
Bull et al [78], (2009)	2	An interactive website, Keep It Real, was developed. It contained 5 modules with 60-90 s long role model stories focusing on condom use and HIV risks. Quizzes were included to test understanding. A booster session at 1-mo follow-up showed the same messages again.	N/A	A computer kiosk at a clinical setting with text-based information on condom use and HIV risks. Quizzes were included to test understanding. A booster session at 1-mo follow-up showed the same messages again.	Proportion of protected sex acts by condoms and condom-use self-efficacy	No significant effects on protected sex acts and condom-use self-efficacy.

XSL•FO
RenderX

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
Bull et al [77], (2012)	2	A Facebook page, "Just/US," was developed and ran for 2 mo. Eight topics focusing on communication skills and condom-use efficacy were delivered through daily contents in different formats, eg, videos, quizzes, games, and discussions.	N/A	A Facebook page called "18-24 News" showing interesting news tailored for people aged 18-24 from 6 PM to midnight. Sexual health contents were avoided.	Condom use at last sex and proportion of protected sex acts by condoms	Condom use at last sex remained stable in intervention group but decreased in the control group.; proportion of protected sex acts remained stable in intervention group but significantly decreased in the control group.
Christensen et al [69], (2013)	2	A simulation video game, SOLVE, was used to simulate common obstacles encountered by YMSM at different scenarios. Depending on participants' decisions, different choice points would be awarded. Sexuality affirmations were included to reduce shame for MSM[d].	N/A	No intervention.	Number of unprotected anal intercourse UAI[e]	Significant indirect decrease in number of UAI caused by shame reduction in the intervention group
Cordova et al [79], (2020)	2	In addition to the clinic's usual services, a mobile app, Storytelling 4 Empowerment, was developed to deliver HIV, tobacco, and substance use knowledge through 3 modules. A clinician version with participants' history and risk assessments was also developed to facilitate communication during in-person counseling sessions.	N/A	The clinic's usual services on sexual health as well as a printed copy of Storytelling 4 Empowerment tobacco module contents.	Sexual risk self-efficacy, sexual risk prevention knowledge, sexual risk behaviors	Small insignificant effect sizes in sexual risk self-efficacy change scores within the intervention group.; larger gains of sexual risk prevention knowledge, including HIV, in intervention group compared to the control group; greater, but insignificant reduction in condomless sex among intervention participants than control participants.
Ezegbe et al [52], (2018)	2	An 8-wk long, twice a week, digital storytelling intervention. The RED-Story intervention contained digital sessions with videos from social media platforms, as well as offline therapist discussion sessions with video narration activities.	N/A	No intervention	HIV knowledge	Significant improvement in HIV knowledge among schoolchildren from intervention group than control group
Fiellin et al [80], (2017)	2	A 2D role-playing adventure game, PlayForward, with interactive features where players can make their own decisions in different social contexts to learn about HIV prevention; 2 one-h sessions per week for 6 wk.	N/A	Twelve irrelevant video games, eg, Angry Birds, with the same number and length of sessions as the experiment group.	Sexual health knowledge	Significant increase in sexual health knowledge in intervention group compared to control group

XSL•FO
RenderX

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
Hightow-Weidman et al [68], (2012)	2	The website, HealthMpowerment.org, contained 7 main sessions with interactive features, including resources, quizzes of different levels, personalized health, and sex journals. Included. four 30-min weekly sessions	NA	A list of 5 Web sites providing general HIV information.	Number of protected sex acts, condom-use self-efficacy, and HIV knowledge	Significant increase in self-reported condom use among participants from both groups at 1-mo follow-up; no significant difference in condom-use self-efficacy and HIV knowledge between the intervention and control group.
Hightow-Weidman et al [67], (2019)	2	Website, HealthMpowerment.org, was revised with a knowledge library on HIV prevention with interactive features, including discussion forums, personal space for uploading information and digital physician consultation.	NA	Website with a smaller-scale knowledge library.	Number of acts of CAI[f]	Significant decline in rate of CAI in the intervention group compared to the control group.
Jones et al [81], (2013)	2	A 12-wk soap opera video series, *Love, Sex, and Choices* was produced and streamed to smartphones. Each episode was 15-20 min long with contents describing different high-risk situations. Principles of HIV risk reduction were modeled by main characters.	N/A	A 12-wk text messaging intervention containing HIV risk reduction written messages delivered by smartphones.	Percentage unprotected vaginal and anal sex acts	Significant decline in unprotected sex acts in both the intervention and control group; intervention group had a lower percentage of unprotected sex acts compared to the control group at follow-up, given they were the same at baseline.
Klein and Card [82], (2011)	2	A computer-delivered multimedia software developed from an in-person intervention SiHLE. It contained 2 1-h sessions, centering on HIV prevention and sexual decision-making skills delivered through multimedia channels, including games, quizzes, and videos. All sessions were narrated by teenage female health educators.	N/A	A general health education session with 2 laptop-delivered videos about healthy diet and nutritious eating. The 2 videos were 65 min long in total.	HIV knowledge, condom-use self-efficacy, and percentage of condom protected vaginal intercourse acts	Significant improvement in HIV knowledge in both intervention and control group, with higher improvement among intervention group; significant increase in condom-use self-efficacy among nonsexually active participants in the intervention group; significant increase in percentage condom use among intervention participants, compared to no change in condom use in the control group.
Lau et al [56], (2016)	3	A website with videos based on STD[g]-related cognitions. Contents of the videos included HIV and syphilis information as well as means of prevention.	In addition to videos based on STD-related cognitions from intervention 1, a fear-arousing video based on STD-related emotions was added in intervention 2, highlighting consequences of social loss due to STDs.	A website with text HIV-related facts and information.	Number of UAI	No significant association between 3 treatment groups and prevalence of UAI was discovered; significant within-group reduction in UAI in all 3 groups; significant reduction in UAI with casual sex partners in intervention group 2 with fear appeal.

XSL•FO
RenderX

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
Levy et al [61], (2021)	2	Psychological inoculation was used to design an interactive website with HIV knowledge as well as 10 challenging sentences regarding condom use barriers and social pressure. Participants had to refute relative statements. Exaggerated versions would be provided for weak refutation.	N/A	A static website with HIV knowledge and 10 additional true or false questions on condom use to control for the experimental group.	Frequency of condom use	Frequencies of condom use increased in control group but not experiment group.
Logie et al [55], (2023)	3	Weekly SMS will be sent to ask how they are doing and have 2 choices of response, “fine” or “not fine.” Participants who responded “not fine” or did not respond will be followed-up by a peer navigator within 2 d and a week, respectively [86].	HIV self-testing kit, along with instruction, condoms, lubricant, information pamphlet, and referral information for confirmatory testing	Information about HIV testing and services at local clinics and a leaflet on HIV-preventive strategies [86].	HIV testing frequency, knowledge of HIV status, condom-use self-efficacy, and consistent condom use [86]	Significant increase in HIV testing, HIV status in knowledge in both intervention groups comparing to control group; significant improvement in condom usage self-efficacy comparing intervention groups and the control group; not significant difference in consistent condom use comparing the intervention groups and control group.
Marsch et al [75], (2011)	2	On top of the traditional intervention (the same as the control measures), a self-directed web-based program which contained 19 modules on HIV risk reduction and relative risks brought by drug abuse would be added. The system provided customized selection of modules based on participants’ risk assessment survey results.	N/A	A 1-h individual or small group educator-delivered sessions by trained HIV prevention specialists on basic HIV-related information as well as drug-related risks. In addition, a 15 min long video was played.	HIV prevention knowledge, condom-use self-efficacy	Significant increase in HIV prevention knowledge in both groups, while participants from intervention group showed larger increases compared to that of control group; significant increase in skills to correctly use condoms in both groups without intergroup differences.
Marsch et al [76], (2015)	2	A self-directed web-based program, the Therapeutic Education System, which contained 19 modules on HIV risk reduction and relative risks brought by drug abuse. The system provided customized selection of modules based on participants’ risk assessment survey results.	N/A	Two 1-h individual or small group educator-delivered sessions by trained HIV prevention specialists on basic HIV-related information as well as drug-related risks. In addition, a 20 min long video was played.	HIV knowledge and condom-use self-efficacy	Increase in HIV knowledge in both groups without significant intergroup differences; increase in condom-use self-efficacy in both groups without significant intergroup differences.

XSL•FO
RenderX

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
McCrimmon et al [83], (2024)	2	A self-directed web-based program adapted Health Education and Relationship Training, which included 6 modules on safe sex motivation, sexual communication skills, HIV and STI knowledge, safer sexual self-efficacy, and sexual norms and attitudes. Interactive activities include games, animated characters and quizzes [87].	N/A	A 45-min attention matched web-based program that targets growth mindsets was delivered [87].	Condom use intention, attitude and norm, HIV and STI knowledge, and safe sex communication self-efficacy	Higher improvement in attitude toward condom and HIV and STI knowledge in intervention group compared to control group; no significant difference in condom use intention and norm; safe sex communication self-efficacy
Mustanski et al [66], (2013)	2	Three digital sessions with 7 interactive modules in the Keep It Up! Intervention website (around 2 h in total). Different contents, including videos, animations, and games, were implemented. Contents were tailored to YMSM. A 6-wk booster session was applied.	N/A	Same number of modules and sessions as the intervention group with didactic noninteractive texts and images on HIV-related information. No tailored contents for YMSM.	Number of unprotected anal sex acts and HIV knowledge	A small decrease in rate of unprotected anal sex acts in intervention group, but an increase in the control group; a large increase in HIV knowledge in both groups, with no significant difference between groups.
Mustanski et al [65], (2018)	2	Three digital sessions with 7 interactive modules in the Keep It Up! Intervention website (around 1 h in total). Different contents, including videos, animations, and games, were implemented. Contents were tailored to YMSM. 3- and 6-mo booster sessions were applied.	N/A	Same number of modules and sessions as the intervention group with static texts and images on HIV-related information. No tailored contents for YMSM.	Incident STI, reporting number of casual CAS[h] acts and CAS partners	Significant reduction in STI and reported CAS for intervention group versus control group; no difference between intervention and control groups regarding incident HIV.
Nelson et al [64], (2022)	2	Interactive website with 9 modules focusing on 4 main topics, including male anatomy, HIV or STI-related knowledge, general sexual health information, and pornography truths. The website included adventure games, questions and answers videos, and illustrations. Participants needed to unlock the first 3 modules before given the access to the remaining 6 modules.	N/A	Website of Centers for Disease Control and Prevention, as well as Web sites of national HIV and STI testing resources.	HIV knowledge and condom knowledge	Insignificant differences of HIV knowledge scores among intervention and control group; no difference on condom knowledge scores among intervention and control group.

XSL•FO
RenderX

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
Newcomb et al [71], (2023)	2	Five sessions regarding sexual health, HIV risk and relationship functioning were conducted. Three videos on these topics were sent to participants each time, then 3 videoconference group sessions were implemented to build skill. Two individual coaching video conference for each coupon were delivered for skill implementation.	N/A	A single 90-min session. HIV-negative or unknown couple: Testing Together protocol (HIV testing and education). HIV-Positive: Medication and Risk Reduction Counseling (explore motivators and barriers to antiretroviral adherence and plan for it). Serodiscordant: receive both of the above.	STI, CAS, HIV testing, and PrEP[i] use	Significantly lower STI positive cases, CAS frequency in intervention group; no difference on HIV testing and PrEP use among intervention and control group.
Peskin et al [73], (2015)	2	An interactive computer-based video game, IYG-Tech, with thirteen 45 min long modules on sexual health. Animated scenarios, videos, quizzes, fact sheets, role-playing and discussion platforms were built-in activities aiming to equip students with necessary sexual health life-skills.	N/A	Standard textbook health education.	Number of condomless sex, knowledge of STIs, and condom-use self-efficacy	No significant differences in the number of condomless sex among the 2 groups; intervention students had greater knowledge about STIs and higher condom-use self-efficacy than control students in the 1-y follow-up.
Santa Maria et al [74], (2021)	2	A mobile app, MY-RID, was delivered to homeless young adults. Participants firstly set a goal of behavioral change for HIV prevention, then answered daily assessment questions (EMA[k]) followed by tailored messages regarding their specific goal and real-time risk predictors.	N/A	The same mobile app and assessment items were applied, but the focus was on general health behaviors. General messages, instead of tailored ones, would be received.	Frequency of condomless sex acts	No significant intervention effect was observed on frequency of condomless sex acts.
Schnall et al [63], (2022)	2	A mobile app, MyPEEPS Mobile, was developed based on a group-based intervention. It focused on the stories of 4 YMSM (peeps) and aimed to build knowledge, self-awareness, and self-efficacy in sexual risk reduction through 21 mobile activities, completed throughout a 3-mo period.	N/A	Delayed intervention receiving MyPEEPS mobile at 9-mo follow-up after data collection.	Number of condomless sex acts	Significant reduction in condomless anal sex acts among intervention group participants, compared to the control group participants; most pronounced and long-lasting effect on Black participants, compared to participants of other races.

XSL•FO
RenderX

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
Swendeman et al [84], (2024)	4	IG[j]1: coaching and peer support and SMS text messaging; coaching: weekly telehealth strength-based coaching, 30-min for first 2 mo then 5-20 min following, covering assessment, linkage to services, goal setting, problem solving, and cognitive and behavioral skills training; peer Support: up to 16 wk to discuss and support on a moderated digital private forum; seed discussion topics were posted twice a week [88]; SMS text messaging: Daily SMS sent for 24 mo, covering sexual health, physical and mental health, substance use, and medication reminders if applicable, and a weekly self-monitoring survey on HIV or STI symptoms, risky sexual behaviors, and medication adherence.	IG2: coaching and SMS text messaging; IG3: peer support and SMS text messaging	SMS text messaging only.	PrEP use and adherence, consistent condom use, and PEP prescription and adherence	All intervention groups had higher increase in PrEP use compared to the control group; intervention group1 had sustained an increase in PrEP use compared to all other groups; no significant differences between PrEP adherence, PrEP prescription and adherence.
Widman et al [72], (2018)	2	A 45 min long digital intervention, HEART for Girls, with 5 modules targeting motivations, HIV knowledge, social norms, safer sex self-efficacy, and sexual communication skills. Sexual assertiveness skills were emphasized throughout all modules.	NA	A 45 min long attention-matching digital intervention, Growing Minds, with 5 modules on academic and social growth mindsets.	HIV knowledge, condom-use self-efficacy, and condom use at last sexual intercourse	Significant increase in HIV knowledge in the intervention group compared to the control group; significant increase in condom-use self-efficacy in the intervention group compared to the control group; better, but insignificant condom use among participants in the intervention group compared to the control group.
Widman et al [85], (2020)	2	A 45 min long digital intervention, HEART for Teens, with 5 modules targeting motivations, HIV knowledge, social norms, safer sex self-efficacy, and sexual communication skills. Sexual assertiveness skills were emphasized throughout all modules. More male characters were added compared to HEART for Girls.	N/A	A 45 min long attention-matching digital intervention, Growing Minds, with 5 modules on academic and social growth mindsets.	HIV knowledge and condom-use self-efficacy	Higher HIV knowledge and condom-use self-efficacy in intervention group compared to control group; similar intervention effects across 2 boys and girls.

XSL•FO
RenderX

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
Wray et al [62], (2019)	2	In addition to the standard posttest counseling services, a mobile app, Game Plan, was provided after HIV testing. Reflective exercises to prompt self-talk and goal setting functions were incorporated into the app to motivate behavioral change. This is a one-off intervention with follow-up surveys for 3 mo.	N/A	Standard post-HIV testing counseling with referral services.	Number of CAS events	Insignificant effect on number of CAS events in participants from both the intervention and control group.
Ybarra et al [53], (2013)	3	The CyberSenga website had 5 1-h intervention modules focusing on information about HIV and healthy sexual relationship, as well as skills needed for decision-making and condom use. Four versions of materials were developed tailoring different needs according to participants' genders and sexual experiences.	Same as the first intervention group but with a booster module at 4 mo after intervention.	Treatment as usual: no extra interventions other than HIV programs currently being provided at schools.	Reporting number of unprotected sex acts in the past 3 mo	No significant differences at the rates of unprotected sex in intervention and control groups in the past 3 mo.
Ybarra et al [54], (2015)	3	The CyberSenga website had 5 1-h intervention modules focusing on information about HIV and healthy sexual relationship, as well as skills needed for decision-making and condom use. Four versions of materials were developed tailoring different needs according to participants' genders and sexual experiences.	Same as the first intervention group but with a booster module at 4 mo after intervention.	Treatment as usual: no extra interventions other than HIV programs currently being provided at schools.	HIV prevention-related information	Intervention group participants answered greater percentages of questions correctly, compared with control group participants; improvement of HIV-related knowledge was the greatest among intervention participants with the booster module.
Yi et al [57], (2024)	2	LGBTQ[k]-affirmative cognitive behavioral therapy: 10 modules covering goal setting, LGBTQ-related stress and its reactions, automatic thoughts, emotion avoidance, emotion-driven behaviors, and behavioral skill training and experiment. Each module contained videos, 4-5 pages of psychoeducation, exercises and homework.	N/A	Weekly self-monitoring survey for 10 wk.	HIV-transmission-risk behaviors, HIV or syphilis result, condom self-efficacy, perceived condom use benefits, and mental and behavioral health	No significant differences in HIV-transmission-risk behaviors and social cognitive mechanisms between intervention group and control group; no positive HIV or syphilis cases reported; intervention group showed greater improvements in depression and anxiety in follow-ups.

XSL•FO
RenderX

Author (year)	Number of groups	Intervention condition 1	Intervention condition 2	Control condition	Outcomes measured	Results
Zhang et al [58], (2024)	2	Different web-based videos based on participants' behavioral assessment results: (1) video for those at lower risk: promote capacity to refuse peers' invite to have sex with female sex workers or nonregular female sex partners. Content includes high-risk nature, sex partner may be asymptomatic, severe interpersonal consequences, transmit to HIV or STI to stable partners; (2) video for people at high risk to promote their HIV testing: the benefits of taking up HIV testing and the procedures of free HIV testing and counseling at local Centers for Disease Control	N/A	Basic HIV-related knowledge.	Sexual intercourse with female sex partners and sex workers, respectively, condomless sex with female sex partner and sex worker, respectively, and uptake of HIV testing	Significant difference was observed in 6-mo follow-up in lower frequency of sexual intercourse with nonregular female sex partners of intervention group, compared with the control group; no significant differences in all other outcomes or time points.

[a]N/A: not applicable.

[b]YMSM: young men who have sex with men.

[c]STI: sexually transmitted infection.

[d]MSM: men who have sex with men.

[e]UAI: unprotected anal intercourse.

[f]CAI: condomless anal intercourse.

[g]STD: sexually transmitted disease.

[h]CAS: condom anal sex.

[i]PrEP: pre-exposure prophylaxis.

[j]IG: intervention group.

[k]LGBTQ: lesbian, gay, bisexual, transgender, and queer.

For the modality type of the 34 digital interventions, 20 (59%) were web-based interventions (eg, conducted through interactive websites [53,54,56,57,59,61,64-68,72,78,81,83,85] or multimedia software [75,76,82]); 7 (21%) were mobile-based interventions (eg, through mobile apps [62,63,70,74,79] or SMS text messaging [55,60]); 5 (15%) were video-related (eg, video games [69,73,80], educational videos [58], and videoconference [71]); 2 (6%) interventions were conducted through social media [52,77], and 1 (3%) through a combination of videoconference, SMS text messaging, and social media [84].

In total, 9 (27%) of the 34 interventions were delivered at in-person sessions under the supervision of research staff [53,54,62,73,75,76,79,80,82], while the remaining interventions (n=25, 73%) were operated and delivered remotely. All 3 interventions using multimedia software required on-site delivery with a relatively small sample size (n=56 to 178) [75,76,82], hinting the limitation of such modality type in large-scale community implementation.

Theories and BCTs

Table 4 provides an overview of theories and BCTs used in all studies. Of the 34 studies, 23 (68%) interventions were theory-based. Cognitive-based theories, including Integrated Behavioral Model [67,68], Information- Motivation- Behavior Skills Model [65,66], Social Cognitive Theory [69,73,82,84], behavioral theory on Dyadic Health model [71], as well as emotion-based approaches, such as Rational Emotive Behavior Therapy theory [55] and Psychological Wise theory [83], were used in these studies. In addition, over half of studies (21/34, 62%) included the BCTs across the 3 hierarchical levels under the Transtheoretical Model (Figure 1), and the other 13 studies included BCTs solely at level 1 and level 2.

XSL•FO
RenderX

Table 4. Theoretical frameworks and behavior change techniques (BCTs) used in included interventions (N=34).

Author (year)	Theory used	BCTs used[a]
Anand et al [59], (2020)	—[b]	• Level 1: 1, 2, 3, 6, 8, 9, 13 • Level 2: 4, 17, 20 • Level 3: 18, 23
Bauermeister et al [70], (2019)	Cognitive- emotional decision-making framework	• Level 1: 1, 2, 6, 8 • Level 2: 4, 5, 17, 20 • Level 3: 12
Brody et al [60], (2022)	—	• Level 1: 1, 2, 8 • Level 2: 4, 20
Bull et al [78], (2009)	Theory of reasoned action and planned behavior and social cognitive theory	• Level 1: 1, 2, 3, 6, 19 • Level 2: 4 • Level 3: 18
Bull et al [77], (2012)	—	• Level 1: 1, 3, 6, 8, 19 • Level 2: 4, 5, 20
Christensen et al [69], (2013)	Theory of planned behavior and social cognitive theory	• Level 1: 1, 2, 6, 8, 13 • Level 2: 4, 5, 7, 17, 22 • Level 3: 24
Cordova et al [79], (2020)	Ecodevelopment theory and empowerment theory	• Level 1: 1, 2, 6, 8 • Level 2: 4, 20
Ezegbe et al [52], (2018)	Rational emotive behavior therapy theory	• Level 1: 1, 2, 6, 19 • Level 2: 4, 17, 21, 22
Fiellin et al [80], (2017)	—	• Level 1: 1, 2, 3, 6, 8, 19 • Level 2: 4, 5, 17
Hightow-Weidman et al [68], (2012)	Integrated behavioral model	• Level 1: 1, 2, 3, 6, 8, 13, 19 • Level 2: 4, 5, 7, 20 • Level 3: 12
Hightow-Weidman et al [67], (2019)	Integrated behavioral model	• Level 1: 1, 2, 3, 6, 8, 13, 19 • Level 2: 4, 5, 7, 20, 21 • Level 3: 12, 18
Jones et al [81], (2013)	Sex script theory and theory of power as knowing participation in change	• Level 1: 6, 8, 19 • Level 2: 4, 5, 22
Klein and Card [82], (2011)	Social cognitive theory and the theory of gender and power	• Level 1: 1, 2, 3, 6, 8, 9, 13, 19 • Level 2: 4, 5, 20
Lau et al [56], (2016)	—	• Level 1: 1, 2, 3, 6, 8 • Level 2: 4, 20
Levy et al [61], (2021)	Cognitive dissonance	• Level 1: 1, 13 • Level 2: 4, 7
Logie et al [55], (2023)	—	• Level 1: 1, 2, 8, 9 • Level 2: 4, 5, 17, 20 • Level 3: 18
Marsch et al, [75], (2011)	—	• Level 1: 1, 2, 6, 8 • Level 2: 4, 5, 20
Marsch et al [76], (2015)	—	• Level 1: 1, 2, 6, 8 • Level 2: 4, 5, 20

XSL•FO
RenderX

Author (year)	Theory used	BCTs used[a]
McCrimmon et al [83], (2024)	Psychological wise theory	• Level 1: 1, 2, 3, 6, 8, 9, 13, 19 • Level 2: 4, 5, 10, 17, 22 • Level 3: 12, 25
Mustanski et al [66], (2013)	Information-motivation-behavior-skills model	• Level 1: 1, 2, 3, 6, 8, 9, 13, 19 • Level 2: 4, 5, 10, 17 • Level 3: 11, 18, 23
Mustanski et al [65], (2018)	Information-motivation-behavior-skills model	• Level 1: 1, 2, 3, 6, 8, 9, 13, 19 • Level 2: 4, 5, 10, 17 • Level 3: 11, 18, 23
Nelson et al [64], (2022)	—	• Level 1: 1, 2, 6, 8, 9, 13, 19 • Level 2: 4, 5, 7
Newcomb et al [71], (2023)	Dyadic health model	• Level 1: 1, 2, 3, 6, 8, 9, 13, 19 • Level 2: 4, 5, 16, 17, 20, 21 • Level 3: 18, 23, 24, 25
Peskin et al [73], (2015)	Social cognitive behavioral theory	• Level 1: 1, 2, 3, 6, 8, 19 • Level 2: 4, 5, 17, 22 • Level 3: 11, 12
Santa Maria et al [74], (2021)	Information-motivation-behavior model	• Level 1: 1, 2, 6, 8, 13 • Level 2: 4, 17 • Level 3: 11, 23
Schnall et al [63], (2022)	Social learning theory	• Level 1: 1, 2, 3, 6, 8, 19 • Level 2: 4, 5, 22 • Level 3: 11, 18, 25
Swendeman et al [84], (2024)	Social cognitive theory	• Level 1: 1, 2, 3, 6, 13, 14, 19 • Level 2: 4, 5, 7, 10, 17, 20, 21 • Level 3: 11, 18, 25
Widman et al [72], (2018)	Reasoned action model and fuzzy trace theory	• Level 1: 1, 2, 3, 6, 8, 9, 13, 19 • Level 2: 4, 5, 17 • Level 3: 12, 25
Widman et al [85], (2020)	Reasoned action model and fuzzy trace theory	• Level 1: 1, 2, 3, 6, 8, 9, 13, 19 • Level 2: 4, 5, 17 • Level 3: 12, 25
Wray et al [62], (2019)	—	• Level 1: 1, 2, 6, 13, 19 • Level 2: 4, 5, 10, 17, 20, 22 • Level 3: 11, 12, 25
Ybarra et al [53], (2013)	Information-motivation-behavior model	• Level 1: 1, 2, 3, 6, 8, 9 • Level 2: 4 • Level 3: 18
Ybarra et al [54], (2015)	Information-motivation-behavior model	• Level 1: 1, 2, 3, 6, 8, 9 • Level 2: 4 • Level 3: 18
Yi et al [57], (2024)	Cognitive behavioral therapy	• Level 1: 1, 2, 8, 9, 13 • Level 2: 4, 5, 10, 17, 20, 22 • Level 3: 11, 12, 18, 24
Zhang et al [58], (2024)	—	• Level 1: 1, 2, 3, 8, 9 • Level 2: 4, 20

[a]Level 1 to level 3 are 3 hierarchical levels of BCTs based on BCT taxonomy developed by Abraham and Michie [29]: level 1, to provide information and resources; level 2, to initiate action; level 3, to manage and maintain health behaviors for a sustainable change.

XSL•FO
RenderX

[b]Not available.

The frequency of each BCT used in all studies is shown in Table 5. The most commonly used BCTs were prompt intention formation (34/34, 100%), provide information about behavior-health link (33/34, 97%), provide information on consequences (31/34, 91%), provide instruction (30/34, 88%), provide general encouragement (29/34, 85%), prompt barrier identification (23/34, 68%), provide information about others' approval (20/34, 59%), and provide opportunities for social comparison (19/34, 56%). Unused BCTs include provide contingent rewards, teach to use prompts or cues, and time management. The proportion of BCTs used under each level was 67%, 40%, and 17% for level 1, level 2, and level 3, respectively. All interventions included BCTs from level 1 and level 2. However, BCTs from level 3 were more frequently used in studies targeting older adolescents (aged ≥18 years) [55,57,62,65-71,74,84] compared to younger adolescents (aged <18 years) [53,54,63,72,73,83,85]. Specifically, level 3 BCTs in studies targeting younger adolescents included follow-up prompts [53-55,57,63,71,84], prompt self-monitoring of behavior [57,72,73,83,85], motivational interviewing [63,71,72,83-85], and prompt review of behavioral goals [57,63,73,84], while those targeting older adolescents included prompt review of behavioral goals [62,65,66,74], prompt self-monitoring of behavior [62,67,68,70], relapse prevention [65,66,71,74], stress management [57,69,71], follow-up prompts [65-67], and motivational interviewing [62,80].

XSL•FO
RenderX

Table 5. Frequencies of behavior change techniques (BCTs) used across all studies and studies with different improved outcomes.

BCT number	BCT	All studies (N=34), n (%)	Studies with improved HIV-related knowledge (n=11), n (%)	Studies with improved condom-use self-efficacy (n=7), n (%)	Studies with improved condom use (n=14), n (%)	Combined average percentage of improved outcomes (%)
4	Prompt intention formation	34 (100)	11 (100)	7 (100)	14 (100)	100
1	Provide information about behavior-health link	33 (97)	11 (100)	7 (100)	13 (93)	97
2	Provide information on consequences	33 (97)	11 (100)	7 (100)	13 (93)	97
8	Provide instruction	33 (97)	10 (91)	7 (100)	14 (100)	97
6	Provide general encouragement	30 (88)	10 (91)	6 (86)	13 (93)	91
5	Prompt barrier identification	27 (79)	8 (73)	7 (10)	11 (79)	81
3	Provide information about others' approval	23 (68)	7 (64)	4 (57)	11 (79)	69
19	Provide opportunities for social comparison	21 (62)	7 (64)	4 (57)	9 (64)	63
20	Plan social support or social change	20 (59)	5 (45)	4 (57)	9 (64)	56
17	Prompt practice	19 (56)	7 (64)	5 (71)	6 (43)	56
13	Provide feedback on performance	16 (47)	4 (36)	3 (43)	8 (57)	48
9	Model or demonstrate the behavior	15 (44)	6 (55)	4 (57)	6 (43)	50
18	Use follow-up prompts	12 (35)	3 (27)	1 (14)	5 (36)	28
12	Prompt self-monitoring of behavior	11 (32)	3 (27)	4 (57)	3 (21)	40
11	Prompt review of behavioral goals	8 (24)	1 (9)	1 (14)	3 (21)	15
22	Prompt self-talk	8 (24)	3 (27)	1 (14)	3 (21)	22
25	Motivational interviewing	8 (24)	3 (27)	2 (29)	2 (14)	22
7	Set graded tasks	6 (18)	0 (0)	0 (0)	3 (21)	9
10	Prompt specific goal setting	6 (18)	1 (9)	0 (0)	2 (17)	9
23	Relapse prevention	5 (15)	0 (0)	0 (0)	4 (29)	12
21	Prompt identification as role model	4 (12)	1 (9)	0 (0)	2 (14)	9
24	Stress management	3 (9)	0 (0)	0 (0)	2 (14)	6
16	Agree behavioral contract	1 (3)	0 (0)	0 (0)	1 (7)	3
14	Provide contingent rewards	0 (0)	0 (0)	0 (0)	0 (0)	0
15	Teach to use prompts or cues	0 (0)	0 (0)	0 (0)	0 (0)	0
26	Time management	0 (0)	0 (0)	0 (0)	0 (0)	0

Results and BCTs per Improved Outcome

Three key outcome measurements were identified among the included 34 studies, including (1) HIV knowledge (n=16, 47%), (2) condom-use self-efficacy (n=15, 44%), and (3) frequency of condom use (n=26, 76%). Over half (20/34, 59%) of the interventions had an intervention completion rate higher than or equal to 85%. It is estimated that the high completion rate might be a result of school-based and clinic-based recruitment. Follow-up measurements were commonly carried out immediately after intervention, at 3 and 6-month follow-ups. Variations existed depending on the nature and duration of studies (Tables 2 and 3).

In total, 11 (69%) of the 16 studies that included HIV knowledge as an outcome reported a substantial improvement in HIV

XSL•FO
RenderX

knowledge. In addition, 7 (47%) of 15 studies that included condom-use self-efficacy as an outcome reported a statistically significant increase in condom-use self-efficacy. Moreover, 14 (54%) of 26 studies that included condom use as an outcome reported a statistically significant increase in condom use frequency. Table 5 summarizes the frequencies of BCTs linked to each outcome indicators in studies with significant improvements. Prompt intention formation (100%), provide information about behavioral-health link (97%), provide information on consequences (97%), provide instruction (97%), and provide general encouragement (91%) were the most commonly used BCTs among all the 3 improved target outcomes. Prompt practice (5/7, 71%) and prompt self-monitoring of behavior (4/7, 57%) were more frequently used in studies to drive improvement in condom-use self-efficacy compared to other outcome indicators, while a more diverse and even use of the 26 BCTs across the 3 levels was observed in studies with an increase in condom use frequency.

Quality Assessment

The quality assessment of included studies based on the Joanna Briggs Institute tool is provided in Table 6. All 34 studies included were randomized. Among these, 10 (29%) studies reported the use of concealed allocation [52,57-59,63,65,66,71,81,84], and 31 (91%) studies demonstrated baseline similarity between intervention and control groups [52-54,56-77,79-83,85]. Blinding was less frequently reported, with 5 (15%) studies documenting participant blinding [52,63,65,66,84], 5 (15%) studies documenting outcome assessor blinding [52,55,56,60,82], and 7 (21%) studies documenting blinding of deliverers [60,63,65,66,72,81,84]. Most studies (31/34, 91%) ensured that the groups were identical except for the intervention [52-54,56,57,59-70,72-85]. All studies reported identical outcome measurements, and 32 (94%) studies used reliable outcome measurements, with 2 (6%) studies lacking this information [55,60]. Follow-up analysis was conducted in 23 (68%) studies [52-54,56,59,60,62,65,67,69,70,72-81,84,86], and 31 (91%) studies performed analysis according to correct groupings [52-56,58-71,73-82,84,85]. The quality score of the included studies ranged from 45% [55,83] to 91% [52,65], with 24 (71%) studies receiving the score <70%, and 10 (29%) studies receiving a quality score >70%.

XSL•FO
RenderX

Table 6. Risk of bias assessment for included studies using the Joanna Briggs Institute appraisal tool (N=34).

Author (year)	Quality (%)	Randomized	Concealed allocation	Similar to baseline	Participants blind to assignment	Deliverers blind to assignment	Identical groups except intervention of interest	Outcome assessors blind to assignment	Identical outcome measurement	Reliable outcome measurement	Analysis of follow-up	Analysis according to correct groupings
Anand et al [59], (2020)	73	Yes	Yes	Yes	No	No	Yes	Unclear	Yes	Yes	Yes	Yes
Bauermeister et al [70], (2019)	64	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Brody et al [60], (2022)	73	Yes	Unclear	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes
Bull et al [78], (2009)	55	Yes	Unclear	No	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Bull et al [77], (2012)	64	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Christensen et al. [69], (2013)	64	Yes	No	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Cordova et al [79], (2020)	64	Yes	Unclear	Yes	Unclear	No	Yes	Unclear	Yes	Yes	Yes	Yes
Ezegbe et al [52], (2018)	91	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Fiellin et al [80], (2017)	64	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Hightow-Weidman et al [68], (2012)	55	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	No	Yes
Hightow-Weidman et al [67], (2019)	64	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Jones et al [81], (2013)	82	Yes	Yes	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Klein and Card [82], (2011)	64	Yes	Unclear	Yes	Unclear	No	Yes	Yes	Yes	Yes	No	Yes
Lau et al [56], (2016)	73	Yes	No	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Levy et al [61], (2021)	55	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	No	Yes
Logie et al [55], (2023)	45	Yes	No	No	Unclear	Unclear	No	Yes	Yes	No	Yes	Yes

XSL•FO
RenderX

Author (year)	Quality (%)	Randomized	Concealed allocation	Similar to baseline	Participants blind to assignment	Deliverers blind to assignment	Identical groups except intervention of interest	Outcome assessors blind to assignment	Identical outcome measurement	Reliable outcome measurement	Analysis of follow-up	Analysis according to correct groupings
Marsch et al [75] (2011)	64	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Marsch et al [76], (2015)	64	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
McCrimmon et al [83], (2024)	45	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	No	No
Mustanski et al [66], (2013)	82	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	Yes
Mustanski et al [65], (2018)	91	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Nelson et al [64], (2022)	55	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	No	Yes
Newcomb et al [71], (2023)	55	Yes	Yes	Yes	Unclear	Unclear	No	Unclear	Yes	Yes	No	Yes
Peskin et al [73], (2015)	64	Yes	Unclear	Yes	Unclear	No	Yes	Unclear	Yes	Yes	Yes	Yes
Santa Maria et al [74], (2021)	55	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Schnall et al [63], (2022)	82	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	No	Yes
Swendeman et al [84], (2024)	82	Yes	Yes	No	Yes	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Widman et al [72], (2018)	73	Yes	Unclear	Yes	Unclear	Yes	Yes	Unclear	Yes	Yes	Yes	Yes
Widman et al [85], (2020)	64	Yes	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Yes	Yes	Yes	Yes
Wray et al [62], (2019)	64	Yes	No	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes
Ybarra et al [53], (2013)	64	Yes	No	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes
Ybarra et al [54], (2015)	64	Yes	No	Yes	No	No	Yes	No	Yes	Yes	Yes	Yes
Yi et al [57], (2024)	55	Yes	Yes	Yes	No	No	Yes	Unclear	Yes	Yes	No	No

XSL•FO
RenderX

Author (year)	Quality (%)	Randomized	Concealed allocation	Similar to baseline	Participants blind to assignment	Deliverers blind to assignment	Identical groups except intervention of interest	Outcome assessors blind to assignment	Identical outcome measurement	Reliable outcome measurement	Analysis of follow-up	Analysis according to correct groupings
Zhang et al [58], (2024)	64	Yes	Yes	Yes	No	No	Yes	Unclear	Yes	Yes	No	Yes

Discussion

Use of BCTs and Theories

From our understanding, this study is the first review on the use of BCTs in digital HIV interventions among adolescents and young people. The findings showed that among the 3 levels of BCTs, BCTs at level 1 are the most frequently used with an average frequency of 67%. BCTs at level 1 help to deepen participants' understanding of HIV and the role of condoms in preventing sexually transmitted diseases. By enhancing and consolidating important knowledge, higher motivation will be generated to encourage change. This is especially useful for people at the precontemplation and contemplation stages. In contrast, level 2 BCTs mark the starting point of turning internal ideas into actual behavioral change, or in other words mobilizing participants to enter the preparation and action stage from precontemplation stage, according to the transtheoretical model. Level 2 BCTs had an average frequency of 40% in this study, which is less commonly used among the 34 studies compared to level 1 BCTs. Among all, level 3 BCTs were at the highest hierarchical level as they are the most complicated to implement though potentially most impactful in causing long-term change. Level 3 BCTs were the least used among all 3 levels with an average frequency of only 17%. The use of BCTs from 3 hierarchical levels in this systematic review matches with the rule of thumb for at-risk populations under the transtheoretical model. Specifically, level 1 BCTs were commonly used for the largest group of recipients from precontemplation (40%) and contemplation (40%) stage, while level 2 BCTs were less frequently used for populations at the preparation stage (20%) [32]. It is also important to note that the majority of the included studies (23/34, 68%) were observed to be supported by or constructed based on behavior change models or theories. No substantial difference was observed in the number of BCTs used under each theory in this review. In other words, the impact of theories on the pattern of BCTs used seems to be minimal.

BCTs Linked to Improved Outcomes

Findings of the study found that level 1 BCTs with a provision nature were the most commonly used in bringing significant improvement in HIV knowledge, indicating that *providing* information and instructions can be an effective method to increase knowledge of participants. Some level 2 BCTs were used for consolidating information, such as prompt practice. Level 3 BCTs were rarely used regarding this outcome, indicating that higher level BCTs concerning larger degrees of self-initiated participation were not considered necessary in enhancing HIV knowledge. In contrast, a higher percentage use of level 2 BCTs was observed in studies with significant increase in condom-use self-efficacy. As suggested by the self-efficacy theory [89,90], self-efficacy can be determined by past experience (performance accomplishment), modeling by others (vicarious experience), coaching and evaluative feedback (social persuasion), as well as physiological and emotional states. Level 2 BCTs, such as prompt practice, plan social support, and motivational interviewing, were more frequently used together with level 1 BCTs, including providing feedback on performance and model the behavior, to provide and initiate experience, and modeling and coaching feedback to participants to effectively enhance one's condom-use self-efficacy.

Finally, studies with a significant increase in condom use frequencies used level 3 BCTs the most among all 3 outcome indicators, suggesting the importance of managing an action for a sustained behavioral change. Involved level 3 BCTs include use follow-up prompts, relapse prevention, and stress management, which are rarely or have never been used in achieving the other 2 outcome indicators. In addition, provide information about others' approval, provide feedback on performance (level 1), and plan social support (level 2) were used more frequently in increasing condom use, showing the importance of perceived social norms on driving actual behavior change. Findings of this paper match with those from relevant systematic reviews of digital health interventions of other health conditions, which show that provide encouragement, provide information about behavior-health links, provide instructions, and provide information on consequences were the most frequently used BCTs [91,92].

Advantages of Digital Platforms for Delivering BCTs in HIV Prevention

Digital platforms offer unique features that allow the application of certain BCTs in HIV prevention interventions. Web-based interventions can reach a wider audience, including hard-to-reach populations (eg, MSM), and offer cost-effective solutions compared to traditional face-to-face interventions [93]. In addition, the interactive and engaging nature of digital platforms helps to deliver BCTs more effectively, making them a powerful tool in HIV prevention efforts. For instance, customizable avatars allow participants to choose characters that share their race and cultural background, which makes it easier for participants to empathize and imagine the cases in reality. This customizable function facilitates the use of BCTs, such as prompt barrier identification and providing opportunities for social comparisons. Anonymous interactions on forums and video-uploading platforms enable participants to express genuine concerns and thoughts freely. These interactions incorporate BCTs like providing information about others' approval and facilitating planning of social change [94]. Such anonymity ensures that the youth feel safe and supported, which is particularly important for sensitive topics like HIV prevention.

XSL•FO
RenderX

Tailor-made content based on risk assessment results and preset goals ensures that messages are accurate and relatable to targeted populations, which can enhance the relevance and impact of the interventions. The process of information through the central route is deepened with the use of follow-up prompts and providing feedback on performance [95], ensuring that participants fully understand and engage with the information provided. Finally, unlocking graded modules, for example, quizzes or chapters with elevating difficulty, helps participants gradually build a solid foundation of knowledge and behavioral change. BCTs like setting graded tasks and providing performance feedback are effectively integrated into digital interventions, ensuring continuous engagement and learning in HIV prevention.

Public Health Contribution and Implications

The youth demographic is particularly vulnerable due to the high level of engagement in risky sex and barriers to accessing traditional health care services. The focus on digital interventions is significant in this underserved and hard-to-reach population. The internet offers a wider yet more complex platform for health education. Its distinctive attributes, such as rapidly evolving trends, a multitude of information sources, and diverse user demographics, can make it challenging to effectively use digital health interventions. To navigate this evolving landscape of health education, it is beneficial to examine successful strategies that incorporate commonly used BCTs in digital interventions.

The findings of this review offer valuable guidance and insights for the design and implementation of future HIV prevention programs aimed at adolescents and young people. The BCTs identified in this review serve as a valuable reference for further studies, enabling researchers to select appropriate BCTs for intervention development and improvement. The most used BCTs identified in this review included prompting intention formation, providing information about behavior-health links, providing information on consequences, providing instructions, and providing general encouragement, which is important in fostering awareness and motivation for behavior change. Interventions showing improvements in HIV knowledge mostly included BCTs with a provision focus, such as providing information on consequences, behavior-health links, and encouragement. This suggests that educational components are essential for building foundational knowledge and promoting initial engagement and readiness for behavior change. In this review, BCTs aimed at initiating actions (eg, prompts for intention formation and barrier identification) were commonly observed in studies reporting a significant increase in condom-use self-efficacy, which provides evidence for bridging the gap between awareness and action. While to improve actual condom use, BCTs focused on managing and sustaining behavior (eg, use of follow-up prompts and relapse prevention) may need to be incorporated into the interventions. Considering the diverse range of outcomes to be achieved, hierarchical levels of BCTs should also be taken into account to design tailored and effective interventions. It is recommended that public health practitioners use behavior change theories as a foundation, using BCTs at different levels to select intervention activities and deliver pertinent messages for maximizing program effectiveness.

In addition, future studies can delve into identifying intrinsic and extrinsic motivational techniques within the current BCT taxonomy. Exploring the effectiveness of emotive techniques in driving behavior change in digital interventions could be another promising avenue for investigation. Moreover, the digital nature of the included interventions may allow for the scalability and adaptability of these findings to other topics beyond HIV prevention, particularly those that share similar characteristics. By leveraging the knowledge gained from this review, we can enhance the effectiveness of HIV prevention programs for adolescents and young people and facilitate border implementation.

Limitations

There were some limitations of the study. First, some BCTs may not be easily observable and thus might have limited the identification of the exact BCTs used. Four BCTs, namely provide contingent rewards, teach to use prompts or cues, agree behavioral contract, and time management were never used in the included studies. Among these, teach to use prompts or cues and time management can be unobservable techniques that are hidden in message contents, rather than being described as intervention activities. The limited description of intervention messages can conceal certain BCTs, leading to bias in analysis and estimation of BCT effects. Also, definitions of some BCTs, such as provide general encouragement and prompt intention formation, may be too general and ambiguous for having a meaningful analysis. Some BCTs, such as prompt intention formation (34/34, 100%), provide information about behavioral-health link (33/34, 97), provide information on consequences (33/34, 97%), and provide instruction (33/34, 97%), were used in almost every study. Their high frequency can be due to their broad definitions, making them almost necessary for all behavioral change interventions.

Second, emotive techniques are proven to be highly effective in affecting one's judgment and decision-making [96] but are not included in the 26-item taxonomy of BCTs. Fear appeal is an emotive technique used in the intervention of one of the included studies but failed to be categorized into any of the BCTs. The omission of important emotion-driven behavior change approaches in the BCTs may have limited the applicability of the taxonomy in other settings.

Third, counting the frequencies of BCTs used may not fully represent the associations between BCTs and significant improvements. Effect size of individual BCT is difficult to ascertain as multiple BCTs are used in one study. Due to the lack of effect size calculation, it is hard to determine whether the high frequency of use of a BCT is due to its effectiveness or its simplicity in application. It also limited meta-analytic evaluations. Hence, it is difficult to fully attribute the significant improvements in results to the use of a specific BCT.

Finally, this review included study samples aged 10 to 30 years, consisting of early adolescents, middle adolescents, and late adolescents or young people. The age difference of targeted populations may influence the effectiveness of the BCTs

XSL•FO
RenderX

provided, as the risk levels of HIV infection and challenges in enacting certain BCTs may differ among early adolescents and young people. Future studies can consider age-specific tailoring to address the potential differences.

Conclusions

This systematic review has identified the most commonly used BCTs used in digital HIV prevention programs targeting adolescents and young people. The findings can serve as important references for future interventions to provide more effective approaches for HIV prevention on the internet.

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Authors' Contributions

PK-hM was responsible for conceptualization, methodology, formal analysis, supervision, project administration, writing (original draft), and writing (review and editing). LX was responsible for methodology, formal analysis, data curation, and writing—review and editing. TCL was responsible for data curation. AYCL was responsible for data curation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File , 17 KB - publichealth_v11i1e59519_app1.docx]

Multimedia Appendix 2
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist.
[PDF File (Adobe PDF File), 85 KB - publichealth_v11i1e59519_app2.pdf]

Multimedia Appendix 3
Eligibility criteria.
[DOCX File , 15 KB - publichealth_v11i1e59519_app3.docx]

References

1. HIV data and statistics. World Health Organization. URL: https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics [accessed 2024-01-01]
2. HIV and AIDS. World Health Organization. 2024 Jul 22. URL: https://www.who.int/news-room/fact-sheets/detail/hiv-aids [accessed 2025-02-07]
3. About HIV. Centers for Disease Control and Prevention. 2025 Jan 14. URL: https://www.cdc.gov/hiv/basics/whatishiv.html [accessed 2025-02-07]
4. Schackman BR, Fleishman JA, Su AE, Berkowitz BK, Moore RD, Walensky RP, et al. The lifetime medical cost savings from preventing HIV in the United States. Med Care 2015 Apr;53(4):293-301 [FREE Full text] [doi: 10.1097/MLR.0000000000000308] [Medline: 25710311]
5. Young people and HIV. Joint United Nations Programme on HIV/AIDS. 2021. URL: https://www.unaids.org/sites/default/files/media_asset/young-people-and-hiv_en.pdf [accessed 2025-04-03]
6. HIV surveillance report: diagnoses of HIV infection in the United States and dependent areas, 2021. Centers for Disease Control and Prevention. 2022. URL: https://stacks.cdc.gov/view/cdc/149071 [accessed 2025-02-07]
7. Estimated HIV incidence and prevalence in the United States, 2015–2019. Centers for Disease Control and Prevention. URL: https://stacks.cdc.gov/view/cdc/112160/cdc_112160_DS1.pdf [accessed 2025-04-03]
8. Countries commit to action to end AIDS by 2030. United Nations. URL: https://www.un.org/en/academic-impact/countries-commit-action-end-aids-2030 [accessed 2025-02-09]
9. Bertoli RS, Scheidmantel CE, De Carvalho NS. College students and HIV infection: a study of sexual behavior and vulnerabilities. Braz J Sex Transm Dis 2016;28(3):90-95.
10. Zhang Kudon H, Mulatu MS, Song W, Heitgerd J, Rao S. Trends in condomless sex among MSM who participated in CDC-funded HIV risk-reduction interventions in the United States, 2012-2017. J Public Health Manag Pract 2022;28(2):170-173. [doi: 10.1097/PHH.0000000000001143] [Medline: 32332483]

XSL•FO
RenderX

11. Szucs LE, Lowry R, Fasula AM, Pampati S, Copen CE, Hussaini KS, et al. Condom and contraceptive use among sexually active high school students - youth risk behavior survey, United States, 2019. MMWR Suppl 2020 Aug 21;69(1):11-18 [FREE Full text] [doi: 10.15585/mmwr.su6901a2] [Medline: 32817600]
12. Kayiki SP, Forste R. HIV/AIDS related knowledge and perceived risk associated with condom use among adolescents in Uganda. Afr J Reprod Health 2011 Mar;15(1):57-63. [Medline: 21987938]
13. Lammers J, van Wijnbergen SJ, Willebrands D. Condom use, risk perception, and HIV knowledge: a comparison across sexes in Nigeria. HIV AIDS (Auckl) 2013;5:283-293 [FREE Full text] [doi: 10.2147/HIV.S31687] [Medline: 24187512]
14. Casey MK, Timmermann L, Allen M, Krahn S, Turkiewicz KL. Response and self-efficacy of condom use: a meta-analysis of this important element of AIDS education and prevention. South Commun J 2009 Feb 02;74(1):57-78. [doi: 10.1080/10417940802335953]
15. Sheeran P, Abraham C, Orbell S. Psychosocial correlates of heterosexual condom use: a meta-analysis. Psychol Bull 1999 Jan;125(1):90-132. [doi: 10.1037/0033-2909.125.1.90] [Medline: 9990846]
16. Yu B, Wang Y, Chen X. Perception of peer condom use buffers the associations between HIV knowledge, self-efficacy, and condom-use intention among adolescents: a moderated mediation model. Prev Sci 2022 Aug 28;23(6):879-888. [doi: 10.1007/s11121-021-01324-6] [Medline: 34962622]
17. Adolescent HIV prevention. UNICEF. URL: https://data.unicef.org/topic/hivaids/adolescents-young-people/ [accessed 2024-01-01]
18. Alarape AI, Olapegba PO, Chovwen CO. Condom use among students: the influence of condom self-efficacy, social norms and affective attitude towards condom. J Soc Sci 2017 Oct 09;17(3):237-241. [doi: 10.1080/09718923.2008.11892657]
19. Moore BA, Fazzino T, Garnet B, Cutter CJ, Barry DT. Computer-based interventions for drug use disorders: a systematic review. J Subst Abuse Treat 2011 Apr;40(3):215-223 [FREE Full text] [doi: 10.1016/j.jsat.2010.11.002] [Medline: 21185683]
20. Gulliver A, Farrer L, Chan JK, Tait RJ, Bennett K, Calear AL, et al. Technology-based interventions for tobacco and other drug use in university and college students: a systematic review and meta-analysis. Addict Sci Clin Pract 2015 Feb 24;10(1):5 [FREE Full text] [doi: 10.1186/s13722-015-0027-4] [Medline: 25928221]
21. Park E, Kwon M. Health-related internet use by children and adolescents: systematic review. J Med Internet Res 2018 Apr 03;20(4):e120 [FREE Full text] [doi: 10.2196/jmir.7731] [Medline: 29615385]
22. Wartella E, Rideout V, Montague H, Beaudoin-Ryan L, Lauricella A. Teens, health and technology: a national survey. Media Commun 2016;4(3) [FREE Full text] [doi: 10.17645/mac.v4i3.515]
23. Mustanski B, Lyons T, Garcia SC. Internet use and sexual health of young men who have sex with men: a mixed-methods study. Arch Sex Behav 2011 Apr 25;40(2):289-300 [FREE Full text] [doi: 10.1007/s10508-009-9596-1] [Medline: 20182787]
24. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. AIDS 2009 Jan 02;23(1):107-115. [doi: 10.1097/QAD.0b013e32831c5500] [Medline: 19050392]
25. Knight R, Karamouzian M, Salway T, Gilbert M, Shoveller J. Online interventions to address HIV and other sexually transmitted and blood-borne infections among young gay, bisexual and other men who have sex with men: a systematic review. J Int AIDS Soc 2017 Nov;20(3):e25017 [FREE Full text] [doi: 10.1002/jia2.25017] [Medline: 29091340]
26. Xin M, Viswanath K, Li AY, Cao W, Hu Y, Lau JT, et al. The effectiveness of electronic health interventions for promoting HIV-preventive behaviors among men who have sex with men: meta-analysis based on an integrative framework of design and implementation features. J Med Internet Res 2020 May 25;22(5):e15977 [FREE Full text] [doi: 10.2196/15977] [Medline: 32449685]
27. Maloney KM, Bratcher A, Wilkerson R, Sullivan PS. Electronic and other new media technology interventions for HIV care and prevention: a systematic review. J Int AIDS Soc 2020 Jan;23(1):e25439 [FREE Full text] [doi: 10.1002/jia2.25439] [Medline: 31909896]
28. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of web-based vs. non-web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
29. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
30. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015 Nov;19(99):1-188 [FREE Full text] [doi: 10.3310/hta19990] [Medline: 26616119]
31. Hagger MS, Cameron LD, Hamilton K, Hankonen N, Lintunen T. The Handbook of Behavior Change. Cambridge, UK: Cambridge University Press; 2020.
32. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
33. Howlett N, Trivedi D, Troop NA, Chater AM. Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med 2019 Jan 01;9(1):147-157 [FREE Full text] [doi: 10.1093/tbm/iby010] [Medline: 29506209]

XSL•FO
RenderX

34. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug 15;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
35. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
36. Bird EL, Baker G, Mutrie N, Ogilvie D, Sahlqvist S, Powell J. Behavior change techniques used to promote walking and cycling: a systematic review. Health Psychol 2013 Aug;32(8):829-838 [FREE Full text] [doi: 10.1037/a0032078] [Medline: 23477577]
37. Van Rhoon L, Byrne M, Morrissey E, Murphy J, McSharry J. A systematic review of the behaviour change techniques and digital features in technology-driven type 2 diabetes prevention interventions. Digit Health 2020;6:2055207620914427 [FREE Full text] [doi: 10.1177/2055207620914427] [Medline: 32269830]
38. Winter SJ, Sheats JL, King AC. The use of behavior change techniques and theory in technologies for cardiovascular disease prevention and treatment in adults: a comprehensive review. Prog Cardiovasc Dis 2016;58(6):605-612 [FREE Full text] [doi: 10.1016/j.pcad.2016.02.005] [Medline: 26902519]
39. Patterson K, Davey R, Keegan R, Kunstler B, Woodward A, Freene N. Behaviour change techniques in cardiovascular disease smartphone apps to improve physical activity and sedentary behaviour: systematic review and meta-regression. Int J Behav Nutr Phys Act 2022 Jul 07;19(1):81 [FREE Full text] [doi: 10.1186/s12966-022-01319-8] [Medline: 35799263]
40. Dugas M, Gao GG, Agarwal R. Unpacking mHealth interventions: a systematic review of behavior change techniques used in randomized controlled trials assessing mHealth effectiveness. Digit Health 2020;6:2055207620905411 [FREE Full text] [doi: 10.1177/2055207620905411] [Medline: 32128233]
41. Yeh PT, Kennedy CE, Minamitani A, Baggaley R, Shah P, Verster A, et al. Web-based service provision of HIV, viral hepatitis, and sexually transmitted infection prevention, testing, linkage, and treatment for key populations: systematic review and meta-analysis. J Med Internet Res 2022 Dec 22;24(12):e40150 [FREE Full text] [doi: 10.2196/40150] [Medline: 36548036]
42. Nguyen LH, Tran BX, Rocha LE, Nguyen HL, Yang C, Latkin CA, et al. A systematic review of eHealth interventions addressing HIV/STI prevention among men who have sex with men. AIDS Behav 2019 Sep;23(9):2253-2272 [FREE Full text] [doi: 10.1007/s10461-019-02626-1] [Medline: 31401741]
43. Wadham E, Green C, Debattista J, Somerset S, Sav A. New digital media interventions for sexual health promotion among young people: a systematic review. Sex Health 2019 Apr;16(2):101-123. [doi: 10.1071/SH18127] [Medline: 30819326]
44. Mwale M, Muula AS. Systematic review: a review of adolescent behavior change interventions [BCI] and their effectiveness in HIV and AIDS prevention in sub-Saharan Africa. BMC Public Health 2017 Sep 18;17(1):718 [FREE Full text] [doi: 10.1186/s12889-017-4729-2] [Medline: 28923040]
45. Laricchia F. Smartphones - statistics and facts. Statista. 2024 Jun 12. URL: https://www.statista.com/topics/840/smartphones/#topicOverview [accessed 2025-02-19]
46. PRISMA 2020 statement. PRISMA. URL: https://www.prisma-statement.org/ [accessed 2025-02-07]
47. Barker TH, Stone JC, Sears K, Klugar M, Tufanaru C, Leonardi-Bee J, et al. The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. JBI Evid Synth 2023 Mar 01;21(3):494-506. [doi: 10.11124/JBIES-22-00430] [Medline: 36727247]
48. Widman L, Nesi J, Kamke K, Choukas-Bradley S, Stewart JL. Technology-based interventions to reduce sexually transmitted infections and unintended pregnancy among youth. J Adolesc Health 2018 Jun;62(6):651-660 [FREE Full text] [doi: 10.1016/j.jadohealth.2018.02.007] [Medline: 29784112]
49. Hergenrather KC, Emmanuel D, Durant S, Rhodes SD. Enhancing HIV prevention among young men who have sex with men: a systematic review of HIV behavioral interventions for young gay and bisexual men. AIDS Educ Prev 2016 Jun;28(3):252-271. [doi: 10.1521/aeap.2016.28.3.252] [Medline: 27244193]
50. Chávez NR, Shearer LS, Rosenthal SL. Use of digital media technology for primary prevention of STIs/HIV in youth. J Pediatr Adolesc Gynecol 2014 Oct;27(5):244-257. [doi: 10.1016/j.jpag.2013.07.008] [Medline: 24332613]
51. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
52. Ezegbe B, Eseadi C, Ede MO, Igbo JN, Aneke A, Mezieobi D, et al. Efficacy of rational emotive digital storytelling intervention on knowledge and risk perception of HIV/AIDS among schoolchildren in Nigeria. Medicine (Baltimore) 2018 Nov;97(47):e12910 [FREE Full text] [doi: 10.1097/MD.0000000000012910] [Medline: 30461604]
53. Ybarra ML, Bull SS, Prescott TL, Korchmaros JD, Bangsberg DR, Kiwanuka JP. Adolescent abstinence and unprotected sex in CyberSenga, an internet-based HIV prevention program: randomized clinical trial of efficacy. PLoS One 2013 Aug 14;8(8):e70083 [FREE Full text] [doi: 10.1371/journal.pone.0070083] [Medline: 23967069]
54. Ybarra ML, Korchmaros JD, Prescott TL, Birungi R. A randomized controlled trial to increase HIV preventive information, motivation, and behavioral skills in Ugandan adolescents. Ann Behav Med 2015 Jun 30;49(3):473-485 [FREE Full text] [doi: 10.1007/s12160-014-9673-0] [Medline: 25633626]

XSL•FO
RenderX

55. Logie CH, Okumu M, Berry I, Hakiza R, Baral SD, Musoke DK, et al. Findings from the Tushirikiane mobile health (mHealth) HIV self-testing pragmatic trial with refugee adolescents and youth living in informal settlements in Kampala, Uganda. J Int AIDS Soc 2023 Oct;26(10):e26185 [FREE Full text] [doi: 10.1002/jia2.26185] [Medline: 37850816]
56. Lau JT, Lee AL, Tse WS, Mo PK, Fong F, Wang Z, et al. A randomized control trial for evaluating efficacies of two online cognitive interventions with and without fear-appeal imagery approaches in preventing unprotected anal sex among Chinese men who have sex with men. AIDS Behav 2016 Sep 22;20(9):1851-1862. [doi: 10.1007/s10461-015-1263-z] [Medline: 26802004]
57. Yi M, Li X, Chiaramonte D, Sun S, Pan S, Soulliard Z, et al. Guided internet-based LGBTQ-affirmative cognitive-behavioral therapy: a randomized controlled trial among sexual minority men in China. Behav Res Ther 2024 Oct;181:104605 [FREE Full text] [doi: 10.1016/j.brat.2024.104605] [Medline: 39029333]
58. Zhang K, Cao B, Fang Y, Liang X, Ye D, Chen YQ, et al. Comparing the efficacy of 2 WeChat mini programs in reducing nonmarital heterosexual contact by male factory workers: randomized controlled trial. J Med Internet Res 2024 Sep 09;26:e49362 [FREE Full text] [doi: 10.2196/49362] [Medline: 39250213]
59. Anand T, Nitpolprasert C, Jantarapakde J, Meksena R, Phomthong S, Phoseeta P, et al. Implementation and impact of a technology-based HIV risk-reduction intervention among Thai men who have sex with men using "Vialogues": a randomized controlled trial. AIDS Care 2020 Mar;32(3):394-405 [FREE Full text] [doi: 10.1080/09540121.2019.1622638] [Medline: 31154811]
60. Brody C, Chhoun P, Tuot S, Fehrenbacher AE, Moran A, Swendeman D, et al. A mobile intervention to link young female entertainment workers in Cambodia to health and gender-based violence services: randomized controlled trial. J Med Internet Res 2022 Jan 04;24(1):e27696 [FREE Full text] [doi: 10.2196/27696] [Medline: 34982716]
61. Levy E, Warner LM, Fleig L, Kaufman MR, Deschepper R, Gidron Y. The effects of psychological inoculation on condom use tendencies and barriers; a randomized controlled trial. Psychol Health 2021 May 05;36(5):575-592. [doi: 10.1080/08870446.2020.1775832] [Medline: 32498613]
62. Wray TB, Kahler CW, Simpanen EM, Operario D. A preliminary randomized controlled trial of game plan, a web application to help men who have sex with men reduce their HIV risk and alcohol use. AIDS Behav 2019 Jun 22;23(6):1668-1679 [FREE Full text] [doi: 10.1007/s10461-019-02396-w] [Medline: 30671682]
63. Schnall R, Kuhns LM, Pearson C, Batey DS, Bruce J, Hidalgo MA, et al. Efficacy of MyPEEPS mobile, an HIV prevention intervention using mobile technology, on reducing sexual risk among same-sex attracted adolescent males: a randomized clinical trial. JAMA Netw Open 2022 Sep 01;5(9):e2231853 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.31853] [Medline: 36129712]
64. Nelson KM, Perry NS, Stout CD, Dunsiger SI, Carey MP. The young men and media study: a pilot randomized controlled trial of a community-informed, online HIV prevention intervention for 14-17-year-old sexual minority males. AIDS Behav 2022 Feb;26(2):569-583 [FREE Full text] [doi: 10.1007/s10461-021-03412-8] [Medline: 34342742]
65. Mustanski B, Parsons JT, Sullivan PS, Madkins K, Rosenberg E, Swann G. Biomedical and behavioral outcomes of keep it up!: an eHealth HIV prevention program RCT. Am J Prev Med 2018 Aug;55(2):151-158 [FREE Full text] [doi: 10.1016/j.amepre.2018.04.026] [Medline: 29937115]
66. Mustanski B, Garofalo R, Monahan C, Gratzer B, Andrews R. Feasibility, acceptability, and preliminary efficacy of an online HIV prevention program for diverse young men who have sex with men: the keep it up! intervention. AIDS Behav 2013 Nov;17(9):2999-3012 [FREE Full text] [doi: 10.1007/s10461-013-0507-z] [Medline: 23673793]
67. Hightow-Weidman LB, LeGrand S, Muessig KE, Simmons RA, Soni K, Choi SK, et al. A randomized trial of an online risk reduction intervention for young Black MSM. AIDS Behav 2019 May;23(5):1166-1177 [FREE Full text] [doi: 10.1007/s10461-018-2289-9] [Medline: 30269231]
68. Hightow-Weidman LB, Pike E, Fowler B, Matthews DM, Kibe J, McCoy R, et al. HealthMpowerment.org: feasibility and acceptability of delivering an internet intervention to young Black men who have sex with men. AIDS Care 2012 Jan 24;24(7):910-920 [FREE Full text] [doi: 10.1080/09540121.2011.647677] [Medline: 22272759]
69. Christensen JL, Miller LC, Appleby PR, Corsbie-Massay C, Godoy CG, Marsella SC, et al. Reducing shame in a game that predicts HIV risk reduction for young adult MSM: a randomized trial delivered nationally over the web. J Int AIDS Soc 2013 Nov 13;16(3 Suppl 2):18716 [FREE Full text] [doi: 10.7448/IAS.16.3.18716] [Medline: 24242264]
70. Bauermeister JA, Tingler RC, Demers M, Connochie D, Gillard G, Shaver J, et al. Acceptability and preliminary efficacy of an online HIV prevention intervention for single young men who have sex with men seeking partners online: the myDEx project. AIDS Behav 2019 Nov;23(11):3064-3077 [FREE Full text] [doi: 10.1007/s10461-019-02426-7] [Medline: 30762190]
71. Newcomb ME, Swann G, Macapagal K, Sarno EL, Whitton SW, Mustanski B. Biomedical and behavioral outcomes of 2GETHER: a randomized controlled trial of a telehealth HIV prevention program for young male couples. J Consult Clin Psychol 2023 Sep;91(9):505-520. [doi: 10.1037/ccp0000823] [Medline: 37141032]
72. Widman L, Golin CE, Kamke K, Burnette JL, Prinstein MJ. Sexual assertiveness skills and sexual decision-making in adolescent girls: randomized controlled trial of an online program. Am J Public Health 2018 Jan;108(1):96-102. [doi: 10.2105/AJPH.2017.304106] [Medline: 29161072]

XSL•FO
RenderX

73. Peskin MF, Shegog R, Markham CM, Thiel M, Baumler ER, Addy RC, et al. Efficacy of it's your game-tech: a computer-based sexual health education program for middle school youth. J Adolesc Health 2015 May;56(5):515-521 [FREE Full text] [doi: 10.1016/j.jadohealth.2015.01.001] [Medline: 25739520]
74. Santa Maria D, Padhye N, Businelle M, Yang Y, Jones J, Sims A, et al. Efficacy of a just-in-time adaptive intervention to promote HIV risk reduction behaviors among young adults experiencing homelessness: pilot randomized controlled trial. J Med Internet Res 2021 Jul 06;23(7):e26704 [FREE Full text] [doi: 10.2196/26704] [Medline: 34255679]
75. Marsch LA, Grabinski MJ, Bickel WK, Desrosiers A, Guarino H, Muehlbach B, et al. Computer-assisted HIV prevention for youth with substance use disorders. Subst Use Misuse 2011 Dec 29;46(1):46-56 [FREE Full text] [doi: 10.3109/10826084.2011.521088] [Medline: 21190405]
76. Marsch LA, Guarino H, Grabinski MJ, Syckes C, Dillingham ET, Xie H, et al. Comparative effectiveness of web-based vs. educator-delivered HIV prevention for adolescent substance users: a randomized, controlled trial. J Subst Abuse Treat 2015 Dec;59:30-37 [FREE Full text] [doi: 10.1016/j.jsat.2015.07.003] [Medline: 26293644]
77. Bull SS, Levine DK, Black SR, Schmiege SJ, Santelli J. Social media-delivered sexual health intervention: a cluster randomized controlled trial. Am J Prev Med 2012 Nov;43(5):467-474 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.022] [Medline: 23079168]
78. Bull S, Pratte K, Whitesell N, Rietmeijer C, McFarlane M. Effects of an internet-based intervention for HIV prevention: the Youthnet trials. AIDS Behav 2009 Jun 27;13(3):474-487 [FREE Full text] [doi: 10.1007/s10461-008-9487-9] [Medline: 19037719]
79. Cordova D, Munoz-Velazquez J, Mendoza Lua F, Fessler K, Warner S, Delva J, Youth Leadership Council, et al. Pilot study of a multilevel mobile health app for substance use, sexual risk behaviors, and testing for sexually transmitted infections and HIV among youth: randomized controlled trial. JMIR Mhealth Uhealth 2020 Mar 17;8(3):e16251 [FREE Full text] [doi: 10.2196/16251] [Medline: 32181747]
80. Fiellin LE, Hieftje KD, Pendergrass TM, Kyriakides TC, Duncan LR, Dziura JD, et al. Video game intervention for sexual risk reduction in minority adolescents: randomized controlled trial. J Med Internet Res 2017 Sep 18;19(9):e314 [FREE Full text] [doi: 10.2196/jmir.8148] [Medline: 28923788]
81. Jones R, Hoover DR, Lacroix LJ. A randomized controlled trial of soap opera videos streamed to smartphones to reduce risk of sexually transmitted human immunodeficiency virus (HIV) in young urban African American women. Nurs Outlook 2013;61(4):205-15.e3 [FREE Full text] [doi: 10.1016/j.outlook.2013.03.006] [Medline: 23743482]
82. Klein CH, Card JJ. Preliminary efficacy of a computer-delivered HIV prevention intervention for African American teenage females. AIDS Educ Prev 2011 Dec;23(6):564-576. [doi: 10.1521/aeap.2011.23.6.564] [Medline: 22201239]
83. McCrimmon J, Widman L, Javidi H, Brasileiro J, Hurst J. Evaluation of a brief online sexual health program for adolescents: a randomized controlled trial. Health Promot Pract 2024 Jul;25(4):689-697 [FREE Full text] [doi: 10.1177/15248399231162379] [Medline: 37013260]
84. Swendeman D, Rotheram-Borus M, Arnold E, Fernández M, Comulada W, Lee SJ, et al. Optimal strategies to improve uptake of and adherence to HIV prevention among young people at risk for HIV acquisition in the USA (ATN 149): a randomised, controlled, factorial trial. The Lancet Digital Health 2024 Mar;6(3):e187-e200 [FREE Full text] [doi: 10.1016/s2589-7500(23)00252-2]
85. Widman L, Kamke K, Evans R, Stewart JL, Choukas-Bradley S, Golin CE. Feasibility, acceptability, and preliminary efficacy of a brief online sexual health program for adolescents. J Sex Res 2020 Feb 09;57(2):145-154 [FREE Full text] [doi: 10.1080/00224499.2019.1630800] [Medline: 31287336]
86. Logie C, Okumu M, Hakiza R, Kibuuka Musoke D, Berry I, Mwima S, et al. Mobile health-supported HIV self-testing strategy among urban refugee and displaced youth in Kampala, Uganda: protocol for a cluster randomized trial (Tushirikiane, Supporting Each Other). JMIR Res Protoc 2021 Feb 02;10(2):e26192 [FREE Full text] [doi: 10.2196/26192] [Medline: 33528378]
87. Javidi H, Widman L, Lipsey N, Brasileiro J, Javidi F, Jhala A. Redeveloping a digital sexual health intervention for adolescents to allow for broader dissemination: implications for HIV and STD prevention. AIDS Educ Prev 2021 Apr;33(2):89-102. [doi: 10.1521/aeap.2021.33.2.89] [Medline: 33821678]
88. Swendeman D, Arnold EM, Harris D, Fournier J, Comulada WS, Reback C, et al. Text-messaging, online peer support group, and coaching strategies to optimize the HIV prevention continuum for youth: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Aug 09;8(8):e11165 [FREE Full text] [doi: 10.2196/11165] [Medline: 31400109]
89. Carey MP, Forsyth AD. Teaching tip sheet: self-efficacy. American Psychological Association. URL: https://www.apa.org/pi/aids/resources/education/self-efficacy [accessed 2023-06-01]
90. Yusoff YM. Self-efficacy, perceived social support, and psychological adjustment in international undergraduate students in a public higher education institution in Malaysia. J Stud Int Educ 2011 Jun 16;16(4):353-371. [doi: 10.1177/1028315311408914]
91. Salas-Groves E, Galyean S, Alcorn M, Childress A. Behavior change effectiveness using nutrition apps in people with chronic diseases: scoping review. JMIR Mhealth Uhealth 2023 Jan 13;11:e41235 [FREE Full text] [doi: 10.2196/41235] [Medline: 36637888]

XSL•FO
RenderX

92. McGrady ME, Schwartz LE, Noser AE, Klages KL, Sweenie R, Breen G, et al. Systematic evaluation of the behavior change techniques and quality of commercially available cancer self-management apps. JCO Oncol Pract 2023 Feb;19(2):e228-e237 [FREE Full text] [doi: 10.1200/OP.22.00491] [Medline: 36446046]
93. Yan J, Zhang A, Zhou L, Huang Z, Zhang P, Yang G. Development and effectiveness of a mobile phone application conducting health behavioral intervention among men who have sex with men, a randomized controlled trial: study protocol. BMC Public Health 2017 Apr 24;17(1):355 [FREE Full text] [doi: 10.1186/s12889-017-4235-6] [Medline: 28438144]
94. Nurmi J, Knittle K, Ginchev T, Khattak F, Helf C, Zwickl P, et al. Engaging users in the behavior change process with digitalized motivational interviewing and gamification: development and feasibility testing of the precious app. JMIR Mhealth Uhealth 2020 Jan 30;8(1):e12884 [FREE Full text] [doi: 10.2196/12884] [Medline: 32003750]
95. Petty RE, Cacioppo JT. The elaboration likelihood model of persuasion. Adv Exp Soc Psychol 1986;19:123-205 [FREE Full text] [doi: 10.1016/S0065-2601(08)60214-2]
96. Lerner JS, Li Y, Valdesolo P, Kassam KS. Emotion and decision making. Annu Rev Psychol 2015 Jan 03;66:799-823 [FREE Full text] [doi: 10.1146/annurev-psych-010213-115043] [Medline: 25251484]

Abbreviations

BCT: behavior change technique
MSM: men who have sex with men
RCT: randomized controlled trial

Edited by A Mavragani; submitted 14.04.24; peer-reviewed by SMK Moghadam, D Murphy, Y Yan; comments to author 31.05.24; revised version received 04.07.24; accepted 05.12.24; published 28.04.25.

Please cite as:
Mo PKH, Xie L, Lee TC, Li AYC
Use of Behavior Change Techniques in Digital HIV Prevention Programs for Adolescents and Young People: Systematic Review
JMIR Public Health Surveill 2025;11:e59519
URL: https://publichealth.jmir.org/2025/1/e59519
doi:10.2196/59519
PMID:

©Phoenix Kit-han Mo, Luyao Xie, Tsz Ching Lee, Angela Yuen Chun Li. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Perceptions of Long-Acting Injectable Pre-Exposure Prophylaxis Among Men Who Have Sex With Men and Transgender Individuals in Europe Using Structural Text Modeling Technique: Cross-Sectional Study

Haoyi Wang[1,2], MPH, MSc; Johann Kolstee[1], PhD; Julio Croce Martinez[1], BA; David van de Vijver[2], PhD; Jonathan Tosh[3], PhD; Melanie Schroeder[3], MSc; Ama Appiah[3], MSc; Hanne ML Zimmermann[1], PhD; Kai J Jonas[1], PhD

[1]Department of Work and Social Psychology, Maastricht University, Maastricht, The Netherlands

[2]Department of Viroscience, Erasmus MC, Rotterdam, The Netherlands

[3]ViiV Healthcare Ltd, London, United Kingdom

Corresponding Author:
Haoyi Wang, MPH, MSc
Department of Work and Social Psychology
Maastricht University
PO Box 616
Universiteitssingel 40
Maastricht, 6200 MD
The Netherlands
Phone: 31 43 38 84531
Email: haoyi.wang@maastrichtuniversity.nl

Abstract

Background: Oral pre-exposure prophylaxis (PrEP) effectively prevents HIV but remains unevenly accessible across Europe. Long-acting PrEP (LA-PrEP), recently approved in Europe, offers new HIV protection options. However, no qualitative evidence is available to inform people's perceptions of this novel modality.

Objective: This study provides the first large-scale, qualitative evidence in English from 20 European countries on how men who have sex with men (MSM) and trans* individuals perceive LA-PrEP.

Methods: We analyzed open-ended responses from 3123 HIV-negative MSM and trans* individuals from 20 European countries who completed the PROTECT survey in English. Participants were asked to describe what LA-PrEP means to them using words or short phrases. We used word clouds for initial insights and structural topic modeling to identify topics and explore their relevance across socioeconomic status, migration background, oral PrEP use, and affordability and the association of these perceptions with LA-PrEP intention.

Results: The responses reflected generally positive associations toward LA-PrEP, with the most frequently mentioned word being "safe," "freedom," and "convenient." However, some ambivalent and negative perceptions, such as "nothing," "unknown," "dunno," and "unnecessary," were also noted. Structural topic modeling identified 5 main response topics: safety, empowerment, convenience/reliability, peace of mind, and concerns/uncertainties. The empowerment offered by LA-PrEP was the most prominent topic, representing one third (28.1%) of the responses, followed by safety (21%), convenience and reliability (16%), and peace of mind (15%), while concerns/uncertainties made up 20%. Variation in the relevance of these topics was found, showing LA-PrEP being seen as more empowering (=.070, 95% CI 0.042-0.097) and convenient (=.057, 95% CI 0.034-0.081) by current oral PrEP users, but less empowering (=–.052, 95% CI –0.087 to 0.017) and convenient (=–.034, 95% CI –0.064 to 0.005) for individuals in countries with limited oral PrEP access and affordability. The topic of safety was more relevant among those with lower levels of education (=.052, 95% CI 0.022-0.083) and those living in a country where PrEP was not reimbursed (=.035, 95% CI 0.002-0.069), but less relevant among current oral PrEP users (=–.094, 95% CI –0.12 to 0.066). We also found that ambivalent and negative perceptions were less relevant among current oral PrEP users (=–.032, 95% CI –0.056 to 0.007) and were negatively associated with a lower intention to use LA-PrEP (=–.075, 95% CI –0.101 to 0.005).

Conclusions: Our research showed that MSM and trans* individuals in Europe generally have a positive outlook on LA-PrEP, suggesting it is likely to be well accepted upon its introduction in Europe. However, a subset of the target population may be

XSL•FO
RenderX

hesitant to adopt LA-PrEP, underscoring the need for alternative HIV prevention strategies tailored to these individuals. To support potential future LA-PrEP implementation in Europe and to maximize its impact, appropriate communication strategies are essential for supporting informed decision-making.

(JMIR Public Health Surveill 2025;11:e72491) doi:10.2196/72491

KEYWORDS

men who have sex with men; trans*; pre-exposure prophylaxis; long-acting injectable PrEP; qualitative; topic modeling; Europe

Introduction

HIV oral pre-exposure prophylaxis (PrEP) has proven to be highly effective in preventing HIV [1,2] and is widely accepted by key populations affected by HIV [3,4]. By 2024, approximately 3.5 million individuals worldwide were using PrEP [5]. Despite efforts to increase its availability and accessibility, PrEP uptake remains lower than expected in many countries and among key populations [6-8], falling short of the United Nations 2025 target of reaching 10 million users globally. The full potential of oral PrEP at the population level has yet to be reached [9-12].

Recently, novel long-acting (LA) PrEP formulations have emerged to address the limitations of oral PrEP [13-17] and enhance PrEP coverage and its impact at the population level [18-21]. Notably, cabotegravir administered every 2 months has demonstrated statistically superior efficacy in preventing HIV compared to oral PrEP across all key populations [22,23]. Recognizing this, the World Health Organization (WHO) endorsed cabotegravir for high-risk groups [24]. Cabotegravir was subsequently approved for use in the United States in 2022 and in Europe in 2023 [25,26]. In addition, the newer LA injectable lenacapavir, administered twice yearly, has recently demonstrated 100% efficacy in preventing HIV among cisgender women [27].

Understanding the views and perceptions of novel LA-PrEP among potential end users is crucial, particularly in Europe, where data on LA-PrEP perceptions and attitudes are scarce [28]. As of early 2025, only one systematic review has synthesized global perspectives on LA-PrEP among key populations [16], reporting that while LA-PrEP is generally perceived as effective, easy to use, and convenient, concerns remain about potential pain, side effects, and logistical challenges such as the need for regular appointments. However, these insights stemmed primarily from qualitative studies with small sample sizes, raising questions about their generalizability. Furthermore, most of the available evidence originates from North America and Africa, with only 5 studies focusing on Europe, none of which provided qualitative data on European key populations' expectations, attitudes, or perceptions regarding LA-PrEP, and none included trans* participants. This gap underscores the need for large-scale qualitative evidence specific to the European context to inform the introduction and scaling of LA-PrEP.

Previous research has highlighted that perceptions and preferences for LA-PrEP vary significantly among subgroups within key populations [16]. Drawing from experiences with oral PrEP in Europe, individuals with lower socioeconomic positions (SEP) and migration backgrounds were less likely to access oral PrEP, a pattern that may similarly influence perceptions and access to LA-PrEP [4,7,29,30]. In addition, past and current experiences with oral PrEP can shape perceptions of LA-PrEP [31]. For example, PrEP-naïve MSM in Europe are less likely to intend to use LA-PrEP compared to current users [28], indicating underlying differences in attitudes and expectations. The LA nature of LA-PrEP, which alleviates the burden of daily pill-taking, may particularly appeal to individuals who experience challenges with adherence to daily regimens [19,22]. However, beyond individual-level considerations, structural factors such as affordability and health system infrastructure are also critical. In countries where oral PrEP is not yet reimbursed or remains unaffordable, LA-PrEP may be viewed differently than in countries with more accessible PrEP options [7,32]. Additional structural barriers—including provider bias, mismatches in provider-client expectations [33], and geographic distance to PrEP clinics [34,35]—may further shape individuals' LA-PrEP intention and uptake. Understanding these subgroup variations emphasizes the importance of analyzing large-scale qualitative data to gain insights into how different populations perceive LA-PrEP and how these perceptions may influence their intention to use it. Such insights are essential for informing European access policies and guidelines to ensure that the introduction and scale-up of LA-PrEP meet the diverse needs of key populations.

Gathering large-scale qualitative data is, however, often challenging and costly, leading to studies with smaller sample sizes [36]. One potential solution is incorporating open-ended responses into large-scale quantitative cross-sectional surveys. Yet, most global HIV prevention surveys, including those on PrEP [24,37-39], rely on closed-ended questions, which limit responses to predefined options and scales. To address these limitations, text mining and topic modeling techniques have emerged as scalable methods to analyze qualitative responses. However, their application in HIV prevention research, particularly for understanding end-user perceptions, remains limited. Most existing work is typically applied to public discourse, such as user reactions on social media platforms such as X (formerly Twitter) or Instagram. While such studies offer insights into general sentiment or advocacy trends, they are constrained by the absence of detailed respondent information, making subgroup analysis difficult. Moreover, topics tend to reflect political or public health narratives rather than individual perspectives or lived experience. Only a few studies have applied text mining and topic modeling to HIV-related content. For instance, BERTopic has been used to examine trends in HIV care services literature [40], and topic modeling has been used to explore public discussions of PrEP on social media [41-43], demonstrating the potential of such tools to synthesize

XSL•FO
RenderX

broad themes. However, none of these approaches has been used to analyze individual-level qualitative data on LA-PrEP, particularly within a European context.

Among topic modeling methods, structural topic modeling (STM) is uniquely suited to social science and public health research [44,45]. Unlike other unsupervised methods such as keyword-assisted topic modeling [46], seeded latent Dirichlet allocation [47], or BERTopic [48], STM uniquely allows the inclusion of metadata, such as socioeconomic status or previous PrEP experience, as covariates that influence topical prevalence and content [45,49,50]. This makes STM especially valuable for hypothesis-driven or exploratory analysis, enabling the exploration of how themes or topics vary across different subgroups—something rarely possible in social media–based or literature-based studies, and ideal for analyzing open-ended survey responses in social science and public health domains [44]. By integrating STM with large-scale, individual-level qualitative data, we offer a novel methodological approach to understand how key populations perceive LA-PrEP, filling an important gap in both computational social science and HIV prevention research.

Taken together, this study aimed to provide the first large-scale qualitative insight into the attitudes and perceptions of LA-PrEP among MSM and trans* individuals in Europe. In addition, we sought to understand whether these perceptions differ by subgroup characteristics and previous oral PrEP use, and whether they are associated with intention to use of LA-PrEP using an STM approach.

Methods

Study Recruitment and Study Population

We conducted an internet-based survey (PROTECT study) across 20 European countries (Austria, Belgium, Cyprus, the Czech Republic, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, and the United Kingdom) among MSM and trans* individuals. The PROTECT survey aimed to better understand the characteristics and behavior of key populations in Europe who may benefit from new PrEP modalities. Participants were recruited using a convenience sampling approach. Ads containing text information about the purpose of the study, photo images, and inbox messages were used to solicit potential participants mainly from gay dating apps (eg, Grindr) and through a social media campaign on Instagram and Facebook during a period of 6 months (October 2023-April 2024). Participants could fill in the survey in 22 different languages. Overall, 63.4% of the participants who started the survey completed at least 95% of it, resulting in 15,428 final survey responses. For this study, we included only HIV-negative participants who responded to the survey in English. The full study procedure, the full survey, and the recruitment approach have been described elsewhere [51-53].

Measures

In the survey, participants were asked to describe what LA-PrEP meant to them using words or short phrases. To analyze European MSM and trans* individuals' perceptions of LA-PrEP, we first processed their responses by normalizing the open response entries—removing punctuation and numbers, converting all words to lowercase, and removing stop words [54]. Since most responses were single words, additional text preparation was required for robust model performance [49,55,56]. In this text normalization preparation process, to reduce dimensionality and improve coherence, we manually standardized entries, aligning them with the most frequently entered term to avoid having 2 or more expressions of the same word stem [57]. For instance, if "safe" appeared more often than "safety," we replaced all instances of "safety" with "safe" to maintain consistency and reflect the most common expressions used by participants. In addition, for short phrases that indicated the same meaning as a single word, we manually replaced these short phrases with the existing single words to ensure key information would not be removed during text cleaning [58]. For example, we replaced "I don't know" with "dunno." Finally, we used abbreviations for multiple-word phrases that would otherwise lose their meaning, such as converting "peace of mind" to "pom" [59]. This text normalization step was conducted by HW and KJJ to ensure consistency, with intercoder agreement established on all final decisions. The raw data and entries after manual cleaning are shown in Multimedia Appendix 2.

To explore potential variations in different meanings and expectations of LA-PrEP between groups, we also measured participants' SEP, migration background, oral PrEP use history, PrEP affordability, and LA-PrEP intention. SEP variables included perceived income, education, and migration background. Oral PrEP use variables included uptake, regimens, and adherence. PrEP affordability was proxied by the current reimbursement policies of the participant's country of residence (Table 1). We considered a country to have the fully-reimbursed status if PrEP was being included in health insurance coverage (Denmark, Finland, France, Ireland, Luxembourg, Norway, Portugal, Spain and the United Kingdom), the partially-reimbursed status if a copayment was needed for PrEP access (Austria, Belgium, Germany, Netherlands, Sweden and Switzerland), and the non-reimbursed status if PrEP was only accessible via out-of-pocket payment (Cyprus, Czech Republic, Greece, Italy, and Poland). The classification of countries was based on policy settings up to and including January 1, 2023. LA-PrEP intention was dichotomized as "higher intention" and lower intention." Variable descriptions can be found in Table 2.

XSL•FO
RenderX

Table 1. Country classification of pre-exposure prophylaxis (PrEP) reimbursement status as of January 1, 2023, which was relevant for the sampling period of October 2023-April 2024.

PrEP reimbursement status	Countries
Fully reimbursed	Denmark, Finland, France, Ireland, Luxembourg, Norway, Portugal, Spain, and the United Kingdom
Partially reimbursed	Austria, Belgium, Germany, the Netherlands, Sweden, and Switzerland
Nonreimbursed	Cyprus, the Czech Republic, Greece, Italy, and Poland

XSL•FO
RenderX

Table 2. Sample characteristics of 3123 HIV-negative men who have sex with men and trans* individuals who completed the PROTECT survey in English across 20 European countries from October 2023-April 2024.

Characteristic	Value, n (%)
Age (years)	
18-24	273 (8.7)
25-29	462 (14.8)
30-39	1037 (33.2)
40-49	702 (22.5)
50-59	374 (12)
60-69	161 (5.2)
70+	114 (3.7)
Education	
Below Bachelor's degree	705 (22.6)
Bachelor's or above	2418 (77.4)
Perceived income	
Lower	1811 (58)
Higher	1312 (42)
Migration status	
Nonmigrant	1467 (47)
Migrant	1656 (53)
Oral PrEP use	
Current users	1710 (54.8)
Noncurrent users	1413 (45.2)
Oral PrEP[a] regimen[b]	
Daily	971 (56.8)
Event-driven	487 (28.5)
Mixed-use	252 (14.7)
Oral PrEP adherence[c]	
Optimal	1056 (72.4)
Suboptimal	402 (27.6)
PrEP affordability	
Fully reimbursed	1758 (56.3)
Partially reimbursed	593 (19)
Nonreimbursed	772 (24.7)
Long-acting PrEP intention	
Yes	2306 (73.8)
No	817 (26.2)

[a]PrEP: pre-exposure prophylaxis.

[b]Available only to participants currently using oral PrEP.

[c]Available only to those who reported currently using daily or event-driven PrEP.

Statistical Analysis

Descriptive Analysis

We performed descriptive statistical analyses on all independent variables. For participants' perceptions of LA-PrEP, we generated word clouds of the 50 most frequently mentioned terms, both for the overall sample and within subgroups based on SEP, migration background, oral PrEP use, PrEP affordability, and LA-PrEP intention, to provide a visual overview.

XSL•FO
RenderX

Structural Topic Modeling

In our STM analysis, the number of topics was not prespecified but instead was determined by statistical evaluation of models with varying numbers of topics, ranging from 3 to 10. After text normalization and data cleaning (described previously in the "Measures" section), we assessed the statistical fit of each K-topic model by examining several diagnostic metrics: held-out likelihood, which reflects the (log-)likelihood of estimating the probability of unseen documents given a trained model and thus indicates predictive performance [60]; residual analysis, which evaluates how well the model captures word co-occurrence patterns by checking for remaining correlation between words after accounting for topics [61]; semantic coherence, which measures the degree to which words within a topic co-occur, reflecting topic interpretability [62]; and the lower bound, which serves as an indicator of model convergence during variational inference [62]. We considered models with higher held-out likelihood, lower residuals, and greater semantic coherence to be better fitting, and those with higher lower bound values to exhibit more robust convergence. Based on these criteria, we selected the model that demonstrated the best balance of performance and convergence across all metrics.

We then manually explored the optimal model with the best fit, evaluating the relevance of each topic in terms of the meanings and expectations of LA-PrEP among MSM and trans* individuals. This included examining the top 6 highest probability words for each topic in a thematic approach, based on frequency and exclusivity (FREX) word profiling—FREX words are weighted based on their overall FREX to a topic [45]. For the sensitivity analysis, we evaluated each topic in three other word profiles, including the highest probability words, lift words, and score words [45]. We also estimated the topic relevance, which refers to the relative importance of these topics in the overall corpus [63], by assessing the proportion of a document devoted to a certain topic [64].

Finally, to explore variations in the relevance of LA-PrEP meanings and expectations across different subgroups of MSM and trans* individuals, we conducted univariable linear topic regressions. These regressions used the topic proportion of each estimated topic as the dependent variable, considering factors such as SEP, migration background, oral PrEP use, PrEP affordability, and LA-PrEP intention based on the text corpus. Such regressions allow us to assess the conditional expectation of topic prevalence given document characteristics [65]. We conducted only univariable topic regression to explore how individual covariates relate to topic prevalence, rather than using a multivariable STM model. Given the exploratory nature of our analysis, this approach allowed us to examine the influence of each covariate independently without introducing potential confounding or multicollinearity. Several covariates in our dataset were correlated (eg, education, income, and migration status or oral PrEP use, adherence, and regimens), and including them simultaneously in a multivariable model could lead to unstable estimates and reduced interpretability. Moreover, univariable models help to avoid overfitting and ensure more stable estimation, especially in the presence of small subgroups and a relatively large number of topics. This strategy provided a clear first step in identifying patterns of interest for future hypothesis-driven analysis. All analyses were performed using R Foundation for Statistical Computing (version 4.3.2).

Ethical Considerations

This study is reported following the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for reporting observational studies (Multimedia Appendix 5) [66]. Ethical approval was obtained from the Ethics Review Committee Psychology and Neuroscience at Maastricht University (OZL_262_08_01_2023_S21). All PROTECT participants provided written informed consent at the time of enrollment. The data collected were fully anonymized and cannot be traced back to individual participants. No financial or other form of compensation was provided to participants in the PROTECT study.

Results

Sample Characteristics

In this study, 3123 HIV-negative MSM and trans* individuals completed the PROTECT survey in English and were included, with a median age of 37 (IQR 30-46) years. Of these participants, most had an educational attainment level of a bachelor's degree or above (n=2418, 77%), more than half perceived a lower income (n=1811, 58%), and more than half reported having a migration background (n=1656, 53%). Among those who reported a migration background, the majority (n=2500, 62%; Table S1 in Multimedia Appendix 1) were living in non-English–speaking countries such as Germany (n=214, 12.9%) and the Netherlands (n=194, 11.7%), while 38% (n=623) were living in English-speaking countries, including Ireland (n=147, 8.9%) and the United Kingdom (n=476, 28.7%).

Overall, more than half of the participants (n=1710, 55%) were currently using oral PrEP, with the majority (n=971, 57%) reporting daily PrEP use, followed by event-driven PrEP (n=487, 28%) and 15% (n=252) reporting a mix of daily and event-driven PrEP use. Notably, a significant proportion of those using daily and event-driven PrEP reported suboptimal adherence (n=402, 28%). Regarding PrEP affordability, the majority lived in countries where oral PrEP was fully reimbursed (n=1758, 56%), while 19% (n=n=772) lived in countries where oral PrEP was not reimbursed. In addition, most participants (n=2306, 74%) indicated their intention to use LA-PrEP if it became available and affordable. Detailed sample characteristics are provided in Table 2.

Word Associations With LA-PrEP

The overall word cloud highlighting the 50 most frequently mentioned words associated with LA-PrEP is shown in Figure S2 in Multimedia Appendix 1. Descriptively, the responses indicated generally positive associations with LA-PrEP. The most frequently mentioned words were "safe," followed by "freedom," and "convenient." However, ambivalent and relatively negative perceptions were also noted, including terms such as "nothing," "unknown," "dunno," and "unnecessary," as well as perceived side effects such as "pain."

Variations in perceptions of LA-PrEP were observed across different SEP subgroups. Lower-income participants frequently

XSL•FO
RenderX

used terms such as “safe” and more ambivalent phrases such as “nothing” or “unknown,” while higher-income participants emphasized “convenient” (Figure S1 in Multimedia Appendix 1). Education also showed notable differences: those with a bachelor’s degree or higher often mentioned “freedom” and “convenient,” whereas individuals with lower education mostly used “safe” and fewer terms overall (Figure S2c and S2d in Multimedia Appendix 1). Similar patterns emerged between participants with and without a migration background (Figure S1 in Multimedia Appendix 1).

Among oral PrEP use subgroups, current users highlighted LA-PrEP as “convenient” and associated it with “freedom,” whereas nonusers focused on “safe” and used more ambivalent terms such as “nothing” and “dunno” (Figure S3A and S3B in Multimedia Appendix 1). Daily PrEP users anticipated greater convenience and expressed fewer negative views compared with nondaily users (Figure S2C and S2D in Multimedia Appendix 1). Furthermore, users with suboptimal adherence expected more “freedom” from LA-PrEP and had fewer negative perceptions than those with optimal adherence (Figure S2e and S2f in Multimedia Appendix 1).

Variations were also noted regarding PrEP affordability and intentions to use LA-PrEP. Participants from countries with full PrEP reimbursement had more positive views and frequently used “freedom,” while those from countries with partial or no reimbursement mentioned it less (Figure S4A-C in Multimedia Appendix 1). A significant difference was observed between participants with high and low intentions to use LA-PrEP. Those with high intentions expressed almost exclusively positive views, while those with low intentions frequently used ambivalent or negative terms such as “nothing,” “unknown,” “uncertainty,” and “unnecessary” (Figure S4D and S4E in Multimedia Appendix 1). Detailed word frequencies are shown in Multimedia Appendix 3.

Structural Topic Modeling

Model Selection

The model with 5 topics was selected based on a comprehensive evaluation of multiple diagnostic metrics, which collectively indicated an optimal balance between model fit, interpretability, and generalizability (Figure S5 in Multimedia Appendix 1). In terms of held-out likelihood, a measure of predictive performance, the top-performing models were those with 3, 5, and 10 topics. Although the models with 3 and 10 topics showed slightly higher held-out likelihoods, the model with 10 topics also had the highest residuals, suggesting potential overfitting and a poorer capture of word co-occurrence patterns. This could introduce greater complexity and increase the risk of thematic redundancy. Conversely, the model with 3 topics, while simpler, had the lowest semantic coherence and weaker convergence based on lower bound values. In contrast, the model with 5 topics offered the best overall trade-off, demonstrating improved held-out likelihood, higher semantic coherence, lower residuals, and robust model convergence. Therefore, it was selected as the most balanced and interpretable solution.

Topic Interpretation

Table 3 presents the FREX word profile, interpretation, and relevance of each topic of the meanings and expectations of LA-PrEP. Topic 1 (safety) captures the expectation that LA-PrEP will be a safe option with long-lasting protection against HIV. Topic 2 (empowerment) emphasizes the perception that LA-PrEP is a suitable method of HIV prevention that increases the sense of agency among future users. Topic 3 (peace of mind [POM]) highlights that LA-PrEP is an innovative option that is practical to use and provides a feeling of calmness and nonagitation. Topic 4 (convenience and reliability) focused on confidence in the effectiveness of LA-PrEP. Topic 5 (concerns and uncertainties), conversely, reflected ambivalent and relatively negative perceptions of LA-PrEP. No major differences were found between different word profiling methods (Table S2 in Multimedia Appendix 1).

Other word profiles based on other methods can be found in Table S1.

Based on our document-topic proportion estimations (Table 2; Figure S6 in Multimedia Appendix 1), the topic of empowerment emerged as the most relevant, accounting for 28% of the associations with and expectations of LA-PrEP. This suggests that nearly one third of MSM and trans* participants anticipated that LA-PrEP would be an empowering HIV prevention option. The second most prominent topic was safety (21%), reflecting the importance placed on the superior safety profile of LA-PrEP. Following this were the topics of concerns and uncertainties (20%), convenience and reliability (16%), and POM (15%).

XSL•FO
RenderX

Table 3. Topic word profiles, relevance, and interpretations of the open responses to the question, "What does LA-PrEP mean to you?" from 3123 HIV-negative men who have sex with men and trans* individuals who completed the PROTECT survey in English across 20 European countries from October 2023-April 2024.

Topic	Relevance	FREX[a] word profile	Interpretation	Interpretation keyword
Topic 1	20.7%	Long, safe, last, protect, access, and worry	LA-PrEP[b] will be a safe option that has long-lasting protection	Safety
Topic 2	28.1%	Freedom, easy, interest, alternative, pain, and better	Even though LA-PrEP may lead to some pain, it is an alternative that can be interesting, easy, and empowering for more freedom	Empowerment
Topic 3	15%	Reassurance, secure, practical, simple, innovative, and use	LA-PrEP as an innovative option may bring POM[c] by being practical	POM
Topic 4	16.2%	Convenient, prevent, reliable, hope, inject, and HIV	LA-PrEP is an injectable PrEP[d] for HIV that can be convenient and reliable	Convenient and reliable
Topic 5	20%	Nothing, option, unknown, NA[e], unnecessary, and dunno	Concerns and uncertainties on LA-PrEP	Concerns and uncertainties

[a]FREX: frequency and exclusivity.

[b]LA-PrEP: long-acting pre-exposure prophylaxis.

[c]POM: peace of mind.

[d]PrEP: pre-exposure prophylaxis.

[e]NA: not available.

Topic Regressions

Multimedia Appendix 4 summarizes the topic regressions between each estimated topic and SEP, migration background, oral PrEP use, PrEP affordability, and LA-PrEP intention subgroups. For the topic of safety, compared to those with a bachelor's degree or higher, participants with education below a bachelor's degree were significantly more likely to discuss this theme, with a 5.2-percentage point higher expected topic probability (=.052, 95% CI 0.022-0.083). Safety concerns were also 3.5 percentage points more prevalent among participants living in countries where PrEP was not reimbursed (=.035, 95% CI 0.002-0.069). In contrast, current oral PrEP users were 9.4 percentage points less likely to discuss safety as a concern (=–.094, 95% CI –0.123 to 0.066). The topic of empowerment was significantly more prevalent among current oral PrEP users, who showed a 7.0 percentage point higher expected topic probability compared to nonusers (=.070, 95% CI 0.042-0.097), and among those with a higher intention to use LA-PrEP (5.1 percentage point increase; =.051, 95% CI 0.017-0.085). In contrast, it was 5.2 percentage points less prevalent among participants living in countries where PrEP was not reimbursed (=–.052, 95% CI –0.087 to 0.017). The topic of LA-PrEP being convenient or reliable showed a similar pattern. It was 5.7 percentage points more prevalent among current oral PrEP users (=.057, 95% CI 0.034-0.081) and 3.2 percentage points more prevalent among those with a higher intention to use LA-PrEP (=.032, 95% CI 0.003-0.062). However, it was 3.7 percentage points less prevalent among individuals currently using event-driven oral PrEP (=–.037, 95% CI –0.078 to 0.003) and 3.4 percentage points less prevalent among participants in countries where PrEP was nonreimbursed (=–.034, 95% CI –0.064 to –0.005). For the topic of concerns or uncertainties, prevalence was 3.2 percentage points lower among current oral PrEP users (=–.032, 95% CI –0.056 to –0.007) and 7.5 percentage points lower among those with higher intention to use LA-PrEP (=–.075, 95% CI –0.101 to –0.005). For the topic of POM, we did not find any significant differences across the subgroups.

Discussion

Principal Findings

To our knowledge, this is the first study to explore views and perceptions of LA-PrEP in Europe and the first to apply STM to large-scale qualitative data from open-ended survey responses on LA-PrEP perceptions. Our findings demonstrate that machine learning methods can effectively analyze large-scale qualitative data in HIV prevention, enhancing the capabilities of human coders. These techniques can be applied to other topics where qualitative evidence is needed. Consistent with a previous systematic review [16], this study found generally positive views of LA-PrEP among MSM and trans* individuals across Europe, suggesting that it is likely to be well received alongside existing oral PrEP options once it becomes available. However, the study also uncovered significant concerns and uncertainties, accounting for 20% of topic relevance, regarding LA-PrEP among key populations. While this proportion reflects meaningful apprehension, it is somewhat lower than levels of concern typically reported in earlier qualitative research on novel biomedical interventions, such as initial PrEP rollout [67,68] or vaccine hesitancy studies [69,70]. Given that LA-PrEP is a new and potentially more invasive prevention method compared with oral regimens [71,72], ongoing attention to these concerns is warranted. We need to acknowledge that LA-PrEP may not be suitable for everyone, and alternative prevention options should remain accessible to maximize the public health impact of the current HIV prevention strategies.

XSL•FO
RenderX

Our models identified 5 main topics within these views and perceptions. Three of these were notably positive. These positive sentiments align with the intended benefits of LA-PrEP, which aim to address some of the challenges of oral PrEP [15,22]. Notably, the topics of empowerment and convenience and reliability were more relevant among current oral PrEP users, who also showed a higher intention to use LA-PrEP [28]. In contrast, these topics were less relevant to individuals in less privileged contexts, such as those in countries where oral PrEP is not or not fully reimbursed. Our findings thus underscore the importance of oral PrEP as a HIV prevention modality—without optimal oral PrEP services, LA-PrEP might not be viewed positively and might not effectively address the remaining unmet needs that key populations may experience. Further, the perception of LA-PrEP as convenient and reliable was less relevant for those using event-driven PrEP. Since this population typically experiences fewer challenges in using oral PrEP [73], LA-PrEP may not be as appealing to them as it is to other groups.

Conversely, the topic of the superior safety profile of LA-PrEP compared with oral PrEP was more relevant to individuals in less privileged circumstances, such as those with lower educational attainment or those living in countries where oral PrEP is not fully reimbursed. However, this topic was significantly less relevant to current oral PrEP users. One possible reason is that our study was conducted after the publication of cabotegravir efficacy data, which may be more accessible to those with greater privileges. Another reason could be the positive safety experiences reported by current oral PrEP users [74], suggesting that the safety profile of LA-PrEP may not be their primary concern.

Interestingly, concerns and uncertainties about LA-PrEP were significantly less relevant to current oral PrEP users, while remaining important for other subgroups. This aligns with previous findings that current oral PrEP users are more likely to adopt LA-PrEP compared with PrEP-naïve MSM [28,31]. Our qualitative evidence suggests that current users, familiar with and positive about PrEP [74], have fewer concerns about LA-PrEP as a new HIV prevention option, which may explain their higher intention to use it. Conversely, individuals with concerns or uncertainties are less likely to intend to use LA-PrEP. To facilitate the adoption of LA-PrEP, it is crucial for HIV prevention authorities, health care providers, and communities to address these concerns, especially among those who show interest. In addition, guidelines and support should be developed to help individuals navigate periods of use and nonuse or to switch between prevention regimens [6], ensuring they can make informed decisions.

Limitations

Our study has some limitations. First, our analysis was based solely on an English-language corpus, which included responses from migrants from non-English–speaking countries who chose to complete the survey in English, as well as participants residing in the United Kingdom and Ireland. This choice was also driven by the lingua franca role of English in Europe. Translating non-English entries into English could potentially create translation or semantic bias and increase the risk of bias due to language-to-country matching in terms of determinants in the regression analysis. The fact that the regression results do not only exclusively reflect the situation in the United Kingdom and Ireland supports broader representation. To mitigate this risk, we prioritized semantic accuracy over broader representativeness. As a result, the generalizability and external validity of our findings may be limited, particularly among speakers of other mother tongues. Future research should explore perceptions of LA-PrEP in local languages to generate more region-specific and culturally specific insights, especially in settings where English is not the primary language. Second, our sample was recruited via an internet-based survey, likely resulting in an overrepresentation of digitally literate individuals with greater health engagement. This may limit the external validity of our findings, particularly for populations with limited internet access or digital literacy. Finally, we did not conduct in-depth interviews and relied solely on responses to open-ended survey questions. While STM is an innovative tool for analyzing large-scale qualitative data, it relies on probabilistic modeling, which may result in the loss of important contexts and meanings, introducing potential bias. Despite this, our findings are consistent with previous research outside Europe [16], supporting their ecological validity. Future work should complement these findings with qualitative interviews to provide deeper insight.

Conclusions

Our research reveals a generally positive outlook on LA-PrEP, suggesting that it is likely to be well-accepted upon its introduction in Europe. However, notable concerns remain, indicating that a subset of the target population may be hesitant to adopt LA-PrEP, highlighting the need for alternative HIV prevention strategies tailored to these individuals. These concerns must be addressed among potential LA-PrEP users to ensure a successful implementation. To achieve broad coverage and maximize public health impact, the introduction of LA-PrEP must include evidence-based, tailored public health communication strategies that promote informed decision-making among diverse populations, thereby contributing to the global effort to end the HIV epidemic.

Acknowledgments

We thank all participants in our study. The results of this study were presented at the AIDSImpact conference in May 2025. The PROTECT survey was partly funded by ViiV Healthcare (grant 219497). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

HW: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Visualization; Writing – original draft; Writing – review & editing
JK: Conceptualization; Data curation; Project administration; Writing – review & editing
JCM: Writing – review & editing
DvdV: Writing – review & editing
JT: Writing – review & editing
MS: Writing – review & editing
AA: Writing – review & editing
HMLZ: Conceptualization; Data curation; Writing – review & editing
KJJ: Conceptualization; Data curation; Investigation; Project administration; Resources; Supervision; Validation; Writing – review & editing

Conflicts of Interest

HW, JK, JCM, HMLZ, and KJJ report no direct competing interests to disclose. JT, MS, and AA were employed by ViiV Healthcare at the time of the study.

Multimedia Appendix 1
Additional figures and tables.
[DOCX File , 6857 KB - publichealth_v11i1e72491_app1.docx]

Multimedia Appendix 2
Raw data and entries after manual cleaning.
[XLSX File (Microsoft Excel File), 62 KB - publichealth_v11i1e72491_app2.xlsx]

Multimedia Appendix 3
Detailed word frequencies.
[XLSX File (Microsoft Excel File), 76 KB - publichealth_v11i1e72491_app3.xlsx]

Multimedia Appendix 4
Topic regression of the open responses of “What LA-PrEP means to you?” question from 3123 HIV-negative MSM and trans* individuals who completed the PROTECT survey in English across 20 European countries, October 2023-April 2024.
[DOCX File , 24 KB - publichealth_v11i1e72491_app4.docx]

Multimedia Appendix 5
STROBE checklist.
[DOC File , 82 KB - publichealth_v11i1e72491_app5.doc]

References

1. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, iPrEx Study Team. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]
2. Molina J, Capitant C, Spire B, Pialoux G, Cotte L, Charreau I, et al. On-Demand Preexposure Prophylaxis in Men at High Risk for HIV-1 Infection. N Engl J Med 2015 Dec 03;373(23):2237-2246. [doi: 10.1056/NEJMoa1506273] [Medline: 26624850]
3. Koechlin FM, Fonner VA, Dalglish SL, O'Reilly KR, Baggaley R, Grant RM, et al. Values and preferences on the use of oral pre-exposure prophylaxis (PrEP) for HIV prevention among multiple populations: A systematic review of the literature. AIDS Behav 2017;21(5):1325-1335 [FREE Full text] [doi: 10.1007/s10461-016-1627-z] [Medline: 27900502]
4. Wang H, Shobowale O, den Daas C, Op de Coul E, Bakker B, Radyowijati A, et al. Determinants of PrEP uptake, intention and awareness in the Netherlands: A socio-spatial analysis. Int J Environ Res Public Health 2022;19(14):8829 [FREE Full text] [doi: 10.3390/ijerph19148829] [Medline: 35886681]
5. UNAIDS. The urgency of now: AIDS at a crossroads. 2024 UNAIDS Global AIDS Update. 2024. URL: https://www.unaids.org/sites/default/files/media_asset/2024-unaids-global-aids-update-summary_en.pdf [accessed 2025-08-14]

XSL•FO
RenderX

6. Bavinton BR, Grulich AE. HIV pre-exposure prophylaxis: scaling up for impact now and in the future. Lancet Public Health 2021 Jul;6(7):e528-e533 [FREE Full text] [doi: 10.1016/S2468-2667(21)00112-2] [Medline: 34087117]
7. Wang H, Molina JM, Dray-Spira R, Schmidt AJ, Hickson F, van de Vijver D, et al. Spatio-temporal changes in pre-exposure prophylaxis uptake among MSM in mainland France between 2016 and 2021: a Bayesian small area approach with MSM population estimation. J Int AIDS Soc 2023;26(5):e26089 [FREE Full text] [doi: 10.1002/jia2.26089] [Medline: 37221971]
8. van Dijk M, de Wit JBF, Guadamuz TE, Martinez JE, Jonas KJ. Slow uptake of PrEP: Behavioral predictors and the influence of price on PrEP uptake among MSM with a high interest in PrEP. AIDS Behav 2021;25(8):2382-2390 [FREE Full text] [doi: 10.1007/s10461-021-03200-4] [Medline: 33611697]
9. Coyer L, Hoornenborg E. Reaching the full preventive potential of HIV pre-exposure prophylaxis. Lancet Public Health 2022 Jun;7(6):e488-e489 [FREE Full text] [doi: 10.1016/S2468-2667(22)00116-5] [Medline: 35660206]
10. Zhang J, Li C, Xu J, Hu Z, Rutstein SE, Tucker JD, et al. Discontinuation, suboptimal adherence, and reinitiation of oral HIV pre-exposure prophylaxis: a global systematic review and meta-analysis. Lancet HIV 2022 Apr;9(4):e254-e268 [FREE Full text] [doi: 10.1016/S2352-3018(22)00030-3] [Medline: 35364026]
11. Wang H, Delnoij M, Zimmermann HML, Jonas KJ. The Disparities of PrEP Adherence Among Men Who Have Sex With Men Between the Global South and the Global North: An Updated Determinantal Global Meta-Analysis. J Acquir Immune Defic Syndr 2025 May 01;99(1):1-8. [doi: 10.1097/QAI.0000000000003635] [Medline: 39878582]
12. Zimmermann HM, Eekman SW, Achterbergh RC, Schim van der Loeff MF, Prins M, de Vries HJ, Amsterdam PrEP Project team in the HIV Transmission Elimination AMsterdam Consortium (H-TEAM). Motives for choosing, switching and stopping daily or event-driven pre-exposure prophylaxis - a qualitative analysis. J Int AIDS Soc 2019;22(10):e25389 [FREE Full text] [doi: 10.1002/jia2.25389] [Medline: 31612621]
13. Celum C, Grinsztejn B, Ngure K. Preparing for long-acting PrEP delivery: building on lessons from oral PrEP. J Int AIDS Soc 2023;26(Suppl 2):e26103 [FREE Full text] [doi: 10.1002/jia2.26103] [Medline: 37439077]
14. Delany-Moretlwe S, Flexner C, Bauermeister JA. Advancing use of long-acting and extended delivery HIV prevention and treatment regimens. J Int AIDS Soc 2023;26(Suppl 2):e26126 [FREE Full text] [doi: 10.1002/jia2.26126] [Medline: 37439079]
15. Haberer JE, Mujugira A, Mayer KH. The future of HIV pre-exposure prophylaxis adherence: reducing barriers and increasing opportunities. Lancet HIV 2023 Jun;10(6):e404-e411. [doi: 10.1016/S2352-3018(23)00079-6] [Medline: 37178710]
16. Lorenzetti L, Dinh N, van der Straten A, Fonner V, Ridgeway K, Rodolph M, et al. Systematic review of the values and preferences regarding the use of injectable pre-exposure prophylaxis to prevent HIV acquisition. J Int AIDS Soc 2023;26(Suppl 2):e26107 [FREE Full text] [doi: 10.1002/jia2.26107] [Medline: 37439057]
17. Pike C, Bekker LG. Interrogating the promise of long-acting HIV pre-exposure prophylaxis. Trends Mol Med 2023;29(2):93-98. [doi: 10.1016/j.molmed.2022.11.003] [Medline: 36435632]
18. Mitchell KM, Boily MC, Hanscom B, Moore M, Todd J, Paz-Bailey G, et al. Estimating the impact of HIV PrEP regimens containing long-acting injectable cabotegravir or daily oral tenofovir disoproxil fumarate/emtricitabine among men who have sex with men in the United States: a mathematical modelling study for HPTN 083. Lancet Reg Health Am 2023 Feb;18:100416 [FREE Full text] [doi: 10.1016/j.lana.2022.100416] [Medline: 36844011]
19. Stansfield SE, Heitner J, Mitchell KM, Doyle CM, Milwid RM, Moore M, et al. Population-level impact of expanding PrEP coverage by offering long-acting injectable PrEP to MSM in three high-resource settings: a model comparison analysis. J Int AIDS Soc 2023;26 Suppl 2(Suppl 2):e26109 [FREE Full text] [doi: 10.1002/jia2.26109] [Medline: 37439080]
20. Wang H, Popping S, van de Vijver D, Jonas KJ. Epidemiological impact and cost-effectiveness analysis of PrEP provision expansion among MSM in the Netherlands. J Int AIDS Soc 2025;28(6):e26516. [doi: 10.1002/jia2.26516] [Medline: 40462500]
21. Wang H, Jonas K, Van DVD. Epidemiological impact of targeting long-acting injectable cabotegravir to MSM with Low PrEP use adherence in the Netherlands: A mathematical modelling study. The Lancet 2023. [doi: 10.2139/ssrn.4415665]
22. Landovitz RJ, Donnell D, Clement ME, Hanscom B, Cottle L, Coelho L, HPTN 083 Study Team. Cabotegravir for HIV Prevention in Cisgender Men and Transgender Women. N Engl J Med 2021 Aug 12;385(7):595-608 [FREE Full text] [doi: 10.1056/NEJMoa2101016] [Medline: 34379922]
23. Delany-Moretlwe S, Hughes JP, Bock P, Ouma SG, Hunidzarira P, Kalonji D, HPTN 084 study group. Cabotegravir for the prevention of HIV-1 in women: results from HPTN 084, a phase 3, randomised clinical trial. Lancet 2022;399(10337):1779-1789 [FREE Full text] [doi: 10.1016/S0140-6736(22)00538-4] [Medline: 35378077]
24. Henderson M. Guidelines on Long-acting Injectable Cabotegravir for HIV Prevention: web annex F: perspectives and preferences regarding long-acting injectable cabotegravir for HIV prevention among pre-exposure prophylaxis (PrEP) providers: a global survey and in-depth Interviews. Geneva: World Health Organization; 2022.
25. FDA approves first injectable treatment for HIV pre-exposure prevention. FDA. 2021. URL: https://www.fda.gov/news-events/press-announcements/fda-approves-first-injectable-treatment-hiv-pre-exposure-prevention [accessed 2025-08-14]
26. ViiV Healthcare's cabotegravir for HIV prevention receives positive CHMP opinion from European Medicines Agency. ViiV Healthcare. 2023. URL: https://www.gsk.com/en-gb/media/press-releases/viiv-healthcare-s-cabotegravir-for-hiv-prevention-receives-positive-chmp-opinion-from-european-medicines-agency/ [accessed 2025-08-14]

XSL•FO
RenderX

27. Bekker LG, Das M, Abdool Karim Q, Ahmed K, Batting J, Brumskine W, PURPOSE 1 Study Team. Twice-Yearly Lenacapavir or Daily F/TAF for HIV Prevention in Cisgender Women. N Engl J Med 2024 Oct 03;391(13):1179-1192. [doi: 10.1056/NEJMoa2407001] [Medline: 39046157]
28. Wang H, Zimmermann HML, van de Vijver D, Jonas KJ. Intention and preference to use long-acting injectable PrEP among MSM in the Netherlands: a diffusion of innovation approach. AIDS Care 2024;36(sup1):89-100 [FREE Full text] [doi: 10.1080/09540121.2024.2307378] [Medline: 38713631]
29. Van Landeghem E, Dielen S, Semaan A, Rotsaert A, Vanhamel J, Masquillier C, et al. Insights into barriers and facilitators in PrEP uptake and use among migrant men and transwomen who have sex with men in Belgium. BMC Public Health 2023;23(1):712 [FREE Full text] [doi: 10.1186/s12889-023-15540-y] [Medline: 37076823]
30. Tieosapjaroen W, Zhang Y, Fairley CK, Zhang L, Chow EPF, Phillips TR, et al. Improving access to oral pre-exposure prophylaxis for HIV among international migrant populations. Lancet Public Health 2023 Aug;8(8):e651-e658 [FREE Full text] [doi: 10.1016/S2468-2667(23)00105-6] [Medline: 37421969]
31. Wang H, Kolstee J, Adriaque LA, Aphami L, Gaetani M, Zimmermann H. Interest in and preference for long-acting injectable PrEP among men who have sex with men, trans* individuals, and cis-gender heterosexual women: a global systematic review and meta-analysis. medRxiv 2024. [doi: 10.1097/QAI.0000000000003754] [Medline: 40862307]
32. Venturelli S, Mezzadri L, Bana N, Comi L, Borghi F, Ouabou A. SC-13-Impact of PrEP cost on prophylaxis initiation in an Italian centre. Sex Transm Infect 2024;100(Suppl 1):A80. [doi: 10.1136/sextrans-ICAR-2024.74]
33. Zimmermann H, Wang H, Schroeder M, Appiah A. Preparing for long-acting PrEP delivery: Provider preferences for the provision of long-acting PrEP differ between MSM and heterosexual individuals in Europe. J Int AIDS Soc 2024;27(S6):e26370 [FREE Full text]
34. Mayer CM, Owaraganise A, Kabami J, Kwarisiima D, Koss CA, Charlebois ED, et al. Distance to clinic is a barrier to PrEP uptake and visit attendance in a community in rural Uganda. J Int AIDS Soc 2019;22(4):e25276 [FREE Full text] [doi: 10.1002/jia2.25276] [Medline: 31037845]
35. Glick JL, Nestadt DF, Sanchez T, Li K, Hannah M, Rawlings MK, et al. Location preferences for accessing long-acting injectable pre-exposure prophylaxis (LA-PrEP) among men who have sex with men (MSM) currently on daily oral PrEP. J Int Assoc Provid AIDS Care 2024;23:23259582241293336 [FREE Full text] [doi: 10.1177/23259582241293336] [Medline: 39698980]
36. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. Soc Sci Med 2022;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
37. Hubert M, Bajos N, Sandfort T. Sexual Behaviour and HIV/AIDS in Europe: Comparisons of National Surveys. Hong Kong: Routledge; 2022.
38. Weatherburn P, Hickson F, Reid DS, Marcus U, Schmidt AJ. European men-who-have-sex-with-men internet survey (EMIS-2017): Design and methods. Sex Res Soc Policy 2019;17(4):543-557. [doi: 10.1007/s13178-019-00413-0]
39. Reyes-Díaz M, Celly A, Folch C, Lorente N, Stuardo V, Veras MA, et al. Latin American internet survey for men who have sex with men (LAMIS-2018): Design, methods and implementation. PLoS One 2022;17(11):e0277518 [FREE Full text] [doi: 10.1371/journal.pone.0277518] [Medline: 36395121]
40. Lee S, Kim L, Shim MS, Kim GS. Identifying health care services offered in the HIV care continuum via a machine learning-based topic modeling approach: Exploratory literature review. JMIR Public Health Surveill 2025;11:e65081-e65081 [FREE Full text] [doi: 10.2196/65081] [Medline: 40632764]
41. Keddem S, Agha A, Morawej S, Buck A, Cronholm P, Sonalkar S, et al. Characterizing Twitter content about HIV pre-exposure prophylaxis (PrEP) for women: Qualitative content analysis. J Med Internet Res 2023;25:e43596 [FREE Full text] [doi: 10.2196/43596] [Medline: 37166954]
42. Erdengasileng A, Tian S, Green SS, Naar S, He Z. Using Twitter data analysis to understand the perceptions, awareness, and Barriers to the wide use of pre-exposure prophylaxis in the United States. Proceedings (IEEE Int Conf Bioinformatics Biomed) 2022;2022:3000-3007 [FREE Full text] [doi: 10.1109/bibm55620.2022.9995568] [Medline: 36818418]
43. Klein AZ, Meanley S, O'Connor K, Bauermeister JA, Gonzalez-Hernandez G. Toward using Twitter for PrEP-related interventions: An automated natural language processing pipeline for identifying gay or bisexual men in the United States. JMIR Public Health Surveill 2022;8(4):e32405 [FREE Full text] [doi: 10.2196/32405] [Medline: 35468092]
44. Roberts ME, Stewart BM, Tingley D, Airoldi EM. The structural topic model and applied social science. 2013 Presented at: International Conference on Neural Information Processing; 2013; New Zealand.
45. Roberts ME, Stewart BM, Tingley D, Lucas C, Leder‐Luis J, Gadarian SK, et al. Structural topic models for open‐ended survey responses. American J Political Sci 2014;58(4):1064-1082. [doi: 10.1111/ajps.12103]
46. Eshima S, Imai K, Sasaki T. Keyword‐Assisted Topic Models. American J Political Sci 2023;68(2):730-750. [doi: 10.1111/ajps.12779]
47. Maier D, Waldherr A, Miltner P, Wiedemann G, Niekler A, Keinert A, et al. Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. CMM 2018;12(2-3):93-118. [doi: 10.1080/19312458.2018.1430754]
48. Abuzayed A, Al-Khalifa H. BERT for Arabic topic modeling: An experimental study on BERTopic technique. Procedia Computer Science 2021;189:191-194. [doi: 10.1016/j.procs.2021.05.096]

XSL•FO
RenderX

49. Grimmer J, Stewart BM. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. Polit anal 2017;21(3):267-297. [doi: 10.1093/pan/mps028]
50. Mourtgos SM, Adams IT. The rhetoric of de-policing: Evaluating open-ended survey responses from police officers with machine learning-based structural topic modeling. J Crim Justice 2019;64:101627. [doi: 10.1016/j.jcrimjus.2019.101627]
51. Kolstee J, Wang H, Zimmermann H, Schroeder M, Appiah A, Brown C. Design and methods of the - understanding pre-exposure prophylaxis modalities for HIV prevention in European communities - (PROTECT) survey. medRxiv 2024:6697. [doi: 10.1101/2024.11.04.24316697]
52. Wang H, Kolstee J, Casalini JL, Hakim S, Zimmermann H, Jonas K. Likelihood of HIV and recent bacterial sexually transmitted infections among transgender and non-binary individuals in 20 European countries, October 2023 to April 2024. Euro Surveill 2024;29(48):2400347. [doi: 10.2807/1560-7917.ES.2024.29.48.2400347] [Medline: 39611207]
53. Wang H, Lozano AA, Kolstee J, Zimmermann HML, Tosh J, Schroeder M, et al. Oral PrEP use and intention to use long-acting PrEP regimens among MSM accessing PrEP via governmental and non-governmental provision pathways, 20 European countries, October 2023 to April 2024. Eurosurveillance 2025 Aug 28;30(34) [FREE Full text] [doi: 10.2807/1560-7917.ES.2025.30.34.2500122]
54. Hickman L, Thapa S, Tay L, Cao M, Srinivasan P. Text preprocessing for text mining in organizational research: Review and recommendations. Organ Res Methods 2020;25(1):114-146. [doi: 10.1177/1094428120971683]
55. Roberts ME, Stewart BM, Tingley D. stm: An R Package for Structural Topic Models. J Stat Soft 2019;91(2):1-40. [doi: 10.18637/jss.v091.i02]
56. Blei DM, Lafferty JD. A correlated topic model of science. Ann Appl Stat 2007;1(1). [doi: 10.1214/07-aoas114]
57. Chakraborty G, Pagolu M, Garla S. Text Mining and Analysis: Practical Methods, Examples, and Case Studies Using SAS. North Carolina: SAS Institute; 2014.
58. Srivastava AN, Sahami M. Text mining: Classification, Clustering, and Applications. New York: CRC press; 2009.
59. Blei D, Lafferty J. Dynamic Topic Models. USA: Princeton University; 2006:113-120.
60. Wallach HM, Murray I, Salakhutdinov R, Mimno D. Evaluation methods for topic models. : Association for Computing Machinery; 2009 Presented at: Proceedings of the 26th Annual International Conference on Machine Learning; 2009; Montreal, Quebec, Canada p. 1105-1112. [doi: 10.1145/1553374.1553515]
61. Taddy M, Neil DL, Mark G. On estimation and selection for topic models. 2012 Presented at: Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics; Proceedings of Machine Learning Research; 2012; New York p. 1184-1193.
62. Mimno D, Wallach H, Talley E, Leenders M, McCallum A. Optimizing semantic coherence in topic models. 2011 Presented at: Proceedings of the Conference on Empirical Methods in Natural Language Processing; 2011 Jul 27; United States p. 262-272.
63. Grajzl P, Murrell P. Characterizing a legal–intellectual culture: Bacon, Coke, and seventeenth-century England. Cliometrica 2020;15(1):43-88. [doi: 10.1007/s11698-020-00202-5]
64. Mostafa MM. A one-hundred-year structural topic modeling analysis of the knowledge structure of international management research. Qual Quant 2022;57(4):1-31 [FREE Full text] [doi: 10.1007/s11135-022-01548-w] [Medline: 36249708]
65. Roberts ME, Stewart BM, Nielsen RA. Adjusting for confounding with text matching. American J Political Sci 2020;64(4):887-903. [doi: 10.1111/ajps.12526]
66. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Bull World Health Organ 2007;85(11):867-872 [FREE Full text] [doi: 10.2471/blt.07.045120] [Medline: 18038077]
67. Daniels B, Chetty T. Assessing the value of oral HIV pre-exposure prophylaxis for men who have sex with men in Benin. Lancet Glob Health 2025 Jun;13(6):e979-e980 [FREE Full text] [doi: 10.1016/S2214-109X(25)00147-0] [Medline: 40412403]
68. Jiang H, Zou H. PrEParing for HIV prevention among men who have sex with men in China: challenges and solutions. Lancet Glob Health 2025. [doi: 10.1016/s2214-109x(25)00234-7]
69. Wang H, Varol T, Gültzow T, Zimmermann HML, Ruiter RAC, Jonas KJ. Spatio-temporal distributions of COVID-19 vaccine doses uptake in the Netherlands: a Bayesian ecological modelling analysis. Epidemiol Infect 2024 Oct 07;152:e119. [doi: 10.1017/S0950268824001249] [Medline: 39370683]
70. Toshkov D. What accounts for the variation in COVID-19 vaccine hesitancy in Eastern, Southern and Western Europe? Vaccine 2023;41(20):3178-3188 [FREE Full text] [doi: 10.1016/j.vaccine.2023.03.030] [Medline: 37059674]
71. Marcus JL, Weddle A, Kelley CF, Agwu A, Montalvo S, Sherman E, et al. Policy recommendations to support equitable access to long-acting injectables for human immunodeficiency virus prevention and treatment: A policy paper of the infectious diseases society of America and the HIV medicine association. Clin Infect Dis 2025:e648. [doi: 10.1093/cid/ciae648] [Medline: 39873495]
72. Henderson M, Schmidt HA, Chitembo L, Peralta H, Alaama AS, Johnson C, et al. The future of pre-exposure prophylaxis (PrEP) for HIV prevention: A global qualitative consultation on provider perspectives on new products and differentiated service delivery. AIDS Behav 2023;27(11):3755-3766 [FREE Full text] [doi: 10.1007/s10461-023-04093-1] [Medline: 37351685]

XSL•FO
RenderX

73. Zimmermann HML, Jongen VW, Boyd A, Hoornenborg E, Prins M, de Vries HJC, Amsterdam PrEP Project Team in the HIV Transmission Elimination Amsterdam Initiative (H-TEAM). Decision-making regarding condom use among daily and event-driven users of preexposure prophylaxis in the Netherlands. AIDS 2020 Dec 01;34(15):2295-2304. [doi: 10.1097/QAD.0000000000002714] [Medline: 33196494]
74. van Dijk M, de Wit JBF, Kamps R, Guadamuz TE, Martinez JE, Jonas KJ. Socio-sexual experiences and access to healthcare among informal PrEP users in the Netherlands. AIDS Behav 2021;25(4):1236-1246 [FREE Full text] [doi: 10.1007/s10461-020-03085-9] [Medline: 33196938]

Abbreviations

FREX: frequency and exclusivity
LA: long-acting
POM: peace of mind
PrEP: pre-exposure prophylaxis
SEP: socioeconomic position
STM: structural topic modeling
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WHO: World Health Organization

Edited by A Mavragani; submitted 11.02.25; peer-reviewed by S Tian, S Kumi, B Ssuna; comments to author 18.07.25; revised version received 05.08.25; accepted 08.08.25; published 12.09.25.

Please cite as:
Wang H, Kolstee J, Martinez JC, van de Vijver D, Tosh J, Schroeder M, Appiah A, Zimmermann HML, Jonas KJ
Perceptions of Long-Acting Injectable Pre-Exposure Prophylaxis Among Men Who Have Sex With Men and Transgender Individuals in Europe Using Structural Text Modeling Technique: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e72491
URL: https://publichealth.jmir.org/2025/1/e72491
doi:10.2196/72491
PMID:40938647

©Haoyi Wang, Johann Kolstee, Julio Croce Martinez, David van de Vijver, Jonathan Tosh, Melanie Schroeder, Ama Appiah, Hanne ML Zimmermann, Kai J Jonas. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Secondary Distribution of Dual HIV/Syphilis Self-Testing Among Men Who Have Sex With Men: Pragmatic Randomized Controlled Trial in China

Min Wang[1,2*], MD; Yun Zhang[1,2*], MD; Weiyi Tian[1], MD; Jinli Mo[2], MD; Haimei Huang[1,3], MD; Jiawen Zhu[1], MD; Sumin Tan[2], MD; Yingqiong Huang[1], MD; Li Jiang[1], PhD; Ping Cen[4], PhD; Guanghua Lan[5], PhD; Hao Wang[6], PhD; Wei Pan[7,8], PhD; Joseph D Tucker[9,10], PhD; Chuanyi Ning[1,11], PhD

[1]Nursing College, Guangxi Medical University, Nanning, China

[2]The Second Affiliated Hospital of Guangxi Medical University, Nanning, China

[3]Department of Infection Disease, The Fourth People's Hospital of Nanning, Nanning, China

[4]Department of AIDS/STD Control and Prevention, Nanning Municipal Center for Disease Prevention and Control, Nanning, China

[5]Institute of HIV/AIDS Prevention and Control, Guangxi Zhuang Autonomous Region Center for Disease Control and Prevention, Nanning, China

[6]Institute of Biomedicine, Department of Infectious Diseases, University of Gothenburg, Gothenburg, Sweden

[7]Duke University School of Nursing, Durham, NC, United States

[8]Department of Population Health Sciences, Duke University School of Medicine, Durham, NC, United States

[9]Department of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[10]Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom

[11]Guangxi Key Laboratory of AIDS Prevention and Treatment, Nanning, China

[*]these authors contributed equally

Corresponding Author:
Chuanyi Ning, PhD
Nursing College
Guangxi Medical University
8 Shuangyong Road
Nanning, 530021
China
Phone: 86 19163959719
Email: ningchuanyi@126.com

Abstract

Background: The World Health Organization recommends dual HIV/syphilis testing, but this approach has not been examined in many low- and middle-income countries. Dual HIV/syphilis self-testing may accelerate secondary distribution of self-test kits. Preliminary studies in Guangdong, China, have demonstrated the feasibility and cost-effectiveness of dual HIV/syphilis self-testing distribution via social media, but evidence comparing dual HIV/syphilis self-testing to single HIV self-testing for secondary distribution within social networks remains limited in resource-limited settings.

Objective: We aimed to evaluate the effectiveness of secondary distribution of dual HIV/syphilis self-testing kits in promoting HIV testing uptake among men who have sex with men (MSM) in China.

Methods: We conducted a pragmatic 3-arm randomized controlled study in the Guangxi Zhuang Autonomous Region, China. MSM aged 18 years or older who were HIV-negative were enrolled and randomly assigned (1:1:1) to either the site-based HIV testing (SBT) group (110/330, 33.3%), single HIV self-testing (SST) group (110/330, 33.3%), or dual HIV/syphilis self-testing (DST) group (110/330, 33.3%). Participants in the SST and DST groups received free finger-prick-based HIV self-testing or HIV/syphilis self-testing kits at enrollment and during the 12-month follow-up. The primary outcome was the mean number of social network members motivated by the participant and the mean frequency of HIV tests per participant within a 3-month period. The data were analyzed using an intention-to-treat analysis.

Results: A total of 330 MSM were recruited, among whom 319 (319/330, 96.7%) completed at least 1 follow-up survey and were subsequently included in the analysis. Among the participants, 245/319 (77%) had a college education or above. Compared to social network members in the SBT group, those in the intervention SST and DST groups were more likely to motivate others

XSL•FO
RenderX

for HIV testing over a 3-month average duration. The mean number of motivated individuals was 0.42 in the SST group versus 0.20 in the SBT group, a mean difference (MD) of 0.22 (95% CI 0.12-0.33; *P*<.001). The mean was 0.51 in the DST group versus 0.20 in the SBT group, with an MD of 0.32 (95% CI 0.20-0.43; *P*<.001). The mean frequency of total HIV tests per participant in the SST group (1.33) was higher than that in the SBT group (0.87), with an MD of 0.46 (95% CI 0.31-0.62; *P*<.001) over 3 months. Over a 3-month period, the mean number of HIV tests per participant was higher in the DST group (1.43) than in the SBT group (0.87), with an MD of 0.57 (95% CI 0.41-0.73; *P*<.001). A total of 4 (1.3%) individuals had a new HIV positive result, while 11 (3.4%) had a new syphilis positive result. All individuals who had positive self-test results underwent laboratory-based confirmation tests. There were no adverse events reported.

Conclusions: Our data demonstrate that the secondary distribution strategy of HIV/syphilis self-testing proves to be an effective means of expanding HIV testing coverage by encouraging the distribution of testing kits within the social networks of MSM.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2100050898; https://www.chictr.org.cn/hvshowproject.html?id=158876

(JMIR Public Health Surveill 2025;11:e70775) doi:10.2196/70775

KEYWORDS

HIV; syphilis; men who have sex with men; self-testing

Introduction

The syndemic relationship between HIV and syphilis among men who have sex with men (MSM) is well-documented, with each infection potentiating the transmission and progression of the other [1,2]. Syphilis increases the risk of HIV acquisition by 2- to 5-fold through mucosal inflammation and immune activation, while HIV accelerates syphilis progression and complicates clinical management [3,4]. The World Health Organization (WHO) recommends dual HIV/syphilis rapid diagnostic tests to screen high-risk populations, thereby reducing undiagnosed infection, untreated morbidity, and mortality [5]. This integrated approach is particularly vital in low- and middle-income countries (LMICs) like China, where key populations face disproportionate burdens. Recently, a systematic review and meta-analysis showed that the prevalence of HIV/syphilis coinfection among MSM in China was 2.7% [6].

Despite WHO guidance, dual testing uptake remains alarmingly low. A cross-sectional study revealed that only 37% (95% CI 34%-40%) of Chinese MSM who had ever tested for HIV received dual screening [7]. Several barriers impede MSM from facility-based HIV testing at hospitals or Centers for Disease Control and Prevention (CDC) clinics, including HIV-related stigma, inconvenient testing times at clinics, and prolonged testing waiting times, all of which decrease HIV test uptake [8,9]. According to the Joint United Nations Programme on HIV/AIDS (UNAIDS) data, 13% of people living with HIV in the world were still unaware of their infection status in 2024 [10]. The population of undiagnosed people living with HIV in China is still large, at an estimated 15.7% in 2023 [11].

HIV self-testing (HIVST) has emerged as a transformative strategy to overcome these barriers. By enhancing privacy and convenience and decreasing stigma, HIVST enables people to test autonomously and expands coverage among hard-to-reach populations [12,13]. Building on this innovation, dual HIV/syphilis self-testing represents a significant advancement. Evidence confirms that dual screening for HIV and syphilis is acceptable to MSM and more cost-effective than separate rapid tests for each infection [14,15]. A European clinic-utility study evaluated dual point-of-care tests for HIV/syphilis across community-based organizations in Spain, Slovenia, Latvia, and Ukraine and reported high acceptability and usability among both MSM users and providers [16]. Considering the known synergy between HIV and syphilis, dual HIV/syphilis self-testing could be useful for clinical and public health reasons.

Research that explores the effect of dual HIV/syphilis self-testing secondary distribution among MSM in China is limited. Social network–based HIV testing approaches involve individuals motivating their sexual or social contacts to take HIV tests [17]. HIVST secondary distribution is a method rooted in social networks, where individuals receive multiple HIVST kits and distribute them to people in their social networks, including sexual partners and friends [18]. The distribution of HIVST kits within social networks can expand HIV testing coverage among MSM. A secondary data analysis reported the distribution of HIVST kits to 647 individuals, 54.6% of whom were first-time testers [19]. In a cohort study, HIVST kits were distributed to 319 Nigerian MSM, 17.9% of them being first-time testers [20]. Given the synergy between HIV and syphilis, dual self-testing could optimize the public health impact. Recent studies in Guangdong, China, highlight the promise of dual HIV/syphilis self-testing for MSM. Wu et al [21] showed that social media-based dual HIV/syphilis self-testing distribution increased testing uptake and case identification, while Wang et al [22] confirmed its cost-effectiveness in a multiarm randomized controlled trial (RCT). However, none have compared dual HIV/syphilis self-testing against single HIVST within a social network or assessed scalability in resource-limited settings. This gap is significant, given the synergy between HIV and syphilis and the potential for secondary distribution to reach hidden high-risk subgroups. To address this gap, we conducted the first pragmatic 3-arm RCT to examine whether the secondary distribution of dual HIV/syphilis self-testing kits promotes HIV testing uptake among Chinese MSM, compared to single HIVST and site-based HIV testing.

XSL•FO
RenderX

Methods

Study Design and Participants

The full study protocol has been previously published [23]. This study was designed as a parallel, open-label, 3-arm RCT and adhered to the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Multimedia Appendix 1). The trial was conducted in collaboration with 4 community-based organizations (CBOs) serving MSM populations across 4 cities in Guangxi, China (Nanning, Guilin, Liuzhou, and Beihai). Each CBO operated within government-supported HIV testing clinics at local CDC facilities and provided MSM-friendly health services, including HIV/STI testing, counselling, and linkage to care. Additionally, the CBOs utilize online platforms (WeChat public accounts and Blued, a gay-friendly social network app) for health promotion. Participants were recruited through convenience sampling between September 1, 2021, and February 28, 2022. The follow-up was completed by February 29, 2023. Our multichannel recruitment strategy included online recruitment through study advertisements (with QR codes) on CBOs' WeChat public accounts and Blued, as well as on-site recruitment at CBO offices, CDC clinics, and MSM-frequented venues (eg, bars and parks) where trained staff approached potential participants.

MSM were eligible if they (1) were biologically male, (2) aged ≥18 years, (3) reported engaging in anal sex with male sex partners, (4) tested negative for HIV and syphilis by rapid screening testing, (5) resided in Guangxi, and (6) were willing to provide personal information and informed consent. Participants were excluded if they met either of the following criteria: (1) presented with evident mental illness or intellectual disability, or (2) encountered language and communication barriers hindering their ability to effectively participate and complete the study.

Randomization and Masking

Participants deemed eligible were assigned, in a 1:1:1 allocation ratio, to one of the 3 groups: the site-based HIV testing (SBT or control) group, the single HIV self-testing (SST) group, and the dual HIV/syphilis self-testing (DST) group. This allocation was achieved through a simple randomization process using the computer-generated tables in SPSS software (version 29.0; IBM Corp). An external research assistant, uninvolved in the trial, securely sealed the study allocations within sequentially numbered opaque envelopes. Recruitment staff and participants were not informed about study arm allocation until the opening of the sealed envelope following informed consent and enrollment. The statisticians responsible for the data analysis were blinded to participant assignment throughout the study.

Procedures

Potential participants completed a 3-step screening process. First, prescreening was conducted online or face-to-face to verify inclusion criteria (biological male, age ≥18 years, reported anal sex with male sex partners, and Guangxi residency). Second, eligible individuals underwent onsite rapid HIV/syphilis testing at CBOs, with only double-negative participants proceeding. Third, these individuals provided written informed consent and completed the baseline survey during the baseline assessment phase.

Upon enrollment, participants assigned to the SBT group underwent HIV testing at local hospitals and free HIV testing and counselling clinics at local CDCs and CBOs. Additionally, they had the option to purchase single HIVST kits or dual HIV/syphilis self-testing kits from nearby pharmacies or through online platforms. For the SBT group, personalized online virtual cards featuring a unique identification number and QR code were provided every 3 months via WeChat. Participants could upload their HIV test results by scanning the QR code embedded in the virtual card. Participants were encouraged to distribute the online cards to their social network members (friends and sexual partners) to undergo HIV testing at local hospitals, CDCs, and CBOs.

Participants in the SST group or DST group paid a refundable deposit (US $4) to receive 2 single HIVST kits or dual HIV/syphilis self-testing kits along with a self-testing instructional video via WeChat. When participants uploaded their self-testing results through WeChat, contacted a research assistant, or scanned the QR code in the kits, the deposit was refunded. Participants could order additional self-testing kits after returning the results from their self-test kits during the 12-month study follow-up period. The self-testing kits were capped at 2 kits each. Participants were also encouraged to distribute self-testing kits to their social network members, including friends and sexual partners. To protect participants' privacy and accommodate their preferences, the test kits were dispatched via a confidential package through postal services, without any reference to HIV.

We used the finger-prick Determine HIV 1/2 early detect test kit (Alere Medical Co) and the SD Bioline HIV/syphilis Duo rapid test kit (Abbott Rapid Diagnostics) in this study. These 2 kits have high specificity and sensitivity. For the Determine HIV 1/2 test, the specificity and sensitivity are 99.75% and 100%, respectively [24]. For the Bioline HIV/syphilis duo, the specificity and sensitivity for HIV-1/2 are 99.67% and 99.91%, and for syphilis, 99.72% and 99.67%, respectively [25]. The syphilis component of the Bioline HIV/syphilis Duo test detects antibodies against *Treponema pallidum* (IgG and IgM), which indicates exposure to syphilis but cannot differentiate between current and past infections. The interpretation of the test results requires a minimum waiting period of 15 minutes, which can extend to 60 minutes. All participants received structured pretest counselling delivered face-to-face or online by trained CBO staff. The pretest counselling covered information on HIV and syphilis transmission and prevention, step-by-step instructions on how to perform the self-test, interpretation of possible results (including the possibility of false positives/negatives), risk reduction strategies, and emergency contact information. For participants and their social network members who got a positive HIV or syphilis self-test result, immediate posttest telephone counselling was provided by trained CBO staff within 2 hours of the result upload. The posttest counselling included emotional support, interpretation of the reactive results, instructions for laboratory confirmation at designated CDC clinics/hospitals within 48 hours, and prebooked appointments. All confirmed

XSL•FO
RenderX

positive cases were linked to treatment through the national HIV/syphilis control program.

Data were collected through face-to-face or online questionnaire interviews that were approximately 15 minutes in length. The baseline questionnaire covered sociodemographic information, past HIV/syphilis testing, HIVST, and the frequency of HIV testing, including both HIVST and site-based HIV testing, among MSM in the past 12 months. Follow-up questionnaires at 3, 6, 9, and 12 months covered the frequency and results of HIV testing, distribution of self-testing kits, and sexual behaviors.

Participants received a maximum subsidy of US $27 for completing study surveys: baseline surveys, follow-up surveys at 3 and 6 months (US $15), and follow-up surveys at 9 and 12 months (US $12).

Outcomes

The primary outcome was the average number of social network members that each participant motivated to test for HIV and the mean frequency of total HIV testing within a 3-month period (HIV testing included both facility-based HIV testing and self-testing). Frequencies were determined by uploading photos of HIV test results and self-reporting test frequency, excluding instances of self-testing with invalid results. The former was determined by uploading photos of HIV test results and self-reporting. The secondary outcomes were HIV and syphilis positive results.

Ethical Considerations

This study was reviewed and approved by the Medical Ethics Committee of Guangxi Medical University, China (20210173). Written informed consent was obtained from all participants before interviews and study procedures. Privacy and confidentiality were protected through deidentification of study data, secure double-data entry using EpiData software (version 3.1; EpiData Association), and dispatch of test kits via confidential postal packages without HIV-related identifiers. Participants received compensation totaling up to US $27 for survey completion (US $15 for baseline, 3-month, and 6-month surveys, and US $12 for 9-month and 12-month surveys). No images of the participants are included in the manuscript or supplementary materials.

Statistical Analysis

The required study sample size was determined using the PASS software (version 15.0; NCSS) based on the HIV testing frequency among MSM. It was assumed that the intervention groups would exhibit a higher frequency of HIV testing compared to the control group. Based on the literature [26], the control group had a mean of 1.40 (SD 1.13) and the intervention group had a mean of 2.18 (SD 1.77), which were adopted as the anticipated effect sizes for this study. With an alpha level of 0.05, a statistical power of 0.90, and assuming a conservative 10% loss to follow-up, it was determined that a minimum of 99 participants per arm, totaling 330 MSM, would provide adequate statistical power.

Data were double-entered into Epidata software. The primary and secondary outcomes were analyzed using intention-to-treat analysis, and participants were included in the complete case analysis if they initiated at least 1 follow-up survey. Baseline characteristics were reported using descriptive statistics. Categorical variables were presented as counts and relative proportions. Continuous variables were expressed as the median and IQR. The mean number of social network members motivated by participant and HIV test frequency across generalized linear mixed models was compared, and effect sizes were estimated using mean difference (MD) and risk difference (RD) for primary and secondary outcomes. A 95% CI for MD and RD was calculated. All statistical analyses were conducted using SPSS software (version 26.0; IBM Corp). For multiple comparisons involving the 3 study arms, Bonferroni correction was applied by dividing the alpha level (0.05) by the number of outcome comparisons (n=3), resulting in a significance threshold of $P<.017$. The study was registered with the Chinese Clinical Trial Registry (ChiCTR2100050898).

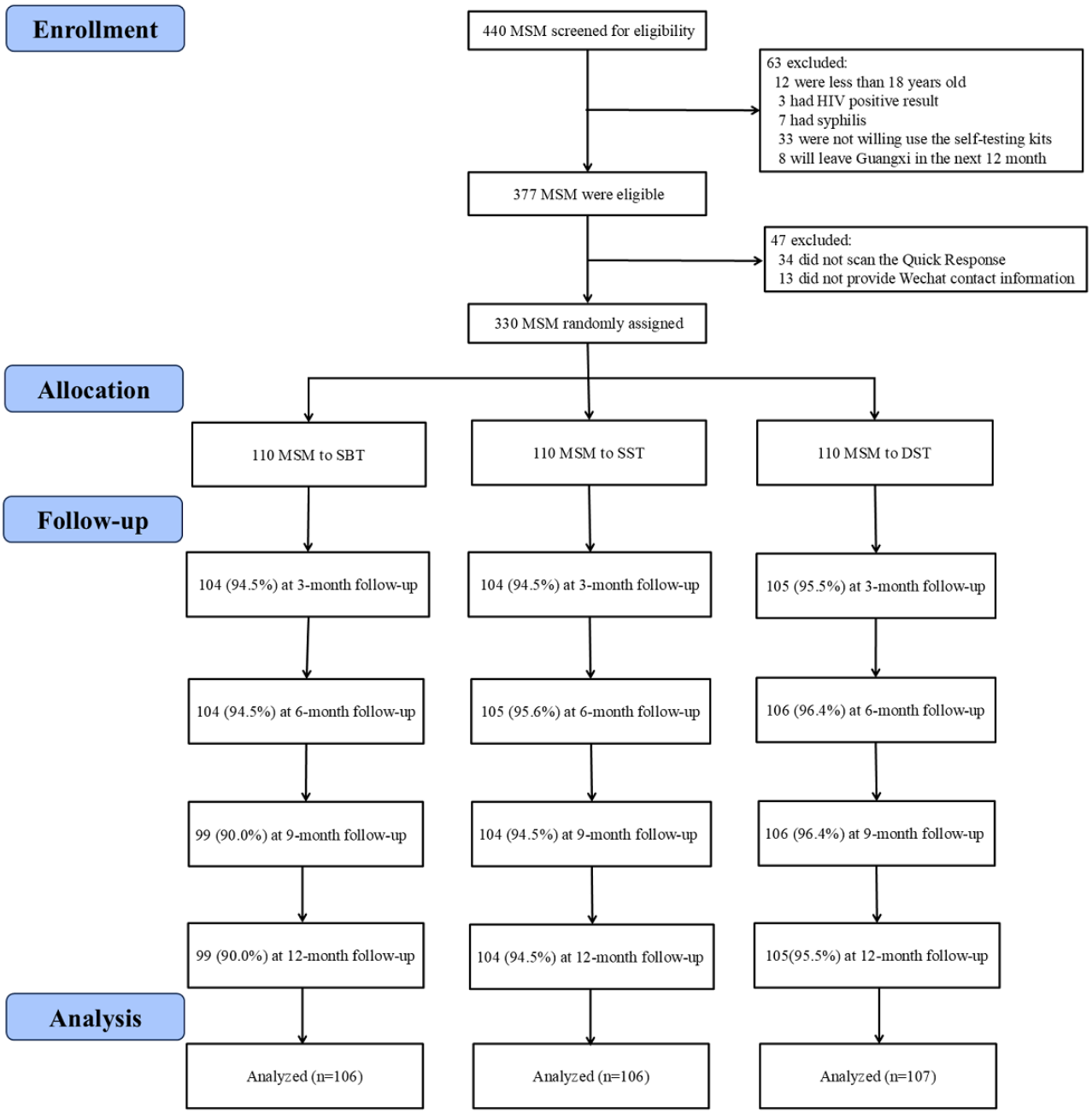
Results

Participant Characteristics

Between September 1, 2021, and February 28, 2022, we assessed the eligibility of 440 MSM. Of these, 63 participants were excluded for not meeting the inclusion criteria, 34 did not scan the Quick Response Code, and 13 were excluded due to incomplete WeChat contact information. Ultimately, 330 eligible MSM were recruited and randomized into the SBT group (110/330, 33.3%), the SST group (110/330, 33.3%), and the DST group (110/330, 33.3%). After the 12-month study period, a total of 106 participants in the SBT group, 106 in the SST group, and 107 in the DST group completed at least 1 follow-up questionnaire and were included in the data analysis. The overall follow-up rate was 96.7% (319/330) (Figure 1).

XSL•FO
RenderX

Figure 1. DST: dual HIV/syphilis self-testing group; MSM: men who have sex with men; SBT: site-based HIV testing group; SST: single HIV self-testing group.



Baseline Sociodemographic and Behavioral Characteristics of MSM

A comprehensive overview of the baseline sociodemographic and behavioral characteristics is described in Table 1. All participants had similar baseline characteristics across the 3 arms. The majority of the cohort, comprising 73.4% (234/319) of participants, were aged between 18 and 30 years, and a substantial proportion (245/319, 76.8%) had a college education or above. Notably, 63.9% (204/319) had previous experience with HIVST. Furthermore, 43.9% (140/319) reported engaging in condomless anal and oral sex with male partners in the past 6 months.

XSL•FO
RenderX

Table 1. Baseline sociodemographic and behavioral characteristics of MSM[a] enrolled in a pragmatic RCT[b] of secondary HIV/syphilis self-testing distribution in Guangxi, China (between September 2021 and February 2022).

Characteristics	Overall (N=319), n (%)	SBT[c] group (n=106), n (%)	SST[d] group (n=106), n (%)	DST[e] group (n=107), n (%)	*P* value
Age, year					.254
18-30	234 (73.4)	71 (67.0)	80 (75.5)	83 (77.6)	
31-40	61 (19.1)	26 (24.5)	16 (15.1）	19 (17.8)	
>40	24 (7.5)	9 (8.5)	10 (9.4）	5 (4. 7)	
Ethnicity					.912
Han majority	207 (64.9)	69 (65.1)	67 (63.2)	71 (66.4)	
Zhuang	96 (30.1)	32 (30.2)	32 (30.2)	32 (29.9)	
Other	16 (5.0)	5 (4.7)	7 (6.6)	4 (3.7)	
Insurance					.797
No	38 (11.9)	13 (12.3)	14 (13.2)	11 (10.3)	
Yes	281 (88.1)	93 (87.7)	92 (86.8)	96 (89.7)	
Highest level of education					.064
High school or below	74 (23.2)	31 (29.2)	26 (24.5)	17 (15. 9)	
College or beyond	245 (76.8)	75 (70.8)	80 (75.5)	90 (84.1)	
Occupation					.623
Student	106 (33.2)	34 (32.1)	39 (36.8)	33 (30.8)	
Nonstudent	213 (66.8)	72 (67.9)	67 (63.2)	74 (69.2)	
Cohabitation					.779
Living with others	143 (44.8)	45 (42.5)	52 (49.1)	46 (43.0)	
Living with family	79 (24.8)	29 (27.4)	25 (23.6)	25 (23.4)	
Living alone	97 (30.4)	32 (30.2)	29 (27.4)	36 (33.6)	
Monthly personal income					.445
＜3000 RMB	135 (42.3)	50 (47.2)	44 (41.5)	41 (38.3)	
3000-5000 RMB	83 (26.0)	29 (27.4)	24 (22.6)	30 (28.0)	
＞5000 RMB	101 (31.7)	27 (25.5)	38 (35.8)	36 (33.6)	
Marital status					.934
Unmarried	287 (90.0)	94 (88.7)	95 (89.6)	98 (91.6)	
Married	23 (7.2)	8 (7.5)	8 (7.5)	7 (6.5)	
Divorced or widowed	9 (2.8)	4 (3.8)	3 (2.8)	2 (1.9)	
Ever used HIV self-testing					.247
No	115 (36.1)	45 (42. 5)	35 (33.0)	35 (32.7)	
Yes	204 (63.9)	61 (57.5)	71 (67.0)	72 (67.3)	
Number of male sex partners in the past 6 months					.389
<2	134 (42.0)	44 (41.5)	51 (48.1)	39 (36.4)	
2-5	157 (49.2)	55 (51.9)	45 (42. 5)	57 (53.3)	
>5	28 (8.8)	7 (6.6)	10 (9.4)	11 (10.3)	
Have male regular sex partners in the past 6 months					.285
No	158 (49.5)	48 (45.3)	59 (55.7)	51 (47.7)	
Yes	161 (50.5)	58 (54.7)	47 (44.3)	56 (52.3)	
Have male casual sex partners in the past 6 months					.415

XSL•FO
RenderX

Characteristics	Overall (N=319), n (%)	SBT[c] group (n=106), n (%)	SST[d] group (n=106), n (%)	DST[e] group (n=107), n (%)	*P* value
No	132 (41.4)	43 (40.6)	49 (46.2)	40 (37.4)	
Yes	187 (58.6)	63 (59.4)	57 (53.8)	67 (62.6)	
Had a one-night stand with a male sex partner in the past 6 months					.178
No	139 (43.6)	46 (43.4)	53 (50.0)	40 (37.4)	
Yes	180 (56.4)	60 (56.6)	53 (50.0)	67 (62.6)	
Had condomless anal and oral sex with male partners in the past 6 months					.625
No	179 (56.1)	63 (59.4)	56 (52.8)	60 (56.1)	
Yes	140 (43.9)	43 (40.6)	50 (47.2)	47 (43.9)	

[a]MSM: men who have sex with men.

[b]RCT: randomized controlled trial.

[c]SBT: site-based HIV testing.

[d]SST: single HIV self-testing.

[e]DST: dual HIV/syphilis self-testing.

Primary Outcomes

Over 3 months, social network members in the intervention SST and DST groups were more likely to motivate others to undergo HIV testing compared with those in the SBT group. The mean number of motivated individuals was 0.42 in the SST group versus 0.20 in the SBT group (MD 0.22, 95% CI 0.12-0.33; *P*<.001), and 0.51 in the DST group versus 0.20 in the SBT group (MD 0.32, 95% CI 0.20–0.43; *P*<.001), both below the Bonferroni-corrected significance threshold of .017. However, there was no significant difference in the mean number of social network members tested per participant between the SST and DST groups over an average follow-up of 3 months (MD 0.09, 95% CI –0.05 to 0.23; *P*=.082), which is above the Bonferroni-corrected significance threshold of .017 (Table 2).

Table 2. Secondary distribution outcomes: mean number of social network members motivated to test for HIV by study arm in a pragmatic RCT[a] of secondary HIV/syphilis self-testing distribution among MSM[b] in Guangxi, China (3-month average duration between 2021 and 2023).

Outcomes	SBT[c] group	SST[d] group	DST[e] group	SST versus SBT		DST versus SBT		DST versus SST	
	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	MD[f] (95% CI)	*P* value[g]	MD (95% CI)	*P* value	MD (95% CI)	*P* value
Total number of social network members	0.20 (0.15-0.26)	0.42 (0.34-0.52)	0.51 (0.42-0.63)	0.22 (0.12-0.33)	<.001	0.32 (0.20-0.43)	<.001	0.09 (–0.05 to 0.23)	.187
Via self-testing kits	0.09 (0.06-0.13)	0.27 (0.20-0.36)	0.35 (0.27-0.46)	0.18 (0.10-0.27)	<.001	0.27 (0.16-0.37)	<.001	0.08 (–0.04 to 0.21)	.175
Via facility-based testing	0.11 (0.07-0.15)	0.14 (0.10-0.18)	0.16 (0.11-0.20)	0.03 (–0.03 to 0.08)	.300	0.05 (–0.01 to 0.10)	.100	0.02 (–0.04 to 0.08)	.541

[a]RCT: randomized controlled trial.

[b]MSM: men who have sex with men.

[c]SBT: site-based HIV testing.

[d]SST: single HIV self-testing.

[e]DST: dual HIV/syphilis self-testing.

[f]MD: mean difference.

[g]*P* values are interpreted against a Bonferroni-corrected threshold of .017 for the 3 pairwise comparisons per outcome.

Using the Bonferroni-corrected threshold of .017, the mean frequency of total HIV tests per participant in the SST group (1.33) was higher than that in the SBT group (0.87; MD=0.46, 95% CI 0.31-0.62; *P*<.001). This disparity was mainly due to the difference in HIVST between the 2 groups (SST 0.88 vs SBT 0.33; MD=0.56, 95% CI 0.41-0.70; *P*<.001). Conversely, the mean frequency of facility-based HIV testing was comparable between SST (0.43) and SBT (0.51) (MD=–0.09,

XSL•FO
RenderX

95% CI –0.18 to 0.01; *P*=.07). Similarly, the mean frequency of total HIV tests per participant in the DST group (1.43) surpassed the mean (0.87) in the SBT group (MD=0.57, 95% CI 0.41-0.73; *P*<.001) within a 3-month average duration. This difference was predominantly attributed to the difference in HIVST between the 2 groups (DST 0.94 vs SBT 0.33; MD=0.61, 95% CI 0.46-0.75; *P*<.001). Notably, the mean frequency of facility-based HIV testing was also comparable between the DST (0.47) and SBT (0.51) group (MD=–0.04; 95% CI –0.14 to 0.06; *P*=.398). However, there was no difference between the SST and DST groups in the mean frequency of either total, self-tests, or site-based tests (all *P*>.017) (Table 3).

Table 3. Frequency of HIV testing per participant by study arm in an RCT[a] of secondary HIV/syphilis self-testing distribution among MSM[b] in Guangxi, China (3-month average duration between 2021 and 2023).

Outcomes	SBT[c] group	SST[d] group	DST[e] group	SST versus SBT		DST versus SBT		DST versus SST	
	Mean (95% CI)	Mean (95% CI)	Mean (95% CI)	MD[f] (95% CI)	*P* value[g]	MD (95% CI)	*P* value	MD (95% CI)	*P* value
Total Frequency of HIV tests	0.87 (0.77-0.97)	1.33 (1.21-1.45)	1.43 (1.31-1.56)	0.46 (0.31-0.62)	<.001	0.57 (0.41-0.73)	<.001	0.10 (–0.07 to 0.28)	.245
Via self-testing kits	0.33 (0.27-0.40)	0.88 (0.77-1.01)	0.94 (0.82-1.07)	0.56 (0.41-0.70)	<.001	0.61 (0.46-0.75)	<.001	0.05 (–0.13 to 0.23)	.555
Via facility-based testing	0.51 (0.44-0.58)	0.43 (0.36-0.49)	0.47 (0.41-0.54)	–0.09 (–0.18 to 0.01)	.070	–0.04 (–0.14 to 0.06)	.398	0.05 (–0.05 to 0.14)	.328

[a]RCT: randomized controlled trial.

[b]MSM: men who have sex with men.

[c]SBT: site-based HIV testing.

[d]SST: single HIV self-testing.

[e]DST: dual HIV/syphilis self-testing.

[f]MD: mean difference.

[g]*P* values are interpreted against a Bonferroni-corrected threshold of .017 for the 3 pairwise comparisons per outcome.

Secondary Outcomes

Over the 12-month follow-up, 26 individuals were diagnosed with HIV, including 4 (15.4%) participants and 22 (84.6%) social network members. Of the 22 social network members diagnosed with HIV, 1 (4.5%) was from the SBT group, 12 (54.5%) were from the SST group, and 9 (40.9%) were from the DST group. Regarding syphilis diagnoses, 11 participants were positive, with 1 (9.1%) from the SBT group, 3 (27.3%) from the SST group, and 7 (63.6%) from the DST group. All individuals identified as living with HIV or syphilis positive underwent laboratory-based confirmation tests and were linked to treatment through the national HIV or syphilis control program. There were no statistically significant differences between participants in the 3 study groups regarding HIV and syphilis positive test outcomes and social network members with an HIV positive test (Table 4).

XSL•FO
RenderX

Table 4. Identified HIV and syphilis infection among participants and their social network members during the 12-month follow-up period in a pragmatic RCT[a] of secondary HIV/syphilis self-testing distribution among MSM[b] in Guangxi, China (2021-2023).

Outcomes	SBT[c]	SST[d]	DST[e]	SST versus SBT		DST versus SBT		DST versus SST	
	n/N (%; 95% CI)	n/N (%; 95% CI)	n/N (%; 95% CI)	RD[f] (95% CI)	*P* value[g]	RD (95% CI)	*P* value	RD (95% CI)	*P* value
HIV positive	1/106 (0.94; –0.93 to 2.81)	1/106 1. –0.93 to 2.81)	2/107 (1.87; –0.74 to 4.48)	—[h]	—	0.93 (–3.16 to 5.01)	.620	0.93 (–3.16 to 5.01)	.620
Syphilis positive	1/106 (0.94; –0.93 to 2.81)	3/106 (2.83; –0.38 to 6.04)	7/107 (6.54; 1.78-11.30)	1.89 (–2.71 to 6.48)	0.624	5.60 (–0.37 to 11.57)	.038	3.71 (–2.88 to 10.30)	.202
HIV positive results of social network members	1/126 (0.79; –0.77 to 2.36)	12/245 (4.90; 2.18-7.62)	9/309 (2.90; 1.03-4.80)	4.10 (0.39-7.82)	0.019	2.12 (–0.87 to 5.11)	.460	–1.99 (–5.64 to 1.67)	.341

[a]RCT: randomized controlled trial.

[b]MSM: men who have sex with men.

[c]SBT: site-based HIV testing.

[d]SST: single HIV self-testing.

[e]DST: dual HIV/syphilis self-testing.

[f]RD: risk (probability) difference (expressed as a percentage).

[g]*P* values are interpreted against a Bonferroni-corrected threshold of .017 for the 3 pairwise comparisons per outcome.

[h]Data not available.

Discussion

Principal Findings

Our pragmatic RCT evaluated whether secondary distribution strategy of dual HIV/syphilis self-testing and HIVST could increase HIV testing coverage by facilitating kit distribution within social networks, thereby addressing a critical evidence gap in LMICs. This approach contributes significantly to achieving the initial "95%" target outlined by UNAIDS [27]. While the absolute increase in motivated social network members per participant was modest at the individual level, prior cost-effectiveness analyses demonstrate that such incremental gains translate to meaningful population-level impacts. Wang et al [22] showed that dual HIV/syphilis self-testing among Chinese MSM is cost-efficient, with a cost per person tested substantially lower for dual HIV/syphilis self-testing (US $26.55) compared to standard care (US $66.19), and an incremental cost of US $17.55 per additional person tested. This confirms the efficiency of scaling such interventions in LMICs.

Notably, previous secondary distribution programs of HIVST in China have demonstrated the effect of this strategy in expanding testing coverage [28]. Similarly, a study conducted in California suggests that a network-based strategy for HIVST distribution is a promising intervention to increase testing uptake and reduce undiagnosed infections among MSM [29]. Moreover, monetary incentives, either independently or in conjunction with peer referrals, have proven effective in promoting the secondary distribution of HIVST among MSM [28]. A previous study in China also suggests that integrating social media with the secondary distribution of self-test kits holds promise to increase HIV/syphilis testing coverage and case identification among MSM [21]. These findings carry significant implications for the expansion of HIVST among key populations and may be applicable to other areas of public health research.

Our results also indicate that dual HIV/syphilis self-testing and HIVST can increase the frequency of total HIV testing in MSM, consistent with findings from other studies [26,30,31]. In contrast to the SBT group, both the SST and DST groups increased the frequency of total HIV testing among MSM in China. As the differences in the average frequency of site-based HIV testing among the 3 groups were similar, the higher average frequency of total HIV testing in both the SST and DST groups can be mainly attributed to their increased average frequency of HIVST compared to the SBT group. This shift did not reduce facility-based testing, indicating self-testing complements rather than replaces clinical services.

Our findings indicate that dual HIV/syphilis self-testing can yield results similar to single HIVST in the average frequency. The integration of HIV and syphilis self-testing has the potential to strengthen public health interventions aimed at addressing HIV and syphilis control among MSM [14]. Despite syphilis being a significant public health challenge, it is often overlooked and underfunded, particularly in LMICs [32,33]. Previous research indicates that only 30% of MSM in China have ever undergone a syphilis test [9]. The anticipated stigma associated with syphilis may impede testing efforts [34]. Therefore, syphilis self-testing emerges as a potentially effective approach to overcoming these barriers [9,35]. Crucially, DST achieved comparable HIV testing outcomes to SST while simultaneously enabling syphilis screening. Given the low HIV and syphilis

XSL•FO
RenderX

testing uptake in LMICs, it is crucial to use dual HIV/syphilis self-testing to increase testing coverage.

A systematic review has demonstrated the acceptability and cost-effectiveness of integrating syphilis self-testing into HIVST services [14]. It also significantly increases syphilis testing among MSM in China, demonstrating cost-effectiveness [22]. Building upon this evidence base, our trial uniquely focuses on the secondary distribution of dual HIV/syphilis self-testing among MSM. To our knowledge, this is the first RCT to directly compare DST with SST within a network-based distribution model. Our findings demonstrate that DST increased HIV testing frequency comparably to SST and led to the identification of more undiagnosed syphilis cases (7 in DST vs 3 in SST), highlighting its potential to address the syndemic transmission of sexually transmitted infections (STIs) in LMICs where syphilis testing is neglected. This highlights the significant added value of DST beyond HIVST alone. However, the DST group's numerical increase in HIV or syphilis positivity among participants and their network members did not reach statistical significance compared to other groups. This lack of statistical significance is likely attributable to insufficient statistical power due to the relatively small number of incident cases and the observation period of this trial. Therefore, further research with larger sample sizes and longer follow-up is needed to robustly assess the impact of DST distribution on HIV and syphilis incidence.

Despite demonstrated efficacy in secondary DST distribution, scaling this intervention across LMICs necessitates addressing practical barriers—particularly cost constraints for sustained kit provision and infrastructure limitations in remote regions. Mitigation strategies include integrating with existing public health systems such as CBOs and CDC clinics to leverage their established networks for efficient distribution, confirmatory testing, and care linkage, thereby reducing operational costs. Further cost efficiency may be achieved through the bulk procurement of dual self-test kits instead of separate HIV/syphilis tests and by establishing public-private partnerships for subsidized pricing. Implementation should prioritize phased geographical scale-up, beginning with urban hubs possessing CBO partnerships before extending to periurban or rural areas.

Limitations

Our study has several limitations. First, this RCT recruited participants mainly through CBO referrals and social media, focusing on individuals who volunteered for HIV testing at CBOs and expressed a willingness to use HIVST. This recruitment approach may have excluded MSM who were averse to attending site-based testing at CBOs or hesitant to adopt HIVST, particularly those with lower educational attainment or limited digital access. Consequently, our findings are primarily generalizable to MSM engaged with CBOs or actively seeking testing services and may not fully represent the broader MSM population in China, including hidden high-risk subgroups with socioeconomic vulnerabilities. Future studies should employ diversified recruitment strategies (eg, venue-based sampling or peer-driven recruitment in rural areas) to include MSM with diverse educational backgrounds. Additionally, hybrid interventions combining online platforms and offline community support may enhance accessibility for populations with limited literacy or digital resources. Second, the implementation of this RCT coincided with the COVID-19 lockdown period in China, particularly during the 6-month follow-up, disrupting HIV services. The impact of this disruption may have enhanced the uptake of self-testing. Third, social desirability and self-report biases may influence the accuracy of self-reported testing frequencies and sexual behaviors. Although we implemented mitigation strategies, including anonymized data collection and photo verification of self-test results, residual bias could persist. The study findings could be susceptible to social desirability reporting bias, as participants were required to self-report sensitive personal information. Fourth, the absence of a limit on the number of testing kits participants could obtain might lead to unnecessary testing during the experiment. Our original intention, allowing individuals to apply for multiple kits, aimed to encourage participants to distribute more kits within their social network. Fifth, although we observed sustained effects over 12 months, the long-term durability beyond 1 year remains uncertain. Future studies with longer follow-up are needed to assess whether the effects persist without the ongoing provision of free self-testing kits. Sixth, the lack of detailed syphilis outcomes—such as treatment initiation, serological outcomes after diagnosis, or longitudinal follow-up—restricts our ability to fully evaluate the comparative effectiveness of DST versus SST in improving syphilis control. Future studies should prioritize comprehensive syphilis outcome metrics to validate the public health utility of dual testing.

Conclusions

The secondary distribution strategy of HIV and dual HIV/syphilis self-testing effectively expanded HIV testing coverage among MSM in China by facilitating kit distribution within social networks. Both strategies increase HIV testing frequency. These findings provide policymakers with evidence to support the incorporation of self-testing into national HIV and STI control programs, particularly in resource-limited settings, as a cost-effective approach to improve testing coverage. Future studies should explore cost-effective implementation methods in high-risk MSM.

Acknowledgments

We sincerely thank all the study participants. This study was funded by the Outstanding Youth Fund of the Natural Science Foundation of Guangxi (2023GXNSFFA026007), Guangxi Key Laboratory of Major Infectious Disease Prevention and Control and Biosafety Emergency Response (GKLKP-KF-202202), Innovation Project of Guangxi Graduate Education (YCSW2023249), and the General Program of the Natural Science Foundation of Guangxi 2023GXNSFAA026287) in China. No generative artificial

XSL•FO
RenderX

intelligence (AI) tools were used in the preparation, drafting, analysis, or revision of this manuscript. All content reflects the original work and intellectual contributions of the human authors.

Data Availability

The deidentified data sets generated during the study, along with the statistical analysis plan and analytic code, can be made available from the corresponding author on reasonable request.

Authors' Contributions

MW was responsible for conceptualization, data curation, formal analysis, investigation, methodology, project administration, validation, visualization, and writing (original draft; review and editing). YZ contributed to conceptualization, data curation, investigation, methodology, resources, validation, and writing (original draft; review and editing). WT was involved in conceptualization, data curation, investigation, methodology, visualization, and writing (review and editing). JM contributed to conceptualization, data curation, methodology, validation, and writing (review and editing). HH was responsible for conceptualization, methodology, visualization, and writing (review and editing). JZ contributed to conceptualization, methodology, and software. ST was involved in conceptualization, methodology, and project administration. YH contributed to conceptualization and methodology. LJ contributed to conceptualization and methodology. PC was responsible for conceptualization, methodology, and supervision. GL contributed to conceptualization, funding acquisition, methodology, and supervision. HW was responsible for methodology, validation, and writing (review and editing). WP contributed to methodology, validation, and writing (review and editing). JDT was responsible for methodology, validation, and writing (review and editing). CN contributed to conceptualization, formal analysis, funding acquisition, methodology, validation, and writing (review and editing). All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CONSORT (Consolidated Standards of Reporting Trials) 2010 checklist.
[PDF File (Adobe PDF File), 214 KB - publichealth_v11i1e70775_app1.pdf]

References

1. Jennings JM, Tilchin C, Meza B, Schumacher C, Fields E, Latkin C, et al. Overlapping transmission networks of early syphilis and/or newly HIV diagnosed gay, bisexual and other men who have sex with men (MSM): opportunities for optimizing public health interventions. AIDS Behav 2020 Oct;24(10):2895-2905. [doi: 10.1007/s10461-020-02840-2] [Medline: 32239359]
2. Vergara-Ortega DN, Tapia-Maltos A, Herrera-Ortíz A, García-Cisneros S, Olamendi-Portugal M, Sánchez-Alemán. High prevalence of syphilis and syphilis/HIV coinfection among men who have sex with men who attend meeting places in Mexico. Pathogens 2023 Feb 21;12(3):356 [FREE Full text] [doi: 10.3390/pathogens12030356] [Medline: 36986278]
3. Zetola NM, Engelman J, Jensen TP, Klausner JD. Syphilis in the United States: an update for clinicians with an emphasis on HIV coinfection. Mayo Clin Proc 2007 Sep;82(9):1091-1102. [doi: 10.4065/82.9.1091] [Medline: 17803877]
4. Lee N, Chen Y, Liu H, Li C, Li C, Ko W, et al. Increased repeat syphilis among HIV-infected patients: A nationwide population-based cohort study in Taiwan. Medicine (Baltimore) 2020 Jul 10;99(28):e21132 [FREE Full text] [doi: 10.1097/MD.0000000000021132] [Medline: 32664143]
5. WHO. Information note on the use of dual HIV/syphilis rapid diagnostic tests (RDT). World Health Organization. 2017. URL: https://www.who.int/publications/i/item/WHO-RHR-17.01 [accessed 2024-11-18]
6. Su R, Liu Y, Shan D, Li P, Ge L, Li D. Prevalence of HIV/syphilis co-infection among men who have sex with men in China: a systematic review and meta-analysis. BMC Public Health 2025 Apr 07;25(1):1297 [FREE Full text] [doi: 10.1186/s12889-025-22499-5] [Medline: 40197258]
7. Ong JJ, Liao M, Lee A, Fu H, Pan SW, Tang W, et al. Bridging the HIV-syphilis testing gap: dual testing among men who have sex with men living in China. Sex Transm Infect 2019 Jun;95(4):251-253. [doi: 10.1136/sextrans-2018-053527] [Medline: 30126948]
8. Liu Y, Sun X, Qian H, Yin L, Yan Z, Wang L, et al. Qualitative assessment of barriers and facilitators of access to HIV testing among men who have sex with men in China. AIDS Patient Care STDS 2015 Sep;29(9):481-489 [FREE Full text] [doi: 10.1089/apc.2015.0083] [Medline: 26186029]
9. Ong JJ, Fu H, Smith MK, Tucker JD. Expanding syphilis testing: a scoping review of syphilis testing interventions among key populations. Expert Rev Anti Infect Ther 2018 May;16(5):423-432 [FREE Full text] [doi: 10.1080/14787210.2018.1463846] [Medline: 29633888]

XSL•FO
RenderX

10. The Joint United Nations Programme on HIV/AIDS. 2024. URL: https://www.unaids.org/en/resources/fact-sheet [accessed 2025-08-18]
11. China National Disease Control and Prevention Administration. Progress report on HIV/AIDS prevention and treatment in China. 2024 Presented at: The 9th National Academic Conference on HIV/AIDS; April 25-28; Qingdao, China.
12. Qin Y, Han L, Babbitt A, Walker JS, Liu F, Thirumurthy H, et al. Experiences using and organizing HIV self-testing. AIDS 2018 Jan 28;32(3):371-381 [FREE Full text] [doi: 10.1097/QAD.0000000000001705] [Medline: 29194120]
13. Conserve DF, Abu-Ba'are GR, Janson S, Mhando F, Munisi GV, Drezgic B, et al. Development and feasibility of the peer and nurse-led HIV Self-Testing Education and Promotion (STEP) intervention among social networks of men in Dar es Salaam, Tanzania: application of the ADAPT-ITT model. BMC Health Serv Res 2024 Oct 02;24(1):1166. [doi: 10.1186/s12913-024-11586-9] [Medline: 39358728]
14. Gliddon HD, Peeling RW, Kamb ML, Toskin I, Wi TE, Taylor MM. A systematic review and meta-analysis of studies evaluating the performance and operational characteristics of dual point-of-care tests for HIV and syphilis. Sex Transm Infect 2017 Dec;93(S4):S3-S15 [FREE Full text] [doi: 10.1136/sextrans-2016-053069] [Medline: 28747410]
15. Coomes D, Green D, Barnabas R, Sharma M, Barr-DiChiara M, Jamil MS, et al. Cost-effectiveness of implementing HIV and HIV/syphilis dual testing among key populations in Viet Nam: a modelling analysis. BMJ Open 2022 Aug 11;12(8):e056887 [FREE Full text] [doi: 10.1136/bmjopen-2021-056887] [Medline: 35953255]
16. Fernàndez-López L, Reyes-Urueña J, Egea L, Chernyshev A, Upmace I, Ćosić M, ProSPeRo group. A clinical utility evaluation of dual HIV/Syphilis point-of-care tests in non-clinical settings for screening for HIV and syphilis in men who have sex with men. BMC Infect Dis 2024 Feb 29;24(Suppl 1):264 [FREE Full text] [doi: 10.1186/s12879-024-09017-5] [Medline: 38419023]
17. WHO recommends social network-based HIV testing approaches for key populations as part of partner services package. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/WHO-CDS-HIV-19.32 [accessed 2024-11-18]
18. Masters SH, Agot K, Obonyo B, Napierala MS, Maman S, Thirumurthy H. Promoting partner testing and couples testing through secondary distribution of HIV self-tests: a randomized clinical trial. PLoS Med 2016 Nov;13(11):e1002166 [FREE Full text] [doi: 10.1371/journal.pmed.1002166] [Medline: 27824882]
19. Eustaquio PC, Figuracion R, Izumi K, Morin MJ, Samaco K, Flores SM, et al. Outcomes of a community-led online-based HIV self-testing demonstration among cisgender men who have sex with men and transgender women in the Philippines during the COVID-19 pandemic: a retrospective cohort study. BMC Public Health 2022 Feb 21;22(1):366 [FREE Full text] [doi: 10.1186/s12889-022-12705-z] [Medline: 35189850]
20. Tun W, Vu L, Dirisu O, Sekoni A, Shoyemi E, Njab J, et al. Uptake of HIV self-testing and linkage to treatment among men who have sex with men (MSM) in Nigeria: A pilot programme using key opinion leaders to reach MSM. J Int AIDS Soc 2018 Jul;21 Suppl 5:e25124 [FREE Full text] [doi: 10.1002/jia2.25124] [Medline: 30033680]
21. Wu D, Zhou Y, Yang N, Huang S, He X, Tucker J, et al. Social media-based secondary distribution of HIV/syphilis self-testing among Chinese men who have sex with men. Clin Infect Dis 2020 Jun 26:e2251-e2257. [doi: 10.1093/cid/ciaa825] [Medline: 32588883]
22. Wang C, Ong JJ, Zhao P, Weideman AM, Tang W, Smith MK, et al. Expanding syphilis test uptake using rapid dual self-testing for syphilis and HIV among men who have sex with men in China: A multiarm randomized controlled trial. PLoS Med 2022 Mar;19(3):e1003930 [FREE Full text] [doi: 10.1371/journal.pmed.1003930] [Medline: 35235573]
23. Zhang H, Wang M, Zhang Y, Tan S, Huang H, Zhu J, et al. Usability of dual HIV/syphilis self-testing among men who have sex with men in China: study protocol for a three-arm randomised controlled trial. BMJ Open 2022 Oct 10;12(10):e059034 [FREE Full text] [doi: 10.1136/bmjopen-2021-059034] [Medline: 36216414]
24. Alere HIV Combo quick reference guide APAC (English). Abbott. URL: https://www.globalpointofcare.abbott/zh_cn/product-details/alere-hiv-combo.html [accessed 2024-11-18]
25. SD Bioline HIV/Syphilis Duo. Abbott. URL: https://www.globalpointofcare.abbott/zh_cn/product-details/sd-bioline-hiv-syphilis-duo.html [accessed 2024-11-18]
26. Zhang C, Koniak-Griffin D, Qian H, Goldsamt LA, Wang H, Brecht M, et al. Impact of providing free HIV self-testing kits on frequency of testing among men who have sex with men and their sexual partners in China: A randomized controlled trial. PLoS Med 2020 Oct;17(10):e1003365 [FREE Full text] [doi: 10.1371/journal.pmed.1003365] [Medline: 33035206]
27. 2025 AIDS targets. United Nations Programme on HIV/AIDS. 2021. URL: https://aidstargets2025.unaids.org [accessed 2024-11-18]
28. Zhou Y, Lu Y, Ni Y, Wu D, He X, Ong JJ, et al. Monetary incentives and peer referral in promoting secondary distribution of HIV self-testing among men who have sex with men in China: A randomized controlled trial. PLoS Med 2022 Feb;19(2):e1003928 [FREE Full text] [doi: 10.1371/journal.pmed.1003928] [Medline: 35157727]
29. Lightfoot MA, Campbell CK, Moss N, Treves-Kagan S, Agnew E, Kang Dufour M, et al. Using a social network strategy to distribute HIV self-test kits to African American and Latino MSM. J Acquir Immune Defic Syndr 2018 Sep 01;79(1):38-45. [doi: 10.1097/QAI.0000000000001726] [Medline: 29771792]

XSL•FO
RenderX

30. Jamil MS, Prestage G, Fairley CK, Grulich AE, Smith KS, Chen M, et al. Effect of availability of HIV self-testing on HIV testing frequency in gay and bisexual men at high risk of infection (FORTH): a waiting-list randomised controlled trial. Lancet HIV 2017 Jun;4(6):e241-e250. [doi: 10.1016/S2352-3018(17)30023-1] [Medline: 28219619]
31. Katz DA, Golden MR, Hughes JP, Farquhar C, Stekler JD. HIV self-testing increases HIV testing frequency in high-risk men who have sex with men: a randomized controlled trial. J Acquir Immune Defic Syndr 2018 Aug 15;78(5):505-512. [doi: 10.1097/QAI.0000000000001709] [Medline: 29697595]
32. Roberts CP, Klausner JD. Global challenges in human immunodeficiency virus and syphilis coinfection among men who have sex with men. Expert Rev Anti Infect Ther 2016 Nov;14(11):1037-1046 [FREE Full text] [doi: 10.1080/14787210.2016.1236683] [Medline: 27626361]
33. Dual HIV/syphilis rapid diagnostic tests can be used as the first test in antenatal care. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/WHO-CDS-HIV-19.38 [accessed 2024-11-18]
34. Turpin R, Rosario AD, Dyer T. Barriers to syphilis testing among men who have sex with men: a systematic review of the literature. Sex Health 2020 Jun;17(3):201-213. [doi: 10.1071/SH19044] [Medline: 32586417]
35. Zhong F, Tang W, Cheng W, Lin P, Wu Q, Cai Y, et al. Acceptability and feasibility of a social entrepreneurship testing model to promote HIV self-testing and linkage to care among men who have sex with men. HIV Med 2017 May;18(5):376-382 [FREE Full text] [doi: 10.1111/hiv.12437] [Medline: 27601301]

Abbreviations

CBO: community-based organization
CDC: Centers for Disease Control and Prevention
DST: dual HIV/syphilis self-testing
HIVST: HIV self-testing
LMIC: low- and middle-income country
MD: mean difference
MSM: men who have sex with men
RCT: randomized controlled trial
RD: risk difference
SBT: site-based HIV testing
SST: single HIV self-testing
STI: sexually transmitted infection
UNAIDS: United Nations Programme on HIV/AIDS
WHO: World Health Organization

Edited by A Mavragani; submitted 07.01.25; peer-reviewed by S Chen, Y Chen; comments to author 28.04.25; revised version received 11.08.25; accepted 19.08.25; published 25.09.25.

Please cite as:
Wang M, Zhang Y, Tian W, Mo J, Huang H, Zhu J, Tan S, Huang Y, Jiang L, Cen P, Lan G, Wang H, Pan W, Tucker JD, Ning C Secondary Distribution of Dual HIV/Syphilis Self-Testing Among Men Who Have Sex With Men: Pragmatic Randomized Controlled Trial in China
JMIR Public Health Surveill 2025;11:e70775
URL: https://publichealth.jmir.org/2025/1/e70775
doi:10.2196/70775
PMID:

©Min Wang, Yun Zhang, Weiyi Tian, Jinli Mo, Haimei Huang, Jiawen Zhu, Sumin Tan, Yingqiong Huang, Li Jiang, Ping Cen, Guanghua Lan, Hao Wang, Wei Pan, Joseph D Tucker, Chuanyi Ning. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 25.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Maternal Use of the Continuum of Care on Complementary Feeding Practices in Bangladesh: Cross-Sectional Study

M A Rifat[1], MPH; Rokibul Islam[2], MSc; Rinath Bintey Didar[2], MSc; Joya Bhowmick[3], MSc; Plabon Sarkar[2], MPH; Md Ruhul Amin[2], PhD; Sanjib Saha[4], PhD

[1]BRAC Institute of Educational Development, BRAC University, Bangladesh

[2]Institute of Nutrition and Food Science, University of Dhaka, Bangladesh

[3]Department of Food and Nutrition, National College of Home Economics, Dhaka, Bangladesh

[4]Lund University, BMC, Sölvegatan 19, Lund, Sweden

Corresponding Author:
Sanjib Saha, PhD
Lund University, BMC, Sölvegatan 19, Lund, Sweden

Abstract

Background: The continuum of care (CoC) for maternal health, which includes ≥4 antenatal care (ANC) visits, delivery assisted by skilled birth attendants, and a postnatal care (PNC) visit within 48 hours of delivery, is a crucial health care package associated with survival and improved health outcomes for children and mothers. In addition, the CoC serves as a platform for delivering messages and counseling on child feeding practices. However, the effect of maternal use of the CoC on complementary feeding practices in Bangladesh remains unexamined.

Objective: This study aimed to estimate the effect of maternal use of CoC on complementary feeding practices among children aged 6 to 23 months in Bangladesh.

Methods: Data from 2 consecutive nationally representative surveys—the Bangladesh Demographic and Health Survey 2017-2018 and 2022—were analyzed. Observations corresponding to the CoC for maternal health and complementary feeding indicators, including (1) timely introduction of solid, semisolid, and soft food (ISSSF); (2) minimum meal frequency (MMF); (3) minimum dietary diversity (MDD); and (4) minimum acceptable diet (MAD), were merged to prepare the analyzed samples. The differences in complementary feeding practice indicators by maternal use of the CoC were observed using chi-square tests. Multivariable logistic regression models were used to observe the associations.

Results: The analysis included 887, 4967, 4967, and 4967 mother-child pairs for the timely ISSSF, MMF, MDD, and MAD indicators, respectively. The status of complementary feeding indicators was significantly different (P<.05) by maternal use of the full CoC, ≥4 ANC visits, status of receiving PNC within 48 hours of birth, maternal educational level, husband's educational level, maternal occupation, wealth index of families, maternal perceived problems with accessing health care, and division of residence. Mothers who received the full CoC were 29% and 32% more likely to meet the MDD (adjusted odds ratio [AOR] 1.29, 95% CI 1.10-1.51; P=.002) and MAD (AOR 1.32, 95% CI 1.13-1.55; P=.001) than those who did not receive the full CoC, respectively. Among the individual components of the CoC, mothers who received ≥4 ANC visits were 23%, 31%, and 34% more likely to meet the MMF (AOR 1.23, 95% CI 1.05-1.45; P=.01), MDD (AOR 1.31, 95% CI 1.13-1.51; P<.001), and MAD (AOR 1.34, 95% CI 1.15-1.56; P<.001) than those with fewer ANC visits, respectively. Furthermore, the effects of delivery assisted by skilled birth attendants and receiving PNC within 48 hours of delivery on complementary feeding indicators were also statistically insignificant (P≥.05).

Conclusions: Maternal use of the CoC appears to be effective in improving MDD and MAD among children aged 6 to 23 months in Bangladesh. The findings highlight the pivotal role of recommended ANC visits in promoting complementary feeding practices and further suggest opportunities to amplify the impact of CoC on timely ISSSF and MMF.

(JMIR Public Health Surveill 2025;11:e76666) doi:10.2196/76666

KEYWORDS

antenatal care; ANC; postnatal care; PNC; complementary feeding; dietary diversity; minimum acceptable diet; Bangladesh; Bangladesh Demographic and Health Survey; BDHS

XSL•FO
RenderX

Introduction

Background

The age of 6 to 23 months constitutes the largest proportion of the first 1000 days of life, which are referred to as the critical window of opportunity [1]. This age group is characterized by rapid body growth, cognitive development, language acquisition, personality development, and outstanding learning skills [1,2]. Therefore, children aged 6 to 23 months critically require proper nutrition. This is primarily met through appropriate complementary feeding, which demonstrates a significant association with childhood growth [3,4]. As infants grow, their requirement for essential micronutrients increases, which is not fulfilled by breast milk alone, requiring quality and frequency of complementary food. Evidence shows that proper complementary feeding can help maintain adequate supply of essential macro- and micronutrients (eg, protein, calories, vitamin A, iron, and zinc), which are essential for linear growth, cognitive development, and protecting young children from deficiency disorders [5-7]. Furthermore, poor complementary feeding results in growth faltering and frequent infection in the short term, whereas the long-term consequences include an elevated risk of diabetes, obesity, and poor neuromuscular development [8,9]. Children aged 6 to 23 months are highly vulnerable to malnutrition; therefore, poor complementary feeding can result in growth faltering, which might be irreversible in later life [10]. Apart from undernutrition, the period of complementary feeding is a crucial opportunity to prevent overweight and obesity during adulthood [11].

Initiation of complementary feeding at the age of 6 months triggers the transition from breastfeeding to family foods; however, the process should take several factors into careful consideration, including children's growth patterns, requirements of essential nutrients, appetite, satiety, and quality of food [12]. The frequency, quantity, and quality of complementary food differ by children's age, so the provision of age-appropriate complementary feeding requires caregivers skilled or educated in complementary feeding practices [13]. One of the most considered strategies to improve caregivers' complementary feeding practices is to provide them with knowledge and motivation, especially through interventions during the antenatal and postnatal periods, which are often integrated with antenatal care (ANC) and postnatal care (PNC) [14].

The continuum of care (CoC) for maternal health offers a package of health care services from pregnancy to the postnatal period that are essential for the survival and well-being of the mother and newborn. These services include at least 4 ANC visits by skilled care providers during pregnancy, delivery assisted by skilled birth attendants (SBAs), and a PNC visit within 48 hours of delivery. These services provide pregnant mothers with opportunities to be in contact with trained health care providers during critical periods. Therefore, ANC and PNC are platforms to deliver nutrition education and sensitize mothers on childcare and complementary feeding. Furthermore, delivery by SBAs indicates mothers' positive health care–seeking behavior, which may be carried forward to the practice of recommended complementary feeding. Evidence shows that interventions to improve complementary feeding practice target pregnant mothers, considering ANC as a valuable platform to provide counseling and education [15-17]. In low- and middle-income countries, nutrition education and complementary feeding interventions appeared to exhibit significant positive results on linear growth among children [4,18].

Objectives

Despite invaluable benefits, complementary feeding practices in Bangladesh demonstrate a frustrating picture, with only 20% of children aged 6 to 23 months receiving a minimum acceptable diet (MAD) in 2024 [19,20]. According to the latest Bangladesh Demographic and Health Survey (BDHS) in 2022, only 30% of children were fed with an MAD, and this unfortunately experienced a drop from 35% in 2017 [21,22]. Addressing malnutrition and micronutrient deficiencies through the promotion of complementary feeding practices is one of the key priorities in the second national plan of action of nutrition in Bangladesh [23]. One of the strategies to achieve nutrition goals is mainstreaming nutrition, where many nutrition services are integrated into health services such as ANC and PNC, with the hypothesis that mothers who undergo maternal health care services will exhibit better performance in child feeding practices [23,24]. However, no previous study has investigated the effect of maternal use of the CoC on complementary feeding practices in Bangladesh. Using 2 nationally representative surveys, this study aimed to fill this research gap through estimating the overall effect of maternal use of the CoC and its components on complementary feeding practices over the survey periods. The findings might help policymakers understand how the benefits of maternal use of CoC translate to complementary feeding practices, identifying priorities and ways forward.

Methods

Study Design and Samples

In this study, the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist was used as the reporting guideline [25]. The STROBE checklist is provided in Checklist 1. We used 2 nationally representative cross-sectional surveys as data sources. These sources were the BDHS of 2017-2018 and 2022. The surveys applied 2-stage stratified cluster random sampling and obtained a response rate of more than 98%. In both surveys, the samples represented mothers of children aged 6 to 23 months in Bangladesh. The sample represented both urban and rural mothers. Families were recruited for the survey based on a random selection process and subject to their consent to participate. The interviews were conducted at the households of selected families. The detailed methods are outlined elsewhere [21,22].

However, in brief, the study only analyzed samples with complete information about maternal use of the CoC and coupled with the respective complementary feeding indicators (ie, introduction to solid, semisolid, and soft food [ISSSF]; minimum meal frequency [MMF]; minimum dietary diversity [MDD]; and MAD). Therefore, women who were interviewed for birth history information (ie, data about women who had given birth) were filtered from the dataset for inclusion in this

XSL•FO
RenderX

study. Women who had a live birth within the last 3 and 2 years before the surveys, respectively, for the BDHS 2017-2018 and 2022 were then retrieved. These retrieved observations served as the source datasets to estimate maternal use of the CoC or the exposure variables. Finally, the maternal CoC variables were merged with variables accommodating the respective complementary feeding indicators, making up the analyzed samples for each complementary feeding indicator.

Exposure Variables

The exposure variables were maternal use of the CoC and its 3 components. The components included (1) at least 4 ANC visits by skilled health care providers during pregnancy, (2) SBA-assisted delivery, and (3) at least one PNC visit by skilled health care providers within 48 hours of delivery. In this study, skilled health care providers were denoted as physicians, nurses, midwives or paramedics, family welfare visitors, community SBAs, and subassistant community medical officers. Mothers were coded as 1 if they received all 3 components of the CoC and as 0 otherwise. Similarly, mothers who used any individual component of the CoC were coded as 1 or as 0 otherwise and analyzed separately.

Outcome Variables

The outcome variables included complementary feeding indicators (eg, ISSSF, MMF, MDD, and MAD). The status of ISSSF was estimated by calculating the proportion of children aged 6 to 8 months who were fed solid, semisolid, or soft food during the previous day or night. MMF refers to the minimum number of meals provided among children aged 6 to 23 months by their age category (6-8 months and 9-23 months) and breastfeeding status (breastfed and nonbreastfed). Furthermore, MDD indicates the minimum number of food groups fed to children aged 6 to 23 months. Finally, the estimation of the MAD considers both the MMF and MDD, indicating that a child had an MAD if both the MMF and MDD were met. Calculation of the outcome variables was based on BDHS guidelines [26]. Positive status of ISSSF, MMF, MDD, and MAD was coded as 1 and as 0 otherwise.

Covariates

Covariates were selected based on their expected association with the exposure variables and evidence from the published literature [27-29]. Therefore, the covariates considered were age at delivery (<19 years, 19-30 years, and 31-49 years); educational level (no education, primary education, secondary education, and higher education); husband's educational level (no education, primary education, secondary education, and higher education); occupation (not working or working); husband's occupation (not working or working); parity (1, 2-3, and 3); whether they had ever terminated a pregnancy (never or no and ever or yes); desired pregnancy (yes or no); wealth index (poorest, poorer, middle, richer, or richest); any exposure to television, radio, or the news (yes or no); whether they were facing problems such as obtaining permission and money and long distance to access health care (big problem or not a big problem); religion (Muslim or other); place of residence (rural or urban); division of residence (Barisal, Chittagong, Dhaka, Khulna, Mymensingh, Rajshahi, Rangpur, or Sylhet), and survey round (BDHS 2017-2018 or 2022).

Statistical Analysis

Cross-tabulation was used to observe the distribution of complementary feeding status by maternal use of the CoC and covariates, and the differences were observed using chi-square tests. The binary nature of the outcome variables led to the consideration of univariable and multivariable logistic regression models to estimate the effect of maternal use of the CoC, further divided into its individual components, on complementary feeding practices among children aged 6 to 23 months. Therefore, the individual effect of maternal use of the full CoC, receiving ≥4 ANC visits, delivery assisted by SBAs, and receiving PNC within 48 hours of delivery on the 4 complementary feeding indicators was observed, resulting in a total of 16 multivariable logistic regression models to observe the associations. The logistic regression models were adjusted for sampling weight, strata, primary sampling unit, and covariates using the *svy* command. Multicollinearity was observed using the SE of the odds ratio and variance inflation factor (VIF), considering an SE of >2 or VIF of >5 as indicators of the presence of multicollinearity [30,31]. All the analyses were conducted considering $P<.05$ as statistically significant. The Stata software (version 17; StataCorp) was used to conduct the statistical analyses.

Ethical Considerations

This study was exempted from receiving approval from an ethical review committee because it was conducted based on publicly available BDHS data, which involved interviewing mothers of children under 2 years of age to collect information. However, the BDHS was approved by the ICF Macro Institutional Review Board (United States) following the overall requirements of Title 45, Part 46 of the Code of Federal Regulations, "Protection of Human Subjects." Moreover, as the data collection procedure was carried out in Bangladesh, the National Research Ethics Committee of the Bangladesh Medical Research Council, Dhaka, Bangladesh, also reviewed and approved the BDHS. Before enrollment in the study, informed consent was obtained verbally from each participant and their intimate partners (for all ever-married women aged 15 - 49 years). Verbal consent was considered the most suitable option to ensure the participation of respondents who were incapable of reading. Very young children who were under 5 years of age and were born in 2014 or later were also considered in the sample when the data collection procedure was conducted by the BDHS. Verbal consent for children's participation in the survey was obtained from their mothers, who acted as legal guardians and provided authorization on behalf of their children. Before using the data for research purposes, the responses were deidentified. No compensation was provided to any participants for their responses or the information used in this research. Furthermore, this paper and its supplementary material do not contain any identifiable information or images of participants whose responses were used to conduct this study. Detailed ethical aspects and consent forms are available in the BDHS website and report.

XSL•FO
RenderX

Results

Figure 1 shows the results of the sample selection process. A total of 8561 observations were included in the analysis (n=887, 10.36% for ISSSF; n=4967, 58.02% for MMF; n=4967, 58.02% for MDD; and n=4967, 58.02% for MAD). The status of timely ISSSF, MMF, MDD, and MAD is shown in Figure 2.

Figure 1. Selection process of the analyzed sample from the Bangladesh Demographic and Health Survey (BDHS) 2017-2018 and 2022 data. CoC: continuum of care; ISSSF: introduction to solid, semisolid, and soft food; MAD: minimum acceptable diet; MDD: minimum dietary diversity; MMF: minimum meal frequency.

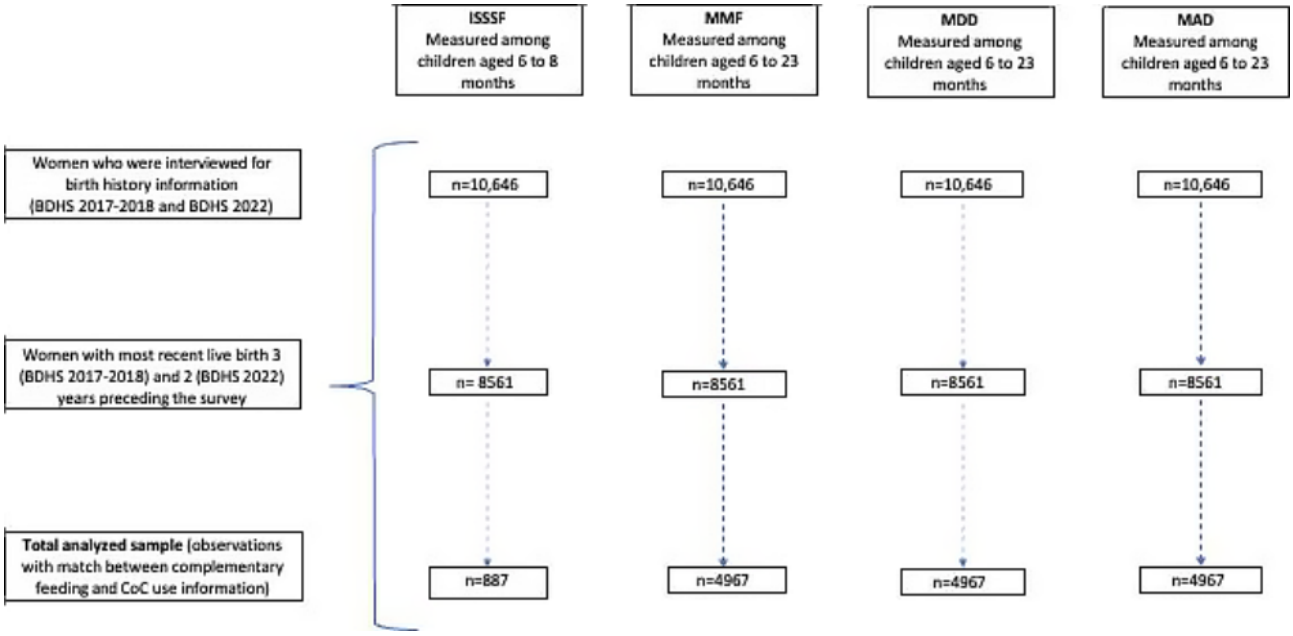
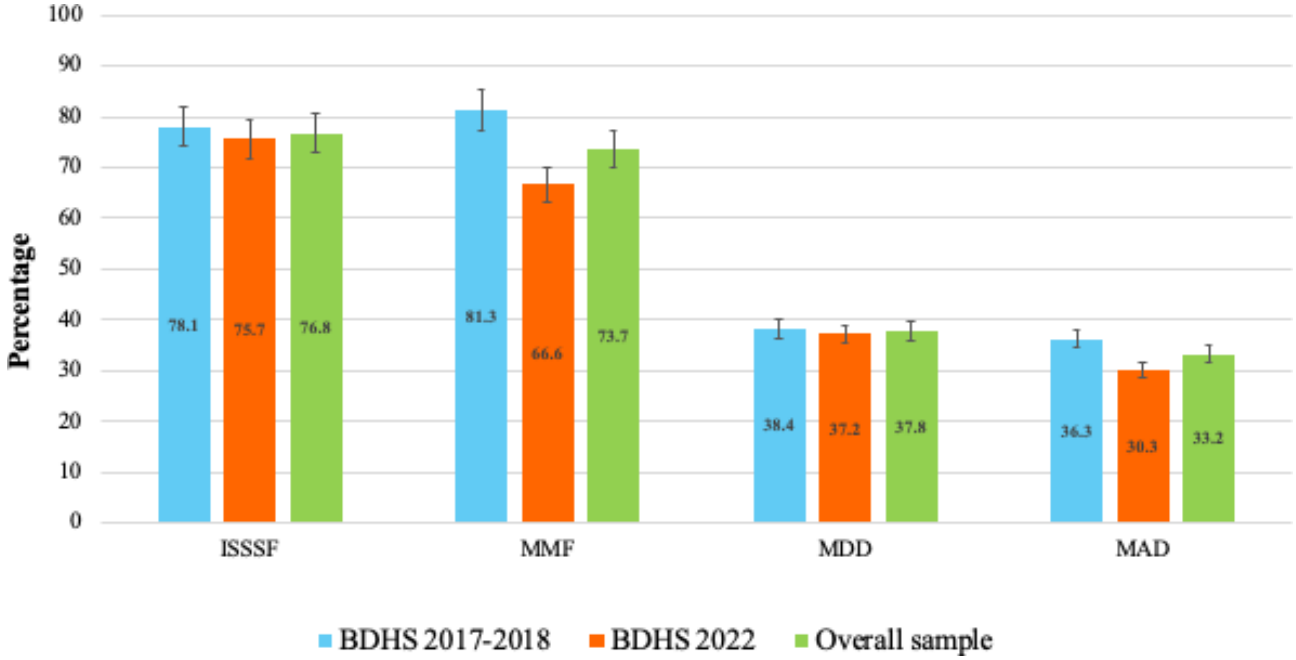


Figure 2. Status of complementary feeding practices in Bangladesh calculated from the analyzed sample. BDHS: Bangladesh Demographic and Health Survey; ISSSF: introduction to solid, semisolid, and soft food; MAD: minimum acceptable diet; MDD: minimum dietary diversity; MMF: minimum meal frequency.



The distribution of outcome variables by maternal use of the CoC, ANC visits, delivery assisted by SBAs, PNC, and maternal socioeconomic characteristics is shown in Table 1. The analysis shows that the distributions of all 4 outcome variables were statistically significantly (P<.05) and differed by maternal use of the CoC, status of receiving ≥4 ANC visits, status of receiving PNC within 48 hours of birth, maternal educational level, husband's educational level, maternal occupation, wealth index, perceived problems with accessing health care, and division of residence. In addition, the status of MMF, MDD, and MAD significantly differed by maternal status of media exposure and place of residence, whereas delivery assisted by SBAs, husband's occupation, and parity were variables that showed significant differences in the distribution of MDD and MAD.

XSL•FO
RenderX

Table . Distribution of complementary feeding indicators by status of maternal use of the continuum of care (CoC) and background characteristics[a].

Independent variable	ISSSF[b], n (%)			MMF[c], n (%)			MDD[d], n (%)			MAD[e], n (%)		
	No	Yes	*P* value	No	Yes	*P* value	No	Yes	*P* value	No	Yes	*P* value
CoC use			.001			<.001			<.001			<.001
Yes	39 (13.93)	241 (86.07)		290 (19.42)	1203 (80.58)		746 (49.97)	747 (50.03)		811 (54.32)	682 (45.6)	
≥4 ANC[f] visits			.02			<.001			<.001			<.001
Yes	70 (17.16)	338 (82.84)		450 (20.39)	1757 (79.61)		1173 (53.15)	1034 (46.85)		1277 (57.86)	930 (42.14)	
Delivery by SBAs[g]			.08			.11			<.001			<.001
Yes	108 (18.95)	462 (81.05)		764 (24.93)	2300 (75.07)		1763 (57.54)	1301 (42.46)		1926 (62.32)	1138 (37.14)	
PNC[h] within 48 hours of birth			.003			<.001			<.001			<.001
Yes	85 (17.10)	412 (82.90)		602 (22.82)	2036 (77.18)		1504 (57.01)	1134 (42.99)		1619 (61.37)	1019 (38.63)	
Age at delivery (y)			.39			.09			.74			.47
<19	20 (17.70)	93 (82.30)		177 (24.05)	559 (75.95)		448 (60.87)	288 (39.13)		482 (65.49)	254 (34.51)	
19-30	138 (21.94)	491 (78.06)		917 (26.60)	2531 (73.40)		2140 (62.06)	1308 (37.94)		2305 (66.84)	1143 (33.15)	
31-49	26 (17.93)	119 (82.07)		183 (23.37)	600 (76.63)		477 (60.92)	306 (39.08)		507 (64.75)	276 (35.25)	
Educational level			.03			<.001			<.001			<.001
No education	12 (26.09)	34 (73.91)		90 (32.85)	184 (67.15)		217 (79.20)	57 (20.80)		227 (82.85)	47 (17.15)	
Primary education	48 (23.41)	157 (76.59)		338 (27.06)	911 (72.94)		903 (72.30)	346 (27.70)		936 (74.94)	313 (25.06)	
Secondary education	104 (21.89)	371 (78.11)		659 (26.31)	1846 (73.69)		1528 (61.00)	977 (39.00)		1657 (66.15)	848 (33.85)	
Higher education	20 (12.42)	141 (87.58)		190 (20.23)	749 (79.77)		417 (44.41)	522 (55.59)		474 (50.48)	465 (49.52)	
Husband's educational level			.047			<.001			<.001			<.001
No education	30 (22.56)	103 (77.44)		230 (32.72)	473 (67.28)		523 (74.40)	180 (25.60)		553 (78.66)	150 (21.34)	
Primary education	66 (23.66)	213 (76.34)		412 (25.99)	1173 (74.01)		1070 (67.51)	515 (32.49)		1116 (70.41)	469 (29.59)	
Secondary education	66 (21.57)	240 (78.43)		426 (26.30)	1194 (73.70)		989 (61.05)	631 (38.95)		1081 (66.73)	539 (33.27)	
Higher education	21 (12.88)	142 (87.12)		189 (18.94)	809 (81.06)		444 (44.49)	554 (55.51)		500 (50.10)	498 (49.90)	
Occupation status			.007			<.001			.02			<.001
Not working	144 (23.19)	477 (76.81)		978 (28.83)	2414 (71.17)		2130 (62.79)	1262 (37.21)		2310 (68.10)	1082 (31.90)	

XSL•FO
RenderX

Independent variable	ISSSF[b], n (%)			MMF[c], n (%)			MDD[d], n (%)			MAD[e], n (%)		
	No	Yes	*P* value	No	Yes	*P* value	No	Yes	*P* value	No	Yes	*P* value
Working	40 (15.09)	225 (84.91)		299 (19.00)	1275 (81.00)		934 (59.34)	640 (40.66)		983 (62.45)	591 (37.55)	
Husband's occupation status			.78			.13			.04			.02
Not working	4 (23.53)	13 (76.47)		32 (32.32)	67 (67.68)		71 (71.72)	28 (28.28)		77 (77.78)	22 (22.22)	
Working	180 (20.71)	689 (79.29)		1236 (25.52)	3608 (74.48)		2980 (61.52)	1864 (38.48)		3200 (66,06)	1644 (33.94)	
Parity			.13			.07			<.001			<.001
1	57 (17.59)	267 (82.41)		445 (23.87)	1419 (76.13)		1067 (57.24)	797 (42.76)		1172 (62.88)	692 (37.12)	
2-3	101 (21.77)	363 (78.23)		694 (26.91)	1885 (73.09)		1628 (63.13)	951 (36.87)		1740 (67.7)	839 (32.53)	
>3	26 (26.26)	73 (73.74)		138 (26.34)	386 (73.66)		370 (70.61)	154 (29.39)		382 (72.90)	142 (27.10)	
Ever terminated a pregnancy			.52			.16			.23			.21
Yes	30 (18.87)	129 (81.13)		208 (23.83)	665 (76.17)		523 (59.91)	350 (40.09)		563 (64.49)	310 (35.51)	
Desired pregnancy			.65			.24			.78			.67
No	41 (21.93)	146 (78.07)		251 (24.27)	783 (75.73)		642 (62.09)	392 (37.91)		680 (65.76)	354 (34.24)	
Wealth index			.03			<.001			<.001			<.001
Poorest	51 (27.27)	136 (72.73)		328 (30.14)	737 (69.86)		780 (73.93)	275 (26.07)		817 (77.44)	238 (22.56)	
Poorer	36 (20.22)	142 (79.78)		285 (27.51)	751 (72.49)		684 (66.02)	352 (33.98)		726 (70.08)	310 (29.92)	
Middle	31 (16.85)	153 (83.15)		235 (25.54)	685 (74.46)		592 (64.35)	328 (35.65)		630 (68.48)	290 (31.52)	
Richer	43 (23.50)	140 (76.50)		244 (24.67)	745 (75.33)		580 (58.65)	409 (41.35)		632 (63.90)	357 (36.10)	
Richest	23 (14.84)	132 (85.16)		195 (20.17)	772 (79.83)		429 (44.36)	538 (55.64)		489 (50.57)	478 (49.43)	
Media exposure[i]			.12			<.001			<.001			<.001
Yes	97 (18.95)	415 (81.05)		657 (21.97)	2334 (78.03)		1674 (55.97)	1317 (44.03)		1807 (60.41)	1184 (39.59)	
Access to health care[j]			.005			<.001			<.001			<.001
Not a big problem	61 (16.27)	314 (83.73)		434 (21.35)	1599 (78.65)		1175 (57.80)	858 (42.20)		1265 (62.22)	768 (37.78)	
Big problem	123 (24.02)	389 (75.98)		843 (28.73)	2091 (71.27)		1890 (64.42)	1044 (35.58)		2029 (69.15)	905 (30.85)	
Religion			.96			.10			.46			.53
Muslim	168 (20.77)	641 (79.23)		1191 (26.00)	3389 (74.00)		2833 (61.86)	1747 (38.14)		3043 (66.44)	1537 (33.56)	
Others	16 (20.51)	62 (79.49)		86 (22.22)	301 (77.78)		232 (59.95)	155 (40.05)		251 (64.86)	136 (35.14)	
Residence			.08			.001			<.001			<.001

XSL•FO
RenderX

Independent variable	ISSSF[b], n (%)			MMF[c], n (%)			MDD[d], n (%)			MAD[e], n (%)		
	No	Yes	*P* value	No	Yes	*P* value	No	Yes	*P* value	No	Yes	*P* value
Urban	46 (17.10)	223 (82.90)		368 (22.63)	1258 (77.37)		909 (55.90)	717 (44.10)		990 (60.89)	636 (39.11)	
Rural	138 (22.33)	480 (77.67)		909 (27.21)	2432 (72.79)		2156 (64.53)	1185 (35.47)		2304 (68.96)	1037 (31.04)	
Division of residence			<.001			<.001			<.001			<.001
Barisal	20 (19.23)	84 (80.77)		168 (31.82)	360 (68.18)		355 (67.23)	173 (32.77)		375 (71.02)	153 (28.98)	
Chittagong	49 (36.84)	84 (63.16)		292 (35.57)	529 (64.43)		541 (65.90)	280 (34.10)		588 (71.62)	233 (28.38)	
Dhaka	30 (24.00)	95 (76.00)		170 (23.55)	552 (76.45)		418 (57.89)	304 (42.11)		452 (62.60)	270 (37.40)	
Khulna	7 (7.29)	89 (92.71)		79 (15.11)	444 (84.89)		279 (53.35)	244 (46.65)		297 (56.79)	226 (43.21)	
Mymensingh	14 (12.50)	98 (87.50)		137 (22.03)	485 (77.97)		369 (59.32)	253 (40.68)		401 (64.470)	221 (35.53)	
Rajshahi	13 (15.85)	69 (84.15)		124 (24.60)	380 (75.40)		307 (60.91)	197 (39.09)		328 (65.08)	176 (34.92)	
Rangpur	15 (13.76)	94 (86.24)		117 (19.70)	477 (80.30)		338 (56.90)	256 (43.10)		371 (62.46)	223 (37.54)	
Sylhet	36 (28.57)	90 (71.43)		190 (29.10)	463 (70.90)		458 (70.14)	195 (29.86)		482 (73.81)	171 (26.19)	
Survey round			.17			<.001			.37			<.001
BDHS[k] 2017-2018	77 (18.73)	334 (81.27)		426 (17.63)	1991 (82.37)		1476 (61.07)	941 (38.93)		1526 (63.14)	891 (36.86)	
BDHS 2022	107 (22.48)	369 (77.52)		851 (33.37)	1699 (66.63)		1589 (62.31)	961 (37.69)		1768 (69.33)	782 (30.67)	

[a]For dichotomous responses (yes or no), only the positive response (yes) is presented.

[b]ISSSF: introduction to solid, semisolid, and soft food.

[c]MMF: minimum meal frequency.

[d]MDD: minimum dietary diversity.

[e]MAD: minimum acceptable diet.

[f]ANC: antenatal care.

[g]SBA: skilled birth attendant.

[h]PNC: postnatal care.

[i]Reading newspapers, watching television, or listening to the radio.

[j]Perception that receiving permission to visit health care facilities, not having money to access health care, or the distance to health care facilities are big problems.

[k]BDHS: Bangladesh Demographic and Health Survey.

The association between complementary feeding practices and maternal use of the CoC and its components is shown in Table 2. The analysis shows that maternal use of either the CoC or its components had no significant association with timely ISSSF among children aged 6 to 8 months. On the other hand, MDD and MAD among children aged 6 to 23 months were statistically significantly associated with maternal use of the CoC and receiving ≥4 ANC visits.

XSL•FO
RenderX

Table . Multivariable regression models demonstrating the association between maternal use of the continuum of care (CoC) and complementary feeding indicators.

Exposure variable	ISSSF[a]		MMF[b]		MDD[c]		MAD[d]	
	AOR[e] (95% CI)	*P* value	AOR[e] (95% CI)	*P* value	AOR[e] (95% CI)	*P* value	AOR[e] (95% CI)	*P* value
Full CoC use								
No	Reference	—[f]	Reference	—	Reference	—	Reference	—
Yes	1.30 (0.77-2.20)	.33	1.13 (0.94-1.37)	.20	1.29[g] (1.10-1.51)	.002	1.32[g] (1.13-1.55)	.001
≥4 ANC[h] visits								
No	Reference	—	Reference	—	Reference	—	Reference	—
Yes	1.00 (0.64-1.57)	.98	1.23[g] (1.05-1.45)	.01	1.31[g] (1.13-1.51)	<.001	1.34[g] (1.15-1.56)	<.001
Delivery by SBAs[i]								
No	Reference	—	Reference	—	Reference	—	Reference	—
Yes	0.98 (0.63-1.52)	.92	0.94 (0.80-1.11)	.49	1.04 (0.89-1.22)	.62	1.03 (0.87-1.21)	.77
PNC[j] within 48 hours of birth								
No	Reference	—	Reference	—	Reference	—	Reference	—
Yes	1.27 (0.83-1.93)	.27	1.05 (0.89-1.24)	.55	0.99 (0.85-1.15)	.88	1.03 (0.88-1.20)	.72

[a]ISSSF: introduction to solid, semisolid, and soft food.

[b]MMF: minimum meal frequency.

[c]MDD: minimum dietary diversity.

[d]MAD: minimum acceptable diet.

[e]AOR: adjusted odds ratio. Adjusted for primary sampling unit, sampling strata, sampling weight, age at last delivery, educational level, husband's educational level, occupation, husband's occupation, parity, whether they had ever terminated a pregnancy, desired pregnancy, wealth index, media exposure, access to health care, religion, residence, division of residence, and survey round.

[f]Not applicable.

[g]*P*<.05.

[h]ANC: antenatal care.

[i]SBA: skilled birth attendant.

[j]PNC: postnatal care.

Mothers who received the full CoC were 29% and 32% more likely to provide an MDD (adjusted odds ratio [AOR] 1.29, 95% CI 1.10-1.51; *P*=.002) and MAD (AOR 1.32, 95% CI 1.13-1.55; *P*<.001) to their children, respectively, than their counterparts. Similarly, maternal status of receiving ≥4 ANC visits was significantly associated with 23%, 31%, and 34% higher likelihood of providing an MMF (AOR 1.23, 95% CI 1.05-1.45; *P*=.01), MDD (AOR 1.31, 95% CI 1.13-1.51; *P*<.001), and MAD (AOR 1.34, 95% CI 1.15-1.56; *P*<.001) to their children, respectively, compared to mothers who did not receive ≥4 ANC.

In all regression models, the SE of the odds ratio was <2, and the VIF was <3, indicating the absence of collinearity among the independent variables. The detailed regression models are provided in Multimedia Appendix 1. There were some observations (n=83) with missing values in covariates in the analyzed sample. Therefore, we also conducted the analysis after excluding observations with missing values. However, we did not find any significant difference in the direction and significance of the association between the analysis including observations with missing values in covariates and the complete case analysis.

Discussion

Principal Findings

Our findings demonstrate significant positive effects of maternal use of the CoC on providing an MDD and MAD to their children. Breaking down the effect of individual CoC components, the status of receiving ≥4 ANC visits exhibited a significant positive association with MMF, MDD, and MAD. However, we did not find any significant effect of the status of maternal use of CoC or its components on ISSSF.

XSL•FO
RenderX

Consistent with our findings, a previous study using the BDHS 2007 data also reported no significant association between ISSSF and the number of ANC and PNC visits [32]. Timely ISSSF is associated with exclusive breastfeeding because children who are not exclusively breastfed are likely to receive family foods before the age of 6 months, which results in disqualifying them from receiving timely ISSSF. In Bangladesh, the exclusive breastfeeding rate declined from 64% in 2017 to 53% in 2022; therefore, this might partly explain the statistically insignificant effect of maternal use of the CoC and its components on timely ISSSF. There is a scarcity of studies explicitly examining the association between maternal CoC use and timely ISSSF. Previous studies highlight that complementary feeding practices are primarily influenced by maternal knowledge, cultural beliefs, family practices, and household factors rather than formal health care service use [19,33]. Similarly, a study in rural Bangladesh revealed that only 19% of mothers knew the appropriate age for introducing complementary foods, and this knowledge was not significantly associated with their feeding practices [34]. This provides an insight that, although an increase in maternal CoC use indicates a better coverage of maternal health care services, it may not ensure better child feeding practices (eg, ISSSF) unless the maternal health care services are strengthened with targeted nutrition education and behavior change interventions. The analysis revealed that mothers who received ≥4 ANC visits and the full CoC had significantly higher odds of providing an MMF than their counterparts. This finding is consistent with those of a previous study conducted using the Bangladesh Integrated Household Survey, which reported a positive association between ≥4 ANC visits and MMF [35]. Indeed, ANC provides platforms for counseling on nutrition as well as the importance of child feeding frequency and dietary adequacy. MAD is a composite indicator of both MMF and MDD. In this study, we found that the status of using the full CoC and receiving ≥4 ANC visits had a significant positive association with MAD. Similar studies based on the Bangladesh Integrated Household Survey 2018 - 2019 and from other South Asian countries have also reported a positive association between ≥4 ANC visits and MAD [28,35-37]. Four or more ANC visits facilitate positive complementary feeding practices such as MDD and MAD through sustained opportunities for nutritional counseling [38]. Mothers who received the full CoC package were 29% more likely to provide foods meeting the MDD for their children. Notably, the trend of MDD in Bangladesh remained relatively consistent over a 10-year period from 2004 to 2014 [20]. Looking into the individual effects of the components of the CoC, the status of receiving ≥4 ANC visits also exhibited a significant association with MDD status. A previous study identified that mothers who received ≥4 ANC visits were 31% more likely to give their babies minimum diversified foods than those who did not have any ANC visits [39]. No significant effect of delivery by SBAs and PNC within 48 hours of delivery was observed on MDD. This could be because these services are more focused on maternal and essential newborn care rather than on complementary feeding practices, which require targeted counseling and behavior change communication. A previous study reported significant associations between dietary diversity score and status of receiving ≥4 ANC visits, which aligns with our findings [40].

This study was conducted using 2 consecutive nationally representative cross-sectional surveys (BDHS 2017 - 2018 and 2022) conducted after the National Nutrition Policy 2015 was adopted in Bangladesh. Therefore, the findings provide a valuable insight into how the benefits of nutrition mainstreaming to maternal and child health care platforms are reflected in complementary feeding practices. The exposure variables included both maternal use of the CoC and its components (ie, ≥4 ANC visits, delivery assisted by SBAs, and PNC within 48 hours of delivery), which provides a holistic overview of their association with recommended complementary feeding practices and which maternal health care component should be further strengthened with policy support to increase the benefits. Furthermore, adjustment of covariates increased the accuracy of the estimates. The surveys use validated measurements for the exposure and outcome variables, strengthening the internal validity. On the other hand, the cross-sectional design limits causal inference between maternal use of the CoC and complementary feeding practices. Key behavioral factors such as maternal nutritional knowledge and attitude, cultural norms, household food security, and intrahousehold food dynamics were not estimated in the survey and, therefore, not included in this analysis. Furthermore, the quality of ANC and PNC, especially nutrition counseling while providing these services, was not considered in the statistical models. These issues should be considered carefully when interpreting and using the findings.

Conclusions

Maternal use of the CoC, particularly through receiving ≥4 ANC visits, appears to be a significant determinant of improved MDD and MAD among children aged 6 to 23 months in Bangladesh. Enhancing maternal knowledge and practices related to exclusive breastfeeding may help facilitate the practice of timely ISSSF. Expanding maternal use of the CoC through policy support and targeted interventions presents a valuable opportunity to fully leverage its positive impact on maternal complementary feeding practices in Bangladesh.

Acknowledgments

The authors received no funding from any public, commercial, or not-for-profit entities to conduct this study. However, the authors acknowledge Lund University, Sweden, for providing article processing charge support for open access publishing of the article. Furthermore, no generative artificial intelligence was used in preparing any part of the paper.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: MAR
Data curation: MAR (lead), RI (supporting), RBD (supporting), JB (supporting)
Formal analysis: MAR (lead), SS (equal)
Investigation: MAR (lead), RI (supporting), RBD (supporting), JB (supporting), PS (supporting)
Methodology: SS (lead), MAR (equal), MRA (equal)
Project administration: MAR
Resources: MAR (lead), SS (equal)
Supervision: SS (lead), MRA (equal), MAR (supporting)
Validation: MAR (lead), RI (supporting), RBD (supporting), JB (supporting)
Visualization: MAR (lead), SS (equal), RI (supporting), RBD (supporting), JB (supporting), PS (supporting), MRA (supporting)
Writing – original draft: MAR (lead), RI (supporting), RBD (supporting), JB (supporting), PS (supporting)
Writing – review & editing: MAR (lead), SS (supporting), MRA (supporting), RI (supporting), RBD (supporting), JB (supporting), PS (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details of the regression model outcomes.
[PDF File, 1074 KB - publichealth_v11i1e76666_app1.pdf]

Checklist 1
STROBE checklist.
[PDF File, 144 KB - publichealth_v11i1e76666_app2.pdf]

References

1. Baye K, Faber M. Windows of opportunity for setting the critical path for healthy growth. Public Health Nutr 2015 Jul;18(10):1715-1717. [doi: 10.1017/S136898001500186X] [Medline: 26017475]
2. Indrio F, Pietrobelli A, Dargenio VN, et al. The key 1000 life-changing days. Global Pediatrics 2023 Jun;4:100049. [doi: 10.1016/j.gpeds.2023.100049]
3. Chang SY, He W, Chen CM. Complementary feeding and growth of infant and young child in China. Biomed Environ Sci 2008 Jun;21(3):264-268. [doi: 10.1016/S0895-3988(08)60040-9] [Medline: 18714827]
4. Onyango AW, Borghi E, de Onis M, Casanovas MDC, Garza C. Complementary feeding and attained linear growth among 6-23-month-old children. Public Health Nutr 2014 Sep;17(9):1975-1983. [doi: 10.1017/S1368980013002401] [Medline: 24050753]
5. Obbagy JE, English LK, Psota TL, et al. Complementary feeding and micronutrient status: a systematic review. Am J Clin Nutr 2019 Mar 1;109(Suppl_7):852S-871S. [doi: 10.1093/ajcn/nqy266] [Medline: 30982869]
6. Imdad A, Yakoob MY, Bhutta ZA. Impact of maternal education about complementary feeding and provision of complementary foods on child growth in developing countries. BMC Public Health 2011 Apr 13;11 Suppl 3(Suppl 3):S25. [doi: 10.1186/1471-2458-11-S3-S25] [Medline: 21501443]
7. Carter RC, Jacobson JL, Burden MJ, et al. Iron deficiency anemia and cognitive function in infancy. Pediatrics 2010 Aug;126(2):e427-e434. [doi: 10.1542/peds.2009-2097] [Medline: 20660551]
8. Przyrembel H. Timing of introduction of complementary food: short- and long-term health consequences. Ann Nutr Metab 2012;60 Suppl 2:8-20. [doi: 10.1159/000336287] [Medline: 22555185]
9. Masuke R, Msuya SE, Mahande JM, et al. Effect of inappropriate complementary feeding practices on the nutritional status of children aged 6-24 months in urban Moshi, Northern Tanzania: cohort study. PLoS One 2021;16(5):e0250562. [doi: 10.1371/journal.pone.0250562] [Medline: 33983950]
10. Dewey KG, Adu‐Afarwuah S. Systematic review of the efficacy and effectiveness of complementary feeding interventions in developing countries. Matern Child Nutr 2008 Apr;4(s1):24-85. [doi: 10.1111/j.1740-8709.2007.00124.x]
11. Michaelsen KF, Grummer-Strawn L, Bégin F. Emerging issues in complementary feeding: global aspects. Matern Child Nutr 2017 Oct;13 Suppl 2(Suppl 2):e12444. [doi: 10.1111/mcn.12444] [Medline: 29032617]
12. Global strategy for infant and young child feeding. World Health Organization (WHO). URL: https://www.who.int/publications/i/item/9241562218 [accessed 2025-09-25]

XSL•FO
RenderX

13. European Society for Paediatric Gastroenterology, Hepatology & Nutrition (ESPGHAN), European Academy of Paediatrics (EAP), European Society for Paediatric Research (ESPR), et al. World Health Organization (WHO) guideline on the complementary feeding of infants and young children aged 6-23 months 2023: A multisociety response. J Pediatr Gastroenterol Nutr 2024 Jul;79(1):181-188. [doi: 10.1002/jpn3.12248] [Medline: 38743631]
14. Nguyen PH, Kachwaha S, Tran LM, et al. Strengthening nutrition interventions in antenatal care services affects dietary intake, micronutrient intake, gestational weight gain, and breastfeeding in Uttar Pradesh, India: results of a cluster-randomized program evaluation. J Nutr 2021 Aug 7;151(8):2282-2295. [doi: 10.1093/jn/nxab131] [Medline: 34038529]
15. Haider R, Thorley V, Mihrshahi S. Improving complementary feeding practices through peer counselling in Bangladesh: moving successfully from project to program mode. Breastfeed Rev 2021;29(2):27-40. [doi: 10.3316/informit.920990521511100]
16. Rahman A, Bhuiyan MB, Das SK. Effect of short-term educational intervention on complementary feeding index among infants in rural Bangladesh: a randomized control trial. BMC Nutr 2022 Aug 2;8(1):73. [doi: 10.1186/s40795-022-00565-0] [Medline: 35918734]
17. Shi L, Zhang J. Recent evidence of the effectiveness of educational interventions for improving complementary feeding practices in developing countries. J Trop Pediatr 2011 Apr;57(2):91-98. [doi: 10.1093/tropej/fmq053] [Medline: 20558381]
18. Panjwani A, Heidkamp R. Complementary feeding interventions have a small but significant impact on linear and ponderal growth of children in low- and middle-income countries: a systematic review and meta-analysis. J Nutr 2017 Nov;147(11):2169S-2178S. [doi: 10.3945/jn.116.243857] [Medline: 28904113]
19. Manikam L, Robinson A, Kuah JY, et al. A systematic review of complementary feeding practices in South Asian infants and young children: the Bangladesh perspective. BMC Nutr 2017;3(1):56. [doi: 10.1186/s40795-017-0176-9] [Medline: 32153836]
20. Na M, Aguayo VM, Arimond M, Narayan A, Stewart CP. Stagnating trends in complementary feeding practices in Bangladesh: an analysis of national surveys from 2004 - 2014. Matern Child Nutr 2018 Nov;14 Suppl 4(Suppl 4):e12624. [doi: 10.1111/mcn.12624] [Medline: 29999230]
21. National Institute of Population Research and Training (NIPORT), ICF. Bangladesh demographic and health survey 2017-18. The DHS Program. 2020 Oct. URL: https://dhsprogram.com/publications/publication-FR344-DHS-Final-Reports.cfm [accessed 2025-05-29]
22. National Institute of Population Research and Training (NIPORT), ICF. Bangladesh demographic and health survey 2022. The DHS Program. 2024 Jan. URL: https://dhsprogram.com/publications/publication-fr386-dhs-final-reports.cfm [accessed 2025-10-02]
23. Ministry of Health and Family Welfare, Government of the People's Republic of Bangladesh. Second national plan of action for nutrition (2016-2025). FAOLEX Database. 2017 Aug. URL: https://faolex.fao.org/docs/pdf/bgd206820.pdf [accessed 2025-05-29]
24. Ministry of Health and Family Welfare, Government of the People's Republic of Bangladesh. National nutrition policy 2015. FAOLEX Database. URL: https://faolex.fao.org/docs/pdf/bgd152517.pdf [accessed 2025-10-02]
25. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD] [Medline: 17947786]
26. Croft TN, Allen CK, Zachary BW, et al. Guide to DHS statistics. The DHS Program. 2023. URL: https://dhsprogram.com/data/Guide-to-DHS-Statistics/index.cfm [accessed 2025-09-25]
27. Chowdhury MRK, Rahman MS, Khan MMH. Levels and determinants of complementary feeding based on meal frequency among children of 6 to 23 months in Bangladesh. BMC Public Health 2016 Sep 7;16(1):944. [doi: 10.1186/s12889-016-3607-7] [Medline: 27604631]
28. Dhami MV, Ogbo FA, Osuagwu UL, Agho KE. Prevalence and factors associated with complementary feeding practices among children aged 6-23 months in India: a regional analysis. BMC Public Health 2019 Aug 1;19(1):1034. [doi: 10.1186/s12889-019-7360-6] [Medline: 31370827]
29. Kothavale A, Meher T. Level of completion along continuum of care for maternal, newborn and child health services and factors associated with it among women in India: a population-based cross-sectional study. BMC Pregnancy Childbirth 2021 Oct 27;21(1):731. [doi: 10.1186/s12884-021-04198-2] [Medline: 34706680]
30. Chan YH. Biostatistics 202: logistic regression analysis. Singapore Med J 2004 Apr;45(4):149-153. [Medline: 15094982]
31. Chowdhury MRK, Khan HTA, Rashid M, et al. Differences in risk factors associated with single and multiple concurrent forms of undernutrition (stunting, wasting or underweight) among children under 5 in Bangladesh: a nationally representative cross-sectional study. BMJ Open 2021 Dec 13;11(12):e052814. [doi: 10.1136/bmjopen-2021-052814] [Medline: 34903543]
32. Kabir I, Khanam M, Agho KE, Mihrshahi S, Dibley MJ, Roy SK. Determinants of inappropriate complementary feeding practices in infant and young children in Bangladesh: secondary data analysis of demographic health survey 2007. Matern Child Nutr 2012 Jan;8 Suppl 1(Suppl 1):11-27. [doi: 10.1111/j.1740-8709.2011.00379.x] [Medline: 22168516]
33. Herman H, Mansur AR, Chang YJ. Factors associated with appropriate complementary feeding: a scoping review. J Pediatr Nurs 2023;71:e75-e89. [doi: 10.1016/j.pedn.2023.04.017] [Medline: 37150632]

XSL•FO
RenderX

34. Owais A, Suchdev PS, Schwartz B, et al. Maternal knowledge and attitudes towards complementary feeding in relation to timing of its initiation in rural Bangladesh. BMC Nutr 2019;5(1):7. [doi: 10.1186/s40795-019-0272-0] [Medline: 32153921]
35. Jubayer A, Nowar A, Islam S, Islam MH, Nayan MM. Complementary feeding practices and their determinants among children aged 6-23 months in rural Bangladesh: evidence from Bangladesh Integrated Household Survey (BIHS) 2018-2019 evaluated against WHO/UNICEF guideline -2021. Arch Public Health 2023 Jun 21;81(1):114. [doi: 10.1186/s13690-023-01131-1] [Medline: 37344900]
36. Dhami MV, Ogbo FA, Diallo TM, et al. Infant and young child feeding practices among adolescent mothers and associated factors in India. Nutrients 2021 Jul 12;13(7):2376. [doi: 10.3390/nu13072376] [Medline: 34371886]
37. Rahman MA, Kundu S, Rashid HO, Tohan MM, Islam MA. Socio-economic inequalities in and factors associated with minimum dietary diversity among children aged 6-23 months in South Asia: a decomposition analysis. BMJ Open 2023 Dec 20;13(12):e072775. [doi: 10.1136/bmjopen-2023-072775] [Medline: 38128933]
38. Wagris M, Seid A, Kahssay M, Ahmed O. Minimum meal frequency practice and its associated factors among children aged 6-23 months in Amibara district, North East Ethiopia. J Environ Public Health 2019;2019:8240864. [doi: 10.1155/2019/8240864] [Medline: 31929807]
39. Khan JR, Awan N, Sheikh MT. A multilevel and spatial analysis of the infant and young child feeding practices and associated factors among the under-2 aged children in Bangladesh. Child Care in Practice 2022 Apr 3;28(2):178-195. [doi: 10.1080/13575279.2019.1680528]
40. Ahmed KT, Karimuzzaman M, Pinky GN, et al. Association of dietary diversity of 6-23 months aged children with prenatal and postnatal obstetric care: evidence from a nationwide cross-sectional study. J Health Popul Nutr 2023 Nov 6;42(1):120. [doi: 10.1186/s41043-023-00470-7] [Medline: 37932856]

Abbreviations

ANC: antenatal care
AOR: adjusted odds ratio
BDHS: Bangladesh Demographic and Health Survey
CoC: continuum of care
ISSSF: introduction to solid, semisolid, and soft food
MAD: minimum acceptable diet
MDD: minimum dietary diversity
MMF: minimum meal frequency
PNC: postnatal care
SBA: skilled birth attendant
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
VIF: variance inflation factor

Edited by A Mavragani, T Sanchez; submitted 30.04.25; peer-reviewed by F Rahman, P Babar; revised version received 16.09.25; accepted 17.09.25; published 07.10.25.

Please cite as:
Rifat MA, Islam R, Didar RB, Bhowmick J, Sarkar P, Amin MR, Saha S
Effects of Maternal Use of the Continuum of Care on Complementary Feeding Practices in Bangladesh: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e76666
URL: https://publichealth.jmir.org/2025/1/e76666
doi:10.2196/76666

© M A Rifat, Rokibul Islam, Rinath Bintey Didar, Joya Bhowmick, Plabon Sarkar, Md Ruhul Amin, Sanjib Saha. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 7.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Use of Cigarettes, Cannabis, and Alcohol Among Asian American, Native Hawaiian, and Pacific Islander Adults: Community-Based National Survey Analysis

Vuong Van Do[1,2], PhD; Van My Ta Park[2,3], PhD; Nhung Nguyen[1,2,4], PhD; Pamela May Ling[1,4], MD, MPH; Marian Tzuang[3], PhD; Bora Nam[3], PhD; Marcelle M Dougan[5], ScD; Oanh L Meyer[2,6], PhD; Janice Y Tsoh[1,2,7], PhD

[1]Center for Tobacco Control Research and Education, University of California San Francisco, San Francisco, CA, United States

[2]Asian American Research Center on Health, University of California San Francisco, San Francisco, CA, United States

[3]Department of Community Health Systems, School of Nursing, University of California San Francisco, San Francisco, CA, United States

[4]Division of General Internal Medicine, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[5]Department of Public Health and Recreation, San José State University, San Jose, United States

[6]Department of Neurology, University of California, Davis, Davis, United States

[7]Department of Psychiatry and Behavioral Sciences, School of Medicine, University of California San Francisco, 675 18th Street, Nancy Friend Pritzker Psychiatry Building, San Francisco, CA, United States

Corresponding Author:
Janice Y Tsoh, PhD
Center for Tobacco Control Research and Education, University of California San Francisco, San Francisco, CA, United States

Abstract

Background: Asian American, Native Hawaiian, and Pacific Islander (AANHPI) populations have diverse cultural, immigration, and sociodemographic characteristics. Aggregated data could mask substantial differences in substance use between cultural subgroups in this population. Yet, studies examining substance use among the AANHPI population are limited.

Objective: This study aimed to describe cigarette, cannabis, and alcohol use among AANHPI adults by cultural subgroup and sex.

Methods: We analyzed data from 3411 AANHPI respondents of a multilingual national survey "COMPASS" during December 2021-May 2022. Primary outcomes were self-report current (every day or some days) use of cigarettes, cannabis, and alcohol. Cultural subgroups included Asian Indian, Ethnic Chinese, Filipino, Japanese, Korean, Native Hawaiian and Pacific Islander, Vietnamese, other cultural groups, and multicultural groups. Other covariates include sex, other sociodemographics, experiences of discrimination (Everyday Discrimination Scale), and mental health (Patient Health Questionnaire 4). Multivariable logistic regressions were used to examine correlates of each substance use among AANHPI adults.

Results: The prevalence of current cigarette, cannabis, and alcohol use was 4.2% (142/3359), 5.5% (184/3235) and 37.6% (1265/3361), respectively. Cigarette use ranged from 1.0% (1/100) in Asian Indian females to 14.8% (10/71) in multicultural males. Cannabis use ranged from 1.9% among Asian Indian (1/53) and Vietnamese males (4/211) to 15.7% (11/70) in multicultural females. Alcohol use varied from 6.6% (4/61) in Native Hawaiian and Pacific Islander females to 56.3% (40/71) among multicultural males. Male participants with elevated depression and anxiety symptoms were more likely to report using all 3 substances than males with minimal symptoms. However, depression and anxiety symptoms were only associated with cannabis and alcohol use among female participants. US-born female participants were more likely to report using all 3 substances compared to foreign-born females, while being US-born was only associated with higher odds of alcohol use among males. Perceived discriminatory experience was associated with higher odds of smoking in both sexes and alcohol drinking in males.

Conclusions: Cigarette smoking, cannabis, and alcohol use varied widely across AANHPI cultural groups, between and within each sex. These findings underscore the necessity to disaggregate data for substance use behaviors to guide health policy and intervention programs for AANHPI adults.

(JMIR Public Health Surveill 2025;11:e76465) doi:10.2196/76465

KEYWORDS

cigarette smoking; cannabis; alcohol; Asian Americans; Native Hawaiians and Pacific Islanders

XSL•FO
RenderX

Introduction

Tobacco, cannabis, and alcohol are the most commonly used substances among US adults, and tobacco and alcohol are leading causes of morbidity and mortality, especially cancer [1]. Among the Asian American, Native Hawaiian, and Pacific Islander (AANHPI) population, cancer has been the leading cause of death since 2000, with lung and liver cancer deaths the two most common among men and ranked first and fourth among women [2,3]. In addition, the 5 - year cancer - specific survival for AANHPI men (62%) was lower compared to non - Latinx White American men (68%) [2]. In addition, people who misuse one substance are also more likely to use other illicit drugs [4,5], including methamphetamine and fentanyl, the two leading causes of drug-related deaths among Asian Americans and Native Hawaiians and other Pacific Islanders (NHPI) [6]. One of the proposed strategies to reduce cancer and drug overdose mortality among the AANHPI population is to improve cancer screening and implement behavioral interventions to reduce substance use among high-risk groups [2]. Yet, studies examining substance use among the AANHPI population are limited.

According to data from the 2022 National Survey on Drug Use and Health, past month substance use prevalence among English-speaking US adults was 15.9% for cigarettes and cannabis and 52.9% for alcohol [7]. AANHPI adults are often combined into one or two groups in national reports on substance use to compare with other race and ethnic groups, which shows a lower prevalence of substance use among this group compared with other racial or ethnic groups [8,9]. In addition, the stereotype that AANHPI groups are model minority populations has negatively impacted data equity [10,11], leading to an underrepresentation of the AANHPI subgroups in clinical research [12]. However, the AANHPI population has many cultural subgroups with diverse cultural, immigration, and sociodemographic characteristics. Therefore, aggregated data could mask substantial differences in substance use between cultural subgroups in this population.

About 20.6 million people self-identify as AANHPI (not in combination with another race), accounting for about 6.2% of the total US population [13]. The Department of Health and Human Services provides guidelines for collecting standard racial and ethnic information such that the Asian American category is divided into seven subgroups (Asian Indian, Chinese, Filipino, Japanese, Korean, Vietnamese, and other Asians), and the NHPI group categorized into four subgroups (Native Hawaiian, Guamanian or Chamorro, Samoan, and Other Pacific Islanders) [14]. Despite these guidelines, AANHPI subgroups are often under-sampled in national surveys [15], preventing the disaggregation of data among this population. A few studies have examined tobacco use among the AANHPI population and found a wide range in the prevalence of use, including two national surveys [16,17], both of which were English-administered surveys and collected data more than a decade ago. In addition, a community-based survey of 306 NHPI adults during the COVID-19 pandemic found a high prevalence of smoking, alcohol, and substance use among this population [18]. Significant sex differences in substance use have also been noted among AANHPI populations, with males generally at higher risk of substance use compared to females [16,17,19]. Thus, it is crucial to monitor tobacco and other substance use within each sex across AANHPI subgroups, including both English and non-English speakers.

Disaggregated data are instrumental to understanding subgroup differences and addressing substance-use disparities that might otherwise go unnoticed. This effort can inform effective health policies and behavior change interventions to address the unique challenges faced by subgroups with a higher rate of substance use within AANHPI communities. To fill the literature gap, this study examined cigarettes, cannabis, and alcohol use among AANHPI adults, disaggregated by cultural subgroups and sexes using data from a national survey administered in English and multiple Asian languages.

Methods

Data Sources and Participants

This study used data from the "COVID-19 Effects on the Mental and Physical Health of Asian Americans & Pacific Islanders Survey Study" (COMPASS), a community-based national survey that assesses COVID-19 effects on AANHPI adults. Data collection of the first survey (COMPASS I) occurred during October 2020 and January 2021, with 5420 participants completing the survey. Details about participant recruitment and study procedures have been reported elsewhere [20]. Briefly, eligibility criteria included self-identification as AANHPI alone or in combination with other races and ethnicities, being able to read English or Chinese (traditional and simplified Chinese) or Korean or Samoan or Vietnamese (the survey languages), being 18 years old or older, and residing in the United States. Participants were recruited through community partners who served AANHPI communities, personal or professional networks, social media, ethnic media, and through the Collaborative Approach for Asian Americans, Native Hawaiians, and Pacific Islanders Research and Education registry [21]. Participants completed the survey by themselves via the study website or were interviewed by research staff over the phone or in person. Data were collected using REDCap (Research Electronic Data Capture; Vanderbilt University) [22].

COMPASS I participants were invited to complete a follow-up survey (COMPASS II) between December 2021 and May 2022. This study analyzed data from the 3411 participants who completed the COMPASS II survey.

Measures

Cigarette, Cannabis, and Alcohol Use

Participants were asked, "Do you now smoke cigarettes?" Similar questions were asked for cannabis and alcohol use. Response options were "every day," "some days," and "not at all." "Every day" or "some days" responses were coded as current use of the substance.

Cultural Subgroups

Cultural subgroups were self-reported and classified into Asian Indian, Ethnic Chinese, Filipino, Japanese, Korean, NHPI, Vietnamese, Other Asian group, and multicultural groups. The

XSL•FO
RenderX

Ethnic Chinese group included mainland Chinese, Hongkongers, Taiwanese, and Huaren. The multicultural group included participants who self-identified with two or more cultural groups.

Other Covariates

Perceived discriminatory experience (yes or no) was measured using a revised 8-item everyday discrimination scale [23]. Experience of discrimination was coded "yes" if participants reported any of the 8 items in the past 6 months. Depression and anxiety symptoms were measured using the Patient Health Questionnaire 4 (PHQ-4) and categorized as normal, mild, and moderate or severe levels [24]. Self-reported English proficiency was coded as limited if participants reported any of their abilities for speaking, reading, or writing English as "some," "a little," or "not at all." Other demographic covariates included age, biological sex (female vs male), education, marital status, census region, and nativity (US-born vs foreign-born).

Ethical Considerations

The institutional review board at the University of California San Francisco reviewed and approved the COMPASS study (UCSF IRB: #20 - 31925). Participants were provided with detailed information about the study, and informed consent was obtained before survey completion. Participants received a US $10 gift card upon survey completion. Personal details of participants were kept confidential, and a unique study identification number was assigned to each participant for data management. Deidentified data were used for the analysis.

Statistical Analysis

Data were analyzed from April 15 to May 30, 2024, using Stata version 17 (StataCorp LLC). Descriptive statistics were used to describe the characteristics of participants and substance use. Multivariable logistic regression models were used to examine the association between each type of substance use and the cultural group and other covariates for each sex. To compare the odds of using a substance between a cultural subgroup and the rest of the AANHPI adults, we repeated regression models for the dummy variable of each cultural subgroup. All statistical tests were 2-sided, and $P<.05$ was considered statistically significant.

Results

Participant Characteristics

Participant characteristics are provided in Table 1. Among 3411 participants, the majority self-identified as ethnic Chinese (n=1176, 34.5%), Korean (n=692, 20.3%), and Vietnamese (n=669, 19.6%). Participants from Japanese, Asian Indian, Filipino, and NHPI culture groups each accounted for less than 5% of the sample. Most participants were female (n=2262, 66.3%), had a bachelor's degree or higher (n=2527, 74.1%), were born abroad (n=2175, 63.8%), and reported recent discriminatory experiences (n=2050, 60.1%).

XSL•FO
RenderX

Table . Characteristics of participants in the COMPASS II study in 2021 (n=3411).

Characteristics	Values n (%)
Cultural group	
Asian Indian	156 (4.6)
Ethnic Chinese[a]	1176 (34.5)
Filipino	115 (3.4)
Japanese	165 (4.8)
Korean	692 (20.3)
Native Hawaiians and Pacific Islanders	95 (2.8)
Vietnamese	669 (19.6)
Other cultural group	118 (3.5)
Multicultural group	225 (6.6)
Age group (years)	
18 - 29	695 (20.4)
30 - 49	1250 (36.6)
50 - 64	979 (28.7)
65+	487 (14.3)
Sex	
Male	1149 (33.7)
Female	2262 (66.3)
Education	
Less than high school	160 (4.7)
Completed high school or equivalent	344 (10.1)
Some college no degree	343 (10.1)
Bachelor’s degree or higher	2527 (74.1)
Missing	37 (1.1)
Marital status	
Single	901 (26.4)
Married or living with a partner	2240 (65.7)
Separated, divorced, or widowed	250 (7.3)
Missing	20 (0.6)
Census Region	
Midwest	289 (8.5)
Northeast	344 (10.1)
South	472 (13.8)
West	2303 (67.6)
Nativity	
US-born	1236 (36.2)
Foreign born	2175 (63.8)
Limited English proficiency	
Yes[b]	657 (19.3)
No	2754 (80.7)
Perceived discriminatory experience	

XSL•FO
RenderX

Characteristics	Values n (%)
Yes	2050 (60.1)
No	1320 (38.7)
Missing	41 (1.2)
Depression and anxiety symptoms[c]	
Normal	2215 (64.9)
Mild	853 (25.0)
Moderate or severe	343 (10.1)

[a]Ethnic Chinese includes mainland Chinese, Hongkonger, Taiwanese, and Huaren.

[b]Self-reported English proficiency was categorized as limited ("yes") if speaking or reading or writing English indicated as "some," "a little," or "not at all."

[c]Depression and anxiety symptoms were measured using 4 items from the Patient Health Questionnaire 4 (PHQ-4).

Prevalence of Cigarette, Cannabis, and Alcohol Use by Cultural Subgroups and Sex

The proportions of current substance use among AANHPI adults by cultural groups and sex are presented in Table 2. The overall prevalence of cigarette smoking among AANHPI adults was 4.2% (142/3359), with 8.4% (94/1123) of males and 2.1% (48/2236) of females reporting current smoking. Among male AANHPI adults, the prevalence of current cigarette smoking varied widely from 3.7% (2/54) among Asian Indians to 9.8% (21/214) among Vietnamese and 14.8% (10/71) among the multicultural group. The prevalence of current cigarette smoking among AANHPI females was low (<3%) in most cultural subgroups, except for NHPI (4/61, 6.6%) and those from the multicultural group (8/153, 5.2%).

XSL•FO
RenderX

Table . Cigarette smoking, cannabis, and alcohol use among Asian Americans, Native Hawaiians, and Pacific Islanders (AANHPI) by cultural subgroups and sex.

		Current use (every day or some days)		
Cultural subgroups		Cigarette (n=3359), n/N (%)	Cannabis (n=3325), n/N (%)	Alcohol (n=3361), n/N (%)
Asian Indian				
	Overall	3/154 (1.9)	5/153 (3.3)	58/153 (37.9)
	Male	2/54 (3.7)	1/53 (1.9)	29/53 (54.7)
	Female	1/100 (1.0)	4/100 (4.0)	29/100 (29.0)
Ethnic Chinese				
	Overall	34/1157 (2.9)	63/1146 (5.5)	415/1159 (35.8)
	Male	23/379 (6.1)	25/375 (6.7)	156/382 (40.8)
	Female	11/778 (1.4)	38/771 (4.9)	259/777 (33.3)
Filipino				
	Overall	5/113 (4.4)	12/113 (10.6)	68/114 (59.7)
	Male[a]	3/31 (9.7)[a]	3/31 (9.7)[a]	24/32 (75.0)[a]
	Female	2/82 (2.4)	9/82 (11.0)	44/82 (53.7)
Japanese				
	Overall	8/161 (5.0)	9/158 (5.7)	71/163 (43.6)
	Male[a]	5/45 (11.1)[a]	3/45 (6.7)[a]	23/47 (48.9)[a]
	Female	3/116 (2.6)	6/113 (5.3)	48/116 (41.4)
Korean				
	Overall	33/676 (4.9)	22/670 (3.3)	316/679 (46.5)
	Male	23/266 (8.6)	8/263 (3.0)	141/265 (53.2)
	Female	10/410 (2.4)	14/407 (3.4)	175/414 (42.3)
NHPI[b]				
	Overall	9/95 (9.5)	5/94 (5.3)	7/95 (7.4)
	Male[a]	5/34 (14.7)[a]	2/34 (5.9)[a]	3/34 (8.8)[a]
	Female	4/61 (6.6)	3/60 (5.0)	4/61 (6.6)
Vietnamese				
	Overall	28/661 (4.2)	29/653 (4.4)	167/658 (25.4)
	Male	21/214 (9.8)	4/211 (1.9)	69/213 (32.4)
	Female	7/447 (1.6)	25/442 (5.7)	98/445 (22.0)
Other cultural group				
	Overall	4/118 (3.4)	10/117 (8.5)	59/118 (50.0)
	Male[a]	2/29 (6.9)[a]	4/29 (13.8)[a]	20/29 (69.0)[a]
	Female	2/89 (2.2)	6/88 (6.8)	39/89 (43.8)
Multicultural group				
	Overall	18/224 (8.0)	29/221 (13.1)	104/222 (46.8)
	Male	10/71 (14.8)	11/70 (15.7)	40/71 (56.3)
	Female	8/153 (5.2)	18/151 (11.9)	64/151 (42.4)
Overall AANHPI		142/3359 (4.2)	184/3235 (5.5)	1265/3361 (37.6)
Male AANHPI		94/1123 (8.4)	61/1111 (5.5)	505/1126 (44.8)
Female AANHPI		48/2236 (2.1)	123/2214 (5.6)	760/2253 (34.0)

XSL•FO
RenderX

[a]Unstable estimates (sample size <50)

[b]NHPI: Native Hawaiians and Pacific Islanders.

The overall prevalence of cannabis use was 5.5% (184/3235) among AANHPI adults, with the prevalence almost equal between males and females. Cannabis use prevalence was highest among Filipinos (12/113, 10.6%) and multicultural participants (29/221, 13.1%). Among males, cannabis use ranged widely from 1.9% in participants of Asian Indian (1/53) and Vietnamese (4/211) cultures to 15.7% (11/70) in those of the multicultural group. Cannabis use was highest among females of the multicultural group (18/151, 11.9%) and Filipino (9/82, 11%). In addition, cannabis use was more prevalent among females than males for participants of Vietnamese (5.7% [25/442] vs 1.9% [4/211]) and Asian Indian (4% [4/100] vs 1.9% [1/53]) cultures.

Alcohol use was common among AANHPI adults (1265/3361, 37.6%) and more common among males (505/1126, 44.8%) than females (760/2235, 34%). Among cultural subgroups, the prevalence of current alcohol use varied from 7.4% (7/95) among NHPIs to 59.7% (68/114) among Filipinos. Male participants of Asian Indian (58/153, 54.7%), Korean (141/265, 53.2%), or a multicultural group (40/71, 56.3%) were most likely to report current alcohol use. Among females, alcohol use was most common among Filipinos (44/82, 53.7%), Koreans (175/414, 42.3%), Japanese (48/116, 41.4%), and the multicultural group (64/151, 42.4%).

Correlates of Substance Use Among AANHPI Males

Among male participants, there were no statistically significant differences in the odds of cigarette smoking between each cultural subgroup compared with the rest of the AANHPI adults (Table 3). However, compared with the rest of the AANHPI males, Vietnamese males were less likely (adjusted odds ratio [AOR] 0.26, 95% CI 0.09 - 0.77) to report cannabis use, whereas those males of the multicultural group had higher odds of reporting cannabis use (AOR 2.39, 95% CI 1.10 - 5.19). In addition, Korean males had higher odds of alcohol use (AOR 1.41, 95% CI 1.02 - 1.95) compared with other males, while Ethnic Chinese males were less likely to report alcohol use (AOR 0.63, 95% CI 0.48 - 0.83) compared with their other male peers.

XSL•FO
RenderX

Table . Associated factors of cigarette smoking, cannabis, and alcohol use among Asian Americans and Native Hawaiians and Pacific Islander males.

	Current use (daily or some days), AOR[a] (95% CI)		
Characteristics	Cigarette (n=1096)	Cannabis (n=1082)	Alcohol (n=1097)
Cultural group[b]			
Asian Indian	0.55 (0.12 - 2.44)	0.47 (0.06 - 3.73)	1.72 (0.95 - 3.12)
Ethnic Chinese	0.66 (0.39 - 1.11)	1.23 (0.69 - 2.18)	0.63[c] (0.48 - 0.83)
Filipino	__[d]	—	—
Japanese	—	—	—
Korean	1.15 (0.65 - 2.05)	0.59 (0.26 - 1.36)	1.41[e] (1.02 - 1.95)
NHPI[f]	—	—	—
Vietnamese	1.32 (0.73 - 2.39)	0.26[e] (0.09 - 0.77)	0.80 (0.57 - 1.14)
Other cultural group	—	—	—
Multicultural group	1.80 (0.83 - 3.90)	2.39[e] (1.10 - 5.19)	1.58 (0.95 - 2.63)
Age group (years)			
18 - 29	Reference	Reference	Reference
30 - 49	1.19 (0.52 - 2.69)	0.32[c] (0.15 - 0.70)	1.58[e] (1.02 - 2.44)
50 - 64	0.87 (0.34 - 2.21)	0.19[c] (0.07 - 0.52)	1.00 (0.62 - 1.62)
65+	0.36 (0.12 - 1.13)	0.10[c] (0.02 - 0.45)	0.75 (0.43 - 1.31)
Education			
Less than high school	Reference	Reference	Reference
Completed high school or equivalent	0.40[e] (0.16 - 0.99)	0.18[e] (0.03 - 0.88)	0.53 (0.25 - 1.15)
Some college no degree	0.32[e] (0.11 - 0.99)	0.10[c] (0.02 - 0.56)	0.77 (0.34 - 1.75)
Bachelor's degree or higher	0.16[g] (0.06 - 0.41)	0.13[c] (0.03 - 0.58)	0.72 (0.35 - 1.49)
Limited English proficiency	1.61 (0.81 - 3.21)	0.69 (0.19 - 2.59)	0.97 (0.63 - 1.49)
Marital status			
Single	Reference	Reference	Reference
Married or living with a partner	1.59 (0.79 - 3.19)	0.96 (0.47 - 1.95)	1.17 (0.81 - 1.67)
Separated, divorced, or widowed	6.19[g] (1.97 - 19.53)	0.74 (0.08 - 7.07)	1.30 (0.58 - 2.93)
US-born vs foreign born	1.14 (0.63 - 2.09)	1.39 (0.73 - 2.66)	1.53[e] (1.11 - 2.11)
Perceived discriminatory experience	2.08[i] (1.19 - 3.61)	1.42 (0.74 - 2.74)	1.31[e] (1.01 - 1.72)
Depression and anxiety symptoms[h]			
Normal	Reference	Reference	Reference
Mild	2.11[c] (1.25 - 3.55)	1.91 (0.97 - 3.76)	1.40[e] (1.02 - 2.44)
Moderate or severe	2.80[c] (1.42 - 5.51)	5.41[g] (2.58 - 11.37)	1.00 (0.64 - 1.59)
Census region			
Midwest	Reference	Reference	Reference
Northeast	1.81 (0.60 - 5.44)	1.20 (0.27 - 5.22)	0.98 (0.55 - 1.75)
South	1.97 (0.71 - 5.49)	1.92 (0.53 - 6.97)	0.67 (0.41 - 1.17)
West	1.24 (0.47 - 3.31)	2.45 (0.75 - 8.03)	0.80 (0.50 - 1.29)

[a]AOR: adjusted odds ratio.

XSL•FO
RenderX

[b]AOR for cultural subgroups were from comparisons between a specific subgroup and all the rest of the participants, using different regressions that controlled for the same covariates.

[c]*P*<.01.

[d]Not available.

[e]*P*<.05 .

[f]NHPI: Native Hawaiians and Pacific Islanders.

[g]*P*<.001.

[h]Depression and anxiety symptoms were measured using 4 items from the Patient Health Questionnaire 4 (PHQ-4). Estimates were not produced for Filipino, Japanese, NHPI, and other Asian American groups due to the small sample size of these groups.

The perceived discriminatory experience was significantly associated with higher odds of smoking (AOR 2.08, 95% CI 1.19 - 3.61) and drinking alcohol (AOR 1.31, 95% CI 1.01 - 1.72) but not using cannabis. Male participants with moderate or severe depression and anxiety symptoms were almost 3 times more likely to report smoking (95% CI 1.42 - 5.51) and 5 times more likely to report cannabis use (95% CI 2.58 - 11.37) compared with those who had a normal level of depression and anxiety symptoms. Being born in the United States was associated with higher odds of drinking alcohol, but not for cigarette and cannabis use. English proficiency was not found to be associated with using any of these 3 substances. In addition, male participants of older age groups were less likely to report cannabis use than young adults aged 18 - 29 years, and those with higher levels of education were also less likely to report smoking and cannabis use compared to those with less than high school education.

Correlates of Substance Use Among AANHPI Females

Factors associated with substance use among female participants are shown in Table 4. The odds of smoking among female Koreans were twice (AOR 2.27, 95% CI 1.04 - 4.98) that of other non-Korean AANHPI females. Female participants of the multicultural group were about 2 times more likely to report cannabis use (AOR 1.89, 95% CI 1.06 - 3.37) compared with the rest. Female participants with Filipino (AOR 1.84, 95% CI 1.15 - 2.93) and Korean (AOR 1.70, 95% CI 1.32 - 2.19) cultural backgrounds were more likely to report alcohol use, while those females with ethnic Chinese, Vietnamese, and NHPI backgrounds were less likely to report alcohol use compared with females of the rest.

XSL•FO
RenderX

Table . Associated factors of cigarette smoking, cannabis, and alcohol use among Asian Americans and Native Hawaiians and Pacific Islander females.

	Current use (daily or some days), AOR[a] (95% CI)		
Characteristics	Cigarette (n=2173)	Cannabis (n=2151)	Alcohol (n=2174)
Cultural group[b]			
Asian Indian	0.61 (0.08 - 4.77)	0.86 (0.30 - 2.51)	0.70 (0.44 - 1.13)
Ethnic Chinese	0.50 (0.25 - 1.04)	0.66 (0.43 - 1.01)	0.78[c] (0.64 - 0.95)
Filipino	1.16 (0.26 - 5.30)	1.46 (0.67 - 3.18)	1.84[c] (1.15 - 2.93)
Japanese	0.74 (0.16 - 3.32)	0.99 (0.40 - 2.46)	1.40 (0.92 - 2.12)
Korean	2.27[c] (1.04 - 4.98)	0.84 (0.46 - 1.55)	1.70[d] (1.32 - 2.19)
Native Hawaiians and Pacific Islanders	2.49 (0.72 - 8.63)	1.06 (0.21 - 5.30)	0.27[c] (0.08 - 0.90)
Vietnamese	0.63 (0.26 - 1.53)	1.09 (0.66 - 1.82)	0.64[e] (0.48 - 0.84)
Other cultural group	0.74 (0.14 - 3.93)	1.60 (0.57 - 4.49)	1.10 (0.67 - 1.80)
Multicultural group	2.11 (0.91 - 4.91)	1.89[c] (1.06 - 3.37)	1.29 (0.90 - 1.83)
Age group (years)			
18 - 29	Reference	Reference	Reference
30 - 49	4.46[e] (1.71 - 11.62)	0.35[d] (0.20 - 0.63)	0.70[c] (0.52 - 0.94)
50 - 64	3.66[c] (1.22 - 10.94)	0.35 (0.93 - 10.54)	0.56[e] (0.40 - 0.79)
65+	2.72 (0.65 - 5.84)	0.06[e] (0.01 - 0.31)	0.45[d] (0.29 - 0.69)
Education			
Less than high school	Reference	Reference	Reference
Completed high school or equivalent	0.22[c] (0.06 - 0.85)	0.12[e] (0.02 - 0.58)	0.63 (0.28 - 1.40)
Some college no degree	0.33 (0.09 - 1.20)	0.45 (0.11 - 1.80)	1.31 (0.62 - 2.75)
Bachelor's degree or higher	0.07[d] (0.02 - 0.26)	0.24[c] (0.06 - 0.93)	1.86 (0.92 - 3.75)
Limited English proficiency	0.75 (0.25 - 2.26)	0.34 (0.10 - 1.12)	0.53[d] (0.37 - 0.75)
Marital status			
Single	Reference	Reference	Reference
Married or living with a partner	0.34[e] (0.16 - 0.72)	1.10 (0.66 - 1.84)	1.10 (0.84 - 1.43)
Separated, divorced, or widowed	0.36 (0.10 - 1.21)	2.61[c] (1.07 - 6.34)	1.19 (0.76 - 1.85)
US-born vs Foreign born	3.73[c] (1.68 - 8.30)	2.73[d] (1.70 - 4.38)	1.42[e] (1.14 - 1.77)
Perceived discriminatory experience	2.21[c] (1.02 - 4.80)	1.25 (0.80 - 1.94)	1.00 (0.82 - 1.23)
Depression and anxiety symptoms[f]			
Normal	Reference	Reference	Reference
Mild	1.35 (0.67 - 2.73)	1.72[c] (1.09 - 2.72)	1.35[e] (1.08 - 1.68)
Moderate or severe	1.78 (0.76 - 4.13)	2.63[d] (1.56 - 4.44)	1.46[c] (1.07 - 1.98)
Census region			
Midwest	Reference	Reference	Reference
Northeast	1.59 (0.31 - 8.11)	6.02[e] (1.81 - 20.02)	1.00 (0.64 - 1.57)
South	2.20 (0.48 - 10.04)	3.14 (0.93 - 10.54)	1.02 (0.67 - 1.56)
West	1.47 (0.37 - 5.84)	3.91[c] (1.29 - 11.83)	0.75 (0.52 - 1.08)

XSL•FO
RenderX

[a]AOR: adjusted odds ratio.

[b]AORs for cultural subgroups were from comparisons between a specific subgroup and all the rest of the participants, using different regressions that controlled for the same covariates.

[c]$P<.05$.

[d]$P<.001$

[e]$P<.01$.

[f]Depression and anxiety symptoms were measured using 4 items from the Patient Health Questionnaire 4 (PHQ-4).

The perceived discriminatory experience was significantly associated with smoking (AOR 2.21, 95% CI 1.02 - 4.80) among female participants but not for other substances, whereas depression and anxiety symptoms were not associated with smoking. However, female participants with elevated depression and anxiety symptoms were more likely to report using cannabis and alcohol than those females with a normal level of depression and anxiety symptoms. Being born in the United States was significantly associated with a higher likelihood of using all 3 substances among AANHPI females. Older females were more likely to smoke cigarettes than young females between ages 18 and 29 years, but the relationship was the opposite for cannabis and alcohol use. Limited English proficiency was significantly associated with lower odds of alcohol drinking (AOR 0.53, 95% CI 0.37 - 0.75) but not cigarette and cannabis use. In addition, females of older age groups were more likely to report smoking but less likely to report cannabis and alcohol use compared to their young adult counterparts aged 18 - 29 years. Female participants with higher levels of education were also less likely to report smoking and cannabis use compared to those with less than high school education.

Discussion

Principal Findings

Given that disaggregated data about substance use by sex among the AANHPI communities are scarce, this study adds several important findings to the literature. First, our study showed different patterns of use of 3 substances, including cigarettes, cannabis, and alcohol, among a nationally recruited sample of AANHPI adults. Second, the study provided disaggregated data for each type of substance use and indicated a wide variation in the prevalence of use by cultural subgroups within and between each sex. For example, smoking was mostly among AANHPI males and ranged from 3.7% among Asian Indians to 14.8% among the multicultural group; whereas, the rates of cannabis use among AANHPI adults were similar between the 2 sexes, and Filipinos reported the highest rate of current cannabis use. Alcohol use was common among males of some subgroups, including Filipino, Indian, and Korean. Third, findings revealed different factors associated with each type of substance use within each sex.

Despite a low overall prevalence of cigarette smoking (4.2%) among AANHPI adults, the prevalence among AANHPI males was almost four times higher than among AANHPI females. The sex disparity in smoking remained consistent with findings from previous studies [16,17]. In addition, smoking prevalence varied widely across cultural subgroups among male AANHPI adults. Particularly, participants of Vietnamese (9.8%) and multicultural subgroups (14.8%) had the highest prevalence of current smoking, which was comparable to estimates for general US adults in 2022 (11.6%) or adults of other racial/ethnic groups such as non-Hispanic White (12.7%), non-Hispanic Black (14.2%), and Hispanic (8.0%) [25] of the National Health Interview Survey. Our finding confirms the need for interventions focusing on smoking among high-risk subgroups of AANHPI males, given that cancer is the leading cause of death and lung cancer is the leading cause of cancer mortality among Asian American men [26].

We found that experience of discrimination was significantly associated with current smoking among both male and female AANHPI adults, which is consistent with previous studies conducted before and during the COVID-19 pandemic [27-29]. Some AANHPI adults might have adopted smoking as it may have been perceived as a way to cope with increased discrimination experience that they faced during the pandemic [30]. In addition, racial discrimination may also contribute to depressive symptoms by increasing internalized negative self-perceptions and lowering self-esteem [31]. However, elevated levels of depression and anxiety symptoms were significantly associated with higher odds of smoking among AANHPI males only but not among AANHPI females in our study. Current literature shows mixed results in the association between sex, depressive symptoms, and smoking, with some studies reporting a significant smoking-depression association in women but not in men [32-34], some studies indicating a significant association in both sexes [35], and others showing a significant association among males but not in females [36]. These mixed findings could be attributed to the differences in study samples and analysis modeling. Future studies should explore the mechanism for sex differences in the association between smoking and depression.

Regarding cannabis use, we found that Filipino and multicultural groups had the highest rate of cannabis use. The finding of a high rate of cannabis use among Filipinos was consistent with the findings from another study which also found Filipinos had the highest rate of past 12-month drug use (eg, cannabis, cocaine, prescription drugs without physician recommendation, and other drugs) among Asian Americans [19]. The overall prevalence of cannabis use among NHPIs in our study was 5.3%. There are limited studies on cannabis use among NHPIs [37], including one study conducted during the COVID-19 pandemic (between April and November 2021) reporting the rate of lifetime use of illicit substance use (eg, cannabis, opioids) of 35% among NHPIs [18]. Given the small sample of the NHPI group and the underrepresentation of NHPI males in our study, future studies should re-estimate the prevalence of cannabis among NHPIs.

Previous studies that do not take into account ethnic subgroups often reported lower rates of drug use (including cannabis) among female Asian Americans than their male counterparts

XSL•FO
RenderX

and females of other race or ethnic groups [19,37]. However, female participants of Asian Indian and Vietnamese groups even seemed to have a higher rate of cannabis use than their male counterparts in our study, indicating the need for taking into account both ethnic subgroups and sex differences in cannabis use among AANHPI adults. We found that younger age (ie, 18 - 29 y) and elevated levels of depression and anxiety symptoms were associated with a higher likelihood of cannabis use in both male and female AANHPI adults, which was also consistent with the previous study [19]. However, in our study, discrimination experience and English proficiency were not associated with cannabis among AANHPI males or females, and being born in the United States was only associated with cannabis use among female AANHPI adults.

Similar to smoking and cannabis use, national data (ie, National Survey on Drug Use and Health) often indicate that Asian Americans have the lowest estimated current alcohol use compared to other racial groups (eg, 40.3% among Asian Americans vs 57% among Whites, 45.9% among African Americans, and 49.2% among Hispanics) [38]. Although the overall prevalence of alcohol use among AANHPI adults in this study (37.7%) was comparable to the rate reported from the national data, our study showed that some subgroups (eg, Filipinos or Koreans) had the current alcohol use rate as high as other racial groups. Previous studies also indicate important correlates of alcohol drinking are racial discrimination experience, depression and anxiety symptoms, and nativity [30,39]. This study added more insights to the variation of these correlates by sex, such as a significant association between alcohol use and discrimination experience among AANHPI males but not among AANHPI females.

Limitations

This study has several limitations. First, despite a large nationally recruited sample of AANHPI adults, the generalizability of the findings to all AANHPI adults across the United States was limited due to a nonprobability sampling approach. In addition, we were unable to estimate the prevalence of substance use among males of some cultural subgroups (ie, Japanese, Filipino, and NHPI) due to a small sample of these groups. Data from this study were collected during the COVID-19 pandemic, so the estimated prevalence of substance use for participants could be lower than when measured at a normal time. Finally, self-reported data are subject to recall and social desirability bias.

Conclusions and Public Health Implications

Our study reconfirmed the need to address the diversity of AANHPI adults in public health programs, including programs that address substance use and health-related disparities to address various needs across cultural subgroups and general AANHPI adults who experience mental health issues and discrimination. A high rate of substance use among some subgroups of the AANHPI adults has contributed to the disparity in their health outcomes, including lung and liver cancers. Also, differences in cultural norms, stigma, and language barriers could prevent AANHPI adults from accessing social support and treatment services [40]. Yet, AANHPI populations remained underrepresented in clinical research, including the development of substance use prevention and treatment programs [12,41]. Thus, it is essential to develop and implement culturally appropriate prevention and cessation programs to address the needs of AANHPI adults. These programs should engage communities in their development to ensure that intervention content, languages, and delivery methods are aligned with the cultural beliefs, values, norms, and preferences of specific AANHPI groups [41].

Acknowledgments

We are grateful for all the COVID-19 Effects on the Mental and Physical Health of Asian Americans & Pacific Islanders Survey Study (COMPASS) participants who completed the study. We would also like to acknowledge our community and academic partners in supporting the outreach and data collection. This study is conducted in collaboration with multiple academic institutions and community partners that serve diverse Asian American, Native Hawaiian, and Pacific Islander (AANHPI) communities: University of California, San Francisco (UCSF), University of California, Davis (UCD), International Children Assistance Network (ICAN), National Asian Pacific Center on Aging (NAPCA), Philip Jaisohn Memorial Foundation, Chinese Community Center, Hmong Community Center of Minnesota, Greater Boston Chinese Golden Age Center, United Sikhs, Memory and Aging Center at UCSF, EPATT, Yu-Ai Kai, India Community Center, and the American Samoa Community Cancer Coalition (ASCCC). The COMPASS research project and CARE Registry are supported by the National Institute on Aging of the National Institutes of Health (awards R24AG063718 and R01AG083926). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

VVD conceptualized the study, conducted the analysis, drafted the manuscript, and contributed to all subsequent drafts of the manuscript. NN, PML, MT, BN, MMD, and OLM interpreted results and reviewed and revised the manuscript. VMTP and JYT

XSL•FO
RenderX

obtained funding, conceptualized the study, planned and supervised the data collection, and reviewed and revised the manuscript. All authors have read and approved the final manuscript for submission.

Conflicts of Interest

None declared.

References

1. US Burden of Disease Collaborators, Mokdad AH, Ballestros K, et al. The state of US health, 1990-2016: burden of diseases, injuries, and risk factors among US States. JAMA 2018 Apr 10;319(14):1444-1472. [doi: 10.1001/jama.2018.0158] [Medline: 29634829]
2. Torre LA, Sauer AMG, Chen MS Jr, Kagawa-Singer M, Jemal A, Siegel RL. Cancer statistics for Asian Americans, Native Hawaiians, and Pacific Islanders, 2016: converging incidence in males and females. CA Cancer J Clin 2016 May;66(3):182-202. [doi: 10.3322/caac.21335] [Medline: 26766789]
3. Lee RJ, Madan RA, Kim J, Posadas EM, Yu EY. Disparities in cancer care and the Asian American Population. Oncologist 2021 Jun;26(6):453-460. [doi: 10.1002/onco.13748] [Medline: 33683795]
4. Bujarski S, Roche DJO, Lunny K, et al. The relationship between methamphetamine and alcohol use in a community sample of methamphetamine users. Drug Alcohol Depend 2014 Sep 1;142:127-132. [doi: 10.1016/j.drugalcdep.2014.06.004] [Medline: 24975477]
5. Crummy EA, O'Neal TJ, Baskin BM, Ferguson SM. One is not enough: understanding and modeling polysubstance use. Front Neurosci 2020;14:569. [doi: 10.3389/fnins.2020.00569] [Medline: 32612502]
6. Zhu DT, Park A. National trends in drug overdose mortality among Asian American, Native Hawaiian, and Pacific Islander populations. Ann Epidemiol 2025 Jan;101:36-41. [doi: 10.1016/j.annepidem.2024.12.005] [Medline: 39706382]
7. 2022 NSDUH annual national report. Substance Abuse and Mental Health Services Administration. 2023. URL: https://www.samhsa.gov/data/report/2022-nsduh-annual-national-report [accessed 2025-08-13]
8. Arrazola RA, Griffin T, Lunsford NB, et al. US cigarette smoking disparities by race and ethnicity - keep going and going!. Prev Chronic Dis 2023 Jun 1;20:E45. [doi: 10.5888/pcd20.220375] [Medline: 37262328]
9. Martins SS, Segura LE, Levy NS, et al. Racial and Ethnic differences in cannabis use following legalization in US states with medical cannabis laws. JAMA Netw Open 2021 Sep 1;4(9):e2127002. [doi: 10.1001/jamanetworkopen.2021.27002] [Medline: 34570205]
10. Chen MS, Hawks BL. A debunking of the myth of healthy Asian Americans and Pacific Islanders. Am J Health Promot 1995;9(4):261-268. [doi: 10.4278/0890-1171-9.4.261] [Medline: 10150729]
11. Tseng W, Kwon SC. Introduction: shining the light on Asian American, Native Hawaiian, and Pacific Islander health. J Health Care Poor Underserved 2015 May;26(2 Suppl):vii-xiv. [doi: 10.1353/hpu.2015.0052] [Medline: 25981100]
12. Đoàn LN, Takata Y, Sakuma KLK, Irvin VL. Trends in clinical research including Asian American, Native Hawaiian, and Pacific Islander participants funded by the US National Institutes of Health, 1992 to 2018. JAMA Netw Open 2019 Jul 3;2(7):e197432. [doi: 10.1001/jamanetworkopen.2019.7432] [Medline: 31339543]
13. 2020 census demographic profile. United States Census Bureau. 2023. URL: https://www.census.gov/data/tables/2023/dec/2020-census-demographic-profile.html [accessed 2025-08-13]
14. HHS implementation guidance on data collection standards for race, ethnicity, sex, primary language, and disability status. Office of the Assistant Secretary for Planning and Evaluation. 2011 Jul 17. URL: https://aspe.hhs.gov/reports/hhs-implementation-guidance-data-collection-standards-race-ethnicity-sex-primary-language-disability [accessed 2025-08-13]
15. Ponce N, Scheitler A, Shimkhada R. Understanding the culture of health for Asian American, Native Hawaiian and Pacific Islanders (AANHPIs): what do population-based health surveys across the nation tell us about the state of data disaggregation for AANHPIs? : Robert Wood Johnson Foundation; 2018.
16. Nguyen AB. Disaggregating Asian American and Native Hawaiian and Other Pacific Islander (AANHOPI) adult tobacco use: findings from Wave 1 of the Population Assessment of Tobacco and Health (PATH) Study, 2013-2014. J Racial Ethn Health Disparities 2019 Apr;6(2):356-363. [doi: 10.1007/s40615-018-00532-1] [Medline: 30610569]
17. Mukherjea A, Wackowski OA, Lee YO, Delnevo CD. Asian American, Native Hawaiian and Pacific Islander tobacco use patterns. Am J Health Behav 2014 May;38(3):362-369. [doi: 10.5993/AJHB.38.3.5] [Medline: 24636032]
18. Subica AM, Guerrero EG, Martin TKK, et al. Native Hawaiian/Pacific Islander alcohol, tobacco and other drug use, mental health and treatment need in the United States during COVID-19. Drug Alcohol Rev 2022 Nov;41(7):1653-1663. [doi: 10.1111/dar.13522] [Medline: 35953887]
19. Bersamira CS, Lin YA, Park K, Marsh JC. Drug use among Asian Americans: differentiating use by acculturation status and gender. J Subst Abuse Treat 2017 Aug;79:76-81. [doi: 10.1016/j.jsat.2017.06.002] [Medline: 28673532]
20. Ta Park VM, Dougan MM, Meyer OL, et al. Discrimination experiences during COVID-19 among a national, multi-lingual, community-based sample of Asian Americans and Pacific Islanders: COMPASS findings. Int J Environ Res Public Health 2022 Jan 14;19(2):35055744. [doi: 10.3390/ijerph19020924] [Medline: 35055744]
21. CARE registry. UCSF. 2024 Jul. URL: https://careregistry.ucsf.edu/ [accessed 2025-08-13]

XSL•FO
RenderX

22. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
23. Shariff-Marco S, Breen N, Landrine H, et al. Measuring everyday racial/ethnic discrimination in health surveys: how best to ask the questions, in one or two stages, across multiple racial/ethnic groups? Du Bois Rev 2011;8(1):159-177. [doi: 10.1017/S1742058X11000129] [Medline: 29354187]
24. Adzrago D, Walker TJ, Williams F. Reliability and validity of the Patient Health Questionnaire-4 scale and its subscales of depression and anxiety among US adults based on nativity. BMC Psychiatry 2024 Mar 18;24(1):213. [doi: 10.1186/s12888-024-05665-8] [Medline: 38500115]
25. Tobacco product use among adults - United States, 2022. US Centers for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/tobacco/media/pdfs/2024/09/cdc-osh-ncis-data-report-508.pdf [accessed 2025-08-13]
26. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66(1):7-30. [doi: 10.3322/caac.21332] [Medline: 26742998]
27. Chae DH, Takeuchi DT, Barbeau EM, Bennett GG, Lindsey J, Krieger N. Unfair treatment, racial/ethnic discrimination, ethnic identification, and smoking among Asian Americans in the National Latino and Asian American Study. Am J Public Health 2008 Mar;98(3):485-492. [doi: 10.2105/AJPH.2006.102012] [Medline: 18235073]
28. Yoo HC, Gee GC, Lowthrop CK, Robertson J. Self-reported racial discrimination and substance use among Asian Americans in Arizona. J Immigr Minor Health 2010 Oct;12(5):683-690. [doi: 10.1007/s10903-009-9306-z] [Medline: 20012204]
29. Lozano P, Peters A, Southworth A, et al. Smoking behavior among Asian Americans during the initial phase of the COVID-19 pandemic: The influence of pandemic stressors and depression. Tob Induc Dis 2024;22:22. [doi: 10.18332/tid/176923] [Medline: 38274001]
30. Bacong AM, Maglalang DD, Tsoh JY, Saw A. Perceived discrimination and coping with substance use among Asian Americans during the COVID-19 pandemic: a cross-sectional analysis. BMC Public Health 2025 Feb 20;25(1):698. [doi: 10.1186/s12889-025-21824-2] [Medline: 39979879]
31. Brondolo E, Ng W, Pierre KLJ, Lane R. Racism and mental health: examining the link between racism and depression from a social cognitive perspective, in the cost of racism for people of color: contextualizing experiences of discrimination.: American Psychological Association; 2016. URL: https://doi.org/10.1037/14852-006 [accessed 2025-08-13]
32. Fan Z, Gong X, Xu H, et al. Gender differences in the associations between tobacco smoke exposure and depressive symptoms among U.S. adults: NHANES 2007-2018. J Psychiatr Res 2022 Feb;146:249-257. [doi: 10.1016/j.jpsychires.2021.11.013] [Medline: 34799126]
33. Luk JW, Tsoh JY. Moderation of gender on smoking and depression in Chinese Americans. Addict Behav 2010 Nov;35(11):1040-1043. [doi: 10.1016/j.addbeh.2010.06.021] [Medline: 20655665]
34. Poulin C, Hand D, Boudreau B, Santor D. Gender differences in the association between substance use and elevated depressive symptoms in a general adolescent population. Addiction 2005 Apr;100(4):525-535. [doi: 10.1111/j.1360-0443.2005.01033.x] [Medline: 15784067]
35. Khaled SM, Bulloch A, Exner DV, Patten SB. Cigarette smoking, stages of change, and major depression in the Canadian population. Can J Psychiatry 2009 Mar;54(3):204-208. [doi: 10.1177/070674370905400309] [Medline: 19321025]
36. Korhonen T, Broms U, Varjonen J, et al. Smoking behaviour as a predictor of depression among Finnish men and women: a prospective cohort study of adult twins. Psychol Med 2007 May;37(5):705-715. [doi: 10.1017/S0033291706009639] [Medline: 17181913]
37. Montgomery L, Dixon S, Mantey DS. Racial and Ethnic Differences in Cannabis Use and Cannabis Use Disorder: Implications for Researchers. Curr Addict Rep 2022 Mar;9(1):14-22. [doi: 10.1007/s40429-021-00404-5] [Medline: 35251891]
38. 2022 NSDUH detailed tables. Substance Abuse and Mental Health Services Administration. 2023. URL: https://www.samhsa.gov/data/report/2022-nsduh-detailed-tables [accessed 2025-08-13]
39. Subica AM, An K, Okamoto SK. Asian American and Native Hawaiian/Pacific Islander substance use and disparities: review of current evidence and recommendations for the field. Curr Addict Rep 2024;11(3):501-510. [doi: 10.1007/s40429-024-00544-4]
40. Kramer EJ, Kwong K, Lee E, Chung H. Cultural factors influencing the mental health of Asian Americans. West J Med 2002 Sep;176(4):227-231. [Medline: 12208826]
41. Wen KY, Liang J, Diep D, Barta J, Juon HS. Smoking cessation interventions among Asian Americans: a scoping review. J Racial Ethn Health Disparities 2024 Oct;11(5):3015-3028. [doi: 10.1007/s40615-023-01760-w] [Medline: 37624538]

Abbreviations

AANHPI: Asian American, Native Hawaiian, and Pacific Islander
AOR: adjusted odds ratio
COMPASS: COVID-19 Effects on the Mental and Physical Health of Asian Americans & Pacific Islanders Survey Study
NHPI: Native Hawaiian and Pacific Islander
PHQ-4: Patient Health Questionnaire 4

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 23.04.25; peer-reviewed by DT Zhu, O Abasilim; revised version received 29.06.25; accepted 01.07.25; published 20.08.25.

Please cite as:
Do VV, Ta Park VM, Nguyen N, Ling PM, Tzuang M, Nam B, Dougan MM, Meyer OL, Tsoh JY
Use of Cigarettes, Cannabis, and Alcohol Among Asian American, Native Hawaiian, and Pacific Islander Adults: Community-Based National Survey Analysis
JMIR Public Health Surveill 2025;11:e76465
URL: https://publichealth.jmir.org/2025/1/e76465
doi:10.2196/76465

© Vuong Van Do, Van My Ta Park, Nhung Nguyen, Pamela May Ling, Marian Tzuang, Bora Nam, Marcelle M Dougan, Oanh L Meyer, Janice Y Tsoh. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between Social Distancing Compliance and Public Place Crowding During the COVID-19 Pandemic: Cross-Sectional Observational Study Using Computer Vision to Analyze Surveillance Footage

Lasse Suonperä Liebst[1,2], PhD; Wim Bernasco[2,3], PhD; Peter Ejbye-Ernst[2], PhD; Nigel van Herwijnen[4,5], MSc; Thomas van der Veen[2], MSc; Dennis Koelma[6], PhD; Cees G M Snoek[6], PhD; Marie Rosenkrantz Lindegaard[2,7], PhD

[1]Department of Sociology, University of Copenhagen, Copenhagen, Denmark

[2]Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam, The Netherlands

[3]Department of Human Geography, Planning and International Development Studies, University of Amsterdam, Amsterdam, The Netherlands

[4]Department of Social and Cultural Anthropology, Universitat Autònoma de Barcelona, Barcelona, Spain

[5]Department of Social and Political Sciences, Università degli Studi di Milano, Milan, Italy

[6]Informatics Institute, Faculty of Science, University of Amsterdam, Amsterdam, The Netherlands

[7]Department of Sociology, Faculty of Social and Behavioural Sciences, University of Amsterdam, Amsterdam, The Netherlands

Corresponding Author:
Marie Rosenkrantz Lindegaard, PhD
Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam, The Netherlands

Abstract

Background: Social distancing behavior has been a critical nonpharmaceutical measure for mitigating the COVID-19 pandemic. For this reason, there has been widespread interest in the factors determining social distancing violations, with a particular focus on individual-based factors.

Objective: In this paper, we examine an alternative and less appreciated indicator of social distancing violations: the situational opportunity for maintaining interpersonal distance in crowded settings. This focus on situational opportunities is borrowed from criminology, where it offers an alternative to individual-based explanations of crime and rule violations. We extend this approach to the COVID-19 pandemic context, suggesting its relevance in understanding distancing compliance behavior.

Methods: Our data comprise a large collection of video clips (n=56,429) from public places in Amsterdam, the Netherlands, captured by municipal surveillance cameras throughout the first year of the pandemic. We automatized the analysis of this footage using a computer vision algorithm designed for pedestrian detection and estimation of metric distances between individuals in the video still frames. This method allowed us to record social distancing violations of over half a million individuals (n=539,127) across more and less crowded street contexts.

Results: The data revealed a clear positive association between crowding and social distancing violations, evident both at the individual level and when aggregated per still frame. At the individual level, the analysis estimated that each additional 10 people present increased the likelihood of a distancing violation by 9 percentage points for a given pedestrian. At the aggregated level, there was an estimated increase of approximately 6 additional violations for every 10 additional individuals present, with a very large R^2 of 0.80. Additionally, a comparison with simulation data indicated that street spaces should, in principle, provide sufficient room for people to pass each other while maintaining a 1.5-meter distance. This suggests that pedestrians tend to gravitate toward others, even when ample space exists to maintain distance.

Conclusions: The direct positive relationship between crowding and distancing violations suggests that potential transmission encounters can be identified by simply counting the number of people present in a location. Our findings thus provide a reliable and scalable proxy measure of distancing noncompliance that offers epidemiologists a tool to easily incorporate real-life behavior into predictive models of airborne contagious diseases. Furthermore, our results highlight the need for scholars and public health agencies to consider the situational factors influencing social distancing violations, especially those related to crowding in public settings.

(JMIR Public Health Surveill 2025;11:e50929) doi:10.2196/50929

KEYWORDS

social distancing; compliance; crowding; urban public spaces; computer vision; surveillance footage; COVID-19 pandemic

XSL•FO
RenderX

Introduction

On the recommendation of the world's leading public health organizations, most countries around the globe installed and enforced social distancing measures to avoid close stranger contact and crowded settings during the COVID-19 pandemic. This is in line with evidence showing that social distancing is an effective strategy to limit the number of cases and deaths of COVID-19 [1]. However, the mitigating effect of these measures obviously depends on public compliance, which might be limited since people often do not adhere to health recommendations [2]. Consequently, public health agencies and researchers have closely monitored adherence to COVID-19 measures throughout the pandemic, tracking both compliance levels and influencing factors.

A primary research instrument for these evaluations has been repeated surveys [3]. This research approach has identified a number of individual-level compliance predictors, including risk perceptions, normative attitudes, and age [4]. This line of research has also identified situational correlates of social distancing compliance, relating to one's opportunity to violate [5]. For example, one survey-based study found that people who reported going outside for nonessential activities were more likely to violate social distancing and stay-at-home measures [6]. However, studies investigating such situational factors are relatively scarce, and therefore, called for in the literature [4]—plausibly reflecting the circumstance that self-reports are less suitable for measuring the situational dimension of people's everyday lives and COVID-19 health routines [7,8] (though it has been successfully done [9]). More broadly, this mirrors the issue that self-reports are known to offer a coarse-grained picture of how people behave in here-and-now situations [10].

This study addresses this research gap by applying a methodological approach that is increasingly recognized as the gold standard for microdetailed examinations of human behavior in everyday situations: video observation [11,12]. This method also addresses a recent call for unobtrusive, naturalistic observations of health behaviors, a technique that has been underused in health research [13]. Specifically, we leveraged a unique dataset of more than 60,000 hours of video of public place behavior captured by security cameras throughout the first year of the COVID-19 pandemic. To analyze the video footage, we used a computer vision algorithm, which we have developed and reliability-tested [14]. Computer vision algorithms are at the frontier of video data analysis research but remain rare in empirical studies of social and health behaviors [13,15].

This analysis, with a sample size exceeding half a million individuals, represents the most comprehensive observational study of social distancing behavior in public settings to date. It is also the first to use an automated, and thus, truly scalable approach. Building on our previous smaller-scale studies using manually coded video data [16,17], we hypothesize a positive association between crowding and noncompliance with social distancing guidelines (note that the video data for these 2 cited studies were also sampled in Amsterdam during the pandemic, although we stress that there is no direct data overlap between the current and these prior studies) [16,17]. As such, this study may be considered a large-scale replication of the prior studies, addressing the call for more replication studies in the social-behavioral sciences [18]. In addition, we mention that the current data and overall findings are cited in a separate paper outlining the technical aspects of the computer vision algorithm [14]. Specifically, we expect that as the number of people in a given space increases, so does the likelihood of close encounters [19]. This is particularly true in highly crowded areas where maintaining the recommended interpersonal distance becomes physically challenging [20]. Our focus on crowding as a situational factor in explaining distancing violations marks a shift from the literature's emphasis on individual-level predictors. This approach aligns with a "situational opportunity" perspective used in criminology to explain the occurrence of rule-breaking and crime [19]. This implies a focus on the opportunities people have to violate rules when engaging in their everyday routines and environments, irrespective of who they are.

In our previous 2 papers, we empirically established a positive correlation between crowd size and social distancing violations [16,17]. However, we did not examine the precise nature of this relationship in detail—whether the functional form of the slope is linear or follows a curved pattern. Moreover, a positive relationship between the number of people present and social distancing is neither obvious nor trivial, as observing the presence of others may change individuals' behavior [21]. A classical illustration of this is research on bystander intervention in emergencies, which shows that the presence of additional bystanders may reduce each bystander's willingness to intervene [22]. If this reduction is substantial, it could outweigh the safety-in-numbers dynamic where a larger crowd offers more opportunities for helping victims (though see Philpot et al [23]). Consequently, an increase in the number of bystanders may counterintuitively reduce the likelihood of any bystander intervening. Similarly, increasing the number of people in a place might not only make it physically more challenging to maintain distance from others but also promote greater vigilance and make individuals more focused on keeping their distance.

Methods

Study Design

Data were video footage of public behavior captured by 49 municipal surveillance cameras located in relatively busy outdoor public areas of Amsterdam, the Netherlands. Data access was provided by the Amsterdam police with permission from the Netherlands Public Prosecution Service (PaG/BJZ/49986).

The footage was recorded on Thursdays and Saturdays between 9 AM and 8 PM during the first year of the COVID-19 pandemic, from the beginning of March 2020 to June 2021. Throughout this period, the Dutch Government recommended that citizens kept a 1.5-meter distance in public spaces. Constrained by limited police resources to convert, store, and transport video recordings, we chose Thursdays and Saturdays to ensure that both the weekday and weekend behavior was included in the sample. The computer vision algorithm used for

XSL•FO
RenderX

the automatic data coding was developed from the convolutional neural network model for pedestrian detection by Hasan et al [24], and in turn, extended with a linear regression model to estimate the metric distance between pedestrians.

We used the algorithm to analyze still frames, in total 56,429, and selected each full hour for all cameras across the available footage. For each frame, we recorded the total number of people present (ie, the people crowding predictor). Note that we excluded all still frames with fewer than 2 persons, given that we are interested in how the presence of other people may influence the likelihood of social distancing violations—that is, if alone, one cannot be involved in any close encounters. Since we did not use facial recognition or other identification software, we do not know if the same individuals reappeared in the footage at different times and on different cameras. For each of the 539,127 persons observed, the computer vision algorithm further estimated the number of other persons within a 1.5-meter radius, which is a social distancing violation as defined in the Netherlands (ie, the outcome: 0=no social distancing violation; 1=at least 1 violation).

It should be emphasized that in the vast majority of situations, crowding levels did not make it practically impossible to avoid social distancing violations. On average, the viewsheds of the cameras captured around 650 m^2 of walkable street surface. While an area of this size could theoretically contain up to 433 individuals without 1.5-meter distancing violations, the median number of individuals observed in a still frame was 7 (SD 7.7), which would give each individual an average personal space of no less than 93 m^2. At this low level of crowding, individuals are not forced to violate social distancing rules for lack of physical space, even if some parts of the space-covered area are not available for pedestrians. It is, therefore, neither obvious that social distancing violations should take place, nor that increasing levels of crowding at the levels we observed should automatically give rise to increasing numbers of social distancing violations.

The reliability of the computer vision algorithm was assessed by comparing automated ratings of a subset of still frames against manual ratings by trained human coders, who had reference objects (eg, street tiles) available to evaluate the metric distance between those present. The human-computer interrater reliability assessment yielded excellent Krippendorff α scores larger than 0.8 (for further details, see Bernasco et al [14] and Krippendorff [25]). For practical reasons, this evaluation was conducted using situational counts of social distancing violations per still frame rather than individual-level records, which would have been preferred given that individuals are used as the unit of analysis in some of the current analyses. However, given the excellent interrater reliability score, we here assume that the corresponding individual-level score for social distancing would be at least acceptable. All statistical analyses were run in Stata 16 (StataCorp) and R (The R Foundation).

Ethical Considerations

Informed consent was neither feasible to obtain nor required under prevailing ethical guidelines, which exempt researchers from obtaining consent for naturalistic observation conducted in public settings (American Psychological Association's Ethical Principles, 2017, Section 8.03 [26]). No personally identifiable information was coded from the footage, and the resulting statistical dataset was fully anonymized. No compensation was offered. The study was approved by the Ethics Review Board of the Faculty of Social and Behavioral Sciences at the University of Amsterdam (2021-AISSR-14225).

Results

Almost half of the persons observed (mean 0.44, SD 0.50) were involved in at least 1 social distancing violation. This may be considered a very high incident rate, given that the close encounters were measured from single moments (1 still frame) rather than across several moments as people moved throughout space. The median of persons present in the still frames was 7 (SD 7.7), ranging between 2 and 67 persons.

Estimated with a linear probability model [27] specified with cluster-corrected standard errors (ie, persons nested in still frames), we found a positive association between people crowding and social distancing violations at a very conservative α level of .0005 (B=.009, 99.95% CI 0.009-0.010). This means that for each additional 10 people present, the likelihood of a pedestrian failing to comply with social distancing increases by 9 percentage points.

The above result remained robust across a number of alternative specifications, including adjustments for recording camera (B=.011, 99.95% CI 0.010-0.011) and noncompliance measured alternatively as close encounters with 2 or more persons (B=.006, 99.95% CI 0.005-0.006). The latter operationalization was meaningful because it plausibly excluded many close encounters with household members with whom the person was walking side-by-side [28], focusing instead on close encounters with strangers (ie, social distancing measures concern close encounters with strangers). Finally, we identified extreme values in our crowding measure as those exceeding the upper outer fence (Q3+3×IQR, or values above 60). After winsorizing these extremes (0.11% of data) by capping them at 60, the above result remained unaltered (B=.009, 99.95% CI 0.009-0.010).

Next, to evaluate the functional form, we visualized the association with binned scatterplots of the unwinsorized raw data [29]. As shown in Figure 1A, the analysis suggests that the association is curvilinear, with an initial steep slope that decreases with larger crowd counts. In other words, the individual probability of violating social distancing recommendations increases when more people are present, but this increase stagnates as the crowd grows larger. From this graph, it may also be established that people crowding has a substantial accumulating effect on noncompliance—in a situation where there are 30 individuals present, each individual has around 60% likelihood of transgressing the social distancing recommendations.

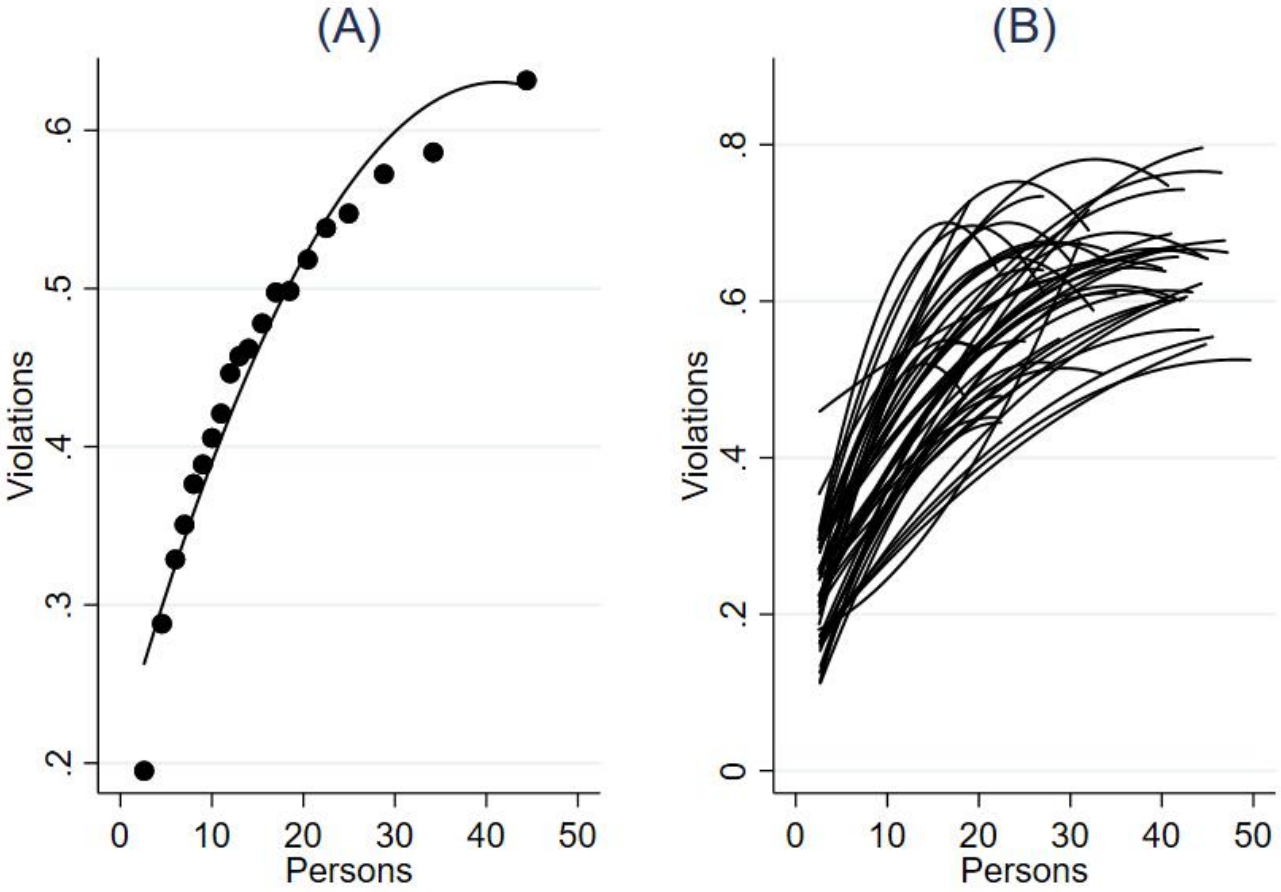
The size and layout of the video-captured streets vary among the cameras in the sample. For example, some cameras capture streets while others capture wide squares. These variations may influence the density of people present, and thus, in turn, the likelihood of violations (eg, if people are more densely crowded

XSL•FO
RenderX

in a small area, it may be more difficult to maintain distance). To account for this variation, we reran the analysis for each camera in the sample to assess whether the link between crowding and social distancing remained consistent across different cameras or settings. This visualization for each camera is presented in Figure 1B. This figure shows that although there is some variation between the cameras, the relationship between crowding and individual violation propensity was consistently positive and stagnant as group size increased across the 49 cameras in the sample.

It should be acknowledged that the above analysis plausibly overestimates the association due to spatial autocorrelation [30]. With 2 persons in a still frame, their potential violation acts are not just interdependent but are inherently linked since the physical distance between the 2 individuals is always the same for both. One solution is to use the still frames, rather than individuals, as the unit of analysis, with the outcome measured alternatively as an aggregate count of individual violations per still frame.

Figure 1. Association between people crowding and the individual-level likelihood of social distancing violations (A) across cameras and (B) per camera. Plots based on video-observational data from Amsterdam, March 2020 to June 2021 (n=539,127).



As we see in Figure 2A, it is noteworthy that the functional form was nearly linear with a slight tendency to increase curvilinearly—rather than decreasing curvilinearly at peak crowdedness (as in Figure 1). Regressing the number of violations on the number of persons present resulted in an unstandardized coefficient of 0.61 (99.5% CI 0.61-0.62), suggesting approximately 6 additional violations for every 10 additional persons present, with a very large R^2 of 0.80. The frequency of violations in a still frame thus grows at a steady positive rate as the number of people present increases. Although the individual and situational analyses suggest different functional forms, both indicated that the association was sizeable and positive in nature.

In Figure 2B, we again estimate the effect of people crowding on the number of social distancing violations within each of the cameras in the sample to see if the heterogeneity of the cameras might influence this relationship. While there was some variation between the cameras in the sample, the graph showed a consistent positive relationship between people crowding and the number of social distancing violations for each of the cameras.

One might argue that the increased risk of social distancing violations and crowding is trivial, as more people inherently raise the risk of close contact. To determine whether the findings are merely a result of this logical deduction or if a behavioral component is also at play, we simulated this logical baseline and compared it to our empirical results. To establish this baseline, we positioned fixed numbers of individuals randomly in a fictive rectangle the size of the average camera viewshed (650 m²), and counted the number of pairs that were within 1.5 meters distance of each other. This setup of random location assignment assumes that when positioning themselves in space, individuals are neither attracted nor repelled by others, and thus, do not have a specific preference for distance from or proximity

XSL•FO
RenderX

to others. To account for the possibility that not all 650 m² in the camera viewsheds could be easily occupied by pedestrians, we repeated the simulation with reduced areas of 300 and 150 m².

Figure 2. Association between people crowding and the situational-level frequency of social distancing violations. (A) Across cameras. (B) Per camera. Plots based on video-observational data from Amsterdam, March 2020 to June 2021 (n=56,429).

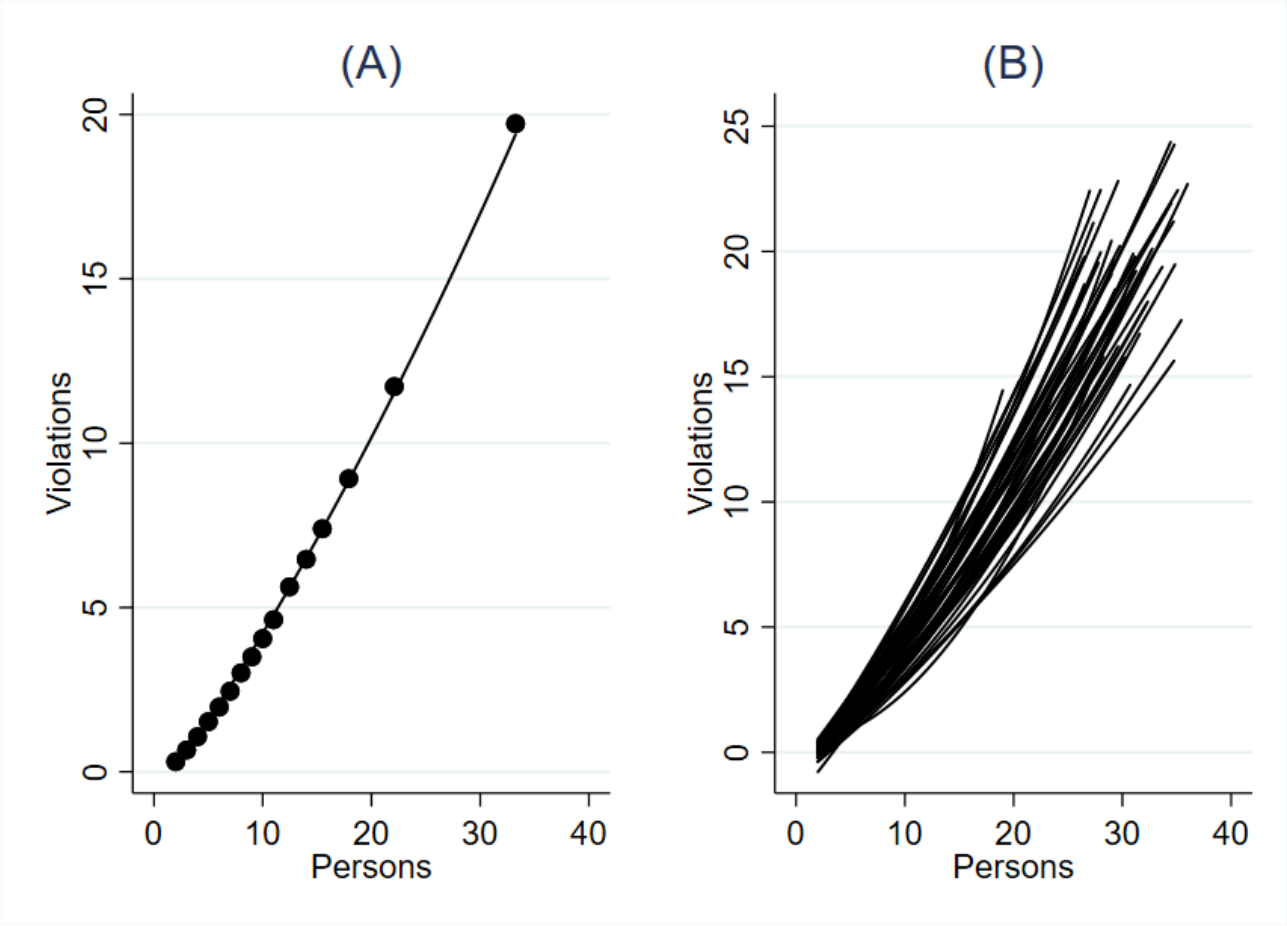
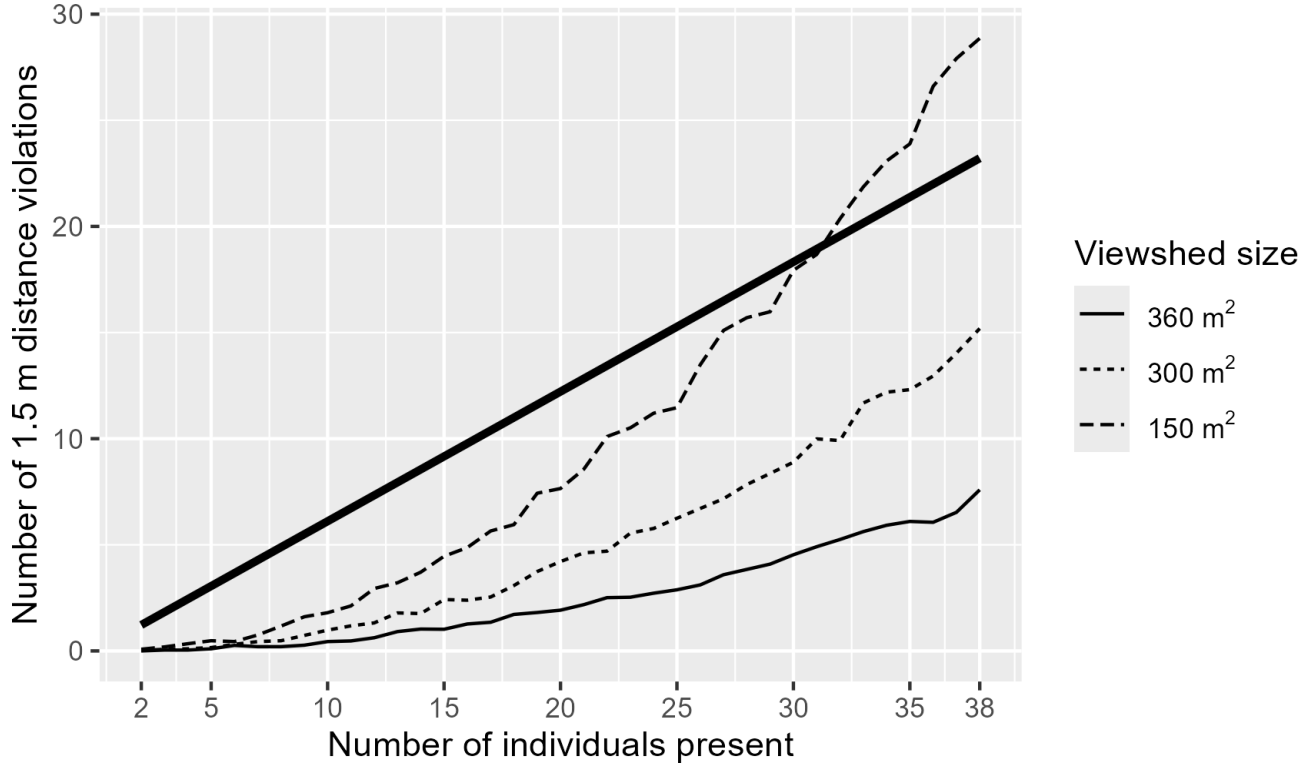


Figure 3 illustrates that under this design of random location assignment, the number of violations increases exponentially with the number of people present. It also shows that, over most of the observed range of crowding, the number of violations (bold straight line) is substantially larger than the number expected if navigational choices were random. The areas of the street should thus, in principle, offer enough space for it to be possible for people to be copresent without being too close. However, people seem to prefer to stay within 1.5 meters of others rather than keep a larger distance—a tendency that reverses only in very high-density settings, with large crowds concentrated in spaces smaller than the real-life average (such crowding levels are very rare in our data: levels above 25 persons occur in less than 5 percent of our still frames, and levels above 38 in less than 1 percent).

Our data does not shed light on the behavioral reason for this nonobvious tendency. However, one explanation could be that people have a tendency to gravitate toward the center of the street, walking in each other's footsteps and proximity. This could mirror cultural norms for appropriate pedestrian behavior [31] or suggest that the midstreet offers pedestrians an optimal visual overview, which is a critical determinant for pedestrian movement flows [32]. Additionally, the tendency for proximity can likely be attributed to people being with their household members [28], who were exempt from social distancing directives.

XSL•FO
RenderX

Figure 3. Observed (bold line) and simulated (3 thin lines) association between people crowding and frequency of social distancing violations.



Discussion

Principal Findings

Using an automated computer vision annotation algorithm, we found robust evidence for a positive association between people crowding and social distancing noncompliance. This result is in line with empirical studies on the link between interpersonal distance and crowding conducted in COVID-19-related [14,16] and other public settings [20].

Our results challenge the common focus on individual factors (like age, gender, and motivation) as primary explanations for distancing noncompliance. Instead, we highlight the importance of situational factors—thus addressing a call for research on situational, rather than individual-level, explanations of social distancing compliance [4]. Specifically, this study demonstrates that distancing noncompliance is influenced by the "situational opportunity" created by crowding [19]. The effect of this crowding rapidly accumulates to a magnitude that leaves little room for individual-level influences [33]. In other words, if you are in a crowded setting, chances are high that you will be involved in close encounters regardless of who you are and your willingness to distance. The situational opportunity perspective was originally developed within the field of criminology as an alternative to individual-centered explanations of crime [19], and our argument adds credence to the view that this theory generalizes to COVID-19-related rule violations (see also Liu et al [5]).

The findings of this study hold the potential to assist epidemiologists in incorporating real-life behavior into predictive models of airborne contagious diseases like COVID-19. The strong correlation between crowding and social distancing violations suggests that crowding may serve as a valid proxy measure for distancing behavior. Further, the number of people present is a straightforward measure that can easily be automated with computer vision across surveillance cameras. This is thus a flexible, low-cost, and scalable measure that epidemiologists can incorporate into prediction models of contagious diseases like COVID-19. While more advanced technology can estimate the distance between people in video footage, this study shows that simply counting the number of people present is a sufficient and reliable proxy measure of social distancing violations during the COVID-19 pandemic. Future research should explore whether this behavioral measure can help predict disease incidence effectively.

Our results also have potential practical implications for how policy makers may increase public adherence to social distancing measures. Public health information has centered around keeping interpersonal distance, while the other aspect of social distancing—to avoid gatherings and crowded places—has arguably received less attention. Our findings question that focus, given that social distancing violations are a direct function of crowding. Thus, at least in urban public settings, the safest way to optimize compliance seems to be avoidance of crowded places altogether, and such a recommendation could counter the misleading idea that one may maneuver through a crowded place while keeping interpersonal distance.

This argument could be taken one step further by simply recommending that policy makers focus less on individually targeted interventions—such as information campaigns—and more on situational interventions, such as crowd management strategies [34]. In tailoring such crowd management, the emphasis may not be on creating more street space for social distancing. Our analysis indicates that existing streets already, in principle, offer enough room for copresent pedestrians to

XSL•FO
RenderX

maintain the requested 1.5-meter distance (see Figure 3). In this view, the challenge seems to lie, in people's tendency to gravitate toward other people when moving through public space. Therefore, successful interventions aimed at increasing distance compliance could consider this crowd dynamic, instead of expanding physical space.

Limitations

It should be acknowledged as a limitation of the current analysis that it did not include any individual-level predictors—either because it was beyond the technological capacity of the computer vision tool developed to detect these properties (eg, age) or because it would have required us to survey the observed persons (eg, risk perceptions), which was not practically feasible. Therefore, we cannot directly compare the effect size of these against the situational effect of people crowding and examine how they may interact. We also did not measure whether individuals wore face masks or other kinds of protective equipment, which could potentially influence the likelihood of people complying with social distancing recommendations. For instance, the risk compensation theory posits that wearing masks might give a false sense of security, potentially causing people to relax their distancing efforts. However, it is important to note that this hypothesis lacks consistent behavioral evidence [17,35].

Further, we acknowledge that our results may not generalize to nonurban rural areas or private, semipublic, and indoor settings. For example, patterns of human presence in indoor places tend to be more socially programmed than outdoors [36], indicating that the crowding-encounter link may be less mechanically direct indoors vis-a-vis outdoors. More broadly, it is important to consider that interpersonal distance is shaped by cultural norms [37], highlighting potential constraints on the applicability of our findings across diverse cultural contexts.

Finally, it should be stressed that our analysis focuses on behavioral compliance without assuming that noncompliance with distancing leads to interpersonal virus transmission. In fact, the high prevalence of violations does not necessarily pose an epidemiological problem, given that the observed encounters are mostly brief, nonverbal, and occur outdoors, thus implying a relatively low risk for coronavirus transmission [38]. While this study introduces a readily available measure to study the spread of airborne contagious diseases, the intention of this paper is not to argue that our measure alone suffices to predict transmission rates. Depending on the disease under scrutiny, a number of other mitigating and aggravating variables should be accounted for in addition to the behavioral measures developed in this paper.

Acknowledgments

We thank the Amsterdam Police for facilitating the collection of the video data, in particular, Maikel van Scheppingen and Ronny van Axel Dongen. We thank Evelien Hoeben, Joska Appelman, Kiki Bijleveld, and Josephine Thomas for their work in collecting, organizing, and coding the video recordings. For manuscript proofreading, we used the generative artificial intelligence (AI) tools ChatGPT 4.0 and Claude 3.5 Sonnet. We confirm that no generative AI was used in content creation or data analysis. The study was supported by a grant provided by the Dutch Organization for Health Research and Innovation (ZonMw 50-56300-98-603).

Data Availability

The scripts and datasets analyzed during this study are available on OSF [39].

Authors' Contributions

LSL, WB, PEE, and MRL conceptualized and designed the study. MRL secured funding for the study with assistance from LSL, CGMS, and WB. DK developed the computer vision algorithm under the supervision of CGMS. The data analysis was conducted by LSL, WB, TV, and NH. The first draft was written by LSL with input from the other authors.

Conflicts of Interest

None declared.

References

1. Haug N, Geyrhofer L, Londei A, et al. Ranking the effectiveness of worldwide COVID-19 government interventions. Nat Hum Behav 2020 Dec;4(12):1303-1312. [doi: 10.1038/s41562-020-01009-0] [Medline: 33199859]
2. Hevey D. Adherence to health recommendations. In: Perk J, Gohlke H, Hellemans I, Sellier P, Mathes P, Monpère C, editors. Cardiovascular Prevention and Rehabilitation: Springer; 2007:293-300. [doi: 10.1007/978-1-84628-502-8_35]
3. Davies R, Mowbray F, Martin AF, Smith LE, Rubin GJ. A systematic review of observational methods used to quantify personal protective behaviours among members of the public during the COVID-19 pandemic, and the concordance between observational and self-report measures in infectious disease health protection. BMC Public Health 2022 Jul 28;22(1):1436. [doi: 10.1186/s12889-022-13819-0] [Medline: 35902818]
4. Kooistra EB, van Rooij B. Pandemic compliance: a systematic review of influences on social distancing behaviour during the first wave of the COVID-19 outbreak. SSRN Journal 2020. [doi: 10.2139/ssrn.3738047]
5. Liu N, Folmer CPR, Lo CH, Rooij B. Situational voluntary compliance. China Rev 2023;23:31-69 [FREE Full text]

XSL•FO
RenderX

6. van Rooij B, de Bruijn AL, Folmer CR, et al. Compliance with COVID-19 mitigation measures in the United States. SSRN J 2020. [doi: 10.2139/ssrn.3582626]
7. Hansen PG, Larsen EG, Gundersen CD. Reporting on one's behavior: a survey experiment on the nonvalidity of self-reported COVID-19 hygiene-relevant routine behaviors. Behav Public Policy 2022 Jan;6(1):34-51. [doi: 10.1017/bpp.2021.13]
8. Jerolmack C, Khan S. Talk is cheap: ethnography and the attitudinal fallacy. Sociol Methods Res 2014;43:178-209. [doi: 10.1177/0049124114523396]
9. Koher A, Jørgensen F, Petersen MB, Lehmann S. Monitoring public behavior during a pandemic using surveys: proof-of-concept via epidemic modelling. 2021 URL: https://raw.githubusercontent.com/mariefly/HOPE/master/Monitoring_Public_Behavior_During_a_Pandemic_Using_Surveys_Proof-of-concept_Via_Epidemic_Modelling_20210831.pdf [accessed 2025-04-03]
10. Deutscher I, Pestello FP, Pestello HFG. Sentiments and Acts: Transaction Publishers; 1993. [doi: 10.1515/9783112417560]
11. Gilmore RO, Adolph KE. Video can make behavioural science more reproducible. Nat Hum Behav 2017 Jul;1(7):1-2. [doi: 10.1038/s41562-017-0128] [Medline: 30775454]
12. Nassauer A, Legewie NM. Video data analysis: a methodological frame for a novel research trend. Sociol Methods Res 2021 Feb;50(1):135-174. [doi: 10.1177/0049124118769093]
13. Benton JS, French DP. Untapped potential of unobtrusive observation for studying health behaviors. JMIR Public Health Surveill 2024 Feb 21;10:e46638. [doi: 10.2196/46638] [Medline: 38381483]
14. Bernasco W, M. Hoeben E, Koelma D, et al. Promise into practice: application of computer vision in empirical research on social distancing. Sociol Methods Res 2023 Aug;52(3):1239-1287. [doi: 10.1177/00491241221099554]
15. Goldstein Y, Legewie NM, Shiffer-Sebba D. 3D social research: analysis of social interaction using computer vision. Sociol Methods Res 2023 Aug;52(3):1201-1238. [doi: 10.1177/00491241221147495]
16. Hoeben EM, Bernasco W, Liebst LS, van Baak C, Lindegaard MR. Social distancing compliance: a video observational analysis. PLOS One 2021;16(3):e0248221. [doi: 10.1371/journal.pone.0248221] [Medline: 33720951]
17. Liebst LS, Ejbye-Ernst P, de Bruin M, Thomas J, Lindegaard MR. No evidence that mask-wearing in public places elicits risk compensation behavior during the COVID-19 pandemic. Sci Rep 2022 Jan 27;12(1):1511. [doi: 10.1038/s41598-022-05270-3] [Medline: 35087100]
18. Munafò MR, Nosek BA, Bishop DVM, et al. A manifesto for reproducible science. Nat Hum Behav 2017 Jan 10;1(1):0021. [doi: 10.1038/s41562-016-0021] [Medline: 33954258]
19. Wilcox P, Cullen FT. Situational opportunity theories of crime. Annu Rev Criminol 2018 Jan 13;1(1):123-148. [doi: 10.1146/annurev-criminol-032317-092421]
20. Kaya N, Erkíp F. Invasion of personal space under the condition of short-term crowding: a case study on an automatic teller machine. J Environ Psychol 1999 Jun;19(2):183-189. [doi: 10.1006/jevp.1999.0125]
21. Latané B. The psychology of social impact. Am Psychol 1981;36(4):343-356. [doi: 10.1037//0003-066X.36.4.343]
22. Fischer P, Krueger JI, Greitemeyer T, et al. The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. Psychol Bull 2011 Jul;137(4):517-537. [doi: 10.1037/a0023304] [Medline: 21534650]
23. Philpot R, Liebst LS, Levine M, Bernasco W, Lindegaard MR. Would I be helped? Cross-national CCTV footage shows that intervention is the norm in public conflicts. Am Psychol 2020 Jan;75(1):66-75. [doi: 10.1037/amp0000469] [Medline: 31157529]
24. Hasan I, Liao S, Li J, Akram SU, Shao L. Generalizable pedestrian detection: the elephant in the room. Presented at: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); Jun 20-25, 2021; Nashville, TN, USA. [doi: 10.1109/CVPR46437.2021.01117]
25. Krippendorff K. Reliability in content analysis: some common misconceptions and recommendations. Hum Commun Res 2004;30:411-433. [doi: 10.1111/j.1468-2958.2004.tb00738.x]
26. Ethical principles of psychologists and code of conduct. American Psychological Association. 2017. URL: https://www.apa.org/ethics/code [accessed 2025-04-15]
27. Breen R, Karlson KB, Holm A. Interpreting and understanding logits, probits, and other nonlinear probability models. Annu Rev Sociol 2018 Jul 30;44(1):39-54. [doi: 10.1146/annurev-soc-073117-041429]
28. Liebst LS, Baggesen L, Dausel KL, Pallante V, Lindegaard MR. Human observers are accurate in judging personal relationships in real-life settings: a methodological tool for human observational research. Field Methods 2023 Nov;35(4):364-377. [doi: 10.1177/1525822X231158888]
29. Starr E, Goldfarb B. Binned scatterplots: a simple tool to make research easier and better. Strategic Manage J 2020 Dec;41(12):2261-2274. [doi: 10.1002/smj.3199]
30. Ward MD, Gleditsch KS. Spatial Regression Models: Sage Publications; 2018. [doi: 10.4135/9781071802588]
31. Jensen OB. Negotiation in motion: unpacking a geography of mobility. Space Cult 2010 Nov;13(4):389-402. [doi: 10.1177/1206331210374149]
32. Turner A, Penn A. Making isovists syntactic: isovist integration analysis. Presented at: 2nd International Symposium on Space Syntax; Mar 29 to Apr 2, 1999; Brazil URL: https://www.researchgate.net/profile/Alan-Penn-2/publication/

XSL•FO
RenderX

242075155_Making_isovists_syntactic_isovist_integration_analysis/links/0046352bbef92ce31e000000/Making-isovists-syntactic-isovist-integration-analysis.pdf [accessed 2025-04-03]
33. Funder DC, Ozer DJ. Evaluating effect size in psychological research: sense and nonsense. Adv Methods Pract Psychol Sci 2019 Jun;2(2):156-168. [doi: 10.1177/2515245919847202]
34. Johansson A, Batty M, Hayashi K, Al Bar O, Marcozzi D, Memish ZA. Crowd and environmental management during mass gatherings. Lancet Infect Dis 2012 Feb;12(2):150-156. [doi: 10.1016/S1473-3099(11)70287-0] [Medline: 22252150]
35. Mantzari E, Rubin GJ, Marteau TM. Is risk compensation threatening public health in the covid-19 pandemic? BMJ 2020 Jul 26;370:m2913. [doi: 10.1136/bmj.m2913] [Medline: 32713835]
36. Hanson J. Decoding Homes and Houses: Cambridge University Press; 2003.
37. Hall ET. The Hidden Dimension: Doubleday; 1966.
38. Bulfone TC, Malekinejad M, Rutherford GW, Razani N. Outdoor transmission of SARS-CoV-2 and other respiratory viruses: a systematic review. J Infect Dis 2021 Feb 24;223(4):550-561. [doi: 10.1093/infdis/jiaa742] [Medline: 33249484]
39. For review: social distancing violations as a situational correlate of crowding. OSF. URL: https://osf.io/5pc3h/?view_only=54ed1fcd074b4f7686567ed3d661fba3 [accessed 2025-04-12]

Edited by A Mavragani; submitted 17.07.23; peer-reviewed by CR Telles, G Seres; revised version received 28.02.25; accepted 04.03.25; published 17.04.25.

Please cite as:
Liebst LS, Bernasco W, Ejbye-Ernst P, van Herwijnen N, van der Veen T, Koelma D, Snoek CGM, Lindegaard MR
Association Between Social Distancing Compliance and Public Place Crowding During the COVID-19 Pandemic: Cross-Sectional Observational Study Using Computer Vision to Analyze Surveillance Footage
JMIR Public Health Surveill 2025;11:e50929
URL: https://publichealth.jmir.org/2025/1/e50929
doi:10.2196/50929

© Lasse Suonperä Liebst, Wim Bernasco, Peter Ejbye-Ernst, Nigel van Herwijnen, Thomas van der Veen, Dennis Koelma, Cees G M Snoek, Marie Rosenkrantz Lindegaard. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Diagnostic Testing Preferences in Rural and Vulnerable Populations During a Pandemic: Discrete Choice Experiment

Eline van den Broek-Altenburg[1*], MA, MSci, PhD; Jamie Benson[2*], BA; Yvonne Jonk[3], PhD; Abimbola Leslie[1], MD; Jan Carney[1,4], MD; Gary Stein[1,5], MD

[1]Department of Radiology, College of Medicine, University of Vermont, 89 Beaumont Avenue, Burlington, VT, United States

[2]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[3]Maine Rural Health Research Center, University of Southern Maine, Portland, OR, United States

[4]Department of Medicine, College of Medicine, University of Vermont, Burlington, VT, United States

[5]Department of Biochemistry, College of Medicine, University of Vermont, Burlington, VT, United States

[*]these authors contributed equally

Corresponding Author:
Eline van den Broek-Altenburg, MA, MSci, PhD
Department of Radiology, College of Medicine, University of Vermont, 89 Beaumont Avenue, Burlington, VT, United States

Abstract

Background: A particular challenge during the COVID-19 pandemic was to provide testing and treatment for already disadvantaged and vulnerable populations. Many states implemented testing in a sporadic and disorganized way, and it is unclear to what extent this disproportionally affected population experienced barriers to accessing care. It is also unclear whether potential barriers to testing were caused by systemic challenges, such as rurality, or by individuals' motivations for not getting tested.

Objective: The objective of this study was to understand the trade-offs individuals in rural and vulnerable populations make between attributes of COVID-19 testing and how these vary across individuals. The study was part of RADx-UP, a consortium of more than 125 projects studying COVID-19 testing patterns in communities across the United States.

Methods: First, we conducted 7 focus groups to identify barriers to COVID-19 testing and optimal strategies to increase testing. These barriers and strategies were then used to develop hypothetical choice scenarios in a discrete choice experiment. Data regarding preferences for testing were collected from an online panel (n=780) and oversampled in rural populations. We used quota sampling for age, gender, household income, and race: 50% of household incomes were above and below the median rural income of $52k per year 2023, and the maximum number of White, non-Hispanic respondents was 615. The data were analyzed using a conditional logit model (CL) and latent class analysis (LCA).

Results: We found that the attributes for testing locations were almost all significant and had the expected signs. As hypothesized, respondents were less likely to choose a test location that had a higher wait time (coefficient −0.183, SE 0.006); more travel time to get tested (coefficient −1.129, SE0.054); that was higher cost (coefficient −0.020, SE 0.000); where someone else would collect the sample (coefficient −0.230, SE 0.036); where it would take more time to receive results (coefficient −0.032, SE 0.006); and where the tests would cause more discomfort (coefficient −0.125, SE 0.007). They were more likely to choose a mail-order option (coefficient 0.494, SE 0.075) and options that had higher test accuracy (coefficient 0.026, SE 0.001). While respondents cared about these structural factors, these were not the primary drivers of choice for testing. Some important covariates were driving preferences, including age, gender, medical vulnerability, insurance status, trust in government organizations, and previous flu vaccination, which may be a proxy for compliance. These covariates helped explain the observed preference heterogeneity.

Conclusions: The results suggest that important social, behavioral, and policy factors affect choice for testing. Contrary to our hypotheses, rurality did not significantly impact preferences for testing; however, attitudes toward government and other beliefs did. Health care interventions intended to reduce rural health disparities that do not reflect the underlying values of individuals in those subpopulations are unlikely to be successful.

(JMIR Public Health Surveill 2025;11:e68734) doi:10.2196/68734

KEYWORDS

access to care; diagnostic testing; discrete choice experiment; individual and surveillance testing; pandemic; preferences; rural populations

XSL•FO
RenderX

Introduction

The COVID-19 pandemic created unprecedented challenges for health care and public health systems. Health care providers had to grapple with sudden changes in care delivery, including potential inpatient bed capacity constraints, delays in care, and the need to remotely manage medically and socially complex patients. Public health agencies had to quickly ramp up testing on an unparalleled scale; additionally, infrastructure and new processes to report these results to the public and health care providers were needed.

A particular challenge has been to provide COVID-19 testing and treatment for already disadvantaged and vulnerable populations [1-5], such as racial/ethnic disparities [6-11]. Media reports suggested socioeconomic and racial disparities in accessing COVID-19 testing. Many states implemented COVID-19 testing in a sporadic and disorganized way, partly because the Centers for Disease Control and Prevention guidelines changed several times over a short time span [12]. It is unclear to what extent disproportionately affected populations experienced barriers to accessing care. Even if historically underserved populations received fewer COVID-19 tests, the question remains whether this was due to systemic challenges or whether underlying individuals' motivations influenced testing decisions.

Previous research has illustrated that many factors affect the formation of new habitual behavior, such as seeking a COVID-19 test. Recent data from a discrete choice experiment (DCE) suggested that the attribute of the highest relative importance was test result turnaround time, followed by the type of test, specimen, and venue [13]. A DCE provides the opportunity to estimate pairwise choices and analyze marginal values or the total value of a health service or good. In simulations by Zimba et al, immediate or same-day test results, both polymerase chain reaction and serology, or oral specimens substantially increased testing uptake over the current typical testing option. The simulated uptake of a hypothetical testing scenario of polymerase chain reaction and serology via a saliva sample at a pharmacy with same-day results was 97.7% in this study, compared to 1.8% opting not to test. This study was performed in 2020 when at-home tests were not on the market yet. Thus, the study did not take preference heterogeneity into account.

It remains unclear what other factors encourage behavior change and incentivize people to seek testing. A review of approaches across different fields reveals several shortcomings in public health policy, most importantly that public health interventions do not take into account various psychological and behavioral factors [13]. Five blocks of factors have been identified with regard to the new (health) behavior in the literature: risk, attitudinal, normative, ability, and self-regulation factors [14]. This aligns with the health belief model, which postulates that a person's belief in a personal threat of an illness or disease, together with a person's belief in the effectiveness of the recommended health behavior or action, such as COVID-19 testing, predicts the likelihood the person will adopt the behavior [15-17].

The aim of this study was to understand the trade-offs individuals in rural and vulnerable populations make between attributes of COVID-19 tests and how these vary across individuals. The study was conducted as part of RADx-UP, a consortium of more than 125 research projects studying COVID-19 testing patterns in communities across the United States and its territories, as well as Tribal Nations. The goal of RADx-UP is to "speed innovation in the development, commercialization, and implementation of technologies for COVID-19 testing." The study was focused on identifying structural, social, behavioral, and policy factors that could be sources of COVID-19 testing disparities. Structural factors generally refer to broader conditions that either increase or decrease an individual's likelihood of getting tested. In this study, the primary structural factor is access to testing, defined in terms of travel time, which has been found to be a strong predictor of satisfaction with access to care [18]. In addition to travel time, costs and waiting times can also be structural factors [13].

Our main hypothesis was that preferences for attributes of COVID-19 testing vary between and within subpopulations, particularly among vulnerable rural populations. We hypothesized that individuals were less likely to get tested if they faced longer travel wait times, more discomfort with testing, higher costs, or longer wait times to hear of test results. We also hypothesize that rurality is an important factor affecting preferences for testing, influencing the travel time and distance to the nearest testing center.

Methods

Overview

For this study, we used a mixed methods approach. To understand which aspects or attributes of testing were important to people, we first conducted a series of focus groups throughout the Northern New England region to identify barriers to COVID-19 testing among medically and socially vulnerable rural adult populations. We then identified optimal strategies to increase testing using hypothetical scenarios fielded within a representative sample of the rural US population by developing a DCE [19,20].

Qualitative Research

We conducted 7 focus groups with 6-8 participants, each lasting approximately 45 - 60 minutes. Interviews and focus groups were held in person or virtually via Zoom. Informed consent was obtained before all sessions, which were audio- or video-recorded and transcribed by a third-party professional transcription service. We used the 4-level classification scheme (Urban, Large Rural, Small Rural, and Isolated Small Rural-Categorization A) for the Rural-Urban Commuting Area codes, which is a census tract–based classification that uses standard census measures of population density, levels of urbanization, and journey-to-work commuting to characterize all US census tracts with respect to their rural/urban status and commuting relationships to other census tracts [21].

Collaboration with community organizations supported the recruitment of rural residents for the focus groups. Participants

XSL•FO
RenderX

included older adults living in congregate housing and parents of school-aged children in rural communities. Focus group discussions explored participants' trusted sources of health information, their decision-making processes for COVID-19 testing, and facilitators and barriers they encountered in accessing testing. These discussions provided insights into how rural communities navigated the information landscape during the COVID-19 public health crisis. Participants in the focus groups were offered a $30 gift card for their participation. Focus group transcripts were coded independently using NVivo v14.23.0 software and analyzed systematically using both deductive and inductive thematic analysis techniques.

The focus group questions addressed topics such as reasons for seeking COVID-19 testing, testing barriers, feelings about getting tested, history of vaccination hesitancy, and social distancing behaviors during the pandemic. We also asked about sources of information that influenced their decision-making process. Focus group discussions also explored region-specific testing logistics, common barriers, swapping methods available, trusted sources of health information, public health campaigns to increase testing, and resources facilitating access to testing.

Experimental Design

The data from the focus groups identified distinct DCE terms and wording for conceptual attribute development [22]. Understanding how people made decisions about getting tested, and what sources of information informed their decisions, ensured the DCE accurately represented a realistic choice scenario. The unique contribution of a DCE is that it allows researchers to analyze the trade-offs that patients are willing to make, including options that may not exist but could in the future. DCEs are preferred to surveys because instead of simply asking "Would you get a COVID-19 test?", this approach asks: "You have the choice between two tests, A and B; they differ in the following ways […]. Which would you prefer?" The set of direct, discrete choice options systematically varies and facilitates identifying and prioritizing the set of attributes that decision-makers care most about (eg, travel time vs swab testing method). DCEs follow a general pattern of describing a COVID-19 scenario, followed by questions to elicit underlying preferences. In this study's DCE, testing is described by its attributes, and the options presented vary by the levels of those attributes. The DCE included 15 choice questions each, with 3 different testing options and an option not to test. The respondents were introduced to a hypothetical choice setting (see Multimedia Appendix 1) and asked to select which testing option they would prefer most.

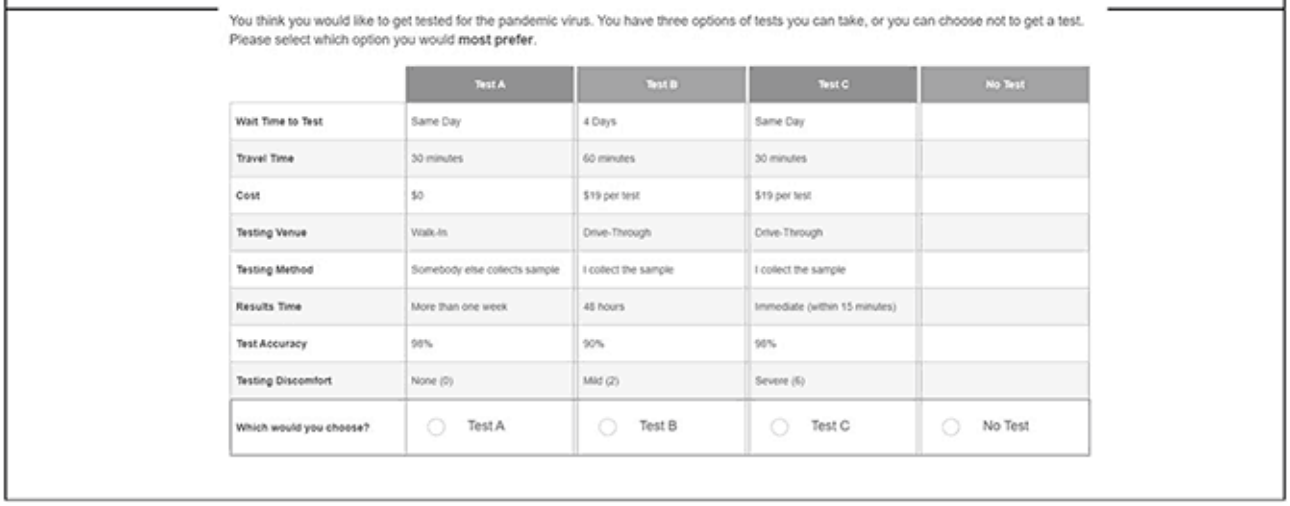
Based on the qualitative research, we included 8 testing attributes: cost, travel time, wait time to results, wait time to test, test accuracy, testing venue, testing methods, and testing discomfort (see Multimedia Appendix 1 for details). We used a DCE in which we collected demographic data, data on housing and employment, vaccination acceptance, testing attitudes, risk-taking behaviors (using the GRiPS score, a validated general risk propensity scale [23]), conspiratorial thinking and anti-expert sentiment [24], trust in public health authorities [25], political preference, and religion. The marginal values of the testing attributes were estimated based on analyzing the set of pairwise choices [26]. Thus, the DCE facilitated analyzing trade-offs people were willing to make, including testing options that may not currently exist in the respondents' local environments but could exist in the future.

The experimental design of the DCE was based on a-priori estimates of the values respondents would give for different attributes of the choice. These numbers were based on pilot data and estimates from the literature. This approach allowed us to avoid using the full factorial design of all the options for attribute levels, but an "efficient" design that would optimize our information with just 15 combinations of attributes and levels. Figure 1 shows the attributes of a choice set and an example of a DCE choice set. The attributes and levels that differed by testing option were explained to participants prior to answering the survey questions with the vignette in Figure 1.

XSL•FO
RenderX

Figure 1. Choice attributes and an example of a choice task.

Attribute	Scenario	Levels
"For the purpose of this study, imagine there is a new variant of the COVID-19 virus spreading, and you think you would like to get tested for it. Recently, many people have been testing using freely available "rapid test" kits, such as BinaxNOW, iHealth, or QuickVue. Though they have been free for most people, they are also available for purchase. In this scenario, you will be asked to choose between tests that may resemble the "rapid tests" you are familiar with, but they may be paid, or free, and they will not be labelled as "rapid tests." Instead, you will be shown different tests, with varying features, and are asked to choose among them. You will have three options of tests you can take, or you can choose not to get a test."		
1. Cost	"This is the amount of money you would need to pay out of pocket to get the test. When purchasing a rapid test kit, two tests will come in the box-- for all options you see, the price shown is for one test."	$0, $19, $120 which were based on reference levels for free tests; rapid test kits and PCR tests.
2. Travel Time	"This would be the maximum time you would need to get to the location for your test."	0 minutes (which meant they would not have to leave their home to get the test), 30 minutes, 60 minutes, 90 minutes (regardless of the type of transportation)
3. Wait Time to Results	"This means how long you would have to wait to find out if you have the disease or not, starting from when you got the test."	Immediate (within 15 minutes), same day, 48 hours, 5 days, more than one week.
4. Wait Time to Test	"Often, a testing appointment is not available right away, or if you have ordered a test online, it may take some time to ship. This means how long you will need to wait to take the test once you have ordered or scheduled it."	Same day, 2 days, 4 days, 8 days.
5. Test Accuracy	"No test is perfect, and sometimes when a test tells you that you don't have the virus, the test is wrong. This is not common, but it happens to some people, and can result in someone accidentally exposing others to the virus."	99% (meaning that out of 100 people who test negative for the disease using this test, one of them actually has the disease, and 99 do not), 90%, 82%.
6. Testing venue	"The location where the respondent would have to go to get the test."	Walk-In, Drive-Through, Mail-Order,Home Visit
7. Testing Method	"How the sample would be collected."	Somebody else collecting the sample or they would do it themselves.
8. Testing Discomfort	"Sometimes getting a test can be uncomfortable, for example a swab may be briefly put deep into your nose to collect a sample."	Discomfort was rated mild (2), moderate (4), severe (6).



Data

The 7 focus groups were conducted between November 2022 and March 2023. The DCE data were sampled from an online Centiment panel from April 29, 2023, through May 5, 2023. Centiment is a survey company that recruits individuals to answer surveys to generate rewards for themselves or to pledge their earnings to a nonprofit of their choice, and it is open to anyone to participate. Centiment has engineered complex systems to manage their respondents and ensure they are providing thoughtful responses. The standard invitation language for all members of the online Centiment panel is as follows: "We've got a survey for you. Reward: XX, Minutes: XX." Centiment does not provide any detail beyond the associated reward and average time to complete to avoid any selection bias. Respondents who meet the targeting criteria will have the

XSL•FO
RenderX

survey unlocked in their dashboard automatically and may come across it without an email. Initially, 1575 respondents initiated the survey but did not complete it. Out of 982 respondents who completed it, 184 were timed out before survey completion, leaving 798 completed surveys. Of the 798 respondents, 13 did not complete all the choice tasks, and 5 were likely robots based on our checks, leaving 780 completed responses. Of those 780 with good data, the mean survey duration was 28.25 minutes (median 20.52).

Centiment oversampled rural populations and used quota sampling for age, gender, household income, and race: 50% of household incomes were above and below the median rural income of $52k per year [27], and the maximum number of White, non-Hispanic respondents was 615. The study population included people who were medically or historically underserved: older adults in congregate housing, parents, and individuals living in Large Rural, Small Rural, and Isolated Small Rural areas nationally. The sample, therefore, included diverse and higher levels of social vulnerability and was nationally representative of rural populations. The online SurveyEngine platform was used to collect data (SurveyEngine GmbH, Berlin, Germany, 2023).

Analytic Approach

We analyzed the data using both a CL model and LCA. The CL, developed by Daniel MacFadden [28], models the expected utilities in terms of characteristics of the alternatives rather than attributes of the individuals. The data from our repeated choice tasks can be treated as panel data in a conditional fixed effects logit model [29]. With this model, we used the respondents as their own controls and controlled for “stable” characteristics that do not change over time even if these were not measured. It is similar to an experiment with random assignment, controlling for omitted variables. With this model, we focused on the estimation of within-individual differences in preferences.

The latent class (LC) model is based on the assumption that individual preferences have a discrete distribution [30,31]. In the latter, we used a weighted probability of class membership and multiplied that with the probability of choosing a particular choice option. LC accounts for serial correlation, which means we model the probability of observing a sequence of choices. This way, we were able to predict the probability of the sequence of choices respondents made.

LC models are based on maximum likelihood estimation, adding information about the preference heterogeneity within our study population. For example, the LC model does not assume that everyone has the same preference for the no-choice option; rather, we analyzed whether some people were more likely to choose a no-choice option. Subgroup analysis can be used to identify heterogeneous treatment effects. LCA structure uses a class allocation model where probabilities vary across individual decision-makers as a function of their observed characteristics. Crucially, we analyzed whether membership in a subgroup may differ by health disparities, socioeconomic status, and different beliefs and attitudes by entering vulnerability, rural location, comorbidities, income, race, and other indicator variables in the class membership function.

We used the consistent Akaike information criterion (CAIC) and Bayesian information criterion (BIC), both statistical measures of fit, to determine the number of classes for the LCA [32]. Often the CAIC and BIC are used to determine the optimal number of classes, but with LC, we cannot directly compare the likelihood functions, as models with more parameters will generally produce better results, even though the model is not statistically considered better [33]. We used Stata 18.0 for the statistical analysis, which uses the expectation-maximization (EM) algorithm to fit a discrete-mixture logit model (StataCorp, College Station, TX, 2024). The EM algorithm is an iterative method to find (local) maximum likelihood or maximum a posteriori estimates of parameters in statistical models where the model depends on unobserved latent variables.

Ethical Considerations

This study (STUDY00002116) has been reviewed by the Institutional Review Board of the University of Vermont and the Institutional Official has authorized the publication of the research data.. Focus group protocols were reviewed by the IRB using the exempt procedures set forth under 45 CFR 46.104, specifically, under Exemption Category: (2)(ii) tests, surveys, interviews, or observation (low risk), with waiver of documentation of consent under 46.117(c)(1). The Discrete Choice Experiment (DCE) was originally deemed not human subjects research by the IRB’s assessment tool; a quality review later determined that the study should have been reviewed and would likely have been approved with a waiver of written consent. In the DCE, survey data were collected without identifiers. Respondents received $5 compensation for their time from the survey company Centiment. Additionally, no identification of individual participants in any images of the manuscript or supplementary material is possible.

Results

Sample Characteristics

As described, data were gathered using a quota sampling approach for age, gender, household income, and race. We compared the sample characteristics of those who had previous COVID-19 tests to those who did not to see what the predicted uptake of testing would be, based on differences in attitudes and beliefs. Some studies have shown that previous choices may affect current choices, both in general terms [34] and in health choices specifically [35]. Within the descriptive analyses, chi-square tests were used to detect statistically significant differences between the two groups. Table 1 shows that those who had a previous COVID-19 test had a significantly higher GRiPS score and were somewhat more leaning toward politically liberal.

XSL•FO
RenderX

Table . Attitudes of respondents toward government and the public health system[a].

Attitude	Previous COVID-19 test	
	No (n=206)	Yes (n=546)
Conspiratorial thinking score, mean (SD)	16.2 (6.3)	16.3 (6.3)
GRiPS risk score[b], mean (SD)	17.1 (7.8)	17.7 (7.8)
Political leaning: conservative (1) to liberal (7)[b], mean (SD)	3.2 (1.9)	3.7 (2.0)
Trust (0% - 100%), mean (SD)		
State's government[b]	39.3 (32.4)	45.3 (30.9)
City/town government[b]	41.9 (32.6)	47.4 (30.4)
Police[b]	45.7 (34.0)	49.3 (33.7)
State health department[b]	49.0 (33.8)	56.8 (31.0)
Public health experts (CDC)[b]	48.3 (37.1)	55.8 (35.0)
Red Cross[b]	50.4 (34.6)	55.2 (33.1)
Health care system[b]	58.6 (32.8)	65.7 (30.7)
WHO[b]	41.4 (36.3)	48.7 (35.5)
Federal government[b]	31.5 (33.6)	37.9 (32.9)
Scientific researchers[b]	50.3 (35.1)	54.4 (34.4)
Governmental effort[b]	33.1 (31.9)	39.2 (32.1)
Religiosity, n (%)		
Atheistic/agnostic[b]	12 (5.8)	53 (9.7)
Organized religion[b]	102 (49.5)	282 (51.6)
Spiritual/non-organized/other[b]	92 (44.7)	211 (38.6)

[a]Data were collected from a discrete choice experiment eliciting preferences for COVID-19 testing from rural and vulnerable populations sampled from an online Centiment panel from April 29, 2023, through May 5, 2023.

[b]Significant difference between groups at P<.05.

Respondents who had a previous COVID-19 test also had more trust in public health organizations and governmental institutions, were more religious, and less spiritual than respondents who did not have a previous COVID-19 test. We did not find a statistical difference for conspiratorial thinking and anti-expert sentiment scores between people who had already gone for testing and those who did not.

Conditional Logit

The results of the choice models can be found in Table 2. In the CL model, the dependent variable was "choice for test location," defined by the attributes in the choice task. Each coefficient shows the probability of picking a particular option versus not picking that option. We found that the attributes were almost all significant and had the expected signs. As hypothesized, respondents were less likely to choose a test location that had a higher wait time (coefficient −0.183, SE 0.006); more travel time to get tested (−1.129, 0.054); that was more costly (−0.020, 0.000); where someone else would collect the sample (−0.230, 0.036); where it would take more time to receive results (−0.032, 0.006); and where tests would have more discomfort (−0.125, 0.007). They were more likely to choose a mail-order option (0.494, 0.075) and options that had higher test accuracy (0.026, 0.001).

XSL•FO
RenderX

Table . Results of the conditional logit model and latent class analysis—preferences for attributes of testing locations[a].

	Conditional logit model β[b] logit modelβ	Latent Class 1 "Compliant"	Latent Class 2 "Evaluative/less-compliant"	Latent Class 3 "Convenience"
Class share (%)	__[c]	0.434	0.159	0.407
Wait time to test (d)	−0.183[d] (0.006)	−0.112[d] (0.008)	−0.114[d] (0.043)	−0.198[d] (0.012)
Travel time to test (h)	−1.129[d] (0.054)	−0.452[d] (0.069)	−1.448[d] (0.442)	−1.965[d] (0.127)
Test cost ($)	−0.020[d] (0.000)	−0.015[d] (0.001)	−0.011[d] (0.003)	−0.034[d] (0.001)
Venue (ref: drive-through)				
Venue: walk-in	0.043 (0.040)	−0.079 (0.052)	0.205 (0.305)	−0.213[d] (0.069)
Venue: home visit	0.061 (0.069)	−0.175(0.131)	0.373 (0.457)	−0.536[d] (0.159)
Venue: mail order	0.494[d]((0.075)	0.179[e] (0.108)	0.871[f]((0.443)	−0.567[d] (0.134)
Test method	−0.230[d] (0.036)	−0.275[d] (0.047)	0.152 (0.254)	−0.130[f] (0.061)
Time to results (d)	−0.032***[d] (0.006)	−0.021***[d] (0.008)	−0.030 (0.037)	−0.036[d] (0.011)
Test accuracy (%)	0.026[d] (0.001)	0.044[d] (0.002)	0.024[d] (0.005)	0.028[d] (0.001)
Test discomfort	−0.125[d] (0.007)	−0.030[d] (0.009)	−0.131[d] (0.046)	−0.163[d] (0.021)

[a]Data were collected from a discrete choice experiment eliciting preferences for COVID-19 testing from rural and vulnerable populations sampled from an online Centiment panel from April 29, 2023, through May 5, 2023.

[b]The results show the coefficients (β) for the different attributes of the choice; standard errors in parenthesis.

[c]not applicable.

[d]*P*<.01.

[e]*P*<.10.

[f]*P*<.05.

Latent Class Analysis

After identifying preferences for attributes of testing for the sample population, we addressed the issue of unobserved preferences of respondents by probabilistically segmenting the sample population into different groups or "classes" based on a latent variable. Class membership was first defined by a membership function including the indicator variables, after which the utility functions of different classes were estimated through maximum likelihood estimation. First, we determined the number of LC. We considered theoretical interpretability and compared the statistical tests of model fit using models for 1-5 possible LC.

There are several approaches to choosing the number of classes in LCA. When looking at the information criteria and the likelihood ratio test, there was a slightly better model fit for the 5-class model. Information criteria start with a computation of fit—the log-likelihood and then penalize this based on the number of classes. Information criteria commonly applied to the selection of number of classes include the BIC and the CAIC. However, other methods that are frequently used include replicability and domain-usefulness. The number of classes that get the most consistent results is considered to be the best. This approach can also be viewed as a form of cross-validation.

We combined this approach with domain usefulness, which means we choose the number of classes that create the solution that is most interesting from an interpretation perspective. This approach is often used when the difference in information criteria is small. The difference in model fit between the 3-, 4-, and 5-class models was not large, and the interpretability of the 3-class model is more straightforward. We therefore report the results of the 3-class model in Table 2.

The covariates we used for the class membership model included gender (female), education, income, age, race (White), currently insured, currently employed, self-assessed health rating, vulnerability, the number of previous tests, rurality, flu vaccination history, risk-taking score (GRiPS), conspiratorial thinking and anti-expert sentiment scales, trust in public health and government organizations, self-rated religiosity, and whether they were politically Republican-leaning.

Based on coefficients for the attributes of the choice for test location, three groups can be identified: testing method does not matter/drive-through locations are preferred, strong preference for at-home test and mail-order options, and prefers self-administering the test. We call the first class the "Compliant" class, since these respondents are more likely to get tested, but they do not care about the type of venue as long as they can take the test themselves. We call Class 2 the "Evaluative/less compliant" since they seem to care most about the ease of testing, not requiring time or travel, and do not care about the test method or how long it takes to get results back. The results implied that members of Class 2 would get tested because they had to or to be compliant, but not because they thought it was important. This suggests they only go for testing

XSL•FO
RenderX

if they have to. Members of Class 3 want to drive somewhere for testing, but close to their homes. We call this class "Convenience."

We found that age (−0.709, *P*<.01), being insured (1.742, *P*<.01), previous flu vaccination (0.580, *P*=.05), and trust in various institutions (0.510, *P*=.02) all significantly predicted membership of Class 1. For example, with every additional point increase on the Likert scale for trust, respondents would be more likely to be in Class 1 versus Class 3 which was the reference. We also found that gender (−2.610, *P*=.06), vulnerability (0.803, *P*=.030), previous flu vaccination (−0.921, *P*<.01), and trust in government institutions (−0.687, *P*=.04) also explained membership of Class 2. For example, females were far less likely than men to be in Class 2 compared to Class 3. Previous flu vaccination, which could be considered a proxy for "compliance," defined Class 1 membership, but those who were vaccinated were less likely to be in Class 2. Where higher levels of trust in government institutions predicted Class 1 membership, these respondents were less likely to be in Class 2.

We found that 43.4% of the respondents in our study fall in Class 1, which is the group that cares significantly about being able to perform the test themselves. The coefficient (β=−.275) is negative and twice as large as for the other two groups, meaning that they would be less likely to pick a test location where someone else would collect the sample.

Class 2 includes 15.9% of respondents, and they seem to care the most about convenience, as they have a significant and positive coefficient for mail-order (β=.871). This seems to be more about not wanting to travel, since they do not feel strongly about who administers the test (β=.152). Indeed, the coefficient for travel time to the test center is large (β=−1.448) and negative, suggesting that with every additional hour of drive time, respondents are less likely to pick that testing option.

Class 3 has a strong preference for drive-through testing options. All the other venue options have a negative sign, meaning they are less likely to pick an option that involves walk-in (−0.213), home visit (−0.536), and mail-order (−0.567). The coefficient for travel time to test is large and negative (−1.965), meaning that with every 1-hour increase of travel time, respondents are disproportionately less likely to pick that testing option.

Marginal Rates of Substitution

Table 3 shows an analysis of the trade-offs respondents were willing to make, known as the marginal rates of substitution. First, we looked at their willingness to pay for different attributes of testing. Members of all three classes were willing to pay significantly to decrease travel time. Class 1 members would pay $30 more to decrease travel time within 1 hour; this was $131 for members in Class 2 and $57 in Class 3. Members of the different classes were willing to pay for different venues but varied widely. For example, members of Class 2 were willing to pay $18 more for a walk-in testing option than drive-through, where they would be willing to pay $33 for an at-home testing option and $79 for mail order. Members of Class 1 would only pay more ($11) for a mail-order option, while members of Class 3 would not pay for any of these options when compared to drive-through. Members of Class 2 would also pay $11 more to receive a test with less discomfort, where this was only $2 for Class 1 and $4 for Class 3.

Table . Marginal rates of substitution when choosing for different diagnostic testing locations.

	WTP[a] Class 1	WTP Class 2	WTP Class 3	Accuracy Class 1	Accuracy Class 2	Accuracy Class 3
Wait time to appointment	−7.445	−10.367	−5.822	2.530	−4.703	7.123
Travel time	−30.065	−131.631	−57.648	10.217	−59.715	70.537
Cost	—[b]	—	—	0.340	−0.454	1.224
Venue_walk	−5.251	18.655	−6.250	1.785	8.463	7.647
Venue_home	−11.639	33.933	−15.733	3.955	15.394	19.250
Venue_mail	11.913	79.194	−16.631	−4.049	35.927	20.350
Method	−18.320	13.845	−3.826	6.226	6.281	4.682
Time to results	1.428	−2.738	−1.048	−0.485	−1.242	1.282
Accuracy	2.943	−2.204	0.817	—	—	—
Discomfort	−2.014	−11.873	−4.788	0.685	−5.386	5.858

[a]WTP: willingness to pay.

[b]not applicable.

The table shows the marginal rates of substitution for costs (willingness to pay) and for the accuracy of the test for the 4 LC. For example, respondents in Class 1 were willing to pay $30 less if they had to travel 10 minutes more to a testing center.

Data were collected from a DCE eliciting preferences for COVID-19 testing from rural and vulnerable populations sampled from an online Centiment panel from April 29, 2023, through May 5, 2023.

In terms of test accuracy and results, only members of Classes 1 and 3 appeared to be willing to trade off wait time, travel time, and even discomfort to get a test that was more accurate. They were not willing to pay significantly more for a more accurate

XSL•FO
RenderX

test, however. Members of Class 2 did not care about the accuracy of the test at all. This was consistent with findings for time to results: members of Class 2 would not trade off travel time, wait time, cost, or discomfort to get results more quickly. This also seems consistent with the covariates predicting class membership: these were respondents who had lower trust in government institutions, and fewer of them had previous flu vaccination.

Discussion

Principal Findings and Comparison With Previous Works

This study used data from a DCE to understand what trade-offs individuals in rural and vulnerable populations make between attributes of COVID-19, and how these vary by different individuals. We found that the attributes for testing locations were almost all significant and had the expected signs. As hypothesized, respondents were less likely to choose a test location that had a higher wait time (coefficient −0.183, SE 0.006); more travel time to get tested (−1.129, 0.054); that was more costly (−0.020, 0.000); where someone else would collect the sample (−0.230, 0.036); where it would take more time to receive results (−0.032, 0.006); and where tests would have more discomfort (−0.125, 0.007). They were more likely to choose a mail-order option (0.494, 0.075) and options that had higher test accuracy (0.026, 0.001). While we found that respondents cared about these structural factors, these were not the primary drivers of choice for testing. Some important covariates were driving preferences, including age, gender, medical vulnerability, insurance status, trust in government organizations, and previous flu vaccination—which may be a proxy for compliance. These covariates helped explain the observed preference heterogeneity.

It is essential to consider our findings regarding testing in the broader context of prior work on access to care, pandemic preparedness, strength of the public health workforce, and unique social and cultural factors potentially impacting health decisions. Several areas of public health practice and preparedness for future pandemics relate to our findings. Rurality in our study did not significantly impact testing preferences; however, the complexities and nuances in rural health differences are well documented [36-40]. Pandemic workforce stress resulted in an “exodus” of public health workers [38], leaving many areas of the United States vulnerable and unprepared for future pandemics. Additional literature further quantified workforce needs to stabilize state and local public health workforces to meet essential needs [39], including the necessity for “surge” capacity [40].

Published data from the Pew Research Center further demonstrate differences in urban, suburban, and rural residents’ political views on social issues, including health [41]. More recent studies highlight the variability of rural geographic settings, distinct rural classes, and the use of community capital to further define differences among rural communities [42].

A 2023 scoping review of COVID-19 preparedness in rural and remote areas showed many differences across broad social health determinants, including local economies and less health system preparation [43]. In contrast, local multisector collaborations, evidence of engaged local cultures, and the strength of local leaders were positive factors [43]. Similarly, the role of primary health care, adequacy of the public health local workforce, and strengths of local collaborations likely all play important roles in pandemic responses but may vary across rural geographic areas in the United States and be challenging to measure [44-46].

Other authors emphasize the importance of culture and history in rural communities, including cultural assets, social cohesion, “community spirit,” and the strengths of community leadership and engagement [47]. Strategies to optimize pandemic preparedness for the future (in addition to adequate health care and public health workforce) must incorporate familiarity with, engagement with, and respect for these factors to fully leverage optimal responses to public health threats, including, but not limited to, COVID-19. Given the complexity of local community culture, it is not surprising that it is challenging to systematically increase access to testing: our study helps identify actionable priorities to help guide future public health interventions.

While the findings of this study add significantly to the existing literature regarding access to diagnostic testing during the pandemic, they are not without limitations. The study was limited by its relatively small sample size and snowball sampling technique. The COVID-19 pandemic is an evolving crisis, and this research was only focused on a particular period of time which we studied at a later period of time. Nevertheless, many study participants said they clearly remembered their considerations at the time, since it was relatively recent. The hypothetical setting of the DCE and the online panel makes it challenging to make out-of-sample predictions, which is why our conclusions cannot be generalized to potential future pandemic scenarios.

It is important to note from the results of the LCA that decision-makers looking to optimize testing strategies, while incorporating patient preferences for attributes of testing, should differentiate between the different subpopulations they are serving. This requires careful consideration of individual characteristics as well as preferences. Overall, we found in this study that subpopulations are characterized by being either primarily compliant or evaluative to some extent non-compliant, or they were driven solely by convenience. We did not find an association between rurality and testing preferences, which means that rurality was not associated with the probability of belonging to a specific class. It does not mean such an association does not exist. An alternative to investigate such interaction is with a mixed logit model, which allows for individual intercepts and more flexible substitution patterns. However, the sample size was not significantly large to analyze elaborate interaction models. This would be important to further analyze in future studies.

Conclusions

This suggests that important social, behavioral, and even policy factors affect choice for testing. Contrary to our hypotheses, rurality did not significantly impact preferences for testing. This study provides a clear message to public health and surveillance systems seeking to increase testing rates during a pandemic like

XSL•FO
RenderX

COVID-19. We conclude that health care interventions intended to reduce rural health disparities that do not reflect the underlying values of individuals in those subpopulations are unlikely to be successful.

Another concrete public health implication from this study is that adding a consistent and "easy" mail-order testing option in a future pandemic may significantly increase testing rates. Both the total sample population in the CL model and the different classes identified in the LCA, while having different underlying tastes and unobserved preferences, have a significant, positive, and relatively large effect size for adding this hypothetical mail-order option. Members of two or three classes, adding up to 84% of all respondents, also had a strong preference for self-administering the test. Compared to other attributes of testing, such as travel time or wait time, mail-order and self-administering are significantly more important aspects of testing for most people.

Acknowledgments

We are grateful to Drs. Kimberley Fox, Carolyn Gray, Elizabeth Woods, and Mercedes Avila for their feedback on the DCE development and qualitative research efforts. Research reported in this RADx® Underserved Populations (RADx-UP) publication was supported by the National Institutes of Health under award number U01 CA271329. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to restrictions in data sharing from the National Institutes of Health and the RADx-UP but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Hypothetical choice setting and choice attributes.
[DOCX File, 23 KB - publichealth_v11i1e68734_app1.docx]

References

1. Winterbauer E, Levy PD, Calhoun D, et al. Qualitative review of promising practices for testing vulnerable populations at off-site COVID-19 testing centers. Healthc (Amst) 2021 Mar;9(1):100519. [doi: 10.1016/j.hjdsi.2021.100519] [Medline: 33454583]
2. Lieberman-Cribbin W, Tuminello S, Flores RM, Taioli E. Disparities in COVID-19 testing and positivity in New York City. Am J Prev Med 2020 Sep;59(3):326-332. [doi: 10.1016/j.amepre.2020.06.005] [Medline: 32703702]
3. Oates GR, Juarez LD, Horswell R, et al. The association between neighborhood social vulnerability and COVID-19 testing, positivity, and incidence in Alabama and Louisiana. J Community Health 2021 Dec;46(6):1115-1123. [doi: 10.1007/s10900-021-00998-x] [Medline: 33966116]
4. Morrone A, Buonomini AR, Sannella A, Pimpinelli F, Rotulo A. Unequal access to testing and vaccination services for the homeless and undocumented population during COVID-19 pandemic. Int J Public Health 2022;67:1604851. [doi: 10.3389/ijph.2022.1604851] [Medline: 35774355]
5. D'Agostino EM, Rosenberg LM, Richmond A, et al. You & me: test and treat study protocol for promoting COVID-19 test and treatment access to underserved populations. BMC Public Health 2023 Oct 28;23(1):2121. [doi: 10.1186/s12889-023-16960-6] [Medline: 37898741]
6. Baymon DE, Vakkalanka JP, Krishnadasan A, et al. Race, ethnicity, and delayed time to COVID-19 testing among US health care workers. JAMA Netw Open 2024 Apr 1;7(4):e245697. [doi: 10.1001/jamanetworkopen.2024.5697] [Medline: 38598239]
7. Abdul-Mutakabbir JC, Tillman F III, Marcelin JR, Saunders IM, Arya V. Slowed progression: the utility of test to treat initiatives in improving the neglected inequities of COVID-19 among racially/ethnically minoritized groups. J Am Pharm Assoc (2003) 2023;63(1):424-429. [doi: 10.1016/j.japh.2022.10.025] [Medline: 36470731]
8. Asabor EN, Warren JL, Cohen T. Racial/ethnic segregation and access to COVID-19 testing: spatial distribution of COVID-19 testing sites in the four largest highly segregated cities in the United States. Am J Public Health 2022 Mar;112(3):518-526. [doi: 10.2105/AJPH.2021.306558] [Medline: 35196059]
9. Khazanchi R, Strumpf A, Essien UR, Powers SD, McManus KA. Geographic accessibility of COVID-19 test to treat sites by race, ethnicity, age, and rurality. JAMA Netw Open 2022 Nov 1;5(11):e2241144. [doi: 10.1001/jamanetworkopen.2022.41144] [Medline: 36350654]

XSL•FO
RenderX

10. Reitsma MB, Claypool AL, Vargo J, et al. Racial/ethnic disparities in COVID-19 exposure risk, testing, and cases at the subcounty level in California: study examines racial/ethnic disparities in COVID-19 risk, testing, and cases. Health Aff 2021 Jun 1;40(6):870-878. [doi: 10.1377/hlthaff.2021.00098] [Medline: 33979192]
11. Mody A, Pfeifauf K, Bradley C, et al. Understanding drivers of coronavirus disease 2019 (COVID-19) racial disparities: a population-level analysis of COVID-19 testing among Black and White populations. Clin Infect Dis 2021 Nov 2;73(9):e2921-e2931. [doi: 10.1093/cid/ciaa1848] [Medline: 33315066]
12. Overview of testing for SARS-CoV-2. Centers for Disease Control and Prevention (CDC). 2020 Jul 2. URL: https://www.cdc.gov/covid/hcp/clinical-care/overview-testing-sars-cov-2.html [accessed 2020-04-13]
13. Zimba R, Kulkarni S, Berry A, et al. SARS-CoV-2 testing service preferences of adults in the United States: discrete choice experiment. JMIR Public Health Surveill 2020 Dec 31;6(4):e25546. [doi: 10.2196/25546] [Medline: 33315584]
14. Mosler HJ. A systematic approach to behavior change interventions for the water and sanitation sector in developing countries: a conceptual model, a review, and a guideline. Int J Environ Health Res 2012;22(5):431-449. [doi: 10.1080/09603123.2011.650156] [Medline: 22292899]
15. Rosenstock IM. Historical origins of the health belief model. Health Educ Monogr 1974 Dec;2(4):328-335. [doi: 10.1177/109019817400200403]
16. Rosenstock IM. The health belief model and preventive health behavior. Health Educ Monogr 1974 Dec;2(4):354-386. [doi: 10.1177/109019817400200405]
17. Janz NK, Becker MH. The health belief model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
18. Penchansky R, Thomas JW. The concept of access: definition and relationship to consumer satisfaction. Med Care 1981 Feb;19(2):127-140. [doi: 10.1097/00005650-198102000-00001] [Medline: 7206846]
19. Kjaer T. A review of the Discrete Choice Experiment—with emphasis on its application in health care.: Syddansk Universitet; 2005. URL: https://www.sdu.dk/-/media/files/om_sdu/centre/cohere/working+papers/20051pdf.pdf [accessed 2025-09-23]
20. van den Broek-Altenburg E, Atherly A. Using discrete choice experiments to measure preferences for hard to observe choice attributes to inform health policy decisions. Health Econ Rev 2020 Jun 11;10(1):18. [doi: 10.1186/s13561-020-00276-x] [Medline: 32529586]
21. Paykin S, Menghaney M, Lin Q. Rural, suburban, urban classification for small area analysis. 2021. URL: https://github.com/spaykin/rural-urban-classification [accessed 2025-08-30]
22. Coast J, Al-Janabi H, Sutton EJ, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. Health Econ 2012 Jun;21(6):730-741. [doi: 10.1002/hec.1739] [Medline: 21557381]
23. Zhang DC, Highhouse S, Nye CD. Development and validation of the general risk propensity scale (GRiPS). Behavioral Decision Making 2019 Apr;32(2):152-167. [doi: 10.1002/bdm.2102]
24. Han H, Blackburn AM, Jeftić A, et al. Validity testing of the conspiratorial thinking and anti-expert sentiment scales during the COVID-19 pandemic across 24 languages from a large-scale global dataset. Epidemiol Infect 2022 Sep 12;150:e167. [doi: 10.1017/S0950268822001443] [Medline: 36093606]
25. Lieberoth A, Lin SY, Stöckli S, et al. Stress and worry in the 2020 coronavirus pandemic: relationships to trust and compliance with preventive measures across 48 countries in the COVIDiSTRESS global survey. R Soc Open Sci 2021 Feb 10;8(2):200589. [doi: 10.1098/rsos.200589] [Medline: 33972837]
26. Viney R, Lancsar E, Louviere J. Discrete choice experiments to measure consumer preferences for health and healthcare. Expert Rev Pharmacoecon Outcomes Res 2002 Aug;2(4):319-326. [doi: 10.1586/14737167.2.4.319]
27. Bishaw A, Posey KG. A comparison of rural and urban America: household income and poverty. Census Social, Economic and Housing Statistics Division. URL: https://www.census.gov/newsroom/blogs/random-samplings/2016/12/a_comparison_of_rura.html#:~:text=According%20to%20the%202015%20American%20Community%20Survey%2C%20median%20family%20incomes%20below%20the%20official%20poverty%20threshold [accessed 2024-01-25]
28. McFadden D. Conditional logit analysis of qualitative choice behavior. : University of California, Berkely URL: https://escholarship.org/content/qt61s3q2xr/qt61s3q2xr.pdf [accessed 2025-09-03]
29. Stammann A, Heiss F, McFadden D. Estimating fixed effects logit models with large panel data. : ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg URL: https://www.econstor.eu/handle/10419/145837 [accessed 2025-09-03]
30. Clogg CC. Latent class models. In: Arminger G, Clogg CC, Sobel ME, editors. In: Handbook of Statistical Modeling for the Social and Behavioral Sciences: Springer US; 1995:311-359. [doi: 10.1007/978-1-4899-1292-3_6]
31. Greene WH, Hensher DA. A latent class model for discrete choice analysis: contrasts with mixed logit. Transp Res B Methodol 2003 Sep;37(8):681-698. [doi: 10.1016/S0191-2615(02)00046-2]
32. Beck MJ, Rose JM, Hensher DA. Consistently inconsistent: the role of certainty, acceptability and scale in choice. Transp Res E Logist Transp Rev 2013 Sep;56:81-93. [doi: 10.1016/j.tre.2013.05.001]
33. Louviere JJ, Hensher DA, Swait JD. Stated Choice Methods: Analysis and Applications: Cambridge University Press; 2000. [doi: 10.1017/CBO9780511753831]

XSL•FO
RenderX

34. Thomadsen R, Rooderkerk RP, Amir O, et al. How context affects choice. Cust Need and Solut 2018 Mar;5(1-2):3-14. [doi: 10.1007/s40547-017-0084-9]
35. Atherly A, Feldman RD, Dowd B, van den Broek-Altenburg E. Switching costs in medicare advantage. Forum Health Econ Policy 2020 Mar 5;23(1):20190023. [doi: 10.1515/fhep-2019-0023] [Medline: 32134731]
36. Long JC, Delamater PL, Holmes GM. Which definition of rurality should I use?: The relative performance of 8 federal rural definitions in identifying rural-urban disparities. Med Care 2021;59(Suppl 5):S413-S419. [doi: 10.1097/MLR.0000000000001612] [Medline: 34524237]
37. Parker K, Horowitz JM, Brown A, Fry R, Cohn D, Igielnik R. Demographic and economic trends in urban, suburban and rural communities. Pew Research Center's Social & Demographic Trends Project. 2018 May 22. URL: https://www.pewresearch.org/social-trends/2018/05/22/demographic-and-economic-trends-in-urban-suburban-and-rural-communities/ [accessed 2025-09-23]
38. Leider JP, Castrucci BC, Robins M, et al. The exodus of state and local public health employees: separations started before and continued throughout COVID-19. Health Aff (Millwood) 2023 Mar;42(3):338-348. [doi: 10.1377/hlthaff.2022.01251] [Medline: 36877909]
39. Leider JP, McCullough JM, Singh SR, et al. Staffing up and sustaining the public health workforce. J Public Health Manag Pract 2023;29(3):E100-E107. [doi: 10.1097/PHH.0000000000001614] [Medline: 36228097]
40. Lee KM, Bosold A, Alvarez C, Dada OO, Porterfield DS. Experiences and lessons learned from surging the governmental public health workforce during the COVID-19 pandemic. J Public Health Manag Pract 2025;31(2):283-290. [doi: 10.1097/PHH.0000000000002051] [Medline: 39248717]
41. Parker K, Horowitz JM, Brown A, Fry R, Cohn D, Igielnik R. How urban, suburban and rural residents' view social and political issues. Pew Research Center's Social & Demographic Trends Project. 2018 May 22. URL: https://www.pewresearch.org/social-trends/2018/05/22/urban-suburban-and-rural-residents-views-on-key-social-and-political-issues/ [accessed 2025-09-23]
42. Cruz JL, Luke DA, Ceballos RM, Ramanadhan S, Emmons KM. Reconceptualizing rurality: exploring community capital to identify distinct rural classes in the United States. SSM Popul Health 2025 Mar;29:101729. [doi: 10.1016/j.ssmph.2024.101729] [Medline: 39759383]
43. Dudley L, Couper I, Kannangarage NW, et al. COVID-19 preparedness and response in rural and remote areas: a scoping review. PLOS Glob Public Health 2023;3(11):e0002602. [doi: 10.1371/journal.pgph.0002602] [Medline: 37967067]
44. O'Sullivan B, Leader J, Couch D, Purnell J. Rural pandemic preparedness: the risk, resilience and response required of primary healthcare. Risk Manag Healthc Policy 2020;13:1187-1194. [doi: 10.2147/RMHP.S265610] [Medline: 32904086]
45. O'Sullivan TL, Phillips KP. From SARS to pandemic influenza: the framing of high-risk populations. Nat Hazards (Dordr) 2019;98(1):103-117. [doi: 10.1007/s11069-019-03584-6] [Medline: 32214659]
46. Hare Bork R, Robins M, Schaffer K, Leider JP, Castrucci BC. Workplace perceptions and experiences related to COVID-19 response efforts among public health workers—Public Health Workforce Interests and Needs Survey, United States, September 2021-January 2022. MMWR Morb Mortal Wkly Rep 2022 Jul 22;71(29):920-924. [doi: 10.15585/mmwr.mm7129a3] [Medline: 35862270]
47. Leveraging culture and history to improve health and equity in rural communities. : The Walsh Center for Rural Health Analysis; 2018 Feb URL: https://www.norc.org/content/dam/norc-org/pdfs/Practice%20Brief_Culture%20and%20History_2-5-18_long.pdf [accessed 2025-08-30]

Abbreviations

BIC: Bayesian information criterion
CAIC: consistent Akaike information criterion
CL: conditional logit
DCE: discrete choice experiment
EM: expectation-maximization
LC: latent class
LCA: latent class analysis

Edited by A Mavragani; submitted 12.11.24; peer-reviewed by K Elechi, S Apelian; revised version received 10.05.25; accepted 19.05.25; published 21.10.25.

Please cite as:
van den Broek-Altenburg E, Benson J, Jonk Y, Leslie A, Carney J, Stein G
Diagnostic Testing Preferences in Rural and Vulnerable Populations During a Pandemic: Discrete Choice Experiment
JMIR Public Health Surveill 2025;11:e68734
URL: https://publichealth.jmir.org/2025/1/e68734
doi:10.2196/68734

XSL•FO
RenderX

© Eline van den Broek-Altenburg, Jamie Benson, Yvonne Jonk, Abimbola Leslie, Jan Carney, Gary Stein. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mapping Key Populations to Develop Improved HIV and AIDS Interventions: Multiphase Cross-Sectional Observational Mapping Study Using a District and City Approach

Pande Putu Januraga[1], MPH, MD, DrPH; Endang Lukitosari[2], MD; Lanny Luhukay[2], MD; Rizky Hasby[2], MSc; Aang Sutrisna[3], MSc

[1]Center for Public Health Innovation, Faculty of Medicine, Udayana University, Denpasar, Indonesia

[2]Ministry of Health of Indonesia, Jakarta, Indonesia

[3]Monitoring, Evaluation, and Learning Platform USAID, Jakarta, Indonesia

Corresponding Author:
Pande Putu Januraga, MPH, MD, DrPH
Center for Public Health Innovation
Faculty of Medicine
Udayana University
Jl PB Sudirman
Denpasar, 80232
Indonesia
Phone: 62 81246180389
Email: januraga@unud.ac.id

Abstract

Background: Indonesia's vast archipelago and substantial population size present unique challenges in addressing its multifaceted HIV epidemic, with 90% of its 514 districts and cities reporting cases. Identifying key populations (KPs) is essential for effectively targeting interventions and allocating resources to address the changing dynamics of the epidemic.

Objective: We examine the 2022 mapping of Indonesia's KPs to develop improved HIV and AIDS interventions.

Methods: In 2022, a district-based mapping of KPs was conducted across 201 districts and cities chosen for their HIV program intensity. This multiphase process included participatory workshops for hotspot identification, followed by direct hotspot observation, then followed by a second direct observation in selected hotspots for quality control. Data from 49,346 informants (KPs) were collected and analyzed. The results from individual hotspots were aggregated at the district or city level, and a formula was used to estimate the population size.

Results: The mapping initiative identified 18,339 hotspots across 201 districts and cities, revealing substantial disparities in hotspot distribution. Of the 18,339 hotspots, 16,964 (92.5%) were observed, of which 1822 (10.74%) underwent a second review to enhance data accuracy. The findings mostly aligned with local stakeholders' estimates, but showed a lower median. Interviews indicated a shift in KP dynamics, with a median decline in hotspot attendance since the pandemic, and there was notable variation in mapping results across district categories. In "comprehensive" areas, the average results for men who have sex with men (MSM), people who inject drugs, transgender women, and female sex workers (FSWs) were 1008 (median 694, IQR 317-1367), 224 (median 114, IQR 59-202), 196 (median 167, IQR 81-265), and 775 (median 573, IQR 352-1131), respectively. "Medium" areas had lower averages: MSM at 381 (median 199, IQR 91-454), people who inject drugs at 51 (median 54, IQR 15-63), transgender women at 101 (median 55, IQR 29-127), and FSWs at 304 (median 231, IQR 118-425). "Basic" areas showed the lowest averages: MSM at 161 (median 73, IQR 49-285), people who inject drugs at 7 (median 7, IQR 7-7), transgender women at 59 (median 26, IQR 12-60), and FSWs at 161 (median 131, IQR 59-188). Comparisons with ongoing outreach programs revealed substantial differences: the mapped MSM population was >50% lower than program coverage; the estimates for people who inject drugs were twice as high as the program coverage.

Conclusions: The mapping results highlight significant variations in hotspots and KPs across districts and cities and underscore the necessity of adaptive HIV prevention strategies. The findings informed programmatic decisions, such as reallocating resources to underserved districts and recalibrating outreach strategies to better match KP dynamics. Developing strategies beyond identified hotspots, integrating mapping data into planning, and adopting a longitudinal approach to understand KP behavior over time are critical for effective HIV and AIDS prevention and control.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e56820) doi:10.2196/56820

KEYWORDS

Indonesia; key population; mapping; pandemic; HIV; AIDS; hotspot

Introduction

Background

Indonesia faces a multifaceted HIV epidemic shaped by its extensive archipelago and significant population size: 90% of its 514 districts and cities, home to >260 million people, have reported HIV and AIDS cases [1]. This vast distribution presents unique challenges in the fight against HIV. It is projected that in 2023, there will be 543,509 people living with HIV aged ≥15 years, and the number of new HIV infections will reach 23,040 in that same year (Ministry of Health of Indonesia, unpublished data, 2023).

Although HIV transmission has been reported throughout the nation, the epidemic remains predominantly concentrated among key populations (KPs), including men who have sex with men (MSM), transgender women, female sex workers (FSWs), and people who inject drugs [2]. However, the dynamics of the HIV epidemic differ among these groups. The MSM group has seen a rise in HIV prevalence, in stark contrast to the other groups that have experienced a decline. The most significant reduction in HIV cases was observed in the FSW group: in 2007, the HIV prevalence was 10%, but this figure fell to 2.1% in 2018-2019. A similar decline was noted among the group of people who inject drugs: approximately 52.4% of HIV cases in 2017 decreased dramatically by almost 60% to 13.6% in 2018-2019. Transgender women experienced a 50% reduction in HIV prevalence—from 24.33% in 2007 to 12% in 2018. Conversely, the MSM group exhibited an upward trend: HIV prevalence surged from 5.33% in 2007 to 17.9% in 2018-2019 [3].

The observed decrease in HIV prevalence among KPs—except for the MSM group—is influenced by multiple factors and should not be interpreted as a significant reduction in risk or transmission at the community level. It is crucial to emphasize that the primary transmission mode in new cases reported in 2021 predominantly involved heterosexual groups, specifically transmission from husbands to wives or long-term partners. This includes long-term partners from among KPs. In addition, there remains a concerning deficiency in the HIV care cascade in Indonesia [4]: in 2022, only 42% of people living with HIV were on antiretroviral therapy (ART; Ministry of Health of Indonesia, unpublished data, 2023).

To effectively combat the HIV epidemic, it is paramount that HIV prevention and control initiatives prioritize KPs in Indonesia [4]. Given the country's unique characteristics, these programs must have precise data about their target populations. KPs were mapped to offer insights into these target groups' size and role in Indonesia's HIV dynamics. The most recent mapping was conducted in 2022. What distinguishes the mapping efforts of 2022 is their execution in the post–COVID-19 pandemic period, a time when the pandemic has profoundly altered the public health landscape [5]. Various Indonesian studies have documented shifts in health care behaviors and access among KPs due to the pandemic. Notably, the COVID-19 era has also driven significant changes in digital technology use, impacting both risky behaviors and access to health services among these populations [6,7]. The mapping conducted during this period considers these pandemic-induced changes, offering insights into the evolving dynamics of health service use and risk behaviors in a world altered by the COVID-19 pandemic. This makes the recent mapping efforts particularly relevant and timely, reflecting the new normal of health care engagement and the needs of KPs in Indonesia. Notably, the findings from this mapping exercise have not been reported in scientific publications.

Objectives

This study examines the outcomes of district- and city-level KP mapping in Indonesia. First, it aims to pinpoint hotspots for KPs, describing the types of places where these individuals gather, meet sexual partners, or obtain drugs for injection. In addition, the study seeks to estimate the size of KPs in selected districts and cities. Finally, it aims to compare the 2022 mapping data for 201 districts and cities with data from the pre–COVID-19 period, shedding light on the changes and continuities in KP distributions influenced by the pandemic's impact on public health and societal norms.

Methods

Mapping Frameworks

The 2022 mapping followed the basic methodology of the Priorities for Local AIDS Control Efforts (PLACE) strategy. PLACE is a cross-sectional data collection approach that recruited informants directly from hotspot locations. This method is a location based and aims to identify specific areas where targeted outreach services could effectively reach populations at high risk of HIV infection [8,9].

The core objective of the 2022 mapping effort using the PLACE methodology was to identify and map locations frequented by individuals engaged in high-risk behaviors, such as meeting sexual partners or obtaining drugs for injection. Once these locations were mapped, informants were selected and recruited from these areas, especially if the number of locations exceeded a predetermined maximum quota. Therefore, PLACE was a comprehensive strategy that involved not only identifying hotspot locations but also extensively mapping these areas. Subsequently, informants from the selected locations were systematically recruited and interviewed.

Definitions

For the purpose of mapping to identify KP sizes in Indonesia, standard definitions of the target groups were needed. FSWs were defined as women who sold sex for money or goods as their primary source of income. These women worked in brothels, streets, or public places where customers came to buy sex. They may or may not have worked for a manager or pimp.

XSL•FO
RenderX

This group also included women who worked in entertainment venues (such as karaoke, bars, massage parlors, etc) and sold sex to the customers encountered in these venues. Sex transactions occurred both within entertainment venues and outside, and the owner or manager of these venues may or may not have facilitated such transactions.

MSM, defined as men who have sex with their male partners, included those who identified as gay, bisexual, or straight, as well as individuals who sold and bought sex from other men (male sex workers). Transgender women included biologically male individuals who identified as female and behaved and dressed like women. People who inject drugs comprised men and women who had injected drugs in the past 12 months, excluding those prescribed by a medical professional. This group also included individuals receiving opiate substitution therapy or participating in abstinence-based programs.

The time variable of the past 12 months was included in the definition of people who inject drugs to ensure accurate and current data on drug injection behaviors, which could fluctuate more frequently and were subject to recent intervention impacts. By contrast, behaviors and identities within the populations of FSWs, MSM, and transgender women were generally more stable over time and did not typically require a recent time frame for accurate classification.

Mapping Procedures

The mapping procedure was carried out in multiple phases, beginning with the selection of districts and cities. For the 2022 KP mapping, 201 districts and cities were chosen proportionally based on a classification system established by the Ministry of Health, which categorized regions into 3 levels of HIV control programs: “basic,” “medium,” and “comprehensive.” Basic regions had minimal HIV testing and treatment services; medium regions provided additional services, including health promotion, HIV prevention, outreach, and condom distribution; and comprehensive regions offered a full range of HIV and AIDS services, including all levels of prevention and advocacy related to health laws in HIV control.

The selection criteria included various parameters, such as the estimated population numbers for MSM, transgender women, FSWs, and people who inject drugs as of 2020. Additional factors considered included the cumulative number of new HIV cases reported from 2011 to 2021, projected new cases for the period from 2022 to 2024, and the number of people living with HIV who were receiving ART in 2021.

The comprehensive category consisted of 100 districts and cities, accounting for 80.59% (42,681/52,955) of the total new HIV cases, individuals receiving ART, and most KPs. Of these 100 districts and cities, 97 (97%) were selected as mapping locations in 2022. The medium category included 138 districts and cities with a significant proportion of new cases and KPs, with 89 (64.5%) selected for mapping in 2022. The basic category comprised 276 districts and cities that had limited interventions in HIV testing services and ART, of which 15 (5.4%) were included in the 2022 mapping. The number of districts and cities selected for KP mapping in 2022 was proportional to the estimated number of KPs in 2020, as well as the district and city categories for HIV control programs. In addition, districts and cities that participated in the integrated biological and behavioral survey of KPs in 2023 were confirmed as locations for mapping.

Overall, the mapping activities covered 201 (39.1%) of the 514 districts and cities in Indonesia. Identifying hotspots within the 201 selected districts and cities was achieved through 1-day participatory workshops held in July and August 2022. These workshops included 4 to 13 stakeholders from each district or city, with 1809 participants across all locations. Stakeholders included representatives from the district health office, social welfare office, local police, outreach nongovernmental organizations (NGOs), and community representatives. During the workshops, working teams were formed in each district and city, consisting of a supervisor from the health office, field officers from health offices and health centers, local NGOs, and a data manager from the health office. The workshops produced a list of hotspots, incorporating insights from stakeholders about key individuals and peak activity times at each identified location.

After the workshop phase, field officers—1 individual each from the health office, local NGO, and health center—received training on the technical implementation of mapping before visiting the hotspots. The training was conducted in groups of districts, which was a cost-effective approach to maximizing the number of trained officers. Topics covered included how to perform mapping and observations at hotspots, conduct interviews, and fill out data using paper-based forms.

The mapping team carried out hotspot observations in the next phase of the project. The team members visited specific locations known as gathering points for KPs during peak hours from July to October 2022. For quality control purposes, 1822 (10.74%) of the 16,964 observed hotspots underwent a second observation. Trained enumerators conducted interviews at these hotspots using predetermined paper-based forms to collect data. The interviews addressed various aspects, such as the number of KPs present and how often they visited the hotspots within a single day. Before the interviews, all participants were asked for their informed consent to participate in the mapping initiative. Informant interviews were conducted at each hotspot with at least 2 individuals from KPs and 2 individuals from outside KPs, such as clients of sex workers or other relevant visitors.

On the basis of the information provided by the informants, the working teams from the districts and the mapping teams assessed the reliability or confidence level of the information obtained, categorizing it as low or high. Only information rated with a high confidence level was used for further analysis. The validation criteria included complete visit times and fully completed forms. These data were subsequently used to calculate the counts of KPs per hotspot and for aggregating data at the district level.

Community Participation

Involving community members in the field team was crucial for observing KPs in hotspots and gaining the trust of key individuals within these populations. To enhance the accuracy

XSL•FO
RenderX

of the geomapping results, particularly for MSM and transgender women, district or city teams predominantly included members from MSM or transgender women communities in their field teams. A mixed approach was adopted for mapping FSWs, combining members from the FSW community with others who had experience with and knowledge about FSWs. When conducting mapping for people who inject drugs, the field team collaborated with the individuals responsible for accurately recording information on hotspot datasheets. This extensive engagement with individuals directly and indirectly involved with the KPs provided a comprehensive and nuanced understanding of the hotspot dynamics, which was crucial for developing effective HIV prevention and care strategies.

Data Management and Analysis

The mapping project used Microsoft Excel as a data management platform. This application allowed the mapping team to input, organize, and analyze information regarding observation times, hotspot characteristics, and informant details.

A careful mathematical and statistical approach was essential for accurately aggregating the number of KPs across districts, cities, and hotspots. By using a formula based on the average number of hotspots visited in the last 24 hours, we were able to estimate KP mapping results on a larger scale. This method ensured high accuracy and minimized the risk of duplication.

The aggregation process involved consolidating the mapping results from multiple hotspots into district and city totals. By focusing on the average hotspots visited, we ensured that the figures reflected the true distribution of KPs while avoiding excessive duplication. The following simple formula was used:

$$\sum [\![\text{in district-city} = (\sum [\![\text{hotspot mapping}]\!]) / (\text{average number of hotspots visited})]\!]$$

In this context, ∑(Hotspot mapping) represents the median number of individuals identified during peak times and Saturdays from all credible informants across all hotspots within a particular district or city. The “Average number of hotspots visited” corresponds to the mean number of distinct hotspots that a member of the KP reported visiting in the past 24 hours.

This simplified formula is the initial step in quantifying the size of KPs, although it does not fully encapsulate the complexities of KP movements and interactions [10-12]. Nonetheless, it provides a valuable baseline for refining the total KP figures at the district level, aiming to minimize the potential for duplication in the mapping process.

Ethical Considerations

The mapping initiative was approved after receiving ethics clearance from Atmajaya University (0010A/III/PPPE.PM.10.05/06/2022). Due to the legal implications associated with activities such as sex work and drug injection in Indonesia and the ongoing stigma and discrimination faced by other KPs, verbal consent was carefully obtained during the interview phase to ensure the protection of participants’ rights and confidentiality. No financial incentives were provided. The data obtained from the mapping initiative are securely stored in a password-protected folder at Udayana University and the Ministry of Health. The stored and analyzed data do not include any personally identifiable information.

Results

Mapping Results Based on Stakeholder-Provided Information During the District- or City-Level Workshops

Through collaborative, multisectoral hotspot identification at the district or city level workshops, 18,339 hotspots were identified nationwide, spanning 201 districts and cities. On the basis of stakeholder-provided information, it was estimated that, of these 18,339 hotspots, 8030 (43.79%) were frequented by FSWs, 6923 (37.75%) by MSM, 2718 (14.82%) by transgender women, and 668 (3.64%) by people who inject drugs.

The analysis of hotspot identification revealed a clear trend: districts and cities categorized as “comprehensive” have a significantly higher average number of hotspots for each KP group than those labeled “medium” and “basic.” Specifically, for MSM, on average, there were 48 hotspots per district or city in the comprehensive category compared to 24 in the medium category and 12 in the basic category; for transgender women, there were 18 hotspots on average in the comprehensive category, 11 in the medium category, and 7 in the basic category; FSWs reported an average of 54 hotspots in the comprehensive category, 30 in the medium category, and 16 in the basic category; while people who inject drugs had the fewest hotspots, averaging 12 in the comprehensive category, 6 in the medium category, and 1 in the basic category.

While the types of hotspots are often similar across districts with comprehensive, medium, and basic classifications for each KP, preferences for specific hotspot types varied among the different groups (Table 1). For MSM and people who inject drugs, open spaces were the predominant hotspot type at all priority levels of districts and cities, with the highest proportions in the comprehensive category. By contrast, salons were the primary hotspot type for transgender women, especially in the comprehensive and medium districts. Although slightly less prevalent in basic districts, salons remained a significant hotspot type. Entertainment venues consistently emerged as the most common hotspot for the FSW population across all categories, underscoring their role as central activity points for FSWs.

Overall, mapping the number of KPs at the district or city level, based on stakeholder-provided information, revealed broader disparities between the priority categories of districts and cities compared to the individual hotspot levels. These variations are depicted in Figure 1, illustrating the differences in the number of hotspots and KP numbers across the priority categories of districts and cities.

XSL•FO
RenderX

Table 1. Types of hotspots identified by the stakeholders (n=17,092).

Types of hotspots	MSM[a] (n=6851), n (%)			People who inject drugs (n=651), n (%)			Transgender women (n=2687), n (%)			FSWs[b] (n=6903), n (%)		
	Comprehensive (n=97)	Medium (n=89)	Basic (n=15)	Comprehensive (n=97)	Medium (n=89)	Basic (n=15)	Comprehensive (n=97)	Medium (n=89)	Basic (n=15)	Comprehensive (n=97)	Medium (n=89)	Basic (n=15)
Open spaces	1178 (26)	562 (27)	49 (28)	336 (60)	62 (58)	1 (100)	353 (21)	169 (19)	20 (19)	562 (11)	276 (11)	43 (18)
Cafes	1127 (24)	529 (26)	53 (30)	42 (8)	6 (6)	—[c]	61 (4)	73 (8)	21 (20)	753 (15)	636 (25)	75 (31)
Recreation areas	1038 (23)	433 (21)	26 (15)	119 (21)	30 (28)	—	91 (5)	53 (6)	5 (5)	225 (4)	173 (7)	16 (7)
Salons	437 (9)	257 (12)	34 (19)	1 (0.2)	1 (1)	—	819 (48)	488 (55)	46 (45)	116 (2)	94 (4)	10 (4)
Entertainment venues	278 (6)	92 (4)	07 (4)	6 (1.1)		—	44 (3)	33 (4)	6 (6)	1,180 (23)	526 (21)	45 (19)
Residences	449 (10)	132 (6)	04 (2)	37 (7)	5 (5)	—	248 (15)	30 (3)	1 (1)	524 (10)	298 (12)	6 (2)
Fixed places	—	—	—	—	—	—	48 (3)	33 (4)	1 (1)	753 (15)	238 (9)	12 (5)
Massage parlors	23 (0.5)	8 (0.4)	1 (0.6)	—	—	—	2 (0.1)	2 (0.2)	—	707 (13.7)	159 (6.3)	14 (5.8)
Hotels	61 (1)	38 (2)	4 (2)	2 (0.4)	—	—	14 (0.8)	8 (0.9)	1 (1)	321 (6.2)	108 (4.3)	20 (8.3)
Malls or markets	6 (0.1)	3 (0.1)	—	—	—	—	—	0 (0)	—	—	7 (0.3)	—
Others	14 (0.3)	8 (0.4)	—	16 (3)	3 (3)	—	13 (0.8)	2 (0.2)	2 (2)	6 (0.1)	—	—

[a]MSM: men who have sex with men.

[b]FSW: female sex worker.

[c]Not applicable.

Figure 1. Differences in the number of key populations (KPs) identified by stakeholders across the priority categories of districts and cities. FSW: female sex worker; MSM: men who have sex with men.

XSL•FO
RenderX

Mapping Results Based on Observations and Interviews

Of the 18,339 hotspots identified across the 201 districts and cities, 16,964 (92.5%) were successfully observed. Despite these efforts, some of the hotspots remained unobserved: 6.4% (443/6923) of the MSM hotspots, 6.9% (46/668) of the hotspots frequented by people who inject drugs, 8.64% (235/2718) of the hotspots frequented by transgender women, and 8.1% (651/8032) of the FSW hotspots.

The primary reason for not observing some hotspots was changes in the field, as reported by the informants. Notably, several open spaces, previously gathering places for KPs, were no longer used. Other hotspot types, including eateries, entertainment venues, recreation areas, salons, fixed locations for selling sex, massage parlors, hotels, and malls or markets, were found to be closed or no longer operational. A few hotspots remained unobserved due to geographic challenges, where the distance from the district or city capital posed a barrier for the mapping officers.

The majority of the hotspot observations (13,832/16,964, 81.53%) for the 4 KPs were conducted from the afternoon into the evening, specifically between noon and 11:59 PM. The most active observation period was noted to be between noon and 5:59 PM, accounting for 46% (7803/16,964) of all observations across the KPs, indicating this as the prime time for data collection. Nighttime observations (6 PM- 11:59 PM) held particular importance for transgender women and FSWs, with nearly one-third (770/2483, 31.01%) to more than one-third (2730/7379, 37%) of their total observations commencing during these hours. Conversely, morning observations (6 AM-11:59 AM) were less frequent across all populations, with people who inject drugs mainly showing the lowest rate of observations during this time period.

The observed data revealed a pattern in the distribution of KP numbers, which generally corresponded with the estimates provided by stakeholders at the district or city level, albeit with a slightly lower median value. This consistency suggested that, despite the lower observed figures, the overall trends and relative differences among districts and cities of varying priority categories still mirrored the stakeholder estimates. The observations verified significant concentrations of KPs at certain hotspots. Nonetheless, the maximum values recorded indicate that while some hotspots were highly populated, their population densities were typically not as high as initially projected by the stakeholders. Figure 2 graphically displays these findings, illustrating the number of KPs per hotspot in different district or city categories.

Figure 2. Distribution of the number of key populations (KPs) observed per hotspot and per district or city. FSW: female sex worker; MSM: men who have sex with men.

XSL•FO
RenderX

Furthermore, the mapping results incorporate data from interviews conducted with the informants during the observations. During the hotspot observation period, the mapping team interviewed 49,346 informants, averaging 2 to 3 informants per hotspot. Of these 49,346 informants, 31,588 (64.01%) were members of KPs, while the remaining 17,758 (35.99%) were not classified as KPs. Specifically, on average, 2.2 informants from among people who inject drugs were interviewed per hotspot (1346 informants in total). In hotspots frequented by MSM and transgender women, on average, 1.9 informants were interviewed per hotspot (12,415 MSM and 4699 transgender women informants in total). In hotspots frequented by FSWs, on average, 1.8 informants were interviewed per hotspot (13,128 FSWs in total).

These data are categorized into 3 groups: hotspot conditions during peak times, hotspot conditions on Saturdays, and hotspot conditions before the COVID-19 pandemic. The aggregated mapping, displaying the distribution of KPs at hotspots during peak times across different districts and cities, revealed a widening gap between districts and cities categorized as comprehensive and those classified as medium or basic. This disparity is attributed to the higher number of hotspots in comprehensive districts and cities, which amplifies the difference in the mapped presence of KPs at these hotspots.

When the mapping of KP numbers at hotspots during peak times is compared with the stakeholders' estimates from Figure 1, it becomes evident that the mapped numbers during peak times are generally lower than the estimates provided by stakeholders. In addition, a comparison of the mapping distribution of KP numbers at peak times with previous data shows a typically lower Saturday attendance.

Using information from informants about the population density at hotspots before the pandemic compared to the current situation, it was found that 63.12% (10,708/16,964) of the hotspots had at least 1 informant who provided data on the number of KPs at peak times before the COVID-19 pandemic. The hotspots with the most prepandemic population data were FSW hotspots (2594/7379, 35.15%), followed by hotspots frequented by transgender women (677/2483, 27.27%) and hotspots frequented by people who inject drugs (129/622, 20.7%). Only 1 of the 6480 MSM hotspots lacked KP count information for peak times before the COVID-19 pandemic.

The analysis results regarding the median difference in the number of KPs at hotspots are detailed in Table 2. These results show a median decrease in the number of KPs present in almost all hotspot types at the time the study was conducted compared to the prepandemic period. For the MSM population, open spaces, cafes or stalls, and recreation or socialization areas exhibited a noticeable median decline: open spaces dropped from 15 (IQR 8-25) to 10 (IQR 7-20), cafes or stalls from 13 (IQR 8-20) to 10 (IQR 7-12), and recreation or socialization places from 15 (IQR 10-25) to 10 (IQR 7-20). The population of people who inject drugs also saw a decline, particularly in entertainment venues, with the prepandemic median of 15 (IQR 8.5-22.5) decreasing to 7.5 (IQR 6.25-8.5) at the time of the study was conducted.

Transgender women experienced a median decline in attendance across all hotspot types, with the most substantial decreases observed in massage parlors (from 15 to 8.5) and sex-selling locations (from 10 to 7.5). Similarly, FSW hotspots exhibited a median decline in attendance in most locations, with the most significant drops in entertainment venues (from 15 to 11) and massage parlors (from 8 to 6).

Finally, interviews with informants were conducted at these hotspot locations to account for KPs who did not frequent hotspots in our population estimates (Figure 3). For the MSM population in comprehensive districts and cities, the median number of individuals who had never visited a hotspot was 13, with a maximum of 270. In the context of the population of people who inject drugs within comprehensive districts and cities, the median was 4, escalating to a maximum of 400. This indicates that many people who inject drugs may not interact with existing hotspots. The median in comprehensive districts and cities for transgender women was 6, peaking at 78. Conversely, the FSW population in comprehensive districts displayed a median of 7, with the highest value recorded being 133.

XSL•FO
RenderX

Table 2. Distribution of key population numbers at hotspots before and during mapping (after the COVID-19 pandemic), categorized by hotspot type and key population.

Hotspot types	MSM[a], median (IQR)		People who inject drugs, median (IQR)		Transgender women, median (IQR)		FSWs[b], median (IQR)	
	Before the pandemic	During mapping	Before the pandemic	During mapping	Before the pandemic	During mapping	Before the pandemic	During mapping
Open spaces	15 (8-25)[c]	10 (7-20)[d]	8 (5-15)[c]	7 (5-13.75)[d]	10 (7-15)[c]	10 (6-13)[d]	12 (7-20)[c]	10 (6-15)[d]
Cafes	13 (8-20)[c]	10 (7-17)[d]	6 (4-10)[c]	5 (4.25-9.75)[d]	8.5 (5-15)[c]	6 (4-10)[d]	10 (6-17)[c]	8 (5-13)[d]
Salons	10 (7-15) [c]	8 (5-13)[d]	9 (9-9)[d]	12 (12-12)[c]	8 (5-12)[c]	6 (4-10)[d]	7 (5-11)[c]	5 (4-8)[d]
Entertainment venues	12 (8-20)[c]	10 (6-18.25)[d]	15 (8.5-22.5)[c]	7.5 (6.25-8.5)[d]	10 (6-13)[c]	6 (4-10.5)[d]	15 (7-28)[c]	11 (6-21)[d]
Recreation areas	15 (10-25)[c]	10 (7-20)[d]	9 (5-15)[c]	8 (5-15)[d]	10 (6-15)[c]	7.5 (5-12.75)[d]	9 (6-15)[c]	7 (5-10)[d]
Residences	12 (7-20)[c]	10 (6-15)[d]	10 (7-12.5)[c]	10 (5.5-12)[d]	10 (6-15)[c]	7 (5-10)[d]	12 (6-20)[c]	10 (6-15)[d]
Fixed places	—[e]	—	—	—	10 (7.75-15)[c]	7.5 (6-10)[d]	11 (6-25)[c]	9 (5-20)[d]
Massage parlors	14 (7.5-20)[c]	10 (6-15.5)[d]	—	—	15 (12.5-7.5)[c]	8.5 (7.75-9.25)[d]	8 (5-15)[c]	6 (4-12)[d]
Hotels	8 (5-20)[c]	9 (5-15)[d]	4 (4-4)[d]	5 (5-5)[c]	9.5 (5.75-15)[c]	8 (5.75-10)[d]	13 (8-20)[c]	10 (6-15)[d]
Malls or markets	7.5 (4.25-13.75)[c]	5.5 (4-13)[d]	—	—	—	—	5 (5-6)[c]	4 (3-6)[d]
Others	7 (6-8)[c]	6 (5-7)[d]	10 (4.75-25)[c]	5.5 (4-21.25)[d]	15 (9-22.5)[c]	12 (4.5-13.5)[d]	20 (10-20)[c]	10 (10-13)[d]

[a]MSM: men who have sex with men.

[b]FSW: female sex worker.

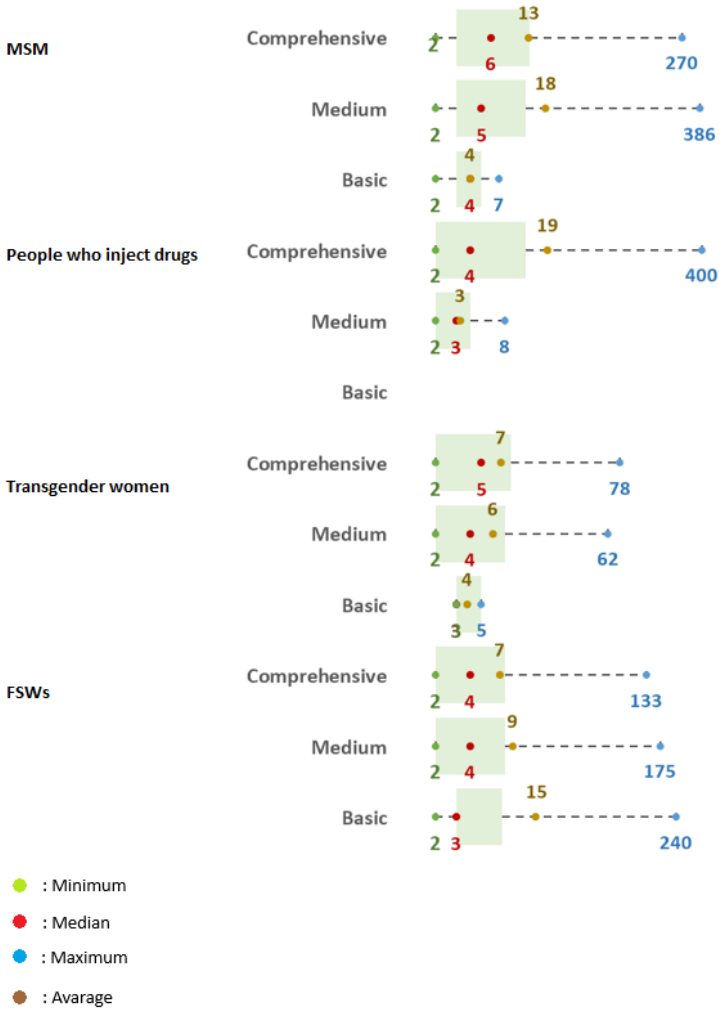
[c]Higher.

[d]Lower.

[e]Not applicable.

XSL•FO
RenderX

Figure 3. Distribution of key populations who never visited hotspots, based on mapping data. FSW: female sex worker; MSM: men who have sex with men.



Aggregation of Mapping Results in Districts and Cities

An aggregation of the mapping results was performed using the data collected from observations and interviews. The aggregations were carried out for each KP and district or city classification. Figure 4 illustrates the distribution of these aggregated mapping results across districts and cities, categorized by KPs and district or city classifications. The average number of mapping results per district or city for the MSM population in the comprehensive category was 1008 (median 694; range 381-6511). The wide range indicates a significant variation in mapping results across districts or cities in this category. By contrast, the medium category showed an average of 381 (median 199) mapping results. The basic category presented even lower figures, with an average of 161 (median 73), exhibiting more limited variation.

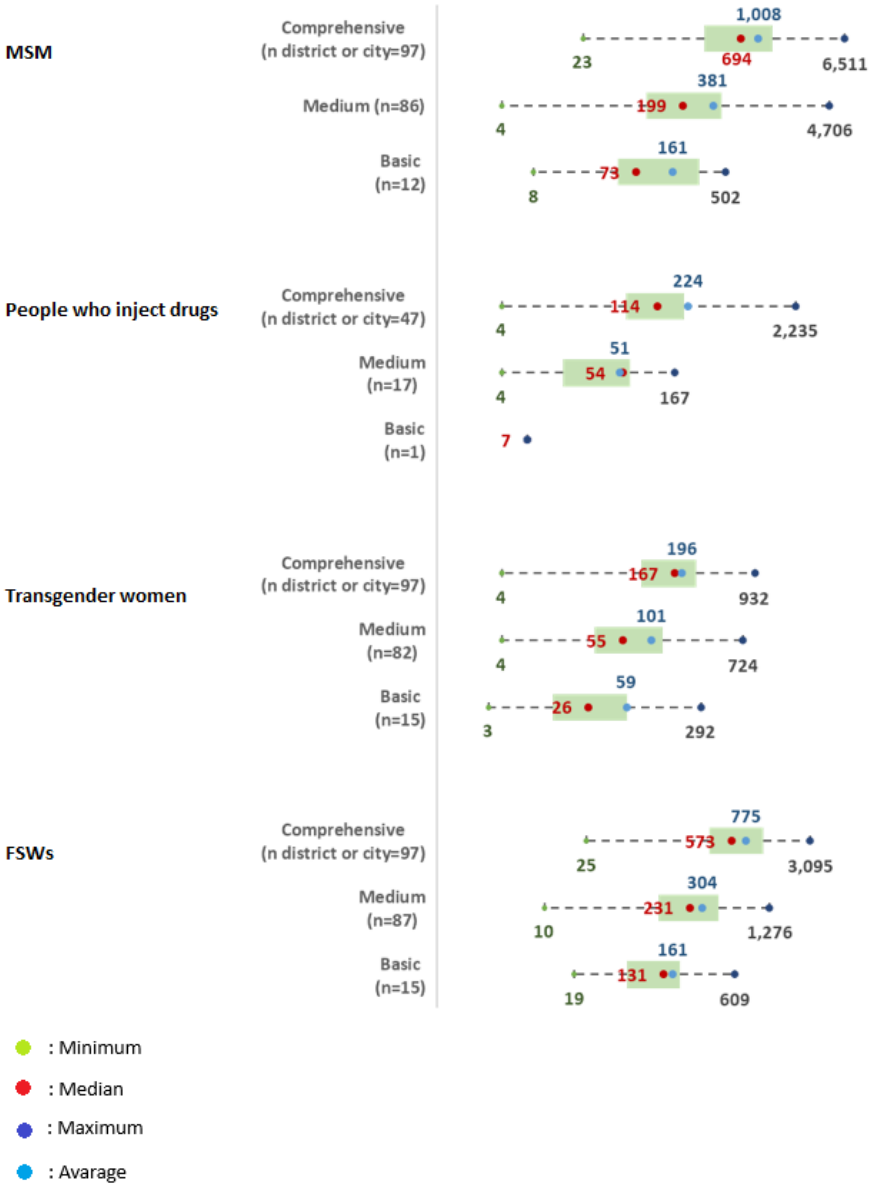
The population of people who inject drugs exhibits a similar trend but with overall lower numbers. The average aggregated mapping results in comprehensive districts and cities was 224 (median 114). The averages in the medium and basic categories were 51 (median 54) and 7, respectively, highlighting this population's disparity in aggregated mapping data.

For transgender women, the comprehensive category had an average of 196 (median 167) mapping results. In the medium category, the average decreased to 101 (median 55). The basic category showed an even lower average of 59 (median 26). Regarding the FSW population, comprehensive districts and cities averaged 775 (median 573) mapping results. The medium and basic categories saw a decrease in averages, dropping to 304 (median 231) and 161 (median 131), respectively.

Subsequently, the mapping results were compared with data from ongoing outreach programs. Table 3 provides the median number of populations mapped successfully and compares these figures with the outreach data. Generally, the mapping results yielded lower numbers than the outreach results collected throughout the program, and there was a notable discrepancy between the estimated numbers of MSM through mapping exercises and the coverage of HIV outreach and testing programs. These estimates indicated that the actual number of MSM was >50% lower than the program coverage. A similar disparity existed for FSWs, with estimates being 13% lower than outreach results. By contrast, for people who inject drugs and transgender women, the estimated numbers exceeded the extent of existing programs. Notably, the estimated figures for people who inject drugs were twice as high as the program coverage.

XSL•FO
RenderX

Figure 4. Distribution of aggregated mapping results across districts and cities, reported by key population type. FSW: female sex worker; MSM: men who have sex with men.



XSL•FO
RenderX

Table 3. Comparison of distribution of key population (KP) mapping data with program data by district or city.

KPs	Informed estimates	Observed	Saturdays	Peak times	Empirical estimates	Coverage by ORWs[a]	Had HIV test
MSM[b]							
Values, median (IQR; range)	238 (105.5-560; 2-4481)	189 (78-384; 2-2905)	275 (109-553; 4-4930)	323 (119-601; 4-5500)	376 (152-822; 4-6511)	2218 (1240-3412106-14,996)	934 (239-1899.5; 2-11,472)
Values, mean (SD)	425 (556)	299 (380)	470 (617)	499 (651)	697 (919)	2778 (2360)	1338 (1580)
People who inject drugs							
Values, median (IQR; range)	55 (21.5-119; 2-535)	36 (17-68; 2-278)	55 (19-97; 2-560)	64 (31-131; 2-555)	85 (36-179; 4-2235)	223.5 (80.25-441.5; 3-1473)	29.5 (24.75-167; 2-1046)
Values, mean (SD)	97 (120)	53 (59)	85 (103)	98 (108)	175 (364)	310 (323)	123 (191)
Transgender women							
Values, median (IQR; range)	63.5 (29-132.25; 1-717)	46 (22-871-369)	63 (31.75-143.5; 3-496)	73 (31-155; 3-507)	98.5 (38.75-208.5; 3-932)	158.5 (67-266.25; 14-814)	59 (16-167.5; 1-723)
Values, mean (SD)	95 (97)	63 (62)	100 (101)	106 (101)	148 (154)	194 (155)	102 (119)
FSWs[c]							
Values, median (IQR; range)	309 (143-538; 22-3603)	211 (119-371; 13-2718)	309 (150-536; 10-3087)	329 (175-551; 19-3099)	364 (180-660; 10-3095)	707 (309-1435; 69-3560)	291 (68.25-592.25; 1-2670)
Values, mean (SD)	465 (573)	317 (357)	436 (468)	465 (481)	528 (515)	956 (774)	431 (480)

[a]ORW: outreach worker.

[b]MSM: men who have sex with men.

[c]FSW: female sex worker.

Discussion

Principal Findings

The findings from district and city workshops showed that comprehensive districts and cities generally had significantly higher hotspot locations than medium and basic areas; for example, the average number of MSM hotspots per district or city in the comprehensive category was 48. This figure contrasts with 24 in the medium category and 12 in the basic category. It is important to target specific areas for HIV and AIDS interventions, such as open spaces for MSM, salons for transgender women, and entertainment venues for FSWs. Although there has been a decrease in hotspot density after the pandemic, the types of hotspots identified remain consistent with those commonly recognized. The preferences of KPs for these hotspots have also shown little change. Similar trends were observed in other mapping studies conducted in different geographic settings [13,14]. Despite varying proportions among district categories, these hotspot types are critical points for intervention across all levels of district priority.

Insights from observation data show that the concentration of KPs in hotspots varied, based on the district or city category; for instance, in the comprehensive category, the median number of MSM per district or city was 302, while for people who inject drugs, transgender women, and FSWs, the median numbers were 46, 65, and 360, respectively. The comprehensive category had a significantly higher concentration of KPs per hotspot than the medium and basic categories. However, these numbers were lower than the estimates from workshop results, where the comprehensive category had higher medians: 422 for MSM, 75 for people who inject drugs, 100 for transgender women, and 471 for FSWs. This indicates a greater focus on comprehensive areas for HIV and AIDS control in Indonesia, with some attention also given to medium category areas. It is worth noting that basic areas also receive attention, including through the Global Fund Program, although this does not imply the absence of HIV prevention programs in these areas. Standard HIV testing and treatment services are available in basic areas, following the concept of allocating interventions based on geographic epidemiological conditions [15]. Moreover, key findings from field observations indicated significant fluctuations in the presence of KPs in hotspots, particularly during peak periods. This highlights the necessity for careful interpretation of the data and underscores the complexity of addressing HIV and AIDS in varied district or city contexts [16].

The comparison between workshops' estimates of KP numbers and informant interview results revealed a consistent trend: stakeholder estimates were generally higher. This discrepancy likely stems from stakeholders basing their numbers on experience or expectations, whereas informant interviews provide a more realistic, albeit lower, representation of the actual situation. In addition, the variance might be attributed to the limited duration of observations, which often spanned only a few hours and may not have captured the full range of KP

XSL•FO
RenderX

arrivals. This suggests that extending observation periods and broadening the time frame could yield a more comprehensive and accurate portrayal of KP activities [17].

The report highlights the significant impact of the COVID-19 pandemic on the mobility of KPs. There has been a decrease in the number of KPs in almost all types of hotspots compared to the prepandemic period. For the MSM population, there has been a noticeable median decline in attendance in open spaces (from 15 to 10), cafes or stalls (from 13 to 10), and recreation or socialization areas (from 15 to 10). The population of people who inject drugs also experienced a decline in attendance, particularly in entertainment venues, with the prepandemic median of 15 decreasing to 7.5 today. This aligns with academic research that documents the pandemic's effects on mobility and social interactions, especially among KPs of individuals living with HIV [6,7,18]. Studies worldwide indicate a postpandemic decline in public space social interactions [19-21]. Understanding these dynamics is crucial for developing effective postpandemic public health interventions, particularly restoring social interactions in hotspots and preventing disease spread. Strategies should evolve in response to the altered social and epidemiological landscapes, drawing on pandemic experiences to ensure that HIV prevention and response are effectively tailored to the new context.

A comparison of the mapping results for KP distribution with the coverage of HIV prevention and testing programs revealed notable discrepancies. The actual number of MSM was >50% lower than the program coverage. A similar disparity existed for FSWs, with estimates being 13% lower than outreach results. Mapping, which uses various methods such as stakeholder inputs and direct observations, often yields lower estimates compared to the broader scope of ongoing, sustainable programs; for instance, the median number of MSM per district identified through mapping was 238, significantly lower than the median recorded by outreach program coverage, which was 2218. Similar patterns were observed for people who inject drugs, transgender women, and FSWs, where the figures from HIV outreach and testing programs consistently exceeded those from mapping exercises. This suggests that mapping provides a temporal "snapshot" and might not fully capture the year-round dynamics of KPs [11]. Therefore, integrating mapping results with data from year-long programs is crucial to gaining a more comprehensive understanding of the needs and effectiveness of HIV intervention programs [14].

Finally, the mapping process conducted across 201 Indonesian districts and cities identified several limitations that are important to acknowledge for accurate analysis and future planning. These include incomplete hotspot identification and focusing only on certain segments of KPs, leading to incomplete representation. The mapping primarily accounts for visible KP members at hotspots during observation times, omitting those who might visit at other times. Another limitation is the heavy reliance on key informants, whose reports may be affected by memory bias and challenges in estimating precise numbers, potentially compromising data accuracy.

In addition, as highlighted previously, data collection was predominantly conducted in a single visit, which fails to consider temporal changes and affects the reliability of the findings. Focusing on high or peak periods could also lead to overestimations that do not represent typical daily conditions. Consequently, these factors create challenges in extrapolating the observed "snapshot" into a more extensive understanding of KP presence over extended periods. Recognizing these limitations also presents opportunities for future refinement of mapping methodologies and strategies. Further research is essential to validate and expand upon these findings. Longitudinal data collection would be particularly beneficial, providing a deeper understanding of the evolving behaviors of KPs over time and in response to various influences, including global pandemics. Despite its limitations, the mapping exercise has yielded valuable insights into the distribution of KPs based on district or city categories, which are instrumental in enhancing HIV control efforts.

Conclusions

In conclusion, the findings from this comprehensive mapping of KPs of people living with HIV across Indonesian districts and cities carry significant public health implications. First, identifying high-risk hotspots and their concentration in comprehensive districts highlights the urgent need for geographically targeted interventions, which can optimize resource allocation and enhance the effectiveness of HIV prevention efforts. Despite the postpandemic changes, the consistency in the hotspots frequented by KPs indicates stable behavioral patterns, which should inform the design of tailored intervention strategies. The discrepancies between stakeholder estimates and direct observations underscore the importance of robust, longitudinal data collection to understand KP dynamics better. In addition, the impact of the COVID-19 pandemic on mobility and social interactions among KPs necessitates the adaptation of intervention strategies to the new social and epidemiological realities. Addressing these challenges through improved mapping methodologies and sustained public health initiatives can significantly enhance HIV control efforts, reduce transmission rates, and ultimately improve health outcomes for KPs in Indonesia.

Acknowledgments

In the mapping effort, the vital involvement of multiple stakeholders was crucial. The authors' deepest thanks go to the Ministry of Health for the consistent guidance and foundational support. The provincial and district health offices were indispensable, contributing to data gathering and contextualizing the research. The authors are deeply grateful to the dedicated supervisors and outreach workers whose tireless work connected research with ground realities. Their firsthand knowledge and engagement deepened the authors' understanding of the key populations. Central to this study were the communities to whom the authors owe their utmost appreciation. Their active participation, insights, and trust formed the cornerstone of this project. Collectively,

XSL•FO
RenderX

these contributors charted a clear course, ensuring that the authors' interventions were both realistic and impactful. The authors would also like to acknowledge the primary recipient of the Global Fund Program in Indonesia, the Ministry of Health, for the support in funding the mapping initiative and the publication cost.

Data Availability

Data sharing is available upon request to the Ministry of Health of Indonesia. The authors do not possess the data; moreover, they are not authorized to share any data with other researchers.

Authors' Contributions

PPJ and AS conceptualized the study. RH, LL, EL, and AS curated the data. PPJ, LL, and AS were responsible for formal analysis. RH, LL, and EL were responsible for funding acquisition. RH, LL, EL, and AS were responsible for investigation. PPJ and AS were responsible for methodology. RH was responsible for project administration. RH, LL, EL, and AS supervised the project. PPJ was responsible for validation and wrote the original draft. All authors reviewed and edited the manuscript.

Conflicts of Interest

None declared.

References

1. Indonesia HIV AIDS quarterly report, fourth quarter 2022. Ministry of Health of Indonesia. 2022. URL: https://hivaids-pimsindonesia.or.id/download/file/Laporan_tahunan_HIV_AIDS_2022.pdf [accessed 2024-04-29]
2. Ministry of Health of Indonesia. HIV AIDS Estimates and Projections in Indonesia Update 2023. HIV AIDS Estimates and Projections in Indonesia Update 2023. Jakarta: Ministry of Health of Indonesia; 2023. URL: https:// [accessed 2024-04-29]
3. Integrated biological and behavioral survey (IBBS) 2018-2019. Ministry of Health of Indonesia. 2018. URL: https://www.unicef.org/indonesia/media/8496/file/IBBS%20Report%20YKAP%202019.pdf [accessed 2024-04-29]
4. Januraga PP, Reekie J, Mulyani T, Lestari BW, Iskandar S, Wisaksana R, et al. The cascade of HIV care among key populations in Indonesia: a prospective cohort study. Lancet HIV 2018 Oct;5(10):e560-e568. [doi: 10.1016/S2352-3018(18)30148-6] [Medline: 30143455]
5. Gedela K, Januraga PP, Luis H, Wignall FS, Irwanto I. COVID-19 lockdown in Indonesia: greater investment will be needed to mitigate the impact on people living with HIV. Asia Pac J Public Health 2020 Nov 19;32(8):461-462. [doi: 10.1177/1010539520962611] [Medline: 33073580]
6. Magnani RJ, Wirawan DN, Sawitri AA, Mahendra IG, Susanti D, Utami Ds NK, et al. The short-term effects of COVID-19 on HIV and AIDS control efforts among female sex workers in Indonesia. BMC Womens Health 2022 Jan 26;22(1):21 [FREE Full text] [doi: 10.1186/s12905-021-01583-z] [Medline: 35081950]
7. Wirawan GB, Wardhani BD, Pradnyani PE, Nurhalina A, Sulaiman N, Sukmaningrum E, et al. Behavioral changes, adaptation, and supports among Indonesian female sex workers facing dual risk of COVID-19 and HIV in a pandemic. Int J Environ Res Public Health 2022 Jan 26;19(3):1361 [FREE Full text] [doi: 10.3390/ijerph19031361] [Medline: 35162384]
8. Priorities for local AIDS control efforts (PLACE) tool kit. MEASURE Evaluation. URL: https://www.measureevaluation.org/place.html [accessed 2024-04-29]
9. Weir S, Ssengooba F, Ssekamatte-Sebuliba J. PLACE in Uganda: monitoring AIDS-prevention programs in Kampala, Uganda using the PLACE method priorities for local AIDS control. The United States Agency for International Development (USAID). URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=dc046dd8d0ae6c2c1bb794268dcf724f8ec0f1e9 [accessed 2024-04-29]
10. Wambura M, Nyato DJ, Makyao N, Drake M, Kuringe E, Casalini C, et al. Programmatic mapping and size estimation of key populations to inform HIV programming in Tanzania. PLoS One 2020 Jan 30;15(1):e0228618 [FREE Full text] [doi: 10.1371/journal.pone.0228618] [Medline: 31999810]
11. Odek WO, Githuka GN, Avery L, Njoroge PK, Kasonde L, Gorgens M, et al. Estimating the size of the female sex worker population in Kenya to inform HIV prevention programming. PLoS One 2014 Mar 3;9(3):e89180 [FREE Full text] [doi: 10.1371/journal.pone.0089180] [Medline: 24595029]
12. Bunjaku DG, Deva E, Gashi L, Kaçaniku-Gunga P, Comins CA, Emmanuel F. Programmatic mapping to estimate size, distribution, and dynamics of key populations in Kosovo. JMIR Public Health Surveill 2019 Mar 05;5(1):e11194 [FREE Full text] [doi: 10.2196/11194] [Medline: 30835241]
13. Lo J, Nwafor SU, Schwitters AM, Mitchell A, Sebastian V, Stafford KA, et al. Key population hotspots in Nigeria for targeted HIV program planning: mapping, validation, and reconciliation. JMIR Public Health Surveill 2021 Feb 22;7(2):e25623 [FREE Full text] [doi: 10.2196/25623] [Medline: 33616537]
14. Cheuk E, Isac S, Musyoki H, Pickles M, Bhattacharjee P, Gichangi P, et al. Informing HIV prevention programs for adolescent girls and young women: a modified approach to programmatic mapping and key population size estimation. JMIR Public Health Surveill 2019 Apr 01;5(2):e11196 [FREE Full text] [doi: 10.2196/11196] [Medline: 30932868]

XSL•FO
RenderX

15. Anderson S, Ghys PD, Ombam R, Hallett TB. HIV prevention where it is needed most: comparison of strategies for the geographical allocation of interventions. J Int AIDS Soc 2017 Dec 08;20(4):e25020 [FREE Full text] [doi: 10.1002/jia2.25020] [Medline: 29220115]
16. Singh SK, Malviya A, Panakadan S, Tshering N, Rehman A, Vishwakarma D. Methodological issues in estimating size of key affected population: illustrations with data from South Asia. Demogr India 2018;47(1):47-60 [FREE Full text]
17. Zhao X, Zhou Y, Chen W, Li X, Li X, Li D. Mapping hourly population dynamics using remotely sensed and geospatial data: a case study in Beijing, China. GISci Remote Sens 2021 Jun 17;58(5):717-732. [doi: 10.1080/15481603.2021.1935128]
18. Hung CC, Banerjee S, Gilada I, Green K, Inoue Y, Kamarulzaman A, et al. Impact of COVID-19 on the HIV care continuum in Asia: insights from people living with HIV, key populations, and HIV healthcare providers. PLoS One 2022 Jul 20;17(7):e0270831 [FREE Full text] [doi: 10.1371/journal.pone.0270831] [Medline: 35857755]
19. Horváth D, Ásványi K, Cosovan A, Csordás T, Faludi J, Galla D, et al. Online only: future outlooks of post-pandemic education based on student experiences of the virtual university. Soc Econ 2022 Mar 04;44(1):2-21. [doi: 10.1556/204.2021.00026]
20. Towner E, Tomova L, Ladensack D, Chu K, Callaghan B. Virtual social interaction and loneliness among emerging adults amid the COVID-19 pandemic. Curr Res Ecol Soc Psychol 2022;3:100058 [FREE Full text] [doi: 10.1016/j.cresp.2022.100058] [Medline: 35937279]
21. Long E, Patterson S, Maxwell K, Blake C, Bosó Pérez R, Lewis R, et al. COVID-19 pandemic and its impact on social relationships and health. J Epidemiol Community Health 2022 Feb;76(2):128-132 [FREE Full text] [doi: 10.1136/jech-2021-216690] [Medline: 34413184]

Abbreviations

ART: antiretroviral therapy
FSW: female sex worker
KP: key population
MSM: men who have sex with men
NGO: nongovernmental organization
PLACE: Priorities for Local AIDS Control Efforts

Edited by A Mavragani; submitted 27.01.24; peer-reviewed by S Jameela, K Iwata, I Ezirim; comments to author 08.05.24; revised version received 22.07.24; accepted 24.12.24; published 30.01.25.

Please cite as:
Januraga PP, Lukitosari E, Luhukay L, Hasby R, Sutrisna A
Mapping Key Populations to Develop Improved HIV and AIDS Interventions: Multiphase Cross-Sectional Observational Mapping Study Using a District and City Approach
JMIR Public Health Surveill 2025;11:e56820
URL: https://publichealth.jmir.org/2025/1/e56820
doi:10.2196/56820
PMID:39883483

©Pande Putu Januraga, Endang Lukitosari, Lanny Luhukay, Rizky Hasby, Aang Sutrisna. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The PrEP Pharmacy Reach Study: Protocol for the Creation of Maps to Visualize the Impact of Expanding Access to HIV Prevention Services Through Pharmacies

Kristin R V Harrington[1], MD, PhD; Chante Hamilton[2], MPH; Daniel I Alohan[2], MPH; Alexis Hudson[2], MPH; Henry N Young[3], PhD; Natalie D Crawford[2], PhD

[1]Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton Rd NE, Atlanta, GA, United States

[2]Department of Behavioral, Social and Health Education Sciences, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[3]Department of Clinical and Administrative Pharmacy, College of Pharmacy, University of Georgia, Athens, GA, United States

Corresponding Author:
Kristin R V Harrington, MD, PhD
Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton Rd NE, Atlanta, GA, United States

Abstract

Background: Despite the proven efficacy of preexposure prophylaxis (PrEP) in reducing the risk of HIV transmission, uptake remains suboptimal among populations with limited access and availability to PrEP-prescribing locations, particularly in the Southern United States. The accessibility of pharmacies positions them as a promising resource for expanding PrEP delivery and access and supporting uptake and adherence through HIV prevention programs to reduce geographic disparities.

Objective: This study outlines a geospatial protocol to assess disparities in national PrEP availability. The aim of this study is to develop a protocol to identify, map, and describe the potential impact of expanding PrEP access and HIV prevention services to pharmacies as alternative PrEP delivery sites across the United States. We propose a reproducible mapping and analytic framework to visualize gaps and inform implementation strategies at the state and local levels.

Methods: We create local and state maps to help visualize the impact of expanding access to HIV prevention services through pharmacies. We obtain data from three main data sources: (1) pharmacy locations from the National Council for Prescription Drug Programs (NCPDP), (2) PrEP-prescribing facility locations from the CDC's National Prevention Information Network (NPIN), and (3) HIV case data per 100,000 persons from AIDSVu. We geocode pharmacies and PrEP-prescribing locations with Google application programming interfaces (APIs) via the *ggmap* package in R software. Maps are created by overlaying several different layers of general maps and aggregated data including base maps, choropleth maps, dot density maps. To quantitatively examine the potential impact of expanding HIV prevention services, we calculate a PrEP facility-to-need ratio (PfnR) as the total number of facilities (PrEP-prescribing or pharmacies) divided by the number of HIV cases per 100,000 persons. Fold-change estimates are computed to quantify the increase in service reach if pharmacies were incorporated into PrEP delivery.

Results: Overall, maps depicted far greater availability of local pharmacies compared to PrEP-prescribing facilities. When considered in the context of HIV cases per 100,000 persons, pharmacies were more prevalent than PrEP-prescribing facilities in areas with higher HIV caseloads. Mean PfnRs for pharmacies ranged from 0.04 (Mississippi) to 1.3 (Alaska), while PrEP-prescribing facility PfnRs were as low as 0.0004 (Puerto Rico). Estimated fold-change increases ranged from 6.4 in Idaho to 120.3 in Puerto Rico, with the greatest increases in the Southern and Midwestern United States.

Conclusions: This protocol provides a scalable and reproducible framework to assess PrEP service distribution and to identify areas which may benefit most from pharmacy integration. Our findings suggest that incorporating pharmacies into PrEP-delivery could substantially reduce geographic access barriers, especially in underserved regions. This work has critical implications for state and national policies focused on avenues to increase PrEP access and uptake and subsequently reduce HIV transmission in their regions.

(JMIR Public Health Surveill 2025;11:e75077) doi:10.2196/75077

KEYWORDS

pre-exposure prophylaxis; HIV prevention access; pharmacies; geospatial; spatial analysis

XSL•FO
RenderX

Introduction

There are over 1.1 million people living with HIV in the United States, with almost 40,000 new cases diagnosed in 2022 [1]. Preexposure prophylaxis, or PrEP, significantly reduces the risk of HIV transmission, with an effectiveness of over 99% when taken consistently [2-4]. Inequities in PrEP uptake span across populations, risk groups and geographic regions. The overall rate of PrEP use in 2021 was 129 per 100,000 persons with the lowest rate of PrEP use among Latino populations and similar rates of use among Black and White populations [5]. People who inject drugs have a significantly increased risk of HIV transmission and consistently have some of the lowest rates of PrEP use compared to other key populations [6]. Further, although most men who have sex with men (MSM) are aware of PrEP, large discrepancies remain in reported PrEP use. For example, despite a greater need, about 15% fewer Black MSM report PrEP use compared to white MSM [7]. Although the Southern United States comprises over half of new HIV diagnoses and almost half of all HIV cases [1], PrEP use is relatively stagnant in this region, reflecting structural, socioeconomic, and health care access barriers leading to difficulties in access and availability of PrEP [6].

The Centers for Disease Control and Prevention, National Institutes of Health, and other public health organizations have recently emphasized the role of pharmacies in achieving the goals of the Ending the HIV Epidemic Initiative (EHE) [8,9]. EHE aims to reduce the number of new HIV infections in the United States by 90% by 2030, with a focus on priority jurisdictions with >50% of new HIV diagnoses [10]. Community pharmacies provide convenient and accessible care for the majority of Americans, with approximately 89% of the population living within 5 miles of one [11-13]. Their extended operating hours, including evenings, weekends, and holidays make them an ideal option for individuals who cannot visit clinics during standard hours [14]. This accessibility positions pharmacies as a vital resource for expanding PrEP access and supporting adherence, advancing public health efforts nationwide but especially in communities with fewer healthcare facilities [15,16]. Pharmacists have demonstrated their capability in supporting the provision of PrEP services from initiation to ongoing management [15,17,18]. Further, previous research highlights the potential of pharmacies as accessible sites for delivering HIV-related services, supported by a body of systematic literature reviews that examine the integration of HIV prevention resources, such as PrEP into pharmacy practice [19-22].

Inequities in access and availability of services within public health often display geographic patterns. given local differences in epidemiology, policies, and resources. Spatial mapping of disease patterns has the ability to identify high-risk regions and inform evidence-based public health interventions [23,24]. The use of maps in policy and decision-making offers substantial benefits, by enabling the identification of underserved regions, guiding the strategic allocation of resources and uncovering patterns and disparities that may not be readily apparent in other data representations [11,25].

Geospatial analyses have been instrumental in identifying inequities in the availability of HIV prevention services in areas with high HIV burden. Previous work has highlighted the limited geographic availability of PrEP clinics, particularly in the Southern United States where HIV prevalence is high, revealing significant gaps in service coverage for regions with substantial unmet need [26]. Other studies have demonstrated the utility of AIDSVu, a publicly available database and interactive mapping tool, in visualizing geographic inequities in HIV prevention and care, providing a useful tool for identifying areas where interventions are needed most [25]. Building on this foundation, recent work has demonstrated the potential of pharmacies to bridge these gaps by mapping their locations alongside PrEP prescribing sites in underserved communities [27].

This study builds on previous research by expanding mapping of pharmacies offering HIV prevention services, specifically PrEP, across the United States. By combining geographic data with key indicators such as HIV prevalence and socioeconomic factors, this study assesses the potential reach of pharmacies in addressing PrEP. This approach identifies geographic gaps in PrEP access, offering valuable information to inform the development of targeted strategies to address these disparities. We aim to identify, map, and describe the potential impact of expanding PrEP access and HIV prevention services to pharmacies across the United States.

The findings from this study have the potential to guide local and state-level policy initiatives aimed at enhancing PrEP accessibility through pharmacies. Additionally, this research may serve as a model for similar efforts in other areas of health care delivery, demonstrating how geographic data and health indicators can be used to optimize resource allocation. Ultimately, this work has broader implications for improving health equity by addressing inequities in HIV prevention and ensuring that at-risk populations have access to necessary healthcare services.

Methods

Overview

This paper outlines a protocol for the creation of maps to help visualize the impact of expanding access to HIV prevention services through community pharmacies. These maps allow for the examination of geographic distributions of current PrEP-prescribing locations compared with pharmacy locations. We further describe the creation of PrEP facility-to-need ratios to quantify the potential reach of pharmacies to expand PrEP access. Similar implementation of this protocol has been described previously [27].

Data Sources

We used three main data sources for the creation of our maps. Pharmacy locations for pharmacies across the United States were obtained from the National Council for Prescription Drug Programs (NCPDP) [28]. This dataset included pharmacy street address, contact information, provider type, accessibility services, and levels of delivery services provided. PrEP-prescribing locations were obtained from the National

XSL•FO
RenderX

Prevention Information Network (NPIN) [29]. This platform, managed by the Centers for Disease Control and Prevention (CDC), provided aggregated street addresses and latitude and longitude coordinates for PrEP-prescribing locations. PrEP-prescribing locations include any health care facility in which a licensed provider can prescribe PrEP, which may be a community health center, primary care clinic, or a specialized clinic. These facilities can provide general and specialized care services depending on their capacity and level of specialization of providers. Data on HIV cases per 100,000 persons were collected at the state, county, and zip code levels from AIDSVu for states, counties, and cities [30]. AIDSVu is a publicly available database and interactive mapping tool that integrates data from the CDC and local health departments [31]. Website users may explore the HIV epidemic using multiple outcome measures at different spatial scales. Datasets are freely available for download for different outcomes at various levels of spatial granularity.

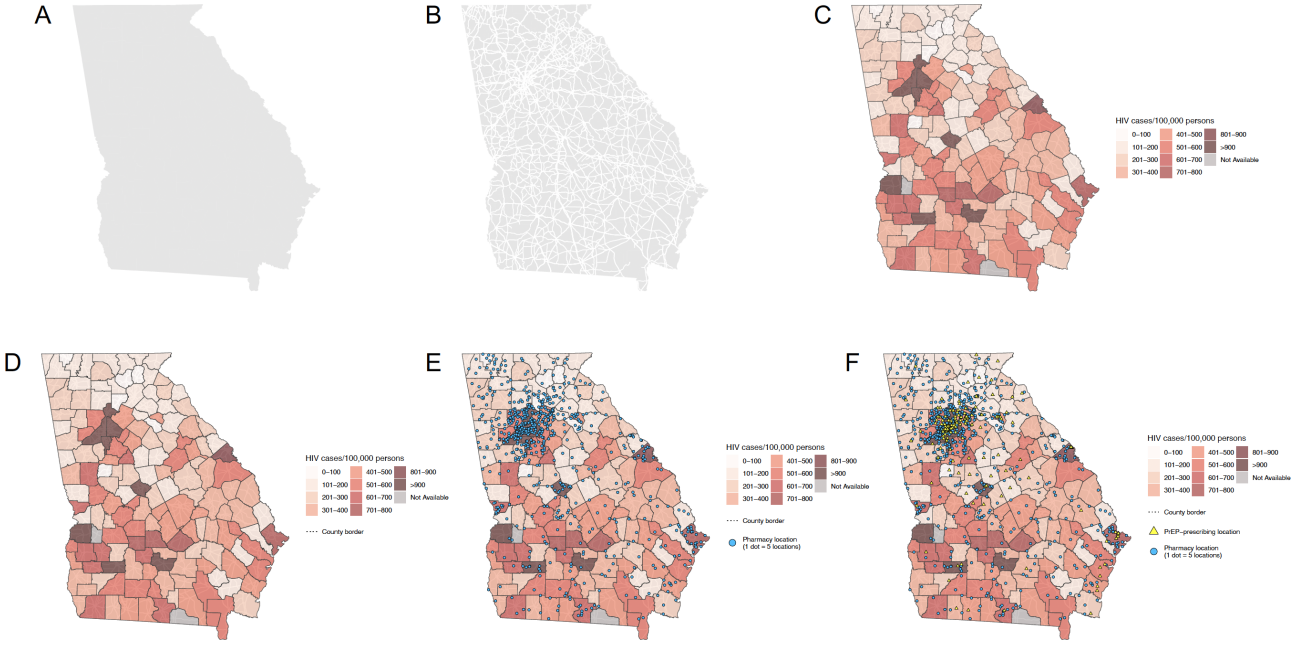
Geocoding

To geocode pharmacies and PrEP-prescribing locations, we used Google application programming interfaces (APIs) via the *ggmap* package in R software (version 4.4.3; R Foundation for Statistical Computing) [32]. This application converts street addresses to latitude and longitude coordinates that can then be plotted on maps. State, county, and zip code boundary files were gathered from the US Census Bureau using the R package *tigris* (version 1.5) [33]. Street addresses for all locations were geocoded to latitude and longitude coordinates. Quality checks were implemented at this stage, including plotting points and confirming coordinates were within correct state and county boundaries. For locations outside of expected boundaries, locations were manually geocoded using provided address.

Map Creation

Maps were created by overlaying several different layers of general maps and aggregated data. Base maps including landmass areas and roads were obtained by using functions from the *sf* package in R [34,35]. Sequential steps in the creation of state and city maps are displayed in Figures 1 and 2, respectively. First, landmass areas were plotted and overlaid with major roadways (Figures 1A, B, 2A). Next, a choropleth map was overlaid displaying the number of HIV cases per 100,000 persons (Figures 1C and 2B). Choropleth maps of HIV cases per 100,000 persons were indexed to the state, county, and zip code levels for either state or city maps depending on the available level of granularity of data from AIDSVu. To better distinguish spatial boundaries, county borders were then layered on top for both states and cities, followed by additional boundaries for zip code borders for cities (Figures 1D, 2C and D). The next layer was a dot density plot for pharmacies, where one point corresponded to five pharmacy locations to better visualize locations given the generally large number of existing pharmacies (Figures 1E and 2E). This was subsequently followed by a point layer depicting the geocoded PrEP-prescribing locations, where each point represented a specific location (Figures 1F and 2F). Last, city labels were applied for further spatial context (Figures 3 and 4).

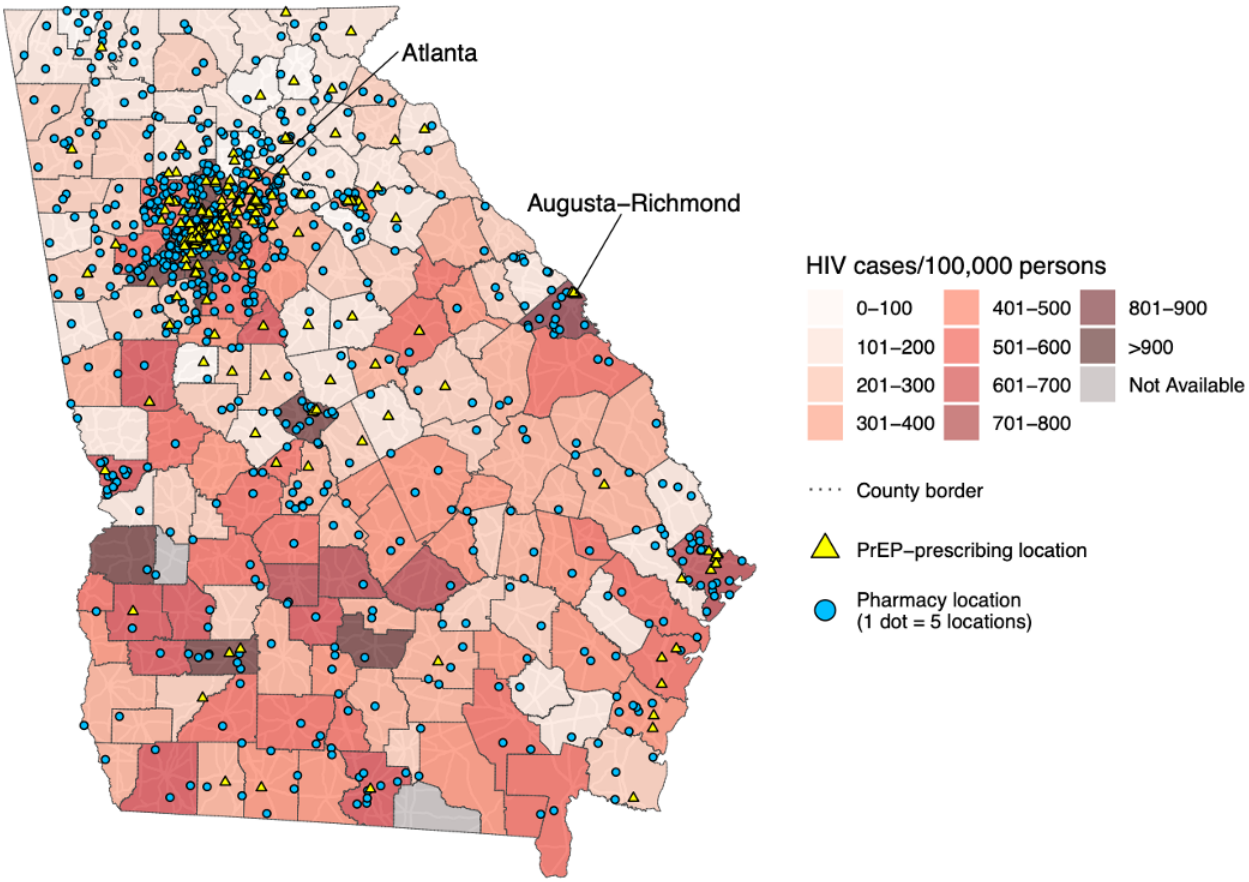
Figure 1. Sequential process of the creation of a state map.



XSL•FO
RenderX

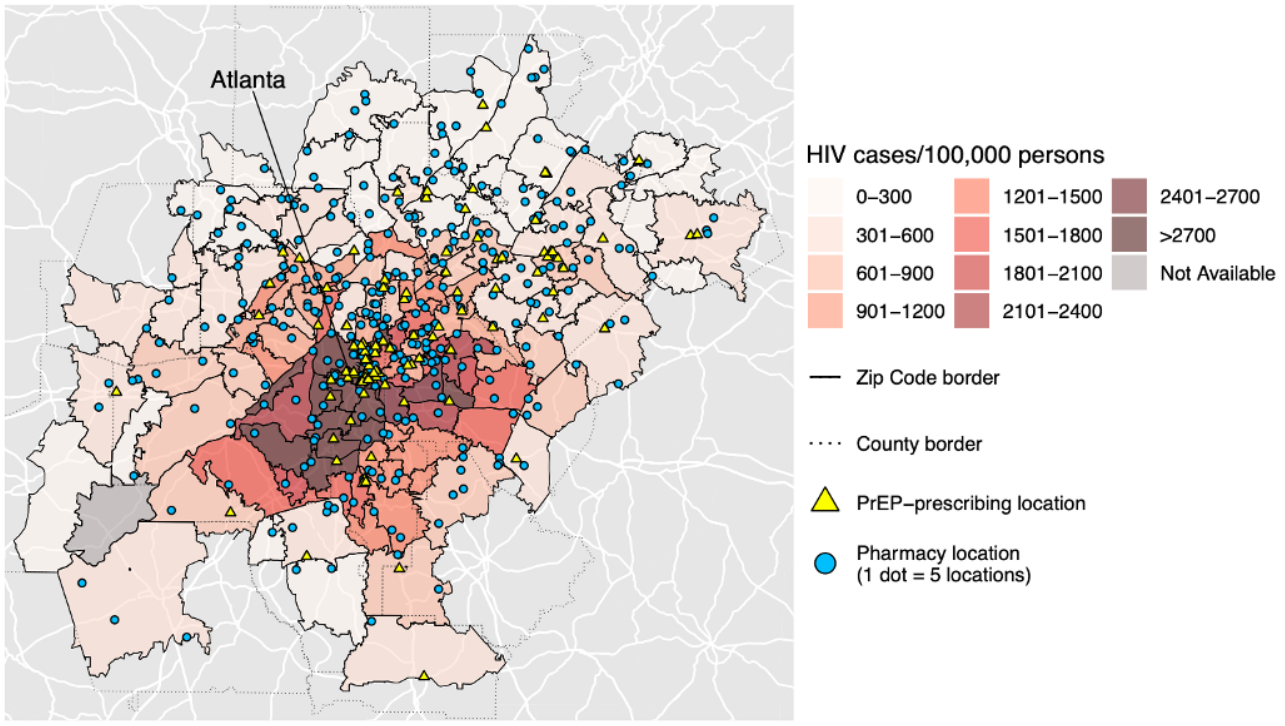
Figure 2. Sequential process of the creation of a city map.

Figure 3. Completed state map of Georgia. PrEP: .



XSL•FO
RenderX

Figure 4. Completed city map of Atlanta. PrEP



Statistical Analysis

To quantitatively examine the potential impact of expanding HIV prevention services, we calculated a PrEP facility-to-need ratio (PfnR) as the total number of facilities (PrEP-prescribing or pharmacies) divided by the number of HIV cases per 100,000 persons. Lower PfnRs indicate lower capacity to meet the needs of the HIV risk population. Additionally, we calculated the fold-change difference between PfnR estimates between PrEP-prescribing facilities and pharmacies within the same spatial area. Higher values indicate greater potential impact of expansion of HIV prevention services to pharmacies.

Ethical Considerations

This study was conducted using publicly available and deidentified datasets. No individual-level protected health information was used, and no human participants were directly involved. Thus, institutional review board approval was not required.

Results

Maps were created for 50 states, 54 cities, and one US territory. In general, maps depicted far greater accessibility to local pharmacies compared to PrEP-prescribing facilities. When taken in the context of HIV cases per 100,000 persons, pharmacies were more prevalent than PrEP-prescribing facilities in areas with higher HIV caseloads. These trends persisted at state, county, and zip code levels across the majority of states. Some regions had more aptly placed PrEP-prescribing locations across their areas including areas with higher HIV caseloads, while there were some states that had more rural regions with a paucity of both pharmacies and PrEP-prescribing facilities.

Mean PfnRs for pharmacy locations at the state level ranged from 0.04 in Mississippi to 1.3 in Alaska, with mean PfnRs for PrEP-prescribing facilities ranging from 0.0004 in Puerto Rico to 0.19 in Alaska. Estimated fold-increase for expansion of HIV prevention services to pharmacies was the lowest in Idaho (6.4) and highest in Puerto Rico (120.3). Overall, fold-increases were greatest for Midwestern and Southern states, and lowest for Western and Northeastern states.

Discussion

This study presents a scalable and reproducible geospatial framework to identify and visualize disparities in PrEP access with a specific focus on evaluating the impact of incorporating pharmacies as a potential solution to improve geographic inequities. We found that pharmacies are significantly more prevalent than PrEP-prescribing facilities in almost all locations evaluated. The greatest impact of expanding HIV prevention services was observed in the Southern and Midwestern regions of the United States where HIV burden is disproportionately high. The calculated PfnRs and associated fold-change estimates highlight the potential for pharmacies to serve as key locations for expanding HIV prevention services.

Our work builds on prior studies demonstrating that PrEP availability often fails to align with population-level need [36], as well as recent literature emphasizing the logistical limitations of clinic-based PrEP delivery models [37,38]. States can address these limitations by leveraging the existing extensive network of community pharmacies to provide more equitable PrEP access [39]. Importantly, the PfnR used in this study serves as a novel indicator of PrEP delivery capacity relative to local HIV burden.

The ideal pharmacy-based PrEP delivery model would include both prescribing and ongoing management in collaboration with

XSL•FO
RenderX

a clinician, spanning baseline eligibility assessments, lab testing, medication administration, adherence counseling, and follow-up monitoring [40]. However, expansion of HIV prevention services to pharmacies across the United States requires navigating state-level regulatory heterogeneity [41]. For states in which pharmacists lack the authority or reimbursement mechanisms to provide PrEP, operationalizing the model described here may be more challenging. Future work could incorporate state-specific policy data to evaluate the readiness of implementation. Additionally, future studies may evaluate the differences in impact of expanding HIV prevention services to all available pharmacies versus a subset optimized to either location or availability of resources.

To our knowledge, this was the first study to spatially examine the distribution of pharmacies and PrEP-prescribing facilities within the context of HIV prevalence across the United States. This work helps identify the areas of greatest need for PrEP availability and HIV prevention services. Further, the methodology used in this study has the potential to be applied to a number of other health conditions, spanning infectious diseases and non-communicable diseases. Other demographic indicators may be substituted in for disease prevalence if examining risk factors.

We acknowledge that this work is subject to a number of limitations. First, maps were limited to the granularity and accuracy of the data provided. Thus, for some states (eg, Alaska), the number of HIV cases per 100,000 persons was unavailable at the county or zip code level, so one estimate was provided for the entire state. This limits the ability to develop interpretations at a level more granular than the state level. Further, these differences in spatial granularity across regions limit our ability to make comparisons across dissimilar spatial scales (ie, comparing a state PfnR to a county PfnR in a different state). Regarding accessibility of data, some states do not provide HIV case data for all counties, resulting in several counties and zip codes with estimates that are unavailable. This data may be withheld by states or censored by AIDSVu for many reasons including privacy concerns if only very few cases live in a spatial area, or uncertainty in counts leading to unstable estimates. Lastly, our estimates do not account for variation in state regulations around pharmacy-based PrEP delivery and thus, these differences may limit implementation feasibility in certain jurisdictions. Despite these limitations, this study protocol can provide guidance around resource allocation and inform more equitable and geographically tailored PrEP expansion strategies.

The findings from this study have important implications for both research and policy. By providing a data-driven framework to visualize PrEP access gaps and estimate the potential of pharmacy-based delivery, this protocol offers a practical tool for public health agencies, policymakers, and health systems to guide resource allocation. The contribution of this project will be a robust examination of pharmacy locations with respect to HIV caseload across the United States. By using NCPDP data, it is likely that we were able to capture all pharmacy locations in the country, resulting in more accurate estimates of the potential impact of expanding HIV prevention services to pharmacies. This work has critical implications for state and national policies focused on avenues to increase PrEP access and uptake and subsequently reduce HIV transmission in their regions.

Acknowledgments

We would like to acknowledge the National Institute of Mental Health for its support under award numbers R34MH119007 and R01MH132470.

Data Availability

The datasets analyzed in the study are publicly available or available via request from respective organizations described. Pharmacy location data were obtained from the National Council for Prescription Drug Programs [28]. PrEP-prescribing facility data were accessed via the CDC National Prevention Information Network [29]. Data on HIV cases were obtained from AIDSVu [31]. All data processing code used for geocoding and mapping can be made available upon reasonable request.

Authors' Contributions

Conceptualization - KRVH, NDC
Data curation - KRVH, CH, AH
Formal analysis - KRVH
Funding acquistion - NDC, HNY
Methodology - KRVH
Project administration - CH, AH
Visualization - KRVH
Writing - original draft - KRVH, CH
Writing - review & editing - CH, DIA, AH, HNY, NDC

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. HIV surveillance report: diagnoses, deaths, and prevalence of HIV in the United States and 6 territories and freely associated states, 2022. CDC Stacks. 2024. URL: https://stacks.cdc.gov/view/cdc/156509 [accessed 2024-12-03]
2. Jourdain H, de Gage SB, Desplas D, Dray-Spira R. Real-world effectiveness of pre-exposure prophylaxis in men at high risk of HIV infection in France: a nested case-control study. Lancet Public Health 2022 Jun;7(6):e529-e536. [doi: 10.1016/S2468-2667(22)00106-2] [Medline: 35660214]
3. Spinner CD, Boesecke C, Zink A, et al. HIV pre-exposure prophylaxis (PrEP): a review of current knowledge of oral systemic HIV PrEP in humans. Infection 2016 Apr;44(2):151-158. [doi: 10.1007/s15010-015-0850-2]
4. Riddell J, Amico KR, Mayer KH. HIV preexposure prophylaxis: a review. JAMA 2018 Mar 27;319(12):1261-1268. [doi: 10.1001/jama.2018.1917] [Medline: 29584848]
5. Sullivan PS, DuBose SN, Castel AD, et al. Equity of PrEP uptake by race, ethnicity, sex and region in the United States in the first decade of PrEP: a population-based analysis. The Lancet Regional Health - Americas 2024 May;33:100738. [doi: 10.1016/j.lana.2024.100738]
6. Kamitani E, Johnson WD, Wichser ME, Adegbite AH, Mullins MM, Sipe TA. Growth in proportion and disparities of HIV PrEP use among key populations identified in the united states national goals: systematic review and meta-analysis of published surveys. J Acquir Immune Defic Syndr 2020;84(4):379-386. [doi: 10.1097/QAI.0000000000002345]
7. Finlayson T, Cha S, Xia M, et al. Changes in HIV preexposure prophylaxis awareness and use among men who have sex with men — 20 urban areas, 2014 and 2017. MMWR Morb Mortal Wkly Rep 2022 Jul;68(27):597-603. [doi: 10.15585/mmwr.mm6827a1]
8. McCree DH, Byrd KK, Johnston M, Gaines M, Weidle PJ. Roles for pharmacists in the “Ending the HIV Epidemic: A Plan for America” initiative. Public Health Rep 2020;135(5):547-554. [doi: 10.1177/0033354920941184] [Medline: 32780671]
9. Pharmacists and Pharmacies Play a Vital Role in Helping to End the HIV Epidemic | HIV.gov. URL: https://www.hiv.gov/blog/pharmacists-play-important-role-on-national-hiv-testing-day [accessed 2024-12-03]
10. EHE overview. HIV.gov. URL: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview [accessed 2024-12-03]
11. Qato DM, Zenk S, Wilder J, Harrington R, Gaskin D, Alexander GC. The availability of pharmacies in the United States: 2007-2015. PLoS One 2017;12(8):e0183172. [doi: 10.1371/journal.pone.0183172] [Medline: 28813473]
12. Sharareh N, Zheutlin AR, Qato DM, Guadamuz J, Bress A, Vos RO. Access to community pharmacies based on drive time and by rurality across the contiguous United States. J Am Pharm Assoc (2003) 2024;64(2):476-482. [doi: 10.1016/j.japh.2024.01.004] [Medline: 38215823]
13. Berenbrok LA, Tang S, Gabriel N, et al. Access to community pharmacies: a nationwide geographic information systems cross-sectional analysis. J Am Pharm Assoc (2003) 2022 Nov;62(6):1816-1822. [doi: 10.1016/j.japh.2022.07.003]
14. National association of chain drug stores, “Re: health care workshop, project no.P131207. supplemental comments. : National Association of Chain Drug Stores (NACDS); 2014 Apr URL: https://www.nacds.org/ceo/2014/0508/supplemental_comments.pdf [accessed 2024-12-03]
15. Havens JP, Scarsi KK, Sayles H, Klepser DG, Swindells S, Bares SH. Acceptability and feasibility of a pharmacist-led HIV pre-exposure prophylaxis (PrEP) program in the Midwestern United States. Open Forum Infect Dis 2019 Oct 1;6(10):fz365. [doi: 10.1093/ofid/ofz365] [Medline: 31412131]
16. Farmer EK, Koren DE, Cha A, Grossman K, Cates DW. The pharmacist’s expanding role in HIV pre-exposure prophylaxis. AIDS Patient Care STDS 2019 May;33(5):207-213. [doi: 10.1089/apc.2018.0294] [Medline: 31067124]
17. Collaborative practice agreements and pharmacists’ patient care services: a resource for pharmacists. CDC Stacks. URL: https://stacks.cdc.gov/view/cdc/49016 [accessed 2024-12-03]
18. Tung E, Thomas A, Eichner A, Shalit P. Feasibility of a pharmacist-run HIV prep clinic in a community pharmacy setting. Presented at: Conference on Retroviruses and Opportunistic Infections; Feb 13-16, 2017; Seattle, Washington URL: https://www.croiconference.org/abstract/feasibility-pharmacist-run-hiv-prep-clinic-community-pharmacy-setting/ [accessed 2021-09-08]
19. Zhao A, Dangerfield DT II, Nunn A, et al. Pharmacy-based interventions to increase use of HIV pre-exposure prophylaxis in the United States: a scoping review. AIDS Behav 2022 May;26(5):1377-1392. [doi: 10.1007/s10461-021-03494-4]
20. Ahmed A, Dujaili JA, Hashmi FK, Awaisu A, Chaiyakunapruk N, Hasan SS. The economic impact of pharmacist care for people living with HIV/AIDS: a systematic review. Explor Res Clin Soc Pharm 2021 Sep;3:100066. [doi: 10.1016/j.rcsop.2021.100066] [Medline: 35480597]
21. Crawford ND, Myers S, Young H, Klepser D, Tung E. The role of pharmacies in the HIV prevention and care continuums: a systematic review. AIDS Behav 2021 Jun;25(6):1819-1828. [doi: 10.1007/s10461-020-03111-w] [Medline: 33386509]
22. Kennedy CE, Yeh PT, Atkins K, Ferguson L, Baggaley R, Narasimhan M. PrEP distribution in pharmacies: a systematic review. BMJ Open 2022 Feb 21;12(2):e054121. [doi: 10.1136/bmjopen-2021-054121] [Medline: 35190430]
23. GIS and public health at CDC. CDC Restored. URL: https://www.restoredcdc.org/www.cdc.gov/gis/php/index.html [accessed 2024-12-03]
24. Kirby RS, Delmelle E, Eberth JM. Advances in spatial epidemiology and geographic information systems. Ann Epidemiol 2017 Jan;27(1):1-9. [doi: 10.1016/j.annepidem.2016.12.001] [Medline: 28081893]

XSL•FO
RenderX

25. Sullivan PS, Woodyatt C, Koski C, et al. A data visualization and dissemination resource to support HIV prevention and care at the local level: analysis and uses of the AIDSVu public data resource. J Med Internet Res 2020 Oct;22(10):e23173. [doi: 10.2196/23173]
26. Siegler AJ, Bratcher A, Weiss KM. Geographic access to preexposure prophylaxis clinics among men who have sex with men in the United States. Am J Public Health 2019 Sep;109(9):1216-1223. [doi: 10.2105/AJPH.2019.305172] [Medline: 31318587]
27. Harrington KRV, Chandra C, Alohan DI, et al. Examination of HIV preexposure prophylaxis need, availability, and potential pharmacy integration in the Southeastern US. JAMA Netw Open 2023 Jul 3;6(7):e2326028. [doi: 10.1001/jamanetworkopen.2023.26028] [Medline: 37498599]
28. NCPDP. URL: https://www.ncpdp.org/ [accessed 2024-12-03]
29. National Prevention Information Network. URL: https://npin.cdc.gov/ [accessed 2024-12-03]
30. Sullivan PS, Woodyatt C, Koski C, et al. A data visualization and dissemination resource to support HIV prevention and care at the local level: analysis and uses of the AIDSVu public data resource. J Med Internet Res 2020;22(10):e23173. [doi: 10.2196/23173]
31. AIDSVu. Emory University and Rollins School of Public Health. URL: https://aidsvu.org/ [accessed 2025-01-23]
32. Kahle D, Wickham H. ggmap: spatial visualization with ggplot2. R J 2013;5(1):144. [doi: 10.32614/RJ-2013-014]
33. Walker K. tigris: An R package to access and work with geographic data from the US Census Bureau. R J 2016;8(2):231. [doi: 10.32614/RJ-2016-043]
34. Pebesma E, Bivand R. Spatial Data Science With Applications in R 2023:1-300. [doi: 10.1201/9780429459016]
35. Pebesma E. Simple features for R: standardized support for spatial vector data. R J 2018 Jul;10(1):439. [doi: 10.32614/RJ-2018-009]
36. Weiss KM, Prasad P, Sanchez T, Goodreau SM, Jenness SM. Association between HIV PrEP indications and use in a national sexual network study of US men who have sex with men. J Int AIDS Soc 2021 Oct;24(10):e25826. [doi: 10.1002/jia2.25826] [Medline: 34605174]
37. Tran J, Mishra A, Zimmermann M, Hansen R. Bridging PrEP access gaps: mapping geospatial accessibility across the United States and leveraging community pharmacies for expansion. J Am Pharm Assoc (2003) 2025;65(2):102274. [doi: 10.1016/j.japh.2024.102274] [Medline: 39461720]
38. Kamitani E, Mizuno Y, DeLuca JB, Collins CB Jr. Systematic review of alternative HIV preexposure prophylaxis care delivery models to improve preexposure prophylaxis services. AIDS 2023 Aug 1;37(10):1593-1602. [doi: 10.1097/QAD.0000000000003601] [Medline: 37199602]
39. Tharumia Jagadeesan C, Wirtz VJ. Geographical accessibility of medicines: a systematic literature review of pharmacy mapping. J of Pharm Policy and Pract 2021 Dec;14(1):1-13. [doi: 10.1186/s40545-020-00291-7]
40. Tun W, Kopeka M, Conserve DF, et al. Development of a pharmacy-based HIV PrEP service delivery intervention for Washington, District of Columbia (DC): a study protocol. PLoS ONE 2024 Oct;19(10):e0311694. [doi: 10.1371/journal.pone.0311694]
41. Le T, Kelly T. State policies on pharmacist-initiated PREP and PREP usage. J Am Pharm Assoc (2003) 2025;65(5):102415. [doi: 10.1016/j.japh.2025.102415] [Medline: 40348187]

Abbreviations

API: application programming interfaces
CDC: Centers for Disease Control and Prevention
EHE: Ending the HIV Epidemic Initiative
MSM: who have sex with men
NCPDP: National Council for Prescription Drug Programs
NPIN: National Prevention Information Network
PfnR: PrEP facility-to-need ratio
PrEP: preexposure prophylaxis

Edited by A Mavragani; submitted 27.03.25; peer-reviewed by K Ortblad, N Sharareh; revised version received 12.09.25; accepted 30.09.25; published 27.10.25.

Please cite as:
Harrington KRV, Hamilton C, Alohan DI, Hudson A, Young HN, Crawford ND
The PrEP Pharmacy Reach Study: Protocol for the Creation of Maps to Visualize the Impact of Expanding Access to HIV Prevention Services Through Pharmacies
JMIR Public Health Surveill 2025;11:e75077
URL: https://publichealth.jmir.org/2025/1/e75077
doi:10.2196/75077

XSL•FO
RenderX

© Kristin R V Harrington, Chante Hamilton, Daniel I Alohan, Alexis Hudson, Henry N Young, Natalie D Crawford. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

An Enhanced Social Network Strategy to Increase the Uptake of HIV Services: Protocol for Type I Hybrid Implementation Study (Carolinas RESPOND)

Ian Dale[1], MPH; Meagan Zarwell[2], PhD; Alicia Diggs[3], MPH; Jesse Strunk Elkins[2], MPH; Eunice Okumu[3], MPH; Sharon Weissman[4], MD; Feng-Chang Lin[5], PhD; Victoria Mobley[6], MD, MPH; Carol Golin[3], MD; Abby Rudolph[7], PhD; Ann M Dennis[8], MD, MS

[1]Institute for Global Health and Infectious Diseases, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]Department of Epidemiology and Community Health, College of Health and Human Services, University of North Carolina at Charlotte, Charlotte, NC, United States

[3]Center for AIDS Research, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[4]School of Medicine, University of South Carolina, Columbia, SC, United States

[5]Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[6]Division of Public Health, Communicable Disease Branch, NC Department of Health and Human Services, Raleigh, NC, United States

[7]College of Public Health, Temple University, Philadelphia, PA, United States

[8]School of Medicine, University of North Carolina at Chapel Hill, 130 Mason Farm Rd, Chapel Hill, NC, United States

Corresponding Author:
Ann M Dennis, MD, MS
School of Medicine, University of North Carolina at Chapel Hill, 130 Mason Farm Rd, Chapel Hill, NC, United States

Abstract

Background: In the United States, persisting new HIV diagnoses among gay, bisexual, and other cisgender men who have sex with men (GBMSM) and transgender women make it unlikely that the United States will meet the Ending the HIV Epidemic's (EHE) goal to reduce new HIV diagnoses by 90% by 2030. Innovative strategies are needed to address this challenge, particularly in the US South, where Black and Latinx GBMSM and transgender women are disproportionately impacted by HIV. Social network approaches have led to increased HIV testing uptake. Social network interventions that are responsive to individuals' needs among disproportionately impacted groups could also increase engagement across the HIV prevention and care continuum.

Objective: This hybrid type 1 effectiveness-implementation study will evaluate an enhanced social network strategy (eSNS) intervention designed to increase engagement in HIV services (HIV testing, pre-exposure prophylaxis [PrEP] use, and HIV care) by groups disproportionately affected by HIV. From 2025 to 2027, eSNS will be delivered in the Charlotte, North Carolina (NC) region, which includes Mecklenburg County, a priority EHE jurisdiction.

Methods: The study's phase 1 was a formative period of mixed methods data collection to operationalize enhancements to the Centers for Disease Control and Prevention's social network strategy (SNS). In Phase 2, the intervention will be integrated into standard NC Partner Services for people diagnosed with HIV and their sexual or social contacts, which is routinely performed by disease intervention specialists (DISs). We will identify network recruiters (ambassadors) who are 18 years and older and are either reached by study team DIS (DIS coaches) performing partner services or referred at community sites. Over 2 - 6 weeks, DIS coaches will guide ambassadors to identify and refer people in their network (peers) for HIV services and will facilitate peers' referrals to HIV services. Finally, Phase 3 will evaluate the eSNS's effectiveness in increasing HIV services uptake compared to standard-of-care partner services in the Raleigh, NC region.

Results: This project was funded by the National Institutes of Health and initially approved by the University of North Carolina at Chapel Hill's Institutional Review Board in 2022. Phase 1 concluded in August 2024. Implementation of eSNS (Phase 2) was launched in March 2025. Based on phase 1 findings, the study was modified to include Ambassadors of any race or ethnicity and gender (originally only Black GBMSM and transgender women) and expand identification of ambassadors through community sites (in addition to partner services).

Conclusions: Substantial reductions in new HIV diagnoses depend on public health approaches that effectively reach people with a higher likelihood of acquiring HIV. Our protocol proposes integrating existing strategies with an innovative intervention (eSNS) to reduce social barriers to disproportionately affected groups' engagement in the full HIV prevention and care continuum.

(JMIR Public Health Surveill 2025;11:e69495) doi:10.2196/69495

XSL•FO
RenderX

KEYWORDS

social networks; HIV-1; public health surveillance; clinical research protocol; implementation science

Introduction

Innovative public health strategies are needed to increase engagement in HIV prevention and care services in the United States, and thereby curb onward HIV transmission. Gay, bisexual, and other cisgender men who have sex with men (GBMSM) and transgender women, particularly Black and Latin individuals residing in the US South, are disproportionately impacted by HIV due to structural barriers to health care and their position in networks with a higher prevalence of HIV [1]. In 2022, 52% of all HIV diagnoses in the United States occurred in the South; of those, 47% were people who identify as Black or African American and 26% among people who identify as Hispanic or Latino, despite representing only 19% and 18% of the region's population, respectively [2]. Cultural and structural barriers such as stigma, medical mistrust, and geographic distance further hinder engagement with care and prevention services for Black and Latinx GBMSM and transgender women [3-5].

Partner services is a longstanding strategy conducted by some local and state health departments to reduce HIV transmission, usually by deploying disease intervention specialists (DISs) to offer testing and linkage to care or prevention services to sex or injecting partners of people newly diagnosed or reported with HIV [6]. While there is evidence that partner services yield a relatively high proportion of new HIV diagnoses among identified partners [7-9], including when compared to interventions where people newly diagnosed with HIV notify partners themselves [10], reaching sex or injecting partners through partner services can be challenging. Historic increases in DIS caseloads and clients unable to recall information about sex partners met on dating apps have reduced the feasibility of partner notification [11], while health departments may not always be able to interview people with new HIV diagnoses to request partners' names or conduct in-person field investigations [12], which have better partner outcomes compared to telephone-based partner services [13]. In addition, partner services frequently fail to identify sources of infection (ie, transmitting partner) [14-16]. Given the disproportionate and concentrated burden of HIV among Black and Latinx GBMSM and transgender women, there is an urgent need for interventions that better meet these populations' specific needs.

Strategies that leverage existing social connections, thereby reducing stigma and increasing trust, may be highly impactful among GBMSM and transgender women, where sexual networks that are disproportionately burdened by HIV are also frequently homogeneous by race and gender [17]. The Centers for Disease Control and Prevention (CDC)'s social network strategy (SNS) has been effective at increasing HIV testing among populations disproportionately affected by HIV [18-22]. The SNS program coaches people from high-incidence social or sexual networks to recruit their peers (eg, friends, not only sex or injecting partners) to engage in HIV testing. By leveraging trust among members of these social networks, SNS operates on the principle that people are more likely to respond positively to health behaviors and messages about engaging in HIV services from people in their social circle rather than unfamiliar people, such as DIS or other public health workers. By supplementing existing partner services programs with SNS, we hypothesize that a combined partner services-SNS intervention package could reach more individuals needing HIV services by implementing it among social networks with recent HIV transmission. A focus on networks involving recent HIV transmission (ie, acute or early HIV) may be optimal for promoting early antiretroviral therapy (ART) and uptake of pre-exposure prophylaxis (PrEP) among network members, thereby reducing further transmission. In North Carolina, individuals diagnosed with acute HIV infection reported having more sexual partners than those diagnosed with chronic HIV, providing an opportunity to intervene in networks with high probability of HIV acquisition [23]. Combining contact network and HIV genotype data with social network strategies can further enhance our ability to identify and intervene in networks whose members have a higher probability of acquiring HIV, potentially preventing onward HIV transmission [14,15]. Molecular epidemiology uses HIV sequences derived from genotypes to identify clusters of highly similar sequences that indicate recent transmission [24,25]. An approach that integrates SNS, standard partner services, and molecular epidemiology may provide a "next generation" public health strategy that better connects disproportionately affected networks with HIV prevention and care services.

In this paper, we present the protocol for such an integrated approach: an enhanced social network strategy (eSNS) intervention, which is part of the Carolinas RESPOND study. Adapted from CDC's SNS, our eSNS enhances SNS in three key ways: (1) Focusing on engagement in PrEP and HIV care for social network members, not only on HIV testing, (2) integration with partner services, and (3) expanding potential networks to include HIV genetic cluster members identified by molecular epidemiology. From 2025 to 2027, we will implement eSNS in the Charlotte region of North Carolina, which includes Mecklenburg County and directly borders the state of South Carolina; both are regions prioritized by the Ending the HIV Epidemic (EHE) initiative [26]. Our eSNS will prioritize networks involving recent HIV transmission, which we anticipate will most frequently be comprised of Black and Latinx GBMSM and transgender women but may include participants of any race or gender (including nonbinary individuals) [27]. The intervention will be delivered alongside standard partner services to better direct eSNS toward networks with recent HIV diagnoses, particularly those with acute or recent HIV infection and clusters with rapid HIV transmission. The study will evaluate the effectiveness of eSNS to (1) increase HIV prevention and treatment uptake among these key populations and (2) identify HIV transmission networks connected via sexual or injecting relationships or characterized by closely related viral infections and ongoing HIV transmission. This protocol describes our formative data collection efforts, intervention procedures, evaluation framework, and key adaptations.

XSL•FO
RenderX

Methods

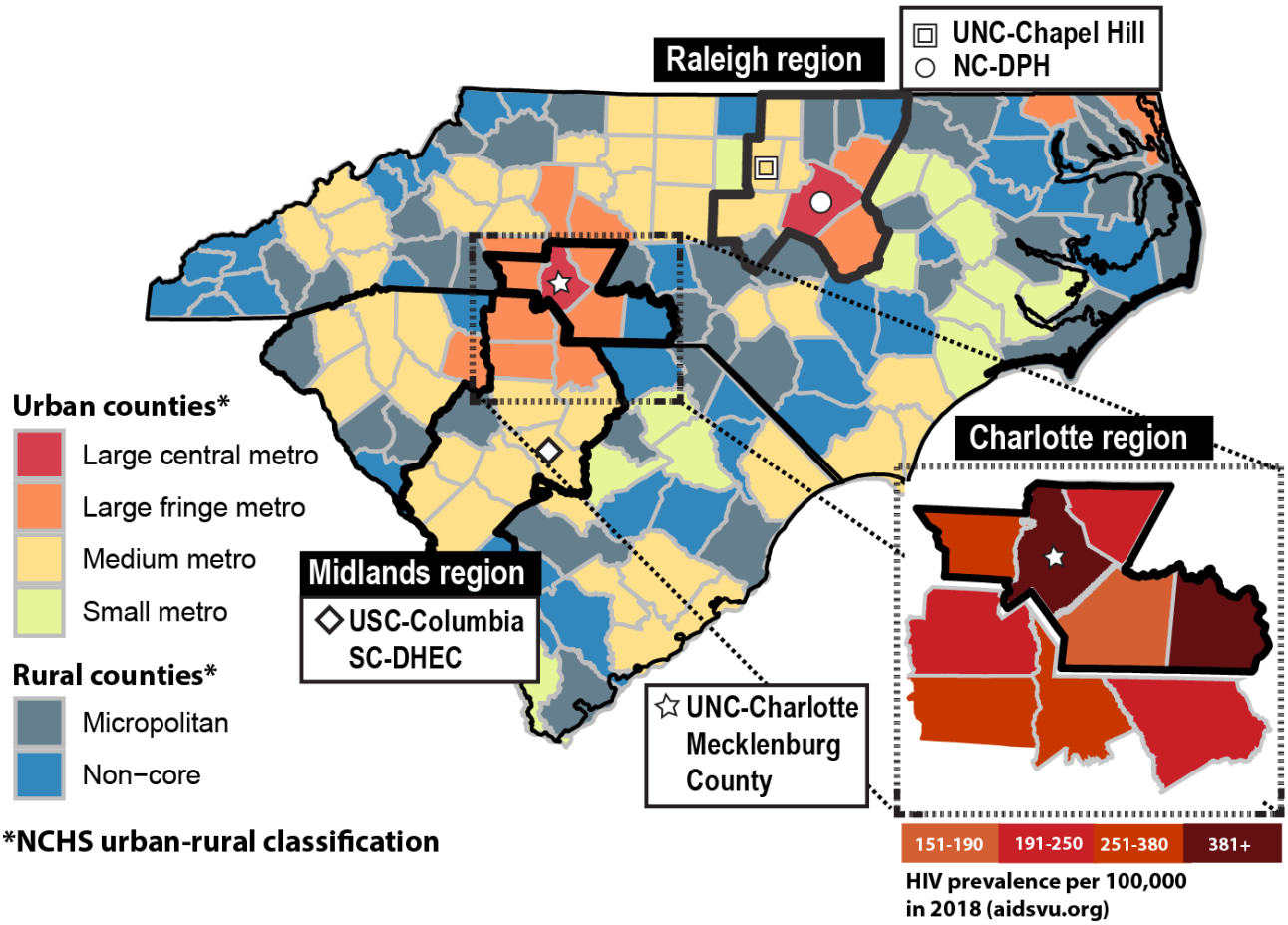
Study Aims

The Carolinas RESPOND study's aims are to: (1) assess how genetic sequence analysis added to standard surveillance impacts spatial and demographic features of HIV networks in juxtaposed high burden regions of North Carolina and South Carolina; (2) identify barriers and facilitators to engagement in HIV prevention and treatment services by individuals from communities disproportionately impacted by HIV who are part of networks with HIV recent transmission; and (3) evaluate an intensified partner services response using eSNS, which leverages HIV genetic cluster analysis-based network characterization. Applying genetic sequence analyses used in cluster detection and response will provide an additional tool to direct eSNS within networks experiencing recent HIV transmission. This protocol describes the plans for eSNS adaptation and implementation in the RESPOND study.

Study Setting

The Carolinas RESPOND study focuses on North Carolina and South Carolina, with collaboration among academic and public health partners in both states (see Figure 1). The eSNS intervention described in this protocol is planned for the Charlotte region, covering a 5-county area in the North Carolina Department and Health and Human Services' (DHHS) Field Services Region II, which includes the EHE priority jurisdiction [28] of Mecklenburg County (with star in Figure 1). This region directly borders the Midlands region of South Carolina, one of 7 US states with a disproportionate burden of HIV in rural areas [26]; the regions are connected by Interstate 77, which runs from Charlotte, North Carolina, to Columbia, South Carolina (see diamond in Figure 1). Although implementation is focused on the Charlotte region, our eSNS will follow networks that may cross both states because of Charlotte's geospatial proximity to South Carolina. Due to comparable demographics and HIV epidemiology [27], North Carolina DHHS's Field Services Region IV, the 11-county area surrounding the Raleigh, North Carolina metropolitan area, was chosen as a "control" region against which outcomes in the Charlotte, North Carolina "intervention" region will be compared.

Figure 1. Map of study setting. The enhanced social network strategy (eSNS) will be implemented in the 5-county area of North Carolina Department of Health and Human Services (DHHS) Field Services Region II (Charlotte region shown bordered in black in the inset at the right). North Carolina DHHS Field Services Region IV (Raleigh, 11 counties) is shown bordered. UNC: University of North Carolina; NC-DPH: North Carolina Division of Public Health; USC: University of South Carolina; SC-DHEC: South Carolina Department of Health and Environmental Control.



XSL•FO
RenderX

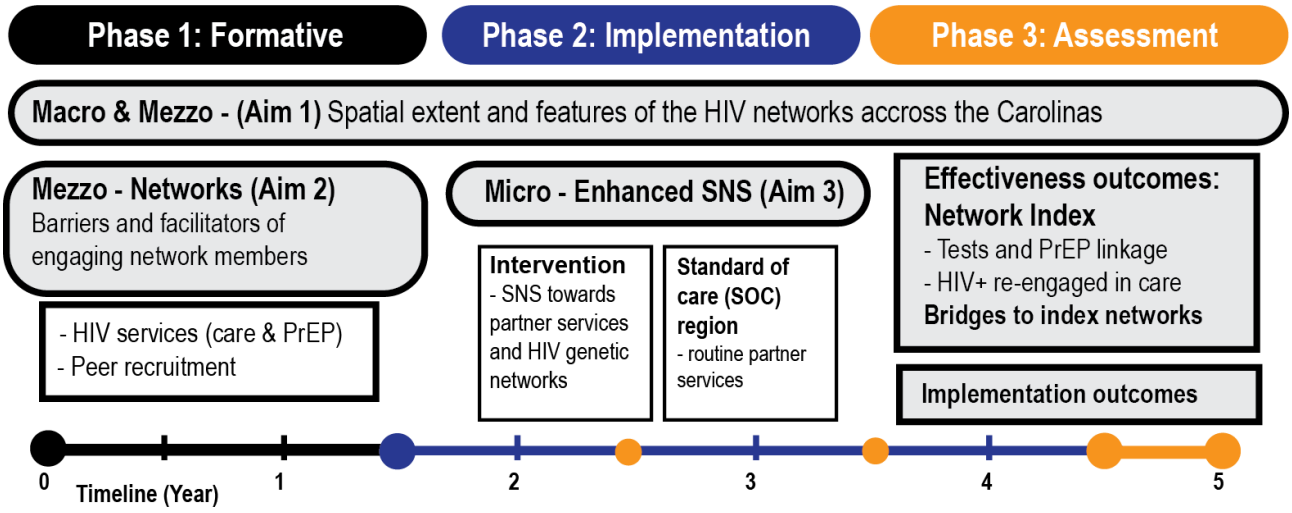
Theoretical Framework

Our study design is guided by the modified social ecological model (MSEM) [29]. This model examines the likelihood of HIV acquisition through multiple levels by situating individual behaviors within network, community, sociostructural, and epidemic contexts. Depending on network structures and underlying prevalence, individual behaviors may either facilitate or hinder HIV acquisition. Understanding behaviors that may protect one from or increase one's chances of acquiring HIV within the context of individuals' connections can inform more effective strategies to prevent ongoing transmission [29]. Within the MSEM model, we draw on the theory of how social networks impact health [30], wherein social networks are mezzo-level features that are shaped by macro-level social structural conditions and capable of influencing micro-level social and interpersonal behaviors. For example, individuals who are connected by sex or injecting partnerships within social networks characterized by a high HIV prevalence (or incidence), or who bridge to other networks by having sex or injecting partnerships with 1 or more individuals in distinct networks, may have an increased likelihood of acquiring HIV. At the behavioral level, networks can also have positive influences, such as the provision of social support, informational support, emotional support, social engagement and attachment, and access to resources and material goods. Just as HIV can be transmitted across networks, health information and behavior changes can similarly diffuse through networks [31,32]. Interventions that leverage existing networks of relationships can promote healthier behavior norms through social influence (eg, increased HIV testing and PrEP uptake) [33].

Study Design

We designed a 3-phase process (see Figure 2) to develop, implement, and evaluate eSNS using a quasi-experimental, nonrandomized hybrid type 1 effectiveness-implementation design [34]. This design allows for evaluation of both eSNS effectiveness and the intervention delivery process. Phase 1 (preimplementation), which involved establishing a community advisory group and collecting mixed methods data to inform adaptations to SNS procedures, concluded in August 2024. Analyses of Phase 1 data, which are ongoing, aim to understand the structural, spatial, and network factors that influence HIV services engagement among the study's priority groups, plus barriers and facilitators to eSNS implementation. In Phase 2, eSNS will be implemented in the Charlotte region over a 3-year period (2025 - 2027). In Phase 3, we will evaluate the impact of eSNS on HIV services use compared to standard-of-care partner services delivered in the Raleigh, North Carolina metropolitan region [27] (see Figure 1).

Figure 2. Schematic overview of the study design. SNS: social network strategy; PrEP: pre-exposure prophylaxis.

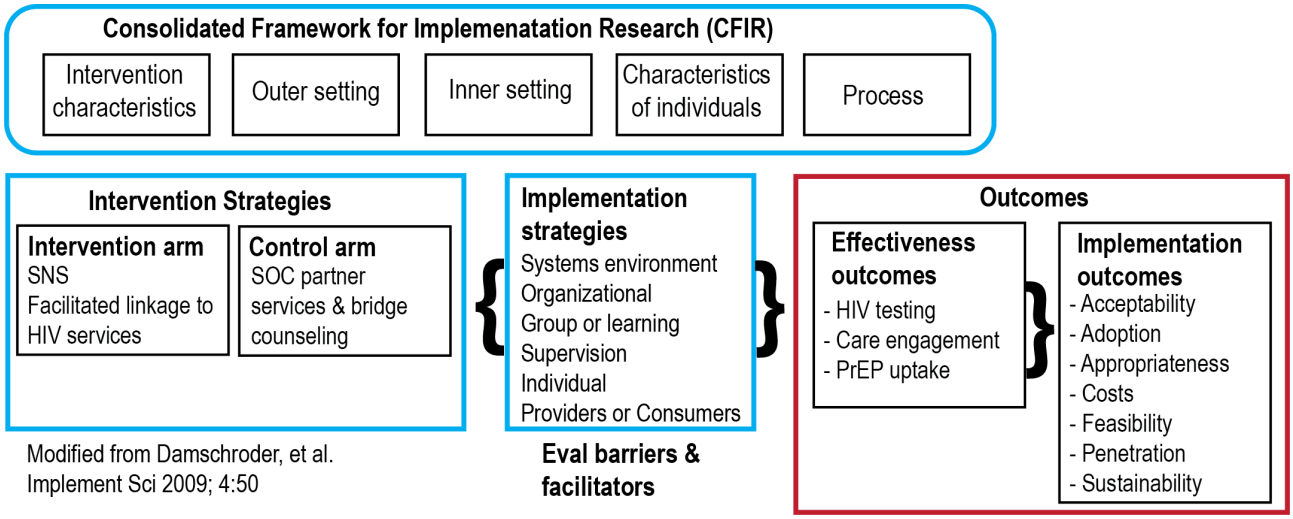


Implementation Research Logic Model

We use the Consolidated Framework for Implementation Research (CFIR) [35] to guide exploration of potential barriers and facilitators to eSNS implementation and to inform a strategy for regional scale-up and sustainability (see Figure 3). The CFIR offers a comprehensive and widely used framework for categorizing potential determinants of intervention implementation. We chose a hybrid type 1 design because it has the advantage of speeding translation of eSNS toward successful implementation by assessing implementation barriers and effectiveness simultaneously [34]. Integrating this design with the CFIR demonstrates where implementation determinants will be assessed in implementation planning.

XSL•FO
RenderX

Figure 3. Conceptual implementation science framework for evaluating enhanced social network strategy (eSNS). PrEP: pre-exposure prophylaxis.



Community Advisory Group

A community advisory group (CAG) led by a co-investigator (AD) with lived experience was first formed in August 2023. Members were initially invited to join from the Charlotte-area researchers' professional networks, community organizations and clinics, and community initiatives. The group meets quarterly, and attendees are provided an honorarium of US $100 after each meeting. We plan to maintain a roster of approximately 10 CAG members at any given time; during phase 2, interested eSNS participants will also be invited to join. Our CAG will be consulted throughout the study to share local resources, to advise about appropriate community engagement strategies and study participant compensation, and to provide feedback on the study's recruitment materials, participant engagement language, and social media presence.

eSNS Intervention Overview

The eSNS intervention will be implemented by DIS coaches toward priority networks, which are networks with recent HIV transmission involving key population members (including Black or African American, Latin GBMSM, and transgender women). DIS coaches are study team members responsible for delivering both the eSNS intervention and routine partner services (see Textbox 1). Priority networks will be identified through routine reporting of positive HIV tests to the North Carolina DHHS HIV surveillance system, which are then assigned for partner services by state and county DIS. The eSNS intervention will initially identify priority network members (index persons) by assigning DIS coaches to reach individuals aged 13 years and older (classified by the CDC as adults for HIV surveillance [36]). DIS coaches will be preferentially assigned to index persons diagnosed with acute HIV infection or less than 30 years old with chronic (antibody-positive) HIV, as younger persons are more likely to have recent HIV infection, but will also provide partner services to index persons who are virally unsuppressed and identified as part of an HIV transmission cluster (cluster member; see Figure 4).

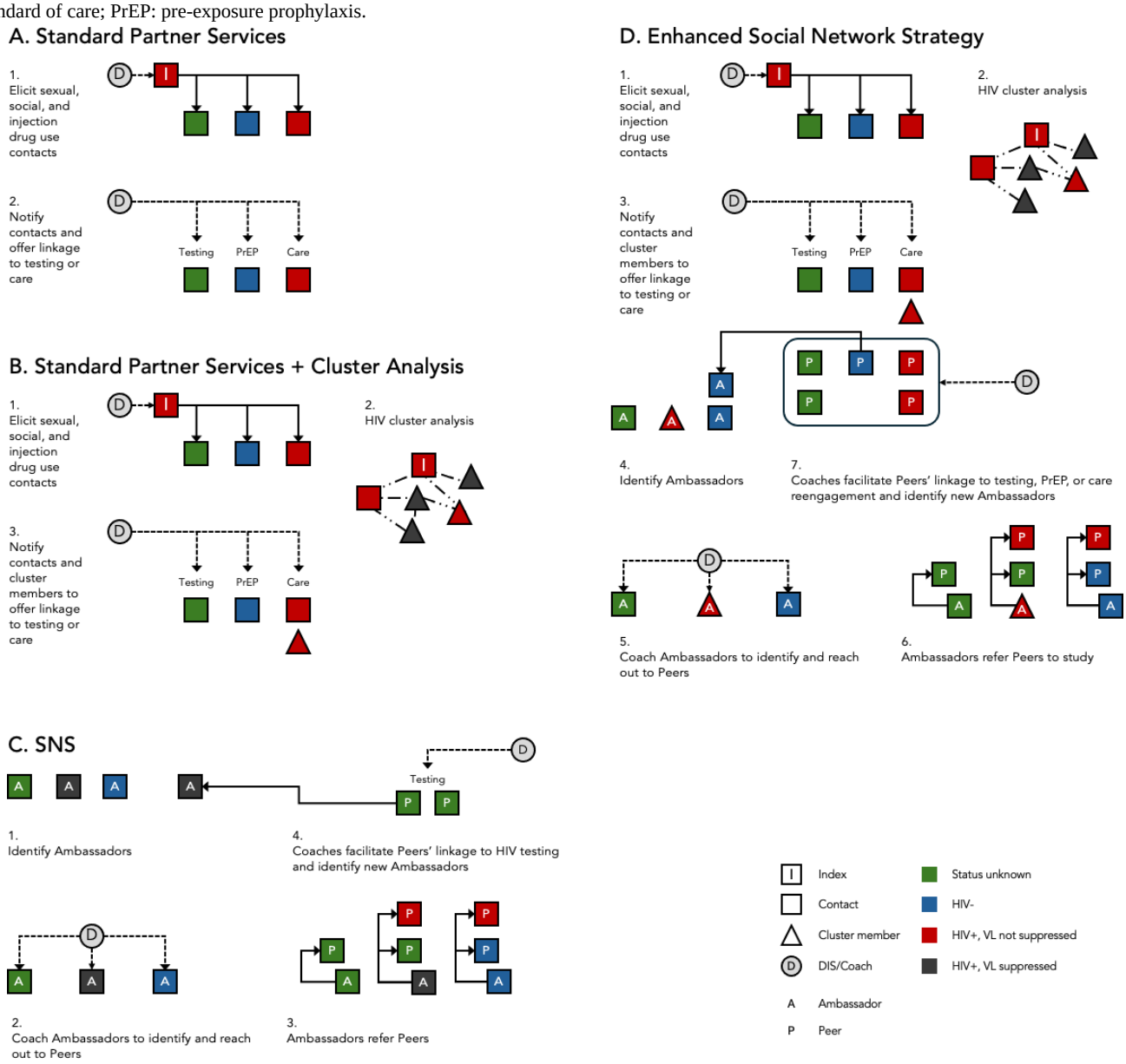
XSL•FO
RenderX

Textbox 1. Definitions of study population and key roles within the enhanced social network strategy (eSNS).

1. Key population:
   - Member of a group disproportionately affected by HIV (eg, Black or African American or Latine gay, bisexual, and other cisgender men who have sex with men or transgender women).
2. Priority network:
   - Sexual (or injecting) network with ≥1 new HIV diagnosis within 12 months.
   - Identified through reporting of new HIV diagnosis.
3. Community ambassador:
   - Key population member or priority network member.
   - Aged 18 years and older.
   - Resides in the Charlotte region.
   - Able to provide informed consent in English and willing to refer peers to HIV services.
4. Peer:
   - Person in the social, sexual, or injecting network of the ambassador, whom the ambassador believes could benefit from HIV services.
   - Aged 18 years and older.
   - Resides in the Charlotte region or nearby such that they can attend an HIV service appointment in the Charlotte region.
5. Disease intervention specialist (DIS) coach:
   - Study team member who conducts routine partner services for members of priority networks.
   - Also delivers eSNS intervention.

XSL•FO
RenderX

Figure 4. Schematic overview of enhanced social network strategy (eSNS) intervention among priority networks identified during routine partner services. A. Partner services as routinely delivered by disease intervention specialists (DIS) eliciting contacts (usually sexual or injecting partners) for notification of exposure to HIV. B. Partner services when combined with HIV cluster response, where genetic cluster members who are not virally suppressed are also linked to care services. C. Standard social network strategy (SNS), where potential network recruiters (ambassadors) are identified from a population of interest, then coached to reach out to network members (peers) for HIV testing. D. eSNS, which combines all 3 strategies. SOC: standard of care; PrEP: pre-exposure prophylaxis.



Following the completion of routine partner services activities (contact tracing and voluntary partner notification), potential community ambassadors will be identified from among members of Priority Networks reached through partner services, including index persons and their social and sexual contacts (see Figure 4). DIS coaches will then lead ambassadors through a semistructured coaching session designed to develop an action plan for ambassadors to reach out to and refer their peers—who may be friends or acquaintances, not necessarily sexual or injecting partners—for HIV prevention and care services. Once peers contact the study team, DIS coaches will assist peers in accessing HIV testing, care, or PrEP services.

Overview of Key Roles

DIS coaches are trained study staff responsible for delivering eSNS with the support of the study coordinator in Charlotte. These staff members have dual roles as public health DIS and coaches for community ambassadors. As DIS, they receive training and credentials from NC DHHS to perform routine partner services: HIV or sexually transmitted infection (STI) notification, counseling, testing, and linkage to care services for persons with new HIV diagnoses and their social and sexual partners. As coaches, they will identify, enroll, and coach community ambassadors for eSNS. DIS coaches will provide peers with referrals to HIV prevention and treatment services and assist with scheduling appointments.

Community ambassadors are individuals from the study's key population, who could be referred to eSNS as an index person recently diagnosed with HIV; as a priority network member, that is, a social, sexual, or viral genetic cluster contact of an index person engaged through partner services; or through local

XSL•FO
RenderX

community-based organizations (CBOs). While an optimal ambassador will have a confirmed link to a recent HIV transmission network, whether through routine contact tracing or genetic cluster analysis, less emphasis is placed on recruiting the index person as an ambassador because a person with a new HIV diagnosis may be experiencing increased social or psychological stress [37]. After modifying inclusion criteria following Phase 1's formative findings, eligible ambassadors will include residents of the Charlotte region who are at least 18 years old, report ever having sex with another person assigned male sex at birth, can provide informed consent in English, and indicate their willingness to refer people in their social networks (peers) who they believe may benefit from HIV prevention or treatment services (see Textbox 1).

Peers are people in the social circle of ambassadors (network members) who may be referred to eSNS. Community ambassadors may refer anyone 18 years and older in their social circle whom they believe could benefit from or has reason to use HIV prevention or treatment services. There are no demographic restrictions (aside from age) on who may qualify as a peer, if identified and referred by an enrolled community ambassador. Peers of any sex, gender, or race can be referred from a community ambassador's social circles if the ambassador reasonably believes they could benefit from HIV prevention or treatment services and the peer lives close enough to attend HIV services in the Charlotte region.

Phase 1: Evaluate Barriers and Facilitators to Engaging Network Members

For formative data collected during phase 1 (2022 - 2024), we used a mixed methods approach to determine the optimal approach to engage members of priority networks and the study's key population. We collected qualitative data from 2 participant pools, HIV service providers and community members from our key population, to explore local barriers and facilitators to HIV services engagement and general perceptions of eSNS. Between December 2022 and February 2023, we held 4 focus groups and conducted 12 in-depth interviews (IDIs) with Charlotte-area HIV care and prevention services providers (clinicians, public health officials, community organization leaders, HIV case managers, and public health DIS). From July 2023 to February 2024, we conducted in-depth interviews with 13 Black cisgender men who have sex with men and 4 Black transgender women who reported engaging with HIV-related services in the previous 12 months. A total of 13 participants were living with HIV, 3 of whom were diagnosed within 12 months of being interviewed. From April to June 2024, we conducted additional follow-up IDIs with 8 of the 17 participants to further explore their perspectives of barriers and facilitators to implementing specific eSNS components. All IDIs and focus group guides probed aspects of eSNS adaptation and implementation. Interview questions about eSNS implementation were mapped to specific CFIR constructs [38] to maintain consistency across qualitative analyses.

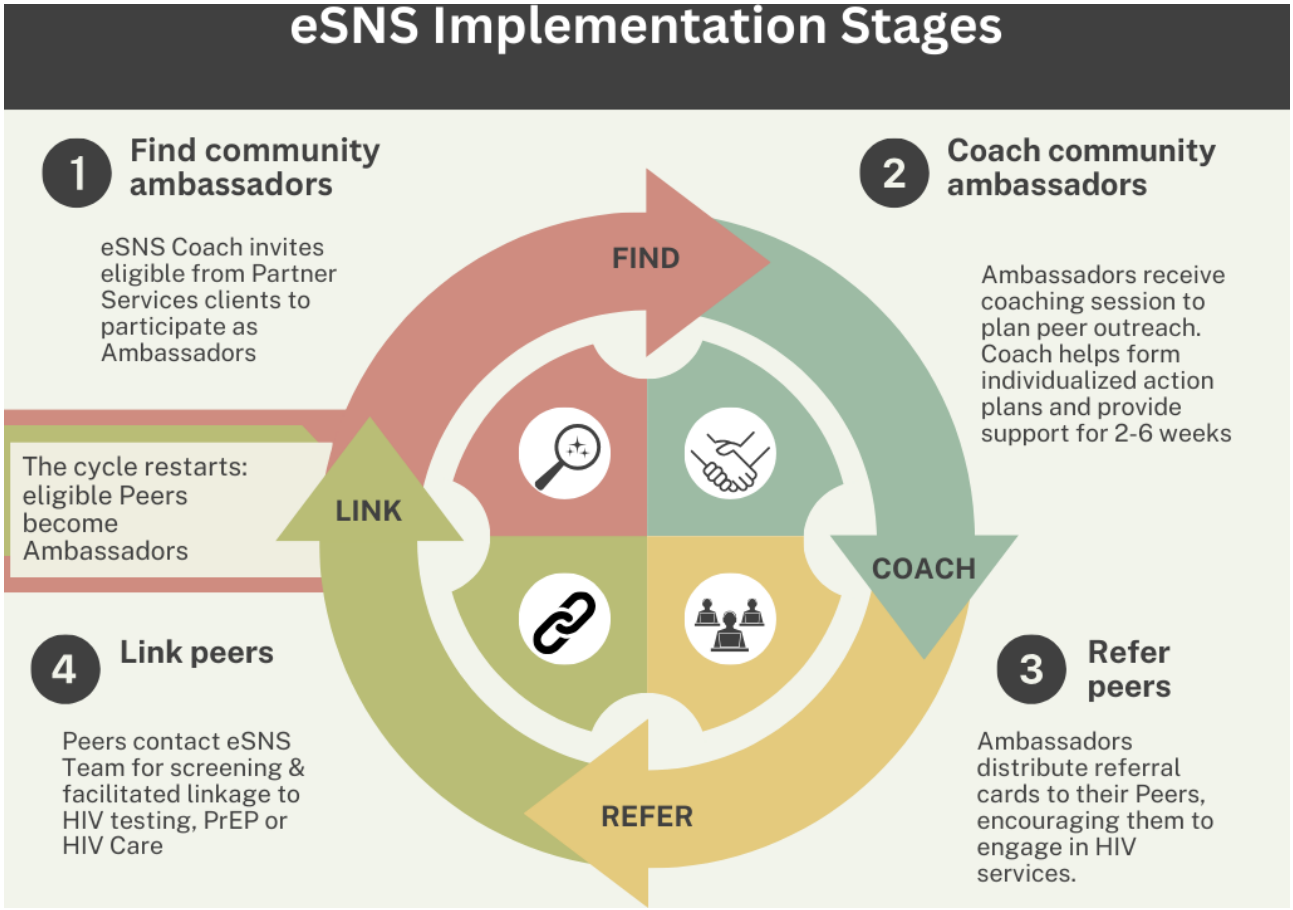
To characterize the social and sexual networks of our key population, including members of priority networks, we created a web-based, self-administered Personal Network Survey (PNS) hosted in REDCap (Research Electronic Data Capture; Vanderbilt University), which we launched in May 2023. A Spanish version was offered beginning in January 2024. Phase 1 survey participants were either referred by a DIS following routine partner services or referred from community organizations and clinics providing HIV services and self-reported engaging in at least one HIV-related service (eg, testing) in the previous 12 months; respondents received a US $20 gift card. The PNS (initial sample n=73, through February 2025) will continue during Phase 2. The PNS was adapted from a quantitative social network assessment survey [39] originally based on the General Social Survey [40]. A total of 3 name generators prompt participants to list pseudonyms of up to 5 individuals (alters) who, in the previous 6 months, provided social support, were sexual partners, or with whom participants used or injected drugs. Respondents can indicate if alters belong to multiple categories. A maximum of 15 unique names is possible. Following the initial inventory, participants are then asked to describe each alter's demographic characteristics, further categorize their relationship type ("partner," "family," "friend," "advisor or mentor," "co-worker," or "other"), indicate the level of closeness with each alter ("especially close" or "not especially close"), and estimate the closeness between alters listed ("stranger," "in-between," "especially close," or "don't know"). The PNS has been adapted for Phase 2 to include more granular data on social factors that could influence transmission risk, including drug use and chemsex.

Phase 2: Enhanced Social Network Strategy (ESNS) Intervention

In adapting the CDC's SNS, we retained the original intervention's 4-step model while renaming the steps with revised terminology that is more person-forward [41] and better meets the needs of our key population (see Figure 5). We will make several adaptations to SNS. Before phase 1, we had planned for eSNS service linkage to include PrEP linkage or care re-engagement (not only HIV testing, unlike SNS) and to embed eSNS within a public health department's partner services program. Following phase 1, we also widened our key population to include English and non–English-speaking groups disproportionately affected by HIV in North Carolina (although all ambassadors must be able to provide informed consent in English), and, guided by formative data and recommendations of CAG members, developed a process for referrals from community sites to engage key population members outside of partner services (described further in the Results section).

XSL•FO
RenderX

Figure 5. Schematic diagram of enhanced social network strategy (eSNS) procedures, displaying the 4-step intervention structure and revised terminology. SNS: social network strategy; DIS: disease intervention specialist; PrEP: pre-exposure prophylaxis.



Step 1: Finding Community Ambassadors

We will use 3 strategies to enroll a planned sample of at least 20 community ambassadors per year: (1) invite eligible priority network members identified during routine partner services, (2) publicize the study with local community organizations and clinics who provide HIV-related services to the study's key population, and (3) invite eligible peers to become ambassadors.

For (1), DIS coaches will inform individuals contacted during partner services investigation of priority networks (recent HIV transmission networks involving our key population) about eSNS. To reach networks with the highest probability of recent HIV transmission [42], we will collaborate with NC DHHS to prioritize index persons diagnosed with acute incident HIV infection [43] for assignment to DIS coaches. Following completion of partner services activities, DIS coaches will invite interested clients who meet inclusion criteria to become ambassadors. Eligible clients include people recently diagnosed with HIV, named partners of people recently diagnosed with HIV, or genetic cluster members re-engaged with HIV care who were identified through cluster detection protocols. Ambassadors do not have to be living with HIV.

Based on recommendations from phase 1, for (2) we will identify potential ambassadors through community-based recruitment in addition to partner services. Through partnerships with local CBOs that offer HIV-related services to our key population, we will post print materials at clinic and community venues and distribute digital recruitment materials through social media platforms. For (3), previously referred peers may become ambassadors, facilitating deeper penetration into networks and recruitment of new peers who have not previously been reached. Eligible peers will be contacted by a DIS coach for a brief program orientation before providing informed consent to serve as ambassadors. Newly identified ambassadors will replenish ambassadors who completed peer outreach (Step 3).

Step 2: Ambassador Coaching

Following orientation and informed consent, DIS coaches will deliver a semistructured coaching session to ambassadors. Coaching is a core component of eSNS designed to foster rapport and to develop an individualized action plan for contacting peers. We adapted CDC's SNS Coaching Guide [44] to accommodate findings from Phase 1 formative data and integrated our guide into Network Canvas (Complex Data Collective), an open-source software interface using intuitive visuals to collect personal network data [45]. DIS coaches are encouraged (but not required) to hold the first coaching session in person. All coaching sessions must take place in a private location. Ambassadors will receive a US $35 gift card for completing the coaching session.

During the first coaching session, ambassadors will be prompted to identify people in their social network (peers) whom they believe would benefit from linkage to HIV prevention or treatment services, including but not limited to sexual or injecting partners. Collectively, these peers comprise each

XSL•FO
RenderX

ambassador's outreach network. In addition to demographic information, ambassadors will be asked to provide each peer's first name and last initial (or nickname or pseudonym). The DIS coach will guide the ambassador through a semistructured conversation to create a unique action plan to contact each peer (see Multimedia Appendix 1). DIS coaches will advise ambassadors that they do not need to disclose their HIV status or medical history and will help devise conversational strategies to manage situations where peers may ask about the ambassador's HIV status. At the first coaching session, ambassadors will be supplied with printed and digital referral cards to distribute to peers in their outreach network (see Figure 6).

Figure 6. Example referral card that ambassadors distribute to peers during the enhanced social network strategy (eSNS) outreach period. Each ambassador has a unique ID and image (bottom right). The QR code and URL link to a contact form uniquely associated with each ambassador. Peers who complete the contact form have unique codes that link to the ambassador's referral card.

Step 3: Peer Outreach

In step 3, ambassadors will approach each peer identified in their outreach network to give them a unique referral card (associated with the ambassador's study identifier) and encourage them to contact the eSNS team for assisted linkage to HIV services (see Figure 6). There will be no predetermined limit on the number of referral cards ambassadors can disperse, although DIS coaches will be trained to guide ambassadors to select peers whom they believe will benefit from one or more HIV services (Step 4). Peers indicate their interest by submitting a contact form to request that the eSNS team reach out to them, calling the eSNS team directly, or meeting the DIS coach in person at a time and location facilitated by the ambassador (eg, for field-based HIV testing).

We anticipate that ambassadors will engage in peer outreach for approximately 2 - 6 weeks, during which DIS coaches will contact ambassadors at least weekly for brief follow-up coaching to encourage them to complete their peer outreach plan or to discuss strategies to adapt it. During follow-up coaching, new peers may be identified by the ambassador. Peer outreach will conclude when an ambassador has saturated all potential peers in their outreach network or 6 weeks have elapsed since the first coaching session, whichever comes first. Our referral period of 2 - 6 weeks is a practical application of ongoing coaching and booster sessions described in implementation guidance, as SNS is based on short-term coaching [44]. Grounds for dismissal earlier than the 6-week cut-off include coercing a peer's participation, ambassadors and peers knowingly misrepresenting their identities to obtain gift cards, repeatedly referring people not part of their network or who are ineligible, or unreliable communication following enrollment.

Step 4: Linking Peers to HIV Services

After a peer initiates contact with the study team, the DIS coach will attempt to follow up with that peer within 24 hours for a study screening visit. The purpose of the screening visit is to explain the study's aims and for the DIS coach to verify whether a peer is eligible for linkage to HIV services. DIS coaches will first ask peers for their ambassador's name to validate against the code and image on that ambassador's referral card. Ambassadors will be compensated with a US $20 gift card per peer in their outreach network who completes screening, even if they do not enroll in the study intervention. Ambassadors will only receive compensation for a maximum of 5 peers' screening visits (US $100 total in gift cards), although more than 5 peers may be referred.

Peers will be offered assisted linkage to HIV prevention and care services available in the Charlotte region. First, informed consent and Health Insurance Portability and Accountability

XSL•FO
RenderX

Act (HIPAA) authorization will be obtained from Peers to permit the eSNS team to obtain data regarding the outcomes of scheduled HIV service appointments at community-based sites. Before scheduling appointments, DIS coaches will assess peers' preferences among local care and prevention service providers and whether they prefer to be accompanied by their ambassador. If peers prefer, DIS coaches may also meet peers at a private location for field-based HIV testing or accompany them to a testing site. For peers who do not present with clinical symptoms of acute HIV infection, DIS coaches may perform rapid HIV testing (OraQuick HIV or bioLytical INSTI HIV dual testing) according to North Carolina DHHS protocols. Peers who test positive or indeterminate following field-based rapid testing will receive counseling by DIS coaches and expedited linkage to HIV care and bridge counseling (see the Ethical Considerations section for more detail on DIS training). Peers who reside in South Carolina and cannot attend appointments in the Charlotte region will be linked to HIV services available in Columbia, South Carolina but will not be eligible to become ambassadors.

For HIV testing, peers will be offered testing if they have never previously tested or if their last HIV test occurred more than three months prior and yielded negative or unknown results. Peers living with HIV will be linked to initial care if they are newly diagnosed or never entered care; or re-engaged with care if previously diagnosed but not currently engaged in care. Care engagement is defined as both having had a clinic visit within 12 months and currently taking ART. Peers will be offered facilitated linkage to PrEP if HIV test results obtained within the prior three months were confirmed non-reactive and they have never been on PrEP, or if they previously took PrEP but have not received a PrEP prescription in the past 6 months.

Peers will receive US $20 gift cards for attending each of up to 2 HIV service appointments, US $40 total in gift cards (eg, completing both HIV testing and attending a PrEP initiation appointment). The study team will obtain and document all HIV test results. For peers choosing to engage in HIV services at a community site unaccompanied by a DIS coach, the eSNS team will contact the site and provide the HIPAA authorization to confirm that the peer attended the appointment and, as applicable, to obtain their HIV test results. Peers may also choose to self-disclose their attendance or results to the eSNS team with a photo or screenshot. Following linkage, peers may be invited to enroll as community ambassadors and initiate the intervention cycle again by participating in coaching (Step 2) and peer outreach (Step 3).

Ethical Considerations

The study was approved by the University of North Carolina at Chapel Hill Institutional Review Board (IRB #22 - 1988). Intervention procedures were determined to present minimal risks to participants, which include potential loss of confidentiality or emotional distress caused by the potentially sensitive nature of discussing HIV or receiving an HIV diagnosis following testing facilitated by the study. Ambassadors and enrolled peer participants will provide informed consent. All DIS credentialed in North Carolina, including DIS coaches, are trained to address psychological consequences resulting from an HIV or STI diagnosis, which includes specific modules on linking people to services (medical, behavior, and social support), in-person training, and ongoing mentorship with senior DIS. The eSNS team will follow applicable North Carolina DHHS and IRB procedures for lawful reporting of safety concerns, logging unanticipated events involving study participants, and referrals to social services agencies.

Research that leverages peer relationships can create ethical dilemmas around how to minimize the risk of coercion among peers and protect the confidentiality of participants who may know each other [46]. Creating a peer outreach plan will require asking for information about individuals who are not yet enrolled—and may never enroll—in the study intervention. To avoid a scenario where identifiable information is indirectly collected about unconsented peers (making them "secondary subjects"), only first names and last initials or pseudonyms will be recorded during coaching. Despite this precaution, peers contacted by ambassadors may have unpredictable reactions to receiving outreach or discussing HIV services, including assuming that sensitive or identifiable information was provided about them. However, as no personal identifiers will be recorded unless peers initiate screening, the study team will have no way of contacting (or identifying) peers until peers contact the study team. Once peers engage in screening, DIS coaches will explain that no data collected by the study team will be shared with the ambassador who referred them.

Compensation for peer outreach was structured to minimize the possibility of creating a financial incentive for ambassadors to coerce peers into research participation. Ambassador gift cards for peer referrals are limited in both number (maximum 5) and amount (US $20). Providing ambassadors gift cards for one of their peers completing a screening visit, rather than for enrolling in the study or completing an HIV service appointment, ensures that ambassadors are fairly compensated for time spent recruiting peers while avoiding a scenario where their compensation is contingent on peers' enrollment or participation in the study's intervention. The screening visit includes a description of the study and eligibility survey performed by a DIS coach, who will be trained to emphasize to peers that participation is voluntary, and peers who choose not to participate but who still wish to receive HIV services (eg, HIV testing) will be provided a referral to a community site.

Data Sources and Management

The REDCap platform hosted at UNC Chapel Hill is a HIPAA-compliant web-based platform that will serve as a central repository for data collected by the study team, including administration of ambassador and peer e-consent forms, participant recruitment and enrollment tracking, and collection of screening, appointment results, and survey data. Public health surveillance data obtained from the North Carolina Electronic Disease Surveillance System (under a data sharing agreement with North Carolina DHHS) will be stored in a secure cloud drive and REDCap. We plan to routinely analyze HIV sequences reported to the state and associated case-level epidemiological data to facilitate timely response to clusters within NC DHHS's Region II (Charlotte).

XSL•FO
RenderX

Data Analysis

Hypothesis

Our principal hypothesis posits that implementing eSNS will increase the uptake of HIV services within the Charlotte region among members of priority networks when compared to the Raleigh region over the same period. By harnessing peer social support networks, eSNS is designed to mitigate structural barriers—including stigma, distrust, limited PrEP knowledge, and difficulties navigating complicated medication logistics—that underlie disparities in HIV prevention and treatment services uptake among members of priority networks engaged by routine partner services (and among our key population). Compared to partner services alone, which is the standard of care intervention in North Carolina and South Carolina, we also hypothesize that enhancing standard SNS with genetic cluster analysis will enable further penetration into priority networks where onward HIV transmission is occurring by identifying a greater number of persons living with undiagnosed HIV that is closely related to HIV in other members of the network.

Effectiveness Evaluation

Intervention effectiveness will be assessed by comparing the 5-county Charlotte region to another urban North Carolina region, Raleigh, with similar demographics (see Figure 1). The metropolitan region surrounding Raleigh has the second-largest population in North Carolina, after Charlotte [47]. The primary effectiveness measures are the number of HIV-negative individuals linked to PrEP and the individuals living with HIV who are linked to care per index person from a priority network. Index persons in the intervention region (Charlotte) priority networks will be compared to index persons in the control region (Raleigh) priority networks, who receive standard-of-care partner services. Partner services are standardized among all North Carolina DHHS Field Service regions [48]. We will use a regression point displacement design (RPDD) [49] to assess the impact of eSNS on trends in our primary effectiveness measures between regions over the intervention period; the effective reproduction number for HIV will also be calculated over this period. Measures will be collected for each region at a preintervention baseline (2 years before), postintervention (6 months after), and during implementation for interim analyses at 6-month intervals. Because state HIV surveillance data is available in both the Charlotte and Raleigh regions from the baseline timepoint to a postintervention timepoint, RPDD provides statistical power to estimate the effect of eSNS when aggregating at a group level (index persons from priority networks) [50].

Although we adapted our original intervention design to enroll eSNS ambassadors through both routine partner services for index persons from priority networks and through ongoing waves of community-based recruitment, we will calculate the same effectiveness measures for ambassadors enrolled through each source. Only outcomes for ambassadors enrolled through partner services will be included in the RPDD analysis. However, for ambassadors enrolled through both sources, secondary effectiveness measures will also assess the extent to which eSNS penetrates priority networks where HIV transmission is ongoing. Among peers referred by all ambassadors, we will use a combination of study data, state HIV surveillance data, and molecular epidemiology to calculate the proportion of peers living with HIV that has close genetic linkages [42] to a cluster. We will also assess penetration into priority networks by calculating the proportion of newly diagnosed peers who initiate care, proportion of previously diagnosed peers who re-engage in care, the overall positivity among peers who test for HIV, and the proportion of peers testing negative who then initiate PrEP. Selected primary and secondary effectiveness measures are described in Table 1.

XSL•FO
RenderX

Table . Selected effectiveness measures of the enhanced social network strategy (eSNS) intervention.

Measures and description	Data source
Network Index	
Number of individuals diagnosed with HIV and linked to care per associated index case.[a]	Surveillance data
Number of HIV-negative individuals linked to PrEP[b] per associated index case.[a]	Surveillance data
Cluster analysis	
Proportion and number of peers living with HIV who are associated with a genetic cluster involving a new HIV diagnosis.	Study and surveillance data
Peers identified with high network bridging scores (betweenness centrality) [51].	Study and surveillance data
Viral suppression	
Number of peers per ambassador previously diagnosed with HIV who are not virally suppressed.	Study and surveillance data
Care and prevention linkage	
Proportion and number of peers newly diagnosed with HIV who initiate care.	Study data
Proportion and number of peers previously diagnosed with HIV who re-engage in care.	Study data
Proportion and number of peers who test negative and initiate PrEP.	Study data
Proportion of positive tests among all peers who test for HIV.	Study data

[a]Primary effectiveness measures.

[b]PrEP: pre-exposure prophylaxis.

Sample Size and Power Estimates

Of approximately 280 new HIV diagnoses reported annually in the Charlotte region, 80 would meet criteria for index persons eligible to be assigned to DIS coaches for routine partner services (including 20 acute HIV infection). From these 80 index persons, we estimate that 25% of contact networks reached through partner services will yield an enrolled ambassador (n=20). Based on partner services data previously reported to North Carolina DHHS, 100 contacts were elicited from 60 persons who were newly diagnosed with HIV in the Charlotte region from 2016 to 2019, yielding a network index of 1.7 (SD 1.8). With a peer outreach network index of 5.0 per ambassador (based on previous reports [22]), this would increase the overall network index in the Charlotte region from 1.7 to 3.0. We estimate an 80% power with alpha=.05 to detect a significant effect of eSNS on network index with 20 ambassadors enrolled per year for 3 years (n=60). Of note, we do expect that ambassadors' referral networks will have a bimodal distribution—with most having a low number of peers while a few will yield a high number of peers.

Process Evaluation

Process measures will be routinely collected during phase 2 to examine implementation facilitators and barriers concurrent to intervention delivery, which will inform future scale-up (see Figure 3). We will use a triangulation protocol to integrate mixed-methods data from qualitative and quantitative sources [52]. Guided by the CFIR [38], we plan to conduct in-depth interviews with 2 ambassadors per intervention year (total of 6 interviews with ambassadors) and annual interviews pre- and postimplementation with the 2 DIS coaches (total of 4 interviews). Interviews will explore DIS coach and ambassador perceptions across three CFIR domains: including (1) the adaptability and complexity of the intervention (intervention characteristics); (2) the structural characteristics, communications, relative priority of eSNS within a state public health agency delivering partner services (inner setting); and (3) the characteristics of DIS coaches' and ambassadors' self-efficacy and knowledge (individuals). To understand additional factors that may influence intervention delivery, we developed a survey instrument administered to ambassadors after concluding peer outreach. This instrument will incorporate items from previously validated scales to quantitatively assess self-efficacy and intervention acceptability [53], HIV-related stigma (for ambassadors living with HIV) [54,55], experiences of discrimination [56], and gender identity and connectedness with the LGBTQ+ community [57,58].

Routine program monitoring data will be analyzed at 6-month intervals to assess whether eSNS is being delivered as intended (fidelity) over the intervention period and the extent to which eSNS as designed mitigates social barriers to HIV services engagement. We will monitor program delivery along each step in eSNS. Fidelity measures include the number of potential ambassadors identified per year (step 1), the length of the initial coaching session (step 2), the number of follow-up coaching contacts each DIS coach holds with each ambassador, and the length of time between peers contacting the study team and screening (step 4). A full set of monitoring measures is provided in Multimedia Appendix 2.

Results

The study was funded in July 2022. The initial protocol for phase 1 was approved by the IRB in October 2022, and a modification detailing eSNS intervention procedures for Phase 2 was approved in June 2024. As of December 2024, we have completed the initial phase 1 formative assessments. We launched participant recruitment for eSNS in March 2025.

XSL•FO
RenderX

Baseline analyses of HIV molecular epidemiology across North Carolina and South Carolina remain pending execution of data sharing agreements in both states.

Our preliminary research findings and guidance from our CAG and phase 1 identified several recommendations for adaptation and implementation of eSNS. An overarching theme in our qualitative data was the importance of fostering trust through partnerships with CBOs. This led us to develop procedures for supplementing participant recruitment of potential ambassadors beyond partner services clients in priority networks (eSNS step 1) and for facilitating HIV services linkage at local CBOs and clinics (eSNS step 4). Formative qualitative analyses also identified recommendations to address the experience of our key populations with research engagement—specifically by considering strategies to alleviate research fatigue and to ensure that our key populations are represented among the study team. Our CAG recommended opening study participation to members of the Latin community, which is increasingly affected by HIV in the Charlotte area. The amount of overlap between Black and Spanish-speaking Latin sexual networks is currently unclear. To address this, we began offering the PNS in Spanish during phase 1 and expanded the study's key population to include both Latin and Black GBMSM and transgender women. In phase 2, English-speaking Latin persons will be included as ambassadors. Future intervention scale-up will be needed to adapt eSNS to Spanish-speaking Latine persons.

From interviews with members of our key population who self-identified as Black or African American, we found that privacy and flexibility in HIV prevention and care services were important themes. We have adopted recommendations to allow appointment scheduling for peers at nontraditional times, having DIS coaches offer peers discreet, field-based HIV rapid testing, ensuring time in between appointments to avoid peers crossing paths, and allowing ambassadors to choose their preferred mode of coaching (in-person or phone call). Finally, confusion arose about intervention roles (DIS coaches vs ambassadors) during qualitative data collection with both HIV services providers and key population members. We will train DIS coaches to clearly communicate the structure of the compensation process, and that peers may choose whether to be accompanied by their ambassador to service appointments, and that under no circumstances will ambassadors be able to access their peers' protected health information or study data.

Discussion

Implications

Our study aims to provide critical information guiding "next-generation" partner services in high HIV burden areas of the Southern US that extend across urban-rural and state boundaries, leveraging active involvement from community partners impacted by HIV. Our initial intervention design incorporated several key adaptations to CDC's SNS procedures. Compared to previous SNS pilot studies [18,20-22], eSNS expands the scope of service engagement to include PrEP and HIV care in addition to HIV testing. Importantly, we aim to specifically engage priority networks experiencing recent HIV transmission, not only demographic groups disproportionately impacted by HIV. We will focus engagement with these networks by integrating eSNS into existing health department–delivered partner services and, to assess the penetration of eSNS into priority networks, propose detailed characterization of ambassadors' network relationships. These analyses will inform continued monitoring and adaptation of the eSNS intervention.

Challenges and Limitations

Our rationale for understanding and leveraging recent HIV transmission networks is based on rigorous previous research [59]. We anticipated several challenges and have incorporated strategies informed by our formative work to mitigate potential problems. Challenges to recruitment could limit study success, which may be heightened within partner services, where participants may be unreachable, unwilling or unable to refer sexual partners, or distrustful of public health. We have used several strategies to mitigate this: routine elicitation of social contacts during partner services, expansion of ambassador recruitment to community sites, and robust engagement of our CAG (the latter 2 were identified by our formative research findings). Ambassadors may not have well-connected networks and may still have difficulty reaching marginalized social communities. We plan to review the composition of Ambassadors' outreach networks on an ongoing basis, which will specifically try to identify communities that are and are not reached. In addition, potential problems can arise during eSNS implementation, such as misinterpretation of messaging around networks, fraudulent activity, and privacy violations. Informed by phase 1 formative work, we established robust community partnerships to optimize trust and foster active input in study implementation. Finally, surveillance data acquisition and management of multiple, dynamic datasets are obstacles for timely study initiation and cluster identification and require careful data management, security measures, and staff training. Obtaining surveillance data across several jurisdictions has stalled the molecular epidemiology components of our study due to legal restrictions on data sharing. As of January 2025, a data use agreement with the South Carolina Department of Public Health remains pending; lack of South Carolina surveillance data will limit identification of HIV clusters that span both states.

Conclusions

We plan to implement an evidence-based social network strategy to increase engagement in HIV services among key populations disproportionately affected by HIV. This paper provides the protocol for the Carolinas RESPOND eSNS intervention, detailing the 3 phases of planned study design, implementation, and evaluation. Our mixed methods approach and active CAG involvement before and during implementation to adapt recruitment strategies will facilitate understanding of barriers and facilitators of peer recruitment success and allow us to make community-appropriate adaptations in real time. eSNS implementation began in March 2025, with final results expected in 2028.

XSL•FO
RenderX

Acknowledgments

We thank the following individuals who conducted interviews in the formative phase of the study: Maria Esposito, Grace Galphin, Jordyn McCrimmon, Ujunwa Onyeama, Matthew Zinck, and Sable Watson. We would like to acknowledge the Community Advisory Group members and public health partners in Mecklenburg County Public Health and Region II Field Services Office, Division of Public Health, North Carolina.

The project is funded by the National Institutes of Health (NIH) through Grant Award Number R01AI169602. The UNC Center for AIDS Research (CFAR) is funded by award number P30 AI50410. REDCap is supported under award UL1TR002489 from the Clinical and Translational Science Award program of the Division of Research Resources, NIH.

Data Availability

The data sets generated and analyzed during phase 1 of the study are available from the corresponding author on reasonable request, upon completion of primary analyses.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Coaching guide embedded within Network Canvas software and used to help ambassadors develop a peer outreach action plan.
[PDF File, 4733 KB - publichealth_v11i1e69495_app1.pdf]

Multimedia Appendix 2
Program monitoring measures that will be routinely collected and analyzed to assess program performance, organized by the 4 “steps” of eSNS that they assess.
[DOCX File, 25 KB - publichealth_v11i1e69495_app2.docx]

References

1. Sullivan PS, Satcher Johnson A, Pembleton ES, et al. Epidemiology of HIV in the USA: epidemic burden, inequities, contexts, and responses. Lancet 2021 Mar 20;397(10279):1095-1106. [doi: 10.1016/S0140-6736(21)00395-0] [Medline: 33617774]
2. Centers for Disease Control and Prevention. Diagnoses, deaths, and prevalence of HIV in the United States and 6 territories and freely associated States, 2022. 2024 Report No.: 35 URL: https://www.cdc.gov/hiv-data/nhss/hiv-diagnoses-deaths-prevalence.html [accessed 2024-01-13]
3. Kerr JC, Valois RF, Diclemente RJ, et al. HIV-related stigma among African-American youth in the Northeast and Southeast US. AIDS Behav 2014 Jun;18(6):1063-1067. [doi: 10.1007/s10461-013-0687-6] [Medline: 24402690]
4. Reif S, Safley D, McAllaster C, Wilson E, Whetten K. State of HIV in the US Deep South. J Community Health 2017 Oct;42(5):844-853. [doi: 10.1007/s10900-017-0325-8] [Medline: 28247067]
5. Carter JW, Flores SA. Improving the HIV prevention landscape to reduce disparities for Black MSM in the South. AIDS Behav 2019 Oct;23(Suppl 3):331-339. [doi: 10.1007/s10461-019-02671-w] [Medline: 31541391]
6. Centers for Disease Control and Prevention (CDC). Recommendations for partner services programs for HIV infection, syphilis, gonorrhea, and chlamydial infection. MMWR Recomm Rep 2008 Nov 7;57(RR-9):1-83. [Medline: 18987617]
7. Hogben M, McNally T, McPheeters M, Hutchinson AB. The effectiveness of HIV partner counseling and referral services in increasing identification of HIV-positive individuals a systematic review. Am J Prev Med 2007 Aug;33(2 Suppl):S89-100. [doi: 10.1016/j.amepre.2007.04.015] [Medline: 17675019]
8. Williams WO, Song W, Huang T, Mulatu MS, Uhl G, Rorie M. HIV diagnoses through partner services in the United States in 2019 and opportunities for improvement. Sexual Trans Dis 2023;50(2):74-78. [doi: 10.1097/OLQ.0000000000001738]
9. Delivery of HIV partner services in the United States and dependent areas 2021. Centers for Disease Control and Prevention (CDC). 2021. URL: https://stacks.cdc.gov/view/cdc/149048 [accessed 2025-06-17]
10. Ferreira A, Young T, Mathews C, Zunza M, Low N, editor. Strategies for partner notification for sexually transmitted infections, including HIV. Cochrane Database Syst Rev 2013 Oct 3;2013(10):CD002843. [doi: 10.1002/14651858.CD002843.pub2] [Medline: 24092529]
11. Cope AB, Mobley VL, Samoff E, O’Connor K, Peterman TA. The changing role of disease intervention specialists in modern public health programs. Public Health Rep 2019;134(1):11-16. [doi: 10.1177/0033354918813549] [Medline: 30500306]

XSL•FO
RenderX

12. Golden MR, AugsJoost B, Bender M, et al. The organization, content, and case-finding effectiveness of HIV assisted partner services in high HIV morbidity areas of the United States. J Acquir Immune Defic Syndr 2022 Apr 15;89(5):498-504. [doi: 10.1097/QAI.0000000000002904] [Medline: 34974472]
13. Martin EG, Myderrizi A, Kim H, et al. Disease intervention specialist-delivered interventions and other partner services for HIV and sexually transmitted infections: a systematic review. Am J Prev Med 2025 Jan;68(1):182-203. [doi: 10.1016/j.amepre.2024.08.004] [Medline: 39142501]
14. Dennis AM, Pasquale DK, Billock R, et al. Integration of contact tracing and phylogenetics in an investigation of acute HIV infection. Sex Transm Dis 2018 Apr;45(4):222-228. [doi: 10.1097/OLQ.0000000000000726] [Medline: 29465708]
15. Green N, Hoenigl M, Chaillon A, et al. Partner services in adults with acute and early HIV infection. AIDS 2017 Jan 14;31(2):287-293. [doi: 10.1097/QAD.0000000000001308] [Medline: 27831950]
16. Pasquale DK, Doherty IA, Sampson LA, et al. Leveraging phylogenetics to understand HIV transmission and partner notification networks. J Acquir Immune Defic Syndr 2018 Aug 1;78(4):367-375. [doi: 10.1097/QAI.0000000000001695] [Medline: 29940601]
17. Hurt CB, Beagle S, Leone PA, et al. Investigating a sexual network of black men who have sex with men: implications for transmission and prevention of HIV infection in the United States. J Acquir Immune Defic Syndr 2012 Dec 1;61(4):515-521. [doi: 10.1097/QAI.0b013e31827076a4] [Medline: 22972020]
18. Centers for Disease Control and Prevention (CDC). Use of social networks to identify persons with undiagnosed HIV infection--seven U.S. cities, October 2003-September 2004. MMWR Morb Mortal Wkly Rep 2005 Sep;54(24):601-605 [FREE Full text] [Medline: 15973240]
19. Lightfoot MA, Campbell CK, Moss N, et al. Using a social network strategy to distribute HIV self-test kits to African American and Latino MSM. J Acquir Immune Defic Syndr 2018 Sep 1;79(1):38-45. [doi: 10.1097/QAI.0000000000001726] [Medline: 29771792]
20. McCree DH, Millett G, Baytop C, et al. Lessons learned from use of social network strategy in HIV testing programs targeting African American men who have sex with men. Am J Public Health 2013 Oct;103(10):1851-1856. [doi: 10.2105/AJPH.2013.301260] [Medline: 23948017]
21. Schumann C, Kahn D, Broaddus M, et al. Implementing a standardized social networks testing strategy in a low HIV prevalence jurisdiction. AIDS Behav 2019 Jan;23(Suppl 1):41-47. [doi: 10.1007/s10461-018-2146-x] [Medline: 29766328]
22. McGoy SL, Pettit AC, Morrison M, et al. Use of social network strategy among young black men who have sex with men for hiv testing, linkage to care, and reengagement in care, Tennessee, 2013-2016. Public Health Rep 2018;133(2_suppl):43S-51S. [doi: 10.1177/0033354918801893] [Medline: 30457951]
23. Moore ZS, McCoy S, Kuruc J, Hilton M, Leone P. Number of named partners and number of partners newly diagnosed with HIV infection identified by persons with acute versus established HIV infection. J Acquir Immune Defic Syndr 2009 Dec 1;52(4):509-513. [doi: 10.1097/QAI.0b013e3181ac12bf] [Medline: 19568174]
24. Oster AM, France AM, Panneer N, et al. Identifying clusters of recent and rapid HIV transmission through analysis of molecular surveillance data. J Acquir Immune Defic Syndr 2018 Dec 15;79(5):543-550. [doi: 10.1097/QAI.0000000000001856] [Medline: 30222659]
25. Oster AM, Lyss SB, McClung RP, et al. HIV cluster and outbreak detection and response: the science and experience. Am J Prev Med 2021 Nov;61(5 Suppl 1):S130-S142. [doi: 10.1016/j.amepre.2021.05.029] [Medline: 34686282]
26. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. JAMA 2019 Mar 5;321(9):844-845. [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
27. 2022 North Carolina HIV surveillance report. North Carolina HIV/STD/Hepatitis Surveillance Unit.: North Carolina Department of Health and Human Services, Division of Public Health, Communicable Disease Branch; 2023. URL: https://epi.dph.ncdhhs.gov/cd/stds/figures/2022-HIVSurveillanceReportSummary.pdf [accessed 2025-08-09]
28. EHE priority jurisdictions. Office of Infectious Disease and HIV/AIDS. 2023. URL: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/jurisdictions/phase-one [accessed 2025-01-14]
29. Baral S, Logie CH, Grosso A, Wirtz AL, Beyrer C. Modified social ecological model: a tool to guide the assessment of the risks and risk contexts of HIV epidemics. BMC Public Health 2013 May 17;13:482. [doi: 10.1186/1471-2458-13-482] [Medline: 23679953]
30. Berkman LF, Glass T, Brissette I, Seeman TE. From social integration to health: Durkheim in the new millennium. Soc Sci Med 2000 Sep;51(6):843-857. [doi: 10.1016/S0277-9536(00)00065-4]
31. Latkin CA, Knowlton AR. Social network assessments and interventions for health behavior change: a critical review. Behav Med 2015;41(3):90-97. [doi: 10.1080/08964289.2015.1034645] [Medline: 26332926]
32. Zhang J, Centola D. Social networks and health: new developments in diffusion, online and offline. Annu Rev Sociol 2019 Jul 30;45(1):91-109 [FREE Full text] [doi: 10.1146/annurev-soc-073117-041421]
33. Tobin KE, Latkin CA. Social networks of HIV positive gay men: their role and importance in HIV prevention. In: Wilton L, editor. Understanding Prevention for HIV Positive Gay Men: Innovative Approaches in Addressing the AIDS Epidemic: Springer New York; 2017:349-366.

XSL•FO
RenderX

34. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226. [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
35. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation Sci 2009 Dec;4:50. [doi: 10.1186/1748-5908-4-50]
36. Selik RM, Mokotoff ED, Branson B. Revised surveillance case definition for HIV infection--United States, 2014. MMWR Recomm Rep 2014 Apr 11;63(RR-03):1-10. [Medline: 24717910]
37. Wells N, Murphy D, Ellard J, Philpot SP, Prestage G, RISE Study Team. HIV diagnosis as both biographical disruption and biographical reinforcement: experiences of HIV diagnoses among recently diagnosed people living with HIV. Qual Health Res 2023 Feb;33(3):165-175. [doi: 10.1177/10497323221146467] [Medline: 36633974]
38. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implementation Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0]
39. Dhand A, White CC, Johnson C, Xia Z, De Jager PL. A scalable online tool for quantitative social network assessment reveals potentially modifiable social environmental risks. Nat Commun 2018 Sep 26;9(1):3930. [doi: 10.1038/s41467-018-06408-6] [Medline: 30258103]
40. Burt RS. Network items and the general social survey. Soc Networks 1984 Dec;6(4):293-339. [doi: 10.1016/0378-8733(84)90007-8]
41. HIV language guide. National Institute of Allergy and Infectious Diseases. 2024 May. URL: https://www.niaid.nih.gov/research/hiv-language-guide [accessed 2025-01-10]
42. Dennis AM, Frost SDW, Enders K, et al. HIV-1 transmission linkages among persons with incident infection to inform public health surveillance. EClinicalMedicine 2021 Jul;37:100968. [doi: 10.1016/j.eclinm.2021.100968] [Medline: 34195581]
43. Cohen MS, Shaw GM, McMichael AJ, Haynes BF. Acute HIV-1 infection. N Engl J Med 2011 May 19;364(20):1943-1954. [doi: 10.1056/NEJMra1011874] [Medline: 21591946]
44. Social network strategy for HIV testing recruitment. CDC. 2023. URL: https://www.cdc.gov/hiv/effective-interventions/diagnose/social-network-strategy/index.html [accessed 2025-08-15]
45. Janulis P, Phillips G 2nd, Melville J, et al. Network canvas: an open-source tool for capturing social and contact network data. Int J Epidemiol 2023 Aug 2;52(4):1286-1291. [doi: 10.1093/ije/dyad036] [Medline: 36944105]
46. Semaan S, Santibanez S, Garfein RS, Heckathorn DD, Des Jarlais DC. Ethical and regulatory considerations in HIV prevention studies employing respondent-driven sampling. Int J Drug Policy 2009 Jan;20(1):14-27. [doi: 10.1016/j.drugpo.2007.12.006] [Medline: 18243679]
47. Annual estimates of the resident population for metropolitan statistical areas in the United States and puerto rico: April 1, 2020 to July 1, 2023. U.S. Census Bureau. 2023. URL: https://www2.census.gov/programs-surveys/popest/tables/2020-2023/metro/totals/cbsa-met-est2023-pop.xlsx [accessed 2024-11-14]
48. Field services for local health departments. North Carolina Department of Health and Human Services. 2023. URL: https://epi.dph.ncdhhs.gov/cd/about/field_svcs.html [accessed 2025-01-13]
49. Linden A, Trochim WMK, Adams JL. Evaluating program effectiveness using the regression point displacement design. Eval Health Prof 2006;29(4):407-423. [doi: 10.1177/0163278706293402] [Medline: 17102063]
50. Wyman PA, Henry D, Knoblauch S, Brown CH. Designs for testing group-based interventions with limited numbers of social units: the dynamic wait-listed and regression point displacement designs. Prev Sci 2015 Oct;16(7):956-966. [doi: 10.1007/s11121-014-0535-6] [Medline: 25481512]
51. Shah NS, Iveniuk J, Muth SQ, et al. Structural bridging network position is associated with HIV status in a younger Black men who have sex with men epidemic. AIDS Behav 2014 Feb;18(2):335-345. [doi: 10.1007/s10461-013-0677-8] [Medline: 24337699]
52. Fielding NG. Triangulation and mixed methods designs: data integration with new research technologies. J Mix Methods Res 2012 Apr;6(2):124-136. [doi: 10.1177/1558689812437101]
53. Sekhon M, Cartwright M, Francis JJ. Acceptability of health care interventions: a theoretical framework and proposed research agenda. Br J Health Psychol 2018 Sep;23(3):519-531. [doi: 10.1111/bjhp.12295] [Medline: 29453791]
54. Wright K, Naar-King S, Lam P, Templin T, Frey M. Stigma scale revised: reliability and validity of a brief measure of stigma for HIV+ youth. J Adolesc Health 2007 Jan;40(1):96-98. [doi: 10.1016/j.jadohealth.2006.08.001] [Medline: 17185215]
55. Beer L, Tie Y, McCree DH, et al. HIV stigma among a national probability sample of adults with diagnosed HIV-United States, 2018-2019. AIDS Behav 2022 Jan;26(Suppl 1):39-50. [doi: 10.1007/s10461-021-03414-6] [Medline: 34374919]
56. Stucky BD, Gottfredson NC, Panter AT, Daye CE, Allen WR, Wightman LF. An item factor analysis and item response theory-based revision of the Everyday Discrimination Scale. Cultur Divers Ethnic Minor Psychol 2011 Apr;17(2):175-185. [doi: 10.1037/a0023356]
57. Frost DM, Meyer IH. Measuring community connectedness among diverse sexual minority populations. J Sex Res 2012;49(1):36-49. [doi: 10.1080/00224499.2011.565427] [Medline: 21512945]

XSL•FO
RenderX

58. Turpin R, Dangerfield Ii DT, Oke T, Thorpe RJ, Hickson DA. Psychometric validation of the connectedness to the LGBT Community Scale among black sexual minority men living with HIV. J Community Health 2025 Feb;50(1):63-70. [doi: 10.1007/s10900-024-01392-z] [Medline: 39190036]
59. Hurt CB, Morrison AS, Guy J, et al. Beyond disease intervention: exploring an expanded role for partner services in the MATRix-NC demonstration project. Sexual Trans Dis 2022;49(2):93-98. [doi: 10.1097/OLQ.0000000000001544]

Abbreviations

ART: antiretroviral therapy
CAG: community advisory group
CBO: community-based organization
CDC: Centers for Disease Control and Prevention
CFIR: Consolidated Framework for Implementation Research
DHHS: Department of Health and Human Services
DIS: disease intervention specialist
EHE: Ending the HIV Epidemic
eSNS: enhanced social network strategy
GBMSM: gay, bisexual, and other cisgender men who have sex with men
HIPAA: Health Insurance Portability and Accountability Act
MSEM: modified social ecological model
PrEP: pre-exposure prophylaxis
REDCap: Research Electronic Data Capture
RPDD: regression point displacement design
SNS: social network strategy

Edited by GJB Sousa; submitted 03.02.25; peer-reviewed by A Algarin, BG Rogers; revised version received 18.07.25; accepted 19.07.25; published 29.08.25.

Please cite as:
Dale I, Zarwell M, Diggs A, Strunk Elkins J, Okumu E, Weissman S, Lin FC, Mobley V, Golin C, Rudolph A, Dennis AM
An Enhanced Social Network Strategy to Increase the Uptake of HIV Services: Protocol for Type I Hybrid Implementation Study (Carolinas RESPOND)
JMIR Public Health Surveill 2025;11:e69495
URL: https://publichealth.jmir.org/2025/1/e69495
doi:10.2196/69495

© Ian Dale, Meagan Zarwell, Alicia Diggs, Jesse Strunk Elkins, Eunice Okumu, Sharon Weissman, Feng-Chang Lin, Victoria Mobley, Carol Golin, Abby Rudolph, Ann M Dennis. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Effectiveness of an Intelligent mHealth Intervention for Child Unintentional Injury Prevention: Protocol for a Cluster Randomized Controlled Trial

Yang Yuan[1], BA; Yuying Jing[1], BA; Jiyu Li[1], BA; Jieyi He[2], PhD; David C Schwebel[3], PhD; Peishan Ning[1,4], PhD; Zhenzhen Rao[1], PhD; Li Li[1], PhD; Guoqing Hu[1], PhD

[1]Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Public Health Bldg, 172 Tongzipo Road, Changsha, China

[2]Clinical Trials Center, The Affiliated Hospital of Guizhou Medical University, Guizhou, China

[3]Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, United States

[4]Hunan Provincial Key Laboratory of Philosophy and Social Sciences of Urban Smart Governance, Changsha, China

Corresponding Author:
Li Li, PhD
Department of Epidemiology and Health Statistics, Xiangya School of Public Health, Central South University, Public Health Bldg, 172 Tongzipo Road, Changsha, China

Abstract

Background: Unintentional injury is a leading cause of childhood morbidity and mortality worldwide. In China, real-world implementation of child injury prevention efforts remains inadequate due to constrained workforce capacity and a lack of operational frameworks.

Objective: This study aims to assess the effectiveness of a mobile health (mHealth) intervention, the Intelligent Child Unintentional-Injury Reduction & Education (iCURE) project, embedded within China's National Basic Public Health Service Program. The intervention relies on a WeChat (Tencent) service account for caregivers and a web-based platform for health care providers to deliver standardized unintentional injury prevention strategies for young children. Key features of the program include interactive questions and answers, injury risk assessment with instant feedback, a tailored injury prevention knowledge disseminator, and regular reminders to caregivers.

Methods: A double-blind, 12-month follow-up, cluster randomized controlled trial will be implemented in Changsha, Hunan Province, China. Caregivers of children aged ≤5 years will be recruited. Randomization will be conducted at the street or town level. The control group will receive routine safety education, while the intervention group will receive both routine safety education and the iCURE mHealth intervention focused on unintentional injury prevention and delivered via WeChat. Data will be collected at baseline and every 3 months during the study period. The primary outcome is 12-month incidence of unintentional injuries among children, including minor injuries and as reported by caregivers. Secondary outcomes include children's injury risk level and caregiver supervision behaviors assessed using a standard questionnaire. Data analysis will be conducted using generalized linear mixed models with a Poisson link function and generalized estimating equations to assess the effectiveness of the iCURE intervention, following intention-to-treat principles. Sensitivity analyses will be conducted with per-protocol principles and excluding participants with missing primary outcomes.

Results: As of May 2025, a total of 6701 participants have been successfully enrolled and baseline data were collected for all participants. Of those enrolled, 87.2% (5842/6,701) completed the first follow-up assessment.

Conclusions: This trial will examine the effectiveness of an intelligent mHealth intervention for child unintentional injury prevention building on China's National Basic Public Health Service Program. If successful, the iCURE intervention may provide a cost-effective strategy for child injury prevention in low- and middle-income countries.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400092698; https://www.chictr.org.cn/showprojEN.html?proj=238156

(JMIR Public Health Surveill 2025;11:e76195) doi:10.2196/76195

KEYWORDS

unintentional injury prevention; intelligent mobile health; cluster randomized controlled trial; children; China's National Basic Public Health Service Program

XSL•FO
RenderX

Introduction

Unintentional injuries remain a leading cause of morbidity and mortality among children worldwide [1]. In 2021, approximately 200,000 children younger than 5 years of age died globally from injuries and 38.7 million experienced nonfatal injuries [2]. In China, where the current research was conducted, injuries were the primary cause of child death in 2021, accounting for 10,827 deaths and 2.15 million incident cases in children younger than 5 years of age [2-4].

Injury prevention frameworks such as the Haddon Matrix and the "4E" model have significantly shaped global policies and interventions over recent decades. Building upon these, the World Health Organization (WHO)'s World Report on Child Injury Prevention (2008) identified key interventions across many domains, including legislation, product safety, environmental modification, education, and emergency care, to reduce major types of child injuries [5]. High-income countries implement many of these strategies through structured programs such as The Injury Prevention Program (TIPP), which offers age-specific safety counseling and pediatrician training to enhance injury prevention practices [6]. However, such resource-intensive approaches require substantial institutional capacity, trained personnel, and long-term funding, posing implementation challenges for low- and middle-income countries (LMICs) such as China.

Chinese health authorities have incorporated injury prevention into the National Basic Public Health Service Program [7], a system whereby health-related information is conveyed to the population through health care providers. Despite its broad coverage, the program faces significant capacity constraints. Primary care providers are often understaffed and overburdened with routine services such as immunizations, chronic disease management, and regular medical examinations [8]. Most practitioners lack training in injury prevention, and the absence of operational guidelines limits implementation [9-11]. Consequently, neither globally recommended programs [5,6] nor technical guidelines [12] developed within China have been implemented in Chinese primary care settings.

Recognizing the potential of digital health as an alternative to deliver health-related messaging to parents and families, both the WHO and China's national health strategies have emphasized the importance of integrating mobile health (mHealth) into public service delivery [13-15]. By leveraging digital technologies, mHealth can bridge gaps between research and practice by overcoming geographical and temporal barriers to deliver scalable, cost-effective interventions, including to resource-limited settings [16].

Further, the rapid advancement of artificial intelligence (AI) technologies has made it increasingly feasible to design interventions that integrate AI into mHealth apps. AI-enabled systems can ease workforce shortages by providing real-time support and interactions, automated risk assessments, and personalized recommendations. For example, a randomized controlled trial reported that an AI-powered digital platform for lifestyle change was effective in supporting weight loss among adults [17].

mHealth apps for child injury prevention have been piloted in high-income countries such as the United States [18-22], South Korea [23,24], and Australia [25], as well as in China [26-30]. Most of these interventions focus on delivering educational content via mobile platforms. However, several limitations have restricted their broader application in China [18-30]: (1) they primarily target preschool-aged children ages 3 - 6 years, overlooking risk among infants and toddlers [18,19,21,24,27-30]; (2) they rely on standalone applications, creating limited user engagement and poor long-term adherence [18-29]; and (3) they typically require dedicated technical support, continuous user assistance, specialized intervention skills, and localized cultural and linguistic appreciation, infrastructure that is not embedded within the National Basic Public Health Service Program. Together, these limitations undermine scalability, sustainability, and integration into routine service delivery in China [18-30].

To address these gaps, we developed the Intelligent Child Unintentional-Injury Reduction & Education (iCURE) program that is integrated into China's National Basic Public Health Service Program. iCURE is designed for health care providers to provide tailored advice on child injury prevention to caregivers via WeChat (Tencent), easing knowledge and temporal burden on providers and still conveying critical knowledge to caregivers. This protocol outlines a 12-month cluster randomized controlled trial to evaluate the effectiveness of iCURE in improving child injury–related knowledge, behaviors, and outcomes among caregivers of young children. If effective, iCURE offers a scalable and sustainable approach to incorporate child injury prevention into primary health care for young children and their parents across LMICs.

Methods

Study Design

A double-blind, 12-month, parallel-group cluster randomized controlled trial with 1:1 allocation ratio will be implemented in Changsha, Hunan Province, China. This study will strictly adhere to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines (Checklist 1)[31,32]. This protocol version is 1.0 and was enacted on August 1, 2024.

Participant Recruitment

A 2-stage recruitment strategy will be conducted. First, 12 streets and towns were selected from 5 administrative districts in Changsha, including 3 urban districts (Kaifu, Yuhua, and Yuelu) and 2 rural areas (Ningxiang and Liuyang). Each district consisted of at least 2 streets and towns that were eligible for study recruitment (Figure 1). Eligible streets and towns had a primary health care institution present that (1) provides basic public health services for a population with over 2000 children younger than 5 years of age and (2) reports they are not involved in any active child injury prevention programs.

One health care provider per institution will facilitate participant recruitment through (1) face-to-face communication, leaflet distribution, and poster display during routine services; (2) digital advertisements and instructional videos shared through

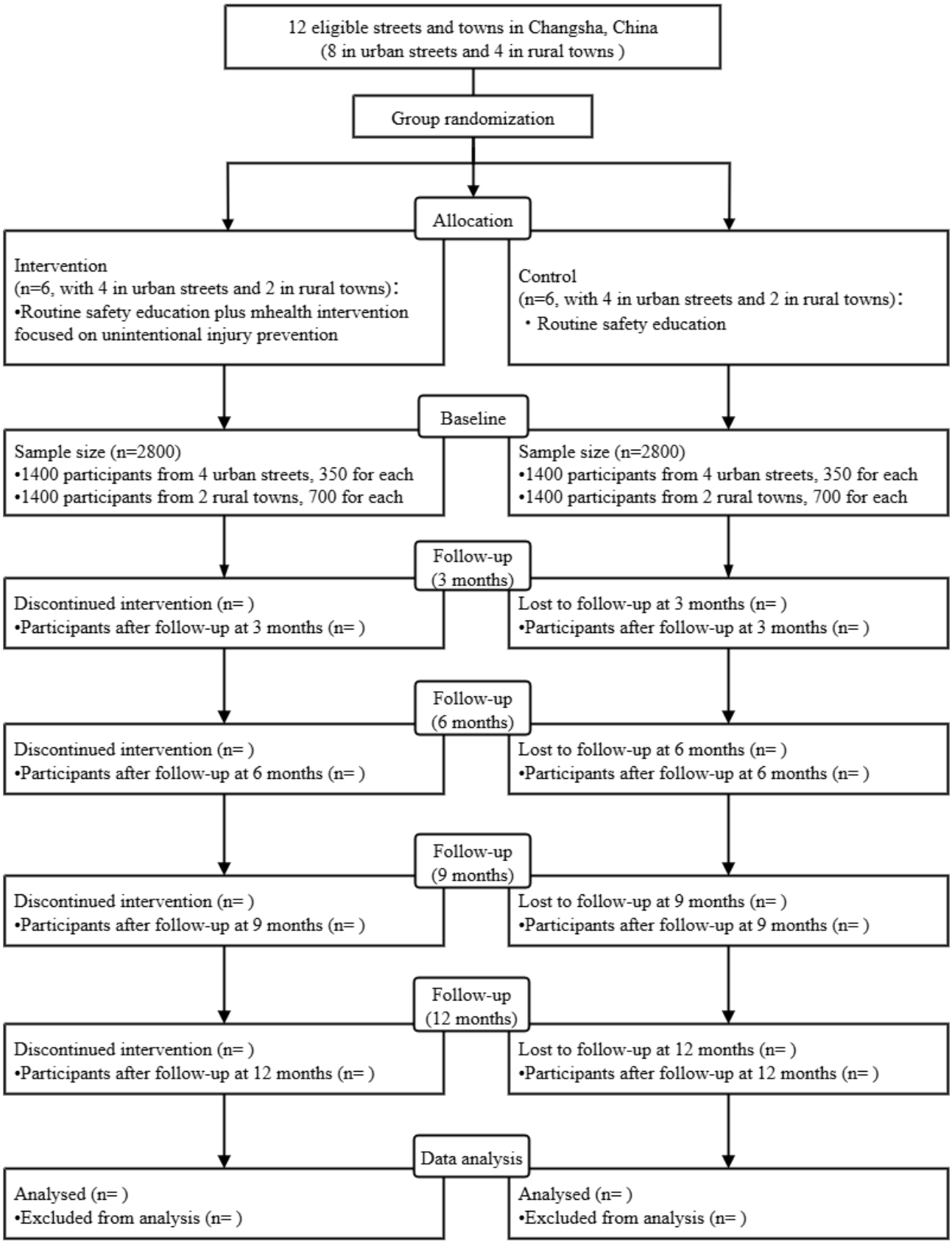
XSL•FO
RenderX

established communication platforms, such as WeChat groups; and (3) collaboration with local schools serving children aged 3 to 5 years, where teachers will distribute flyers to caregivers.

Families will be eligible if children (1) are aged ≤5 years at the time of baseline survey; (2) live with a primary caregiver who provides daily support; (3) reside in the selected study site; (4) have a caregiver aged ≥18 years who is literate, owns a smartphone, and provides informed consent; and (5) are not enrolled in other structured injury prevention programs. Children with severe illnesses requiring specialized care or hospitalization will be excluded.

XSL•FO
RenderX

Figure 1. Study flow diagram.



Sample Size

Previous research reported a 12-month incidence of unintentional injury of 15.5% among children younger than 6 years of age in Changsha [33]. We hypothesize that the proposed mHealth intervention will reduce the injury incidence rate by 25% [34], with a design effect of 1.5 for the cluster sampling [35]. We also very conservatively assume 30% loss to follow-up [36]. Given these assumptions, a minimum sample size of 5178 children is needed to achieve 80% statistical power at a .05

XSL•FO
RenderX

significance level with a 1:1 allocation ratio between intervention and control groups (G*Power 3.1.9.7). Given that the actual values of certain parameters (eg, incidence rate and effect size) may be somewhat lower than assumed [28,34], we empirically expanded the sample size to recruit approximately 5600 children, with (2800 children from urban areas and 2800 from rural areas, including at least 350 children from each selected urban street and 700 from each rural town) within the constraints of available research resources. The 1:2 recruitment ratio of urban streets and rural towns is based on prior field experience, which suggests greater recruitment challenges in urban areas compared to rural ones.

Intervention

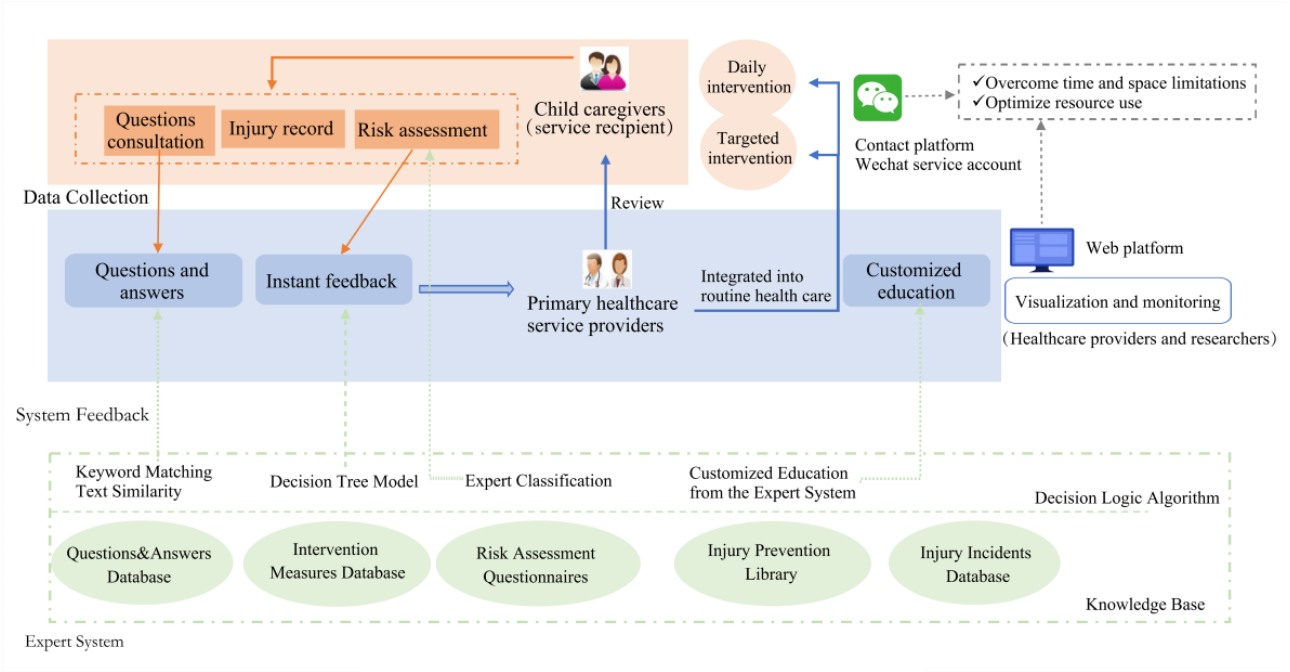
The iCURE program is designed as an expert system, defined as a computer-based system that simulates the decision-making ability of a human expert, integrating the latest evidence-based knowledge on child injury prevention and applying rules of logic to provide context-specific guidance and tailored education to caregivers [37]. Its development is guided by expert system methodologies and the intervention mapping framework [38]. The design incorporates core principles from injury prevention and behavior change theories, along with empirical evidence and key stakeholders' perspectives [39,40]. These components were integrated through a systematic development process involving literature review, needs assessment, and expert consultation [38]. The program applies algorithms such as keyword matching, text similarity analysis, decision tree models, and rule-based logic to deliver personalized recommendations and support [37,40-42].

A pilot study was conducted to assess the feasibility and operational performance of the iCURE program. The findings from the pilot were used to optimize its design and implementation, including improvements in system stability, registration flow, reward mechanisms, and data quality control.

Key Components and Content of the iCURE Program

The iCURE program consists of 2 components, a WeChat service account for caregivers and a web-based platform for health care providers (Figure 2 and Multimedia Appendices 1 and 2). This dual-platform approach facilitates information delivery and, crucially, supports caregiver-provider interaction.

Figure 2. The functional modules and operational flowchart diagram of the mHealth intervention for child injury prevention.



The WeChat service account serves as the main channel for delivering intervention content to caregivers. The web-based platform is only accessible to health care providers, who support the system by managing and customizing educational content, monitoring caregiver engagement, and tracking self-reported injury events. All data are saved securely once collected and can be visualized on the platform. Providers and researchers can also export relevant data, including risk assessment outcomes, injury events, and intervention engagement, as needed (Multimedia Appendix 1).

The WeChat service account for caregivers comprises 5 key modules:

1. Injury risk assessment: a comprehensive child injury risk assessment questionnaire, developed based on validated tools [43-47], to assess child age, caregiver and child behaviors, and environmental risk.
2. Instant feedback: upon completing the risk assessment, participants receive immediate feedback and tailored recommendations generated by an integrated expert classification system and a decision tree model. This system stratifies children into low or moderate risk or high risk, which supports further personalized prevention strategies. For research purposes, participants who are randomly assigned to the control group will complete the assessment but will not receive feedback. Caregivers of children in the intervention group who are identified by the system as "high risk" will receive item-specific risk reports. This individualized risk profile will guide the dissemination of AI-tailored education materials for each child.

XSL•FO
RenderX

3. Injury reporting: caregivers are encouraged to report any injury events their children experienced during the study period. The module facilitates timely documentation and monitoring of injury events.
4. Interactive questions and answers (Q&A): caregivers may submit injury prevention questions through the platform's real-time Q&A interface, which provides instant computer-generated and automated responses. The iCURE platform uses a hybrid natural language processing approach that combines keyword matching and text similarity analysis to retrieve and deliver relevant, automated responses from a curated injury prevention knowledge base created by the research group.
5. Tailored education: personalized educational content on child injury prevention is delivered in various formats (eg, text, images, long-form infographic post, flyers, and short videos). Delivery frequency is determined by a rule-based decision logic algorithm and tailored to the child's risk level; caregivers of low-risk children receive content once per week, while those with higher-risk children receive it 2 or 3 times per week.

Implementation Protocol

During routine consultations, primary health care providers will assist caregivers in registering for the iCURE WeChat service account. Upon registration, all caregivers will complete a questionnaire providing demographic information, their child's injury history, and an injury risk assessment survey.

Caregivers randomly assigned to the intervention group will then receive the full version of the iCURE program, including real-time feedback, tailored educational content, and access to the Q&A module. The control group will access a simplified version involving completion of assessment surveys that are used only for data collection; they will receive no feedback, education, or supplemental materials. Both groups will continue to receive routine safety education provided by preschool teachers and primary health providers. Group allocation is detailed in Table 1. Details regarding the timeline for recruitment and intervention period are shown in Figure 3.

Table . Functions of iCURE (Intelligent Child Unintentional-Injury Reduction & Education) used by intervention group and control group.

Functions	Intervention group	Control group
Data collection	✓	✓
Injury risk assessment	✓	✓[a]
Instant feedback	✓	
Tailored education	✓	
Interactive questions and answers	✓	

[a]Control group will complete the injury risk assessment for data collection purposes, with no risk evaluation reports provided.

XSL•FO
RenderX

Figure 3. Timeline for the schedule of enrollment, interventions, and assessments. iCURE: Intelligent Child Unintentional-Injury Reduction & Education.

	Study Period					
	Allocation	Enrollment	Postenrollment			
Timepoint	-1 month	0 month	3 month	6 month	9 month	12 month
Enrollment:						
Allocation	√					
Eligibility screen		√				
Follow and register WeChat service account		√				
Informed consent		√				
Intervention:						
iCURE system and routine safety education		●	—	—	—	●
Control:						
Rountine safety education		●	—	—	—	●
Assessments:						
Demographic information and previous injury history		√				
Unintentional injury incidence among children in the past 3 months			√	√	√	√
Caregiver supervision behavior		√				√
Children's injury risk assessment		√				√
Compliance data of using iCURE		●	—	—	—	●

Outcomes

The primary outcome is unintentional injury incidence, calculated as:

$$\frac{\text{Number of children who injured during the 12-month study period}}{\text{Total number of children}} \times 100\%$$

Injury events will be identified based on established criteria from previous studies [48], defined as any of the following: (1) receiving medical attention at a health care facility and being diagnosed with a specific injury; (2) receiving first aid from a nonmedical responder (eg, family member and teacher), such as wound dressing or bandaging; and (3) missing school for more than half a day due to an injury. We will also compare injury rates at 3-month intervals to assess the effect over time.

In addition, several secondary outcomes will be assessed:

(1) Change in child injury risk level: A 46-item questionnaire will assess children's risk of unintentional injury, classifying children as low-to-moderate or high risk. The questionnaire was adapted from existing validated instruments and covers 8 common types of injuries: drowning (5 items), road traffic injuries (6 items), choking (6 items), falls (6 items), burns and scalds (6 items), poisoning (5 items), animal bites (5 items), and sharp and blunt object or foreign body injuries (7 items) [43-47].

(2) Change in caregiver supervision behavior score: Caregivers' caregiver supervision behavior will be measured using a simplified version of a validated 5-item questionnaire, with responses rated on a 5-point Likert scale [49]. Items overlapping with the child injury risk assessment questionnaire were excluded to avoid redundancy.

Randomization and Blinding

Randomization will be conducted at the street or town level to minimize potential contamination. Each street or town will be randomly assigned to either the intervention or control group, stratified by administrative division, using a random number sequence generated by R 4.4.1 software (R Core Team). The group allocation will remain concealed from both participants and the researcher conducting data analysis.

Strategies to Enhance Engagement

To maximize caregiver engagement with iCURE, we will implement 4 strategies: (1) during registration, caregivers will be instructed to disable the "Do Not Disturb" mode in their WeChat service account; (2) caregivers in the intervention group will earn lottery entries for completing questionnaires and digital points for viewing educational materials, with both entries and points available to be redeemed in ranking competitions, as supported by prior research [28,50] and findings from a needs assessment; (3) safety-themed activities will be organized on designated special days, such as World Drowning Prevention Day, to reinforce injury prevention knowledge; and (4) iCURE materials will be shared through existing WeChat groups by teachers and health care providers in intervention areas to leverage community networks and expand outreach.

Data Collection

Primary outcome data will be collected every 3 months to minimize recall bias. Secondary outcomes, which are less sensitive to recall bias, will be assessed annually. The baseline survey includes demographic information, child injury history over the past 12 months, caregiver supervision behaviors, and

XSL•FO
RenderX

injury risk assessment, selected based on established associations with child injury and potential confounders [51]. All data will be self-reported, collected via the internet-based iCURE system, and securely stored on encrypted databases and hard drives. Caregivers will receive automated reminders followed by telephone follow-up if necessary to improve response rates.

The iCURE system will automatically record user engagement indicators, including (1) whether educational content was viewed, (2) frequency and duration of viewing, (3) messages to iCURE research team, and (4) online subscription to the iCURE WeChat service account.

Data Analysis

Data will be analyzed following intention-to-treat principles and including all participants as originally assigned, regardless of adherence. Descriptive analysis (mean, SD, median, IQR, and proportion) will be used to describe baseline characteristics by groups. Differences in unintentional injury incidence across 5 timepoints (baseline, 3, 6, 9, and 12 mo) and other outcomes at 2 timepoints (baseline and 12 mo) will be examined using rank-sum tests for ordinal outcomes, chi-square tests for categorical outcomes, and independent 2-sample *t* tests for continuous outcomes. Within-group changes over time will also be evaluated.

Generalized linear mixed models with a Poisson link function will be used to compare the 12-month injury incidence between groups, adjusting for baseline covariates and accounting for clustering at the street or town level. Risk ratios and 95% CIs will be reported. Generalized estimating equations will also be used to assess repeated measures of injury incidence across follow-up timepoints, providing population-averaged estimates while accounting for within-subject and within-cluster correlations. Secondary outcomes, including changes in injury risk levels and caregiver supervision behaviors, will be analyzed using regression models appropriate to the type and distribution of the variable. User engagement indicators will be included as covariates of the effectiveness analysis model.

Two sets of sensitivity analyses will be performed to assess the robustness of findings: (1) per-protocol analyses and (2) models excluding participants with missing primary outcome data. Missing data will be documented and addressed using appropriate statistical methods as needed.

This study will adhere to the CONSORT (Consolidated Standards of Reporting Trials) 2025 [52] statement to perform data analysis and report results.

Ethical Considerations

This study was reviewed and approved by the Ethics Committee of Xiangya School of Public Health, Central South University (XYGW-2024-71). All participants will provide electronic informed consent via the iCURE WeChat service account prior to participation, in accordance with national and institutional ethical guidelines. To ensure privacy and confidentiality, all data will be securely stored on encrypted databases and password-protected hard drives. Data will be de-identified before analysis, and no personally identifiable information will be retained in the analytic dataset. Participants will receive a modest digital monetary incentive ranging from 5 to 50 Chinese Yuan through a lottery-based mechanism via WeChat for their time and participation. No identifiable information will be included in the manuscript or any supplementary materials.

Results

The study received funding in January 2023. Participant enrollment began in August 2024 and was completed in February 2025. A total of 6701 individuals completed baseline data, with 3110 allocated to the intervention group and 3591 to the control group (Table 2). The two groups somewhat differed in child age, place of residence, caregiver's age, and education—the control group had higher proportions of children aged 3 to 5 years (69.7% vs 57.6%), residing in urban areas (61.3% vs 57.4%), caregivers aged 30 years and older (81.2% vs 74.9%) and receiving high or vocational school education and higher (85.7% vs 80.2%) than the intervention group (P<.05). The baseline 12-month injury incidence reported by caregivers was 8.86% (7.96-9.84) in the control group and 9.04% (8.06-10.11) in the intervention group.

By May 2025, 5,842 (87.2%) participants had completed the first follow-up survey.

XSL•FO
RenderX

Table . Number of recruited caregivers and children by demographic characteristics (%).

Variables	Overall, n (%)	Control, n (%)	Intervention, n (%)	*P* value
Total	6701 (100.0)	3591 (100.0)	3110 (100.0)	
Child gender				>.99
Girl	3147 (47.0)	1686 (47.0)	1461 (47.0)	
Boy	3554 (53.0)	1905 (53.0)	1649 (53.0)	
Child age (years)				<.001
1	1047 (15.6)	468 (13.0)	579 (18.6)	
1-2	1358 (20.3)	619 (17.2)	739 (23.8)	
3-5	4296 (64.1)	2504 (69.7)	1792 (57.6)	
Place of residence				.001
Urban	3988 (59.5)	2203 (61.3)	1785 (57.4)	
Rural	2713 (40.5)	1388 (38.7)	1325 (42.6)	
Caregiver's gender				.09
Female	5400 (80.6)	2922 (81.4)	2478 (79.7)	
Male	1301 (19.4)	669 (18.6)	632 (20.3)	
Caregiver's age (years)				<.001
＜30	1458 (21.8)	677 (18.9)	781 (25.1)	
30-39	4343 (64.8)	2408 (67.1)	1935 (62.2)	
≥40	900 (13.4)	506 (14.1)	394 (12.7)	
Number of children in the family				.51
1	3010 (44.9)	1627 (45.3)	1383 (44.5)	
≥2	3691 (55.1)	1964 (54.7)	1727 (55.5)	
Caregiver's education				<.001
Junior high school or lower	1131 (16.9)	514 (14.3)	617 (19.8)	
High or vocational school	1493 (22.3)	816 (22.7)	677 (21.8)	
College or higher	4077 (60.8)	2261 (63.0)	1816 (58.4)	

Discussion

Anticipated Findings

This protocol outlines a 12-month cluster randomized controlled trial designed to evaluate the effectiveness of iCURE, an intelligent, WeChat-based mHealth intervention for child unintentional injury prevention. The intervention is structured to align with the current primary health care service model in China. If proven effective, iCURE can be generalized and implemented nationwide as a low-cost and easy-to-use unintentional injury prevention intervention.

Childhood injuries result in high morbidity and mortality rates worldwide, including in China. Although injury prevention has been integrated into China's National Basic Public Health Service system, implementation has been limited due to workforce shortages, high workloads, inadequate training, and the absence of operational guidelines. This study addresses these gaps through the development and evaluation of iCURE. Similar mHealth initiatives have demonstrated effectiveness in child unintentional injury prevention [18-30]. However, unlike traditional app-based programs, iCURE operates via a widely used platform (WeChat) without requiring app installation and incorporates real-time, caregiver-reported injury monitoring. These features are particularly suited to resource-constrained settings and may support scalable implementation in LMICs.

The iCURE program adopts a structured and theory-informed design that is rooted in behavioral science and an injury prevention framework. By integrating an AI-assisted dissemination strategy that is guided by personalized risk assessment reports, it sends individualized and tailored education materials to each caregiver. Leveraging mHealth technologies, iCURE offers cost-effective, easily implementable, and real-time interactive solutions. In addition, the evaluation uses a rigorous cluster-randomized controlled trial design, with blinded outcome assessment and control for potential confounders. It includes clearly defined outcome measures and a standardized data collection process, enhancing the validity of findings.

Potential challenges to the study include inadequate caregiver enrollment, loss to follow-up, and reporting bias. To address these concerns, the study includes a multifaceted engagement

XSL•FO
RenderX

strategy that includes incentives, safety-themed campaigns, and active follow-up via automated reminders and telephone outreach. The involvement of health care providers in participant recruitment and follow-up is crucial to maintaining engagement throughout the study. Data will be collected every 3 months to minimize recall bias.

To minimize potential recall bias from caregivers' reports, we purposively adopted follow-up assessments every 3 months and strictly implemented standardized definitions of unintentional injury and prompts for minor events (eg, bandaging and disinfection). Nevertheless, such strategies cannot fully overcome recall biases as caregivers may interpret and treat injury events differently, and some might not report minor injuries that were not specified by the operational definition or were not perceived as injuries [53].

If proven effective, iCURE could enhance the capacity of China's primary health care system to deliver standardized, accessible child injury prevention services. It also offers a practical model for integrating digital tools into national child health programs in LMICs, where injury prevention resources are typically scarce.

Conclusions

This study responds to the urgent need for practical, evidence-based strategies to reduce childhood injury in China and other LMICs. Through a structured development process and rigorous evaluation, it aims to generate high-quality evidence on the effectiveness of an intelligent, system-integrated injury prevention program. If successful, iCURE could serve as a cost-effective, scalable solution for child injury prevention and surveillance, with potential for adaptation and dissemination through national public health platforms globally.

Acknowledgments

The authors acknowledge the support of the local government departments for coordinating the participating schools. The study was funded by the National Natural Science Foundation of China (Young Scientist Fund; 82204165).

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

LL and YY contributed to the conceptualization and design of the protocol and drafted the initial manuscript. YL, JL, JH, DCS, PN, and ZR contributed to the review and editing of the manuscript. GH contributed to the study design, provided supervision, and finalized the manuscript. All authors read and approved the final version of the protocol. The funder had no role in the study design, interpretation, or writing of the manuscript.

Conflicts of Interest

The authors declare that the research was and will be conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Multimedia Appendix 1
Home page of web platform.
[DOCX File, 359 KB - publichealth_v11i1e76195_app1.docx]

Multimedia Appendix 2
Home page and modules of WeChat service account. Note: Eight images within the figure were derived from the WeChat service account "iCURE" (Intelligent Child Unintentional-Injury Reduction & Education) that was developed by the research team for unintentional injury prevention among children aged 0-6 and will be tested in this trial. (a) Home page of WeChat service account, (b) questionnaire filling interface, (c) injury risk assessment questionnaire, (d) feedback of risk assessment results, (e) feedback with tailored recommendations, (f) injury reporting module, (g) interactive Q&A (questions and answers) module, and (h) tailored education module.
[DOCX File, 621 KB - publichealth_v11i1e76195_app2.docx]

Checklist 1
SPIRIT checklist
[DOC File, 130 KB - publichealth_v11i1e76195_app3.doc]

References

1. Injuries and violence: the facts 2014. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789241508018 [accessed 2025-04-17]

XSL•FO
RenderX

2. GBD results. Institute for Health Metrics and Evaluation. 2024. URL: https://vizhub.healthdata.org/gbd-results/ [accessed 2025-06-28]
3. Zhou J, Zhao M, Huang H, et al. Injury mortality of children and adolescents aged 0-19 years - China, 2010-2021. China CDC Wkly 2024 Apr 5;6(14):294-299. [doi: 10.46234/ccdcw2024.057] [Medline: 38634100]
4. National Center for Chronic and Non-communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention. China Death Cause Surveillance Datasets 2021: Beijing: China Science and Technology Press; 2022.
5. World report on child injury prevention. World Health Organization. 2008. URL: https://www.who.int/publications/i/item/9789241563574 [accessed 2025-06-28]
6. Perrin EM, Skinner AC, Sanders LM, et al. The injury prevention program to reduce early childhood injuries: a cluster randomized trial. Pediatrics 2024 May 1;153(5):e2023062966. [doi: 10.1542/peds.2023-062966] [Medline: 38557871]
7. Sleet DA. The global challenge of child injury prevention. Int J Environ Res Public Health 2018 Sep 4;15(9):1921. [doi: 10.3390/ijerph15091921] [Medline: 30181462]
8. Feng YH, Li XH, Zhao J, Shen HL, Xia JW, Lv J. Health management policies for 0–6 year old children in shanghai city from 2009 to 2023: an analysis of relevant documents issued by government agencies. Chin J Public Health 2024;40(3):353-357. [doi: 10.11847/zgggws1142668]
9. Technical specification for health management of children aged 0-6 years (ws/t 479-2015). National Health Commission of China. 2015. URL: https://www.nhc.gov.cn/wjw/c100311/201507/baec7f22eee148b29554e7e48ef7e92b.shtml [accessed 2025-04-17]
10. Xiao F, Li RL, Wang LH, Yin DL, Zheng XG, Yin T, et al. Analysis on the implementation situation of health management service for 0-6-year-old children in china. Chin Matern Child Health Care 2014;29(35):5741-5744. [doi: 10.7620/zgfybj.j.issn.1001-4411.2014.35.01]
11. Wang HS, Xia HB, Wei G. Study on a novel health management model for children aged 0-6 in longgang district of shenzhen. J Qiqihar Med Univ 2020;41(6):733-735. [doi: 10.3969/j.issn.1002-1256.2020.06.033]
12. Notice on issuing the series of technical guidelines for injury intervention. 2011. URL: https://www.nhc.gov.cn/jkj/c100063/201109/1396921204e34c7eb11e99c22bc9a8da.shtml [accessed 2025-04-17]
13. Global strategy on digital health 2020-2025. World Health Organization. 2021. URL: https://iris.who.int/handle/10665/344249 [accessed 2025-06-28]
14. Han WP, Lei Y, Yin XM, Chen YJ, Yang JJ. The application and prospects of healthcare big data in children health management. Chin J Health Inform Manag 2023;20(4):501-507. [doi: 10.3969/j.issn.1672-5166.2023.04.01]
15. Healthy children action improvement plan (2021-2025). National Health Commission of China. 2021. URL: https://www.gov.cn/zhengce/zhengceku/2021-11/05/content_5649019.htm [accessed 2025-04-17]
16. Cortez NG, Cohen IG, Kesselheim AS. FDA regulation of mobile health technologies. N Engl J Med 2014 Jul 24;371(4):372-379. [doi: 10.1056/NEJMhle1403384] [Medline: 25054722]
17. Khokhar S, Holden J, Toomer C, Del Parigi A. Weight loss with an AI-powered digital platform for lifestyle intervention. Obes Surg 2024 May;34(5):1810-1818. [doi: 10.1007/s11695-024-07209-1] [Medline: 38573389]
18. Gielen AC, Bishai DM, Omaki E, et al. Results of an RCT in two pediatric emergency departments to evaluate the efficacy of an m-Health educational app on car seat use. Am J Prev Med 2018 Jun;54(6):746-755. [doi: 10.1016/j.amepre.2018.01.042] [Medline: 29656914]
19. Dixon CA, Ammerman RT, Johnson BL, et al. A randomized controlled field trial of iBsafe-a novel child safety game app. Mhealth 2019;5:3. [doi: 10.21037/mhealth.2019.01.02] [Medline: 30842951]
20. McKenzie LB, Roberts KJ, McAdams RJ, et al. Efficacy of a mobile technology-based intervention for increasing parents' safety knowledge and actions: a randomized controlled trial. Inj Epidemiol 2021 Oct 1;8(1):56. [doi: 10.1186/s40621-021-00350-w] [Medline: 34593040]
21. Omaki E, Shields WC, McDonald E, et al. Evaluating a smartphone application to improve child passenger safety and fire safety knowledge and behaviour. Inj Prev 2017 Feb;23(1):58. [doi: 10.1136/injuryprev-2016-042161] [Medline: 27597399]
22. Arbogast H, Burke RV, Muller V, Ruiz P, Knudson MM, Upperman JS. Randomized controlled trial to evaluate the effectiveness of a video game as a child pedestrian educational tool. J Trauma Acute Care Surg 2014 May;76(5):1317-1321. [doi: 10.1097/TA.0000000000000217] [Medline: 24747467]
23. Choi Y, Ahn HY. Developing and evaluating a mobile-based parental education program for preventing unintentional injuries in early childhood: a randomized controlled trial. Asian Nurs Res (Korean Soc Nurs Sci) 2021 Dec;15(5):329-336. [doi: 10.1016/j.anr.2021.12.001] [Medline: 34915186]
24. Park IT, Oh WO, Jang GC, Han J. Effectiveness of mHealth-Safe Kids Hospital for the prevention of hospitalized children safety incidents: a randomized controlled trial. J Nurs Scholarsh 2021 Sep;53(5):623-633. [doi: 10.1111/jnu.12693] [Medline: 34396685]
25. Burgess J, Watt K, Kimble RM, Cameron CM. Combining technology and research to prevent scald injuries (the Cool Runnings Intervention): randomized controlled trial. J Med Internet Res 2018 Oct 10;20(10):e10361. [doi: 10.2196/10361] [Medline: 30305263]

XSL•FO
RenderX

26. Chow CB, Wong WHS, Leung WC, et al. Effectiveness of a technology-based injury prevention program for enhancing mothers' knowledge of child safety: protocol for a randomized controlled trial. JMIR Res Protoc 2016 Oct 31;5(4):e205. [doi: 10.2196/resprot.6216] [Medline: 27799138]
27. He J, Wang W, Ning P, et al. Assessing the effectiveness of an app-based child unintentional injury prevention intervention for caregivers of rural Chinese preschoolers: protocol for a cluster randomized controlled trial. BMC Public Health 2021 Nov 20;21(1):2137. [doi: 10.1186/s12889-021-12156-y] [Medline: 34801006]
28. Ning P, Cheng P, Schwebel DC, et al. An app-based intervention for caregivers to prevent unintentional injury among preschoolers: cluster randomized controlled trial. JMIR Mhealth Uhealth 2019 Aug 9;7(8):e13519. [doi: 10.2196/13519] [Medline: 31400105]
29. Wong RS, Tung KT, Wong HT, et al. A mobile game (Safe City) designed to promote children's safety knowledge and behaviors: protocol for a randomized controlled trial. JMIR Res Protoc 2020 Jun 12;9(6):e17756. [doi: 10.2196/17756] [Medline: 32530436]
30. Yan S, Yang J, Fu J, et al. Assessing an app-based child restraint system use intervention in China: an RCT. Am J Prev Med 2020 Sep;59(3):e141-e147. [doi: 10.1016/j.amepre.2020.02.003] [Medline: 32334955]
31. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346(jan08 15):e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
32. Butcher NJ, Monsour A, Mew EJ, et al. Guidelines for reporting outcomes in trial protocols: the SPIRIT-Outcomes 2022 extension. JAMA 2022 Dec 20;328(23):2345-2356. [doi: 10.1001/jama.2022.21243] [Medline: 36512367]
33. Wang ZY, Hu M, Yu TL, Yang J. Incidence of injuries and its influencing factors among children aged 0-6 years in urban area of changsha city. Chin J Health Stat 2020;37(1):14-17. [doi: 10.3969/j.issn.1002-3674.2020.01.004]
34. Li L, Wang GY, Zhao DM, Qu JS. Effectiveness evaluation for pre-school children with health education to reduce unintentional injuries. Chin J Child Health Care 2011;19(11):1056-1058.
35. Gianini RJ, Caneto NF, Junqueira NM, et al. Precision assessment of a hypertension prevalence survey. BMC Public Health 2024 Aug 12;24(1):2188. [doi: 10.1186/s12889-024-19626-z] [Medline: 39135026]
36. Li J, Ning P, Cheng P, et al. Factors associated with dropout of participants in an app-based child injury prevention study: secondary data analysis of a cluster randomized controlled trial. J Med Internet Res 2021 Jan 29;23(1):e21636. [doi: 10.2196/21636] [Medline: 33512318]
37. Liao SH. Expert system methodologies and applications—a decade review from 1995 to 2004. Expert Syst Appl 2005 Jan;28(1):93-103. [doi: 10.1016/j.eswa.2004.08.003]
38. Li RT, Ning PS, Li J, Wang WH, Hu GQ. Review on the progress regarding the intervention mapping protocol. Zhonghua Liu Xing Bing Xue Za Zhi 2019 Apr 10;40(4):488-492. [doi: 10.3760/cma.j.issn.0254-6450.2019.04.022] [Medline: 31006214]
39. Li GH, Baker SP. Injury Research: Theories, Methods, and Approaches: Springer Science+Business Media; 2012. [doi: 10.1007/978-1-4614-1599-2]
40. Gielen AC, Sleet DA, DiClemente RJ. Injury and violence prevention: behavioral science theories, methods, and applications. Inj Prev 2007 Oct;13(5):360-360. [doi: 10.1136/ip.2007.016279] [Medline: 17446251]
41. Li G, Baker SP. Injury Research: Springer; 2012.
42. National basic public health service project management platform (child health management for 0–6 years). Department of Primary Health, National Health Commission of China. 2017. URL: http://www.nbphsp.org.cn/jbgw/06rt/ [accessed 2025-04-17]
43. Sutchritpongsa S, Sangwisit S, Sonjaipanich S. Parental awareness of household injury prevention: adequacy of anticipatory guidance for well childcare. J Med Assoc Thai 2013 Dec;96(12):1531-1535. [Medline: 24511716]
44. Wang ZY, Ran Y, Nian HY, Shao K, Yu TL, Hu M. Development and performance test of the environment scale of unintentional injury in the home for children aged 0-6 years old in urban area of China. Zhonghua Yu Fang Yi Xue Za Zhi 2020 Feb 6;54(2):139-143. [doi: 10.3760/cma.j.issn.0253-9624.2020.02.005] [Medline: 32074699]
45. Song J, Shao K, Zhang BL, Wan SQ, Hu M. Development and validation of risk behavior scale for in-home unintentional injury in urban children aged 0-6 years in China. Zhonghua Liu Xing Bing Xue Za Zhi 2021 Sep 10;42(9):1662-1669. [doi: 10.3760/cma.j.cn112338-20201105-01310] [Medline: 34814598]
46. Wang X, Ran N, Li KN, Sun YR, Niu YS, Yang XL. Quality evaluation and content analysis of guidelines and expert consensus for prevention of unintentional suffocation in children under three. Mod Prev Med 2022;49(10):1780-1786.
47. Cao XL. Accident risk assessment tool to research in preschool children [master's thesis]. : Nanhua University; 2017.
48. Wang SY. Injury Epidemiology: People's Medical Publishing House; 2003.
49. Yang GT, Huang BL, Liu ZJ, Li WY, Lin JL, Chen JF, et al. Research on the knowledge, attitude and practice in parents of preschool children regarding the occurrence of unintentional injury. Inj Med (Electron Ed) 2019;8(3):16-23. [doi: 10.3868/j.issn.2095-1566.2019.03.004]
50. Ning P, Gao D, Cheng P, et al. Needs analysis for a parenting app to prevent unintentional injury in newborn babies and toddlers: focus group and survey study among Chinese caregivers. JMIR Mhealth Uhealth 2019 Apr 30;7(4):e11957. [doi: 10.2196/11957] [Medline: 31038465]

XSL•FO
RenderX

51. Schwebel DC, Gaines J. Pediatric unintentional injury: behavioral risk factors and implications for prevention. J Dev Behav Pediatr 2007 Jun;28(3):245-254. [doi: 10.1097/01.DBP.0000268561.80204.2a] [Medline: 17565295]
52. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. PLoS Med 2025 Apr;22(4):e1004587. [doi: 10.1371/journal.pmed.1004587] [Medline: 40228477]
53. Zonfrillo MR, Myers RK, Durbin DR, Curry AE. Validation of parent-reported injuries to their children. Clin Pediatr (Phila) 2015 Sep;54(10):983-986. [doi: 10.1177/0009922814566931] [Medline: 25573947]

Abbreviations

AI: artificial intelligence
CONSORT: Consolidated Standards of Reporting Trials
iCURE: Intelligent Child Unintentional-Injury Reduction & Education
LMIC: low- and middle-income country
mHealth: mobile health
Q&A: questions and answers
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TIPP: The Injury Prevention Program
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 20.04.25; peer-reviewed by SCL Au, T Amusa; revised version received 10.06.25; accepted 11.06.25; published 18.07.25.

Please cite as:
Yuan Y, Jing Y, Li J, He J, Schwebel DC, Ning P, Rao Z, Li L, Hu G
Evaluating the Effectiveness of an Intelligent mHealth Intervention for Child Unintentional Injury Prevention: Protocol for a Cluster Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e76195
URL: https://publichealth.jmir.org/2025/1/e76195
doi:10.2196/76195

© Yang Yuan, Yuying Jing, Jiyu Li, Jieyi He, David C Schwebel, Peishan Ning, Zhenzhen Rao, Li Li, Guoqing Hu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Substance-Related Acute Toxicity Deaths in Canada From 2016 to 2017: Protocol for a Retrospective Chart Review Study of Coroner and Medical Examiner Files

Jenny Rotondo[1], MHSc; Amanda VanSteelandt[1], PhD; Fiona Kouyoumdjian[2], MPH, MD, PhD; Matthew J Bowes[3], MD; Tanya Kakkar[1], MPH; Graham Jones[4], PhD; Brandi Abele[5], Addictions Counsellor Diploma; Regan Murray[6], MPH; Emily Schleihauf[3,7], MAE; Jessica Halverson[8], MPH, MSW; Jennifer Leason[9], PhD; Dirk Huyer[10], MD; Beth Jackson[11], PhD; Songul Bozat-Emre[12,13], PhD; Devanshi Shah[1], MPH; Erin E Rees[14], PhD

[1]Substance-Related Harms Division, Health Promotion and Chronic Disease Prevention Branch, Public Health Agency of Canada, Ottawa, ON, Canada
[2]Ontario Ministry of Health, Toronto, ON, Canada
[3]Nova Scotia Medical Examiner Service, Department of Justice, Dartmouth, NS, Canada
[4]Department of Laboratory Medicine and Pathology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada
[5]Board of Directors, Canadian Association of People Who Use Drugs, Saskatoon, SK, Canada
[6]Collaboration on Death Investigation Data, Health Promotion and Chronic Disease Prevention Branch, Public Health Agency of Canada, Ottawa, ON, Canada
[7]Canadian Public Health Service, Emergency Management Branch, Public Health Agency of Canada, Ottawa, ON, Canada
[8]Migration Health Branch, Immigration, Refugees and Citizenship Canada, Ottawa, ON, Canada
[9]Department of Anthropology and Archaeology, University of Calgary, Calgary, AB, Canada
[10]Office of the Chief Coroner for Ontario, Ontario Forensic Pathology Service, Toronto, ON, Canada
[11]Health Equity Policy Division, Strategic Policy Branch, Public Health Agency of Canada, Ottawa, ON, Canada
[12]Provincial Information Management and Analytics, Manitoba Health, Winnipeg, MB, Canada
[13]Community Health Sciences, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada
[14]Public Health Risk Sciences Division, National Microbiology Laboratory, Public Health Agency of Canada, Saint-Hyacinthe, QC, Canada

Corresponding Author:
Jenny Rotondo, MHSc
Substance-Related Harms Division
Health Promotion and Chronic Disease Prevention Branch
Public Health Agency of Canada
785 Carling Avenue
Ottawa, ON, K1A 0K9
Canada
Phone: 1 613 614 9487
Email: jenny.rotondo@phac-aspc.gc.ca

Abstract

Background: Canada continues to experience a national overdose crisis. While studies are available at the regional, provincial, and territorial levels, detailed national data regarding the burden and context of substance-related acute toxicity deaths are limited, particularly in subpopulations. In response to the overdose crisis, the Public Health Agency of Canada, in collaboration with provincial and territorial ministries of health and chief coroner and chief medical examiner offices, has undertaken a national chart review study.

Objective: This study was conducted to describe and compare the characteristics of substance-related acute toxicity deaths that occurred in Canada between 2016 and 2017, including descriptions of those who died, the substances involved, and the circumstances surrounding their deaths. This paper describes the study methodology in detail.

Methods: This retrospective, population-based, and cross-sectional study involved the review of coroner and medical examiner files for deaths that met the study case definition. Data were collected on demographic and socioeconomic characteristics, medical and substance use history, proximal circumstances surrounding the death, and toxicology findings using a standardized data collection tool that underwent 2 pilot studies. Data abstractors underwent training and regular intrarater reliability exercises with

XSL•FO
RenderX

a fictitious death investigation file. Data quality was assessed based on the consistency of abstractor intrarater reliability scores and the completeness of core variables and variables for key concepts. Data were linked to national datasets to allow for the examination of area-level geographic and socioeconomic characteristics. Descriptive analyses will examine differences across subpopulations and the general Canadian population. Latent class, spatiotemporal, qualitative, and premature death analyses are also planned. Where possible, analyses will be stratified by the manner of death and sex.

Results: The study began in the summer of 2018, and abstraction was delayed due to the COVID-19 pandemic. All activities are expected to be completed by early 2025. A total of 9414 coroner and medical examiner files met the study case definition. Most abstractors (25/26, 96%) met the established threshold for consistency throughout abstraction without the need for remedial training. In general, core study variables, including geographic variables and substances contributing to death, had very good availability. Study variables related to the person's health, history of substance use, and events surrounding the acute toxicity event were available for most records. Socioeconomic variables and variables describing socially constructed identities and potentially traumatic life events were mostly unavailable.

Conclusions: This study provides the most detailed national information on substance-related acute toxicity deaths in Canada to date and can serve as a pre–COVID-19 pandemic baseline for assessing the evolution of the overdose crisis. Results can inform policies and programs to address the overdose crisis, the development of common approaches to medicolegal death investigations, and future research activities.

(JMIR Public Health Surveill 2025;11:e49981) doi:10.2196/49981

KEYWORDS

acute toxicity; Canada; chart review study; coroner; death investigations; drug overdose; medical examiner; mortality; poisoning; protocol

Introduction

Background

Canada continues to experience a national overdose crisis, with a large burden of harms that are inequitably distributed throughout the population. Between January 2016 and June 2023, a total of 40,642 people died of opioid toxicity alone, with associated mortality rates increasing over time [1]. Compared to other countries where data are available, the United States and Canada have seen rapid escalations and persistently high accidental acute toxicity mortality rates [2]. Information on the characteristics and circumstances of those who have died is critical to informing and evaluating strategies aimed at preventing harms.

High-quality data are particularly needed at the national level to estimate burden and explore acute toxicity deaths among subpopulations, thereby enhancing our understanding of inequities and risk factors [3-6]. Although numerous studies and reports have described substance-related acute toxicity deaths at the regional, provincial, and territorial levels [7-24], they used a variety of methods and case definitions, making it difficult to aggregate findings at the national level and compare them across provinces and territories. Vital statistics data from medical certificates of death are standardized and routinely collated at the national level in Canada [25]. While this is a common source for acute toxicity mortality data [26,27], it lacks detailed information on the characteristics of those who died and the specific substances that contributed to their deaths [6,28,29].

By contrast, coroners and medical examiners (CMEs) collect demographic, socioeconomic, health, and other risk factor data during death investigations, as well as information on the circumstances of the death and toxicological findings. Although national opioid toxicity mortality surveillance uses CME data reported by provincial and territorial ministries of health to the Public Health Agency of Canada (PHAC), data are submitted at an aggregate level by some ministries, with stratification by a small number of variables of interest [1]. As demonstrated in similar studies, additional valuable information can be gleaned from complete CME files [30-32].

Objectives

Our study objective was to describe and compare the characteristics of substance-related acute toxicity deaths that occurred in Canada in 2016 and 2017, including descriptions of those who died, the substances involved, and the circumstances surrounding their death. This paper describes the study methodology in detail, as well as the process for developing study material and findings regarding variable completion. It may be used as a reference document and to support the development of similar studies in the future.

Methods

Study Design

This retrospective, population-based, and cross-sectional study involved a review of the CME files of people who died of substance-related acute toxicity in Canada between 2016 and 2017. Upon identifying CME files that met the study case definition, data were abstracted using a standardized data collection tool.

Personnel

At the study outset, the PHAC study team assembled a coinvestigator team to inform the development and rollout of the study protocol, analysis planning, interpretation of findings, and production of knowledge translation products. The coinvestigator team consisted of 14 people with expertise in the CME system, death investigations, toxicology, pharmacology, social determinants of health, surveillance, epidemiology, harm

XSL•FO
RenderX

reduction, the health of Indigenous people, qualitative research, corrections, social services, medicine, provincial and territorial health authorities, the federal Health Portfolio, public health in Canada, and lived experience with substance use. Coinvestigators were also well connected with other substance-related surveillance, research, and prevention activities occurring in Canada.

In addition to the PHAC study and coinvestigator teams, data abstractors and students were recruited to support data collection, preparation, and analysis activities. Additional subject matter experts were invited to participate in analysis-specific project teams. For example, people with lived experience and people with expertise in youth and older adult health, suicide, emergency medicine, substance use, pain research, race and ethnicity studies, addiction medicine, and corrections have joined specific projects focused on these topics. Finally, a team of Indigenous researchers worked with coinvestigators and project team members to develop products focused on First Nations, Métis, and Inuit people who died due to acute toxicity, with a focus on protective factors.

Literature Review

To inform the development of the study protocol, the data collection tool, and planned analyses, a literature review was conducted to identify knowledge gaps with respect to substance-related acute toxicity deaths at the national level in Canada. Collaborating with a librarian, a search strategy with inclusion and exclusion criteria and filters was developed. The MEDLINE, Embase, and SCOPUS databases were searched, and 2 team members reviewed titles and abstracts for relevance. A gray literature search was also conducted to identify relevant reports not published in peer-reviewed journals. Finally, additional articles were identified by reviewing the reference lists of relevant articles and reports. Information was abstracted using a standardized data collection tool and synthesized based on themes.

Although the review concentrated on the Canadian context, evidence from other countries was drawn on when limited Canadian data were available. The review focused primarily on substance use; pharmaceutical prescribing; risk factors; proximal circumstances; and severe harms related to substance use, including emergency department visits, hospitalizations, and deaths.

Consultations

In addition to the literature review and expertise of the coinvestigator team, the study's design and rollout were informed by consultations with various groups, including chief coroners and chief medical examiners (CCCMEs), CCCME office staff, PHAC public health officers, other government of Canada departments (including Health Canada and Indigenous Services Canada), and staff within provincial and territorial ministries of health. The PHAC study team also consulted with national Indigenous organizations during the planning and analysis phases of the study, as well as a Government of Canada council of people with lived and living experience during the analysis phase.

Study Population

The target population for this study included all people who died in Canada in 2016 and 2017 due to acute toxicity. The accessible population included people whose deaths (1) were investigated by a CME, (2) had a CME report available for data abstraction, and (3) were identified as meeting the case definition. Given the resource intensity of conducting a nationwide chart review study, only 2 years' worth of data were collected; 2016 and 2017 were selected, as these were the latest 2 years for which data were likely available when work on the study began in 2018.

Case Definition

The study case definition was any individual who died in Canada between January 1, 2016, and December 31, 2017, after an acute intoxication or toxicity resulting from the direct effects of the administration of exogenous substances where ≥1 of the substances was a drug or alcohol. The study case definition was based on the national CME case definition for an apparent opioid-related death (as of March 1, 2017) [1]. "Acute" refers to adverse health effects that occur within a short period (measured in minutes, hours, or days) following dosage [33]. Substances included alcohol; controlled, illegal, and prescription drugs; new psychoactive substances [34]; over-the-counter pharmaceutical products; and chemicals not intended for human use (eg, nonpharmaceutical inhalants, industrial or household chemicals, or veterinary drugs). Inclusion and exclusion criteria are described in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for our national chart review study of substance-related acute toxicity deaths.

Inclusion criteria

- The death occurred in the reporting province or territory.
- The death occurred in 2016 or 2017.
- The investigation is either ongoing (open, preliminary, or active) or closed (certified or completed).
- The manner of death was deemed an accident (unintentional), suicide (intentional), or undetermined.
- According to the (1) death certificate, (2) autopsy summary, or (3) coroner or medical examiner report, acute toxicity resulting from the direct effects of the administration of exogenous substances where ≥1 of the substances was a drug or alcohol was identified as having caused or contributed to the death.

Exclusion criteria

- The death was solely caused by any of the following: chronic substance use (while people who died only due to the effects of chronic substance use were excluded from this study, people who experienced chronic substance use but died due to acute toxicity were included), medical assistance in dying, palliative or comfort care, homicide, occupational exposure, trauma where an intoxicant contributed to the circumstances of the injury (eg, death caused by a motor vehicle collision involving a driver impaired by an intoxicant), other adverse drug effects (eg, anaphylactic shock), or acute toxicity due to products of combustion (eg, carbon monoxide, carbon dioxide, sulfur dioxide, and nitrogen oxides).

Data Sources

Legislation exists in all Canadian provinces and territories requiring that deaths believed to be the result of violence, accident, suicide, negligence, misconduct, or malpractice be reported to a CME. These acts also require that a CME investigate these deaths to establish the identity of the deceased; the date, time, and place of death; the circumstances under which the death occurred; the cause of death; and the manner of death. Case files generally include the CME's medical certificate of death, the CME's summary report, toxicology test results, an autopsy or external examination report (if performed), and continuation notes (file communication records). In addition, they may include a police report, an emergency medical service record, the decedent's recent medical history (the period differs across CMEs, provinces, and territories), scene notes and photographs, and correspondence with individuals relevant to the case (eg, family physicians or other professionals who knew the person who died). Although there are similarities in how CME investigations are conducted across Canada, the information collected varies over time across provinces and territories, the investigating CME, and the circumstances of the death [35].

The primary source of data for this study was physical and electronic CME files. While every effort was made to obtain complete information for each case, physical files were not always accessible, and the COVID-19 pandemic resulted in additional challenges to the review of case files. For all British Columbia cases and some Ontario cases, data were only obtained from electronic coroner databases and mapped to corresponding variables in the study data collection tool where variables were equivalent. Additional population size and area-level data were obtained from Statistics Canada's 2016 Census [36], Statistics Canada's Postal Code Conversion File Plus [37], Statistics Canada's Index of Remoteness [38], Statistics Canada's Canadian Index of Multiple Deprivation [39], and Environment and Climate Change Canada [40] to support specific analyses.

Data Collection Tool

Potential variables of interest were identified from (1) the study's literature review, (2) CME investigative tools already in use by provinces and territories, and (3) consultations with key stakeholders. Potential variables' usefulness, relevance to research questions, comparability with other studies, and privacy concerns were considered before their inclusion in the data collection tool. As one of the study goals was to report on data availability within CME files, variables of interest with low expected availability were not excluded. Data were collected on demographic, socioeconomic, and risk factors; drug, medical, and substance use history; proximal circumstances surrounding the death; and toxicology findings for each person who died (Multimedia Appendix 1 provides the variable list). Identifying variables, such as the name and address of the person who died, were collected only if requested by the CCCME office for internal use and were not sent to the PHAC study team.

A standardized data collection tool was built using Microsoft Access (Multimedia Appendix 2 provides screenshots of the tool), and an accompanying detailed user manual and data dictionary were created to enhance data quality and consistency. The data collection tool and accompanying material were pilot-tested twice in 6 provinces and further refined before the official start of the study.

Data Collection

To facilitate data access and sharing between CCCME office staff and the PHAC study team, 13 data-sharing or research agreements were developed between the PHAC and each CCCME office. Data sharing agreements were also developed between the PHAC and several provincial and territorial ministries of health if required.

Data collection was conducted by individuals approved by collaborators within each CCCME office, including public health officers, CCCME office staff, nurses, epidemiologists, health information analysts, and health sciences and epidemiology students. All abstractors completed a privacy course; received extensive standardized training on how to collect data from CME investigation files; and were given time

XSL•FO
RenderX

to review the data collection tool, user manual, and data dictionary. Abstractors were also required to demonstrate their ability to abstract data using a fictitious test case file before data collection began.

Case identification was done using available information systems. Initial broad searches were conducted by CCCME offices or data abstractors to identify anyone who could possibly meet the study case definition. Data abstractors then determined whether each case met the case definition while reviewing the CME files. In offices where it was possible to identify potential cases based on an *International Classification of Disease, 10th Revision*, underlying cause of death code, relevant categories included X40 to X45, X47, and X49 (accidental poisoning); X60 to X65, X67, and X69 (intentional self-poisoning); and Y10-Y15, Y17, and Y19 (poisoning with undetermined intent) [41].

As part of the data collection process, abstractors were instructed to regularly run built-in queries to check data completion. This supported quality assurance activities and decreased the need to return to CCCME offices for further data collection.

Ethical Considerations

Study data management activities were developed to ensure data confidentiality. All data collected for the study existed electronically in password-protected datasets that were transferred to the PHAC study team via a secure data transfer method. Received data were stored in a restricted PHAC network folder accessible only to study staff responsible for data management or analysis. All study coinvestigators and data abstractors were asked to sign confidentiality agreements.

To ensure confidentiality in knowledge translation products, random rounding to base 3 has been applied to the raw counts of all tables displaying results, counts <10 will be suppressed, and variable categories will be collapsed where appropriate. In random rounding to base 3, values that are multiples of 3 do not change, and those that are not multiples of 3 have a two-thirds chance of rounding to the nearest multiple of 3 and a one-third chance of rounding to the second nearest multiple of 3 [42]. This will perturb the count to within 2 units of the true value. As individual counts, subtotals, and totals are independently rounded, they may not add up when summed. Percentages, which are calculated using the rounded values, also may not add up to 100%. All study products will be audited to ensure that the same cell value is rounded in the same direction across tables.

As the study was part of a time-limited, exploratory research activity, research ethics board review and approval were sought. The study protocol was reviewed and approved by the Health Canada and PHAC Research Ethics Board (REB 2018-027P), the University of Manitoba Health Research Ethics Board (HS22710), and the Newfoundland and Labrador Health Research Ethics Board (20200153). It also underwent privacy assessments by the PHAC; the Manitoba Health, Seniors, and Active Living Health Information Privacy Committee; and the British Columbia Office of the Chief Information Officer.

Data Preparation

Collected data underwent multiple stages of preparation, including case confirmation, comment review, redundancy checks, and variable cleaning. In situations where abstractors were unsure whether a case met the study case definition, they were asked to provide a nonidentifying description of the situation, which was shared with a subset of the coinvestigator team for decision-making. Similarly, when abstractors were unsure of how to record information from a case file, they were instructed to leave detailed comments in the database explaining the issue. These comments were later reviewed by the PHAC study team, and corresponding variables were adjusted as required.

Because the data collection tool had to allow abstractors to abstract data in a variety of formats, the database contained redundant variables that captured the same concept in slightly different ways. The study team reviewed the logic of redundant variables to ensure that variables with missing information were completed with information from redundant variables where possible. For example, if a person was documented to have been enrolled in an opioid agonist treatment program, their medical history was updated to include a history of substance use disorder.

Finally, the dataset underwent extensive data cleaning activities, particularly for values reported in free-text fields. Free-text fields were reviewed by the study team to correct spelling and match with existing variable values missed by abstractors. For more complex variables, the entered values were categorized according to a framework to simplify future analyses (eg, specific medical conditions were categorized according to types of medical conditions relevant to acute toxicity). Finally, some free-text variables were left “as is” for future thematic analysis.

Data Quality Indicators

Consistency of Data Collection

To ensure that data were collected in a consistent manner across data abstractors and throughout the data collection period, abstractors were required to complete an intrarater reliability (IaRR) exercise. After completing training activities and before beginning data collection, each abstractor was asked to abstract data from a fictitious death investigation file, and the abstracted data were compared to a gold standard. This was done again at the midpoint and at the end of data collection, allowing for the IaRR to be assessed at these points. Additional IaRR assessments were also conducted between any breaks in data collection, if required.

To develop the gold standard, study coinvestigators and PHAC study team members created a fictitious CME test file. Next, several PHAC study team members not involved in data collection independently used the data collection tool to collect data from the CME test file. The data collected were reviewed by members of the coinvestigator and PHAC study teams to identify inconsistencies in what was collected and to come to a consensus regarding what should have been captured in the data collection tool’s fields. Because the data collection tool includes hundreds of fields, it was not optimal to base IaRR measurements on all data variables. Instead, the IaRR

XSL•FO
RenderX

measurements were based on a subset of variables that were deemed to be of critical importance, as well as variables for which data collection was inconsistent during the pilot phase.

As variables were mostly categorical, IaRR was measured using Cohen κ, which accounts for the extent of agreement possible, having controlled for consensus by chance [43]. In instances where variables were not categorical (eg, dates), κ was based on whether the abstracted fields were consistent with the gold standard. The abstracted field was considered consistent only if blank variables in the gold standard were also blank in the data abstracted during the IaRR exercise or if the IaRR exercise data captured the same information in the same variable field as the gold standard. All other scenarios would deem the field to be discordant with the gold standard. This approach dichotomized nonnominal data so that they could contribute to κ coefficient calculations.

The baseline measurement for each data abstractor was taken before starting data collection for the study. No specific feedback on their performance was provided, as they were asked to abstract the same file again over the course of data collection for the IaRR assessments. Generally, a minimum value of 0.60 was expected for κ values calculated against the gold standard, as per the approach recommended by Landis and Koch [44]; values below this threshold resulted in remedial abstraction training (Table 1). IaRR κ values calculated against previous time point measurements were considered alongside the IaRR κ values calculated against the gold standard to determine the need for remedial training at each time point. This was to ensure that any decreases to the κ value were appropriately interpreted in cases where discordances between time points resulted from incorrect responses being changed to correct responses.

Table 1. How data abstractor intrarater reliability κ values were interpreted for our national chart review study of substance-related acute toxicity deaths.

κ		Actions
Test vs gold standard	Test between time points	
Low (<0.60)	High (≥0.60)	Abstractor is consistent across time points, but accuracy is problematic. Remedial action is required.
High (≥0.60)	Low (<0.60)	Abstractor is exhibiting improved accuracy. Provided that the gold standard κ is >0.60, no remedial action is required.
Low (<0.60)	Low (<0.60)	Abstractor is not exhibiting acceptable accuracy. Further investigation is needed to determine why IaRR κ remains low and why remedial action is required.
High (≥0.60)	High (≥0.60)	Abstractor exhibits consistent, high-quality abstraction. No remedial action is required.

Data Completeness

The availability of information for core variables and a set of variables capturing key concepts in the data collection tool was used as an indicator of data completeness. For each data source, this was measured as the proportion of records missing information for each core variable and key concept.

Data Analysis

The variables selected for analysis were based on either 1) hypothesized relationships with substance use, acute toxicity, and acute toxicity mortality, or 2) as theoretical intervention points to prevent acute toxicity deaths (eg, contacts with the health care system). Descriptive statistics were used to compare subgroups among people who died of acute toxicity (within the dataset), as well as to compare people who died of acute toxicity with the general Canadian population. Data obtained through linkage to Statistics Canada's Postal Code Conversion File Plus [37] and the Canadian Index of Multiple Deprivation [39] based on the postal code or municipality of residence, acute toxicity event, and death allows for the assessment of area-level geographic and socioeconomic characteristics.

Many of the study's variables have missing data because only information that was available in CME files was used for data abstraction. As cases with missing data will not be removed from analyses unless they are systematically unavailable for a province or territory, most of the descriptive analyses will present minimum proportions or rates; that is, the minimum number of people who died of acute toxicity with a given characteristic. This is an important consideration for comparisons with the general population, as proportions or rates from the study population that are lower than those observed in the general population may be due to missing data. The true value could be lower than, equal to, or higher than the national statistic. However, if the study population's proportion or rate is higher than that of the general population, we expect that the true value is at least as high as that observed in the study. In situations where it is known that a variable was only available for select provinces and territories, subnational analyses may be performed.

More complex analyses of the study dataset are also planned. These include using latent class analysis to identify subpopulations among those who died of acute toxicity, spatiotemporal analysis to determine whether geographic and temporal clustering exists in the distribution of acute toxicity deaths, and a case-crossover analysis of weather patterns and acute toxicity deaths. Several free-text variables will also be used for thematic qualitative analyses. Finally, data from the study will be compared with other sources of Canadian mortality data (national surveillance data on apparent opioid-related deaths and Vital Statistics data) [1,25] to measure case ascertainment and premature mortality.

Where possible, analyses of the study dataset have been stratified by the manner of death and used the "sex- and gender-based analysis plus" approach to understand how intersectional identities, histories, and the distribution of resources contribute to acute toxicity deaths [45].

XSL•FO
RenderX

Results

Study Timeline

Work on the study began in the summer of 2018, and all activities are expected to be completed in 2025 (Textbox 2). The extended lag between the start of the study and the dissemination of findings is related to several factors. This ambitious project was the first research collaboration between the PHAC and all CCCME offices across the country. As these were new partnerships, time was invested in developing data-sharing agreements and seeking approvals from research ethics boards, privacy committees, and provincial and territorial officials. The nature of the study itself, a chart review, required significant personnel and time resources to review files, abstract, and prepare data, with each case file taking approximately 60 (range 30-100) minutes to review. Finally, study progress was also impacted by the COVID-19 pandemic, which diverted resources and significantly interrupted data collection, particularly where abstractors were working with physical files in office locations.

Textbox 2. Study timeline for our national chart review study of substance-related acute toxicity deaths.

- August 2018 to October 2019: Consultations, planning, data tool and documentation development and piloting, and research approvals (including departmental, scientific, security, research ethics board, and privacy reviews)
- February 2019 to March 2022: Development of data-sharing agreements with chief coroner and chief medical examiner offices
- March 2019 to October 2022: Abstractor training and data collection
- October 2019 to November 2022: Data preparation
- May 2022 to December 2024: Data analysis and knowledge translation
- June 2022 to September 2022: Release of preliminary results [46,47]
- December 2022: National summary report [48]
- January 2023 to December 2025: Peer-reviewed publications [49-52]
- May 2023 to March 2025: Evaluation

Available Files

A total of 9414 CME files met the case definition and were accessible for the study (Table 2). The records available from British Columbia included only deaths due to “street drugs” or diverted prescriptions that were accidental or had an undetermined manner of death. As such, data for people who experienced acute toxicity deaths by suicide or due solely to prescribed substances or alcohol were not available. On the basis of a report released by the British Columbia Coroners Service, there were approximately 199 deaths due to suicide by poisoning in 2016 and 2017 that may have met this study’s case definition [53].

XSL•FO
RenderX

Table 2. Study population size[a] and data sources by jurisdiction and year for our national chart review study of substance-related acute toxicity deaths.

Province or territory	2016 deaths (n=4164), n (%)	2017 deaths (n=5247), n (%)	Total deaths (N=9414), n (%)	Data sources
British Columbia[b]	993 (23.85)	1494 (28.47)	2487 (26.42)	Electronic database
Alberta	807 (19.38)	951 (18.12)	1758 (18.67)	Physical files
Saskatchewan	123 (2.95)	123 (2.34)	246 (2.61)	Physical and electronic files
Manitoba	180 (4.32)	198 (3.77)	378 (4.02)	Physical files (46 files were partially abstracted)
Ontario	1311 (31.48)	1710 (32.59)	3021 (32.09)	Physical files and electronic database (data for 681 cases from 2017 came from the electronic database only)
Quebec	534 (12.82)	537 (10.23)	1068 (11.34)	Physical files for most cases and electronic coroner reports for 426 cases from 2017
Newfoundland and Labrador	30 (0.72)	42 (0.8)	75 (0.79)	Physical files
New Brunswick	66 (1.58)	63 (1.2)	126 (1.34)	Physical files
Nova Scotia	84 (2.02)	99 (1.89)	183 (1.94)	Physical and electronic files
Prince Edward Island	15 (0.36)	15 (0.28)	30 (0.32)	Physical files
Yukon	Suppressed	Suppressed	21 (0.22)	Physical files
Northwest Territories	Suppressed	Suppressed	Suppressed	Physical files
Nunavut	Suppressed	Suppressed	Suppressed	Physical files

[a]To protect privacy, counts have been randomly rounded to base 3 and numbers <10 have been suppressed. As counts have been randomly rounded to base 3, they may not add up when summed.

[b]Data from British Columbia were available only for people who experienced accidental or undetermined acute toxicity deaths involving "street drugs" or pharmaceutical substances not prescribed to them. As such, data for people who experienced acute toxicity deaths by suicide or due solely to prescribed substances or alcohol were not available.

Due to COVID-19–related public health measures, the ability to abstract physical case files was limited, resulting in data being collected from electronic files or databases in 2 provinces. Altogether, 681 of the 1710 Ontario cases from 2017 (22.54% of all Ontario cases) were mapped from an electronic database, and 426 of the 537 Quebec cases from 2017 (39.89% of all Quebec cases) were abstracted from electronic coroner reports rather than physical files.

Data Quality Indicators

Consistency of Data Collection

Baseline measurements for the fictitious CME test file were within the established threshold for 96% (25/26) of all abstractors, and IaRR measurements taken thereafter over the course of abstraction remained within that threshold for 96% (25/26) of all abstractors. Abstractors who did not meet this threshold received remedial training and were reassessed. All obtained satisfactory κ values through this process, allowing them to start or resume abstraction.

Availability of Data on Key Concepts

Information collected from CME files was available when (1) it was systematically collected during death investigations, (2) it was collected during death investigations but not systematically, or (3) it was not specifically sought by the CME during death investigations but was serendipitously captured in the case file. As investigation protocols, forms, and records available to CMEs varied by jurisdiction, the availability of information also varied. Where the data source was an electronic file or database, less information was often available, except for electronic files from Saskatchewan, which included scanned copies of all paper documents in the case file.

Core study variables had good availability overall and across data sources (Table 3). Using a combination of postal codes and municipalities, the study was able to collect the census subdivision of residence for 95.44% (n=8985) of people who died during the study period. Full toxicology report results were available for 90.36% (n=8586) of people who died but were unavailable for 434 (4.61% of all people who died) Ontario 2017 cases that were mapped from an electronic database where only partial toxicology results were available. Toxicology information was also unavailable in instances where there was decomposition of the person before they were found (<1% of cases), a prolonged hospital stay before death (<1% of cases), or their hospital admission blood sample was not available (<1% of cases).

XSL•FO
RenderX

Table 3. Availability of information for core variables and key concepts among 9414 CME[a] case files for our national chart review study of substance-related acute toxicity deaths.

Concept	Cases with unknown or unavailable variables, n (%)	Data sources with missing information for >90% of records[b,c]
Core variables		
Age	0 (0)	None
Sex	0 (0)	None
Manner of death	0 (0)	None
Date of death	0 (0)	None
Residence census subdivision	429 (4.56)	None
Residence postal code	2127 (22.59)	Quebec electronic files
Residence location type	1011 (10.74)	None
Acute toxicity event census subdivision	1218 (12.94)	None
Acute toxicity event postal code	1797 (19.09)	Quebec electronic files
Acute toxicity event location type	678 (7.2)	None
Death census subdivision	129 (1.37)	None
Death postal code	1581 (16.79)	Quebec electronic files
Death location type	69 (0.73)	None
Toxicology report available	828 (8.8)	None
Specific substances contributing to death	852 (9.05)	None
Classes of substances contributing to death	792 (8.41)	None
Key concepts		
Race	5304 (56.34)	Partially abstracted Manitoba cases, Quebec electronic files, and Nova Scotia
Ethnicity	8520 (90.5)	Alberta, Manitoba, Ontario electronic database, Quebec, New Brunswick, Nova Scotia, Prince Edward Island, Newfoundland and Labrador, and Yukon
Who the person who died lived with	4722 (50.16)	British Columbia, Ontario electronic database, and Prince Edward Island
Education	9012 (95.73)	British Columbia, Alberta, Saskatchewan, Manitoba, Ontario, Quebec electronic files, New Brunswick, Prince Edward Island, and Yukon
Income source	4572 (48.57)	British Columbia and Prince Edward Island
Occupation classification	7050 (74.89)	Saskatchewan, Quebec electronic files, and Prince Edward Island
Industry classification	7419 (78.81)	Saskatchewan, Quebec electronic files, and Prince Edward Island
Potentially traumatic events	5424 (57.62)	Prince Edward Island
Had an accessible family physician	6510 (69.15)	British Columbia and Prince Edward Island
Contact with the health system in the year before death	2124 (22.56)	None
Medical history	2775 (29.48)	None
Mental health history	2322 (24.67)	None
Prescription history	4536 (48.18)	British Columbia
History of substance use	1443 (15.33)	None
Known substances used	3174 (39.15)	British Columbia
Frequency of substance use (excluding alcohol) in the year before death	3114 (44.22)	Partially abstracted Manitoba cases, Ontario electronic database, Prince Edward Island, and Yukon

XSL•FO
RenderX

Concept	Cases with unknown or unavailable variables, n (%)	Data sources with missing information for >90% of records[b,c]
Substance use in the presence of others before the acute toxicity event	4011 (42.61)	Ontario electronic files
Witnesses to the acute toxicity event	3087 (32.79)	British Columbia
Apparent mode of substance use	5100 (54.17)	Partially abstracted Manitoba cases and Ontario electronic database
Presence of substances at the scene	4377 (46.49)	British Columbia
Tools of substance use at the scene	4275 (45.41)	British Columbia
Actions taken by witnesses	2544 (40.95)	None
First responder actions	888 (23.44)	British Columbia
Whether naloxone was administered	1713 (45.22)	Prince Edward Island

[a]CME: coroner and medical examiner.

[b]This table excludes data from Nunavut and the Northwest Territories due to small numbers.

[c]Denominators for specific variables may not be the sum of all cases due to skip patterns in the data collection tool.

In general, information on socioeconomic variables (including education, occupation, and industry), potentially traumatic life events, and socially constructed identities (such as gender [not shown], ethnicity, and race) were often unavailable. Conversely, variables related to the person's health, history of substance use, and events surrounding the acute toxicity event were available for most records. Some exceptions include whether the person had an accessible family physician (missing: n=6510, 69.15%), the apparent mode of substance use at the time of the acute toxicity event leading to death (missing: n=5100, 54.17%), and whether naloxone was administered (missing: n=1713, 45.22%).

Discussion

Principal Findings

As Canada continues to grapple with an ongoing overdose crisis, there is an increased need for high-quality data that can support clinicians, public health experts, and advocates in their efforts to refine interventions. This can be supported by relevant, comparable, and timely CME data.

Our study confirms that CME files are a rich source of information that is highly relevant to public health practitioners. Work has begun on knowledge translation products [46-52], including a summary report of findings from this study that provides insight into national trends over time and by geography, sociodemographic factors, health and substance use history, the circumstances of the acute toxicity event, and substances involved [48]. Improving the usefulness of CME files as sources for acute toxicity mortality data will depend upon how the Canadian CCCME offices meet 3 broad challenges. First, CMEs do not primarily collect data for the purpose of public health intervention but rather to meet their investigation mandate needs. As illustrated in Table 3, the availability of variables of public health interest varies across provinces and territories. Second, CME offices have different approaches to the collection, storage, and transfer of data. With respect to data infrastructure, some offices have electronic systems that can be readily searched, and some keep mostly paper-based files. Most of the electronic systems that do exist are not compatible with each other. Third, as CMEs use classification systems and terminology that suit the needs of local internal and external stakeholders, systems and terms may not be comparable across provinces and territories, necessitating resource-intensive studies such as the one described here to allow for comparable data.

Recognizing some of these challenges and opportunities, the PHAC, all 13 Canadian CCCME offices, and Statistics Canada established the CCCME and Public Health Collaborative in 2021, with a mandate to support the development of common approaches to death investigations and data infrastructure requirements. The collaborative has identified substance-related acute toxicity as one of its prioritized causes of death and will consider the knowledge gaps and lessons learned from this study in its development of core and minimum data elements to improve the comparability, usefulness, and accessibility of national mortality data. Increased knowledge transfer and collaboration across public health and CCCME jurisdictions in the development and implementation of common practices in death investigations will enhance the availability of timely and comparable data for public health use.

The authors are aware of how critical timeliness is to public health practitioners and community members who work on interventions. This study took longer than anticipated, partly due to challenges encountered during the global COVID-19 pandemic as well as navigating legislative and administrative requirements aimed at ensuring good data governance across multiple departments within 14 federal, provincial, and territorial governments. We anticipate that future similar national studies and surveillance activities can now occur more efficiently using the pathways created for this study. While data collected by this study may no longer reflect current trends in acute toxicity deaths, they provide an important baseline early in Canada's overdose crisis that can be used to measure future progress.

Limitations

Although efforts were made to identify all CME files that met the study case definition, it is possible that some people who died of acute toxicity were not included in this study if (1) their

XSL•FO
RenderX

death was not reported to the CME office or (2) their death was not identified as a potential case. Systematic differences in study data sources (Table 2); how data were abstracted; and data availability associated with differences in death investigation processes, death classification methods, and toxicology testing methods (Table 3) may result in an underestimation of burden and could bias findings. Extensive standardized training and documentation were provided to abstractors, all of whom underwent IaRR assessments throughout data collection. Nonetheless, as data from each case file was collected from a single abstractor, error and bias may have resulted from differences in how abstractors assessed whether cases met eligibility requirements and abstracted data.

As most variables collected by our study had a nontrivial amount of missing data, study findings will be reported as minimum counts, proportions, and mortality rate estimates that likely underestimate the true population prevalence of reported characteristics. Results from variables with a high percentage of missing data should be interpreted with caution, particularly because the distribution of values where the information was known may not be the same as the distribution of values where information was unknown. The lack of information on demographic and socioeconomic variables and socially constructed identities (eg, gender) will hinder our ability to apply a sex- and gender-based analysis plus approach to our analyses, although efforts will be made to describe the distribution of people who died by sex as well as area-level characteristics where possible.

Conclusions

This study provides the most detailed national information on substance-related acute toxicity deaths in Canada to date. It is anticipated that results will allow for comparisons across provinces and territories, provide evidence to inform programs and policies at all levels, inform the expansion of national surveillance activities on substance-related harms, supply evidence to inform the development of common approaches to medicolegal death investigations and routine postmortem testing, and serve as a platform for future research activities. It also provides a baseline of acute toxicity mortality before the COVID-19 pandemic for assessing the evolution of the overdose crisis. Finally, the partnerships and processes initiated by this study have contributed to fruitful new collaborations between Canada's public health and CME communities.

Acknowledgments

This study was led by the Public Health Agency of Canada (PHAC) and the Federal, Provincial, and Territorial Special Advisory Committee on the Epidemic of Opioid Overdoses. The authors would like to acknowledge their collaborators within the chief coroner and chief medical examiner offices across Canada for providing access to their death investigation files and identifying potential cases. The authors would also like to acknowledge the contributions of their dedicated data abstractors and the many people who supported the rollout of the study.

This study was funded by the PHAC. Opinions and conclusions in this report are the authors' own and not endorsed or approved by data providers or funders. The authors would also like to acknowledge Raahyma Ahmad, Noushon Farmanara, Diana George, Aganeta Enns, Jennifer Pennock, and Jeyasakthi Venugopal for their feedback on the earlier versions of this report.

Data Availability

The study dataset is not publicly available due to provisions in the data-sharing agreements with provincial and territorial data providers. The corresponding author can assist in directing inquiries about data access to the original data providers.

Authors' Contributions

JR, BA, MJB, SB-E, JH, DH, BJ, GJ, FK, JL, RM, EER, and ES conceptualized and designed the study protocol and data collection tool. JR and AV were responsible for project administration and supervision. JR, TK, RM, ES, and AV participated in the acquisition of data. JR, TK, and AV curated and analyzed the data. JR, MJB, TK, FK, DS, and AV drafted the manuscript, and all authors revised the paper. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Variable list.
[DOCX File , 123 KB - publichealth_v11i1e49981_app1.docx]

Multimedia Appendix 2
Database screenshots.
[DOCX File , 6131 KB - publichealth_v11i1e49981_app2.docx]

XSL•FO
RenderX

References

1. Federal, Provincial, and Territorial Special Advisory Committee on the Epidemic of Opioid Overdoses. Opioid- and stimulant-related harms in Canada. Public Health Agency of Canada. URL: https://health-infobase.canada.ca/substance-related-harms/opioids-stimulants/ [accessed 2024-03-19]
2. Snowdon J. Drug overdose death rates in different countries: who should be alarmed? Australas Psychiatry 2022 Feb 02;30(1):26-30. [doi: 10.1177/10398562221075192] [Medline: 35236130]
3. Federal approach on Canada's opioid crisis. Health Canada. 2018 Feb. URL: https://www.canada.ca/en/health-canada/services/publications/healthy-living/federal-approach-canada-opioids-crisis.html [accessed 2022-12-01]
4. CCENDU bulletin: deaths involving fentanyl in Canada, 2009-2014. Canadian Centre on Substance Abuse, Canadian Community Epidemiology Network on Drug Use. 2015 Aug. URL: https://www.ccsa.ca/sites/default/files/2019-05/CCSA-CCENDU-Fentanyl-Deaths-Canada-Bulletin-2015-en.pdf [accessed 2019-02-01]
5. Fischer B, Goldman B, Rehm J, Popova S. Non-medical use of prescription opioids and public health in Canada. Can J Public Health 2008 May 1;99(3):182-184. [doi: 10.1007/bf03405469]
6. Gladstone E, Smolina K, Morgan SG, Fernandes KA, Martins D, Gomes T. Sensitivity and specificity of administrative mortality data for identifying prescription opioid-related deaths. CMAJ 2016 Mar 01;188(4):E67-E72 [FREE Full text] [doi: 10.1503/cmaj.150349] [Medline: 26622006]
7. Substance use surveillance data. Alberta Health. 2024. URL: https://www.alberta.ca/substance-use-surveillance-data [accessed 2024-03-13]
8. Statistical reports on deaths in British Columbia. British Columbia Coroners Service. 2024 Feb 29. URL: https://www2.gov.bc.ca/gov/content/life-events/death/coroners-service/statistical-reports [accessed 2024-03-13]
9. Unregulated drug poisoning emergency dashboard. British Columbia Centre for Disease Control. 2024. URL: http://www.bccdc.ca/health-professionals/data-reports/substance-use-harm-reduction-dashboard [accessed 2024-03-13]
10. Palis H, Bélair MA, Hu K, Tu A, Buxton J, Slaunwhite A. Overdose deaths and the COVID-19 pandemic in British Columbia, Canada. Drug Alcohol Rev 2022 May 15;41(4):912-917. [doi: 10.1111/dar.13424] [Medline: 34908203]
11. Toxic drug crisis data. First Nations Health Authority. 2024. URL: https://www.fnha.ca/what-we-do/mental-wellness-and-substance-use/harm-reduction-and-the-toxic-drug-crisis/toxic-drug-crisis-data [accessed 2024-03-13]
12. Opioid-related morbidity and mortality in Ontario: interactive opioid tool. Public Health Ontario. 2024. URL: https://www.publichealthontario.ca/en/data-and-analysis/substance-use/interactive-opioid-tool [accessed 2024-03-13]
13. Campbell TJ, Kitchen SA, Tadrous M, Damba C, Johnson CH, Smoke A, et al. Varying circumstances surrounding opioid toxicity deaths across ethno-racial groups in Ontario, Canada: a population-based descriptive cross-sectional study. BMJ Public Health 2024 Mar 13;2(1):e000480. [doi: 10.1136/bmjph-2023-000480]
14. Holton A, Gomes T, Leece P, Yang J, Hamzat B, Ledlie S, et al. Prescribing patterns, substance use disorder diagnoses, and access to treatment prior to substance-related toxicity deaths in Ontario. The Ontario Drug Policy Research Network and Public Health Ontario. URL: https://www.publichealthontario.ca/-/media/Documents/O/24/odprn-prescribing-substance-treatment-substance-ontario.pdf?rev=9f316c59e2314cb092f2b06c202f1e5b&sc_lang=en [accessed 2024-03-13]
15. MacDonald M, Cheng C, Wang T, McCormack D, Kolla G, Cahill TM, Members of the Opioid Drug Observatory Lived Experience Group. Trends in varying modes of drug use in opioid toxicity deaths in Ontario from 2017 to 2021. Int J Drug Policy 2023 Sep 22:104197. [doi: 10.1016/j.drugpo.2023.104197] [Medline: 37741701]
16. Gomes T, Leece P, Iacono A, Yang J, Kolla G, Cheng C, et al. Characteristics of substance-related toxicity deaths in Ontario: stimulant, opioid, benzodiazepine, and alcohol-related deaths. The Ontario Drug Policy Research Network and Public Health Ontario. 2023. URL: https://odprn.ca/wp-content/uploads/2023/09/Substance-Toxicity-Report-Final.pdf [accessed 2024-03-13]
17. Parent S, Buttemer S, Philpott J, Moore K. Opioid-related deaths in Kingston, Frontenac, Lennox and Addington in Ontario, Canada: the shadow epidemic. Health Promot Chronic Dis Prev Can 2023 Feb;43(2):62-72 [FREE Full text] [doi: 10.24095/hpcdp.43.2.02] [Medline: 36794823]
18. Substance related harms surveillance report. Government of Manitoba, Epidemiology and Surveillance Unit. 2024 Jan 22. URL: https://www.gov.mb.ca/mh/srh-public-report.html#Deaths [accessed 2024-03-13]
19. Confirmed and suspected drug toxicity deaths (2016 – most recent). Saskatchewan Coroners Service. 2023. URL: https://publications.saskatchewan.ca/#/products/90505 [accessed 2024-03-13]
20. Surveillance of apparent opioid overdoses, 2023 Q2. Public Health New Brunswick. 2023 Dec. URL: https://www2.gnb.ca/content/gnb/en/corporate/promo/opioids/public_health_surveillance.html [accessed 2024-03-13]
21. Opioid use and overdose strategy. Province of Nova Scotia. 2024 Mar 2. URL: https://novascotia.ca/opioid/ [accessed 2024-03-13]
22. Preventing opioid-related overdoses. Government of Prince Edward Island. 2023 Jun 27. URL: https://www.princeedwardisland.ca/en/information/health-and-wellness/preventing-opioid-related-overdoses [accessed 2024-03-13]
23. Bertrand-Deschênes A, Parent AA, Bergeron-Longpré M. Surdoses au Québec et inégalités sociales de santé. Drogues, Santé et Société 2023 Jan 06;20(2):48-71. [doi: 10.7202/1095346ar]

XSL•FO
RenderX

24. Schleihauf E, Bowes MJ. Suicide and drug toxicity mortality in the first year of the COVID-19 pandemic: use of medical examiner data for public health in Nova Scotia. Health Promot Chronic Dis Prev Can 2022 Feb 16;42(2):60-67 [FREE Full text] [doi: 10.24095/hpcdp.42.2.02] [Medline: 34757897]
25. Canadian vital statistics - death database (CVSD). Statistics Canada. 2023 Mar 9. URL: http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3233 [accessed 2023-03-30]
26. Australia's annual overdose report. Penington Institute. 2023. URL: https://www.penington.org.au/wp-content/uploads/2023/08/PEN_Annual-Overdose-Report-2023_FINAL.pdf [accessed 2024-04-29]
27. Spencer MR, Garnett MF, Miniño AM. Drug overdose deaths in the United States, 2002–2022: NCHS data brief No. 491. NCHS Data Brief. 2024. URL: https://www.cdc.gov/nchs/data/databriefs/db491.pdf [accessed 2024-04-29]
28. Konefal S, Sherk A, Maloney-Hall B, Young M, Kent P, Biggar E. Polysubstance use poisoning deaths in Canada: an analysis of trends from 2014 to 2017 using mortality data. BMC Public Health 2022 Feb 10;22(1):269 [FREE Full text] [doi: 10.1186/s12889-022-12678-z] [Medline: 35144586]
29. Snowdon J, Choi N. Unanticipated changes in drug overdose death rates in Canada during the opioid crisis. Int J Ment Health Addict 2022 Oct 10:1-19 [FREE Full text] [doi: 10.1007/s11469-022-00932-9] [Medline: 36248265]
30. Roxburgh A, Pilgrim JL, Hall WD, Burns L, Degenhardt L. Accurate identification of opioid overdose deaths using coronial data. Forensic Sci Int 2018 Jun;287:40-46. [doi: 10.1016/j.forsciint.2018.03.032] [Medline: 29627712]
31. State unintentional drug overdose reporting system (SUDORS). Centers for Disease Control and Prevention. 2024 Feb 26. URL: https://www.cdc.gov/drugoverdose/fatal/dashboard [accessed 2024-03-22]
32. Lewer D, Brothers TD, Harris M, Rock KL, Copeland CS. Opioid-related deaths during hospital admissions or shortly after discharge in the United Kingdom: a thematic framework analysis of coroner reports. PLoS One 2023 Apr 19;18(4):e0283549 [FREE Full text] [doi: 10.1371/journal.pone.0283549] [Medline: 37075021]
33. 'acute toxicity': compendium of chemical terminology. International Union of Pure and Applied Chemistry. 2019. URL: https://goldbook.iupac.org/ [accessed 2024-04-29]
34. Global SMART Programme. The challenge of new psychoactive substances. United Nations Office on Drugs and Crime. 2013 Mar. URL: https://www.unodc.org/documents/scientific/NPS_Report.pdf [accessed 2024-03-13]
35. Kelsall D, Bowes MJ. No standards: medicolegal investigation of deaths. CMAJ 2016 Feb 16;188(3):169 [FREE Full text] [doi: 10.1503/cmaj.160041] [Medline: 26833736]
36. Dictionary, census of population, 2016, census subdivision (CSD). Statistics Canada. 2016 Nov 16. URL: https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/geo012-eng.cfm [accessed 2019-02-01]
37. Postal code conversion file plus (PCCF+). Statistics Canada. 2017 Mar 6. URL: https://www150.statcan.gc.ca/n1/en/catalogue/82F0086X [accessed 2022-07-18]
38. Subedi R, Roshanafshar S, Greenberg TL. Developing meaningful categories for distinguishing levels of remoteness in Canada (catalogue no. 11-633-X-026). Statistics Canada. 2020 Aug 11. URL: https://www150.statcan.gc.ca/n1/pub/11-633-x/11-633-x2020002-eng.htm [accessed 2022-07-18]
39. The Canadian index of multiple deprivation. Statistics Canada. 2019 Jun 12. URL: https://www150.statcan.gc.ca/n1/en/catalogue/45200001 [accessed 2022-07-18]
40. Historical data on climate. Environment and Climate Change Canada. 2021 Jan 31. URL: https://climate.weather.gc.ca/historical_data/search_historic_data_e.html [accessed 2023-03-05]
41. International statistical classification of diseases and related health problems (10th revision). World Health Organization. 2019. URL: https://icd.who.int/browse10/2019/en [accessed 2019-12-01]
42. How we keep data private. New Zealand Government Ministry of Social Development. 2022. URL: https://www.msd.govt.nz/about-msd-and-our-work/tools/how-we-keep-data-private.html[accessed [accessed 2022-11-18]
43. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012 Feb 01;8(1):23-34 [FREE Full text] [doi: 10.20982/tqmp.08.1.p023] [Medline: 22833776]
44. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159. [doi: 10.2307/2529310]
45. Health portfolio sex- and gender-based analysis plus policy: advancing equity, diversity and inclusion. Health Canada. 2023. URL: https://www.canada.ca/en/health-canada/corporate/transparency/corporate-management-reporting/heath-portfolio-sex-gender-based-analysis-policy.html [accessed 2023-03-05]
46. Special Advisory Committee on the Epidemic of Opioid Overdoses. Homelessness and substance-related acute toxicity deaths: a descriptive analysis of a national chart review study of coroner and medical examiner data. Health Canada. 2022 Jun. URL: https://www.canada.ca/en/health-canada/services/opioids/data-surveillance-research/homelessness-substance-related-acute-toxicity-deaths.html [accessed 2022-11-18]
47. Special Advisory Committee on the Epidemic of Opioid Overdoses. Area-level characteristics of substance-related acute toxicity deaths: a descriptive analysis of a national chart review study of coroner and medical examiner data. Health Canada. 2022 Sep. URL: https://www.canada.ca/en/health-canada/services/opioids/data-surveillance-research/area-level-characteristics-substance-related-acute-toxicity-deaths-analysis-coroner-medical-examiner-data.html [accessed 2022-11-18]

XSL•FO
RenderX

48. Substance-related acute toxicity deaths in Canada from 2016 to 2017: a review of coroner and medical examiner files. Public Health Agency of Canada. 2017. URL: https://www.canada.ca/content/dam/hc-sc/documents/services/opioids/data-surveillance-research/substance-related-acute-toxicity-deaths-canada-2016-2017-review-coroner-medical-examiner-files/substance-related-acute-toxicity-deaths-canada-2016-2017-review-coroner-medical-examiner-files.pdf [accessed 2024-04-29]
49. Baddeliyanage R, Enns A, Vansteelandt A, Abele B, Kouyoumdjian F, Schleihauf E, et al. Substance-related acute toxicity deaths by area-based characteristics: a descriptive analysis of a national chart review study of coroner and medical examiner data. Int J Ment Health Addiction 2024 Feb 23:1-24. [doi: 10.1007/s11469-024-01259-3]
50. VanSteelandt A, Rotondo J. Using death investigation data to better understand the overdose crisis. Health Promot Chronic Dis Prev Can 2024 Mar;44(3):75-76 [FREE Full text] [doi: 10.24095/hpcdp.44.3.01] [Medline: 38501678]
51. Chang GY, VanSteelandt A, McKenzie K, Kouyoumdjian F. Accidental substance-related acute toxicity deaths among youth in Canada: a descriptive analysis of a national chart review study of coroner and medical examiner data. Health Promot Chronic Dis Prev Can 2024 Mar;44(3):77-88 [FREE Full text] [doi: 10.24095/hpcdp.44.3.02] [Medline: 38501679]
52. Ha JH, Burt J, Randell S, VanSteelandt A. Accidental substance-related acute toxicity deaths in older adults in 2016 and 2017: a national chart review study. Health Promot Chronic Dis Prev Can 2024 Mar;44(3):89-100 [FREE Full text] [doi: 10.24095/hpcdp.44.3.03] [Medline: 38501680]
53. Suicide deaths in BC: 2008-2018. British Columbia Coroners Service. 2020. URL: https://www2.gov.bc.ca/assets/gov/birth-adoption-death-marriage-and-divorce/deaths/coroners-service/statistical/suicide.pdf [accessed 2020-06-18]

Abbreviations

CCCME: chief coroner and chief medical examiner
CME: coroner and medical examiner
IaRR: intrarater reliability
PHAC: Public Health Agency of Canada

Edited by A Mavragani; submitted 15.06.23; peer-reviewed by R Fujiwara, N Pollock; comments to author 27.02.24; revised version received 23.03.24; accepted 07.05.24; published 10.04.25.

Please cite as:
Rotondo J, VanSteelandt A, Kouyoumdjian F, Bowes MJ, Kakkar T, Jones G, Abele B, Murray R, Schleihauf E, Halverson J, Leason J, Huyer D, Jackson B, Bozat-Emre S, Shah D, Rees EE
Substance-Related Acute Toxicity Deaths in Canada From 2016 to 2017: Protocol for a Retrospective Chart Review Study of Coroner and Medical Examiner Files
JMIR Public Health Surveill 2025;11:e49981
URL: https://publichealth.jmir.org/2025/1/e49981
doi:10.2196/49981
PMID:40208661

©Jenny Rotondo, Amanda VanSteelandt, Fiona Kouyoumdjian, Matthew J Bowes, Tanya Kakkar, Graham Jones, Brandi Abele, Regan Murray, Emily Schleihauf, Jessica Halverson, Jennifer Leason, Dirk Huyer, Beth Jackson, Songul Bozat-Emre, Devanshi Shah, Erin E Rees. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.04.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Trends in Prescription of Stimulants and Narcoleptic Drugs in Switzerland: Longitudinal Health Insurance Claims Analysis for the Years 2014-2021

Tamara Scharf[1,2,3*], MD; Carola A Huber[4,5*], MPH, PhD; Markus Näpflin[4*]; Zhongxing Zhang[3,6*], PhD; Ramin Khatami[3,6,7*], Prof Dr Med

[1]Graduate School of Health Sciences, University of Bern, Bern, Switzerland
[2]Institute of Primary Health Care (BIHAM), University of Bern, Bern, Switzerland
[3]Centre of Sleep Medicine and Epileptology Barmelweid, Klinik Barmelweid AG, Aargau, Switzerland
[4]Department of Health Sciences, Helsana Group, Zürich, Switzerland
[5]Institute of Primary Care, University of Zürich, University Hospital Zürich, Zürich, Switzerland
[6]Barmelweid Academy, Klinik Barmelweid AG, Barmelweid, Switzerland
[7]Department of Neurology, Inselspital Bern University Hospital, University of Bern, Bern, Switzerland
[*]all authors contributed equally

Corresponding Author:
Ramin Khatami, Prof Dr Med
Centre of Sleep Medicine and Epileptology Barmelweid, Klinik Barmelweid AG, Aargau, Switzerland

Abstract

Background: Stimulants are potent treatments for central hypersomnolence disorders or attention-deficit/hyperactivity disorders/attention deficit disorders but concerns have been raised about their potential negative consequences and their increasing prescription rates.

Objective: We aimed to describe stimulant prescription trends in Switzerland from 2014 to 2021. Second, we aimed to analyze the characteristics of individuals who received stimulant prescriptions in 2021 and investigate the link between stimulant prescriptions and hospitalization rates in 2021, using hospitalization as a potential indicator of adverse health outcomes.

Methods: Longitudinal and cross-sectional data from a large Swiss health care insurance were analyzed from all insureds older than 6 years. The results were extrapolated to the Swiss general population. We identified prescriptions for methylphenidate, lisdexamfetamine, modafinil, and sodium oxybate and calculated prevalences of each drug prescription over the period from 2014 to 2021. For 2021 we provide detailed information on the prescribers and evaluate the association of stimulant prescription and the number and duration of hospitalization using logistic regression models.

Results: We observed increasing prescription rates of all stimulants in all age groups from 2014 to 2021 (0.55% to 0.81%, 43,848 to 66,113 insureds with a prescription). In 2021, 37.1% (28,057 prescriptions) of the medications were prescribed by psychiatrists, followed by 36.1% (n=27,323) prescribed by general practitioners and 1% (n=748) by neurologists. Only sodium oxybate, which is highly specific for narcolepsy treatment, was most frequently prescribed by neurologists (27.8%, 37 prescriptions). Comorbid psychiatric disorders were common in patients receiving stimulants. Patients hospitalized in a psychiatric institution were 5.3 times (odds ratio 5.3, 95% CI 4.63 - 6.08, P<.001) more likely to have a stimulant prescription than those without hospitalization. There were no significant associations between stimulant prescription and the total length of inpatient stay (odds ratio 1, 95% CI 1 - 1, P=.13).

Conclusions: The prescription of stimulant medication in Switzerland increased slightly but continuously over years, but at lower rates compared to the estimated prevalence of central hypersomnolence disorders and attention-deficit/hyperactivity disorders/attention deficit disorders. Most stimulants are prescribed by psychiatrists, closely followed by general practitioners. The increased odds for hospitalization to psychiatric institutions for stimulant receivers reflects the severity of disease and the higher psychiatric comorbidities in these patients.

(JMIR Public Health Surveill 2025;11:e53957) doi:10.2196/53957

KEYWORDS

prescription trends; claims data; cross-sectional data; narcolepsy; prescribers; prescribing practices; medical care; stimulants; stimulant medication

XSL•FO
RenderX

Introduction

Stimulants such as methylphenidate, lisdexamfetamine, and modafinil are highly potent pharmacologic treatment options for hypersomnolence disorders, including narcolepsy or attention-deficit/hyperactivity disorders/attention deficit disorders (ADHDs/ADDs).

ADHDs/ADDs are some of the most common diagnosed psychiatric disorders worldwide with a prevalence of 5.3% worldwide in the years of 1978 to 2010 among people aged 18 years or younger [1]. Prevalence can be different due to varying diagnostic methods per country. In Switzerland, a prevalence of 5.2% was found in children aged 7 to 17 years and in adult men a prevalence of 4% [2,3].

Prescriptions of stimulant drugs have been increasing at various rates over the past decades, with methylphenidate showing an 8.2 fold increase from 1996 to 2013 [4,5]. In the United States, amphetamine and methylphenidate increased from 5.6% to 6.1% in adults aged 20 years or older between 2014 and 2019 [4,5]. In New Zealand, 1.06% of adolescents received stimulants in 2016, an increase of 41.3% from 0.75% in 2011 [6]. Denmark observed a trend in stimulant prescriptions rising from 0.31 per 1000 person-years in 1996 to 7.29 per 1000 person-years in 2010 [7]. In Switzerland, data on stimulant prescription rates is scarce. A Swiss study from 2015 found a lifetime prevalence of stimulants and other substances enhancing cognitive abilities of about 1.4% in employees and students [8]. No sufficient data have, however, been collected in a nationwide study.

Stimulant medication is used for diseases often manifesting during childhood or adolescence and in many cases long-term pharmacological treatment throughout adulthood is needed. Prescription of stimulant agents in this age group is therefore of special interest to balance the need of medication and potential risk by over prescription or under prescription. Prescription of stimulants in young age groups increased from 0.02% to 0.26% over time in Asia, Australia, Europe, and North America in children and 0.003% to 1.48% in adults [9]. Among Swiss school children between 2002 and 2005, methylphenidate prescription increased from 0.74% to 1.02% for children aged 5 to 14 years [10].

This rapid prescription increase may indicate over prescription or even misuse. Misuse of stimulants is common, with up to 17% prevalence in US college students according to a meta-analysis. Misuse can lead to a range of negative consequences such as decreased appetite, insomnia, and increases in heart rate and blood pressure with increased long-term cardiovascular risk and possibly lead to increased hospitalization rates [11-13].

Properly identifying current stimulant prescription rates and discovering the prescription patterns or circumstances of their prescription (prescriber, package size, and comorbidities) may help to further identify alarming prescription increases and potential misuse in Switzerland, and the possible causes.

We therefore had 2 objectives: first, we aim to describe the rate of stimulant prescriptions in Switzerland from 2014 to 2021, focusing on both minors and adults. We hypothesize that the rate of stimulant prescriptions in Switzerland, similar to international trends, has increased over the last decade, with the highest prescription rates being recorded in 2021.

Second, we aimed at analyzing the characteristics of individuals who received stimulant prescriptions in 2021. This analysis will include factors such as comorbidities, age, and prescription details such as package size and health care providers most frequently issuing these prescriptions. Our hypotheses are that recipients likely mirror the characteristics of ADHDs/ADDs and central hypersomnolence disorders, most are minors with few comorbidities, and the likelihood of receiving a prescription decreases with age. Specialists, particularly psychiatrists and neurologists, are expected to be the primary prescribers.

Lastly, this study aims to investigate the link between stimulant prescriptions and hospitalization rates in 2021, using hospitalization as a potential indicator of adverse health outcomes. Given the expected increase in stimulant prescriptions until 2021, the hypothesis is that individuals prescribed stimulants may have a higher risk of hospitalization, particularly if prescription rates are on the rise.

Methods

Study Design

This study is a longitudinal and cross-sectional analysis of the Helsana health care insurance data of around 1.5 million people in Switzerland insured over the period of 2014 to 2021. Helsana belongs to a group of the biggest insurance companies in Switzerland and insures 14% of the Swiss population, with insureds in 26/26 cantons of Switzerland. Data describes general information on the insured persons and all their invoices for health services directed to the insurance. These invoices are representative of all health care costs of the insureds, except for the costs that were not sent as invoices to the insurance and paid by the insureds themselves (ie, over the counter drug costs and dental costs). We decided to provide more detailed descriptive statistics for data in 2021, since it was the most recent data available to us and presumably with the highest rate of stimulant prescription.

Identification of Drugs

We identified the drug invoices through their Anatomical Therapeutic Chemical code, which classifies chemical substances based on their therapeutical properties. The identification was performed for the following drugs: methylphenidate (N06BA04), lisdexamfetamine (N06BA12), modafinil (N06BA07), and sodium oxybate (N07XX04) which is a specific medication for narcolepsy treatment and used to estimate the treatment prevalence of narcolepsy patients in the dataset. Pitolisant (N07XX11) was identified too as a stimulant with specific use for narcolepsy but was excluded from further analysis, as it was only authorized for use in 2020 but had neglectable low prescription rates. We also identified the Swissmedic code of the medications, which is specific not only for the chemical substance but also the producer of the medication and package size. These drugs are only accessible through prescription by a medical professional and reimbursed by the insurance company. Overlapping prescriptions of the 4

XSL•FO
RenderX

drugs was defined as 1 prescription of one of the 4 drugs invoiced with at least one of the other drugs once or multiple times during the year of 2021.

Variables

The dataset consisted of all insureds aged ≥6 years with information on their age, sex, region of language, and region of residence. We categorized 5 age groups (6 - 17 y, 18 - 35 y, 36 - 65 y, 66 - 75 y, and 76+ y). We divided the insureds` residential regions into "rural," "intermediate," and "urban" subgroups according to the Swiss federal office for statistics.

The insureds had various health care plans including standard care and managed care models (eg family physician model). These health insurance plans were identified and categorized into standard care and managed care (ie, the combination of telemedical care and general practitioner [GP] care).

Chronic health condition status was identified by substance prescriptions related to chronic diseases. This was carried out according to approaches developed in previous research on the dataset [14]. We classified 22 different chronic conditions and categorized them into psychiatric, cardiologic, rheumatologic, respiratory comorbidities, or all other.

All invoices for hospitalization (ie, hospitals of all sizes providing acute care and psychiatric clinics) and the length of stay were included in our analysis.

Based on the medical prescriber who issued the invoices, several prescriber categories were defined: GP, psychiatrist, neurologist, other specialists (combining all other prescribers, such as nonspecific group practices, cardiologists, pulmonologists, rheumatologists, etc). Only health care personnel in Switzerland are allowed to prescribe medication. We further grouped them into "only prescriber" of the medication when a describer prescribed all medication exclusively for single individuals, ">50% prescriber" meaning more than 50% of the prescriptions for single individuals were invoiced by the prescriber, and "rest" with all other prescription proportions.

Within the prescriber categories (inpatient psychiatry, inpatient acute, and rehabilitation, nursing) we differentiated between inpatient (during a hospital stay) and outpatient invoices.

Statistical Analysis

For our first objective we provide descriptive statistics for prescription trends among different age groups across the years 2014 and 2021 by identifying individuals in the dataset with at least one prescription of the predefined drugs. To obtain representative data for Swiss population we extrapolated these data by current residency numbers and populations statistics of the Federal Statistical Office.

For our second objective we restricted the descriptive statistics to the year 2021 and provide detailed information on age, sex, region of residence, health insurance status, comorbidities, prescribers, and package size. Logistic regression models were performed to evaluate the association of at least one stimulant prescription versus no prescription (ie, the dependent variable is prescription yes or no), number of hospitalization, and length of stay, adjusted for age, sex, region of residence, health insurance status, and number of chronic diseases for the year 2021.

All data management, graphic generation, and analysis was performed with the statistics program R (version 4.2.1; R Foundation).

Data Availability

The authors were permitted access to the data by collaboration with the insurance companies research team. The datasets generated and analyzed during this study are available from the corresponding author on reasonable request. AI was not used in any way in data generation, analysis, and presentation of results.

Ethical Considerations

According to ethical and legal regulations in Switzerland no ethical approval or patient consent was needed for this study, as all data complied with privacy regulations and personal data protection, data was anonymized when presented to the research team. The Swiss Human Research Act (REQ-2017 - 00280) did not apply to this project. The exploratory statistical analyses of the feasibility test complied with the Swiss Federal Law on data protection. All data were anonymized and deidentified prior to the performed analysis to protect the privacy of patients, physicians, and hospitals. According to the national ethical and legal regulation, an ethical approval was not needed because the data were pre-existing and deidentified. Since data was anonymized, no consent of patients was required.

Results

Trend of Stimulant Prescription Per Year

As baseline we refer to the prescription period in 2014. Between baseline and 2021, on average 14% of Swiss people were at any time insured with the Helsana Group.

In the year 2014, 0.55% (42,848 insureds) of insured people of any age received at least 1 stimulant agent (or stimulant prescription). This number increased up to 0.8% (66,113 insureds) in the year of 2021. The largest growing percentage of stimulus prescriptions were in the youngest age group with an increase of prescription of 0.6 percentage points (17,972, 1.8% to 24,982, 2.4%) between 2014 and 2021 (Table 1).

Between 2014 and 2021, 32,2418 packages of methylphenidate were the most prescribed stimulant, followed by 46,074 packages of lisdexamfetamine, 8797 packages of modafinil, and 3115 packages of sodium oxybate. We found a prescription increase of all identified medications in younger and middle age groups over the years of 2014 to 2021, extrapolated by the Swiss population. The increase in prescription was steady in age groups aged 36 - 65 years, whereas in other age groups prescription stagnated from 2018 to 2020, with a steep increase from 2020 to 2021.

Methylphenidate prescription increased overall in all age groups with a steady increase in insureds aged 36 - 65 years. All other age groups experienced a steep increase after 2020. Only insureds aged 66 - 75 years and 76+ years ever experienced a

XSL•FO
RenderX

smaller prescription rate than at baseline in 2014, with a drop to 89% (664/746) in 2017 (Figure 1).

Lisdexamfetamine prescription was low at baseline and increased steadily in all age groups with great increase from 0 prescriptions at baseline to 39 in insureds aged 60 - 65 years. Smaller increase in prescription was seen in age groups aged 66 - 75 years and 76+ years (Figure 1).

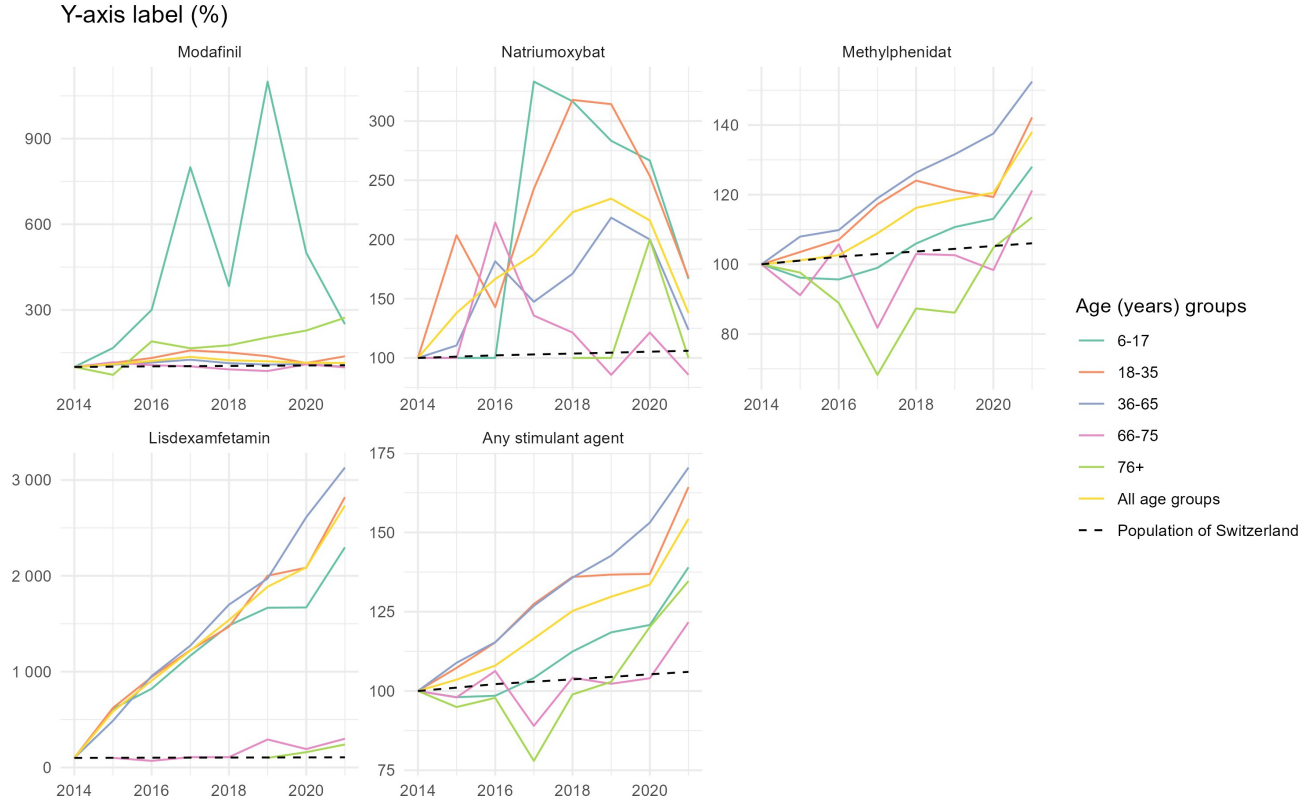
Prescription rates of modafinil—only prescribed to few—increased the most in the youngest age group (6 - 17 y). The prescribing trend in this age group strongly fluctuated. From its highest peak in 2019 rates decreased from around 1100% (66/6) to slightly less than 300% (15/6) prescription compared to the baseline in 2014. In 2021, modafinil was most frequently prescribed in age groups aged 6 - 17 years and 76+ years. All other age groups had only a moderate increase in modafinil prescriptions over the years 2014 to 2021.

Sodium oxybate overall increased the most in insureds aged 18 - 35 years and 6 - 17 years from 100% to 175% in 2021. Prescription rates for sodium oxybate only decreased overall in insureds aged 66 - 75 years. Age groups aged 36 - 65 years and 76+ experienced a similar prescription trend over the years with overall increase but declining prescription rates after 2020.

Table . Proportions of Swiss insureds with at least 1 stimulant agent or other narcolepsy treatment prescription within age groups for each year between 2014 and 2021.

Age (years)	2014, n (%)	2015, n (%)	2016, n (%)	2017, n (%)	2018, n (%)	2019, n (%)	2020, n (%)	2021, n (%)	*P* value (chi-square test for trend in proportions)
6 - 17	17,972 (1.8)	17,624 (1.8)	17,702 (1.8)	18,711 (1.9)	20,209 (2)	21,296 (2.1)	21,714 (2.1)	24,982 (2.4)	<.001
18 - 35	13,145 (0.68)	14,103 (0.72)	15,159 (0.77)	16,755 (0.85)	17,874 (0.91)	17,969 (0.92)	17,998 (0.92)	21,599 (1.10)	<.001
36 - 65	10,708 (0.31)	11,661 (0.34)	12,355 (0.35)	13,584 (0.38)	14,537 (0.41)	15,277 (0.43)	16,394 (0.45)	18,251 (0.50)	<.001
66 - 75	746 (0.1)	731 (0.1)	793 (0.1)	664 (0.1)	777 (0.1)	763 (0.1)	776 (0.1)	908 (0.1)	.10
76+	277 (0)	263 (0)	271 (0)	216 (0)	274 (0)	285 (0)	333 (0)	373 (0.10)	.05
All age groups	42,848 (0.55)	44,382 (0.57)	46,280 (0.59)	49,931 (0.63)	53,671 (0.67)	55,590 (0.69)	57,216 (0.70)	66,113 (0.81)	<.001

Figure 1. Trends of stimulant prescription from 2014 to 2021 per age group and active ingredient (indexed, base year=2014).



XSL•FO
RenderX

Factors Associated With Prescription in 2021

Characteristics of Stimulant Users

Most stimulant receivers in the year of 2021 were male representing 61% (42,803/70,396) of our population. Only modafinil was more often prescribed to women than to men with 55% (970/1776) female receivers. The highest proportion of prescriptions was provided to people living in urban areas with a proportion of 67% (46,968/70,396) of all stimulants compared to intermediate and rural area residents. Managed care was the preferred health care plan for patients receiving stimulants with 72% (50,513/70,396) receiving. A total of 45% (798/1776) of all modafinil receivers had 3+ chronic illnesses, 55% (775/1399) of those had psychiatric comorbidities, followed by cardiological and rheumatological diseases. A total of 42% (49/121) of sodium oxybate users had no comorbidities. Most common chronic illness in sodium oxybate users was psychiatric (43/72, 60%) or cardiologic (31/72, 43%). Additionally, half of the methylphenidate users had comorbidities (28,619/57,128, 50%, 28,509 had no chronic illness identified) with psychological (14,396/28,619, 50%) and other chronic conditions (7034/28,619, 25%) as the most common identified chronic diseases. Similar results were found for lisdexamfetamine users, that is, 43% (4908/11,371) users had no comorbidities and psychological (3556/6463, 55%) and other (1989/6463, 31%) comorbidities were the most common chronic diseases (Table 2).

XSL•FO
RenderX

Table . Characteristics of Swiss insureds receiving any stimulant prescription in the year 2021.

Characteristic	Modafinil (n=1776)	Sodium oxybate (n=121)	Methylphenidate (n=57,128)	Lisdexamfetamine (n=11,371)
Sex, n (%)				
Male	806 (45.4)	63 (52.1)	35,038 (61.3)	6896 (60.7)
Female	970 (54.6)	58 (48)	22,090 (38.7)	4475 (39.4)
Age (years)				
Median (IQR)	45 (33-6)	38 (29-5)	22 (14-4)	26 (16-4)
Mean (SD)	46 (16)	40 (18)	27 (16)	28 (14)
Age (years, in groups), n (%)				
6 - 17	16 (0.9)	10 (8.3)	22,948 (40.2)	3355 (29.5)
18 - 35	568 (32)	47 (38.8)	18,002 (31.5)	4401 (38.7)
36 - 65	976 (55)	48 (39.7)	15,157 (26.5)	3565 (31.4)
66 - 75	139 (7.8)	12 (9.9)	742 (1.3)	38 (0.3)
76+	77 (4.3)	4 (3.3)	279 (0.5)	12 (0.1)
Region of residence, n (%)				
Urban	1060 (59.7)	80 (66.1)	38,196 (66.9)	7632 (67.1)
Intermediate	449 (25.3)	29 (24)	11,554 (20.2)	2422 (21.3)
Rural	267 (15)	12 (9.9)	7378 (12.9)	1317 (11.6)
Health insurance status, n (%)				
Managed care	1064 (59.9)	82 (67.8)	41,178 (72.1)	8189 (72)
Standard care	712 (40.1)	39 (32.2)	15,950 (27.9)	3182 (28)
Comorbidities [14]**, n (%)**				
0	377 (21.2)	49 (40.5)	28,509 (49.9)	4908 (43.2)
1	340 (19.1)	27 (22.3)	12,617 (22.1)	2601 (22.9)
2	261 (14.7)	25 (21.7)	6678 (11.7)	1619 (14.2)
3+	798 (44.9)	20 (16.5)	9324 (16.3)	2243 (19.7)
Most frequent, n (%)	psyd:[a] 775 (43.6)	psyd: 43 (36)	psyd: 14,396 (25.2)	psyd: 3556 (31.3)
2nd most frequent, n (%)	card:[b] 592 (33.3)	card: 31 (26)	ther:[c] 7034 (12.3)	ther: 1989 (17.5)
3rd most frequent, n (%)	rheu:[d] 430 (24.2)	rheu: 17 (14)	rheu: 6588 (11.5)	resp:[e] 1435 (12.6)

[a] psyd: psychiatric.

[b]card: cardiological.

[c]ther: other.

[d]rheu: rheumatological.

[e]resp: respiratory.

Package Size

All medication was predominantly prescribed more than once within a year, with ≥5 packages prescribed in 40.4% (715/1768) of all modafinil prescriptions, 95.8% (115/120) of all sodium oxybate prescriptions, 47.5% (27,124/57,093) of all methylphenidate prescriptions, and 59.1% (6727/11,392) of all lisdexamfetamine prescriptions (Table 3).

XSL•FO
RenderX

Table . Number of Swiss insureds and their stimulant prescriptions by number of packages and by prescriber (profession of the physician) in the year 2021.

		At least 1 stimulant agent (total)	At least 1 modafinil use	At least 1 sodium oxybate use	At least 1 methylphenidate use	At least 1 lisdexamfetamin use
Total patients (N)		66,113	1768	120	57,093	11,392
Number of packages, n (%)						
1		12,235 (17.4)	536 (30.3)	0 (0)	9984 (17.5)	1715 (15.1)
2		9301 (13.2)	228 (12.9)	5 (4.2)	7968 (14)	1099 (9.6)
3		7435 (10.6)	113 (6.4)	0 (0)	6300 (11)	1022 (9)
4		6721 (9.6)	176 (10)	0 (0)	5716 (10)	829 (7.3)
≥5		34,681 (49.3)	715 (40.4)	115 (95.8)	27,124 (47.5)	6727 (59.1)
Package sizes, median (IQR)[a]						
1		—[b]	30 (30-90)	—	30 (30-100)	30 (30-30)
2		—	90 (60-90)	—	50 (30-100)	30 (30-30)
3		—	90 (67.5 - 90)	—	50 (30-83)	30 (30-30)
4		—	90 (90-90)	—	50 (35-72)	30 (30-30)
≥5		—	90 (90-90)	—	45 (30-60)	30 (30-30)
Prescriber of the issued prescriptions, n (%)						
General practitioner (GP)		27,323 (41.3)	639 (36.1)	10 (8.3)	24,127 (42.3)	3529 (31)
	Only	22,578 (82.6)	476 (74.5)	10 (100)	19,521 (80.9)	1762 (49.9)
	>50	1911 (7)	63 (9.9)	0 (0)	1802 (7.5)	717 (20.3)
	Rest	2834 (10.4)	100 (15.6)	0 (0)	2804 (11.6)	1050 (29.8)
Psychiatrist		28,057 (42.4)	336 (19)	5 (4.2)	23,553 (41.3)	6222 (54.6)
	Only	23,267 (82.9)	220 (65.5)	0 (0)	18,167 (77.1)	3204 (51.5)
	>50	2247 (8)	31 (9.2)	5 (100)	2260 (9.6)	1308 (21)
	Rest	2543 (9.1)	85 (25.3)	0 (0)	3126 (13.3)	1710 (27.5)
Neurologist		748 (1.1)	181 (10.2)	37 (30.8)	545 (1)	63 (0.6)
	Only	558 (74.6)	97 (53.6)	11 (29.7)	384 (70.5)	26 (41.3)
	>50	99 (13.2)	31 (17.1)	20 (54.1)	48 (8.8)	25 (39.7)
	Rest	91 (12.2)	53 (29.3)	6 (16.2)	113 (20.7)	12 (19)
Other specialists		7962 (12)	286 (16.2)	15 (12.5)	6680 (11.7)	1300 (11.4)
	Only	4599 (57.8)	176 (61.5)	10 (66.7)	3782 (56.6)	437 (33.6)
	>50	955 (12)	28 (9.8)	5 (33.3)	801 (12)	234 (18)
	Rest	2408 (30.2)	82 (28.7)	0 (0)	2097 (31.4)	629 (48.4)
Inpatient psychiatry		6000 (9.1)	70 (4)	6 (5)	5027 (8.8)	1234 (10.8)
	Only	3429 (57.2)	27 (38.6)	0 (0)	2857 (56.8)	362 (29.3)
	>50	898 (15)	11 (15.7)	0 (0)	756 (15)	214 (17.3)
	Rest	1673 (27.9)	32 (45.7)	6 (100)	1414 (28.1)	658 (53.3)
Inpatient acute, rehabilitation, and nursing		5549 (8.4)	496 (28.1)	60 (50)	4251 (7.4)	1110 (9.7)
	Only	3312 (59.7)	284 (57.3)	11 (18.3)	2376 (55.9)	415 (37.4)
	>50	756 (13.6)	47 (9.5)	32 (53.3)	699 (16.4)	136 (12.3)
	Rest	1481 (26.7)	165 (33.3)	17 (28.3)	1176 (27.7)	559 (50.4)

[a]Median (IQR) package size (number of units) per number of prescribed packages. Not shown for sodium oxybate, since this substance is a liquid.

XSL•FO
RenderX

[b]Not applicable.

Prescribers

In 2021, psychiatrists were the most frequent prescribers with 42.4% (28,057/66,172) of all the prescribed medications. Lisdexamfetamine (6222/11,396, 54.6%), followed by methylphenidate with 41.3% (23,553/57,038) were mainly prescribed by them. By contrast modafinil and sodium oxybate were rarely prescribed by psychiatrists as 19% and 4.2% (336/1769 and 5/120), respectively. If chosen as a prescriber, psychiatrists are often the only source of prescription for 82.9% (23,267/28,057) insured individuals of all medication of 2021.

The second most frequent prescribers were GPs, with a similar high proportion of 41.3% (27,323/66,172) of all drug prescriptions. More specifically 42.3% (24,127/57,038) methylphenidate was prescribed by GPs followed by modafinil (639/1769, 36.1%), lisdexamfetamine (3529/11,396, 31%), and smaller proportions for sodium oxybate (10/120, 8.3%). GPs were most often the exclusive prescribers of the medications (22,578/27,323, 82.6%). Only 10.4% (2834/27,323) of prescriptions by GPs shared the prescribing job with other medical specialists.

Neurologists were rarely prescribers of stimulants or narcolepsy treatments, as only 1.1% (748/66,172) of all prescriptions were invoiced by them. Only sodium oxybate was the most frequent (37/120, 30.8%) medication prescribed by neurologists. If chosen as the prescriber, they are often the only source (558/748, 74.6%) from which individuals received the prescriptions in 2021 (Table 3).

Concerning invoices handed in by hospitals or psychiatric clinics or rehabilitation facilities, they only made up a small part of overall stimulant invoices. An exception is sodium oxybate, of which 45.1% (60/133) of invoices are issued by an acute clinic, rehabilitation clinic, or nursing home (Table 3).

Association of Stimulant Use and Outcomes

We found an association with patients receiving a stimulant or narcolepsy treatment prescription and increased hospitalizations in a psychiatric facility (odds ratio [OR] 5.30, 95% CI 4.63 - 6.08, *P*<.001). In contrast, there was a negative association between stimulant prescription and hospitalization in an acute medical care facility (OR 0.77, 95% CI 0.73 - 0.82, *P*<.001). There were no significant associations between stimulant prescription and the total length of inpatient stay (OR 1, 95% CI 1 - 1, *P*=.13; Table 4).

Table . Regression model of predicting the outcomes of hospitalization and length of stay in Swiss insureds, who received stimulant prescription in the year 2021.

Characteristic		Odds ratio (95% CI)	*P* value
Sex			
	Male	—[a] (—)	—
	Female	0.89 (0.86 - 0.92)	<.001
Age (years, in groups)			
	6-17	1.17 (1.13-1.21)	<.001
	18 - 35	— (—)	—
	36 - 65	0.62 (0.60 - 0.64)	<.001
	66 - 75	0.31 (0.27 - 0.34)	<.001
	76+	0.24 (0.20 - 0.28)	<.001
Region of residence			
	Urban	— (—)	—
	Intermediate	1.09 (1.05 - 1.13)	<.001
	Rural	0.97 (0.94 - 1.02)	.2
Health insurance status			
	Standard	— (—)	—
	Managed care	0.69 (0.67 - 0.71)	<.001
Number of comorbidities		1.56 (1.54 - 1.59)	<.001
Total inpatient length of stay		1 (1 - 1)	.13
Hospitalization acute (yes or no)		0.77 (0.73 - 0.82)	<.001
Hospitalization psychiatry (yes or no)		5.3 (4.63 - 6.08)	<.001

[a]Not applicable.

XSL•FO
RenderX

When only focusing on methylphenidate or lisdexamfetamine compared to a balanced sample of nonstimulant users, we found a positive association between their prescription and hospitalization in a psychiatric facility (OR 6.85, 95% CI 5.89 - 7.99, *P*<.001). No significant association was found between the prescription rate and the total inpatient length of stay (OR 1, 95% CI 1 - 1, *P*=.1). Hospitalization in an acute medical care facility was less likely (OR 0.76, 95% CI 0.71 - 0.8, *P*<.001) to prescribe methylphenidate or lisdexamfetamine (Table 5).

Table . Regression models of predicting the outcomes of hospitalization and length of stay in Swiss insureds who received stimulant prescriptions in the year 2021. Prescriptions of methylphenidate or lisdexamfetamine are shown in the left columns, prescriptions of modafinil or sodium oxybate are shown in the right columns.

Characteristic		Methylphenidate- or lisdexamfetamine-users versus nonusers		Modafinil- or sodium oxybate-users versus nonusers	
		OR[a] (95% CI)	*P* value	OR[a] (95% CI)	*P* value
Sex					
	Male	— (—)	—	— (—)	—
	Female	0.9 (0.87 - 0.93)	<.001	0.84 (0.72 - 0.97)	.02
Age (years, in groups)					
	6-17	1.17 (1.13-1.21)	<.001	1.17 (0.73-1.88)	.5
	18 - 35	— (—)	—	— (—)	—
	36 - 65	0.63 (0.6 - 0.65)	<.001	0.57 (0.48 - 0.68)	<.001
	66 - 75	0.31 (0.28 - 0.36)	<.001	0.25 (0.19 - 0.34)	<.001
	76+	0.26 (0.21 - 0.31)	<.001	0.15 (0.09 - 0.23)	<.001
Region of residence					
	Urban	— (—)	—	— (—)	—
	Intermediate	1.11 (1.07 - 1.15)	<.001	0.96 (0.8 - 1.15)	.7
	Rural	0.95 (0.91 - 0.99)	.019	1.11 (0.89 - 1.38)	.3
Health insurance status					
	Standard	— (—)	—	— (—)	—
	Managed care	0.69 (0.67 - 0.71)	<.001	0.58 (0.5 - 0.68)	<.001
Number of comorbidities		1.55 (1.53 - 1.57)	<.001	1.62 (1.54 - 1.71)	<.001
Total inpatient length of stay		1 (1-1)	.10	1 (1-1)	.8
Hospitalization acute (yes or no)		0.76 (0.71 - 0.8)	<.001	1.23 (0.96 - 1.56)	.10
Hospitalization psychiatry (yes or no)		6.85 (5.89 - 7.99)	<.001	2.15 (1.17 - 4.23)	.02

[a]Odds ratio

[b]Not applicable.

When only focusing on modafinil or sodium oxybate compared to a balanced sample of nonstimulant users, we found a positive association between their prescription and hospitalization in a psychiatric facility (OR of 2.15, 95% CI 1.17-4.23, *P*=.02). We found no significant association between total inpatient length of stay (OR 1, 95% CI 1 - 1, *P*=.80) and hospitalization in an acute medical care facility (OR 1.23, 95% CI 0.96 - 1.56, *P*=.10; Table 5).

Discussion

Summary

We found an increasing trend of stimulant and narcoleptic drug prescriptions in Switzerland from 0.6% to 0.8% over the years 2014 to 2021. Most stimulants are prescribed continuously with more than 5 packages in 1 year per insured to underaged individuals with no comorbidities. Psychiatrists and GPs are often the prescribers of stimulants, much more frequently than neurologists.

Context in Research

Compared with data on global stimulant prescription rates, our results are notably lower, showing a smaller increase in prescription rates than in the United States (5.6% - 6.1%), New Zealand (0.75% - 1.06%), and Denmark (0.03% - 0.73%) [5-7]. Given the ADHDs/ADDs prevalence rates of 4% in adult men and 5.3% in children in Switzerland [2,3], compared to a worldwide prevalence of between 5% and 11.4% [1,15,16], it is reasonable to interpret the lower Swiss prescription rates as either an indication of underprescription or as lower stimulant misuse rates in Switzerland [15]. Even when taking all stimulants together, current prescription rates do not reach the prevalence rate of ADHDs/ADDs in Switzerland. The fact that

XSL•FO
RenderX

ADHDs/ADDs are also treated nonpharmacologically is another argument for assuming that our result of low prescription prevalence is lower than the disease prevalence. Unfortunately, there are no data that quantifies the extent of drug or nonpharmacological treatment for ADHDs/ADDs and thus could help define the normal gap between prescription rate and disease prevalence. Our distribution of prescription rates corresponds to the disease distribution in different age groups, with prescriptions and prevalence of ADHDs/ADDs being higher in the younger age groups [17]. The distribution of prescriptions by gender reflects the current disease prevalence of ADHDs/ADDs, as we found a slightly lower proportion of females than males, in line with another summary by Thapar and Cooper [17]

Stimulants were predominantly prescribed in urban areas. A higher population compared to rural areas, a higher density of prescribing physicians, and a higher number of hospitals specialized in diagnosis and treatment of these diseases in urban settings may account for this predominant prescription pattern. Vice versa underprescription in rural areas could be due to reduced access to adequate health care services.

Methylphenidate was by far the most prescribed stimulant agent, followed by lisdexamfetamine, a new stimulant showing promising results for treatment of ADHDs/ADDs [18]. The prescription characteristics were very similar between the two stimulants, with lisdexamfetamine more often prescribed to older adults than methylphenidate. This shows that lisdexamfetamine is possibly used as a second medication after methylphenidate was prescribed in the young and did not give continuous results while the patients aged.

Prescriptions were most frequently issued by psychiatrists. Since methylphenidate and lisdexamfetamine are standard treatment for the psychiatric diseases ADHDs and ADDs, this prescription pattern makes sense from a health provider perspective and is confirmed by other findings about prescribers of stimulants [19].

Prescription characteristics of sodium oxybate should reflect its specificity for narcolepsy because it is not indicated for any other disease. Here we find most of the prescriptions in middle aged groups and fewer in children and teenagers, with a nearly even distribution between women and men. This grouping does not match with the expected narcolepsy features of young patients with a possible second peak in the late forties [20]. The lack of prescription in young patients is best explained by a missed or severely delayed diagnosis of narcolepsy, which is in line with the recently published delayed diagnosis for Switzerland and other European countries [21,22]. Surprisingly, neurologists who diagnose and treat narcolepsy, are rarely the patients' prescription source, even for highly specific and not easy to handle medication, such as sodium oxybate. Multiple reasons may account for this prescription practice, among them is the lower barrier for receiving an appointment with GPs before prescriptions expires, compared to neurologists.

We found no significant change in the odds of at least one stimulant prescription in patients hospitalized in an acute hospital but a significant increase in the odds of patients hospitalized in a psychiatric facility. This reflects the fact that ADHDs/ADDs are often overlapping or comorbid with other psychiatric diseases which can lead to hospitalization, such as addiction, disruptive disorders, anxiety disorders, or bipolar disorders [23-26], and by the severity of disease.

We assume that our data is representative for Switzerland, since we analyzed claims data extrapolated to the entire population from one of the biggest insurance companies in Switzerland, with a nearly equal distribution across the country. As stimulant agents are only accessed through prescription, we were able to register all invoices for the medication in question in real world; therefore, our data minimize sampling bias and recall bias that frequently influence the accuracy and reliability of retrospective studies.

Limitations

This study is an analysis of health insurance claims data, which does not contain any information on the clinical reason of why a medication was indicated. We therefore could not distinguish between prescriptions according to current treatment guidelines and prescriptions of pharmacological treatment for diseases without proper diagnosis.

We identified prescribers by Zahlsteller register (registered number for medical personel allowed to bill insurance companies) number and medication prescription pattern, which may lead in some cases to misclassification, as some physicians share Zahlsteller register numbers in group practices and GPs go sometimes through additional training as psychiatrists or neurologists.

Conclusion

The prescription of stimulants and sodium oxybate in Switzerland increased slightly but continuously over the past years, but at lower rates compared to the estimated prevalence of central hypersomnolence disorders and ADHDs/ADDs. Most stimulants are prescribed by psychiatrists, closely followed by GPs. The increased odds for hospitalization to psychiatric institutions for stimulant receivers reflects the severity of disease and the higher psychiatric comorbidities in these patients.

Conflicts of Interest

None declared.

References

1. Polanczyk GV, Willcutt EG, Salum GA, Kieling C, Rohde LA. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. Int J Epidemiol 2014 Apr;43(2):434-442. [doi: 10.1093/ije/dyt261] [Medline: 24464188]

XSL•FO
RenderX

2. Steinhausen HC, Metzke CW, Meier M, Kannenberg R. Prevalence of child and adolescent psychiatric disorders: the Zürich epidemiological study. Acta Psychiatr Scand 1998 Oct;98(4):262-271. [doi: 10.1111/j.1600-0447.1998.tb10082.x] [Medline: 9821446]
3. Estévez N, Eich-Höchli D, Dey M, Gmel G, Studer J, Mohler-Kuo M. Prevalence of and associated factors for adult attention deficit hyperactivity disorder in young Swiss men. PLoS ONE 2014;9(2):e89298. [doi: 10.1371/journal.pone.0089298] [Medline: 24586672]
4. Safer DJ. Recent trends in stimulant usage. J Atten Disord 2016 Jun;20(6):471-477. [doi: 10.1177/1087054715605915] [Medline: 26486603]
5. Board AR, Guy G, Jones CM, Hoots B. Trends in stimulant dispensing by age, sex, state of residence, and prescriber specialty - United States, 2014-2019. Drug Alcohol Depend 2020 Dec 1;217:108297. [doi: 10.1016/j.drugalcdep.2020.108297] [Medline: 32961454]
6. Barczyk ZA, Rucklidge JJ, Eggleston M, Mulder RT. Psychotropic medication prescription rates and trends for New Zealand children and adolescents 2008-2016. J Child Adolesc Psychopharmacol 2020 Mar;30(2):87-96. [doi: 10.1089/cap.2019.0032] [Medline: 31633377]
7. Steinhausen HC, Bisgaard C. Nationwide time trends in dispensed prescriptions of psychotropic medication for children and adolescents in Denmark. Acta Psychiatr Scand 2014 Mar;129(3):221-231. [doi: 10.1111/acps.12155] [Medline: 23738593]
8. Maier LJ, Haug S, Schaub MP. Prevalence of and motives for pharmacological neuroenhancement in Switzerland--results from a national Internet panel. Addiction 2016 Feb;111(2):280-295. [doi: 10.1111/add.13059] [Medline: 26189457]
9. Raman SR, Man KKC, Bahmanyar S, et al. Trends in attention-deficit hyperactivity disorder medication use: a retrospective observational study using population-based databases. Lancet Psychiatry 2018 Oct;5(10):824-835. [doi: 10.1016/S2215-0366(18)30293-1] [Medline: 30220514]
10. Gumy C, Huissoud T, Dubois-Arber F. Prevalence of methylphenidate prescription among school-aged children in a Swiss population: increase in the number of prescriptions in the Swiss Canton of Vaud, from 2002 to 2005, and changes in patient demographics. J Atten Disord 2010 Nov;14(3):267-272. [doi: 10.1177/1087054709356386] [Medline: 20484708]
11. Benson K, Flory K, Humphreys KL, Lee SS. Misuse of stimulant medication among college students: a comprehensive review and meta-analysis. Clin Child Fam Psychol Rev 2015 Mar;18(1):50-76. [doi: 10.1007/s10567-014-0177-z] [Medline: 25575768]
12. Charach A, Ickowicz A, Schachar R. Stimulant treatment over five years: adherence, effectiveness, and adverse effects. J Am Acad Child Adolesc Psychiatry 2004 May;43(5):559-567. [doi: 10.1097/00004583-200405000-00009]
13. Lakhan SE, Kirchgessner A. Prescription stimulants in individuals with and without attention deficit hyperactivity disorder: misuse, cognitive impact, and adverse effects. Brain Behav 2012 Sep;2(5):661-677. [doi: 10.1002/brb3.78] [Medline: 23139911]
14. Huber CA, Szucs TD, Rapold R, Reich O. Identifying patients with chronic conditions using pharmacy data in Switzerland: an updated mapping approach to the classification of medications. BMC Public Health 2013 Oct 30;13(1):1-10. [doi: 10.1186/1471-2458-13-1030] [Medline: 24172142]
15. Danielson ML, Claussen AH, Bitsko RH, et al. ADHD prevalence among U.S. children and adolescents in 2022: diagnosis, severity, co-occurring disorders, and treatment. J Clin Child Adolesc Psychol 2024;53(3):343-360. [doi: 10.1080/15374416.2024.2335625] [Medline: 38778436]
16. Thomas R, Sanders S, Doust J, Beller E, Glasziou P. Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. Pediatrics 2015 Apr;135(4):e994-1001. [doi: 10.1542/peds.2014-3482] [Medline: 25733754]
17. Thapar A, Cooper M. Attention deficit hyperactivity disorder. The Lancet 2016 Mar;387(10024):1240-1250. [doi: 10.1016/S0140-6736(15)00238-X]
18. Roskell NS, Setyawan J, Zimovetz EA, Hodgkins P. Systematic evidence synthesis of treatments for ADHD in children and adolescents: indirect treatment comparisons of lisdexamfetamine with methylphenidate and atomoxetine. Curr Med Res Opin 2014 Aug;30(8):1673-1685. [doi: 10.1185/03007995.2014.904772] [Medline: 24627974]
19. Karanges EA, Stephenson CP, McGregor IS. Longitudinal trends in the dispensing of psychotropic medications in Australia from 2009–2012: focus on children, adolescents and prescriber specialty. Aust N Z J Psychiatry 2014 Oct;48(10):917-931. [doi: 10.1177/0004867414538675] [Medline: 24927734]
20. Dauvilliers Y, Montplaisir J, Molinari N, et al. Age at onset of narcolepsy in two large populations of patients in France and Quebec. Neurology (ECronicon) 2001 Dec 11;57(11):2029-2033. [doi: 10.1212/wnl.57.11.2029] [Medline: 11739821]
21. Oberle D, Drechsel-Bäuerle U, Schmidtmann I, Mayer G, Keller-Stanislawski B. Incidence of narcolepsy in Germany. Sleep 2015 Oct 1;38(10):1619-1628. [doi: 10.5665/sleep.5060] [Medline: 25902804]
22. Zhang Z, Dauvilliers Y, Plazzi G, et al. Idling for decades: a european study on risk factors associated with the delay before a narcolepsy diagnosis. Nat Sci Sleep 2022;14:1031-1047. [doi: 10.2147/NSS.S359980] [Medline: 35669411]
23. van de Glind G, Van Emmerik‐van Oortmerssen K, Carpentier PJ, et al. The International ADHD in Substance Use Disorders Prevalence (IASP) study: background, methods and study population. Int J Methods Psych Res 2013 Sep;22(3):232-244 [FREE Full text] [doi: 10.1002/mpr.1397]

XSL•FO
RenderX

24. Strålin P, Hetta J. First episode psychosis and comorbid ADHD, autism and intellectual disability. Eur Psychiatry 2019 Jan;55:18-22. [doi: 10.1016/j.eurpsy.2018.09.007] [Medline: 30384107]
25. Jensen PS, Hinshaw SP, Kraemer HC, et al. ADHD comorbidity findings from the MTA study: comparing comorbid subgroups. J Am Acad Child Adolesc Psychiatry 2001 Feb;40(2):147-158. [doi: 10.1097/00004583-200102000-00009] [Medline: 11211363]
26. Marangoni C, De Chiara L, Faedda GL. Bipolar disorder and ADHD: comorbidity and diagnostic distinctions. Curr Psychiatry Rep 2015 Aug;17(8):604. [doi: 10.1007/s11920-015-0604-y] [Medline: 26084666]

Abbreviations

ADHD/ADD: attention-deficit/hyperactivity disorder/attention deficit disorder
GP: general practitioner
OR: odds ratio

Edited by A Mavragani, T Sanchez; submitted 25.10.23; peer-reviewed by H Demir, K Karimova; revised version received 29.09.24; accepted 30.09.24; published 07.01.25.

Please cite as:
Scharf T, Huber CA, Näpflin M, Zhang Z, Khatami R
Trends in Prescription of Stimulants and Narcoleptic Drugs in Switzerland: Longitudinal Health Insurance Claims Analysis for the Years 2014-2021
JMIR Public Health Surveill 2025;11:e53957
URL: https://publichealth.jmir.org/2025/1/e53957
doi:10.2196/53957

© Tamara Scharf, Carola A Huber, Markus Näpflin, Zhongxing Zhang, Ramin Khatami. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 7.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessment of Bidirectional and Threshold-Response Associations Between Cognitive Function and Physical Performance: Nationwide Cross-Sectional Study

Huixiu Hu[1]; Lanying Xie[2*]; Yuqing Hao[3*]; Yajie Zhao[4]; Huanhuan Luo[5]; Kang Yu[6]; Chao Sun[1]

[1]Department of Nursing, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, Beijing Hospital, 1 Da Hua Road, DongDan, Dongcheng District, Beijing, China

[2]School of Nursing, Beijing University of Chinese Medicine, Beijing, China

[3]Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[4]Department of Cardiology, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, Beijing Hospital, Beijing, China

[5]School of nursing, Peking University, Beijing, China

[6]Department of Clinical Nutrition, Chinese Academy of Medical Science and Peking Union Medical College, Peking Union Medical College Hospital, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Chao Sun
Department of Nursing, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, Beijing Hospital, 1 Da Hua Road, DongDan, Dongcheng District, Beijing, China

Related Article:

This is a corrected version. See correction statement: https://publichealth.jmir.org/2025/1/e88582

Abstract

Background: The global aging population faces increasing risks of cognitive and physical decline, yet the bidirectional and nonlinear dynamics between these domains remain poorly understood.

Objective: This study aims to investigate the bidirectional and nonlinear associations between cognitive function and physical performance in older adults and to identify threshold effects and population subgroups at elevated risk.

Methods: This multicenter study of 20,671 older adults in China analyzed the bidirectional asymmetric and threshold-response associations between cognitive and physical function using logistic or linear regression and restricted cubic spline, with subgroup and interaction analyses.

Results: The study revealed an asymmetrical bidirectional relationship between cognitive function and physical performance, with cognitive impairment having a stronger association with physical function (Short Physical Performance Battery [SPPB], odds ratio [OR] 2.45, *P*<.05 vs 1.16,*P*=.051; gait speed: OR 2.36,*P*<.05 vs 1.89, *P*<.050). Restricted cubic spline analysis identified 3 key inflection points at Mini-Mental State Examination (MMSE) scores of 19, 24, and 29. When MMSE <24, cognitive improvement significantly protected physical performance (SPPB: OR 0.86, 95% CI 0.84 - 0.87; gait speed: OR 0.84, 95% CI 0.82 - 0.85). The critical intervention window was between MMSE 19 - 24 (SPPB, OR 0.66, 95% CI 0.57-0.77). At MMSE ≥24, balance protection remained significant (OR 0.69, 95% CI 0.68 - 0.70), but overall physical performance showed diminished effects, with SPPB demonstrating a weaker, marginal association (*P*=.051). A paradoxical reversal effect was observed for gait speed (OR 1.05, 95% CI 1.04 - 1.07). High-income (OR 5.24, 95% CI 3.91 - 7.02) and high-education subgroups (OR 3.98, 95% CI 1.48 - 10.70) exhibited heightened vulnerability.

Conclusions: This study demonstrates a bidirectional, asymmetric relationship between cognitive function and physical performance, identifies a nonlinear threshold effect, and highlights population heterogeneity.

(JMIR Public Health Surveill 2025;11:e80575) doi:10.2196/80575

KEYWORDS

cognitive function; physical performance; threshold effect; aging; bidirectional association

XSL•FO
RenderX

Introduction

The global population is aging at an accelerated rate, with the proportion of individuals aged ≥60 years projected to reach 2.1 billion by 2050 [1]. This demographic shift amplifies the dual burden of cognitive decline and physical disability, which synergistically exacerbate risks of falls, hospitalization, and dependency [2-4]. Cognitive function, encompassing memory, executive control, and information processing, deteriorates progressively from mild cognitive impairment to dementia. Concurrently, physical performance, as measured by the Short Physical Performance Battery (SPPB) [5], declines linearly with age, marked by reduced balance, gait speed, and muscle strength [6,7]. Critically, lower SPPB scores predict mortality and institutionalization [8-11], while cognitive impairment further compounds these risks through bidirectional pathways [4].

Extensive research [12-14] has demonstrated that impaired physical performance can significantly affect cognitive function, with slow gait speed being recognized as an early indicator for predicting cognitive decline and dementia. In addition, cognitive decline can also lead to impaired physical performance, with studies highlighting a bidirectional relationship between physical and cognitive health [15,16]. However, critical gaps persist. First, balance and muscle strength, key components of physical function, are rarely examined in relation to cognition. Second, evidence on bidirectionality is contradictory; one study [17] reported cognitive effects on gait speed but not balance, another linked balance deficits to cognitive impairment [18], and longitudinal data suggest cognitive decline drives physical deterioration more strongly than the reverse [15,19]. Third, prior studies predominantly assume linear relationships, neglecting potential thresholds where cognitive deterioration disproportionately escalates physical disability—a limitation obscuring precision intervention targets.

Our nationwide multicenter study explores bidirectional, nonlinear links between cognition and physical performance in older adults, identifying critical cognitive thresholds that increase physical risks. These findings guide precision interventions for dual decline, offering scalable strategies to reduce disability in aging populations.

Methods

Study Design

This study adopts a nationwide cross-sectional design to explore the relationship between cognitive function and physical performance, following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [20].

Participants

A multistage stratified sampling method was used to recruit 20,868 community-dwelling adults aged 60+ years from 6 provinces in China (Fujian, Guangdong, Nanjing, Sichuan, Xinjiang, and Guizhou) between January and August 2022. The required sample size was calculated using the formula , yielding a sample size of 13,830 for a 95% CI with a margin of error of 0.05 and SD of 3 [21]. To account for invalid questionnaires, a 20% adjustment was made, resulting in a final sample size of 17,288. Inclusion criteria: participants aged 60+ with adequate communication and complete assessments. Exclusion criteria: participants with severe physical disabilities preventing SPPB testing or acute medical conditions (eg, postoperative recovery).

Measures

Physical Performance

Physical performance was evaluated using the SPPB, a validated tool from the National Institute on Aging [22]. The SPPB comprises 3 subtests: balance, gait speed, and muscle strength, each component scored 0 - 4, with a total score ranging from 0 to 12. Participants were classified as having normal (10–12) or poor (0 - 9) physical performance [8,23]. In addition, a score of 4 in any subtest indicated normal function, while 0 - 3 suggested impairment [24].

Balance Test

Three standing positions were feet-together, semi-tandem, and tandem. Point distribution was as follows: participants held the position for >10 seconds in feet-together or semi-tandem, earning them 1 point; standing for 3 - 9.9 seconds in the tandem position earned 1 point, while standing for ≥10 seconds earned 2 points.

Gait Speed Test

The 2.44-meter walking speed test assigned scores based on speed: <0.43 m/s (1 point), 0.44 - 0.60 m/s (2 points), 0.61 - 0.77 m/s (3 points), and ≥0.78 m/s (4 points).

Muscle Strength Test

The Five-Times-Sit-to-Stand Test (FTSST) was used to assess lower limb muscle strength, with scores based on the time taken: 16.70 - 60 seconds (1 point), 13.70 - 16.69 seconds (2 points), 11.20 - 13.69 seconds (3 points), and ≤11.19 seconds (4 points).

Cognitive Function

Cognitive function was evaluated using the Mini-Mental State Examination (MMSE) [25], with the Chinese version validated by Wang et al [26] showing high test-retest reliability (intraclass correlation coefficient=0.91). The MMSE assesses five cognitive domains: orientation (time and place), memory (immediate and delayed recall), attention and calculation, language, and visuospatial ability. Scores ranged from 0 to 30, with higher scores indicating better cognitive function. The criteria for normal cognitive function vary by educational level: (1) middle school or above: score >24 (normal), ≤24 (impairment); (2) primary school: score >20 (normal), ≤20 (impairment); (3) illiteracy: score >17 (normal), ≤17 (impairment).

Covariates

Sociodemographic characteristics and health behaviors were considered covariates. Sociodemographic characteristics included age, sex (male and female); BMI (<18.5, 18.5 - 23.9, ≥24 kg/m^2), educational level (primary school or below, middle school, high school or vocational school, and college or above), marital status (married vs other, including divorced, widowed,

XSL•FO
RenderX

or single), number of children (none, 1, 2, 3, or more), residential area (rural, urban), living arrangement (living alone, living with relatives, living with a spouse), geographic region (eastern or western), data source (home-based or community health service center), and monthly household income (<5000 or ≥5000).

Health behaviors included the presence of chronic diseases (defined as having at least one: hypertension, coronary heart disease, cerebrovascular disease, or diabetes), long-term medications use (none, 1‐2, or ≥3), alcohol consumption (never, formerly, or currently), smoking status (never, formerly, or currently), weekly social participation frequency (none, 1‐3, or ≥4 d), weekly exercise frequency (none, 1‐3, or ≥4 d), self-reported social support (insufficient, inadequate, or sufficient), and self-reported visual and hearing impairment (no or yes).

Data Collection

A designated assessor was assigned to each research site to oversee data collection using the "Jingyice Elderly Function Assessment Platform," a WeChat (Tencent) mini-program developed by the research team. All assessors, registered nurses from community health centers, nursing homes, or hospitals, received uniform online training on the platform and study procedures. Before assessment, participants were informed of the study's aims and procedures, and written informed consent was obtained. Assessors conducted face-to-face interviews and functional tests, entering responses and test results (eg, cognitive assessments and physical performance measures) into the mini-program, which synchronized data to a secure web-based platform for real-time monitoring. To ensure data quality, the backend recorded the time and duration of each evaluation, and the research team reviewed data weekly, focusing on submissions completed between 10 PM and 8 AM or within 15 minutes. Questionnaires and test records with incomplete scale data, >15% missing demographic information, patterned responses, or logical inconsistencies were excluded.

Statistical Analysis

Continuous variables are presented as mean SD, while categorical variables are summarized as frequencies (%) via Pearson chi-square tests. To assess bidirectional associations, cognitive function (MMSE score) was modeled as both continuously and categorically (impaired or normal), whereas physical performance (SPPB total score) was analyzed as a continuous variable in linear regression and dichotomized (poor or normal) in logistic regression. All models were adjusted for covariates listed in the "Measures" section.

Nonlinear relationships were assessed using restricted cubic splines (RCS), with the optimal model fit chosen based on the Akaike Information Criterion from models with 3 to 6 knots. A 2-stage segmented linear regression approach identified threshold effects using the likelihood ratio test, and the optimal threshold was determined via an iterative grid search.

Subgroup analyses by sociodemographic factors (eg, age, education, and income) were conducted, and interaction effects were assessed using product terms (eg, MMSE×income). All analyses were performed with R (version 4.2; R Core Team) and SPSS (version 25.0; IBM Corp). Statistical significance was set at a 2-tailed P value <.05.

Ethical Considerations

This study was approved by the Ethics Committee of Beijing Hospital (approval number 2023BJYYEC-446‐02). Written informed consent was obtained from all participants prior to participation. All data were anonymized, participants received no compensation, and no identifiable images are included in the article or supplementary materials. The study was conducted in accordance with the ethical standards of the Declaration of Helsinki.

Results

Baseline Characteristics

A total of 20,671 participants were included in the final analysis, with a mean age of 72.19 (SD 12.238) years; 48% (9931/20,671) were male and 52% (10,740/20,671) were female (Table 1). The majority of participants (12,703/20,671, 61.5%) had low education levels. Cognitive impairment was identified in 14.4% (2975/20,671), and 62.2% (12,857/20,671) exhibited poor physical performance with a mean SPPB score of 8.4 (SD 2.94). Domain-specific impairments were prevalent: 26.7% (5521/20,671) had balance deficits, 67.3% (13,906/20,671) showed reduced gait speed, and 58.4% (12,071/20,671) demonstrated muscle weakness, with gait speed showing the most pronounced deficit (mean score 2.34, SD 1.36; Figure 1). Furthermore, there were also differences in SPPB physical performance across various covariate subgroups (Table 1).

XSL•FO
RenderX

Table . Baseline characteristics.

Characteristics	Overall (N=20,671)	SPPB[a]		*P* value
		Normal (n=7804)	Poor (n=12,867)	
Age (years), mean (SD)	72.19 (12.238)	70.14 (11.791)	73.43 (12.337)	<.001
Sex, n (%)				.23
Male	9931 (48)	3791 (48.6)	6140 (47.7)	
Female	10,740 (52)	4013 (51.4)	6727 (52.3)	
BMI (kg/m^2), n (%)				<.001
<18.5	1218 (5.9)	328 (4.2)	890 (6.9)	
18.5 - 24	11,559 (55.9)	4326 (55.4)	7233 (56.2)	
≥24	7894 (38.2)	3150 (40.4)	4744 (36.9)	
Education, n (%)				<.001
Primary school or below	12,703 (61.5)	4229 (54.2)	8474 (65.9)	
Middle school	4401 (21.3)	1912 (24.5)	2489 (19.3)	
High school or vocational school	2635 (12.7)	1247 (16)	1388 (10.8)	
College or above	932 (4.5)	416 (5.3)	516 (4)	
Marital status, n (%)				<.001
Other	2271 (11)	708 (9.1)	1563 (12.1)	
Married	18,400 (89)	7096 (90.9)	11,304 (87.9)	
Children, n (%)				<.001
None	304 (1.5)	125 (1.6)	179 (1.4)	
One	4436 (21.5)	2123 (27.2)	2313 (18)	
Two	7164 (34.7)	2907 (37.3)	4257 (33.1)	
Three and more	8767 (42.4)	2649 (33.9)	6118 (47.5)	
Residential category, n (%)				<.001
Rural	9729 (47.1)	3369 (43.2)	6360 (49.4)	
Urban	10,942 (52.9)	4435 (56.8)	6507 (50.6)	
Living arrangement, n (%)				<.001
Living alone	1001 (4.8)	356 (4.6)	645 (5)	
Living with relatives	5833 (28.2)	2338 (30)	3495 (27.2)	
Living with spouse	13,837 (66.9)	5110 (65.5)	8727 (67.8)	
Region, n (%)				.62
Eastern	12,666 (61.3)	4765 (61.1)	7901 (61.4)	
Western	8005 (38.7)	3039 (38.9)	4966 (38.6)	
Source, n (%)				<.001
Home	11,920 (57.7)	4380 (56.1)	7540 (58.6)	
Community health service center	8751 (42.3)	3424 (43.9)	5327 (41.4)	
Income (yuan/month)[b], n (%)				<.001
<5000	16,762 (81.1)	6032 (77.3)	10,730 (83.4)	
≥5000	3909 (18.9)	1772 (22.7)	2137 (16.6)	
Chronic diseases, n (%)				.79
None	7020 (34)	2659 (34.1)	4361 (33.9)	

XSL•FO
RenderX

Characteristics	Overall (N=20,671)	SPPB[a]		*P* value
		Normal (n=7804)	Poor (n=12,867)	
Yes	13,651 (66)	5145 (65.9)	8506 (66.1)	
Types of long-term medications, n (%)				.14
None	8998 (43.5)	3453 (44.2)	5545 (43.1)	
1 - 2	8321 (40.3)	3075 (39.4)	5246 (40.8)	
≥3	3352 (16.2)	1276 (16.4)	2076 (16.1)	
Alcohol, n (%)				<.001
Never	14,871 (71.9)	5415 (69.4)	9456 (73.5)	
Formerly	3761 (18.2)	1553 (19.9)	2208 (17.2)	
Currently	2039 (9.9)	836 (10.7)	1203 (9.3)	
Smoking, n (%)				<.001
Never	14,650 (70.9)	5419 (69.4)	9231 (71.7)	
Formerly	3195 (15.5)	1228 (15.7)	1967 (15.3)	
Currently	2826 (13.7)	1157 (14.8)	1669 (13)	
Social activity (d/wk), n (%)				<.001
None	5578 (27)	2021 (25.9)	3557 (27.6)	
1 - 3	7538 (36.5)	2624 (33.6)	4914 (38.2)	
≥4	7555 (36.5)	3159 (40.5)	4396 (34.2)	
Exercise (d/wk), n (%)				<.001
None	5958 (28.8)	1966 (25.2)	3992 (31)	
1 - 3	6443 (31.2)	2358 (30.2)	4085 (31.7)	
≥4	8270 (40)	3480 (44.6)	4790 (37.2)	
Social support, n (%)				.87
Insufficient	803 (3.9)	307 (3.9)	496 (3.9)	
Inadequate	1079 (5.2)	414 (5.3)	665 (5.2)	
Sufficient	18,789 (90.9)	7083 (90.8)	11,706 (91)	
Visual impairment, n (%)				.76
No	13,654 (66.1)	5165 (66.2)	8489 (66)	
Yes	7017 (33.9)	2639 (33.8)	4378 (34)	
Hearing impairment, n (%)				.27
No	17,696 (85.6)	6807 (87.2)	11,155 (86.7)	
Yes	2975 (14.4)	997 (12.8)	1712 (13.3)	
MMSE[c], mean (SD)	25.51 (4.629)	26.54 (3.62)	24.88 (5.044)	<.001
Cognitive impairment, n (%)				<.001
No	17,696 (85.6)	7204 (92.3)	10,492 (81.5)	
Yes	2975 (14.4)	600 (7.7)	2375 (18.5)	

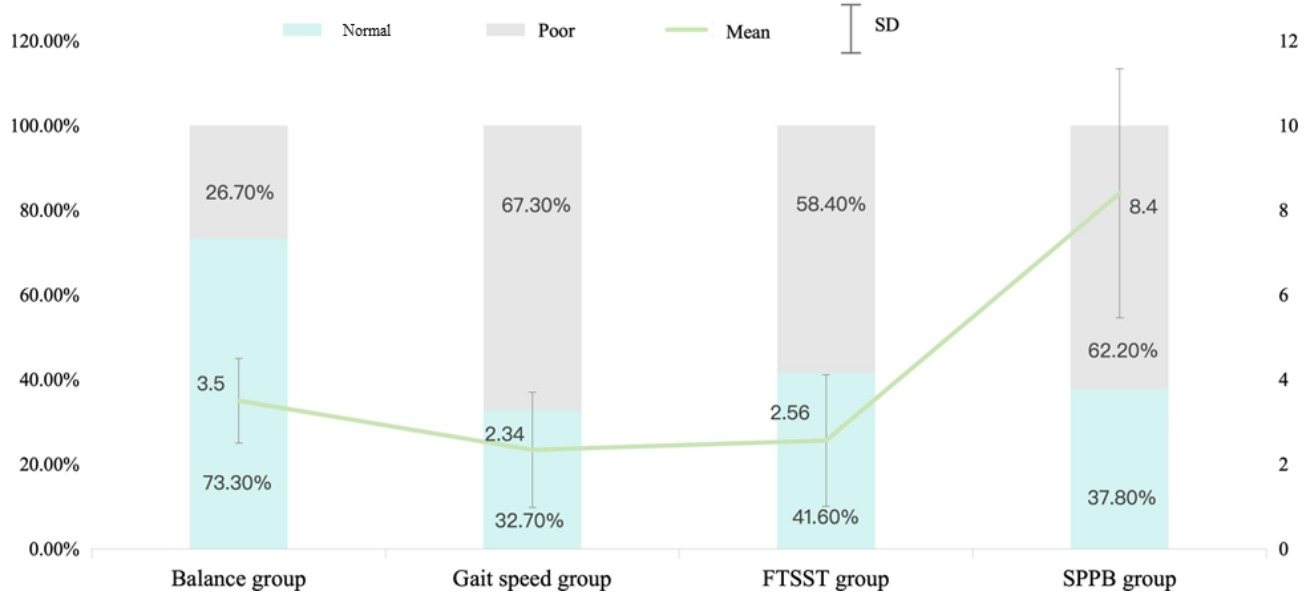
[a]SPPB: Short Physical Performance Battery.

[b]The exchange rate at the time of the study was US $1 = ¥7.113.

[c]MMSE: Mini-Mental State Examination.

XSL•FO
RenderX

Figure 1. Physical function characteristics. FTSST: Five-Times-Sit-to-Stand Test; SPPB: Short Physical Performance Battery.



Bidirectional Relationship Between Cognitive Function and Physical Performance

Association of Cognitive Function on Physical Performance

Logistic regression models, adjusted for sociodemographic and behavioral covariates, demonstrated graded protective effects of cognitive function on physical performance. For each 1-point increase in the MMSE score, the risk of poor SPPB decreased by 7% (odds ratio [OR] 0.93, 95% CI 0.92 - 0.93). Improvements were also observed in specific domains, including balance (OR 0.90, 95% CI 0.89 - 0.90), gait speed (OR 0.94, 95% CI 0.94 - 0.95), and muscle strength (OR 0.90, 95% CI 0.90 - 0.91). Individuals with cognitive impairment had nearly double the risk of poor physical performance compared to those without (OR 2.45, 95% CI 2.22 - 2.70), with gait speed (OR 2.36, 95% CI 2.13 - 2.62) and muscle weakness (OR 2.23, 95% CI 2.03 - 2.44) more strongly linked to cognitive impairment (Table 2).

XSL•FO
RenderX

Table . Logistic regression analysis of cognitive function on physical performance.

Models	SPPB[a] group, OR (95% CI)	Balance group, OR (95% CI)	Gait speed group, OR (95% CI)	FTSST[b] group, OR (95% CI)
Crude model[c]				
MMSE[d]	0.92(0.91-0.92)[e]	0.89 (0.88-0.89)[e]	0.94 (0.93-0.95)[e]	0.89 (0.89-0.90)[e]
Cognitive impairment				
No	Reference	Reference	Reference	Reference
Yes	2.72 (2.47-2.99)[e]	1.90 (1.75-2.06)[e]	2.50 (2.27-2.76)[e]	2.61 (2.38-2.85)[e]
Model 1[f]				
MMSE	0.93 (0.92-0.94)[e]	0.90 (0.89-0.90)[e]	0.95 (0.94-0.95)[e]	0.90 (0.90-0.91)[e]
Cognitive impairment				
No	Reference	Reference	Reference	Reference
Yes	2.43 (2.21-2.68)[e]	1.63 (1.49-1.78)[e]	2.36 (2.13-2.61)[e]	2.24 (2.04-2.46)[e]
Model 2[g]				
MMSE[d]	0.93 (0.92-0.93)[e]	0.90 (0.89-0.90)[e]	0.94 (0.94-0.95)[e]	0.90 (0.90-0.91)[e]
Cognitive impairment				
No	Reference	Reference	Reference	Reference
Yes	2.45 (2.22-2.70)[e]	1.62 (1.48-1.76)[e]	2.36 (2.13-2.62)[e]	2.23 (2.03-2.44)[e]

[a]SPPB: Short Physical Performance Battery.

[b]FTSST: Five-Times-Sit-to-Stand-Test.

[c]Crude model: unadjusted covariates.

[d]MMSE: Mini-Mental State Examination.

[e]*P*<.001.

[f]Model 1: adjusted for sociodemographic characteristics adjusted for sociodemographic characteristics (age, gender, BMI, education level, marital status, number of children, residential category, living arrangement, geographic region, data source, monthly household income).

[g]Model 2: adjusted for Model 1 variables plus health and behavioral factors(chronic diseases, types of long-term medication, alcohol consumption, smoking, frequency of weekly social activities, frequency of weekly exercise, self-reported social support, and self-reported vision and hearing impairments).

Linear regression analyses corroborated these findings. After full adjustment, each MMSE point increase predicted higher physical performance scores (SPPB: β=0.18, 95% CI 0.17 - 0.19; balance: β=0.05, 95% CI 0.05 - 0.06; gait speed: β=0.04, 95% CI 0.04 - 0.05; FTSST: β=0.08, 95% CI 0.07 - 0.08; all *Ps*<.001). Cognitive impairment persisted as a robust predictor of physical decline (SPPB: β=−1.35, 95% CI −1.46 to −1.24, *P*<.001) (Table S1 in Multimedia Appendix 1).

Association of Physical Performance on Cognitive Function

After adjusting for all covariates, the association of physical performance on cognitive function was further examined, revealing a significant impact on cognitive health (Table 3). Better physical performance, particularly better gait speed scores (OR 0.82, 95% CI 0.79 - 0.85), was associated with a reduced risk of cognitive impairment. The SPPB impairments showed a marginal association with cognitive status (OR 1.16, 95% CI 1 - 1.34; *P*=.051), while balance impairments (OR 1.71, 95% CI 1.57 - 1.88), gait speed limitations (OR 1.89, 95% CI 1.66 - 2.16), and muscle weakness (OR 1.66, 95% CI 1.47 - 1.86) were significantly associated with cognitive impairment (*P*<.001).

XSL•FO
RenderX

Table . Logistic regression analysis of physical performance on cognitive function.

Models	Cognitive impairment, OR (95%CI)
Crude model[a]	
Balance	0.84 (0.81 - 0.87)[b]
Gait speed	0.83 (0.80 - 0.85)[b]
FTSST[c]	0.88 (0.86 - 0.91)[b]
SPPB[d]	0.85 (0.84 - 0.86)[b]
Balance group	1.84 (1.69 - 2.00)[b]
Gait speed group	1.88 (1.66 - 2.14)[b]
FTSST group	1.86 (1.66 - 2.07)[b]
SPPB group	1.18 (1.03 - 1.36)[e]
Model 1[f]	
Balance	0.88 (0.84 - 0.91)[b]
Gait speed	0.82 (0.79 - 0.90)[b]
FTSST	0.86 (0.831 - 0.89)[b]
SPPB	0.85 (0.84 - 0.86)[b]
Balance group	1.74 (1.59 - 1.90)[b]
Gait speed group	1.89 (1.66 - 2.15)[b]
FTSST group	1.69 (1.50 - 1.90)[b]
SPPB group	1.14 (0.98 - 1.31)
Model 2[g]	
Balance	0.88 (0.85 - 0.92)[e]
Gait speed	0.82 (0.79 - 0.85)[b]
FTSST	0.86 (0.83 - 0.89)[b]
Balance group	1.71 (1.57 - 1.88)[b]
Gait speed group	1.89 (1.66 - 2.16)[b]
FTSST group	1.66 (1.47 - 1.86)[b]
SPPB group	1.16 (1.00 - 1.34)

[a]Crude model: unadjusted covariates.

[b]*P*<.001.

[c]FTSST: Five-Times-Sit-to-Stand-Test.

[d]SPPB: Short Physical Performance Battery.

[e]*P*<.05.

[f]Model 1: adjusted for sociodemographic characteristics adjusted for sociodemographic characteristics (age, gender, BMI, education level, marital status, number of children, residential category, living arrangement, geographic region, data source, monthly household income).

[g]Model 2: adjusted for Model 1 variables plus health and behavioral factors (chronic diseases, types of long-term medication, alcohol consumption, smoking, frequency of weekly social activities, frequency of weekly exercise, self-reported social support, and self-reported vision and hearing impairments).

Linear regression analysis confirmed our findings. In the fully adjusted model, higher scores in the SPPB, balance, gait speed, and muscle strength were associated with higher MMSE scores, while individuals with impaired physical performance had lower MMSE scores (*P*<.001), with balance ability showing the strongest association (β=1.01, 95% CI 0.96 - 1.07). Those with balance impairments had an average MMSE score 2.02 points lower than those with normal balance (95% CI −2.15 to −1.89; Table S2 in Multimedia Appendix 1). These findings highlight

XSL•FO
RenderX

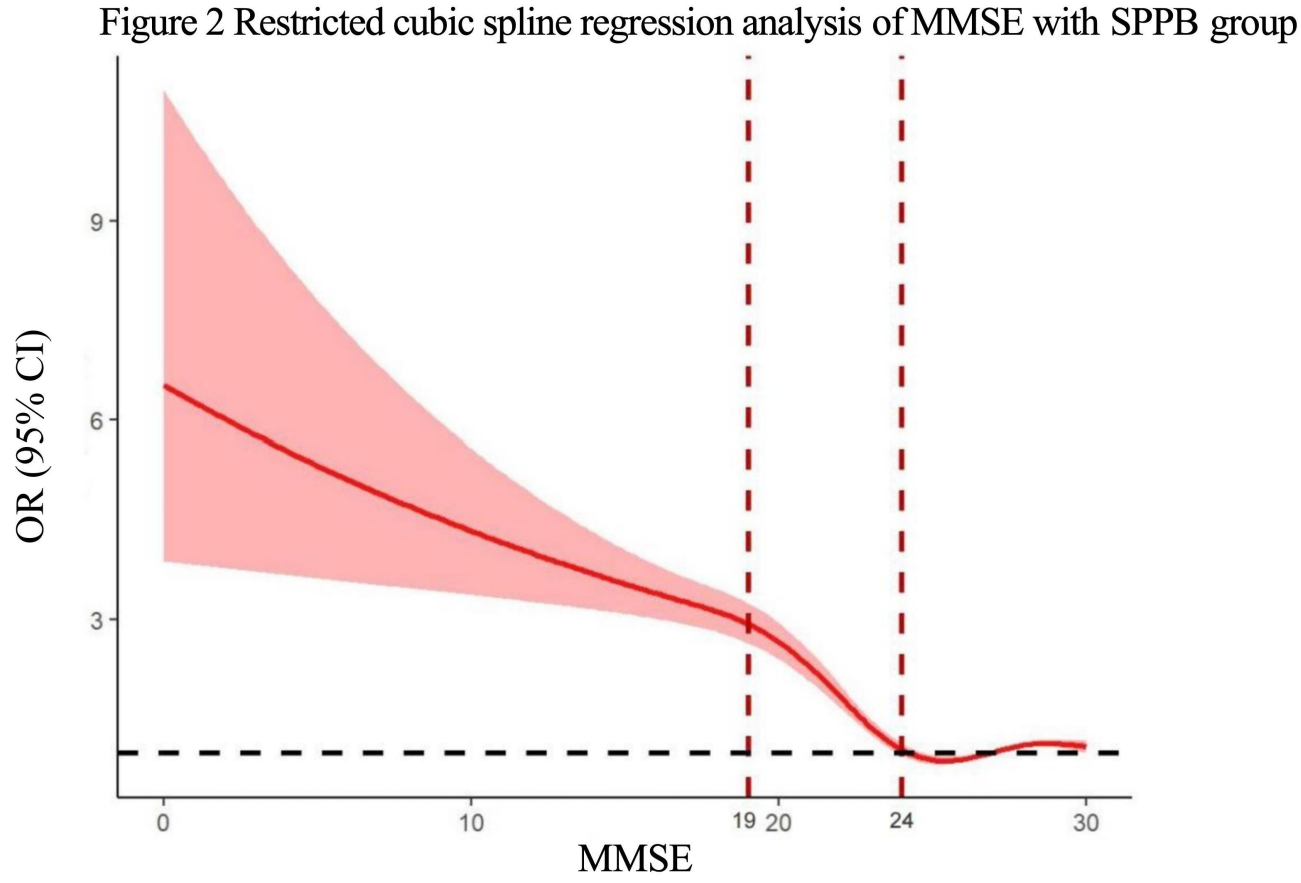
gait speed and balance as key biomarkers for cognitive risk stratification.

This study shows a significant bidirectional and asymmetric relationship between cognitive function and physical performance. Cognitive impairment strongly affects physical performance, increasing the risk of SPPB (OR=2.45, 95% CI 2.22 - 2.70), gait speed (OR=2.36, 95% CI 2.13 - 2.62), muscle strength deficits (OR=2.32, 95% CI 2.03 - 2.44), and balance impairments (OR=1.16, 95% CI 1 - 1.34). In contrast, physical performance has a weaker association with cognitive function, with balance ability (OR 1.71, 95% CI 1.57 - 1.88) being the strongest, and gait speed (OR 1.89, 95% CI 1.66 - 2.16), muscle strength (OR 1.66, 95% CI 1.47 - 1.86), and SPPB showing weaker, marginal associations (*P*=.051). This asymmetry suggests that maintaining cognitive health can better support physical function in older adults, offering dual benefits for both.

RCS and Threshold Effect Analysis of Cognitive Function on Physical Performance

RCS analyses showed significant nonlinear associations between cognitive function (MMSE) and physical performance measures (SPPB, gait speed, balance, and muscle strength; Figure 2 ; Figures S1-S3 in Multimedia Appendix 1). These relationships remained consistent in gender and education-stratified analyses (Figures 3 and 4; Figures S1A-S3B in Multimedia Appendix 1), confirming robustness across demographic subgroups.

Figure 2. Restricted cubic spline regression analysis of Mini-Mental State Examination with Short Physical Performance Battery group. MMSE: Mini-Mental State Examination; SPPB: Short Physical Performance Battery.



XSL•FO
RenderX

Figure 3. Gender stratification. MMSE: Mini-Mental State Examination.

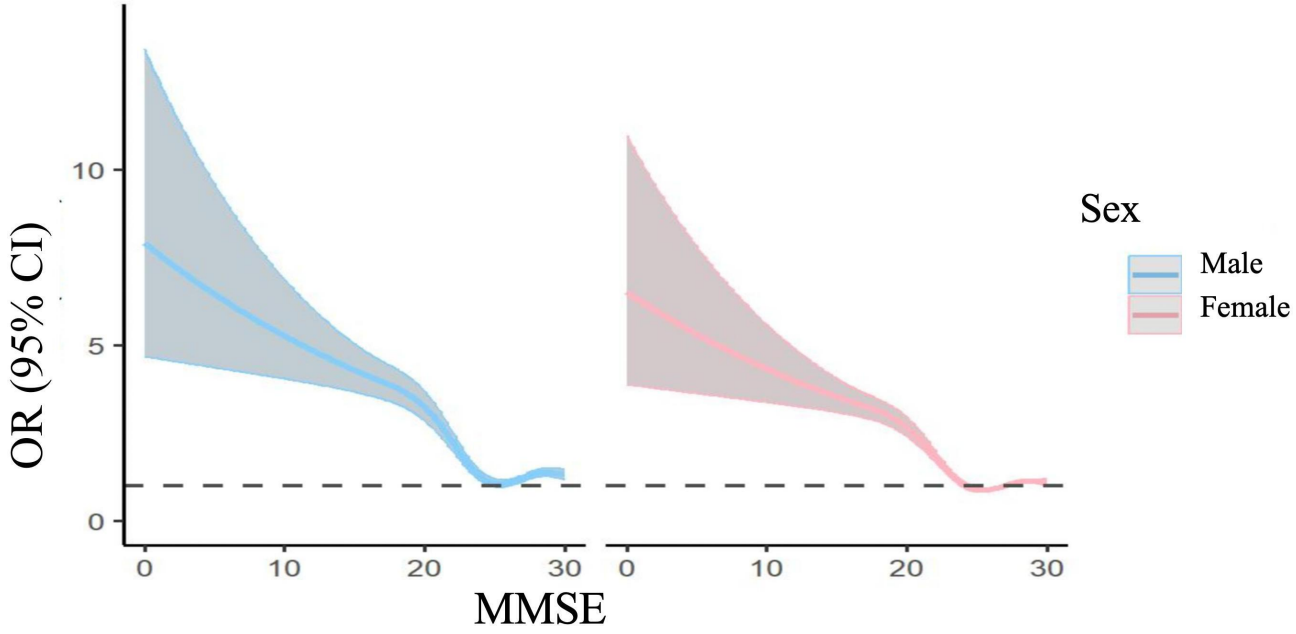
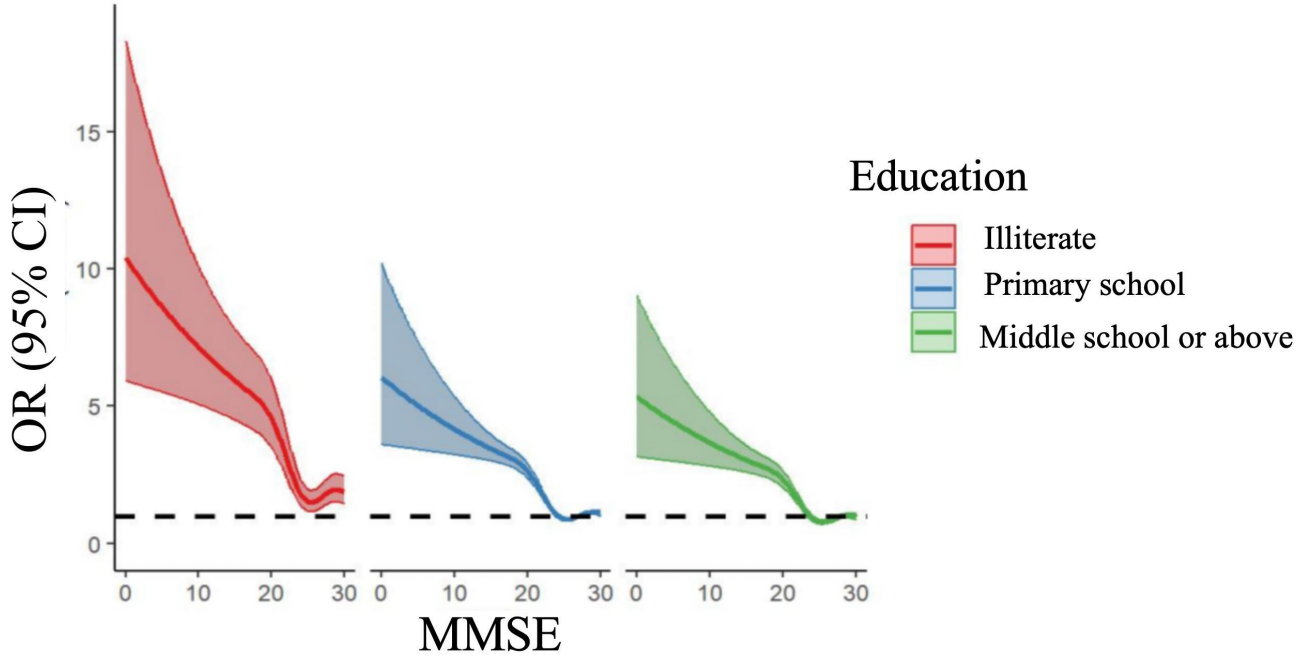


Figure 4. Education level stratification. MMSE: Mini-Mental State Examination.



This study revealed threshold effects and dynamic changes between cognitive function and physical performance using RCS analyses. Significant inflection points were found at MMSE=24 for overall performance (SPPB), balance, and gait speed, and at MMSE=29 for FTSST (Table 4). In the impaired cognitive period (MMSE <24), cognition improvements protected physical performance, reducing the risk of SPPB by 14% (OR 0.86, 95 CI 0.84 - 0.87), gait speed by 16% (OR 0.84, 95% CI 0.82 - 0.85), and balance by 4.8% (OR 1.05, 95% CI 1.03 - 1.06). In the optimal cognitive period (MMSE ≥24), the protective association remained significant for balance (risk reduced by 31%, OR 0.69, 95% CI 0.68 - 0.70) but weakened for overall performance (SPPB,*P*=.054), and a marginal positive association was observed for gait speed (OR 1.05, 95% CI 1.04-1.07), likely because cognitively intact individuals were already near the upper limit of normal performance, rather than reflecting a harmful relationship. The protective association on muscle strength persisted, with the most significant impact at MMSE >29 (risk reduced by 32%; OR 0.68, 95% CI 0.63 - 0.74).

XSL•FO
RenderX

Table . Threshold effect analysis of cognitive function, MMSE[a] on physical function.

Groups[b]	OR (95% CI)	*P* value
SPPB[c] group		
Fitting by the standard linear model	0.92 (0.91 - 0.93)	<.001
Fitting by the two-piecewise linear model	__[d]	—
Inflection point (score=24)		
MMSE[a]<24	0.86 (0.84-0.87)	<.001
MMSE≥24	1.00 (0.98-1.01)	.54
P for likelihood ratio	<.001	—
Balance group		
Fitting by the standard linear model	0.90 (0.89-0.90)	<.001
Fitting by the two-piecewise linear model	—	—
Inflection point (score=24)		
MMSE<24	1.05 (1.03-1.06)	<.001
MMSE≥24	0.69 (0.68-0.70)	<.001
P for likelihood ratio	<.001	—
Gait speed group		
Fitting by the standard linear model	0.94 (0.94-0.95)	<.001
Fitting by the two-piecewise linear model	—	—
Inflection point (score=24)		
MMSE<24	0.84 (0.82-0.85)	<.001
MMSE≥24	1.05 (1.04-1.07)	<.001
P for likelihood ratio	<.001	—
FTSST[e] group		
Fitting by the standard linear model	0.90 (0.90-0.91)	<.001
Fitting by the two-piecewise linear model	—	—
Inflection point (score=29)		
MMSE<29	0.92 (0.91-0.93)	<.001
MMSE≥29	0.68 (0.63-0.74)	<.001
P for likelihood ratio	<.001	—

[a]MMSE: Mini-Mental State Examination.

[b]Adjusted for age, gender, body mass index (BMI), education level, marital status, number of children, residential category, living arrangement, geographic region, data source, monthly household income, chronic diseases, types of long-term medication, alcohol consumption, smoking, frequency of weekly social activities, frequency of weekly exercise, self-reported social support, and self-reported vision and hearing impairments.

[c]SPPB: Short Physical Performance Battery.

[d]Not applicable.

[e]FTSST: Five-Times-Sit-to-Stand-Test.

In addition, we identified a second inflection point in over physical performance (SPPB) at MMSE=19 using second derivative analysis (Figure S4 in Multimedia Appendix 1). After adjusting for confounding factors, the MMSE 19 - 24 group showed a 34% lower risk of physical performance impairment (SPPB) compared to the MMSE <19 group (OR 0.66, 95% CI 0.57 - 0.77) (Table S3 in Multimedia Appendix 1).

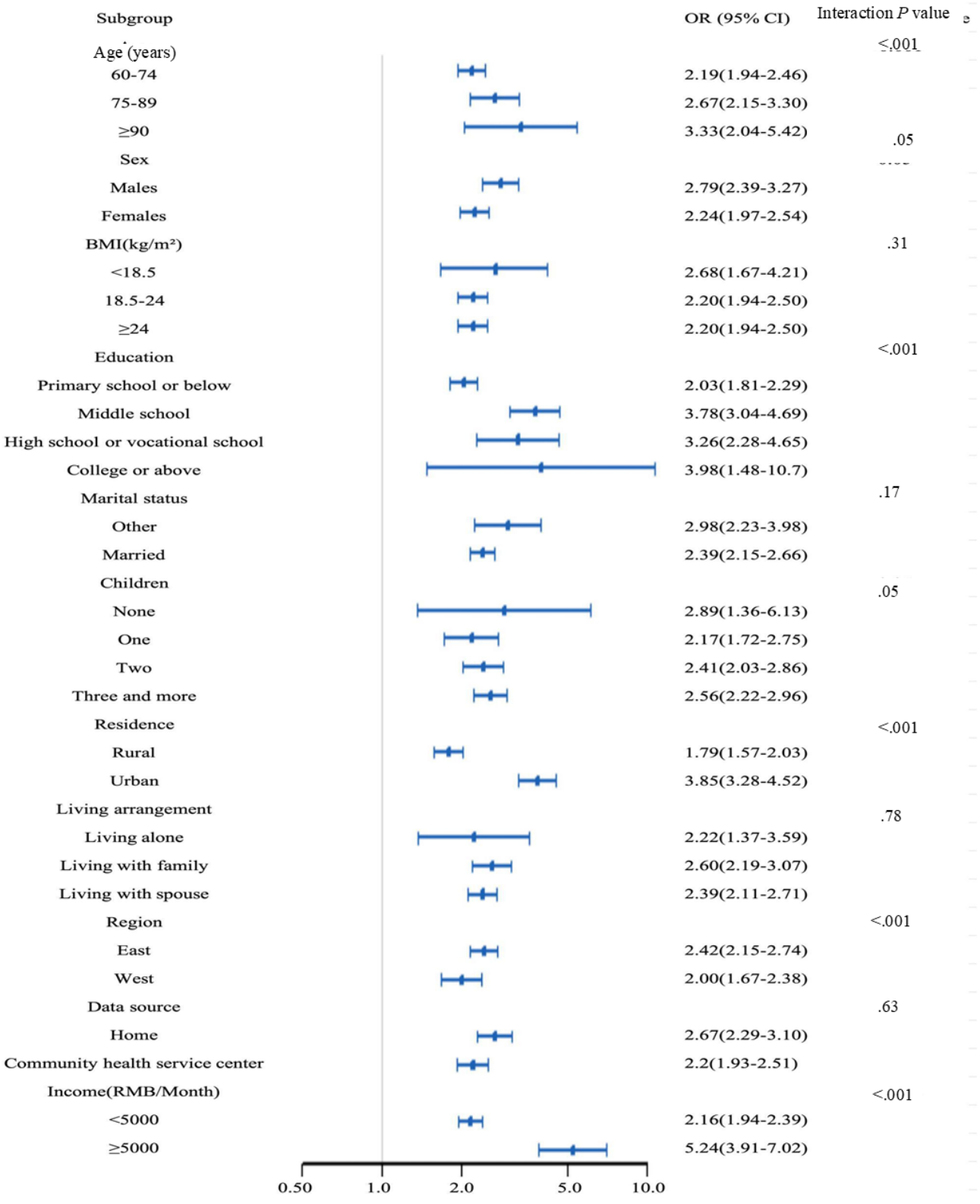
Subgroup Analysis and Interaction Analysis

Subgroup analyses revealed significant heterogeneity in the association between cognitive impairment and physical performance across demographic and socioeconomic strata (Figure 5). The strongest associations were observed in older adults aged ≥90 years (OR 3.33, 95% CI 2.04 - 5.42), those with college or above level education (OR 3.98, 95% CI 1.48 - 10.70), high-income households (≥5000 yuan per month [US $1 = ¥7.113]: OR 5.24, 95% CI 3.91 - 7.02), and urban

XSL•FO
RenderX

residents (OR 3.85, 95% CI 3.28 - 4.52). Interaction models further demonstrated graded effects: the detrimental impact of cognitive impairment increased with advancing age (β=0.45 - 0.71 per age group), higher education (β=0.55 - 0.77), and socioeconomic advantage (β=0.64 - 0.85 for urban residency and high income), suggesting these factors may amplify vulnerability to physical decline (Table S4 in Multimedia Appendix 1).

Figure 5. Subgroup and interaction analyses between cognitive function and physical performance across various subgroups.



XSL•FO
RenderX

Discussion

Principal Findings

This study advances the understanding of the cognitive-physical performance relationship through 3 key innovations. First, we demonstrated a bidirectional yet asymmetric relationship: cognitive impairment exerted a stronger detrimental association on physical function than vice versa. Second, our nonlinear threshold analysis revealed critical inflection points (MMSE=19, 24, 29), defining distinct intervention phases. The MMSE 19‐24 range marked a critical intervention window where cognitive improvements maximally protected physical function, whereas MMSE >24 signaled a ceiling effect for global physical performance and a paradoxical reversal in gait speed, aligning with the cautious gait hypothesis. Finally, subgroup analyses uncovered a socioeconomic paradox: higher socioeconomic status (SES), indicated by advanced education, income, and living in an urban area, exhibited heightened vulnerability to physical functional decline. These innovations combine mechanistic research with clinical applications, offering strategies to address the disability burden in an aging society.

Our findings corroborate and extend the evidence on cognitive-physical interdependencies, while challenging the prevailing assumption of symmetrical reciprocity. Longitudinal studies have shown cognitive function as a predictor of physical decline, particularly fall risk [27] and gait deterioration [28]. A cross-sectional study [29] found that increased cognitive impairment severity leads to worsened balance control, consistent with our observation of bidirectional associations. Notably, while Li et al [28] reported a bidirectional relationship between cognitive function and gait speed, our analysis revealed a critical asymmetry: cognitive impairment exerted a disproportionately stronger impact on physical disability than the reverse. This divergence challenges the notion of equivalent bidirectional association and positions cognitive preservation as a pivotal strategy to disrupt the disability cascade.

This study found a nonlinear relationship between cognitive function and physical performance across individuals of various genders and educational levels. It identified 3 protective thresholds at scores of 19, 24, and 29, offering key evidence for developing precise intervention strategies to maintain functionality in older individuals. At the stage where cognitive function may be impaired (MMSE <24), it significantly protects all physical performance dimensions, with the greatest association observed in the 19‐24 score range, suggesting it should be a target for clinical intervention. These results align with previous research [30], which found a positive correlation between cognitive and physical function, suggesting cognitive intervention can delay physical decline. Our study further quantifies cognitive function threshold points. Thus, it is recommended to assess functional levels at this stage and conduct adaptive dual-task training targeting both cognitive and physical functions to enhance both [31].

In the optimal cognitive function stage (MMSE ≥24), physical performance exhibits dimension-specific changes: overall physical performance (SPPB) shows a clear “ceiling effect,” suggesting that further improvements in physical performance may have reached physiological limits, leaving limited room for cognitive function to contribute. Gait speed, however, exhibits an unusual reversal phenomenon, consistent with George et al [32] inverted U-shaped gait-cognition curve, attributed to the “cautious gait hypothesis.” This hypothesis suggests that older adults with higher cognitive function consciously regulate their movements, resulting in a more cautious and controlled walking pattern [33]. Cognitive function exerts a continuous protective association on balance ability, most pronounced when MMSE >29. This finding supports the view of Kuan et al [34] that early-stage mild cognitive impairment, regardless of musculoskeletal function, can impact balance, potentially due to the weakening of the cortical-vestibular network in individuals with cognitive impairment, particularly in key brain areas involved in integrating visual, auditory, and vestibular signals, correlating with the decline in balance ability [35].

Interestingly, our subgroup and interaction analyses revealed a socioeconomic paradox:

older individuals with higher SES, as measured by education and income, exhibited amplified risks of concurrent cognitive-physical decline, contradicting the conventional SES-health gradient. This counterintuitive association necessitates a multidimensional interpretation rooted in the Cognitive Reserve Hypothesis. While enhanced neurocompensatory mechanisms in high-SES populations may initially delay clinical manifestations of cognitive impairment [36], prolonged compensatory demands could deplete neural resources, precipitating accelerated physical deterioration once pathological thresholds are breached—a” double-edged sword” effect [37]. Moreover, behavior patterns related to occupational characteristics warrant attention. This group may have maintained an imbalanced “high cognitive load-low physical activity” state throughout their careers, leading to accelerated degeneration of motor function neural circuits [38]. These insights advocate for dual-target interventions prioritizing simultaneous cognitive reinforcement and physical activation in high-SES older adults, particularly during the identified critical window (MMSE 19‐24), to counteract reserve depletion trajectories.

Of course, our study has several limitations. First, as this study uses a cross-sectional design, future research should adopt longitudinal approaches to explore causal relationships and developmental trajectories. Second, due to the feasibility constraints of large-scale studies, many variables were assessed via self-report measures, potentially introducing reporting bias. Third, we did not directly measure occupational type or physical activity levels using validated instruments. Future studies are recommended to further investigate the “high cognitive load–low physical activity” hypothesis. Finally, this study evaluated overall cognitive function but did not assess specific cognitive domains, such as memory or executive function. Future research should investigate these domains to better understand the relationship between cognitive function and physical performance.

XSL•FO
RenderX

Conclusions

This study demonstrates a bidirectional, asymmetric relationship between cognitive function and physical performance, identifies a nonlinear threshold effect, and highlights substantial population heterogeneity. These findings offer novel insights into the interaction between cognitive decline and physical disability, particularly in older individuals, and provide guidance for developing targeted interventions for high-risk subgroups in an aging society.

Acknowledgments

The authors would like to thank all participants and staff involved in the data collection and management of this study.

Funding

This study was supported by the following funding sources: National High Level Hospital Clinical Research Funding (BJ-2024-199; BJ-2024-267); National Natural Science Foundation of China (72304046; 72474032), and National Key R&D Program of China (2020YFC2008504).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy and ethical restrictions protecting participant confidentiality, but may be available from the corresponding authors upon reasonable request and subject to institutional approvals.

Authors' Contributions

HH conceived and designed the study, performed the statistical analysis, and drafted the article. LX assisted in the study design, contributed to data collection, and revised the article critically. YH conducted statistical analysis and interpreted the data. YZ and HL provided feedback and assisted in result interpretation. CS and KY supervised the study and made major contributions to article revision. All authors read and approved the final article.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Multiple linear and restricted cubic spline regression analyses of cognitive function and physical performance, including inflection points, interaction effects, gender- and education-stratified analyses, and a second derivative analysis of the spline curve.
[PDF File, 526 KB - publichealth_v11i1e80575_app1.pdf]

References

1. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-03-10]
2. Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-03-10]
3. Disability. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/disability-and-health [accessed 2025-03-10]
4. Means KM, O'Sullivan PS. Modifying a functional obstacle course to test balance and mobility in the community. J Rehabil Res Dev 2000;37(5):621-632. [Medline: 11322160]
5. Freiberger E, de Vreede P, Schoene D, et al. Performance-based physical function in older community-dwelling persons: a systematic review of instruments. Age Ageing 2012 Nov;41(6):712-721. [doi: 10.1093/ageing/afs099] [Medline: 22885845]
6. Coelho-Junior HJ, Uchida MC, Gonçalves IO, et al. Age- and gender-related changes in physical function in community-dwelling Brazilian adults aged 50 to 102 years. J Geriatr Phys Ther 2021;44(2):E123-E131. [doi: 10.1519/JPT.0000000000000246] [Medline: 31693536]
7. Seino S, Shinkai S, Fujiwara Y, et al. Reference values and age and sex differences in physical performance measures for community-dwelling older Japanese: a pooled analysis of six cohort studies. PLoS ONE 2014;9(6):e99487. [doi: 10.1371/journal.pone.0099487] [Medline: 24923425]
8. de Fátima Ribeiro Silva C, Ohara DG, Matos AP, Pinto A, Pegorari MS. Short physical performance battery as a measure of physical performance and mortality predictor in older adults: a comprehensive literature review. Int J Environ Res Public Health 2021 Oct 10;18(20):10612. [doi: 10.3390/ijerph182010612] [Medline: 34682359]
9. Fujita K, Nakashima H, Kako M, et al. Short physical performance battery discriminates clinical outcomes in hospitalized patients aged 75 years and over. Arch Gerontol Geriatr 2020;90(104155). [doi: 10.1016/j.archger.2020.104155] [Medline: 32585555]

XSL•FO
RenderX

10. Lauretani F, Ticinesi A, Gionti L, et al. Short-Physical Performance Battery (SPPB) score is associated with falls in older outpatients. Aging Clin Exp Res 2019 Oct;31(10):1435-1442. [doi: 10.1007/s40520-018-1082-y] [Medline: 30515724]
11. Ramírez-Vélez R, López Sáez de Asteasu M, Morley JE, Cano-Gutierrez CA, Izquierdo M. Performance of the short physical performance battery in identifying the frailty phenotype and predicting geriatric syndromes in community-dwelling elderly. J Nutr Health Aging 2021;25(2):209-217. [doi: 10.1007/s12603-020-1484-3] [Medline: 33491036]
12. Tolea MI, Rosenfeld A, Roy SV, Besser LM, O'Shea DM, Galvin JE. Gait, balance, and physical performance as markers of early Alzheimer's disease and related dementia risk. J Alzheimers Dis 2025 Jan 26. [doi: 10.1177/13872877241313144] [Medline: 39865686]
13. Meunier CC, Smit E, Fitzpatrick AL, Odden MC. Balance and cognitive decline in older adults in the cardiovascular health study. Age Ageing 2021 Jun 28;50(4):1342-1348. [doi: 10.1093/ageing/afab038] [Medline: 33693525]
14. Stephan Y, Sutin AR, Luchetti M, Aschwanden D, Karakose S, Terracciano A. Balance, strength, and risk of dementia: findings from the health and retirement study and the english longitudinal study of ageing. J Gerontol A Biol Sci Med Sci 2024 Aug 1;79(8):glae165. [doi: 10.1093/gerona/glae165] [Medline: 38918945]
15. Tolea MI, Morris JC, Galvin JE. Longitudinal associations between physical and cognitive performance among community-dwelling older adults. PLoS ONE 2015;10(4):e0122878. [doi: 10.1371/journal.pone.0122878] [Medline: 25875165]
16. Yang S, Li J, Fu P, Sun Y, Liu Y, Zhou C. Bidirectional associations of grip strength-gait speed with mild cognitive impairment and specific cognitive abilities among older adults: a longitudinal analysis. Arch Gerontol Geriatr 2025 Apr;131:105733. [doi: 10.1016/j.archger.2024.105733] [Medline: 39742819]
17. Biasin F, Ceolin C, Celli S, et al. Interrelation between functional decline and dementia: the potential role of balance assessment. Hum Mov Sci 2023 Jun;89:103095. [doi: 10.1016/j.humov.2023.103095] [Medline: 37120906]
18. Greene BR, Kenny RA. Assessment of cognitive decline through quantitative analysis of the timed up and go test. IEEE Trans Biomed Eng 2012 Apr;59(4):988-995. [doi: 10.1109/TBME.2011.2181844] [Medline: 22207634]
19. Su D, Liu Y, Su Y, Zhang X, Chan P. Cognitive impairment is a risk factor for decreased physical performance in the elderly. Heliyon 2024 Jun 15;10(11):e32132. [doi: 10.1016/j.heliyon.2024.e32132] [Medline: 38867948]
20. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
21. Ortega-Pérez de Villar L, Martínez-Olmos FJ, Junqué-Jiménez A, et al. Test-retest reliability and minimal detectable change scores for the short physical performance battery, one-legged standing test and timed up and go test in patients undergoing hemodialysis. PLoS ONE 2018;13(8):e0201035. [doi: 10.1371/journal.pone.0201035] [Medline: 30133445]
22. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
23. Chen LK, Woo J, Assantachai P, et al. Asian working group for sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020 Mar;21(3):300-307. [doi: 10.1016/j.jamda.2019.12.012] [Medline: 32033882]
24. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991 Feb;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
25. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
26. Wang ZY, Zhang MY. "Application of the Chinese version of Mini-Mental State Examination (MMSE)" [中文版简易智能状态检查(MMSE)的应用] [Article in Chinese]. Shanghai Journal of Psychiatry 1989;7(3):108-111 [FREE Full text]
27. Amini R, Counseller Q, Taylor R, Fayyad D, Naimi R. Short physical performance battery and mediation of the effect of mild cognitive impairment on falls by community-dwelling older adults. J Neuropsychiatry Clin Neurosci 2023;35(1):59-68. [doi: 10.1176/appi.neuropsych.21050145] [Medline: 35686347]
28. Li H, Zhang J, Zou X, et al. The bidirectional association between cognitive function and gait speed in Chinese older adults: longitudinal observational study. JMIR Public Health Surveill 2023 Mar 14;9:e44274. [doi: 10.2196/44274] [Medline: 36917163]
29. Tangen GG, Engedal K, Bergland A, Moger TA, Mengshoel AM. Relationships between balance and cognition in patients with subjective cognitive impairment, mild cognitive impairment, and Alzheimer disease. Phys Ther 2014 Aug;94(8):1123-1134. [doi: 10.2522/ptj.20130298] [Medline: 24764071]
30. Handing EP, Leng XI, Kritchevsky SB, Craft S. Association between physical performance and cognitive function in older adults across multiple studies: a pooled analysis study. Innov Aging 2020;4(6):igaa050. [doi: 10.1093/geroni/igaa050] [Medline: 33241126]
31. Duckworth SC, Higginbotham CS, Pederson JA, et al. Physical and cognitive performance during upper-extremity versus full-body exercise under dual tasking conditions. Percept Mot Skills 2021 Feb;128(1):338-352. [doi: 10.1177/0031512520945088] [Medline: 32741251]

XSL•FO
RenderX

32. George KM, Gilsanz P, Peterson RL, et al. Physical performance and cognition in a diverse cohort: Kaiser Healthy Aging and Diverse Life Experiences (KHANDLE) study. Alzheimer Dis Assoc Disord 2021;35(1):23-29. [doi: 10.1097/WAD.0000000000000428] [Medline: 33629977]
33. Ellmers TJ, Kal EC, Richardson JK, Young WR. Short-latency inhibition mitigates the relationship between conscious movement processing and overly cautious gait. Age Ageing 2021 May 5;50(3):830-837. [doi: 10.1093/ageing/afaa230] [Medline: 33951155]
34. Kuan YC, Huang LK, Wang YH, et al. Balance and gait performance in older adults with early-stage cognitive impairment. Eur J Phys Rehabil Med 2021 Aug;57(4):560-567. [doi: 10.23736/S1973-9087.20.06550-8] [Medline: 33258361]
35. Xia R, Ren J, Li X, et al. Alterations in corticocortical vestibular network functional connectivity are associated with decreased balance ability in elderly individuals with mild cognitive impairment. Brain Sci 2022 Dec 29;13(1):63. [doi: 10.3390/brainsci13010063] [Medline: 36672045]
36. Tucker AM, Stern Y. Cognitive reserve in aging. Curr Alzheimer Res 2011 Jun;8(4):354-360. [doi: 10.2174/156720511795745320] [Medline: 21222591]
37. Stern Y, Albert S, Tang MX, Tsai WY. Rate of memory decline in AD is related to education and occupation: cognitive reserve? Neurology (ECronicon) 1999 Dec 10;53(9):1942-1947. [doi: 10.1212/wnl.53.9.1942] [Medline: 10599762]
38. Moreno-Llamas A, García-Mayor J, De la Cruz-Sánchez E. The socioeconomic paradox of physical activity and sedentary behavior in Europe. J Phys Act Health 2023 Mar 1;20(3):193-203. [doi: 10.1123/jpah.2022-0036] [Medline: 36669502]

Abbreviations

MMSE: Mini-Mental State Examination
OR: odds ratio
RCS: restricted cubic spline
SES: socioeconomic status
SPPB: Short Physical Performance Battery

Edited by A Mavragani, T Sanchez; submitted 13.07.25; peer-reviewed by S Cao, Y Li; accepted 13.10.25; published 13.11.25.

Please cite as:
Hu H, Xie L, Hao Y, Zhao Y, Luo H, Yu K, Sun C
Assessment of Bidirectional and Threshold-Response Associations Between Cognitive Function and Physical Performance: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e80575
URL: https://publichealth.jmir.org/2025/1/e80575
doi:10.2196/80575

© Huixiu Hu, Lanying Xie, Yuqing Hao, Yajie Zhao, Huanhuan Luo, Kang Yu, Chao Sun. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Knowledge, Perception, and Attitude Toward Voluntary Counseling and Testing for HIV Among Secondary School Students in Iringa Rural District: Descriptive Cross-Sectional Study

Maiko Charles Mkwambe[1*], MD, MMed; Shakilu Jumanne Kayungo[2*], MD, MMed, MSc; Youping Deng[1*], MD

[1]Department of Pediatrics, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuchang District, Wuhan, Hubei, China

[2]Department of Paediatrics & Child Health, University of Dodoma, Dodoma, United Republic of Tanzania

[*]all authors contributed equally

Corresponding Author:
Youping Deng, MD
Department of Pediatrics, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuchang District, Wuhan, Hubei, China

Abstract

Background: Voluntary counseling and testing (VCT) for HIV/AIDS is characterized by several key components, including pretest and posttest counseling, as well as the formulation of individualized risk reduction plans. Adolescents, including secondary school students, represent a population particularly vulnerable to HIV infection due to various biological, psychological, and social factors.

Objective: The present study aimed to assess the knowledge, perception, attitude, and barriers toward VCT for HIV/AIDS among secondary school students in Iringa Rural District, Tanzania.

Methods: A descriptive cross-sectional study was conducted in Iringa Rural District, targeting secondary school students. A random sampling technique was employed to select the participating schools. Data were collected through self-administered questionnaires, which were completed solely by students who voluntarily consented to participate in the study. Data management and analysis were carried out using Epi Info™ 7.2 software.

Results: The study involved 127 secondary school students aged 15 - 25 years from three schools in Iringa Rural District, with 69 (54.3%) female and 58 (45.7%) male participants. All students were aware of the VCT services. The primary source of VCT information was school-based education (92, 33.3%), followed by radio/TV (65, 23.6%), friends/family (46, 16.7%), and magazines (35, 12.7%). Some students also cited health centers, hospitals, and religious seminars (38, 13.8%) as sources. Knowledge of VCT increased with education level, with Form IV students showing the highest awareness (67, 55.4%). Most students understood VCT's main purpose: 88 (50.6%) of them linked it to knowing one's HIV status, 58 (33.3%) to HIV prevention, and 28 (16.1%) to preparing for test results. No misconceptions about VCT's purpose were reported. Attitudes toward HIV testing showed that stigma concerns increased with age, particularly among those aged 17 - 18 years, where 26 (60.5%) were unwilling to be identified as HIV-positive. Nonetheless, 65 (51.2%) students supported knowing their HIV status and 86 (54.4%) said they would inform others and change their behavior after testing. Key barriers to VCT uptake included fear of stigma (71, 50.6%), lack of confidentiality, insufficient trained personnel, and poor infrastructure.

Conclusions: The study found that secondary school students in Iringa Rural District had generally high awareness and positive attitudes toward VCT, mainly informed through schools and media. Awareness increased with education level, but willingness to disclose test results remained low due to stigma and confidentiality concerns. While many students were ready to take responsible actions after testing, barriers such as fear, misinformation, and limited access to services persisted. The study emphasizes the need for improved school-based VCT programs, better-trained staff, community education, and strategies to address stigma and infrastructural gaps.

(JMIR Public Health Surveill 2025;11:e66739) doi:10.2196/66739

KEYWORDS

knowledge; perception; attitudes; secondary school students; voluntary counselling and testing; HIV; AIDS; sexually transmitted diseases; questionnaire; quantitative; cross-sectional

XSL•FO
RenderX

Introduction

Since the year 2000, approximately 38.1 million individuals have contracted HIV, and about 25.3 million have lost their lives due to illnesses related to AIDS. Despite better access to antiretroviral therapy and health care in many parts of the world, the AIDS epidemic still claimed 1.2 million lives in 2014 alone, and globally, youth aged 15 - 24 years account for nearly one-third of new HIV infections [1,2]. Despite this, uptake of voluntary counseling and testing (VCT) services among youth remains low. The majority of these cases occurred in sub-Saharan Africa, which is home to 66.6% of all people living with HIV, and most HIV transmissions happen through heterosexual sex, from mother to child during childbirth or breastfeeding, and through unsafe blood transfusions [1,3]. Moreover, since there is no effective vaccine or cure, VCT for HIV has served as a key entry point, providing access to a range of prevention and care services, including antiretroviral therapy.

Sub-Saharan Africa's burden is particularly evident in countries like Tanzania, where rural regions face distinct challenges. HIV/AIDS remains a major public health concern in Tanzania, with approximately 1.6 million people living with HIV as of 2023 [4]. While the national prevalence stands at 4.7% among adults aged 15 - 49 years [5], some regions, particularly in the southern highlands, report significantly higher rates. Iringa Region, including its rural districts, has one of the highest HIV prevalence rates in the country, reported at 11.1% after Njombe region, which is the leading region with highest HIV infection of 12.3% [6]. This high prevalence underscores the need for targeted interventions, especially among vulnerable populations such as secondary school students. Rural communities in Tanzania experience unique challenges related to HIV, including limited access to health care services, stigma, and lower awareness levels [7]. These issues are compounded by traditional gender norms, low literacy rates, and poverty, which hinder HIV prevention and testing efforts [8].

VCT is a critical component in HIV prevention and control strategies. However, its uptake among secondary school students in Tanzania remains suboptimal. Studies have shown that while awareness of VCT services is relatively high, actual utilization is low due to factors such as fear of positive results, stigma, and lack of perceived risk [9,10]. In Iringa Rural District, specific data on VCT uptake among secondary school students is limited, highlighting the need for focused research in this demographic to inform targeted interventions. Adolescents and young people, particularly those in secondary schools, represent a vulnerable group in the HIV epidemic. In rural Tanzania, multiple structural, social, and cultural factors inhibit youth from accessing VCT.

A predominant barrier to VCT uptake is the fear of receiving a positive HIV result, which is often associated with social stigma. Students express concerns about being ostracized by peers and community members if they tested positive, leading to reluctance in seeking testing services. This fear is compounded by misconceptions about HIV transmission and the belief that a positive result equates to imminent death [11-13]. Students often worry about the confidentiality of their test results, fearing that health workers may disclose their status to others in the community. This concern is particularly acute in tight-knit rural communities where privacy is hard to maintain. Ensuring strict confidentiality protocols and training for health workers is crucial. In tight-knit rural communities, concerns about privacy can discourage individuals from utilizing local testing services [11].

The absence of youth-friendly VCT services tailored to adolescents' needs and concerns contributes to low uptake. Health facilities may not provide an environment that is welcoming or comfortable for young people, deterring them from seeking services. Implementing adolescent-focused approaches and involving youth in service design can enhance engagement. Many health facilities are not tailored to meet the specific needs of adolescents, making them less likely to seek testing [11]. Geographical barriers and inadequate health care infrastructure in rural areas hinder access to testing facilities [14]. In rural districts like Iringa, access to VCT services is hindered by geographical barriers, including long distances to health facilities and lack of transportation. Additionally, the operating hours of VCT centers may not align with students' schedules, further limiting accessibility. A significant number of youths underestimate their risk of HIV infection, leading to lower testing rates. Many secondary school students in rural Tanzania perceive themselves to be at low risk of HIV infection, even when engaging in high-risk behaviors. This low-risk perception diminishes the perceived need for testing, thereby reducing VCT uptake [11,15]. Similar youth-focused barriers have been reported in other low-resource settings, underscoring the global challenge of VCT uptake [16-18].

Employing diverse strategies to fight against HIV and minimize its effects on individuals, families, and society is crucial. These approaches encompass VCT, provider-initiated counseling and testing, diagnosing HIV in infants and young children, providing family care and partner testing, detecting and managing sexually transmitted infections, offering counseling on safer sex and risk reduction, promoting male circumcision, and implementing targeted interventions for sex workers and individuals who identify as homosexual [19].

While there is some data on the awareness and usage of VCT services in a few regions in Tanzania, there is a noticeable gap in research regarding the knowledge, attitudes, and perceptions of secondary school students toward VCT services. This is worrisome because the secondary school setting is conducive to engaging in risky HIV behaviors, like unprotected sex. Secondary school students are especially susceptible to these behaviors due to their inclination toward sexual experimentation, involvement with multiple partners, and inconsistent condom use [20].

To our knowledge, there is a lack of studies assessing knowledge, attitudes, and perceptions among secondary school students in Iringa Rural District. Also, the provision of VCT services in Iringa has been inconsistent, and even when these services are accessible, the number of people utilizing these, especially youths, has been relatively low. To ensure effective VCT services, it was essential to evaluate the awareness, usage, and willingness of secondary school students toward VCT

XSL•FO
RenderX

services. Consequently, this study focused on examining the knowledge, attitudes, and perceptions of secondary school students toward VCT services in the Iringa Rural District in the Iringa region.

Methods

Study Setting and Study Design

This study employed a cross-sectional descriptive study to assess the knowledge, perception, and attitude toward VCT among secondary school students in Iringa Rural District. This study was done at three secondary schools (Lipuli, Kalenga, and Kidamali) from September to December 2023. Iringa Rural District is one of the four districts of the Iringa Region of Tanzania, East Africa. It is bordered to the north by the Dodoma Region, to the east by Kilolo District and encircles Iringa Urban District, to the south by the Mufindi District, to the southwest by the Mbeya Region, and to the northwest by the Singida Region. Iringa district has a cold climate, which favors agricultural activities and livestock keeping; thus, the majority of the people are peasant farmers. Original occupants are the Hehe tribe, but due to the migration of people, the district has been occupied by people of different ethnicity.

Study Population

The target population was students from three secondary schools and comprised both males and females aged between 15 and 25 years. This age group was chosen due to their vulnerability to HIV infection and the relevance of VCT in their health-seeking behavior.

Sample Size

Sample size was calculated from the formula $N=Z^2PQ/E^2$, where N=minimum sample size; Z=standard normal deviation, usually set at 1.96, which corresponds to 95% CI; Q=(1–P); P=the prevalence of HIV/AIDS in Iringa region (9.1%) [Tanzania HIV Impact Survey (THIS 2022)–National AIDS Control Programme, Tanzania [21], and US President'S Emergency Plan for AIDS Relief (PEPFAR) Tanzania Country Operational Plan (COP 2023)–PEPFAR [22]], E=standard error of estimate, 0.05. Therefore, the sample size required was 127.

Sampling Technique

A multistage sampling technique was used. Stage 1: Three wards were randomly selected from Iringa Rural District. Stage 2: From each ward, one public secondary school was selected using simple random sampling. Stage 3: Within each school, a list of eligible students was obtained and systematic random sampling was used to select participants proportional to the school size.

Inclusion Criteria

The inclusion criteria for the study were secondary school students, aged between 15 and 25 years, willing to be interviewed and participate in the study, enrolled in Form I-IV.

Exclusion Criteria

This study excluded students who were below 15 years and above 25 years of age, who declined to participate in the study, who were absent on the day of data collection, and who had cognitive impairments that limited their ability to participate.

Data Collection

Data Collection Tools

The data were collected from secondary school students using self-administered questionnaires. The questionnaire was developed based on validated tools from similar studies and included the following sections: demographics (age, sex, grade level), knowledge on HIV and VCT, perception of risk and benefits of VCT, attitudes toward VCT services, and barriers to accessing VCT services. The questionnaire was pretested among 20 students from a school not included in the main sample to assess clarity and reliability.

Data Collection Procedure

Data collection was conducted by trained research assistants fluent in both Kiswahili and English. Prior to administration, participants were informed about the study objectives and provided with instructions, allowing them to make an informed decision about their own participation. Their questions were addressed. Participants were informed that their involvement in the study was voluntary. To ensure anonymity, participants were instructed not to provide their names on the questionnaire forms. Confidentiality was also assured and confidentiality guarantees given.

Data Entry and Analysis

Coding, entry, cleaning, and analysis of data were done using Epi Info™ 7.2 software (Centers for Disease Control and Prevention). Descriptive statistics (frequencies and percentages) were computed.

Ethical Considerations

Ethical approval of the study was obtained from the Review Board of the University of Dodoma (DJ.232/238/0 - 28) and Iringa district council (FA.255/265/01/PART 'C;/72), with a waiver of written informed consent. To maintain confidentiality, unique identification codes were assigned to the study subjects. The principal investigators securely stored and controlled access to the data, limiting it to authorized personnel only. The study was performed in accordance with the University of Dodoma ethical guidelines for research involving human subjects and the ethical standards of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Results

Characteristics of the Study Participants

A total of 127 students from three secondary schools aged 15 - 25 years were enrolled in the study. Out of the 127 students, 69 (54.3%) and 58 (45.7%) students were females and males, respectively. Participants were mostly within the adolescent age group. A majority of 72 (56.7%) students were aged 17 - 18 years, followed by 41 (32.3%) aged 15 - 16 years, and only 14 (11%) aged 19 - 20 years. This age range is particularly significant, as adolescents are a key population for HIV prevention strategies due to their increased vulnerability and evolving risk behaviors. In terms of academic level, the largest

XSL•FO
RenderX

proportion of students 70 (55.1%) were in Form IV, followed by Form III 48 (37.8%), and a small group 9 (7.1%) in Form II. This distribution indicates that most participants were senior students who are likely to have greater exposure to sexual and reproductive health education. All the students were aware of VCT services. The source of information on VCT for HIV/AIDS among students was determined. The primary source of information that involved describing whether they have heard of VCT or not was their schools, accounting for 92 (33.3%) participants. Radios or TV were the second common source, with 65 (33.3%) participants obtaining information from this medium. Friends or family were another significant source, with 46 (16.7%) participants relying on them. Magazines were the least utilized source, with only 35 (12.7%) participants obtaining information from them. On the other hand, approximately 38 (13.8%) participants reported obtaining information from health centers, hospitals, mosques, and church seminars (Table 1).

Table . Demographic characteristics of the study group.

Characteristics	n (%)
Gender	
Female	69 (54.3)
Male	58 (45.7)
Age (years)	
15-16	41 (32.3)
17-18	72 (56.7)
19-20	14 (11.0)
Grade	
II	9 (7.1)
III	48 (37.8)
IV	70 (55.1)
Student knowledge toward VCT[a] services	
Yes	127 (100)
No	-
Source of information about VCT services	
Friends/family	46 (16.7)
Taught in school	92 (33.3)
Radios/TV	65 (33.3)
Magazine	35 (12.7)
Others	38 (13.8)

[a]VCT: voluntary counseling and testing.

Knowledge Toward the Importance of VCT Services

Table 2 and Table 3 present two key elements: (1) the level of students' awareness of the importance of VCT services by education level and (2) students' understanding of specific reasons why VCT services are important.

Table . The students' knowledge about the importance of VCT[a] services (N=127). The categorical variables are represented in numbers and percentages.

Level of education	Do you know the importance of VCT services? n (%)	
	Yes	No
Form II	8 (6.6)	1 (16.7)
Form III	46 (38.0)	2 (33.3)
Form IV	67 (55.4)	3 (50.0)

[a]VCT: voluntary counseling and testing.

XSL•FO
RenderX

Table . Perceived importance of VCT[a] services (N=127). The categorical variables are represented in numbers and percentages.

Importance of VCT services	n (%)
To know HIV status	88 (50.6)
Prepares you to accept HIV test results	28 (16.1)
Important in HIV prevention	58 (33.3)
Stigmatization to infected	-
Confuse infected person	-

[a]VCT: voluntary counseling and testing.

Awareness of the Importance of VCT by Education Level

The distribution of responses by form level indicates a trend of increasing awareness with educational progression. Among Form II students, 8 out of 9 (88.9%) reported knowing the importance of VCT services, while 1 (11.1%) student did not. In Form III, 46 (95.8%) students demonstrated awareness, with only 2 (4.2%) students responding “No.” Among Form IV students, the highest level of awareness was observed, with 67 (95.7%) students, and only 3 (4.3%) students reporting no awareness (Table 2 and Table 3).

Perceived Importance of VCT Services

When asked about why VCT is important, students identified three major reasons: (1) To know HIV status (88, 50.6%); this was the most frequently cited reason. (2) Important in HIV prevention (58, 33.3%), significant proportion recognized VCT’s preventive value. (3) Prepares you to accept HIV test results (28, 16.1%); this shows that some students understood the emotional or psychological support role that VCT provides, though this was less frequently mentioned. No students associated VCT with negative outcomes such as stigmatization or confusing the infected person (Table 2 and Table 3).

Attitudes of Participants Towards VCT Services of Secondary School Students

Regarding attitudes of secondary school students toward uptake of VCT, Table 4 shows three dimensions: (1) readiness to be known for their HIV status by age, (2) general acceptability of HIV testing, and (3) willingness to disclose HIV test results.

Table . Attitudes of secondary school students toward HIV tests (N=127). The categorical variables were represented in numbers and percentages.

	Acceptance level, n (%)				
Attitudes	Strongly disagree	Disagree	Agree	Strongly agree	Not sure
Are you ready to be known for your HIV status?					
Age (years)					
15 - 16	12 (27.9)	11 (40.7)	5 (22.7)	9 (33.3)	4 (50.0)
17 - 18	26 (60.5)	12 (44.4)	15 (68.2)	15 (55.6)	4 (50.0)
19 - 20	5 (11.6)	4 (14.8)	2 (9.1)	3 (11.1)	—[a]
Acceptability of secondary school students towards HIV test	6 (4.7)	8 (6.3)	40 (31.5)	65 (51.2)	8 (6.3)
Willingness of secondary school students to disclose the HIV test results	43 (33.9)	27 (21.3)	22 (17.3)	27 (21.3)	8 (6.3)

[a]not applicable.

Readiness to Be Known for Their HIV Status By Age

Responses varied across age groups, with the 17 - to 18-year-old group showing the highest level of acceptance; among those who agreed or strongly agreed to being known for their HIV status, the 17 - to 18-year-old group made up the majority, accounting for 15 (68.2%) agreeing and 15 (55.6%) strongly agreeing. In contrast, the 15 - to 16-year-old group had relatively higher percentages in disagreement or uncertainty, where 12 (27.9%) strongly disagreed, 11 (40.7%) disagreed, and 4 (50%) were not sure. The 19 - to 20-year age group showed lower engagement across all categories (Table 4).

Acceptability of HIV Testing

Overall, the acceptability of HIV testing among students was encouraging, where a large majority either agreed (40, 31.5%) or strongly agreed (65, 51.2%) with the idea of HIV testing. Only 14 (11%) (6 strongly disagree +8 disagree) rejected the idea of testing, and 8 (6.3%) were unsure. This positive attitude toward testing reflects a generally favorable view of VCT among

XSL•FO
RenderX

the student population, showing potential for increased VCT uptake (Table 4).

Willingness to Disclose HIV Test Results

When it comes to disclosing HIV status, students were less willing; out of 127 students, a significant proportion of 43 (33.9%) strongly disagreed and 27 (21.3%) disagreed with the idea of disclosing their HIV test results. However, only 22 (17.3%) agreed and 27 (21.3%) strongly agreed to disclose their results, while 8 (6.3%) remained unsure, indicating a substantial level of concern over confidentiality, stigma, or social repercussions associated with a known HIV status—even among those who accept testing (Table 4).

Perception of Participants Toward VCT Services of Secondary School Students

Insights into students' anticipated actions after receiving HIV test results and their expected reactions if they tested HIV-positive were determined. These perceptions reflected both their understanding of the implications of testing and their emotional and behavioral readiness to cope with outcomes.

Actions to Be Taken After Testing for HIV

Out of 150 responses, 86 (54.4%) students responded that they will tell their parents or close friends, indicating a relatively positive perception of social support. Also, 50 (31.7%) students responded to changing behavior. This significant proportion of students reported that they would modify their behavior after an HIV test, which is a highly encouraging sign. While 14 (8.9%) students reported using condoms, reflecting that awareness or acceptance of specific protective practices like condom use may still be limited. Cultural taboos or misinformation may be contributing factors. However, 8 (5.1%) students responded not telling anyone (Table 5). A minority of students indicated they would remain silent about their results. This may reflect concerns about stigma, fear of discrimination, or lack of trust in confidentiality.

Table . Perception of secondary school students toward VCT[a] services for HIV testing (N=127). The categorical variables were represented in numbers and percentages.

Variable	n (%)
Action to be taken by secondary school students after having tested for HIV	
I will tell my parents/close friends	86 (54.4)
I won't tell anybody	8 (5.1)
I will change my behavior	50 (31.7)
I will start using condom	14 (8.9)
Students' reaction if tested positive for HIV	
Kill myself	1 (0.5)
Return to God	14 (7.4)
Follow doctor's order	122 (64.2)
Inform my parents/close friends	23 (12.1)
Continue with my normal life	14 (7.4)
Others	—[b]

[a]VCT: voluntary counseling and testing.

[b]not applicable.

Reactions if Tested Positive for HIV

Out of 174 students' responses, 122 (64.2%) students responded to follow doctor's orders, the overwhelmingly dominant response showing a strong level of health-seeking behavior and trust in medical professionals. Also, 23 (12.1%) students showed that they will inform their parents or close friends, showing a willingness by some to seek emotional support and maintain open communication. Returning to God was reported by 14 (7.4%) students, and 14 (7.4%) students reported continuing with normal life, indicating spiritual and psychological coping mechanisms.

Suicidal ideation was reported by 1 (0.5%) student; the fact that even one student reported this highlights the psychological distress and fear that can be associated with an HIV-positive diagnosis. It underscores the importance of counseling services and mental health support within VCT programs (Table 5).

Barriers Toward Not Using VCT Services Among Secondary School Students

Barriers among secondary school students toward utilization of VCT services were determined. Out of 140 students' responses, most students (71, 50.6%) responded that stigmatization from society was the great obstacle to having an HIV test. While 26 (18.6%) students said that they will die early, reflecting fear-based misconceptions about HIV. On the other hand, 25 (17.86%) students said that a lack of confidentiality still exists among health care workers, making it harder for them to know their HIV statuses. Thirteen (9.3%) students believed that they are too young to get HIV, reflecting low perceived susceptibility. However, 5 (3.6%) students reported perceived cost of VCT

XSL•FO
RenderX

services, reflecting that VCT services are usually free in Tanzania (Table 6).

Table . Barriers among secondary school students towards utilization of VCT[a] services (N=127). The categorical variables were represented in numbers and percentages.

Variable	n (%)
Barriers to students not to use VCT services	
Stigmatization	71 (50.6)
I'm very young; therefore, I can't get HIV	13 (9.3)
I will die early	26 (18.6)
VCT services are expensive	5 (3.6)
Lack of confidentiality among health workers	25 (17.9)
Reason for lack of coverage of VCT services for secondary school students	
Poor infrastructure	17 (21.1)
Lack of personnel for VCT services	46 (56.8)
I don't want this service	2 (2.5)
It brings conflicts among people	9 (11.1)
I don't believe in this service	2 (2.5)
Others	5 (6.2)

[a]VCT: voluntary counseling and testing.

Reasons for Lack of VCT Service Coverage for Students

Reasons for lack of coverage for VCT services for secondary school students were determined. Out of 81 total responses of students, 46 (56.8%) students reported lack of personnel response, revealing a critical shortage of trained staff to provide VCT in schools or nearby health centers. Poor infrastructure was reported by 17 (21.1%) students, reflecting a broader health systems challenge. Conflicts among people were reported by 9 (11.1%) students, and mistrust and disinterest were reported by 2 (2.5% each) students, suggesting low health literacy or resistance to external influence (Table 6).

Discussion

The Main Findings

Accurately measuring VCT service uptake among young adults is imperative and essential, not only due to their heightened vulnerability to HIV infection but also because this population frequently encounters significant structural and socioeconomic impediments to accessing these critical services. The findings reveal a critical paradox; although student awareness of VCT services was widespread and views on its importance were predominantly positive, alongside strong endorsement for the acceptability of HIV testing itself, significant barriers persist that account for the low utilization rates observed in Iringa Rural District. Among these barriers are a pronounced reluctance to disclose HIV status, rooted primarily in fear of societal stigmatization, concerns about confidentiality breaches, and prevalent misconceptions regarding HIV/AIDS and VCT outcomes. Furthermore, systemic challenges, notably severe shortages of trained personnel and inadequate infrastructure, directly contribute to the inconsistent VCT service provision in Iringa Rural District. These findings collectively demonstrate that high awareness is necessary but insufficient for effective VCT uptake.

Participant Characteristics and Implications for VCT Uptake

This study enrolled 127 secondary school students aged 15 - 25 years from three secondary schools, with a relatively balanced gender distribution, consistent with school-based demographic trends in rural Tanzania. The majority were aged 17 - 18 years, suggesting that mid-adolescence is the key period for interventions promoting VCT services. This age group typically experiences increased autonomy, sexual debut, and curiosity—factors that elevate HIV risk and underline the importance of early preventive strategies. Similar demographic trends were found in a study conducted by Mkumbo and Ingham in Mwanza [23], which emphasized that students aged 16 - 18 years were more receptive to sexual and reproductive health interventions due to greater maturity and exposure to sexual reproductive health topics in school curricula. This supports the current finding that students in higher academic levels (Form III and IV) showed greater awareness and more positive attitudes toward VCT, emphasizing the educational system's critical role in shaping health behaviors.

Knowledge of VCT and Information Sources

All students reported awareness of VCT services, and schools were identified as the main source of this information, followed by radio/TV and friends/family. This aligns with findings by Anastasia et al in Uganda, where school-based interventions were the leading source of HIV-related knowledge among adolescents [24]. The prominence of mass media as a secondary source supports the notion that integrating media campaigns with school-based programs can reinforce HIV prevention

XSL•FO
RenderX

messages. However, the low reliance on interpersonal sources like family or health centers suggests a missed opportunity for community-based health promotion. This is consistent with Mahmoud et al who reported that youth often perceive HIV as a "school topic" rather than a family concern, emphasizing the need to involve parents and local health care workers in sensitization campaigns to bridge this communication gap [25]. Our study findings were different from the study conducted by Shedura et al, who found that most of the students did not know about VCT services for HIV/AIDS. Moreover, 157 (68.0%) students have never heard of VCT, while only 74 (32.0%) students have heard of VCT, but higher findings from mass media by the study conducted by Sisay et al found that the most frequently mentioned sources of information for HIV/AIDS were mass media 211 (62.2%) followed by health professionals 177 (52.2%) [26,27]. This signifies that our study found a very low rate of VCT services information from mass media by secondary school students. Our study showed that there is a need to develop a policy that will be beneficial to have VCT training programs and seminars in all mass media in Tanzania on VCT education, so as to facilitate the awareness of VCT services for HIV/AIDS among youths.

Awareness of the importance of VCT improved with educational level, with Form IV students demonstrating the highest levels of awareness. This echoes findings from Mboya et al in Southern Highlands, Tanzania, where older students and those in upper forms exhibited significantly higher knowledge and appreciation of VCT's role in HIV prevention and care [28]. Students primarily cited knowing one's HIV status and preventing HIV as key reasons for seeking VCT, indicating a sound understanding of VCT's core objectives. Nevertheless, the relatively low mention of emotional preparedness reflects a gap in psychological aspects of VCT counseling. Comprehensive VCT services should therefore include components that prepare youth emotionally, not just medically, for their HIV test results. Similarly, a study by Mshana et al in rural Tanzania found that upper-secondary students (Forms III–IV) had significantly higher VCT knowledge than Form I students, attributing this to cumulative exposure to HIV education in school curricula [10]. In Kenya, Maitino et al observed that Form IV students were twice as likely to accurately describe VCT benefits compared to Form I peers, linking this to peer-led interventions in higher grades [29]. A Ugandan study by Nabaggala et al reported similar trends, with senior students (aged 17 - 19 years) showing 95% awareness of VCT versus 68% among juniors, due to repeated sensitization in advanced biology and life-skills classes [30]. Banda et al. in their study showed that, in terms of peer influence, senior students act as information conduits [31]. The consistency of these findings suggests that scaling up age-appropriate, curriculum-based VCT education—particularly in lower secondary levels—could bridge awareness gaps. Pairing this with peer-led initiatives (eg, Form IV ambassadors) may amplify impact.

In terms of anticipated posttest behavior, the most common actions were informing parents or close friends and changing risky behavior. These findings indicate a positive perception of VCT as a preventive and supportive service. However, the low number of students indicating condom use as a posttest action suggests persistent gaps in knowledge or cultural taboos around condom use. A study by Culyba et al in rural Morogoro found that misconceptions and sociocultural norms significantly deterred condom acceptance, even among informed youth [32]. The vast majority of students said they would follow doctors' orders if tested HIV-positive, which is encouraging. However, the fact that one student reported suicidal ideation highlights the urgent need for school-based mental health services and counseling support integrated into VCT programs. Furthermore, in our study, many respondents cited fear and anxiety as the main reasons for their reluctance to disclose their HIV test results. They worried about the emotional distress and disruption to their lives that could result from a positive test result. This reflects the reason why there is still a prevalence of HIV infection in Iringa, accounting for a second leading region after Njombe region, as most young population who are sexually active continue their inclination toward sexual experimentation, involvement with multiple partners, and inconsistent condom use while fearing to disclose their HIV statuses to their counter sexual partners. The lack of knowledge and confidence in dealing with the outcome, as well as uncertainty about posttest support and care, contributed to their concerns. These findings are consistent with a study conducted in Ethiopia, which also identified fear and anxiety as influential factors in the uptake of VCT among young people [33,34].

The majority of students expressed positive attitudes toward HIV testing, with 51.2% strongly agreeing and 31.5% agreeing to be tested. However, only 38.6% were willing to disclose their HIV test results, while 55.2% disagreed or strongly disagreed. Furthermore, the fact that even one student reported suicidal ideation highlights the psychological distress and fear that can be associated with an HIV-positive diagnosis. It reveals an important barrier to open dialog and follow-up care and highlights the need for strengthening counseling services, promoting confidential testing environments, and conducting anti-stigma campaigns within schools and communities and highlighting the need for follow-up care. It underscores the importance of counseling services and mental health support within VCT programs. This dichotomy reflects a tension between individual health behavior and social repercussions, likely fueled by ongoing stigma. Comparable concerns were found in Wandera et al in Kenya, where stigma was a major deterrent to disclosure despite high testing willingness [35]. Similarly, students who perceived high community stigma were 5 times more likely to refuse disclosure, even if they tested [31]. In Zimbabwe, Fulton et al, in their study, reported that school-based stigma reduction programs increased willingness to disclose by 22%, highlighting the need for psychosocial support in VCT services [36].

The 17 - to 18-year-old group exhibited the highest acceptance of being known for their HIV status, suggesting this group may serve as peer educators or ambassadors in youth-centered VCT campaigns. On the contrary, the 15 - to 16-year-old group showed higher levels of uncertainty and fear—highlighting the importance of age-specific sensitization. Older students (19 - 25 y) may exhibit greater reluctance due to heightened awareness of social stigma or fear of impacts on future opportunities, necessitating targeted interventions such as private counseling

XSL•FO
RenderX

sessions. The findings from this study align with a study in South Africa by Owen et al, who found that 17‐to 19-year-olds were more likely to test for HIV and disclose their status than younger adolescents, attributing this to increased maturity and sexual health awareness [37]. Research in Kenya by Maitino et al highlighted that peer educators aged 16‐18 years significantly improved VCT uptake in schools, supporting the idea that this age group can effectively influence testing behaviors [29]. Similarly, studies have reported that 15‐to 16-year-olds often exhibit higher HIV testing hesitancy due to fear of parental disclosure or lack of autonomy. Younger adolescents required school-based sensitization programs to reduce misconceptions, mirroring this study's call for age-tailored interventions [10,38]. Older youth (19‐25 y) frequently avoid testing due to anticipated stigma, particularly in settings where HIV status can affect employment or relationships [39]. Confidentiality-guaranteed services, such as mobile or after-hours clinics, have proven effective for this group [40].

Barriers Toward Utilization of VCT Services by Secondary School Students

In this study, Stigma emerged as the most reported barrier, consistent with nearly all recent adolescent HIV studies in sub-Saharan Africa. For instance, Mugisha et al in Uganda reported that perceived stigma was the leading reason adolescents avoided VCT services, despite widespread awareness [41]. In the current study, students also cited fear of dying and concerns about confidentiality, indicating that both misinformation and systemic trust issues with health care workers remain significant challenges, and the belief that they are expensive may indicate poor communication about service accessibility. These findings support the need for community-wide destigmatization campaigns and robust training for health care providers in confidentiality. Similar findings were found by Shedura et al [26], Anaba et al [9], and Adulo et al [42]. Lack of confidentiality among health care providers was also the hindering factor for secondary school students toward VCT services utilization, while others have the notion that they are too young; therefore, they cannot get HIV, which in fact contributes to the continuous prevailing of high spread of HIV infection among young population in the Iringa region.

Our study also identified a lack of trained personnel and poor infrastructure as major constraints to VCT accessibility. These are structural barriers that mirror findings from the Joint United Nations for Acquired Immune Deficiency Syndrome (UNAIDS) (2023), which stressed that youth-friendly HIV services are severely under-resourced in rural areas. Efforts to decentralize and strengthen VCT services in school environments are therefore paramount [43].

Implications of the Study Findings

School-Centered Interventions Are Effective

The high levels of awareness and acceptance of VCT among Form III and IV students affirm the value of integrating VCT education into school curricula. Education ministries should consider expanding such programs to lower forms to reach younger adolescents early.

Addressing Stigma and Confidentiality

The reluctance to disclose HIV status and the perceived breach of confidentiality point to urgent needs for stigma-reduction campaigns and the implementation of trusted, youth-friendly VCT centers with strict confidentiality policies.

Targeted Behavioral Change Communication

The limited mention of condoms as a preventive method and misconceptions about HIV's fatality indicate a need for targeted behavioral interventions that challenge myths and promote accurate knowledge, especially around prevention and healthy living.

Engaging Families and Communities

Since interpersonal sources like families and health workers play a minor role in VCT education, strategies should involve parents, religious leaders, and community health workers to normalize discussions about HIV and reduce shame.

Mental Health Integration

The instance of suicidal ideation underscores the need to include psychosocial support services within VCT frameworks. Adolescents require reassurance and coping mechanisms to handle positive diagnoses constructively.

Resource Allocation and Infrastructure

The findings highlight that for VCT programs to scale up in rural districts, investments are needed in human resources, mobile clinics, and infrastructure to bridge current service delivery gaps.

Limitations of the Study

Although this study provides valuable insights, it recognizes several limitations. This study was conducted in only three secondary schools within Iringa Rural District. The findings may not be generalizable. Future longitudinal research is needed to cover these areas in large geographical populations of multiple regions along the country, compounded with large sample size.

Conclusion

This study explored the knowledge, perceptions, and attitudes of secondary school students in Iringa Rural District regarding VCT for HIV. The findings revealed a generally high level of awareness and positive attitudes toward VCT services among students, with schools and mass media being the primary sources of information. Awareness of the importance of VCT increased with educational level, reflecting the significant role of school-based health education. Most students recognized the value of VCT in knowing one's HIV status and preventing the spread of HIV, although fewer acknowledged its psychological support functions. The overall acceptability of VCT was encouraging; however, willingness to disclose test results remained low, underscoring ongoing concerns around confidentiality and stigma. Students demonstrated a readiness to act responsibly upon receiving HIV test results, including informing trusted individuals and adopting safer behaviors. However, some also expressed reluctance to disclose results or misconceptions about HIV, pointing to lingering stigma and misinformation. Key barriers to VCT utilization included societal stigma, fear of early death, perceived lack of

XSL•FO
RenderX

confidentiality, and insufficient knowledge. Structural challenges such as a shortage of trained personnel and inadequate infrastructure further limit access to VCT services within schools. These findings provide a roadmap for designing youth-centered VCT programs that tackle stigma, enhance confidentiality, and improve rural access, significantly advancing HIV prevention in Tanzania.

Acknowledgments

We would like to express our sincere gratitude to Dr. Shakilu Jumanne Kayungo, our first clinical instructor, for his invaluable guidance, unwavering support, and dedicated mentorship throughout the development of our research proposal and the compilation of the final report. His commitment and encouragement were instrumental to the successful completion of this study.

Our heartfelt appreciation also goes to Professor Youping Deng, our second clinical instructor, whose intellectual insights and constructive feedback significantly contributed to shaping the foundation of this research. Her guidance has been vital to our academic growth and understanding.

We are deeply grateful to the District Medical Officer of Iringa Rural District and the District Education Officer, Mr. Chiza D. N., for granting us permission to carry out this study within the district. We extend special thanks to the headmasters of Lipuli Secondary School, Kalenga Secondary School, and Kidamali Secondary School for their continued cooperation and support throughout the approval process, data collection phase, and our departure from the schools.

Finally, we extend our sincere appreciation to all secondary students who participated in this study. Their willingness to contribute made this research possible, and their input is the cornerstone of our findings.

Data Availability

Data used to support the findings for this study are included in the article.

Authors' Contributions

The authors MCM, SJK, and YD had full access to all the data and take responsibility of the integrity of the data and accuracy of the data analysis.
Concept and design: MCM and SJK.
Acquisition, analysis, and interpretation of data: MCM.
Drafting of the manuscript: MCM, SJK, and YD.
Manuscript review with critical intellectual contents: All authors.
Administrative, technical and supportive supervision, and cornerstone intellectual contents: SJK and YD.

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[PDF File, 173 KB - publichealth_v11i1e66739_app1.pdf]

References

1. World AIDS day 2015, UNAIDS fact sheet 2015. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6446a1.htm] [accessed 2025-09-10]
2. Adolescents, HIV and AIDS. UNICEF. 2022. URL: https://www.unicef.org/press-releases/despite-progress-adolescent-girls-continue-bear-brunt-hiv-epidemic-98000-new [accessed 2025-09-10]
3. Tan X, Pan J, Zhou D, Wang C, Xie C. HIV/AIDS knowledge, attitudes and behaviors assessment of Chinese students: a questionnaire study. Int J Environ Res Public Health 2007 Sep;4(3):248-253. [doi: 10.3390/ijerph2007030009] [Medline: 17911665]
4. Al Awaidy S, Ghazy RM, Mahomed O. Correction: progress of the Gulf Cooperation Council (GCC) countries towards achieving the 95-95-95 UNAIDS targets: a review. J Epidemiol Glob Health 2023 Sep;13(3):589-589. [doi: 10.1007/s44197-023-00129-w] [Medline: 37347453]
5. Wang Y, Kinsler JJ, Kiwuwa-Muyingo S. Factors associated with HIV testing among youth in Tanzania based on the 2016-2017 Tanzania HIV Impact Survey (THIS). PLOS Glob Public Health 2022;2(11):e0000536. [doi: 10.1371/journal.pgph.0000536] [Medline: 36589732]

XSL•FO
RenderX

6. Kerrigan D, Mbwambo J, Likindikoki S, et al. Project Shikamana: community empowerment-based combination HIV prevention significantly impacts HIV incidence and care continuum outcomes among female sex workers in Iringa, Tanzania. J Acquir Immune Defic Syndr 2019 Oct 1;82(2):141-148. [doi: 10.1097/QAI.0000000000002123] [Medline: 31513552]
7. Sanga ES, Mukumbang FC, Mushi AK, Lerebo W, Zarowsky C. Understanding factors influencing linkage to HIV care in a rural setting, Mbeya, Tanzania: qualitative findings of a mixed methods study. BMC Public Health 2019 Apr 5;19(1):383. [doi: 10.1186/s12889-019-6691-7] [Medline: 30953503]
8. Mcharo RD, Olomi W, Mayaud P, Msuya SE. Risky sexual behaviours among young adults attending higher learning institutions in Mbeya, Tanzania: implications for STIs and HIV preventive programs. Open Res Africa 2021;3:41. [doi: 10.12688/aasopenres.13123.2]
9. Anaba EA, Buabeng Z, Okai GA. Voluntary HIV counselling and testing services: knowledge, attitudes, and correlates of utilisation among young people in the Tema Metropolis, Ghana. Youth 2022;2(4):469-478. [doi: 10.3390/youth2040034]
10. Sanga Z, Kapanda G, Msuya S, Mwangi R. Factors influencing the uptake of voluntary HIV counseling and testing among secondary school students in Arusha City, Tanzania: a cross sectional study. BMC Public Health 2015 May 2;15(452):452. [doi: 10.1186/s12889-015-1771-9] [Medline: 25933806]
11. Kidman R, Waidler J, Palermo T, Tanzania Adolescent Cash Plus Evaluation Team. Uptake of HIV testing among adolescents and associated adolescent-friendly services. BMC Health Serv Res 2020 Sep 17;20(1):881. [doi: 10.1186/s12913-020-05731-3] [Medline: 32943066]
12. Danielle A W, Kyle L J, Jacen S M. Stigma related to fear and shame restricts access to HIV testing and treatment in Tanzania. J Fam Med Dis Prev 2019;5(1) [FREE Full text] [doi: 10.23937/2469-5793/1510099]
13. Mhode M, Nyamhanga T. Experiences and impact of stigma and discrimination among people on antiretroviral therapy in Dar es Salaam: a qualitative perspective. AIDS Res Treat 2016;2016:7925052. [doi: 10.1155/2016/7925052] [Medline: 27110395]
14. Clausen A, Stephenson RB, Sullivan PS, et al. Distance as a barrier to HIV testing among sexual and gender minority populations in the rural southern US: a cross-sectional study. Rural Remote Health 2023 Nov;23(4):8227. [doi: 10.22605/RRH8227] [Medline: 37988704]
15. Sukari O. Barriers and attitudes towards HIV Voluntary Counseling and Testing among Secondary School Pupils of Sengerema in Mwanza. Dar Es Salaam Medical Students' Journal 2008;15(1):20-23. [doi: 10.4314/dmsj.v15i1.49596]
16. Jackson F, T Haile Z. Association between educational attainment and risky sexual behaviour among Ghanaian female youth. Afr H Sci 2023 Mar;23(1):301-308. [doi: 10.4314/ahs.v23i1.32]
17. AIDSinfo UNAIDS, author. AIDSinfo. URL: https://aidsinfo.unaids.org [accessed 2020-01-23]
18. AIDS deaths decline, but not HIV cases: UNAIDS report. CBS NEWS. URL: https://www.cbsnews.com/news/aids-deaths-decline-but-not-hiv-cases-unai [accessed 2020-01-23]
19. Addis Z, Yalew A, Shiferaw Y, et al. Knowledge, attitude and practice towards voluntary counseling and testing among university students in North West Ethiopia: a cross sectional study. BMC Public Health 2013 Aug 2;13(1):714. [doi: 10.1186/1471-2458-13-714] [Medline: 23914738]
20. Vulstan S, Christine P. Knowledge, attitude and perception on utilization of voluntary counselling and testing services: experiences from secondary school students in northern Tanzania. . Preprint posted online on Oct, 2022 URL: https://tinyurl.com/5259msn6 [doi: 10.14293/S2199-1006.1.SOR-.PP99AQQ.v1]
21. Tanzania HIV impact survey (THIS 2022): summary sheet – regional findings. National AIDS Control Programme (NACP) Tanzania.: Ministry of Health; 2022. URL: https://www.nacp.go.tz/ [accessed 2025-09-10]
22. Tanzania country operational plan (COP 2023): strategic direction summary. US President'S Emergency Plan for AIDS Relief (PEPFAR). 2023. URL: https://www.state.gov/pepfar/ [accessed 2025-09-10]
23. Mkumbo K, Ingham R. Sexual and reproductive health knowledge among adolescent secondary school students in Tanzania. Afr J Reprod Health 2010;14(1):84-96. [doi: 10.2307/27805862]
24. Anastasia A, Nanyonjo A, Tumwesigye NM, Kabwama SN. Adolescent knowledge and utilization of HIV services in Uganda. BMC Public Health 2022;22(1):1534. [doi: 10.1186/s12889-022-13867-9]
25. Mahmoud A, Bakr A, Elsaid A, Wahba Y. Prevalence of Helicobacter pylori infection among children with primary nephrotic syndrome: a cross-sectional study. Afr H Sci 2020;20(4):1624-1631. [doi: 10.4314/ahs.v20i4.15]
26. Shedura VJ, Paul CH. Knowledge, attitude and perception on utilization of voluntary counselling and testing services: experiences from secondary school students in northern tanzania. ScienceOpen. . [doi: 10.14293/S2199-1006.1.SOR-.PP99AQQ.v1]
27. Sisay S, Erku W, Medhin G, Woldeyohannes D. Perception of high school students on risk for acquiring HIV and utilization of voluntary counseling and testing (VCT) service for HIV in Debre-berhan Town, Ethiopia: a quantitative cross-sectional study. BMC Res Notes 2014 Aug 12;7(1):518. [doi: 10.1186/1756-0500-7-518] [Medline: 25112147]
28. Mboya MM, Leyna GH, Mmbaga EJ. School-level predictors of HIV testing among adolescents in Southern Tanzania. Int J Adolesc Youth 2019;24(2):212-224. [doi: 10.1080/02673843.2018.1491806]
29. Maitino EM, Shafir SC, Beymer MR, et al. Age at first HIV test for MSM at a community health clinic in Los Angeles. AIDS Care 2020 Feb;32(2):186-192. [doi: 10.1080/09540121.2019.1683806] [Medline: 31663365]

XSL•FO
RenderX

30. Nabaggala MS, Parkes-Ratanshi R, Kasirye R, et al. Age and education as predictors of HIV testing literacy among Ugandan adolescents. J Adolesc Health 2022;70(3):412-420. [doi: 10.1016/j.jadohealth.2021.10.008]
31. Banda JM, Mburu G, Nyamurungi KN, et al. Peer networks and HIV knowledge diffusion in Malawian schools. PLoS ONE 2019;14(6):e0218742. [doi: 10.1371/journal.pone.0218742]
32. Culyba AJ, Sigel E, Westers N, Barondeau J, Trent M, Coyne-Beasley T. Freedom to play, learn, live, and thrive: a youth-serving professional call to action to address firearm violence. J Adolesc Health 2023 Feb;72(2):177-178. [doi: 10.1016/j.jadohealth.2022.10.014] [Medline: 36351863]
33. Gatta AA, Thupayagale-Tshweneagae G. Knowledge of, and attitudes towards, voluntary HIV counselling and testing services amongst adolescent high school students in Addis Ababa, Ethiopia. Curationis 2012 Dec 6;35(1):103. [doi: 10.4102/curationis.v35i1.103] [Medline: 23327777]
34. Tsegay G, Edris M, Meseret S. Assessment of voluntary counseling and testing service utilization and associated factors among Debre Markos University Students, North West Ethiopia: a cross-sectional survey in 2011. BMC Public Health 2013 Mar 19;13(1):243. [doi: 10.1186/1471-2458-13-243] [Medline: 23509888]
35. Wandera B, Tumushabe G, Musiimenta A. Youth attitudes towards HIV testing and disclosure in rural Kenya. PLoS ONE 2021;16(4):e0249871. [doi: 10.1371/journal.pone.0249871]
36. Fulton BD, Hwong AR, Tran TD, et al. School-based interventions to reduce HIV stigma in Zimbabwe. AIDS Behav 2021;25(8):2450-2460. [doi: 10.1007/s10461-021-03207-x]
37. Owen KB, Parker PD, Astell-Burt T, Lonsdale C. Regular physical activity and educational outcomes in youth: a longitudinal study. J Adolesc Health 2018 Mar;62(3):334-340. [doi: 10.1016/j.jadohealth.2017.09.014] [Medline: 29229462]
38. Denison JA, Burke VM, Miti S, et al. HIV testing hesitancy among adolescents in sub-Saharan Africa. AIDS Behav 2019;23(5):1250-1262. [doi: 10.1007/s10461-018-02385-3]
39. Wringe A, Renju J, Seeley J, et al. HIV-related stigma among older adolescents in Eastern Africa. Soc Sci Med 2020;245:112686. [doi: 10.1016/j.socscimed.2019.112686]
40. Hlongwa M, Mashamba-Thompson TP, Makhunga S, Hlongwana K. Confidentiality concerns and HIV testing preferences among youth in South Africa. PLoS ONE 2022;17(3):e0265121. [doi: 10.1371/journal.pone.0265121]
41. Mugisha R, Baryamutuma R, Atuhaire L, Karamagi Y. Perceived stigma and its effects on HIV testing among youth in Uganda. BMC Infect Dis 2020;20:650. [doi: 10.1186/s12879-020-05358-y]
42. Adulo LA, Hassen SS, Kontuab AM. Utilization of voluntary counseling and testing experience among Mizan-Tepi University students in Southwestern Ethiopia. AIDS Res Treat 2022;2022:7911385. [doi: 10.1155/2022/7911385] [Medline: 35898517]
43. Adolescent and youth HIV services report: sub-Saharan Africa. UNAIDS. 2023. URL: https://www.unaids.org/en/resources/documents/2023/youth-hiv-services-sub-saharan-africa [accessed 2025-09-16]

Abbreviations

UNAIDS: The Joint United Nations for Acquired Immune Deficiency Syndrome
VCT: voluntary counseling and testing

Edited by A Mavragani, T Sanchez; submitted 21.09.24; peer-reviewed by OS Igunma, SS Msigwa; revised version received 04.08.25; accepted 06.08.25; published 26.09.25.

Please cite as:
Mkwambe MC, Kayungo SJ, Deng Y
Knowledge, Perception, and Attitude Toward Voluntary Counseling and Testing for HIV Among Secondary School Students in Iringa Rural District: Descriptive Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e66739
URL: https://publichealth.jmir.org/2025/1/e66739
doi:10.2196/66739

© Maiko Charles Mkwambe, Shakilu Jumanne Kayungo, Youping Deng. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Severity Classification of Anxiety and Depression Using Generalized Anxiety Disorder Scale and Patient Health Questionnaire: National Cross-Sectional Study Applying Classification and Regression Tree Models

Andre Faro[1*], PhD; Julian Tejada[1*], PhD; Wael Al-Delaimy[2*], MD, PhD

[1]Departament of Psychology, Federal University of Sergipe, Aracaju, Brazil

[2]School of Medicine, Department of Family Medicine, Herbert Wertheim School of Public Health & Human Longevity Science, University of California San Diego, 9500 Gilman Dr, La Jolla, San Diego, CA, United States

[*]all authors contributed equally

Corresponding Author:
Wael Al-Delaimy, MD, PhD
School of Medicine, Department of Family Medicine, Herbert Wertheim School of Public Health & Human Longevity Science, University of California San Diego, 9500 Gilman Dr, La Jolla, San Diego, CA, United States

Abstract

Background: Scalable and accurate screening tools are critical for public mental health strategies, especially in low- and middle-income countries (LMICs). While the Generalized Anxiety Disorder Scale (GAD-7) and Patient Health Questionnaire (PHQ-9) are widely used, their full application in large-scale programs can pose feasibility challenges. By contrast, shorter versions like GAD-2 and PHQ-2 reduce burdens but fail to capture symptom diversity.

Objective: This study aimed to optimize screening for anxiety and depression severity using classification and regression tree (CART) models, identifying concise and high-performing decision rules based on the GAD-7 and PHQ-9 items, and to test their reproducibility in 5 independent datasets.

Methods: A cross-sectional, nonprobabilistic study was conducted with 20,585 Brazilian adults from all 27 states and more than 3,000 cities, collected using digital outreach. Anxiety and depression symptoms were assessed using the GAD-7 and PHQ-9. CART models were trained and tested on bootstrapped samples (70% training, 30% testing), totaling 45,000 trees per scale. Each model used combinations of scale items and sociodemographic predictors. Robustness was evaluated via 10-fold cross-validation and evaluation across 3 hyperparameter configurations (*minsplit* and *minbucket*=500, 1000, 2000). Performance metrics included accuracy, sensitivity, specificity, precision, F1-score, and area under the curve (AUC).

Results: The CART models produced concise, high-performing decision rules—using only 2 items for the GAD-7 and 3 for the PHQ-9. No sociodemographic variable appeared in the final classification paths. For GAD-7, the models achieved an accuracy of 86.1% for minimal or mild severity and 85.1% for severe cases, with both categories showing AUC values above 0.900. By contrast, the moderate severity class had lower performance, with accuracy around 51% and an AUC of 0.728. For PHQ-9, the models achieved 81.7% accuracy for minimal or mild cases and 78.8% for severe cases, with AUCs again exceeding 0.900 for the extreme classes; the moderate or moderately severe class showed 66.9% accuracy and an AUC of 0.776. The most frequently repeated rules included the following: “GAD2<2 and GAD4<2” for identifying minimal or mild anxiety and “GAD2≥2 and GAD4=3” for severe anxiety; for depression, “PHQ2<2and PHQ4<2” for minimal or mild cases and “PHQ2≥2 and PHQ8≥2” for severe cases. These rule-based models demonstrated stable performance across thousands of bootstrapped replications and showed reproducibility in 5 independent datasets through external validation.

Conclusions: CART models enabled simplified, symptom-specific pathways for stratifying anxiety and depression severity with high precision and minimal item burden. These rule-based shortcuts offer an efficient alternative to fixed short forms (eg, GAD-2, PHQ-2) by preserving symptom diversity and severity discrimination. The findings support and lay the groundwork for adaptive, cost-effective screening and intervention models, especially in resource-limited settings and LMICs.

(JMIR Public Health Surveill 2025;11:e72591) doi:10.2196/72591

KEYWORDS

anxiety; depression; machine learning; decision tree models; public mental health

XSL•FO
RenderX

Introduction

Public mental health has become a central focus in the 21st century due to the global mental illness epidemic [1,2]. The extensive global impact of the COVID-19 pandemic has increased the urgency for public mental health [3,4]. Common mental disorders (CMDs), particularly anxiety and depression, are significant challenges for public health institutions in both high-income and low- and middle-income countries (LMICs) for the foreseeable future, with the latter bearing a disproportionately higher burden [5-9]. Even before the pandemic, high CMD prevalence underscored the need for tailored national and global strategies [1,10-12].

At national levels, comprehensive mental health strategies must include regular screening and a care model capable of addressing the full spectrum of needs—from prevention and basic psychological care to outpatient interventions and emergency support during crises [1,13]. Such an approach should be routine, not limited to critical periods; include health professional training to provide basic mental care, equipping specialists and nonspecialists to address anxiety and depression symptoms (at least regarding basic psychological care); and ensure resource availability during high-pressure periods [2,14]. A robust mental health surveillance system capable of detecting varying CMD risk levels can play a vital role in building an effective care system [6,8,15]. Using brief protocols or simple instruments to stratify risks can provide actionable insights and support targeted intervention planning [16-19].

A care model with tailored strategies based on the presence and severity of health outcomes is more feasible if it is brief; moreover, reliable tools are available for both individual assessments and widespread screenings [20-24]. The Generalized Anxiety Disorder Scale-7 (GAD-7) and Patient Health Questionnaire-9 (PHQ-9) are widely used for measuring anxiety and depression symptoms, respectively, and match several important qualities [25-27]. However, they may pose a burden on participants in large-scale mental health initiatives, especially in resource-limited settings, in which the GAD-2 and PHQ-2 are used as alternatives. Notwithstanding, these shorter versions do not differentiate mental health condition severity and revert to the first 2 questions after a binary screening diagnostic [28-30], rather than more specifically screening the severity or specific symptoms of anxiety or depression.

By focusing on only two symptoms of each disorder—nervousness and worrying in GAD-2 and anhedonia and hopelessness in PHQ-2—these scales overlook other relevant symptoms, such as fatigue, sleep disturbances, difficulty concentrating, and irritability (common in anxiety disorders) or changes in appetite, guilt, and suicidal ideation (common in depression) [31]. Consequently, they are limited in capturing important symptom profiles across diverse populations, which is crucial for tailoring interventions. This restricts a comprehensive understanding of anxiety and depression when compared to the broader symptomatology outlined in the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [31].

Machine learning (ML), a subfield of artificial intelligence (AI), identifies hidden patterns in datasets through data-driven algorithms, enabling explanatory or predictive model creation [32-35]. In the mental health field, ML models provide evidence of predicting risks and streamlining assessments [36-40]. ML algorithms can support shorter, more accurate tool development, addressing challenges such as participant fatigue and low response rates while maintaining scalability and cost-effectiveness as end points [41,42]. Furthermore, within ML algorithms, decision tree models offer a distinct approach to streamlining the path from individual items to diagnostic classification (eg, by identifying the most relevant features, thereby preserving predictive accuracy and improving efficiency). By incorporating additional variables (eg, sociodemographic data, psychological constructs, and other health outcomes), decision trees allow for adaptive and personalized analyses [43-46]. ML-based scale reduction techniques, combined with logical “shortcuts” (or rules) for case classification, offer a promising approach to balancing precision and brevity in public mental health.

This study aims to leverage ML algorithms, particularly decision tree models, to enhance the efficiency and precision of GAD-7 and PHQ-9 assessments by identifying the most predictive components across severity levels. To further examine the robustness and generalizability of the derived rules, their performance was evaluated across 5 independent datasets, including 1 collected nationwide in Brazil and 4 from international samples. We hypothesized that classification and regression tree (CART) models can identify concise, high-performing decision rules using less items from the GAD-7 and PHQ-9 scales, enabling accurate stratification of disorders into minimal/mild, moderate, and severe categories, thereby optimizing mental health screening in public health applications.

Methods

Study Design

This study adopted a cross-sectional, nonprobabilistic design and used a digital convenience sampling method. Data were collected between March and June 2024 through online recruitment via social media, primarily Instagram and Facebook. Digitally boosted posts promoted participation across all regions of Brazil. A university-based research laboratory specializing in public mental health managed the campaign through its official page. Upon accessing the study link, participants were directed to a landing page describing the study objectives and procedures. Participants completed the survey independently using their own devices in home or personal settings, without any direct researcher supervision. Only individuals aged 18 years or older could participate. The final sample included 20,585 respondents from all 27 Brazilian states and more than 3,000 cities, encompassing both rural and urban regions. The average completion time for the questionnaire was approximately 5 minutes. Incomplete responses were excluded, resulting in a withdrawal rate of approximately 8%.

Ethical Considerations

This study was approved by Brazil’s National Council for Ethics in Research (Conselho Nacional de Ética em Pesquisa [CONEP];

XSL•FO
RenderX

protocol number 30485420.6.0000.0008). All procedures complied with Brazilian and international regulations for research involving human participants. Informed consent was obtained electronically from all participants before survey participation. Participants were first presented with the study objectives, procedures, and data protection measures. Only individuals who explicitly agreed to participate were then granted access to the survey instruments. All data were collected anonymously, and no personally identifiable information was recorded. The dataset was anonymized before analysis to ensure confidentiality and data protection, and no supplementary materials contain any information that could lead to the identification of individual participants. In accordance with Brazilian research ethics policies involving human subjects, no monetary or material compensation was provided to participants. No generative AI tools were used in the manuscript writing, ideation, or drafting.

Instruments

The GAD-7 is a 7-item measure for assessing anxiety [47], with each item rated on a scale from 0 (not at all) to 3 (nearly every day) and the final score ranging from 0 to 21. It has demonstrated satisfactory psychometric properties and validity evidence in Brazil [48]. Specifically, studies using the Brazilian-Portuguese version have reported adequate internal consistency (eg, Cronbach α values typically above .80) and construct validity, supporting its use for assessing generalized anxiety symptoms in this population. The PHQ-9 is a 9-item scale for assessing depression [49] with the same Likert response format as the GAD-7 and the total score ranging from 0 to 27. It has also shown satisfactory psychometrics in the Brazilian population [50] with good internal consistency (eg, Cronbach α often exceeding .85) and strong evidence of both construct and criterion validity, making it a reliable tool for depression assessment [51]. Additionally, a sociodemographic questionnaire was administered, including items on sex/gender, age, education level, and self-reported skin color/ethnicity.

A common cutoff score of 10 is often used for GAD-7 and PHQ-9 to distinguish clinical from nonclinical cases. However, alternative stratifications based on severity levels offer a more nuanced differentiation between subclinical and extreme cases (see [27]). One alternative involves classifying noncases (“minimal or mild” severity, scores <10 on both GAD-7 and PHQ-9) and cases, including the stratification of 2 levels of severity: “moderate” (“moderate” [GAD-7, scores 10 - 14] and “moderate and moderately severe” [PHQ-9, scores 10 - 19]) and “severe” (GAD-7, score ≥ 15 and PHQ-9, score ≥ 20). We adopted this tripartite stratification due to its parsimonious approach to prioritizing individuals as it supports different care strategies under a straightforward interpretation.

Decision Tree

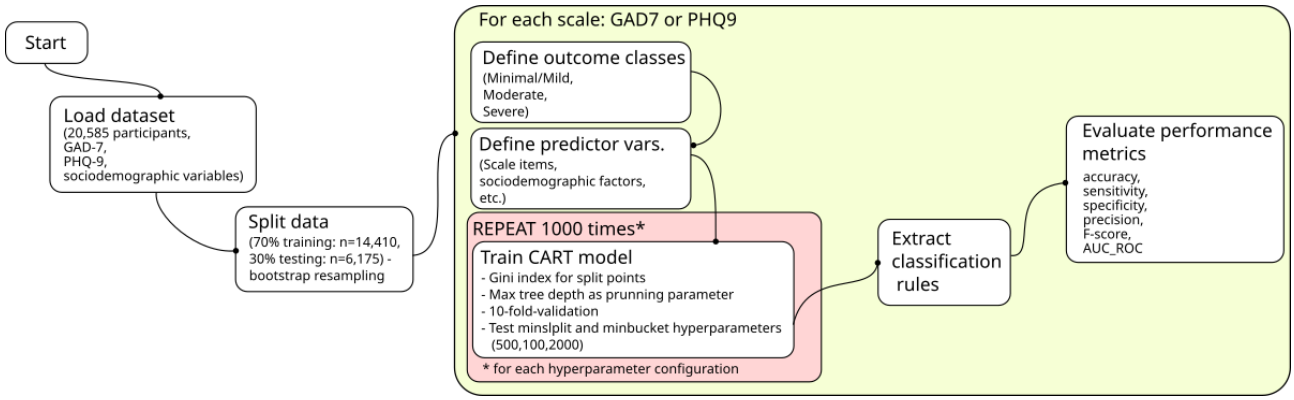
The CART method, originally proposed by Breiman, Friedman, Olshen, and Stone [52], is one of the most widely used recursive partitioning decision tree algorithms and implemented through the rpart package in R. This ML algorithm begins with the complete dataset and recursively partitions it into smaller subsets based on both the predictor values and the outcome classes (see [53-56]). In this study, the classes were defined by the trichotomized GAD-7 and PHQ-9 scores, whereas the variables included both the scale items and sociodemographic factors.

Data Modeling

A combination of bootstrapping and cross-validation was used to mitigate overly optimistic estimations. In this, the entire dataset was resampled multiple times, with 70% of the samples (n=14,410) used for training the ML model and the remaining 30% (n=6175) for testing its performance. This approach helps incorporate sample variability while reducing the risk of overfitting by ensuring the model is not trained on a single dataset. Generalization performance was assessed by estimating accuracy and area under the receiver operating characteristic curves (AUC-ROC) for each repetition. Classification trees were constructed using the Gini index method to determine split points, with maximum tree depth as a pruning parameter and a 10-fold cross-validation strategy implemented via the rpart2 command from the caret package. To test model robustness, the following values for the hyperparameters minsplit (minimum number of observations required in a node) and minbucket (minimum number of observations in a terminal node) were systematically explored: 500, 1000, and 2000. These values were chosen to represent a range from more flexible (500, allowing for smaller nodes and potentially more complex trees) to more constrained (2000, leading to larger nodes and simpler trees), ensuring that the identified rules were not overly specific to a single set of parameters and thus enhancing generalizability (see Figure 1). Each hyperparameter configuration was tested 1000 times to assess the stability and consistency of the derived rules across multiple bootstrapped samples (see [57-59]).

XSL•FO
RenderX

Figure 1. Flowchart of the data modeling procedure for generating decision tree rules to classify anxiety and depression severity in a national sample of Brazilian adults (2024). CART: classification and regression tree; GAD-7: Generalized Anxiety Disorder Scale; PHQ-9: Patient Health Questionnaire (PHQ-9).



Each scale was analyzed separately, with GAD-7 and PHQ-9 scores as predictor variables (features) and severity levels from each scale as outcomes (targets). Four additional models for each scale incorporated demographic variables as additional features. From the 45,000 decision trees generated for each scale, classification rules for each class were extracted. The most frequent rules for each class were evaluated using common performance metrics, including accuracy, sensitivity, specificity, precision, F-score, and AUC-ROC curve. These metrics are widely recognized for classifier performance evaluation (see [60]).

The confusion matrix compared actual classifications (cutoffs) with model predictions. For example, GAD-7 scores categorized anxiety as minimum/mild (0 - 9), moderate (10-14), or severe (15+). Correct classification of a severe case (eg, score 16) is a true positive, and misclassification is a false negative. Similarly, if a nonsevere case (eg, score 6) is correctly classified, it is a true negative and otherwise, a false positive. Confusion matrices summarize classification accuracy, with higher accuracy reflecting better alignment between predicted and actual severity levels. Sensitivity measures the model's ability to detect true positives, while specificity assesses correct identification of true negatives. Higher sensitivity ensures better detection of severe cases, and higher specificity improves nonsevere classification. Precision minimizes false positives, with the F-score balancing precision and recall. The AUC-ROC curve evaluates performance by plotting the true-positive versus false-positive rates, with higher values indicating better classification (see [60]). Variable importance analysis examined demographic predictors' impact on model performance.

External Validation

To evaluate the stability and generalizability of the rules derived from the main sample, we performed an external validation with 5 datasets: 4 collected by other research groups and available in open repositories and 1 collected between November and December 2024 in Brazil using the same recruitment procedures as in the primary study. A description of these datasets is provided in Table 1 of the Supplementary Material. The rules identified in the main analyses were applied unchanged to these datasets to assess reproducibility and to verify that the models were not overfitted to the initial sample but could generalize to a broader population. Model performance was assessed using the same metrics used in the main analyses: accuracy, sensitivity, specificity, precision, and F1-score for each class of each rule.

Software

R (R Development Core Team, 2024) plus RStudio were used to perform all analyses. The decision trees were built using the combination of the *rpart* and *caret* R packages [61] to perform cross-validation as a pruning criterion. Additionally, the values of 500, 1000, and 2000 and their combinations were used to set the *minsplit* and *minbucket* hyperparameters. The R package *pROC* [62] was used to estimate the AUC-ROC curve. All R scripts used for the data analysis are available at the Open Science Framework (OSF) repository [63].

Results

Overview

Based on the total sample, more than 90% were women (n=18,844, 91.5%). Approximately half of the participants declared having white skin color (n=10,405, 50.5%) and high school level education (n=10,083, 49%). The most common age groups were 40 - 49 (n=6071, 29.5%) and 30 - 39 years (n=5927, 28.8%). The mean age was 41.1 years (SD=12.9, range=18 - 80) (Table 1).

XSL•FO
RenderX

Table . Sociodemographic profile of participants in a cross-sectional study on anxiety and depression symptoms in Brazilian adults (n=20,585), 2024.

Variables	Main sample (n=20,585)
Sex/gender, n (%)	
Male	7.6 (1564)
Female	91.5 (18,844)
Nonbinary	0.9 (177)
Skin color/ethnicity, n (%)	
White	50.5 (10,405)
Black	10.5 (2153)
Parda (mixed race)	36.7 (7575)
Other	2.2 (452)
Education level, n (%)	
Up to high school	36.3 (7062)
Undergraduate students	16.7 (3440)
Graduate school	49.0 (10,083)
Age group (years), n (%)	
18 - 29	18.1 (3730)
30 - 39	28.8 (5927)
40 - 49	29.5 (6071)
50 - 60	14.9 (3062)
More than 60	8.7 (1795)

Generalized Anxiety Disorder Scale-7

All trees for the GAD-7 scale showed satisfactory accuracy with a mean of 0.742 and over 0.850 for the “minimal or mild” and “severe” levels, respectively. In total, 159 distinct rules were identified, collectively repeated 206,900 times across the trees, and no demographic variable appeared as part of any rule. The most repeated rules from each GAD-7 class (see Table 2) accounted for more than 59% of all rule repetitions. Table 2 presents the rules for each GAD-7 severity class along with the mean and SD of several classification performance metrics. The most frequently repeated rules for the “minimal or mild” and “severe” severity levels exhibited the highest performance across the metrics. The “moderate” level demonstrated a lower performance than others but with satisfactory indices. Figure 2 displays the ROC curves estimated for all repetitions of the most frequently repeated rules.

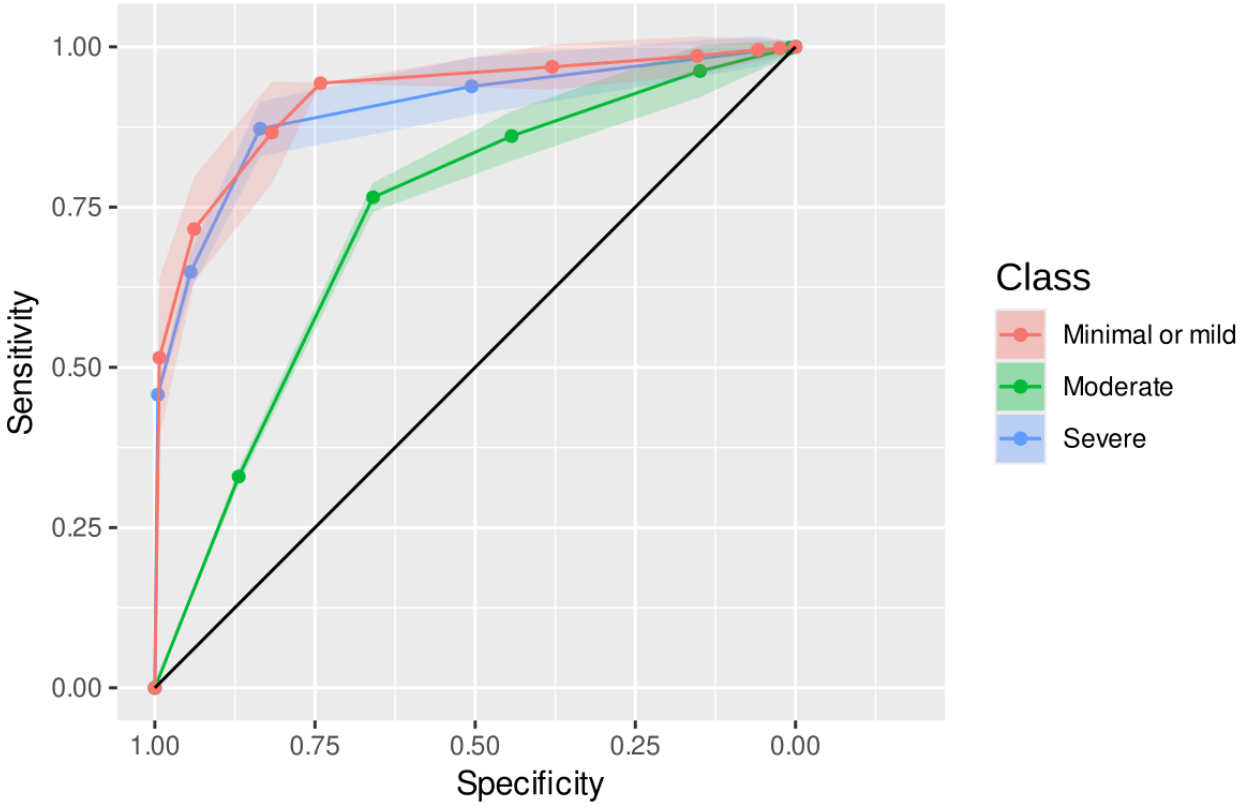
Table . Performance metrics for the most frequently repeated decision tree rules for each severity level of the Generalized Anxiety Disorder (GAD) Scale-7 in a cross-sectional study of Brazilian adults (n=20,585), 2024.[a]

Class	Rule	Repetition	Accuracy	Sensitivity	Specificity	Precision	F-score	AUC[b]
Minimal or mild	GAD2<2 and GAD4<2	43155	0.861(0.003)	0.744 (0.023)	0.942 (0.009)	0.857 (0.015)	0.796 (0.008)	0.923 (0.004)
Moderate	GAD2<2 and GAD4≥2	41780	0.515 (0.066)	0.606 (0.054)	0.798 (0.031)	0.534 (0.021)	0.566 (0.019)	0.728 (0.013)
	GAD2≥2 and GAD4<3	9270	0.507 (0.005)	0.680 (0.009)	0.749 (0.005)	0.507 (0.008)	0.580 (0.007)	0.728 (0.013)
Severe	GAD2≥2 and GAD4=3	27800	0.851 (0.033)	0.831 (0.063)	0.877 (0.030)	0.824 (0.025)	0.825 (0.022)	0.902 (0.009)

[a]Values for accuracy, sensitivity, specificity, precision, F1-score, and AUC are reported as mean (SD) across 45,000 bootstrapped decision tree models.
[b]AUC: area under the curve.

XSL•FO
RenderX

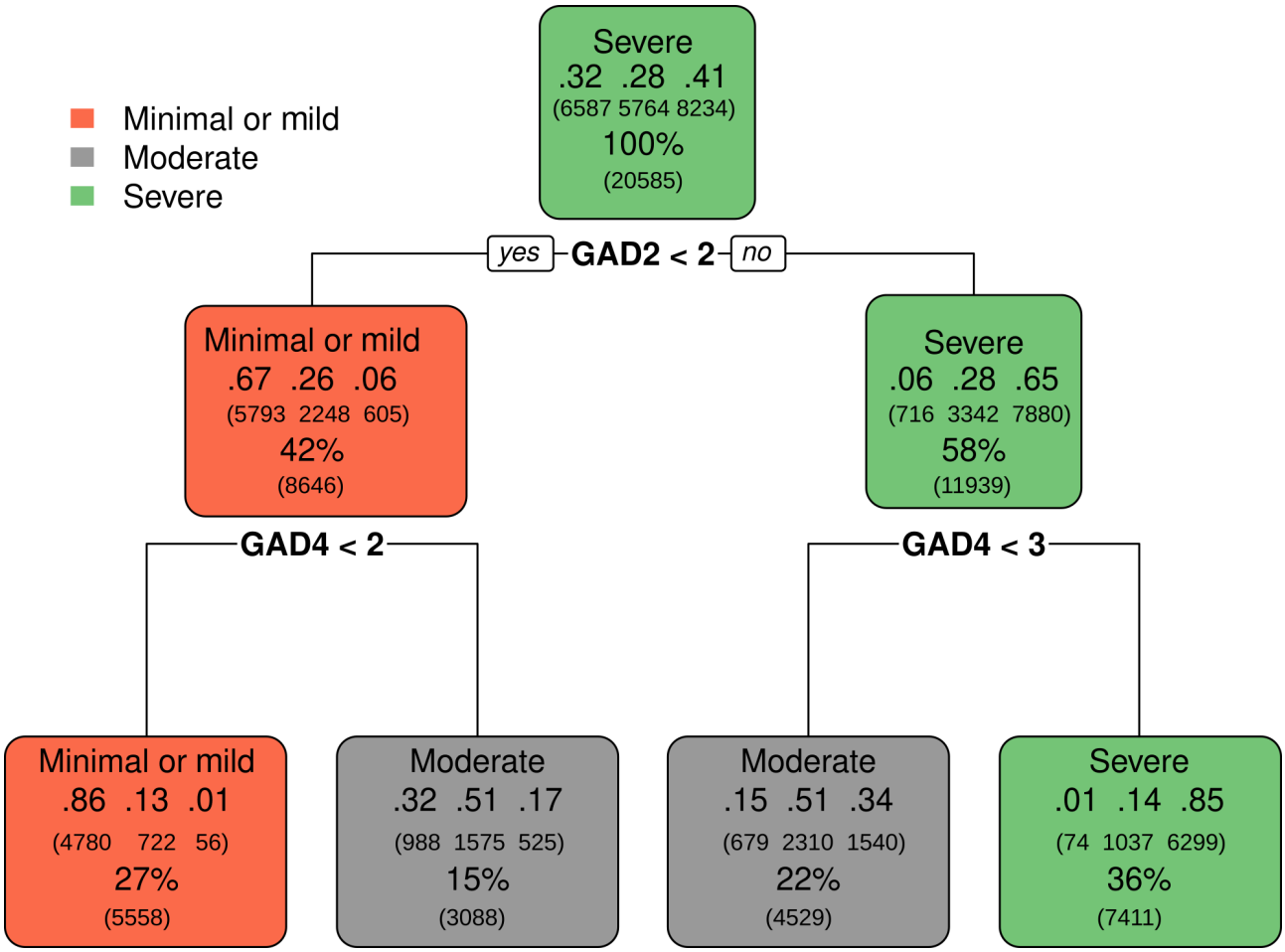
Figure 2. ROC curves for the most frequent decision tree rules classifying GAD-7 severity levels in a cross-sectional study of Brazilian adults (n=20,585), 2024. Each ROC curve represents the classification performance of the most frequently repeated decision tree rule for a given GAD-7 severity level (minimal, mild, moderate, or severe). The colored ribbons around the curves indicate the SD across 45,000 bootstrapped replications. GAD-7: Generalized Anxiety Disorder scale; ROC: receiver operating characteristic.



The most frequently repeated schematic rules for each GAD-7 class as prototypical decision trees are illustrated in Figure 3. These include the "minimal or mild" level and are characterized by the rule "GAD2<2 and GAD4<2." This means that a person who scored 0 or 1 in GAD2 ("not being able to stop or control worrying") and GAD4 ("trouble relaxing") was correctly classified as having "minimal or mild" level anxiety in 86% (0.86) of the cases, within a total of 27% of cases that matched with this class. Of the 27% classified as "minimal or mild," 13% were misclassified as "moderate" (0.13) and almost none as "severe" (0.01).

XSL•FO
RenderX

Figure 3. Prototypical decision trees representing the most frequently repeated classification rules for each severity level of the GAD-7 scale in a cross-sectional study of Brazilian adults (n=20,585), 2024. Each node displays the classification output of a decision rule for a specific GAD-7 severity level (minimal or mild, moderate, and severe). The middle line of each rectangle shows decimal values representing the proportion of participants in each true severity class, ordered left to right: minimal or mild, moderate, and severe. These values indicate the actual distribution of GAD-7 scores among those classified by the rule. The percentage on the bottom line indicates the proportion of the total sample that followed the corresponding decision path. GAD-7: Generalized Anxiety Disorder scale.



The rules were similar for cases from moderate severity levels or positive cases at any level up to significant anxiety symptoms. For example, to classify the "severe" level, the rule was "GAD2≥2 and GAD4=3." Of the participants scoring more than or equal to 2 in GAD2, 58% were included in the "severe" level, with 65% (.65) correctly classified. By adding GAD4 equal to 3, the percentage of cases classified as "severe" reduced to 36%, but the proportion of correctly classified cases increased to 85% (.85), with almost no cases misclassified as "minimal or mild" (.01). This means that the additional item reduced false-positive cases while increasing the accuracy level. The "moderate" levels showed 2 rules and 2 paths, implying more complexity: (a) "GAD2<2 and GAD4≥2," with 15% of total cases, of which 51% (.51) were correctly classified and (b) "GAD2≥2 and GAD4<3" with 22% of the total cases, of which 51% (.51) were correctly classified through the other path.

Beyond statistical performance, the clinical interpretation of these decision rules offers valuable insights. For the "minimal or mild" anxiety level, the rule "GAD2<2 and GAD4<2" suggests that individuals who report "not being able to stop or control worrying" (GAD2) and "trouble relaxing" (GAD4) "not at all" or "several days" (scores of 0 or 1) are highly likely to be correctly classified as having minimal or mild anxiety. By contrast, for the "severe" anxiety level, the rule "GAD2 ≥2 and GAD4=3" indicates that individuals who report "not being able to stop or control worrying" "more than half the days" or "nearly every day" (GAD2 score ≥2) and "trouble relaxing" "nearly every day" (GAD4 score=3) are highly indicative of severe anxiety. This specific combination underscores a significant and pervasive functional impairment related to anxiety symptoms, where both cognitive (worry) and somatic/behavioral (relaxation) aspects are severely affected. The addition of GAD4=3 to the rule for severe cases importantly reduces false positives, indicating that a high level of trouble relaxing is a critical differentiating factor for severe anxiety.

Patient Health Questionnaire-9

The decision trees for the PHQ-9 scores also demonstrated satisfactory accuracy, with a mean of 0.757 and "minimal or mild" and "severe" levels close to 0.800. In total, 126 distinct rules were identified and repeated 230,630 times across the trees. Similar to the GAD-7, no demographic variable appeared as part of any rule. The most repeated rules for each PHQ-9 class (Table 3) accounted for more than 51% of all rule repetitions. Table 3 presents such rules for each PHQ-9 severity class and the classification performance metrics. The "minimal or mild" and "severe" severity levels of the PHQ-9 also

XSL•FO
RenderX

exhibited the highest performance across the estimated metrics. The fit was lower in the "moderate level," although with acceptable indicators. Figure 4 displays the ROC curves for all rule repetitions.

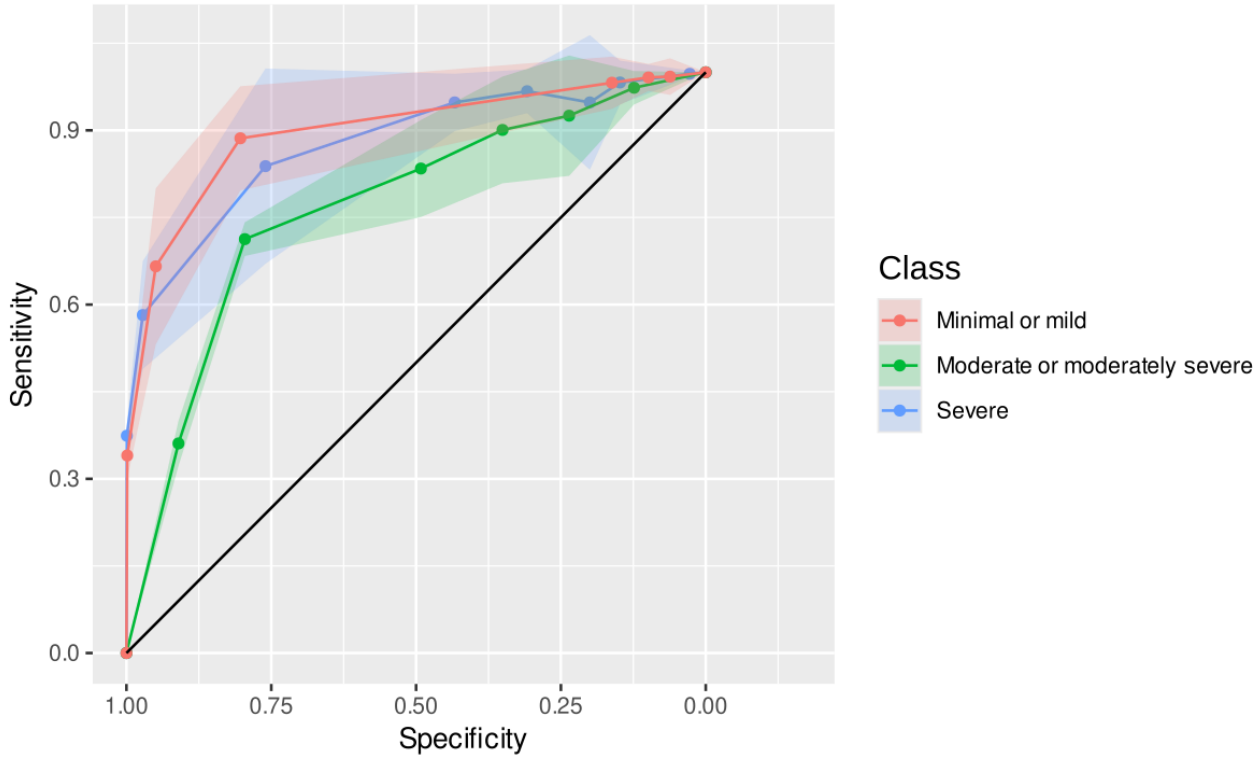
Table . Performance metrics for the most frequently repeated decision tree rules for each severity level of the PHQ-9 scale in a cross-sectional study of Brazilian adults (n=20,585), 2024.[a]

Class	Rule	Repetition	Accuracy	Sensitivity	Specificity	Precision	F1-score	AUC
Minimal or mild	PHQ2<2 and PHQ4<2	28740	0.817 (0.004)	0.767 (0.008)	0.930 (0.003)	0.817 (0.007)	0.792 (0.006)	0.921 (0.011)
Moderate or moderately severe	PHQ2≥2 and PHQ8<2	32125	0.669 (0.004)	0.783 (0.041)	0.723 (0.023)	0.677 (0.012)	0.725 (0.016)	0.776 (0.029)
	PHQ2<2 and PHQ4≥2	29680	0.688 (0.005)	0.761 (0.039)	0.733 (0.032)	0.679 (0.019)	0.717 (0.011)	0.776 (0.029)
Severe	PHQ2≥2 and PHQ8≥2	26405	0.788 (0.003)	0.700 (0.041)	0.921 (0.010)	0.782 (0.012)	0.738 (0.017)	0.903 (0.011)

[a]Values for accuracy, sensitivity, specificity, precision, F1-score, and AUC are reported as mean (SD) across 45,000 bootstrapped decision tree models. Precision, F1-score, and AUC are reported as mean (SD) across 45,000 bootstrapped decision tree models.

[b]AUC: area under the curve.

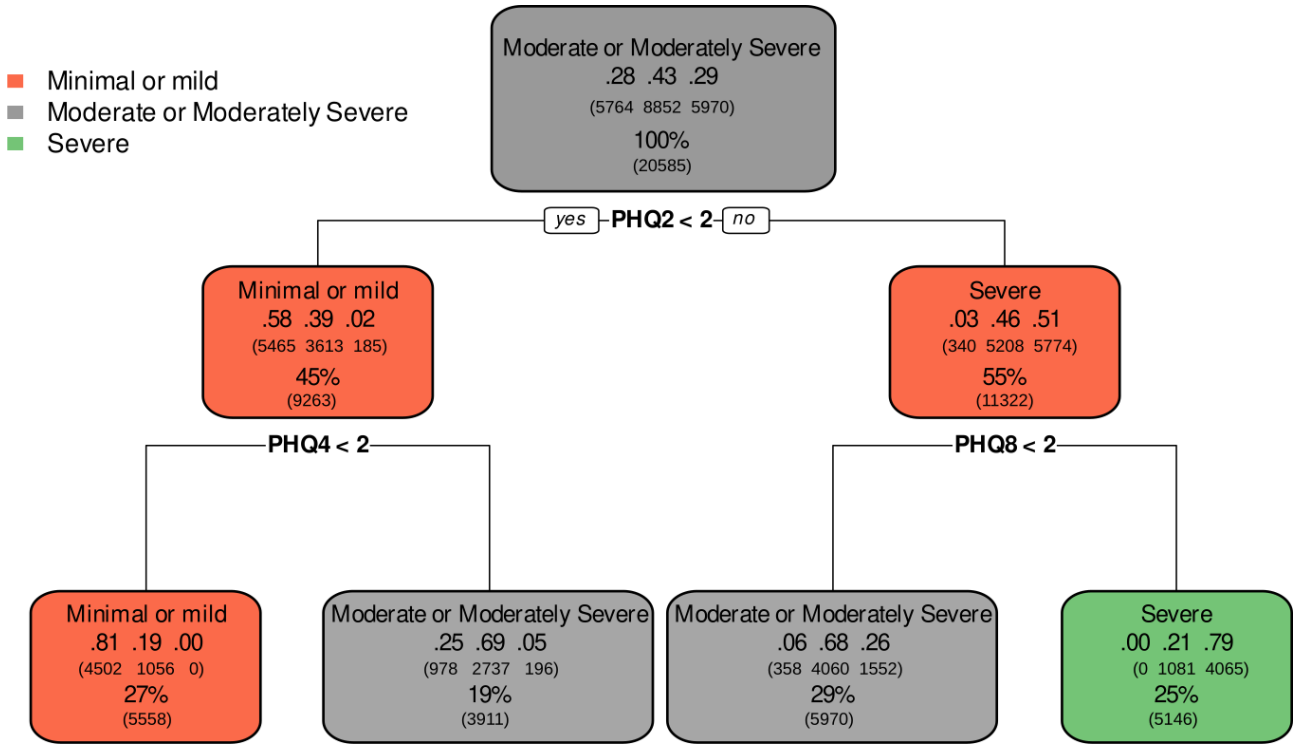
Figure 4. ROC curves for the most frequent decision tree rules classifying PHQ-9 severity levels in a cross-sectional study of Brazilian adults (n=20,585), 2024. Each ROC curve represents the classification performance of the most frequently repeated decision tree rule for a given PHQ-9 severity level (minimal, mild, moderate, moderately severe, or severe). The colored ribbons around the curves indicate the SD across 45,000 bootstrapped replications. PHQ-9: Patient Health Questionnaire-9; ROC: receiver operating characteristic.



The highest repeated rules for each PHQ-9 class are illustrated as prototypical decision trees in Figure 5. The "minimal or mild" level was characterized by the rule "PHQ2<2 and PHQ4<2," which means that a person who scored 0 or 1 on PHQ2 ("feeling down, depressed, or hopeless") and PHQ4 ("feeling tired or having little energy") was correctly classified as having a "minimal or mild" level depression in 81% (.81) of the 27% total cases that fit within this class. Concerning misclassifications, 10% were misclassified as "moderate or moderately severe" (.10) and none as "severe" (.0).

XSL•FO
RenderX

Figure 5. Prototypical decision trees representing the most frequently repeated classification rules for each severity level of the PHQ-9 scale in a cross-sectional study of Brazilian adults (n=20,585), 2024. Each node shows the classification output of a decision rule for a specific PHQ-9 severity level (minimal or mild, moderate or moderately severe, and severe). The decimal values on the middle line of each rectangle represent the proportion of participants in each true severity class, ordered from left to right: minimal or mild, moderate or moderately severe, and severe. These values indicate the actual distribution of PHQ-9 scores among participants classified by the rule. The percentage on the bottom line indicates the proportion of the total sample that followed the corresponding decision path. PHQ-9=Patient Health Questionnaire-9.



Considering only the "moderate" level, which can imply positive cases at some level to significant depression symptoms, the rules were based on a few items. For instance, to classify the "severe" level, the rule was "PHQ2≥2 and PHQ8≥2." Considering the response higher or equal to 2 in PHQ2, 55% of the cases were at the severe level, with 51% (0.51) correctly classified. When the PHQ8 item ("moving or speaking so slowly that other people could have noticed. Or the opposite—being so fidgety or restless that you have been moving around a lot more than usual") greater than or equal to 2 was added, the proportion of cases classified as "severe" decreased to 25%; however, the proportion of correctly classified cases increased to 79% (0.79), with none as "minimal or mild" (0.0). The "moderate or moderately severe" level showed 2 rules, indicating a more complex classification: (a) "PHQ2≥2 and PHQ8<2" with 29% of total cases and shared some fit with the "minimal or mild" (0.06) and "severe" levels (0.26), although 68% (0.68) of them were correctly classified and (b) "PHQ2<2 and PHQ4≥2," with 19% of total cases, of which 69% (0.69) were correctly classified in the other route, which shared some fit with "minimal or mild" and "severe" cases (0.25 and 0.05, respectively). Finally, item 9, which addresses suicidal ideation, did not align with the final rules for the PHQ9 scale.

The clinical interpretation of the PHQ-9 decision paths reveals key depressive symptom profiles. For the "minimal or mild" depression level, the rule "PHQ2<2 and PHQ4<2" suggests that individuals reporting "feeling down, depressed, or hopeless" (PHQ2) and "feeling tired or having little energy" (PHQ4) "not at all" or "several days" (scores of 0 or 1, respectively) were accurately classified as having minimal or mild depression. This highlights that the absence of frequent anhedonia/hopelessness and fatigue is a strong indicator of lower severity. For the "severe" level, the rule "PHQ2≥2 and PHQ8≥2" was particularly revealing. It indicated that individuals experiencing "feeling down, depressed, or hopeless" "more than half the days" or "nearly every day" (PHQ2≥2) and significant psychomotor agitation or retardation ("moving or speaking so slowly.or being so fidgety or restless." [PHQ8≥2]) are highly indicative of severe depression. The inclusion of PHQ8 points to a clinically significant functional impairment often associated with severe depressive episodes, with observable changes in motor activity.

External Validation

When applied to the independent datasets, the rules identified in the main analyses showed consistent performance across both instruments. For the GAD-7, the rules showed high accuracy (range=0.852 - 0.902, M=0.875, SD=0.025), sensitivity (0.792 - 0.903, M=0.847, SD=0.047), specificity (0.916 - 0.951, M=0.937, SD=0.015), precision (0.925 - 0.983, M=0.961, SD=0.033), and F1-score (0.853 - 0.942, M=0.900, SD=0.036) for the minimal anxiety class. For the severe class, the results were also satisfactory, with accuracy (0.865 - 0.969, M=0.919, SD=0.055), sensitivity (0.585 - 0.797, M=0.722, SD=0.094), specificity (0.929 - 0.986, M=0.965, SD=0.025), precision (0.687 - 0.859, M=0.788, SD=0.073), and F1-score (0.673 - 0.793, M=0.745, SD=0.053). The moderate class, however, presented somewhat lower performance, with accuracy (0.761 - 0.877, M=0.816, SD=0.055), sensitivity (0.676 - 0.829, M=0.758, SD=0.071), specificity (0.788 - 0.898, M=0.828,

XSL•FO
RenderX

SD=0.047), precision (0.449 - 0.546, M=0.485, SD=0.046), and F1-score (0.549 - 0.646, M=0.588, SD=0.*041)*.

For the PHQ-9, the rules showed balanced performance across severity levels. For the minimal depression class, the results were as follows: accuracy (range=0.803-0.927, M=0.859, SD=0.051), sensitivity (0.711-0.804, M=0.769, SD=0.044), specificity (0.916-0.981, M=0.939, SD=0.030), precision (0.878-0.935, M=0.899, SD=0.027), and F1-score (0.808-0.839, M=0.828, SD=0.014). Performance for the severe class was also balanced, with accuracy (0.862-0.889, M=0.879, SD=0.012), sensitivity (0.513-0.724, M=0.668, SD=0.104), specificity (0.945-0.958, M=0.948, SD=0.006), precision (0.598-0.864, M=0.752, SD=0.111), and F1-score (0.552-0.782, M=0.707, SD=0.105). The moderate/moderately severe class also showed reasonable performance, with accuracy (0.730-0.808, M=0.768, SD=0.031), sensitivity (0.749-0.892, M=0.803, SD=0.068), specificity (0.677-0.778, M=0.741, SD=0.048), precision (0.591-0.752, M=0.665, SD=0.066), and F1-score (0.688-0.816, M=0.726, SD= 0.059).

Overall, the PHQ-9 rules demonstrated more uniform performance across severity classes, whereas the GAD-7 rules performed strongly for minimal and severe cases but less so for moderate cases. The performance metrics obtained from the validation datasets were consistently aligned with those observed during model training (see Tables 2 and 3, as well as Supplementary Tables 2 and 3 in Multimedia Appendix 1).

Discussion

Principal Findings

This study investigated whether CART models could generate simplified yet accurate decision rules for stratifying anxiety and depression severity using GAD-7 and PHQ-9 items. The findings confirmed that a small number of items—2 for anxiety and 3 for depression—could produce high-performing rules, particularly for identifying minimal/mild and severe cases. Using bootstrapped resampling and cross-validation on a large nonclinical sample, we identified consistent and robust classification patterns, which were also reproduced in an independent dataset collected in Brazil with the same recruitment strategy, and also other four datasets in international contexts and different samples, providing evidence of external validity in relation to the target population. The resulting decision tree rules provided concise symptom pathways with strong performance across multiple metrics, supporting previous findings regarding simplified assessment methods [42]. The most frequent and best-performing rules involved only 2 or 3 items per scale, with no sociodemographic variable contributing to the classification. These findings underscore the potential of brief, rule-based models to enhance screening efficiency while maintaining classification accuracy [53,54].

These findings can be compared with those of other studies that used decision trees to analyze PHQ-9 and GAD-7 [43-46]. However, the models used in this study had the advantage of being less complex and using a large sample size. For example, in two studies that administered PHQ-9 to 2830 Japanese participants, the final model comprised all scale items and more than 20 rules [44,45], while this study had 4 rules. Another study that also applied PHQ-9 to 10,179 participants in Hong Kong found a final model with 8 items and more than 8 rules [43]. Furthermore, some studies with decision tree models included the total GAD-7 and PHQ-9 scores [46], with no straightforward severity stratification. To the best of our knowledge, this study combines the largest sample size with the simplest model. The findings suggest that abbreviated paths could optimize strategies for large-scale healthcare management [41].

Innovative approaches, such as logical rules for the GAD-7 and PHQ-9, as seen in this study, may serve as “shortcuts” for interpreting score combinations, enhancing the precision of large-scale screenings and targeted interventions. The approach leveraged ML models to identify adaptive conditional pathways within the full GAD-7 and PHQ-9 scales. This method preserved the comprehensiveness of the full symptom set while creating shortened, symptom-specific paths—similar to the GAD-2 and PHQ-2—but with the added ability to integrate other symptoms and severity levels in classifications. This approach balances respondent burden with a more nuanced and individualized evaluation of anxiety and depression symptoms, enhancing both screening precision and clinical utility [42]. By identifying the key items and logical rules, these results can support the development of scalable, cost-effective mental health screening models, addressing critical gaps in public health care, particularly in resource-limited settings.

The decision rules identified in this study offer a practical framework for rapid mental health screening, particularly in low-resource settings. By reducing the number of items needed for severity classification, the models allow nonspecialists to quickly identify minimal/mild, moderate, or severe cases. This can support timely decisions in primary care, schools, and community programs—settings where time and training are limited. The high performance for classifying extreme severity levels enables confident triage, guiding decisions from basic psychoeducation to referral for specialized care. These rules are also well-suited for integration into digital tools, expanding access to scalable and efficient assessment.

The current proposal offers significant advantages over GAD-2 and PHQ-2. First, the established ML rules enable the screening of symptom presence or absence and identification of severity levels, while providing detailed information on specific symptoms or sets of symptoms across the full scales. This approach moves beyond the fixed focus on 2 predetermined items in each scale. Second, through such rules, instruments as concise as the GAD-2 and PHQ-2 could be used while retaining the ability to detect other relevant symptoms, facilitating intervention design. The adaptiveness of these rules—the algorithms—allows for replication in diverse contexts, accommodating cultural and local specificities and facilitating the tracking of specific symptom trajectories across various scenarios. This adaptability is important as mental health expressions and symptomatology can vary significantly across cultural backgrounds (eg, [64-66]). Third, the rules can function as a preliminary screening step using if/else conditionals, reducing the length of the full scales, minimizing respondent burden, and lowering associated costs. Finally, it enables tailored

XSL•FO
RenderX

interventions focused on specific symptoms and severity levels. For example, different strategies could be used for addressing PHQ-9 item 4 (fatigue) at minimal or mild versus moderate severity levels.

Our results demonstrated robustness when the procedure was repeated with different cutoff values for the scale scores, consistently identifying rules based on the same scale items. For example, in the PHQ-9 classifications, items 2 and 8 were most relevant for characterizing "severe" cases, whereas items 2 and 4 were most relevant for characterizing "minimal or mild" cases. However, a combination of these 3 items can show the path to the "moderate" levels: more related to the "minimal or mild" class while being closer to the "severe" level. GAD-7 findings could be interpreted similarly, where the moderate level shared some overlapping with "minimum or mild" and "severe" cases. Moreover, common limitations of decision trees, such as overfitting and instability, were not detected in this study [67]. Overfitting, often characterized by overly complex rules [54], was avoided as no rules contained more than 3 elements. Similarly, instability, which refers to changes in the tree structure due to variations in training data, was mitigated by identifying consistent patterns through resampling [68].

The decision tree rules demonstrated robust classification performance for the "minimal or mild" and "severe" categories; however, noticeably lower accuracy was observed for the "moderate" severity class (GAD-7: 51%; PHQ-9: 67%). This finding suggests that the symptom profiles for moderate anxiety and depression may be inherently more nuanced and overlapping with those of minimal/mild and severe cases, leading to increased classification complexity. As illustrated by the prototypical decision trees, the "moderate" levels often required more complex rules and paths compared to the more distinct "minimal or mild" and "severe" levels, indicating a less clear-cut symptomatic boundary. This partial overlap with other severity levels points to an area for future methodological refinement. Future studies can improve the mid-severity level classification using alternative ML approaches. For instance, ensemble models (eg, random forests, gradient boosting machines), which combine multiple decision trees to reduce overfitting and enhance predictive accuracy, might be more adept at capturing the subtle distinctions within moderate symptom presentations. Additionally, exploring probabilistic thresholds rather than strict categorical cutoffs, or incorporating a fuzzy logic approach to severity classification, could provide a more nuanced and potentially more accurate assignment to the moderate range.

While the introduction highlighted the potential for decision trees to incorporate additional variables for adaptive and personalized analyses, the specific analyses revealed that sociodemographic factors (sex/gender, age, education level, and self-reported skin color/ethnicity) had minimal predictive value and did not emerge as part of any of the most frequently repeated classification rules for either the GAD-7 or PHQ-9. This suggests that within the context of our CART models and the specific severity stratification, the core symptom items were overwhelmingly more influential in classification. Consequently, detailed subgroup analyses based on these sociodemographic characteristics were not the most impactful predictive features for streamlined screening. Future research, perhaps using different ML approaches or focusing on distinct research questions, could more thoroughly investigate the interplay between sociodemographic factors and symptom presentation. By contrast, the lower impact of sociodemographic variables reinforces the principle that a broad approach to mental health literacy—including brief, self-guided interventions—can be directed at the general population with a reasonable margin of safety in reaching the intended audience. In other words, it offers a cost-effective strategy by enabling a concise and focused implementation to support scalable public policies.

These findings have significant practical implications for public mental health. The concise, rule-based models derived from the GAD-7 and PHQ-9 can be efficient "shortcuts" for large-scale mental health screening, particularly in resource-limited settings where extensive assessments are often impractical. For instance, in primary care settings, these simplified rules can empower health professionals, including nonspecialists, to rapidly identify disorders at varying risk levels (minimal/mild, moderate, and severe). This rapid stratification can facilitate tailored interventions, ranging from basic psychoeducation and self-guided resources for minimal/mild cases to immediate referral for more intensive support (eg, in-person appointments, specialized care) for moderate and severe cases. This directly supports the development of scalable and cost-effective mental health policies, enabling more efficient allocation of resources and improved access to care for a wider population.

While previous research has used similar methodologies to examine the PHQ-9 and GAD-7 scales using decision trees [43-46], our approach provides several key advantages, warranting a more critical comparison. For instance, some studies [44,45] are focused on individual PHQ-9 items; however, they often used different hyperparameters and, in some cases, dichotomized PHQ-9 scores, leading to substantially more complex decision trees with over 20 rules. By contrast, our study achieved highly accurate classifications with a significantly smaller set of rules (4 for PHQ-9), demonstrating a superior level of parsimony. Furthermore, some prior work [46] used total PHQ-9 and GAD-7 scores and incorporated sociodemographic variables directly into their models but, unlike our study, did not aim for severity stratification based on minimal item sets.

A critical advantage of our study lies in its large sample size, encompassing more than 20,000 participants from across all Brazilian states. This extensive and geographically diverse dataset significantly enhances the external validity and generalizability of our findings compared to those of previous studies with smaller cohorts. This large sample also allowed us to rigorously use bootstrapping and cross-validation, enabling robust estimation of rule performance and providing valuable dispersion measures that attest to the stability of the identified rules. While digital convenience sampling presents a limitation regarding full representativeness of the general population and potential selection bias toward individuals with digital access, the sheer scale and geographic breadth offer a substantial improvement in sample diversity compared to that in several previous studies in this domain. Thus, our work represents a significant advancement in optimizing mental health screening by combining the largest known sample size with the

XSL•FO
RenderX

development of the most parsimonious and highly accurate decision tree models for GAD-7 and PHQ-9 severity classification.

Unlike fixed short scales such as the GAD-2 and PHQ-2, which rely on 2 specific symptoms and offer only binary outcomes, our models generate brief but adaptive rules that capture severity levels across a wider symptom range. While prior decision tree studies often retained most items and many rules, our models achieved high accuracy using only 2 - 3 items and 4 rules per scale. These findings combine the simplicity of ultra-brief tools with the precision of full assessments, offering a more scalable and clinically informative alternative for public mental health applications.

The models identified in this study are concise and can be implemented in primary care by training health professionals in basic psychological care and developing specific intervention tools (eg, self-guided booklets, telephone consultations, and remote support) to manage "minimal or mild" symptoms. For moderate and severe cases, they serve as pathways for the rapid identification of individuals requiring more intensive support, such as in-person appointments, home visits, outpatient care, or even higher-level interventions such as inpatient treatment (eg, day hospitals or long-term hospitalization) [14,69,70]. These initiatives may be more feasible with straightforward instruments for evaluation and decision-making, improving the cost-benefit ratio and adding value to low-cost solutions for most demanding realities, such as mental health care in LMICs [8,40,71-75]. Thus, this study is aligned with the principles of a precision approach in the mental health field [76-78]—specifically, a precision health psychology approach.

Limitations

This study has some limitations. First, the sample presented unbalanced sociodemographic characteristics, with more than 90% of participants identifying as female and an uneven distribution of skin color/ethnicity and education levels. Although variable importance analysis showed that sociodemographic factors—including sex/gender—contributed little to the final classification models (as none appeared in the most frequent rules), this representation pattern may limit the generalizability of the decision rules, particularly for male populations. Future studies should aim for a more balanced sample, especially with respect to sex/gender, which remains a persistent challenge across research contexts [79,80]. Although the CART models prioritized symptomatic items, it is possible that gender-specific symptom expressions or reporting styles may affect rule performance in different demographic contexts.

Second, although the dataset was large and included individuals from both rural and urban areas across Brazil, it is not representative of the general population. The recruitment was based on digital convenience sampling, using social media and electronic invitations, which may have introduced selection bias. Participants with limited internet access or low digital literacy may be underrepresented. Future studies should consider alternative recruitment methods to increase sample heterogeneity and representativeness.

Third, although robustness was supported by cross-validation, bootstrapping, and reproducibility in an independent national dataset, the reliance on digital convenience sampling may still limit representativeness. Future studies should extend external validation to different recruitment strategies and other nonclinical and clinical populations to strengthen generalizability. In addition, the choice of CART models was based on their interpretability and feasibility for large-scale use, but other machine learning techniques—such as ensemble models or deep learning—may offer alternative trade-offs between performance and complexity. Comparative studies could support future applications, especially in resource-limited settings.

Conclusions

This study applied machine learning techniques—specifically decision trees combined with bootstrapping and cross-validation—to improve the classification of anxiety and depression severity based on GAD-7 and PHQ-9 items. Our findings show that just 2 items for anxiety and 3 for depression could generate highly accurate rules for identifying individuals with minimal or severe symptoms. Although the models performed well in distinguishing the extremes, their accuracy was lower for moderate cases, likely due to overlapping symptom patterns. These results suggest that caution is warranted when interpreting classifications in the middle severity ranges.

Simplified, rule-based models like those tested here can contribute to more efficient mental health care strategies. By using fewer items without compromising accuracy, they allow for faster assessments, better allocation of clinical resources, and quicker identification of cases requiring closer attention. In large-scale settings, this may support early interventions for mild cases and timely referrals for individuals showing more severe symptoms. Reducing the length and complexity of screening tools can also ease the burden of national data collection efforts, making it easier to integrate screening into population-based initiatives. Personalized guidance can then be offered based on specific symptom patterns—from psychoeducational materials and mental health literacy to behavioral interventions and clinical support. In this way, simple algorithms may support scalable and adaptive mental health surveillance strategies.

Further research should test these models in more diverse populations, using broader recruitment strategies beyond digital outreach to improve generalizability. While decision trees offer transparency and efficiency, comparative studies involving other machine learning methods—such as ensemble models or deep learning—could shed light on different cost-benefit trade-offs, especially in resource-limited settings. Longitudinal studies are also needed to evaluate the stability of these simplified models over time and their potential to monitor symptom progression or support personalized care. Finally, integrating these tools into digital health platforms or mobile applications could further enhance their accessibility and usefulness in real-world settings.

XSL•FO
RenderX

Acknowledgments

AF is a recipient of a Research Fellowship from the National Council for Scientific and Technological Development (Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, Brazil) and the Coordination for the Improvement of Higher Education Personnel (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, Brazil). The author acknowledges the support provided by these institutions, which contributed to the development of this research. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request. All data are anonymized and managed in accordance with national and international ethical guidelines for research involving human participants. Files are stored in the Open Science Framework (OSF) repository and may be shared with qualified researchers, provided that such access aligns with the terms approved by the research ethics committee. Access details are available [81].

Conflicts of Interest

None declared.

Multimedia Appendix 1
External validation of decision tree models for classifying anxiety and depression severity using GAD-7 and PHQ-9 scores: procedures and results.
[PDF File, 260 KB - publichealth_v11i1e72591_app1.pdf]

References

1. GBD. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3]
2. Patel V, Saxena S, Lund C, et al. Transforming mental health systems globally: principles and policy recommendations. Lancet 2023 Aug 19;402(10402):656-666. [doi: 10.1016/S0140-6736(23)00918-2] [Medline: 37597892]
3. Ahmed N, Barnett P, Greenburgh A, et al. Mental health in Europe during the COVID-19 pandemic: a systematic review. Lancet Psychiatry 2023 Jul;10(7):537-556. [doi: 10.1016/S2215-0366(23)00113-X] [Medline: 37321240]
4. Blasco-Belled A, Tejada-Gallardo C, Fatsini-Prats M, Alsinet C. Mental health among the general population and healthcare workers during the COVID-19 pandemic: A meta-analysis of well-being and psychological distress prevalence. Curr Psychol 2024 Mar;43(9):8435-8446. [doi: 10.1007/s12144-022-02913-6]
5. Castro-de-Araujo LFS, Machado DB. Impact of COVID-19 on mental health in a low and middle-income country. Cien Saude Colet 2020 Jun;25(suppl 1):2457-2460. [doi: 10.1590/1413-81232020256.1.10932020] [Medline: 32520289]
6. Santomauro DF, Mantilla Herrera AM, Shadid J, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. The Lancet 2021 Nov;398(10312):1700-1712. [doi: 10.1016/S0140-6736(21)02143-7]
7. Kola L, Kohrt BA, Hanlon C, et al. COVID-19 mental health impact and responses in low-income and middle-income countries: reimagining global mental health. Lancet Psychiatry 2021 Jun;8(6):535-550. [doi: 10.1016/S2215-0366(21)00025-0] [Medline: 33639109]
8. Moitra M, Owens S, Hailemariam M, et al. Global mental health: where we are and where we are going. Curr Psychiatry Rep 2023 Jul;25(7):301-311. [doi: 10.1007/s11920-023-01426-8] [Medline: 37256471]
9. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022 Dec;54:101675. [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
10. Pringle B, Williams M, Collins PY. Evidence for action: stimulating research to address the global mental health treatment gap. In: Okpaku S, editor. Innovations in Global Mental Health Treatment: Springer International Publishing; 2019:1-26. [doi: 10.1007/978-3-319-70134-9_94-1]
11. Thom J, Mauz E, Peitz D, et al. Establishing a mental health surveillance in Germany: development of a framework concept and indicator set. Robert Koch-Institut 2021 Dec 8. [doi: 10.25646/8861]
12. Wainberg ML, Scorza P, Shultz JM, et al. Challenges and opportunities in global mental health: a research-to-practice perspective. Curr Psychiatry Rep 2017 May;19(5):28. [doi: 10.1007/s11920-017-0780-z] [Medline: 28425023]
13. Peitz D, Kersjes C, Thom J, Hoelling H, Mauz E. Indicators for public mental health: a scoping review. Front Public Health 2021;9:714497. [doi: 10.3389/fpubh.2021.714497] [Medline: 34646802]
14. Mudiyanselage KWW, De Santis KK, Jörg F, et al. The effectiveness of mental health interventions involving non-specialists and digital technology in low-and middle-income countries - a systematic review. BMC Public Health 2024 Jan 3;24(1):77. [doi: 10.1186/s12889-023-17417-6] [Medline: 38172713]

XSL•FO
RenderX

15. Shakeri Hossein Abad Z, Kline A, Sultana M, et al. Digital public health surveillance: a systematic scoping review. NPJ Digit Med 2021 Mar 3;4(1):41. [doi: 10.1038/s41746-021-00407-6] [Medline: 33658681]
16. Ali GC, Ryan G, De Silva MJ. Validated screening tools for common mental disorders in low and middle income countries: a systematic review. PLOS ONE 2016;11(6):e0156939. [doi: 10.1371/journal.pone.0156939] [Medline: 27310297]
17. Miller P, Newby D, Walkom E, Schneider J, Li SC, Evans TJ. The performance and accuracy of depression screening tools capable of self-administration in primary care: a systematic review and meta-analysis. Eur J Psychiatry 2021 Jan;35(1):1-18. [doi: 10.1016/j.ejpsy.2020.10.002]
18. Terman JM, Rohan KJ, Cruz RC, Greenberger E. Ultra-brief cognitive-behavioral therapy (for routine primary care visits: feasibility and acceptability of a brief provider training workshop. Fam Syst Health 2024 Dec;42(4):559-569. [doi: 10.1037/fsh0000862] [Medline: 38032676]
19. Bond RR, Mulvenna MD, Potts C, O'Neill S, Ennis E, Torous J. Digital transformation of mental health services. NPJ Mental Health Res 2023 Aug 22;2(1):1-9. [doi: 10.1038/s44184-023-00033-y]
20. Makivić I, Kragelj A, Lasalvia A. Quantitative needs assessment tools for people with mental health problems: a systematic scoping review. Curr Psychol 2024 Jun;43(22):19996-20013. [doi: 10.1007/s12144-024-05817-9]
21. Mughal S, Salmon A, Churchill A, Tee K, Jaouich A, Shah J. Guiding principles for implementing stepped care in mental health: alignment on the bigger picture. Community Ment Health J 2023 Aug;59(6):1035-1042. [doi: 10.1007/s10597-023-01116-y] [Medline: 37002486]
22. Summanwar D, Ropert C, Barton J, et al. Agile implementation of alcohol screening in primary care. BMC Prim Care 2024 Jul 11;25(1):251. [doi: 10.1186/s12875-024-02500-7] [Medline: 38992572]
23. Arias de la Torre J, Vilagut G, Ronaldson A, et al. Reconsidering the use of population health surveys for monitoring of mental health. JMIR Public Health Surveill 2023 Nov 23;9(1):e48138. [doi: 10.2196/48138] [Medline: 37995112]
24. Kost RG, Correa da Rosa J. Impact of survey length and compensation on validity, reliability, and sample characteristics for Ultrashort-, Short-, and Long-Research Participant Perception Surveys. J Clin Trans Sci 2018 Feb;2(1):31-37. [doi: 10.1017/cts.2018.18]
25. Shevlin M, Butter S, McBride O, et al. Measurement invariance of the Patient Health Questionnaire (PHQ-9) and Generalized Anxiety Disorder scale (GAD-7) across four European countries during the COVID-19 pandemic. BMC Psychiatry 2022 Mar 1;22(1):154. [doi: 10.1186/s12888-022-03787-5] [Medline: 35232409]
26. Stochl J, Fried EI, Fritz J, et al. On Dimensionality, measurement invariance, and suitability of sum scores for the PHQ-9 and the GAD-7. Assessment 2022 Apr;29(3):355-366. [doi: 10.1177/1073191120976863] [Medline: 33269612]
27. Villarreal-Zegarra D, Barrera-Begazo J, Otazú-Alfaro S, Mayo-Puchoc N, Bazo-Alvarez JC, Huarcaya-Victoria J. Sensitivity and specificity of the Patient Health Questionnaire (PHQ-9, PHQ-8, PHQ-2) and General Anxiety Disorder scale (GAD-7, GAD-2) for depression and anxiety diagnosis: a cross-sectional study in a Peruvian hospital population. BMJ Open 2023 Sep 15;13(9):e076193. [doi: 10.1136/bmjopen-2023-076193] [Medline: 37714674]
28. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
29. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 6;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
30. Kroenke K, Spitzer RL, Williams JBW, Löwe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. Gen Hosp Psychiatry 2010;32(4):345-359. [doi: 10.1016/j.genhosppsych.2010.03.006] [Medline: 20633738]
31. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5: American Psychiatric Publ; 2013.
32. Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion 2020 Jun;58:82-115. [doi: 10.1016/j.inffus.2019.12.012]
33. Emmert‐Streib F, Yli‐Harja O, Dehmer M. Explainable artificial intelligence and machine learning: A reality rooted perspective. WIREs Data Min & Knowl 2020 Nov;10(6):e1368 [FREE Full text] [doi: 10.1002/widm.1368]
34. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning: With Applications in R: Springer; 2021.
35. Sarker IH. Machine learning: algorithms, real-world applications and research directions. SN Comput Sci 2021;2(3):160. [doi: 10.1007/s42979-021-00592-x] [Medline: 33778771]
36. Jiang T, Gradus JL, Rosellini AJ. Supervised machine learning: a brief primer. Behav Ther 2020 Sep;51(5):675-687. [doi: 10.1016/j.beth.2020.05.002] [Medline: 32800297]
37. Montag C, Hall B. Enhancing real-time digital surveillance can guide evidence-based policymaking to improve global mental health. Nat Mental Health 2023 Oct;1(10):697-698. [doi: 10.1038/s44220-023-00129-6]
38. Shatte ABR, Hutchinson DM, Teague SJ. Machine learning in mental health: a scoping review of methods and applications. Psychol Med 2019 Jul;49(9):1426-1448. [doi: 10.1017/S0033291719000151] [Medline: 30744717]
39. Skaik R, Inkpen D. Using social media for mental health surveillance. ACM Comput Surv 2021 Nov 30;53(6):1-31. [doi: 10.1145/3422824]

XSL•FO
RenderX

40. Olawade DB, Wada OZ, Odetayo A, David-Olawade AC, Asaolu F, Eberhardt J. Enhancing mental health with artificial intelligence: current trends and future prospects. Journal of Medicine, Surgery, and Public Health 2024 Aug;3:100099. [doi: 10.1016/j.glmedi.2024.100099]
41. Gonzalez O. Psychometric and machine learning approaches to reduce the length of scales. Multivariate Behav Res 2021;56(6):903-919. [doi: 10.1080/00273171.2020.1781585] [Medline: 32749158]
42. Pargent F, Schoedel R, Stachl C. Best practices in supervised machine learning: a tutorial for psychologists. Advances in Methods and Practices in Psychological Science 2023 Jul;6(3). [doi: 10.1177/25152459231162559]
43. Chin WY, Wan EYF, Dowrick C, Arroll B, Lam CLK. Tree analysis modeling of the associations between PHQ-9 depressive symptoms and doctor diagnosis of depression in primary care. Psychol Med 2019 Feb;49(3):449-457. [doi: 10.1017/S0033291718001058] [Medline: 29697038]
44. Colledani D, Anselmi P, Robusto E. Machine learning-decision tree classifiers in psychiatric assessment: an application to the diagnosis of major depressive disorder. Psychiatry Res 2023 Apr;322:115127. [doi: 10.1016/j.psychres.2023.115127] [Medline: 36842398]
45. Colledani D, Robusto E, Anselmi P. Shortening and personalizing psychodiagnostic assessments with decision tree-machine learning classifiers: an application example based on the Patient Health Questionnaire-9. Int J Ment Health Addiction 2024 May 23. [doi: 10.1007/s11469-024-01332-x]
46. Wy S, Choe S, Lee YJ, et al. Decision tree algorithm-based prediction of vulnerability to depressive and anxiety symptoms in caregivers of children with glaucoma. Am J Ophthalmol 2022 Jul;239:90-97. [doi: 10.1016/j.ajo.2022.01.025] [Medline: 35172169]
47. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder. Arch Intern Med 2006 May 22;166(10):1092. [doi: 10.1001/archinte.166.10.1092]
48. Moreno AL, DeSousa DA, Souza AMFLP, et al. Factor structure, reliability, and item parameters of the Brazilian-Portuguese version of the GAD-7 Questionnaire. Temas Psicol 2016;24(1):367-376. [doi: 10.9788/TP2016.1-25]
49. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
50. de Lima Osório F, Vilela Mendes A, Crippa JA, Loureiro SR. Study of the discriminative validity of the PHQ-9 and PHQ-2 in a sample of Brazilian women in the context of primary health care. Perspect Psychiatr Care 2009 Jul;45(3):216-227. [doi: 10.1111/j.1744-6163.2009.00224.x]
51. Faro A, Nunes D, Falk D. Depressive symptomatology in Brazil: perspectives of statistical and psychometrics analyses of the PHQ-9 at four time-points (2020-2023) in the COVID-19 pandemic. Front Psychol 2025;16:1440054. [doi: 10.3389/fpsyg.2025.1440054] [Medline: 39973955]
52. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and Regression Trees: Taylor & Francis; 1984.
53. Loh WY. Classification and regression trees. WIREs Data Min & Knowl 2011 Jan;1(1):14-23. [doi: 10.1002/widm.8]
54. Loh WY. Fifty years of classification and regression trees. Int Statistical Rev 2014 Dec;82(3):329-348. [doi: 10.1111/insr.12016]
55. Steinberg D. CART: Classification and Regression Trees The Top Ten Algorithms in Data Mining: Chapman and Hall; 2009. URL: https://theswissbay.ch/pdf/Gentoomen%20Library/Algorithms/Top%20Ten%20Data%20Mining%20Algos%20Chapters/%231.pdf [accessed 2025-09-22]
56. Cleophas TJ, Zwinderman AH. Regression trees. In: Regression Analysis in Medical Research: For Starters and 2nd Levelers 2021. URL: https://www.scribd.com/document/899891356/Regression-Analysis-In-Medical-Research-For-Starters-And-2nd-Levelers-Ton-J-Cleophas-instant-download [accessed 2025-09-22]
57. Grubinger T, Zeileis A, Pfeiffer KP. evtree: evolutionary learning of globally optimal classification and regression trees in R. J Stat Soft 2011;61(1). [doi: 10.18637/jss.v061.i01]
58. Bischl B, Binder M, Lang M, et al. Hyperparameter optimization: foundations, algorithms, best practices, and open challenges. WIREs Data Min & Knowl 2023 Mar;13(2):e1484. [doi: 10.1002/widm.1484]
59. Mantovani RG, Horvath T, Cerri R, Vanschoren J, de Carvalho A. Hyper-parameter tuning of a decision tree induction algorithm. Presented at: 2016 5th Brazilian Conference on Intelligent Systems (BRACIS); Recife p. 37-42. [doi: 10.1109/BRACIS.2016.018]
60. M H, M.n S. A review on evaluation metrics for data classification evaluations. IJDKP 2015 Mar 31;5(2):01-11. [doi: 10.5121/ijdkp.2015.5201]
61. Kuhn M. Building predictive models in R using the caret package. J Stat Softw 2008 Nov 10;28:1-26. [doi: 10.18637/jss.v028.i05]
62. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011 Mar 17;12(1):77. [doi: 10.1186/1471-2105-12-77] [Medline: 21414208]
63. Tejada J, Faro A. Optimizing anxiety and depression screening through a precision psychology approach. OSF 2024 Dec 20 [FREE Full text]

XSL•FO
RenderX

64. Harper Shehadeh M, Heim E, Chowdhary N, Maercker A, Albanese E. Cultural adaptation of minimally guided interventions for common mental disorders: a systematic review and meta-analysis. JMIR Ment Health 2016 Sep 26;3(3):e44. [doi: 10.2196/mental.5776]
65. Rathod S, Gega L, Degnan A, et al. The current status of culturally adapted mental health interventions: a practice-focused review of meta-analyses. Neuropsychiatr Dis Treat 2018;14:165-178. [doi: 10.2147/NDT.S138430] [Medline: 29379289]
66. Zhao Y, Summers R, Gathara D, English M. Conducting cross-cultural, multi-lingual or multi-country scale development and validation in health care research: a 10-step framework based on a scoping review. J Glob Health 2024 Jul 19;14:04151. [doi: 10.7189/jogh.14.04151] [Medline: 39024643]
67. Hawkins DM. Recursive partitioning. WIREs Computational Stats 2009 Nov;1(3):290-295. [doi: 10.1002/wics.44]
68. Rokach L. Decision forest: twenty years of research. Information Fusion 2016 Jan;27:111-125. [doi: 10.1016/j.inffus.2015.06.005]
69. Purgato M, Uphoff E, Singh R, Thapa Pachya A, Abdulmalik J, van Ginneken N. Promotion, prevention and treatment interventions for mental health in low- and middle-income countries through a task-shifting approach. Epidemiol Psychiatr Sci 2020 Aug 3;29:e150. [doi: 10.1017/S204579602000061X] [Medline: 32744223]
70. Whitfield J, Owens S, Bhat A, Felker B, Jewell T, Chwastiak L. Successful ingredients of effective collaborative care programs in low- and middle-income countries: a rapid review. Camb prisms Glob ment health 2023;10:e11. [doi: 10.1017/gmh.2022.60]
71. Freeman M. Investing for population mental health in low and middle income countries-where and why? Int J Ment Health Syst 2022 Aug 11;16(1):38. [doi: 10.1186/s13033-022-00547-6] [Medline: 35953845]
72. Lattie EG, Stiles-Shields C, Graham AK. An overview of and recommendations for more accessible digital mental health services. Nat Rev Psychol 2022 Feb;1(2):87-100. [doi: 10.1038/s44159-021-00003-1] [Medline: 38515434]
73. Le PD, Eschliman EL, Grivel MM, et al. Barriers and facilitators to implementation of evidence-based task-sharing mental health interventions in low- and middle-income countries: a systematic review using implementation science frameworks. Implementation Sci 2022 Jan 12;17(1):4. [doi: 10.1186/s13012-021-01179-z]
74. Lund C. Global mental health and its social determinants: how should we intervene? Behav Res Ther 2023 Oct;169:104402. [doi: 10.1016/j.brat.2023.104402] [Medline: 37677893]
75. Montag C, Ali R, Al-Thani D, Hall BJ. On artificial intelligence and global mental health. Asian J Psychiatr 2024 Jan;91:103855. [doi: 10.1016/j.ajp.2023.103855] [Medline: 38113698]
76. Bickman L, Lyon AR, Wolpert M. Achieving precision mental health through effective assessment, monitoring, and feedback processes : introduction to the special issue. Adm Policy Ment Health 2016 May;43(3):271-276. [doi: 10.1007/s10488-016-0718-5] [Medline: 26887937]
77. DeRubeis RJ. The history, current status, and possible future of precision mental health. Behav Res Ther 2019 Dec;123:103506. [doi: 10.1016/j.brat.2019.103506] [Medline: 31706160]
78. Hekler E, Tiro JA, Hunter CM, Nebeker C. Precision health: the role of the social and behavioral sciences in advancing the vision. Ann Behav Med 2020 Nov 1;54(11):805-826. [doi: 10.1093/abm/kaaa018] [Medline: 32338719]
79. Borg DJ, Haritopoulou-Sinanidou M, Gabrovska P, et al. Barriers and facilitators for recruiting and retaining male participants into longitudinal health research: a systematic review. BMC Med Res Methodol 2024 Feb 22;24(1):46. [doi: 10.1186/s12874-024-02163-z] [Medline: 38389065]
80. Ryan J, Lopian L, Le B, et al. It's not raining men: a mixed-methods study investigating methods of improving male recruitment to health behaviour research. BMC Public Health 2019 Jun 24;19(1):814. [doi: 10.1186/s12889-019-7087-4] [Medline: 31234825]
81. OSF home. URL: https://osf.io/ntsuf/ [accessed 2025-09-22]

Abbreviations

AI: artificial intelligence
AUC: area under the curve
CART: classification and regression tree
CMDs: common mental disorders
CONEP: Conselho Nacional de Ética em Pesquisa
GAD-2: Generalized Anxiety Disorder Scale-2
GAD-7: Generalized Anxiety Disorder Scale-7
LMICs: low- and middle-income countries
ML: machine learning
OSF: Open Science Framework
PHQ-2: Patient Health Questionnaire-2
PHQ-9: Patient Health Questionnaire-9
ROC: receiver operating characteristic

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 12.02.25; peer-reviewed by M Cavus, O Valentim; revised version received 23.08.25; accepted 28.08.25; published 30.09.25.

Please cite as:
Faro A, Tejada J, Al-Delaimy W
Severity Classification of Anxiety and Depression Using Generalized Anxiety Disorder Scale and Patient Health Questionnaire: National Cross-Sectional Study Applying Classification and Regression Tree Models
JMIR Public Health Surveill 2025;11:e72591
URL: https://publichealth.jmir.org/2025/1/e72591
doi:10.2196/72591

©Andre Faro, Julian Tejada, Wael Al-Delaimy. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Trends in Avoidable Hospitalizations Before and During the COVID-19 Pandemic: Multiple Cross-Sectional Study Using Administrative Data From Beijing, China

Xiangzhen Wang[1*], MPH; Yin Chen[2*], MM; Yuqi Ta[3], PhD; Moning Guo[2], MPH; Hongqiao Fu[1], PhD

[1]School of Public Health, Peking University, No. 38 Xueyuan Road, Beijing, China

[2]Beijing Municipal Health Big Data and Policy Research Center, Beijing, China

[3]School of Public Finance and Taxation, Central University of Finance and Economics, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Hongqiao Fu, PhD
School of Public Health, Peking University, No. 38 Xueyuan Road, Beijing, China

Abstract

Background: Avoidable hospitalizations (AHs) have been widely used in high-income countries as a proxy indicator for the quality of primary care. However, it is rarely evaluated in low- and middle-income countries such as China. Studies examining changes in AHs before and during the COVID-19 pandemic are also limited. The appropriateness of AHs as an indicator measuring primary care quality under pandemic conditions has not been well discussed.

Objective: This study aims to describe trends in AHs in Beijing, China, during both the prepandemic (2016-2019) and pandemic (2020-2021) periods and examine factors associated with AH rates.

Methods: We used hospital discharge data of Beijing residents between January 1, 2016, and December 31, 2021. We identified AH cases from all discharge cases and calculated AH rates each year, adjusting for population structure changes. We performed regression analyses to explore factors associated with AH rates, where the COVID-19 outbreak, health care resources, and socioeconomic characteristics were used as the main explanatory variables.

Results: Before the COVID-19 pandemic, the total number of hospital discharges in Beijing increased steadily from 2016 to 2019 but decreased sharply in 2020 and partially rebounded in 2021. The sex- and age-standardized AH rate per 100,000 population rose from 514.7 (95% CI 511.4 - 517.9) in 2016 to 552.8 (95% CI 549.4 - 556.1) in 2019. Then it declined to 331.2 (95% CI 328.6 - 333.8) in 2020 and rebounded to 465.1 (95% CI 462.1 - 468.1) in 2021, which was still below the prepandemic level. Regression analyses show that the presence of newly confirmed COVID-19 cases was significantly associated with a lower AH rate. As for other factors, higher densities of primary physicians were linked to lower AH rates. Moreover, AH rates were also associated with population structure, the level of economic development, and demographic variables.

Conclusions: The AH rate in Beijing exhibited a consistent upward trend before the pandemic and remained higher than in many high-income countries. These characteristics suggest a potential overuse of tertiary care and highlight the necessity for health care system reforms in Beijing, particularly a transition from the hospital-centered model to a primary care–focused delivery system. In addition, the observed associations between AH rates and factors, such as pandemic shock and socioeconomic variables, indicate that AH should be interpreted with appropriate controls when it is used as an indicator of primary care performance.

(JMIR Public Health Surveill 2025;11:e69768) doi:10.2196/69768

KEYWORDS

avoidable hospitalization; COVID-19; primary care; accessibility; quality of care

Introduction

In the past 2 decades, improving access to high-quality primary or ambulatory care has been a global policy priority [1]. It has been regarded as a means to improve population health and control health expenditures. One key measure that policy makers and researchers use to evaluate the accessibility to high-quality primary care is hospitalizations related to ambulatory care sensitive conditions (ACSCs), as they could be potentially avoided with timely and effective health care management [2,3]. These ACSCs-related hospitalizations are often referred to as avoidable hospitalizations (AHs) that may impose substantial burdens on individuals, the health care system, and the whole society.

AHs have been widely used to evaluate primary care performance in high-income countries, such as the United States,

XSL•FO
RenderX

Europe, Australia, and South Korea [4-9]. A series of studies have also identified factors associated with AHs in high-income countries, such as access to primary care, organizational arrangements, and individual socioeconomic status [10-12]. However, to the best of our knowledge, few studies have investigated the trends in AHs and their associated factors in low- and middle-income countries [13,14], where health care resources are generally insufficient and the quality of care is usually unsatisfactory. More evidence on AHs from low-income countries contributes to a better understanding of their health care system performance through international comparisons.

Furthermore, our understanding of trends in AHs during the COVID-19 pandemic is limited, especially in low-income countries. A series of studies on high-income countries have shown the impacts of the COVID-19 pandemic on disrupting access to non–COVID-19–related medical services, such as decreasing hospitalizations and preventive service usage [15-17]. However, little is known about changes in AHs in low- and middle-income countries before and during the pandemic, although existing studies have found substantial declines in outpatient visits and hospital admissions [18-20]. A few studies have also examined changes in admissions for ACSCs during the pandemic but results have been inconsistent [21-24]. On the one hand, lockdown policies during the pandemic may reduce hospitalizations related to ACSCs, resulting in a decrease in the AH rate. On the other hand, delayed health care usage may exacerbate ACSCs, leading to an increased AH rate.

In this study, we aim to use Beijing as an example to examine AHs in the setting of the Chinese health care system. We used hospital discharge data from Beijing between 2016 and 2021 to describe trends in AHs, which cover periods prior to and during the COVID-19 outbreak. We also performed regression analysis to explore the factors influencing AH rates, with COVID-19 cases and socioeconomic characteristics as the main explanatory variables. As the capital of China, Beijing leads in socioeconomic development and population health in China. It is the medical hub of North China with abundant medical resources. The social health insurance schemes in Beijing, especially in the coverage of outpatient care, are more generous than those in most other Chinese cities. The hospital discharge rate among Beijing's local residents in 2019 was approximately 11.1% (2,431,274/21,901,000), which was lower than both the national average in China (18.6%) and the average across 35 Organisation for Economic Cooperation and Development (OECD) countries (14.9%) [25,26]. However, similar to other Chinese cities, Beijing has a profit-driven, hospital-centered health care system with weak primary care [27-29]. In 2019, more than 50.0% of the medical revenues of public hospitals in Beijing came from pharmaceuticals and diagnostic tests, while less than 25.0% were from medical services [30]. Approximately 172.1 million outpatient visits were provided by hospitals, while primary care institutions accounted for only 68.3 million visits [31]. Moreover, 62.6% of total health expenditures went to hospitals, while only 10.2% were allocated to primary health care institutions [32]. All these figures suggested the system's heavy reliance on hospital-based care. Due to low perceived quality, patients do not trust primary care facilities much and often bypass the primary care system. Consequently, they often seek care at tertiary hospitals for simple conditions, leading to overcrowded tertiary hospitals and rising health expenditures. Therefore, results about AHs in Beijing not only inform us of the health system performance in Beijing but also provide implications for the challenges in the Chinese health care system.

One of the prominent events during our study period was the outbreak of COVID-19 pandemic, which may profoundly affect the trend and level of AHs. After the outbreak of the COVID-19 pandemic in January 2020, the Beijing municipal government took strict measures to control the viral transmission, including restrictions on human mobility, body temperature monitoring, and contact tracing [33]. Although there were only sporadic outbreaks and a small number of confirmed cases in the following 2 years (as shown in Figures S1 and S2 in Multimedia Appendix 1), these strict measures may generate substantial effects on people's daily lives, including consumption behaviors, willingness to work from home, investment decisions, and health care usage [34-37]. As some studies have already examined the complex drivers of AHs (eg, primary care quality, socioeconomic conditions, hospital admission policy, and health insurance types) and questioned the appropriateness of AHs in measuring the quality of primary care [38,39], trends in AHs before and during the COVID-19 pandemic could provide important implications for the effectiveness and limitations of the usage of AH as a proxy for access and quality of primary care, given that large shocks to both the demand and supply sides of health care services were caused by the COVID-19 pandemic.

Methods

Study Population and Data Source

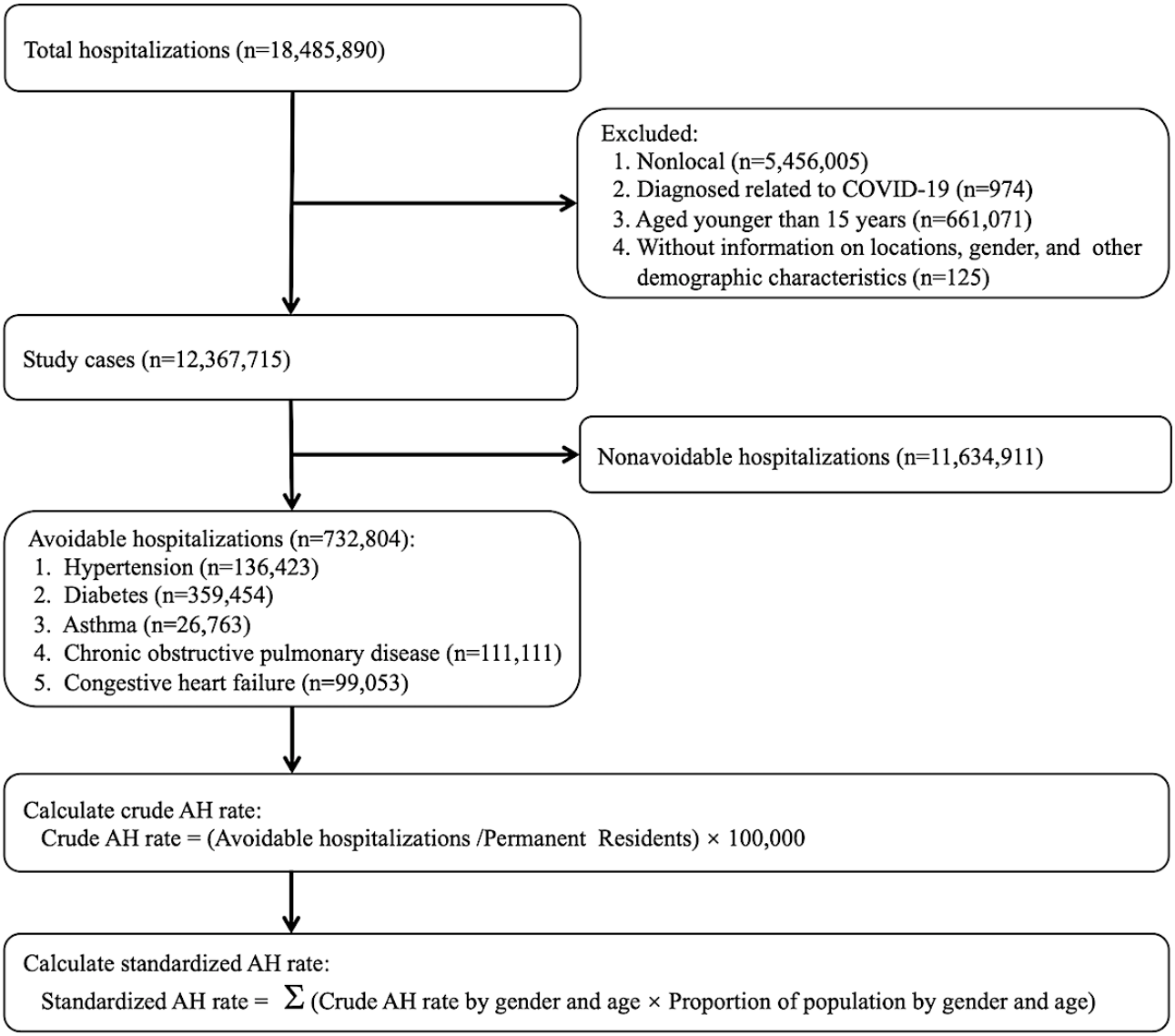
In this study, we mainly used hospital discharge data in Beijing between January 1, 2016, and December 31, 2021, to investigate trends in AHs. This dataset was acquired from the Beijing Municipal Health Big Data and Policy Research Center. The dataset included all hospital discharge records in 267 hospitals in Beijing, except for 13 military hospitals. It contained extensive information on medical diagnoses (the *ICD-10* [*International Statistical Classification of Diseases, Tenth Revision*] code), procedures (the *ICD-9-CM-3* [Volume 3 of the *International Classification of Diseases, Ninth Revision, Clinical Modification*] code), length of stay, health expenditures, health insurance types, and patient demographic characteristics [40]. As this dataset encompassed both the prepandemic and pandemic periods, it allowed us to scrutinize how the COVID-19 pandemic affected health care usage among residents in Beijing.

According to the Health Care Quality Indicator Definitions proposed by the OECD [41], discharges with a principle of *ICD-10-CM* (*International Classification of Diseases, Tenth Revision, Clinical Modification*) code for any 1 of the following 6 ACSCs were identified as AHs: hypertension, diabetes, asthma, chronic obstructive pulmonary disease (COPD), congestive heart failure (CHF), and diabetes-related lower limb amputation. Because of their small volume (less than 200 cases per year) and the overlap with diabetes cases, diabetes-related lower limb amputations were merged into the scope of diabetes

XSL•FO
RenderX

for analysis. This classification allowed us to compare our results with international studies as well as some other China-specific findings [13,42,43]. As depicted in Figure 1, we identified a total of 732,804 cases of AHs, adhering to the specific inclusion and exclusion criteria outlined in Table S1 in Multimedia Appendix 1. Among these AHs, there were 136,423 cases of hypertension, 359,454 cases of diabetes, 26,763 cases of asthma, 111,111 cases of COPD, and 99,053 cases of CHF.

Figure 1. Flowchart for identifying AHs and calculating AH rates. The initial dataset contained 18,485,890 discharge cases during the period between 2016 and 2021. Then, we narrowed our sample to local patients because this study aimed to examine AHs among permanent residents in Beijing. Beijing is the largest medical hub in China, serving a large share of inpatient cases from various provinces nationwide. Since these nonlocal inpatients' conditions tend to be more severe and are not related to primary care in Beijing, nonlocal inpatients were excluded to avoid biasing the estimation of AH rates. We also excluded inpatient cases related to COVID-19, cases related to age less than 15 years [41], and cases lacking information on residential locations, sex, and other demographic characteristics. Finally, we obtained a sample of 12,367,715 cases for our analysis. AH: Avoidable hospitalization.



Variables

In this study, we defined key, dependent variables at both case and district levels. At the case level, the dependent variable is a binary indicator denoting whether a given hospital admission was an AH. At the district level, the dependent variable is the monthly AH rate for each district. As shown in Figure 1, we conducted calculations by dividing the count of AHs by the total resident population and then multiplied the result by 100,000 to derive the crude AH rates per 100,000 population. In addition, we calculated sex- and age-standardized AH rates, referencing the population structure of Beijing in 2010, to examine the consistency and stability of our findings.

Regarding independent variables, we also collected data at both the individual and district levels. First, we extracted demographic and clinical characteristics from hospital discharge records, including sex, age, marital status, health insurance types, and the Charlson Comorbidity Index (CCI) score. Second, we collected data on socioeconomic and health resources at the district level to examine the associated factors of AH rates. These variables included densities of hospital physicians (ie, the number of hospital physicians per 1000 population), primary physicians (ie, the number of primary physicians per 1000

XSL•FO
RenderX

population), proportion of males, proportion of the population aged older than 65 years, education level, and economic development, which were calculated based on information from the Beijing annual reports of medical institutions and Beijing statistical yearbooks. Finally, we obtained data on the daily confirmed COVID-19 cases in Beijing from the Chinese Center for Disease Control and Prevention to construct 2 dummy variables to indicate the presence of the COVID-19 pandemic in the current month or the previous month within each district of Beijing. Table S2 in Multimedia Appendix 1 shows the definitions of all variables used in this study.

Statistical Analysis

We first conducted descriptive analyses. Categorical variables are described with numbers and percentages, and continuous variables are described with means and SDs. We reported crude and standardized AH rates with 95% CIs for years between 2016 and 2021 and used chi-square tests to examine the statistical significance of differences between different years. In addition, we calculated AH rates across various districts in Beijing to show geographic variations. There are 16 administrative districts in Beijing, consisting of 6 urban districts located at the center of the city and 10 suburban districts. Compared with the suburban districts, the central districts have better socioeconomic developments, higher population density, and more abundant medical resources.

Furthermore, we performed linear regressions with robust SEs to explore the associated factors of AHs. First, we used district-level AH data to explore the association between AHs and the COVID-19 outbreak, health care resources, and socioeconomic characteristics. The model can be written as follows:

$$(1) Yit=\beta 0+\beta 1COVIDit+\beta 2Resourceit+\beta 3Controlsit+\lambda year+ it$$

where the subscript *i* denotes the district in Beijing and *t* denotes each month from January 2016 to December 2021. The dependent variable Y_{it} is the crude or standardized monthly AH rate for each district. The independent variable $COVID_{it}$ is a dummy variable that takes 1 if the district experienced 1 or more newly confirmed COVID-19 cases in the current month. Otherwise, it equals 0. $Resource_{it}$ denotes variables measuring the density of medical resources at the district level, including the number of hospital physicians per 1000 population and the number of primary physicians per 1000 population. Because of collinearity, other variables measuring the density of medical resources, such as hospital beds per 1000 population, are not included in the analysis. $Controls_{it}$ are a series of socioeconomic variables, including population structure, average education level, and economic developments. λ_{year} denotes the year fixed effects to control unobservable variables that change over time but are constant over districts. $_{it}$ denotes the robust error term.

In addition, we used individual-level data to perform logistic regressions to explore individual-level factors associated with AHs. The model can be written as follows:

$$(2) P(Yijst=1)=11+e-(\beta 0+\beta 1COVIDjst+\beta 2Xijst+\beta 3Hospitalj+\beta 4Districtst+\lambda t)$$

where the subscript *i* denotes the hospitalized case, *j* denotes the hospital, *s* denotes the district in Beijing, and *t* indicates the year. The dependent variable Y_{ijst} is a binary indicator that takes 1 if the hospitalized patient case *i* who resides in district *s* was admitted to hospital *j* in year *t* for ACSCs, and 0 if the admission was due to other conditions. The independent variable $COVID_{ijst}$ is a dummy variable that takes 1 if the district where the patient lived had 1 or more new confirmed COVID-19 cases within a period before the patient was admitted. Otherwise, it equals 0. X_{ijst} denotes a series of individual-level variables, including sex, age, marital status, CCI score, and health insurance types. $Hospital_{jt}$ denotes the ranking of the hospital, which is classified into 3 groups: primary, secondary, or tertiary hospitals. $District_{st}$ denotes variables measuring the density of medical resources in the district where the patient lived, including the number of hospital physicians per 1000 population, the number of primary physicians per 1000 population, and other socioeconomic variables. λ_t denotes the year fixed effects to control for unobservable variables that change over time but are constant over the patient. Robust SEs are clustered at the district level. All statistical analyses were carried out in Stata 17.0 (StataCorp LLC). All *P* values were from 2-sided tests and results were deemed statistically significant at $P<.1$.

Ethical Considerations

This study was approved as a secondary use of administrative data with a waiver of informed consent by the Peking University institutional review board. All data used in the study were deidentified prior to analysis to protect participants' privacy and confidentiality. The study was conducted in accordance with institutional ethical standards and followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guideline for cross-sectional studies.

Results

Summary Statistics

Table 1 shows the summary statistics for variables in the discharge data. From 2016 to 2019, the number of discharge cases for local inpatients aged older than 15 years increased from 1,931,230 to 2,303,308, while it decreased to 1,697,228 in 2020 and rebounded to 2,282,250 in 2021. The proportion of females was consistently higher than that of males and it decreased modestly over the years. The share of inpatients older than 60 years increased from 46.0% (887,617/1,931,230) in 2016 to 50.5% (1,151,718/2,282,250) in 2021. The share of inpatients without comorbidities as measured by the CCI (Table S3 in Multimedia Appendix 1) decreased by about 6 percentage points during this period. Patients receiving treatment at tertiary hospitals took a larger share after 2020, which implies a significant change in the patient composition during the COVID-19 pandemic. Additionally, these findings were in line with rising expenditures per admission after 2020, compared with those between 2016 and 2019.

Tables S4 and S5 in Multimedia Appendix 1 show the summary statistics for AHs and non-AHs, respectively. Similar to the trends in all inpatient cases, both the numbers of AHs and non-AHs experienced large declines during the pandemic. However, the reduction in AHs (38.4%, 54,499/141,811) in 2020 was markedly greater than that in non-AHs (25.5%,

XSL•FO
RenderX

555,581/2,161,497). More than half of AHs were males, contrary to the sex share among non-AHs. The proportion of AHs aged older than 60 years was consistently larger than that of non-AHs during our study period. Most AHs had comorbidities and the group of AHs had much higher CCI scores than non-AHs. However, the average expenditures for AHs were consistently lower than those for non-AHs.

XSL•FO
RenderX

Table . Summary statistics of all study cases[a].

Variables	2016	2017	2018	2019	2020	2021
Total study cases	1,931,230	2,010,201	2,143,498	2,303,308	1,697,228	2,282,250
AHs[b], n (%)	118,764 (6.1)	125,712 (6.3)	134,928 (6.3)	141,811 (6.2)	87,312 (5.1)	124,277 (5.4)
Non-AHs[b], n (%)	1,812,466 (93.9)	1,884,489 (93.7)	2,008,570 (93.7)	2,161,497 (93.8)	1,609,916 (94.9)	2,157,973 (94.6)
Sex, n (%)						
Male	804,176 (41.6)	857,130 (42.6)	928,838 (43.3)	1,001,061 (43.5)	743,233 (43.8)	1,022,609 (44.8)
Female	1,127,054 (58.4)	1,153,071 (57.4)	1,214,660 (56.7)	1,302,247 (56.5)	953,995 (56.2)	1,259,641 (55.2)
Age (years), n (%)						
15 - 39	552,136 (28.6)	535,717 (26.7)	532,211 (24.8)	556,681 (24.2)	417,628 (24.6)	506,023 (22.2)
40 - 59	491,477 (25.5)	515,863 (25.7)	550,768 (25.7)	584,500 (25.4)	437,957 (25.8)	624,509 (27.4)
60 - 79	647,796 (33.5)	697,979 (34.7)	775,375 (36.2)	856,302 (37.2)	639,299 (37.7)	899,392 (39.4)
≥80	239,821 (12.4)	260,642 (13.0)	285,144 (13.3)	305,825 (13.3)	202,344 (11.9)	252,326 (11.1)
Marital status, n (%)						
Unmarried[c]	194,987 (10.1)	197,506 (9.8)	224,798 (10.5)	259,816 (11.3)	203,233 (12.0)	296,146 (13.0)
Married	1,736,243 (89.9)	1,812,695 (90.2)	1,918,700 (89.5)	2,043,492 (88.7)	1,493,995 (88.0)	1,986,104 (87.0)
CCI score[d], n (%)						
0	1,019,696 (52.8)	1,028,803 (51.2)	1,054,523 (49.2)	1,110,311 (48.2)	810,073 (47.7)	1,067,047 (46.8)
1 - 2	494,200 (25.6)	524,864 (26.1)	574,434 (26.8)	620,420 (26.9)	446,545 (26.3)	609,370 (26.7)
3 - 4	262,134 (13.6)	282,375 (14.1)	314,127 (14.7)	343,646 (14.9)	245,645 (14.5)	337,759 (14.8)
≥5	155,200 (8.0)	174,159 (8.7)	200,414 (9.4)	228,931 (9.9)	194,965 (11.5)	268,074 (11.8)
Health insurance types, n (%)						
UEBMI[e]	1,130,824 (58.6)	1,214,368 (60.4)	1,339,584 (62.5)	1,484,728 (64.5)	1,112,266 (65.5)	1,547,479 (67.8)
URBMI[f] or NCRMS[g]	314,991 (16.3)	326,698 (16.3)	314,739 (14.7)	319,374 (13.9)	228,482 (13.5)	312,375 (13.7)
Others[h]	485,415 (25.1)	469,135 (23.3)	489,175 (22.8)	499,206 (21.7)	356,480 (21.0)	422,396 (18.5)
Hospital ranking, n (%)						
Primary	9297 (0.5)	13,231 (0.7)	10,220 (0.5)	7547 (0.3)	3988 (0.2)	3906 (0.2)
Secondary	516,936 (26.8)	503,479 (25.1)	521,391 (24.3)	515,976 (22.4)	375,445 (22.1)	308,888 (13.5)
Tertiary	1,404,997 (72.8)	1,493,491 (74.3)	1,611,887 (75.2)	1,779,785 (77.3)	1,317,795 (77.6)	1,969,456 (86.3)
Expenditure, mean (SD), ¥[i]	17,469 (27,212)	18,192 (28,105)	18,756 (29,171)	18,944 (28,808)	22,060 (34,955)	21,015 (33,301)
LoS[j] (day), mean (SD)	10.8 (47.8)	11.2 (84.7)	10.5 (41.6)	9.6 (30.9)	10.1 (36.6)	9.2 (44.0)

[a]Categorical variables are described with numbers and percentages (%), and continuous variables are described with means and SDs.

[b]AHs: avoidable hospitalizations.

[c]Unmarried status includes never married, widowed, and divorced.

[d]CCI score: the Charlson Comorbidity Index score. It was included to measure the severity of comorbidities, which was calculated for each case based on the information from the hospital discharge data (Table S3 in Multimedia Appendix 1). It was further categorized into 4 subgroups (0, 1 - 2, 3 - 4, and ≥5) [44].

[e]UEBMI: urban employee basic medical insurance.

[f]URBMI: urban resident basic medical insurance.

[g]NCRMS: new rural cooperative medical system. URBMI and NCRMS, mainly provided for minors and unemployed persons, were put together for

XSL•FO
RenderX

analysis because they had been merged since 2018.

[h]Other types of health insurance schemes include commercial medical insurance, other social insurance, poverty relief, public free health, and without insurance coverage.

[i]All expenditures are reported in Chinese yuan (CNY). The average exchange rate in 2021 was 6.4515 CNY to 1 USD.

[j]LoS: length of stay.

Trends in AHs

The first row of Table 2 shows crude changes in overall AHs in Beijing between 2016 and 2021, suggesting that the COVID-19 pandemic had significant impacts on AHs. Before the pandemic (2016-2019), the overall crude AH rate per 100,000 population exhibited a consistent upward trend, rising annually from the lowest of 633.4 (95% CI 629.8 - 637.0) in 2016 to the highest of 740.0 (95% CI 736.2 - 743.8) in 2019. The difference in the crude AH rate between 2016 and 2019 was statistically significant (difference=106.6; P<.001). However, during the pandemic, the crude AH rate experienced a sharp decline to 452.4 (95% CI 449.4 - 455.4) in 2020 and rebounded to 639.3 (95% CI 635.8 - 642.9) in 2021, but the crude AH rate in 2021 was still lower than the average level in the prepandemic period. Sex- and age-standardized AH rates followed a similar trend to crude AH rates before and during the pandemic. As for regional differences, Figure S3 and Table S6 in Multimedia Appendix 1 report AH rates at the district level from 2016 to 2021. Overall, suburban districts consistently have higher AH rates than central districts, showing similar trends. Figure S4 in Multimedia Appendix 1 also illustrates trends in mean expenditure and length of stay for avoidable hospitalizations from 2016 to 2021.

XSL•FO
RenderX

Table . Trends of avoidable hospitalization rates in Beijing, 2016 - 2021.

Variable	AH[a] rates per 100,000 population (95% CI)						2016 versus 2019		2019 versus 2021		2016 versus 2021	
	2016	2017	2018	2019	2020	2021	Difference[b]	*P* value	Difference	*P* value	Difference	*P* value
Crude rates												
Total AHs	633.4 (629.8 - 637.0)	665.5 (661.9 - 669.2)	709.2 (705.4 - 712.9)	740 (736.2 - 743.8)	452.4 (449.4 - 455.4)	639.3 (635.8 - 642.9)	106.6	<.001	–100.7	<.001	5.9	.02
Hypertension	119 (117.4 - 120.5)	126.5 (124.8 - 128.1)	132.9 (131.3 - 134.6)	138.6 (137.0 - 140.3)	77.9 (76.6 - 79.1)	120.1 (118.5 - 121.6)	19.7	<.001	–18.6	<.001	1.1	.32
Diabetes	301.8 (299.3 - 304.3)	319.5 (316.9 - 322.0)	348 (345.4 - 350.7)	365.4 (362.7 - 368.1)	222.3 (220.2 - 224.4)	326 (323.5 - 328.5)	63.5	<.001	–39.4	<.001	24.2	<.001
Asthma	28.2 (27.4 - 29.0)	26.5 (25.7 - 27.2)	26.1 (25.4 - 26.8)	25 (24.3 - 25.7)	14.9 (14.4 - 15.4)	19.7 (19.1 - 20.4)	–3.2	<.001	–5.3	<.001	–8.5	<.001
COPD[c]	111.1 (109.6 - 112.6)	112.3 (110.8 - 113.8)	108.7 (107.2 - 110.2)	111.8 (110.3 - 113.3)	63 (61.9 - 64.1)	76.2 (75.0 - 77.4)	0.7	.50	–35.6	<.001	–34.9	<.001
CHF[d]	73.3 (72.1 - 74.5)	80.9 (79.6 - 82.2)	93.4 (92.0 - 94.8)	99.1 (97.7 - 100.6)	74.3 (73.1 - 75.5)	97.3 (95.9 - 98.7)	25.9	<.001	–1.8	.07	24.0	<.001
Standardized rates[e]												
Total AHs	514.7 (511.4 - 517.9)	524.7 (521.5 - 528.0)	542.2 (538.9 - 545.5)	552.8 (549.4 - 556.1)	331.2 (328.6 - 333.8)	465.1 (462.1 - 468.1)	38.1	<.001	–87.7	<.001	–49.6	<.001
Hypertension	97.9 (96.5 - 99.3)	101.2 (99.8 - 102.6)	103.7 (102.2 - 105.1)	106 (104.6 - 107.5)	58.3 (57.2 - 59.4)	90.2 (88.9 - 91.6)	8.1	<.001	–15.8	<.001	–7.7	<.001
Diabetes	254.3 (252.0 - 256.6)	262.2 (259.9 - 264.5)	277.8 (275.4 - 280.1)	286.3 (283.9 - 288.7)	172.2 (170.3 - 174.0)	249.6 (247.4 - 251.9)	32.0	<.001	–36.7	<.001	–4.7	.004
Asthma	23.3 (22.6 - 24.0)	21.3 (20.7 - 22.0)	20.5 (19.8 - 21.1)	19.3 (18.7 - 20.0)	11.4 (10.9 - 11.8)	14.9 (14.4 - 15.4)	–4.0	<.001	–4.4	<.001	–8.4	<.001
COPD	84.2 (82.9 - 85.6)	82.2 (80.9 - 83.5)	76.5 (75.3 - 77.8)	75.8 (74.5 - 77.0)	41.3 (40.4 - 42.2)	49.4 (48.4 - 50.3)	–8.5	<.001	–26.4	<.001	–34.9	<.001
CHF	54.9 (53.8 - 55.9)	57.9 (56.8 - 58.9)	63.8 (62.6 - 64.9)	65.3 (64.2 - 66.5)	48 (47.0 - 49.0)	61 (59.9 - 62.1)	10.5	<.001	–4.4	<.001	6.1	<.001

[a]AH: avoidable hospitalization.

[b]Chi-square tests were used to test the statistical significance of differences between different survey years.

[c]COPD: chronic obstructive pulmonary disease.

[d]CHF: congestive heart failure.

[e]Standardized AH rates were standardized by the sex and age population structure of Beijing in 2010.

Impact of the COVID-19 Pandemic and Associated Factors of AHs

Table 3 reports linear regression results that investigated the impact of the COVID-19 pandemic on the monthly AH rates for each district as well as associated factors of AH rates at the district level. In these regression analyses, we included the key variables $COVID_{this\ month}$ and $COVID_{last\ month}$, representing the existence of newly confirmed COVID-19 cases in the district for the current month and the previous month, respectively. We also included variables related to medical resources, population structure, and other control variables. In column 1 of Table 3, results show that the existence of the COVID-19 pandemic in the current month was significantly associated with a decrease in the crude AH rate for that month. Considering that the existence of the COVID-19 pandemic in the previous month

XSL•FO
RenderX

might influence the change in AHs in the current month, we further included $COVID_{last\ month}$ to examine the impact of the pandemic on AHs. In column 2, it is evident that the existence of the COVID-19 pandemic was correlated with a significant decrease in the AH rate in the following month. The magnitude of change caused by the pandemic in the last month was larger than that in the current month. Regarding district-level medical resources, a higher density of hospital physicians was positively associated with AH rates. In contrast, a higher density of primary physicians was associated with lower AH rates. To account for demographic changes in Beijing, we changed the dependent variable from the crude AH rate to the standardized AH rate by the Beijing reference population. As shown in columns 3 and 4 of Table 3, the coefficients for various variables remained largely unchanged when using the standardized AH rate as the dependent variable. When we controlled for month-year fixed effects, our results remained largely robust, as shown in Table S8 in Multimedia Appendix 1.

We performed logistic regression analyses using individual-level data to further test the robustness of our findings. As shown in the first 2 columns of Table S7 in Multimedia Appendix 1, the presence of newly confirmed COVID-19 cases within 14 days or 1 month before hospital admission, as well as a greater number of primary care physicians in the patient's residential district, was associated with a decreased likelihood of the hospitalization being classified as an AH. These results are consistent with the findings based on district-level data. In terms of individual-level factors, hospitalized patients who were male, older, and had comorbid conditions had a higher probability of being classified as AH cases than their female, younger, and noncomorbid counterparts. Inpatients covered by urban resident basic medical insurance or new rural cooperative medical scheme were more likely to be AHs than those covered by urban employee basic medical insurance. Additionally, patients admitted to primary hospitals were more likely to be AHs than those treated in secondary or tertiary hospitals. Table S7 in Multimedia Appendix 1 also shows regression results using expenditure and length of stay as dependent variables to examine potential changes in case composition.

XSL•FO
RenderX

Table . The impact of the COVID-19 pandemic on avoidable hospitalization.

Variables	Crude AH[a] rate[b]				Standardized AH rate[c]			
	Model 1		Model 2		Model 3		Model 4	
	β (RSE)[d]	*P* value	β (RSE)	*P* value	β (RSE)	*P* value	β (RSE)	*P* value
$COVID_{this\ month}$[e] (reference no.)								
Yes	−7.752 (3.1325)	.01	−4.234 (2.6500)	.11	−6.928 (2.1914)	.002	−4.151 (1.8445)	.03
$COVID_{last\ month}$[e] (reference no.)								
Yes	—[f]	—	−18.873 (2.7642)	<.001	—	—	−14.452 (1.9310)	<.001
Medical resource								
Hospital physicians/1000	2.792 (0.5864)	<.001	2.718 (0.5747)	<.001	3.723 (0.4351)	<.001	3.589 (0.4311)	<.001
Primary physicians/1000	−14.900 (3.3889)	<.001	−13.769 (3.3667)	<.001	−8.638 (2.7418)	.002	−8.153 (2.7280)	.003
Population structure								
Proportion of males	−8.081 (1.0309)	<.001	−8.530 (1.0156)	<.001	—	—	—	—
Proportion of those aged ≥65 years	1.788 (0.6637)	.007	1.401 (0.6516)	.03	—	—	—	—
Mean educational years	−3.082 (1.7056)	.07	−2.940 (1.6636)	.08	−0.890 (1.1810)	.45	−0.461 (1.1639)	.69
$GDP_{district}$	−0.004 (0.0005)	<.001	−0.004 (0.0005)	<.001	−0.002 (0.0004)	<.001	−0.002 (0.0004)	<.001
Constant	505.081 (75.5535)	<.001	531.931 (74.0819)	<.001	59.399 (15.3988)	<.001	54.670 (15.1962)	<.001
Year fixed effect	Yes	—	Yes	—	Yes	—	Yes	—
Observations	1152	—	1152	—	1152	—	1152	—

XSL•FO
RenderX

[a]AH: avoidable hospitalization.

[b]Crude AH rate in models (1) and (2) indicates the monthly crude AH rate of each district.

[c]Standardized AH rate in models (3) and (4) indicates the standardized monthly AH rate of each district according to the population structure of Beijing in 2010.

[d]RSE: robust standard error.

[e]Key variables $COVID_{this\ month}$ and $COVID_{last\ month}$ indicate whether new COVID-19 cases were recognized in the district this or last month.

[f]Not included in this model.

Discussion

Principal Findings

Using discharge data from Beijing, this study found that both the crude and the sex- and age-standardized AH rates in Beijing displayed an upward trend before the COVID-19 pandemic (2016 - 2019) and experienced a sharp decline after the outbreak of the pandemic. This trend suggests substantial impacts of the COVID-19 pandemic on health care usage among residents in Beijing. Moreover, regression results showed that AHs were significantly associated with multiple factors, including pandemic shocks, densities of health care resources, socioeconomic variables, and various demographic characteristics.

We found that both the crude and the sex- and age-standardized AH rates in Beijing displayed an upward trend before the COVID-19 pandemic (2016 - 2019) and experienced a sharp decline after the outbreak of the pandemic. This trend suggests substantial impacts of the COVID-19 pandemic on health care usage among residents in Beijing.

Based on the results regarding trends in AHs before and during the COVID-19 pandemic, we further compared the AH rate in Beijing with that in some high-income countries, including the United States, the United Kingdom, France, Canada, and Korea, by calculating sex- and age-standardized AH rates based on the OECD 2015 reference population [45]. As shown in Table 4, the overall sex- and age-standardized AH rate per 100,000 population in Beijing was 904.1 in 2019, higher than that in the United Kingdom (530.5 in 2019), France (636.0 in 2015), Canada (537.3 in 2019), and Korea (641.0 in 2019), but slightly lower than that in the United States (961.0 in 2019). This comparison suggests that it is necessary for Beijing to strengthen its primary care system and restructure its hospital-centered delivery system.

In terms of specific diseases, the sex- and age-standardized AH rates for hypertension and diabetes in Beijing were much higher than those in all comparison countries. That finding may be due to the relatively low awareness and control rates of these conditions among the Chinese population [29], which often results in delayed diagnosis and more severe symptoms requiring hospitalization. In addition, medical insurance policies in China provide higher reimbursement rates for inpatient care than for outpatient care, which may incentivize hospitalization over outpatient care. By contrast, the standardized AH rates for asthma, COPD, and CHF in Beijing were at a relatively low level, compared with those in other countries. Given the high proportion of hypertension and diabetes among all AHs in Beijing, more attention should be paid to the prevention and control of hypertension and diabetes.

Moreover, this study showed that relatively high-income districts with abundant primary care medical resources generally had lower AH rates than suburban districts. This was consistent with the regression result that a higher density of primary physicians was associated with a lower AH rate. It also implied that the AH rate could at least partially reflect the accessibility to primary care. In addition, we found that the AH rate was significantly associated with demographic and socioeconomic variables, such as population structure and economic development. These supported the previous finding that AHs should be interpreted in conjunction with metrics of demographic composition and the level of economic development [39,46].

More importantly, we found that the COVID-19 pandemic was associated with the reduction of the AH rates in Beijing. This association remained robust across alternative regression specifications, including using various definitions of the COVID-19 pandemic, changing the dependent variable from crude rates to standardized rates, and using data at the individual level. These results highlight that we cannot interpret solely the reduction in AH rates as the result of the changing performance of primary health care, as the pandemic has caused many changes in health care providers' and patients' behaviors that could be related to the AH rates. During the COVID-19 pandemic, many countries took large-scale public health measures such as stay-at-home orders, closing schools, cancellation of mass gatherings, and even lockdowns of endemic areas [20,47-50]. In Beijing, despite low COVID-19 mortality and a limited number of confirmed cases, the government implemented a stringent "zero-COVID" policy, including movement restrictions, community lockdowns, testing mandates, and hospital visitor bans. These measures created substantial barriers to hospital access for patients with no COVID-19 cases. Compared to non-AHs, AH patients typically present with milder symptoms and some of them seek hospitalization primarily because the reimbursement rate for inpatient care is higher than for outpatient care in the nonpandemic period. As a result, during the pandemic period, AH patients became more sensitive to the demand-side factors (eg, patients with ACSC avoiding hospitals unless their symptoms became severe) and the supply-side factors (eg, providers restricting admissions for mild cases to prevent in-hospital viral transmission), which collectively contributed to the larger reductions in AHs.

Taken together, our findings highlight the complexity of using AH as a proxy for access to primary care, especially when there are systemwide changes. The determinants of AHs are multifactorial, including factors at the patient level, regional level, and national system level. These factors may not be independent but intertwined, especially when there are changes that simultaneously occur in both the demand and supply for

XSL•FO
RenderX

medical care. As a result, when we observe a change in AH rates, it may be inappropriate and inaccurate to interpret them solely as the consequence of changes in the accessibility and quality of primary health care or outpatient care.

Table . International comparisons of sex- and age-standardized avoidable hospitalization rate[a].

Variables	Beijing (2019)	United States[b] (2019)	United Kingdom[b] (2019)	France[b] (2015)	Canada[b] (2019)	Korea[b] (2019)
Total AHs[c]	904.1	961.0	530.5	636.0	537.3	641.0
Hypertension	167.5	60.6	18.7	34.8	14.6	79.2
Diabetes	405.1	248.5	81.7	155.6	98.9	237.3
Asthma	29.4	37.3	74.9	29.8	13.9	68.5
COPD[d]	153.5	186.8	239.2	128.8	229.5	161.5
CHF[e]	148.5	427.8	116.0	287.0	180.4	94.5

[a]All the sex- and age-standardized AH rates per 100,000 population were standardized by the OECD (Organisation for Economic Co-operation and Development) 2015 reference population.

[b]These countries' sex- and age-standardized AH rates were obtained from the OECD website.

[c]AHs: avoidable hospitalizations.

[d]COPD: chronic obstructive pulmonary disease.

[e]CHF: congestive heart failure.

Limitations

There are several limitations to this study. First, we can only roughly measure the impact of primary health care access but not directly measure the effect of primary health care quality on AHs due to data availability. Second, we cannot evaluate the gains in social welfare from reductions in AHs during the pandemic. We also cannot clearly identify the mechanisms through which the pandemic significantly reduced the AH rate. Third, we used the definition of ACSCs proposed by OECD, which might not fully capture China's epidemiological features and health system characteristics [51]. Finally, due to the limited time span of the data, we were unable to explore long-term changes in AH rates after the COVID-19 pandemic, which may help us better understand the long-lasting effects of pandemic-related disruptions in health care usage. These limitations should be addressed in future research.

Conclusions

Using administrative discharge data, we found that Beijing had a relatively high AH rate compared with other high-income countries. Trend analysis showed that the overall AH rate in Beijing displayed an upward trend between 2016 and 2019 but experienced a sharp decrease in 2020. The increasing AH rates in Beijing before the pandemic highlight the necessity to strengthen its primary care system and restructure its hospital-centered delivery system in Beijing. Potential strategies include improving the capacity of primary care physicians, creating incentives for health care providers to provide preventive and primary care, and expanding the benefit coverage of health insurance schemes for outpatient services. Regression results showed that AHs were significantly associated with multiple factors, including the COVID-19 pandemic shocks, the regional density of health care resources, and demographic variables. These results imply that changes in AHs cannot solely be interpreted as changes in access and quality of primary care. Appropriate controls accounting for the multifactorial drivers of all hospital admissions are needed when we use this indicator as a tool to monitor health care system performance, especially when there are system-wide changes such as the COVID-19 pandemic.

Acknowledgments

The authors thank all seminar participants at Peking University for helpful comments. This study was financially supported by the Beijing Social Science Fund (grant 20GLC036).

Data Availability

The dataset used in this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional statistical analyses, data sources, and inclusion and exclusion criteria for study cases.

XSL•FO
RenderX

[DOCX File, 2902 KB - publichealth_v11i1e69768_app1.docx]

References

1. Hone T, Macinko J, Millett C. Revisiting Alma-Ata: what is the role of primary health care in achieving the sustainable development goals? Lancet 2018 Oct 20;392(10156):1461-1472. [doi: 10.1016/S0140-6736(18)31829-4] [Medline: 30343860]
2. Pappas G, Hadden WC, Kozak LJ, Fisher GF. Potentially avoidable hospitalizations: inequalities in rates between US socioeconomic groups. Am J Public Health 1997 May;87(5):811-816. [doi: 10.2105/ajph.87.5.811] [Medline: 9184511]
3. Rosano A, Loha CA, Falvo R, et al. The relationship between avoidable hospitalization and accessibility to primary care: a systematic review. Eur J Public Health 2013 Jun;23(3):356-360. [doi: 10.1093/eurpub/cks053] [Medline: 22645236]
4. Figueroa JF, Burke LG, Horneffer KE, Zheng J, John Orav E, Jha AK. Avoidable hospitalizations and observation stays: shifts in racial disparities. Health Aff (Millwood) 2020 Jun;39(6):1065-1071. [doi: 10.1377/hlthaff.2019.01019] [Medline: 32479235]
5. Kozak LJ, Hall MJ, Owings MF. Trends in avoidable hospitalizations, 1980-1998. Health Aff (Millwood) 2001;20(2):225-232. [doi: 10.1377/hlthaff.20.2.225] [Medline: 11260947]
6. Thygesen LC, Christiansen T, Garcia-Armesto S, et al. Potentially avoidable hospitalizations in five European countries in 2009 and time trends from 2002 to 2009 based on administrative data. Eur J Public Health 2015 Feb;25 Suppl 1:35-43. [doi: 10.1093/eurpub/cku227] [Medline: 25690128]
7. Mercier G, Georgescu V, Bousquet J. Geographic variation in potentially avoidable hospitalizations in France. Health Aff (Millwood) 2015 May;34(5):836-843. [doi: 10.1377/hlthaff.2014.1065] [Medline: 25941286]
8. Hutchison A, Ambrose S, Glover J, Hetzel D. Atlas of Avoidable Hospitalisations in Australia: Ambulatory Care-Sensitive Conditions: Commonwealth of Australia: Public Health Information Development Unit; 2007.
9. Chun SY, Kim W, Park EC. Disparities in avoidable hospitalization by income in South Korea: data from the National Health Insurance cohort. Eur J Public Health 2019 Apr 1;29(2):225-231. [doi: 10.1093/eurpub/cky198] [Medline: 30260368]
10. Allegri C, Banks H, Devillanova C. Avoidable hospitalizations and access to primary care: comparisons among Italians, resident immigrants and undocumented immigrants in administrative hospital discharge records. EClinicalMedicine 2022 Apr;46:101345. [doi: 10.1016/j.eclinm.2022.101345] [Medline: 35295899]
11. Daly MR, Mellor JM, Millones M. Do avoidable hospitalization rates among older adults differ by geographic access to primary care physicians? Health Serv Res 2018 Aug;53(Suppl 1):3245-3264. [doi: 10.1111/1475-6773.12736] [Medline: 28660679]
12. Bindman AB, Chattopadhyay A, Osmond DH, Huen W, Bacchetti P. The impact of Medicaid managed care on hospitalizations for ambulatory care sensitive conditions. Health Serv Res 2005 Feb;40(1):19-38. [doi: 10.1111/j.1475-6773.2005.00340.x] [Medline: 15663700]
13. Chen S, Fu H, Jian W. Trends in avoidable hospitalizations in a developed City in eastern China: 2015 to 2018. BMC Health Serv Res 2022 Jul 4;22(1):856. [doi: 10.1186/s12913-022-08275-w] [Medline: 35788227]
14. Poblano Verástegui O, Torres-Arreola LDP, Flores-Hernández S, Nevarez Sida A, Saturno Hernández PJ. Avoidable hospitalization trends from ambulatory care-sensitive conditions in the public health system in México. Front Public Health 2021;9:765318. [doi: 10.3389/fpubh.2021.765318] [Medline: 35127618]
15. Alexander GC, Tajanlangit M, Heyward J, Mansour O, Qato DM, Stafford RS. Use and content of primary care office-based vs telemedicine care visits during the COVID-19 pandemic in the US. JAMA Netw Open 2020 Oct 1;3(10):e2021476. [doi: 10.1001/jamanetworkopen.2020.21476] [Medline: 33006622]
16. Cantor J, Sood N, Bravata DM, Pera M, Whaley C. The impact of the COVID-19 pandemic and policy response on health care utilization: evidence from county-level medical claims and cellphone data. J Health Econ 2022 Mar;82:102581. [doi: 10.1016/j.jhealeco.2022.102581] [Medline: 35067386]
17. Birkmeyer JD, Barnato A, Birkmeyer N, Bessler R, Skinner J. The impact of the COVID-19 pandemic on hospital admissions in the United States. Health Aff (Millwood) 2020 Nov;39(11):2010-2017. [doi: 10.1377/hlthaff.2020.00980] [Medline: 32970495]
18. Xiao H, Liu F, He Y, et al. Unequal impact of the COVID-19 pandemic on paediatric cancer care: a population-based cohort study in China. Lancet Reg Health West Pac 2022 Feb;19:100347. [doi: 10.1016/j.lanwpc.2021.100347] [Medline: 35005670]
19. Sevalie S, Youkee D, van Duinen AJ, et al. The impact of the COVID-19 pandemic on hospital utilisation in Sierra Leone. BMJ Glob Health 2021 Oct;6(10):34635552. [doi: 10.1136/bmjgh-2021-005988] [Medline: 34635552]
20. He L, Lu F, Du X, et al. Impact of COVID-19 pandemic on hospital admissions of acute coronary syndrome: a Beijing inpatient database study. Lancet Reg Health West Pac 2022 Feb;19:100335. [doi: 10.1016/j.lanwpc.2021.100335] [Medline: 34927111]
21. Rennert-May E, Leal J, Thanh NX, et al. The impact of COVID-19 on hospital admissions and emergency department visits: a population-based study. PLoS One 2021;16(6):e0252441. [doi: 10.1371/journal.pone.0252441] [Medline: 34061888]
22. Leuchter RK, Villaflores CWA, Norris KC, Sorensen A, Vangala S, Sarkisian CA. Racial disparities in potentially avoidable hospitalizations during the COVID-19 pandemic. Am J Prev Med 2021 Aug;61(2):235-239. [doi: 10.1016/j.amepre.2021.01.036] [Medline: 33820665]

XSL•FO
RenderX

23. Kaneko M, Shimizu S, Oishi A, Fushimi K. Impact of COVID-19 infection rates on admissions for ambulatory care sensitive conditions: nationwide difference-in-difference design in Japan. Fam Med Community Health 2022 Oct;10(4):36241252. [doi: 10.1136/fmch-2022-001736] [Medline: 36241252]
24. Becker NV, Karmakar M, Tipirneni R, Ayanian JZ. Trends in hospitalizations for ambulatory care-sensitive conditions during the COVID-19 pandemic. JAMA Netw Open 2022 Mar 1;5(3):e222933. [doi: 10.1001/jamanetworkopen.2022.2933] [Medline: 35297972]
25. China health statistics yearbook 2020. China NHCotPsRo. 2021. URL: http://www.nhc.gov.cn/mohwsbwstjxxzx/tjtjnj/202112/982dd4fa7f9f4b7f8ba9b752aa8aa2e3/files/1740021960667_53749.pdf [accessed 2025-06-25]
26. OECD. Health at a Glance 2023: OECD Indicators. : OECD Publishing; 2023. [doi: 10.1787/7a7afb35-en]
27. Yip W, Fu H, Jian W, et al. Universal health coverage in China part 1: progress and gaps. Lancet Public Health 2023 Dec;8(12):e1025-e1034. [doi: 10.1016/S2468-2667(23)00254-2] [Medline: 38000882]
28. Yip W, Fu H, Chen AT, et al. 10 years of health-care reform in China: progress and gaps in universal health coverage. Lancet 2019 Sep 28;394(10204):1192-1204. [doi: 10.1016/S0140-6736(19)32136-1] [Medline: 31571602]
29. Li Y, Zhang C, Zhan P, Fu H, Yip W. Trends and projections of universal health coverage indicators in China, 1993-2030: an analysis of data from four nationwide household surveys. Lancet Reg Health West Pac 2023 Feb;31:100646. [doi: 10.1016/j.lanwpc.2022.100646] [Medline: 36419465]
30. Gao L, Shi L, Meng Q, Kong X, Guo M, Lu F. Effect of healthcare system reforms on public hospitals' revenue structures: evidence from Beijing, China. Soc Sci Med 2021 Aug;283:114210. [doi: 10.1016/j.socscimed.2021.114210] [Medline: 34274783]
31. 2019 Beijing health development statistical bulletin. Commission BMH. 2020. URL: http://www.phic.org.cn/tjsj/wstjgb/202408/P020240819587365019195.pdf [accessed 2025-05-28]
32. 2020 Beijing health development statistical bulletin. Commission BMH. 2021. URL: http://www.phic.org.cn/tjsj/wstjgb/202408/P020240819587205272843.pdf [accessed 2025-05-28]
33. Chen Q, Rodewald L, Lai S, Gao GF. Rapid and sustained containment of covid-19 is achievable and worthwhile: implications for pandemic response. BMJ 2021 Dec 1;375:e066169. [doi: 10.1136/BMJ-2021-066169] [Medline: 34852997]
34. Rizvi SKA, Mirza N, Naqvi B, Rahat B. Covid-19 and asset management in EU: a preliminary assessment of performance and investment styles. J Asset Manag 2020 Jul;21(4):281-291. [doi: 10.1057/s41260-020-00172-3]
35. Chang HH, Meyerhoefer CD. COVID‐19 and the demand for online food shopping services: empirical evidence from Taiwan. Am J Agri Econ 2021 Mar;103(2):448-465. [doi: 10.1111/ajae.12170]
36. Hu J, Xu H, Yao Y, Zheng L. Is Working from home here to stay? Evidence from job posting data after the Covid-19 Shock. 2021. URL: https://iceanet.org/wp-content/uploads/2021/11/Hu.pdf [accessed 2025-05-28]
37. Huang F, Liu H. The impact of the COVID-19 pandemic and related policy responses on non-COVID-19 healthcare utilization in China. Health Econ 2023 Mar;32(3):620-638. [doi: 10.1002/hec.4636] [Medline: 36397307]
38. Lynch B, Fitzgerald AP, Corcoran P, Buckley C, Healy O, Browne J. Drivers of potentially avoidable emergency admissions in Ireland: an ecological analysis. BMJ Qual Saf 2019 Jun;28(6):438-448. [doi: 10.1136/bmjqs-2018-008002] [Medline: 30314977]
39. Hodgson K, Deeny SR, Steventon A. Ambulatory care-sensitive conditions: their potential uses and limitations. BMJ Qual Saf 2019 Jun;28(6):429-433. [doi: 10.1136/bmjqs-2018-008820] [Medline: 30819838]
40. Lai Y, Fu H, Li L, Yip W. Hospital response to a case-based payment scheme under regional global budget: the case of Guangzhou in China. Soc Sci Med 2022 Jan;292:114601. [doi: 10.1016/j.socscimed.2021.114601] [Medline: 34844079]
41. Healthcare quality and outcomes (HCQO) indicators 2022-23 definitions. OECD. 2024. URL: https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/measuring-health-care-quality/definitions-of-health-care-quality-outcomes.pdf [accessed 2025-06-24]
42. Quan J, Zhang H, Pang D, et al. Avoidable hospital admissions from diabetes complications In Japan, Singapore, Hong Kong, and communities outside Beijing. Health Aff (Millwood) 2017 Nov;36(11):1896-1903. [doi: 10.1377/hlthaff.2017.0479] [Medline: 29137504]
43. Chen T, Pan J. The effect of spatial access to primary care on potentially avoidable hospitalizations of the elderly: evidence from Chishui City, China. Soc Indicator Res 2022 Apr;160(2-3):645-665. [doi: 10.1007/s11205-020-02413-9] [Medline: 35359349]
44. Brusselaers N, Lagergren J. The Charlson Comorbidity Index in registry-based research. Methods Inf Med 2017;56(5):401-406. [doi: 10.3414/ME17-01-0051] [Medline: 29582935]
45. Health care quality indicators. OECD. 2023. URL: https://stats.oecd.org/Index.aspx?DataSetCode=HEALTH_HCQI [accessed 2025-05-28]
46. Saxena S, George J, Barber J, Fitzpatrick J, Majeed A. Association of population and practice factors with potentially avoidable admission rates for chronic diseases in London: cross sectional analysis. J R Soc Med 2006 Feb;99(2):81-89. [doi: 10.1177/014107680609900221] [Medline: 16449782]
47. Kranjac AW, Kranjac D. County-level factors that influenced the trajectory of COVID-19 incidence in the New York City area. Health Secur 2021 Jun;19(S1):S27-S33. [doi: 10.1089/hs.2020.0236] [Medline: 33956531]

XSL•FO
RenderX

48. Berlemann M, Haustein E. Right and yet wrong: A spatio-temporal evaluation of Germany's COVID-19 containment policy. SSRN J 2020. [doi: 10.2139/ssrn.3662054]
49. Davies NG, Barnard RC, Jarvis CI, et al. Association of tiered restrictions and a second lockdown with COVID-19 deaths and hospital admissions in England: a modelling study. Lancet Infect Dis 2021 Apr;21(4):482-492. [doi: 10.1016/S1473-3099(20)30984-1] [Medline: 33357518]
50. Li Z, Chen Q, Feng L, et al. Active case finding with case management: the key to tackling the COVID-19 pandemic. Lancet 2020 Jul 4;396(10243):63-70. [doi: 10.1016/S0140-6736(20)31278-2] [Medline: 32505220]
51. Wang J, Xu DR, Zhang Y, et al. Development of the China's list of ambulatory care sensitive conditions (ACSCs): a study protocol. Glob Health Res Policy 2024 Mar 19;9(1):11. [doi: 10.1186/s41256-024-00350-5] [Medline: 38504369]

Abbreviations

ACSC: ambulatory care sensitive condition
AH: avoidable hospitalization
CCI: Charlson Comorbidity Index
CHF: congestive heart failure
COPD: chronic obstructive pulmonary disease
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
***ICD-10-CM*:** *International Classification of Diseases, Tenth Revision, Clinical Modification*
***ICD-9-CM-3*:** Volume 3 of the *International Classification of Diseases, Ninth Revision, Clinical Modification*
OECD: Organisation for Economic Cooperation and Development
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by C Argyropoulos; submitted 07.12.24; peer-reviewed by K Liu, L Rodewald, S Lai; revised version received 27.04.25; accepted 05.05.25; published 03.07.25.

Please cite as:
Wang X, Chen Y, Ta Y, Guo M, Fu H
Trends in Avoidable Hospitalizations Before and During the COVID-19 Pandemic: Multiple Cross-Sectional Study Using Administrative Data From Beijing, China
JMIR Public Health Surveill 2025;11:e69768
URL: https://publichealth.jmir.org/2025/1/e69768
doi:10.2196/69768

© Xiangzhen Wang, Yin Chen, Yuqi Ta, Moning Guo, Hongqiao Fu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 3.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Influencing the Maintenance of Public Health Behaviors After an Epidemic: Cross-Sectional Study

Xingmin Wang[1,2*], MPH; Hongli Yan[1,2,3*], MRES; Lushaobo Shi[2,4], PhD; Ting Li[5], PhD; Yi Xia[6], PhD; Dong Wang[1,2,3,4], PhD

[1]School of Public Health, Southern Medical University, Guangzhou, China

[2]School of Health Management, Southern Medical University, No. 1023 Shatai Road, Guangzhou, China

[3]Key Laboratory of Philosophy and Social Sciences of Guangdong Higher Education Institutions for Health Policies Research and Evaluation, Guangzhou, China

[4]Research Base for Development of Public Health Service System of Guangzhou, Guangzhou, China

[5]School of Humanities and Management, Guangdong Medical University, Dongguan, China

[6]Center for Faculty Development and Research, Guangzhou Medical University, Guangzhou, China

[*]these authors contributed equally

Corresponding Author:
Dong Wang, PhD
School of Public Health, Southern Medical University, Guangzhou, China

Abstract

Background: The maintenance of public health behaviors stands as a critical issue within the realm of public health. COVID-19, a major global emergency, has profoundly impacted the sustainability of public health behaviors. However, there is currently a gap of empirical studies examining the status and influencing mechanisms of public health behavior maintenance after the pandemic, especially those adopting a multifactorial and integrated approach.

Objective: This study aims to investigate the current status of public health behavior maintenance in China after the COVID-19 pandemic. It integrates the complementary advantages of the Multi-Theory Model (MTM) and the Protective Action Decision Model (PADM) to conduct a comprehensive analysis of the multi-factorial mechanisms influencing the maintenance of public health behaviors. The findings are expected to provide empirical evidence and strategic recommendations for the formulation of public health policies and the promotion of national health.

Methods: An integrated model was developed based on the MTM and the PADM. Data were collected from the Chinese public between October and November 2023 via the online survey panel Sojump (Changsha Ranxing Information Technology Co., Ltd.). The questionnaire included items on health behavior maintenance, variables from the MTM and the PADM, sociodemographic and personal disease, and health characteristics. Univariate analysis, correlation analysis, multivariate regression analysis, and structural equation modeling were performed to explore the determinants of health behavior maintenance.

Results: This study collected 1216 valid samples, including 726 females and 490 males, with an average age of 27.38 (SD 8.52) years, and most of them had been infected with COVID-19 at least once (1054/1216, 86.68%). The public maintenance of health behaviors was at a fairly low level (Mean 2.88, SD 0.45). Multivariate regression analysis revealed that those with high monthly incomes, married individuals, and who were more concerned about their health after the COVID-19 pandemic had higher levels of health behavior maintenance. These variables, along with others from the MTM and the PADM, accounted for 45.5% of the variance in health behavior maintenance. Structural equation modeling indicated that efficacy perception had the most significant positive influence on health behavior maintenance (β=.386, *P*<.001), followed by emotional transformation and practical changes (both β=.213, *P*<.001). Risk perception had a slightly negative effect on health behavior maintenance (β=−.099, *P*=.013). Variables such as social cues, warning messages, and information sources also indirectly influenced the public maintenance of health behaviors.

Conclusions: This study indicates a slight decline in public health behavior maintenance following the COVID-19 pandemic, and our analysis has explored some of its influencing factors. Attention should be given to broadening information channels and appropriately explaining the risks of unhealthy behaviors. In addition, integrating external support and bolstering the public's efficacy in maintaining health behaviors can promote sustainable healthy practices.

(JMIR Public Health Surveill 2025;11:e66535) doi:10.2196/66535

KEYWORDS

health behaviors; maintenance; Multi-Theory Model; Protective Action Decision Model; influencing factors

XSL•FO
RenderX

Introduction

Health behaviors have a direct impact on an individual's health status and are crucial for maintaining health and prolonging life. Gochman defines health behaviors as patterns, actions, and habits related to health maintenance, recovery, and improvement [1]. However, many health-promoting behaviors need to be practiced consistently over time to significantly affect health [2]. This leads to the important concept of health behavior maintenance, which refers to the sustained adoption of health behaviors to maintain overall health and well-being [3]. Engaging in long-term health behaviors, such as regular physical activity and a healthy diet, can greatly reduce the risk of severe acute infectious diseases and chronic noncommunicable conditions, including coronary heart disease, type 2 diabetes mellitus, and cancer [4,5]. Large-scale and sustained health behavior practices are vital as they help mitigate future pandemic risks and enhance preparedness [6].

The COVID-19 pandemic, as a global public health emergency, has exerted significant changes in public health awareness, behavioral patterns, and social support mechanisms, impacting the maintenance of public health behaviors. Studies reveal a divergence of health behaviors caused by the pandemic: some individuals persistently adopt protective measures, such as mask-wearing, due to heightened risk perception [7], while others exhibit behavioral relaxation attributed to "pandemic fatigue" or socioeconomic pressures [8]. Psychological traits, such as conspiracy beliefs, have been found to significantly influence compliance with protective behaviors [9], whereas social trust and cultural values, such as collectivism, regulate the sustainability of behaviors through group identity [10]. Although existing studies in China have addressed the impact of pandemic-related knowledge and anxiety on health behaviors [11,12], and there have also been studies on changes in health behaviors before and during the pandemic in European countries [13], there remains a notable gap in theoretical exploration regarding the dynamic mechanisms of health behavior maintenance and multifactorial interactions.

Numerous studies have explored health behaviors, but most existing studies mainly focus on the static characteristics of health behaviors rather than their continuity [14]. Current studies often concentrate on changes in specific types of health behaviors, such as dietary habits or disease prevention measures, with less attention given to integrating a variety of health-related behaviors. In addition, while psychological and cognitive factors, such as risk perception and efficacy perception, have been emphasized, there has been insufficient comprehensive analysis that incorporates both internal and external factors, such as health-related message and peer support [15]. Based on theories of health behavior change, some studies have been conducted on health education and promotion interventions for patients with chronic disease [16]. Although these studies have identified factors related to the maintenance of health behaviors within the context of interventions, whether health behaviors can be sustained after removing external interventions is worthy of further study.

Studies on health behavior change have typically relied on single theoretical frameworks, such as the Theory of Planned Behavior or the Protection Motivation Theory, often overlooking social and environmental factors and behavioral continuity [17]. These approaches struggle to fully explain the complexity of maintaining health behaviors. In contrast, the Multi-Theory Model (MTM) proposed by Sharma integrates behavioral intention, practical skills, and environmental changes to predict the dynamic maintenance of health behaviors [18,19]. The Protective Action Decision Model (PADM) proposed by Lindell and Perry [20] emphasizes the interaction between risk perception, resource assessment, and adaptive decision-making, which is suitable for analyzing postcrisis behavior adaptation [21]. The MTM helps understand the impact of an individual's emotional and behavioral reactions on behavior maintenance, while the PADM focuses on behaviors in risk environments and complements the perceived factors not addressed in the MTM. The combination of these 2 models can provide a more comprehensive analytical framework for postpandemic health behavior study.

In conclusion, maintaining health behaviors is crucial for enhancing individual well-being amid the pandemic-driven public health challenges, and these challenges also provide a natural opportunity for us to investigate the maintenance of public health behaviors after the pandemic. However, public health behaviors may change due to personal psychological perceptions and external social factors after the pandemic, a topic that has not been systematically and comprehensively explored. Therefore, this study aims to investigate the status of public health behavior maintenance in the postpandemic era. By integrating the MTM and the PADM, we seek to uncover the key factors and mechanisms that influence the maintenance of public health behaviors. This study stands to enrich health behavior knowledge, and the strategies and recommendations proposed based on the findings may offer a theoretical foundation for policy implementation, inform effective public health interventions, and help build a more resilient and health-conscious society after the pandemic.

Methods

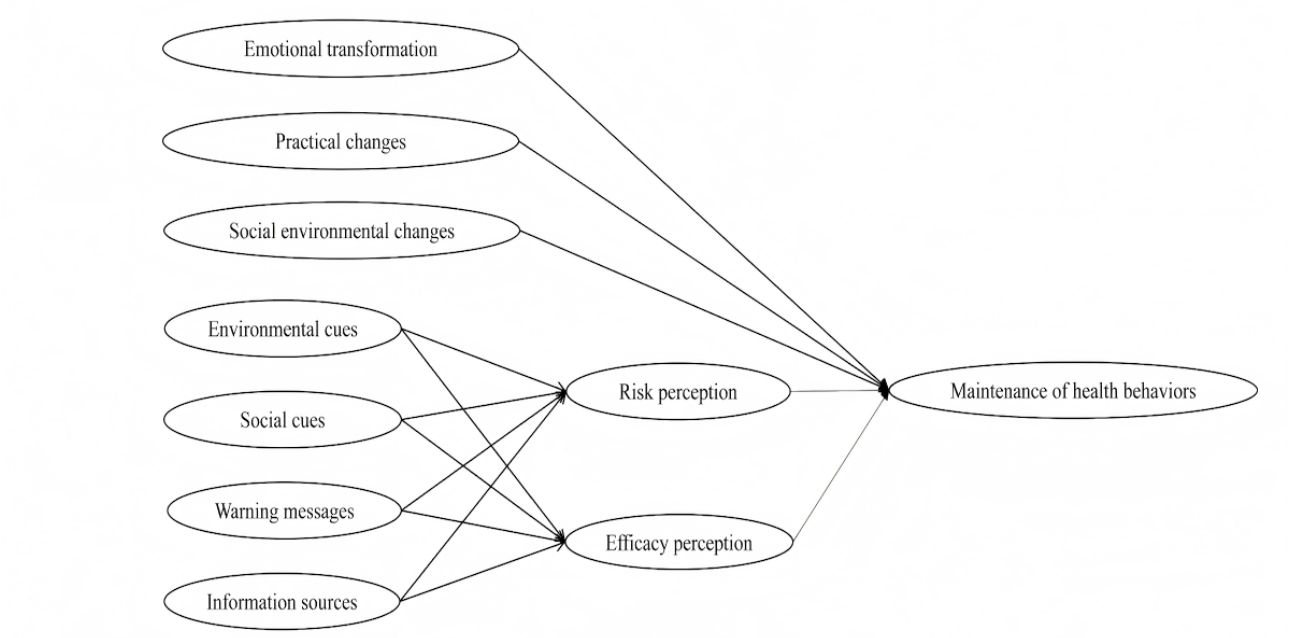
Framework of the Study

In reference to the study by Atanasova et al [22], this study combined 2 models to explore the key factors and mechanisms influencing the maintenance of public health behaviors after the COVID-19 pandemic. Based on the previous context, these 2 models are the MTM and the PADM, both widely used to explain and predict behaviors, and in our case, we target public maintenance of health behaviors. The MTM posits that 3 factors—emotional transformation, practical changes, and social environmental changes—influence long-term behavioral maintenance [18]. Meanwhile, the PADM emphasizes that environmental cues, social cues, warning messages, and information sources trigger core perceptions, including risk perception and efficacy perception, driving protective behavioral responses [20]. In the context of health behavior maintenance after the pandemic, the MTM helps predict the continuation of these behaviors, while the PADM aids in understanding how

XSL•FO
RenderX

people adapt their decisions in response to evolving risks and available resources. Combining these 2 models offers a comprehensive analysis of the factors influencing behavior maintenance from cognitive, motivational, and environmental perspectives as pandemic risks decline (Figure 1).

Figure 1. Framework of predictors for public maintenance of health behaviors based on the MTM and the PADM. MTM: Multi-Theory Model; PADM: Protective Action Decision Model.



Study Setting and Participants

A cross-sectional survey was conducted. From October to November 2023, we recruited participants from the general population in China using convenience sampling. The survey was conducted anonymously via the Sojump platform, a widely used online survey tool in China with a national user base. The eligibility criteria for participants were (1) voluntary completion of the questionnaire and informed consent to participate and (2) proficiency in using electronic devices.

The sample size was calculated using a formula guided by the Cochran technique [23], which is effective for large-population studies [24]. The calculation variables were n (sample size), z=1.96 (for a 95% confidence level), p=50% (the probability of selecting a choice, suitable for highly dispersed population), q=(1–p), and d=3.272% (allowed error percentage). According to the calculation, a sample size of 898 was determined.

$$n=z2pqd=1.962\times0.5\times(1-0.5)0.032722=897.0684\approx898$$

Recruitment Procedures

We distributed the survey links and recruitment advertisements on social media via our research team members, using the Sojump to provide survey links and widely invite eligible participants who met the inclusion criteria. To ensure data quality, all questionnaire items were set as mandatory, and each IP address, device, and account was restricted to one submission only. Also, a time limit was imposed to exclude respondents finishing the survey in less than 2 minutes. To encourage participation, those who successfully completed the survey were rewarded with 3 - 5 RMB (US$ 0.42-0.70). Ultimately, 1216 valid data were obtained.

Measurement Instruments

Dependent Variable

In consideration of the fact that there is currently no mature scale for assessing the maintenance of health behaviors, we adopted the Health Lifestyle and Personal Control Questionnaire developed by Darviri et al [25] to evaluate the public maintenance of health behaviors. This questionnaire comprises 26 items across 5 dimensions: healthy eating behaviors, unhealthy eating behaviors, daily life behaviors, organized physical exercise, and social and psychological balance. Respondents self-reported their sustained engagement in these behaviors during and after the COVID-19 pandemic. The scoring was as follows: “1” indicated the lowest level of maintenance (health behaviors never taken), “2” indicated a lower level of maintenance (behaviors adopted less frequently after the pandemic), “3” indicated unchanged maintenance, and “4” indicated a significant improvement in maintenance (behaviors adopted more frequently after the pandemic). Confirmatory factor analysis (CFA) indicated that the standardized factor loadings for each dimension ranged from 0.684 to 0.772, except for the second dimension (unhealthy eating behaviors) at −0.228. However, as eating behavior is also an important aspect of health behaviors, this dimension was retained. The Cronbach’s alpha (α=0.89) of this questionnaire demonstrated good internal consistency.

Independent Variables

Referring to the study by Li and Wang [26], environmental cues were assessed using 4 items with “yes” or “no” responses, scored as 1 for “no” and 2 for “yes.” A higher score indicated a stronger public perception of external environmental cues.

XSL•FO
RenderX

Based on Saghafi-Asl et al's study [27], social cues were evaluated using 5 items rated on a 5-point scale from 1 (strongly disagree) to 5 (strongly agree). A higher score suggested that the public received more social cues about maintenance of health behaviors. These items showed good internal consistency (α=0.88) with standardized factor loadings ranging from 0.596 to 0.773.

Following Sun et al [28] and Wang et al [29], information sources were measured with 8 items on a 5-point Likert scale (1=never, 2=rarely, 3=sometimes, 4=often, 5=very often). A higher score indicated that the public would obtain more information about health behaviors. The CFA revealed standardized factor loadings between 0.662 and 0.775, and Cronbach's α=0.86 demonstrated good internal consistency.

Warning messages were assessed through 3 items on the same 5-point Likert scale as for information sources, which were developed by Guo et al [21]. A higher score implied that the public received more information prompting them to engage in health behaviors. The CFA showed good validity (standardized factor loadings ranged from 0.837 to 0.853), and the Cronbach's α=0.83 demonstrated good internal consistency.

Referring to the study by Liu et al [30], risk perception was measured with 4 items. A higher score denoted a greater perceived risk associated with unhealthy behaviors. The CFA revealed standardized factor loadings from 0.781 to 0.832. The Cronbach's α=0.81 indicated good internal consistency.

Efficacy perception was assessed using a revised 5-item scale based on the study by Witte [31]. A higher score suggested a greater perception of the benefits of maintaining health behaviors. The CFA showed standardized factor loadings between 0.805 and 0.882, and Cronbach's α=0.85 demonstrated good internal consistency.

According to the study by Sharma et al [19], emotional transformation, practical changes, and social environmental changes in the MTM were evaluated using 3, 2, and 3 items, respectively. Higher scores in each variable indicated a greater likelihood of maintaining health behaviors. The CFA revealed standardized factor loadings ranging from 0.746 to 0.800 for emotional transformation, 0.784 to 0.832 for practical changes, and 0.709 to 0.767 for social environmental changes. The Cronbach's α values 0.81, 0.79, 0.78, respectively demonstrated acceptable internal consistency. These measurements, based on 5 variables, were consistent with the 5-point Likert scale used for social cues.

Control Variables

The maintenance of health behaviors is influenced by multiple factors. Based on previous literature review, we have summarized the following additional factors related to the public that may influence health behavior maintenance: (1) sociodemographic characteristics: gender, age, education, career, average monthly income, settlement type, marital status, and living condition. (2) Personal disease and health characteristics: whether they have a chronic disease, times that they have been infected with the COVID-19, self-assessment of health status, and whether they are more concerned about health after the COVID-19 pandemic.

Statistical Analysis

The data analysis was performed using R Studio (R 4.2.3; R Core Team). The CFA was used to test the validity of the measurements, with standardized factor loadings and average variance extracted as indicators of convergent validity. The reliability of the measurements was examined using Cronbach's α and composite reliability. Results of validity and reliability of measurements are shown in the Multimedia Appendix 1. Categorical variables were presented as frequency and proportion, and continuous variables were expressed as mean and SD. In the univariate analysis, normally distributed data were analyzed using *t* tests or ANOVA, and non-normally distributed data were analyzed using Wilcoxon rank-sum tests or rank correlation tests. Spearman correlation analysis was conducted to explore relationships between variables. Multivariate regression analysis was applied to examine impacts of demographic characteristics and key factors on the maintenance level of health behaviors. Structural equation modeling (SEM) was used to analyze pathways of factors influencing the public maintenance of health behaviors. A 2-sided *P* value below .05 was considered statistically significant.

Ethical Considerations

During the data collection, it was clearly stated to ensure that the survey is anonymous and all personal information would remain confidential. Participants were informed that their involvement is voluntary and they can withdraw from the study at any time. This study was approved by the Ethics Committee of Southern Medical University in Guangzhou (approval number: [2022] no. 65) and was conducted in accordance with the ethical principles outlined in the World Medical Association's Declaration of Helsinki for medical research involving human subjects. Participants who successfully completed the survey were rewarded with 3 - 5 RMB (US $0.42–0.70).

Results

Characteristics of Participants

A total of 726 females and 490 males participated in this survey, with an average age of 27.38 (SD 8.52) years. Among the participants, 12.83% (156/1216) reported having a chronic disease, and most of them had been infected with COVID-19 at least once (1054/1216, 86.68%). Over half rated their health as good (746/1216, 61.35%) and 73.27% (891/1216) indicated an increased concern for their health after the COVID-19 pandemic. Details are presented in Table 1.

XSL•FO
RenderX

Table . Characteristics of participants and univariate analysis of public maintenance of health behaviors.

Variables	Values, n (%)	Maintenance of health behaviors, mean (SD)	*F* test (*df*)	*P* value
Sex			0.05 (1214)	.819
Male	490 (40.30)	2.88 (0.43)		
Female	726 (59.70)	2.87 (0.47)		
Age groups (years)			46.01 (1193)[a]	<.001
<21	112 (9.21)	2.83 (0.47)		
21-31 (excluding 31)	836 (68.75)	2.84 (0.45)		
31-41 (excluding 41)	156 (12.83)	3.09 (0.42)		
≥41	93 (7.65)	2.90 (0.42)		
Education			64.12 (1211)[a]	<.001
Junior high school and below	31 (2.55)	2.78 (0.43)		
High school or junior college	83 (6.83)	2.94 (0.47)		
Three-year college	171 (14.06)	3.00 (0.52)		
Undergraduate	658 (54.11)	2.91 (0.43)		
Graduate students and above	273 (22.45)	2.72 (0.41)		
Career			61.34 (1208)[a]	<.001
Party and government officials	26 (2.14)	2.82 (0.36)		
Career employees	188 (15.46)	2.94 (0.39)		
Workers in enterprises	319 (26.23)	2.99 (0.43)		
Agricultural, forestry, livestock, or fishing water resource management personnel	39 (3.21)	2.94 (0.39)		
Self-employed	142 (11.68)	2.89 (0.50)		
Students	454 (37.34)	2.77 (0.44)		
Military personnel	2 (0.16)	3.48 (0.41)		
Others	46 (3.78)	2.79 (0.53)		
Average monthly income (yuan)[b]			91.84 (1210)[a]	<.001
<3000	451 (37.09)	2.74 (0.43)		
3000-5000 (excluding 5000)	208 (17.11)	2.87 (0.45)		
5000-8000 (excluding 8000)	312 (25.66)	3.00 (0.44)		
8000-10,000 (excluding 10,000)	145 (11.92)	3.01 (0.41)		
≥ 10,000	99 (8.14)	2.91 (0.47)		
Settlement type			0.65 (1214)	<.421
Urban	943 (77.55)	2.88 (0.45)		
Rural	273 (22.45)	2.86 (0.46)		
Marital status			85.57 (1211)[a]	<.001
Unmarried	833 (68.50)	2.80 (0.44)		
Married	374 (30.76)	3.05 (0.43)		
Divorcee	6 (0.49)	2.74 (0.52)		

XSL•FO
RenderX

Variables	Values, n (%)	Maintenance of health behaviors, mean (SD)	*F* test (*df*)	*P* value
Widowhood	2 (0.16)	2.81 (0.11)		
Living condition			8.53 (1214)[c]	.003
Living alone	346 (28.45)	2.82 (0.43)		
Not living alone	870 (71.55)	2.90 (0.46)		
Have a chronic disease			4.78 (1214)[d]	.029
No	1060 (87.17)	2.87 (0.45)		
Yes	156 (12.83)	2.93 (0.47)		
Times infected with COVID-19			4.54 (1212)	.209
Unknown	58 (4.77)	2.76 (0.50)		
No	104 (8.55)	2.87 (0.46)		
Once	750 (61.68)	2.89 (0.45)		
Twice or more	304 (25)	2.86 (0.44)		
Self-assessed health condition			59.01 (1211)[a]	<.001
Very poor	8 (0.66)	2.58 (0.57)		
Poor	59 (4.85)	2.77 (0.51)		
Acceptable	403 (33.14)	2.78 (0.44)		
Good	555 (45.64)	2.90 (0.42)		
Very good	191 (15.71)	3.05 (0.46)		
Concerned more about health after the COVID-19 pandemic			226.02 (1209)[a]	<.001
Strongly disagree	17 (1.40)	2.45 (0.50)		
Disagree	31 (2.55)	2.70 (0.46)		
Undecided	275 (22.62)	2.63 (0.41)		
Agree	551 (45.31)	2.86 (0.41)		
Strongly agree	340 (27.96)	3.14 (0.39)		

[a]*P*<.001

[b]1 yuan= US $0.14.

[c]*P*<.01

[d]*P*<.05

Univariate Analysis and Correlation Among Variables

For all participants, the overall maintenance score of health behaviors was 2.88 (SD 0.45) (Table 2). Factors such as age, education, career, average monthly income, marital status, living condition, whether one has a chronic illness, self-assessed health status, and the level of attention to personal health after the COVID-19 pandemic showed statistically significant associations with the public maintenance of health behaviors. Specifically, the factors of settlement type and time of COVID-19 infection showed no statistical differences with the public maintenance of health behaviors (Table 1). The results of the correlation analysis between variables reveal that all correlation coefficients were statistically significant, and the specific values of the correlation coefficients are shown in Table 3.

XSL•FO
RenderX

Table . Public maintenance level of health behaviors.

Dimension	Mean (SD)
Maintenance of health behaviors	2.88 (0.45)
Maintenance of healthy eating behaviors	2.86 (0.70)
Maintenance of unhealthy eating behaviors	2.59 (0.73)
Maintenance of daily life behaviors	2.97 (0.60)
Maintenance of organized physical exercise	2.92 (0.93)
Maintenance of social and psychological balance	2.98 (0.59)

Table . Correlations of all latent variables.

	Maintenance of health behaviors	Environmental cues	Information sources	Warning messages	Social cues	Emotional transformation	Practical changes	Social environmental changes	Risk perception	Efficacy perception
Maintenance of health behaviors	1.00									
Environmental cues	0.11[a]	1.00								
Information sources	0.48[a]	0.19[a]	1.00							
Warning messages	0.43[a]	0.15[a]	0.67[a]	1.00						
Social cues	0.49[a]	0.10[a]	0.55[a]	0.49[a]	1.00					
Emotional transformation	0.56[a]	0.09 (P=.001)[b]	0.49[a]	0.45[a]	0.64[a]	1.00				
Practical changes	0.58[a]	0.11[a]	0.47[a]	0.42[a]	0.57[a]	0.73[a]	1.00			
Environmental changes	0.50[a]	0.09 (P=.003)[b]	0.44[a]	0.42[a]	0.56[a]	0.67[a]	0.64[a]	1.00		
Risk perception	0.32[a]	0.06 (P=.043)[c]	0.31[a]	0.27[a]	0.53[a]	0.48[a]	0.35[a]	0.47[a]	1.00	
Efficacy perception	0.56[a]	0.08 (P=.005)[b]	0.49[a]	0.41[a]	0.62[a]	0.67[a]	0.69[a]	0.62[a]	0.44[a]	1.00

[a]P<.001

[b]P<.01

[c]P<.05

Multivariate Regression Analysis

The demographic factors found to be statistically significant in the univariate analysis, along with core study variables, were sequentially incorporated into the multivariate regression analysis using a stepwise forward method. Specifically, the inclusion order prioritized (1) statistically significant demographic factors identified in univariate analysis, (2) variables from the MTM, and (3) variables from the PADM. By the way, 2-tailed t test for a regression coefficient was conducted and the degrees of freedom was 1175. The results indicated that average monthly income, marital status, increased health concern after the COVID-19, emotional transformation, practical changes, information sources, warning messages, and efficacy perception influenced the public maintenance of health behaviors, with an adjusted R^2 of 0.455 (Table 4).

XSL•FO
RenderX

Table . Multivariate regression analysis on factors influencing the maintenance of public health behaviors (after full adjustment).

Influencing factors (control condition)	Estimate	SE	*t* test (*df*), 2-tailed	*P* value
Intercept	16.42	5.17	3.18 (1175)[b]	.002
Age, years (<21)				
21-31 (excluding 31)	–0.48	1.02	–0.47 (1175)	.639
31-41 (excluding 41)	0.47	1.35	0.35 (1175)	.725
≥41	0.87	1.58	0.55 (1175)	.581
Education (junior high school and below)				
High school or junior college	–0.53	1.92	–0.28 (1175)	.782
Three-year college	–1.27	1.83	–0.69 (1175)	.490
Undergraduate	–1.75	1.80	–0.97 (1175)	.332
Graduate students and above	–3.68	1.93	–1.90 (1175)	.057
Career (party and government officials)				
Career employees	1.09	1.89	0.58 (1175)	.562
Workers in enterprises	1.43	1.84	0.78 (1175)	.438
Agricultural, forestry, livestock, or fishing water resource management personnel	0.07	2.29	0.03 (1175)	.975
Self-employed	–0.26	1.94	–0.13 (1175)	.895
Students	2.49	2.06	1.21 (1175)	.227
Military personnel	12.23	6.50	1.88 (1175)	.060
Others	<0.01	2.25	<0.01 (1175)	.999
Average monthly income (yuan ; <3000)[c]				
3000-5000 (excluding 5000)	2.35	1.02	2.29 (1175)[d]	.022
5000-8000 (excluding 8000)	3.73	1.13	3.31 (1175)[b]	.001
8000-10,000 (excluding 10,000)	3.97	1.26	3.15 (1175)[b]	.002
≥10,000	4.07	1.34	3.04 (1175)[b]	.002
Marital status (unmarried)				
Married	1.81	0.83	2.18 (1175)[d]	.029
Divorced	1.51	3.71	0.41 (1175)	.683
Widowed	8.58	6.32	1.36 (1175)	.175
Living condition (living alone)				
Not living alone	0.75	0.64	1.18 (1175)	.237
Have a chronic disease (no)				
Yes	0.28	0.79	0.35 (1175)	.725
Self-assessed health status (very poor)				
Poor	5.06	3.36	1.51 (1175)	.132
Acceptable	3.82	3.19	1.20 (1175)	.231
Good	4.45	3.18	1.40 (1175)	.162
Very good	5.08	3.22	1.58 (1175)	.115
Concerned more about health after the COVID-19 pandemic (strongly disagree)				

XSL•FO
RenderX

Influencing factors (control condition)	Estimate	SE	*t* test (*df*), 2-tailed	*P* value
Disagree	8.68	2.70	3.22 (1175)[b]	.001
Undecided	6.44	2.26	2.85 (1175)[b]	.004
Agree	8.74	2.23	3.93[a] (1175)	<.001
Strongly agree	10.46	2.26	4.63[a] (1175)	<.001
Emotional transformation	0.64	0.22	2.98 (1175)[b]	.003
Practical changes	0.99	0.24	4.09[a] (1175)	<.001
Social environmental changes	0.19	0.19	1.03 (1175)	.306
Environmental cues	0.36	0.31	1.17 (1175)	.241
Information source	0.17	0.07	2.44 (1175)[b]	.015
Warning messages	0.48	0.14	3.51[a] (1175)	<.001
Social cues	0.09	0.13	0.72 (1175)	.474
Risk perception	0.03	0.15	0.23 (1175)	.819
Efficacy perception	0.56	0.12	4.83[a] (1175)	<.001

[a]*P*<.001.

[b]*P*<.01.

[c]1 yuan=US $0.14.

[d]*P*<.05.

Structural Equation Modeling and Hypothesis Testing

In the SEM, the path coefficients could reflect the strength and direction of causality between variables. Compared with regression coefficients, path coefficients capture the net effect of one variable on another after controlling variables [32]. Standardized path coefficients convert these into a unitless value for direct comparison [33]. All standardized path coefficients in this study were within the normal range, as shown in Table 5. The modified model shows good overall fit to the actual sample data, with χ^2/df of 4.92 (χ^2_{353}=1737.201) and the goodness-of-fit index, comparative fit index, standardized root mean square residual, and root mean square error of approximation of 0.894, 0.923, 0.055, and 0.057, respectively. All indicators meet the requiremexcept for thet for goodness-of-fit index, which is slightly below 0.90.

XSL•FO
RenderX

Table . Path coefficients of variables in the SEM.

Hypothesis	Unstandardized coefficient	SE	*z*	*P* value	Standardized coefficient
Maintenance of health behaviors ← emotional transformation	0.296	0.045	6.635	<.001	0.213
Maintenance of health behaviors ← practical changes	0.296	0.045	6.635	<.001	0.213
Maintenance of health behaviors ← social environmental changes	0.036	0.087	0.418	.676	0.026
Maintenance of health behaviors ← efficacy perception	0.313	0.038	8.320	<.001	0.386
Maintenance of health behaviors ← risk perception	–0.095	0.038	–2.484	.013	–0.099
Risk perception ← environmental cues	0.015	0.043	0.356	.721	0.009
Risk perception ← social cues	1.130	0.081	13.939	<.001	0.789
Risk perception ← warning messages	–0.091	0.032	–2.859	.004	–0.064
Risk perception ← information sources	–0.091	0.032	–2.859	.004	–0.064
Efficacy perception ← environmental cues	–0.007	0.044	–0.149	.882	–0.003
Efficacy perception ← social cues	1.334	0.092	14.562	<.001	0.782
Efficacy perception ← warning messages	0.038	0.032	1.184	.236	0.022
Efficacy perception ← information sources	0.038	0.032	1.184	.236	0.022

The analysis shows that social cues have a positive influence on efficacy perception (β=.782, *P*<.001), which in turn has the most significant effect on the maintenance of public health behaviors (β=.386, *P*<.001). Emotional transformation and practical changes from the MTM also positively affect public health behavior maintenance (both β=.213, *P*<.001). Risk perception is found to have a negative impact on public health behaviors (β=−.099, *P*=.013). Social cues are the most influential factor for risk perception with a positive effect (β=.789, *P*<.001), while warning messages and information sources are negatively associated with risk perception (both β=−.064, *P*=.004).

Discussion

Principal Findings

This study found that the public maintenance of health behaviors was at a fairly low level (mean 2.88, SD 0.45) after the COVID-19 pandemic. Similar to the study of Wang and Gago [34], children exhibited health behavior maintenance issues after the pandemic, such as poor sleep, increased screen time, and changed daily activities. When examining the dimensions of health behavior maintenance, the public surveyed in this study maintained relatively stable social and psychological balance, daily life behaviors, and organized physical exercise. However, the maintenance of dietary behaviors showed significant fluctuation. This echoes the study by van Reekum et al [35], which highlighted the difficulty in sustaining dietary patterns formed during the pandemic once normal life resumed. Moreover, as the pandemic progressed, pandemic fatigue set in, leading to a decline in health behaviors due to the sustained high level of vigilance required [36]. Even a short-term decrease in health behavior maintenance may adversely affect health. Therefore, it is crucial to focus on public health behavior maintenance and its influencing factors after the pandemic and to implement effective and timely interventions.

In this study, factors such as average monthly income, marital status, and personal health concern after the COVID-19 pandemic were found to be associated with the public maintenance of health behaviors. Wambiya et al [37] similarly found that low-income groups, burdened by economic hardships and work demands, may neglect the role of health behaviors in promoting physical health. Meanwhile, spousal encouragement,

XSL•FO
RenderX

support, and supervision can aid in maintaining health behaviors [38]. In addition, individuals with higher health concerns tend to have better health cognition, enabling them to practice health behaviors more effectively [39]. Notably, this study found no correlation between settlement type or whether individuals have been infected with COVID-19 and public health behavior maintenance. This is inconsistent with the study of Wollast et al [40]. The discrepancy may arise because most participants surveyed in this study had been infected with COVID-19 (86.68%). With frequent social interactions and population movements between urban and rural areas, resulting in widespread COVID-19 infections [41], which led to no significant differences in population distribution.

This study confirms that emotional transformation and practical changes in the MTM are positive predictors of health behavior maintenance, aligning with prior study [42]. Emotional transformation emphasizes redirecting negative emotions associated with health behaviors, such as reluctance and resistance, toward the goal of maintaining health behaviors [43]. Practical changes focus on identifying and correcting barriers and incorrect practices in implementing health behaviors through recording or reflection [44]. Throughout these 2 processes, individuals typically combine intrinsic self-motivation with extrinsic emotional support, seeking solutions that use both internal and external resources to overcome psychological barriers in maintaining health behaviors.

This study reveals that warning messages and information sources in the PADM negatively impact risk perception, which in turn undermines the public maintenance of health behaviors. This contradicts Cao and Liu's [45] findings in their study on internet users. Generally, access to public health information is positively correlated with health behaviors [46]. However, in this study, the public might have encountered an abundance of conflicting and dubious health information, resulting in confusion, information overload, and distrust [47]. This aligns with Pedro et al's [48] study, which could diminish sensitivity to health risks. Some studies point out that risk perception is a positive predictor of health behaviors [49,50], but excessive risk perception may trigger impatience or rebelliousness, decreasing individuals' willingness and frequency of action [51]. Therefore, the bidirectional effects of information reception and risk perception on behavior warrant further exploration.

This study indicates a positive correlation between efficacy perception and the public maintenance of health behaviors, consistent with existing studies [52,53]. Individuals with stronger beliefs in their ability to maintain health behaviors exhibit higher self-control and willingness, highlighting the beneficial role of confidence and motivation in adhering to health practices [54]. Moreover, a greater perception of health behavior effects strengthens their belief in maintaining these behaviors [55]. However, this study found that information sources and warning messages do not influence health behavior maintenance through efficacy perception, conflicting with the hypothesis in the PADM [20]. This might be due to asymmetric information transmitted through these channels, leading to varying efficacy perceptions. When individuals perceive insufficient benefits, they tend to adopt a conservative behavior pattern and are less likely to persist in health behaviors [56].

This study found that social cues positively influence risk perception and efficacy perception, in turn affecting public health behavior maintenance, consistent with the PADM [20]. Individuals with greater exposure to social cues show heightened awareness of potential behavioral risks [57] and acquire more health-related information. This increased awareness and information boost their confidence in maintaining health behaviors and motivate them to sustain these behaviors. However, unlike previous studies [58,59], no statistical link was found between environmental cues and health behavior maintenance through risk and efficacy perception, nor any impact of social environmental changes on health behavior maintenance. This might be because most participants had COVID-19–related experiences during the study period. Due to frequent social interactions and population movements, people's sensitivity to environmental cues and social environmental changes decreased compared to the pandemic period. Similar to a "gray rhino" event, individuals tend to overlook health risk warnings when they are accustomed to a situation [60].

Strengths and Limitations

The MTM and the PADM are both useful for studying behavior change mechanisms. The MTM focuses on individual psychology and emotional shifts, while the PADM emphasizes external environmental information. Despite their different focal points, they have not been integrated into previous studies. This study combines their strengths to analyze factors influencing public health behavior maintenance. It takes into account the public real-time health risk-benefit perception and behavioral responses after the pandemic, offering a new perspective on health behavior study and providing references for promoting public health behavior maintenance.

This study offers initial insights into the health behavior maintenance and its influencing factors among the Chinese public, but it has limitations that need further exploration. The results of this study rely on self-reporting by the participants, which may be prone to recall bias. Future research could use smart wearable devices to collect objective data on health behavior maintenance, such as daily step counts and sleep duration. In addition, the online survey with convenience sampling leans toward younger participants who are more tech-savvy, while they can provide unique perspectives on this study. To enhance data representativeness, future studies can combine online and offline recruitment channels, with regional stratified sampling to recruit more diverse participants, and embrace a longitudinal design to better establish causal relationships between factors.

Implications

Drawing on this study, we propose the following suggestions to enhance public health behavior maintenance. First, government, health departments, and medical institutions should be actively engaged as authoritative sources to promote public health behaviors, leveraging their authority and credibility. Second, integrate traditional, online, and social media to broaden information-sharing channels, innovate dissemination methods, and embed health concepts into public daily life. While propagandizing the benefits of health behaviors, it is also

XSL•FO
RenderX

necessary to educate the public about the dangers of unhealthy behaviors and corresponding health warning messages. It is crucial to avoid false information and exaggerated information. Finally, enhance family and social support to help individuals manage negative emotions related to health behavior maintenance and encouraging self-motivation for effective health behavior management.

Conclusions

This study, integrating the MTM and the PADM, explores the status and influencing mechanisms of public health behavior maintenance after the COVID-19 pandemic. Our findings indicate that enhancing emotional transformation, practical changes, social cues, and warning messages can boost health behavior maintenance. However, attention should be given to pandemic-related fatigue, and the risk information should be communicated properly. Also, external support should be integrated to boost the public's sense of efficacy in maintaining health behaviors. This study attempts to integrate the MTM and the PADM within the Chinese cultural context and offers perspectives on health behavior studies. It provides some strategies and recommendations for promoting the public maintenance of health behaviors.

Acknowledgments

We are grateful to all participants in this study. This work was financially supported by the Guangdong Basic and Applied Basic Research Foundation (grant number: 2023A1515011124) and the Guangdong Basic and Applied Basic Research Foundation Joint Fund Project (grant number: 2023A1515110723).

Disclaimer

The views expressed in this article are our own, with generative artificial intelligence subsequently employed for linguistic enhancements.

Data Availability

The datasets generated or analyzed during this study can be obtained from the corresponding author on reasonable request.

Authors' Contributions

XW and DW designed the study. Data collection was carried out by XW, HY, LS, TL, and YX. Data analysis was conducted by XW, LS, and DW. The manuscript was written by XW, HY, and DW, with all authors contributing to its revision. All authors approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Validity and reliability of measurements.
[DOCX File, 30 KB - publichealth_v11i1e66535_app1.docx]

Checklist 1
STROBE Checklist for reporting cross-sectional studies.
[PDF File, 136 KB - publichealth_v11i1e66535_app2.pdf]

References

1. Gochman DS. Handbook of Health Behavior Research: Plenum Press; 1997. [doi: 10.1007/978-1-4899-1757-7]
2. Gardner B, Arden MA, Brown D, et al. Developing habit-based health behaviour change interventions: twenty-one questions to guide future research. Psychol Health 2023 Apr;38(4):518-540. [doi: 10.1080/08870446.2021.2003362] [Medline: 34779335]
3. Health behavior maintenance. Office of Behavioral and Social Sciences Research. URL: https://obssr.od.nih.gov/about/health-behavior-maintenance [accessed 2025-04-22]
4. Dunton GF. Sustaining health-protective behaviors such as physical activity and healthy eating. JAMA 2018 Aug 21;320(7):639-640. [doi: 10.1001/jama.2018.6621] [Medline: 29852046]
5. Dominguez LJ, Di Bella G, Veronese N, Barbagallo M. Impact of Mediterranean diet on chronic non-communicable diseases and longevity. Nutrients 2021 Jun 12;13(6):2028. [doi: 10.3390/nu13062028] [Medline: 34204683]
6. Michie S, West R. Sustained behavior change is key to preventing and tackling future pandemics. Nat Med 2021 May;27(5):749-752. [doi: 10.1038/s41591-021-01345-2] [Medline: 33972794]

XSL•FO
RenderX

7. Voltmer E, Köslich-Strumann S, Walther A, Kasem M, Obst K, Kötter T. The impact of the COVID-19 pandemic on stress, mental health and coping behavior in German University students - a longitudinal study before and after the onset of the pandemic. BMC Public Health 2021 Jul 13;21(1):1385. [doi: 10.1186/s12889-021-11295-6] [Medline: 34256717]
8. Yang M, He Z, Zhang Y, Liu T, Ming WK. Pandemic fatigue and preferences for COVID-19 public health and social measures in China: nationwide discrete choice experiment. JMIR Public Health Surveill 2024 Jun 27;10:e45840. [doi: 10.2196/45840] [Medline: 38935420]
9. Pilch I, Turska-Kawa A, Wardawy P, Olszanecka-Marmola A, Smołkowska-Jędo W. Contemporary trends in psychological research on conspiracy beliefs. A systematic review. Front Psychol 2023;14:1075779. [doi: 10.3389/fpsyg.2023.1075779] [Medline: 36844318]
10. Xing S, Mohamed Zainal SR. Collectivism-based organizational culture, green empowerment, environmental self-identity and workplace green behavior: the stimulus-organism-response perspective. Discov Sustain 2024;5(1):1-14. [doi: 10.1007/s43621-024-00417-w]
11. Liu Y, Li P, Lv Y, et al. Public awareness and anxiety during COVID-19 epidemic in China: a cross-sectional study. Compr Psychiatry 2021 May;107:152235. [doi: 10.1016/j.comppsych.2021.152235] [Medline: 33765493]
12. Zhang N, Kou Y. Implicit theories of health, consideration of future consequences, and engagement in health protective behaviors during the COVID-19 pandemic in China. J Health Psychol 2022 May;27(6):1462-1469. [doi: 10.1177/13591053211017191] [Medline: 33983050]
13. Skotnicka M, Karwowska K, Kłobukowski F, Wasilewska E, Małgorzewicz S. Dietary habits before and during the COVID-19 epidemic in selected European countries. Nutrients 2021 May 16;13(5):1690. [doi: 10.3390/nu13051690] [Medline: 34065655]
14. Ory MG, Lee Smith M, Mier N, Wernicke MM. The science of sustaining health behavior change: the health maintenance consortium. Am J Health Behav 2010;34(6):647-659. [doi: 10.5993/ajhb.34.6.2] [Medline: 20604691]
15. Marks B, Sisirak J, Magallanes R, Krok K, Donohue-Chase D. Effectiveness of a HealthMessages peer-to-peer program for people with intellectual and developmental disabilities. Intellect Dev Disabil 2019 Jun;57(3):242-258. [doi: 10.1352/1934-9556-57.3.242] [Medline: 31120402]
16. Kanai M, Miki T, Sakoda T, Hagiwara Y. The effect of combining mHealth and Health professional-led intervention for improving health-related outcomes in chronic diseases: systematic review and meta-analysis. Interact J Med Res 2025 Jan 20;14:e55835. [doi: 10.2196/55835] [Medline: 39832160]
17. Wilderink L, Bakker I, Schuit AJ, Seidell JC, Pop IA, Renders CM. A theoretical perspective on why socioeconomic health inequalities are persistent: building the case for an effective approach. Int J Environ Res Public Health 2022 Jul 8;19(14):8384. [doi: 10.3390/ijerph19148384] [Medline: 35886234]
18. Sharma M. Multi-theory model (MTM) for health behavior change. ResearchGate. 2015 Sep 23. URL: http://www.webmedcentral.com/article_view/4982 [accessed 2024-11-28]
19. Sharma M, Khubchandani J, Nahar VK. Applying a new theory to smoking cessation: case of multi-theory model (MTM) for health behavior change. Health Promot Perspect 2017;7(2):102-105. [doi: 10.15171/hpp.2017.18] [Medline: 28326291]
20. Lindell MK, Perry RW. The protective action decision model: theoretical modifications and additional evidence. Risk Anal 2012 Apr;32(4):616-632. [doi: 10.1111/j.1539-6924.2011.01647.x] [Medline: 21689129]
21. Guo Y, An S, Comes T. From warning messages to preparedness behavior: the role of risk perception and information interaction in the Covid-19 pandemic. Int J Disaster Risk Reduct 2022 Apr 15;73:102871. [doi: 10.1016/j.ijdrr.2022.102871] [Medline: 35261877]
22. Atanasova S, Kamin T, Perger N. Predictors of COVID-19 vaccination intention and behavior among young people in a European Union country with low COVID-19 vaccination rates: cross-sectional study. JMIR Public Health Surveill 2025 Feb 21;11:e64653. [doi: 10.2196/64653] [Medline: 39983109]
23. Borg DN, Bach AJE, O'Brien JL, Sainani KL. Calculating sample size for reliability studies. PM R 2022 Aug;14(8):1018-1025. [doi: 10.1002/pmrj.12850] [Medline: 35596122]
24. Rahman S, Munam AM, Hossain A, Hossain A, Bhuiya RA. Socio-economic factors affecting the academic performance of private university students in Bangladesh: a cross-sectional bivariate and multivariate analysis. SN Soc Sci 2023;3(2):26. [doi: 10.1007/s43545-023-00614-w] [Medline: 36685659]
25. Darviri C, Alexopoulos EC, Artemiadis AK, et al. The Healthy Lifestyle and Personal Control Questionnaire (HLPCQ): a novel tool for assessing self-empowerment through a constellation of daily activities. BMC Public Health 2014 Sep 24;14(1):995. [doi: 10.1186/1471-2458-14-995] [Medline: 25253039]
26. Li M, Wang W. Educational sisparities in COVID-19 prevention in China: the role of contextual danger, perceived risk, and interventional context. Int J Environ Res Public Health 2021 Mar 24;18(7):3383. [doi: 10.3390/ijerph18073383] [Medline: 33805222]
27. Saghafi-Asl M, Aliasgharzadeh S, Asghari-Jafarabadi M. Factors influencing weight management behavior among college students: an application of the Health Belief Model. PLoS ONE 2020;15(2):e0228058. [doi: 10.1371/journal.pone.0228058] [Medline: 32032376]

XSL•FO
RenderX

28. Sun Y, Huang SK, Arlikatti S, Lindell MK. What attributes influence rural household's willingness to get vaccinated for COVID-19? Perspectives from six Chinese townships. Vaccine (Auckl) 2023 Jan 16;41(3):702-715. [doi: 10.1016/j.vaccine.2022.11.062] [Medline: 36535824]
29. Wang F, Wei J, Shi X. Compliance with recommended protective actions during an H7N9 emergency: a risk perception perspective. Disasters 2018 Apr;42(2):207-232. [doi: 10.1111/disa.12240] [Medline: 28799670]
30. Liu BB, Zhang FF, Cheng L, Zhang XH. Influencing factors of consumer nutritional food choice: based on an expanded model of health beliefs. Journal of HuaZhong 2022;56(6):1074-1084 [FREE Full text] [doi: 10.19603/j.cnki.1000-1190.2022.06.019]
31. Witte K, Cameron KA, McKeon JK, Berkowitz JM. Predicting risk behaviors: development and validation of a diagnostic scale. J Health Commun 1996;1(4):317-341. [doi: 10.1080/108107396127988] [Medline: 10947367]
32. Li YT, Sawhney R, Luz Tortorella G. Empirical analysis of factors impacting turnover intention among manufacturing workers. IJBM 2019;14(4):1. [doi: 10.5539/ijbm.v14n4p1]
33. Kozak M, Kang MS. Note on modern path analysis in application to crop science. Commun Biom Crop Sci 2006;1(1):32-34 [FREE Full text]
34. Wang ML, Gago CM. Shifts in child health behaviors and obesity after COVID-19. JAMA Pediatr 2024 May 1;178(5):427-428. [doi: 10.1001/jamapediatrics.2024.0027] [Medline: 38436952]
35. van Reekum EA, Woo JJ, Petropoulos JA, Samaan Z, Mbuagbaw L. Association between the COVID-19 pandemic and psychiatric symptoms in people with preexisting obsessive-compulsive, eating, anxiety, and mood disorders: a systematic review and meta-analysis of before-after studies. Psychiatry Clin Neurosci 2023 Nov;77(11):583-591. [doi: 10.1111/pcn.13582] [Medline: 37565691]
36. Lai DWL, Jin J, Yan E, Lee VWP. Predictors and moderators of COVID-19 pandemic fatigue in Hong Kong. J Infect Public Health 2023 May;16(5):645-650. [doi: 10.1016/j.jiph.2023.03.003] [Medline: 36913768]
37. Wambiya EOA, Donfouet HPP, Kipruto S, et al. Patterns, socioeconomic inequalities and determinants of healthy eating in Kenya: results from a national cross-sectional survey. BMJ Open 2025 Apr 14;15(4):e090698. [doi: 10.1136/bmjopen-2024-090698] [Medline: 40228849]
38. Choi SH, Templin T, Glenn D. Preferences for couple-based smoking cessation interventions among couples who smoke: online discrete choice experiment. Health Educ Res 2025 Apr 16;40(3):cyaf010. [doi: 10.1093/her/cyaf010] [Medline: 40242933]
39. Tang L, Wang J. Effects of new media use on health behaviors: a case study in China. Iran J Public Health 2021 May;50(5):949-958. [doi: 10.18502/ijph.v50i5.6112] [Medline: 34183953]
40. Wollast R, Schmitz M, Bigot A, Brisbois M, Luminet O. Predicting health behaviors during the COVID-19 pandemic: a longitudinal study. PLoS ONE 2024;19(3):e0299868. [doi: 10.1371/journal.pone.0299868] [Medline: 38489308]
41. Franceschi VB, Santos AS, Glaeser AB, et al. Population-based prevalence surveys during the Covid-19 pandemic: a systematic review. Rev Med Virol 2021 Jul;31(4):e2200. [doi: 10.1002/rmv.2200] [Medline: 34260777]
42. Sharma M, Largo-Wight E, Kanekar A, Kusumoto H, Hooper S, Nahar VK. Using the multi-theory model (MTM) of health behavior change to explain intentional outdoor nature contact behavior among college students. Int J Environ Res Public Health 2020 Aug 21;17(17):6104. [doi: 10.3390/ijerph17176104] [Medline: 32825734]
43. Sharma A, Jain M, Nahar VK, Sharma M. Determining predictors of change in sugar sweetened beverage consumption behaviour among university students in India. Int J Adolesc Med Health 2020 Jan 25;34(1). [doi: 10.1515/ijamh-2019-0078] [Medline: 31981451]
44. Kapukotuwa S, Nerida T, Batra K, Sharma M. Utilization of the multi-theory model (MTM) of health behavior change to explain health behaviors: a systematic review. Health Promot Perspect 2024;14(2):121-135. [doi: 10.34172/hpp.42887] [Medline: 39291044]
45. Cao P, Liu R. Exploring the cognitive factors of corrective health information adoption intention in infodemic using SEM and fsQCA. 2022 Presented at: WHICEB 2022 Proceedings URL: https://aisel.aisnet.org/whiceb2022/93
46. Wang E, An N, Gao Z, Kiprop E, Geng X. Consumer food stockpiling behavior and willingness to pay for food reserves in COVID-19. Food Secur 2020;12(4):739-747. [doi: 10.1007/s12571-020-01092-1] [Medline: 32837661]
47. Santomauro DF, Mantilla Herrera AM, Shadid J. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021 Nov 6;398(10312):1700-1712. [doi: 10.1016/S0140-6736(21)02143-7] [Medline: 34634250]
48. Pedro LC, Fernandes JL, Arent CO, et al. Frequency and method of seeking for information about COVID-19 and its relationship with psychological symptoms and stress levels. An Acad Bras Cienc 2025;97(2):e20231050. [doi: 10.1590/0001-3765202520231050] [Medline: 40243805]
49. Pietrobelli A, Pecoraro L, Ferruzzi A, et al. Effects of COVID-19 lockdown on lifestyle behaviors in children with obesity living in Verona, Italy: a longitudinal study. Obesity (Silver Spring) 2020 Aug;28(8):1382-1385. [doi: 10.1002/oby.22861] [Medline: 32352652]
50. Jin L, Zhang N, Zhu J. Healthy lifestyle changes and mental health of healthcare workers during the COVID-19 pandemic in China. Curr Psychol 2022 Aug 13:1-10. [doi: 10.1007/s12144-022-03562-5] [Medline: 35990210]

XSL•FO
RenderX

51. Yan AF, Sun X, Zheng J, et al. Corrigendum to "Perceived risk, behavior changes and Health-related outcomes during COVID-19 pandemic: Findings among adults with and without diabetes in China". [Diabetes Res. Clin. Practice 167 (2020) 108350]. Diabetes Res Clin Pract 2021 Jul;177:108881. [doi: 10.1016/j.diabres.2021.108881] [Medline: 34120759]
52. Dai B, Fu D, Meng G, Liu B, Li Q, Liu X. The effects of governmental and individual predictors on COVID-19 protective behaviors in China: a path analysis model. Public Adm Rev 2020;80(5):797-804. [doi: 10.1111/puar.13236] [Medline: 32836438]
53. Kakaire CN, Christofides N. The role of perceived threat and self-efficacy in the use of Insecticide Treated Bednets (ITNs) to prevent malaria among pregnant women in Tororo District, Uganda. PLoS ONE 2023;18(7):e0289097. [doi: 10.1371/journal.pone.0289097] [Medline: 37494377]
54. Melnyk BM, Hsieh AP, Tan A, et al. State of mental health, healthy behaviors, and wellness support in Big 10 University Nursing and Health Sciences faculty, staff, and students During COVID-19. J Prof Nurs 2023;48:152-162. [doi: 10.1016/j.profnurs.2023.07.006] [Medline: 37775230]
55. Khaira N, Faisal TI, Magfirah M, Veri N, Usrina N, Andy Rias Y. Effectiveness of a health belief model-based education program on self-efficacy and preparedness for infectious disasters in rural populations. Creat Nurs 2024 May;30(2):165-172. [doi: 10.1177/10784535241245436] [Medline: 38591961]
56. Bults M, Beaujean D, Richardus JH, Voeten H. Perceptions and behavioral responses of the general public during the 2009 influenza A (H1N1) pandemic: a systematic review. Disaster Med Public Health Prep 2015 Apr;9(2):207-219. [doi: 10.1017/dmp.2014.160] [Medline: 25882127]
57. Zhou R, Qin K, Huang R. A study of the influence of destination social cues on tourists' protective decision-making in a risky environment. Tour Manag Technol Econ 2024;7(1):143-153. [doi: 10.23977/tmte.2024.070118]
58. Meng G, Li Q, Yuan X, et al. The roles of risk perception, negative emotions and perceived efficacy in the association between COVID-19 infection cues and preventive behaviors: a moderated mediation model. BMC Public Health 2023 Jan 16;23(1):109. [doi: 10.1186/s12889-022-14870-7] [Medline: 36647034]
59. Doria MDF, Pidgeon N, Hunter PR. Perceptions of drinking water quality and risk and its effect on behaviour: a cross-national study. Sci Total Environ 2009 Oct 15;407(21):5455-5464. [doi: 10.1016/j.scitotenv.2009.06.031] [Medline: 19664799]
60. Robertson J. Planning for grey rhino risks: how to prepare for the "unforeseeable". J Bus Contin Emer Plan 2024 Jan 1;17(4):383-394. [Medline: 38736160]

Abbreviations

CFA: confirmatory factor analysis
MTM: Multi-Theory Model
PADM: Protective Action Decision Model
SEM: structural equation modeling

Edited by A Mavragani; submitted 16.09.24; peer-reviewed by A Hossain, X Yao; revised version received 29.05.25; accepted 28.05.25; published 23.07.25.

Please cite as:
Wang X, Yan H, Shi L, Li T, Xia Y, Wang D
Factors Influencing the Maintenance of Public Health Behaviors After an Epidemic: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e66535
URL: https://publichealth.jmir.org/2025/1/e66535
doi:10.2196/66535

© Xingmin Wang, Hongli Yan, Lushaobo Shi, Ting Li, Yi Xia, Dong Wang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Perceived Unfairness, Uncertainty, and Life Stress on Urban Residents' Health in China: Cross-Sectional Study

Jinsong Chen[1], PhD; Shuhan Jiang[2], PhD; Mingli Pang[3], MPH; Huan Zhou[4], PhD; Weifang Zhang[4], PhD

[1]School of Law, Hangzhou City University, Hangzhou, China

[2]School of Humanities and Management, Zhejiang Chinese Medical University, NO.548 Binwen Rd, Binjiang District, Hangzhou, China

[3]National Institute for Health Innovation, School of Population Health, University of Auckland, Auckland, New Zealand

[4]Clinical Research Center for Oral Diseases of Zhejiang Province, Key Laboratory of Oral Biomedical Research of Zhejiang Province, Cancer Center of Zhejiang University, Stomatology Hospital, School of Stomatology, School of Medicine, Zhejiang University, Hangzhou, China

Corresponding Author:
Shuhan Jiang, PhD
School of Humanities and Management, Zhejiang Chinese Medical University, NO.548 Binwen Rd, Binjiang District, Hangzhou, China

Abstract

Background: The prevalence of stress factors in urban settings has increased, with significant potential impact on health outcomes.

Objective: This study aimed to explore the effect of perceived unfairness, uncertainty, and life stress on self-reported health status, short-term illness, and noncommunicable chronic diseases (NCDs).

Methods: A cross-sectional study using multistage stratified sampling was carried out in Xi'an, Hangzhou, Guangzhou, and Guiyang in China in July 2021, and 2851 participants were included in data analysis. Self-reported health status, short-term illness, and NCDs were assessed using self-administered questionnaires. Perceived life stress, unfairness stress, and uncertainty stress were measured using a standard scale. Descriptive statistics were used to analyze participants' demographic characteristics, and chi-square tests clarified statistical differences in self-reported health status, short-term illness, and NCDs based on these characteristics. Multiple logistic regression models were constructed to measure the effects of perceived unfairness stress, uncertainty stress, and life stress on self-reported health status, short-term illness, and NCDs.

Results: Severe uncertainty stress (odds ratio [OR] 1.230, 95% CI 1.007-1.503) and severe life stress (OR 1.728, 95% CI 1.411-2.118) were associated with a higher likelihood of poor self-reported health status. Severe uncertainty stress (OR 1.565, 95% CI 1.270-1.929) and severe life stress (OR 1.404, 95% CI 1.136-1.731) increased the odds of short-term illness. In addition, severe unfairness stress (OR 1.306, 95% CI 1.053-1.620), severe uncertainty stress (OR 1.542, 95% CI 1.248-1.905), and severe life stress (OR 1.344, 95% CI 1.084-1.667) were linked to a higher prevalence of NCDs.

Conclusions: To conclude, severe uncertainty and life stress were both associated with increased odds of poor health outcomes (self-reported health status, short-term illness, and NCDs), while severe unfairness only affected NCDs. The findings of this study may serve as an empirical reference for the improvement of self-reported illness among Chinese urban residents.

(JMIR Public Health Surveill 2025;11:e72094) doi:10.2196/72094

KEYWORDS

self-reported illness; self-reported health status; short-term illness; noncommunicable chronic diseases; unfairness stress; uncertainty stress; life stress

Introduction

Stressors typically arise from an imbalance between an individual's perceived needs and their ability to meet those needs. Consequently, stress occurs when individuals are unable to effectively manage or regulate these demands [1,2]. Increasing evidence suggests that stress levels among populations have reached unprecedented heights during natural disasters [3] and the COVID-19 pandemic [4], driven primarily by stressors such as a sharp economic downturn, widespread unemployment, and geopolitical instability [5].

Current research examining the relationship between stress and health has faced criticism on several fronts. According to the stress-diathesis theory, further classifications of general stress are necessary, as different types of stress exert distinct effects on health [6]; however, many studies have yet to differentiate between various forms of stress. Life stress is defined as persistent daily worries, such as poor living conditions and strained personal relationships [7,8]. These stressors have been

XSL•FO
RenderX

shown to negatively affect health. A review article highlighted that life stress, especially stress arising from adverse social environments, serves as a key indicator of aging [9]. Previous studies have also demonstrated a positive association between life stress and conditions such as insomnia and depression [10,11]. Furthermore, life stress has been linked to physical health issues, including hypertension and respiratory illness [12-14].

Perceived unfairness has been identified as a significant source of stress, encompassing violations of rights or psychological infringements linked to group membership or personal experiences [15]. As a psychological construct shaped by social dynamics, perceived unfairness has profound implications for physical health. According to the perceived unfairness model, perceived injustice triggers stress responses, which can be detrimental to physical well-being [16]. Specifically, perceived unfairness contributes to anxiety and insecurity, and prolonged exposure to such psychological challenges can severely damage individuals' social lives and well-being, ultimately impacting their overall health [17]. Substantial evidence supports the adverse effects of unfairness on various health indicators [18,19]. For instance, a study among multiethnic adults shown that perceived unfairness was related to self-reported history of depression and high blood pressure [20], Similarly, another study identified unfairness as a risk factor for metabolic syndrome and its components [21].

Uncertainty stress is a subjective experience stemming from perceptions of unpredictability and ambiguity [22]. Early studies have confirmed that uncertainty stress is a significant stressor in the context of illness and hospitalization [23], and uncertainty stress has been linked to the exacerbation of health conditions, with individuals facing health-related uncertainties reporting greater symptom severity and functional limitations [24]. Experimental research has further demonstrated that exposure to uncertainty stress leads to increases in both systolic and diastolic blood pressure [25]. Besides, chronic exposure to daily uncertainty has also been associated with a poorer quality of life and reduced well-being [26]. Given the pervasive nature of uncertainty in modern society, recent studies have shifted their focus to uncertainty stress in broader social contexts, particularly social change uncertainty and social value uncertainty [22], which reflect deeper societal and cultural transformations that shape individuals' stress responses and health outcomes. Research has found that uncertainty stress is linked to a range of adverse health outcomes and unhealthy behaviors, including unintentional injuries [27], deliberate self-harm [28], suicidal ideation [29], problematic alcohol use [30], and poorer self-rated health [31,32]. Following the outbreak of COVID-19 pandemic, studies have demonstrated that uncertainty stress is positively associated with disease-related fear [33], depression, and anxiety symptoms [34]. However, few studies have specifically examined the relationship between uncertainty stress and different types of illnesses, highlighting a critical gap in the literature.

The novelty of this study lies in 2 key aspects. First, most previous studies on this topic have been conducted outside of China, with limited research addressing the unique characteristics of the Chinese population. Despite profound societal changes following the Reform and Opening-up, including a widening income gap and increasing disparities in resource distribution the overall perception of inequity among China's residents appears relatively weak, with a high level of acceptance [35]. This context underscores the urgency of exploring whether perceived unfairness is still associated with adverse health outcomes and identifying the specific health outcomes involved. Second, this study uniquely explores the effects of 3 distinct stressors on health outcomes. It is critical to examine whether perceived uncertainty stress and perceived unfairness stress remain associated with health outcomes, even after controlling for life stressors. The results would be crucial for the development of future stress management strategy.

To conclude, very few studies have examined the relationship between perceived unfairness, uncertainty, and life stress in terms of urban residents' health among the Chinese population. The aim of this study was to examine the effect of perceived unfairness, uncertainty, and life stress on urban residents' health among the Chinese population, including self-reported illness, short-term illness, and noncommunicable chronic diseases (NCDs).

Methods

Participants

A cross-sectional study was conducted across 4 provincial capital cities in China—Hangzhou (East China), Xi'an (Northwest China), Guangzhou (South China), and Guiyang (West China)—in July 2021. A multistage stratified sampling approach was used. In the first stage, 4 provincial capital cities were selected to represent different geographical regions of China. In the second stage, 2 districts in the main urban area were randomly selected within each city. In the third stage, 4 communities in each selected district were chosen randomly. In the fourth stage, due to the restriction caused by the COVID-19 epidemic, convenience sampling was used within communities to recruit participants. Urban residents aged ≥18 years were eligible to enroll, and a total of 2851 participants completed the survey and were included in data analysis. Informed consent was obtained from all participants before their inclusion in the study.

Assessment and Measurements

Self-Reported Health Status

Self-reported health status was measured by the item "What do you think about your health status when compared with your majority of colleagues/classmates?" The response was recoded dichotomously, with responses as excellent or good=0 and fair, poor, or bad=1 during data analysis [36].

Short-Term Illness

Short-term illness was assessed by the question "Have you experienced short-term illnesses such as diarrhea, colds and others at least three times in the last 6 months?" The answer was coded as "yes=1" and "no=0" [37].

XSL•FO
RenderX

NCDs

NCDs were assessed by the question "Are you suffering from any chronic diseases, such as digestive trouble, hypertension, other cardiovascular diseases, and others," with the responses coded as "yes=1" and "no=0."

Perceived Life Stress

Perceived life stress was measured through standard questionnaires designed by Yang and colleagues [38]: life stress from (1) unstable work and income, lack of security in life; (2) high work pressure, heavy tasks, and a fast pace; (3) tension with leaders or coworkers; (4) friction between family members; (5) difficulty in children's education or employment; and (6) poor health of family members or oneself. These questions were measured by a 5-point Likert scale ranging from 1 (no stress) to 5 (severe stress). The average score of >3 was classified as severe life stress.

Perceived Unfairness Stress

Perceived unfairness stress was measured by 7 dimensions: income unfairness, education unfairness, promotion unfairness, health care unfairness, workplace benefit unfairness, pension unfairness, and other unfairness. Responses were recorded on a 5-point Likert scale from 1 to 5: never, rarely, sometimes, often, and very often. An average score more than 3 was classified as severe perceived unfairness stress.

Perceived Uncertainty Stress

Perceived uncertainty stress was assessed by standard questionnaires designed by Yang and colleagues [38]. The perceived uncertainty items included the following: life is unstable and cannot be controlled, uncertain about what will happen in the future, uncertain about how to achieve goals, and cannot follow social values. The score was based on a 5-point scale (very low, low, medium, high, and very high). The higher the total score, the higher the perceived uncertainty stress. An average score greater than 3 was classified as severe uncertainty stress. Previous studies have illustrated the good reliability and validity of the scale [29,30,39].

Sociodemographic Information

Sociodemographic characteristics of participants were measured by the following items: hukou (urban and rural; a population management policy in China), district (Xi'an, Hangzhou, Guangzhou, and Guiyang), sex (male and female), age (<25, 25‐34, 35‐44, 45‐54, >55 years), ethnicity (Han and minority), marital status (never married, married and living together, married without living together or divorced, or widowed), education (junior high school and below, senior high school, university, and postgraduate or above), occupation (government agencies, public institutions, or state-owned enterprises; private enterprise; self-employed people; freelance work; unemployed, student, or retired), average household income (1 RMB=US $0.15 at the time of the study; <50,000 RMB, 50,001‐100,000 RMB, 100,001‐150,000 RMB, 150,001‐200,000 RMB, and >200,000 RMB).

Statistical Analysis

Descriptive statistics were used to analyze the demographic characteristics of the participants. Chi-square test was used to clarify the statistical differences of self-reported health status, short-term illness, and NCDs in terms of demographic characteristics, respectively. Multiple logistic regression models were constructed to examine the effects of perceived unfairness, uncertainty, and life stress on self-reported health status, short-term illness, and NCDs. Variables that showed statistical significance in chi-square tests were selected and then simultaneously entered into the model using the "Enter" method in SPSS (IBM Corp).

Ethical Considerations

The study was approved by the Ethics Committee at the Zhejiang Chinese Medical University Medical Center (approval number 20210622‐3). The procedures followed were in accordance with the Declaration of Helsinki. Verbal consent was obtained from all participants before data collection, and each participant received a small gift valued at 3 yuan (1 RMB=US $0.15 at the time of the study). The study ensured that all participant data were anonymized and deidentified to protect privacy. The data were securely stored and accessible only to the research team.

Results

As shown in Table 1, the study included 2851 participants, with a balanced representation of rural and urban hukou, a nearly equal gender distribution, and a predominance of Han ethnicity. The largest age group was participants aged <25 years (771/2851, 27.1%) and over half (1622/2851, 56.9%) were married or living together. University-educated individuals account for 46.8% (1334/2851), and 34.9% (994/2851) of them were unemployed, students, or retired. Household income levels varied, with the most common range being 150,001‐200,000 RMB (1 RMB=US $0.15 at the time of the study) per year. Notable disparities in health outcomes were observed across different sociodemographic participants. Rural hukou individuals reported poorer self-reported health status (485/1382, 35.1%) compared to urban hukou individuals (420/1469, 28.6%), while urban hukou individuals exhibited a slightly higher prevalence of NCDs. Significant regional differences were also noted, with Xi'an reporting the highest prevalence of both poor self-reported health status and NCDs. Age was a strong determinant, with older participants (>55 y) showing higher prevalence of poor self-reported health status and NCDs. Other factors, such as ethnicity, education level, and employment status, were similarly associated with variations in health outcomes.

XSL•FO
RenderX

Table . Sociodemographic characteristics of urban residents in Xi'an, Hangzhou, Guangzhou, Guiyang, China, in July 2021.

Variables	Values, n (%)	Experiences and prevalence		Chi-square (*df*), *P* value	
		Poor self-reported health status, n (%)	NCDs[a], n (%)	Self-reported health status	NCDs
Hukou				13.899 (1), *P*<.001	1.414 (1), *P*=.234
Urban	1469 (51.5)	420 (28.6)	479 (32.6)		
Rural	1382 (48.5)	485 (35.1)	422 (30.5)		
District				29.876 (3), *P*<.001	34.093 (3), *P*<.001
Xi'an	770 (27)	267 (34.7)	292 (37.9)		
Hangzhou	699 (24.5)	223 (31.9)	242 (34.6)		
Guangzhou	835 (29.3)	293 (35.1)	215 (25.7)		
Guiyang	547 (19.2)	122 (22.3)	152 (27.8)		
Sex				1.849 (1), *P*=.174	2.123 (1), *P*=.145
Male	1411 (49.5)	431 (30.5)	464 (32.9)		
Female	1440 (50.5)	474 (32.9)	437 (30.3)		
Age (years)				42.824 (4), *P*<.001	292.324 (4), *P*<.001
<25	771 (27.1)	236 (30.6)	142 (18.4)		
25 - 34	443 (15.5)	112 (25.3)	83 (18.7)		
35 - 44	581 (20.4)	162 (27.9)	166 (28.6)		
45 - 54	485 (17)	153 (31.5)	177 (36.5)		
>55	571 (20)	242 (42.4)	333 (58.3)		
Ethnicity				17.564 (1), *P*<.001	4.299 (1), *P*=.038
Han	2436 (85.4)	810 (33.3)	788 (32.3)		
Minority	415 (14.6)	95 (22.9)	113 (27.2)		
Marital status				34.018 (3), *P*<.001	171.839 (3), *P*<.001
Never married	1000 (35.1)	293 (29.3)	189 (18.9)		
Married, and living together	1622 (56.9)	506 (31.2)	588 (36.3)		
Married without living together or divorced	126 (4.4)	48 (38.1)	50 (39.7)		
Widowed	103 (3.6)	58 (56.3)	74 (71.8)		
Education				45.357 (3), *P*<.001	170.152 (3), *P*<.001
Junior high school and below	675 (23.7)	280 (41.5)	347 (51.4)		
Senior high school	683 (24)	172 (25.2)	204 (29.9)		
University	1334 (46.8)	410 (30.7)	308 (23.1)		
Postgraduate or above	159 (5.5)	43 (27)	42 (26.4)		
Occupation				50.543 (4), *P*<.001	39.179 (4), *P*<.001
Government agencies, public institutions, or state-owned enterprises	467 (16.4)	110 (23.6)	175 (37.5)		
Private enterprise	469 (16.5)	135 (28.8)	99 (21.1)		
Self-employed people	404 (14.1)	118 (29.2)	130 (32.2)		
Freelance work	517 (18.1)	146 (28.2)	148 (28.6)		
Unemployed, student, or retired	994 (34.9)	396 (39.8)	349 (35.1)		

XSL•FO
RenderX

Variables	Values, n (%)	Experiences and prevalence		Chi-square (*df*), *P* value	
		Poor self-reported health status, n (%)	NCDs[a], n (%)	Self-reported health status	NCDs
Average household income (RMB)[b]				52.368 (4), *P*<.001	25.876 (4), *P*<.001
<50,000	793 (27.8)	324 (40.9)	294 (37.1)		
50,001 - 100,000	865 (30.4)	277 (32.0)	288 (33.3)		
100,001 - 150,000	608 (21.3)	155 (25.5)	168 (27.6)		
150,001 - 200,000	280 (9.8)	68 (24.3)	75 (26.8)		
>200,000	305 (10.7)	224 (26.6)	76 (24.9)		
	81				

[a]NCDs: noncommunicable chronic diseases.

[b]1 RMB=US $0.15 at the time of the study.

As shown in Table 2, multiple logistic regression models were constructed to estimate the odds ratios (ORs) for self-reported health status, adjusting for unfairness stress, uncertainty stress, and life stress, as well as a full model including all variables. Key findings revealed that rural hukou individuals had higher odds of poor self-reported health compared to urban hukou individuals. Significant regional differences were also observed, with Hangzhou and Guangzhou showing higher odds of poor health compared to Xi'an, while the difference between Guiyang and Xi'an was not statistically significant. Age was also a significant factor, with individuals aged >55 years having the highest odds of poor health. Sociodemographic factors, such as ethnicity, marital status, education level, occupation, and income, also played critical roles.

XSL•FO
RenderX

Table . Multiple logistic regression models of self-reported health status among urban residents in Xi'an, Hangzhou, Guangzhou, or Guiyang, China, in July 2021.

Group	Model 1: Unfairness model	Model 2: Uncertainty stress model	Model 3: Life stress model	Model 4: Full model, OR
Hukou, OR[a] (95% CI)				
Urban	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Rural	1.243 (1.036-1.491)[b]	1.229 (1.024-1.475)[b]	1.204 (1.002-1.446)[b]	1.205(1.003-1.449)[b]
District, OR (95% CI)				
Xi'an	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Hangzhou	1.291 (1.008-1.655)[b]	1.296 (1.011-1.661)[b]	1.428 (1.110-1.837)[c]	1.442 (1.120-1.856)[c]
Guangzhou	1.693 (1.335-2.146)[c]	1.705 (1.344-2.162)[c]	1.738 (1.369-2.208)[c]	1.733 (1.364-2.202)[c]
Guiyang	0.770 (0.566-1.049)	0.755 (0.555-1.026)	0.760 (0.558-1.033)	0.798 (0.584-1.091)
Age (years), OR (95% CI)				
<25	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
25 - 34	0.921 (0.648-1.310)	0.957 (0.673-1.363)	0.932 (0.653-1.330)	0.934 (0.654-1.333)
35 - 44	1.038 (0.713-1.511)	1.075 (0.737-1.568)	1.081 (0.740-1.578)	1.070 (0.732-1.564)
45 - 54	1.214 (0.821-1.793)	1.258 (0.850-1.861)	1.231 (0.830-1.825)	1.242 (0.837-1.843)
>55	1.625 (1.107,2.385)[b]	1.758 (1.195,2.587)[c]	1.860 (1.261,2.742)[c]	1.859 (1.259,2,746)[c]
Ethnicity, OR (95% CI)				
Han	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Minority	0.819 (0.609-1.102）	0.803 (0.596-1.082)	0.799 (0.593-1.078)	0.786 (0.582-1.060)
Marital status, OR (95% CI)				
Never married	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Married, and living together	1.457 (1.060-2.003)[b]	1.425 (1.037-1.960)[b]	1.437 (1.043-1.980)[b]	1.422 (1.032-1.960)[b]
Married without living together or divorced	2.293 (1.419-3.705)[c]	2.219 (1.373-3.584)[c]	2.184 (1.349-3.534)[c]	2.160 (1.334-3.496)[c]
Widowed	3.021 (1.786-5.112)[c]	2.917 (1.718-4.953)[c]	2.933 (1.723-4.993)[c]	2.909 (1.706-4.960)[c]
Education, OR (95% CI)				
Junior high school and below	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Senior high school	0.727 (0.561-0.942)[b]	0.743 (0.573-0.964)[b]	0.720 (0.555-0.934)[b]	0.730 (0.562-0.947)[b]
University	1.114 (0.853-1.456）	1.129 (0.864-1.476)	1.097 (0.838-1.435)	1.099 (0.839-1.439)
Postgraduate or above	1.047 (0.664-1.651)	1.074 (0.681-1.695)	1.045 (0.661-1.654)	1.054 (0.666-1.669)
Occupation, OR (95% CI)				
Government agencies, public institutions, or state-owned enterprises	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Private enterprise	1.370 (0.997-1.881)	1.344 (0.979-1.845)	1.371 (0.997-2.106)	1.383 (1.005-1.904)[b]
Self-employed people	1.467 (1.052-2.046)[b]	1.433 (1,027-1.999)[b]	1.426 (1.020-1.994)[b]	1.429 (1.022-1.998)[b]
Freelance work	1.247 (0.911-1.707)	1.229 (0.897-1.684)	1.248 (0.909-1.713)	1.241 (0.904-1.704)
Unemployed, student, or retired	2.112 (1.559-2.861)[c]	2.005 (1.479-2.718)[c]	2.083 (1.533-2.829)[c]	2.064 (1.518-2.806)[c]
Average household income (RMB)[d], OR (95% CI)				
<50,000	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)

XSL•FO
RenderX

Group	Model 1: Unfairness model	Model 2: Uncertainty stress model	Model 3: Life stress model	Model 4: Full model, OR
50,001 - 100,000	0.798 (0.640-0.996)[b]	0.805 (0.645-1.004)	0.808 (0.647-1.010)	0.812 (0.650-1.015)
100,001 - 150,000	0.619 (0.475-0.806)[c]	0.624 (0.479-0.812)[c]	0.618 (0.474-0.807)[c]	0.626 (0.480-0.817)[c]
150,001 - 200,000	0.548 (0.387-0.775)[c]	0.549 (0.388-0.776)[c]	0.553 (0.390-0.784)[c]	0.561 (0.396-0.795)[c]
>200,000	0.552 (0.397-0.769)[c]	0.556 (0.399-0.775)[c]	0.561 (0.402-0.782)[c]	0.565 (0.405-0.790)[c]
Unfairness stress, OR (95% CI)				
Low unfairness stress	1.00 (reference)	—[e]	—	1.00 (reference)
Severe unfairness stress	1.490 (1.248-1.779)[c]	—	—	1.069 (0.871-1.313)
Uncertainty stress, OR (95% CI)				
Low uncertainty stress	—	1.00 (reference)	—	1.00 (reference)
Severe uncertainty stress	—	1.620 (1.368-1.919)[c]	—	1.230 (1.007-1.503)[b]
Life stress, OR (95% CI)				
Low life stress	—	—	1.00 (reference)	1.00 (reference)
Severe life stress	—	—	1.959 (1.648-2.330)[c]	1.728 (1.411-2.118)[c]
Fixed parameters, β	0.180[c]	0.163[c]	0.149[c]	0.138[c]

[a]OR: odds ratio.

[b]*P*<.05.

[c]*P*<.01.

[d]1 RMB=US $0.15 at the time of the study.

[e]Not applicable.

Stress factors were important determinants, and individuals who were at severe uncertainty stress and life stress had significantly higher the odds of poor health. Both were statistically significant.

Multiple logistic regression models (short-term illness) were constructed to estimate the ORs, adjusting for unfairness stress, uncertainty stress, life stress, and a full model incorporating all factors. As shown in Table 3, key findings highlighted that rural hukou individuals had slightly higher odds (OR 1.206, 95% CI 0.996-1.460) of short-term illness compared to urban hukou individuals, and regional disparities were evident. Compared to Xi'an, Hangzhou showed lower odds of short-term illness, while Guangzhou and Guiyang exhibited little to no difference. Sociodemographic factors such as age, ethnicity, marital status, education, and income also played roles.

XSL•FO
RenderX

Table . Multiple logistic regression models of short-term illness among urban residents in Xi'an, Hangzhou, Guangzhou, or Guiyang, China, in July 2021.

Group	Model 1: Unfairness model	Model 2: Uncertainty stress model	Model 3: Life stress model	Model 4: Full model
Hukou, OR[a] (95% CI)				
Urban	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Rural	1.240 (1.026-1.500)[b]	1.224 (1.011-1.481)[b]	1.208 (0.998-1.462)	1.206 (0.996-1.460)
District, OR (95% CI)				
Xi'an	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Hangzhou	0.613 (0.469-0.799)[c]	0.627 (0.480-0.819)[c]	0.660 (0.505-0.863)[c]	0.669 (0.511-0.876)[b]
Guangzhou	0.992 (0.777-1.267)	0.999 (0.781-1.277)	1.011 (0.791-1.293)	1.008 (0.788-1.289)
Guiyang	0.993 (0.732-1.384)	1.016 (0.749-1.379)	0.973 (0.719-1.318)	1.046 (0.768-1.427)
Age (years), OR (95% CI)				
<25	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
25 - 34	0.869 (0.612-1.235)	0.896 (0.629-1.276)	0.877 (0.616-1.248)	0.884 (0.619-1.260)
35 - 44	0.801 (0.548-1.170)	0.820 (0.560-1.203)	0.829 (0.567-1.214)	0.823 (0.561-1.208)
45 - 54	0.681 (0.454-1.019)	0.706 (0.470-1.060)	0.690 (0.459-1.036)	0.702 (0.467-1.056)
>55	0.999 (0.677-1.474)	1.097 (0.741-1.626)	1.123 (0.758-1.662)	1.144 (0.770-1.699)
Ethnicity, OR (95% CI)				
Han	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Minority	0.869 (0.646-1.167)	0.829 (0.614-1.118)	0.849 (0.630-1.143)	0.817 (0.604-1.103)
Marital status, OR (95% CI)				
Never married	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Married, and living together	0.981 (0.714-1.347)	0.951 (0.691-1.308)	0.963 (0.700-1.324)	0.942 (0.684-1.297)
Married without living together or divorced	1.435 (0.882-2.334)	1.378 (0.846-2.245)	1.371 (0.842-2.233)	1.345 (0.825-2.193)
Widowed	1.905 (1.138-3.189)[b]	1.823 (1.084-3.066)[b]	1.835 (1.092-3.082)[b]	1.795 (1.065-3.024)[b]
Education, OR (95% CI)				
Junior high school and below	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Senior high school	0.612 (0.466-0.804)[c]	0.626 (0.476-0.824)[c]	0.605 (0.460-0.794)[c]	0.619 (0.470-0.814)[c]
University	0.752 (0.570-0.991)[b]	0.761 (0.577-1.004)	0.742 (0.562-0.979)[b]	0.747 (0.566-0.987)[b]
Postgraduate or above	0.672 (0.418-1.080)	0.687 (0.426-1.108)	0.676 (0.420-1.088)	0.683 (0.423-1.102)
Occupation, OR (95% CI)				
Government agencies, public institutions, or state-owned enterprises	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Private enterprise	1.006 (0.725-1.395)	0.986 (0.709-1.369)	1.002 (0.722-1.390)	1.001 (0.720-1.391)
Self-employed people	1.191 (0.849-1.670)	1.159 (0.825-1.628)	1.160 (0.826-1.630)	1.153 (0.820-1.622)
Freelance work	0.841 (0.607-1.165)	0.818 (0.589-1.136)	0.838 (0.604-1.162)	0.822 (0.591-1.142)
Unemployed, student, or retired	1.105 (0.812-1.504)	1.035 (0.759-1.411)	1.083 (0.794-1.476)	1.047 (0.767-1.430)
Average household income (RMB)[d], OR (95% CI)				
<50,000	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
50,001 - 100,000	1.015 (0.804-1.282)	1.028 (0.813-1.299)	1.027 (0.812-1.298)	1.036 (0.819-1.310)

XSL•FO
RenderX

Group	Model 1: Unfairness model	Model 2: Uncertainty stress model	Model 3: Life stress model	Model 4: Full model
100,001 - 150,000	1.018 (0.774-1.337)	1.035 (0.786-1.361)	1.016 (0.772-1.336)	1.037 (0.787-1.366)
150,001 - 200,000	1.055 (0.740-1.505)	1.069 (0.748-1.526)	1.055 (0.739-1.507)	1.078 (0.754-1.541)
>200,000	0.912 (0.642-1.295)	0.929 (0.653-1.323)	0.921 (0.648-1.310)	0.939 (0.659-1.337)
Unfairness stress, OR (95% CI)				
Low unfairness stress	1.00 (reference)	—[e]	—	1.00 (reference)
Severe unfairness stress	1.436 (1.194-1.727)[c]	—	—	1.012 (0.817-1.253)
Uncertainty stress, OR (95% CI)				
Low uncertainty stress	—	1.00 (reference)	—	1.00 (reference)
Severe uncertainty stress	—	1.834 (1.535-2.191)[c]	—	1.565 (1.270-1.929)[c]
Life stress, OR (95% CI)				
Low life stress	—	—	—	1.00 (reference)
Severe life stress	—	—	1.738 (1.453-2.078)[c]	1.404 (1.136-1.731)[c]
Fixed parameters, β	0.429[c]	0.358[c]	0.377[c]	0.327[c]

[a]OR: odds ratio.
[b]*P*<.05.
[c]*P*<.01.
[d]1 RMB=US $0.15 at the time of the study.
[e]Not applicable.

Individuals experiencing severe uncertainty stress had a 56.5% higher odds of reporting short-term illness compared to those with low uncertainty stress. Individuals experiencing severe life stress had 40.4% higher odds of reporting short-term illness compared to those with low life stress. Both were statistically significant.

Multiple logistic regression models (NCD) were constructed to estimate ORs, adjusting for unfairness stress, uncertainty stress, life stress, and a full model incorporating all factors.

Illustrated in Table 4, key findings showed that rural hukou individuals had a slightly lower odds of NCDs compared with urban hukou individuals. Significant regional differences were observed, with Hangzhou reporting higher odds (OR 1.290, 95% CI 0.994 - 1.674) of NCDs compared to Xi'an, while Guiyang showed lower odds. Age was also a significant factor, with older participants, especially those over 55 years old, having the highest OR of NCDs. Other sociodemographic variables like ethnicity, marital status, education level, occupation, and income also influenced outcomes.

XSL•FO
RenderX

Table . Multiple logistic regression models of noncommunicable chronic diseases among urban residents in Xi'an, Hangzhou, Guangzhou, or Guiyang, China, in July 2021.

Group	Model 1: Unfairness model	Model 2: Uncertainty model	Model 3: Life stress model	Model 4: Full model
Hukou, OR[a] (95% CI)				
Urban	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Rural	0.825 (0.684-0.995)	0.797 (0.659-0.963)	0.795 (0.659-0.960)	0.797 (0.659-0.963)
District, OR (95% CI)				
Xi'an	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Hangzhou	1.092 (0.858-1.391)	1.095 (0.859-1.394)	1.155 (0.904-1.475)	1.290 (0.994-1.674)
Guangzhou	0.983 (0.769-1.256)	0.996 (0.779-1.274)	1.012 (0.792-1.294)	1.054 (0.820-1.356)
Guiyang	0.735 (0.542-0.996)	0.708 (0.523-0.958)	0.670 (0.496-0.904)	0.803 (0.587-1.098)
Age (years), OR (95% CI)				
<25	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
25 - 34	0.999 (0.678-1.471)	1.045 (0.707-1.543)	1.030 (0.698-1.522)	1.015 (0.687-1.501)
35 - 44	1.440 (0.968-2.141)	1.494 (1.001-2.229)[b]	1.533 (1.029-2.284)[b]	1.476 (0.988-2.204)
45 - 54	1.821 (1.211-2.738)[b]	1.899 (1.258-2.866)[b]	1.858 (1.232-2.800)[b]	1.892 (1.253-2.857)[b]
>55	3.479 (2.330-5.194)[c]	3.879 (2.586-5.821)[c]	3.972 (2.659-5.956)[c]	3.965 (2.637-5.963)[c]
Ethnicity, OR (95% CI)				
Han	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Minority	0.956 (0.712-1.284)	0.932 (0.693-1.254)	0.956 (0.712-1.285)	0.918 (0.682-1.234)
Marital status, OR (95% CI)				
Never married	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Married, and living together	1.363 (0.979-1.897)	1.339 (0.960-1.866)	1.356 (0.973-1.889)	1.331 (0.953-1.857)
Married without living together or divorced	2.161 (1.328-3.516)[b]	2.089 (1.283-3.401)[b]	2.108 (1.298-3.425)[b]	2.049 (1.258-3.339)[b]
Widowed	3.413 (1.949-5.977)[c]	3.322 (1.889-5.844)[c]	3.348 (1.906-5.880)[c]	3.336 (1.894-5.877)[c]
Education, OR (95% CI)				
Junior high school and below	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Senior high school	0.615 (0.479-0.791)[c]	0.636 (0.494-0.819)[c]	0.609 (0.473-0.783)[c]	0.628 (0.487-0.809)[c]
University	0.549 (0.423-0.713)[c]	0.562(0.433-0.731)[c]	0.545 (0.420-0.708)[c]	0.550 (0.423-0.716)[c]
Postgraduate or above	0.508 (0.325-0.795)[b]	0.534 (0.341-0.836)	0.509 (0.325-0.7981)[c]	0.525 (0.334-0.825)
Occupation, OR (95% CI)				
Government agencies, public institutions, or state-owned private enterprises	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Private enterprise	0.596 (0.435-0.817)[c]	0.574 (0.419-0.788)[c]	0.574 (0.419-0.787)[c]	0.589 (0.429-0.810)[c]
Self-employed people	0.753 (0.550-1.030)	0.725 (0.529-0.994)[b]	0.726 (0.530-0.995)[b]	0.738 (0.537-1.013)
Freelance work	0.597 (0.442-0.807)[c]	0.594 (0.438-0.805)[c]	0.603 (0.445-0.816)	0.595 (0.439-0.807)[c]
Unemployed, student, or retired	1.047 (0.779-1.408)	0.966 (0.716-1.302)	1.021 (0.758-1.375)	0.989 (0.733-1.336)
Average household income (RMB)[d], OR (95% CI)				
<50,000	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
50,001 - 100,000	0.586 (0.426-0.806)[b]	0.567 (0.412-0.780)[c]	0.564 (0.410-0.776)[c]	0.582 (0.422-0.802)[b]

XSL•FO
RenderX

Group	Model 1: Unfairness model	Model 2: Uncertainty model	Model 3: Life stress model	Model 4: Full model
100,001 - 150,000	0.762 (0.556-1.044)	0.733 (0.534-1.006)	0.734 (0.535-1.006)	0.740 (0.539-1.017)
150,001 - 200,000	0.617 (0.456-0.835)[b]	0.605 (0.446-0.820)[b]	0.614 (0.453-0.832)[b]	0.606 (0.446-0.823)[b]
>200,000	1.036 (0.769-1.396)	0.957 (0.709-1.292)	1.006 (0.745-1.357)	0.985 (0.728-1.331)
Unfairness stress, OR (95% CI)				
Low unfairness stress	1.00 (reference)	—[e]	—	1.00 (reference)
Severe unfairness stress	1.813 (1.505-2.183)[c]	—	—	1.306 (1.053-1.620)[b]
Uncertainty stress, OR (95% CI)				
Low uncertainty stress	—	1.00 (reference)	—	1.00 (reference)
Severe uncertainty stress	—	1.971 (1.648-2.357)[c]	—	1.542 (1.248-1.905)[c]
Life stress, OR (95% CI)				
Low life stress	—	—	1.00 (reference)	1.00 (reference)
Severe life stress	—	—	1.848 (1.540-2.216)[c]	1.344 (1.084-1.667)[c]
Fixed parameters, β	0.376[c]	0.334[c]	0.353[c]	0.293[c]

[a]OR: odds ratio.

[b]$P<.05$.

[c]$P<.01$.

[d]1 RMB=US $0.15 at the time of the study.

[e]Not applicable.

Individuals experiencing severe unfairness stress had 30.6% higher odds of developing NCDs compared to those with low uncertainty stress. Those with severe uncertainty stress had 54.2% higher odds and those with severe life stress had 34.4% higher odds. All findings were statistically significant.

Discussion

Principal Findings

The study provided valuable insights into self-reported illness among urban residents in China, particularly concerning perceptions of unfairness, uncertainty, and life stress. Furthermore, this research pioneered an exploration of how perceived unfairness, uncertainty, and life stress relate to self-reported illness, encompassing health conditions such as short-term illness and NCDs.

The Effect of Perceived Unfairness, Uncertainty, and Life Stress on Health Status

Individuals experiencing severe uncertainty (OR 1.230, 95% CI 1.007-1.503) and life stress (OR 1.728, 95% CI 1.411-2.118) were more likely to report health problems, due to the physiological and psychological impacts of prolonged stress [40]. Uncertainty intensified anxiety levels, impeding effective coping strategies and exacerbating stress responses [41]. Chronic life stress contributed to dysregulation of the hypothalamic-pituitary-adrenal axis [42], leading to prolonged secretion of stress hormones like cortisol, increase vulnerability to illnesses [43], contributing to poor self-reported health. Moreover, prolonged stress could disrupt sleep patterns [44], appetite regulation [45], and cardiovascular health [46], further compromising overall well-being.

The Effect of Perceived Unfairness, Uncertainty, and Life Stress on Short-Term Illness

Individuals experiencing severe uncertainty stress (OR 1.565, 95% CI 1.270-1.929) and life stress (OR 1.404, 95% CI 1.136-1.731) were more likely to report higher odds of short-term illness. Uncertainty stress significantly elevated anxiety, disrupted coping mechanisms, and impaired immune function. The heightened stress led to prolonged activation of the hypothalamic-pituitary-adrenal axis [42], resulting in elevated cortisol levels, then altering the gut microbiota composition, weakening the gut barrier and increasing the risk of infections. Furthermore, chronic life stress could exacerbate gut inflammation [47], dysbiosis [48], and disrupt gut motility and transit time, all of these were associated with short-term illness.

The Effect of Perceived Unfairness, Uncertainty, and Life Stress on NCDs

Individuals experiencing severe unfairness (OR 1.306, 95% CI 1.053-1.620), uncertainty (OR 1.542, 95% CI 1.248-1.905), and life stress (OR 1.344, 95% CI 1.084-1.667) exhibited significantly higher odds of NCDs. These stressors contributed to the development of NCDs through a variety of physiological and psychological mechanisms [49]. Unfairness stress could trigger chronic systemic inflammation by activating the release of pro-inflammatory cytokines, which were implicated in the pathogenesis of conditions like cardiovascular disease and diabetes [50]. Moreover, unfairness stress could disrupt metabolic homeostasis, leading to dysregulated glucose metabolism and insulin resistance, further predisposing individuals to diabetes and metabolic syndrome [51].

XSL•FO
RenderX

Uncertainty stress would lead to prolonged stress response: it not only impaired immune function but also promoted oxidative stress and inflammation, contributing to the development of NCDs. Dysregulation of the immune system under uncertainty stress could also impair immune surveillance against cancer cells and increase susceptibility to infections [52], potentially influencing the onset and progression of various NCDs. Additionally, chronic stress disrupted sleep patterns [44] and might lead to unhealthy coping behaviors [53], such as poor diet choices, physical inactivity, and substance abuse, all of which are risk factors for NCDs.

Although some public health professionals and members of the public recognize that perceived stress is associated with adverse health outcomes [54,55], few studies have specifically examined the impact of different stressors on various types of illness. A key finding of this study is that even after controlling for life stress and perceived uncertainty stress, perceived unfairness stress remains significantly associated with a higher prevalence of NCDs. We believe this is closely linked to China's unique sociocultural context. Hofstede's cultural dimension theory identifies China as one of the nations with the highest power distance indices globally [56]. In such a culture, the unequal distribution of power and resources is often regarded as natural and legitimate, leading individuals to internalize power disparities and normalize unfair social relations [57]. It is reported that Chinese residents tend to have a limited perception of unfairness and a high level of acceptance [35]. However, despite this apparent adaptation to unfairness, such dignitary harm [58] threatens the fundamental basis of social respect [59]. Living in a society with low levels of fairness negatively affects psychological well-being by diminishing feelings of social justice, reducing social trust, and weakening reciprocity [60]. As a result, individuals face greater constraints in accessing coping resources, which, according to Conservation of Resources theory, may lead to chronic and long-term harm to physical health [61].

Implications

First, implement programs that provide psychosocial support tailored to address perceived unfairness, uncertainty, and life stress among urban residents, include counseling services, stress management workshops, and community support groups to enhance mental resilience. Second, encourage employers to establish stress management initiatives in workplaces, focusing on mitigating factors such as perceived unfairness and uncertainty, like promoting work-life balance, improving communication channels, and offering employee assistance programs to support mental well-being. Third, community mental health programs should address uncertainty stress through targeted interventions, such as digital and online resources, public awareness campaigns, tailored interventions targeting high-risk groups, as well as crisis helplines.

Limitations

First, this study was a cross-sectional study, and could not clearly indicate the causal relationship between the variables. However, this study was large-scale, and results could provide some implications regarding the important associations between different types of stress and urban residents' health. Second, in this study, an important outcome variable was "Are you suffering from chronic diseases?" No in-depth research was conducted to investigate the type and severity of NCDs, but the data collected still offered a general overview of the prevalence of chronic conditions among the surveyed population, offering valuable insights for policy makers and contributing to a better understanding of the health outcomes of stress. Third, as this study focused on specific urban residents in China, the results might not fully capture the experiences or health outcomes of other regions, which would limit the broader applicability of the conclusions, but it is still useful for other countries at a similar stage of development as China. Finally, the reliance on self-reported health outcomes, which were not validated using clinical measures, may introduce reporting bias due to participants' subjective interpretations or recall errors; however, the use of self-reported data is a widely used method in public health research for capturing individual health awareness and perceived well-being.

Conclusions

In summary, the cross-sectional study clarified the effect of perceived unfairness, uncertainty, and life stress on self-reported illness among Chinese urban residents. The study found significant associations between several stress factors and self-reported health status. Severe uncertainty and life stress were both associated with increased odds of poor health outcomes, while severe unfairness was only associated with NCDs. Future longitudinal research is needed to better understand the temporal dynamics and causal pathways underlying these associations and to further inform effective interventions that decrease stress and promote well-being in urban populations. The findings of this study emphasize the importance of policies that strive for a fairer society, with a great focus on well-being, and may serve as an empirical reference for improving population health among Chinese urban residents.

Acknowledgments

(1) The authors would like to acknowledge that Weifang Zhang was a co-corresponding author of this manuscript. Contact information: Weifang Zhang, PhD, Clinical Research Center for Oral Diseases of Zhejiang Province, Key Laboratory of Oral Biomedical Research of Zhejiang Province, Cancer Center of Zhejiang University, Stomatology Hospital, School of Stomatology, School of Medicine, Zhejiang University, 866 Yuhangtang Road, Hangzhou, Zhejiang, China. chzwf@zju.edu.cn. 86 15000090389.

(2) This study was partly funded by Nature Science Foundation of Zhejiang Province of China (LQ20G030014) and Social Science Foundation of Ministry of Education of China (23YJCZH090).

(3) The research team greatly appreciates the funding support, and the research participants for their cooperation and support.

XSL•FO
RenderX

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy restrictions but are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

References

1. Frankenhaeuser M. A psychobiological framework for research on human stress and coping. In: Appley MH, Trumbull R, editors. Dynamics of Stress: Physiological, Psychological and Social Perspectives: Springer US; 1986:101-116.
2. Östberg V, Almquist YB, Folkesson L, Låftman SB, Modin B, Lindfors P. The complexity of stress in mid-adolescent girls and boys. Child Ind Res 2015 Jun;8(2):403-423. [doi: 10.1007/s12187-014-9245-7]
3. Neyazi A, Mohammadi B, Griffiths MD. The impact of the 2023 earthquakes on the Afghan health-care system. Lancet 2023 Nov 18;402(10415):1829-1830. [doi: 10.1016/S0140-6736(23)02344-9] [Medline: 37980084]
4. Wolf K, Schmitz J. Scoping review: longitudinal effects of the COVID-19 pandemic on child and adolescent mental health. Eur Child Adolesc Psychiatry 2024 May;33(5):1257-1312. [doi: 10.1007/s00787-023-02206-8] [Medline: 37081139]
5. Neyazi A, Mohammadi AQ, Razaqi N, et al. Health survey on anxiety, depression, and stress in Afghanistan: a large-scale cross-sectional study amid ongoing challenges. Discov Ment Health 2024 Sep 20;4(1):38. [doi: 10.1007/s44192-024-00090-5] [Medline: 39302527]
6. Monroe SM, Simons AD. Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. Psychol Bull 1991 Nov;110(3):406-425. [doi: 10.1037/0033-2909.110.3.406] [Medline: 1758917]
7. Yang T, Yang XY, Yu L, Cottrell RR, Jiang S. Individual and regional association between socioeconomic status and uncertainty stress, and life stress: a representative nationwide study of China. Int J Equity Health 2017 Jul 5;16(1):118. [doi: 10.1186/s12939-017-0618-7] [Medline: 28679409]
8. Wang H, Yang XY, Yang T, et al. Socioeconomic inequalities and mental stress in individual and regional level: a twenty one cities study in China. Int J Equity Health 2015 Mar 7;14:25. [doi: 10.1186/s12939-015-0152-4] [Medline: 25889251]
9. Lyons CE, Razzoli M, Bartolomucci A. The impact of life stress on hallmarks of aging and accelerated senescence: connections in sickness and in health. Neurosci Biobehav Rev 2023 Oct;153:105359. [doi: 10.1016/j.neubiorev.2023.105359] [Medline: 37586578]
10. Yang Y, Liu X, Liu ZZ, Tein JY, Jia CX. Life stress, insomnia, and anxiety/depressive symptoms in adolescents: a three-wave longitudinal study. J Affect Disord 2023 Feb;322:91-98. [doi: 10.1016/j.jad.2022.11.002]
11. Liu X, Yang Y, Liu ZZ, Jia CX. Life stress and suicidality mediated by insomnia and depressive symptoms in adolescents: a three-wave longitudinal study. Sleep 2024 Mar 11;47(3):zsad121. [doi: 10.1093/sleep/zsad121] [Medline: 37075813]
12. Jacobs MA, Spilken AZ, Norman MM, Anderson LS. Life stress and respiratory illness. Psychosom Med 1970 May;32(3):233-242. [doi: 10.1097/00006842-197005000-00003]
13. Tennant C. Life stress and hypertension. Eur J Cardiovasc Prev Rehabil 2001 Feb 1;8(1):51-56. [doi: 10.1177/174182670100800107]
14. Neyazi A, Mohammadi AQ, Neyazi M, Timilsina S, Padhi BK, Griffiths MD. Hypertension, depression, and health-related quality of life among hospitalized patients in Afghanistan. J Hum Hypertens 2024 Jun;38(6):529-537. [doi: 10.1038/s41371-024-00914-5] [Medline: 38684866]
15. Lee MA, Kawachi I. Perceived unfairness and psychological distress: less harmful as age increases? Soc Just Res 2019 Mar;32(1):1-25. [doi: 10.1007/s11211-019-00325-9]
16. Jackson B, Kubzansky LD, Wright RJ. Linking perceived unfairness to physical health: the Perceived Unfairness Model. Rev Gen Psychol 2006 Mar;10(1):21-40. [doi: 10.1037/1089-2680.10.1.21]
17. Walton E. The inner level: how more equal societies reduce stress, restore sanity and improve everyone's well-being: inequality as it affects individuals. Br J Gen Pract 2019 Sep;69(686):452-453. [doi: 10.3399/bjgp19X705377] [Medline: 31467020]
18. Greenberg J. Organizational injustice as an occupational health risk. Ann R Coll Physicians Surg Can 2010 Jan;4(1):205-243. [doi: 10.5465/19416520.2010.481174]
19. Robbins JM, Ford MT, Tetrick LE. Perceived unfairness and employee health: a meta-analytic integration. J Appl Psychol 2012 Mar;97(2):235-272. [doi: 10.1037/a0025408] [Medline: 21928872]
20. Resnicow K, Patel M, Green M, et al. The association of unfairness with mental and physical health in a multiethnic sample of adults: cross-sectional study. JMIR Public Health Surveill 2021 May 10;7(5):e26622. [doi: 10.2196/26622] [Medline: 33970121]
21. De Vogli R, Brunner E, Marmot MG. Unfairness and the social gradient of metabolic syndrome in the Whitehall II Study. J Psychosom Res 2007 Oct;63(4):413-419. [doi: 10.1016/j.jpsychores.2007.04.006] [Medline: 17905050]
22. Tingzhong Y. Health Research: Social and Behavioral Theory and Methods: People's Medical Publishing House; 2018.

XSL•FO
RenderX

23. Mishel MH. Perceived uncertainty and stress in illness. Res Nurs Health 1984 Sep;7(3):163-171. [doi: 10.1002/nur.4770070304] [Medline: 6567948]
24. Small SP, Graydon JE. Perceived uncertainty, physical symptoms, and negative mood in hospitalized patients with chronic obstructive pulmonary disease. Heart Lung 1992;21(6):568-574. [Medline: 1447004]
25. Greco V, Roger D. Uncertainty, stress, and health. Pers Individ Dif 2003 Apr;34(6):1057-1068. [doi: 10.1016/S0191-8869(02)00091-0]
26. Nasr S, Heneiti A, Hammoudeh W. Life conditions and quality of life of Palestinians living in the Biddu Enclave: a qualitative study. Lancet 2021 Jul;398 Suppl 1:S41. [doi: 10.1016/S0140-6736(21)01527-0] [Medline: 34227975]
27. Peng S, Yang T, Rockett IRH. Life stress and uncertainty stress: which is more associated with unintentional injury? Psychol Health Med 2020 Jul;25(6):774-780. [doi: 10.1080/13548506.2019.1687913] [Medline: 31684773]
28. Wu D, Rockett IRH, Yang T, Feng X, Jiang S, Yu L. Deliberate self-harm among Chinese medical students: a population-based study. J Affect Disord 2016 Sep 15;202:137-144. [doi: 10.1016/j.jad.2016.05.030] [Medline: 27262635]
29. Wu D, Yang T, Rockett IR, Yu L, Peng S, Jiang S. Uncertainty stress, social capital, and suicidal ideation among Chinese medical students: findings from a 22-university survey. J Health Psychol 2021 Feb;26(2):214-225. [doi: 10.1177/1359105318805820] [Medline: 30318917]
30. Yang T, Barnett R, Peng S, Yu L, Zhang C, Zhang W. Individual and regional factors affecting stress and problem alcohol use: A representative nationwide study of China. Health Place 2018 May;51:19-27. [doi: 10.1016/j.healthplace.2018.02.008] [Medline: 29518714]
31. Wu D, Yang T. Late bedtime, uncertainty stress among Chinese college students: impact on academic performance and self-rated health. Psychol Health Med 2023 Dec;28(10):2915-2926. [doi: 10.1080/13548506.2022.2067337] [Medline: 35437084]
32. Zhu Z, Wu D, Wei K, et al. Uncertainty stress and self-rated health during the early stage of the COVID-19 outbreak. Health Psychol Behav Med 2023;11(1):2173202. [doi: 10.1080/21642850.2023.2173202] [Medline: 36818392]
33. Peng S, Yang XY, Yang T, Zhang W, Cottrell RR. Uncertainty stress, and its impact on disease fear and prevention behavior during the COVID-19 epidemic in China: a panel study. Am J Health Behav 2021 Mar 1;45(2):334-341. [doi: 10.5993/AJHB.45.2.12] [Medline: 33888193]
34. Ben Salah A, DeAngelis BN, al'Absi M. Uncertainty and psychological distress during COVID-19: what about protective factors? Curr Psychol 2022 May 28:1-8. [doi: 10.1007/s12144-022-03244-2] [Medline: 35669208]
35. Whyte M. Fair versus unfair: how do Chinese citizens view current inequalities? In: Oi J, Rozelle S, Zhou X, editors. Growing Pains: Tensions and Opportunity in China's Transformation: Stanford: Shorenstein Center; 2010. URL: https://martinwhyte.scholars.harvard.edu/sites/g/files/omnuum2491/files/martinwhyte/files/fair_versus_unfair.pdf [accessed 2025-06-20]
36. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav 1997 Mar;38(1):21-37. [Medline: 9097506]
37. Yang T, Jiang S, Yu L, Cottrell RR, Si Q. Life stress, uncertainty stress and self-reported illness: a representative nationwide study of Chinese students. J Public Health 2018 Apr;26(2):205-209. [doi: 10.1007/s10389-017-0837-9]
38. Yang T, Barnett R, Fan Y, Li L. The effect of urban green space on uncertainty stress and life stress: a nationwide study of university students in China. Health Place 2019 Sep;59:102199. [doi: 10.1016/j.healthplace.2019.102199] [Medline: 31514059]
39. Wu D, Yang T, Herold F, et al. Validation of the 4-item and 10-item uncertainty stress scale in a community-based sample of Chinese adults. Psychol Res Behav Manag 2022;15:2803-2813. [doi: 10.2147/PRBM.S379180] [Medline: 36193332]
40. Kaplan GA, Roberts RE, Camacho TC, Coyne JC. Psychosocial predictors of depression. Prospective evidence from the human population laboratory studies. Am J Epidemiol 1987 Feb;125(2):206-220. [doi: 10.1093/oxfordjournals.aje.a114521] [Medline: 3812429]
41. Germani A, Buratta L, Delvecchio E, Gizzi G, Mazzeschi C. Anxiety severity, perceived risk of COVID-19 and individual functioning in emerging adults facing the pandemic. Front Psychol 2020;11:567505. [doi: 10.3389/fpsyg.2020.567505] [Medline: 33364996]
42. Poitras M, Lebeau M, Plamondon H. The cycle of stress: a systematic review of the impact of chronic psychological stress models on the rodent estrous cycle. Neurosci Biobehav Rev 2024 Jul;162:105730. [doi: 10.1016/j.neubiorev.2024.105730] [Medline: 38763179]
43. Feng X, Ma X, Li J, et al. Inflammatory pathogenesis of post-stroke depression. Aging Dis 2024 Feb 9;16(1):209-238. [doi: 10.14336/AD.2024.0203] [Medline: 38377025]
44. Liu Z, Liu R, Zhang Y, et al. Association between perceived stress and depression among medical students during the outbreak of COVID-19: the mediating role of insomnia. J Affect Disord 2021 Sep 1;292:89-94. [doi: 10.1016/j.jad.2021.05.028] [Medline: 34107425]
45. Pani L, Porcella A, Gessa GL. The role of stress in the pathophysiology of the dopaminergic system. Mol Psychiatry 2000 Jan;5(1):14-21. [doi: 10.1038/sj.mp.4000589] [Medline: 10673764]

XSL•FO
RenderX

46. Natt Och Dag Y, Mehlig K, Rosengren A, Lissner L, Rosvall M. Negative emotional states and negative life events: Consequences for cardiovascular health in a general population. J Psychosom Res 2020 Feb;129:109888. [doi: 10.1016/j.jpsychores.2019.109888] [Medline: 31835155]
47. Madison A, Kiecolt-Glaser JK. Stress, depression, diet, and the gut microbiota: human–bacteria interactions at the core of psychoneuroimmunology and nutrition. Curr Opin Behav Sci 2019 Aug;28:105-110. [doi: 10.1016/j.cobeha.2019.01.011]
48. Simpson CA, Diaz-Arteche C, Eliby D, Schwartz OS, Simmons JG, Cowan CSM. The gut microbiota in anxiety and depression - a systematic review. Clin Psychol Rev 2021 Feb;83:101943. [doi: 10.1016/j.cpr.2020.101943] [Medline: 33271426]
49. Norris SA, Anuar H, Matzen P, Cheah JCH, Jensen BB, Hanson M. The life and health challenges of young Malaysian couples: results from a stakeholder consensus and engagement study to support non-communicable disease prevention. BMC Public Health 2014;14(Suppl 2):S6. [doi: 10.1186/1471-2458-14-S2-S6] [Medline: 25080995]
50. Littrell J. New developments in understanding cardiovascular disease and the implications for social work. Soc Work Health Care 2008;46(2):35-49. [doi: 10.1300/j010v46n02_03] [Medline: 18192196]
51. Kaufman D, Banerji MA, Shorman I, et al. Early-life stress and the development of obesity and insulin resistance in juvenile bonnet macaques. Diabetes 2007 May;56(5):1382-1386. [doi: 10.2337/db06-1409] [Medline: 17470564]
52. Kivimäki M, Bartolomucci A, Kawachi I. The multiple roles of life stress in metabolic disorders. Nat Rev Endocrinol 2023 Jan;19(1):10-27. [doi: 10.1038/s41574-022-00746-8] [Medline: 36224493]
53. Jackson JS, Knight KM, Rafferty JA. Race and unhealthy behaviors: chronic stress, the HPA axis, and physical and mental health disparities over the life course. Am J Public Health 2010 May;100(5):933-939. [doi: 10.2105/AJPH.2008.143446] [Medline: 19846689]
54. López-Del-Hoyo Y, Fernández-Martínez S, Perez-Aranda A, et al. Effectiveness of a web-based self-guided intervention (MINDxYOU) for reducing stress and promoting mental health among health professionals: results from a stepped-wedge cluster randomized trial. J Med Internet Res 2025 Feb 3;27:e59653. [doi: 10.2196/59653] [Medline: 39899345]
55. Sasaki N, Watanabe K, Egawa M, et al. Job stress, psychological distress, and menstruation-related symptoms in female workers: a cross-sectional study. BJOG 2025 Mar 28. [doi: 10.1111/1471-0528.18153] [Medline: 40151925]
56. Country comparison tool. The Culture Factor Group. URL: https://www.hofstede-insights.com/country-comparison-tool [accessed 2025-05-22]
57. Hofstede G. Culture′s Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations: Sage Publications; 2001.
58. Marmot M. Dignity and inequality. Lancet 2004;364(9439):1019-1021. [doi: 10.1016/S0140-6736(04)17075-X] [Medline: 15380951]
59. Ford MT. Perceived unfairness at work, social and personal resources, and resting blood pressure. Stress Health 2014 Feb;30(1):12-22. [doi: 10.1002/smi.2491] [Medline: 23625544]
60. You JS. Social trust: fairness matters more than homogeneity. Polit Psychol 2012 Jun 18;33(5):701-721 [FREE Full text] [doi: 10.1111/j.1467-9221.2012.00893.x]
61. Hobfoll SE. Conservation of resources theory: its implication for stress, health, and resilience. In: The Oxford Handbook of Stress, Health, and Coping: Oxford University Press; 2011:127-147.

Abbreviations

NCD: noncommunicable chronic diseases
OR: odds ratio

Edited by A Mavragani; submitted 03.02.25; peer-reviewed by A Neyazi, E Hannah; revised version received 31.03.25; accepted 02.04.25; published 14.08.25.

Please cite as:
Chen J, Jiang S, Pang M, Zhou H, Zhang W
Impact of Perceived Unfairness, Uncertainty, and Life Stress on Urban Residents' Health in China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e72094
URL: https://publichealth.jmir.org/2025/1/e72094
doi:10.2196/72094

© Jinsong Chen, Shuhan Jiang, Mingli Pang, Huan Zhou, Weifang Zhang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly

XSL•FO
RenderX

cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Knowledge and Beliefs Toward Mammography Screening Among Jordanian Women: Cross-Sectional Study

Ahmad Shaker Abuabed[1], MPH; Hana Taha[2,3], PhD; Luz María Garcia-Valdes[1], PhD; Mohammad Al-Share[4], MD; Hadi Al-shar'e[5], MD; Carmen Amezcua-Prieto[1,6,7], PhD

[1]Department of Preventive Medicine and Public Health, Faculty of Medicine, University of Granada, Granada, Spain

[2]Department of Family and Community Medicine, School of Medicine, University of Jordan, Amman, Jordan

[3]Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Nobels väg 6, Stockholm, Sweden

[4]General Surgery Department, General Surgery Department, Al-Bashir Hospital, Amman, Jordan

[5]General Practitioner, Private Clinic, Amman, Jordan

[6]Madrid, CIBER of Epidemiology and Public Health, Carlos III Health Institute, Granada, Spain

[7]Instituto de Investigación Biosanitaria de Granada (IBS.GRANADA), Granada, Spain

Corresponding Author:
Hana Taha, PhD
Department of Family and Community Medicine, School of Medicine, University of Jordan, Amman, Jordan

Abstract

Background: Breast cancer (BC) is the most commonly diagnosed cancer and the leading cause of cancer-related deaths among women globally. Despite the significance of mammography screening rate for early BC detection among Jordanian women, it remains low, mainly due to various cognitive, psychosocial, and behavioral barriers. Understanding these factors is essential for developing effective interventions.

Objective: This study aims to assess the BC knowledge and beliefs about mammography screening among Jordanian women aged 40 years and older based on the Health Belief Model (HBM) as a theoretical framework.

Methods: A cross-sectional design was used with a convenience sample of eligible women from a Jordanian public hospital. Data were collected through face-to-face interviews using a validated Arabic structured questionnaire consisting of 3 sections: sociodemographic data, knowledge about BC, and health beliefs about mammography. Descriptive statistics and multivariate analysis of variance (MANOVA) were conducted using IBM SPSS version 28.

Results: A total of 405 women completed the study, with an average (SD) age of 52.4 (8.57) years. Findings revealed a notably low knowledge level, as participants scored an average (SD) of 5.80 (2.64) out of 12. The average (SD) scores for the health beliefs section (out of 5) were also low: perceived benefits, 2.59 (0.59); perceived barriers, 2.48 (0.71); and health motivation, 2.51 (0.71). Significant associations (P<.001) with medium to large effect sizes (ηp^2>0.06) were observed between participants' age and education level in relation to BC knowledge and health beliefs regarding mammography screening. Participants cited several reasons for their reluctance to undergo mammography, including a lack of knowledge (72.8%), cultural beliefs (63%), and religious factors (29.4%).

Conclusions: A significant gap exists in BC knowledge and beliefs about mammography among Jordanian women aged 40 years and older. Policy makers and health care providers should prioritize the development of tailored strategies and context-specific, sensitive educational interventions. These efforts should address the unique needs, cultural beliefs, and awareness levels of this population to improve mammogram screening practices in Jordan.

(JMIR Public Health Surveill 2025;11:e75384) doi:10.2196/75384

KEYWORDS

breast cancer; mammogram screening; beliefs; champion's health belief model; knowledge

Introduction

Background

Cancer has become a significant global public health concern, surpassing cardiovascular diseases as the leading cause of death in many high-income countries [1]. Breast cancer (BC) is the most diagnosed, with around 2.3 million new cases and 670,000 deaths in 2022 [2-4], making it the leading cause of cancer-related death among women globally [3,4]. Although the incidence of BC is lower in high-income countries, it is often diagnosed at advanced stages, leading to higher mortality rates [5].

XSL•FO
RenderX

As a low-income country, the World Bank classifies Jordan as a lower-middle-income country (LMIC) [6]. It confronts substantial health care resource management challenges, particularly for noncommunicable diseases like cancer [5,7]. BC is the most common cancer among women in Jordan, accounting for 39.2% of female cancer cases in 2022. The age-standardized rate (ASR) of BC is 60 per 100,000 population, which is higher than in some neighboring Arab countries but lower than in Western countries. The global ASR average is 46.8 [2,8,9]. The rising incidence of BC in LMICs is linked to increased life expectancy, improved health care services, and increased health insurance coverage. In addition, the adoption of Western lifestyles and changes in reproductive behaviors, such as delayed childbirth and shorter breastfeeding durations, have also contributed to this trend [10,11].

The American College of Radiology (ACR) recommends that all women at average risk of BC should begin annual mammograms at the age of 40 years. Moreover, women who are at higher-than-average risk due to various factors are recommended to undergo more intensive screening [12]. Ultimately, mammography screening can reduce BC mortality by 40% or more in women aged 40 years and older [12]. Early detection of BC through mammography can lead to significantly improved survival rates [12,13]. In response to this, Jordan established the Jordan Cancer Registry in 1996, the King Hussein Cancer Foundation and Center in 2001, and the Jordan Breast Cancer Program (JBCP) in 2007 to coordinate detection efforts, provide screening, and raise public awareness [11]. It also recommends that women begin mammogram screenings at the age of 40 years [9,14].

Despite these national initiatives, participation in mammography screening remains low. According to the Jordan Cancer Registry, only 18.6% of women aged 40-69 years have undergone mammographic screening in the past 2 years [15]. Another survey indicated that just 14% of women aged 40-49 years have ever had a mammogram [16]. These rates are comparable to those in neighboring countries, such as Qatar (22.5%) and Saudi Arabia (27.7%) [17,18]. However, they are still significantly lower than the rates reported in high-income countries, such as France (48.5%) and the United States (61%) [19,20]. This low participation persists despite the expansion of screening services and public awareness campaigns, indicating ongoing barriers to preventive care [7,11,14].

One of the main challenges and barriers is that many Jordanian women are unaware of their eligibility for screening due to a lack of awareness, absence of health problems, fear of diagnosis, and the influence of cultural and religious beliefs [16,21]. Although educational interventions have been implemented, evidence suggests that simply increasing awareness is insufficient to effect meaningful change in screening behaviors [21-23]. Deeply rooted beliefs and attitudes, shaped by cultural and individual factors, continue to hinder adherence to breast cancer screening (BCS) guidelines [24,25]. Furthermore, health beliefs regarding mammography among this age group in Jordan are underresearched [26,27].

To address these gaps, this study aims to investigate how knowledge and health beliefs about mammography screening practices based on the Health Belief Model (HBM) among Jordanian women aged 40 years and older. The study also seeks to identify the sociodemographic characteristics that influence knowledge and beliefs regarding screening.

Theoretical Framework

This study is guided by the HBM, a theoretical framework that explains preventive health behaviors in various contexts, such as vaccination and smoking [28,29]. The HBM suggests that individuals' actions are influenced by their perceptions of susceptibility, severity, benefits, barriers, cues for action, and self-efficacy [28-30].

Victoria Champion developed the Champion Health Belief Model (CHBM) as an adaptation of the original HBM to better understand and predict women's participation in BCS behaviors [31,32]. Champion initially created separate scales to measure constructs such as perceived susceptibility, benefits, and barriers related to BCS [31,32]. These scales were later combined and revised to form the Champion Health Belief Model Scale (CHBMS) [33,34]. This comprehensive instrument emphasizes health motivation and self-efficacy in women's decisions to participate in BCS, while also taking into account cultural factors [33,34]. The CHBMS has been adapted and validated in various cultural contexts to assess the beliefs and attitudes that influence BCS behaviors [22,26,34,35].

In this study, we used 3 constructs of the CHBMS: perceived benefits, barriers, and health motivation. These constructs were selected based on the study conducted by Petro-Nustas [26], who revised these specific aspects of the CHBMS [22,33]. This psychometrically validated tool has been used to assess the beliefs of Jordanian women regarding mammography.

Methods

Study Design

This study implemented a cross-sectional design to assess the knowledge and beliefs about BC and mammography among Jordanian women aged 40 years and older. It adhered to the Strengthening of the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (Checklist 1) [36]. The HBM served as the theoretical framework for this study, as it has been identified as the most frequently used and suitable theoretical health behavior model in previous studies that explain women's BCS practices [37,38].

Study Setting

This study was conducted at Albashir Hospital in Amman, Jordan, the largest public medical complex in the country. It is a national referral center comprising multiple specialized hospitals and centers. The sample included women from different socioeconomic and clinical backgrounds. In 2022, the hospital's outpatient clinics received a total of 632,352 patients who visited the hospital and the outpatient clinics. Participant recruitment and data collection took place from women attending various outpatient clinics between May and July 2024 during regular clinic hours (8:00 AM to 4:00 PM, Sunday through Thursday).

XSL•FO
RenderX

Participants

Eligibility Criteria

Women of Jordanian nationality who are 40 years of age and older were eligible to participate in this study, provided they met specific criteria. Participants were required to demonstrate sufficient literacy in Arabic, have not received mammography screening within the previous 12 months, and be free from any physical or mental disabilities that could impede independent participation or reliable self-reporting. To maintain the study's focus on screening behaviors, individuals with a personal history of BC or those who have a first-degree relative (parent, sibling, or child) diagnosed with BC were excluded from participation.

Selection and Recruitment

A convenience sampling strategy was used, with a multiphase approach to uphold ethical standards, ensure participant privacy, and maintain methodological rigor. Before the recruitment process began, all health care providers involved in the recruitment and data collection team participated in a 3-day training in May 2024. This training addressed the study's methodology, objectives, questionnaire elements, and ethical considerations to ensure consistent data collection and minimize interviewer bias.

Each morning, the appointment schedules at the various outpatient clinics were reviewed to identify women eligible for the study, thereby ensuring a diverse population and minimizing selection bias. Team members approached these women in the waiting areas, explaining the research and inviting them to participate while ensuring that they minimized any disruption to their time. Interested women were then escorted to a private meeting room, where they received detailed information and reassurances that their participation would not impact their health care.

Women had the right to decline or provide written consent if they agreed to participate. After obtaining consent, face-to-face interviews were conducted by the same team member who initiated recruitment, ensuring consistency and rapport with the participants. To support women with low literacy levels, each interview followed a structured questionnaire format. The interviewer offered additional clarification and read the questions aloud as needed, ensuring that all participants, regardless of their literacy level, could fully understand and answer each item. The interviews, conducted in Arabic, typically lasted between 10 and 15 minutes to reduce information bias.

Data Sources and Measurement

The variables of interest in this study—sociodemographic data, knowledge of BC, and health beliefs about mammography screening—were clearly defined and systematically measured. Data for all variables were collected using a questionnaire. The questionnaire consisted of 3 sections: sociodemographic characteristics, knowledge about BC, and health beliefs, measured using 3 constructs from the CHBMS.

The sociodemographic data section included questions about participants' age, education, marital status, health insurance status, and mammography screening history in the past 12 months (these questions were used to assess participants' eligibility). In addition, participants were asked an open-ended question to provide up to 4 reasons for not having a mammogram, offering qualitative insight into the barriers to screening, with no specific order required.

The knowledge section featured 12 true, false, or do not know questions about BC, with 7 correct answers marked as "true" and 5 as "false." Participants scored 1 point for each correct answer and zero for each incorrect or "do not know" answer. The maximum possible score was 12, indicating excellent knowledge, while a score of zero indicated limited knowledge.

The health belief section consisted of 3 constructs from the CHBMS, comprising 18 statements: 6 related to perceived benefits, 5 to perceived barriers, and 7 to health motivation. The scale is positively associated with screening behaviors, except for perceived barriers, which are negatively associated. Each statement was rated on a 5-point Likert scale (1=strongly disagree to 5=strongly agree), with reversed scoring applied to the barriers. Scores ranged from 18 (highly negative beliefs) to 90 (highly positive beliefs).

In the previous section on health beliefs, we used a culturally adapted Arabic version of the CHBMS (Multimedia Appendix 1), initially translated and validated for BCS among Jordanian women by Mikhail and Petro-Nustas in 2001 [35]. Petro-Nustas later revised and validated the entire questionnaire to better reflect the health beliefs of Jordanian women regarding mammography, focusing on 3 constructs from the CHBMS: perceived benefits, perceived barriers, and health motivation [22]. Our study followed this methodology and obtained permission from Petro-Nustas and Champion to use the questionnaire (Multimedia Appendix 2).

Study Size

The study size was determined based on previous relevant studies [14,24,27] 2727 and calculated using the Raosoft sample size calculator (Raosoft, Inc) [39]. The target population consisted of Jordanian women aged 40 years or older, totaling approximately 1,460,300 in 2023. Based on a 95% confidence level, a 5% margin of error, and a response distribution of 50%, the minimum required sample size was determined to be 384 participants. A total of 405 women were recruited to mitigate potential nonresponses, attrition, and data exclusion while preserving statistical power.

Statistical Analysis

Descriptive statistics were calculated, including the mean and SD for continuous data and frequencies and percentages for categorical data. A Multivariate Analysis of Variance (MANOVA) was conducted to examine the association between sociodemographic independent variables: age, educational level, marital status, and health insurance, and the dependent variables: knowledge and health beliefs based on the HBM, specifically in terms of perceived benefits, perceived barriers, and health motivation. Following significant results from the MANOVA (P<.05), indicating an association between the dependent and independent variables, Bonferroni post hoc tests were conducted to determine which specific groups within each dependent variable showed statistically significant differences (P<.001); to quantify the magnitude of effects beyond statistical

XSL•FO
RenderX

significance, partial eta-squared (ηp^2) values were calculated for each independent and dependent variable relationship. All statistical analyses were performed using IBM SPSS version 28 [40].

Ethical Considerations

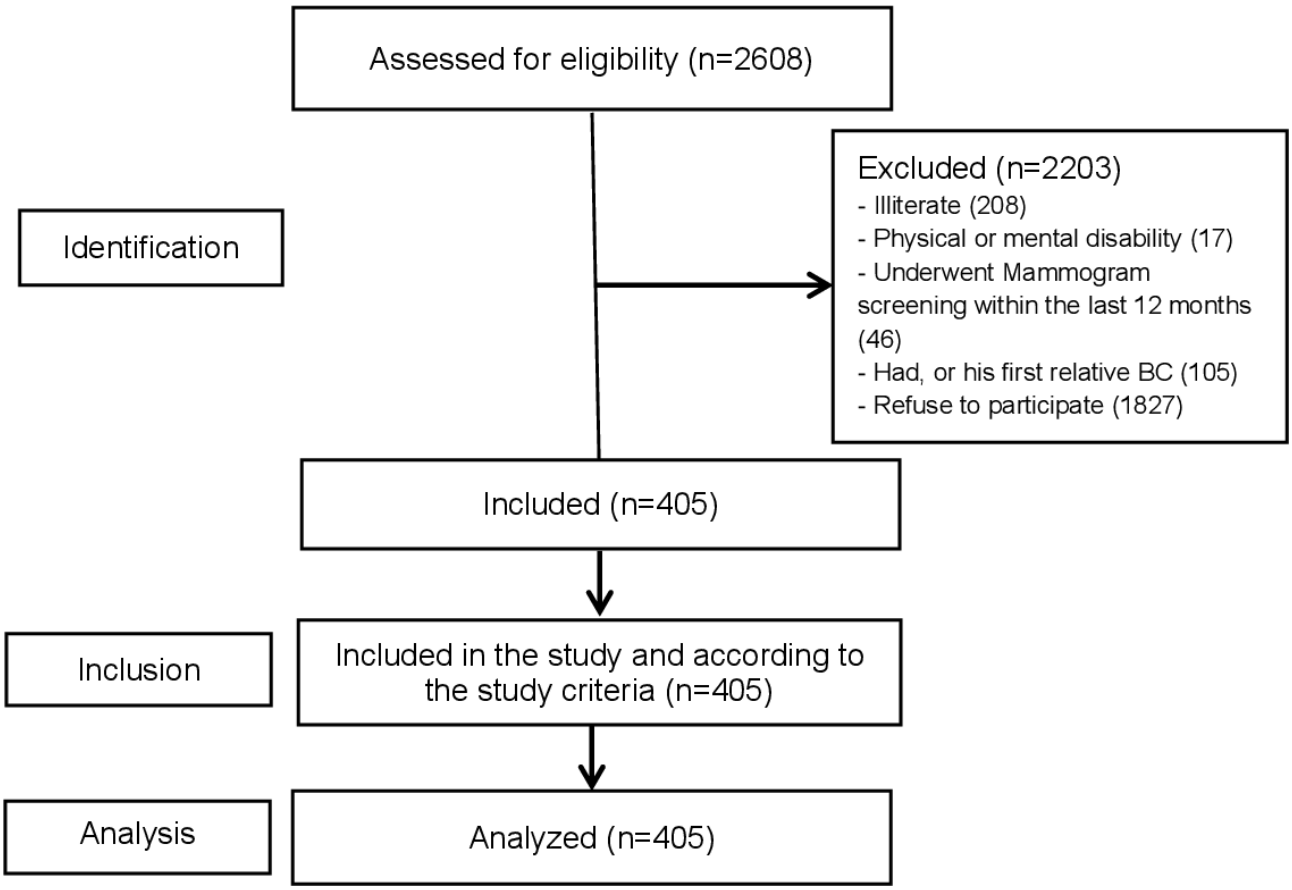
This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Board (IRB) of the Ministry of Health and Albashir Hospital in Jordan (IRB-7790) in May 2024 (Multimedia Appendix 3). Written consent was obtained from all participants before their participation (Multimedia Appendix 4). They were provided with detailed information about the study's aims and significance, as well as their right to withdraw at any time without any obligations. To ensure privacy, no personal identifiers were used; participants were assigned unique code numbers to maintain confidentiality. No financial compensation or incentives were provided to participants.

Results

Overview

Out of 2608 women screened at Albashir Hospital outpatient clinics, 376 were excluded, and 1827 refused to participate, resulting in 405 eligible participants who consented and completed the study. This process is outlined in the STROBE flow diagram [36] as shown in Figure 1.

Figure 1. Strengthening the Reporting of Observational studies in Epidemiology flow diagram



Sociodemographic Characteristics

A total of 405 Jordanian women participated in the study, aged 40-76 years. The vast majority, 187 (46.2%), fell within the 40 - 49 years age group, 173 (42.7%) completed secondary education, 213 (52.6%) were married, 372 (91.9%) reported having health insurance, and none had undergone mammogram screening in the past 12 months (Table 1).

XSL•FO
RenderX

Table . Sociodemographic characteristics of the study participants (N=405).

Variables	Frequency, n (%)
Age groups	
40 - 49 years	187 (46.2)
50 - 59 years	140 (34.6)
60 years and older	78 (19.2)
Education level	
Primary	123 (30.4)
Secondary	173 (42.7)
University	109 (26.9)
Marital status	
Single	104 (25.7)
Married	213 (52.6)
Divorced	37 (9.1)
Widowed	51 (12.6)
Do you have any health insurance?	
Yes	372 (91.9)
No	33 (8.1)
Mammogram screening within the last 12 months	
Yes	0 (0)
No	405 (100)

The participants' responses to the open-ended questions were analyzed and categorized into 8 main reasons that may negatively impact screening practices. An example of the main reasons reported among the 405 participants was a lack of knowledge, cited by 72.8% (295). This was followed by fear of results, reported by 66.4% (269), cultural beliefs at 63% (255), and religious reasons, mentioned by 29.4% (119) (Multimedia Appendix 5).

Knowledge About Breast Cancer

Table 2 shows that the overall knowledge score about BC was relatively low (48.3%), with a mean (SD) score of 5.80 (2.64). Regarding knowledge aspects, approximately out of 405, three-quarters of the sample, 74.3% (301), correctly recognized that BC is the most common cancer among women. In addition, half of the sample, 206 (50.9%), identified excess weight and having a first child after age 30 years as risk factors for BC. In contrast, only 145 (35.8%) correctly identified that most breast tumors are not cancerous.

XSL•FO
RenderX

Table . Breast cancer knowledge questions with correct answers (N=405).

Item	Correct answer	Correct answer, n (%)
Myths and misconceptions		
A strong blow to the chest can cause breast cancer after a period of time.	False	192 (47.4)
The constant discomfort of a tight bra may cause breast cancer after a period of time.	False	176 (43.5)
Some types of fibrous sacs in the breast (noncancerous tumors) increase the risk of breast cancer.	False	176 (43.5)
Women over the age of 70 rarely have breast cancer.	False	186 (45.9)
Most breast tumors are cancerous.	False	145 (35.8)
Risk factors		
Excess weight may cause breast cancer to some women.	True	195 (48.1%)
A woman who conceives her first child after the age of thirty is more likely to have breast cancer than a woman who conceives her first child before the age of thirty.	True	206 (50.9%)
Women in the United States are more likely to have breast cancer than Asian or African women.	True	167 (41.2)
Breast cancer is more common among women at age 65 than at age 40.	True	188 (46.4)
The most common cancer among women is breast cancer.	True	301 (74.3)
Physical exercises reduce the risk of breast cancer.	True	202 (49.9)
Diet affects the risk of breast cancer.	True	214 (52.8)

Beliefs About Mammography

Scores for each construct and question within the CHBM scale ranged from 1 to 5. The perceived benefits of mammography had a relatively low mean (SD) score of 2.59 (0.59). Among the perceived benefits, women identified that mammography helps in the early detection of lumps, with a mean (SD0 score of 2.62 (1.08). The perception that mammography decreases the chances of dying from BC had a mean (SD0 score of 2.60 (1.13). Furthermore, the item related to reducing worry about BC was rated the lowest, with a mean (SD) score of 2.45 (1.19) (Table 3).

XSL•FO
RenderX

Table . Breast cancer knowledge and health beliefs scores based on the Champion Health Belief Model scale.

Items	Mean (SD)
Knowledge	5.80[a] (2.64)
CHBM[b] construct	
Perceived benefits	2.59[c] (0.59)
When I get a recommended mammogram, I feel good about myself.	2.62 (1.10)
When I get a mammogram, I don't worry as much about breast cancer.	2.45 (1.19)
Having a mammogram of the breast will help me find lumps early.	2.62 (1.08)
Having a mammogram of the breast will decrease my chances of dying from breast cancer.	2.60 (1.13)
Having a mammogram of the breast will decrease my chances of requiring radical or disfiguring surgery if breast cancer occurs.	2.60 (1.14)
Having a mammogram will help me find a lump before it can be felt by myself or a health care professional.	2.62 (1.03)
Perceived barriers[d]	2.48 (0.71)
Having a routine mammogram of the breast would make me worry about breast cancer.	2.52 (0.92)
Having a mammogram of the breast would be embarrassing.	2.38 (1.00)
Having a mammogram of the breast would take too much time.	2.49 (0.96)
Having a mammogram of the breast would be painful.	2.48 (1.02)
Having a mammogram of the breast would cost too much money.	2.54 (1.05)
Health Motivation	2.51 (0.71)
I want to discover health problems early.	2.64 (1.09)
Maintaining good health is extremely important to me.	2.62 (1.12)
I search for new information to improve my health.	2.64 (1.12)
I feel it is important to carry out activities which will improve my health.	2.63 (1.12)
I eat well-balanced meals.	2.61 (1.13)
I exercise at least 3 times a week.	2.31 (0.95)
I have regular health check-ups, even when I'm not sick.	2.11 (1.08)

[a]Mean score for the knowledge item out of 12.

[b]CHBM: Champion Health Belief Model

[c]Mean score for all items of perceived benefits, barriers, and health motivation out of 5.

[d]Reversed score: a lower score indicates higher perceived barriers.

The perceived barriers to undergoing mammograms were assessed with a mean (SD) score of 2.48 (0.71), indicating that women have various concerns about this procedure. Worries related to BC (item 7) had a mean (SD) score of 2.52 (0.92), while fears of pain had a mean (SD) score of 2.48 (1.02). Item 8, "Having a mammogram of the breast would be embarrassing," was the highest perceived barrier, with a mean (SD) score of 2.38 (1.00).

Participants reported a low health motivation score regarding mammography, with a mean (SD) of 2.51 (0.71). Among the health motivation items, the desire to detect health problems early (item 12) had the highest mean (SD) score of 2.64 (1.09). In contrast, lower scores were observed for exercising 3 times a week (item 17), with a mean (SD0 of 2.31 (0.95), and for undergoing regular health check-ups even when not sick (item 18), with a mean (SD) of 2.11 (1.08).

Knowledge and Beliefs Regarding BC and Mammography Across Sociodemographic Characteristics

The statistical analysis, conducted using MANOVA, revealed that age and educational level had a significant influence on knowledge and beliefs, as measured by the CHBM scale (P<.05 for all). Moreover, the Bonferroni post hoc tests for age groups and educational levels revealed that women aged 60 years and older with primary or secondary education demonstrated lower levels of knowledge, perceived benefits, and health motivation. In addition, they reported higher perceived barriers compared with women in the 40 - 49 years and 50 - 59 years age groups who had a university education (P<.001). In contrast, marital status and health insurance had no significant effect on the variables related to BC and mammography (P>.05 for all) (Table 4).

XSL•FO
RenderX

Table . Mean differences in knowledge and health beliefs.

Variables	Knowledge			Perceived benefits			Perceived barriers[a]			Health motivation		
	Mean (SD)	*P* value[b]	ηp² [c]	Mean (SD)	*P* value	ηp²	Mean (SD)	*P* value	ηp²	Mean (SD)	*P* value	ηp²
Age groups (years)		<.001	0.10		<.001	0.122		<.001	0.088		<.001	0.080
40 - 49	6.43 (2.70)			2.86 (0.97)			2.65 (0.73)			2.68 (0.77)		
50 - 59	5.86 (2.44)			2.54 (0.70)			2.48 (0.63)			2.49 (0.62)		
60 years and older	4.17 (2.13)			2.02 (0.68)			2.08 (0.64)			2.14 (0.57)		
Education level		<.001	0.471		<.001	0.503		<.001	0.184		<.001	0.371
Primary	3.42 (1.69)			1.77 (0.47)			2.09 (0.66)			1.94 (0.47)		
Secondary	5.98 (1.92)			2.63 (0.60)			2.50 (0.60)			2.56 (0.58)		
University	8.19 (2.15)			3.44 (0.81)			2.89 (0.69)			3.08 (0.64)		
Marital status		.05	0.019		.54	0.005		.41	0.007		.71	0.003
Single	5.75 (2.39)			2.61 (0.92)			2.48 (0.76)			2.45 (0.76)		
Married	6.05 (2.73)			2.62 (0.92)			2.49 (0.70)			2.54 (0.72)		
Divorced	5.70 (2.27)			2.51 (0.76)			2.62 (0.63)			2.51 (0.60)		
Widowed	4.92 (2.84)			2.43 (0.78)			2.36 (0.69)			2.51 (0.66)		
Health insurance		.72	0.000		.94	0.000		.87	0.000		.59	0.000
Yes	5.81 (2.69)			2.58 (0.89)			2.48 (0.71)			2.51 (0.72)		
No	5.64 (1.97)			2.60 (0.85)			2.46 (0.69)			2.53 (0.62)		

[a]Reversed score: a lower score indicates higher perceived barriers.

[b]MANOVA test. A *P* value less than .05 was considered significant.

[c]Partial eta square (ηp²): ηp²<0.01:Negligible, 0.01≤ηp²<0.06:Small, 0.06≤η²<0.14:Medium, ηp²≥0.14: Large.

The results showed that age significantly influenced knowledge (F_{2402}=22.44; *P*<.001). Participants aged 40 - 49 years had a mean (SD) score of 6.43 (2.70), while those aged 50 - 59 years had a mean (SD) score of 5.86 (2.44). These groups had significantly higher mean scores than participants aged 60 years and older, who had a mean (SD) score of 4.17 (2.13, *P*<.001). In addition, the educational level also significantly influenced knowledge (F_{2402}=179.05; *P*<.001). Participants with a university education had mean (SD) scores of 8.19 (2.15), which were higher than those with secondary education, 5.98 (1.92), and primary education, 3.42 (1.69), *P*<.001. In addition, the partial eta squared (ηp²) results for age showed a medium effect size for knowledge (ηp²=0.10), while education level displayed a large effect size for knowledge (ηp²=0.471).

Similarly, perceived benefits were significantly influenced by age (F_{2402}=27.82; *P*<.001), and had a medium effect size (ηp²=0.122). The mean (SD) score for the participants aged 40 - 49 years was 2.86 (0.97), and for those aged 50 - 59 years, it was 2.54 (0.70). Both mean scores were higher than those of participants aged 60 years and older, who had a mean (SD) score of 2.02 (0.68), *P*<.001. In addition, the educational level significantly influenced perceived benefits (F_{2402}=203.11; *P*<.001), and had a medium effect size (ηp²=0.503). Participants with a university education had higher mean (SD) scores of 3.44 (0.81) compared with those with secondary education, 2.63 (0.60), and primary education, 1.77 (0.47), *P*<.001.

Moreover, the perceived barriers were significantly influenced by age (F_{2402}=19.32; *P*<.001) with a medium effect size (ηp²=0.088). Participants aged 60 years and older displayed significantly higher barrier scores, with a mean (SD) score of 2.08 (0.64), compared with those in the 40 - 49 years age group, 2.65 (0.73), and the 50 - 59 years age group, 2.48 (0.63),

XSL•FO
RenderX

P<.001. In addition, the educational level had a significant influence (F_{2402}=45.24; P<.001), and a large effect size (ηp^2=0.184). Participants with primary education had higher barrier mean (SD) scores of 2.09 (0.66) compared to those with secondary education, 2.50 (0.60), and university education, 2.89 (0.69), P<.001.

Ultimately, health motivation was significantly influenced by age (F_{2402}=17.53; P<.001). The participants' mean (SD) score for both ages 40 - 49 years was 2.68 (0.77), and that for 50 - 59 years was 2.49 (0.62), which was higher than those of participants 60 years and older, 2.14 (0.57), P<.001. The educational level also had a significant influence (F_{2402}=118.54; P<.001), with women holding a university education 3.08 (0.64) scoring higher than those with secondary education 2.56 (0.58) and primary education 1.94 (0.47), P<.001. Furthermore, the partial eta squared results for age revealed a medium effect size for health motivation (ηp^2=0.080). The effect size was large for both educational level and health motivation (ηp^2=0.371).

Regarding marital status, the effect sizes were small or negligible for all dependent variables: knowledge (ηp^2=0.019), perceived benefits (ηp^2=0.005), perceived barriers (ηp^2=0.007), and health motivation (ηp^2=0.003). Finally, there was no effect (ηp^2=0.000) on any dependent variables for health insurance status.

Discussion

Principal Findings

This cross-sectional study examines the knowledge and beliefs of Jordanian women aged 40 years and older regarding BC and mammography, using the HBM as a theoretical framework. The findings offer valuable insights into the knowledge, perceived benefits, barriers, and health motivations that influence mammogram screening practices among this population.

Participants in this study reported several barriers that adversely affected their participation in screening practices, such as cost, fear of results, pain, cultural beliefs, and religious factors. These findings were consistent with those observed by Abu-Helalah et al in Jordan [14], which identified the main reasons for avoiding screening as fear of results (63.8%), lack of support (59.7%), cost (53.4%), and religious beliefs (51.1%). Another study conducted by AlAbdulKader et al [41] in Saudi Arabia found that the main barriers to receiving BCS include feeling that it is unnecessary (24.2%), believing it to be painful (22.1%), and fearing abnormal results (18.6%). In addition, other regional and global studies support these findings, highlighting fear of pain and fear of diagnosis as common reasons for avoiding BCS [42,43].

Furthermore, the prevalence of cultural beliefs (63%) and religious factors (29.4%) is a significant barrier to mammography screening, as observed in our study and supported by Albadawi et al [44], who found that these beliefs influence BCS behaviors among women in Jordan. Similarly, Taha et al [45] and Salman et al [46] revealed through qualitative interviews that many Jordanian women hold fatalistic views shaped by the religious concept of "Qadaa Wa Qadar" (predestination).

These beliefs lead many to view BC as "from Allah (God)" and thus unavoidable, reducing their perceived benefits of early detection. This religious fatalism, as highlighted in studies by Petro-Nustas et al and Donnelly et al [22,47], serves as a significant barrier to adopting preventive health behaviors. Moreover, Taha et al [45] also discussed this in a study conducted in Jordan, where men viewed BC as an unavoidable act of God, detached from their families. This belief is linked to notions of improper behavior and a readiness to confront the cultural stigma of Eib (shame) [48]. In addition, research by Al-Naggar and Bobryshev [49] found that religious beliefs accounted for 42.5% of screening avoidance among Malaysian Muslim women, closely aligning with our finding of 37.5%. Their study also emphasized that support from religious leaders can boost participation rates, a factor worth exploring in the Jordanian context.

We observed a significant lack of knowledge about BC among the participants in this study, with a low average score of 5.80 out of 12. Furthermore, 72.8% of participants indicated that their lack of knowledge was the primary reason for not undergoing mammogram screenings. This finding aligns with Kawar's study [50], which emphasized the importance of health care advice in improving awareness and encouraging participation in screenings. In addition, a 2023 national survey conducted among Palestinian women revealed that 53.2% believed the idea of "having a physical trauma" was a myth associated with BC, while 19.4% regarded "wearing a tight bra" as a misconception [51]. In contrast, our study found that only 43.5% of participants understood that not all fibrous sacs increase the risk of BC, and only 35.8% recognized that not all breast tumors are cancerous. These findings suggest a substantial gap in understanding breast abnormalities, which is similar to the health literacy deficiencies commonly observed in low-resource settings, where limited access to accurate health information contributes to poor health literacy [7,52,53].

The findings from the CHBM scale in this study revealed that perceived barriers had the lowest mean (SD) score of 2.48 (0.71) out of 5, followed by health motivation at 2.51 (0.59) out of 5 and perceived benefits at 2.59 (0.71) out of 5. This finding is consistent with previous studies that have identified perceived barriers as the most significant predictor of mammogram screening [27,33,37]. Participants in our study showed a negative perception regarding the benefits of mammography, as evidenced by their lower scores on statements such as: "When I get a mammogram, I do not worry as much about BC" and "Having a mammogram will decrease my chances of dying from BC." These results were compatible with findings from other studies [22,33], as reported by Al-Naggar and Bobryshev [49], which indicated that Malaysian women with a limited understanding of the benefits of mammography were less likely to participate in regular screenings; only 15% had ever undergone a mammogram. Together, these findings underscore the importance of educational interventions in addressing misconceptions and promoting awareness of the benefits of mammography [54].

XSL•FO
RenderX

In our study, participants reported higher perceived barriers to mammogram screening, such as the belief that "having a mammogram would be embarrassing" or "having a mammogram would be painful." Previous studies support these findings, as Amin et al [55] found that embarrassment and anxiety about pain significantly deterred Bangladeshi women from mammography screening, especially where modesty is prioritized. In addition, Al-Mousa et al [56] found that Jordanian women often avoid mammography due to feelings of embarrassment and fear of the results, as they believe they do not need a mammogram if they have no family history of BC, which is deeply rooted in cultural beliefs. These studies emphasize the need for culturally sensitive interventions to increase the screening rate.

Moreover, participants in our study had a negative perception of health motivation, with low mean scores for attending regular check-ups, maintaining a balanced diet, and exercising regularly. These findings suggest that targeted health interventions should increase awareness about the importance of routine check-ups, provide accessible and culturally appropriate dietary guidance, and promote exercise through structured programs to foster sustained engagement in preventive health behaviors [22,23,56]. The motivation for health and awareness was significantly associated with screening behavior [27].

Our study's findings, related to the association between sociodemographic characteristics and knowledge and beliefs regarding BC and mammography, reveal that participants' age and education have a significant influence on their knowledge and beliefs. Older and less educated women showed lower awareness and more negative perceptions regarding the screening. These results align with previous studies and underscore the significant influence of age and education on shaping women's perspectives toward mammography screening [23,27,56,57].

Specifically, our study's findings revealed a significant association between knowledge and beliefs, as measured by the CHBMS, and the ages and educational levels of the participants (P<.05). Participants aged 40 - 49 years and 50 - 59 years scored higher on knowledge, perceived benefits, and health motivation compared with those aged 60 years and older (P<.001). This was consistent with previous studies suggesting that younger women have better access to health information, are more exposed to BC awareness campaigns, and benefit from higher literacy rates and greater access to digital platforms [58,59]. Younger women appreciate early detection, which is influenced by targeted screening and health care recommendations. Moreover, older women may underestimate the benefits of mammography, possibly due to rooted cultural beliefs, as found by Mah et al [60] that older women felt less vulnerable to BC than younger women and were less positive about screening. This suggests that motivation declines with age due to health priorities and misconceptions about the effectiveness of screening in older adults [61,62].

Similarly, women with a university education scored higher on knowledge, perceived benefits, and health motivation compared to those with secondary and primary education (P<.001). This supports the notion that individuals with higher education are more proactive about their health and preventive care and are more likely to follow medical advice [63]. This suggests that highly educated women are more likely to recognize the value of early detection and preventive care [64]. This suggests a more substantial health motivation for the importance of screening among women with higher levels of education [63,64].

Conversely, the perceived barriers were significantly more negative among participants aged 60 years and older compared with those aged 40 - 49 years and 50 - 59 years (P<.001), aligning with the logistical and psychological challenges faced by older women and their traditional beliefs [65,66]. Furthermore, participants with primary education reported more barriers than those with secondary and university education (P<.001), suggesting that lower education is associated with higher perceived obstacles, such as lack of knowledge, fear, embarrassment, financial concerns, and cultural beliefs [67,68].

Neither marital status nor health insurance had a significant impact on knowledge and beliefs (P>.05). This finding contrasts with studies suggesting that married women benefit from spousal support for mammography [69,70]. While health insurance facilitates access to screening, its impact may diminish in areas with few financial barriers [71,72].

Limitations

The study has several limitations. Its cross-sectional design uses convenience sampling, which prevents the determination of causal relationships between beliefs and screening behaviors and limits the generalizability of the results. The reliance on self-reported data may introduce information bias. Recruiting participants from a single hospital through convenience sampling may limit the generalizability of the findings to all Jordanian women aged 40 years and older, as it will not encompass all Jordanian women from different geographical locations, particularly those from rural areas. Furthermore, excluding women who had a mammogram within the last 12 months, as well as those with personal or family histories of BC, may introduce selection bias and limit insights into the influence of hereditary factors on screening behaviors.

Implications for Practice

Developing effective educational interventions for Jordanian women requires a multifaceted approach that addresses both individual and systemic barriers. These interventions must account for varying levels of awareness, health beliefs, and the influence of religious factors, such as fatalism, which affect screening behaviors. Involving religious and community leaders in program content can boost motivation and reduce participation barriers.

Educational initiatives should include different community activities, such as workshops, home visits, and outreach sessions with community leaders, religious figures, and BC survivors. This approach fosters trust and delivers culturally appropriate information. Expanding mobile mammography units and organizing workplace screening campaigns are crucial, especially for underserved rural populations.

Public health officials and health care providers need ongoing training in culturally competent communication and patient

XSL•FO
RenderX

counseling to ensure sensitive and supportive care. To improve accessibility, financial barriers can be addressed by offering free or subsidized mammograms for low-income women and advocating for broader insurance coverage.

Delivery methods should use digital health tools, such as mobile apps and SMS text messaging reminders, which can enhance personalized education and appointment alerts. In addition, providing culturally appropriate support services, including flexible scheduling and female-only environments, is important for addressing concerns about modesty and privacy.

These interventions should be integrated within a robust policy framework that updates national screening guidelines and incorporates educational efforts into a national registry and monitoring system. By tackling awareness levels, health beliefs, and religious perspectives, Jordan can implement effective interventions to improve BCS and outcomes for women nationwide. Ultimately, ongoing research to monitor screening trends is crucial for evaluating and enhancing health initiatives, providing valuable insights to inform future policies and programs.

Conclusion

This study contributes to the growing body of research on BCS in Jordan and similar contexts. It highlights significant barriers to mammogram screening among Jordanian women aged 40 years and older, including a lack of knowledge, cultural beliefs, and psychological and financial challenges. The study emphasizes the importance of establishing a solid foundation for culturally appropriate and sensitive interventions that are tailored to the local context and targeted to improve health awareness, accessibility, and policy reforms. Health care providers and policy makers can adopt an evidence-based approach to enhance early detection, reduce BC mortality rates, and promote a healthier future. Future research should focus on assessing the long-term effects of these initiatives.

Acknowledgments

We gratefully acknowledge Karolinska Institute for funding the Article Processing Charge (APC) and Albashir Hospital in Jordan for their invaluable administrative and resource assistance. We greatly appreciate Dr. Mohammad Al-Share's technical assistance and contributions to the data collection process. Finally, we acknowledge all participants whose engagement was essential to the success of this study.

The authors declare that the publication fees for this research article were covered under the national open-access agreement between Karolinska Institutet, as a member of Sweden's academic consortium Bibsam, and JMIR journals.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical and privacy considerations, but are available from the corresponding author on reasonable request.

Authors' Contributions

AA contributed to conceptualization, data curation, formal analysis, investigation, methodology, visualization, writing of the original draft, and review and editing of the manuscript. HT contributed to methodology, formal analysis, funding acquisition, supervision, and writing – review and editing. LGV contributed to methodology, formal analysis, supervision, and writing – review and editing. MA contributed to data curation and investigation. HA contributed to investigation. CAP contributed to methodology, formal analysis, project administration, and writing – review and editing.

Conflicts of Interest

The Karolinska Institute funded the Article publication charge (APC) for this article. Still, it was not involved in the study's design, data collection, analysis, interpretation, writing, or the decision to submit it for publication. Additionally, Dr. Hana Taha, one of the study's authors, serves as an editor for the JMIR Public Health and Surveillance journal.

Multimedia Appendix 1
Mammography Screening Questionnaire.
[DOCX File, 106 KB - publichealth_v11i1e75384_app1.docx]

Multimedia Appendix 2
Permission to view and modify the Health Belief Model.
[PDF File, 61 KB - publichealth_v11i1e75384_app2.pdf]

Multimedia Appendix 3
Ethical Approval.
[DOCX File, 1153 KB - publichealth_v11i1e75384_app3.docx]

XSL•FO
RenderX

Multimedia Appendix 4
Consent form.
[DOCX File, 98 KB - publichealth_v11i1e75384_app4.docx]

Multimedia Appendix 5
Table: Reasons for not getting Mammogram screening within the last 12 months
[DOCX File, 17 KB - publichealth_v11i1e75384_app5.docx]

Checklist 1
Strengthening the Reporting of Observational studies in Epidemiology checklist.
[DOCX File, 45 KB - publichealth_v11i1e75384_app6.docx]

References

1. Hastings KG, Boothroyd DB, Kapphahn K, et al. Socioeconomic differences in the epidemiologic transition from heart disease to cancer as the leading cause of death in the United States, 2003 to 2015: an observational study. Ann Intern Med 2018 Dec 18;169(12):836-844. [doi: 10.7326/M17-0796] [Medline: 30422275]
2. Population factsheets. International Agency for Research on Cancer. URL: https://gco.iarc.fr/today/en/fact-sheets-populations [accessed 2025-03-18]
3. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
4. Breast cancer. World health organization.: World Health Organization; 2024. URL: https://www.who.int/news-room/fact-sheets/detail/breast-cancer [accessed 2025-03-18]
5. Francies FZ, Hull R, Khanyile R, Dlamini Z. Breast cancer in low-middle income countries: abnormality in splicing and lack of targeted treatment options. Am J Cancer Res 2020;10(5):1568-1591. [Medline: 32509398]
6. Country classifications by income. World Bank.: World Bank; 2025. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [accessed 2025-03-18]
7. Abdel-Razeq H, Mansour A. Challenges and opportunities in breast cancer care in low-resourced countries, Jordan as an example. Cancers (Basel) 2024 Apr 30;16(9):1751. [doi: 10.3390/cancers16091751] [Medline: 38730701]
8. Cancer J. Registry. Jordan Cancer Registry annual report. Jordan Ministry of Health. 2022. URL: https://www.moh.gov.jo [accessed 2025-03-18]
9. Promoting Breast Cancer Awareness and Early Detection. Jordan breast cancer program. URL: https://www.jbcp.jo [accessed 2025-03-18]
10. Lambertini M, Santoro L, Del Mastro L, et al. Reproductive behaviors and risk of developing breast cancer according to tumor subtype: a systematic review and meta-analysis of epidemiological studies. Cancer Treat Rev 2016 Sep;49:65-76. [doi: 10.1016/j.ctrv.2016.07.006] [Medline: 27529149]
11. Abdel-Razeq H, Mansour A, Jaddan D. Breast cancer care in Jordan. JCO Glob Oncol 2020 Feb;6:260-268. [doi: 10.1200/JGO.19.00279] [Medline: 32083950]
12. Monticciolo DL, Newell MS, Moy L, Lee CS, Destounis SV. Breast cancer screening for women at higher-than-average risk: updated recommendations from the ACR. J Am Coll Radiol 2023 Sep;20(9):902-914. [doi: 10.1016/j.jacr.2023.04.002] [Medline: 37150275]
13. Cancer facts & figures for african american/black people 2022-2024. American Cancer Society. 2022. URL: https://www.cancer.org [accessed 2025-03-18]
14. Abu-Helalah MA, Alshraideh HA, Al-Serhan AA, Kawaleet M, Nesheiwat AI. Knowledge, barriers and attitudes towards breast cancer mammography screening in jordan. Asian Pac J Cancer Prev 2015;16(9):3981-3990. [doi: 10.7314/apjcp.2015.16.9.3981] [Medline: 25987073]
15. Cancer incidence in jordan. Jordan Ministry of Health. 2018. URL: https://tinyurl.com/mzx2r7vb [accessed 2025-03-18]
16. Obeidat RF, Lally RM. Determinants of breast cancer screening mammography uptake: Results from Jordan population and family health survey. Indian J Cancer 2023;60(2):266-274. [doi: 10.4103/ijc.IJC_622_20] [Medline: 37530252]
17. Bener A, El Ayoubi HR, Moore MA, Basha B, Joseph S, Chouchane L. Do we need to maximise the breast cancer screening awareness? Experience with an endogamous society with high fertility. Asian Pac J Cancer Prev 2009;10(4):599-604. [Medline: 19827877]
18. Al-Zalabani AH, Alharbi KD, Fallatah NI, Alqabshawi RI, Al-Zalabani AA, Alghamdi SM. Breast cancer knowledge and screening practice and barriers among women in Madinah, Saudi Arabia. J Cancer Educ 2018 Feb;33(1):201-207. [doi: 10.1007/s13187-016-1057-7] [Medline: 27271153]
19. Quintin C, Chatignoux E, Plaine J, Hamers FF, Rogel A. Coverage rate of opportunistic and organised breast cancer screening in France: Department-level estimation. Cancer Epidemiol 2022 Dec;81:102270. [doi: 10.1016/j.canep.2022.102270] [Medline: 36215917]

XSL•FO
RenderX

20. Melvin CL, Jefferson MS, Rice LJ, Cartmell KB, Halbert CH. Predictors of participation in mammography screening among non-Hispanic Black, non-Hispanic White, and Hispanic women. Front Public Health 2016;4:188. [doi: 10.3389/fpubh.2016.00188] [Medline: 27656640]
21. Gupta A, Shridhar K, Dhillon PK. A review of breast cancer awareness among women in India: Cancer literate or awareness deficit? Eur J Cancer 2015 Sep;51(14):2058-2066. [doi: 10.1016/j.ejca.2015.07.008] [Medline: 26232859]
22. Taha H, Halabi Y, Berggren V, et al. Educational intervention to improve breast health knowledge among women in Jordan. Asian Pac J Cancer Prev 2010;11(5):1167-1173. [Medline: 21198258]
23. Alkhasawneh IM. Knowledge and practice of breast cancer screening among Jordanian nurses. Oncol Nurs Forum 2007 Nov;34(6):1211-1217. [doi: 10.1188/07.ONF.1211-1217] [Medline: 18024348]
24. Areej Othman RN, Ahram M, Obeidat RF, Obeidat N, Rasoul Al-Tarawneh M. Barriers for mammography among non-adherent women in Jordan: a national survey. Life Sci 1097;10(4):8135 [FREE Full text]
25. Ferreira CS, Rodrigues J, Moreira S, Ribeiro F, Longatto-Filho A. Breast cancer screening adherence rates and barriers of implementation in ethnic, cultural and religious minorities: A systematic review. Mol Clin Oncol 2021 Jul;15(1):139. [doi: 10.3892/mco.2021.2301] [Medline: 34055354]
26. Petro-Nustas W. Young Jordanian women's health beliefs about mammography. J Community Health Nurs 2001;18(3):177-194. [doi: 10.1207/S15327655JCHN1803_04] [Medline: 11560110]
27. Othman AK, Kiviniemi MT, Wu YW, Lally RM. Influence of demographic factors, knowledge, and beliefs on Jordanian women's intention to undergo mammography screening. J Nurs Scholarsh 2012 Mar;44(1):19-26. [doi: 10.1111/j.1547-5069.2011.01435.x] [Medline: 22339890]
28. Rosenstock IM. The health belief model and preventive health behavior. Health Educ Monogr 1974 Dec;2(4):354-386. [doi: 10.1177/109019817400200405]
29. Alyafei A, Easton-Carr R. The health belief model of behavior change. In: StatPearls: StatPearls Publishing; 2025. URL: https://www.ncbi.nlm.nih.gov/books/NBK606120 [accessed 2025-03-18]
30. Becker MH, Maiman LA. Strategies for enhancing patient compliance. J Community Health 1980;6(2):113-135. [doi: 10.1007/BF01318980] [Medline: 7204635]
31. Champion VL. Instrument development for health belief model constructs. ANS Adv Nurs Sci 1984 Apr;6(3):73-85. [doi: 10.1097/00012272-198404000-00011] [Medline: 6426380]
32. Champion VL. Instrument refinement for breast cancer screening behaviors. Nurs Res 1993;42(3):139-143. [doi: 10.1097/00006199-199305000-00003] [Medline: 8506161]
33. Champion VL. Revised susceptibility, benefits, and barriers scale for mammography screening. Res Nurs Health 1999 Aug;22(4):341-348. [doi: 10.1002/(sici)1098-240x(199908)22:4<341::aid-nur8>3.0.co;2-p] [Medline: 10435551]
34. Champion VL, Springston J. Mammography adherence and beliefs in a sample of low-income African American women. Int J Behav Med 1999;6(3):228-240. [doi: 10.1207/s15327558ijbm0603_2] [Medline: 16250677]
35. Mikhail BI, Petro-Nustas WI. Transcultural adaptation of champion's health belief model scales. J Nurs Scholarsh 2001;33(2):159-165. [doi: 10.1111/j.1547-5069.2001.00159.x] [Medline: 11419312]
36. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008]
37. Saei Ghare Naz M, Simbar M, Rashidi Fakari F, Ghasemi V. Effects of model-based interventions on breast cancer screening behavior of women: a systematic review. Asian Pac J Cancer Prev 2018 Aug 24;19(8):2031-2041. [doi: 10.22034/APJCP.2018.19.8.2031] [Medline: 30139040]
38. Matlabi H, Asgari Z, Morsali Asl S, Mousavi S, Rezakhani Moghaddam H. The effectiveness of health belief model initiative in breast cancer screening behaviors among women health volunteers. Soc Work Public Health 2021 Apr 3;36(3):365-376. [doi: 10.1080/19371918.2021.1896407] [Medline: 33849403]
39. Sample size calculator. Raosoft, inc. URL: http://www.raosoft.com/samplesize.html [accessed 2025-03-18]
40. SPSS statistics for windows version 280. IBM corporation. 2021. URL: https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-28 [accessed 2025-08-14]
41. AlAbdulKader A, Gari D, Al Yousif G, et al. Perceived Barriers and Facilitators to Breast Cancer Screening Among Women in Saudi Arabia. Breast Cancer (Dove Med Press) 2023;15:505-513. [doi: 10.2147/BCTT.S406029] [Medline: 37520408]
42. Ayaad O, Al Ajmi AA, Al Baimani K, et al. Breast Cancer Awareness, Screening Practices, Barriers, and Educational Interventions in Middle Eastern Countries: Challenges and Successes. Asian Pac J Cancer Biol 2025;10(2):251-260. [doi: 10.31557/apjcb.2025.10.2.251-260]
43. Kawar LN. Barriers to breast cancer screening participation among Jordanian and Palestinian American women. Eur J Oncol Nurs 2013 Feb;17(1):88-94. [doi: 10.1016/j.ejon.2012.02.004] [Medline: 22459258]
44. Albadawi RS, Alsharawneh A, Othman EH. Determinants and barriers to women's participation in breast cancer screening activities in Jordan: an in-depth study. BMC Public Health 2025 Apr 10;25(1):1339. [doi: 10.1186/s12889-025-22611-9] [Medline: 40211231]

XSL•FO
RenderX

45. Taha H, Al-Qutob R, Nyström L, Wahlström R, Berggren V. “Voices of fear and safety” women’s ambivalence towards breast cancer and breast health: a qualitative study from Jordan. BMC Womens Health 2012 Jul 26;12:21. [doi: 10.1186/1472-6874-12-21] [Medline: 22834874]
46. Salman K, Zoucha R, Nawafleh H. Understanding Jordanian women’s values and beliefs related to breast cancer: a focused ethnography. J Transcult Nurs 2018 Mar;29(2):139-145. [doi: 10.1177/1043659616681424] [Medline: 28826339]
47. Donnelly TT, Al Khater AH, Al-Bader SB, et al. Beliefs and attitudes about breast cancer and screening practices among Arab women living in Qatar: a cross-sectional study. BMC Womens Health 2013 Dec 13;13:49. [doi: 10.1186/1472-6874-13-49] [Medline: 24330708]
48. Taha H, Al-Qutob R, Nyström L, Wahlström R, Berggren V. “Would a man smell a rose then throw it away?” Jordanian men’s perspectives on women’s breast cancer and breast health. BMC Womens Health 2013 Oct 25;13:41. [doi: 10.1186/1472-6874-13-41] [Medline: 24160268]
49. Al-Naggar RA, Bobryshev YV. Practice and barriers of mammography among Malaysian women in the general population. Asian Pac J Cancer Prev 2012;13(8):3595-3600. [doi: 10.7314/apjcp.2012.13.8.3595] [Medline: 23098439]
50. Kawar LN. Knowledge about breast cancer and negative influences affecting breast cancer screening among women in Jordan. IJHSS 2012;2(18):1-10. [doi: 10.30845/ijhss]
51. Elshami M, Ismail IO, Alser M, et al. Common myths and misconceptions about breast cancer causation among Palestinian women: a national cross-sectional study. BMC Public Health 2023 Nov 29;23(1):2370. [doi: 10.1186/s12889-023-17074-9] [Medline: 38031084]
52. Atrooz F, Aljararwah SM, Acquati C, Khabour OF, Salim S. Breast cancer beliefs and screening practices among Syrian refugee women and Jordanian women. Int J Environ Res Public Health 2023 Feb 18;20(4):3645. [doi: 10.3390/ijerph20043645] [Medline: 36834339]
53. Othman A, Ahram M, Al-Tarawneh MR, Shahrouri M. Knowledge, attitudes and practices of breast cancer screening among women in Jordan. Health Care Women Int 2015;36(5):578-592. [doi: 10.1080/07399332.2014.926900] [Medline: 24911030]
54. Rezaeian M, Sharifirad G, Mostafavi F, Moodi M, Abbasi MH. The effects of breast cancer educational intervention on knowledge and health beliefs of women 40 years and older, Isfahan, Iran. J Educ Health Promot 2014;3:43. [doi: 10.4103/2277-9531.131929] [Medline: 25013836]
55. Amin MN, Uddin MG, Uddin MN, et al. A hospital based survey to evaluate knowledge, awareness and perceived barriers regarding breast cancer screening among females in Bangladesh. Heliyon 2020 Apr;6(4):e03753. [doi: 10.1016/j.heliyon.2020.e03753] [Medline: 32382673]
56. Al-Mousa DS, Spuur K, Attar R, Kleib I, Alakhras M. Knowledge, attitudes, and practices related to breast cancer screening among female Jordanian university employees: A cross-sectional study. Radiography (Lond) 2024 Jan;30(1):258-264. [doi: 10.1016/j.radi.2023.11.004] [Medline: 38035443]
57. Damiani G, Basso D, Acampora A, et al. The impact of level of education on adherence to breast and cervical cancer screening: Evidence from a systematic review and meta-analysis. Prev Med 2015 Dec;81:281-289. [doi: 10.1016/j.ypmed.2015.09.011] [Medline: 26408405]
58. Eun Y, Lee EE, Kim MJ, Fogg L. Breast cancer screening beliefs among older Korean American women. J Gerontol Nurs 2009 Sep;35(9):40-50. [doi: 10.3928/00989134-20090731-09] [Medline: 19715259]
59. Friedman LC, Neff NE, Webb JA, Latham CK. Age-related differences in mammography use and in breast cancer knowledge, attitudes, and behaviors. J Cancer Educ 1998;13(1):26-30. [doi: 10.1080/08858199809528507] [Medline: 9565858]
60. Mah Z, Bryant H. Age as a factor in breast cancer knowledge, attitudes and screening behaviour. CMAJ 1992 Jun 15;146(12):2167-2174. [Medline: 1308756]
61. Dolan NC, Lee AM, McDermott MM. Age-related differences in breast carcinoma knowledge, beliefs, and perceived risk among women visiting an academic general medicine practice. Cancer 1997;80(3):413-420. [doi: 10.1002/(sici)1097-0142(19970801)80:3<413::aid-cncr9>3.0.co;2-s] [Medline: 9241075]
62. Safizade H, Amirzadeh N, Mangolian Shahrbabaki P. Motivational factors for breast cancer screening behaviors in Iranian women: a qualitative study. Asian Pac J Cancer Prev 2020 Oct 1;21(10):3109-3114. [doi: 10.31557/APJCP.2020.21.10.3109] [Medline: 33112574]
63. Gürdal S, Saraçoğlu GV, Oran E, Yankol Y, Soybir GR. The effects of educational level on breast cancer awareness: a cross-sectional study in Turkey. Asian Pac J Cancer Prev 2012;13(1):295-300. [doi: 10.7314/APJCP.2012.13.1.295] [Medline: 22502688]
64. Rasu RS, Rianon NJ, Shahidullah SM, Faisel AJ, Selwyn BJ. Effect of educational level on knowledge and use of breast cancer screening practices in Bangladeshi women. Health Care Women Int 2011 Mar;32(3):177-189. [doi: 10.1080/07399332.2010.529213] [Medline: 21337241]
65. Abdel-Aziz SB, Amin TT, Al-Gadeeb MB, et al. Perceived barriers to breast cancer screening among Saudi women at primary care setting. J Prev Med Hyg 2018 Mar;59(1):E20-E29. [doi: 10.15167/2421-4248/jpmh2018.59.1.689] [Medline: 29938236]
66. Joho AA, Mdoe MB, Masoi TJ, Yahaya JJ. Perceived barriers and factors influencing uptake of breast cancer screening among women: a population-based cross-sectional study. Sci Rep 2024 May 29;14(1):12291. [doi: 10.1038/s41598-024-62218-5] [Medline: 38811672]

XSL•FO
RenderX

67. Abugu LI, Nwagu EN, Okeke AI, Odo AN. Knowledge of breast cancer, willingness and barriers to mammography screening among rural women in Enugu State, Nigeria. Afr Health Sci 2023 Sep;23(3):280-290. [doi: 10.4314/ahs.v23i3.34] [Medline: 38357180]
68. Olsson SE, Shah S, Haase E, Butler E, Amado I, Pagidas K. Perceptions and barriers to screening mammography and clinical breast examination: a survey study of underserved populations in North Texas. Public Health Chall 2025 Mar;4(1):e70032. [doi: 10.1002/puh2.70032] [Medline: 40496116]
69. Hanske J, Meyer CP, Sammon JD, et al. The influence of marital status on the use of breast, cervical, and colorectal cancer screening. Prev Med 2016 Aug;89:140-145. [doi: 10.1016/j.ypmed.2016.05.017]
70. Hinyard L, Wirth LS, Clancy JM, Schwartz T. The effect of marital status on breast cancer-related outcomes in women under 65: A SEER database analysis. Breast 2017 Apr;32:13-17. [doi: 10.1016/j.breast.2016.12.008] [Medline: 28012410]
71. Bitler MP, Carpenter CS. Health insurance mandates, mammography, and breast cancer diagnoses. Am Econ J Econ Policy 2016 Aug;8(3):39-68. [doi: 10.1257/pol.20120298] [Medline: 29527253]
72. Levy AR, Bruen BK, Ku L. Health care reform and women's insurance coverage for breast and cervical cancer screening. Prev Chronic Dis 2012;9:E159. [doi: 10.5888/pcd9.120069] [Medline: 23098646]

Abbreviations

ACR: American College of Radiology
ASR: age-standardized rate
BC: breast cancer
BCS: breast cancer screening
CHBM: Champions Health Belief Model
CHBMS: Champions Health Belief Model Scale
HBM: Health Belief Model
IRB: Institutional Review Board
JBCP: Jordan Breast Cancer Program
LMIC: lower-middle-income country
MANOVA: multivariate analysis of variance
STROBE: STrengthening the Reporting of OBservational studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 02.04.25; peer-reviewed by O Allela, SH Sayed, T Kayode, T Amusa; revised version received 05.07.25; accepted 06.07.25; published 21.08.25.

Please cite as:
Abuabed AS, Taha H, Garcia-Valdes LM, Al-Share M, Al-shar'e H, Amezcua-Prieto C
Knowledge and Beliefs Toward Mammography Screening Among Jordanian Women: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e75384
URL: https://publichealth.jmir.org/2025/1/e75384
doi:10.2196/75384

©Ahmad Shaker Abuabed, Hana Taha, Luz María Garcia-Valdes, Mohammad Al-Share, Hadi Al-shar'e, Carmen Amezcua-Prieto. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Health Indicators as Measures of Individual Health Status, Their Perceived Importance, and Associated Factors: Cross-Sectional Observational Study

Temiloluwa Sokoya[1], DrPH; Yuchun Zhou[2], PhD; Sebastian Diaz[3], JD, PhD; Timothy D Law[4], DO; Lina Himawan[5], MSci; Francisca Lekey[6], MPH; Lu Shi[7], PhD; Sarah F Griffin[7], PhD; Ronald W Gimbel[7], PhD; Xia Jing[7], MD, PhD

[1]Department of Behavioral and Environmental Health, College of Health Sciences, Jackson State University, Jackson, MS, United States

[2]Gladys W and David H Patton College of Education, Ohio University, Athens, OH, United States

[3]College of Medicine, Northeast Ohio Medical University, Rootstown, OH, United States

[4]Ohio Musculoskeletal and Neurological Institute, Ohio University, Athens, OH, United States

[5]Department of Psychology, College of Arts and Sciences, Ohio University, Athens, OH, United States

[6]College of Health Sciences and Professions, Ohio University, Athens, OH, United States

[7]Department of Public Health Sciences, College of Behavioral, Social, and Health Sciences, Clemson University, Edwards Hall 519, PHS, CBSHS, Clemson, SC, United States

Corresponding Author:
Xia Jing, MD, PhD
Department of Public Health Sciences, College of Behavioral, Social, and Health Sciences, Clemson University, Edwards Hall 519, PHS, CBSHS, Clemson, SC, United States

Abstract

Background: Health indicators can facilitate the assessment of individual and population health status and are valuable in health care and public health settings. However, it is challenging to collect all measures at once. There is no single, objective, comprehensive, accurate, easy-to-use measure for individual health status yet, and we aim to develop one. We identified 29 health indicators via literature review and ranked their importance through public ratings to start the effort.

Objective: This study aims to explore the effects of 4 demographic variables (age, sex, professional group, and educational level) on the importance ratings of the 29 indicators.

Methods: This was a cross-sectional study. Online surveys were administered on 2 US college campuses and through ResearchMatch, an online platform specialized in finding clinical research participants. Only individuals 18 years or older were considered for inclusion as respondents. A 4-way multivariate analysis of variance (MANOVA) with post hoc testing was conducted to determine the effects of 4 demographic factors on the 29 health indicators. Descriptive statistics and bivariate correlation analysis were reported, along with the results of MANOVA with post hoc testing.

Results: The study included 1153 participants from the United States of America with 1144 complete responses. Our study found significant correlations among the 29 health indicators. Age, education, sex, and professional groups showed correlations with multiple health indicators. Correlation analysis indicated that age was correlated with 16 of the 29 health indicators, followed by education (9 indicators), sex (8 indicators), and professional group (7 indicators). MANOVA modeling revealed that age (partial eta squared=0.058; $P<.001$), sex (partial eta squared=0.05; $P<.001$), professional group (partial eta squared=0.049; $P<.001$), and educational levels (partial eta squared=0.047; $P<.001$), as well as the interactions between sex and professional group (partial eta squared=0.044; $P<.001$) and sex and age (partial eta squared=0.041; $P<.001$), showed significant effects on the combination of the 29 health indicators. In general, female participants provided significantly higher ratings on 10 health indicators than male participants ($P<.05$), whereas older participants provided significantly higher ratings on 20 health indicators than younger participants ($P<.05$).

Conclusions: Age emerged as a critical factor influencing the ratings of the 29 health indicators and showed the largest effect size along with sex, professional group, and educational level. These findings can be used as a quantitative foundation for the development of an objective measure for individual health status, for grouping health indicators, for conducting demographic factor adjustments during data analysis, for designing customized behavioral intervention studies based on demographic factors, or for health policy-making or program development.

(JMIR Public Health Surveill 2025;11:e65616) doi:10.2196/65616

XSL•FO
RenderX

KEYWORDS

health indicators; demographic factors; survey; public perspectives; correlation; multivariate analysis

Introduction

Contexts for Health Indicators

Health indicators can be used in various health care scenarios and public health settings, for example, measuring individual health status [1], measuring population health, community health, tracking longitudinal changes in individual or population health [2-4], tracking the outcomes of preventive services or programs over time, evaluating outcomes of interventions implemented, and comparing health statuses among more homogeneous subgroups [2,5-7]. The importance of health indicators has been recognized and studied since 1972 [8]. However, most of the effort has focused on group or population level measurements, for example, life expectancy and mortality [3,4,9,10]. Disease stages (eg, different stages of cancer) have long been used to help physicians determine more appropriate treatment plans and establish relatively more homogeneous patient groups for better comparability. For example, the Charlson Comorbidity Index (CCI), a comprehensive measurement that predicts long-term mortality, is one of the landmark tools used broadly to help stratify patients with chronic conditions into more homogeneous subgroups for better comparability and more accurate analysis [11,12].

Limitations of the Current Health Indicators

The most used measurement of overall individual health status is self-rated health status, which has been shown to have good validity among the US population [5,13,14]. Although this measure had initially been used to predict mortality among older adults, published studies show its use in measuring outcomes of primary care [15], people with chronic conditions [16], and general adults [17]. Self-rated health status has also been broadly used by the World Health Organization (WHO) and the Organisation for Economic Co-operation and Development (OECD). Nevertheless, given the lack of clear criteria for rating each level of self-rated health status, the obtained ratings may be considered subjective. This subjectivity has been recognized as a potential limitation that could affect comparisons among different populations [5]. Numerous economists have focused on well-being studies and their measurements [18], using gross domestic product (GDP), GDP per capita, life expectancy, or mortality [19] as measures. Meanwhile, some researchers have focused on population health and socioeconomic determinants [18]. However, a literature search suggests the lack of a comprehensive, objective, and consistent measurement for healthy individuals or populations, that is, a counterpart of the CCI, albeit for healthy individuals instead of patients with chronic conditions, the primary target of the CCI.

Needs of a Single, Objective, Comprehensive, and Easy-to-Use Health Indicator

Although life expectancy, morbidity, and mortality have long been used to measure and compare group health outcomes [20] within and among individual countries by the United Nations or WHO, more objective, comprehensive, and accurate methods for measuring a healthy individual's health status are certainly needed, given the lack of existing measures. Some individual health indicators, such as BMI, blood pressure, drug and substance abuse, alcohol abuse, smoking, and tobacco use, may not be obtained easily all at once. Measurement of individual health status could provide important evidence to stratify the healthy population into more homogeneous subgroups, which can make comparisons meaningful.

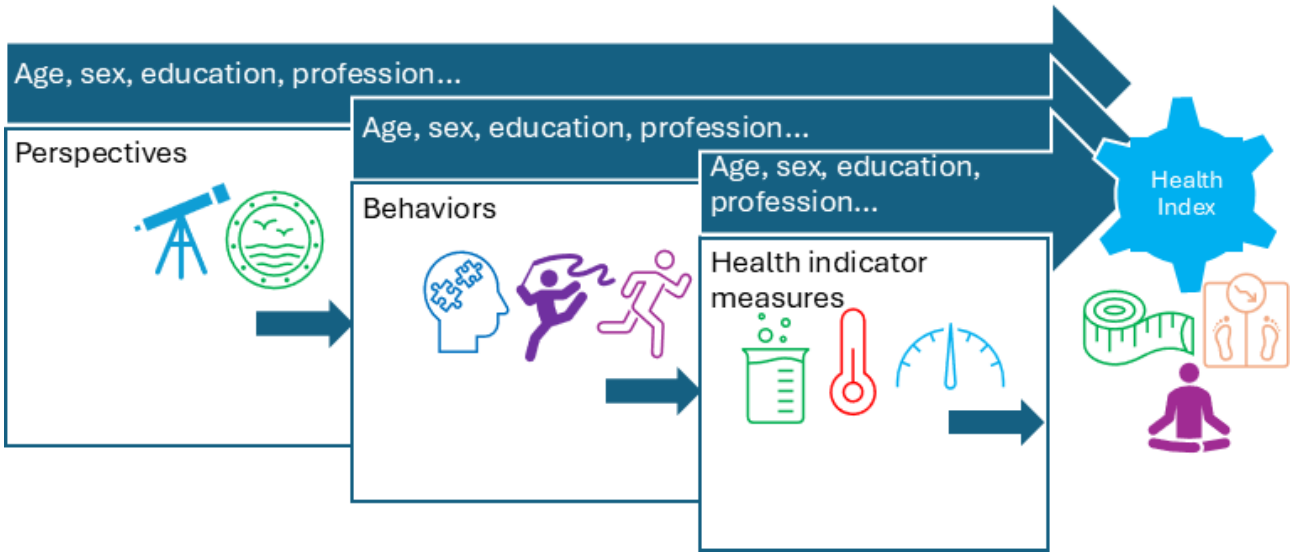
Our Previous Studies on Current Health Indicators

We started by conducting a literature review [1,5,21-25] to compile 29 health indicators [26] measuring an individual's health status, including mental, physical, and social well-being. These health indicators include (in descending order based on public perceived importance rating): drug or substance abuse, smoking and tobacco use, alcohol abuse, major depression, diet and nutrition, blood sugar level, physical inactivity, immunization and vaccination, hypertension screening, low-density lipoprotein (LDL) cholesterol, blood triglycerides, high-density lipoprotein (HDL) cholesterol, having a sense of purpose in one's life, cancer screening detection, total cholesterol, health literacy rate, personal care needs, air quality index > 100, family history of cancer, self-rated health status, HIV testing, insurance coverage, BMI, supply of dentists, sun protection, unemployed individual, engagement in life, high school diploma as a health indicator, and race and ethnicity [26].

Thereafter, we examined 4 commercial electronic health record (EHR) systems, and the results showed that none of the systems included all 29 health indicators [27]. Practically, collecting all 29 health indicators can be burdensome for EHR users [28]. Then we conducted a survey wherein the public was invited to rate the importance of each health indicator. Such ratings may correspondingly impact their behaviors, though not necessarily consciously. Literature has documented people's risk perception and its core impact on one's health behavior [29,30]. The initial analysis of the importance ratings showed that the respondents agreed most on 13 of the 29 health indicators [26]. Figure 1 presents a conceptual framework illustrating the relationship between perspectives, health behaviors, health indicator measures, and associated factors, such as age, sex, education, and profession. This paper is a follow-up and in-depth analysis of the same survey results, focusing on the relationship between importance ratings and demographic information (respondents' age, sex, education level, and professional group).

XSL•FO
RenderX

Figure 1. Conceptual framework between perspectives, health behaviors, health indicator measures, a comprehensive health index (our goal), and the contributing factors, eg, age, sex, education, profession, etc.



Methods

Study Design and Period

This was a cross-sectional study via 3 online surveys administered (Multimedia Appendix 1) at Ohio University (2017), ResearchMatch (2018), and Clemson University (2020). No compensation was provided to any respondents.

Study Population

The only inclusion criterion in this study was that the survey respondents were required to be 18 years or older. The survey was administered via 2 public universities and ResearchMatch [31], which is an online platform for recruiting clinical research participants. Among the 1153 participants who responded, 1144 who provided complete responses for all question items were included as the final total sample.

Data Collection Instrument

The survey was developed via different stages of testing and validation. The health indicators were grouped into 5 categories in the survey to provide a better experience for participants: health risks and behaviors (8 indicators), health care (3 indicators), health care provider supply (3 indicators), blood tests in physical exams (5 indicators), and others (10 indicators) [26]. The survey requested the respondents to rate the importance of 29 health indicators, with each health indicator being defined in the survey. The details related to the survey (see Multimedia Appendix 1) and definitions of each health indicator have all been published [26].

Recruitment and Data Collection Procedures

Qualtrics surveys were used to administer the survey at Ohio University and Clemson University. The importance of each health indicator was rated from 1 (least important) to 10 (most important). The ResearchMatch survey was administered via REDCap (Research Electronic Data Capture; Vanderbilt University), with the ratings ranging from 1 (least important) to 100 (most important). The rating scale was converted into a 1 to 10 scale during the analysis.

The survey link was anonymous, and we explicitly stated in the invitations that the participants were welcome to share the link with others. Multimedia Appendix 2 shows the invitation email used in all 3 settings. All participants consented to the study electronically before they responded to the survey. At Ohio University and Clemson University, the survey was shared via an email broadcast. In ResearchMatch, the survey invitation was sent out randomly. All respondents were volunteers.

Research Question for This Analysis

The following research question was used for this analysis: “What are the effects of age, sex, profession, and education on the perceived importance of 29 health indicators, individually or in combination, among the public?”

Statistical Analysis Plan

This analysis examined all 29 health indicators according to age, sex, educational level, and professional group. Our analysis did not include race or ethnicity, mainly due to imbalanced data [26]. Descriptive statistics were reported to present the demographic profile of the participants (see Table 1).

XSL•FO
RenderX

Table . Descriptive statistics of all independent variable groups.

Independent variable	Results (N=1144), n (%)
Sex	
Female	854 (74.65)
Male	273 (23.86)
Transgender	9 (0.79)
Prefer not to answer	8 (0.7)
Age group (years)	
≤35	460 (40.21)
36-45	161 (14.07)
46-55	151 (13.2)
56-65	200 (17.48)
>65	172 (15.04)
Professional group	
Health care providers	227 (19.84)
Public health professional	70 (6.12)
Researcher who uses health indicator data	78 (6.82)
Other researchers	203 (17.74)
Other professional groups, specify	566 (49.48)
Educational qualification	
High school	216 (18.88)
Associate degree	95 (8.3)
College degree	364 (31.82)
Master's degree	325 (28.41)
Doctoral degree	144 (12.59)

Specifically, to examine bivariate relationships among all variables, a correlation analysis was conducted among the 4 demographic factors and the 29 health indicators. To examine group mean differences, a complete multivariate analysis of variance (MANOVA) with the 4 demographic factors as the independent variables (IVs) and the 29 health indicators as the dependent variables (DVs) was performed. MANOVA was selected to test 29 health indicators simultaneously because it can test the combined effect of all IVs and all DVs at once rather than running separate tests. MANOVA helps control multiplicity by addressing the inflated type 1 error, and it is powerful for detecting multivariate patterns by considering the covariance structure of the 29 dependent variables.

Before we ran MANOVA, we tested the assumptions of normality and homogeneity of variances. Both assumptions were violated. The Box test results indicated significance. Levene test results for 13 of 29 dependent variables were significant. To address the violation of homogeneity, we used Pillai trace values and proceeded with Games-Howell post hoc testing. Regarding the violation of normality, we used MANOVA with no concerns because it is robust to violations of normality when the sample size is large (n=1144). For a sample size larger than 30, the central limit theorem says that the sample mean vectors are approximately multivariate normally distributed, even if the individual observations are not.

The null hypothesis of MANOVA was that a combination of the ratings for the 29 health indicators would have no significant mean difference between the groups. The alpha level for the 4-way MANOVA analysis was set as .05. As such, we identified significant 2-way interaction effects of age by sex and sex by professional group. Therefore, the final MANOVA model included 4 main effects of age, sex, education, professional group, and 2 interaction effects (ie, age by sex and sex by professional group). All statistical analyses were conducted via SPSS (version 29; IBM Corp).

Ethical Considerations

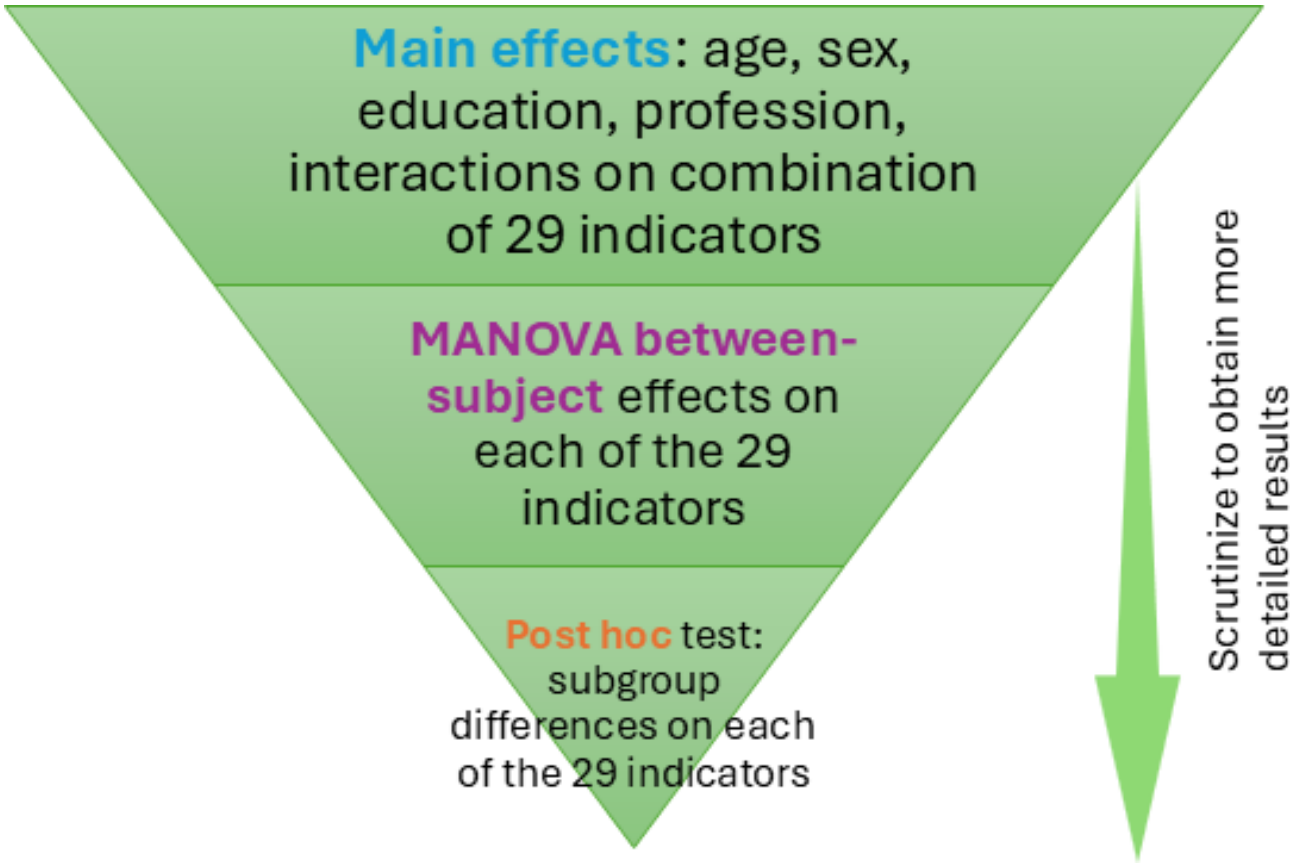
This study was approved by the Ohio University Institutional Review Board (IRB; 17-X-142) and Clemson University IRB (IRB2019-441). All participants completed the consent form remotely before responding to the online survey. This study only uses aggregated results during analysis and in the communication of results; no identifiable information was used in any analysis or publication. The online survey is purely voluntary; no incentives were used except for a warm invitation email.

XSL•FO
RenderX

Results

Figure 2 organizes the results, excluding the participants' profiles and the bivariate correlation results, to help readers navigate the Results section.

Figure 2. Organizational overview of the multivariate analysis results. MANOVA: multivariate analysis of variance.



Respondent Profile

Table 1 summarizes the percentages and numbers of all 4 independent variables (ie, age, sex, profession, and education) and their respective levels (ie, subgroups) to provide a basic profile of the survey respondents.

Bivariate Correlations Between Demographic Factors and the 29 Health Indicators

The bivariate correlation analysis for age, sex, professional group, and education level, and the 29 health indicators (DV) indicated that all the dependent variables were significantly correlated with each other (Pearson correlation coefficient, r=0.058-0.838). The independent variable, age, was correlated with 21 of the dependent variables, except for the following 8 indicators: diet and nutrition, health literacy, high school diploma, HIV testing, immunization, insurance coverage, personal care needs, and total cholesterol level. Sex was correlated with 20 of the DVs, except for the following 9 indicators: air quality index >100, cancer screening, having a sense of purpose, immunization, high school diploma, insurance coverage, hypertension screening, personal care needs, and unemployment. Education level was correlated with these 11 health indicators, namely alcohol abuse, BMI, drug or substance abuse, physical inactivity, smoking, cancer screening, HIV testing, total cholesterol, engaged people, race and ethnicity, and unemployment. The professional group was correlated with 8 health indicators, namely diet and nutrition, engaged people, having a sense of purpose, high school diploma, personal care needs, physical inactivity, total cholesterol levels, and unemployment. The complete correlation matrix is provided in Multimedia Appendix 3.

Main Effects of Demographic Factors on the Combination of the 29 Health Indicators

MANOVA modeling indicated that age, sex, professional group, and education level had significant main effects on the combination of 29 health indicators after controlling for other predictor effects. In addition, significant interaction effects between sex and professional group, as well as sex and age group, were observed on the 29 dependent variables after controlling for other predictor effects. The age group had a relatively increased effect size (ie, partial eta squared=0.058), whereas the interaction between sex and age group had a relatively small effect size (ie, partial eta squared=0.041).

Table 2 details the MANOVA results, including Pillai trace value, df, F, P values, and partial eta squared values. Only statistically significant interactions are presented in Table 2.

XSL•FO
RenderX

Table . Multivariate analysis of variance (MANOVA) results for the main effects of the 4 demographic factors on the combination of the 29 health indicators (only presents the combinations of demographic factors that are statistically significant, N=1114).

Effects	Pillai trace value	*F* test (*df*)	*P* value	Partial eta squared
Age group	0.232	2.304 (116)	<.001	0.058
Sex	0.149	1.951 (87)	<.001	0.050
Professional group	0.198	1.947 (116)	<.001	0.049
Educational qualification	0.189	1.859 (116)	<.001	0.047
Sex and professional group	0.352	1.731 (232)	<.001	0.044
Sex and age group	0.248	1.616 (174)	<.001	0.041

Significant MANOVA Results for Between-Subject Effects

Four-way MANOVA with post hoc testing (Games-Howell) was conducted to examine group mean differences in the 29 health indicators according to the 4 demographic variables (age, sex, professional group, and education level). Table 3 presents the significant results of MANOVA modeling for between-subject effects. Accordingly, age showed significant effects on 16 health indicators, educational level on 9 health indicators, sex on 8 health indicators, and professional group on 7 health indicators. Meanwhile, the interactions between sex and age showed significant effects on 7 health indicators, whereas the interactions between sex and professional group showed significant effects on 5 health indicators.

XSL•FO
RenderX

Table . Significant multivariate analysis of variance (MANOVA) results for between-subject effects.

Independent and dependent variables	*F* test (*df*)	*P* value	Partial eta squared
Sex			
Alcohol abuse	4.304 (3)	.005	0.011
BMI	4.071 (3)	.007	0.011
Diet and nutrition	3.973 (3)	.008	0.011
Drug or substance abuse	3.597 (3)	.013	0.010
Physical inactivity	2.636 (3)	.048	0.007
Smoking or tobacco use	3.675 (3)	.012	0.010
Sun protection	2.642 (3)	.048	0.007
High school diploma as a health indicator	3.914 (3)	.009	0.010
Age group			
Alcohol abuse	2.912 (4)	.021	0.010
BMI	5.336 (4)	.000	0.019
Diet and nutrition	6.697 (4)	.000	0.024
Drug or substance abuse	4.189 (4)	.002	0.015
Family history of cancer	4.201 (4)	.002	0.015
Physical inactivity	4.103 (4)	.003	0.015
Smoking or tobacco use	2.444 (4)	.045	0.009
Sun protection	8.438 (4)	.000	0.029
Hypertension screening	4.240 (4)	.002	0.015
Self-rated health status	3.810 (4)	.004	0.014
Blood sugar level	2.914 (4)	.021	0.010
LDL[a] cholesterol	2.663 (4)	.031	0.009
Air quality index >100	5.507 (4)	.000	0.019
Dentist supply	3.565 (4)	.007	0.013
Engaged people (people’s attention or efforts are occupied)	7.213 (4)	.000	0.025
Having a sense of purpose in one’s life	6.750 (4)	.000	0.024
Professional group			
Smoking or tobacco use	2.579 (4)	.036	0.009
Insurance coverage	3.284 (4)	.011	0.012
Personal care needs	3.616 (4)	.006	0.013
Self-rated health status	2.776 (4)	.026	0.010
Blood triglycerides	2.826 (4)	.024	0.010
Race and ethnicity	3.072 (4)	.016	0.011
Unemployed individual	2.981 (4)	.018	0.011
Educational qualification			
Alcohol abuse	2.511 (4)	.040	0.009
Family history of cancer	2.462 (4)	.044	0.009
Physical inactivity	2.831 (4)	.024	0.010
Smoking or tobacco use	3.041 (4)	.017	0.011
Cancer screening detection	3.413 (4)	.009	0.012
HIV testing	3.562 (4)	.007	0.013

XSL•FO
RenderX

Independent and dependent variables	*F* test (*df*)	*P* value	Partial eta squared
Engaged people (people's attention or efforts are occupied)	2.769 (4)	.026	0.010
Race and ethnicity	7.699 (4)	.000	0.027
Unemployed individual	4.897 (4)	.001	0.017
Sex and age group			
BMI	2.341 (6)	.030	0.012
Diet and nutrition	4.136 (6)	.000	0.022
Physical inactivity	3.502 (6)	.002	0.019
Smoking or tobacco use	2.195 (6)	.041	0.012
Air quality index >100	3.920 (6)	.001	0.021
Having a sense of purpose in one's life	2.896 (6)	.008	0.015
Race and ethnicity	2.703 (6)	.013	0.014
Sex and professional group			
Smoking or tobacco use	2.149 (8)	.029	0.015
Insurance coverage	3.302 (8)	.001	0.023
Personal care needs	2.953 (8)	.003	0.021
Race and ethnicity	4.464 (8)	.000	0.031
Unemployed individual	2.026 (8)	.04	0.014

[a]LDL: low-density lipoprotein.

Table 3 details all significant MANOVA results for between-subject effects. The complete set of the MANOVA testing results of between-subject effects, including both significant and insignificant results, can be found in Multimedia Appendix 4.

Post Hoc Testing Results

Of the 29 health indicators, 13 of them did not have homogenous variances. Due to this violation of the MANOVA assumption, we ran Games-Howell post hoc testing for each of them using ANOVA in SPSS. The other 16 indicators that met the equal variances assumption were tested using Bonferroni post hoc in SPSS. A close examination beyond the main effects of demographic factors on health indicator ratings through post hoc analysis revealed some interesting findings. More specifically, the MANOVA results for between-subject effects indicated significant differences in ratings between female and male respondents on the 10 health indicators (see Table 4, post hoc results and sex). All mean differences between these 10 health indicators showed that female ratings were higher than male ratings. The full set of post hoc testing results is available upon request.

XSL•FO
RenderX

Table . The significant multivariate analysis of variance (MANOVA) results and post hoc testing results.

	Main effects and interaction effects						Post hoc results			
Health indicators	Sex	Age	Professional	Education	Sex by age	Sex by professional	Sex[a]	Age[b]	Profession[c]	Education[d]
Games-Howell post hoc testing results for the 13 health indicators with heterogeneous variances										
HIV testing[e]	—[f]	—[g]	—	✓[h]	—	—	1>2[h]	—	—	1>4, 1>5[h]
Blood sugar[e] level	—[f]	✓[h,f]	—	—	—	—	1>2[h]	3>1, 4>1, 5>1, 4>2, 5>2[h]	—	—
Alcohol abuse	✓[h]	✓[h]	—	✓[h]	—	—	1>2[h]	3>1, 4>1, 5>1, 5>2[h]	—	3>1, 5>1[h]
BMI	✓[h]	✓[h]	—	—	✓[h]	—	—	3>1, 4>1, 5>1, 4>2, 5>2[h]	—	—
Diet and nutrition	✓[h]	✓[h]	—	—	✓[h]	—	1>2[h]	—	—	—
Drug and substance abuse	✓[h]	✓[h]	—	—	—	—	1>2[h]	3>1, 4>1, 5>1, 4>2, 5>2[h]	—	5>1[h]
Physical inactivity	✓[h]	✓[h]	—	✓[h]	✓[h]	—	—	2>1, 4>1[h]	1>5[h]	3>1, 4>1, 5>1[h]
Smoking	✓[h]	✓[h]	✓[h]	✓[h]	✓[h]	✓[h]	1>2[h]	2>1, 3>1, 4>1, 5>1[h]	3>4, 3>5[h]	3>1, 4>1, 5>1[h]
Insurance coverage	—	—	✓[h]	—	—	✓[h]	1>2[h]	—	2>1, 2>5[h]	—
High School	✓[h]	—	—	—	—	—	—	—	—	—
Engaged people	—	✓[h]	—	✓[h]	—	—	—	3>1, 4>1, 5>1, 4>2, 5>2, 5>3[h]	1>3[h]	3>1, 4>1, 5>1[h]
Major depression	—	—	—	—	—	—	—	4>1[h]	—	—
Have a sense of purpose in life	—	✓[h]	—	—	✓[h]	—	—	4>1, 5>1, 4>2, 5>2[h]	1>4, 1>5[h]	—
Bonferroni post hoc testing results for the 16 health indicators with homogenous variances										
Immunization[e]	—	—	—	—	—	—	—	—	—	—
Personal care needs[e]	—	—	✓[h]	—	—	✓[h]	—	—	—	—
Cancer screening[e]	—	—	—	✓[h,f]	—	—	—	—	—	—
Self-rated health status[e]	—	✓[h,f]	✓[h]	—	—	—	—	5>1, 5>2[h]	—	—
Air quality index[e]	—	✓[h,f]	—	—	✓[h]	—	—	4>1, 5>1, 4>2[h]	—	—

XSL•FO
RenderX

	Main effects and interaction effects						Post hoc results			
Dentist supply[e]	—[f]	✓[h,f]	—	—	—	—	1>2[h]	4>1, 5>1, 4>2[h]	—	—
Health literacy rate[e]	—[f]	—	—	—	—	—	—	—	—	—
Blood triglycerides[e]	—[f]	—[f]	✓[h]	—	—	—	—	5>1, 5>2[h]	—	—
HDL[i] cholesterol[e]	—[f]	—	—	—	—	—	—	5>1[h]	—	—
LDL[j] cholesterol[e]	—[f]	✓[h,f]	—	—	—	—	—	5>1, 5>2[h]	—	—
Total cholesterol[e]	—	—	—	—[f]	—	—	—	—	—	—
Family history of cancer	—	✓[h]	—	✓[h]	—	—	—	3>1, 4>1, 5>1, 4>2[h]	—	—
Sun protection	✓[h]	✓[h]	—	—	—	—	1>2[h]	3>1, 4>1, 5>1, 4>2, 5>2[h]	5>4[h]	—
Hypertension screening	—	✓[h]	—	—	—	—	—	3>1, 4>1, 5>1, 4>2, 5>2[h]	—	—
Race and ethnicity	—	—	✓[h]	✓[h]	✓[h]	✓[h]	1>2[h]	4>1, 5>1[h]	2>4, 2>5, 3>4, 3>5[h]	2>1, 3>1, 4>1, 5>1[h]
Unemployed individual	—	—	✓[h]	✓[h]	—	✓[h]	—	5>1[h]	—	3>1, 4>1, 5>1[h]

[a]1: Female; 2: Male; 1>2 indicates female rated higher.

[b]1:≤35 years; 2: 36 - 45 years; 3: 46 - 55 years; 4: 56 - 65 years; 5:>65 years.

[c]1: Healthcare providers; 2: Public health professional; 3: Researcher who uses health indicator data; 4: Other researchers; and 5: Other professional groups, specify.

[d]1: High school; 2: Associate degree; 3: College degree; 4: Master's degree; 5: Doctoral degree.

[e]The 13 health indicators with more consistent ratings.

[f]*P*<.05 when examining the 13 health indicators with more consistent ratings.

[g]Nonsignificant result.

[h]*P*<.05 when examining 29 indicators.

[i]HDL: high-density lipoprotein.

[j]LDL: low-density lipoprotein.

Regarding age, post hoc testing results showed that the ratings for the following 20 health indicators all showed a similar trend of significantly higher ratings among older participants than among younger participants: self-rated health status, air quality index >100, dentist supply, blood sugar level, blood triglycerides, HDL, LDL, alcohol abuse, BMI, drug or substance abuse, family history of cancer, physical inactivity, smoking, sun protection, hypertension screening, engaged people, major depression, having a sense of purpose, race and ethnicity, and unemployment (see Table 4, post hoc results and age).

Post hoc testing results also showed that health care providers rated 3 health indicators significantly higher than other professional groups: physical inactivity, engaged people, and having a sense of purpose in life (see Table 4, post hoc results, and profession). Meanwhile, public health professionals rated 2 health indicators significantly higher than other professional groups: insurance coverage and race and ethnicity. More complete results are presented in Table 4.

Post hoc testing results showed that participants with higher educational levels provided significantly and consistently higher

XSL•FO
RenderX

ratings for the following 7 health indicators than those with lower educational levels: alcohol abuse, drug and substance abuse, physical inactivity, smoking, engaged people, race and ethnicity, and unemployment (see Table 4, post hoc results and education). Meanwhile, participants with high school education rated significantly higher for HIV testing than those with Master's or Doctoral degrees.

Overlapping Results According to Different Approaches

The left part of Table 4 (see Main effects and interaction effects) presents an alternative view of a portion of the results in Table 3. The left side of Table 4 also combines the results when the same approach was used, but only examined the 13 more consistent health indicators per our previous publication [26] (Table 4). After considering the results from all 29 health indicators and the 13 more consistently rated health indicators, we found that age apparently showed consistent and significant effects on the ratings of the following 15 health indicators: blood sugar level, alcohol abuse, BMI, drug and substance abuse, physical inactivity, smoking, engaged people, having a sense of purpose in life, self-rated health status, air quality index >100, dentist supply, LDL, family history of cancer, sun protection, and hypertension screening. In other words, older participants provided higher ratings for these indicators than did younger participants. Meanwhile, the educational level also showed a significant effect on the ratings of HIV testing (inversely), alcohol abuse, physical inactivity, smoking, engaged people, race and ethnicity, and unemployment. Except for HIV testing, all 6 other health indicators were observed such that participants with higher educational levels provided higher ratings for them.

In addition, we also noticed the following interesting findings from Table 4:

1. After examining all 29 health indicators, we found that 5 indicators did not show any significant effects according to demographic factors or their interactions: immunization, health literacy rate, HDL, total cholesterol, and major depression (see Table 4).
2. Meanwhile, there are 3 health indicators showing the significant effects of 4 or 5 demographic factors and their interactions: smoking (5), physical inactivity (4), and race and ethnicity (4).
3. Among the 29 indicators, 13 that received more consistent ratings showed much fewer significant effects according to demographic factors and their interactions:
    - A total of 12 significant effects for 13 indicators versus 40 significant effects for the rest of the 16 indicators, that is, the "Main effects and interaction effects" in Table 4.
    - A total of 11 significant post hoc results for 13 indicators versus 34 significant post hoc results for the rest of 16 indicators, that is, the "Post hoc results" in Table 4.

Discussion

Interpretation of Results

In summary, female participants rated 10 health indicators higher than male participants, and older participants rated 20 health indicators higher than younger participants, based on an overlap analysis of different statistical methods. The demographic factors showed the least significant effects on these 5 health indicators: immunization, health literacy rate, HDL, total cholesterol, and major depression. Meanwhile, demographic factors show significant impacts on ratings of smoking, physical activity, and race and ethnicity.

An alternative presentation of the MANOVA analysis is provided in Table 4 (ie, the "Main effects and interaction effects" from Table 3, whereas the "Post hoc results" details the significant results from post hoc testing); we also noticed the following points. First, the main effects may not be significant when significant post hoc test results were obtained, possibly due to cancellation effects. Second, post hoc results are always necessary when using a complicated MANOVA model for comparisons among subgroups, which cannot be demonstrated via the main effects. Third, among the 29 health indicators, 13 that showed more consistent importance ratings had much fewer main effects and fewer significant post hoc test results than did the remaining 16 health indicators. This third point validates the results across different analytic strategies.

Significance of the Work

This study has endeavored to prioritize the 29 commonly used health indicators via public perspective surveys, thereby providing a foundation for the prioritization of these 29 health indicators, 13 of which received relatively consistent ratings across all 3 samples. The detailed results can be used to determine which health indicators should be included in EHR or PHR systems, given the burden associated with collecting all 29 health indicators. The survey results [26] and further analysis presented in this study provide a foundation for developing a comprehensive health indicator formula that can measure individual health and outcomes of preventive medicine services over time [21,32-34]. Preventive services have been recognized as important for controlling the increasing costs of health care [35,36]. However, the accurate measurement of such services can be challenging given the complexity of measuring most of their outcomes within a short time period.

This paper further explored demographic factors, namely age, sex, educational level, and professional group, and their effects on public ratings of these 29 health indicators. Our results showed that age was a critical factor, with older participants rating 20 health indicators significantly higher than younger ones, indicating the wisdom accumulated based on experience and time. However, more carefully designed experiments are required to determine the actual relationship between age and importance ratings of the 20 health indicators, which is beyond the scope of this paper. However, our results can provide quantitative evidence for age-based adjustments during health indicator-related data analysis.

XSL•FO
RenderX

Sex also played a critical role in ratings of 10 health indicators, with female participants providing significantly higher ratings than male participants. Considering the longer life expectancy among females than males in many countries (5.4 y difference in the United States) [37] and the relationship between perception and health behavior, our public perspective results in the importance ratings of the 10 health indicators aligning well with the life expectancy difference between females and males. Our understanding is that females may pay more attention to their health in general, as evidenced by more female participants in our study and in other studies, higher ratings of health indicators, more female participants in fitness group exercises, and longer life expectancies. However, the actual and accurate relationship may not be as linear, given that females perceived these health indicators to be more important than males, which could lead to healthier behaviors and, eventually, a longer life expectancy. Nonetheless, the observed data points seemingly demonstrate a trend. In addition, we also believe that this point validates our public perspective survey results. Regardless of whether such gender-related differences have physiological roots or purely sociocultural effects [38,39], similar observations were documented between females and males. Our results could be used in making health-related policy, especially in the public health field, for example, conducting more intensive campaigns focusing on males or promoting healthy lifestyles among males. Perhaps new ways of delivering health-related information need to be explored to better engage males during their lifespan.

The detailed comparisons and results presented in this paper can be used as a quantitative foundation to develop a more objective and comprehensive measure of individual health status. In addition, our results can be used as a foundation for designing future health behavioral studies, considering the published demonstration of relationships between perceptions and health behavior [29,30]. The public's perspectives on the studied health indicators have been assumed to be influenced by their health-related awareness or behaviors, either consciously or subconsciously. Our study results could shed some light on the adjustment of demographic-related factors during data analysis or modeling of health-related factors.

Comparison With Other Similar Measures to Provide Context for Our Results

County Health Ranking & Roadmaps, a typical population-level program established by the University of Wisconsin Population Health Institute and supported by the Robert Wood Johnson Foundation [40], compares 2 categories (health outcomes and health factors), 5 classes (health outcomes, health behaviors, clinical care, social & economic factors, and physical environment), and 15 measures and demographic information for each county in the United States [40]. Several similarities in the measures used have been noted between our study (to measure individual health status) and the mentioned program. In fact, a comparison showed a 60% (9/15) overlap in the measures used, which provides a sense of validation and methodological alignment for our study, considering County Health Ranking & Roadmaps is an established and well-respected project, and their results were used and cited broadly in developing population-level health programs and policies. Meanwhile, we also noticed that the main differences were that County Health Rankings used more measures in the social and economic factors, such as income, housing, family and social support, and community safety, whereas our study used more individual-level measures, such as blood sugar level and total cholesterol. Considering that focused groups vary quite significantly (ie, individuals versus county-level populations), the difference in measurement selection is understandable despite similarly attempting to measure health. Our study and results provide a complementary role to the County Health Ranking & Roadmaps, which focuses on the health and comparison between counties, a basic unit, while our study focuses on individuals' health. Considering our different application purposes, we think our study and County Ranking & Roadmaps can be used in complementary ways to study individual or their collective health status.

The UK Health Index, which consists of 3 categories, namely healthy people, healthy lives, and healthy places, has been used to track changes over time and compare the health of residents across different areas of England [2,6,41]. The healthy lives category is very similar to many of our health indicators, also referred to as modifiable risk factors, at the individual level. The healthy people category measures health at the population level, whereas the healthy places category measures health at the system level or beyond individual control (eg, air quality) [42].

A closer comparison of the UK Health Index [6,41] and our 29 health indicators [26] showed a large overlap in 15 health indicators between the two. Over half of the health indicators used in our study were also used by the UK Health Index, which provides additional validity for these health indicators, considering that the UK Health Index is a government-funded project.

However, after closely examining the methodology used in developing the UK Health Index, we noticed the following differences between our project and the UK Health Index project. First, the UK Health Index was developed to measure a nation's health status at both the population and individual levels, whereas our health indicators were used to measure individual health status only. Collectively, however, such indicators can be aggregated to compare health outcomes across different groups or track changes over time. Second, in terms of the long-term objectives, our goal was to use the survey results, analysis results, and real-world data to develop a formula that could calculate a more objective, data-driven health indicator to measure an individual's overall health status. The UK Health Index project used factor analysis to group health indicators and determine the weight of each health indicator. Third, as a next-stage effort, our project will focus on formula development and validation of health indicators, whereas the UK Health Index project uses existing and available data sources and already provides Health Index comparisons among various areas in England, as well as comparisons for the same area over time. Although the references provided by the UK Health Index project are exceedingly helpful [42,43], our study's focus and trajectory differ from that of the UK Health Index project. Fourth, given our initial goal of measuring individual health status, measurement is our core focus, which will be necessary

XSL•FO
RenderX

for the long-term use of our study within EHR or PHR systems. The UK Health Index, however, focuses on the availability and continuity of data sources and compares the same or various locations across different time points.

Strengths and Weaknesses of the Study

Our study explores strategies for measuring individual health more objectively, accurately, and comprehensively. Currently, an objective and accurate measure of individual health status is missing. Such a measure can track the performance of health care systems, preventive services, or intervention programs for individuals, as well as the health and outcomes of the population over time. In addition, such measures can be used to establish more comparable and homogeneous subgroups for conducting epidemiological or health service studies. Although our study constitutes the first step in the development of a larger project, its direction is critical for improving the health of the whole population and the use and promotion of preventive care services. In addition, we used multiple statistical methods during the results analysis, and we emphasized the overlap results across different methods, which provides additional confidence in the results. We feel this is an important step to mitigate the potential limitations of observational studies like ours.

Our work is in the preliminary stages, which is the main limitation of our study. Although our survey included a validation question and only included complete and valid responses in the analysis, this analysis was based on public perspectives. Despite this reality, we believe that our deep analysis of the demographic variables of individual respondents and their effects on and interactions with the rating results are nevertheless helpful for thoroughly understanding the public's perspectives on the 29 health indicators. In addition, our analysis and results can provide insights and quantitative evidence for the future development of a comprehensive individual health index.

Another limitation was that our results were derived from survey results. Despite the relatively large sample sizes and in-depth analysis, our study is an observational study. Nonetheless, our results can shed light on the correlations between various demographic subgroups and ratings. However, given that they are not as convincing as those from a hypothesis-driven randomized controlled study, we urge readers to interpret the results within context. In addition, other factors, such as pain or sleep, had not been considered in our study [10]. Meanwhile, we acknowledge that, due to the voluntary nature of the survey, the respondents have a higher percentage of female participants with relatively higher educational levels compared to the general public. Therefore, users need to put the results into these contexts. In addition, the use of different scales (1 - 10 in Qualtrics and 1 - 100 in REDCap) may also lead to unintended effects in ratings.

Potential Future Developments

Our next step is to explore longitudinal individual health data, possibly through EHR, to establish comparable cohorts to develop, refine, and validate the new individual health measure. All of Us, a National Institutes of Health—initiated project, can be a potential platform and data source to develop and validate the measure [44,45]. Despite numerous studies exploring a measure of health status for decades [8,9,46-49], no consensus has yet been reached regarding such a measure. In addition, others can use our results during data analysis, especially when they need to conduct age- or sex-based adjustments. Meanwhile, our results can also be used as evidence to develop more customized public health programs or interventional studies related to the demographic factors we analyzed in this study. In addition, designing further studies to understand how demographic factors affect the importance ratings of health indicators is certainly worthwhile to explore.

Conclusion

Our analysis showed that all 29 health indicators were significantly intercorrelated. The MANOVA modeling results indicated that age, sex, professional group, educational level, and the interaction between sex and professional group and sex and age had significant main effects on the combination of 29 health indicators. Our findings could provide a quantitative foundation for (1) developing a formula for an overall health index, (2) providing evidence to guide the design of customized behavioral interventional studies for subgroups of the public, and (3) providing evidence for demographic factor-related adjustments during data analysis, modeling, policy making, and program development.

Acknowledgments

We sincerely thank all participants for their responses and for sharing their perceptions with us. Without their participation, none of the work would be possible. Authors have not used generative AI tools in preparing any part of the manuscript at any stage of the writing process. Only Grammarly was used to check and correct grammar errors during writing and revision.

Data Availability

The data sets, statistical analysis strategies, codes, and full sets of testing results for this paper are available upon request from the corresponding author.

Authors' Contributions

XJ, TS, FL, and SD contributed to the initial conceptualization and design of the project. XJ and YZ determined the initial analytic strategies. YZ conducted data analysis and drafted the Methods and Results sections. TS, XJ, and YZ drafted the initial manuscript, and TS revised it significantly. TS cleaned the data and organized it for better usability, with guidance from YZ and XJ. XJ, TS,

XSL•FO
RenderX

YZ, SD, TL, LH, LS, SG, and RWG agreed on the data analysis and interpretation of the results and substantially contributed to revising the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A copy of the online survey used to obtain public perspectives on health indicators.
[PDF File, 187 KB - publichealth_v11i1e65616_app1.pdf]

Multimedia Appendix 2
The invitation message used to invite participants to complete the online survey.
[PDF File, 101 KB - publichealth_v11i1e65616_app2.pdf]

Multimedia Appendix 3
The correlation matrix among the 4 demographic factors and 29 health indicators.
[PDF File, 120 KB - publichealth_v11i1e65616_app3.pdf]

Multimedia Appendix 4
The full version of Table 3 (statistically significant and insignificant results). Multivariate analysis of variance (MANOVA) results for the main effects and interaction effect of demographic factors on the 29 individual health indicators.
[PDF File, 185 KB - publichealth_v11i1e65616_app4.pdf]

References

1. Healthy people 2020. : National Center for Health Statistics; 2020 URL: https://www.cdc.gov/nchs/healthy_people/hp2020/hp2020_indicators.htm [accessed 2025-08-14]
2. Health index explorer. UK Office for National Statistics (ONS). 2018. URL: https://healthindex.lcp.uk.com [accessed 2025-08-14]
3. Hendrikx RJP, Spreeuwenberg MD, Drewes HW, Ruwaard D, Baan CA. How to measure population health: an exploration toward an integration of valid and reliable instruments. Popul Health Manag 2018 Aug;21(4):323-330. [doi: 10.1089/pop.2017.0097] [Medline: 29211631]
4. Kind P, Gudex CM. Measuring health status in the community: a comparison of methods. J Epidemiol Community Health 1994 Feb;48(1):86-91. [doi: 10.1136/jech.48.1.86] [Medline: 8138776]
5. Institute of Medicine. Vital Signs: Core Metrics for Health and Health Care Progress: The National Academies Press; 2015. [doi: 10.17226/19402]
6. UK census 2021-health and social care. Office for National Statistics UK. 2022. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare [accessed 2025-08-14]
7. Jette AM. Health status indicators: their utility in chronic-disease evaluation research. J Chronic Dis 1980;33(9):567-579. [doi: 10.1016/0021-9681(80)90051-x] [Medline: 6447159]
8. Goldsmith SB. The status of health status indicators. Health Serv Rep 1972 Mar;87(3):212-220. [Medline: 4554009]
9. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. Describing health states. Med Care 1984 Jun;22(6):543-552. [doi: 10.1097/00005650-198406000-00005]
10. Brown JH, Kazis LE, Spitz PW, Gertman P, Fries JF, Meenan RF. The dimensions of health outcomes: a cross-validated examination of health status measurement. Am J Public Health 1984 Feb;74(2):159-161. [doi: 10.2105/ajph.74.2.159] [Medline: 6691531]
11. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson Comorbidity Index: a critical review of clinimetric properties. Psychother Psychosom 2022;91(1):8-35. [doi: 10.1159/000521288] [Medline: 34991091]
12. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
13. Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. J Health Soc Behav 1997 Mar;38(1):21-37. [Medline: 9097506]
14. Mossey JM, Shapiro E. Self-rated health: a predictor of mortality among the elderly. Am J Public Health 1982 Aug;72(8):800-808. [doi: 10.2105/ajph.72.8.800] [Medline: 7091475]
15. Rohrer JE, Arif A, Denison A, Young R, Adamson S. Overall self‐rated health as an outcome indicator in primary care. Evaluation Clinical Practice 2007 Dec;13(6):882-888. [doi: 10.1111/j.1365-2753.2006.00766.x]
16. Gill T, Broderick D, Avery J, Dal Grande E, Taylor A. Self reported overall health status: Implications for intervention strategies. AMJ 2008 Nov 1;1(8):44-57. [doi: 10.4066/AMJ.2009.79]

XSL•FO
RenderX

17. Luo N, Johnson JA, Shaw JW, Feeny D, Coons SJ. Self-reported health status of the general adult U.S. population as assessed by the EQ-5D and Health Utilities Index. Med Care 2005 Nov;43(11):1078-1086. [doi: 10.1097/01.mlr.0000182493.57090.c1] [Medline: 16224300]
18. Spitzer S, di Lego V, Greulich A, Muttarak R. A demographic perspective on human wellbeing: concepts, measurement and population heterogeneity. Populationyearbook 2021 Dec 12;19:1-11 [FREE Full text] [doi: 10.1553/populationyearbook2021.int01]
19. Life expectancy at birth. In: Health at a Glance 2017: OECD (Organisation for Economic Co-operation and Development); 2017. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2017/11/health-at-a-glance-2017_g1g800d8/health_glance-2017-en.pdf
20. Kaplan MS, Berthelot JM, Feeny D, McFarland BH, Khan S, Orpana H. The predictive validity of health-related quality of life measures: mortality in a longitudinal population-based study. Qual Life Res 2007 Nov;16(9):1539-1546. [doi: 10.1007/s11136-007-9256-7] [Medline: 17899447]
21. Health indicators: a review of reports currently in use (2008). 2009. URL: http://www.cherylwold.com/images/Wold_Indicators_July08.pdf [accessed 2025-08-14]
22. National health interview survey. National Center for Health Statistics. 2022. URL: https://www.cdc.gov/nchs/nhis [accessed 2025-08-14]
23. Institute of Medicine. State of the USA Health Indicators: Letter Report: The National Academies Press; 2009. URL: https://nap.nationalacademies.org/read/12534/chapter/1 [accessed 2025-08-29]
24. State of global well-being: results of the gallup-healthways global well-being index. : Gallup; 2015 URL: https://wellbeingindex.sharecare.com/wp-content/uploads/2017/12/State-of-Global-Well-Being-2014.pdf [accessed 2025-08-14]
25. Leading Health Indicators 2030: Advancing Health, Equity, and Well-Being: The National Academies Press; 2020:172 URL: https://nap.nationalacademies.org/catalog/25682/leading-health-indicators-2030-advancing-health-equity-and-well-being [accessed 2025-08-29]
26. Sokoya T, Zhou Y, Diaz S, et al. Health indicators as measures of individual health status and their public perspectives: cross-sectional survey study. J Med Internet Res 2022 Jun 21;24(6):e38099. [doi: 10.2196/38099] [Medline: 35623051]
27. Jing X, Lekey F, Kacpura A, Jefford K. Health Indicators within EHR Systems in Primary Care Settings: Availability and Presentation: Springer; 2016. [doi: 10.1007/978-3-319-48335-1_18]
28. McPeek-Hinz E, Boazak M, Sexton JB, et al. Clinician burnout associated with sex, clinician type, work culture, and use of electronic health records. JAMA Netw Open 2021 Apr 1;4(4):e215686. [doi: 10.1001/jamanetworkopen.2021.5686] [Medline: 33877310]
29. Brewer NT, Chapman GB, Gibbons FX, Gerrard M, McCaul KD, Weinstein ND. Meta-analysis of the relationship between risk perception and health behavior: the example of vaccination. Health Psychol 2007 Mar;26(2):136-145. [doi: 10.1037/0278-6133.26.2.136] [Medline: 17385964]
30. Ferrer RA, Klein WMP. Risk perceptions and health behavior. Curr Opin Psychol 2015 Oct;5:85-89. [doi: 10.1016/j.copsyc.2015.03.012]
31. About us. Research Match. 2021. URL: https://www.researchmatch.org/about [accessed 2025-08-14]
32. Hensrud DD. Clinical preventive medicine in primary care: background and practice: 2. Delivering primary preventive services. Mayo Clin Proc 2000 Mar;75(3):255-264. [doi: 10.4065/75.3.255] [Medline: 10725952]
33. Hensrud DD. Clinical preventive medicine in primary care: background and practice: 3. Delivering preventive screening services. Mayo Clin Proc 2000 Apr;75(4):381-385. [doi: 10.4065/75.4.381] [Medline: 10761493]
34. Hensrud DD. Clinical preventive medicine in primary care: background and practice: 1. Rationale and current preventive practices. Mayo Clin Proc 2000 Feb;75(2):165-172. [doi: 10.4065/75.2.165] [Medline: 10683656]
35. Ozieh MN, Bishu KG, Dismuke CE, Egede LE. Trends in healthcare expenditure in United States adults with chronic kidney disease: 2002-2011. BMC Health Serv Res 2017 May 22;17(1):368. [doi: 10.1186/s12913-017-2303-3] [Medline: 28532412]
36. Cannon A, Handelsman Y, Heile M, Shannon M. Burden of illness in type 2 diabetes mellitus. J Manag Care Spec Pharm 2018 Sep;24(9-a Suppl):S5-S13. [doi: 10.18553/jmcp.2018.24.9-a.s5] [Medline: 30156443]
37. Life expectancy for men and women. WorldData. 2024. URL: https://www.worlddata.info/life-expectancy.php [accessed 2024-07-01]
38. Courtenay WH. Constructions of masculinity and their influence on men's well-being: a theory of gender and health. Soc Sci Med 2000 May;50(10):1385-1401. [doi: 10.1016/S0277-9536(99)00390-1]
39. Mairey I, Bjerregaard P, Brønnum-Hansen H. Gender difference in health expectancy trends in Greenland. Scand J Public Health 2014 Dec;42(8):751-758. [doi: 10.1177/1403494814550174] [Medline: 25260639]
40. County Health Rankings & Roadmaps-Building a Culture of Health, County by County. 2023. URL: https://www.countyhealthrankings.org [accessed 2025-08-14]
41. Health index indicators and definitions. Office for National Statistics. 2022. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/methodologies/healthindexindicatorsanddefinitions [accessed 2025-08-14]

XSL•FO
RenderX

42. Health index methods and development: 2015 to 2019. Office for National Statistics. 2022. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/methodologies/healthindexmethodsanddevelopment2015to2019 [accessed 2025-08-14]
43. OECD (Organisation for Economic Co-operation and Development), The Applied Statistics and Econometrics Unit of the Joint Research Centre (JRC), The European Commission. Handbook on Constructing Composite Indicators: Methodology and User Guide: OECD; 2008.
44. The All of Us Research Program Investigators. The "All of Us" research program. N Engl J Med 2019 Aug 15;381(7):668-676. [doi: 10.1056/NEJMsr1809937]
45. All of us research program. NIH. URL: https://allofus.nih.gov [accessed 2021-03-02]
46. Hallett C, Power R, Scott D. Summarisation and visualisation of e-health data repositories. 2006 Presented at: E-Science All-Hands Meeting; Sep 18-21, 2006 URL: http://www.allhands.org.uk/2006/proceedings/papers/652.pdf [accessed 2025-08-26]
47. Horner-Johnson W, Krahn G, Andresen E, Hall T. RRTC expert panel on health status measurement. Developing summary scores of health-related quality of life for a population-based survey. Public Health Rep 2009;124:103-110 [FREE Full text]
48. Verrips GHW, Stuifbergen MC, Bonsel GJ, Gemke R, Paneth N, et al. Measuring health status using the Health Utilities Index: agreement between raters and between modalities of administration. J Clin Epidemiol 2001;54:475-481. [doi: 10.1016/S0895-4356(00)00317-6]
49. VanderZee KI, Sanderman R, Heyink J. A comparison of two multidimensional measures of health status: the Nottingham Health Profile and the RAND 36-Item Health Survey 1.0. Qual Life Res 1996 Feb;5(1):165-174. [doi: 10.1007/BF00435982] [Medline: 8901380]

Abbreviations

CCI: Charlson Comorbidity Index
DV: dependent variable
EHR: electronic health record
GDP: gross domestic product
HDL: high-density lipoprotein
IV: independent variable
LDL: low-density lipoprotein
MANOVA: multivariate analysis of variance
OECD: Organization for Economic Co-operation and Development
WHO: World Health Organization

Edited by A Mavragani; submitted 20.08.24; peer-reviewed by J Ahmad, R Carvalho-Sauer; revised version received 11.05.25; accepted 19.05.25; published 08.09.25.

Please cite as:
Sokoya T, Zhou Y, Diaz S, Law TD, Himawan L, Lekey F, Shi L, Griffin SF, Gimbel RW, Jing X
Health Indicators as Measures of Individual Health Status, Their Perceived Importance, and Associated Factors: Cross-Sectional Observational Study
JMIR Public Health Surveill 2025;11:e65616
URL: https://publichealth.jmir.org/2025/1/e65616
doi:10.2196/65616

© Temiloluwa Sokoya, Yuchun Zhou, Sebastian Diaz, Timothy D Law, Lina Himawan, Francisca Lekey, Lu Shi, Sarah Griffin, Ronald W Gimbel, Xia Jing. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Migrant-Local Differences in the Relationship Between Oral Health, Social Support, and Loneliness Among Older Adults in Weifang, China: Cross-Sectional Study

Hui Liu[1,2,3*], MPH; Jieru Wang[1,2,3,4*], MPH; Rui Chen[1,2,3], MPH; Xixing Xu[1,2,3], MPH; Mingli Pang[1,2,3], MPH; Kaiyuan Feng[1,2,3], MPH; Bingsong Li[1,2,3], MPH; Qinling Li[1,2,3], MPH; Ziwei Qin[1,2,3], MPH; Shuyi Yan[1,2,3], MPH; Nabila Ibn Ziyat[1,2,3], MPH; Fanlei Kong[1,2,3], PhD

[1]Department of Social Medicine and Health Management, School of Public Health, Cheeloo College of Medicine, Shandong University, 44 Wenhuaxi Road, Jinan, China

[2]NHC Key Lab of Health Economics and Policy Research, Shandong University, Jinan, China

[3]Center for Health Management and Policy Research, Shandong University (Shandong Provincial Key New Think Tank), Jinan, China

[4]Editorial Office of Chinese Rural Health Service Administration of Publishing Center, Anhui Medical University, Hefei, China

[*]these authors contributed equally

Corresponding Author:
Fanlei Kong, PhD
Department of Social Medicine and Health Management, School of Public Health, Cheeloo College of Medicine, Shandong University, 44 Wenhuaxi Road, Jinan, China

Abstract

Background: Increased aging and accelerated urbanization have led to the migration of older adults within China. Migrant older adults (MOAs) may experience physical and psychological discomfort in influx cities, and they are a vulnerable group that has emerged in the course of fast urbanization. Previous studies have confirmed the association between oral health and loneliness as well as the relationship between social support and loneliness; however, no research has been done to clarify the underlying mechanisms and the migrant-local difference between oral health, social support, and loneliness.

Objective: This study aimed to test the association between oral health, social support, and loneliness among Chinese older adults, as well as the migrant-local difference on the above relationship.

Methods: Multistage cluster random sampling was used to enroll a total of 1205 participants, including 613 MOAs and 592 local older adults (LOAs). Loneliness was assessed by the 6-item short-form UCLA Loneliness Scale; oral health was measured via the Chinese version Geriatric Oral Health Assessment Index (GOHAI); social support was evaluated by the Social Support Rating Scale (SSRS). Descriptive analysis, χ^2 tests, and *t* tests were conducted. Multigroup structural equation modeling (SEM) was employed to clarify the migrant-local difference on the association between oral health, social support, and loneliness among MOAs and LOAs.

Results: The mean score of loneliness was 8.58 (SD 3.032) for MOAs and 8.00 (SD 2.790) for LOAs. Oral health and social support were found to be negatively related to loneliness among MOAs and LOAs; the standardized direct effects for MOAs were −0.168 and −0.444 (P<.001), and they were −0.243 and −0.392 (P<.001) for LOAs, respectively. Oral health generated a direct positive effect on social support, and the direct effect was 0.186 for MOAs (P<.001) and 0.247 for LOAs (P<.001).

Conclusions: Loneliness was fairly low among older adults in Weifang, China, while MOAs showed higher loneliness than LOAs. Oral health had both direct and indirect negative effects on loneliness among MOAs and LOAs, with no significant path differences between MOAs and LOAs. Social support was found to be negatively associated with loneliness for both MOA and LOA, while the association was stronger among MOAs than LOAs. Oral health exerted a significantly positive effect on social support for both MOAs and LOAs, while no significant difference existed between them. Measures should be taken by the government, society, and families to increase social support, improve oral health, and further reduce loneliness among MOAs and LOAs.

(JMIR Public Health Surveill 2025;11:e66061) doi:10.2196/66061

KEYWORDS

loneliness; oral health; social support; migrant older adults; local older adults

XSL•FO
RenderX

Introduction

China has one of the fastest growing aging populations in the world [1]. According to the data of China's Seventh National Census, the number of people aged 60 years or above was 264.02 million, accounting for 18.70% (the number of people aged 65 years or above was 190.64 million, accounting for 13.50%). Compared with 2010, the proportion of people aged 60 years or above increased by 5.44% [2]. As the population ages, concern regarding older adults' health status has also increased.

In the past few decades, China has experienced a rapid increase in economic level and an acceleration of urbanization, which also caused an economic gap across different cities and may further lead to population migration. Due to the traditional Chinese culture, which highlights the family union, more and more older adults move to live with their adult children. Existing studies refer to the older adults who leave their hometown to migrate with their children as migrant older adults (MOAs) [3]. Previous studies found that MOAs had difficulties in developing social networks and social integration, which would further affect their mental health [4,5], life satisfaction [6], and quality of life [7]. Based on "adaptive theory" [8] and "social support buffering model" [9], MOAs face increased adaptation stress and social isolation, which may lead to higher reliance on social support networks to cope with the challenges posed by migration. Thus, it is also important to pay more attention to the health status of MOAs, in addition to that of local older adults (LOAs).

Loneliness is defined by Weiss [10] as a subjective feeling formed when individuals perceive that they lack satisfying interpersonal relationships and there is a gap between their desire to connect and the actual situation of connection. Donovan et al [11] found that 17.6% of US older adults felt lonely much of the time during the preceding week. Guo et al's [12] nationwide study found that 24.3% of older people reported they were lonely in China. Another study found that nearly one-fourth of Chinese older adult participants felt lonely [13]. In terms of MOAs, one longitudinal study showed that migrants from non–English-speaking countries reported higher levels of loneliness, as compared with the native-born, non-Indigenous Australians [14]. Some studies indicated that loneliness was associated with some negative health outcomes [15,16], such as poor physical health [17] and poor mental health [18,19]. Moreover, another national study in China clarified that a high degree of loneliness could further reduce the life satisfaction of older adults [20]. A cohort study from northern and southern Europe showed that loneliness was associated with a decreased quality of life among older adults [21]. Therefore, it is important to pay more attention to older adults' loneliness and find possible ways to reduce their loneliness.

Oral health has been defined by the World Health Organization (WHO) as a state of being free from chronic mouth and facial pain, oral and throat cancer, oral infection and sores, periodontal (gum) disease, tooth decay, tooth loss, and other diseases and disorders that limit an individual's capacity in biting, chewing, smiling, speaking, and psychosocial well-being [22]. The Fourth Chinese Oral Health Epidemiological Survey Report demonstrated that the prevalence of dental caries in older adults' permanent teeth was fairly high, implying serious oral health conditions [23]. Oral health, as a part of general health, is an important determinant of loneliness. One study examined the relationship between oral health and mental health issues and found that oral health was negatively associated with depression [24]. However, few studies have examined the association between oral health and loneliness. A study among English older adults demonstrated that oral health–related quality of life was identified as an independent risk factor for loneliness both cross-sectionally and longitudinally, and older adults with oral impacts had a significantly higher risk of being lonely than their counterparts without any oral impacts [25]. A previous study also found that the fewer number of teeth, the more likely both Japanese and British older adults were to report social isolation or loneliness [26]. Studies also clarified that the degenerative changes of oral function could lead to eating difficulties and imbalanced nutritional intake, which in turn affects physical and mental health and may increase the loneliness of older adults [27,28]. Thus, it is important to test the empirical association between oral health and loneliness among older adults.

Social support refers to the material and moral help provided by various parties in society, including family, relatives, friends, colleagues, organizations, and trade unions [29]. A meta-analysis has pointed out that increased social support would be supportive of successful aging among older adults [30]. Existing research also found that social support could not only provide the necessary resources (such as financial help and emotional help) for older adults to cope with challenges, but also had a huge positive impact on their physical and mental health [30,31]. Conversely, previous studies clarified that a lack of social support would cause adverse outcomes, such as limitations on activities of daily living [32] and poor quality of life [33]. The association between social support and loneliness has been widely explored. A previous study pointed out that social support was negatively associated with loneliness among older adults [34]. Chung and Kim [35] found that social support was critical in lowering loneliness in middle-aged and older adults. Studies from different countries indicated that the changes in older people's social networks and social support resulted in a significant onset of loneliness during the COVID-19 pandemic [36,37].

A previous study indicated that social support was related to older adults' oral health [38]. A nationwide study in Britain demonstrated social support was associated with the oral health status and oral health behavior of older people [39]. A study conducted in Germany clarified that impaired oral health–related quality of life was positively correlated with lower social support among older seniors [40]. A study had shown that oral condition was closely related to facial appearance, which can affect social image and confidence and negatively affect relationships [41]. A cross-national comparative study found that oral health was strongly associated with social isolation, with poorer oral health accelerating social isolation and thus reducing individuals' perceived social support [26]. Therefore, oral health has a critical impact on social support in older adults, and it is essential to study the effects of both in MOAs and LOAs.

XSL•FO
RenderX

Based on the above literature review, no study was found that determined the association between oral health, social support, and loneliness, and no study has ever compared the difference between MOAs and LOAs regarding the above relationship. Thus, this study aimed to (1) clarify the association between oral health, social support, and loneliness and (2) test whether a statistically significant difference exists for the above relationship.

Methods

Sample Collection

The data were collected in Weifang City, Shandong Province, China in August 2021. The gross domestic product of Weifang City was 701.06 billion Chinese yuan in 2021 (US $96.9 billion at 2025 conversion rate) [42]. Weifang governed 10 districts and 2 counties (59 subdistricts and 59 towns) until July 2020 [43]. The total population of Weifang City was 9.3 million by the end of 2020 according to the Seventh National Census. In 2020, nearly 2.38 million of its whole population constituted migrants from other counties, cities, or provinces, with a variety of sociodemographic and cultural backgrounds [44]. Thus, two groups of older adults were recruited in this study. For MOAs, the inclusion criteria were: (1) aged ≥60 years; (2) their Hukou is not in the present place (Hukou is one of China's oldest tools for population control; it is essentially a household registration permit, similar to an internal passport, which defines where people are registered, not where they currently reside [45]); and (3) an ability and willingness to communicate with surveyors. For LOAs, the inclusion criteria were (1) and (3) from the previous list.

Multistage cluster random sampling was conducted to select the samples of MOAs and LOAs. In the first stage, 4 of 12 districts were selected as the primary sampling units (PSUs), considering the economic development and geographic location. In the second stage, 1 subdistrict was selected from each district (PSU), and a total of 4 subdistricts were taken as the secondary sampling units (SSUs). In the last stage, 4 communities were chosen from the SSUs as the tertiary sampling units (TSUs); that is, 1 community was selected from each of the subdistricts chosen previously. All the MOAs as well as the LOAs who met the above criteria constituted the total study sample. Initially, 616 MOAs and 592 LOAs were selected for interviews. However, 3 MOAs were excluded for answering the questionnaire incorrectly or incompletely. Ultimately, 613 MOAs and 592 LOAs were included in the database. The detailed sample selection process is shown in Multimedia Appendix 1.

Measurement

Sociodemographic Characteristics

The section on sociodemographic characteristic traits included the following: gender (man, woman); age group (60 - 64, 65 - 69, 70 - 74, 75 - 79, ≥80 years); Hukou (rural, urban); marital status (married, single); education level (illiterate, primary school, junior high school, high school and above); pension (Yes, No).

Oral Health

The Chinese version of the Geriatric Oral Assessment Index (GOHAI) was used to measure participants' oral health status. This tool is primarily used to assess the self-reported oral health of older adults and is widely used in China [46]. The Chinese version of the GOHAI is divided into 12 items and 3 subdimensions designed to assess different aspects of oral health: (1) physical functioning (four items), (2) psychosocial functioning (five items), and (3) pain or discomfort (three items). GOHAI scores could be divided into the following 3 categories: (1) 50 and below are defined as low oral health, (2) 51 - 56 as fair-to-moderate oral health, and (3) 57 - 60 as high oral health. In the previous study, the GOHAI scores also had good reliability and validity [47]. In this study, the Cronbach α coefficient was 0.853, indicating that the scale had good reliability.

Social Support

The Social Support Rating Scale (SSRS) was used to measure the social support of MOAs and LOAs, including 10 types of support: friends, residents, neighbors, colleagues, family members, financial, comfort, conversation, help, and activities [48]. The social support scale has 10 items, including 3 dimensions of objective support (3 items), subjective support (4 items), and utilization of social support (3 items). The range of the total score of the scale is 12 - 66. A higher total social support score means the subject received more social support. The SSRS has been proven to have good reliability and validity in practice and has been widely used in China [49,50]. The Cronbach α coefficient was 0.822, implying this scale had good reliability.

Loneliness

The 6-item version of the UCLA Loneliness Scale (ULS-6) was used to assess the loneliness of participants, which excluded 2 reverse-scored items from the 8-item UCLA Loneliness Scale. The scale measures loneliness caused by the discrepancy between the level of desired social engagement and that which is actually experienced. The options are on a Likert scale: l=never, 2=rarely, 3=sometimes, and 4=always, with a total score ranging from 6 - 24. Previous studies have demonstrated good reliability and validity in a Chinese population [51,52]. In this study, the Cronbach α coefficient of the ULS-6 was 0.82, indicating the scale had good reliability.

Statistical Analysis

This study used SPSS (version 24.0; IBM Corp) and AMOS (version 24.0; IBM Corp) to perform the data analysis.

The samples' sociodemographic characteristics in this study were characterized by using descriptive statistics, including frequency (%) and mean (SD). The difference between sociodemographic characteristics of MOAs and LOAs was determined using the χ^2 test, while the *t* test was used to explore the difference in oral health (3 dimensions), life satisfaction (5 items), and loneliness (6 questions) of MOAs and LOAs. *P* values less than .05 were considered statistically significant. The above analyses were performed using SPSS (version 24.0; IBM Corp).

XSL•FO
RenderX

A hypothesized structural equation model was set to analyze the relationship between oral health, social support, and loneliness among MOAs and LOAs in Shandong Province. The maximum likelihood estimation method was used to evaluate the hypothesized model's fit. The structural equation modeling (SEM) process model consisted of 2 categories of variables: latent variables and observed variables. The latent variables were oral health, life satisfaction, and loneliness. The 3 observed variables of oral health included physical functioning, psychosocial functioning, and pain and discomfort. The 3 observed variables of social support were objective support, subjective support, and utilization of social support. The 6 observed variables of loneliness were the 6 items of the ULS-6. All SEM analyses were performed using AMOS (version 24.0; IBM Corp).

The following model fitness indexes were used to judge the fit of the hypothesized model: CMIN (Chi-square value, χ^2), degrees of freedom (*df*), *P* value of the χ^2 test, root mean square error of approximation (RMSEA), the goodness-of-fit index (GFI), the adjusted goodness-of-fit index (AGFI), and the comparative fit index (CFI). In this study, the models would be regarded to be well-fitted when *P*>.05, GFI>0.90, AGFI>0.90, and RMSEA<0.05 [53]. *P* value is easily influenced by sample size under many conditions, so it was only demonstrated in this study and not used as a criterion for judgment [54].

In the multigroup analysis, various parameters are restricted to find the most suitable path model. Five models were displayed in this study, namely M1 (MOA model), M2 (LOA model), M3 (unconstrained model), M4 (measurement weights model), and M5 (structural weights model). M1 and M2 were models fitted based on the sample data of two groups, and M3-M5 were obtained by adding conditions gradually restricted from the initial unconstrained model [55]. The multigroup model invariance was determined before the discussion of the MOA and LOA difference in the structural model of SEM. The change in CFI (ΔCFI) and the change in RMSEA (ΔRMSEA) were used to assess the measurement invariance between unconstrained and constrained multigroup analyses [55]. ΔCFI is independent of both model complexity and sample size, as well as the overall fit measurements. ΔCFI<0.010 indicates that we obtained measurement invariance across groups [56]. For the ΔRMSEA, with more than 300 samples, ΔRMSEA less than 0.015 implies that measurement invariance has been successfully established [57]. After the multigroup model invariance test was passed, we determined whether there were path differences between the different groups based on the model outputs, and paths with an absolute value of the critical ratio greater than 1.96 indicate a significant difference in the coefficients between the two groups (*P*<.05) [58].

Ethical Considerations

The research program was reviewed and approved by the ethical committee of Shandong University (number 20180225). For the original data collected, all participants had given informed consent to our study and were well aware of their right to withdraw from the study at any time. Our data have been completely anonymized and there is no information to identify the participants.

Results

Sample Characteristics

Table 1 shows the demographic characteristics of the participants, with a total of 1205 older adults included in the data analysis, of which 613 were MOAs and 592 were LOAs. Overall, 885 (73.4%) of the total sample were women while 320 (26.6%) were men; more than half (64.7%) of participants belonged to the 60 - 69 year old group; nearly half of older adults had a rural Hukou; 971 of 1205 participants (80.6%) were married; approximately four-fifths of older adults (n=975, 80.6%) were educated; and 785 (65.1%) older adults have a pension.

XSL•FO
RenderX

Table . Characteristics of participants and disparity between MOAs[a] and LOAs[b].

Variables		Total (n=1205), n (%)	MOA (n=613), n (%)	LOA (n=592), n (%)	Chi-square (*df*)	*P* value
Gender					0.083 (1)	.79
	Men	320 (26.6)	165 (26.9)	155 (26.1)		
	Women	885 (73.4)	448 (73.1)	437 (73.8)		
Age group (years)					139.631 (4)	<.001
	60 - 64	436 (36.2)	271 (44.2)	165 (27.9)		
	65 - 69	344 (28.5)	215 (35.1)	129 (21.8)		
	70 - 74	175 (14.5)	76 (12.4)	99 (16.7)		
	75 - 79	105 (8.7)	27 (4.4)	78 (13.2)		
	≥80	145 (12)	24 (3.9)	121 (20.4)		
Hukou					507.268 (1)	<.001
	Rural	649 (53.9)	525 (85.6)	124 (20.9)		
	Urban	556 (46.1)	88 (14.4)	468 (79.1)		
Marital status					43.045 (1)	<.001
	Married	971 (80.6)	539 (87.9)	432 (73)		
	Single	234 (19.4)	74 (12.1)	160 (27)		
Education level					41.675 (3)	<.001
	Illiterate	230 (19.1)	161 (26.2)	69 (11.7)		
	Primary school	403 (33.4)	185 (30.2)	218 (36.8)		
	Junior high school	338 (28)	158 (25.8)	180 (30.4)		
	High school and above	234 (19.4)	109 (17.8)	125 (21.1)		
Pension					106.504 (1)	<.001
	Yes	785 (65.1)	314 (51.2)	471 (79.6)		
	No	420 (34.9)	299 (48.8)	121 (20.4)		

[a]MOAs: migrant older adults.

[b]LOAs: local older adults.

The disparity between MOAs and LOAs was statically significant for age (*P*<.001), Hukou (*P*<.001), marital status (*P*<.001), education level (*P*<.001), and pension (*P*<.001). Specifically, nearly four-fifths of MOA participants (486/613, 79.3%) were aged 60 - 69 years, while less than 50% (294/592, 48.9%) of LOAs were in that age group. In total, 525 (85.6%) of MOAs were rural Hukou while 468 (79.1%) of LOAs were urban Hukou; in addition, there were more than twice as many single LOAs than single MOAs (n=160, 27% vs n=74, 12.1%); over one-fourth of MOAs (n=161, 26.2%) were illiterate, while approximately one-tenth of LOAs (n=69, 11.7%) were illiterate. Finally, 471 (79.6%) LOAs had a pension while only 121 (51.2%) MOAs had one.

Oral Health, Social Support, and Loneliness of the Participants

Table 2 illustrates the general characteristics of participants' oral health, life satisfaction, and loneliness, and the difference between MOAs and LOAs for the above variables. The total scores of GOHAI, SSRS, and ULS-6 for MOAs were 54.95 (SD 6.469), 38.89 (SD 6.629), 8.58 (SD 3.032), and 54.40 (SD 7.024), 39.51 (SD 6.856), and 8.00 (SD 2.790) for LOAs. Statistical differences between MOAs and LOAs were found in total ULS-6 score (t_{1203}=3.442, *P*=.001), SSRS score (objective support*:* t_{1203}=4.545, *P*<.001; subjective support*:* t_{1203}=−3.608, *P*<.001), and GOHAI score (psychosocial function*:* t_{1203}=2.028, *P*=.04). It is noted that there were no statistically significant differences in Q4 (t_{1203}=1.760, *P*=.08) and Q6 (t_{1203}=1.265, *P*=.21) of the total score of loneliness between MOAs and LOAs.

XSL•FO
RenderX

Table . General characteristics of the loneliness, social support, and oral health of MOA[a] and LOA[b] participants.

Variables	Total (n=1205), mean (SD)	MOA (n=613), mean (SD)	LOA (n=592), mean (SD)	*t* test (*df*=1203)	*P* value
Loneliness (ULS-6 [c])					
Total	8.29 (2.929)	8.58 (3.032)	8.00 (2.790)	3.442	.001
Often feel a lack of friends	1.45 (0.757)	1.53 (0.789)	1.36 (0.714)	3.704	<.001
Feel no one can be trusted	1.42 (0.722)	1.47 (0.728)	1.37 (0.713)	2.326	.02
Often feel left out	1.32 (0.587)	1.37 (0.622)	1.28 (0.547)	2.672	.008
Feel separated from others	1.34 (0.654)	1.37 (0.670)	1.30 (0.636)	1.760	.08
Often feel shy	1.31 (0.583)	1.36 (0.636)	1.25 (0.518)	3.365	.001
Surrounded but no one cares	1.46 (0.714)	1.49 (0.720)	1.43 (0.707)	1.265	.21
Social support (SSRS [d])					
Total	39.20 (6.746)	38.89 (6.629)	39.51 (6.856)	−1.612	.11
Objective support	8.22 (2.005)	8.47 (1.636)	7.95 (2.297)	4.545	<.001
Subjective support	23.94 (4.627)	23.47 (4.789)	24.43 (4.404)	−3.608	<.001
Utilization of support	7.04 (2.347)	6.94 (2.257)	7.14 (2.435)	−1.432	.15
Oral health (GOHAI [e])					
Total	54.68 (6.750)	54.95 (6.469)	54.40 (7.024)	1.408	.16
Physical function	17.14 (3.567)	17.35 (3.442)	16.93 (3.683)	2.028	.04
Psychological function	24.05 (2.217)	24.10 (2.061)	23.99 (2.369)	0.832	.41
Pain and discomfort	13.49 (2.162)	13.50 (3.115)	13.48 (2.211)	0.196	.84

[a]MOA: migrant older adults.

[b]LOA: local older adults.

[c]ULS-6: 6-item short-form UCLA Loneliness Scale.

[d]SSRS: Social Support Rating Scale.

[e]GOHAI: Geriatric Oral Health Assessment Index.

Structural Equation Model

Measurement Invariance Across Migration Status

Table 3 shows 5 selected models, which revealed related fit statistics of the measurement invariance across migration status and the fitness indexes. The fitness indexes of MOAs and LOAs should be compared to check whether the variable "migration state" was suitable for the group comparison.

In this study, χ^2, *df*, *P* value, GFI, AGFI, CFI, and RMSEA were the fitness indexes. As shown in Table 3, the fitness indexes of the MOAs were GFI=0.969, AGFI=0.952, CFI=0.971, and RMSEA=0.045 (M1), while for the LOAs, they were GFI=0.964, AGFI=0.944, CFI=0.940, and RMSEA=0.051 (M2). All fitness indexes showed values over 0.90 and very slight differences between the MOA and LOA groups, implying that we could further compare the differences between the MOA and LOA groups with the other models. Although the RMSEA value of M2 was more than 0.05, these variables were mainly used to calculate the change of RMSEA, not to assess the model's fitness. Then, the ΔCFI and ΔRMSEA between M3 (unconstrained model), M4 (measurement weights model), and M5 (structural weights model) were used to evaluate the measurement invariance. The M3 did not restrict any coefficient in the model, the M4 assumed the indicator loadings for the corresponding construct of each group are equal, and the M5 constrained the indicator loadings of the corresponding construct and the structural coefficients between the groups.

As seen in Table 3, the ΔCFI between M4 and M3 was 0.002; between M5 and M4, it was 0. All of the ΔCFI values were less than 0.010, indicating that measurement invariance was established between the models of M1, M2, M3, M4, and M5 between the MOA and LOA groups. The ΔRMSEA between M4 and M3 was 0.001, and it was 0 between M5 and M4. All of the ΔRMSEA values were less than 0.015, also indicating that measurement invariance was established between the models of M1, M2, M3, M4, and M5 across the MOA and LOA groups.

XSL•FO
RenderX

Table . Multigroup model invariance. The variables in 5 models were oral health, social support, and loneliness among MOAs[a] and LOAs[b] (n=1205).

Model	Chi-square (*df*)	*P* value	Chi-square/*df*	GFI[c]	AGFI[d]	CFI[e]	RMSEA[f]	ΔCFI[g]	ΔRMSEA[h]
M1[i]	112.152 (50)	<.001	2.243	0.969	0.952	0.971	0.045	—[j]	—
M2[k]	127.660 (50)	<.001	2.553	0.964	0.944	0.940	0.051	—	—
M3[l]	239.812 (100)	<.001	2.398	0.967	0.948	0.967	0.034	—	—
M4[m]	255.289 (109)	<.001	2.682	0.965	0.950	0.965	0.033	0.002	0.001
M5[n]	258.639 (112)	<.001	2.309	0.964	0.950	0.965	0.033	0.000	0.000

[a]MOAs: migrant older adults.

[b]LOAs: local older adults.

[c]GFI: goodness of fit index.

[d]AGFI: adjusted goodness of fit index.

[e]CFI: comparative fitness index.

[f]RMSEA: root mean square error of approximation.

[g]ΔCFI: change of CFI.

[h]ΔRMSEA: change of RMSEA.

[i]M1: MOA model.

[j]Not applicable.

[k]M2: LOA model.

[l]M3: unconstrained model.

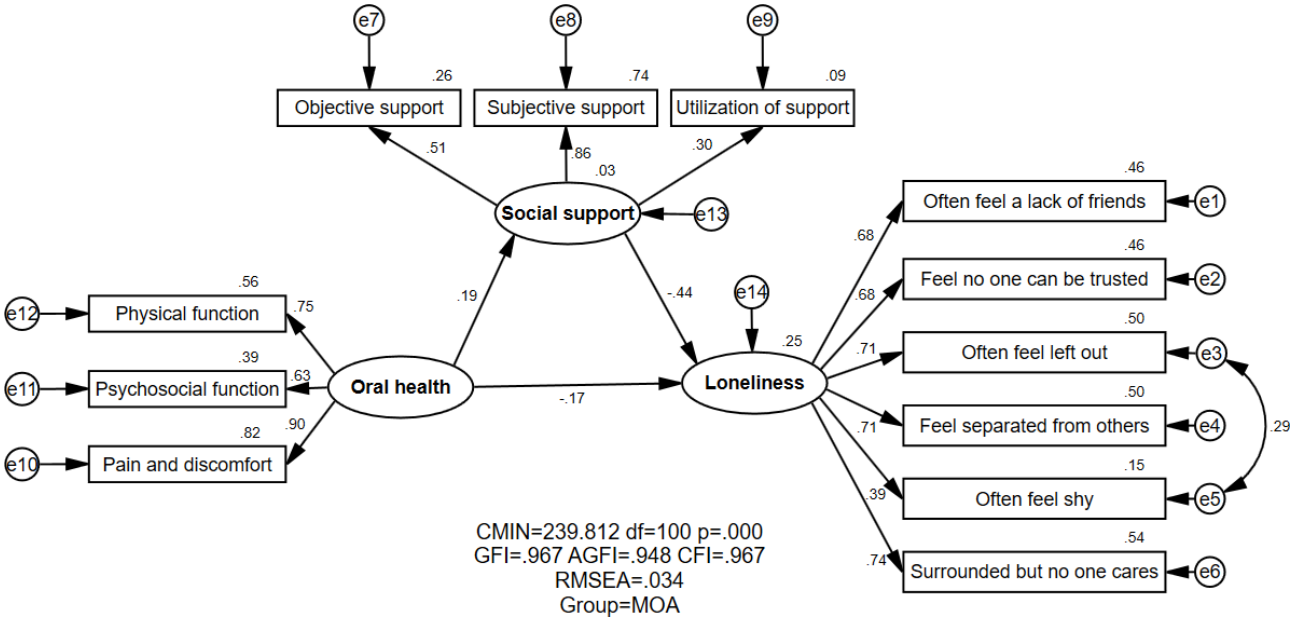
[m]M4: measurement weights model.

[n]M5: structural weights model.

Model Fitness Indexes

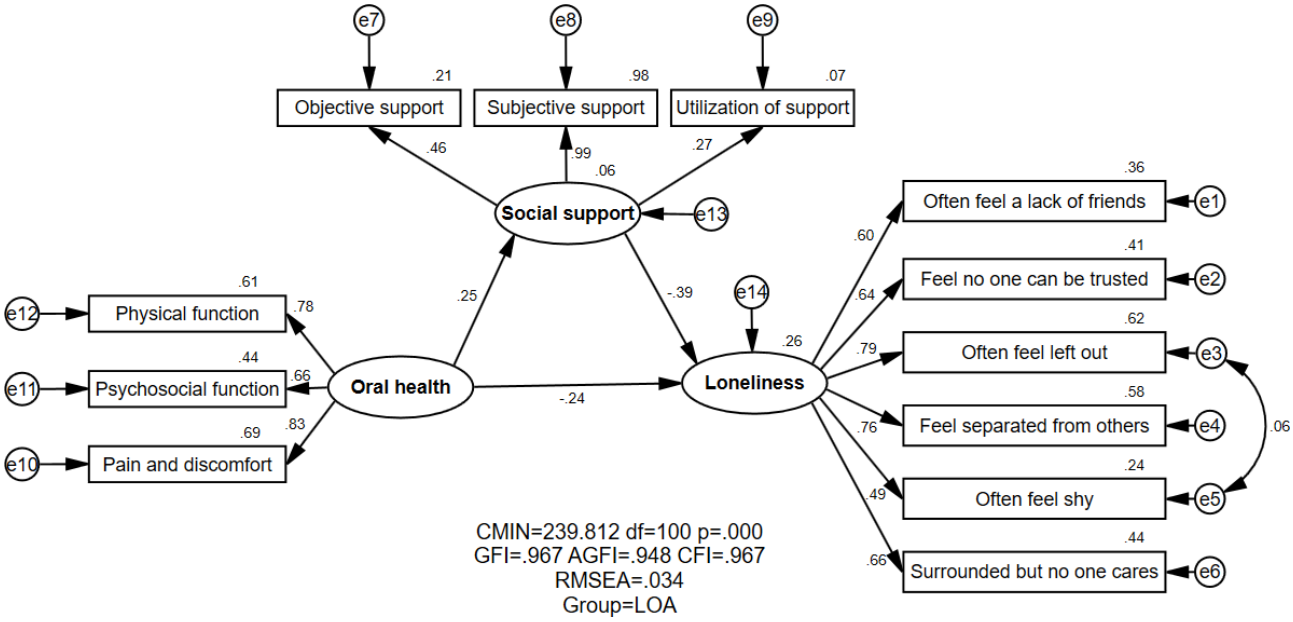
Figures 1 and 2 show the proposed models for MOAs and LOAs, respectively, which contained 3 variables: oral health, social support, and loneliness. Table 3 demonstrates the model fitness indexes for variable models (M1=MOA, M2=LOA). The MOA and LOA groups both had the same estimated value for model fitness: GFI=0.967, AGFI=0.948, CFI=0.967, and RMSEA=0.034. All fitness indexes implied that the theoretical model perfectly matched the empirical data for both the MOA and LOA groups.

Figure 1. Structural equation modeling analysis of the association between oral health, social support, and loneliness of MOAs (n=613). AGFI: adjusted goodness of fit index; CFI: comparative fitness index; CMIN: chi-square value; e: residual variables; GFI: goodness of fit index; MOA: migrant older adults; RMSEA: root mean square error of approximation. All parameter estimates were statistically significant (*P*<.05).



XSL•FO
RenderX

Figure 2. Structural equation modeling analysis of the association between oral health, social support, and loneliness of LOAs (n=592). AGFI: adjusted goodness of fit index; CFI: comparative fitness index; CMIN: chi-square value; e: residual variables; GFI: goodness of fit index; LOA: local older adult; RMSEA: root mean square error of approximation. All parameter estimates were statistically significant (*P*<.05).



Relationship Between Oral Health, Social Support, and Loneliness Assessed by SEM

Association Between Oral Health and Loneliness of Participants

The association between oral health, social support, and loneliness was shown in Figures 1 and 2 and Table 4. A negative and direct association was observed between oral health and loneliness among the MOAs (standardized direct effect=−0.168) and LOAs (standardized direct effect=−0.243). Moreover, oral health could exert a negative effect on loneliness indirectly via social support (standardized indirect effect=−0.083 for MOAs; standardized indirect effect=−0.097 for LOAs). It is noted that oral health was negatively associated with loneliness, which meant that MOAs and LOAs with higher oral health would generally have lower loneliness. A statistically significant relationship between oral health and loneliness was found in both the MOA and LOA groups.

Table . Standardized effects between oral health, social support, and loneliness among MOAs[a] and LOAs[b].

Path	Direct effect		Indirect effect		Total effect		Difference (critical ratio)
	MOA	LOA	MOA	LOA	MOA	LOA	
Oral health → Loneliness	−0.168[c]	−0.243[c]	−0.083[c]	−0.097[c]	−0.251[c]	−0.340[c]	−0.538
Oral health → Social support	0.186[c]	0.247[c]	—[d]	—	0.186[c]	0.247[c]	1.145
Social support → Loneliness	−0.444[c]	−0.392[c]	—	—	−0.444[c]	−0.392[c]	2.741[e]

[a]MOAs: migrant older adults.

[b]LOAs: local older adults.

[c]*P*<.001.

[d]Not applicable.

[e]*P*<.01.

Association Between Social Support and Loneliness of Participants

As for the relationship between social support and loneliness, a negative and direct effect was demonstrated among both MOAs (standardized direct effect=−0.444) and LOAs (standardized direct effect=−0.392), which meant less social support among both MOAs and LOAs would generally indicate higher loneliness. Concerning the group difference, a significantly negative correlation was slightly stronger in the MOA group than in the LOA group (critical ratio=2.741, *P*<.01).

Association Between Oral Health and Social Support of Participants

Oral health had a positive and direct effect on social support for both MOAs and LOAs (standardized direct effect=0.186 for MOAs; standardized direct effect=0.247 for LOAs), indicating that the higher the oral health of MOAs and LOAs, the higher

XSL•FO
RenderX

their social support. It was found that a statistically significant relationship existed between oral health and social support among MOAs and LOAs.

Discussion

Principal Findings

This study examined the severity of loneliness as well as the association between social support, oral health, and loneliness among older adults in Weifang. The results further showed a statistical difference in loneliness between MOAs and LOAs; the empirical associations between oral health, social support, and loneliness (including the local-migrant difference) were also clarified.

Loneliness Among MOAs and LOAs

The mean score of loneliness among MOAs (8.58) and LOAs (8.00) was lower than in a previous study conducted among rural empty-nest older adults in China (16.19) [51], indicating a lower level of loneliness among the MOAs and LOAs in this study. Moreover, this study found loneliness in MOAs was higher than in LOAs, which was similar to a study that showed that immigrant groups were lonelier than older adults born in Canada [59]. This may be due to the fact that, because of MOAs' migration, they need more time to adapt to their new environment while LOA have been in a familiar environment for a long time.

Association Between Oral Health and Loneliness

A negative association between oral health and loneliness was found among both MOAs and LOAs, which was similar to one existing study, which reported that older people who had a poor oral health status had higher odds of experiencing loneliness [60]. Ma and Chen [27] found that masticatory function, swallowing function, tooth loss, tooth function, and toothache were the influencing factors of loneliness among Chinese older people in the community. Another cross-sectional study in Indonesia showed similar findings, where older adults who had a poor oral status had a higher chance of feeling lonely [61]. Some studies pointed out that poor oral health increases psychological stress in communication among older adults, consequently limiting social interaction with others and causing loneliness [62,63]. However, the results showed that the effect of oral health on loneliness was not statistically different between MOAs and LOAs. This may be due to the fact that the relationship between oral health and loneliness is generalizable across older populations [25]. Another possible reason is that when facing oral health problems, both MOAs and LOAs will adopt different coping styles to decrease their loneliness. Some studies suggested that people with higher socioeconomic status and health literacy tend to proactively utilize medical resources to mitigate the negative impacts of oral problems [64]. In contrast, MOAs had poorer oral health services in their hometowns [65] and have experienced and adapted to more oral problems, which were more common and socially acceptable in their lives [66]. These further result in fewer psychological changes and lower loneliness even though oral problems occurred in the inflow cities.

Association Between Social Support and Loneliness

A negative relationship between social support and loneliness was found among both MOAs and LOAs, indicating that older adults with higher social support could reduce their loneliness. This was consistent with a previous nationwide cohort study, which also showed that social support decreased the odds of loneliness incidence among older adults in China [67]. The effect of social support on loneliness could be explained through the interaction theory, which views loneliness as a response to a lack of satisfying social networks and attachment partners [68]. Moreover, this study showed that the negative association between social support and loneliness was stronger among the MOAs than the LOAs. The economic development gap between urban and rural areas led to higher medical and education services in urban areas, which further resulted in population mobility mainly from rural to urban areas in China [69,70]. As for the MOAs, their migration may lead to a decline in the quantity and quality of social connections [71], which may further result in higher levels of loneliness among them than in the LOAs.

Association Between Oral Health and Social Support

The SEM results illustrated that oral health and social support were positively correlated, implying the better the oral health status, the higher the social support. This was consistent with previous findings, which suggested the experience of oral pain was associated with physical discomfort [72] and social barriers [73]. In addition, people with poor oral health may face additional mental health challenges due to problems such as halitosis, creating social anxiety and exacerbating barriers to interaction [74]. The results also noted that no significant difference was found between LOAs and MOAs when examining the impact of oral health on loneliness. This meant that the gradual loss of teeth and changes in facial morphology in older adults as they age, whether migrating or not, could lead to similar social limitations and increased psychological stress, which can affect their social participation and social support.

Implications

In order to reduce older adults' loneliness, the following measures should be taken. First, this study found that MOAs had higher loneliness than LOAs. Hence, strengthening social integration and promoting equality may be beneficial for reducing the loneliness of MOAs. Second, the results indicated that loneliness was negatively correlated with oral health; thus, it is important to aid older adults in maintaining their oral health. The government could speed up the process of including oral health services in health insurance reimbursements for older patients and enhance oral health education for them. Third, the findings showed that high social support would also reduce the loneliness of older adults. In particular, for MOAs, their social networks will be affected and they may face challenges such as language barriers and cultural differences after migration, making it more urgent for them to implement effective measures to alleviate their loneliness. Therefore, there is a need to enhance social support through the efforts of family members, communities, and the government. For instance, the government could implement policies to support young people whose families include older adults since older adults may receive

XSL•FO
RenderX

financial and emotional support from their children. It is necessary to promote the equalization of social security and welfare benefits for older adults who migrate with their families and further enhance their belonging in the inflow city.

Limitations

Some limitations existed in this study and should be addressed in future research. First, this study used cross-sectional data, therefore causal relationships could not be made and a longitudinal design is needed for follow-up research. Second, oral health status was assessed by a self-report scale with a lack of clinical evidence. In the future, clinical oral health examinations could be used to evaluate the oral health of older adults. Third, in addition to social support, there may be other variables that exert an indirect effect on loneliness among older adults and the variables used in this study may also be influenced by other confounding factors; therefore, more research is needed to verify their association.

Conclusion

Loneliness levels were fairly low among older adults in Weifang, China, while MOAs showed higher loneliness than LOAs. Oral health had both direct and indirect negative effects on loneliness among MOAs and LOAs, with no significant path differences between MOAs and LOAs. Social support was found to be negatively associated with loneliness for both MOAs and LOAs, while the association was stronger among MOAs than LOAs. Oral health exerted a significantly positive effect on social support for both MOAs and LOAs, while no significant difference existed between LOAs and MOAs.

Acknowledgments

The research team greatly appreciates the funding support obtained and thanks the research participants for their cooperation and support. This study was supported and funded by the National Natural Science Foundation of China (number 71804094), China Postdoctoral Science Foundation (number 2016M592161), Natural Science Foundation of Shandong Province (number ZR2016GB02), Postdoctoral Science Foundation of Shandong Province (number 201603021), and Fundamental Research Funds of Shandong University (number 2015HW002, number 2018JC055).

Data Availability

Under reasonable requirements, the data and material of this study can be obtained from the corresponding author. The data are not publicly available due to privacy restrictions.

Authors' Contributions

HL and JW conceptualized and designed the study; analyzed the data; and wrote the first draft of the manuscript. RC, XX, and MP contributed to surveys, data curation, and software and methodology usage. KF, BL, QL, ZQ, and SY contributed to data curation and visualization and validation of results. NIZ provided several valuable comments on the draft and corrected some issues related to grammar, vocabulary, and other aspects. FK applied for the funding used to support this study; designed the study; completed the questionnaire design; supervised and combined the data collected; instructed the writing, statistical analysis, and data processing; and provided comments on the modification of the manuscript; FK also polished the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
A flow diagram of the sample selection process.
[PNG File, 199 KB - publichealth_v11i1e66061_app1.png]

References

1. Ageing and health in China. World Health Organization. URL: https://www.who.int/china/health-topics/ageing [accessed 2025-02-28]
2. The seventh national population census [Website in Chinese]. National Bureau of Statistics. 2021. URL: https://www.stats.gov.cn/zt_18555/zdtjgz/zgrkpc/dqcrkpc [accessed 2025-02-28]
3. Pang M, Wang J, Zhao M, et al. The migrant-local difference in the relationship between social support, sleep disturbance, and loneliness among older adults in China: cross-sectional study. JMIR Public Health Surveill 2024 Jan 9;10:e49253. [doi: 10.2196/49253] [Medline: 38194253]
4. Xu J, Liu G, Li H, et al. Association between social integration and loneliness among the female migrant older adults with children: the mediating effect of social support. BMC Geriatr 2024 Jan 3;24(1):4. [doi: 10.1186/s12877-023-04569-8] [Medline: 38172722]

XSL•FO
RenderX

5. Schoenmakers D, Lamkaddem M, Suurmond J. The role of the social network in access to psychosocial services for migrant elderly-a qualitative study. Int J Environ Res Public Health 2017 Oct 11;14(10):1215. [doi: 10.3390/ijerph14101215] [Medline: 29019961]
6. Wang J, Pang M, Kong F. Association between self-reported oral health and life satisfaction among China's migrant elderly following children: the mediating effect of social support. Front Public Health 2023;11:950942. [doi: 10.3389/fpubh.2023.950942] [Medline: 36866087]
7. Wang J, Zhao J, Tian T, et al. The effects of oral health and social support on health-related quality of life of migrant older with children in Weifang, China. BMC Public Health 2022 Aug 6;22(1):1505. [doi: 10.1186/s12889-022-13843-0] [Medline: 35933414]
8. Deutsch S, House JS. Work stress and social support. Contemp Sociol 1983 May;12(3):329. [doi: 10.2307/2069001]
9. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [doi: 10.1037/0033-2909.98.2.310] [Medline: 3901065]
10. Weiss RS. Loneliness: The Experience of Emotional and Social Isolation: MIT Press; 1973. [doi: 10.2307/2062224]
11. Donovan NJ, Wu Q, Rentz DM, Sperling RA, Marshall GA, Glymour MM. Loneliness, depression and cognitive function in older U.S. adults. Int J Geriatr Psychiatry 2017 May;32(5):564-573. [doi: 10.1002/gps.4495] [Medline: 27162047]
12. Guo L, An L, Luo F, Yu B. Social isolation, loneliness and functional disability in Chinese older women and men: a longitudinal study. Age Ageing 2021 Jun 28;50(4):1222-1228. [doi: 10.1093/ageing/afaa271] [Medline: 33352582]
13. Dong T. Association between social support and loneliness in the elderly in China: based on the 2014 Chinese Longitudinal Elderly Social Survey. The World of Survey and Research 2017(8):3-9. [doi: 10.13778/j.cnki.11-3705/c.2017.08.001]
14. Lam J. Neighborhood characteristics, neighborhood satisfaction, and loneliness differences across ethnic-migrant groups in Australia. J Gerontol B Psychol Sci Soc Sci 2022 Nov 23;77(11):2113-2125. [doi: 10.1093/geronb/gbab219] [Medline: 34888639]
15. Wang J, Mann F, Lloyd-Evans B, Ma R, Johnson S. Associations between loneliness and perceived social support and outcomes of mental health problems: a systematic review. BMC Psychiatry 2018 May 29;18(1):156. [doi: 10.1186/s12888-018-1736-5] [Medline: 29843662]
16. Guo L, Wang W, Shi J, Zheng X, Hua Y, Lu C. Evaluation of social isolation trajectories and incident cardiovascular disease among middle-aged and older adults in China: national cohort study. JMIR Public Health Surveill 2023 Jun 30;9:e45677. [doi: 10.2196/45677] [Medline: 37389914]
17. Yu K, Wu S, Jang Y, et al. Longitudinal assessment of the relationships between geriatric conditions and loneliness. J Am Med Dir Assoc 2021 May;22(5):1107-1113. [doi: 10.1016/j.jamda.2020.09.002] [Medline: 33071157]
18. Paquet C, Whitehead J, Shah R, et al. Social prescription interventions addressing social isolation and loneliness in older adults: meta-review integrating on-the-ground resources. J Med Internet Res 2023 May 17;25:e40213. [doi: 10.2196/40213] [Medline: 37195738]
19. Margolis R, Chai X, Verdery AM, Newmyer L. The physical, mental, and social health of middle-aged and older adults without close kin in Canada. J Gerontol B Psychol Sci Soc Sci 2022 Jul 5;77(7):1350-1360. [doi: 10.1093/geronb/gbab222] [Medline: 34864999]
20. Jiang C, Jiang S. Elder mistreatment and life satisfaction of older adults: mediating roles of emotional closeness with children and loneliness. J Elder Abuse Negl 2021 Oct 20;33(5):351-367. [doi: 10.1080/08946566.2021.1980171]
21. Beridze G, Ayala A, Ribeiro O, et al. Are loneliness and social isolation associated with quality of life in older adults? Insights from Northern and Southern Europe. Int J Environ Res Public Health 2020 Nov 20;17(22):8637. [doi: 10.3390/ijerph17228637] [Medline: 33233793]
22. World Health Organization. Oral health. 2024. URL: https://www.afro.who.int/health-topics/oral-health [accessed 2025-02-28]
23. Feng X. Oral health status of chinese residents: report of the fourth oral health epidemiological survey in china. Presented at: 18th Academic Annual Meeting of the Preventive Dentistry of Chinese Stomatological Association; 2018; Xi'an, Shaanxi Province, China p. 14-17.
24. Kunrath I, Silva AER. Oral health and depressive symptoms among older adults: longitudinal study. Aging Ment Health 2021 Dec;25(12):2265-2271. [doi: 10.1080/13607863.2020.1855104] [Medline: 33291966]
25. Rouxel P, Heilmann A, Demakakos P, Aida J, Tsakos G, Watt RG. Oral health-related quality of life and loneliness among older adults. Eur J Ageing 2017 Jun;14(2):101-109. [doi: 10.1007/s10433-016-0392-1] [Medline: 28579932]
26. Koyama S, Saito M, Cable N, et al. Examining the associations between oral health and social isolation: a cross-national comparative study between Japan and England. Soc Sci Med 2021 May;277:113895. [doi: 10.1016/j.socscimed.2021.113895] [Medline: 33882441]
27. Ma X, Chen C. The relationship between the degenerative changes of visual, auditory, and oral functions and loneliness in the elderly in the community. Chin J Gerontol 2021;41(11):2403-2405. [doi: 10.13778/j.cnki.11-3705/c.2017.08.001]
28. Song J, Wu H, Lan H, Xu D, Wang W. The influence of disease status on loneliness of the elderly: evidence from rural China. Int J Environ Res Public Health 2022 Mar 4;19(5):3023. [doi: 10.3390/ijerph19053023] [Medline: 35270713]
29. Langford CPH, Bowsher J, Maloney JP, Lillis PP. Social support: a conceptual analysis. J Adv Nurs 1997 Jan;25(1):95-100. [doi: 10.1046/j.1365-2648.1997.1997025095.x]

XSL•FO
RenderX

30. Qian LFW, Xiaoling S, ling L, Jingyi H, Lunwei L. Correlation between successful aging and social support for the elderly in China: a meta-analysis. Chin J Geriat Care 2022;20(5):46-50.
31. Bai Y, Bian F, Zhang L, Cao Y. The impact of social support on the health of the rural elderly in China. Int J Environ Res Public Health 2020 Mar 18;17(6):2004. [doi: 10.3390/ijerph17062004] [Medline: 32197415]
32. Mahmud NA, Shahein NA, Yoep N, et al. Influence of social support on limitation in daily living among older persons in Malaysia. Geriatrics Gerontology Int 2020 Dec;20(S2):26-32 [FREE Full text] [doi: 10.1111/ggi.14029]
33. Wang J, Xue J, Jiang Y, Zhu T, Chen S. Mediating effects of depressive symptoms on social support and quality of life among rural older Chinese. Health Qual Life Outcomes 2020 Jul 20;18(1):242. [doi: 10.1186/s12955-020-01490-1] [Medline: 32690060]
34. Zhao X, Zhang D, Wu M, et al. Loneliness and depression symptoms among the elderly in nursing homes: a moderated mediation model of resilience and social support. Psychiatry Res 2018 Oct;268:143-151. [doi: 10.1016/j.psychres.2018.07.011] [Medline: 30025285]
35. Chung S, Kim M. Age-friendly environment, social support, sense of community, and loneliness among middle-aged and older adults in Korea. Aging Ment Health 2023;27(7):1352-1359. [doi: 10.1080/13607863.2022.2116395] [Medline: 36036282]
36. Vlachantoni A, Evandrou M, Falkingham J, Qin M. The impact of changing social support on older persons' onset of loneliness during the COVID-19 pandemic in the United Kingdom. Gerontologist 2022 Sep 7;62(8):1147-1159. [doi: 10.1093/geront/gnac033] [Medline: 35235949]
37. Sugaya N, Yamamoto T, Suzuki N, Uchiumi C. Loneliness and social isolation factors under the prolonged COVID-19 pandemic in Japan: 2-year longitudinal study. JMIR Public Health Surveill 2024 Sep 9;10:e51653. [doi: 10.2196/51653] [Medline: 39250195]
38. Wen C, Zhang Q. The role of intergenerational support in shaping oral healthcare-seeking behavior among older adults in China. Front Public Health 2023;11:1234539. [doi: 10.3389/fpubh.2023.1234539] [Medline: 37744500]
39. McGrath C, Bedi R. Influences of social support on the oral health of older people in Britain. J Oral Rehabil 2002 Oct;29(10):918-922. [doi: 10.1046/j.1365-2842.2002.00931.x] [Medline: 12421323]
40. Weber S, Hahnel S, Nitschke I, Schierz O, Rauch A. Older seniors during the COVID-19 pandemic-social support and oral health-related quality of life. Healthcare (Basel) 2021 Sep 7;9(9):1177. [doi: 10.3390/healthcare9091177] [Medline: 34574950]
41. Larsson P, Bondemark L, Häggman-Henrikson B. The impact of oro-facial appearance on oral health-related quality of life: a systematic review. J Oral Rehabil 2021 Mar;48(3):271-281. [doi: 10.1111/joor.12965] [Medline: 32196720]
42. Statistical bulletin of Weifang national economic and social development 2021 [Website in Chinese]. People's Goverment of Weifang City. 2022. URL: http://www.weifang.gov.cn/162/55356/15069075717458534 40.html [accessed 2025-02-28]
43. Weifang overview [Website in Chinese]. People's Goverment of Weifang City. 2023. URL: http://www.weifang.gov.cn/rwyd/wfgl [accessed 2025-02-28]
44. Main data of the seventh national population census in Weifang [Website in Chinese]. People's Government of Weifang City. 2022. URL: http://www.wfkjcs.com/ZiXun/XwDetail/dd4132ba55f743329c60b575b29d452e [accessed 2025-02-28]
45. Chan KW. Post - Mao China: a two - class urban society in the making. Int J Urban Regional Res 1996 Mar;20(1):134-150 [FREE Full text] [doi: 10.1111/j.1468-2427.1996.tb00305.x]
46. A-Dan W, Jun-Qi L. Factors associated with the oral health-related quality of life in elderly persons in dental clinic: validation of a Mandarin Chinese version of GOHAI. Gerodontology 2011 Sep;28(3):184-191. [doi: 10.1111/j.1741-2358.2009.00360.x] [Medline: 21692833]
47. Shao R, Hu T, Zhong YS, et al. Socio-demographic factors, dental status and health-related behaviors associated with geriatric oral health-related quality of life in Southwestern China. Health Qual Life Outcomes 2018 May 21;16(1):98. [doi: 10.1186/s12955-018-0925-8] [Medline: 29784008]
48. Xiao S. Theoretical foundations and research applications of the Social Support Rating Scale. J Clin Psychiatry 1994;02:98-100 [FREE Full text]
49. Chu H, Yang Y, Zhou J, et al. Social support and suicide risk among Chinese university students: a mental health perspective. Front Public Health 2021;9:566993. [doi: 10.3389/fpubh.2021.566993] [Medline: 33681117]
50. Zhao S, Long M, Diao Y, et al. Culture-related grief beliefs and social support influence depressive symptoms of Shidu parents in rural China. Eur J Psychotraumatol 2021;12(1):1945748. [doi: 10.1080/20008198.2021.1945748] [Medline: 34367525]
51. Wang G, Hu M, Xiao SY, Zhou L. Loneliness and depression among rural empty-nest elderly adults in Liuyang, China: a cross-sectional study. BMJ Open 2017 Oct 6;7(10):e016091. [doi: 10.1136/bmjopen-2017-016091] [Medline: 28988166]
52. Niu L, Jia C, Ma Z, Wang G, Yu Z, Zhou L. The validity of proxy-based data on loneliness in suicide research: a case-control psychological autopsy study in rural China. BMC Psychiatry 2018 May 1;18(1):116. [doi: 10.1186/s12888-018-1687-x] [Medline: 29716552]
53. Wu M. Structural Equation Modeling: Operation and Application of AMOS: Chongqing University Press; 2010.
54. Fan X, Thompson B, Wang L. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. Struct Equ Modeling 1999 Jan;6(1):56-83. [doi: 10.1080/10705519909540119]

XSL•FO
RenderX

55. Kline RB. Principles and Practice of Structural Equation Modeling, 4th edition: The Guilford Press; 2016. URL: https://api.semanticscholar.org/CorpusID:164967627 [accessed 2025-03-09]
56. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling 2002 Apr;9(2):233-255. [doi: 10.1207/S15328007SEM0902_5]
57. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct Equ Modeling 2007 Jul 31;14(3):464-504. [doi: 10.1080/10705510701301834]
58. Rong T. Amos and Research Methodology: Chongqing University Press; 2009.
59. De Jong Gierveld J, Van der Pas S, Keating N. Loneliness of older immigrant groups in Canada: effects of ethnic-cultural background. J Cross Cult Gerontol 2015 Sep;30(3):251-268. [doi: 10.1007/s10823-015-9265-x] [Medline: 25982532]
60. Kisely S. No mental health without oral health. Can J Psychiatry 2016 May;61(5):277-282. [doi: 10.1177/0706743716632523] [Medline: 27254802]
61. Susanty S, Chung MH, Chiu HY, et al. Prevalence of loneliness and associated factors among community-dwelling older adults in Indonesia: a cross-sectional study. Int J Environ Res Public Health 2022 Apr 18;19(8):4911. [doi: 10.3390/ijerph19084911] [Medline: 35457781]
62. Bernabé E, de Oliveira C, de Oliveira Duarte YA, Bof de Andrade F, Sabbah W. Social participation and tooth loss, vision, and hearing impairments among older Brazilian adults. J Am Geriatr Soc 2023 Oct;71(10):3152-3162. [doi: 10.1111/jgs.18423] [Medline: 37227109]
63. Igarashi A, Aida J, Yamamoto T, Hiratsuka Y, Kondo K, Osaka K. Associations between vision, hearing and tooth loss and social interactions: the JAGES cross-sectional study. J Epidemiol Community Health 2021 Feb;75(2):171-176. [doi: 10.1136/jech-2020-214545] [Medline: 32972921]
64. Yuan L, Yuan Y, Ren H, et al. Decomposition analysis of the prevalence of denture use between rural and urban older individuals with edentulism in China: cross-sectional study. Interact J Med Res 2024 Dec 13;13:e48778. [doi: 10.2196/48778] [Medline: 39671593]
65. Xu J, Zheng J, Xu L, Wu H. Equity of health services utilisation and expenditure among urban and rural residents under universal health coverage. Int J Environ Res Public Health 2021 Jan 12;18(2):593. [doi: 10.3390/ijerph18020593] [Medline: 33445637]
66. Olofsson H, Ulander EL, Gustafson Y, Hörnsten C. Association between socioeconomic and health factors and edentulism in people aged 65 and older - a population-based survey. Scand J Public Health 2018 Nov;46(7):690-698. [doi: 10.1177/1403494817717406] [Medline: 28768459]
67. Yang F, Gu D. Predictors of loneliness incidence in Chinese older adults from a life course perspective: a national longitudinal study. Aging Ment Health 2020 Jun;24(6):879-888. [doi: 10.1080/13607863.2018.1558174] [Medline: 30621448]
68. Beal C. Loneliness in older women: a review of the literature. Issues Ment Health Nurs 2006;27(7):795-813. [doi: 10.1080/01612840600781196] [Medline: 16849264]
69. Zhang K, Burr JA, Mutchler JE, Lu J. Internet use and loneliness among urban and non-urban Chinese older adults: the roles of family support, friend support, and social participation. The Journals of Gerontology, Series B 2024 Jul 1;79(7):gbae081. [doi: 10.1093/geronb/gbae081]
70. Zhang X, Dong S. The relationships between social support and loneliness: a meta-analysis and review. Acta Psychol (Amst) 2022 Jul;227:103616. [doi: 10.1016/j.actpsy.2022.103616] [Medline: 35576818]
71. Jia W, Liu L, Peng G. The impact of social media on users' self-efficacy and loneliness: an analysis of the mediating mechanism of social support. Psychol Res Behav Manag 2024;17:593-612. [doi: 10.2147/PRBM.S449079] [Medline: 38379636]
72. Jönsson H, Michaud M, Neuman N. What is commensality? A critical discussion of an expanding research field. Int J Environ Res Public Health 2021 Jun 9;18(12):6235. [doi: 10.3390/ijerph18126235] [Medline: 34207626]
73. Masood M, Newton T, Bakri NN, Khalid T, Masood Y. The relationship between oral health and oral health related quality of life among elderly people in United Kingdom. J Dent 2017 Jan;56:78-83. [doi: 10.1016/j.jdent.2016.11.002] [Medline: 27825838]
74. Goergen J, Costa RSA, Rios FS, et al. Oral conditions associated with oral health related quality of life: a population-based cross-sectional study in Brazil. J Dent 2023 Feb;129:104390. [doi: 10.1016/j.jdent.2022.104390] [Medline: 36526085]

Abbreviations

AGFI: adjusted goodness of fit Index
CFI: comparative fitness index
CMIN: chi-square value
GFI: goodness of fit index
GOHAI: Geriatric Oral Health Assessment Index
LOA: local older adults
MOA: migrant older adults
RMSEA: root mean square error of approximation

XSL•FO
RenderX

SSRS: Social Support Rating Scale
ULS-6: 6-item short-form UCLA Loneliness Scale

Edited by A Mavragani; submitted 03.09.24; peer-reviewed by AK Chun Chau, J Zhao; revised version received 17.02.25; accepted 18.02.25; published 13.03.25.

Please cite as:
Liu H, Wang J, Chen R, Xu X, Pang M, Feng K, Li B, Li Q, Qin Z, Yan S, Ibn Ziyat N, Kong F
Migrant-Local Differences in the Relationship Between Oral Health, Social Support, and Loneliness Among Older Adults in Weifang, China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e66061
URL: https://publichealth.jmir.org/2025/1/e66061
doi:10.2196/66061

© Hui Liu, Jieru Wang, Rui Chen, Xixing Xu, Mingli Pang, Kaiyuan Feng, Bingsong Li, Qinling Li, Ziwei Qin, Shuyi Yan, Nabila Ibn Ziyat, Fanlei Kong. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multidimensional Evaluation of the Process of Constructing Age-Friendly Communities Among Different Aged Community Residents in Beijing, China: Cross-Sectional Questionnaire Study

Yingchun Peng[1*], PhD; Zhiying Zhang[2*], MS; Ruyi Zhang[3], MS; Yiyao Zhang[1], BS; Runying Wang[1], BS; Jiaying Zhang[1], BS; Shaoqi Zhai[1], MS; Qilin Jin[4], MS; Jiaojiao Zhou[5], MS; Jingjing Chen[6], MS

[1]School of Medical Humanities, Capital Medical University, No.10, Xitoutiao, You An Men Wai, Beijing, China

[2]Outpatient Office, Beijing Luhe Hospital, Capital Medical University, Beijing, China

[3]Ethics Committee Office, Beijing Ditan Hospital, Capital Medical University, Beijing, China

[4]Cardiac Surgery Department, People's Hospital of Beijing Daxing District, Beijing, China

[5]Medical Department, Fengtai District Xiluoyuan Community Health Service Center, Beijing, China

[6]Huairou District Liulimiao Community Health Service Center, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Yingchun Peng, PhD
School of Medical Humanities, Capital Medical University, No.10, Xitoutiao, You An Men Wai, Beijing, China

Abstract

Background: The World Health Organization (WHO) has made significant efforts to promote age-friendly community initiatives (AFCI) to address the challenges of population aging. Previous studies have discussed the construction of age-friendly communities (AFC) in urban cities, evaluating AFCs often rooted in the WHO's Checklist and focused on a single group, namely older adults, overlooking the role of other age groups in community development.

Objective: This study aims to evaluate AFCs from multidimensional aspects, particularly the positive living experiences of older adults, summarize the deficiencies in both hardware and software aspects in the process of constructing AFCs in China, and provide some recommendations to promote AFCIs worldwide.

Methods: Using a multistage sampling strategy, 470 community residents from urban and suburban areas participated in this study. A self-designed questionnaire was designed to use a standardized method to evaluate older adults' living experiences across five dimensions, including the degree of age-friendliness in the community, social support, sense of gain, sense of happiness, and sense of security. Respondents rated each dimension on a 10-point scale. This study defined community residents into 3 groups: residents younger than 45 years(Group 1: youth), those aged 45-59 years (Group 2: middle-aged), and those aged ≥60 years (Group 3: old-age).

Results: In this study, 382 (81.3%) community residents were unaware of the relevant concepts of AFCs. Most participants highlighted the importance of community support and health services, followed by respect and social inclusion, and outdoor spaces and buildings. The findings showed that the highest-rated dimension was the sense of security. The mean scores for the degree of the sense of security in urban and suburban areas were 7.88 (SD 1.776) and 7.73 (SD 1.853), respectively. For Group 2, the mean scores were 7.60 (SD 2.070) and 8.03 (SD 1.662), while Group 3 had mean scores of 7.34 (SD 2.004) and 7.91 (SD 1.940). The lowest-rated dimension was social support; the mean scores for Group 1 for the degree of social support in urban and suburban areas were 7.63 (SD 1.835) and 7.48 (SD 1.918), respectively. For Group 2, the mean scores were 6.94 (SD 2.087) and 7.36 (SD 2.228), while those for Group 3 were 6.37 (SD 2.299) and 6.84 (SD 2.062). Further, there were significant differences in the scores of residents among different age groups in urban areas regarding age-friendliness (P<.001), social support, (P<.001), and sense of gain (P=.01).

Conclusions: China is in the early stages of developing AFCs. We further highlight the importance of continued research on the collaboration and participation among multiple stakeholders. These outcomes have a direct and positive impact on the well‐being of older adults.

(JMIR Public Health Surveill 2025;11:e66248) doi:10.2196/66248

KEYWORDS

age-friendly; positive living experience; active aging; healthy aging; community residents; older adults; age-friendly communities; multiple stakeholders; evaluating age-friendliness; urban and suburban areas

XSL•FO
RenderX

Introduction

The History and Imperative of Implementing Age-Friendly Communities

Population aging is poised to become one of the most significant problems in the 21st century, impeding the development of many countries globally. Data from *World Populations Prospects* [1] predicated that the number of persons aged ≥65 years in countries and areas with populations that have already peaked is projected to reach 409 million by 2027, surpassing the number of children under 18. The number of persons aged ≥80 years is increasing at an even faster rate and is likely to triple, from 85 million in 2024 to approximately 267 million by 2100. With the increasing trend of population aging, health issues in older adults have become a major concern for scholars and policymakers worldwide. As people age, many experience a myriad of changes impacting their health and functional ability to manage independent living in their homes and communities [2]. Meanwhile, age-related changes such as physiological and cognitive decline may lead to negative health outcomes such as frailty, mobility limitations, and disability. Moreover, many social and physical environmental factors also affect the health status and quality of life of older adults [3,4]. Therefore, finding accessible and affordable measures to help older adults live longer and healthier will be crucial.

To address the challenges of population aging and achieve the goal of healthy aging, the World Health Organization(WHO) has made significant efforts to promote age-friendly community initiatives (AFCI). The concept of age-friendly communities (AFCs) originated from the Ecology Theory of Aging in 1973 [5,6], which indicated that aging was a dynamic process at both the individual and community levels, revealing that a positive surrounding environment can promote well-being among older adults. In 2000, the American Association of Retired Persons first defined AFCs as communities with affordable housing, adequate infrastructure and services, and diverse options of transportation. Based on the report *A Guide: Global Age-Friendly Cities* [7], the AFCs incorporated three aspects and eight key domains. Three aspects included service support, space environment, social and humanistic environment [8], and the eight key domains consisted of outdoor spaces and buildings, transportation, housing, social participation, respect and social inclusion, civic participation and employment, communication and information, and community and health services [9,10]. In the current conceptualization of AFCs, the importance of social environment and physical environment are placed equally as the determinants of health and older adults are considered to be effective social resources involved in community construction [11]. The AFCs are designed to account for the wide diversity among older people and promote their autonomy, inclusion, and contributions in all areas of community life [12].

Meanwhile, the implementation of aging-in-place policies highlights the imperative for AFCIs [13]. Supporting older adults to continue living safely and comfortably within their communities is a shared vision of many governments and international organizations. Due to their continuous efforts, the WHO Global Network for AFCIs currently includes 1705 cities and communities across 60 countries, covering over 330 million people worldwide. However, the lack of measurability of the actual age-friendliness of AFCs has been a major weakness in moving AFCIs forward. As age-friendliness is a complex, dynamic, multidimensional, and highly context-dependent concept, it does not easily lend itself to standardization of measurement [14-16].

Evaluating the Age-Friendliness of Communities

While constructing AFCs, it is crucial to evaluate the age-friendliness of communities [17]. The previous studies have mainly focused on measuring the age-friendliness of a city or community in which older adults live, based on the WHO's Checklist [18]. Jo-Ying Huang et al [19] developed indicators of age-friendliness for communities in Taiwan province that conform to international standards by referring to the WHO's Checklist and Taiwan's existing indicators. Yu, Wong, and Woo [20] examined the relationships between perceptions of neighborhood environment, sense of community, and self-rated health by relying on the WHO Checklist. Wood et al [21] conducted a systematic scoping review of 23 articles using citizen science or participatory approaches. An inductive and deductive thematic analysis was performed to identify local urban barriers and facilitators and map them against the WHO's Checklist. Kim et al [22] used data from the American Association of Retired Persons AFC surveys to evaluate the reliability and concurrent validity of AFC assessment tools. They found that higher levels of perceived availability of community and health services were associated with worse health outcomes.

Additionally, the Age Friendly Cities and Communities Questionnaire (AFCCQ) is designed by Dikken for measuring age-friendliness. Compared with WHO's Checklist, the Age Friendly Cities and Communities Questionnaire only has 23 items, covers eight domains of the WHO's AFC model, and has an additional ninth domain of financial situation [23]. In addition to the above-mentioned measurement tools of age-friendliness, sets of priority indicators, frameworks, and conceptual surveys are proposed by other studies. Jiravanichkul et al [24] identified minimum standard indicators and explored the development of a well-being environment and AFC assessment criteria by using the analytic hierarchy process. Ide et al [17] developed a community-level AFC indicator based on the WHO AFC guidelines by incorporating dementia-friendly elements and tested its validity and reliability.

Beijing as a Case Example

Beijing is the capital of China, and a municipality directly governed by the Central Government. It is famous as China's political, cultural, international communication, and science and technology innovation center. Beijing currently has entered a moderately aging society. By the end of 2023, the permanent resident population aged 60 years and above was 4.948 million, accounting for 22.6% of the total permanent resident population. The resident population aged 65 years and above was 3.469 million, accounting for 15.9% of the total resident population [25]. To enhance older adults' sense of gain, happiness, and security in communities, the National Health Commission (National Office for the Elderly, 2020) launched an initiative

XSL•FO
RenderX

to construct AFCs across the country. By the end of 2023, a total of 93 communities in Beijing had been named " National Model AFCs," which were at the forefront of AFC development nationwide.

Focus of the Current Study

Although many scholars have attempted to evaluate the age-friendliness of communities in various ways, there is limited research emphasizing the positive aspects of health, such as the sense of gain, happiness, and security that older adults experience in their communities. Additionally, most studies have focused on the construction of AFCs in urban cities [26,27], and evaluating the age-friendliness of communities was only based on assessments from older adults [22,28], which overlooked the perspectives of other community residents. Therefore, the current evaluation of constructing AFCs may lack objectivity and accuracy.

This study aims to explore the cognition of AFCs among the different aged communities' residents in Beijing, evaluate AFCs from multidimensional aspects, especially the positive living experience of older adults, summarize the deficiencies in hardware and software infrastructure in the process of constructing AFCs in Beijing, and provide a reference for promoting AFCIs across the world.

Methods

Study Design

This is a cross-sectional study, designed to understand the awareness and perception of AFCs among different aged community residents, explore deficiencies in the process of constructing AFCs, and propose recommendations for improvement.

Study Setting and Sample

Considering the level of economic development and construction status of AFCs among 16 districts in Beijing, a cross-sectional quantitative survey was conducted on community residents between urban and suburban areas. This study adopted a multistage sampling strategy [29,30]. In the first stage, Xicheng District, Fengtai District, Daxing District, and Huairou District were selected based on the functional orientation of the 16 districts in Beijing and the current situation of constructing AFCs. To implement urban strategic positioning, promoting sustainable development, and considering the practical needs of the historical pattern of the ancient capital, Beijing issued the Beijing Urban Master Plan in 2017. According to the Beijing Urban Master Plan, there are six central urban areas (including Dongcheng District, Xicheng District, Chaoyang District, Haidian District, Fengtai District, and Shijingshan District), five new towns in plain areas (including Shunyi District, Daxing District, Yizhuang District, Fangshan District, and Changping District), and six ecological conservation areas (including Mentougou District, Pinggu District, Huairou District, Miyun District, and Yanqing District). For this study, plain new towns and ecological conservation areas were unified as suburban areas, while the central urban areas were unified as urban areas. Among the 16 districts, Xicheng District, Fengtai District, Daxing District, and Huairou Districts ranked second, third, fifth, and eighth, respectively. In the second stage, 3 - 5 communities were randomly picked from each district, amounting to 15 communities in the list. In the last stage, trained investigators visited each sampling community to randomly invite community residents to participate in our investigation. Potential participants were recruited via the invitation of the investigators, and we sought to balance the sample based on population characteristics such as gender, age, education, and region. To better understand the construction of AFCs, our goal was to recruit residents who lived in the communities and who could freely and voluntarily express their insights.

Measures

A self-designed questionnaire was developed to use a standardized method for assessing the residents' living experience and their community's age-friendliness in Beijing. It was designed in consultation with a group of AFC experts [31].

To explore the cognition of community residents regarding the concept of AFCs, the questionnaire included questions such as whether they had heard of or learned about AFCs. Based on our literature review and WHO guidelines, we asked residents to choose the most important domain among the eight domains (outdoor spaces and buildings, transportation, housing, social participation, respect and social inclusion, civic participation and employment, communication and information, and community and health services) in the process of constructing AFCs.

The positive living experience of older adults that was obtained within their community was measured across five dimensions: (1) the degree of age-friendliness in the community, (2) the level of social support for older adults, (3) the sense of gain in community, (4) the sense of happiness in the community, and (5) the sense of security in the community. Respondents rated each dimension on a 10-point scale (0 - 1: does not exist, 2 - 3: poor, 4 - 5: fair, 6 - 7: good, 8 - 9: very good, 10: excellent).

To understand the shortcomings in the current process of constructing AFCs, the deficiencies were categorized into hardware and software aspects. Hardware aspects included increasing the green spaces and beautiful buildings, building more nursing institutions and medical institutions for older adults, building more leisure and entertainment places, installing more age-friendly facilities for older adults, and expanding traffic roads, among others. Software aspects consisted of conducting more abundant recreational activities, creating an age-friendly atmosphere, providing personalized elderly care services and community and neighborhood support, providing community volunteer services, and more personalized medical services. Participants were asked to choose the items that required urgent improvement in their communities.

Data Analysis

Based on the latest WHO age definition criteria [32], this study defined community residents into three age groups: (1) Group 1 as youth (younger than 45 years) (2) Group 2 as middle-aged adults (45-59 years), and (3) Group 3 as old-aged adults (above 60 years).

XSL•FO
RenderX

All the gleaned data were recorded into EpiData software (version 3.1; EpiData Association) and analyzed by SPSS software (version 21.0; IBM Corp). The mean and standard deviation were used to describe continuous data, while the categorical data were presented by composition ratio, frequency distributions, and parity arrangement. Frequency and rank analyses were used to summarize the quantitative data of community residents including demographic characteristics such as gender, age, region, and education level. The rank-sum test was used to analyze the most important domains of AFCs and deficiencies in the construction of the AFCs. One-way ANOVA was used to explore the five dimensions of the community living experiences of older adults.

Ethical Considerations

The study was approved by the Medical Ethics Committee of Capital Medical University, Beijing, China (Reference number Z2023SY048). All participants provided informed consent for the collection, handling, and storage of their personal and health data. All procedures were performed in accordance with relevant guidelines and regulations. All participants joined the study voluntarily, and no compensation was provided to them. All data were kept confidential, deidentified, and anonymous.

Results

A total of 477 residents voluntarily agreed to participate in our survey. We collected 470 valid questionnaires were collected, resulting in a validity rate of 98.53%. Table 1 presents the demographic characteristics of community residents. Among the 470 community residents in our survey, the majority were female (n=339, 72.1%). Approximately 44.5% (n=209) of the participants were younger than 44 years, and the average age of the 470 community residents was 46.74 (SD 18.63) years. A total of 259 (55.1%) community residents were from urban areas. Additionally, 113 participants (24%) had lived in their current community for more than 30 years and 178 (37.9%) participants had obtained a bachelor's degree.

Table . Demographic characteristics of community residents.

Items	Surveys (N=470), n (%)
Gender	
Male	131 (27.9)
Female	339 (72.1)
Age (years)	
≤44	209 (44.5)
45-59	132 (28.1)
≥60	129 (27.4)
Regions	
Urban areas	259 (55.1)
Suburban areas	211 (44.9)
Length of residence (years)	
≤10	157 (33.4)
11-30	200 (42.6)
31-50	53 (11.3)
51-70	42 (8.9)
≥71	18 (3.8)
Education	
Junior high school or below	129 (27.4)
High school or Junior college	99 (21.1)
Bachelor's degree	178 (37.9)
Master's degree or above	64 (13.6)

Table 2 shows community residents' cognition of AFCs. Almost 81.3% of the community residents were not aware of the concept of AFCs; only 88 (18.7%) community residents had heard of AFCs.

XSL•FO
RenderX

Table . Community residents' cognition of the Age-Friendly Communities.

Items	Results
Community residents' cognition of the concept of the Age-Friendly Communities (N=470), Surveys, n (%)	
Know	88 (18.7)
Not know	382 (81.3)
The most important domain of the Age-Friendly Communities from the perspective of community residents, total points[a] (rank)	
Outdoor spaces and buildings	453 (3)
Transportation	260 (6)
Housing	293 (5)
Social participation	364 (4)
Respect and social inclusion	559 (2)
Job opportunities and civic participation	84 (8)
Communication and information	135 (7)
Community support and health services	666 (1)

[a]Note: total points: number of people selected for the first important domain ×3 + number of people selected for the second important domain ×2 +number of people selected for the third important domain ×1.

With regard to the eight important domains of creating an AFC, community support and health services were the most important domains from the perspective of residents when they lived in the community, followed by respect and social inclusion, and outdoor spaces and buildings. Job opportunities and civic participation were considered relatively less important for residents when they lived in the community.

Table 3 displays an evaluation of the living experiences of older adults, categorized by different aged community residents. Among the five dimensions of older adults' living experience, the scores of the sense of security in the community were relatively high, whereas the scores for the degree of social support for older adults were relatively low. There were significant differences in the scores of residents between different age groups in urban areas in terms of age-friendliness (P<.001), social support (P<.001), and sense of gain (P=.01), whereas no significant differences were found in terms of happiness and security. In the evaluation of age-friendliness, there were significant differences between young and middle-aged residents, and between young residents and elderly residents. In the evaluation of the degree of social support, there were significant differences between young residents and middle-aged residents, and between young residents and elderly residents. Additionally, the sense of gain differed significantly between young residents and elderly residents.

XSL•FO
RenderX

Table . Evaluation of living experience of older adults from the perspectives of different aged community residents.

Regions, items, and groups	Scores[a] , mean (SD）	*F* test (*df*)	*P* value
Urban areas			
The degree of age-friendliness		9.687（2, 256）	<.001
Group 1（age <45 years)	7.82 (1.807)[a]		
Group 2 (age 45-60 years)	7.34 (1.689)[a]		
Group 3 （age ≥60 years)	6.58 (2.208)[b]		
The degree of social support		8.914（2, 256）	<.001
Group 1（age <45 years)	7.63 (1.835)[a]		
Group 2 (age 45-60 years)	6.94 (2.087)[b]		
Group 3 (age ≥60 years)	6.37 (2.299)[b]		
The sense of gain in community		4.445（2, 256）	.01
Group 1 (age <45 years)	7.58 (1.693)[a]		
Group 2 (age 45-60 years)	7.10 (1.995)[ab]		
Group 3 (age ≥60 years)	6.79 (1.880)[b]		
The sense of happiness in community		2.741（2, 256）	.07
Group 1 (age <45 years)	7.72 (1.777)		
Group 2 (age 45-60 years)	7.27 (1.933)		
Group 3 (age ≥60 years)	7.09 (2.079)		
The sense of security in community		1.792（2, 256）	.17
Group 1 (age <45 years)	7.88 (1.776)		
Group 2 (age 45-60 years)	7.60 (2.070)		
Group 3 (age ≥60 years)	7.34 (2.004)		
Suburban areas			
The degree of age-friendliness		1.290（2, 208）	.28
Group 1 (age <45 years)	7.61 (1.914)		
Group 2 (age 45-60 years)	7.44 (2.232)		
Group 3 (age ≥60 years)	7.05 (1.910)		
The degree of social support		1.732（2, 208）	.18
Group 1 (age <45 years)	7.48 (1.918)		
Group 2 (age 45-60 years)	7.36 (2.228)		
Group 3 (age ≥60 years)	6.84 (2.062)		
The sense of gain in community		0.138（2, 208）	.87
Group 1 (age <45 years)	7.29 (1.923)		
Group 2 (age 45-60 years)	7.41 (2.045)		
Group 3 (age ≥60 years)	7.21 (2.145)		
The sense of happiness in community		0.519(2, 208）	.60
Group 1 (age <45 years)	7.53 (1.719)		
Group 2 (age 45-60 years)	7.51 (2.081)		
Group 3 (age ≥60 years)	7.82 (1.724)		
The sense of security in community		0.561（2, 208）	.57

XSL•FO
RenderX

Regions, items, and groups	Scores[a] , mean (SD)	*F* test (*df*)	*P* value
Group 1 (age <45 years)	7.73 (1.853)		
Group 2 (age 45-60 years)	8.03 (1.662)		
Group 3 (age ≥60 years)	7.91 (1.940)		

[a]Different letters within the same row indicate statistically significant differences between groups based on the Student-Newman-Keuls (SNK) multiple comparisons (P<.05). Groups sharing the same letter do not differ significantly.

Table 4 presents the shortcomings in the current process of constructing AFCs. Regarding the aspects of community hardware facilities, most residents expressed the need for building more nursing institutions for older adults, followed by building additional medical institutions and installing more suitable age-friendly facilities. Expanding traffic roads was an aspect that concerned fewer people. Regarding the community software facilities, creating an age-friendly atmosphere was most commonly reported by many community residents, followed by providing personalized elderly care services, more personalized medical services, and more recreational activities.

Table . The shortcomings in the current process of constructing age-friendly communities.

Items	Frequency	Rank
The aspects of community hardware facilities that should be improved		
Increase the green area and beautiful buildings	217	5
Build more nursing institutions for older adults	319	1
Build more medical institutions for older adults	257	2
Build more leisure and entertainment places	221	4
Install more suitable facilities for older adults	252	3
Expand traffic roads	78	6
Others	11	7
The aspects of community software facilities that should be improved		
Hold more abundant recreational activities	207	4
Create an age-friendly atmosphere	323	1
Provide personalized elderly care services	317	2
Provide community and neighborhood support	117	6
Provide community volunteer services	170	5
Provide more intimate medical services	241	3
Others	13	7

Discussion

Principal Findings

This study is the first to comprehensively evaluate AFCs from multidimensional aspects in China. It particularly examines the positive living experiences that residents derive from their communities, compares the cognition and perception of AFCs among different aged community residents in urban and suburban areas, and explores the deficiencies in the process of constructing AFCs from both hardware and software perspectives. The community residents' suggestions for improvement for constructing AFCs reflect the needs and interests of older adults and other age groups, providing a direction for constructing AFCs in the future.

Increasing Age-Friendly Awareness and Sense of Belonging Among Community Residents

China has only recently begun promoting AFCIs; compared to other developed countries, there remains significant progress to be made. Consequently, the public is not aware of the concepts of AFCs, which is consistent with our findings as only 88 (18.7%) residents were aware of AFCs. Recognizing the age-friendliness of communities is the premise to engage residents in the construction of AFCs. Relevant studies show that perceived age-friendliness of the community is positively associated with a sense of belonging to the community [33,34]. However, a gap exists between family belonging and community belonging. Given the changes in China's economic and social structures, the sense of community belonging has been weakened [35], which may have led to a lack of initiative and consciousness among residents in the process of creating AFCs. Therefore, it is vital to improve the consciousness of

XSL•FO
RenderX

age-friendliness and strengthen the sense of belonging among community residents in the process of constructing AFCs.

Establishing an Age-Friendly Atmosphere and Improving Communities' Software Facilities

Previous studies have shown that constructing AFCs mainly focuses on eight domains, based on the WHO guideline [14,36]; however, little is known about which domains are considered most important for community residents. In our study, most participants highlighted the importance of community support and health services; they wish to build more medical institutions and provide more personalized medical services in their communities, which reflects the high demand for medical services by community residents in Beijing. However, a major challenge faced by both high-income and low-income countries today is inequitable access to health care resources [37-39]. For example, in suburban or medically underprivileged communities, residents may often be unable to access desired health care services. Telemedicine in China has been improving the dissemination of high-quality health care resources from urban cities to remote suburban areas, providing services such as teleconsultations and specialist diagnoses [37]. Meanwhile, respect and social inclusion are also key components for promoting AFCIs. Many community residents identified that creating an age-friendly atmosphere was the most important step in the process of AFCs to improve the aspects of community software facilities, which is consistent with other studies. Lui et al [13] noted that a supportive environment characterized by positive relations, engagement, and inclusion is a core prerequisite for aging well. Rémillard et al [40] have explored eleven case studies of AFCs worldwide and reported that the need to shift perception, change mindsets, and promote a more positive view of aging were identified as key priorities. Respect and promoting social participation of older adults is also seen as a way to challenge ageism. In addition, the range of recreational activities and social support for community residents within their communities should be increased. Some studies found a positive association between employment, social participation, and healthy aging [41]. These observations could be due to the potential of these activities to enhance individuals' social status, promote psychological well-being, and foster a sense of dignity.

Improving the Living Environment and Strengthening Communities' Hardware Facilities

Our results indicate that many community residents wish to improve the physical infrastructure of their communities; outdoor spaces and buildings were the third most important domain of AFCs. This finding is in line with the reality of communities' infrastructure in China. Most older communities in Beijing lack elevators, which causes major inconvenience for older adults when going out. In addition, there is limited public space for residents to participate in activities and exercise. A common phenomenon in many old communities is that public spaces meant for residents' activities are occupied by parked cars. As shown in Figure 1, limited public space, scarcity of age-appropriate fitness equipment, and the absence of elevators are key physical environment factors that may hinder older adults from participating in community activities.

XSL•FO
RenderX

Figure 1. Overview of hardware facilities in a surveyed community in the Daxing District, showing limited public space, scarcity of age-appropriate fitness equipment, and no elevators.

Actively Participating in Community Activities to Enhance Community Residents' Happiness

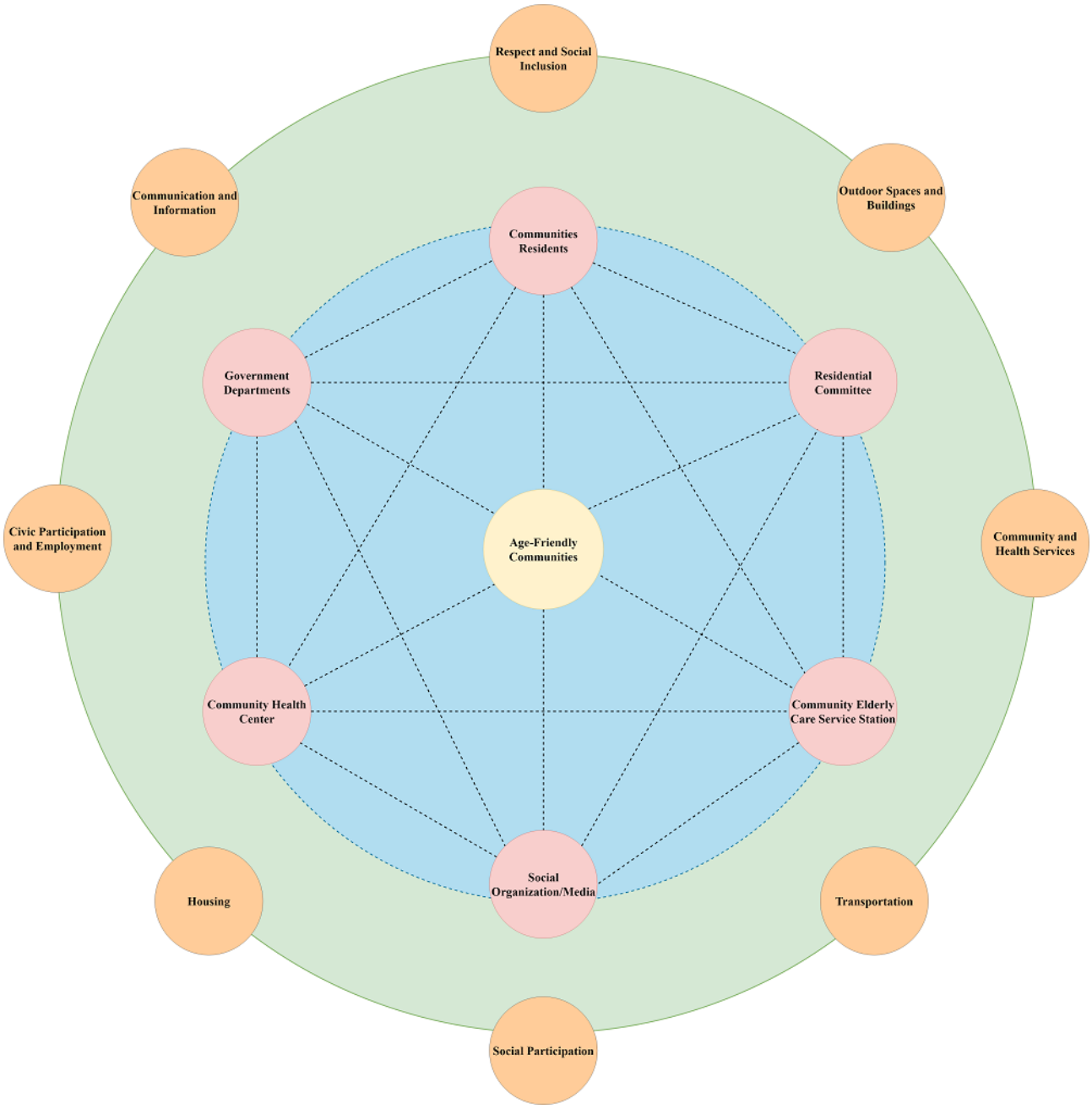
In the process of constructing AFCs, the notion of healthy aging should be considered the final goal. Older adults in good health can continue to be vital societal resources, experience greater job satisfaction and independence in their lives, and engage more actively as community residents. This study aimed to evaluate the positive living experiences that older adults obtained in their communities. Our findings show that among the five dimensions assessed, the sense of security scored the highest, while social support scored the lowest, particularly among urban older adults. Social support is a well-recognized social determinant of health, and it is obvious that obtaining social support from others is crucial throughout life, including old age [42]. Most older adults consistently prefer aging in place, which requires a high level of social support in the process of constructing AFCs [11]. Therefore, improving the degree of social support for older adults should be a key focus while constructing AFCs in the future. Further, there were significant differences in the evaluation of older adults' living experiences among different aged community residents in urban areas. Young residents rated the degree of social support, age-friendliness, sense of gain, and happiness significantly higher than older adults' self-evaluations. This disparity may stem from younger residents having greater physical capacity and healthier habits, which enables them to afford expenses related to transportation, health care, and retirement [41]. Consequently, they have more energy and capacity to actively engage in paid work, social participation, and environmental preservation, thereby fostering their sense of happiness and sense of gain. Besides, elderly residents in suburban areas have the highest scores for the sense of happiness, which suggests that older adults in suburban communities experience higher levels of happiness compared to their urban elderly residents. These findings may be attributed to both the physical environment (eg, accessible public facilities) and the social environment (eg, active engagement in volunteer activities), providing sufficient support to older adults in suburban areas [41]. Suburban seniors typically lived in the rural society for a long time. At the heart of rural society lies a deep-rooted connection to the land. Life in these communities is often dictated by the cycles of nature, with agricultural activities and environmental stewardship forming the backbone of their livelihoods. Therefore, compared with urban elderly residents, suburban seniors may engage in agricultural activities within their capacity, establish trust-based relationships with neighbors, and be more likely to participate in communities' activities.

XSL•FO
RenderX

Participation of Multiple Stakeholders in the Promotion of Constructing AFCs

The task of constructing AFCs is complex, dynamic, and multistage, which involves multiple stakeholders' cooperation. To address the challenges of an aging population and promote AFCIs, collaboration is essential between government departments, residential committees, community health centers, community elderly care service stations, social organizations, media, and community residents (Figure 2).

Figure 2. Constructing age-friendly communities as a central goal involves collaboration and participation from multiple stakeholders, including government departments, residential committees, community health centers, community elderly care service stations, social organizations, media, and community residents. Effective cooperation among these stakeholders is necessary to expand the reach of constructing age-friendly communities across all eight domains.



Government departments should focus on constructing AFCs, by formulating and improving relevant laws and regulations on AFCs, designing scientific and accurate tools for evaluating AFCs, and establishing long-term supervision and feedback mechanisms. Residential committees should first design medium- and long-term plans for promoting AFCIs based on three overarching areas [43]: (1) physical infrastructure (eg, outdoor spaces and buildings, transportation, housing); (2) social opportunities (eg, civic participation, social participation, respect, and social inclusion); and (3) supportive services (eg, communication and information, community and health services). Residential committees should conduct a comprehensive assessment of their respective communities based on AFC policies and tools developed by the government, and identify deficiencies within communities. Our findings show that community residents have urgent needs for old-age care and medical services. Therefore, promoting continuous and integrated community care that combines medical and

XSL•FO
RenderX

old-age care is necessary. Additionally, residential committees should explore smart community care services with the help of information technology, treating the community as the basic unit. Besides, it is important to build a smart service platform that connects "Internet+community home care", with community health centers and elderly care service stations, that can establish a connection between the elderly needs and community service resources. Finally, residential committees should create an atmosphere of age-friendliness within the communities, build more spaces where different aged residents can interact together, and encourage more older adults to participate in the construction of AFCs. Social organizations should adopt age-friendly practices or provide more activities catering to the needs of older adults. Social media should strengthen public awareness of AFCs, ensuring more residents understand the concept of AFCs and their benefits and actively participate in the construction of AFCs. Community residents should also actively cooperate with the work of the communities, improve their cognition of AFCs, and make suggestions for the construction of AFCs.

Limitations

The findings generated from this study should be interpreted considering several limitations. First, our survey included healthy and cognitively intact older adults, while those who were suffering from disabilities or dementia were excluded. However, understanding the living experience of this excluded group would be vital to help construct AFCs. Second, the representativeness of our study was limited, since we surveyed only a sample from 4 out of 16 districts in Beijing. Future studies should consider conducting similar studies in smaller cities across China and including more related subjects. Finally, our study followed a voluntary participation principle when conducting the questionnaire survey. Female respondents showed more willingness to participate in the survey, resulting in the gender ratio of female respondents (over 70%). Additionally, many surveyed residents never left their living communities, and they did not have a deep understanding of other communities. Future work should further explore the influence of gender, education level, and length of residence on community residents' perceptions of AFCs.

Conclusions

China is in the early stages of becoming an age-friendly society. Findings from this study show that community residents in Beijing have a low level of cognition of AFCs, and face many barriers from the perspective of hardware facilities and software facilities. Therefore, it is imperative to take measures to (1) increase age-friendly awareness and sense of belonging among community residents; (2) establish an age-friendly atmosphere and improve communities' software facilities; (3) improve the living environment and strengthen communities' hardware facilities; (4) increase the active participation of older persons in the community to enhance their happiness; and (5) and promote AFCIs through joint efforts of multiple stakeholders, including government departments, residential committees, community health centers, community elderly care service stations, social organizations, media, and community residents. Our study provides suggestions for tangible central government policy and practice initiatives and resource allocation. These outcomes have a direct and positive impact on the well‐being of older adults.

Acknowledgments

The author would like to thank all the participants, experts, and researchers who participated in this study. This study was funded by the Beijing Social Science Foundation Project (Funding Number 22SRBO08). The funding organization had no further role in the study design, data collection and analysis, interpretation of the data, writing the paper, or the decision to submit the paper for publication.

Data Availability

Data are available upon reasonable request. Transcripts will not be shared for online access to protect the anonymity of the participants. Readers who wish to gain access to the data can write to the corresponding author.

Authors' Contributions

Conceptualization: YP, ZZ

Data curation: ZZ, RZ, SZ, JZ[1],YP

Formal analysis: YZ, RW, JZ[2], YP, ZZ

Methodology: YP, ZZ

Writing – original draft: YP, ZZ

Writing – review & editing: YP, ZZ, RZ, YZ, RW, JZ[1], SZ, JZ[2], QJ, JC

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. World population prospects 2024: summary of results. United Nations. 2024. URL: https://desapublications.un.org/publications/world-population-prospects-2024-summary-results [accessed 2025-01-01]
2. Stenholm S, Westerlund H, Salo P, et al. Age-related trajectories of physical functioning in work and retirement: the role of sociodemographic factors, lifestyle and disease. J Epidemiol Community Health 2014 Jun;68(6):503-509. [doi: 10.1136/jech-2013-203555] [Medline: 24534071]
3. Hong A, Welch-Stockton J, Kim JY, Canham SL, Greer V, Sorweid M. Age-friendly community interventions for health and social outcomes: a scoping review. Int J Environ Res Public Health 2023 Jan 31;20(3):36767920. [doi: 10.3390/ijerph20032554] [Medline: 36767920]
4. Fang ML, Sixsmith J, Hamilton-Pryde A, et al. Co-creating inclusive spaces and places: towards an intergenerational and age-friendly living ecosystem. Front Public Health 2022;10(996520). [doi: 10.3389/fpubh.2022.996520] [Medline: 36685002]
5. Lawton MP, Nahemow L. Ecology and the aging process. In: The Psychology of Adult Development and Aging: American Psychological Association; 1973:619-674. [doi: 10.1037/10044-020]
6. Menec VH, Means R, Keating N, Parkhurst G, Eales J. Conceptualizing age-friendly communities. Can J Aging 2011 Sep;30(3):479-493. [doi: 10.1017/S0714980811000237] [Medline: 21745427]
7. Global age-friendly cities: a guide. World Health Organization. 2007. URL: https://www.who.int/publications/i/item/9789241547307 [accessed 2025-01-02]
8. Lai S, Zhou Y, Yuan Y. Associations between community cohesion and subjective wellbeing of the elderly in Guangzhou, China-a cross-sectional study based on the structural equation model. Int J Environ Res Public Health 2021 Jan 22;18(3):953. [doi: 10.3390/ijerph18030953] [Medline: 33499175]
9. Kim K, Buckley TD, Burnette D, Huang J, Kim S. Age-friendly communities and older adults' health in the United States. Int J Environ Res Public Health 2022 Jul 29;19(15):9292. [doi: 10.3390/ijerph19159292] [Medline: 35954648]
10. Hsu HC. Associations of city-level active aging and age friendliness with well-being among older adults aged 55 and over in Taiwan. Int J Environ Res Public Health 2020 Jun 16;17(12):4304. [doi: 10.3390/ijerph17124304] [Medline: 32560170]
11. Jeste DV, Blazer DG 2nd, Buckwalter KC, et al. Age-friendly communities initiative: public health approach to promoting successful aging. Am J Geriatr Psychiatry 2016 Dec;24(12):1158-1170. [doi: 10.1016/j.jagp.2016.07.021] [Medline: 27742528]
12. National programmes for age-friendly cities and communities: a guide. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240068698 [accessed 2025-01-01]
13. Lui CW, Everingham JA, Warburton J, Cuthill M, Bartlett H. What makes a community age-friendly: a review of international literature. Australas J Ageing 2009 Sep;28(3):116-121. [doi: 10.1111/j.1741-6612.2009.00355.x] [Medline: 19845650]
14. Dikken J, van den Hoven RFM, van Staalduinen WH, Hulsebosch-Janssen LMT, van Hoof J. How older people experience the age-friendliness of their city: development of the age-friendly cities and communities questionnaire. Int J Environ Res Public Health 2020 Sep 20;17(18):6867. [doi: 10.3390/ijerph17186867] [Medline: 32962226]
15. Measuring the age-friendliness of cities: a guide to using core indicators. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789241509695 [accessed 2025-01-01]
16. Meeks S. Age-friendly communities: introduction to the special issue. Gerontologist 2022 Jan 14;62(1):1-5. [doi: 10.1093/geront/gnab163] [Medline: 35029660]
17. Ide K, Jeong S, Tsuji T, et al. Suggesting indicators of age-friendly city: social participation and happiness, an ecological study from the JAGES. Int J Environ Res Public Health 2022 Apr 22;19(9):5096. [doi: 10.3390/ijerph19095096] [Medline: 35564490]
18. Checklist of essential features of age-friendly cities. World Health Organization. 2007. URL: https://extranet.who.int/agefriendlyworld/wp-content/uploads/2018/04/Age-Friendly-Checklist-WHOedit.pdf [accessed 2025-01-01]
19. Huang JY, Hsu HC, Hsiao YL, et al. Developing indicators of age-friendliness in Taiwanese communities through a modified Delphi method. Int J Environ Res Public Health 2022 Nov 4;19(21):14430. [doi: 10.3390/ijerph192114430] [Medline: 36361310]
20. Yu R, Wong M, Woo J. Perceptions of neighborhood environment, sense of community, and self-rated health: an age-friendly city project in Hong Kong. J Urban Health 2019 Apr;96(2):276-288. [doi: 10.1007/s11524-018-00331-3] [Medline: 30511137]
21. Wood GER, Pykett J, Daw P, et al. The role of urban environments in promoting active and healthy aging: a systematic scoping review of citizen science approaches. J Urban Health 2022 Jun;99(3):427-456. [doi: 10.1007/s11524-022-00622-w] [Medline: 35587850]
22. Kim K, Buckley T, Burnette D, Kim S, Cho S. Measurement indicators of age-friendly communities: findings from the AARP age-friendly community survey. Gerontologist 2022 Jan 14;62(1):e17-e27. [doi: 10.1093/geront/gnab055] [Medline: 33909074]
23. van Hoof J, Marston HR, Kazak JK, Buffel T. Ten questions concerning age-friendly cities and communities and the built environment. Build Environ 2021 Jul;199:107922. [doi: 10.1016/j.buildenv.2021.107922]
24. Jiravanichkul S, Pinich S, Sreshthaputra A, Jarutach T. The development of a well-being environment and age-friendly communities assessment criteria using the analytic hierarchy process: a case of Thailand. NJEDP 2024 Dec;23(3):416. [doi: 10.54028/NJ202423416]

XSL•FO
RenderX

25. Beijing municipal bureau of statistics [Chinese]. Beijing Statistical YearBook. 2024. URL: https://nj.tjj.beijing.gov.cn/nj/main/2024-tjnj/zk/indexch.htm [accessed 2025-01-01]
26. Buckner S, Mattocks C, Rimmer M, Lafortune L. An evaluation tool for age-friendly and dementia friendly communities. Work Older People 2018;22(1):48-58. [doi: 10.1108/WWOP-11-2017-0032] [Medline: 30166937]
27. Loos E, Sourbati M, Behrendt F. The role of mobility digital ecosystems for age-friendly urban public transport: a narrative literature review. Int J Environ Res Public Health 2020 Oct 14;17(20):33066528. [doi: 10.3390/ijerph17207465] [Medline: 33066528]
28. Molina-Martínez M, Marsillas S, Sánchez-Román M, Del Barrio E. Friendly residential environments and subjective well-being in older people with and without help needs. Int J Environ Res Public Health 2022 Nov 28;19(23):15832. [doi: 10.3390/ijerph192315832] [Medline: 36497900]
29. Wu D, Wang Y, Lam KF, Hesketh T. Health system reforms, violence against doctors and job satisfaction in the medical profession: a cross-sectional survey in Zhejiang Province, Eastern China. BMJ Open 2014 Dec 31;4(12):e006431. [doi: 10.1136/bmjopen-2014-006431] [Medline: 25552614]
30. Zhang T, Wang X. Association of continuity of general practitioner care with utilisation of general practitioner and specialist services in China: a mixed-method study. Healthcare (Basel) 2021 Sep 13;9(9):1206. [doi: 10.3390/healthcare9091206] [Medline: 34574980]
31. Plouffe L, Kalache A. Towards global age-friendly cities: determining urban features that promote active aging. J Urban Health 2010 Sep;87(5):733-739. [doi: 10.1007/s11524-010-9466-0] [Medline: 20549569]
32. Nemoto Y, Nonaka K, Kuraoka M, et al. Effects of intergenerational contact on social capital in community-dwelling adults aged 25-84 years: a non-randomized community-based intervention. BMC Public Health 2022 Sep 24;22(1):1815. [doi: 10.1186/s12889-022-14205-6] [Medline: 36153514]
33. Choi YJ. Understanding aging in place: home and community features, perceived age-friendliness of community, and intention toward aging in place. Gerontologist 2022 Jan 14;62(1):46-55. [doi: 10.1093/geront/gnab070] [Medline: 34043782]
34. Chaudhury H, Oswald F. Advancing understanding of person-environment interaction in later life: one step further. J Aging Stud 2019 Dec;51(100821):100821. [doi: 10.1016/j.jaging.2019.100821] [Medline: 31761094]
35. Lai S, Lu L, Zhou Z, et al. The effects of family physician-contracted service on health-related quality of life and equity in health in China. Int J Equity Health 2021 Jan 6;20(1):15. [doi: 10.1186/s12939-020-01348-4] [Medline: 33407523]
36. Menec VH, Nowicki S. Examining the relationship between communities' "age-friendliness" and life satisfaction and self-perceived health in rural Manitoba, Canada. Rural Remote Health 2014;14(2594):24437338 [FREE Full text] [Medline: 24437338]
37. Gao J, Fan C, Chen B, et al. Telemedicine Is becoming an increasingly popular way to resolve the unequal distribution of healthcare resources: evidence from China. Front Public Health 2022;10(916303):916303. [doi: 10.3389/fpubh.2022.916303] [Medline: 35874991]
38. Jiang J, Wang P. Health status in a transitional society: urban-rural disparities from a dynamic perspective in China. Popul Health Metr 2018 Dec 27;16(1):22. [doi: 10.1186/s12963-018-0179-z] [Medline: 30591053]
39. Ismail M. Regional disparities in the distribution of Sudan's health resources. East Mediterr Health J 2020 Sep 24;26(9):1105-1114. [doi: 10.26719/emhj.20.056] [Medline: 33047802]
40. Rémillard-Boilard S, Buffel T, Phillipson C. Developing age-friendly cities and communities: eleven case studies from around the world. Int J Environ Res Public Health 2020 Dec 27;18(1):133. [doi: 10.3390/ijerph18010133] [Medline: 33375466]
41. Cheng Y, Chen ZL, Wei Y, Gu N, Tang SL. Examining dynamic developmental trends: the interrelationship between age-friendly environments and healthy aging in the Chinese population-evidence from China Health and Retirement Longitudinal Study, 2011-2018. BMC Geriatr 2024 May 15;24(1):429. [doi: 10.1186/s12877-024-05053-7] [Medline: 38750429]
42. Gudlavalleti MVS, John N, Allagh K, et al. Access to health care and employment status of people with disabilities in South India, the SIDE (South India Disability Evidence) study. BMC Public Health 2014 Nov 1;14(1125):1125. [doi: 10.1186/1471-2458-14-1125] [Medline: 25361570]
43. Scharlach AE. Aging in context: individual and environmental pathways to aging-friendly communities-the 2015 Matthew A. Pollack award lecture. Gerontologist 2017 Aug 1;57(4):606-618. [doi: 10.1093/geront/gnx017] [Medline: 28460019]

Abbreviations

AFC: Age-Friendly Communities
AFCI: Age-Friendly Communities Initiatives
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Mavragani; submitted 07.09.24; peer-reviewed by JK Kazak, Z Chen; revised version received 27.01.25; accepted 28.01.25; published 21.03.25.

Please cite as:
Peng Y, Zhang Z, Zhang R, Zhang Y, Wang R, Zhang J, Zhai S, Jin Q, Zhou J, Chen J
Multidimensional Evaluation of the Process of Constructing Age-Friendly Communities Among Different Aged Community Residents in Beijing, China: Cross-Sectional Questionnaire Study
JMIR Public Health Surveill 2025;11:e66248
URL: https://publichealth.jmir.org/2025/1/e66248
doi:10.2196/66248

© Yingchun Peng, Zhiying Zhang, Ruyi Zhang, Yiyao Zhang, Runying Wang, Jiaying Zhang, Shaoqi Zhai, Qilin Jin, Jiaojiao Zhou, Jingjing Chen. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Long-Term Exposure to Fine Particulate Matter (PM2.5) Components and Precocious Puberty Among School-Aged Children: Cross-Sectional Study

Xuelian Zhou[1*], MD; Xiaochi Zhang[2,3*], BS; Guannan Bai[4], PhD; Guanping Dong[1], MD; Xinyi Li[2,3], BS; Ruimin Chen[5], MM; Shaoke Chen[6], MM; Rongxiu Zheng[7], MD; Chunlin Wang[8], MD; Haiyan Wei[9], MM; Bingyan Cao[10], MD; Yan Liang[11], MD; Hui Yao[12], MD; Zhe Su[13], MD; Mireguli Maimaiti[14], MD; Feihong Luo[15], MD; Pin Li[16], MD; Min Zhu[17], MM; Hongwei Du[18], MD; Yu Yang[19], MD; Lanwei Cui[20], MD; Jinling Wang[1], MD; Jinna Yuan[1], MD; Zhuang Liu[21], MD; Wei Wu[1], MD; Qi Zhao[2,3,22], PhD; Junfen Fu[1], PhD

[1]Department of Endocrinology, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, National Children's Regional Medical Center, 3333 Binsheng Road, Hangzhou, China

[10]Department of Endocrinology, Beijing Children's Hospital, Capital Medical University, National Medical Center for Children's Health, Beijing, China

[11]Department of Pediatrics, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[12]Department of Pediatrics, Wuhan Children's Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan, China

[13]Department of Endocrinology, Shenzhen Children's Hospital, Shenzhen, China

[14]Department of Pediatrics, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, China

[15]Department of Pediatric Endocrinology and Inherited Metabolic Diseases, Children's Hospital of Fudan University, Shanghai, China

[16]Department of Endocrinology, Children's Hospital of Shanghai Jiaotong University, Shanghai, China

[17]Department of Endocrinology, Children's Hospital of Chongqing Medical University, Chongqing, China

[18]Department of Pediatric Endocrinology, The First Bethune Hospital of Jilin University, Changchun, China

[19]Department of Endocrinology, Jiangxi Provincial Children's Hospital, Nanchang, China

[2]Department of Epidemiology, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, China

[20]Department of Pediatrics, The First Affiliated Hospital of Harbin Medical University, Harbin, China

[21]Department of Reproductive Medicine, Hospital of Jining Medical University, Jining, China

[22]Faculty of Health, Deakin University, Melbourne, Australia

[3]Shandong University Climate Change and Health Center, Shandong University, Jinan, China

[4]Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China

[5]Department of Endocrinology, Children's Hospital of Fuzhou, Fujian Province, Fuzhou, China

[6]Department of Pediatrics, Maternal and Child Health Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

[7]Department of Pediatrics, Tianjin Medical University Central Hospital, Tianjin, China

[8]Department of Pediatrics, The First Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China

[9]Department of Endocrinology, Zhengzhou Children's Hospital, Zhengzhou, China

[*]these authors contributed equally

Corresponding Author:
Junfen Fu, PhD
Department of Endocrinology, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, National Children's Regional Medical Center, 3333 Binsheng Road, Hangzhou, China

Abstract

Background: The increasing incidence of precocious puberty is a major health challenge for Chinese children, while related risk factors remain less well explored. Exposure to ambient fine particulate matter ($PM_{2.5}$) is a leading environmental hazard in China. Although certain components of $PM_{2.5}$ have been reported to be endocrine disruptors for sex hormones, population-based evidence is still lacking on the association between $PM_{2.5}$ exposure and precocious puberty in China.

Objective: Based on a cross-sectional survey covering 30 cities in 2017 to 2019, this study was designed to explore the association between long-term exposure to $PM_{2.5}$ and its 5 major components with precocious puberty in China and to check the potential modifying effects of family-related and personal factors.

XSL•FO
RenderX

Methods: We included 34,105 children aged 6 to 9 years. We collected the 5-year average concentrations of $PM_{2.5}$ and its 5 major components (sulfate, nitrate, ammonium, organic matter, and black carbon) in the area (at a spatial resolution of 0.1° × 0.1°) where each school was located. We used mixed effect logistic regression to estimate the effect sizes of the total mass of $PM_{2.5}$ and each of its components on precocious puberty, and we examined the modifying effects of family-related and personal factors using an additional interactive term. A weighted quantile sum (WQS) regression model was applied to identify the weights of each component in explaining the effect size of the total mass of $PM_{2.5}$.

Results: We found that the odds ratio (OR) for precocious puberty per IQR increase in the concentration of total $PM_{2.5}$ mass was 1.27 (95% CI 0.92-1.75) for the whole population, 2.12 (95% CI 1.27-3.55) for girls, and 0.90 (95% CI 0.62-1.30) for boys. Similarly, the effect sizes of the 5 major components were all substantial for girls but minimal for boys. Results of the WQS analysis showed that organic matter could explain the highest proportion of the effect of $PM_{2.5}$, with the weight of its contribution being 0.71. Modification effects of family income and dietary habits were only observed in certain population subgroups.

Conclusions: Long-term exposure to total $PM_{2.5}$ mass was significantly associated with precocious puberty in girls, with organic matter identified as the major effect contributor. The results add evidence on the detrimental effects of $PM_{2.5}$ on children's development and growth.

(JMIR Public Health Surveill 2025;11:e62861) doi:10.2196/62861

KEYWORDS

fine particulate matter; PM2.5; PM2.5 components; air pollution; precocious puberty; children; long-term exposure

Introduction

Precocious puberty is an endocrine and metabolic disorder that may promote the development of pathophysiological or psychosocial outcomes in childhood and later life, including stunted growth, obesity, hypertension, type 2 diabetes, ischemic heart disease, stroke, and breast cancer [1]. Over the past decades, there has been a notable increase in the incidence of precocious puberty across countries. For instance, in Denmark, the annual incidence of precocious puberty among girls younger than 8 years rose from 2.6 to 14.6 cases per 10,000 individuals between 1998 and 2017 [2]. Similar increasing trends have also been observed in a few other countries, such as South Korea, where the incidence among girls younger than 9 years rose 4.7 times between 2008 and 2014 [3], and Spain, where the incidence among girls younger than 8 years increased 16.7 times between 1997 and 2006 [4]. Among these countries, China stands out because the incidence is increasing even more rapidly and there is a larger population base [5]. This raises the urgent necessity of clarifying risk factors to facilitate efficient health promotion strategies. Certain factors or preexisting health conditions are speculated to explain the rising incidence of precocious puberty, including genetic predisposition, obesity, environmental exposures, and socioeconomic conditions [1,6-8]. However, specific risk factors involved in the pathogenesis of precocious puberty and their potential interactions still remain unclear and underexplored.

Air pollution is the second leading cause of global disease burden, and exposure to fine particulate matter ($PM_{2.5}$) could underlie most cases of air pollution–related disease, considering the complex toxicity of its components and the fact that over 90% of the world's population lives in highly polluted locations [9]. Some $PM_{2.5}$ components, such as phthalate, bisphenol analogs, and polycyclic aromatic hydrocarbons (PAHs), have been recognized as endocrine-disrupting chemicals (EDCs), affecting the normal production and signaling of sex hormones [10-12]. Some previous studies explored the association between long-term exposure to $PM_{2.5}$ and the risk of precocious puberty or related symptoms in children, but the conclusions remained inconsistent [7,13-15]. Explanations may be related to differing susceptibility across populations and the varied concentration of the essential components of $PM_{2.5}$ across countries and regions.

Ambient $PM_{2.5}$ pollution is a significant environmental hazard in China. Although air pollution has decreased dramatically in the past decade due to a series of new laws and advancements in environmental protection, the annual mean $PM_{2.5}$ concentration in 2023 was still 5 times higher than the recommended threshold of the World Health Organization (WHO) Air Quality Guidelines [9,16]. Considering the substantial number of children and the increasingly early age at onset of puberty, the adverse impact of $PM_{2.5}$ on precocious puberty, if it indeed exists, could stand out as a significant public health concern in China. However, to date, limited studies have clarified this research question.

Using data from a survey covering 30 cities of 11 provinces, our study aims to estimate the association between long-term exposure to $PM_{2.5}$ components and precocious puberty among school-aged children and to identify the components with the highest effect size in China. This will be indispensable for understanding the harmful effects of $PM_{2.5}$ and for health care promotion and environmental protection in this country.

Methods

Study Population

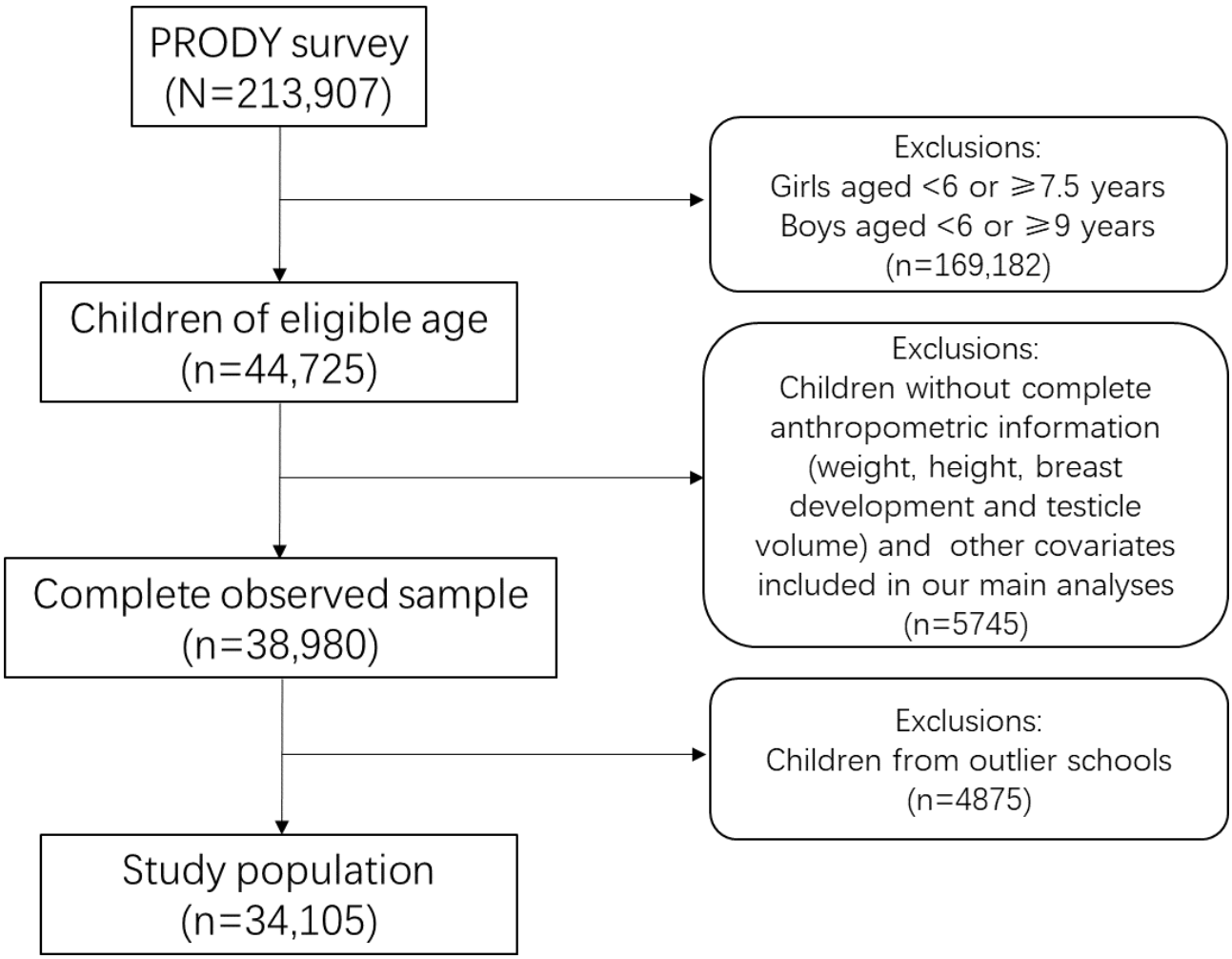
We analyzed data from eligible-aged children in the Prevalence and Risk Factors for Obesity and Diabetes in Youth (PRODY) study. A detailed profile of the PRODY data has been described previously [17]. Briefly, PRODY is a multicenter cross-sectional study conducted among children aged 3 to 18 years in China from 2017 to 2019. Participants were recruited using a stratified cluster random sampling strategy by considering their

XSL•FO
RenderX

geographic region, urbanization level, and economic level. Questionnaires and physical examinations were used to collect demographic characteristics, family information, dietary and behavioral habits, growth and development status, and other information, such as weight, height, parental educational level (low: high school or below; high: bachelor or higher), annual family income (first quartile [Q1]: <50,000 yuan [1 yuan=US $0.14]; Q2: 50,000 - 100,000 yuan; Q3: 100,000 - 200,000 yuan; Q4: >200,000 yuan), physical activity time (Q1:<90 min/week; Q2: 90 - 150 min/week; Q3:>150 min/week), frequency of consumption of dairy products, meat, fried foods, soy products, and junk foods (Q1:<1 times/month; Q2: 1 - 2 times/month; Q3: 3 - 4 times/month; Q4:>4 times/month), whole grains (Q1:<1 times/week; Q2: 1 - 2 times/week; Q3: 3 - 4 times/week; Q4:>4 times/week), and the consumption of sugary drinks, sweet fruit, and acidic fruit (yes or no). BMI (in kg/m^2) was calculated according to the growth reference of the WHO [18] and then categorized into 2 groups: low (normal weight or below) and high (overweight or obese).

This study initially included 213,907 participants across 11 provinces, autonomous regions, or municipalities in China (Xinjiang Uyghur Autonomous Region, Henan, Hubei, Guangdong, Guangxi Zhuang Autonomous Region, Zhejiang, Shanghai, Fujian, Jilin, Beijing, and Tianjin). Among the participants, 44,725 were at the eligible age for studying precocious puberty, and 38,980 of these (87.2%) accepted the related physical examination. Inclusion and exclusion criteria for the participants are shown in Figure 1. Briefly, based on previous studies and the diagnostic criteria for precocious puberty in China [5,19], we selected girls aged 6 to 7.5 years and boys aged 6 to 9 years as the study population. Children without complete anthropometric information or necessary covariates were excluded. To ensure the robustness of the analysis, schools were excluded as outliers if their prevalence rate of precocious puberty was 1.5 times greater or less than the IQR of the prevalence rates across all schools. The sample size for final analysis was 34,105. Differences in the main demographic characteristics of the study population and the original population were not statistically significant (Table S1 in Multimedia Appendix 1).

Figure 1. Flowchart of the inclusion and exclusion of study population. PRODY: Prevalence and Risk Factors for Obesity and Diabetes in Youth.



Ethical Considerations

This study protocol was approved by the Institutional Review Board of the Children's Hospital of Zhejiang University School of Medicine (approval number 2016-JRB-018). Written informed consent to participate in this study was provided by the participants' legal guardians or next of kin. We intended to use survey weights, but this possibility was denied by the data custodian due to data confidentiality concerns. All data used in this study were nonidentifiable.

XSL•FO
RenderX

Outcome Measurements

The children's breast development, testicle volume, and pubic hair were measured by a well-trained endocrinologist. Pubertal stages were assessed following the 5-stage scale described by Marshall and Tanner [20]. Precocious puberty was defined as the initiation of secondary sexual characteristics, designated as Tanner stage 2, before the age of 7.5 years for girls and 9 years for boys [19]. Previous studies have demonstrated that over 98% of Chinese children younger than 6 years have not initiated pubertal development [5]. In line with previous studies, we confined the study population to girls aged 6 to 7.5 years and boys aged 6 to 9 years to improve modeling stability.

Exposure Measurements

Data on annual average concentrations of $PM_{2.5}$ and its 5 major components, sulfate (SO_4^{2-}), nitrate (NO_3^-), ammonium (NH_4^+), organic matter (OM), and black carbon (BC), were collected from the Tracking Air Pollution in China database [21] at a spatial resolution of 0.1° × 0.1° for the entire area of China for the period 2012 to 2019. The concentration at each school location was then calculated. Tracking Air Pollution developed a 2-stage machine learning model to predict daily $PM_{2.5}$ concentration using data from $PM_{2.5}$ measurements, satellite aerosol optical depth retrievals, online community multiscale air quality simulations, meteorological reanalysis data, land use information, and population distribution. Methodological details have been described elsewhere, with the data quality also warranted as high [22,23]. The 5-year average concentrations of $PM_{2.5}$ and its 5 major components, calculated as the average over the 0 to 4 years preceding the children's physical examinations, were assigned to each school address as a surrogate for personal exposure.

Statistical Methods

Single Exposure Model

Mixed-effects logistic regression was applied to quantify the association between $PM_{2.5}$ concentration and each of its 5 components with the binary status of precocious puberty, with the "province" indicator entered as the random intercept. Considering that certain dietary variables were correlated (Figure S1 in Multimedia Appendix 1), we used a least absolute shrinkage and selection operator regression model with the minimum mean squared error to select important dietary confounders related to precocious puberty (Table S2 in Multimedia Appendix 1). The main models were adjusted for gender, age, parental educational level, annual family income, physical activity time, BMI, and dietary habits, including frequency of consumption of dairy products, meat, fried food, and whole grains, as well as the intake of sugary drinks as a snack, fruit as a snack, sweet fruit, and acidic fruit. The results were described as odds ratios (ORs) for precocious puberty with the 95% CI per IQR increase in the concentration of each air pollutant. In this study, $PM_{2.5}$ and its 5 major components were not entered into a single model at the same time to avoid potential multicollinearity problems. The potential interaction effects were examined individually by adding an additional interaction term between $PM_{2.5}$ and each of these variables, including parental educational level, annual family income, physical activity time, BMI, and dietary habits.

Joint Exposure Model

Given the high correlations and uncertain relationships among the various components of $PM_{2.5}$, we used a weighted quantile sum (WQS) regression model to assess the joint exposure effects. The WQS model was used to construct a cumulative linear index by grouping $PM_{2.5}$ components into quartiles, with the corresponding weight of each component indicating its contribution to the overall index. WQS regression assumes that all exposures included in the index are associated with the outcome in the same direction. Therefore, we first evaluated the direction of the association between each pollutant and precocious puberty through the single exposure model. As suggested previously, the training and validation datasets were split 40% and 60%, and the bootstrap was set to 1000 times [24].

Sensitivity Analyses

A series of sensitivity analyses were conducted to examine the robustness of our main findings. First, the main models were expanded by additionally adjusting for other dietary variables selected from the literature that failed the selection procedure. Second, we modified the personal exposure from 0 - 4 years to 1 - 4 years and 0 - 2 years to explore whether the results might differ with exposure duration. Third, we further adjusted for gaseous pollutants, such as SO_2, CO_2, and O_3. The quantile-based g-computation (QGC) model was used for sensitivity analysis of the joint exposure model, as it allows for the evaluation of combined effects of exposure in different directions [25]. Meta regression was used to test the significance of differences in the results of the main analyses and sensitivity analyses.

R (version 4.3.0; R Foundation for Statistical Computing) was used for analyses. *P* values <.05 were considered statistically significant.

Results

Table 1 presents the demographic characteristics of the study participants. A total of 34,105 children (n=23,683, 69.44% boys) with an average age of 7.25 (SD 0.82) years were enrolled in this study. Among them, 62.39% (21,279/34,105) had an annual family income exceeding 100,000 yuan (US $14,327), 63.25% (21,570/34,105) had parents with a college degree or higher, and 26.01% (8872/34,105) were categorized as overweight or obese. A total of 401 of the 34,105 children (1.18%) were diagnosed with precocious puberty, of which 214 were boys (0.9%) and 187 were girls (1.79%). The average individual exposure over a 5-year period was 47.93 (SD 16.52) μg/m^3 for $PM_{2.5}$, 2.31 (SD 0.59) μg/m³ for BC, 6.34 (SD 2.25) μg/m³ for NH_4^+, 11.74 (SD 3.75) μg/m³ for OM, 9.52 (SD 3.95) μg/m³ for NO_3^- and 8.74 (SD 2.11) μg/m³ for SO_4^{2-}, respectively (Table S3 in Multimedia Appendix 1). A map of $PM_{2.5}$ concentration and the study population during the study period is shown in Figure S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Basic characteristics of participants enrolled during 2017 and 2019 in 30 cities in China.

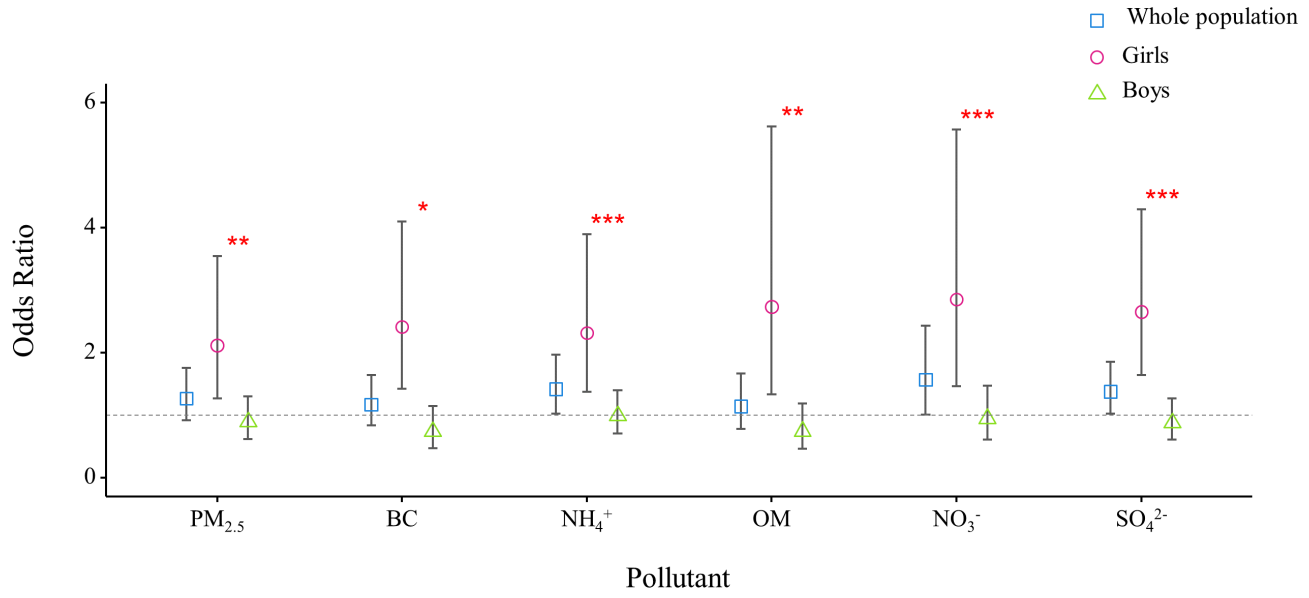
Variables		Precocious puberty (n=401)	Nonprecocious puberty (n=33,704)
$PM_{2.5}$ concentration (μg/m^3), mean (SD)		47.92 (16.79)	47.93 (16.52)
Age (years), mean (SD)		7.56 (0.85)	7.24 (0.82)
Gender, n (%)			
	Female	187 (46.63)	10,235 (30.37)
	Male	214 (53.37)	23,469 (69.63)
Parental educational level, n (%)			
	High school or below	180 (44.89)	12,355 (36.66)
	Bachelor or higher	221 (55.11)	21,349 (63.34)
Annual family income (yuan; 1 yuan=US $0.14), n (%)			
	<50,000	67 (16.71)	5530 (16.41)
	50,000 - 10,0000	98 (24.44)	7131 (21.16)
	100,000 - 200,000	120 (29.93)	10,902 (32.35)
	>200,000	116 (28.93)	10,141 (30.09)
Physical activity (minutes/week) , n (%)			
	<90	154 (38.4)	13,700 (40.65)
	90 - 150	168 (41.9)	12,634 (37.49)
	>150	79 (19.7)	7370 (21.87)
BMI, n (%)			
	Underweight or normal	211 (52.62)	25,022 (74.24)
	Overweight or obese	190 (47.38)	8682 (25.76)
Consumed sugary drinks, n (%)			
	No	283 (70.57)	25,630 (76.04)
	Yes	118 (29.43)	8074 (23.96)
Consumed sweet fruit, n (%)			
	No	71 (17.71)	4709 (13.97)
	Yes	330 (82.29)	28,995 (86.03)
Consumed acidic fruit, n (%)			
	No	197 (49.13)	14,884 (44.16)
	Yes	204 (50.87)	18,820 (55.84)
Junk food frequency (times/month), n (%)			
	<1	354 (88.28)	30,226 (89.68)
	1 - 2	27 (6.73)	2462 (7.30)
	3 - 4	15 (3.74)	750 (2.23)
	>4	5 (1.25)	266 (0.79)
Whole grains frequency (times/week), n (%)			
	<1	63 (15.71)	5196 (15.42)
	1 - 2	248 (61.85)	20,788 (61.68)
	3 - 4	71 (17.71)	5636 (16.72)
	>4	19 (4.74)	2084 (6.18)
Dairy product frequency (times/day), mean (SD)		0.66 (0.54)	0.72 (0.57)
Meat frequency (times/day), mean (SD)		2.60 (1.46)	2.73 (1.46)
Fried food frequency (times/day), mean (SD)		0.18 (0.55)	0.14 (0.46)

XSL•FO
RenderX

Variables	Precocious puberty (n=401)	Nonprecocious puberty (n=33,704)
Soy product frequency (times/day), mean (SD)	0.93 (0.88)	0.98 (0.89)

The analysis of the single exposure model indicated that the ORs for precocious puberty for the whole study population were 1.27 (95% CI 0.92-1.75) for total $PM_{2.5}$ mass, 1.42 (95% CI 1.03 - 1.97) for NH_4^+, 1.57 (95% CI 1.01 - 2.43) for NO_3^-, and 1.38 (95% CI 1.03 - 1.86) for SO_4^{2-} per IQR increase in their 5-year average concentrations (Figure 2). A clear gender variation in the effect size of $PM_{2.5}$ exposure was observed: per IQR increase in $PM_{2.5}$ mass, the ORs were 2.12 (95% CI 1.27 - 3.55) for girls and 0.90 (95% CI 0.62 - 1.30) for boys (P=.01). In addition, the effect sizes of all 5 $PM_{2.5}$ components were also only significant for girls.

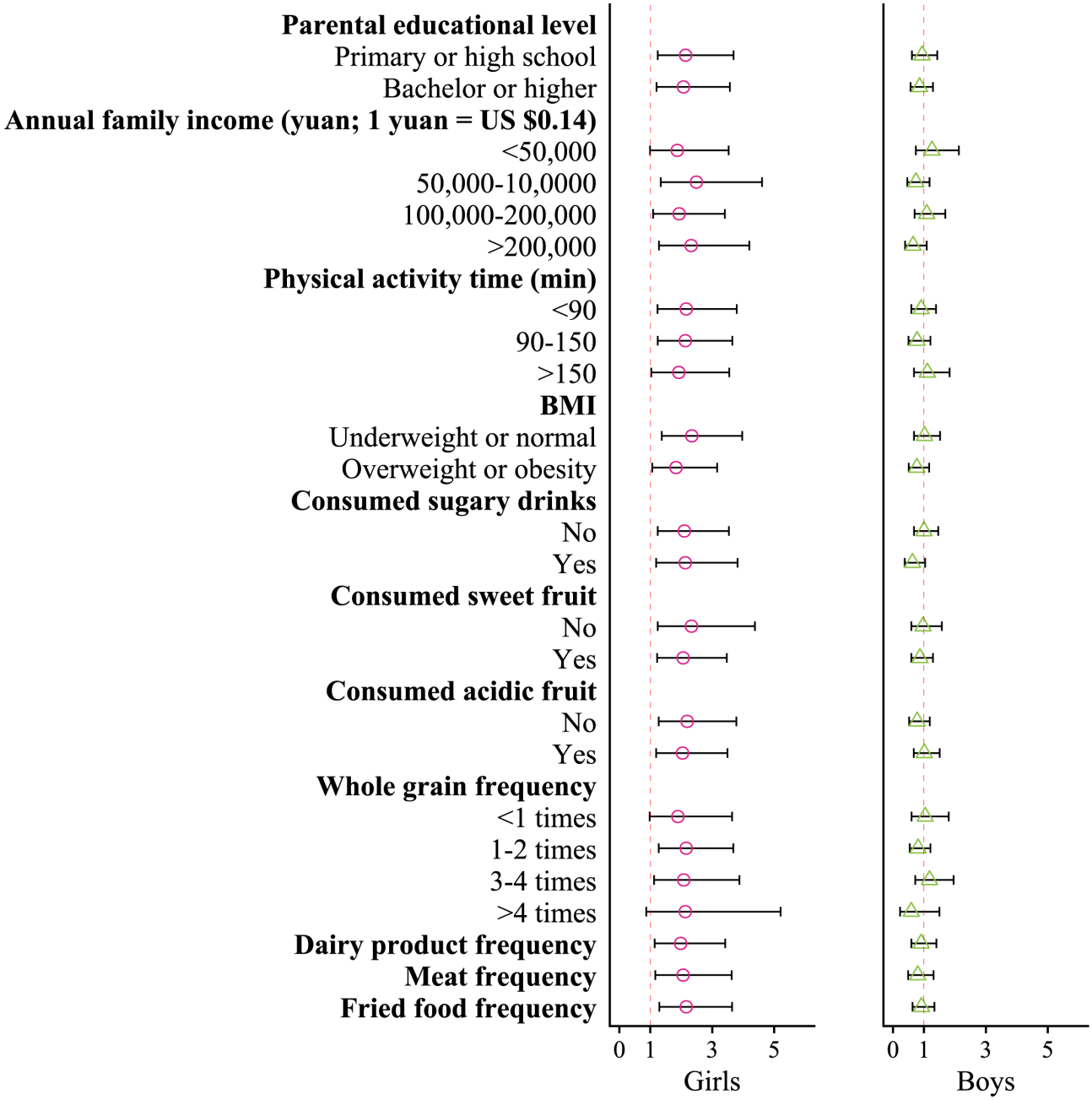
Figure 2. Odds ratio for precocious puberty per IQR increase in $PM_{2.5}$ mass and its 5 major components. $PM_{2.5}$:fine particulate matter; BC: black carbon; OM: organic matter; NO_3^-: nitrate; NH_4^+: ammonium; SO_4^{2-}: sulfate. Asterisks indicate the significance of differences between boys and girls: $^*P$<.05, $^{**}P$<.01, $^{***}P$<.001.



The effect size of $PM_{2.5}$ was nonsignificantly higher in groups with certain demographic characteristics, family information, and behavioral habits (Figure 3). For example, the OR for precocious puberty was 2.49 (95% CI 1.34 - 4.61) for girls with a family income at the Q2 level and 1.87 (95% CI 0.99 - 3.53) for those with family income at the Q1 level. Similarly, the ORs were 2.34 (95% CI 1.37 - 3.97) and 1.83 (95% CI 1.06 - 3.16) for groups with low BMI and high BMI, respectively. Certain modifying effects were observed in boys, such as family income (<50,000 yuan group vs >200,000 yuan group; P=.03) and sugary drink consumption (no vs yes; P=.03).

XSL•FO
RenderX

Figure 3. The modification effect of individual characteristics and dietary habits on the association between $PM_{2.5}$ exposure and precocious puberty.

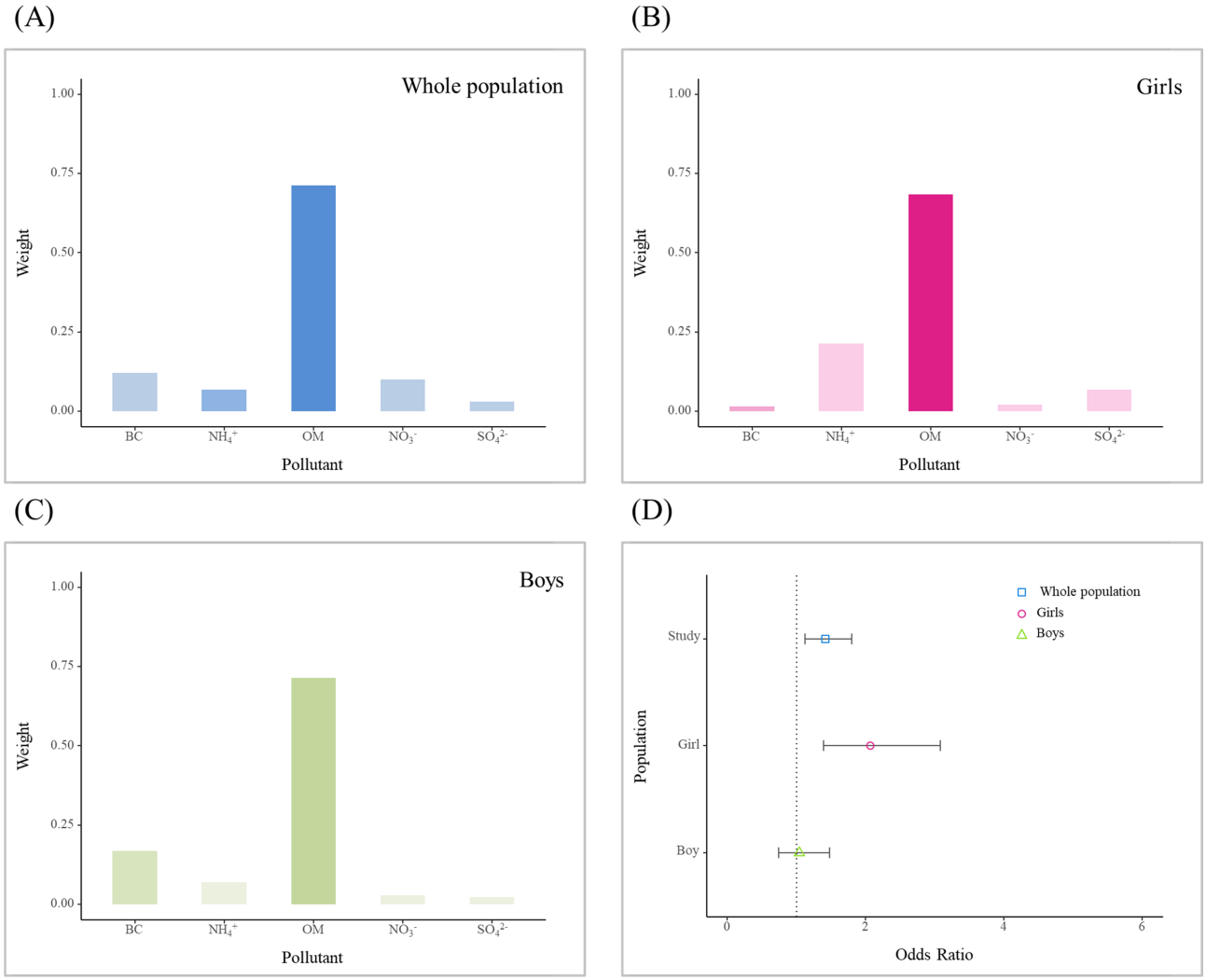


As shown in Figure 4, the WQS index was significantly associated with precocious puberty (OR 1.42, 95% CI 1.12 - 1.80). The major contributors were OM, BC, and NO_3^-, with their weights in the overall effect size being 0.71, 0.12, and 0.10, respectively. After stratification by gender, a significant association was observed only in girls.

The strength of the association was stable after adjusting for additional covariables and gaseous pollutants or changing the exposure period in both the single-air-pollutant models and WQS models. The association between $PM_{2.5}$ components and precocious puberty remained unchanged in the QGC model, with OM continuing to be the major positive contributor to the overall effect of total $PM_{2.5}$ mass (Tables S4-5 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 4. Weights of the contributions of the 5 major components of fine particulate matter ($PM_{2.5}$) to its total effect size. Mean weights for (**A**) the whole study population; (**B**) girls; (**C**) boys. (**D**) Odds ratios for precocious puberty per IQR increase in the weighted quantile sum regression (WQS) index. WQS models were adjusted for gender; age; maximum level of parental education; annual family income; weekly physical activity time; BMI; consumption frequency of dairy products, meat, fried foods, and coarse food grain; intake of sugary drinks, sweet fruit, and acidic fruit; and province. BC: black carbon; OM: organic matter; NO_3^-: nitrate; NH_4^+: ammonium; SO_4^{2-}: sulfate.



Discussion

Principal Results

This study explored the association between exposure to $PM_{2.5}$ and its major components with precocious puberty based on a large survey of school-aged children in China. We observed that $PM_{2.5}$ mass was significantly associated with precocious puberty only in girls. OM was detected as the major effective component when the 5 major components of $PM_{2.5}$ were jointly exposed. Modification effects of certain factors at the individual or family levels were observed (eg, family income and sugary drink consumption), especially for boys.

Comparison With Prior Work

Recent evidence suggests that there is significant variation in the prevalence of precocious puberty across regions and study periods; the prevalence ranges from <0.05% to 2.53% in Denmark, Korea, and China [15,26,27]. To our knowledge, our PRODY study is to date the only large-scale survey on precocious puberty in China. We found that the prevalence in the study areas was 0.9% in boys and 1.79% in girls. Previously, only one study (the predecessor of PRODY) investigated precocious puberty in 5 Chinese provinces and reported a prevalence of 0.43% [26]. The differing prevalence across studies may be partly explained by differences in diagnostic criteria, population characteristics, the years of investigation, and certain limitations in sampling.

Previously, some studies have explored the effect of long-term $PM_{2.5}$ exposure on pubertal development or related sexual hormones and obtained inconsistent results. A cohort study in Poland revealed that children exposed to an average annual concentration of $PM_{2.5}$ exceeding 25 $\mu g/m^3$ were at high risk of menarche occurring before the age of 11 years [14]. Similarly, Jung et al [28] found that each 1 $\mu g/m^3$ increase in the 3-year average of annual mean PM_{10} concentration was associated with age acceleration at menarche by 0.031 years in 639 girls aged 13 to 17 years in South Korea. Additionally, a recent large-scale (n=1,205,784) nationwide retrospective cohort study reported that 4-year exposures to $PM_{2.5}$ and PM_{10} were

XSL•FO
RenderX

associated with precocious puberty in girls in South Korea [15]. By contrast, in Hong Kong's cohort of children born in 1997, Huang et al [7] found a nonsignificant association between childhood PM_{10} exposure and precocious puberty. A German study reported a weak association between $PM_{2.5}$ exposure and decreased estradiol level in girls in Munich [13].

There may be several explanations for the inconsistent findings across studies. For example, in comparison to the Hong Kong study, which used the age of 11 years as the diagnostic threshold for pubertal stage, our study applied a younger diagnostic age (<7.5 years for girls and <9 years for boys), following the latest official recommendation in mainland China. This lower threshold may more effectively capture the early onset of pubertal development in children and the related effect size of $PM_{2.5}$. In addition, the inconsistent results of these studies may also be attributed to the differences in study area, population characteristics, outcome definitions, and other potential confounders like socioeconomic status, as well as exposure period and levels. This study provides supporting evidence for the adverse effect of childhood $PM_{2.5}$ exposure on precocious puberty among a nationwide sample in China.

The potential mechanisms underpinning the adverse effect of $PM_{2.5}$ on precocious puberty have been rarely discussed. One suggested hypothesis is the endocrine disruption caused by $PM_{2.5}$. Certain EDCs, such as phthalates, bisphenol analogs, and PAHs, have been detected as $PM_{2.5}$ components in several major Chinese cities [29-31]. These chemicals may target the estrogen receptor or kisspeptin in the hypothalamic nuclei and exhibit estrogenic or antiandrogenic activity, thereby promoting hypothalamic maturation and consequently causing precocious puberty [8,32-34]. In addition, some heavy metals found in $PM_{2.5}$, especially lead and cadmium, have also been proven to disrupt estrogen levels in female individuals and influence pubertal development [35,36].

The health effects of $PM_{2.5}$ depend on its chemical components. An increasing number of epidemiological studies have paid attention to the association between $PM_{2.5}$ chemical components and a range of health outcomes, such as pregnancy disorders, lung function, mortality, and body weight [37-40]. To the best of our knowledge, our study is the first to report associations of $PM_{2.5}$ components with precocious puberty. We found that OM is the most important component of $PM_{2.5}$ associated with precocious puberty in China. OM is produced by the incomplete combustion of solid fuels, which have higher toxicity than noncombustion aerosols [41]. In China, OM is the primary component of $PM_{2.5}$ in several major cities [42], posing a significant health risk to the population. Although the differential impact of components of $PM_{2.5}$ on precocious puberty has not yet been adequately examined, some underlying mechanisms may explain the observed effects. As previously mentioned, the chemical components in $PM_{2.5}$ that exert endocrine-disrupting effects are predominantly OM, such as PAHs. In addition, OM has the potential to induce mutagenic, inflammatory, and lipid metabolic disturbances in vivo [43-45]. However, without data at the cellular and molecular levels, it is beyond the scope of this study to speculate further about the potential pathways mediating the impact of OM on pubertal development.

Previous studies have suggested that the effects of $PM_{2.5}$ on children's pubertal development may differ by gender [13,15]. Consistent with this, our study observed a more significant association between $PM_{2.5}$ and precocious puberty in girls than in boys. It has been speculated that gender differences are due to the varying effects of PM exposure on different hormone levels in vivo [46], as well as the fact that prepubertal girls are particularly sensitive to low levels of exogenous estrogens [47]. In addition, gender differences may be related to the criteria for measuring secondary sexual characteristics, which depend on physical examinations of breast development and pubic hair in girls and testicle volume and pubic hair in boys, rather than on sex hormones and gonad B-ultrasounds.

Significant modifying effects of family income and sugary drink consumption were observed. Previously, several studies explored potential modifiers of the adverse effect of long-term $PM_{2.5}$ exposure, although few have focused on pubertal development. For example, a large cohort study in Canada found that the impact of $PM_{2.5}$ on mortality was greater for low-income earners than for high-income earners [48]. A study in the United States indicated that the cardiovascular risks associated with $PM_{2.5}$ exposure were higher among low-income groups compared to high-income groups [49]. The inequalities in $PM_{2.5}$ exposure between high-income populations and low-income populations may partly explain the modifying effect of family income.

Pubertal timing is sensitive to nutritional regulation; for example, girls with the highest levels of dietary isoflavone intake may experience early onset of breast development [50]. Some animal studies have found that fructose diet–fed rats had lower plasma testosterone levels, which can lead to a high sensitivity of the pituitary-gonadal axis in Leydig cells [51]. It can be speculated that dietary factors may potentially involve pathways that overlap with $PM_{2.5}$ and promote or hinder the effects of $PM_{2.5}$ on precocious puberty. However, there is still a lack of evidence to explain why significant modifying effects of family income and sugary drink consumption were observed more prominently in boys than in girls. Due to the absence of detailed information, further exploration is not feasible. Future studies may be able to elucidate the underlying mechanisms.

Strengths and Limitations

This study has several strengths that can be mentioned. First, our findings came from a survey covering 30 cities in 11 provinces, improving the generalizability when compared to studies with a smaller geographical scale, such as a single city. Second, the PRODY survey incorporated measurements of breast development and testicle volume conducted by medical professionals, resulting in a more accurate diagnosis of precocious puberty compared to studies using self-reported or parent-reported data. Third, we analyzed the effects of joint exposure to $PM_{2.5}$ components on precocious puberty using WQS regression, which can identify major contributors and therefore provide targeted recommendations for air quality–related health improvement.

XSL•FO
RenderX

We acknowledge some limitations of our analysis. In this study, air pollutant concentrations at the school level may not fully reflect all sources of air pollution, such as indoor air pollution within households, potentially leading to exposure misclassification. Because of the large sample size in this study and the limitations imposed by survey costs, we were unable to measure children's air pollution exposure at a finer scale. Future studies are strongly encouraged to consider a more holistic approach by incorporating direct measurements of indoor air quality, possibly using portable air monitoring devices. Due to ethical considerations, only children willing to undergo physical examinations were included in the study, which might have introduced a certain sampling bias. Without access to data on appropriate survey weights, we were unable to evaluate the accurate generalization of our findings to the whole target population of the country, which warrants further exploration when data are available in the future. Due to the cross-sectional study design, the inference of causality was limited. Furthermore, owing to constraints in data availability, our study did not account for the adverse effects of some other toxic components (ie, heavy metals) or confounding variables such as parental history of precocious puberty, sleep patterns, and psychological status. We suggest that future research should include these additional variables to better understand their potential impacts on the results.

Conclusion

Our study suggests that OM may be a major contributor to the association between long-term exposure to $PM_{2.5}$ and precocious puberty. Our findings further endorse the call for reducing fossil fuel emissions, concurrently advocating the development of specific and effective strategies to protect children's developmental health.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2021YFC2701900 and 2016YFC1305300), Key R&D Program of Zhejiang (2023C03047), National Natural Science Foundation of China (82370863, 81570759, and 81270938), Natural Science Foundation of Shandong Province in China (ZR2021QH318), Shandong Excellent Young Scientists Fund Program (Overseas) (2022HWYQ-055), Taishan Young Scholar Program of Shandong Province (tsqn202306063), and Key R&D Projects of Jining Science and Technology Bureau (2022YXNS086).

Data Availability

The data sets generated and analyzed during this study are available from the corresponding authors upon reasonable request.

Authors' Contributions

X Zhou contributed to data curation, conceptualization, and validation. X Zhang contributed to methodology, formal analysis, visualization, and writing (original draft). GB and WW contributed to conceptualization. KH, GD, and XL contributed to data curation. RC, SC, RZ, CW, HW, BC, YL, HY, ZS, MM, FL, PL, MZ, HD, YY, LC, JW, and JY contributed to investigation. ZL contributed to funding acquisition and conceptualization. QZ and JF contributed to conceptualization, supervision, and writing (review and editing). All authors have read and approved the final version of the manuscript. All authors have final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary materials showing the distribution of air pollutant concentrations and the study population during the study period, a comparison of characteristics between the original recruited population and the study population, and the results of the sensitivity analysis.

[DOCX File, 615 KB - publichealth_v11i1e62861_app1.docx]

References

1. Latronico AC, Brito VN, Carel JC. Causes, diagnosis, and treatment of central precocious puberty. Lancet Diabetes Endocrinol 2016 Mar;4(3):265-274. [doi: 10.1016/S2213-8587(15)00380-0] [Medline: 26852255]
2. Bräuner EV, Busch AS, Eckert-Lind C, Koch T, Hickey M, Juul A. Trends in the incidence of central precocious puberty and normal variant puberty among children in Denmark, 1998 to 2017. JAMA Netw Open 2020 Oct 1;3(10):e2015665. [doi: 10.1001/jamanetworkopen.2020.15665] [Medline: 33044548]
3. Kim YJ, Kwon A, Jung MK, et al. Incidence and prevalence of central precocious puberty in Korea: an epidemiologic study based on a national database. J Pediatr 2019 May;208:221-228. [doi: 10.1016/j.jpeds.2018.12.022] [Medline: 30857777]
4. Soriano-Guillén L, Corripio R, Labarta JI, et al. Central precocious puberty in children living in Spain: incidence, prevalence, and influence of adoption and immigration. J Clin Endocrinol Metab 2010 Sep;95(9):4305-4313. [doi: 10.1210/jc.2010-1025] [Medline: 20554707]

XSL•FO
RenderX

5. Liang X, Huang K, Dong G, et al. Current pubertal development in Chinese children and the impact of overnutrition, lifestyle, and perinatal factors. J Clin Endocrinol Metab 2023 Aug 18;108(9):2282-2289. [doi: 10.1210/clinem/dgad102] [Medline: 36881937]
6. Shi L, Jiang Z, Zhang L. Childhood obesity and central precocious puberty. Front Endocrinol (Lausanne) 2022;13:1056871. [doi: 10.3389/fendo.2022.1056871] [Medline: 36465655]
7. Huang JV, Leung GM, Schooling CM. The association of air pollution with pubertal development: evidence from Hong Kong's "children of 1997" birth cohort. Am J Epidemiol 2017 May 15;185(10):914-923. [doi: 10.1093/aje/kww200] [Medline: 28444106]
8. Brito VN, Canton APM, Seraphim CE, et al. The congenital and acquired mechanisms implicated in the etiology of central precocious puberty. Endocr Rev 2023 Mar 4;44(2):193-221. [doi: 10.1210/endrev/bnac020] [Medline: 35930274]
9. WHO global air quality guidelines: particulate matter (PM25 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240034228 [accessed 2025-01-14]
10. Fang B, Bravo MA, Wang H, et al. Polycyclic aromatic hydrocarbons are associated with later puberty in girls: a longitudinal study. Sci Total Environ 2022 Nov 10;846:157497. [doi: 10.1016/j.scitotenv.2022.157497] [Medline: 35868395]
11. Zhou Q, Chen J, Zhang J, et al. Toxicity and endocrine-disrupting potential of PM2.5: association with particulate polycyclic aromatic hydrocarbons, phthalate esters, and heavy metals. Environ Pollut 2022 Jan 1;292(Pt A):118349. [doi: 10.1016/j.envpol.2021.118349] [Medline: 34653588]
12. Gore AC, Chappell VA, Fenton SE, et al. EDC-2: the Endocrine Society's second scientific statement on endocrine-disrupting chemicals. Endocr Rev 2015 Dec;36(6):E1-E150. [doi: 10.1210/er.2015-1010] [Medline: 26544531]
13. Zhao T, Triebner K, Markevych I, et al. Outdoor air pollution and hormone-assessed pubertal development in children: results from the GINIplus and LISA birth cohorts. Environ Int 2021 Jul;152:106476. [doi: 10.1016/j.envint.2021.106476] [Medline: 33714142]
14. Wronka I, Kliś K. Effect of air pollution on age at menarche in Polish females, born 1993-1998. Sci Rep 2022 Mar 21;12(1):4820. [doi: 10.1038/s41598-022-08577-3] [Medline: 35315430]
15. Oh J, Choi JE, Lee R, et al. Long-term exposure to air pollution and precocious puberty in South Korea. Environ Res 2024 Jul 1;252(Pt 2):118916. [doi: 10.1016/j.envres.2024.118916] [Medline: 38614201]
16. Xiao Q, Geng G, Liu S, Liu J, Meng X, Zhang Q. Spatiotemporal continuous estimates of daily 1 km PM 2.5 from 2000 to present under the Tracking Air Pollution in China (TAP) framework. Atmos Chem Phys 2022;22(19):13229-13242. [doi: 10.5194/acp-22-13229-2022]
17. Huang K, Si S, Chen R, et al. Preterm birth and birth weight and the risk of type 1 diabetes in Chinese children. Front Endocrinol (Lausanne) 2021;12:603277. [doi: 10.3389/fendo.2021.603277] [Medline: 33935963]
18. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ 2007 Sep;85(9):660-667. [doi: 10.2471/blt.07.043497] [Medline: 18026621]
19. Subspecialty Group of Endocrinologic, Hereditary and Metabolic Diseases, the Society of Pediatrics, Chinese Medical Association, Editorial Board, Chinese Journal of Pediatrics. Expert consensus on the diagnosis and treatment of central precocious puberty(2022). Zhonghua Er Ke Za Zhi 2023 Jan 2;61(1):16-22. [doi: 10.3760/cma.j.cn112140-20220802-00693] [Medline: 36594116]
20. Emmanuel M, Bokor BR. Tanner stages. In: StatPearls: StatPearls Publishing; 2024. [Medline: 29262142]
21. TAP: tracking air pollution in China. TAPData. URL: http://tapdata.org.cn [accessed 2025-01-08]
22. Liu S, Geng G, Xiao Q, et al. Tracking daily concentrations of PM 2.5 chemical composition in China since 2000. Environ Sci Technol 2022 Nov 15;56(22):16517-16527. [doi: 10.1021/acs.est.2c06510] [Medline: 36318737]
23. Geng G, Xiao Q, Liu S, et al. Tracking air pollution in China: near real-time PM2.5 retrievals from multisource data fusion. Environ Sci Technol 2021 Sep 7;55(17):12106-12115. [doi: 10.1021/acs.est.1c01863] [Medline: 34407614]
24. Carrico C, Gennings C, Wheeler DC, Factor-Litvak P. Characterization of weighted quantile sum regression for highly correlated data in a risk analysis setting. J Agric Biol Environ Stat 2015 Mar;20(1):100-120. [doi: 10.1007/s13253-014-0180-3] [Medline: 30505142]
25. Keil AP, Buckley JP, O'Brien KM, Ferguson KK, Zhao S, White AJ. A quantile-based g-computation approach to addressing the effects of exposure mixtures. Environ Health Perspect 2020 Apr;128(4):47004. [doi: 10.1289/EHP5838] [Medline: 32255670]
26. Zhu M, Fu J, Liang L, et al. Epidemiologic study on current pubertal development in Chinese school-aged children. Zhejiang Da Xue Xue Bao Yi Xue Ban 2013 Jul;42(4):396-402. [doi: 10.3785/j.issn.1008-9292.2013.04.005] [Medline: 24022927]
27. Teilmann G, Pedersen CB, Jensen TK, Skakkebaek NE, Juul A. Prevalence and incidence of precocious pubertal development in Denmark: an epidemiologic study based on national registries. Pediatrics 2005 Dec;116(6):1323-1328. [doi: 10.1542/peds.2005-0012] [Medline: 16322154]
28. Jung EM, Kim HS, Park H, Ye S, Lee D, Ha EH. Does exposure to PM10 decrease age at menarche? Environ Int 2018 Aug;117:16-21. [doi: 10.1016/j.envint.2018.04.020] [Medline: 29704753]

XSL•FO
RenderX

29. Duan L, Yu H, Wang Q, et al. PM2.5-bound polycyclic aromatic hydrocarbons of a megacity in eastern China: Source apportionment and cancer risk assessment. Sci Total Environ 2023 Apr 15;869:161792. [doi: 10.1016/j.scitotenv.2023.161792] [Medline: 36702280]
30. Zhang X, Wang Q, Qiu T, et al. PM2.5 bound phthalates in four metropolitan cities of China: concentration, seasonal pattern and health risk via inhalation. Sci Total Environ 2019 Dec 15;696:133982. [doi: 10.1016/j.scitotenv.2019.133982] [Medline: 31470327]
31. Munyaneza J, Qaraah FA, Jia Q, Cheng H, Zhen H, Xiu G. Seasonal trends, profiles, and exposure risk of pm2.5-bound bisphenol analogs in ambient outdoor air: a study in Shanghai, China. Aerosol Air Qual Res 2022;22(9):210324. [doi: 10.4209/aaqr.210324]
32. Lopez-Rodriguez D, Franssen D, Heger S, Parent AS. Endocrine-disrupting chemicals and their effects on puberty. Best Pract Res Clin Endocrinol Metab 2021 Sep;35(5):101579. [doi: 10.1016/j.beem.2021.101579] [Medline: 34563408]
33. Zhang Y, Dong S, Wang H, Tao S, Kiyama R. Biological impact of environmental polycyclic aromatic hydrocarbons (ePAHs) as endocrine disruptors. Environ Pollut 2016 Jun;213:809-824. [doi: 10.1016/j.envpol.2016.03.050] [Medline: 27038213]
34. Lopez-Rodriguez D, Franssen D, Bakker J, Lomniczi A, Parent AS. Cellular and molecular features of EDC exposure: consequences for the GnRH network. Nat Rev Endocrinol 2021 Feb;17(2):83-96. [doi: 10.1038/s41574-020-00436-3] [Medline: 33288917]
35. Gerald AC, Ganapathy S, Zhu J, Wei Y. Exposure to endocrine-disrupting metals and serum estrogen levels among US women. Reprod Toxicol 2023 Jun;118:108392. [doi: 10.1016/j.reprotox.2023.108392] [Medline: 37150308]
36. Li F, Yan J, Wei Y, et al. PM2.5-bound heavy metals from the major cities in China: spatiotemporal distribution, fuzzy exposure assessment and health risk management. J Clean Prod 2021 Mar;286:124967. [doi: 10.1016/j.jclepro.2020.124967]
37. Shen Y, Yu G, Liu C, et al. Prenatal exposure to PM2.5 and its specific components and risk of hypertensive disorders in pregnancy: a nationwide cohort study in China. Environ Sci Technol 2022 Aug 16;56(16):11473-11481. [doi: 10.1021/acs.est.2c01103] [Medline: 35914180]
38. Zhang Y, Yin Z, Zhou P, et al. Early-life exposure to PM2.5 constituents and childhood asthma and wheezing: findings from China, Children, Homes, Health study. Environ Int 2022 Jul;165:107297. [doi: 10.1016/j.envint.2022.107297] [Medline: 35709580]
39. Guo Q, Zhang K, Wang B, et al. Chemical constituents of ambient fine particulate matter and obesity among school-aged children: a representative national study in China. Sci Total Environ 2022 Nov;849:157742. [doi: 10.1016/j.scitotenv.2022.157742] [Medline: 35917963]
40. Raaschou-Nielsen O, Antonsen S, Agerbo E, et al. PM2.5 air pollution components and mortality in Denmark. Environ Int 2023 Jan;171:107685. [doi: 10.1016/j.envint.2022.107685] [Medline: 36502699]
41. Park M, Joo HS, Lee K, et al. Differential toxicities of fine particulate matters from various sources. Sci Rep 2018 Nov 19;8(1):17007. [doi: 10.1038/s41598-018-35398-0] [Medline: 30451941]
42. Weagle CL, Snider G, Li C, et al. Global sources of fine particulate matter: interpretation of PM2.5 chemical composition observed by SPARTAN using a global chemical transport model. Environ Sci Technol 2018 Oct 16;52(20):11670-11681. [doi: 10.1021/acs.est.8b01658] [Medline: 30215246]
43. DeMarini DM, Warren SH, Brooks LR. Mutagenicity of the organic fraction of World Trade Center dust. Environ Mol Mutagen 2023 Jan;64(1):16-25. [doi: 10.1002/em.22519] [Medline: 36433931]
44. Jiang X, Han Y, Qiu X, et al. Organic components of personal PM2.5 exposure associated with inflammation: evidence from an untargeted exposomic approach. Environ Sci Technol 2021 Aug 3;55(15):10589-10596. [doi: 10.1021/acs.est.1c02023] [Medline: 34297563]
45. Zhang SY, Shao D, Liu H, et al. Metabolomics analysis reveals that benzo[a]pyrene, a component of PM2.5, promotes pulmonary injury by modifying lipid metabolism in a phospholipase A2-dependent manner in vivo and in vitro. Redox Biol 2017 Oct;13:459-469. [doi: 10.1016/j.redox.2017.07.001] [Medline: 28715731]
46. Wang L, Luo D, Liu X, et al. Effects of PM2.5 exposure on reproductive system and its mechanisms. Chemosphere 2021 Feb;264(Pt 1):128436. [doi: 10.1016/j.chemosphere.2020.128436] [Medline: 33032215]
47. Aksglaede L, Juul A, Leffers H, Skakkebaek NE, Andersson AM. The sensitivity of the child to sex steroids: possible impact of exogenous estrogens. Hum Reprod Update 2006;12(4):341-349. [doi: 10.1093/humupd/dml018] [Medline: 16672247]
48. Chen H, Quick M, Kaufman JS, et al. Impact of lowering fine particulate matter from major emission sources on mortality in Canada: a nationwide causal analysis. Proc Natl Acad Sci U S A 2022 Dec 6;119(49):e2209490119. [doi: 10.1073/pnas.2209490119] [Medline: 36442082]
49. Luo J, Craver A, Jin Z, et al. Contextual deprivation, race and ethnicity, and income in air pollution and cardiovascular disease. JAMA Netw Open 2024 Aug 1;7(8):e2429137. [doi: 10.1001/jamanetworkopen.2024.29137] [Medline: 39158908]
50. Cheng G, Buyken AE, Shi L, et al. Beyond overweight: nutrition as an important lifestyle factor influencing timing of puberty. Nutr Rev 2012 Mar;70(3):133-152. [doi: 10.1111/j.1753-4887.2011.00461.x] [Medline: 22364156]

XSL•FO
RenderX

51. Hsia SM, Chiang YF, Chen HY, Ali M, Wang PS, Wang KL. Effect of high-fructose diet-induced metabolic syndrome on the pituitary-gonadal axis in male rats. Biomedicines 2022 Nov 23;10(12):3009. [doi: 10.3390/biomedicines10123009] [Medline: 36551765]

Abbreviations

BC: black carbon
EDC: endocrine-disruptor chemical
NH_4^+: ammonium
NO_3^-: nitrate
OM: organic matter
OR: odds ratio
PAH: polycyclic aromatic hydrocarbon
$PM_{2.5}$: fine particulate matter
PRODY: Prevalence and Risk Factors for Obesity and Diabetes in Youth
Q: quartile
QGC: quantile-based g-computation
SO_4^{2-}: sulfate
WHO: World Health Organization
WQS: weighted quantile sum

Edited by A Mavragani, T Sanchez; submitted 03.06.24; peer-reviewed by J Ma, P Tang; revised version received 12.12.24; accepted 12.12.24; published 07.02.25.

Please cite as:
Zhou X, Zhang X, Bai G, Dong G, Li X, Chen R, Chen S, Zheng R, Wang C, Wei H, Cao B, Liang Y, Yao H, Su Z, Maimaiti M, Luo F, Li P, Zhu M, Du H, Yang Y, Cui L, Wang J, Yuan J, Liu Z, Wu W, Zhao Q, Fu J
Long-Term Exposure to Fine Particulate Matter (PM2.5) Components and Precocious Puberty Among School-Aged Children: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e62861
URL: https://publichealth.jmir.org/2025/1/e62861
doi:10.2196/62861

© Xuelian Zhou, Xiaochi Zhang, Guannan Bai, Guanping Dong, Xinyi Li, Ruimin Chen, Shaoke Chen, Rongxiu Zheng, Chunlin Wang, Haiyan Wei, Bingyan Cao, Yan Liang, Hui Yao, Zhe Su, Mireguli Maimaiti, Feihong Luo, Pin Li, Min Zhu, Hongwei Du, Yu Yang, Lanwei Cui, Jinling Wang, Jinna Yuan, Zhuang Liu, Wei Wu, Qi Zhao, Junfen Fu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 7.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Interpreting the Influence of Using Blood Donor Residual Samples for SARS-CoV-2 Seroprevalence Studies in Japan: Cross-Sectional Survey Study

Ryo Kinoshita[1], PhD; Sho Miyamoto[2], DVM, PhD; Tadaki Suzuki[2], MD, PhD; Motoi Suzuki[1], MD, PhD; Daisuke Yoneoka[1], PhD

[1]Center for Surveillance, Immunization, and Epidemiologic Research, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-Ku, Tokyo 162-0052 Japan, Tokyo, Japan

[2]Department of Pathology, National Institute of Infectious Diseases, Tokyo, Japan

Corresponding Author:
Daisuke Yoneoka, PhD
Center for Surveillance, Immunization, and Epidemiologic Research, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-Ku, Tokyo 162-0052 Japan, Tokyo, Japan

Abstract

Background: Residual blood donor samples are commonly used in SARS-CoV-2 seroepidemiological studies; however their use may introduce bias due to the healthy donor effect, wherein blood donors are generally healthier than the general population. This potential bias is critical for accurately interpreting seroepidemiological data, as blood donors might not fully represent broader population-level infection rates. **Objective:** This study aims to assess the potential bias in SARS-CoV-2 seroprevalence estimates derived from blood donor samples in Japan by examining the association between blood donation history and COVID-19 diagnosis. By quantifying the healthy donor effect, we seek to refine the interpretation of SARS-CoV-2 seroepidemiological studies using residual blood donor samples. **Methods:** We conducted a web-based survey from December 14 to 28, 2023, recruiting 10,781 Japanese residents aged 16 - 69, stratified by demographic factors to match national representation. Participants provided information on demographics, socioeconomic status, COVID-19 vaccination history, comorbidities, and blood donation experience. A logistic regression model adjusting for confounders such as age, sex, education, occupation, comorbidities, and vaccination status, was used to estimate the odds of COVID-19 diagnosis among blood donors compared to nondonors. **Results:** Of the 10,781 participants, 3583 (33.2%) reported a history of COVID-19 diagnosis, and 5015 (46.5%) indicated they had donated blood at least once in their lifetime, and 1128 (10.5%) donated within the last year. Blood donors had mean of 13.5 (SD 43.6) donations and were older, with a mean age of 46.4 (SD 13.9) years, compared to 38.5 (SD 14.1) years for nondonors. Among blood donors, 39.9% had comorbidities (95% CI 38.5 - 41.2) compared to 27.9% (95% CI 26.7 - 29.0) of nondonors. Blood donors had 1.62 (95% CI: 1.48 - 1.78) times higher odds of COVID-19 diagnosis compared to nondonors. The higher diagnosis rate among blood donors likely reflects increased social interactions and health-seeking behaviors, a phenomenon we refer to as the inverse healthy donor effect. This suggests that blood donor samples could overestimate SARS-CoV-2 seroprevalence when generalized to the broader Japanese population. **Conclusions:** Higher COVID-19 diagnosis rates among blood donors may reflect increased community involvement and health-seeking behaviors, suggesting an inverse healthy donor effect. This pattern indicates that in terms of SARS-CoV-2 infection, blood donors might not represent the healthiest segment of the population. Consequently, seroprevalence studies using blood donor samples could overestimate SARS-CoV-2 infection rates in the general Japanese population. For more accurate public health surveillance, the development of statistical methods to adjust for this bias is recommended.

(JMIR Public Health Surveill 2025;11:e60467) doi:10.2196/60467

KEYWORDS

SARS-CoV-2; COVID-19; seroprevalence; blood donor; selection bias; healthy donor effect; coronavirus; pandemic; Japan; cross-sectional study; residual blood; epidemiology; blood donation; web-based; logistic regression; social economic; comorbidity; COVID-19 vaccination; public health

Introduction

The COVID-19 pandemic, triggered by the rapid dissemination of SARS-CoV-2, has highlighted the critical need for robust epidemiological surveillance to decipher the virus's transmission dynamics and shape effective public health strategies. As the pandemic intensified, the exponential increase in cases coupled with the presence of mild and asymptomatic infections strained traditional case-based surveillance systems. These systems primarily capture symptomatic cases seeking healthcare, leaving

XSL•FO
RenderX

a substantial proportion of the infected population—particularly those with mild or asymptomatic infections—undetected. This complication led to the adoption of seroprevalence studies, which estimate the extent of infection by measuring the presence of nucleocapsid or spike antibodies in populations to provide a broader estimation of infection spread, inclusive of undiagnosed cases.

While seroprevalence studies are invaluable for assessing SARS-CoV-2 exposure at the population level, they are subject to inherent biases [1,2]. For example, antibody levels can wane over time, and variations in assay sensitivity or specificity may underestimate or overestimate the seroprevalence. When residual blood samples from blood donors are used, an additional potential bias known as the healthy donor effect is introduced [3]. This effect arises because blood donors are typically healthier than the general population [4-7], potentially resulting in an underestimation of true seroprevalence. Although blood donor–based SARS-CoV-2 seroepidemiological studies have been widely conducted across various countries [8-11], these studies have not fully accounted for the extent of this effect. Recognizing and addressing such bias is critical to improving the reliability of seroprevalence estimates and their use in public health planning.

Despite recognition of the healthy donor effect, its quantitative impact on SARS-CoV-2 seroprevalence estimates remains largely unexplored, introducing uncertainty in interpreting findings derived from blood donor samples [12]. Our study addresses this gap by focusing on the quantification of the healthy donor effect within Japan's blood donor population, with the goal of improving the interpretation of SARS-CoV-2 seroprevalence estimates and providing a more accurate assessment of the population's exposure to SARS-CoV-2.

Methods

Study Participants and Questionnaire

The study surveyed Japanese residents aged 16 - 69 years—the age group eligible for blood donation in Japan—from all 47 prefectures who could respond in Japanese. The sample size was set at 10,829, based on an approximately 6.1% national blood donation rate reported in a 2022 Japanese Red Cross survey [13], with a 5% alpha error and 80% power. To ensure national representation, a quota-controlled sampling approach was applied, stratified by age, sex, and region, with quotas informed by the 2020 national census. Participants were recruited from Cross Marketing Inc.'s (Japan) panel of 5 million active and diverse members as of March 2024, who volunteered and were incentivized with point-based rewards redeemable for goods and services from affiliated businesses upon completion of the questionnaire. Panel members were accepted on a first-come-first-served basis until quotas for each age and prefecture category were met, ensuring coverage across all 47 prefectures, including both urban and rural populations. The survey was conducted from November 14 to 28, 2023, in alignment with a national seroepidemiological survey among blood donors in Japan. To ensure data completeness, respondents were requested to answer all questions.

The online questionnaire was developed based on a comprehensive literature review on blood donation and epidemiological surveys. For questions related to blood donation, we referenced validated items from governmental surveys [14]. The survey items were divided into three modules: (1) sociodemographic characteristics such as sex, age at the survey, residential prefecture, occupation, education level, and household income; (2) blood donation–related characteristics such as blood donation experience (yes or no) and the total number of blood donations; and (3) COVID-19-related and clinical questions such as the number and timing of prior COVID-19 infections, the number of COVID-19 vaccinations, and medical history. The definitions for COVID-19 vaccination, blood donation, and comorbidity are provided in Multimedia Appendix 1. A binary variable was created from medical history records to identify comorbidities that rendered participants ineligible for blood donation and medically unable to receive the COVID-19 vaccine (see Multimedia Appendix 2). These comorbidities included current treatment for periodontal disease or dental caries, a diagnosis of high blood pressure, diabetes, asthma, or chronic obstructive pulmonary disease, and a history of angina, myocardial infarction, stroke, or cancer.

Statistical Analysis

First, sociodemographic data were tabulated separately for participants with and without blood donation experience to provide basic information on the study population. Continuous variables were analyzed using Mann-Whitney U test and categorical variables were analyzed using Fisher exact test. To estimate the association between COVID-19 infection and blood donation experience, a logistic regression model was used and the associated odds ratio was estimated. Based on a directed acyclic graph, these confounders were adjusted: age groups (16-29, 30-39, 40-49, 50-59, 60-69 years), sex, residential region, education level, occupation, number of blood donations, number of vaccinations, and comorbidities (Multimedia Appendix 3). As behavioral changes induced by vaccination are expected to differ by age group, an interaction term between age groups, and the number of vaccinations was also adjusted in the model. Further, to explore the model robustness, we conducted subgroup analyses using Firth logistic regression to account for potential bias due to small sample sizes or sparse data. Model performance was evaluated by calculating c-statistics, performing the Hosmer-Lemeshow test, and assessing multicollinearity using the adjusted generalized variance inflation factor (GVIF), where GVIF $^{1/(2\times df)}$, with *df* representing the degrees of freedom. All data analyses were conducted using R software (version 4.3.2; R Foundation for Statistical Computing).

Ethical Considerations

This study was reviewed and approved by the Institutional Review Board at the National Institute of Infectious Diseases (authorization no. 1579). Informed consent was obtained from all participants within the web survey, with only those providing consent allowed to proceed. All data used in the analysis were fully anonymized by Cross Marketing Inc. before being provided to the researchers. Participants received point-based rewards as

XSL•FO
RenderX

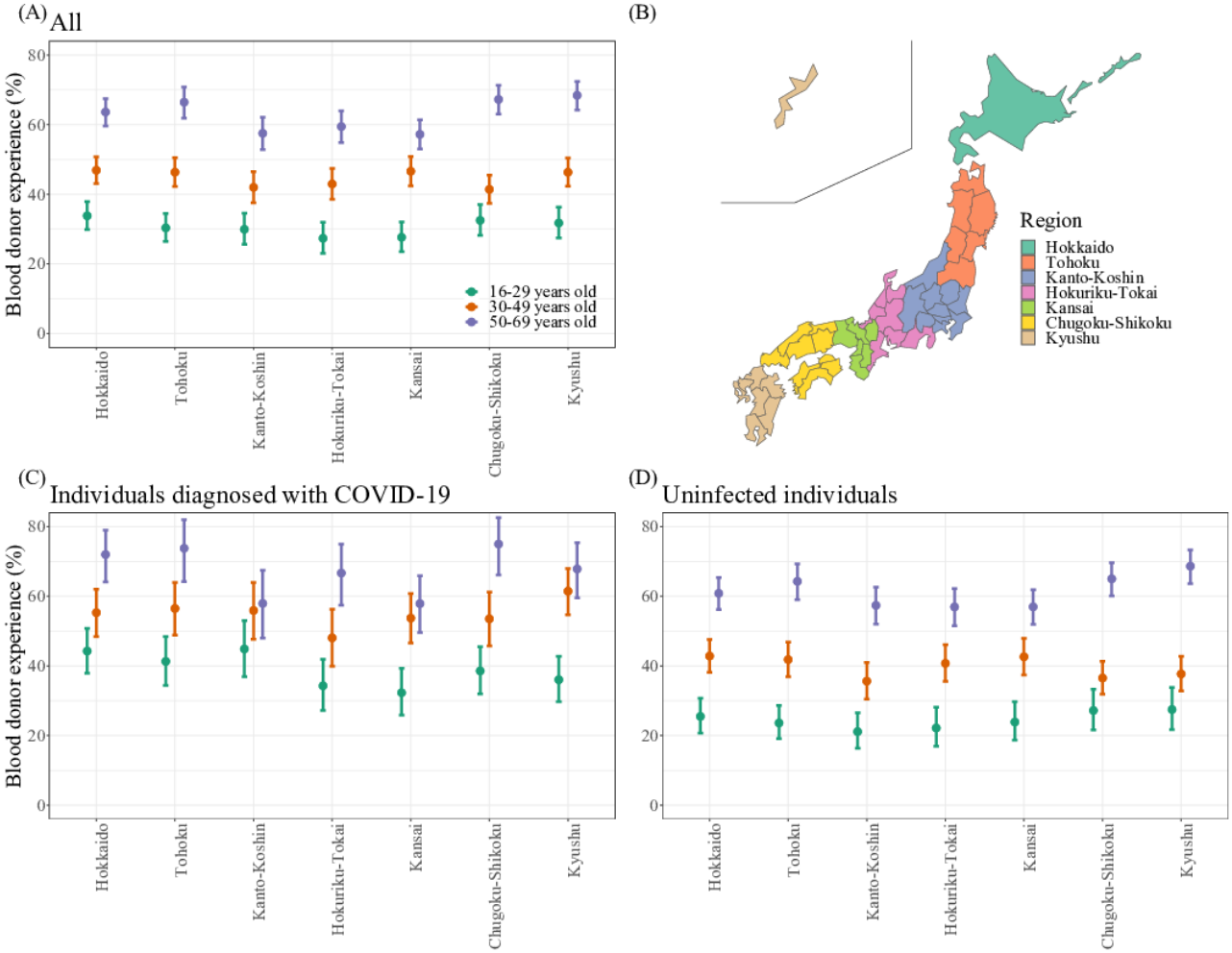
compensation for survey completion, in line with institutional review board–approved ethical guidelines.

Results

The main characteristics of the study participants are presented in Multimedia Appendix 4. A total of 10,781 participants were enrolled in Japan during the study period from December 14 to 28, 2023. Of these, 3583 (33.2%) had a history of COVID-19 diagnosis, and 5015 (46.5%) reported having donated blood at least once in their lifetime. Furthermore, 1128 (10.5%) had donated blood within the past year (Multimedia Appendix 5). Blood donors (defined as those with a history of blood donation) reported mean 13.5 (SD 43.6) donations. . Multimedia Appendix 6 shows the detailed distribution of blood donation frequency in our web survey. The mean age of blood donors was 46.4 (SD 13.9) years, whereas nondonors were younger, with a mean age of 38.5 (SD 14.1) years. Among blood donors, 39.9% (95%CI 38.5 - 41.2) had comorbidities, compared to 27.9% (95% CI 26.7 - 29.0) of nonblood donors (see Multimedia Appendix 7).

Figure 1A depicts the proportion of blood donors categorized by age group and geographic region in Japan. The corresponding geographic regions are illustrated in Figure 1B. In general, individuals who have been diagnosed with COVID-19 had a higher prevalence of prior blood donation compared to those without a history of diagnosis. The estimated results of the logistic regression model shown in Table 1. After excluding individuals with comorbidities that render them ineligible to donate blood or receive the COVID-19 vaccine, and after adjusting for potential confounders, we found that the odds of infection among blood donors were 1.62 times greater (95%CI 1.48 - 1.78) than nondonors. Overall, the positive odds ratio was similar across most subgroup analysis with different stratifications, with most results being statistically significant (see Multimedia Appendix 8). Model validation resulted in a c-statistic of 0.66, a Hosmer-Lemeshow test value of χ^2=6.78 (*P*=.56), and adjusted GVIF values below 5, with a mean of 1.34.

Figure 1. Proportion of blood donors and COVID-19 diagnosis status among Japanese residents aged 16 - 69 years, based on a web-based survey conducted across all 47 prefectures in Japan from December 14 to 28, 2023. (A) Overall proportion of participants with blood donation experience. (B) Map showing geographic regions represented in the study. (C) Proportion of blood donors among individuals diagnosed with COVID-19. (D) Proportion of blood donors among uninfected (undiagnosed) individuals.



XSL•FO
RenderX

Table . Logistic regression results examining the association between COVID-19 diagnosis (outcome) and blood donor experience (main variable of interest) among Japanese residents. Odds ratios with 95% CI and *P* values are provided. Additional variables, including demographic, socioeconomic, and health-related factors, were included as adjustment variables based on a directed acyclic graph (DAG) to control for potential confounding. Vaccination status, household income, and blood donor frequency were included as continuous variables.

Variables	Odds Ratio (OR)	(95% CI)	*P* value
Age group (years)			
16 - 29 (Ref)[a]	1.00	_[b]	–
30 - 39	0.88	(0.67 - 1.14)	.37
40 - 49	0.59	(0.48 - 0.74)	<.001
50 - 59	0.59	(0.46 - 0.77)	<.001
60 - 69	0.55	(0.40 - 0.77)	<.001
Vaccination status	1.05	(1.01-1.10)	.04
Sex			
Female (Ref)[a]	1.00	–	–
Male	0.90	(0.82 - 0.99)	.05
Region			
Hokkaido	1.13	(0.97 - 1.32)	.14
Tohoku	0.97	(0.82 - 1.14)	.73
Kanto-Koshin (Ref)[a]	1.00	–	–
Hokuriku-Tokai	1.11	(0.94 - 1.31)	.27
Kansai	1.28	(1.09 - 1.50)	.004
Chugoku-Shikoku	1.07	(0.91 - 1.26)	.47
Kyushu	1.39	(1.19 - 1.62)	<.001
Highest level of education			
Middle school / High school	1.01	(0.92 - 1.11)	.85
Junior. college / Vocational school / University (Ref)	1.00	–	–
Graduate school (Master / PhD)	0.72	(0.58 - 0.89)	.004
Occupation			
Commerce	1.72	(1.44 - 2.06)	<.001
Construction / Manufacturing / Transportation	1.75	(1.46 - 2.09)	<.001
Education / Student	2.39	(1.96 - 2.92)	<.001
Food / Beverage / Accommodation	1.51	(1.10 - 2.06)	.01
Homemaker	1.88	(1.49 - 2.37)	<.001
Information / Communication	1.92	(1.52 - 2.44)	<.001
Medical / Social welfare	2.53	(2.05 - 3.13)	<.001
Primary industries	1.75	(1.15 - 2.65)	.01
Public servant	2.22	(1.77 - 2.80)	<.001
Other	1.38	(1.10 - 1.75)	.01
Unemployed (Ref)	1.00	–	–
Household Income	1.06	(1.04 - 1.07)	<.001
Blood donor experience			
No (Ref)[a]	1.00	–	–
Yes	1.62	(1.48 - 1.78)	<.001

XSL•FO
RenderX

Variables	Odds Ratio (OR)	(95% CI)	*P* value
Blood donor frequency	1.00	(1.00 - 1.00)	.85
Comorbidity			
No (Ref)[a]	1.00	–	–
Yes	1.40	(1.28 - 1.53)	<.001
Age group (years) * Vaccination Status			
Age group 16 - 29 * Vaccination status (Ref)[a]	1.00	–	–
Age group 30 - 39 * Vaccination status	0.93	(0.86 - 1.01)	.12
Age group 40 - 49 * Vaccination status	0.95	(0.89 - 1.02)	.16
Age group 50 - 59 * Vaccination status	0.87	(0.81 - 0.94)	<.001
Age group 60 - 69 * Vaccination status	0.86	(0.79 - 0.93)	<.001

[a]Ref: Reference variable.

[b]–: Not available.

Discussion

Our study showed a significant association between blood donation experience and COVID-19 diagnosis. However, the diagnosis rates did not increase with the frequency of donations, suggesting that the act of donating blood itself did not contribute to the transmission risk. Instead the higher incidence of COVID-19 among donors may be attributed to increased social interactions and greater health awareness, which could be characterized as an 'inverse' healthy donor effect [15]. Donors' community involvement could raise exposure risks, while their health vigilance could lead to more frequent SARS-CoV-2 testing, leading to higher COVID-19 diagnosis. Further studies are needed to understand the potential underlying mechanism behind this association.

We found that using blood donor samples could potentially bias seroprevalence estimates in Japan. The observed positive association suggests that these samples may overestimate the actual rates compared to the general population. Previous efforts to adjust for demographic variances, such as age, gender, and prefecture, have not completely resolved the bias inherent in using blood donor samples [11]. This highlights the need to measure the healthy donor effect for accurate SARS-CoV-2 seroprevalence interpretations and emphasizes the challenge of correctly determining incidence from donor data, underscoring the need for improved research methodologies. Development of a statistical method to adjust for this bias of blood donor samples in estimating the seroprevalence of the general population of Japan using blood donor samples remains an ongoing focus of our research group.

There are a few limitations to note. First, the identification of COVID-19 cases was based on clinical diagnoses rather than antibody-confirmed infection rates, suggesting that the study may reflect differences in health-seeking behavior rather than actual SARS-CoV-2 infection rates. Second, although we applied quota sampling based on demographic distributions from the national census to improve representation, the sample population may not fully represent the general population in Japan, as the study was conducted via a web-based survey. This approach may introduce potential selection bias, particularly excluding individuals without internet access or those less comfortable with technology. Third, the study relied on self-reported data, which may introduce recall bias or social desirability bias. Fourth, we may not have fully adjusted for residual confounding from variables that were not included in the survey. Fifth, the cross-sectional design has inherent limitations, including difficulty determining temporal relationships and tracking changes over time. Finally, our findings may be specific to the Japanese context and might not necessarily apply to other countries due to differences in demographics, healthcare systems, and social behavior. Further evaluation is needed to assess the external validity across different countries or regions.

In conclusion, our study has identified a potential bias in SARS-CoV-2 seroprevalence studies that use residual samples from blood donors. Quantifying the direction and magnitude of this bias is essential for accurately interpreting seroprevalence surveys, especially when these surveys are used to guide infectious disease surveillance.

XSL•FO
RenderX

Acknowledgments

This study was supported by the Health Labour Sciences Research Grant from the Ministry of Health, Labour and Welfare (23HA2017, 24HA2016). RK and DY received funding from the Japan Society for the Promotion of Science (JSPS) KAKENHI (21K17307 and 21K17292).

Data Availability

The data sets generated during and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: RK, DY
Data curation: RK, DY
Formal analysis: RK, DY
Writing – original draft: RK, DY
Writing – review and editing: RK, SM, TS, MS, DY

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description of COVID-19 vaccination and blood donation-related variables included in the model.
[DOCX File, 14 KB - publichealth_v11i1e60467_app1.docx]

Multimedia Appendix 2
Description of comorbidities that affect eligibility for blood donation in Japan, used to create the binary variable Comorbidity.
[DOCX File, 15 KB - publichealth_v11i1e60467_app2.docx]

Multimedia Appendix 3
Directed acyclic graph (DAG) illustrating the relationship between blood donor experience (exposure) and COVID-19 infection (outcome), with other variables included as potential confounders.
[DOCX File, 324 KB - publichealth_v11i1e60467_app3.docx]

Multimedia Appendix 4
Demographic, socioeconomic, and health characteristics of study participants in the web survey, categorized by COVID-19 diagnosis status (uninfected or infected) and blood donation status (nonblood donor or blood donor).
[DOCX File, 28 KB - publichealth_v11i1e60467_app4.docx]

Multimedia Appendix 5
Comparison of blood donation rates within 1 year among the sampled population (circle points) and Japanese Red Cross Society data for 2022 (star points), categorized by age group and region. The figure displays blood donation proportions across different age groups (16–29, 30–49, 50–69 years) and regions, with whiskers representing 95% confidence intervals.
[DOCX File, 58 KB - publichealth_v11i1e60467_app5.docx]

Multimedia Appendix 6
Frequency distribution of blood donation experience among web survey participants. (A) Number of blood donations over participants' entire lifetime. (B) Number of blood donations within the past year. The X-axis is presented on a logarithmic scale to better illustrate the range and variability in donation frequencies across the study population.
[DOCX File, 43 KB - publichealth_v11i1e60467_app6.docx]

Multimedia Appendix 7
Proportion of comorbidities among participants by blood donor status (0 = nondonor, 1 = donor). Whiskers represent 95% confidence intervals.
[DOCX File, 35 KB - publichealth_v11i1e60467_app7.docx]

Multimedia Appendix 8

XSL•FO
RenderX

Supplementary Table 4.
[DOCX File, 24 KB - publichealth_v11i1e60467_app8.docx]

References

1. Accorsi EK, Qiu X, Rumpler E, et al. How to detect and reduce potential sources of biases in studies of SARS-CoV-2 and COVID-19. Eur J Epidemiol 2021 Feb;36(2):179-196. [doi: 10.1007/s10654-021-00727-7] [Medline: 33634345]
2. Takahashi S, Greenhouse B, Rodríguez-Barraquer I. Are seroprevalence estimates for severe acute respiratory syndrome coronavirus 2 biased? J Infect Dis 2020 Nov 9;222(11):1772-1775. [doi: 10.1093/infdis/jiaa523] [Medline: 32856712]
3. Atsma F, Veldhuizen I, Verbeek A, de Kort W, de Vegt F. Healthy donor effect: its magnitude in health research among blood donors. Transfusion 2011 Aug;51(8):1820-1828. [doi: 10.1111/j.1537-2995.2010.03055.x] [Medline: 21342203]
4. Brodersen T, Rostgaard K, Lau CJ, et al. The healthy donor effect and survey participation, becoming a donor and donor career. Transfusion 2023 Jan;63(1):143-155. [doi: 10.1111/trf.17190] [Medline: 36479702]
5. O'Brien SF, Drews SJ, Lewin A, et al. How do we decide how representative our donors are for public health surveillance? Transfusion 2022 Dec;62(12):2431-2437. [doi: 10.1111/trf.17140] [Medline: 36193865]
6. Saeed S, Uzicanin S, Lewin A, et al. Current challenges of severe acute respiratory syndrome coronavirus 2 seroprevalence studies among blood donors: a scoping review. Vox Sang 2022 Apr;117(4):476-487. [doi: 10.1111/vox.13221] [Medline: 34862614]
7. Su S, Sun Y, Gu X, et al. Exploration of the healthy donor effect among 0.6 million blood donors in China: longitudinal study. JMIR Public Health Surveill 2024 Feb 22;10:e48617. [doi: 10.2196/48617] [Medline: 38386403]
8. Jones JM, Stone M, Sulaeman H, et al. Estimated US Infection- and Vaccine-Induced SARS-CoV-2 Seroprevalence Based on Blood Donations, July 2020-May 2021. JAMA 2021 Oct 12;326(14):1400-1409. [doi: 10.1001/jama.2021.15161] [Medline: 34473201]
9. Chang L, Hou W, Zhao L, et al. The prevalence of antibodies to SARS-CoV-2 among blood donors in China. Nat Commun 2021 Mar 2;12(1):1383. [doi: 10.1038/s41467-021-21503-x] [Medline: 33654063]
10. Erikstrup C, Laksafoss AD, Gladov J, et al. Seroprevalence and infection fatality rate of the SARS-CoV-2 Omicron variant in Denmark: A nationwide serosurveillance study. Lancet Reg Health Eur 2022 Oct;21:100479. [doi: 10.1016/j.lanepe.2022.100479] [Medline: 35959415]
11. Kinoshita R, Arashiro T, Kitamura N, et al. Infection-Induced SARS-CoV-2 Seroprevalence among Blood Donors, Japan, 2022. Emerg Infect Dis 2023 Sep;29(9):1868-1871. [doi: 10.3201/eid2909.230365] [Medline: 37506681]
12. Patel EU, Bloch EM, Tobian AAR. Severe acute respiratory syndrome coronavirus 2 serosurveillance in blood donor populations. J Infect Dis 2022 Jan 5;225(1):1-4. [doi: 10.1093/infdis/jiab517] [Medline: 34626464]
13. Blood donation statistics. Japanese Red Cross Society. 2022. URL: https://www.jrc.or.jp/donation/blood/data/ [accessed 2024-06-01]
14. Survey report on blood donation awareness among young people: questionnaire (in japanese). Ministry of Health Labour and Welfare. 2006. URL: https://www.mhlw.go.jp/new-info/kobetu/iyaku/kenketsugo/7n-04.html [accessed 2024-10-15]
15. Studte S, Clement M, Soliman M, Boenigk S. Blood donors and their changing engagement in other prosocial behaviors. Transfusion 2019 Mar;59(3):1002-1015. [doi: 10.1111/trf.15085] [Medline: 30549293]

Abbreviations

GVIF: generalized variance inflation factor

Edited by A Mavragani; submitted 15.05.24; peer-reviewed by L Aljerf, MJ Miller; revised version received 11.11.24; accepted 12.11.24; published 10.02.25.

Please cite as:
Kinoshita R, Miyamoto S, Suzuki T, Suzuki M, Yoneoka D
Interpreting the Influence of Using Blood Donor Residual Samples for SARS-CoV-2 Seroprevalence Studies in Japan: Cross-Sectional Survey Study
JMIR Public Health Surveill 2025;11:e60467
URL: https://publichealth.jmir.org/2025/1/e60467
doi:10.2196/60467

© Ryo Kinshita, Sho Miyamoto, Tadaki Suzuki, Motoi Suzuki, Daisuke Yoneoka. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly

XSL•FO
RenderX

cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Among BMI, Self-Esteem, and Nonsuicidal Self-Injury in Young Adults to Understand the Influence of Socioenvironmental Factors: Longitudinal Study

Yi Zhang[1,2*], MD; Ruixue Ying[1*], MD; Wan Lu[1*], MD; Xuemeng Liu[1*], MD; Keyan Hu[3], MD; Qing Feng[1], MD; Zixiang Yu[4], MD; Zhen Wang[1], MD; Fangting Lu[1], MD; Yahu Miao[1], MD; Nanzhen Ma[5], MD; Fangbiao Tao[1], MD; Tian Jiang[1], MD; Qiu Zhang[1], PhD

[1]Department of Endocrinology, First Affiliated Hospital of Anhui Medical University, No 81 Meishan Road, Anhui, Hefei, China

[2]Department of Maternal, Child, and Adolescent Health, School of Public Health, Anhui Medical University, Hefei, China

[3]The First Affiliated Hospital and College of Clinical Medicine of Henan, University of Science and Technology, Luoyang, China

[4]Department of Pediatrics, The First Affiliated Hospital of Anhui Medical University, Hefei, China

[5]Hospital of Anhui Medical University, Hefei, China

[*]these authors contributed equally

Corresponding Author:
Qiu Zhang, PhD
Department of Endocrinology, First Affiliated Hospital of Anhui Medical University, No 81 Meishan Road, Anhui, Hefei, China

Abstract

Background: Nonsuicidal self-injury (NSSI) is a major public health problem leading to psychological problems in adolescents and young adults, similar to disorders such as depression and anxiety.

Objective: The aims of this study were to investigate (1) the interaction between BMI and socioenvironmental factors (including chronotype and mental health) that contribute to NSSI, and (2) whether self-esteem plays a mediating role in this association.

Methods: From May to June 2022, the multistage cluster sampling method was used to sample college students in four grades, including freshmen and seniors. The baseline participants were followed up 6 months later, excluding those who did not qualify, and the participants included 1772 college students. Socioenvironmental factors (chronotype/mental health), self-esteem, and NSSI were measured using a questionnaire. Multivariate linear regression models and chi-square analysis were used to evaluate the linear relationship between BMI, socioenvironmental factors, and self-esteem and the NSSI status. We use a process approach (mediation-moderation analysis) to explore the complex relationships between these variables.

Results: The mean age of the participants was 20.53 (SD 1.65) years at baseline. A significant association was revealed, suggesting that a high BMI (β=.056, 95% CI 0.008 - 0.086, *P*=.018) was associated with a higher NSSI. There was also an interaction among BMI, socioenvironmental factors, and NSSI. Socioenvironmental factors played both moderating and mediating roles in the relationship between BMI and NSSI, whereas self-esteem only played a mediating role.

Conclusions: Paying attention to factors such as overweight and obesity is important for early BMI control to identify other potential risk factors for NSSI and to evaluate how self-esteem can be improved considering multiple perspectives to improve the effect of BMI on NSSI in adolescents.

(JMIR Public Health Surveill 2025;11:e52928) doi:10.2196/52928

KEYWORDS

nonsuicidal self-injury; chronotype; BMI; self-esteem; body mass index; adolescent; young adult; teenager; social environmental factor; self-injury; sampling method; undergraduate; college student; linear regression; regression; regression model

Introduction

Nonsuicidal self-injury (NSSI) refers to physical harm, not the intention to commit suicide, and is not a socially recognized behavior [1]. NSSI, while being distinct from suicidal behavior, is a strong predictor of suicide and is also a significant mental health problem among adolescents worldwide [2]. NSSI is gradually increasing in the world; it seriously damages people's physical and mental health and is a psychopathological behavior, which has aroused people's attention and concern. The *DSM-5* identifies NSSI as a separate behavioral category and lists it as "an entry for further research" [3]. Adolescence and college years are important periods of psychological development and changes in life. In a meta-study, the reporting rate of NSSI was 22.5% (95% CI 17.2% - 28.9%) among adolescents [2]. The prevalence of NSSI ranged from 3.2% to 44.8% among

XSL•FO
RenderX

nonclinical populations [4]. For clinical populations, the incidence of NSSI in the past 12 months ranged from 5% to 16.4% [4]. The harm of NSSI goes far beyond that, and after the emergence of NSSI in adolescence, it has maintained a certain "trajectory" into adulthood, including a continuous increase or fluctuation between the increase and decrease of NSSI, as well as an increase in psychobehavioral problems that largely affect people's lives [5,6]. A systematic review and Bayesian meta-analysis concluded that in the entire developmental stage of adolescents, the NSSI frequency (but not the frequency of occurrence) in younger adolescents increased, that in middle adolescents remained stable, and that in older adolescents decreased [3]. On proceeding into the aftermath of the COVID-19 outbreak and trying to recover from its consequences, preventing NSSI and suicidal behaviors has continued to be a priority. In general, the current NSSI epidemic in China is serious [7], and it has become a major public health problem that severely endangers adolescents' physical and mental health. Being an ongoing social health problem, NSSI is a major concern for mental health and it deserves our attention.

With such a high incidence rate of NSSI that has not been effectively controlled thus far, it is important to explore the potential factors of and prevention measures for NSSI, including what factors affect NSSI, how they interact with each other, and what needs to be done to protect those with high NSSI frequency. The specific factors that affect NSSI are as follows: (1) individual psychological factors, including self-abasement, impulsivity, early traumatic experience, self-esteem, and other characteristics; (2) external environmental factors, including negative life events (events that individuals feel unpleasant); (3) neurobiological factors, including the influence of the metabolism of substances such as 5-hydroxytryptamine, glutamate, and dopamine, which may also be related to genetics; (4) family factors, including family environment, family atmosphere, and family economic conditions; and (5) school factors, including school environment, teacher-student relationship, and classmate relationship, which may also have an impact on children and adolescents' NSSI. The occurrence of NSSI behavior in adolescents is not the result of a single factor, but the result of the interaction of multiple factors. In addition, obesity has attracted considerable attention as a cause of many diseases, including physical diseases as well as psychological and behavioral problems [8]. Changes in BMI are correlated with suicide [9], and there is a significant association with suicide attempts in underweight and perceived overweight groups [10]. There is an association between psychological problems and NSSI [10,11]. A previous study demonstrated that while posttraumatic stress disorder leads to suicidal ideation, BMI plays a significant moderating role, indicating that the association between posttraumatic stress disorder and suicidal ideation is the strongest in individuals with high BMI values [11]. Another study has reported a correlation between obesity and NSSI and verified the correlation between BMI and major depressive disorder [12], but no precise research has explored the correlation between BMI and NSSI. NSSI is quite different from suicidal behavior in many aspects such as motivation, lethality, and manifestation; however, it remains a crucial risk factor for suicidal ideation and suicide attempts. This finding also suggests that psychological problems play an essential role in determining how BMI leads to NSSI.

Although some studies identified BMI as a risk factor for self-injury or suicide, other studies have not found positive results, which means that not everyone with an increased BMI necessarily shows NSSI behavior [13], as other factors may also be at play. To identify other possible psychological mechanisms that support this relationship, it is necessary to understand complex associations and mechanisms beyond the current one. In addition to the association between BMI and NSSI, it is significant to understand the mediators and moderators of BMI and NSSI for public health prevention and clinical practice. Research on borderline personality disorder has revealed that sleep deprivation, disrupted sleep, and prolonged sleep duration are associated with suicidal behavior or NSSI [14]. Sleep occupies a very important position in the circadian rhythm, especially the lack of sleep; the circadian rhythm refers to the 24-hour physiological and behavioral rhythm of individuals, which is divided into endogenous rhythm and exogenous rhythm. When the endogenous rhythm and exogenous rhythm are "mismatched," circadian rhythms can be disrupted, which can lead to sleep disturbances [15]. Chronotype is a self-description of an individual's circadian preference, reflecting the differences in their daily activity pattern and sleep-wake cycle; it can be generally divided into three types: morning, intermediate, and evening [16]. Moreover, the evening type has been reported to be strongly associated with suicidal behavior [17].

These studies provide evidence that NSSI moderates the relationship between BMI and socioenvironmental factors. A higher BMI has been associated with a lower risk of suicide in a large prospective study [18]; however, the mechanisms underlying this link require elucidation. Some studies have verified the correlation between self-esteem and NSSI [17,19]. However, little is known about how self-esteem moderates the relationship between BMI and NSSI. Socioenvironmental factors underlying these associations also remain unclear. Moreover, only limited research has been conducted on the relationship between BMI and socioenvironmental factors, and the outcome of NSSI during adolescence. A systematic review and meta-analysis reported that treatment of childhood obesity resulted in increased self-esteem after intervention [19]. The results of studies evaluating the self-esteem and NSSI suggest that although low self-esteem and self-injury are associated with both clinical and nonclinical populations, there are many factors influencing this relationship [20]. Therefore, based on a previous study [19], we hypothesize that there was an association among SE, BMI, and NSSI.

From a management and prevention perspective, these modifiable risk factors are crucial, and how they interact with each other requires further exploration. This study aimed to investigate the influencing factors and possible mechanisms of NSSI in college students and to observe the influence of some possible factors on BMI that could result in NSSI. Therefore, this study aimed to (1) examine the correlation among BMI, socioenvironmental factors, self-esteem, and NSSI and (2) determine whether socioenvironmental factors and self-esteem

XSL•FO
RenderX

would play direct and indirect roles in the relationships between BMI and NSSI among Chinese college students.

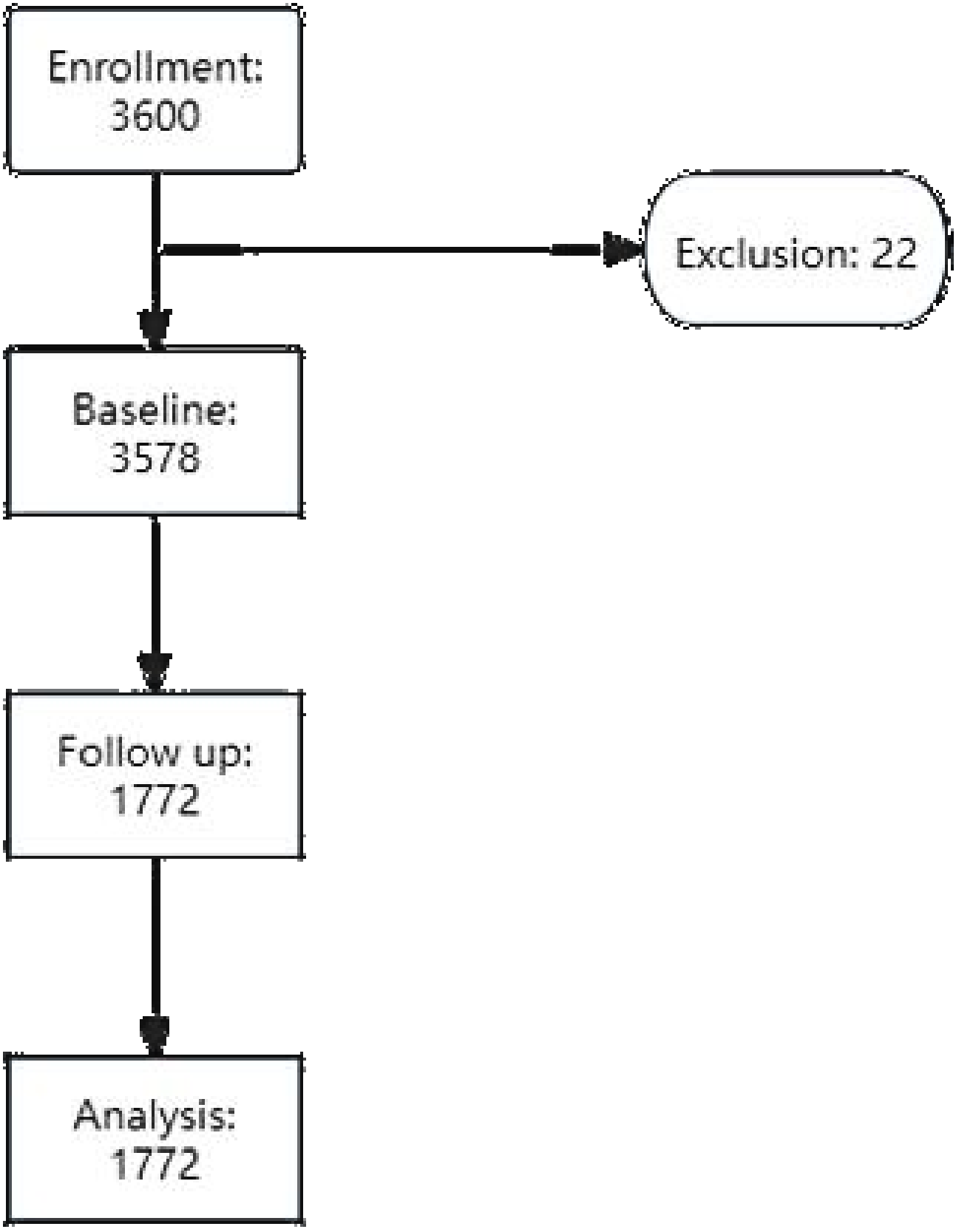
Methods

Study Design and Study Setting for Recruitment

This was a longitudinal study, and the research methods have been described in previous studies [21,22]. The study setting for recruitment was as follows: we first selected a university in a city in the Anhui Province, and then selected college students in four consecutive grades according to the cluster sampling method. A total of 3600 college students were initially recruited using an anonymous electronic questionnaire survey about their health at baseline. The cluster sampling method was used for the schools surveyed, and most of the students from each grade were included in the study. The survey was conducted from May to June 2022. Some participants were excluded from the study owing to unwillingness to answer the questionnaire, absence from class, high levels of missing data (questionnaires with missing values greater than 5% were eliminated), or false responses [23-25] at baseline. The baseline group was followed up with a questionnaire after 6 months. Finally, a total of 1772 participants were included in the follow-up survey (Figure 1).

XSL•FO
RenderX

Figure 1. A flowchart for the enrollment and follow-up of participants.



Inclusion Criteria

Participants were included if (1) informed consent from them and their guardians were obtained, (2) they were college students aged 15 - 26 years, (3) they had no history of mental illness, and (4) they were attending the chosen school.

Exclusion Criteria

Participants were excluded if (1) their informed consent was not obtained, (2) they were junior and high school students, (3) they failed to submit the questionnaire, or (4) they had congenital or acquired immunodeficiency.

XSL•FO
RenderX

Exposure

Chronotype

The morningness-eveningness questionnaire is used to measure college students' sleep habits. According to the previous analysis results, personal sleep habits is classified into three types: morning, intermediate, and evening [26]. The suggested demarcations were further categorized as follows: sleeping from 4 AM to 7 AM is definitely eveningness, sleeping from 8 AM to 11 AM is medium eveningness, sleeping from 12 PM to 5 PM is intermediate type, sleeping from 6 PM to 9 PM is medium morningness, and sleeping from 10 PM to 12 AM is definitely morningness.

Self-Esteem

The self-esteem scale was first developed by Ronsenberg to evaluate an individual's overall feelings of self-worth and self-acceptance [27]. It was later translated and revised into Chinese by Ji Fuyi and Yu Xin [28]. There were 10 items on the scale: 5 items on self-denial and 5 items on self-affirmation. A 4-point Likert scale was used, ranging from 1 indicating "very inconsistent" to 4 indicating "very consistent." The 5 self-denial items were scored inversely. The higher the score, the higher the self-esteem.

Outcome

NSSI

One of the most important question in the questionnaire was "Have you intentionally hurt yourself in the past 12 months, but not for suicide?" The questionnaire also listed several NSSI methods: hit yourself with a fist or palm, pulled your hair, hit a hard object with your head or fist, pinched or scratched yourself, bit yourself, cut yourself, and stabbed yourself. Those who had an NSSI were asked about their NSSI frequency. The participants owning up to 5 or more methods were considered to be prone to NSSI [29].

Statistics Analysis

Conventional means and SDs were used for continuous variables. Univariate ANOVA was used to describe the association between independent and outcome variables. Frequency and percentage tests were used for categorical variables, and chi-square tests were used to explore the association between independent and outcome variables. The PROCESS method was used to conduct a mediation-moderation analysis to explore the relationships among BMI, chronotype, self-esteem, and NSSI [30]. The bootstrap method was used to resample 1000 samples, and the 95% CI was calculated. All data were analyzed using SPSS software (Windows version 23.0; IBM Corp).

Ethical Considerations

The current study is designed and reported according to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist. The design and data collection procedures were approved by the Ethics Committee of the First Affiliated Hospital of Anhui Medical University (review number PJ 2024-06-06). Written informed consent was obtained from the parents or guardians of all the students.

Results

General Demographic Characteristics

Table 1 presents the general distribution of the NSSI variables examined in this study considering the demographic characteristics, while Figure 2 is a directed acyclic graph showing the covariates. Among the 1772 participants included in the questionnaire analysis, 908 (51.1%) were male students; the mean age was 20.53 (SD 1.65) years at baseline; and 853 (48.1%) dwelled in rural areas, 548 (30.9%) in towns, and 371 (20.9%) in urban areas. The prevalence of NSSI was 5.6% (N=1772).

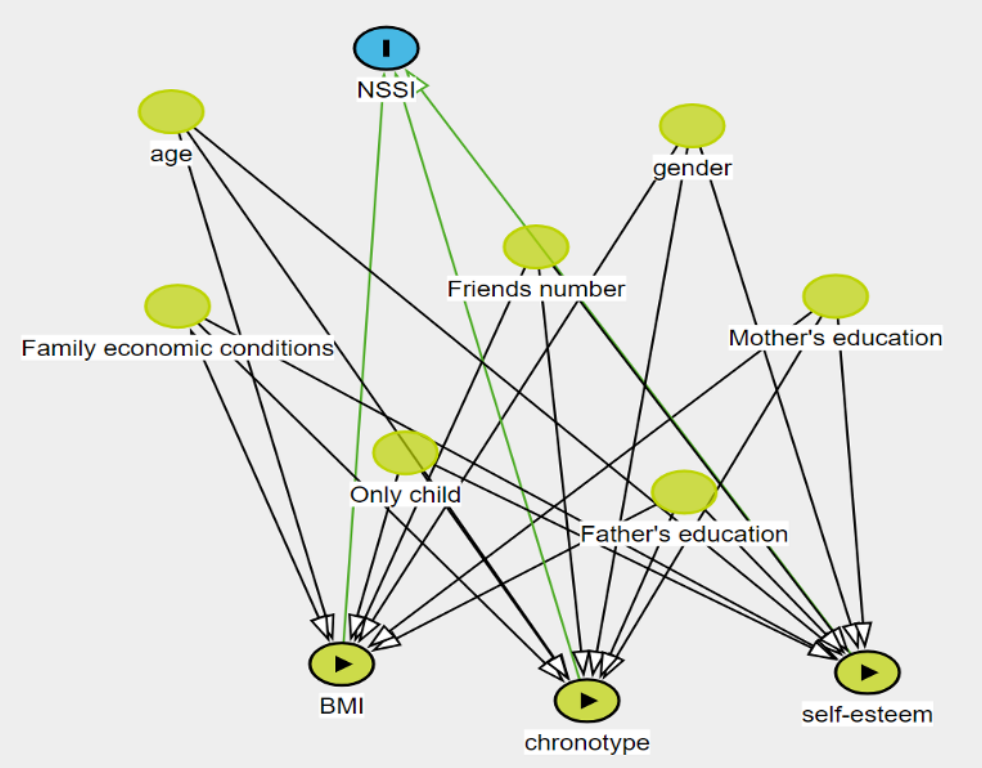
XSL•FO
RenderX

Table . The distribution of the nonsuicidal self-injury (NSSI) variables considering demographic characteristics.

Variables	Total, n (%)	Absence of NSSI, n (%)	Presence of NSSI, n (%)	*t* test/chi-square (*df*)	*P* value
Sex				0.46 (1)	.54
Male	908 (51.1)	854 (51.0)	54 (54.5)		
Female	864 (48.9)	819 (49.0)	45 (45.5)		
Residential areas				0.53 (2)	.76
Rural	788 (44.5)	744 (44.5)	44 (44.4)		
Town	401 (22.6)	376 (22.5)	25 (25.3)		
Urban	583 (32.9)	553 (33.1)	30 (30.3)		
Only child				0.29 (1)	.66
Yes	581 (32.8)	551 (32.9)	30 (30.3)		
No	1191 (67.2)	1122 (67.1)	69 (69.7)		
Father's education				2.31 (5)	.81
Not known	55 (3.1)	4 (4.0)	51 (3.0)		
Below primary	153 (8.6)	7 (7.1)	146 (8.7)		
Primary	205 (11.6)	11 (11.1)	194 (11.6)		
Junior high	661 (37.3)	32 (32.3)	629 (37.6)		
High school or technical secondary school	379 (21.4)	24 (24.2)	355 (21.2)		
Junior college or above	319 (18.0)	21 (21.2)	298 (17.8)		
Mother's education				3.14 (5)	.68
Not known	67 (3.8)	63 (3.8)	4 (4.0)		
Below primary	345 (19.5)	323 (19.3)	22 (22.2)		
Primary	316 (17.8)	299 (17.9)	17 (17.2)		
Junior high	535 (30.2)	504 (30.1)	31 (31.3)		
High school or technical secondary school	301 (17.0)	290 (17.3)	11 (11.1)		
Junior college or above	208 (11.7)	194 (11.6)	14 (14.1)		
Family economic conditions				3.15 (4)	.53
Bad	93 (5.2)	90 (5.4)	3 (3.0)		
Worse	428 (24.2)	402 (24.0)	26 (26.3)		
Medium	1142 (64.4)	1075 (64.3)	67 (67.7)		
Better	100 (5.6)	97 (5.8)	3 (3.0)		
Good	9 (0.5)	9 (0.5)	0 (0)		
Number of friends				4.92 (3)	.18
0	36 (2.0)	32 (1.9)	4 (4.0)		
1 - 2	587 (33.1)	548 (32.8)	39 (39.4)		
3 - 5	854 (48.2)	810 (48.4)	44 (44.4)		
6 or more	295 (16.6)	283 (16.9)	12 (12.1)		

XSL•FO
RenderX

Figure 2. A directed acyclic graph showing the covariates. NSSI: nonsuicidal self-injury.



Association Between Independent Variables and Adolescent Psychological and Behavioral Problems

Table 2 presents the correlation between independent variables and depression symptoms through multilevel linear regression analyses. There was a significant relationship between different BMI and NSSI (β=.056); after controlling for covariates, these relationships were also significant. Furthermore, there was a significant relationship between self-esteem and NSSI (β=−.077); after controlling for covariates, this relationship was also significant. There was no significant correlation between chronotype and NSSI.

Table . Multilevel linear regression between independent variables and nonsuicidal self-injury (NSSI).

	NSSI					
	R^2	β	*t* test (*df*)	*P* value	*F* test (*df*)	95% CI
BMI						
Model 1[a]	0.003	.056	2.359 (1)	.02	5.566 (1)	0.008 to 0.086
Model 2[b]	0.008	.054	2.191 (9)	.03	1.551 (9)	0.005 to 0.085
Chronotype						
Model 1	0.00	.022	0.938 (1)	.99	0.88 (1)	−0.033 to 0.094
Model 2	0.006	.025	1.055 (9)	.29	1.139 (9)	−0.03 to 0.099
Self-esteem						
Model 1	0.006	−.077	−3.234 (1)	<.001	10.461 (1)	−0.136 to −0.033
Model 2	0.01	−.078	−2.896 (9)	.004	1.951 (9)	−0.130 to −0.025

[a]Model 1: crude model.

[b]Model 2: controlled for parent educational level, gender, economic level, whether only child, number of friends, residential area, and age.

Impact of the Mediation Analysis Among BMI, Chronotype, and Self-Esteem on Adolescent NSSI

Mediation-moderation analyses were performed for BMI, self-esteem, chronotype, and NSSI scores. The results are listed in Tables 3 and 4. Among NSSI, findings have suggested the following: a positive association between BMI exposure and NSSI, chronotype was associated with NSSI, and chronotype played a moderating role in the increased risk of NSSI induced by BMI mediated by self-esteem.

XSL•FO
RenderX

Table . Testing the mediation-moderation effects of BMI[a] and chronotype on the nonsuicidal self-injury (NSSI)[b] among college students.

Variables	Self-esteem			NSSI		
	ß	*t* test (*df*)	*P* value	ß	*t* test (*df*)	*P* value
BMI	.0775	−0.4799 (3)	.54	.2262	2.5466 (5)	.01
Chronotype	−.0387	0.6206 (3)	.63	.8102	2.9399 (5)	.003
BMI* chronotype[c]	−.0005	−0.0869 (3)	.93	−.0127	−2.1149 (5)	.04
Self-esteem[d]	N/A[e]	N/A	N/A	.1613	1.3426 (5)	.18
Self-esteem* chronotype	N/A	N/A	N/A	−.0167	−2.0839 (5)	.04
R^2	0.0065	N/A	N/A	.0140	N/A	N/A
F test (*df*)	3.8342	N/A	N/A	5.0212	N/A	N/A

[a]BMI: independent variables.

[b]NSSI: dependent variables.

[c]chronotype: moderated variables.

[d]Self-esteem: mediate variables.

[e]N/A: not applicable.

Table . Bootstrap method showing the conditional direct and indirect effects between self-esteem, chronotype, and nonsuicidal self-injury (NSSI).

	Nonsuicidal self-injury		
	Effect of chronotype	Self-esteem	Lower limit to upper limit
Direct effect	0.0818	0.0270	0.0288 to 0.1347
Predictor	0.0409	0.0199	0.0020 to 0.0799
Moderator	0.0001	0.024	−0.0556 to 0.0558
Indirect effect	11.3542	0.0013	−0.0011 to 0.0053
Predictor	14.5643	0.0038	0.0006 to 0.008
Mediator	17.7744	0.004	−0.0002 to 0.0179

Impact of Mediation Analysis Among BMI, Mental Health, and Self-Esteem on Adolescent NSSI

Mediation-moderation analyses were performed for BMI, self-esteem, mental health, and NSSI. The results are presented in Tables 5 and 6. In NSSI, results have suggested a positive association between BMI exposure and NSSI, where depression was associated with NSSI, and chronotype played a moderating role in the increased risk of NSSI induced by BMI and mediated by self-esteem (Tables 5 and 7). Similar results were obtained for stress and NSSI (Tables 6 and 8).

XSL•FO
RenderX

Table . Testing the mediation-moderation effects of BMI[a] and anxiety on the nonsuicidal self-injury (NSSI)[b] among college students.

Variables	Self-esteem			NSSI		
	ß	*t* test (*df*)	*P* value	ß	*t* test (*df*)	*P* value
BMI	−.0099	−0.1054 (3)	.91	−.2475	−2.2641 (5)	.02
Anxiety	−.1302	−2.5744 (3)	.01	.0233	0.222 (5)	.82
BMI* anxiety[c]	−.0007	−0.2931 (3)	.77	.0070	2.6944 (5)	.007
Self-esteem[d]	N/A[e]	N/A	N/A	.0233	1.2468 (5)	.21
Self-esteem* anxiety	N/A	N/A	N/A	−.0051	−1.8281 (5)	.07
R^2	.1014	N/A	N/A	.0180	N/A	N/A
F test (*df*)	66.5210	N/A	N/A	6.4595	N/A	N/A

[a]Independent variables: BMI.

[b]Dependent variables: NSSI.

[c]Moderated variables: depression.

[d]Mediate variables: self-esteem.

[e]N/A: not applicable.

Table . Bootstrap method showing the conditional direct and indirect effects between self-esteem, anxiety, and nonsuicidal self-injury (NSSI).

	Nonsuicidal self-injury		
	Effect of anxiety	Self-esteem	Lower limit to upper limit
Direct effect	−0.0242	0.0343	−0.0914 to 0.0430
Predictor	0.02390	0.0212	−0.0177 to 0.0655
Moderator	0.0719	0.0238	0.0252 to 0.1186
Indirect effect	0.0001	0.0018	−0.0026 to 0.0051
Predictor	0.0016	0.0017	−0.0006 to 0.0063
Mediator	0.0044	0.0028	0.0001 to 0.0121

Table . Bootstrap method showing the conditional direct and indirect effects self-esteem, depression, and nonsuicidal self-injury (NSSI).

	Nonsuicidal self-injury		
	Effect of depression	Self-esteem	Lower limit to upper limit
Direct effect	−0.0273	0.0325	−0.0911 to 0.0364
Predictor	0.0328	0.0201	−0.0066 to 0.0722
Moderator	0.0929	0.0273	0.0393 to 0.1465
Indirect effect	0.0005	0.0020	−0.0016 to 0.0081
Predictor	0.0022	0.0017	0.00 to 0.0078
Mediator	0.0044	0.0035	−0.000 to 0.0148

XSL•FO
RenderX

Table . Testing the mediation-moderation effects of BMI[a] and depression on the nonsuicidal self-injury (NSSI)[b] among college students.

Variables	Self-esteem			NSSI		
	ß	*t* test (*df*)	*P* value	ß	*t* test (*df*)	*P* value
BMI	.0549	0.7267	.47	−.1616	−1.8212	.07
Depression[c]	−.0766	−2.3582	.02	−.0008	−0.0106	.99
BMI* depression	−.0016	−1.1464	.25	.0040	2.3457	.02
Self-esteem[d]	N/A[e]	N/A	N/A	.0516	0.5076	.61
Self-esteem* depression	N/A	N/A	N/A	−.0026	−1.2582	.21
R^2	.1236	N/A	N/A	.0130	N/A	N/A
F test (*df*)	83.1489	N/A	N/A	4.6602	N/A	N/A

[a]Independent variables: BMI.

[b]Dependent variables: NSSI.

[c]Moderated variables: depression.

[d]Mediate variables: self-esteem.

[e]N/A: not applicable.

Discussion

Principal Findings

In our study, the prevalence of NSSI in a sample of college students was 5.6%. Here, we examined the combination of chronotype and mental health as well as the interaction effect among BMI, self-esteem, and NSSI. BMI was positively correlated with the development of NSSI during adolescence, that is, the higher the BMI, the higher was the incidence of NSSI. Consistent with our hypothesis, our study provided evidence that self-esteem mediates the relationship between BMI and NSSI; this relationship was also moderated by chronotype. The relationship among BMI, self-esteem, and NSSI is complex and may involve a moderating effect of chronotype. Therefore, this study developed and evaluated a mediation-moderation model to elucidate the role of multilevel factors, including BMI, self-esteem, and chronotype, in relation to NSSI [31].

Comparison With Other Studies

With the change in social living environment, adolescents are facing increasing pressure and challenges, and adolescent NSSI has become a problem that cannot be ignored. NSSI is a common mental health problem among adolescents [32]. In a college student–related survey, approximately 1 in 5 (20.4%) students reported lifelong NSSI, and the frequency of persistent NSSI was estimated at 56.4%, with 15.6% of students reporting high-frequency repeat patterns (≥5 times per year) [33]. This high prevalence is worth considering, and NSSI is not only associated with suicidal behavior but is also a specific risk factor for suicide. In addition to the burden on individuals and families, including NSSI and self-harm with suicidal intent, the increased incidence of NSSI requires significant health care and economic costs and can lead to death in severe cases, further emphasizing the need to study why people should be concerned about NSSI [34]. Therefore, it is necessary to pay attention to the factors that affect NSSI; some factors have been explored [35-38], such as adverse life events, stressful life events, negative coping styles, problematic internet use, sleep disorders, traumatic experiences, problematic parent-child relationships, and mental health issues. In the same way that obesity causes physical diseases (eg, cardiovascular disease and diabetes), obesity has a significant triggering effect on the number of times of suicide attempts [39]. However, a population-based telephone survey of US adults conducted by the Centers for Disease Control and Prevention found that traditional suicide risk factors do not show any correlation with BMI; therefore, there is unlikely to be any relationship between BMI and a lower risk of suicide [18]. That said, in some cases, it is not entirely the case that a higher BMI is associated with a greater number of risk factors. Further research into the relationship between BMI and suicide may lead to new modifiable risk factors that can result in or prevent this key cause of death [18]. This also encourages us to look beyond the known causes of suicidal behavior, as there are other factors at play in psychological research.

In the aforementioned research, individual factors are very important, and self-esteem is an important individual factor [20]. Such studies might shed light on how obesity interacts with chronotype factors and can affect self-esteem in people with NSSI [40]. Regarding the association between obesity and suicide-related variables, 3 out of 5 studies investigating these suicide attempts showed a positive association between these attempts and obesity, and this relationship may be related to the reduction in self-worth and self-esteem caused by obesity-related stigma, suggesting that self-esteem plays a key role in BMI-induced suicide attempts [20, 41]; the results suggest that although low self-esteem and NSSI are correlated in both clinical and nonclinical populations, many factors influencing this relationship still exist [20], which needs to be confirmed by further research. According to the developmental psychopathological model, individuals' negative understanding and evaluation of themselves will lead to the emergence of NSSI, so low self-esteem is an important variable that triggers NSSI. Therefore, another study demonstrated the mediating

XSL•FO
RenderX

role of self-esteem in the relationship between BMI and depression or suicidality, as well as the moderating role of sex in the mediated pathways [20], which also provides an effective theoretical basis for our research. It is noteworthy that self-esteem is a crucial factor given its significance in depression and NSSI. This study, thus, converges with the existing literature and extends it. Moreover, we noticed that chronotype patterns could moderate these relationships, and sleep deprivation itself could influence an increase in psychobehavioral problems [42]. Another review concluded that people with bipolar disorder are characterized by extreme mood swings, high rates of suicide, sleep problems, and dysfunction in psychological characteristics, such as self-esteem (a feeling ranked below depression and above mania), which further provides a theoretical basis for understanding the relationship between sleep, self-esteem, and metabolism [43]. The current findings highlight the importance of assessing self-esteem in mediating BMI and NSSI, perhaps targeting self-esteem to mitigate self-esteem–related mental health problems (eg, NSSI and suicide). We also considered the effect of the moderator mental health on BMI and NSSI. The stigma associated with being overweight has increased in recent decades, which may have exacerbated the harmful effects of being overweight on mental health [44], thereby further elevating the occurrence of NSSI.

Another study explored the relationship between adolescents' BMI and mental health issues, including social phobia, depression, suicidal tendencies, and low self-esteem, and the mediating role of sex in participation in bullying [45]. Therefore, this provides a theoretical basis for our study that BMI is closely related to self-esteem, mental health, and suicidal behavior, and the correlation among the variables is also affected by other variables. According to the US Health and Retirement Study, baseline BMI, BMI transition, and BMI trajectory play significant moderating roles in the positive association between adverse childhood experience with new-onset all-cause dementia and Alzheimer disease [46]. The possible mechanisms include increased insulin and glucagon levels in the plasma and cerebrospinal fluid [47]. Additionally, our findings highlight the need for a multidisciplinary, collaborative, and integrated approach to optimize patient care. Different mechanisms have been identified, including gut-brain mechanisms, inflammatory responses (changes in the neutrophil-lymphocyte ratio, monocyte-lymphocyte ratio, and platelet-lymphocyte ratio), and changes in cytokine levels (interferon-γ, interleukin-1β, interleukin-6, interleukin-8, monocyte chemoattractant protein-1, tumor necrosis factor-α, and transforming growth factor-β1) [48].

Clinical Practices

Overweight and obese adolescents often experience prejudice due to the surrounding environment, which causes them to have low self-esteem, depression, and other negative emotions; engage in bad behaviors; cause self-injury; and harbor suicidal thoughts. As sociality is an essential attribute of adolescents, overweight and obese adolescents should communicate with others and obtain support, thereby effectively reducing the occurrence of NSSI. By further exploring the two variables of chronotype and self-esteem, this study aimed to explore the sociological pathway of BMI leading to NSSI.

Strengths and Limitations

The strengths of this study are its longitudinal design and large number of participants. The purpose of mediating is to strengthen the quasi-causal inference about the mechanism by which independent variables influence outcomes, whereas the purpose of moderating is to examine variables that influence the strength and direction of the predictor-outcome relationship. The use of moderated mediation could help identify the interdependence between the BMI and chronotypes. We also explored the role of the processes between early BMI and NSSI later in life.

However, this study has some limitations. First, although this was a follow-up study, only one follow-up survey was conducted, and changes in BMI were not evaluated during this process. These data are self-reported survey data, which can inevitably produce potential recall bias; therefore, we will further correct for the bias caused by recall bias in the future. Second, the study examined results in only one city; it is not clear how representative this sample is. Therefore, follow-up surveys are needed with samples from different regions and cultures across the country. Third, although this study mainly focuses on the relationship between BMI and NSSI under the influence of different factors, biological indicators will be included for further discussion in the future. Finally, we must consider critical attitudes toward the potential mechanism. As only some variables were evaluated in this study, additional variables with psychosocial factors should be considered in the future [48].

Conclusion

In this study, there was a combination effect of BMI, self-esteem, and NSSI, as well as an interaction effect among BMI, self-esteem, and NSSI. NSSI prevention programs should focus on evidence-based approaches relevant to public health and address them in conjunction with relevant influencing factors, including a wide range of interventions such as self-help programs, education, policy change, and legislation. Stronger inferences about the role of BMI in the etiology of adolescent NSSI can be made by exploring sex and ameliorating unfavorable mental health in prospective cohort studies. Future studies should expand upon our small subset of measured variables to paint a broader picture of adolescent NSSI in response to early-life social, psychological, and environmental stimuli.

Acknowledgments

We would like to acknowledge all hospital action teams, the staff and participants from the participating hospitals, and our cooperators for assistance in data collection. This research was funded by the China Postdoctoral Science Foundation (grant 2023M740022), the Research Fund of the Anhui Institute of Translational Medicine (grant 2023 hyx-C33), the National Natural

XSL•FO
RenderX

Science Foundation of China (grant 82370836), the Anhui Medical University Research Fund Project Youth Science Fund (grant 2022xkj307), the National Natural Science Foundation of China Youth Science 2018 Fund Training Program (grant 2022xkj307), The First Affiliated Hospital of Anhui Medical University (grant 2018kj20), and the Hunan Sanuo Diabetes Charity Foundation "Sweet Doctor Cultivation" project (grant 2021SD08).

Data Availability

The datasets generated for this study are available on request to the corresponding author.

Authors' Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work

Conflicts of Interest

None declared.

References

1. Nock MK. Self-injury. Annu Rev Clin Psychol 2010;6:339-363. [doi: 10.1146/annurev.clinpsy.121208.131258] [Medline: 20192787]
2. Knipe D, Padmanathan P, Newton-Howes G, Chan LF, Kapur N. Suicide and self-harm. Lancet 2022 May 14;399(10338):1903-1916. [doi: 10.1016/S0140-6736(22)00173-8] [Medline: 35512727]
3. De Luca L, Pastore M, Palladino BE, Reime B, Warth P, Menesini E. The development of non-suicidal self-injury (NSSI) during adolescence: a systematic review and Bayesian meta-analysis. J Affect Disord 2023 Oct 15;339:648-659. [doi: 10.1016/j.jad.2023.07.091] [Medline: 37479039]
4. Haregu T, Chen Q, Arafat SMY, Cherian A, Armstrong G. Prevalence, correlates and common methods of non-suicidal self-injury in South Asia: a systematic review. BMJ Open 2023 Nov 22;13(11):e074776. [doi: 10.1136/bmjopen-2023-074776] [Medline: 37993150]
5. Tilton-Weaver L, Latina D, Marshall SK. Trajectories of nonsuicidal self-injury during adolescence. J Adolesc 2023 Apr;95(3):437-453. [doi: 10.1002/jad.12126] [Medline: 36437557]
6. Adrian M, Zeman J, Erdley C, Whitlock K, Sim L. Trajectories of non-suicidal self-injury in adolescent girls following inpatient hospitalization. Clin Child Psychol Psychiatry 2019 Oct;24(4):831-846. [doi: 10.1177/1359104519839732] [Medline: 30947520]
7. Ribeiro JD, Franklin JC, Fox KR, et al. Self-injurious thoughts and behaviors as risk factors for future suicide ideation, attempts, and death: a meta-analysis of longitudinal studies. Psychol Med 2016 Jan;46(2):225-236. [doi: 10.1017/S0033291715001804] [Medline: 26370729]
8. Stefanovics EA, Grilo CM, Potenza MN, Pietrzak RH. Obesity in Latinx and White U.S. military veterans: mental health, psychosocial burden, non-suicidal self-injury and suicidal behavior. Psychiatry Res 2024 May;335:115844. [doi: 10.1016/j.psychres.2024.115844] [Medline: 38484606]
9. Hecht LM, Yeh HH, Braciszewski JM, et al. Weighing the association between BMI change and suicide mortality. Psychiatr Serv 2021 Aug 1;72(8):920-925. [doi: 10.1176/appi.ps.202000475] [Medline: 33882679]
10. Kim B, Kim HS, Park S, Kwon JA. BMI and perceived weight on suicide attempts in Korean adolescents: findings from the Korea youth risk behavior survey (KYRBS) 2020 to 2021. BMC Public Health 2023;23(1):1107. [doi: 10.1186/s12889-023-16058-z]
11. Geoffroy MC, Bouchard S, Per M, et al. Prevalence of suicidal ideation and self-harm behaviours in children aged 12 years and younger: a systematic review and meta-analysis. Lancet Psychiatry 2022 Sep;9(9):703-714. [doi: 10.1016/S2215-0366(22)00193-6] [Medline: 35907406]
12. Freitas-Rosa M, Gonçalves S, Antunes H. Is being overweight associated with engagement in self-injurious behaviours in adolescence, or do psychological factors have more "weight"? Eat Weight Disord 2016 Sep;21(3):493-500. [doi: 10.1007/s40519-015-0251-7] [Medline: 26753879]
13. BMI and suicide. Harv Ment Health Lett 2006;22(12).
14. Fitzpatrick S, Varma S, Ip J. The impact of homeostatic and circadian sleep processes on non-suicidal self-injury and suicide urges in borderline personality disorder. Arch Suicide Res 2022;26(3):1556-1571. [doi: 10.1080/13811118.2021.1932647] [Medline: 34348588]
15. Jagannath A, Taylor L, Wakaf Z, Vasudevan SR, Foster RG. The genetics of circadian rhythms, sleep and health. Hum Mol Genet 2017 Oct 1;26(R2):R128-R138. [doi: 10.1093/hmg/ddx240] [Medline: 28977444]

XSL•FO
RenderX

16. Walsh NA, Repa LM, Garland SN. Mindful larks and lonely owls: the relationship between chronotype, mental health, sleep quality, and social support in young adults. J Sleep Res 2022 Feb;31(1):e13442. [doi: 10.1111/jsr.13442] [Medline: 34272788]
17. Park C, Rosenblat JD, Brietzke E, et al. Stress, epigenetics and depression: a systematic review. Neurosci Biobehav Rev 2019 Jul;102:139-152. [doi: 10.1016/j.neubiorev.2019.04.010] [Medline: 31005627]
18. Mukamal KJ, Miller M. BMI and risk factors for suicide: why is BMI inversely related to suicide? Obesity (Silver Spring) 2009 Mar;17(3):532-538. [doi: 10.1038/oby.2008.538] [Medline: 19219065]
19. Wang S, Xu H, Zhang S, Wan Y, Tao F. Mediating effects of self-esteem in the relationship between childhood maltreatment and non-suicidal self-injury among adolescents: the roles of sex and only-child status. Soc Sci Med 2020 Feb 14;249(3):112847. [doi: 10.1016/j.socscimed.2020.112847] [Medline: 32086153]
20. Forrester RL, Slater H, Jomar K, Mitzman S, Taylor PJ. Self-esteem and non-suicidal self-injury in adulthood: a systematic review. J Affect Disord 2017 Oct 15;221:172-183. [doi: 10.1016/j.jad.2017.06.027] [Medline: 28647667]
21. Li S, Wang R, Thomas E, et al. Patterns of adverse childhood experiences and depressive symptom trajectories in young adults: a longitudinal study of college students in China. Front Psychiatry 2022;13:918092. [doi: 10.3389/fpsyt.2022.918092] [Medline: 35958653]
22. Song X, Wang S, Wang R, et al. Mediating effects of specific types of coping styles on the relationship between childhood maltreatment and depressive symptoms among Chinese undergraduates: the role of sex. Int J Environ Res Public Health 2020 Apr 30;17(9):3120. [doi: 10.3390/ijerph17093120] [Medline: 32365812]
23. Wang S, Xu H, Li S, Jiang Z, Wan Y. Sex differences in the determinants of suicide attempt among adolescents in China. Asian J Psychiatr 2020 Mar;49:101961. [doi: 10.1016/j.ajp.2020.101961] [Medline: 32078954]
24. Li S, Wang S, Gao X, et al. Patterns of adverse childhood experiences and suicidal behaviors in adolescents: a four-province study in China. J Affect Disord 2021 Apr;285:69-76. [doi: 10.1016/j.jad.2021.02.045]
25. Jiang Z, Xu H, Wang S, et al. Parent-child relationship quality, childhood maltreatment, and psychological symptoms in Chinese adolescent. J Interpers Violence 2022 Oct;37(19-20):18130-NP18151. [doi: 10.1177/08862605211035869] [Medline: 34330168]
26. Patterson F, Malone SK, Grandner MA, Lozano A, Perkett M, Hanlon A. Interactive effects of sleep duration and morning/evening preference on cardiovascular risk factors. Eur J Public Health 2018 Feb 1;28(1):155-161. [doi: 10.1093/eurpub/ckx029] [Medline: 28371850]
27. Rosenberg M. Rosenberg self-esteem scale (RSE). Acceptance and commitment therapy. Meas Pack 1965;61:52.
28. Yifu J, Xin Y. The self-esteem scale. Chin J Ment Health 1999;Supplement:318-320.
29. Wan Y, Chen R, Ma S, et al. Associations of adverse childhood experiences and social support with self-injurious behaviour and suicidality in adolescents. Br J Psychiatry 2019 Mar;214(3):146-152. [doi: 10.1192/bjp.2018.263] [Medline: 30477603]
30. Jiang T, Zhang Y, Dai F, Liu C, Hu H, Zhang Q. Advanced glycation end products and diabetes and other metabolic indicators. Diabetol Metab Syndr 2022 Jul 25;14(1):104. [doi: 10.1186/s13098-022-00873-2] [Medline: 35879776]
31. Zhang Y, Suo X, Zhang Y, et al. The relationship between negative life events and nonsuicidal self-injury among Chinese adolescents: a moderated-mediation model. Neuropsychiatr Dis Treat 2022;18:2881-2890. [doi: 10.2147/NDT.S386179] [Medline: 36540672]
32. Brown RC, Plener PL. Non-suicidal self-injury in adolescence. Curr Psychiatry Rep 2017 Mar;19(3):20. [doi: 10.1007/s11920-017-0767-9] [Medline: 28315191]
33. Kiekens G, Claes L, Hasking P, et al. A longitudinal investigation of non-suicidal self-injury persistence patterns, risk factors, and clinical outcomes during the college period. Psychol Med 2023 Oct;53(13):6011-6026. [doi: 10.1017/S0033291722003178] [Medline: 36325723]
34. Hepp J, Carpenter RW, Störkel LM, Schmitz SE, Schmahl C, Niedtfeld I. A systematic review of daily life studies on non-suicidal self-injury based on the four-function model. Clin Psychol Rev 2020 Dec;82:101888. [doi: 10.1016/j.cpr.2020.101888] [Medline: 32949907]
35. Herbison CE, Allen K, Robinson M, Newnham J, Pennell C. The impact of life stress on adult depression and anxiety is dependent on gender and timing of exposure. Dev Psychopathol 2017 Oct;29(4):1443-1454. [doi: 10.1017/S0954579417000372] [Medline: 28397629]
36. Fan YY, Liu J, Zeng YY, Conrad R, Tang YL. Factors associated with non-suicidal self-injury in Chinese adolescents: a meta-analysis. Front Psychiatry 2021;12:747031. [doi: 10.3389/fpsyt.2021.747031] [Medline: 34916971]
37. Baetens I, Greene D, Van Hove L, et al. Predictors and consequences of non-suicidal self-injury in relation to life, peer, and school factors. J Adolesc 2021 Jul;90:100-108. [doi: 10.1016/j.adolescence.2021.06.005] [Medline: 34182197]
38. Wang YJ, Li X, Ng CH, Xu DW, Hu S, Yuan TF. Risk factors for non-suicidal self-injury (NSSI) in adolescents: a meta-analysis. EClinMed 2022 Apr;46:101350. [doi: 10.1016/j.eclinm.2022.101350] [Medline: 35330803]
39. Powell-Wiley TM, Poirier P, Burke LE, et al. Obesity and cardiovascular disease: a scientific statement from the American Heart Association. Circulation 2021 May 25;143(21):e984-e1010. [doi: 10.1161/CIR.0000000000000973] [Medline: 33882682]

XSL•FO
RenderX

40. Chauhan S, Pandey R, Vakani K, Norbury R, Ettinger U, Kumari V. Sleep quality mediates the association between chronotype and mental health in young Indian adults. Npj Ment Health Res 2024 Jun 24;3(1):31. [doi: 10.1038/s44184-024-00076-9] [Medline: 38914742]
41. Ye Z, Xiong F, Li W. A meta-analysis of co-occurrence of non-suicidal self-injury and suicide attempt: Implications for clinical intervention and future diagnosis. Front Psychiatry 2022;13:976217. [doi: 10.3389/fpsyt.2022.976217] [Medline: 36032240]
42. Yeo SC, Jos AM, Erwin C, et al. Associations of sleep duration on school nights with self-rated health, overweight, and depression symptoms in adolescents: problems and possible solutions. Sleep Med 2019 Aug;60:96-108. [doi: 10.1016/j.sleep.2018.10.041] [Medline: 30611714]
43. Rantala MJ, Luoto S, Borráz-León JI, Krams I. Bipolar disorder: an evolutionary psychoneuroimmunological approach. Neurosci Biobehav Rev 2021 Mar;122:28-37. [doi: 10.1016/j.neubiorev.2020.12.031] [Medline: 33421542]
44. Daly M, Robinson E, Sutin AR. Perceived overweight and suicidality among US adolescents from 1999 to 2017. Int J Obes (Lond) 2020 Oct;44(10):2075-2079. [doi: 10.1038/s41366-020-0620-9] [Medline: 32546863]
45. Yen CF, Liu TL, Ko CH, Wu YY, Cheng CP. Mediating effects of bullying involvement on the relationship of body mass index with social phobia, depression, suicidality, and self-esteem and sex differences in adolescents in Taiwan. Child Abuse Negl 2014 Mar;38(3):517-526. [doi: 10.1016/j.chiabu.2013.07.015] [Medline: 24011767]
46. Ren Z, Su B, Du Y, Zhou T, Zheng X, Liu J. Effect modifications of BMI transition and trajectory in the associations of adverse childhood experiences with new-onset dementia and its subtypes in older US adults. Gen Psychiatr 2023;36(4):e101092. [doi: 10.1136/gpsych-2023-101092] [Medline: 37622031]
47. Bendix M, Uvnäs-Moberg K, Petersson M, Kaldo V, Åsberg M, Jokinen J. Insulin and glucagon in plasma and cerebrospinal fluid in suicide attempters and healthy controls. Psychoneuroendocrinology 2017 Jul;81:1-7. [doi: 10.1016/j.psyneuen.2017.03.019]
48. Ninla-Aesong P, Puangsri P, Kietdumrongwong P, Jongkrijak H, Noipha K. Being overweight and obese increases suicide risk, the severity of depression, and the inflammatory response in adolescents with major depressive disorders. Front Immunol 2023;14:1197775. [doi: 10.3389/fimmu.2023.1197775] [Medline: 38022570]

Abbreviations

NSSI: Nonsuicidal self-injury

Edited by A Mavragani; submitted 19.09.23; peer-reviewed by D Li, D Di, T Zhang, Y Liu; revised version received 28.08.24; accepted 30.08.24; published 21.02.25.

Please cite as:
Zhang Y, Ying R, Lu W, Liu X, Hu K, Feng Q, Yu Z, Wang Z, Lu F, Miao Y, Ma N, Tao F, Jiang T, Zhang Q
Association Among BMI, Self-Esteem, and Nonsuicidal Self-Injury in Young Adults to Understand the Influence of Socioenvironmental Factors: Longitudinal Study
JMIR Public Health Surveill 2025;11:e52928
URL: https://publichealth.jmir.org/2025/1/e52928
doi:10.2196/52928

© Yi Zhang, Ruixue Ying, Wan Lu, Xuemeng Liu, Keyan Hu, Qing Feng, Zixiang Yu, Zhen Wang, Fangting Lu, Yahu Miao, Nanzhen Ma, Fangbiao Tao, Tian Jiang, Qiu Zhang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Mpox Prevention Self-Efficacy and Associated Factors Among Men Who Have Sex With Men in China: Large Cross-Sectional Study

Quyige Gao[1,2,3*]; Shangbin Liu[2*], MPH; Muzaibaier Tuerxunjiang[3], BPH; Huifang Xu[3], BPH; Jiechen Zhang[1], PhD; Gang Xu[3], PhD; Jianyu Chen[4], PhD; Yong Cai[5], PhD; Fan Hu[6], PhD; Ying Wang[2,3], PhD

[1]Dermatology Department, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Public Health Research Center, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, 1111 Xianxia Road, Changning District, Shanghai, China

[3]School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[4]College of Public Health, Shanghai University of Medicine and Health Sciences, Shanghai, China

[5]Institute of Community Medical Research, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[6]Hongqiao International Institute of Medicine, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Ying Wang, PhD
Public Health Research Center, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, 1111 Xianxia Road, Changning District, Shanghai, China

Abstract

Background: Self-efficacy in mpox (formerly known as monkeypox) prevention plays a pivotal role in promoting preventive behaviors by fostering a sense of control and motivation, especially among men who have sex with men (MSM), the population most affected by mpox in many countries.

Objective: This study aims to assess the mpox prevention self-efficacy among MSM in China and identify factors influencing it, using a validated mpox prevention self-efficacy scale.

Methods: From October 2023 to March 2024, a nationwide cross-sectional study was conducted among MSM (aged ≥18 years) across 6 geographic regions in China using a snowball sampling method. The recruited participants (effective response rate=2403/2481, 96.9%) were asked to complete an anonymous questionnaire designed based on prior knowledge of mpox and social cognitive theory. The mpox prevention self-efficacy scale was evaluated for construct validity using exploratory factor analysis and confirmatory factor analysis, and its reliability was assessed using the Cronbach α coefficient. Univariate and multivariable logistic regression analyses were used to examine the factors associated with mpox prevention self-efficacy among MSM.

Results: A total of 2403 MSM participants were included, with a mean age of 29 (IQR 19 - 39) years. Of these, 1228 (51.1%) were aged 25 - 34 years, 1888 (78.6%) held a college degree or higher, and 2035 (84.7%) were unmarried. The median mpox prevention self-efficacy score was 23 (IQR 18 - 28). Exploratory factor analysis retained 6 items of the mpox prevention self-efficacy scale. Confirmatory factor analysis confirmed a strong model fit (χ^2 =32.1, n=1225; *P*<.001; comparative fit index=0.991; root mean square error of approximation=0.067; standardized root mean square residual=0.02; goodness-of-fit index=0.992; normed fit index=0.990; incremental fit index=0.991; Tucker-Lewis index=0.974), with all indices within acceptable ranges. The scale demonstrated good internal consistency, with a Cronbach α of 0.859. The positive factors associated with mpox prevention self-efficacy were mpox-related knowledge (OR 1.107, 95% CI 1.070 - 1.146), perceived risk awareness (OR 1.338, 95% CI 1.132 - 1.583), and mpox risk perception (OR 1.154, 95% CI 1.066 - 1.250), while the negative factor was age, with individuals aged 25 years and older exhibiting lower self-efficacy in mpox prevention (25 - 34 years: OR 0.789, 95% CI 0.642 - 0.970; 35 - 44 years: OR 0.572, 95% CI 0.444 - 0.736; 45 years and older: OR 0.569, 95% CI 0.394 - 0.823).

Conclusions: These findings highlight the critical role of targeted interventions to enhance mpox prevention self-efficacy, particularly through increasing knowledge, perceived risk awareness, and risk perception. Such interventions are especially important for middle-aged and older MSM, who may experience a decline in self-efficacy. Strengthening self-efficacy in these areas is essential for promoting sustained preventive behaviors, improving mental well-being, and contributing to more effective mpox prevention and control within the MSM community.

(JMIR Public Health Surveill 2025;11:e68400) doi:10.2196/68400

XSL•FO
RenderX

KEYWORDS

mpox; self-efficacy; men who have sex with men; MSM; monkeypox; cross-sectional study; mpox prevention self-efficacy scale

Introduction

Mpox (formerly known as monkeypox) is a viral illness caused by the *Monkeypox virus*, typically characterized by a skin rash or mucosal lesions lasting 2‐4 weeks. Other symptoms include headache, fever, back pain, muscle aches, fatigue, and swollen lymph nodes [1]. The World Health Organization declared the mpox outbreak a Public Health Emergency of International Concern twice: in July 2022 and again in August 2024 [2,3]. As of June 30, 2024, there were 99,176 confirmed cases and 208 deaths reported across 116 countries, including 2460 cases in China [4]. The outbreak predominantly affects men who have sex with men (MSM), who account for 85.8% of cases with known sexual behavior data [4,5]. China, with the world's largest MSM population—estimated at 8.3 million, nearly 1.8 times the size of the United States' MSM population—faces unique public health challenges in controlling mpox transmission [6]. In response, the National Disease Control Bureau of China, in collaboration with the National Health Commission, implemented a prevention and control strategy for mpox [7]. This includes educating high-risk populations, promoting self-protection among MSM, and encouraging timely medical consultations. Although the outbreak in China has been effectively controlled since September 2023, the high mobility and risk behaviors of the MSM population require continued vigilance and proactive measures [8].

According to social cognitive theory, self-efficacy, defined as the belief in one's ability to perform actions necessary to achieve desired outcomes [9], is a key factor in initiating and maintaining health-related behaviors [10]. Self-efficacy has been found to mediate the relationship between HIV self-testing provision and testing frequency among MSM [11], while it also directly influences the uptake of pre-exposure prophylaxis in this population [12], suggesting that enhancing self-efficacy may serve as an effective strategy to improve preventive health behaviors in at-risk populations. In the context of mpox prevention, self-efficacy is particularly critical for MSM, as it influences their confidence in adopting preventive measures, such as consistent condom use, vaccination, testing, and seeking timely medical care [13-16]. High self-efficacy has been associated with greater motivation to engage in mpox prevention behaviors and better adherence to preventive strategies, even among MSM who may continue to engage in high-risk behaviors despite awareness of mpox [17,18]. Given these findings, understanding mpox prevention self-efficacy in MSM is essential for informing targeted interventions and enhancing prevention strategies by identifying priority groups and potential barriers to mpox prevention and control.

A significant challenge in studying self-efficacy in mpox prevention is the lack of standardized measurement tools. Previous studies have primarily relied on single-item measures to assess the association between self-efficacy and mpox vaccination intentions [16,19] or used generalized instruments such as the General Self-Efficacy Scale or isolated questions [19,20]. Although these tools provide a general assessment, they fail to capture the nuanced dimensions of self-efficacy specific to mpox prevention. This limitation is particularly pronounced in the MSM population, where tailored assessments are essential to address unique behavioral and psychosocial factors. The lack of specificity and comprehensiveness in these existing measures not only undermines the validity of research findings, but also hampers the development of targeted interventions. Therefore, there is a clear need for tools specifically tailored to mpox prevention, integrating established scales with the unique characteristics of mpox.

Existing research indicates that the self-efficacy of MSM in preventing mpox may be influenced by multiple factors at various levels, including sociodemographic characteristics (eg, age, education level) [21], behavioral factors (eg, engagement in unsafe sexual practices) [22], and disease-related factors (eg, disease-related knowledge, risk perception) [23]. Research indicates that disease-related knowledge forms the basis for accurate health beliefs and enhances confidence in adopting safe sexual behaviors. A positive correlation has been observed between the level of disease-related knowledge and self-efficacy [24,25]. Additionally, mpox perceived risk awareness refers to individuals' recognition of the potential harm or likelihood of mpox in their surroundings [26,27], while risk perception pertains to the subjective judgment of one's personal susceptibility to mpox [28]. Evidence suggests that risk perception is a key trigger for both preventive awareness and action [29]. Along with disease-specific knowledge and awareness of environmental health risks, risk perception significantly influences mpox prevention awareness and behaviors [30]. However, the relationship between mpox perceived risk awareness, mpox risk perception, and self-efficacy, particularly among MSM, remains an underexplored area.

Given the critical role of self-efficacy in sustaining preventive motivation and goal-directed behavior among MSM, it is essential to assess it within the context of mpox prevention. While existing research on mpox prevention in MSM has primarily focused on vaccination willingness and testing behavior [15,31-33], there is a notable gap in understanding the self-efficacy of MSM in mpox prevention and the factors that influence it. Identifying these factors can help guide the development of targeted interventions. This study aims to evaluate mpox prevention self-efficacy among MSM in China, examining the impacts of sociodemographic characteristics, mpox-related knowledge, perceived risk awareness, and risk perception. By identifying these factors, the study seeks to provide a theoretical basis for designing more effective mpox prevention strategies tailored to this vulnerable population, while simultaneously contributing to improved behavioral outcomes and mental well-being within the MSM community.

XSL•FO
RenderX

Methods

Study Design

From October 2023 to March 2024, a nationwide cross-sectional study was conducted among MSM (age ≥18 years) across 6 geographic regions in China. Survey sites were located in the Northwest (Xinjiang Uyghur Autonomous Region), Northeast (Liaoning Province), Central (Shaanxi Province), Southwest (Yunnan Province), Southeast (Guangdong Province), and Eastern coastal region (Shanghai Municipality). Details of the program have been described previously [34].

Participant Recruitment and Data Collection

Eligibility criteria for participants included: (1) born male; (2) aged ≥18 years old; (3) have ever engaged in sex with men within the last six months; (4) primarily reside in one of the selected locations during the same period; and (5) agree to participate in the survey. Simultaneously, participants were excluded if they: (1) completed the questionnaire within 300 seconds; (2) failed the quality control questions (eg, Is Guangzhou, Shanghai, Beijing, or Shenzhen the capital city in China?); and (3) had an IP address indicating a location outside the target regions.

Participants were recruited from 6 cities in China with the assistance of local governmental centers for disease control and non-governmental organizations (NGOs), using a snowball sampling method. The NGOs routinely engaged with the MSM population in the region and were well-acquainted with local disease prevention and control initiatives. Initially, NGO staff, serving as primary investigators in each region, underwent training on recruitment procedures. They were tasked with recruiting 5‐10 "seed" participants (MSM individuals with good outreach potential) who met the inclusion criteria from each region. These seed participants were then asked to recruit additional participants from their social networks until the target sample size was achieved. Data collection was conducted entirely anonymously through an online survey. Participants accessed the electronic questionnaire via QR code scanning at designated survey locations. NGO staff assisted participants in completing the online survey, providing technical support and non-directive guidance as necessary, in accordance with the study protocol outlined previously [34].

After thorough verification, a total of 2481 questionnaires were collected, of which 2403 were deemed valid, yielding an effective response rate of 96.9%. The collected data underwent thorough cleaning and processing, including checks for missing data, outliers, and logical inconsistencies, with necessary adjustments and corrections made accordingly.

Ethical Considerations

The study was approved by Shanghai University of Medicine and Health Sciences (approval number 2023-MSMMPOX-22‐310222197604080237) on October 8, 2023. All participants signed electronic informed consent forms. Data collection was conducted anonymously using an electronic questionnaire administered through a secure, privacy-compliant platform (Wenjuanxing) via QR code scanning. To encourage participation, each individual who completed the survey and passed quality checks received Chinese yuan (CNY) 80 (about US $10.96). To ensure confidentiality, deidentified data were used for subsequent analysis.

Variables

Sociodemographic Characteristics

A range of sociodemographic characteristics (including age, education level, marital status, and monthly income) and disease diagnoses (including hypertension, diabetes, and hyperlipidemia) were assessed via self-report.

Mpox Prevention Self-Efficacy

Based on Bandura's self-efficacy theory and established scales [19], such as the HIV Prevention Self-Efficacy Scale [35] and the Self-Efficacy for HIV Prevention Behaviors Scale [36], we developed a 6-item mpox prevention self-efficacy scale tailored to the characteristics of MSM and mpox prevention. The scale aimed to assess individuals' confidence in their ability to effectively prevent mpox. Each item was rated on a 5-point Likert scale, ranging from "1 strongly disagree" to "5 strongly agree," as described in Table S1 in Multimedia Appendix 1. The composite variable was calculated by summing the scores of these items (range: 6‐30, Cronbach α=0.859), with a higher score indicating higher mpox prevention self-efficacy. A total score was categorized into 2 groups based on the median (median=23, skewness=−0.94): no/low self-efficacy (<23) and self-efficacy (≥23). When the independent variables are uncorrelated and exhibit a skewed distribution, the median is considered an effective tool for segmentation, which is the case for the variables in our study [37,38].

Mpox-Related Knowledge

Mpox-related knowledge included 12 yes/no questions designed to evaluate participants' understanding of mpox, covering various aspects such as pathogenesis (eg, "Mpox is a viral infectious disease"), epidemiological characteristics (eg, "Mpox can be transmitted through mucous membranes and broken skin"), clinical manifestations (eg, "The only symptom of mpox is a rash"), and preventive measures (eg, "Smallpox vaccination can be used to prevent mpox"). The overall score was calculated by counting the number of correct answers, with each correct response scoring 1 point. Higher scores indicated greater knowledge of mpox, with a possible range of 0‐12 (Cronbach α=0.80).

Mpox Perceived Risk Awareness

Perceived risk awareness, primarily utilized in public environmental health research, refers to the level of participants' attentiveness to potential risks in their environment [26,27]. We developed a composite variable for mpox perceived risk awareness based on 4 statements [28], each rated on a 5-point Likert scale (from "1=strongly agree" to "5=strongly disagree"). A composite scale score was calculated by summing the item scores, with higher scores indicating greater awareness of risks associated with mpox (range: 0‐20; Cronbach α=0.906). A total score was categorized into 2 groups based on the median: low perceived risk awareness (<17.00) and high perceived risk awareness (≥17.00).

XSL•FO
RenderX

Mpox Risk Perception

Risk perception refers to how people perceive the likelihood of experiencing harm or loss due to a disease and represents a subjective judgment of disease susceptibility [28]. It is a core construct in many health behavior theories, such as the Protection Motivation Theory and the Health Belief Model [39]. In the Health Belief Model, risk perception encompasses perceived susceptibility (belief in the likelihood of contracting a disease) and perceived severity (the perceived seriousness of the disease) [40]. In this study, considering the substantial variability in the clinical manifestations of mpox among individuals, mpox risk perception specifically focuses on participants' beliefs regarding the likelihood of contracting mpox. The statement "I believe I am someone who is likely to contract mpox" was used to assess personal mpox risk perception on a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree" [28].

Statistical Analysis

First, descriptive statistics were used to characterize the distribution of the variables. Continuous variables, which were non-normally distributed, were presented as medians and IQRs, while categorical variables were described using frequencies and percentages. Second, the content validity of the questionnaire was evaluated using the critical ratio method and Spearman correlation coefficient (N=2403). The total scores of the mpox prevention self-efficacy scale were ranked from low to high, with the bottom 27% classified as the low-score group and the top 27% as the high-score group. The discrimination of each item was tested using an independent samples *t* test between these 2 groups. Additionally, the total sample was randomly split into 2 subsets using a random number method: one subset (n=1178) was used for exploratory factor analysis (EFA), and the other (n=1225) for confirmatory factor analysis (CFA), to cross-validate the scale. For EFA, the feasibility of factor analysis was assessed using the Kaiser-Meyer-Olkin test and Bartlett test of sphericity. For CFA, model fit was assessed using several fit indices: the chi-square goodness-of-fit (χ^2/df), goodness-of-fit index (GFI), root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), normed fit index (NFI), comparative fit index (CFI), Tucker-Lewis index (TLI), and incremental fit index (IFI). The average variance extracted and composite reliability were also computed to assess the scale's convergent and discriminant validity. Third, univariate logistic regression was employed to examine the crude associations between the outcome variable and the variables of interest (ie, mpox-related and demographic variables). Fourth, a multivariable logistic regression analysis was performed to identify factors associated with mpox prevention self-efficacy, using the Enter method. Odds ratios (OR) and 95% CIs were estimated. Variance inflation factors were calculated to ensure the independence of each variable in the model, as described in Table S2 in Multimedia Appendix 2. The Hosmer-Lemeshow test indicated a good model fit (χ^2_8=6.047, *P*=.642). Descriptive, univariate, and multivariable analyses were conducted using SPSS software (version 22.0; IBM Corp). The reliability and validity of the questionnaire were assessed using AMOS software (version 28.0; IBM Corp). Forest plots were generated using GraphPad Prism software (version 9.5; GraphPad Software Inc).

Results

Demographic Characteristics

The median age of participants was 29 (IQR 19 - 39) years old. In the 30-point self-efficacy score, the median mpox prevention self-efficacy score was 23 (IQR 18 - 28). The majority held a college degree or higher (n=1888, 78.6%), were unmarried (n=2035, 84.7%), and did not have diabetes (n=2334, 97.1%), hypertension (n=2248, 93.6%), or hyperlipidemia (n=2181, 90.8%). Among the participants, those with a college education or higher (1017/1888, 53.9%) and those who were divorced or widowed (62/113, 54.9%) demonstrated higher levels of mpox prevention self-efficacy (Table 1).

XSL•FO
RenderX

Table . Background characteristics for all included MSM[a] in 6 Chinese cities according to mpox prevention self-efficacy (N=2403), from a survey conducted from October 2023 to March 2024.

Characteristic	Overall	No/low self-efficacy	Self-efficacy	*P* value
Age (years)				
Median (IQR)	29 (19 - 39)	30 (19 - 41)	28 (19 - 37)	<.001
18 - 24, n (%)	564 (23.5)	225 (20)	339 (26.5)	<.001
25 - 34, n (%)	1228 (51.1)	563 (50.1)	665 (52)	
35 - 44, n (%)	459 (19.1)	249 (22.2)	210 (16.4)	
≥45, n (%)	152 (6.3)	86 (7.7)	66 (5.2)	
Education level, n (%)				
Junior high school and below	157 (6.5)	85 (54.1)	72 (45.9)	.16
Senior high school	358 (14.9)	167 (46.7)	191 (53.4)	
College or above	1888 (78.6)	871 (46.1)	1017 (53.9)	
Marital status, n (%)				
Unmarried	2035 (84.7)	937 (46)	1098 (54)	.11
Married	255 (10.6)	135 (52.9)	120 (47.1)	
Divorced or widowed	113 (7.4)	51 (45.1)	62 (54.9)	
Income (CNY[b]), n (%)				
≤3000 (US $414)	502 (20.9)	254 (50.6)	248 (49.4)	.01
3001 - 6000 (US $414-$828)	851 (35.4)	403 (47.4)	448 (52.6)	
6001 - 12,000 (US $828-$1656)	802 (33.4)	340 (42.4)	462 (57.6)	
>12,000 (US $1656)	248 (10.3)	126 (50.8)	122 (49.2)	
Hypertension, n (%)				
No	2248 (93.6)	1049 (46.7)	1199 (53.3)	.80
Yes	155 (6.5)	74 (47.7)	81 (52.3)	
Diabetes, n (%)				
No	2334 (97.1)	1086 (46.5)	1248 (53.5)	.24
Yes	69 (2.9)	37 (53.6)	32 (46.4)	
Hyperlipemia, n (%)				
No	2181 (90.8)	991 (45.4)	1190 (54.6)	<.001
Yes	222 (9.2)	132 (59.5)	90 (40.5)	
Mpox-related knowledge, median (IQR)	8 (5-11)	8 (5-11)	9 (6 - 12)	<.001
Perceive risk awareness, n (%)				
Low	1173 (48.8)	609 (51.9)	564 (48.1)	<.001
High	1230 (52.2)	514 (41.8)	716 (58.2)	
Mpox risk perception, median (IQR)	3 (2-4)	3 (2-4)	4 (3-5)	<.001

[a]MSM: men who have sex with men.

[b]CNY: Chinese yuan. An exchange rate of CNY 1=US $0.14 has been applied.

XSL•FO
RenderX

Reliability and Validity Assessment of the Mpox Prevention Self-Efficacy Scale

Content Validity of the Scale

The results of the independent samples *t* test revealed statistically significant differences between the high-score group (top 27%) and the low-score group (bottom 27%) across all items (*P*<.001), as described in Table S1 in Multimedia Appendix 1. The critical ratio was 65.957, indicating that the scale possesses strong discriminatory power. Furthermore, Pearson correlation analysis demonstrated that the correlation coefficients between the total questionnaire score and the scores of the 6 items ranged from 0.746 to 0.817 (*P*<.001), suggesting that each item is highly representative of the overall construct.

Construct Validity of the Scale

The Kaiser-Meyer-Olkin measure of sampling adequacy was 0.817, and the Bartlett test yielded a χ^2_{15} value of 3442.5 (*P*<.001), indicating that the data were suitable for factor analysis. EFA of the baseline data extracted a single factor with an eigenvalue greater than 1.00, accounting for 60.02% of the total variance. The factor loadings for each item on the common factor ranged from 0.630 to 0.799 (Table 2).

Table . Factor loadings of the mpox prevention self-efficacy scale for men who have sex with men (n=1225).

Path	Standardized factor loadings	SE	*P* value	CR[a]	AVE[b]
C1_1<--Self_efficacy	0.630			0.845	0.478
C1_2<--Self_efficacy	0.640	0.043	<.001		
C1_3<--Self_efficacy	0.645	0.070	<.001		
C1_4<--Self_efficacy	0.664	0.058	<.001		
C1_5<--Self_efficacy	0.752	0.056	<.001		
C1_6<--Self_efficacy	0.799	0.055	<.001		

[a]CR: composite reliability.

[b]AVE: average variance extracted.

We then conducted a CFA on the baseline data to evaluate the global model fit of the factor structure suggested by the EFA, as described in Figure S1 in Multimedia Appendix 3. The overall model fit was good (χ^2_5=32.1, n=1225; *P*<.001; CFI=0.991; RMSEA=0.067; SRMR=0.02; GFI=0.992; NFI=0.990; IFI=0.991; TLI =0.974). All indices were within acceptable ranges, indicating that the model was well-aligned with the data (Table 3).

Table . Model fit index of the mpox prevention self-efficacy scale for men who have sex with men (n=1225).

Model fitting index	Chi-square (*df*)[a]	RMSEA[b]	SRMR[c]	GFI[d]	CFI[e]	NFI[f]	TLI[g]	IFI[h]
Reference value	<3.000	<0.080	<0.050	>0.900	>0.900	>0.900	>0.900	>0.900
Model value	6.416	0.067	0.0184	0.992	0.991	0.990	0.974	0.991

[a]Chi-square (*df*): χ^2 goodness-of-fit.

[b]RMSEA: root mean square error of approximation.

[c]SRMR: standardized root mean square residual.

[d]GFI: goodness-of-fit index.

[e]CFI: comparative fit index.

[f]NFI: normed fit index.

[g]TLI: Tucker-Lewis index.

[h]IFI: incremental fit index.

Reliability Analysis of the Scale

The Cronbach α coefficient for the mpox prevention self-efficacy scale was 0.859, indicating good internal consistency (Cronbach α>0.80).

Factors Associated With Mpox Prevention Self-Efficacy

Univariate Logistic Regression Analysis

Univariate analysis revealed that higher levels of mpox-related knowledge, greater perceived risk awareness, and increased mpox risk perception were significantly associated with higher mpox prevention self-efficacy. Conversely, older MSM demonstrated lower self-efficacy in mpox prevention. All these differences were statistically significant (Table 4).

XSL•FO
RenderX

Table . Univariate logistic regression analysis of mpox prevention self-efficacy among men who have sex with men in 6 Chinese cities (N=2403), from a survey conducted from October 2023 to March 2024.

Characteristic	Odds ratio	95% CI	*P* value
Age group, years (reference: 18 - 24)			
25 - 34	0.784	0.640 - 0.960	.02
35 - 44	0.560	0.436 - 0.718	<.001
≥45	0.509	0.355 - 0.732	<.001
Education level (reference: junior high school and below)			
Senior high school	1.350	0.927 - 1.967	.12
College or above	1.378	0.994 - 1.911	.05
Marital status (reference: unmarried)			
Married	0.759	0.584 - 0.985	.04
Divorced or widowed	1.037	0.709 - 1.518	.85
Income, CNY[a] (reference: ≤3000 [US $414])			
3001 - 6000 (US $414-$828)	1.139	0.913 - 1.420	.25
6001 - 12,000 (US $828-$1656)	1.392	1.112 - 1.741	.004
>12,000 (US $1656)	0.992	0.732 - 1.344	.96
Disease-related factor			
Mpox-related knowledge	1.124	1.088 - 1.163	<.001
Perceived risk awareness (high versus low)	1.338	1.132 - 1.583	.001
Mpox risk perception	1.157	1.070 - 1.252	<.001

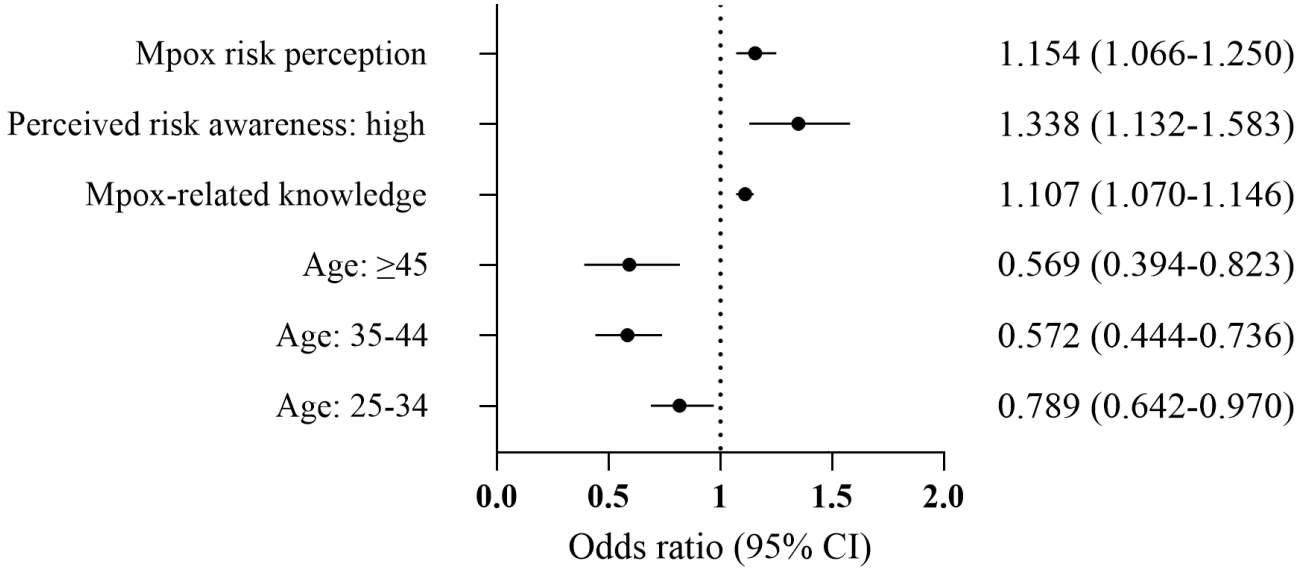
[a]CNY: Chinese yuan. The exchange rate used is 1 CNY =0.14USD.

Multivariable Logistic Regression Analysis

Variables with statistically significant differences from the univariate analysis were included in the multivariable logistic regression model (Figure 1). In this analysis (Y: mpox prevention self-efficacy, 0=no/low self-efficacy, 1=self-efficacy), the positive factors associated with mpox prevention self-efficacy were mpox-related knowledge (OR 1.107, 95% CI 1.070 - 1.146), perceived risk awareness (OR 1.338, 95% CI 1.132 - 1.583), and mpox risk perception (OR 1.154, 95% CI 1.066 - 1.250). Individuals over the age of 25 years may exhibit lower self-efficacy in mpox prevention (25 - 34 years: OR 0.789, 95% CI 0.642 - 0.970; 35 - 44 years: OR 0.572, 95% CI 0.444 - 0.736; 45 years and older: OR 0.569, 95% CI 0.394 - 0.823).

XSL•FO
RenderX

Figure 1. Multivariable logistic regression analysis of mpox prevention self-efficacy among MSM in 6 Chinese cities (N=2403), from a survey conducted from October 2023 to March 2024. The categories "18 - 24 years old" and "low perceived risk awareness" were used as references in the model. MSM: men who have sex with men.



Discussion

Principal Findings

This study represents one of the earliest investigations into mpox prevention self-efficacy and its influencing factors among MSM. Given that the mpox outbreak was once again declared a Public Health Emergency of International Concern in 2024, concerns about the potential spread of the virus—particularly within the MSM community—have intensified. This highlights the critical need for effective preventive strategies tailored to this high-risk group. Considering the crucial role of self-efficacy in health behaviors and disease prevention, our findings offer valuable public health insights for managing mpox outbreaks and enhancing the psychological and behavioral well-being of the MSM population.

Comparison With Prior Work

The study found that the median mpox prevention self-efficacy score was 23 (IQR 18 - 28), with a negatively skewed distribution of self-efficacy scores. This suggests that most participants demonstrated relatively high self-efficacy scores in preventing mpox, though a small proportion reported lower self-efficacy levels. These results indicate that, while the MSM community generally exhibits strong preparedness for mpox prevention, targeted interventions may be necessary for individuals with lower self-efficacy. Establishing and reinforcing high levels of mpox prevention self-efficacy is essential, as it plays a critical role in identifying at-risk populations and implementing targeted interventions [41,42]. Strengthening self-efficacy, particularly among those with lower initial scores, can improve both individual and community-level responses to prevention efforts, thereby contributing to more effective mpox control strategies. Therefore, public health programs should prioritize efforts to improve self-efficacy, focusing on education and risk awareness to increase confidence in preventive actions.

The study found that MSM with an education level of junior high school or below exhibited lower levels of self-efficacy. These findings are consistent with the demographic characteristics observed in previous studies of MSM with mpox [43,44]. Higher education levels may be associated with increased awareness of self-protection measures, which may contribute to higher self-efficacy and a reduction in risky behaviors [45]. This underscores the potential effectiveness of educational interventions aimed at enhancing self-efficacy, particularly for individuals with lower education levels.

An interesting finding of this study is that the proportion of married MSM exhibiting high self-efficacy in mpox prevention was lower compared to their unmarried, divorced, or widowed counterparts, despite the lack of statistical significance (P>.05). This trend may be explained by the fact that unmarried, divorced, or widowed MSM tend to receive less familial support, which may drive them to be more proactive in seeking preventive and self-protective measures to avoid infection [46]. Previous studies have also shown that divorced or widowed MSM tend to have higher levels of mpox-related knowledge compared to married individuals [47], possibly because the latter may be less inclined to seek out information or self-protective behaviors related to mpox [48]. This observation warrants further exploration, as it may suggest that marital status could still influence self-efficacy in mpox prevention, albeit not significantly in this study.

Accurate measurement of self-efficacy is fundamental for developing effective mpox prevention strategies [49]. In this study, the mpox prevention self-efficacy scale was validated and demonstrated strong reliability, validity, and practical utility. Although the General Self-Efficacy Scale, originally developed by Schwarzer and Jerusalem [50], is widely used due to its excellent internal consistency and test-retest reliability, it lacks the specificity required for assessing self-efficacy in the MSM population and in the context of mpox prevention, limiting its applicability in this context. Previous research on mpox predominantly has largely relied on isolated questions to assess self-efficacy, without employing a comprehensive or systematic scale tailored to the MSM community [51]. To fill this gap, this

XSL•FO
RenderX

study designed and evaluated the mpox prevention self-efficacy scale specifically tailored for MSM. The scale offers a more precise tool for assessing self-efficacy and its related factors among MSM, providing valuable insights for future research and interventions aimed at improving mpox prevention strategies.

Univariate analysis revealed a significant negative association between age and mpox prevention self-efficacy, a relationship that remained significant in the multivariable regression model. These findings indicate that older MSM may exhibit lower levels of self-efficacy in mpox prevention compared to their younger counterparts (aged 18 - 24 years). This association may be partially attributed to older MSM having reduced access to mpox-related information. Age has been identified as a key determinant of mpox-related knowledge [52], with older individuals potentially having less exposure to relevant health information, which may contribute to lower awareness, confidence, and self-efficacy in preventing mpox. In contrast, younger MSM, being more familiar with using the internet, are more likely to access mpox-related information. Furthermore, previous studies have shown that older MSM are less willing to change their sexual behavior compared to their younger counterparts [37]. Specifically, older MSM are more likely to maintain risky sexual behaviors even after becoming aware of the health risks associated with mpox, highlighting the need for tailored interventions aimed at this population to enhance their self-efficacy in preventing mpox. Therefore, age-specific communication strategies should be considered in mpox prevention efforts, particularly targeting middle-aged and older MSM through effective media channels. Given that MSM extensively use the internet for socializing, seeking sexual partners, and accessing sexual health information [53,54], eHealth technologies have emerged as promising tools for delivering prevention interventions to this population [55]. Research has shown that enhanced behavioral interventions, such as tailored strategies, can lead to reductions in unprotected sex, while advanced eHealth technologies, including social media platforms, are associated with increased health screening in the MSM community [56]. Therefore, eHealth-based interventions, incorporating targeted communication strategies and educational tools, can be effectively implemented on widely used social platforms or websites to enhance mpox prevention self-efficacy, particularly in populations with lower levels of self-efficacy. Tailored interventions focusing on improving self-efficacy among older and less educated MSM may be particularly beneficial, as these groups tend to exhibit lower self-efficacy, which may increase their susceptibility to infection. These findings underscore the importance of demographic factors in informing targeted mpox prevention strategies, especially in guiding surveillance and health education practices.

This study found that mpox-related knowledge, perceived risk awareness, and risk perception were positively correlated with self-efficacy in mpox prevention. Previous research has demonstrated that attitudes towards a disease are closely linked to disease-related knowledge, suggesting that lower levels of knowledge among MSM are associated with more negative attitudes and decreased intentions to engage in mpox prevention behaviors [57]. Mpox perceived risk awareness reflects beliefs about the potential harm or likelihood of mpox in their surroundings, while risk perception represents a subjective judgment of personal susceptibility to the disease. According to social cognitive theory, individuals who recognize the severity of a negative health condition, believe that it can be avoided, maintain a positive attitude toward recommended actions, and have confidence in their ability to successfully execute these actions are more likely to engage in health-related behaviors [22,58]. Therefore, mpox prevention efforts among MSM should include further public health education and training to enhance perceived risk awareness, provide targeted risk communication, and ensure the privacy and confidentiality of services, which are crucial for effective mpox prevention within this population.

This research has several strengths. First, as one of the earliest studies in China examining mpox prevention self-efficacy among MSM, the findings carry significant public health implications for the prevention and control of mpox outbreaks in the country, while also enhancing the psychological and behavioral well-being of the MSM community. Furthermore, the study's multicenter design contributes to the generalizability of the results, as it incorporates a diverse sample from various regions across China.

Limitations

Despite the valuable insights provided by the study, several limitations must be acknowledged. First, participant recruitment was facilitated by government centers for disease control departments and NGOs using a snowball sampling method, which was used to improve the efficiency of identifying and recruiting members of the hidden MSM population [59]. Although this approach is effective for accessing hard-to-reach groups, it inherently introduces selection bias, as the sample composition is influenced by the characteristics of initial recruits (seeds) and the subsequent limited recruitment chains. Consequently, this may restrict the generalizability of our findings. Second, the sensitive nature of some inquiries, combined with self-reported data, may lead to information bias. Third, the cross-sectional design restricts the ability to infer causality, highlighting the need for future longitudinal studies to further explore the temporal relationships between mpox-related knowledge, perceived risk awareness, risk perception, and self-efficacy. Such prospective research could also evaluate the impact of targeted interventions aimed at enhancing self-efficacy in mpox prevention among MSM.

Conclusion

This study examined the mpox prevention self-efficacy of MSM and its influencing factors, based on the design and evaluation of the mpox prevention self-efficacy scale. The findings demonstrated significant associations between mpox prevention self-efficacy and sociodemographic characteristics, as well as mpox-related factors. These results underscore the importance of targeted interventions to enhance mpox prevention self-efficacy, particularly through efforts to increase mpox-related knowledge, perceived risk awareness, and risk perception. Special attention should be given to middle-aged and older MSM, who may experience a decline in self-efficacy. Future research should adopt longitudinal designs to better

XSL•FO
RenderX

understand the temporal relationships between self-efficacy and related factors, thereby providing more robust evidence to guide the development of effective mpox prevention strategies. Ultimately, these efforts are crucial for improving mpox prevention outcomes and promoting the overall health of the MSM community.

Acknowledgments

The authors would like to express their gratitude to Shanghai Qingai Health Center (Shanghai), Trans China (Yunnan), Consultation Center of AIDS Aid and Health Service (Liaoning), Xinjiang Rong'ai Health Consulting Co Ltd (Xinjiang), Lingnan Partners Community Support Center (Guangdong), and Xi'an Zhitong AIDS Prevention and Health Center (Shaanxi) for their contributions to data collection. We also extend our thanks to all the participants involved in this study. This study was funded by Key discipline projects of Shanghai Three-Year Action Plan for Public Health under Grant (GWVI-11.1-29), Public Health Research Project from Shanghai Disease Control and Prevention Administration (2024GKM30), Changning District Medical and Health Research Special Project (CNKW2022Y08), Tongren Hospital Talent Launch Project (TR2023rc11, TR2024RC10), Science and Technology Commission Shanghai Municipality (number 20JC1410204) for the Seroepidemiological Study of Novel Coronavirus Pneumonia in Key Populations, national health commission of the PRC (grant number Z155080000004), and the Research Fund of Shanghai Tongren Hospital, Shanghai Jiaotong University School of Medicine (TRYJ2024LC13, TRYJ2022LC10).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy and ethical considerations, but are available from the corresponding author on reasonable request.

Authors' Contributions

QG and SL performed the analyses, interpreted the data, and developed the manuscript; MT and HX conceptualized and validated data analyses; JZ and GX contributed to the data collection and provided critical insights into the methodology; JC reviewed and edited the manuscript; YC, FH, and YW designed and supervised this study, and reviewed and edited the manuscript. All authors have made a significant contribution to this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Independent samples *t* test discrimination results of the mpox prevention self-efficacy scale (N=2403).
[DOC File, 18 KB - publichealth_v11i1e68400_app1.doc]

Multimedia Appendix 2
Results of collinearity diagnosis for the multivariable logistic regression analysis of mpox prevention self-efficacy among MSM in 6 Chinese cities (N=2403), from a survey conducted from October 2023 to March 2024.
[DOC File, 18 KB - publichealth_v11i1e68400_app2.doc]

Multimedia Appendix 3
Standardized one-factor structural model of the mpox prevention self-efficacy scale (n=1225).
[PNG File, 101 KB - publichealth_v11i1e68400_app3.png]

References

1. Lu J, Xing H, Wang C, et al. Mpox (formerly monkeypox): pathogenesis, prevention, and treatment. Signal Transduct Target Ther 2023 Dec 27;8(1):458. [doi: 10.1038/s41392-023-01675-2] [Medline: 38148355]
2. Mpox (monkeypox) outbreak. World Health Organization. 2022. URL: https://www.who.int/emergencies/situations/monkeypox-oubreak-2022 [accessed 2024-08-20]
3. WHO Director-General's opening remarks at the media briefing following the IHR emergency committee meeting on the upsurge of mpox 2024-14. World Health Organization. 2024. URL: https://www.who.int/zh/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-following-the-ihr-emergency-committee-meeting-on-the-upsurge-of-mpox-2024—14-august-2024 [accessed 2024-08-20]
4. 2022-24 mpox (monkeypox) outbreak: global trends. World Health Organization. 2024. URL: https://worldhealthorg.shinyapps.io/mpx_global/ [accessed 2024-08-20]

XSL•FO
RenderX

5. Vaughan AM, Cenciarelli O, Colombe S, et al. A large multi-country outbreak of monkeypox across 41 countries in the WHO European Region, 7 March to 23 August 2022. Euro Surveill 2022 Sep;27(36):2200620. [doi: 10.2807/1560-7917.ES.2022.27.36.2200620] [Medline: 36082686]
6. Hu M, Xu C, Wang J. Spatiotemporal analysis of men who have sex with men in Mainland China: social app capture-recapture method. JMIR mHealth uHealth 2020 Jan 24;8(1):e14800. [doi: 10.2196/14800] [Medline: 32012086]
7. Notice of the General Office of the National Health Commission on the issuance of the Technical Guidelines for Monkeypox Prevention and Control (2022 Edition) [Website in Chinese]. Gov.cn. 2022. URL: https://www.gov.cn/zhengce/zhengceku/2022-07/01/content_5698880.htm [accessed 2024-09-22]
8. Chen HJ, Tang HL, Li PL, et al. Survey on the awareness rate of mpox knowledge and related factors among men who have sex with men in China [Article in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2024 Apr 10;45(4):559-565. [doi: 10.3760/cma.j.cn112338-20231030-00257] [Medline: 38678353]
9. Bandura A. Self-efficacy mechanism in human agency. American Psychologist 1982;37(2):122-147. [doi: 10.1037//0003-066X.37.2.122]
10. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Advances in Behaviour Research and Therapy 1978 Jan;1(4):139-161. [doi: 10.1016/0146-6402(78)90002-4]
11. Zhang C, Huang D, Goldsamt LA, et al. Self-efficacy mediates the effect of providing HIV self-testing kits on HIV testing frequency among Chinese men who have sex with men: results from a randomised controlled trial. Sex Health 2023 Apr;20(2):118-125. [doi: 10.1071/SH22110] [Medline: 36848626]
12. Liu H, Lai G, Shi G, Zhong X. The influencing factors of HIV-preventive behavior based on health belief model among HIV-negative MSMs in Western China: a structural equation modeling analysis. Int J Environ Res Public Health 2022 Aug 17;19(16):10185. [doi: 10.3390/ijerph191610185] [Medline: 36011822]
13. DiIorio C, Dudley WN, Kelly M, Soet JE, Mbwara J, Sharpe Potter J. Social cognitive correlates of sexual experience and condom use among 13- through 15-year-old adolescents. J Adolesc Health 2001 Sep;29(3):208-216. [doi: 10.1016/s1054-139x(00)00200-7] [Medline: 11524220]
14. Safren SA, Blashill AJ, Lee JS, et al. Condom-use self-efficacy as a mediator between syndemics and condomless sex in men who have sex with men (MSM). Health Psychol 2018 Sep;37(9):820-827. [doi: 10.1037/hea0000617] [Medline: 29927272]
15. Luo S, Jiao K, Zhang Y, et al. Behavioral intention of receiving monkeypox vaccination and undergoing monkeypox testing and the associated factors among young men who have sex with men in China: large cross-sectional study. JMIR Public Health Surveill 2024 Mar 19;10:e47165. [doi: 10.2196/47165] [Medline: 38502181]
16. Lawrence A. Assessing vaccine intentions, knowledge, self-efficacy, and trust: a cross-sectional study on perceptions of monkeypox vaccination and public health risk awareness in Makurdi, Benue State, Nigeria. Cureus 2024 Oct;16(10):e72131. [doi: 10.7759/cureus.72131] [Medline: 39440166]
17. Tan F, Oka P, Dambha-Miller H, Tan NC. The association between self-efficacy and self-care in essential hypertension: a systematic review. BMC Fam Pract 2021 Feb 22;22(1):44. [doi: 10.1186/s12875-021-01391-2] [Medline: 33618661]
18. Wang SD, Loftus P, Pang RD, Kirkpatrick MG. Impact of self-efficacy on daily intention to not smoke. Addict Behav 2021 Jul;118:106877. [doi: 10.1016/j.addbeh.2021.106877] [Medline: 33714032]
19. Winters M, Malik AA, Omer SB. Attitudes towards Monkeypox vaccination and predictors of vaccination intentions among the US general public. PLoS One 2022;17(12):e0278622. [doi: 10.1371/journal.pone.0278622] [Medline: 36454991]
20. Lee J, Lee EH, Chae D. Self-efficacy instruments for type 2 diabetes self-care: a systematic review of measurement properties. J Adv Nurs 2020 Aug;76(8):2046-2059. [doi: 10.1111/jan.14411] [Medline: 32400902]
21. Takahashi LM, Magalong MG, Debell P, Fasudhani A. HIV and AIDS in suburban Asian and Pacific Islander communities: factors influencing self-efficacy in HIV risk reduction. AIDS Educ Prev 2006 Dec;18(6):529-545. [doi: 10.1521/aeap.2006.18.6.529] [Medline: 17166079]
22. Khumsaen N, Stephenson R. Beliefs and perception about HIV/AIDS, self-efficacy, and HIV sexual risk behaviors among young Thai men who have sex with men. AIDS Educ Prev 2017 Apr;29(2):175-190. [doi: 10.1521/aeap.2017.29.2.175] [Medline: 28467158]
23. Villegas Rodríguez N, Ferrer Lagunas LM, Cianelli Acosta R, Miner S, Lara Campos L, Peragallo N. Knowledge and self efficacy related with HIV prevention among Chilean women. Invest Educ Enferm 2011;29(2):222-229. [Medline: 25284914]
24. Carter G, Brown L, Mahnke B, Ohmit A, Woodward B. Sexual communication self-efficacy and PrEP literacy helps to meet HIV prevention outcomes among Black and Latinx individuals. J Prim Care Community Health 2022;13:21501319221099789. [doi: 10.1177/21501319221099789] [Medline: 35578769]
25. Black DS, Sun P, Rohrbach LA, Sussman S. Decision-making style and gender moderation of the self-efficacy-condom use link among adolescents and young adults: informing targeted STI/HIV prevention programs. Arch Pediatr Adolesc Med 2011 Apr;165(4):320-325. [doi: 10.1001/archpediatrics.2011.17] [Medline: 21464380]
26. Lavallee S, Hynds PD, Brown RS, Majury A. Classification of sub-populations for quantitative risk assessment based on awareness and perception: a cross-sectional population study of private well users in Ontario. Sci Total Environ 2023 Jan 20;857(Pt 3):159677. [doi: 10.1016/j.scitotenv.2022.159677] [Medline: 36302430]

XSL•FO
RenderX

27. Gray-Scholz D, Haney TJ, MacQuarrie P. Out of sight, out of mind? Geographic and social predictors of flood risk awareness. Risk Anal 2019 Nov;39(11):2543-2558. [doi: 10.1111/risa.13357] [Medline: 31173665]
28. Jiao K, Xu Y, Huang S, et al. Mpox risk perception and associated factors among Chinese young men who have sex with men: Results from a large cross-sectional survey. J Med Virol 2023 Aug;95(8):e29057. [doi: 10.1002/jmv.29057] [Medline: 37635623]
29. Wiedemann PM, Schütz H. The precautionary principle and risk perception: experimental studies in the EMF area. Environ Health Perspect 2005 Apr;113(4):402-405. [doi: 10.1289/ehp.7538] [Medline: 15811829]
30. Renner B, Schupp H, Vollmann M, Hartung FM, Schmälzle R, Panzer M. Risk perception, risk communication and health behavior change. Zeitschrift für Gesundheitspsychologie 2008 Jul;16(3):150-153. [doi: 10.1026/0943-8149.16.3.150]
31. Li Y, Peng X, Fu L, et al. Monkeypox awareness and low vaccination hesitancy among men who have sex with men in China. J Med Virol 2023 Feb;95(2):e28567 [FREE Full text] [doi: 10.1002/jmv.28567]
32. Huang X, Lin Z, Qin J, et al. Willingness to accept monkeypox vaccine and its correlates among men who have sex with men in Southern China: a web-based online cross-sectional study. Front Public Health 2024;12:1289918. [doi: 10.3389/fpubh.2024.1289918] [Medline: 38384873]
33. Zheng M, Qin C, Qian X, et al. Knowledge and vaccination acceptance toward the human monkeypox among men who have sex with men in China. Front Public Health 2022;10:997637. [doi: 10.3389/fpubh.2022.997637] [Medline: 36388271]
34. Gao Y, Liu S, Xu H, et al. The Mpox vaccine hesitancy scale for Mpox: links with vaccination intention among men who have sex with men in six cities of China. Vaccines (Basel) 2024 Sep 3;12(9):1009. [doi: 10.3390/vaccines12091009] [Medline: 39340039]
35. Ravelo GJ, Sanchez M, Cyrus E, De La Rosa M, Peragallo N, Rojas P. Associations between gender norms and HIV self-efficacy among Latina immigrants in a farmworker community. Ethn Health 2022 Jan;27(1):27-39. [doi: 10.1080/13557858.2019.1659234] [Medline: 31450962]
36. Cianelli R, Ferrer L, Norr KF, et al. Mano a Mano-Mujer: an effective HIV prevention intervention for Chilean women. Health Care Women Int 2012;33(4):321-341. [doi: 10.1080/07399332.2012.655388] [Medline: 22420675]
37. Chan ZYS, Chong SY, Niaupari S, et al. Receptiveness to monkeypox vaccines and public health communication strategies among gay, bisexual and other men who have sex with men in Singapore: cross-sectional quantitative and qualitative insights. Sex Transm Infect 2024 Aug 19;100(6):362-367. [doi: 10.1136/sextrans-2024-056230] [Medline: 38821878]
38. Bien-Gund CH, Shaw PA, Agnew-Brune C, et al. HIV self-testing and risk behaviors among men who have sex with men in 23 US cities, 2017. JAMA Netw Open 2022 Dec 1;5(12):e2247540. [doi: 10.1001/jamanetworkopen.2022.47540] [Medline: 36534398]
39. Gellman M, Turner J. Encyclopedia of Behavioral Medicine 2013. [doi: 10.1007/978-1-4419-1005-9]
40. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
41. Villegas N, Cianelli R, Gonzalez-Guarda R, Kaelber L, Ferrer L, Peragallo N. Predictors of self-efficacy for HIV prevention among Hispanic women in South Florida. J Assoc Nurses AIDS Care 2013;24(1):27-37. [doi: 10.1016/j.jana.2012.03.004] [Medline: 22795758]
42. Sheeran P, Maki A, Montanaro E, et al. The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: a meta-analysis. Health Psychol 2016 Nov;35(11):1178-1188. [doi: 10.1037/hea0000387] [Medline: 27280365]
43. Du M, Zheng M, Liu J, et al. The prevalence of mpox and its association with sexual behavior among Chinese men who have sex with men in early August 2023. J Med Virol 2023 Dec;95(12):e29320. [doi: 10.1002/jmv.29320] [Medline: 38102917]
44. Zheng M, Du M, Yang G, et al. Lower rate of mpox vaccination hesitancy and medical consultation among Chinese men who have sex with men living with HIV in comparison with those living without HIV: a national observational study. Hum Vaccin Immunother 2023 Dec 15;19(3):2290788. [doi: 10.1080/21645515.2023.2290788] [Medline: 38054460]
45. Wei L, Tian J, Guo M, et al. Trajectories of sexual risk behaviors and the associated factors among young men who have sex with men in China. Front Public Health 2022;10:854616. [doi: 10.3389/fpubh.2022.854616] [Medline: 35387185]
46. Wu S, Zhu S, Yan X, et al. Willingness to receive the COVID-19 vaccine among HIV positive men who have sex with men in China: a cross-sectional study. BMC Public Health 2023 Jan 10;23(1):64. [doi: 10.1186/s12889-022-14961-5] [Medline: 36627582]
47. Zheng M, Chen W, Qian X, et al. Awareness of mpox-related knowledge among men who have sex with men in China. BMC Public Health 2023 Mar 30;23(1):600. [doi: 10.1186/s12889-023-15503-3] [Medline: 36997990]
48. Liu Z, Chen Y, Yao T, et al. Factors related to HIV testing frequency in MSM based on the 2011-2018 survey in Tianjin, China: a hint for risk reduction strategy. BMC Public Health 2021 Oct 20;21(1):1900. [doi: 10.1186/s12889-021-11948-6] [Medline: 34670542]
49. Fisher WA, Fisher JD, Harman JJ. The information - motivation - behavioral skills model: a general social psychological approach to understanding and promoting health behavior. In: Social Psychological Foundations of Health and Illness 2009. [doi: 10.1002/9780470753552]

XSL•FO
RenderX

50. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs: NFER-NELSON; 1995. URL: https://www.religiousforums.com/data/attachment-files/2014/12/22334_285ebbd019ec89856493442b1b6f9154.pdf [accessed 2025-02-11]
51. Bakare D, Salako J, Sogbesan A, Olojede OE, Bakare AA. Assessment of the level of awareness, knowledge, and risk perception of community members about mpox infection in Nigeria. Ann Ib Postgrad Med 2024 Apr 30;22(1):76-87. [Medline: 38939883]
52. Jairoun AA, Al-Hemyari SS, Abdulla NM, et al. Awareness and preparedness of human monkeypox outbreak among university student: Time to worry or one to ignore? J Infect Public Health 2022 Oct;15(10):1065-1071. [doi: 10.1016/j.jiph.2022.08.015] [Medline: 36087548]
53. Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. AIDS Behav 2014 Feb;18(2):285-296. [doi: 10.1007/s10461-013-0671-1] [Medline: 24292281]
54. Zou H, Fan S. Characteristics of men who have sex with men who use smartphone geosocial networking applications and implications for HIV interventions: a systematic review and meta-analysis. Arch Sex Behav 2017 May;46(4):885-894. [doi: 10.1007/s10508-016-0709-3] [Medline: 27043837]
55. Carpenter KM, Stoner SA, Mikko AN, Dhanak LP, Parsons JT. Efficacy of a web-based intervention to reduce sexual risk in men who have sex with men. AIDS Behav 2010 Jun;14(3):549-557. [doi: 10.1007/s10461-009-9578-2] [Medline: 19499321]
56. Xin M, Viswanath K, Li AYC, et al. The effectiveness of electronic health interventions for promoting HIV-preventive behaviors among men who have sex with men: meta-analysis based on an integrative framework of design and implementation features. J Med Internet Res 2020 May 25;22(5):e15977. [doi: 10.2196/15977] [Medline: 32449685]
57. Opoku MP, Agyei-Okyere E, Nketsia W, Torgbenu EL, Kumi EO. Perceived self-efficacy of students and its influence on attitudes and knowledge about HIV/AIDS in Ghana. Int J Health Plann Manage 2022 Mar;37(2):755-769. [doi: 10.1002/hpm.3371] [Medline: 34697830]
58. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q 1988;15(2):175-183. [doi: 10.1177/109019818801500203] [Medline: 3378902]
59. Kendall C, Kerr L, Gondim RC, et al. An empirical comparison of respondent-driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil. AIDS Behav 2008 Jul;12(4 Suppl):S97-104. [doi: 10.1007/s10461-008-9390-4] [Medline: 18389357]

Abbreviations

CFA: confirmatory factor analysis
CFI: comparative fit index
EFA: exploratory factor analysis
GFI: goodness-of-fit index
IFI: incremental fit index
MSM: men who have sex with men
NFI: normed fit index
NGO: nongovernmental organization
OR: odds ratio
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
TLI: Tucker-Lewis index

Edited by A Mavragani; submitted 07.11.24; peer-reviewed by B Foroutan, O Oni, S Luo, TF Ojo; revised version received 05.01.25; accepted 05.01.25; published 28.02.25.

Please cite as:
Gao Q, Liu S, Tuerxunjiang M, Xu H, Zhang J, Xu G, Chen J, Cai Y, Hu F, Wang Y
Mpox Prevention Self-Efficacy and Associated Factors Among Men Who Have Sex With Men in China: Large Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e68400
URL: https://publichealth.jmir.org/2025/1/e68400
doi:10.2196/68400

© Quyige Gao, Shangbin Liu, Muzaibaier Tuerxunjiang, Huifang Xu, Jiechen Zhang, Gang Xu, Jianyu Chen, Yong Cai, Fan Hu, Ying Wang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.2.2025. This is an

XSL•FO
RenderX

open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Preferred Features of Influenza Vaccination Programs Among Chinese Clinicians Practicing Traditional Chinese Medicine and Western Medicine: Discrete Choice Experiment

Liuren Zhang[1,2*], MPH; Linchen Chu[1*], MPH; Maria E Sundaram[3*], PhD; Yi Zhou[1], MPH; Xiu Sun[4], MPH; Zheng Wei[5], MPH; Chuanxi Fu[1], PhD

[1]Institute of Infectious Disease and Vaccine, School of Public Health, Zhejiang Chinese Medical University, Hangzhou, China

[2]The Third People's Hospital of Hubei Province, Wuhan, China

[3]Center for Clinical Epidemiology and Population Health, Marshfield Clinic Research Institute, Marshfield, WI, United States

[4]The Third Affiliated Hospital of Zhejiang Chinese Medical University, Hangzhou, China

[5]The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Chuanxi Fu, PhD
Institute of Infectious Disease and Vaccine, School of Public Health, Zhejiang Chinese Medical University, Hangzhou, China

Abstract

Background: Achieving high vaccine coverage among clinicians is crucial to curb the spread of influenza. Traditional Chinese medicine (TCM), rooted in cultural symbols and concepts without direct parallels in modern Western medicine, may influence perspectives on vaccination. Therefore, understanding the preferences of TCM clinicians towards influenza vaccines is of great importance.

Objective: To understand preferences for features of influenza vaccination programs and identify the optimal influenza vaccination program among clinicians practicing TCM and Western medicine.

Methods: We conducted a discrete choice experiment with a national sample of 3085 Chinese clinicians from various hospital levels (n=1013 practicing TCM) from January to May 2022. Simulations from choice models using the experimental data generated the coefficients of preference and predicted the uptake rate of different influenza vaccination programs. Clinicians were grouped by vaccine preference classification through a latent class analysis.

Results: All included attributes significantly influenced clinicians' preferences for choosing an influenza vaccination program. An approximate hypothetical 60% increase of vaccine uptake could be obtained when the attitude of the workplace changed from "no notice" to "encouraging of vaccination"; there was an approximate hypothetical 35% increase of vaccine uptake when vaccination campaign strategies changed from "individual appointment" to "vaccination in a workplace setting." In the entire sample, about 30% (946/3085) of clinicians preferred free vaccinations, while 26.5% (819/3085) comprehensively considered all attributes, except vaccination campaign strategies, when making a decision about choosing an influenza vaccination program. Clinicians who practiced TCM, worked in tertiary hospital, or had at least a postgraduate degree exhibited a lower preference for free vaccinations. Clinicians who practiced Western medicine, worked in primary hospital, or had at most a bachelor's degree had a higher preference for vaccinations in workplace settings.

Conclusions: Offering a range of influenza vaccination programs targeting the preferred attributes of different clinician groups could potentially encourage more clinicians, including those practicing TCM, to participate in influenza vaccination programs.

(JMIR Public Health Surveill 2025;11:e63314) doi:10.2196/63314

KEYWORDS

influenza vaccination program; traditional Chinese medicine; clinicians; vaccine; health care worker; hospital-acquired; effectiveness; antiviral; cross-sectional study

Introduction

Annual seasonal influenza epidemics result in 290,000-650,000 deaths worldwide and over 88,000 preventable deaths in China annually [1,2]. During influenza seasons, health care workers (HCWs) are at a high risk of infection, leading to economic loss for the institutions they work for and potentially causing staff shortages [3]. HCWs may be asked or pressured to work despite having an influenza infection, contributing to clinic- or hospital-acquired influenza transmission [4]. Previous studies

XSL•FO
RenderX

have indicated that clinicians play a crucial role in influencing vulnerable groups. Clinicians who were vaccinated were more likely to recommend vaccinations to their patients, highlighting the critical role of this group in improving vaccine uptake [5]. Influenza vaccination is the most cost-effective measure to protect high-risk groups against severe influenza-associated diseases and deaths, and the World Health Organization recommends the influenza vaccination in particular for specific target groups, including HCWs [6].

However, despite efforts to promote influenza vaccination among HCWs, uptake rates remain unsatisfactory in China. In the 2019-2020 season, only 61% of Chinese clinicians, who were recruited at a platform for respiratory medical professionals, received vaccines, which was lower than the 87.9% in Finland [7,8]. In the 2021-2022 season, the coverage was even higher in the United States, which had 91.3% clinicians who received vaccines [9]. Barriers to influenza vaccine uptake among HCWs in China include out-of-pocket costs and insufficient workplace regulations [7]. For example, workplaces did not require or encourage clinicians to get vaccinated before the flu season. Most importantly, previous studies have primarily treated Chinese clinicians as practitioners of modern medicine, overlooking the significant presence of clinicians practicing traditional Chinese medicine (TCM) in China [7,10].

Modern Western medicine, originating in Europe, dominates medical practices globally, including in China, and relies on biomedical science, genetics, and advanced medical technologies to diagnose, treat, and prevent illness [11]. In contrast, TCM, rooted in ancient China, follows a distinct philosophy with unique principles, like Yin-Yang and the Five Elements, and employs practices such as acupuncture and herbal remedies [12]. The theoretical and diagnostic foundation of TCM is difficult to interpret through the framework of Western medicine anatomy and physiology; it is based on the philosophy, logic, and beliefs of a distinct civilization, offering a view of health and disease that differs from modern scientific thinking [13]. Ancient texts such as *Huangdi Nei Jing* and *Shanghan Lun* remain core guides for TCM students today [14]. TCM was categorized as a kind of complementary and alternative medicine by the World Health Organization and has been adopted by over 100 countries [15]. In 2020, the United States had over 30,000 licensed acupuncturists, while Germany and Australia also had thriving TCM practitioner communities, with more than 3000 and 4000 practitioners, respectively [12]. In 2020, there were 682,770 registered TCM clinicians, accounting for 16.7% of all clinicians, in mainland China, who provided medical services for over 1 billion outpatients and 35 million hospitalizations [16]. Considering the massive medical work undertaken by TCM clinicians, it is important to understand their influenza vaccination preferences.

We conducted a nationwide survey using a discrete choice experiment (DCE) among both TCM and Western medicine clinicians from January to May 2022 in China, to capture their stated preferences regarding influenza vaccination programs. This study's objectives were: (1) to investigate how different characteristics of influenza vaccination programs influenced clinicians' vaccination decisions, including TCM and Western medicine clinicians; (2) to estimate the expected influenza vaccine coverage and identify the optimal influenza vaccination program among TCM and Western medicine clinicians; and (3) to identify different subgroups that share certain outward preferences using latent class analysis (LCA).

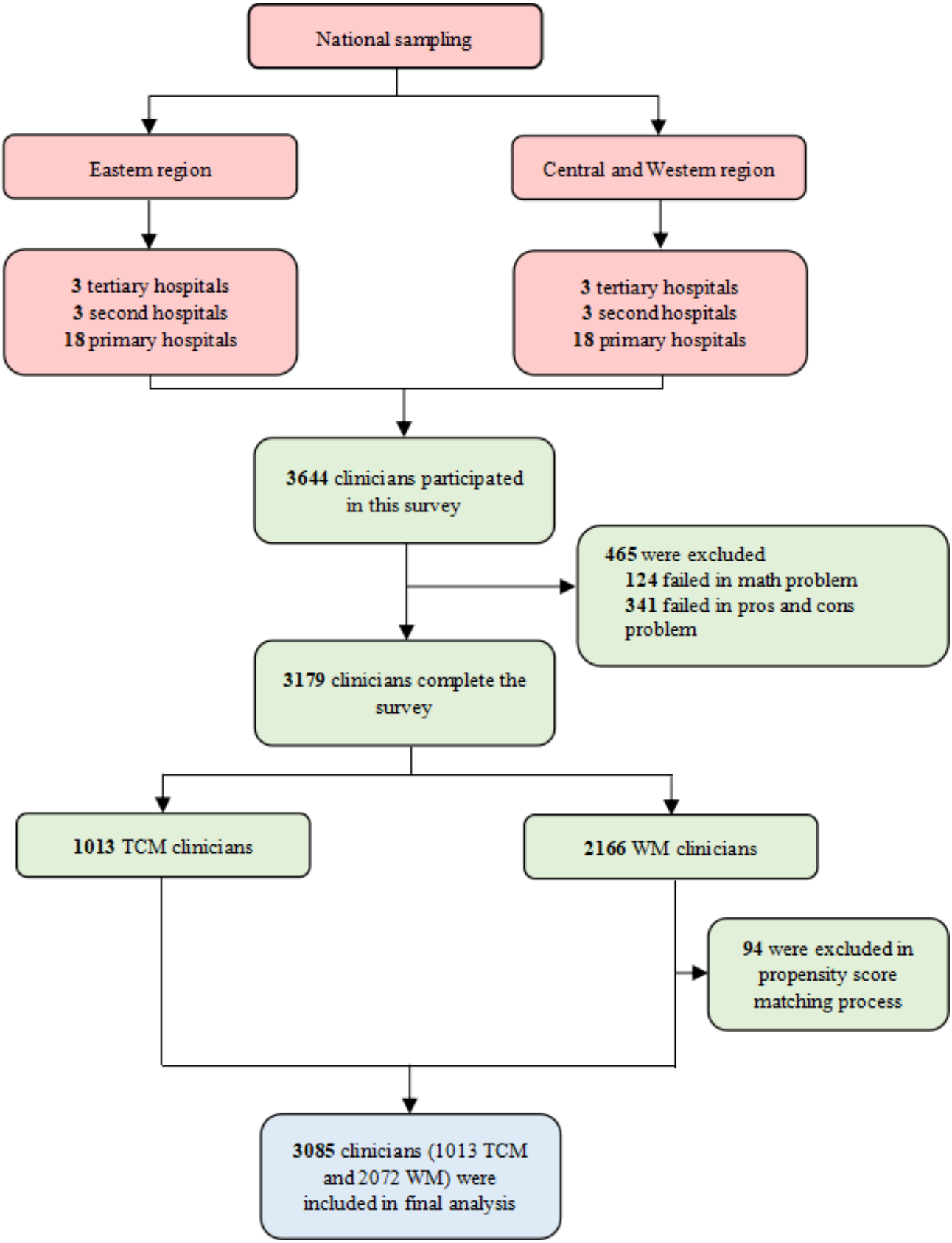
Methods

Participants

Clinicians (both TCM and Western medicine) holding a current practicing license in a hospital or community health service center, as of the date of survey administration, were included as potential participants. Data were collected through stratified cluster sampling between January 10 and May 10, 2022. In China, the coastal region (Eastern) and inland region (Central and Western) were stratified according to geographical location and socioeconomic development. Three tertiary hospitals, 3 secondary hospitals, and 18 primary hospitals were selected as survey spots in both the Eastern region and the Central and Western regions. The names and locations of these tertiary and secondary hospitals are listed in Table S1 in Multimedia Appendix 1. The primary hospitals were selected with the help of local health bureaus. They selected these hospitals through internal networks, and the name of each primary hospital was unknown to the researchers. We confirmed the location of primary hospital clinicians using their internet protocol address when they accessed the online questionnaire. The geographic distribution of all participants is presented in Figure S1 in Multimedia Appendix 1. Each hospital established a messaging group through WeChat (a free messaging and calling app) that included all the clinicians who worked in the hospital. Hospital administrative staff and local health bureaus delivered the online questionnaire link to these WeChat groups. We provided a brief statement for every participant before they answered the questionnaire to introduce the objective of our survey and to obtain informed consent. All participants were anonymized. Each WeChat account was allowed to submit only 1 online questionnaire. In China, real-name registration is required to apply for a phone number, and each phone number can be linked to only 1 WeChat account. Thus, we can reasonably assume that each individual could complete only 1 questionnaire. Participants who completed the questionnaire and correctly answered 2 quality control questions (a simple math problem and a pros-and-cons question where 1 choice is better than the other) were rewarded with a modest monetary incentive. Before the questionnaire was programmed, it was pilot tested by experts as well as 30 clinicians, who were contacted through our personal network, and adapted based on the feedback received. After sampling, we adjusted the Western medicine clinician sample using the propensity score matching method to ensure comparability between TCM clinicians and Western medicine clinicians (Figure 1).

XSL•FO
RenderX

Figure 1. Sample flow chart. TCM: traditional Chinese medicine; WM: Western medicine.



Demographic Variables

We used a 2-part questionnaire with demographic questions and DCE options. The demographic variables assessed are listed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Demographic variables assessed.

- Age in years, divided into groups (≤29, 30–39, 40–49, or ≥50 years)
- Sex (male or female)
- Educational level (junior college degree or lower, bachelor's degree, postgraduate degree, or doctorate)
- Medical license type (traditional Chinese medicine, Western medicine, or traditional Chinese medicine integrated with Western medicine)
- Annual personal income (CNY <¥100,000, ¥100,000-¥200,000, or >¥200,000; US <$13,649.18, $13,649.18-$27,298.35, or >$27,298.35, respectively)
- Length of time practicing medicine, in years (≤3 , 4-9, 10-19, 20-29, or ≥30 years)
- Professional (job) title (resident, fellow, or attending)
- Hospital level (primary, secondary, or tertiary)

Discrete Choice Experiment

A DCE is a technique used widely in the health care sector that elicits people's preferences by prompting them to make specific choices [17]. In this study, participants were asked to select a series of preferable vaccination programs. Each program was defined by a set of attributes, such as cost of the vaccine and risk of mild adverse events, and the variation in each attribute was referred to as a level. By making a series of choices between a vaccination program with one set of attribute levels and another with alternative attribute levels, participants implicitly revealed the degree to which each attribute is more important to them and the value they place on each attribute level.

This DCE was designed according to health-based principles, ranging from technicalities, such as design efficiency, to participant-centered considerations, such as checking if experimental tasks were as clear as possible to participants. In our survey, individuals made 8 choices, and for each choice they needed to choose a preferred influenza vaccination program between 2 alternative influenza vaccination programs. The influenza vaccination programs were described by attributes and levels representing different characteristics of vaccination programs.

Attributes and Levels

Five attributes described the characteristics of vaccination programs in the choice tasks, summarized in Table 1. We selected attributes and levels including: variable cost of vaccine (CNY ¥0; CNY ¥50 [approximately US $7]; or CNY ¥100 [approximately US $14]), variable vaccine effectiveness against infection (20%, 50%, or 80%), variable risk of mild adverse events such as redness, swelling, or pain at the injection site (1%, 3%, or 5%), different vaccination campaign strategies (individual appointment or vaccination in a workplace setting), and different workplace attitudes (no notice or encouragement of vaccination).

Table . Attributes and levels of discrete choice experiment.

Attribute	Levels
1. Cost of vaccine	• CNY ¥0 • CNY ¥50 (approximately $7 USD) • CNY ¥100 (approximately $15 USD)
2. Vaccine effectiveness against infection	• Prevent 20% infection • Prevent 50% infection • Prevent 80% infection
3. Risk of mild adverse events (such as redness, swelling, or pain at the injection site)	• 1% risk • 3% risk • 5% risk
4. Vaccination campaign strategies	• Individual appointment • Vaccination in a workplace setting
5. Workplace attitudes	• No notice • Encouragement of vaccination

We selected attributes and levels related to influenza vaccination based on several sources of evidence: a rapid review of existing research [18-21], one-on-one interviews with 4 TCM clinicians and 4 Western medicine clinicians, and a consultation with subject matter experts. An orthogonal experimental design was used to reduce these choice tasks down to 8 choice tasks with 16 hypothetical vaccination programs. Levels were randomly assigned with equal probabilities and repeatedly recombined, with all combinations being plausible and realistic. Priors were obtained from a pilot study of 30 clinicians. Each clinician answered 8 choice tasks, reporting no understanding problem and respondent fatigue. The full choice tasks, as presented to

XSL•FO
RenderX

respondents, are shown in Table S2 in Multimedia Appendix 1.

Statistical Analyses

The rule of thumb, as proposed by Johnson and Orme [22], suggests that the sample size required for the main effects of a DCE analysis depends on the number of choice tasks (*t*), the number of alternatives for each choice task (*a*), and the largest number of levels for any of the attributes (*c*), according to the following equation [23]:

$$N>1000c/(t\times a)$$

For our study, there were 8 choice tasks; each choice task had 2 alternatives, and the largest number of levels was 3 for any of the attributes. Our sample, which exceeded the minimum sample size calculated, of approximately at least 376 clinicians (or 188 of both TCM and Western medicine clinicians), was sufficient for the purposes of our study.

Sex, age groups, and hospital levels were identified as confounders in propensity score matching, and the propensity score of the Western medicine sample and TCM sample was calculated. We then used a nearest-neighbor matching algorithm to match 3 Western medicine clinicians with 1 TCM clinician based on the propensity score. Descriptive statistics were used to summarize characteristics of TCM and Western medicine clinicians. For categorical variables, frequencies were reported, and Pearson χ^2 tests were used to test for differences between the 2 groups. For continuous variables, mean and standard deviation were reported after Shapiro-Wilk tests indicated that the continuous variables were normally distributed. We used *t* tests to compare differences across the 2 groups.

To estimate the relative impact of influenza vaccine attributes in the DCE, we conducted a mixed logit regression model (MLM) to compute preference weight. An MLM is based on the assumption that random error has a normal distribution, taking heterogeneity as well as correlation between the choice task completed by each participant into account. We validated the linear continuous effects of chosen attributes, and the variables of cost, vaccine efficacy, and risk of mild adverse events were considered as continuous, while the vaccination campaign strategies and workplace attitudes were considered as categorized in the analysis. The MLM allowed for the calculation of compensatory effects across any 2 attributes. For example, a 5% decrease of risk of mild adverse events can compensate the negative impact of a 10% reduction of vaccine effectiveness. The willingness to pay (WTP) refers to the compensatory effect between cost and any other attribute. We calculated the expected vaccine coverage for the base case and the change of coverage when the level of one attribute was changed.

Subsequently, we conducted an LCA to classify individuals based on their preference characteristics; an individual was assigned to the class with the highest posterior probability. For example, the LCA presented that 31.1% (644/2072) of Western medicine clinicians preferred free vaccinations over any other attribute. Our choice in the number of subgroups was based on model fit and interpretability of results. We compared classes across demographic characteristics among the estimated classes using χ^2 tests. The preference subgroups were named based on the most preferred attribute levels in each class.

All analyses were based on 2-sided *P* values, with $P<.05$ indicating statistical significance. Stata (Version 16.0, Stata Corp LLC) was used for analysis.

Ethical Considerations

The Zhejiang Chinese Medical University Ethics Committee reviewed and approved this protocol (no. 20221021 - 1). Prior to participation, all study participants were required to sign an informed consent form, thereby confirming their voluntary engagement in the survey process. The study data were anonymized. All participants who completed the questionnaire and successfully submitted it received a reward of CNY ¥20 (US $2.7), distributed via WeChat.

Results

Demographics

A total of 3644 clinicians from 6 tertiary hospitals, 6 secondary hospitals, and 36 primary hospitals were invited to participate in this survey. Among them, 559 (15.3%) were ultimately excluded for failing the basic math tests (n=124, 3.4%) or the pros-and-cons test (n=341, 9.4%) or in the post-hoc during propensity score matching (n=94, 2.6%). Finally, 3085 subjects, consisting of 1013 (32.8%) TCM clinicians and 2072 (67.2%) Western medicine clinicians, were included in the analysis (Figure 1).

The sample matching adjustment achieved the expected distributions of gender, age groups, and hospital level (Figures S2-S4 in Multimedia Appendix 1). Educational attainment showcased a significant discrepancy, with a notable proportion of Western medicine clinicians holding a bachelor's degree (1106/2072, 53.4%) compared to TCM clinicians (409/1013, 40.4%). Conversely, a higher percentage of TCM clinicians had attained a master's degree (274/1013, 27%) or a doctorate (110/1013, 10.9%) compared to their Western medicine counterparts. In terms of annual personal income, a higher percentage of TCM clinicians reported incomes of CNY ¥100,000 - ¥200,000 (US$13,649.18-$27,298.35) (345/1013, 34.1%) compared to Western medicine clinicians (570/2072, 27.5%). Conversely, a higher proportion of Western medicine clinicians reported incomes above CNY ¥200,000 (US $27,298.35) (413/2072, 19.9%) compared to TCM clinicians (133/1013, 13.2%). In terms of professional title, 50.4% (511/1013) of TCM clinicians held the title of resident compared to Western medicine clinicians (916/2072, 44.2%). Conversely, a higher proportion of Western medicine clinicians held the title of fellow (827/2072, 39.9%) compared to TCM clinicians (346/1013, 34.2%). For the length of time practicing medicine, there was no significant difference between TCM and Western medicine clinicians ($P=.44$). (Table 2).

XSL•FO
RenderX

Table . Demographic data of clinicians practicing traditional Chinese medicine (TCM) and Western medicine (WM) between January and May 2022 in China.

	TCM clinicians (n=1013), n (%)	WM clinicians (n=2072), n (%)	*P* value
Sex			.31
Male	420 (41.5)	899 (43.4)	
Female	593 (58.5)	1173 (56.6)	
Age group (year)			.63
0-29	319 (31.5)	618 (29.8)	
30-39	431 (42.5)	893 (43.1)	
40-49	183 (18.1)	407 (19.6)	
≥50	80 (7.9)	154 (7.5)	
Hospital level			.55
Primary	295 (29.1)	637 (30.7)	
Secondary	240 (23.7)	498 (24)	
Tertiary	478 (47.2)	937 (45.2)	
Educational level			<.001
Junior college degree or lower	220 (21.7)	584 (28.2)	
Bachelor's degree	409 (40.4)	1106 (53.4)	
Master's degree	274 (27)	289 (13.9)	
Doctorate	110 (10.9)	93 (4.5)	
Annual personal income (CNY[a])			<.001
0-100,000	535 (52.8)	1089 (52.6)	
100,000-200,000	345 (34.1)	570 (27.5)	
>200,000	133 (13.2)	413 (19.9)	
Length of time practicing medicine (year)			.44
0-3	287 (28.3)	543 (26.2)	
4–9	331 (32.7)	677 (32.7)	
10–19	242 (23.9)	540 (26.1)	
20–29	106 (10.5)	232 (11.2)	
≥30	47 (4.6)	80 (3.9)	
Professional title			.003
Resident	511 (50.4)	916 (44.2)	
Fellow	346 (34.2)	827 (39.9)	
Attending	156 (15.4)	329 (15.9)	

[a]CNY: Chinese yuan. CNY ¥1=US $0.14.

Preference Weight and Compensatory Effects (Including WTP)

All attributes and levels included were significantly associated with clinicians' preferences for influenza vaccination programs. Both TCM and Western medicine clinicians' preferences for influenza vaccination programs increased with higher vaccine effectiveness, with the vaccination campaign strategies changing from "individual appointment" to "vaccination in a workplace setting," or with the workplace attitude changing from "no notice" to "encouraging of vaccination." Nevertheless, the increase in the risk of mild adverse events and cost could undermine the utility of vaccination towards these 2 groups. The WTPs of TCM clinicians were calculated and showed that each 1% increase in risk of mild adverse events could be compensated by a CNY ¥6.97 (US $1.37) decrease of cost. Similarly, changing the mode of administration from "individual appointment" to "vaccination in a workplace setting" could compensate for the negative effect of a CNY ¥78.84 (US $15.44) increase of cost. Compensatory effects can be calculated between any attributes; for example, a hypothetical 5% increase in risk of mild adverse events could be compensated by a

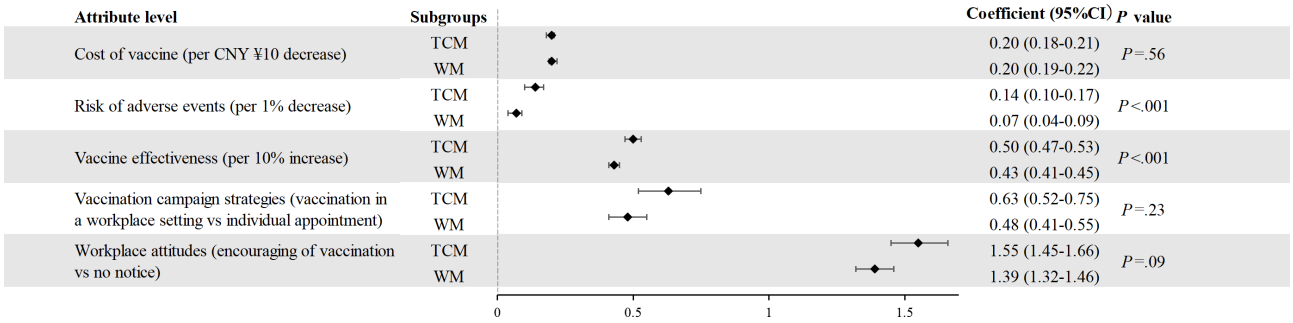
XSL•FO
RenderX

simultaneous hypothetical 13.7% increase in vaccine effectiveness. WTP and the compensatory effect of Western medicine clinicians were calculated in the same way. The preference weights and WTP of attributes and levels of the 2 groups are shown in Table S3 in Multimedia Appendix 1.

Comparison of Preference Weights Among TCM and Western Medicine Clinicians

Significant differences were noted for the coefficients of the risk of mild adverse events and vaccine effectiveness between clinicians practicing TCM and Western medicine. TCM clinicians (odds ratio [OR] 0.14, 95% CI 0.10 - 0.17) expressed a stronger preference for an influenza vaccination program per 1% decrease of mild adverse events risk than Western medicine clinicians (OR 0.07, 95% CI 0.04 - 0.09; *P*<.001). Similarly, TCM clinicians (OR 0.50, 95% CI 0.47 - 0.53) expressed a stronger preference for an influenza vaccination program per 10% increase of vaccine effectiveness than Western medicine clinicians (OR 0.43, 95% CI 0.41 - 0.45; *P*<.001). Other attributes (cost, vaccination campaign strategies, and workplace attitudes) did not present significant differences between the 2 groups (Figure 2).

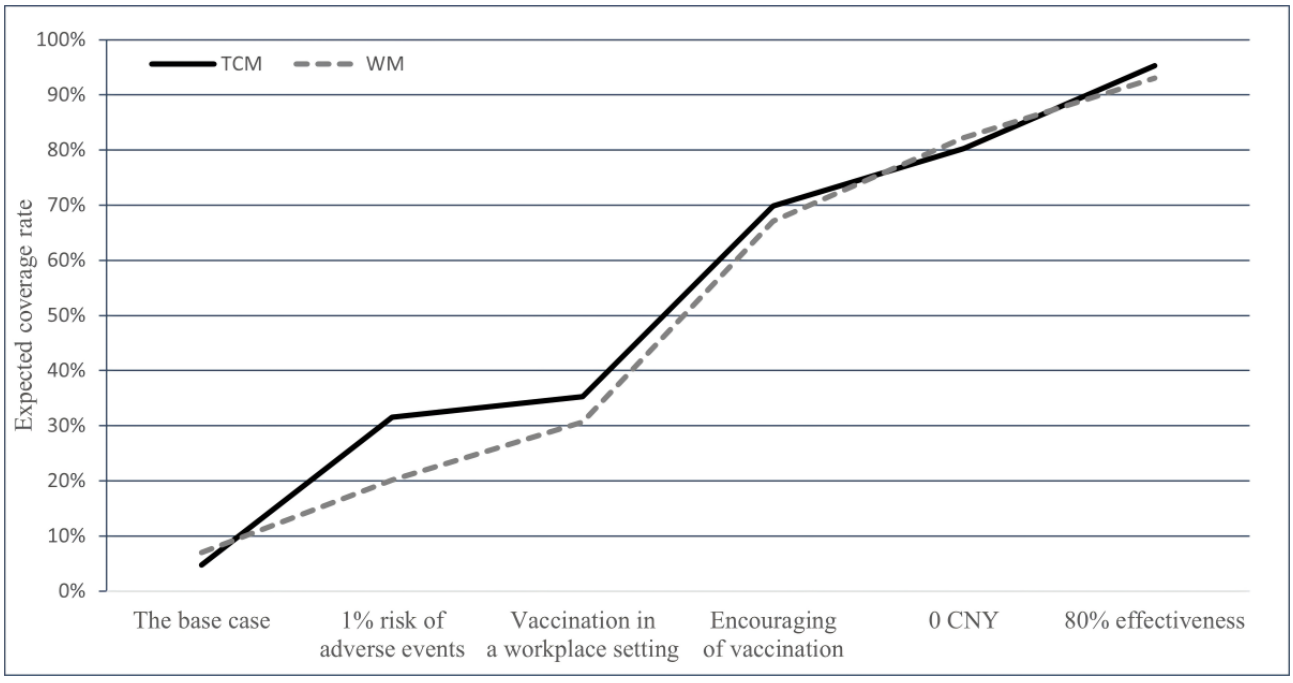
Figure 2. Comparison of preferences in clinicians practicing TCM and WM between January and May 2022 in China. TCM: traditional Chinese medicine; WM: Western medicine.



Expected Influenza Vaccine Coverage Under Various Influenza Vaccination Programs

We set the base case as having a cost of CNY ¥100 (US $13.64), 5% risk of mild adverse events, 20% effectiveness, vaccination by individual arrangement, and no notice from the workplace. The expected coverage of the base case to TCM clinicians was 4.75% and to Western medicine clinicians was 6.99%. Expected coverage increased by approximately 30% with a reduction in risk of mild adverse events from 5% to 1%; greater coverage increases could be achieved by either enhancing effectiveness from 20% to 80% or reducing costs from CNY ¥100 to CNY ¥0. An approximate hypothetical 60% increase could be obtained when the attitude of the workplace changed from “no notice” to “encouraging of vaccination,” and there was an approximate hypothetical 35% increase when vaccination campaign strategies changed from “individual appointment” to “vaccination in a workplace setting.” The changing of expected coverage due to improvements in the influenza vaccination program is depicted in Figure 3 (additional details available in Table S4 in Multimedia Appendix 1).

Figure 3. Changes in expected coverage rate, as attributes in influenza vaccination programs improve, between January and May 2022 in China. We set the base case as having a cost of CNY ¥100 (US $13.64), 5% risk of mild adverse events, 20% effectiveness, vaccination by individual arrangement, and no notice from the workplace. TCM: traditional Chinese medicine; WM: Western medicine; CNY: Chinese yuan.



XSL•FO
RenderX

Preference of Different Subgroups

Table 3 details the demographic composition variations across various subgroups, while the fitting and categorization process of these subgroups is recorded in Table S5 in Multimedia Appendix 1. In the entire sample, about 30% (946/3085) of clinicians preferred free vaccinations, while approximately 26% (819/3085) comprehensively considered all attributes except for vaccination campaign strategies. Notable trends were observed within 2 specific subgroups: clinicians preferring free vaccinations and those preferring vaccinations in a workplace setting. Among clinicians with different medical licenses, TCM practitioners exhibited a lower preference for free vaccinations (302/1013, 29.8%) compared to Western medicine clinicians (644/2072, 31.1%) and a higher preference for workplace vaccinations (186/1013, 18.4%) compared to Western medicine clinicians (351/2072, 17%) (P=.03). Differences were also noted across sex, with female clinicians indicating a lower preference for free vaccinations (533/1766, 30.2%) compared to male clinicians (414/1319, 31.3%) and a higher preference for workplace vaccinations (317/1766, 18%) compared to male clinicians (219/1319, 16.6%) (P=.04). Educational level revealed a U-shaped relationship, where the highest preference for free vaccinations was among those with junior college degrees or lower (270/804, 33.6%); however, this preference decreased significantly among those with postgraduate degrees (158/563, 28%; P<.001). Conversely, there was an inverted U-shaped relationship between educational level and preference for vaccinations in a workplace setting. A strong monotonic relationship was present between hospital level and clinicians' preference for free vaccinations, with 32.5% (303/932) of primary hospital workers preferring free vaccinations compared to 29.4% (416/1415) of those in tertiary hospitals (P<.001). On the other hand, clinicians in tertiary hospitals (262/1415, 18.5%) significantly preferred vaccinations in a workplace setting compared to those in primary (145/932, 15.6%) or secondary hospitals (131/738, 17.7%) (P<.001). No significant differences were found based on age group, annual personal income, length of time practicing medicine, or professional title.

XSL•FO
RenderX

Table . Demographic characteristics by influenza vaccination preference from a latent class analysis between January and May 2022 in China.

Demographics and categories	Prefer encouragement of vaccinations in a workplace setting, n (%)	Prefer free vaccinations, n (%)	Prefer higher vaccine effectiveness, n (%)	Consider all attributes except vaccination campaign strategies, n (%)	Prefer vaccinations in a workplace setting, n (%)	*P* value
Medical practicing license						.03
TCM[a] (n=1013)	144 (14.2)	302 (29.8)	113.46 (11.2)	267 (26.4)	186 (18.4)	
WM[b] (n=2072)	300 (14.5)	644 (31.1)	223.78 (10.8)	552 (26.7)	351 (17)	
Age group (year)						.54
0-29 (n=937)	137 (14.6)	292 (31.2)	102 (11)	242 (25.8)	164 (17.5)	
30-39 (n=1324)	188 (14.2)	408 (30.8)	144.32 (10.9)	357 (27)	226 (17.1)	
40-49 (n=590)	87 (14.7)	173 (29.3)	64.31 (10.9)	158 (26.7)	109 (18.4)	
≥50 (n=234)	34 (14.5)	73 (31.2)	24.57 (10.5)	63 (27)	39 (16.8)	
Sex						.04
Male (n=1319)	193 (14.6)	414 (31.3)	142.45 (10.8)	351 (26.6)	219 (16.6)	
Female (n=1766)	253 (14.3)	533 (30.2)	193 (11)	470 (26.6)	317 (18)	
Educational level						<.001
Junior college degree or lower (n=804)	118 (14.7)	270 (33.6)	85.22 (10.6)	211 (26.3)	119 (14.8)	
Bachelor's degree (n=1515)	218 (14.4)	459 (30.3)	166 (11)	402 (26.6)	270 (17.8)	
Postgraduate degree (n=563)	80 (14.2)	158 (28)	63.62 (11.3)	151 (26.9)	110 (19.6)	
Doctorate (n=203)	28 (14)	59 (29.3)	21.92 (10.8)	54 (26.7)	39 (19.2)	
Annual personal income (CNY[c])						.63
0-100,000 (n=1624)	235 (14.5)	505 (31.1)	175.39 (10.8)	431 (26.5)	277 (17)	
100,000-200,000 (n=915)	131 (14.3)	278 (30.4)	99.74 (10.9)	243 (26.6)	163 (17.8)	
>200,000 (n=546)	78 (14.2)	163 (29.9)	60.61 (11.1)	145 (26.6)	99 (18.2)	
Length of time practicing medicine, in years						.79
0-3 (n=830)	120 (14.5)	262 (31.6)	91.3 (11)	221 (26.6)	135 (16.4)	
4–9 (n=1008)	147 (14.6)	305 (30.3)	107.86 (10.7)	265 (26.3)	182 (18.1)	
10–19 (n=782)	111 (14.2)	236 (30.2)	86 (11.1)	210 (26.8)	139. (17.8)	
20–29 (n=338)	49 (14.4)	103 (30.5)	37.52 (11.1)	91 (27)	58 (17.1)	
≥30 (n=127)	18 (14.3)	39 (30.7)	13.21 (10.4)	34 (26.9)	22 (17.7)	
Professional (job) title						.28
Residents (n=1427)	207 (14.5)	448 (31.4)	154.12 (10.8)	378 (26.5)	240 (16.8)	
Fellows (n=1173)	167 (14.2)	354 (30.2)	129.03 (11)	314 (26.8)	209 (17.8)	

XSL•FO
RenderX

Demographics and categories	Prefer encouragement of vaccinations in a workplace setting, n (%)	Prefer free vaccinations, n (%)	Prefer higher vaccine effectiveness, n (%)	Consider all attributes except vaccination campaign strategies, n (%)	Prefer vaccinations in a workplace setting, n (%)	*P* value
Attending (n=485)	71 (14.6)	144 (29.7)	53.84 (11.1)	128 (26.3)	89 (18.3)	
Hospital level						<.001
Primary (n=932)	133 (14.3)	303 (32.5)	99.72 (10.7)	251 (26.9)	145 (15.6)	
Secondary (n=738)	107 (14.5)	227 (30.7)	80.44 (10.9)	193 (26.2)	131 (17.7)	
Tertiary (n=1415)	204 (14.4)	416 (29.4)	157.07 (11.1)	376 (26.6)	262 (18.5)	

[a]TCM: traditional Chinese medicine.

[b]WM: Western medicine.

[c]CNY: Chinese yuan. CNY ¥1=US $0.14.

Discussion

Principal Findings

We found that (1) TCM clinicians were more sensitive to changes to the risk of mild adverse events and vaccine effectiveness than Western medicine clinicians; (2) decreasing the price of influenza vaccinations, offering convenient vaccinations in workplace settings, and issuing notifications encouraging clinicians to get vaccinated could increase influenza vaccination coverage of both groups; and (3) compared to Western medicine clinicians, a larger proportion of TCM clinicians preferred vaccinations in workplace settings over any other attribute.

The findings indicate a higher probability of vaccines being received when paired with increased vaccine effectiveness and decreased likelihood of vaccine-related adverse events, which aligns with prior research conducted through questionnaire surveys [24]. In our survey, TCM clinicians placed significantly higher preference weights on the risk of mild adverse events and vaccine effectiveness, indicating that fluctuations in the safety and effectiveness of influenza vaccines could have a greater impact on them. A qualitative study shows that patients find vaccination information transmitted by complementary and alternative medicine providers to be more understandable, useful, and trustworthy [25]. A TCM clinician who receives and shares information regarding defective vaccine incidents could cause a chain reaction in the broad population. A systematic review indicated that the provision of free vaccinations, easy access to vaccinations, and modification through educational activities and reminders can effectively increase the influenza vaccine uptake [26]. In the simulation scenario, decreasing the price of the influenza vaccination from CNY ¥100 (US $13.64) to being free resulted in a significant increase in projected coverage. Offering free vaccinations to clinicians can provide additional benefits, such as mitigating absenteeism during the influenza season and preventing hospital-acquired influenza infections. We also found that other workplace-based interventions, such as providing vaccination opportunities at the workplace and promoting vaccination initiatives more broadly, demonstrated a favorable effect on the projected uptake of influenza vaccine. Given the feasibility and efficacy of these measures, we propose that health care facilities should (1) provide free annual influenza vaccinations, (2) furnish accessible vaccination services for employees, and (3) issue notifications encouraging clinicians to avail themselves of vaccination opportunities. In our LCA, the majority of clinicians prioritized free vaccinations, followed by a substantial group who took into account all attributes except vaccination campaign strategies when deciding whether to accept an influenza vaccination program. The results of our survey suggest that a higher proportion of TCM clinicians favored workplace-based vaccinations over any other attribute, whereas a greater number of Western medicine clinicians preferred free vaccinations over any other attribute. This finding can lend support to the creation of a vaccination policy that is more flexible and adaptable. Sex, educational level, and hospital level were all associated with latent classification. Among these considerations, the most significant distinction emerged between clinicians who preferred free vaccinations most and those who preferred vaccinations in a workplace setting most. This differentiation suggests that the other 3 categories (favoring workplace encouragement of vaccinations, prioritizing higher vaccine effectiveness, and considering attributes except vaccination campaign strategies) were evenly distributed across all demographics in our study. As a result, measures like releasing notices in health care workplaces to encourage clinicians to get vaccinated may achieve similar effects across all groups of Chinese clinicians. A previous study showed that women's perceptions of inflation (consumer prices increasing) were higher than men's, but fewer female clinicians preferred free vaccinations over vaccinations in a workplace setting, compared to male clinicians in our study [27]. A policy offering free vaccinations may hold greater appeal for clinicians with lower levels of education (bachelor's degree or below) or those employed in primary health care institutions. Conversely, providing vaccinations in a workplace setting was found to be more favorable among clinicians with higher levels of education (postgraduate degree or doctorate) or those working in tertiary hospitals. In China, clinicians employed in tertiary hospitals serve as the cornerstone of the nation's health care provision, handling extensive medical responsibilities that may make it challenging for them to get vaccinated during their

XSL•FO
RenderX

private leisure time. Consequently, providing convenient vaccination services for clinicians should be prioritized as a primary measure for these tertiary hospitals.

Limitations

Our study is subject to several limitations. First, our sample of clinicians did not perfectly mirror the sociodemographic characteristics of the entire nation. Additionally, there may be biases between individuals who completed the survey and those who did not. Second, as a cross-sectional study, potential recall bias is inevitable. Moreover, DCEs are susceptible to hypothetical bias, meaning that responses in surveys may not perfectly align with real-life behaviors [28]. This discrepancy could limit the accuracy of measured preferences. To mitigate this, we designed the experiment based on real-world influenza vaccination programs and ensured that the details and definitions of each attribute were comprehensible to participants.

Conclusions

In summary, while changes in mild adverse event rates and vaccine effectiveness had a greater impact on TCM clinicians, other attributes, such as workplace-provided free vaccinations, issuing notifications to encourage vaccinations, and providing accessible vaccination services, are practical methods to increase the uptake of influenza vaccination programs among both groups. Policymakers can utilize these insights to formulate flexible intervention measures that facilitate greater accessibility of influenza vaccination programs for clinicians, especially TCM clinicians, worldwide.

Acknowledgments

We thank Gang Cai, Limin Chen, Yu Du, Jinglin Han, Conghua Ji, Xia Jin, Fangyou Li, Xikun Li, Peizhen Li, Chunxiao Liu, Jianping Li, Miner Qian, Baoshan Sun, Yongshou Yang, Yuan Lv, Tianxing Zhang for their help with data collection. This work was supported by the National Social Science Fund of China (22BGL316).

Data Availability

The deidentified cross-sectional data and the discrete choice models from the analysis are available from the corresponding author on reasonable request.

Authors' Contributions

LZ and CF conceptualized the study, contributed to the overall design of the survey experiment, analysis, and interpretation of the data, and are the guarantors for the study. LZ cleaned the data, performed the analyses, drew the figures, conducted the literature searches, and wrote the first draft of the manuscript. LZ, LC, YZ, XS, ZW, MES, and CF discussed the results and interpretation of the data. MES and CF reviewed and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and figures.
[DOCX File, 176 KB - publichealth_v11i1e63314_app1.docx]

References

1. WHO launches new global influenza strategy. World Health Organization. 2019. URL: https://www.who.int/news/item/11-03-2019-who-launches-new-global-influenza-strategy [accessed 2025-01-08]
2. Li L, Liu Y, Wu P, et al. Influenza-associated excess respiratory mortality in China, 2010-15: a population-based study. Lancet Public Health 2019 Sep;4(9):e473-e481. [doi: 10.1016/S2468-2667(19)30163-X] [Medline: 31493844]
3. Gianino MM, Politano G, Scarmozzino A, et al. Cost of sickness absenteeism during seasonal influenza outbreaks of medium intensity among health care workers. Int J Environ Res Public Health 2019 Mar 1;16(5):747. [doi: 10.3390/ijerph16050747] [Medline: 30832264]
4. Chiu S, Black CL, Yue X, et al. Working with influenza-like illness: presenteeism among US health care personnel during the 2014-2015 influenza season. Am J Infect Control 2017 Nov 1;45(11):1254-1258. [doi: 10.1016/j.ajic.2017.04.008] [Medline: 28526310]
5. Blendon RJ, Benson JM, Hero JO. Public trust in physicians--U.S. medicine in international perspective. N Engl J Med 2014 Oct 23;371(17):1570-1572. [doi: 10.1056/NEJMp1407373] [Medline: 25337746]
6. Influenza (seasonal). World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/influenza-(seasonal) [accessed 2025-01-08]
7. Yi H, Yang Y, Zhang L, et al. Improved influenza vaccination coverage among health-care workers: evidence from a web-based survey in China, 2019/2020 season. Hum Vaccin Immunother 2021 Jul 3;17(7):2185-2189. [doi: 10.1080/21645515.2020.1859317] [Medline: 33497309]

XSL•FO
RenderX

8. Hammer CC, Lyytikäinen O, Arifulla D, Toura S, Nohynek H. High influenza vaccination coverage among healthcare workers in acute care hospitals in Finland, seasons 2017/18, 2018/19 and 2019/20. Euro Surveill 2022 Apr;27(17):2100411. [doi: 10.2807/1560-7917.ES.2022.27.17.2100411] [Medline: 35485269]
9. Razzaghi H, Srivastav A, de Perio MA, Laney AS, Black CL. Influenza and COVID-19 vaccination coverage among health care personnel - United States, 2021-22. MMWR Morb Mortal Wkly Rep 2022 Oct 21;71(42):1319-1326. [doi: 10.15585/mmwr.mm7142a2] [Medline: 36264832]
10. Saunier F, Berthelot P, Mottet-Auselo B, et al. Impact of a decision-aid tool on influenza vaccine coverage among HCW in two French hospitals: a cluster-randomized trial. Vaccine 2020 Aug 10;38(36):5759-5763. [doi: 10.1016/j.vaccine.2020.07.011] [Medline: 32684500]
11. Loudon I, editor. Western Medicine: An Illustrated History: Oxford University Press; 1997.
12. Traditional medicine. World Health Organization. 2023. URL: https://www.who.int/news-room/questions-and-answers/item/traditional-medicine [accessed 2025-01-08]
13. Tang JL, Liu BY, Ma KW. Traditional Chinese medicine. Lancet 2008 Dec 6;372(9654):1938-1940. [doi: 10.1016/S0140-6736(08)61354-9] [Medline: 18930523]
14. Cao X. Immunology in China: the past, present and future. Nat Immunol 2008 Apr;9(4):339-342. [doi: 10.1038/ni0408-339] [Medline: 18349809]
15. WHO international standard terminologies on traditional Chinese medicine. World Health Organization. URL: https://www.who.int/publications/b/55892 [accessed 2025-01-08]
16. China's traditional Chinese medicine service capacity has been significantly improved, with the total number of TCM visits reaching 1.2 billion [Article in Chinese]. WeChat. 2023. URL: https://mp.weixin.qq.com/s/eqgzswqtGiQC7NSLgsDgLA [accessed 2025-01-08]
17. Hensher DA, Rose JM, Greene WH. Applied Choice Analysis: Cambridge University Press; 2015.
18. Walsh JM, McPhee SJ. A systems model of clinical preventive care: an analysis of factors influencing patient and physician. Health Educ Q 1992;19(2):157-175. [doi: 10.1177/109019819201900202] [Medline: 1618625]
19. Vasilevska M, Ku J, Fisman DN. Factors associated with healthcare worker acceptance of vaccination: a systematic review and meta-analysis. Infect Control Hosp Epidemiol 2014 Jun;35(6):699-708. [doi: 10.1086/676427] [Medline: 24799647]
20. Wicker S, Marckmann G, Poland GA, Rabenau HF. Healthcare workers' perceptions of mandatory vaccination: results of an anonymous survey in a German University Hospital. Infect Control Hosp Epidemiol 2010 Oct;31(10):1066-1069. [doi: 10.1086/656242] [Medline: 20804415]
21. Pielak KL, McIntyre CC, Tu AW, Remple VP, Halperin B, Buxton JA. Identifying attitudes, beliefs and reported practices of nurses and doctors as immunization providers. J Adv Nurs 2010 Jul;66(7):1602-1611. [doi: 10.1111/j.1365-2648.2010.05326.x] [Medline: 20492025]
22. Johnson R, Orme B. Getting the most from CBC. : Sawtooth Software; 2003 URL: https://sawtoothsoftware.com/resources/technical-papers/getting-the-most-from-cbc [accessed 2025-01-13]
23. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA. Sample size requirements for discrete-choice experiments in healthcare: a practical guide. Patient 2015 Oct;8(5):373-384. [doi: 10.1007/s40271-015-0118-z] [Medline: 25726010]
24. Alshammari TM, Yusuff KB, Aziz MM, Subaie GM. Healthcare professionals' knowledge, attitude and acceptance of influenza vaccination in Saudi Arabia: a multicenter cross-sectional study. BMC Health Serv Res 2019 Apr 15;19(1):229. [doi: 10.1186/s12913-019-4054-9] [Medline: 30992004]
25. Deml MJ, Notter J, Kliem P, et al. "We treat humans, not herds!": a qualitative study of complementary and alternative medicine (CAM) providers' individualized approaches to vaccination in Switzerland. Soc Sci Med 2019 Nov;240:112556. [doi: 10.1016/j.socscimed.2019.112556] [Medline: 31563005]
26. Hollmeyer H, Hayden F, Mounts A, Buchholz U. Review: interventions to increase influenza vaccination among healthcare workers in hospitals. Influenza Other Respir Viruses 2013 Jul;7(4):604-621. [doi: 10.1111/irv.12002] [Medline: 22984794]
27. Plane M, Vermersch G. Inflation and counter-inflationary policy measures: the case of France. Hans-Böckler-Stiftung. 2022. URL: https://nbn-resolving.de/urn:nbn:de:101:1-2023011910431794423230%0A [accessed 2025-01-08]
28. Catalogue of Bias Collaboration, Buckell J, Buchanan J, et al. Hypothetical bias. Catalogue of Bias. URL: https://catalogofbias.org/biases/hypothetical-bias/ [accessed 2025-01-08]

Abbreviations

DCE: discrete choice experiment
HCW: health care worker
MLM: mixed logit regression model
OR: odds ratio
TCM: traditional Chinese medicine
WTP: willingness to pay

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 17.06.24; peer-reviewed by N Jiang, S Jimenez-Jorge; revised version received 15.12.24; accepted 19.12.24; published 20.01.25.

Please cite as:
Zhang L, Chu L, Sundaram ME, Zhou Y, Sun X, Wei Z, Fu C
Identifying Preferred Features of Influenza Vaccination Programs Among Chinese Clinicians Practicing Traditional Chinese Medicine and Western Medicine: Discrete Choice Experiment
JMIR Public Health Surveill 2025;11:e63314
URL: https://publichealth.jmir.org/2025/1/e63314
doi:10.2196/63314

© Liuren Zhang, Linchen Chu, Maria E Sundaram, Yi Zhou, Xiu Sun, Zheng Wei, Chuanxi Fu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Attitudes Toward Psychotherapeutic Treatment and Health Literacy in a Large Sample of the General Population in Germany: Cross-Sectional Study

Rebekka Schröder[1], Dr phil; Tim Hamer[1], MSc; Ralf Suhr[1], PD, Dr med; Lars König[1], Dr rer nat

Stiftung Gesundheitswissen, Berlin, Germany

Corresponding Author:
Lars König, Dr rer nat
Stiftung Gesundheitswissen
Friedrichstr. 134
Berlin, 10117
Germany
Phone: 49 30 4195492 62
Email: lars.koenig@stiftung-gesundheitswissen.de

Abstract

Background: Prevalences of mental disorders are increasing worldwide. However, many people with mental health problems do not receive adequate treatment. An important factor preventing individuals from seeking professional help is negative attitudes toward psychotherapeutic treatment. Although a positive shift in attitudes has been observed in recent years, there is still substantial stigma surrounding psychotherapeutic treatment. First studies have linked higher health literacy with more positive attitudes toward psychotherapy, but more research is needed in this area.

Objective: This study aimed to examine how general and mental health literacy are associated with attitudes toward psychotherapeutic treatment in Germany. Additionally, associations between sociodemographic factors, experience with psychotherapy, and attitudes toward psychotherapy were explored.

Methods: A random sample was drawn from a panel representative of the German-speaking population with internet access in Germany and invited to participate in the study via email. Overall, 2000 individuals aged ≥16 years completed the web-based survey with standardized questionnaires in September and October 2022. Attitudes toward psychotherapy and both general and mental health literacy were assessed using the Questionnaire on Attitudes Towards Psychotherapeutic Treatment (QAPT) with 2 subscales ("positive attitudes" and "non-acceptance of society"), the European Health Literacy Survey instrument (HLS-EU-Q16) and the Mental Health Literacy Tool for the Workplace (MHL-W-G). Associations between the questionnaire scales were assessed with Pearson correlations. Additionally, basic sociodemographic information and information on personal and family experiences with psychotherapy were collected. Pearson correlations (age), ANOVAs (level of education and subjective social status), and *t* tests (experience with psychotherapy, gender, and migration background) were used to analyze how these relate to attitudes toward psychotherapy.

Results: More favorable attitudes toward psychotherapy and lower perceived societal nonacceptance were found in those with higher general (r=0.14, P<.001; r=−0.32, P<.001, respectively) and mental health literacy (r=0.18, P<.001; r=−0.23, P<.001, respectively). Participants with treatment experience for mental health problems ($t_{1260.12}$=−10.40, P<.001, Cohen d=−0.49; $t_{1050.95}$=3.06, P=.002, Cohen d=0.16) and who have relatives with treatment experience ($t_{1912.06}$=−5.66, P<.001, Cohen d=−0.26; t_{1926}=4.77, P<.001, Cohen d=0.22) reported more positive attitudes and higher perceived societal acceptance than those without treatment experience. In terms of sociodemographic differences, being a woman (t_{1992}=−3.60, P<.001, Cohen d=−0.16), younger age (r=−0.11, P<.001), higher subjective social status ($F_{2,1991}$=5.25, P=.005, η^2=.005), and higher levels of education ($F_{2,1983}$=22.27, P<.001, η^2=.021) were associated with more positive attitudes toward psychotherapeutic treatment. Being a man (t_{1994}=5.29, P<.001, Cohen d=0.24), younger age (r=−0.08, P<.001), and lower subjective social status ($F_{2,1993}$=7.71, P<.001, η^2=.008) were associated with higher perceived nonacceptance of psychotherapy.

Conclusions: Positive associations between attitudes toward psychotherapy and both general and mental health literacy were delineated. Future studies should investigate whether targeted health literacy interventions directed at individuals with lower

XSL•FO
RenderX

general and mental health literacy might also help to improve attitudes toward psychotherapeutic treatment and help-seeking behavior.

(JMIR Public Health Surveill 2025;11:e67078) doi:10.2196/67078

KEYWORDS

health literacy; mental health literacy; attitude to health; cross-sectional studies; Germany; adults; representative; psychotherapy

Introduction

In Germany, almost 30% of the population are estimated to have a mental disorder [1] but only about 10% have received treatment in the past year [2]. This means that about two-thirds of those with mental health problems do not receive adequate treatment. While the number of individuals not seeking help for mental disorders has decreased in recent years, the majority of individuals with mental disorders still do not seek treatment [3]. With a global rise in the prevalence of mental disorders [4], it becomes even more important to understand the factors that influence whether an individual seeks help or not.

Psychotherapy is considered effective for treating many mental disorders [5]. However, there are several reasons that prevent individuals from seeking psychotherapy. These encompass structural challenges such as long waiting periods and underprovision, especially in rural areas [2,6,7], as well as individual reasons often related to stigma against mental disorders and psychotherapy [8,9]. Specifically, in self-reports, about one-third of respondents report negative attitudes and stigma toward psychotherapy, with more negative attitudes in male individuals and in individuals with lower education [10]. Similarly, almost a third of the respondents believed that professional care was worse than or equal to no help when faced with serious emotional problems in a large European study [8]. In recent years, attitudes toward professional help, particularly toward psychotherapeutic and psychiatric treatment, have improved both on a global level and in Germany [11,12]. In addition, perceived public stigma toward people with mental disorders has substantially decreased in the past 30 years [13]. At the same time, the individual wish for social distance toward people with depression slightly decreased but increased toward people with schizophrenia [11,14]. Furthermore, in many non-Western countries, there are still substantial negative attitudes toward individuals with mental illnesses [15,16]. Critically, it has been shown that negative attitudes toward help seeking and stigmatizing attitudes toward people with mental illness predict help-seeking behavior [7,17-20]. In summary, this means that although some progress has been made in recent years, there is still considerable stigma associated with both mental illness and its treatment, which needs to be addressed to increase the number of individuals who receive treatment in addition to reducing structural barriers.

Health literacy refers to “people’s knowledge, motivation and competencies to access, understand, appraise, and apply health information in order to make judgments and take decisions in everyday life concerning health care, disease prevention and health promotion to maintain or improve quality of life during the life course” [21]. Low health literacy has been associated with several negative health-related outcomes, including higher rates of hospitalization, lower medical treatment adherence, decreased quality of life, fewer health-promoting behaviors, and higher mortality [22-27]. Mental health literacy was developed from general health literacy and applies the concept to mental health. It is defined as “understanding how to obtain and maintain positive mental health; understanding mental disorders and their treatments; decreasing stigma related to mental disorders; and, enhancing help-seeking efficacy” [28]. High mental health literacy has a number of positive correlates such as mental well-being, physical activity, life satisfaction, healthy eating habits, and self-efficacy [29-31].

First studies have linked increased health literacy with lower stigma and more positive attitudes toward mental disorders and psychotherapy [12,19,30,32-34]. For example, Svensson and Hansson [32] have shown that a higher degree of mental health literacy is related to less stigma and a lower wish for social distance toward individuals with depression and to a lesser extent also toward individuals with schizophrenia. However, this pattern of results was not found in all studies. In a Swiss community sample of young adults, no direct association between mental health literacy and the wish for social distance was delineated. Instead, it was shown that increased mental health literacy may also lead to more stigma by enhancing stereotypes of individuals with mental disorders (eg, perceived dangerousness) due to higher endorsement of biogenetic causal explanation of mental disorders [35]. It has been demonstrated in diverse cultural contexts that specific interventions may enhance both general and mental health literacy [36-41]. Moreover, it has been shown in longitudinal studies that more positive treatment attitudes can predict more frequent mental health service use in the future [19,20]. Along with significantly reducing structural barriers, indirectly improving attitudes toward psychotherapeutic treatment via increased health literacy may therefore be an important starting point to ensure that more people with mental health problems receive professional help.

Although some studies have started to investigate the relationship between general and mental health literacy and attitudes toward psychotherapeutic treatments, there are still substantial gaps in the literature as well as some limitations of previous studies that need to be addressed in order to get a better understanding of how to increase the number of individuals seeking psychotherapeutic treatment. First, in past investigations, mental health literacy was often assessed via unlabeled vignettes [32,35] describing the symptoms of a specific disorder (eg, depression or schizophrenia), and high mental health literacy was assigned to those individuals who correctly identified these vignettes. While this covers a small aspect of the broad definition of mental health literacy, that is, the recognition of mental disorders, other important aspects, such as promoting positive mental health, are not included [28]. This limitation in

XSL•FO
RenderX

previous methodological approaches is why we decided to use an instrument with a broader scope to assess mental health literacy in the present investigation [42,43]. Second, the literature on general health literacy and how it relates to treatment attitudes is relatively sparse and there is a need to replicate and extend previous findings [34]. Third, there is some evidence to suggest that there are substantial interindividual differences in attitudes toward psychotherapy in the community [19,34]. However, the specific determinants of these differences need to be explored more in depth to identify potentially vulnerable segments of the population that could particularly benefit from interventions. For example, it has been shown that gender, age, level of education, and social status play an important role in accessing psychiatric and psychotherapeutic services and might therefore also be related to attitudes [2,10]. In addition, having a migration background or being part of an ethnic minority has been associated with access to mental health care [44-48]. Finally, negative attitudes might be improved by hearing about the positive experiences of others or making these experiences themselves [19,49-51], which is why we decided to include 2 questions on personal and family experience with psychotherapy.

To address these previous limitations and gaps in the literature, in this study, we aimed to explore the relationship between attitudes toward psychotherapy and both general and mental health literacy with comprehensive questionnaires and in a large sample. Furthermore, we aimed at exploring how additional factors including basic sociodemographic factors and experience with psychotherapy relate to attitudes toward psychotherapy in order to inform the future development of target group-specific interventions with the aim of reducing individual barriers to seeking professional help.

Methods

Ethical Considerations

The ethics committee of the Berlin Medical Association had no ethical or professional objections to the study protocol (reference Eth-39/22). Participants gave informed consent to take part in the study. They were not directly compensated for their participation by the independent nonprofit foundation Stiftung Gesundheitswissen. Data collection was conducted by the market research institute forsa (forsa Gesellschaft für Sozialforschung und statistische Analysen mBH), which was commissioned and financially compensated for this work by the nonprofit foundation Stiftung Gesundheitswissen. Importantly, Stiftung Gesundheitswissen did not have any influence on the data collected and the direction of the results. Forsa provided only anonymized data to the Stiftung Gesundheitswissen.

Data Acquisition

Data collection was conducted by forsa using the forsa.omninet panel. Forsa.omninet is a representative panel for the German-speaking population with internet access aged 14 years and older in Germany and currently has around 100,000 participants. Panelists are recruited exclusively offline, that is, by telephone within the framework of forsa.omniTel, a telephone omnibus survey of forsa, in which randomly selected individuals aged 14 years and older are interviewed daily. The composition of the panel is continuously monitored on the basis of key characteristics (eg, region, age, and gender) and recruitment is adjusted accordingly. In addition, approximately 1000 new individuals are recruited monthly. There is no possibility to apply for participation in the panel so the selection of the participants is solely controlled by a random sampling procedure. Inclusion criteria for this investigation were the age of 16 years or older; sufficient German knowledge; and being able to use a computer, laptop, or mobile device for the survey. The only exclusion criterion was not fulfilling the inclusion criteria. With the aim of collecting data from at least 2000 individuals, 3927 panelists aged 16 years and older were randomly drawn from the panel and invited to take part in the survey by email. Exactly 2000 participants completed the study, equaling a response rate of 50.9% (2000/3927). The invitation email contained brief information on the overall topic of the survey as well as the incentive. If necessary, the selected panelists were reminded about their participation in 2 further emails. Recruitment and data collection took place in September and October 2022 with a standardized questionnaire presented in the form of computer-assisted web interviews. The total questionnaire consisted of 155 items, of which not all are relevant to this investigation. For economic reasons, data collection for different research objectives was combined. Results are published in separate and independent publications [31,52].

To minimize the effects of nonresponse and potential bias due to systematically missing data from certain population segments, survey weights were calculated by forsa using an iterative proportional fitting approach with the following weight variables and combinations: (1) gender × age (16-29 years, 30-45 years, 46-64 years, or ≥65 years) × region (West Germany and Berlin, or East Germany) and (2) federal state. This procedure resulted in a single weighting factor for each individual applied before further statistical analyses. The weighting was informed by the population update of the German Federal Statistical Office (as of December 31, 2020).

Measures

Attitudes Toward Psychotherapy

Attitudes toward psychotherapy were assessed with the German version of the Questionnaire on Attitudes Towards Psychotherapeutic Treatment (QAPT) in its revised 2-factorial form [44,53,54]. The questionnaire consists of 11 items answered on a 4-point Likert scale ranging from “do not agree” to “agree.” The items address how the respondents perceive psychotherapeutic treatments. The 2 subscales are “positive attitudes” (6 items) and “non-acceptance of society” (5 items). The positive attitude scale reflects the individuals’ beliefs about the positive effects of psychotherapeutic treatment and the competence of psychotherapists, whereas the nonacceptance scale reflects the fear of stigma and social isolation from others. The second scale is originally labeled “acceptance of society” (“Akzeptanz der Gesellschaft” in German [54]). However, as the items of this scale are negatively coded, that is, higher values indicate lower acceptance, we decided to label the scale “nonacceptance of society” for easier interpretability. For each

XSL•FO
RenderX

scale, a mean value is calculated across all items. The mean value reaches from 1 to 4 with higher values indicating more positive attitudes and higher nonacceptance (eg, lower acceptance), respectively.

Health Literacy

General health literacy was assessed with the German translation of the short version of the European Health Literacy Survey instrument (HLS-EU-Q16 [55-57]). This 16-item questionnaire addresses the subjective difficulty in accessing, understanding, appraising, and applying information in the fields of health care, disease prevention, and health promotion. Respondents are asked to answer each item on a 4-point Likert scale with the options “very easy,” “fairly easy,” “fairly difficult,” and “very difficult.” An overall sum score is calculated for the dichotomized items (eg, 1=“fairly easy” and “very easy,” 0=“fairly difficult” and “very difficult”). The sum score ranges from 0 to 16, with higher values indicating higher health literacy. Groups of individuals with varying degrees of health literacy were determined using this sum score: inadequate health literacy (scores 0-8), problematic health literacy (scores 9-12), and adequate health literacy (scores 13-16).

Mental Health Literacy

Mental health literacy was assessed with the German version of the Mental Health Literacy Tool for the Workplace (MHL-W-G [42,43]). Instructions were slightly adapted so that they also apply to people who are not currently working. Specifically, participants were instructed to imagine they worked if this was not the case. The instrument presents 4 case vignettes, in each of which current symptoms and the circumstances of one individual are described. The vignettes differ in the nature of the mental health issue, its impact on workplace performance, and the gender of the individual. Importantly, the nature of the mental health issue is not explicitly labeled. For each vignette, the same 4 questions are answered covering important aspects of mental health literacy, namely the ability to recognize specific mental disorders, knowledge and beliefs about risk factors and prevention, knowledge and attitudes to facilitate help seeking, and knowledge and beliefs about mental health interventions. After reading each of the vignettes, participants are asked to rate their knowledge about what might be happening, how they could prevent the situation from worsening, what to say or do in the situation, and which resources and services might be helpful. Answers are given on a 5-point Likert scale ranging from “very low” to “very high”. A sum score is calculated by adding up the ratings of each of the 4 questions in each of the 4 vignettes (ranging from a minimum of 16 to a maximum of 80 points), with higher scores reflecting higher levels of self-reported mental health literacy. So far, no generally accepted cut-off scores have been developed to compare groups of individuals with varying levels of mental health literacy using this questionnaire, which is why we do not present stratified data here.

Experience with Psychotherapy

In addition, the participants were presented with 2 simple yes-or-no-questions to assess whether they have ever received any treatments for mental disorders themselves and whether any of their close family members have ever received treatment for mental disorders.

Sociodemographic Information

Participants provided basic sociodemographic information including gender (men, women) and age.

In addition, participants were asked to provide their highest level of formal education. Participants were then categorized into 3 groups: low, middle, and high level of education. The low level of education is equivalent to no formal education or basic secondary school (ohne Haupt-/Volksschulabschluss, Haupt-/Volksschulabschluss); the middle level of education is equivalent to intermediate secondary school (Mittlere Reife, Realschulabschluss, Fachschulreife, Abschluss der Polytechnischen Oberschule, Fachhochschulreife, Abschluss einer Fachoberschule); and the high level of education is equivalent to most advanced secondary school, for example, grammar schools to obtain a general or specialized university entrance qualification or a university degree (Abitur, allgemeine oder fachgebundene Hochschulreife, Fach-/Hochschulstudium).

Subjective social status was assessed with the German version of the MacArthur scale [58,59], in which the participants rank themselves relative to other members of society. They are presented with a metaphor, in which a ladder represents the social structure of society with the highest ladder rung representing the individuals in society with the highest status and the bottom rung presenting the individuals in society with the lowest status. The respondents are asked to identify the rung on which they place themselves. Three categories of subjective social status were determined according to the respondents’ answers, that is, low subjective social status (scores 1-4), middle subjective social status (scores 5-7), and high subjective social status (scores 8-10).

Migration status was assessed with a simple yes-or-no-question. The participants were asked to indicate if they have a migration background. They were instructed to answer with yes if they or one of their parents were not born in Germany.

Statistical Analysis

Statistical analyses were performed with the software SPSS (version 29.0.2.0; IBM). All analyses reported in the main manuscript were conducted with the weighted dataset (details on weighting can be found in the “data acquisition” paragraph). Results for the unweighted dataset can be found in the Multimedia Appendix 1. Cronbach α was calculated to determine the internal consistency of the attitudes and nonacceptance subscales of the QAPT as well as the HLS-EU-Q16 and MHL-W-G total scores. The resulting reliability coefficients were interpreted using the following rule of thumb: >.90 excellent, >.80 good, >.70 acceptable, >.60 questionable, >.50 poor, and <.50 unacceptable [60]. For all inferential analyses, a threshold of α=.05 was used as the significance level and 2-sided *P* values are reported. Associations between the 2 attitudes toward psychotherapy scales and both mental and general health literacy were assessed with Pearson correlations, including 95% CI. Likewise, for the associations between the 2 attitudes toward psychotherapy scales and the metric sociodemographic measure of age, Pearson

XSL•FO
RenderX

correlations were calculated. For each of the dichotomous sociodemographic measures and the 2 experiences with treatment for mental disorder questions (ie, gender, migration status, personal experience, and the experience of a close family member) separate 2-tailed *t* tests were performed for the 2 attitudes toward psychotherapy scales. Effect sizes for the *t* tests are reported as Cohen *d*, with the values of Cohen *d*=0.20, Cohen *d*=0.50, and Cohen *d* = 0.80 corresponding to small, medium, and large effects, respectively [61]. If the Levene test for equality of variances indicated that variances were not homogeneous, degrees of freedom were adjusted accordingly. For each categorial sociodemographic measure with more than 2-factor levels (ie, level of education and subjective social status), separate one-way ANOVAs were conducted. These included the sociodemographic measures as predictors and the attitudes toward psychotherapy scales as the dependent variables. Effect sizes for the ANOVAs are reported as η^2, with the values η^2=.01, η^2=.06, and η^2=.14 corresponding to small, medium, and large effects, respectively [62]. Significant main effects in the ANOVAs were followed up with Bonferroni-corrected pairwise *t* tests. Uncorrected *P* values and corrected α-thresholds are reported for these post hoc *t* tests. Participants could opt out of any question, resulting in a missing value for that item. Individuals with missing data for one item were excluded from all analyses including this item, for example, a scale or mean sum score was not calculated for an individual if a specific item relevant to that scale or mean sum score was missing, but it was calculated for other scales of the same questionnaire that do not contain that item if all other items were complete.

Results

Sample Characteristics and Descriptive Statistics

A sample of 2000 individuals was recruited. The sample characteristics can be found in Table 1 (cumulative percentages may not add up to precisely 100% and sample sizes may vary due to weighting and rounding). Descriptive statistics for the QAPT, HLS-EU-Q16, and MHL-W-G questionnaires across the entire sample are in Table 2. Detailed descriptive statistics for the QAPT attitudes and nonacceptance subscales are in Tables 3 and 4, respectively, and detailed descriptive statistics for each QAPT item can be found in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Characteristics of the weighted and unweighted sample.

Variable	Unweighted sample (n=2000), n (%)	Weighted sample (n=2000), n (%)
Gender		
Men	957 (47.9)	980 (49.0)
Women	1043 (52.1)	1020 (51.0)
Age group (years)		
16-29	461 (23.1)	355 (17.8)
30-45	433 (21.6)	475 (23.7)
46-64	599 (29.9)	655 (32.7)
>65	507 (25.4)	515 (25.8)
Level of education		
Low	349 (17.4)	349 (17.5)
Middle	847 (42.4)	861 (43.0)
High	792 (39.6)	782 (39.1)
Missing	12 (0.6)	8 (0.4)
Social status		
Low	312 (15.6)	315 (15.7)
Middle	1397 (69.8)	1389 (69.4)
High	291 (14.5)	296 (14.8)
Migration background		
No	1871 (93.5)	1876 (93.8)
Yes	129 (6.5)	124 (6.2)
Own experience with psychotherapy		
No	1376 (68.8)	1379 (68.9)
Yes	603 (30.1)	596 (29.8)
Missing	21 (1.1)	25 (1.2)
Family experience with psychotherapy		
No	1062 (53.1)	1053 (52.6)
Yes	870 (43.5)	879 (43.9)
Missing	68 (3.4)	68 (3.4)
General health literacy categories		
Inadequate	192 (9.6)	190 (9.5)
Problematic	540 (27.0)	545 (27.3)
Adequate	1255 (62.7)	1252 (62.6)
Missing	13 (0.7)	13 (0.6)

XSL•FO
RenderX

Table 2. Descriptive statistics of the QAPT[a], HLS-EU-Q16[b], and MHL-W-G[c] questionnaires.

Variable	Sample, n	Score, mean (SD)	Score, 95% CI of the mean	Score, median (range)
QAPT				
Positive attitudes toward psychotherapy	1994	3.28 (0.54)	3.26-3.31	3.33 (1-4)
Nonacceptance of society	1996	2.06 (0.61)	2.03-2.08	2.00 (1-4)
HLS-EU-Q16	1987	12.92 (3.03)	12.79-13.05	14.00 (0-16)
MHL-W-G	1994	52.03 (11.02)	51.55-52.51	52.00 (16-80)

[a]QAPT: Questionnaire on Attitudes Towards Psychotherapeutic Treatment.

[b]HLS-EU-Q16: European Health Literacy Survey instrument.

[c]MHL-W-G: Mental Health Literacy Tool for the Workplace (German version).

Table 3. Descriptive statistics for the QAPT[a] attitudes subscale across the sociodemographic groups.

Variable	Score, mean (SD)	Score, 95% CI of the mean	Score, median
Gender			
Men	3.24 (0.55)	3.20-3.27	3.33
Women	3.33 (0.52)	3.29-3.36	3.33
Level of education			
Low	3.20 (0.56)	3.14-3.25	3.17
Middle	3.23 (0.54)	3.19-3.27	3.17
High	3.38 (0.50)	3.25-3.42	3.50
Social status			
Low	3.22 (0.55)	3.16-3.28	3.17
Middle	3.28 (0.52)	3.25-3.31	3.33
High	3.35 (0.57)	3.30-3.43	3.50
Migration background			
No	3.29 (0.54)	3.25-3.31	3.33
Yes	3.25 (0.53)	3.15-3.34	3.33
Own experience with psychotherapy			
No	3.21 (0.54)	3.18-3.23	3.17
Yes	3.46 (0.48)	3.42-3.50	3.50
Family experience with psychotherapy			
No	3.22 (0.56)	3.19-3.25	3.17
Yes	3.36 (0.50)	3.32-3.39	3.33

[a]QAPT: Questionnaire on Attitudes Towards Psychotherapeutic Treatment.

XSL•FO
RenderX

Table 4. Descriptive statistics for the QAPT[a] nonacceptance subscale across the sociodemographic groups.

Variable	Score, mean (SD)	Score, 95% CI of the mean	Score, median
Gender			
Men	2.13 (0.60)	2.09-2.17	2.20
Women	1.99 (0.61)	1.95-2.02	2.00
Level of education			
Low	2.05 (0.61)	1.99-2.12	2.00
Middle	2.05 (0.60)	2.01-2.09	2.00
High	2.06 (0.61)	2.02-2.10	2.00
Social status			
Low	2.16 (0.59)	2.09-2.22	2.20
Middle	2.05 (0.60)	2.02-2.09	2.00
High	1.96 (0.63)	1.89-2.04	1.80
Migration background			
No	2.06 (0.60)	2.03-2.08	2.00
Yes	2.07 (0.64)	1.96-2.19	2.00
Own experience with psychotherapy			
No	2.08 (0.59)	2.05-2.11	2.00
Yes	1.99 (0.64)	1.94-2.04	2.00
Family experience with psychotherapy			
No	2.12 (0.60)	2.08-2.15	2.20
Yes	1.98 (0.61)	1.94-2.02	2.00

[a]QAPT: Questionnaire on Attitudes Towards Psychotherapeutic Treatment.

Reliability Analyses

The reliability analyses revealed an internal consistency of α=.83 (good) for the attitudes subscale and α=.74 (acceptable) for the nonacceptance subscale of the QAPT. The internal consistency was α=.80 (good) for the HLS-EU-Q16 and α=.92 (excellent) for the MHL-W-G.

Attitudes Toward Psychotherapy and General and Mental Health Literacy

The attitudes subscale of the QAPT correlated negatively with the nonacceptance subscale questionnaire (r=−0.17, 95% CI −0.21 to −0.13; P<.001). General health literacy correlated positively with the attitudes subscale (r=0.14, 95% CI 0.10-0.18; P<.001) and negatively with the nonacceptance subscale (r=−0.32, 95% CI −0.36 to −0.28; P<.001). Mental health literacy correlated positively with the attitudes subscale (r=0.18, 95% CI 0.13-0.22; P<.001) and negatively with the nonacceptance subscale (r=−0.23, 95% CI −0.27 to −0.19; P<.001).

Attitudes Toward Psychotherapy and Experience with Psychotherapy

Participants who have received treatment for a mental disorder reported significantly more positive attitudes toward psychotherapy than participants without treatment experience ($t_{1260.12}$=−10.40, P<.001, Cohen d=−0.49). Likewise, participants who have received treatment for a mental disorder reported significantly less nonacceptance regarding psychotherapy than participants without treatment experience ($t_{1050.95}$=3.06, P=.002, Cohen d=0.16).

Similarly, participants who have a close family member who received treatment for a mental disorder reported significantly more positive attitudes toward psychotherapy than participants without family treatment experience ($t_{1912.06}$=−5.66, P<.001, Cohen d=−0.26). Likewise, participants who have a close family member who received treatment for a mental disorder reported significantly less nonacceptance regarding psychotherapy than participants without family treatment experience (t_{1926}=4.77, P<.001, Cohen d=0.22).

Attitudes Toward Psychotherapy and Sociodemographic Measures

Age correlated negatively with the attitudes subscale (r=−0.11, 95% CI −0.16 to −0.07, P<.001) and negatively with the nonacceptance subscale (r=−0.08, 95% CI −0.12 to −0.03, P<.001).

There were significant gender differences for both subscales. Women reported higher positive attitudes toward psychotherapy than men (t_{1992}=−3.60, P<.001, Cohen d=−0.16). Women also reported lower nonacceptance than men (t_{1994}=5.29, P<.001, Cohen d=0.24).

XSL•FO
RenderX

Subjective social status had a significant effect on both the attitude subscale ($F_{2,1991}$=5.25, P=.005, η^2=.005) and the nonacceptance subscale ($F_{2,1993}$=7.71, P<.001, η^2=.008). Post hoc Bonferroni-corrected t tests for the attitudes subscale (Bonferroni-corrected α=.017) revealed significant differences between the low and high subjective social status groups (t_{607}=−3.08, P=.002, Cohen d=−0.25). Differences between the middle and high (t_{1678}=−2.34, P=.02, Cohen d=−0.15) and the low and middle subjective social status groups (t_{1698}=−1.82, P=.07, Cohen d=−0.11) were not significant after Bonferroni correction. Descriptively, positive attitudes were lowest in individuals with low subjective social status, followed by middle and high subjective social status. Post hoc Bonferroni-corrected t tests for the nonacceptance subscale (Bonferroni-corrected α=.017) revealed significant differences between the low and middle (t_{1699}=2.72, P=.007, Cohen d=0.17) and low and high subjective social status groups (t_{608}=3.88, P<.001, Cohen d=0.32). The difference between the middle and high subjective social status groups was not significant after Bonferroni correction (t_{1679}=2.32, P=.02, Cohen d=0.15). Descriptively, nonacceptance was highest in individuals with low subjective social status, followed by middle and high subjective social status.

Level of education had a significant effect on the attitudes subscale ($F_{2,1983}$=22.27, P<.001, η^2=.021). Post hoc Bonferroni-corrected t tests for the attitudes subscale (Bonferroni-corrected α=.017) revealed significant differences between the low and high educational levels (t_{1126}=−5.52, P<.001, Cohen d=−0.36) and the middle and high educational levels ($t_{1635.23}$=−5.82, P<.001, Cohen d=−0.29), but not between the low and middle educational levels (t_{1204}=−1.01, P=.32, Cohen d=−0.06). Descriptively, positive attitudes were highest in individuals with high educational levels, followed by middle and low educational levels.

There was no significant effect of level of education on the nonacceptance subscale ($F_{2,1984}$=0.02, P=.99, η^2<.001). Descriptively, participants with low, middle, and high levels of education reached similar scores.

There were no significant differences for the attitude subscale between individuals with and without a migration background (t_{1992}=0.73, P=.46, Cohen d=0.07). Likewise, there were no significant differences for the nonacceptance subscale between individuals with and without a migration background (t_{1994}=−0.29, P=.77, Cohen d=−0.03).

Results of the inferential analyses for the unweighted dataset can be found in the Multimedia Appendix 1. Importantly, weighting had no effect on the inferential conclusions except for one analysis. The post hoc t test for the difference between the low and high subjective social status groups did not reach significance in the unweighted dataset after the Bonferroni correction.

Discussion

Overview

This study investigated the associations between attitudes toward psychotherapeutic treatment and both general and mental health literacy in a large and representative sample of the general population in Germany. In addition, attitudes toward psychotherapeutic treatment were analyzed concerning interindividual differences in terms of experience with mental health treatment and sociodemographic factors.

Principal Findings

Overall, participants expressed largely positive attitudes toward psychotherapeutic treatment and reported relatively low perceived nonacceptance. However, significant interindividual differences were observed and are discussed below.

Importantly, while our results demonstrate several statistically significant results, the practical significance of these findings has to be considered before interpretation. All effect sizes (Cohen d and η^2) reported here fall into the small to medium range [61,62]. In addition, some of the CI of the means overlap. This means that the practical impact of our findings in real-world settings may be limited. Nonetheless, these findings point to potentially vulnerable sociodemographic groups in need of specific interventions to promote positive attitudes toward psychotherapy. The small effect sizes observed here suggest that while tailored interventions for specific groups may be more effective due to a better fit between the specific needs of a sociodemographic group and the intervention, there is also a potential overall benefit of broader general-population interventions [63,64].

The questionnaire scales used in this study reached acceptable to excellent internal consistency coefficients speaking for a high degree of interrelatedness of the scale items and their overall reliability which is an important prerequisite for the validity of the results [60]. Furthermore, weighting did not have a marked effect on the conclusions drawn from the statistical findings which highlights the generalizability of the findings.

General and Mental Health Literacy

Both higher general and higher mental health literacy were associated with more positive attitudes and lower nonacceptance of psychotherapeutic treatment. For mental health literacy, this finding aligns well with previous research [12,32,33,35]. Of note, these previous studies mostly used the simple identification of vignettes as a measure of mental health literacy. Here, we used a much broader instrument for mental health literacy more closely mirroring the definition of mental health literacy by also covering aspects of support, prevention, and use of resources [28,53]. For the first time, we show here that by including these broader aspects of mental health literacy positive associations with positive attitudes and negative associations with nonacceptance can be obtained, significantly extending previous observations in studies using vignette identification as a proxy for mental health literacy. Another novel contribution of our investigation is that we replicated initial findings of the associations between the QAPT and MHL-W-G scales

XSL•FO
RenderX

previously observed only in a small sample of working individuals in a substantially larger general population sample with similar effect sizes [43]. Notably, here, we also add to the literature by showing a positive association between general health literacy and the attitudes measure, which to our knowledge was previously only reported in one study [34]. Here, we replicate these initial findings in a larger sample and generalize them to another measure of general health literacy, namely the widely used HLS-EU-Q16. This suggests that individuals with higher levels of both general and mental health literacy are more likely to exhibit positive attitudes toward psychotherapy and perceive societal acceptance to be higher. Consequently, they might also be more likely to seek treatment when faced with mental problems [19]. Crucially, however, it has to be noted that our cross-sectional study design does not allow to infer causal associations between the investigated constructs (also refer to limitations below; [65]). Therefore, we cannot unequivocally conclude whether inadequate or problematic health literacy is indeed the antecedent and origin of less positive attitudes toward psychotherapy. We recommend future studies with appropriate study designs to test for causal effects, for example, randomized controlled trials. In addition, other potentially confounding variables should be carefully assessed and controlled for.

The present findings should also be interpreted in the context of mental health stigma in general. Although some progress in reducing stigma has been made in the past years substantial stigma against people with mental health problems remains [13,14]. Crucially, individuals with more stigmatizing attitudes show less active help-seeking behavior [17]. The promotion of health literacy with targeted interventions along with far-reaching antistigma campaigns could therefore also lead to higher rates of individuals seeking treatment [66]. Importantly, antistigma interventions may be more effective when focusing on psychosocial instead of biogenetic causes for mental disorders [35]. In addition to these interventions focusing on individual barriers, structural barriers to psychotherapeutic treatment (although not the focus of the present research) should not be disregarded and must be addressed as well. For example, waiting periods should be significantly reduced and provision should be increased, especially in rural areas [2,6,7].

Experience with Psychotherapy

Attitudes were more positive, and nonacceptance was lower in participants who have received treatment for mental disorders in the past or those who have a close relative with treatment experience. This is consistent with previous research [8,35]. It is noteworthy, that the treatment experience of a close family member may already improve the treatment attitudes of their relatives although with lower effect sizes than personal experience. Speaking openly about mental health treatment experiences and thus providing information about treatments might therefore be a valuable path to lowering psychotherapy stigma [49]. However, as no causal conclusions can be drawn from our study design, it is possible that positive treatment attitudes are not the result of personal or family treatment experience but that those with a priori more positive treatment attitudes are more likely to seek treatment.

Sociodemographic Factors

Women exhibited significantly more positive attitudes and reported lower perceived societal nonacceptance (ie, higher acceptance) of psychotherapeutic treatment compared to men. This finding aligns with previous research [10,67] and may be attributed to masculinity ideals and self-stigma, preventing men from seeking help for mental health problems [68,69].

Regarding age, more positive attitudes toward psychotherapeutic treatment were found in younger individuals, consistent with some [8] but not all previous investigations [10,70]. However, surprisingly, younger individuals also displayed higher levels of nonacceptance compared to older individuals, suggesting a discrepancy between perceived societal acceptance and individual attitudes. This pattern of results might be interpreted as higher awareness of the persisting mental health stigma in the community in younger individuals, although a generational shift toward less stigma is slowly taking place [14]. Notably, in our data, correlations with age were relatively small for both subscales. Integrating our results with previous literature, an overall heterogeneous picture emerges [8,10,70]: age seems to have no or a very small effect on attitudes toward psychotherapy and how it is perceived to be accepted by society. Future interventions to improve help-seeking for mental problems should therefore address all age groups.

Attitudes toward psychotherapy were also related to the subjective social status of the participants. Specifically, positive attitudes increased, and nonacceptance decreased with higher subjective social status, although the post hoc test between the high and low social status groups for the attitudes subscale narrowly missed significance in the unweighted dataset (refer to Multimedia Appendix 1). Social status has been identified as a major determinant of help-seeking behavior in the past [71,72]. However, more research is needed to determine the role of social status in attitudes toward psychotherapy. Critically, social status also plays a role in access to psychotherapy. Individuals with lower social status are less likely to be offered treatment than individuals with higher social status suggesting that structural barriers are also higher for these individuals than those with higher social status [73]. Potential financial barriers to access to psychotherapeutic treatment should also be considered but do not apply to the majority of our participants as the German public health insurance that insures about 90% of the population covers psychotherapeutic treatment for many diagnoses [74-76]. Treatment attitudes and treatment seeking should be carefully addressed in help-seeking interventions for individuals with lower social status and treatment should be provided equally to individuals of all social status groups. As general and mental health literacy were found to be low in individuals with lower subjective social status, these groups might particularly benefit from interventions [31,77]. There is first evidence to suggest that individuals with low social status might particularly benefit from culturally appropriate and tailored interventions that are based on scientific theories and incorporate different modalities, for example, using print materials and offering direct contact with the interventionists [63,64,78].

XSL•FO
RenderX

Higher educational levels were associated with more positive attitudes toward psychotherapy. However, there was no association between education and perceived acceptance of psychotherapy. In the literature, there is heterogeneous evidence concerning the role of education in attitudes toward mental health help-seeking [8,10,53]. Therefore, more research in this area is recommended to determine the role of education in help-seeking attitudes. Our results carefully suggest that programs directed at individuals with lower levels of education could be particularly beneficial.

Migration background was not related to neither positive attitudes nor perceived nonacceptance in our sample. This result is surprising given previous evidence of more negative attitudes toward psychotherapeutic treatment, and simultaneous higher endorsement of pharmacotherapy, in people with a migration background [44-46]. However, as we only included individuals with sufficient knowledge of the German language, it cannot be ruled out that our sample of individuals with a migration background was not representative of all individuals with a migration background in Germany (refer to *Limitations* section). Future studies could try to overcome this problem by providing translated questionnaires to reach a broader segment of the population with a migration background. In addition, substantial cultural differences have been delineated in the past, which is why we recommend investigating more thoroughly how the cultural backgrounds of the individuals with migration backgrounds are associated with their endorsement of psychotherapeutic treatment [44,79,80], but our data does not allow more fine-grained analyses. Notably, however, interventions to increase the use of mental health care may be more effective when tailored to the specific needs of minority groups and different cultures [47,63,64,78].

Overall, the findings presented here imply that more interventions to improve attitudes toward psychotherapy might be necessary and might be achieved by explicitly targeting both general and health literacy in vulnerable segments of the community. Our findings suggest that male individuals, those with higher age, lower social status, and lower levels of education might particularly benefit from interventions to promote positive attitudes toward psychotherapy. Of note, while a large number of interventions have already been developed and evaluated in the past, systematic reviews of the literature show substantial heterogeneity in the efficacy of these interventions and call for more high-quality evaluations [36-39]. In particular, more effort needs to be put into interventions in different cultures and across ethnic minority groups [40,41,80]. Notably, there are first promising results, for example concerning digital mental health interventions that are both cost-effective and easy to implement [39]. Web-based interventions appear to be generally efficacious when they include certain active components such as interactivity, target-group specificity, evidence-based content, and a sequential modular concept [81], for example as part of an e-learning platform [82,83]. While potential improvements in health literacy may not directly translate to increased help-seeking behavior, further research is needed to explore this connection [81].

Limitations

The findings presented here have to be interpreted in the light of some limitations. First, both attitudes and health literacy were assessed via self-reports. Future studies could focus on more applied measures rather than self-reports for a broader scope and higher ecological validity. Second, 6.2% of the participants in the weighted sample reported having a migration background. This is substantially lower than recent census data from Germany [84], suggesting that our sample may not be representative in terms of the proportion of individuals with a migration background. Third, giving our participants the option to opt-out from specific items might have resulted in systematically missing data which could have affected the generalizability of the results. However, the number of individuals using that option was relatively low, with the highest number (3.4%) for the item asking about family experience with psychotherapy which we speculate might have been the result of lack of knowledge. Importantly, the proportions of missing values were smaller for the main questionnaires used here (eg, 0.3% and 0.2% for the QAPT scales, refer to Table 2). As these proportions are substantially lower than the recommended threshold of 5%, no imputation analyses for missing values were conducted and all analyses were conducted with all available data [85]. Fourth, we relied on data collected on the web which might have resulted in lower data quality [86,87]. However, with this interview mode, a substantially larger sample size could be achieved, simultaneously reducing error variance. In addition, by working with the forsa.omninet panel a few issues common to web-based surveys could be mitigated. For example, participants were randomly selected and could not apply to take part in the study, people who rarely use the internet were also represented due to the panel recruitment via telephone, and multiple registrations of the same individuals were prevented. However, the sample is only representative of the population with internet access as data collection took place exclusively on the internet, which is why results may not be generalizable to individuals without internet access. Finally, as noted above, the present investigation used a cross-sectional study design. Consequently, the results cannot be interpreted in terms of causal relationships [65], which is why further research is needed with more sophisticated study designs including longitudinal and experimental studies that control for potentially confounding factors.

Summary and Conclusion

In this study, positive attitudes toward psychotherapy and perceived societal acceptance of psychotherapeutic treatment were investigated in a large general population sample in Germany. Our results highlight a positive association between these factors and both general and mental health literacy. Individuals who have been treated for mental problems or who have relatives with treatment experience reported more favorable attitudes and higher perceived acceptance of society than those without treatment experience. In terms of sociodemographic differences, it was shown that being a woman, younger age, higher social status, and higher levels of education were associated with more positive attitudes toward psychotherapeutic treatment. Being a woman, older age, and higher subjective social status were associated with higher perceived societal

XSL•FO
RenderX

acceptance of psychotherapy. Since previous studies suggest that target group-specific interventions have the potential to improve health literacy, future studies should focus on establishing whether tailored interventions designed to increase general or mental health literacy and aimed at individuals with deficits in this area may improve attitudes toward seeking help for mental problems.

Acknowledgments

This research was funded by the independent, nonprofit foundation Stiftung Gesundheitswissen. forsa Gesellschaft für Sozialforschung und statistische Analysen mBH was responsible for the recruitment of study participants and data acquisition.

Data Availability

The datasets generated and analyzed during this study are available from the Stiftung Gesundheitswissen on reasonable request.

Authors' Contributions

R Schröder contributed to methodology, formal analysis, writing - original draft, review, and editing. TH managed validation, writing, review, and editing. R Suhr handled supervision, writing reviews, and editing. LK contributed to conceptualization; methodology; supervision; writing, review, and editing.

Conflicts of Interest

All authors are employees of the independent, nonprofit foundation Stiftung Gesundheitswissen.

Multimedia Appendix 1
Supplementary analyses and tables.
[PDF File (Adobe PDF File), 273 KB - publichealth_v11i1e67078_app1.pdf]

References

1. Jacobi F, Höfler M, Siegert J, Mack S, Gerschler A, Scholl L, et al. Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: the mental health module of the German health interview and examination survey for adults (DEGS1-MH). Int J Methods Psychiatr Res 2014;23(3):304-319 [FREE Full text] [doi: 10.1002/mpr.1439] [Medline: 24729411]
2. Inanspruchnahme psychiatrischer und psychotherapeutischer Leistungen – individuelle Determinanten und regionale Unterschiede [Article in German]. Robert Koch-Institut. 2017. URL: https://www.rki.de/DE/Content/Gesundheitsmonitoring/Gesundheitsberichterstattung/GBEDownloadsJ/Focus/JoHM_04_2017_Psychiatrische_Psychotherapeutische_Leistungen.html [accessed 2025-01-11]
3. Brandstetter S, Dodoo-Schittko F, Speerforck S, Apfelbacher C, Grabe H, Jacobi F, et al. Trends in non-help-seeking for mental disorders in Germany between 1997-1999 and 2009-2012: a repeated cross-sectional study. Soc Psychiatry Psychiatr Epidemiol 2017;52(8):1005-1013. [doi: 10.1007/s00127-017-1384-y] [Medline: 28456871]
4. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022;9(2):137-150 [FREE Full text] [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
5. Huhn M, Tardy M, Spineli LM, Kissling W, Förstl H, Pitschel-Walz G, et al. Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: a systematic overview of meta-analyses. JAMA Psychiatry 2014;71(6):706-715. [doi: 10.1001/jamapsychiatry.2014.112] [Medline: 24789675]
6. Singer S, Maier L, Paserat A, Lang K, Wirp B, Kobes J, et al. Wartezeiten auf einen Psychotherapieplatz vor und nach der Psychotherapiestrukturreform. Psychotherapeut 2022;67(2):176-184. [doi: 10.1007/s00278-021-00551-0]
7. Tomczyk S, Schomerus G, Stolzenburg S, Muehlan H, Schmidt S. Ready, willing and able? an investigation of the theory of planned behaviour in help-seeking for a community sample with current untreated depressive symptoms. Prev Sci 2020;21(6):749-760 [FREE Full text] [doi: 10.1007/s11121-020-01099-2] [Medline: 32140825]
8. ten Have M, de Graaf R, Ormel J, Vilagut G, Kovess V, Alonso J, et al. Are attitudes towards mental health help-seeking associated with service use? Results from the European Study of Epidemiology of Mental Disorders. Soc Psychiatry Psychiatr Epidemiol 2010;45(2):153-163 [FREE Full text] [doi: 10.1007/s00127-009-0050-4] [Medline: 19381427]
9. Osman N, Michel C, Schimmelmann BG, Schilbach L, Meisenzahl E, Schultze-Lutter F. Influence of mental health literacy on help-seeking behaviour for mental health problems in the Swiss young adult community: a cohort and longitudinal case-control study. Eur Arch Psychiatry Clin Neurosci 2023;273(3):649-662 [FREE Full text] [doi: 10.1007/s00406-022-01483-9] [Medline: 36088495]
10. Petrowski K, Hessel A, Körner A, Weidner K, Brähler E, Hinz A. Attitudes toward psychotherapy in the general population. Psychother Psychosom Med Psychol 2014;64(2):82-85. [doi: 10.1055/s-0033-1361155] [Medline: 24515849]

XSL•FO
RenderX

11. Angermeyer MC, Matschinger H, Schomerus G. Attitudes towards psychiatric treatment and people with mental illness: changes over two decades. Br J Psychiatry 2013;203(2):146-151. [doi: 10.1192/bjp.bp.112.122978] [Medline: 23787060]
12. Angermeyer MC, van der Auwera S, Carta MG, Schomerus G. Public attitudes towards psychiatry and psychiatric treatment at the beginning of the 21st century: a systematic review and meta-analysis of population surveys. World Psychiatry 2017;16(1):50-61 [FREE Full text] [doi: 10.1002/wps.20383] [Medline: 28127931]
13. Angermeyer MC, Matschinger H, Carta MG, Schomerus G. Changes in the perception of mental illness stigma in Germany over the last two decades. Eur Psychiatry 2014;29(6):390-395. [doi: 10.1016/j.eurpsy.2013.10.004] [Medline: 24321774]
14. Schomerus G, Schindler S, Sander C, Baumann E, Angermeyer MC. Changes in mental illness stigma over 30 years - improvement, persistence, or deterioration? Eur Psychiatry 2022;65(1):e78 [FREE Full text] [doi: 10.1192/j.eurpsy.2022.2337] [Medline: 36328960]
15. Ghanean H, Nojomi M, Jacobsson L. A community study on attitudes to and knowledge of mental illness in Tehran. Open J Psychiatry 2015;05(01):26-30.
16. Atilola O. Level of community mental health literacy in sub-Saharan Africa: current studies are limited in number, scope, spread, and cognizance of cultural nuances. Nord J Psychiatry 2015;69(2):93-101. [doi: 10.3109/08039488.2014.947319] [Medline: 25151995]
17. Schnyder N, Panczak R, Groth N, Schultze-Lutter F. Association between mental health-related stigma and active help-seeking: systematic review and meta-analysis. Br J Psychiatry 2017;210(4):261-268. [doi: 10.1192/bjp.bp.116.189464] [Medline: 28153928]
18. Villatoro AP, DuPont-Reyes MJ, Phelan JC, Link BG. 'Me' vs. 'Them': how mental illness stigma influences adolescent help-seeking behaviors for oneself and recommendations for peers. Stigma Health 2022;7(3):300-310 [FREE Full text] [doi: 10.1037/sah0000392] [Medline: 36776352]
19. Bonabi H, Müller M, Ajdacic-Gross V, Eisele J, Rodgers S, Seifritz E, et al. Mental health literacy, attitudes to help seeking, and perceived need as predictors of mental health service use: a longitudinal study. J Nerv Ment Dis 2016;204(4):321-324 [FREE Full text] [doi: 10.1097/NMD.0000000000000488] [Medline: 27015396]
20. Mojtabai R, Evans-Lacko S, Schomerus G, Thornicroft G. Attitudes toward mental health help seeking as predictors of future help-seeking behavior and use of mental health treatments. Psychiatr Serv 2016;67(6):650-657. [doi: 10.1176/appi.ps.201500164] [Medline: 26876662]
21. Sørensen K, van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
22. Zheng M, Jin H, Shi N, Duan C, Wang D, Yu X, et al. The relationship between health literacy and quality of life: a systematic review and meta-analysis. Health Qual Life Outcomes 2018;16(1):201 [FREE Full text] [doi: 10.1186/s12955-018-1031-7] [Medline: 30326903]
23. Miller TA. Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis. Patient Educ Couns 2016;99(7):1079-1086 [FREE Full text] [doi: 10.1016/j.pec.2016.01.020] [Medline: 26899632]
24. Fan Z, Yang Y, Zhang F. Association between health literacy and mortality: a systematic review and meta-analysis. Arch Public Health 2021;79(1):119 [FREE Full text] [doi: 10.1186/s13690-021-00648-7] [Medline: 34210353]
25. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. Ann Intern Med 2011;155(2):97-107. [doi: 10.7326/0003-4819-155-2-201107190-00005] [Medline: 21768583]
26. Ranjbaran S, Chollou KM, Babazadeh T. Assessment of health literacy and health promoting behaviors among the urban adult population. Ethiop J Health Sci 2022;32(5):985-992 [FREE Full text] [doi: 10.4314/ejhs.v32i5.14] [Medline: 36262704]
27. Rezakhani Moghaddam H, Ranjbaran S, Babazadeh T. The role of e-health literacy and some cognitive factors in adopting protective behaviors of COVID-19 in Khalkhal residents. Front Public Health 2022;10:916362 [FREE Full text] [doi: 10.3389/fpubh.2022.916362] [Medline: 35942262]
28. Kutcher S, Wei Y, Coniglio C. Mental health literacy: past, present, and future. Can J Psychiatry 2016;61(3):154-158 [FREE Full text] [doi: 10.1177/0706743715616609] [Medline: 27254090]
29. Bjørnsen HN, Espnes GA, Eilertsen MB, Ringdal R, Moksnes UK. The relationship between positive mental health literacy and mental well-being among adolescents: implications for school health services. J Sch Nurs 2019;35(2):107-116 [FREE Full text] [doi: 10.1177/1059840517732125] [Medline: 28950750]
30. Suwanwong C, Jansem A, Intarakamhang U, Prasittichok P, Tuntivivat S, Chuenphittayavut K, et al. Modifiable predictors of mental health literacy in the educational context: a systematic review and meta-analysis. BMC Psychol 2024;12(1):378 [FREE Full text] [doi: 10.1186/s40359-024-01878-4] [Medline: 38965633]
31. König L, Hamer T, Suhr R. Die psychische Gesundheitskompetenz der Bevölkerung in Deutschland [Article in German]. Präv Gesundheitsf 2023;19(4):521-526. [doi: 10.1007/s11553-023-01079-8]
32. Svensson B, Hansson L. How mental health literacy and experience of mental illness relate to stigmatizing attitudes and social distance towards people with depression or psychosis: a cross-sectional study. Nord J Psychiatry 2016;70(4):309-313. [doi: 10.3109/08039488.2015.1109140] [Medline: 26643359]

XSL•FO
RenderX

33. Cheng H, Wang C, McDermott RC, Kridel M, Rislin JL. Self-stigma, mental health literacy, and attitudes toward seeking psychological help. Jour of Counseling & Develop 2018;96(1):64-74. [doi: 10.1002/jcad.12178]
34. Fleary SA, Joseph PL, Gonçalves C, Somogie J, Angeles J. The relationship between health literacy and mental health attitudes and beliefs. Health Lit Res Pract 2022;6(4):e270-e279 [FREE Full text] [doi: 10.3928/24748307-20221018-01] [Medline: 36350237]
35. Doll CM, Michel C, Betz LT, Schimmelmann BG, Schultze-Lutter F. The important role of stereotypes in the relation between mental health literacy and stigmatization of depression and psychosis in the community. Community Ment Health J 2022;58(3):474-486 [FREE Full text] [doi: 10.1007/s10597-021-00842-5] [Medline: 34037914]
36. Visscher BB, Steunenberg B, Heijmans M, Hofstede JM, Devillé W, van der Heide I, et al. Evidence on the effectiveness of health literacy interventions in the EU: a systematic review. BMC Public Health 2018;18(1):1414 [FREE Full text] [doi: 10.1186/s12889-018-6331-7] [Medline: 30594180]
37. Walters R, Leslie SJ, Polson R, Cusack T, Gorely T. Establishing the efficacy of interventions to improve health literacy and health behaviours: a systematic review. BMC Public Health 2020;20(1):1040 [FREE Full text] [doi: 10.1186/s12889-020-08991-0] [Medline: 32605608]
38. Ma KKY, Anderson JK, Burn A. Review: school-based interventions to improve mental health literacy and reduce mental health stigma - a systematic review. Child Adolesc Ment Health 2023;28(2):230-240. [doi: 10.1111/camh.12543] [Medline: 35005830]
39. Yeo G, Reich SM, Liaw NA, Chia EYM. The effect of digital mental health literacy interventions on mental health: systematic review and meta-analysis. J Med Internet Res 2024;26:e51268 [FREE Full text] [doi: 10.2196/51268] [Medline: 38421687]
40. Fernández-Gutiérrez M, Bas-Sarmiento P, Albar-Marín MJ, Paloma-Castro O, Romero-Sánchez JM. Health literacy interventions for immigrant populations: a systematic review. Int Nurs Rev 2018;65(1):54-64. [doi: 10.1111/inr.12373] [Medline: 28449363]
41. Meherali S, Punjani NS, Mevawala A. Health literacy interventions to improve health outcomes in low- and middle-income countries. Health Lit Res Pract 2020;4(4):e251-e266 [FREE Full text] [doi: 10.3928/24748307-20201118-01] [Medline: 33313935]
42. Moll S, Zanhour M, Patten SB, Stuart H, MacDermid J. Evaluating mental health literacy in the workplace: development and psychometric properties of a vignette-based tool. J Occup Rehabil 2017;27(4):601-611. [doi: 10.1007/s10926-017-9695-0] [Medline: 28120136]
43. Wulf IC, Ruhle SA. Psychische Gesundheitskompetenz am Arbeitsplatz [Article in German]. Diagnostica 2020;66(1):50-61. [doi: 10.1026/0012-1924/a000237]
44. Schulz W, Shin MA, Schmid-Ott G. Attitudes towards psychotherapy in South Korea and Germany : a cross-cultural comparative study. Nervenarzt 2018;89(1):51-57. [doi: 10.1007/s00115-016-0252-4] [Medline: 27909737]
45. Brandl EJ, Dietrich N, Mell N, Winkler JG, Gutwinski S, Bretz HJ, et al. Attitudes towards psychopharmacology and psychotherapy in psychiatric patients with and without migration background. BMC Psychiatry 2020;20(1):176 [FREE Full text] [doi: 10.1186/s12888-020-02585-1] [Medline: 32303265]
46. Verhulp EE, Stevens GWJM, Pels TVM, van Weert CMC, Vollebergh WAM. Lay beliefs about emotional problems and attitudes toward mental health care among parents and adolescents: Exploring the impact of immigration. Cultur Divers Ethnic Minor Psychol 2017;23(2):269-280. [doi: 10.1037/cdp0000092] [Medline: 27709973]
47. Lee-Tauler SY, Eun J, Corbett D, Collins PY. A systematic review of interventions to improve initiation of mental health care among racial-ethnic minority groups. Psychiatr Serv 2018;69(6):628-647. [doi: 10.1176/appi.ps.201700382] [Medline: 29716446]
48. Lehti V, Kieseppä V, Gissler M, Suvisaari J, Markkula N. Psychotherapy use among migrants: a register-based longitudinal study. J Epidemiol Community Health 2024;79(1):49-55 [FREE Full text] [doi: 10.1136/jech-2024-222330] [Medline: 39147569]
49. Braun-Koch K, Rief W, Teige-Mocigemba S. Changing attitudes towards psychotherapy via social observations: are similarities more important than discrepancies? BMC Psychol 2022;10(1):286 [FREE Full text] [doi: 10.1186/s40359-022-00952-z] [Medline: 36461123]
50. Braun-Koch K, Rief W. Maintenance vs. change of negative therapy expectation: an experimental investigation using video samples. Front Psychiatry 2022;13:836227 [FREE Full text] [doi: 10.3389/fpsyt.2022.836227] [Medline: 35444567]
51. Albarracin D, Shavitt S. Attitudes and attitude change. Annu Rev Psychol 2018;69:299-327. [doi: 10.1146/annurev-psych-122216-011911] [Medline: 28841390]
52. König L, Kuhlmey A, Suhr R. Digital health literacy of the population in Germany and its association with physical health, mental health, life satisfaction, and health behaviors: nationally representative survey study. JMIR Public Health Surveill 2024;10:e48685 [FREE Full text] [doi: 10.2196/48685] [Medline: 38381497]
53. Schmid-Ott G, Reibold S, Ernst G, Niederauer H, Künsebeck H, Schulz W, et al. Development of a questionnaire to assess attitudes towards psychotherapeutic treatment. Dermatol Psychosom 2004;4(4):187-193. [doi: 10.1159/000075905]

XSL•FO
RenderX

54. Ditte D, Schulz W, Schmid-Ott G. Attitude towards psychotherapy in the Russian population and in the population with a Russian/Soviet cultural background in Germany. A pilot study. Nervenarzt 2006;77(1):64-72. [doi: 10.1007/s00115-005-1889-6] [Medline: 15776258]
55. Pelikan JM, Röthlin F, Ganahl K. Measuring comprehensive health literacy in general populations: validation of instrument, indices and scales of the HLS-EU study. Bethesda, MD: Hyatt Regency Bethesda; 2014.
56. Sørensen K, van den Broucke S, Pelikan JM, Fullam J, Doyle G, Slonska Z, et al. Measuring health literacy in populations: illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). BMC Public Health 2013;13:948 [FREE Full text] [doi: 10.1186/1471-2458-13-948] [Medline: 24112855]
57. Jordan S, Hoebel J. Health literacy of adults in Germany: findings from the German Health Update (GEDA) study. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015;58(9):942-950. [doi: 10.1007/s00103-015-2200-z] [Medline: 26227894]
58. Hoebel J, Müters S, Kuntz B, Lange C, Lampert T. Measuring subjective social status in health research with a German version of the MacArthur Scale. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2015;58(7):749-757. [doi: 10.1007/s00103-015-2166-x] [Medline: 25986532]
59. Adler NE, Epel ES, Castellazzo G, Ickovics JR. Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. Health Psychol 2000;19(6):586-592. [doi: 10.1037//0278-6133.19.6.586] [Medline: 11129362]
60. Gliem J, Gliem R. Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales. Indiana University. URL: https://hdl.handle.net/1805/344 [accessed 2025-01-11]
61. Cohen J. A power primer. Psychol Bull 1992;112(1):155-159. [doi: 10.1037//0033-2909.112.1.155] [Medline: 19565683]
62. Ellis PD. The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results 1st Edition. Cambridge, England: Cambridge University Press; 2010.
63. Noar S, Benac C, Harris M. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
64. Lustria MLA, Noar SM, Cortese J, van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
65. Wang X, Cheng Z. Cross-sectional studies: strengths, weaknesses, and recommendations. Chest 2020;158(1S):S65-S71. [doi: 10.1016/j.chest.2020.03.012] [Medline: 32658654]
66. Hansson L, Stjernswärd S, Svensson B. Changes in attitudes, intended behaviour, and mental health literacy in the Swedish population 2009-2014: an evaluation of a national antistigma programme. Acta Psychiatr Scand 2016;134 Suppl 446:71-79. [doi: 10.1111/acps.12609] [Medline: 27426648]
67. Nam SK, Chu HJ, Lee MK, Lee JH, Kim N, Lee SM. A meta-analysis of gender differences in attitudes toward seeking professional psychological help. J Am Coll Health 2010;59(2):110-116. [doi: 10.1080/07448481.2010.483714] [Medline: 20864437]
68. Yousaf O, Popat A, Hunter MS. An investigation of masculinity attitudes, gender, and attitudes toward psychological help-seeking. Psychol. Men Masculinity 2015;16(2):234-237. [doi: 10.1037/a0036241]
69. Levant RF, Stefanov DG, Rankin TJ, Halter MJ, Mellinger C, Williams CM. Moderated path analysis of the relationships between masculinity and men's attitudes toward seeking psychological help. J Couns Psychol 2013;60(3):392-406. [doi: 10.1037/a0033014] [Medline: 23668728]
70. Pepin R, Segal DL, Coolidge FL. Intrinsic and extrinsic barriers to mental health care among community-dwelling younger and older adults. Aging Ment Health 2009;13(5):769-777. [doi: 10.1080/13607860902918231] [Medline: 19882416]
71. Choi N, Miller MJ. Social class, classism, stigma, and college students‘ attitudes toward counseling. Couns. Psychol 2018;46(6):761-785. [doi: 10.1177/0011000018796789]
72. Eisenberg D, Golberstein E, Gollust SE. Help-seeking and access to mental health care in a university student population. Med Care 2007;45(7):594-601. [doi: 10.1097/MLR.0b013e31803bb4c1] [Medline: 17571007]
73. Niemeyer H, Knaevelsrud C. Socioeconomic status and access to psychotherapy. J Clin Psychol 2023;79(4):937-953. [doi: 10.1002/jclp.23449] [Medline: 36251952]
74. Epping J, Muschik D, Geyer S. Social inequalities in the utilization of outpatient psychotherapy: analyses of registry data from German statutory health insurance. Int J Equity Health 2017;16(1):147 [FREE Full text] [doi: 10.1186/s12939-017-0644-5] [Medline: 28814311]
75. Sozialleistungen: Angaben zur Krankenversicherung (Ergebnisse des Mikrozensus) [Article in German]. Fachserie 13 Reihe 1.1. Wiesbaden: Statistisches Bundesamt; 2019. URL: https://www.statistischebibliothek.de/mir/servlets/MCRFileNodeServlet/DEHeft_derivate_00014767/2130110119004_Rev2015.pdf [accessed 2025-01-11]
76. Sareen J, Jagdeo A, Cox BJ, Clara I, ten Have M, Belik S, et al. Perceived barriers to mental health service utilization in the United States, Ontario, and the Netherlands. Psychiatr Serv 2007;58(3):357-364. [doi: 10.1176/ps.2007.58.3.357] [Medline: 17325109]
77. Schaeffer D, Berens E, Vogt D. Health literacy in the German population. Dtsch Arztebl Int 2017;114(4):53-60 [FREE Full text] [doi: 10.3238/arztebl.2017.0053] [Medline: 28211318]

XSL•FO
RenderX

78. Stormacq C, Wosinski J, Boillat E, van den Broucke S. Effects of health literacy interventions on health-related outcomes in socioeconomically disadvantaged adults living in the community: a systematic review. JBI Evid Synth 2020;18(7):1389-1469. [doi: 10.11124/JBISRIR-D-18-00023] [Medline: 32813388]
79. Balkir Neftçi N, Barnow S. One size does not fit all in psychotherapy: understanding depression among patients of turkish origin in Europe. Noro Psikiyatr Ars 2016;53(1):72-79 [FREE Full text] [doi: 10.5152/npa.2016.12666] [Medline: 28360770]
80. Selkirk M, Quayle E, Rothwell N. A systematic review of factors affecting migrant attitudes towards seeking psychological help. J Health Care Poor Underserved 2014;25(1):94-127. [doi: 10.1353/hpu.2014.0026] [Medline: 24509015]
81. Brijnath B, Protheroe J, Mahtani KR, Antoniades J. Do web-based mental health literacy interventions improve the mental health literacy of adult consumers? Results from a systematic review. J Med Internet Res 2016;18(6):e165 [FREE Full text] [doi: 10.2196/jmir.5463] [Medline: 27323907]
82. König L, Suhr R. The effectiveness of publicly available web-based interventions in promoting health app use, digital health literacy, and media literacy: pre-post evaluation study. J Med Internet Res 2023;25:e46336 [FREE Full text] [doi: 10.2196/46336] [Medline: 38048146]
83. König L, Marbach-Breitrück E, Engler A, Suhr R. The development and evaluation of an e-Learning course that promotes digital health literacy in school-age children: pre-post measurement study. J Med Internet Res 2022;24(5):e37523 [FREE Full text] [doi: 10.2196/37523] [Medline: 35576572]
84. Bevölkerung mit Migrationshintergrund - Ergebnisse des Mikrozensus 2021: Fachserie 1 Reihe 2.2 [Article in German]. Endergebnisse. 2021. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Publikationen/Downloads-Migration/migrationshintergrund-endergebnisse-2010220217004.pdf?__blob=publicationFile [accessed 2025-01-11]
85. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol 2017;17(1):162 [FREE Full text] [doi: 10.1186/s12874-017-0442-1] [Medline: 29207961]
86. Zhang X, Kuchinke L, Woud ML, Velten J, Margraf J. Survey method matters: online/offline questionnaires and face-to-face or telephone interviews differ. Comput. Hum. Behav 2017;71:172-180. [doi: 10.1016/j.chb.2017.02.006]
87. Heerwegh D, Loosveldt G. Face-to-face versus web surveying in a high-internet-coverage population: differences in response quality. Public Opin. Q 2008;72(5):836-846.

Abbreviations

HLS-EU-Q16: European Health Literacy Survey instrument
MHL-W-G: Mental Health Literacy Tool for the Workplace (German version)
QAPT: Questionnaire on Attitudes Towards Psychotherapeutic Treatment

Edited by A Mavragani; submitted 04.10.24; peer-reviewed by C Timm, S Ranjbaran, J Rosario; comments to author 04.11.24; revised version received 18.12.24; accepted 05.01.25; published 24.01.25.

Please cite as:
Schröder R, Hamer T, Suhr R, König L
Attitudes Toward Psychotherapeutic Treatment and Health Literacy in a Large Sample of the General Population in Germany: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e67078
URL: https://publichealth.jmir.org/2025/1/e67078
doi:10.2196/67078
PMID:39854696

©Rebekka Schröder, Tim Hamer, Ralf Suhr, Lars König. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Nationwide Trends in Screen Time and Associated Risk Factors by Family Structures Among Adolescents, 2008-2022: Nationwide Cross-Sectional Study

Seokjun Kim[1,2*], MD; Hyesu Jo[2,3*], MS; Yejun Son[2,4*], MS; Min Kyung Shin[5*], MD; Kyeongmin Lee[2,3], MS; Jaeyu Park[2,4], PhD; Hayeon Lee[2,6], PhD; Lee Smith[7], PhD; Elena Dragioti[8], PhD; Guillaume Fond[9], MD, PhD; Laurent Boyer[9], MD, PhD; Guillermo F López Sánchez[10], PhD; Mark A Tully[11], PhD; Masoud Rahmati[9,12,13], PhD; Damiano Pizzol[14,15], PhD; Selin Woo[1,2], PhD; Dong Keon Yon[1,2,3,4,16], MD, PhD

[1]Department of Medicine, Kyung Hee University College of Medicine, Seoul, Republic of Korea

[2]Center for Digital Health, Medical Science Research Institute, Kyung Hee University College of Medicine, Seoul, Republic of Korea

[3]Department of Regulatory Science, Kyung Hee University, Seoul, Republic of Korea

[4]Department of Precision Medicine, Kyung Hee University College of Medicine, Seoul, Republic of Korea

[5]Department of Medicine, Wonkwang University College of Medicine, Iksan, Republic of Korea

[6]Department of Electronics and Information Convergence Engineering, Kyung Hee University, Yongin, Republic of Korea

[7]Centre for Health, Performance and Wellbeing, Anglia Ruskin University, Cambridge, United Kingdom

[8]Research Laboratory Psychology of Patients, Families, and Health Professionals, Department of Nursing, School of Health Sciences, University of Ioannina, Ioannina, Greece

[9]CEReSS-Health Service Research and Quality of Life Center, Assistance Publique-Hôpitaux de Marseille, Aix-Marseille University, Marseille, France

[10]Division of Preventive Medicine and Public Health, Department of Public Health Sciences, School of Medicine, University of Murcia, Murcia, Spain

[11]School of Medicine, Ulster University, Londonderry, United Kingdom

[12]Department of Physical Education and Sport Sciences, Faculty of Literature and Human Sciences, Lorestan University, Khoramabad, Iran

[13]Department of Physical Education and Sport Sciences, Faculty of Literature and Humanities, Vali-E-Asr University of Rafsanjan, Rafsanjan, Iran

[14]Health Unit, Eni, San Donato Milanese, Italy

[15]Health Unit, Eni, Maputo, Mozambique

[16]Department of Pediatrics, Kyung Hee University Medical Center, Kyung Hee University College of Medicine, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Dong Keon Yon, MD, PhD
Department of Medicine
Kyung Hee University College of Medicine
23 Kyungheedae-ro
Dongdaemun-gu
Seoul, 02447
Republic of Korea
Phone: 82 269352476
Fax: 82 5044780201
Email: yonkkang@gmail.com

Abstract

Background: Although understanding long-term trends in adolescent screen time and the influence of family structure is essential, there is a lack of research addressing these issues comprehensively.

Objective: This study aimed to conduct comprehensive investigations into adolescent screen time before and during the COVID-19 pandemic, with a particular focus on family structures.

Methods: This study used nationwide, large-scale data from the Korea Youth Risk Behavior Web-Based Survey from South Korea. We aimed to indicate the changes in adolescent screen time over 15 years from 2008 to 2022. Weighted linear regression was used to analyze annual trends in screen time before and during the pandemic, and stratified analyses were conducted to examine associated risk factors across different family structures.

XSL•FO
RenderX

Results: This study used data from a total of 836,972 individuals (n=403,456, 48.2% women), with an age range of 12-18 years. The analysis revealed an overall increase in screen time prepandemic (β=8.06, 95% CI 7.74-8.39), with a notable increase observed at the onset of the pandemic (β=162.06, 95% CI 159.49-164.64). Among diverse family structures, the orphanage group showed the most substantial increase in screen time during the pandemic (β_{diff}=221.90, 95% CI 159.62-284.17). Risk factors associated with screen time during the pandemic varied by family structure. Notably, the nuclear family group presented distinct screen time–related risk factors, including grade, region of residence, physical activity frequency, sadness and despair, and the highest education level of parents.

Conclusions: There has been a notable increase in average screen time among adolescents since the onset of the pandemic, with the orphanage group exhibiting a pronounced trend. The risk factors associated with screen time during the pandemic varied for each family structure. Findings from this study suggest that the implementation of individualized measures tailored to each family structure should be adopted to effectively address the increased issue of adolescent screen time since the pandemic.

(JMIR Public Health Surveill 2025;11:e57962) doi:10.2196/57962

KEYWORDS

adolescents; family type; pandemic; screen time; South Korea; sedentary activity; risk factor; mobile phone

Introduction

There was a declaration of COVID-19 as a global pandemic by the World Health Organization in March 2020 [1]. With the proclamation, countries worldwide have advocated for indoor living, remote work, and web-based learning to prevent the spread of COVID-19. For example, the National Health Service in England advised individuals to remain at home and participate in remote work unless necessary, while the Chinese government implemented restrictions on public transport and postponed the reopening of schools following the vacation period [2]. Globally, universities and educational institutions were closed, prompting a rapid digital transition in the education system [3,4]. The increase in indoor time since the pandemic has likely induced a concurrent increase in screen time [5,6]. Throughout the pandemic, social distancing and economic hardships restricted daily activities and social interactions, negatively impacting individuals' mental health worldwide [7,8]. Among preschool children, psychological resilience has been identified as a factor that amplifies the positive outcomes associated with the use of digital technologies [9]. The need for discussing how the utilization of digital technologies evolves in response to mental health changes brought about by the pandemic is increasingly evident. In such circumstances, specific focus is placed on the increased screen time of adolescents, due to the association with adolescent depression and mental health [10,11]. Global guidelines, such as Canada's 24-Hour Movement Guidelines for Children and Youth, recommend limiting recreational screen time to 2 hours or less per day, highlighting the potential risks of excessive screen use [12]. With the onset of the COVID-19 pandemic, there is growing global attention to screen time among adolescents.

In South Korea, concerns about excessive screen time among adolescents predate the pandemic. The "Shutdown Law," implemented in 2011, restricted web-based gaming for those younger than 16 years old during late-night hours to prevent gaming addiction but was abolished in 2021 [13]. A recent study highlighted a significant increase in screen time among Korean adolescents, with teenagers spending an average of 303.66 minutes daily during the pandemic, a substantial rise from 133.24 minutes in the prepandemic period [14]. The screen time of adolescents appears to be influenced by their family structure, likely due to variations in parental monitoring, restrictions on screen use, and the promotion of outdoor activities [15]. However, studies on this topic are limited by small sample sizes, a narrow focus on specific family structures, and a lack of consideration for the significant changes brought about by the COVID-19 pandemic [15,16]. This underscores the need for a comprehensive analysis of long-term trends in adolescent screen time, particularly about diverse family structures, both before and during the pandemic.

Therefore, this study investigated the 15-year trends in adolescent screen time and potential risk factors based on family structures, using a nationwide dataset of Korean adolescents surveyed between 2008 and 2022. The primary aim of this study was to assess the changes in adolescent screen time trends and associated risk factors in each family structure since the onset of the pandemic.

Methods

Study Design and Population

This study used the Korea Youth Risk Behavior Web-Based Survey (KYRBS). KYRBS is a large-scale, nationwide survey covering health-related behaviors of Korean adolescents [17]. The KYRBS is designed to collect data from middle and high school students (grades 7 to 12) aged 12 to 18 years enrolled in public and private schools throughout South Korea. To achieve national representativeness, the survey uses a stratified, multistage cluster sampling approach [18]. Schools are chosen as the primary sampling units based on regional and school-type stratifications, followed by a random selection of classes within these schools [18]. This methodology ensures a balanced representation of students from both urban and rural areas. Administered annually by the Korea Disease Control and Prevention Agency, the survey has consistently maintained an average response rate of 95.7% over the past 15 years. Students voluntarily completed a web-based anonymous questionnaire following a standardized protocol in a computer lab in each school [19].

To ensure the data's representativeness, sampling weights were determined and applied. The weighting process considers the

XSL•FO
RenderX

probability of selection at each stage of sampling (schools and classes), adjusts for nonresponse, and incorporates poststratification adjustments to align with the overall student population distributions by grade, sex, and school type for each survey year. To avoid overrepresentation, extreme weights were capped using interquartile range techniques. These adjustments ensure that the results reliably represent the health behaviors of the adolescent population across South Korea. In this study, an initial cohort of 992,702 respondents completed the questionnaires from 2008 to 2022. Following the exclusion of participants with missing data, the final dataset comprised 836,972 samples (84.3% of the initial raw data). The exclusion of missing data was conducted systematically to maintain dataset robustness and minimize bias.

Ethical Considerations

The study protocol was approved by the Institutional Review Board of the Korea Disease Control and Prevention Agency (2014-06EXP-02-P-A), and all participants provided written informed consent. Additionally, the study complied with the Population Health Promotion Act 19 (117058) as mandated by the Korean government. This study was conducted following the principles of the Declaration of Helsinki.

Endpoints

We aimed to investigate trends in screen time and explore associated risk factors, with a focus on the diverse family structures of adolescents. The COVID-19 pandemic period was defined as 2020-2022, starting with the first observed COVID-19 case in South Korea on January 20, 2020, and aligning with widespread public health measures [20,21]. The prepandemic years were grouped into 6-year intervals, each spanning two consecutive years, as follows: 2008-2009, 2010-2011, 2012-2013, 2014-2015, 2016-2017, and 2018-2019. This classification allows for a clear comparison of trends before and during the pandemic.

Screen time was defined as the average daily time spent on the internet or using smartphones for noneducational purposes. Participants were asked: “What was your average daily internet or smartphone use on weekdays in the past 30 days?” and “What was your average daily internet or smartphone use on weekends in the past 30 days?” The overall average screen time was calculated by combining weekday and weekend use, weighted to reflect their relative proportions within a week [19].

Family structures were categorized into 4 groups: nuclear family, living with relatives, living alone, and orphanage. A nuclear family was defined as adolescents living with one or both parents, regardless of biological relationships. The “living with relatives” category included adolescents residing with extended family members, such as grandparents or uncles or aunts, excluding parents. “Living alone” referred to those managing households independently without the presence of any family members. The “orphanage” category encompassed adolescents living in child welfare facilities. These classifications are based on established definitions in prior studies on family dynamics and adolescent health [22].

Covariate Definitions

A total of 11 covariates were included in the analytical models: sex, BMI group (underweight, normal, overweight, and obese), grade (7th to 9th [middle-school] and 10th to 12th [high-school]), region of residence (urban and rural) [23,24], smoking status, alcoholic consumption, school performance (low, middle-low, middle, middle-high, and high), sexual experience, physical activity frequency (lower activity, moderate activity, and higher activity), sadness and despair, and highest education level of parents (high school or lower and college or higher). BMI was calculated using self-reported height and weight data. Conforming to the 2017 Korean National Growth Charts, BMI was categorized into the following four groups: underweight (<5%), normal (5% to 84 %), overweight (85% to 94%), and obese (≥95%) [19]. School performance was categorized into the following five groups according to the self-reports of students: low (<20%), middle-low (20% to 39%), middle (40% to 59%), middle-high (60% to 79%), and high (≥80%). Physical activity frequency was classified into the following three groups based on the engagement in vigorous aerobic and resistance training more than three days per week: lower activity (neither activity is done for more than three days per week), moderate activity (either one activity), and higher activity (both activities). The definitions of variables were extracted from peer-reviewed literature [19].

Statistical Analysis

We used a weighted linear regression model to compute β coefficient and β difference (β_{diff}) along with a 95% CI [25]. This strategy was selected for trend analysis and associated factor analysis. In the analysis of screen time trends, the β coefficient expressed the annual trend of screen time among adolescents in prepandemic or pandemic. β_{diff} was used to indicate the trend difference between the divided year groups based on the onset of the pandemic. In the analysis of screen time-associated factors, the β coefficient represents the correlation between screen time and each biological, social, and familial factor. β_{diff} was also used to reveal the difference in correlation between the year groups divided by the onset of the pandemic. In addition, the associations between screen time and each variable were recalculated after adjusting for potential influencing factors, including age, sex, BMI group, grade, region of residence, smoking status, alcoholic consumption, school performance, sexual experience, physical activity frequency, sadness and despair, and highest education level of parents. Statistical significance, as applied in all the above analyses, was determined by a 2-sided *P* value less than .05. Statistical analysis in this study was conducted using SAS software (version 9.4; SAS Institute Inc).

Results

This study was conducted using KYRBS spanning from 2008 to 2022. A total of 836,972 individuals (n=403,456, 48.20% women) with an average age of 15.01 (SD 1.75) years were finally selected for the analysis following the missing data handling. Baseline characteristics of the raw data are indicated in Table 1.

XSL•FO
RenderX

Table 1. General characteristics of participants in KYRBS[a], 2008-2022 (n=836,972).

	Total	Prepandemic						Pandemic
		2008-2009	2010-2011	2012-2013	2014-2015	2016-2017	2018-2019	2020-2022
Overall, n (%)	836,972 (100)	118,457 (14.15)	121,079 (14.47)	111,432 (13.31)	109,654 (13.1)	115,169 (13.76)	101,771 (12.16)	159,410 (19.05)
Age (years), mean (SD)	15.01 (1.75)	15.03 (1.75)	15.08 (1.75)	14.90 (1.76)	14.95 (1.75)	15.00 (1.74)	15.00 (1.78)	15.10 (1.74)
Sex, n (%)								
Men	433,516 (51.8)	63,372 (53.5)	63,168 (52.17)	58,694 (52.67)	56,442 (51.47)	58,685 (50.96)	51,467 (50.57)	81,688 (51.24)
Women	403,456 (48.2)	55,085 (46.5)	57,911 (47.83)	52,738 (47.33)	53,212 (48.53)	56,484 (49.04)	50,304 (49.43)	77,722 (48.76)
BMI group[b], n (%)								
Underweight	65,467 (7.82)	10,586 (8.94)	10,485 (8.66)	8649 (7.76)	8390 (7.65)	7972 (6.92)	6850 (6.73)	12,535 (7.86)
Normal	608,013 (72.64)	89,246 (75.34)	91,583 (75.64)	83,935 (75.32)	81,264 (74.11)	82,875 (71.96)	71,846 (70.6)	107,264 (67.29)
Overweight	68,622 (8.2)	8068 (6.81)	8428 (6.96)	8496 (7.62)	8849 (8.07)	10,046 (8.72)	9315 (9.15)	15,420 (9.67)
Obese	69,758 (8.33)	6281 (5.3)	6708 (5.54)	7146 (6.41)	7783 (7.1)	10,878 (9.45)	10,875 (10.69)	20,087 (12.6)
Grade, n (%)								
7th	140,196 (16.75)	19,698 (16.63)	20,046 (16.56)	18,516 (16.62)	17,531 (15.99)	18,718 (16.25)	16,875 (16.58)	28,812 (18.07)
8th	145,377 (17.37)	20,953 (17.69)	21,030 (17.37)	19,272 (17.29)	18,982 (17.31)	18,990 (16.49)	17,324 (17.02)	28,826 (18.08)
9th	148,190 (17.71)	21,475 (18.13)	21,644 (17.88)	19,804 (17.77)	19,702 (17.97)	19,492 (16.92)	17,771 (17.46)	28,302 (17.75)
10th	132,805 (15.87)	18,780 (15.85)	19,190 (15.85)	17,531 (15.73)	17,055 (15.55)	19,018 (16.51)	15,835 (15.56)	25,396 (15.93)
11th	135,789 (16.22)	19,028 (16.06)	19,731 (16.3)	17,996 (16.15)	17,843 (16.27)	19,633 (17.05)	16,415 (16.13)	25,143 (15.77)
12th	134,615 (16.08)	18,523 (15.64)	19,438 (16.05)	18,313 (16.43)	18,541 (16.91)	19,318 (16.77)	17,551 (17.25)	22,931 (14.38)
Region of residence, n (%)								
Urban	541,928 (64.75)	71,150 (60.06)	75,778 (62.59)	74,454 (66.82)	72,549 (66.16)	76,221 (66.18)	67,554 (66.38)	104,222 (65.38)
Rural	295,044 (35.25)	47,307 (39.94)	45,301 (37.41)	36,978 (33.18)	37,105 (33.84)	38,948 (33.82)	34,217 (33.62)	55,188 (34.62)
Smoking status, n (%)								
No	798,876 (95.45)	102,201 (86.28)	111,177 (91.82)	108,595 (97.45)	105,948 (96.62)	113,177 (98.27)	99,929 (98.19)	157,849 (99.02)
Yes	38,096 (4.55)	16,256 (13.72)	9902 (8.18)	2837 (2.55)	3706 (3.38)	1992 (1.73)	1842 (1.81)	1561 (0.98)
Alcoholic consumption, n (%)								
No	693,100 (82.81)	90,006 (75.98)	94,937 (78.41)	91,157 (81.81)	91,744 (83.67)	97,895 (85)	86,087 (84.59)	141,274 (88.62)
Yes	143,872 (17.19)	28,451 (24.02)	26,142 (21.59)	20,275 (18.2)	17,910 (16.33)	17,274 (15)	15,684 (15.41)	18,136 (11.38)
School performance[c], n (%)								
Low	95,889 (11.46)	15,897 (13.42)	15,551 (12.84)	14,929 (13.4)	12,301 (11.22)	11,507 (9.99)	10,056 (9.88)	15,648 (9.82)

XSL•FO
RenderX

	Total	Prepandemic						Pandemic
		2008-2009	2010-2011	2012-2013	2014-2015	2016-2017	2018-2019	2020-2022
Middle-low	199,758 (23.87)	30,782 (25.99)	31,193 (25.76)	28,149 (25.26)	25,943 (23.66)	25,751 (22.36)	22,255 (21.87)	35,685 (22.39)
Middle	235,371 (28.12)	31,917 (26.94)	32,365 (26.73)	29,912 (26.84)	30,137 (27.48)	32,783 (28.47)	29,891 (29.37)	48,366 (30.34)
Middle-high	203,302 (24.29)	26,665 (22.51)	28,639 (23.65)	26,227 (23.54)	27,243 (24.84)	29,369 (25.5)	25,768 (25.32)	39,391 (24.71)
High	102,652 (12.26)	13,196 (11.14)	13,331 (11.01)	12,215 (10.96)	14,030 (12.79)	15,759 (13.68)	13,801 (13.56)	20,320 (12.75)
Sexual experience, n (%)								
No	751,024 (89.73)	93,933 (79.3)	97,202 (80.28)	98,844 (88.7)	104,074 (94.91)	109,715 (95.26)	96,079 (94.41)	151,177 (94.84)
Yes	85,948 (10.27)	24,524 (20.7)	23,877 (19.72)	12,588 (11.3)	5580 (5.09)	5454 (4.74)	5692 (5.59)	8233 (5.16)
Physical activity frequency[d], n (%)								
Lower activity	502,005 (59.98)	71,137 (60.05)	72,357 (59.76)	65,536 (58.81)	61,683 (56.25)	64,707 (56.18)	59,804 (58.76)	106,781 (66.99)
Moderate activity	207,881 (24.84)	31,427 (26.53)	32,204 (26.6)	30,804 (27.64)	30,540 (27.85)	31,667 (27.5)	25,261 (24.82)	25,978 (16.3)
Higher activity	127,086 (15.18)	15,893 (13.42)	16,518 (13.64)	15,092 (13.54)	17,431 (15.9)	18,795 (16.32)	16,706 (16.42)	26,651 (16.72)
Sadness and despair, n (%)								
No	585,417 (69.94)	72,615 (61.3)	77,976 (64.4)	76,928 (69.04)	81,804 (74.6)	86,103 (74.76)	73,455 (72.18)	116,536 (73.1)
Yes	251,555 (30.06)	45,842 (38.7)	43,103 (35.6)	34,504 (30.96)	27,850 (25.4)	29,066 (25.24)	28,316 (27.82)	42,874 (26.9)
Highest educational level of parents, n (%)								
High school or lower	424,689 (50.74)	49,612 (41.88)	56,851 (46.95)	55,905 (50.17)	60,585 (55.25)	67,515 (58.62)	52,916 (52)	81,305 (51)
College or higher	412,283 (49.26)	68,845 (58.12)	64,228 (53.05)	55,527 (49.83)	49,069 (44.75)	47,654 (41.38)	48,855 (48)	78,105 (49)
Family structure, n (%)								
Nuclear family	396,035 (47.32)	67,622 (57.09)	64,799 (53.52)	50,829 (45.61)	52,840 (48.19)	52,794 (45.84)	40,550 (39.84)	66,601 (41.78)
Living with relatives	14,831 (1.77)	3160 (2.67)	2862 (2.36)	2328 (2.09)	1843 (1.68)	1729 (1.50)	1269 (1.25)	1640 (1.03)
Living alone	6889 (0.82)	645 (0.54)	629 (0.52)	535 (0.48)	611 (0.56)	772 (0.67)	1322 (1.3)	2375 (1.49)
Orphanage	2786 (0.33)	398 (0.34)	431 (0.36)	372 (0.33)	417 (0.38)	379 (0.33)	327 (0.32)	462 (0.29)

[a]KYRBS: Korea Youth Risk Behavior Web-Based Survey.

[b]BMI was divided into four groups according to the 2017 Korean National Growth Charts: underweight (<5%), normal (5% to 84%), overweight (85% to 94%), and obese (≥95%).

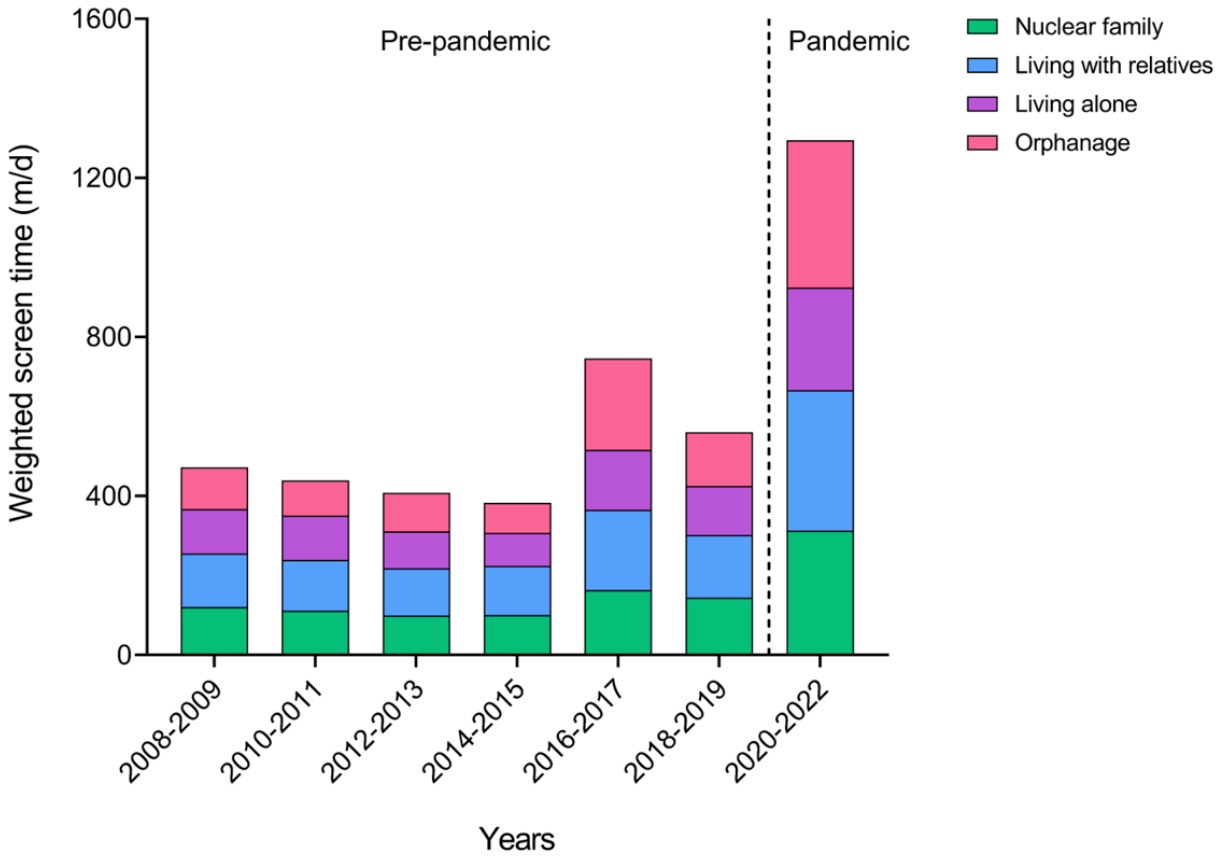
[c]School performance was divided into five groups: low (<20%), middle-low (20% to 39%), middle (40% to 59%), middle-high (60% to 79%), and high (≥80%).

[d]Physical activity frequency was divided into three groups based on the engagement in vigorous aerobic and resistance training more than three days per week: lower activity (neither activity is done for more than three days per week), moderate activity (either one activity), and higher activity (both activities).

The average weekly screen time and the trend change before and during the pandemic for each demographic over 15 years are exhibited in Multimedia Appendix 1 and Tables S1-S4 in Multimedia Appendix 2. Particularly, the trend in average screen time for each family structure is illustrated in Figure 1. Screen time showed a consistent increase from 2008 to 2022. Starting at 119.80 minutes per day (m/d; 95% CI 118.63-120.98) in 2008-2009, it temporarily decreased to 99.64 m/d (95% CI 98.45-100.83) in 2012-2013, then increased to 165.68 m/d (95% CI 163.27-168.08) in 2016-2017, and finally reached 306.80 m/d (95% CI 304.30-309.30) in 2020-2022, reflecting a sharp increase during the pandemic.

XSL•FO
RenderX

Figure 1. Trends in weighted screen time among adolescents in South Korea by family type, 2008-2022.



The pandemic had a significant impact on screen time, as detailed in Tables S1-S4 in Multimedia Appendix 2. The analysis showed a modest increase in screen time during the prepandemic period (β=8.06, 95% CI 7.74-8.39) and a substantial rise during the pandemic (β=162.06, 95% CI 159.49-164.64), with a significant difference in growth rate (β_{diff}=149.43, 95% CI 132.79-166.07). Among family structures, students in orphanages exhibited the largest increase (β_{diff}=221.90, 95% CI 159.62-284.17).

Table 2 highlights associations between screen time and biological, social, and familial factors. Women (β_{diff}=74.91, 95% CI 71.52-78.30) and higher BMI (β_{diff}=2.11, 95% CI 0.80-3.42) were identified as significant biological vulnerabilities. Among social factors, smoking (β_{diff}=79.09, 95% CI 63.13-95.06), alcohol consumption (β_{diff}=54.37, 95% CI 50.40-58.35), and sexual experience (β_{diff}=46.73, 95% CI 40.78-52.68) had the strongest associations with increased screen time. Academic achievement (β_{diff}=28.05, 95% CI 27.04-29.06]), rural residence (β_{diff}=–11.61, 95% CI –15.49 to –7.73), sadness and despair (β_{diff} 10.71, 95% CI 8.34-13.07), higher physical activity (β_{diff}=5.35, 95% CI 3.95-6.75), and higher grade (β_{diff}=3.99, 95% CI 1.52-6.46) showed varying impacts (Figures 2 and 3).

XSL•FO
RenderX

Table 2. Factors associated with weighted average screen time among adolescents before and during the COVID-19 pandemic in KYRBS[a].

Factors	Unadjusted model			Adjusted model[b]		
	Prepandemic screen time, β (95% CI)	Pandemic screen time, β (95% CI)	β_{diff} (95% CI)	Prepandemic screen time, β (95% CI)	Pandemic screen time, β (95% CI)	β_{diff} (95% CI)
Biological factors						
Sex						
Men	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Women	–5.51 (–6.82 to –4.19)[c]	61.18 (57.75 to 64.62)[c]	66.69 (63.01 to 70.38)[c]	–9.12 (–10.33 to –7.90)[c]	65.79 (62.63 to 68.95)[c]	74.91 (71.52 to 78.30)[c]
BMI group[d]						
Underweight	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Obese	5.13 (4.65 to 5.61)[c]	2.61 (1.31 to 3.91)[c]	–2.52 (–3.90 to –1.13)[c]	3.51 (3.04 to 3.98)[c]	5.62 (4.40 to 6.85)[c]	2.11 (0.80 to 3.42)[c]
Social factors						
Grade						
7th	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
12th	–0.09 (–0.40 to 0.22)	9.43 (8.18 to 10.68)[c]	9.52 (8.24 to 10.81)[c]	–5.65 (–6.31 to –4.98)[c]	–1.66 (–4.04 to 0.72)	3.99 (1.52 to 6.46)[c]
Region of residence						
Rural	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Urban	–0.79 (–2.09 to 0.51)	–16.20 (–21.03 to –11.38)[c]	–15.41 (–20.41 to –10.41)[c]	1.08 (–0.13 to 2.29)	–10.53 (–14.22 to –6.84)[c]	–11.61 (–15.49 to –7.73)[c]
Smoking status						
No	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Yes	9.10 (7.27 to 10.93)[c]	174.49 (158.20 to 190.79)[c]	165.39 (148.99 to 181.79)[c]	6.46 (4.55 to 8.38)[c]	85.55 (69.70 to 101.41)[c]	79.09 (63.13 to 95.06)[c]
Alcoholic consumption						
No	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Yes	14.79 (13.75 to 15.83)[c]	95.80 (91.51 to 100.08)[c]	81.01 (76.60 to 85.42)[c]	13.89 (12.85 to 14.93)[c]	68.26 (64.42 to 72.10)[c]	54.37 (50.40 to 58.35)[c]
School performance[e]						
Low	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
High	10.15 (9.84 to 10.46)[c]	43.33 (42.33 to 44.33)[c]	33.18 (32.14 to 34.23)[c]	9.29 (8.97 to 9.60)[c]	37.34 (36.38 to 38.30)[c]	28.05 (27.04 to 29.06)[c]
Sexual experience						
No	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Yes	–0.31 (–1.55 to 0.94)	88.53 (82.21 to 94.86)[c]	88.84 (82.40 to 95.28)[c]	0.92 (–0.30 to 2.14)	47.65 (41.83 to 53.47)[c]	46.73 (40.78 to 52.68)[c]
Physical activity frequency[f]						

XSL•FO
RenderX

Factors	Unadjusted model			Adjusted model[b]		
	Prepandemic screen time, β (95% CI)	Pandemic screen time, β (95% CI)	β_{diff} (95% CI)	Prepandemic screen time, β (95% CI)	Pandemic screen time, β (95% CI)	β_{diff} (95% CI)
Lower activity	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Higher activity	–3.96 (–4.47 to –3.44)[c]	–11.69 (–13.17 to –10.22)[c]	–7.74 (–9.30 to –6.17)[c]	–7.37 (–7.84 to –6.91)[c]	–2.02 (–3.34 to –0.71)[c]	5.35 (3.95 to 6.75)[c]
Sadness and despair						
No	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
Yes	6.91 (6.17 to 7.66)[c]	44.01 (41.61 to 46.42)[c]	37.10 (34.58 to 39.62)[c]	7.84 (7.11 to 8.57)[c]	18.55 (16.29 to 20.80)[c]	10.71 (8.34 to 13.07)[c]
Familial factor						
Highest education level of parents						
High school or lower	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
College or higher	–14.11 (–14.88 to –13.34)[c]	–41.55 (–43.79 to –39.30)[c]	–27.43 (–29.81 to –25.06)[c]	–11.22 (–11.97 to –10.46)[c]	–26.31 (–28.45 to –24.17)[c]	–15.10 (–17.36 to –12.83)[c]

[a]KYRBS: Korea Youth Risk Behavior Web-Based Survey.

[b]Adjustment for age, sex, BMI group, grade, region of residence, smoking status, alcoholic consumption, school performance, sexual experience, physical activity frequency, sadness and despair, and highest education level of parents.

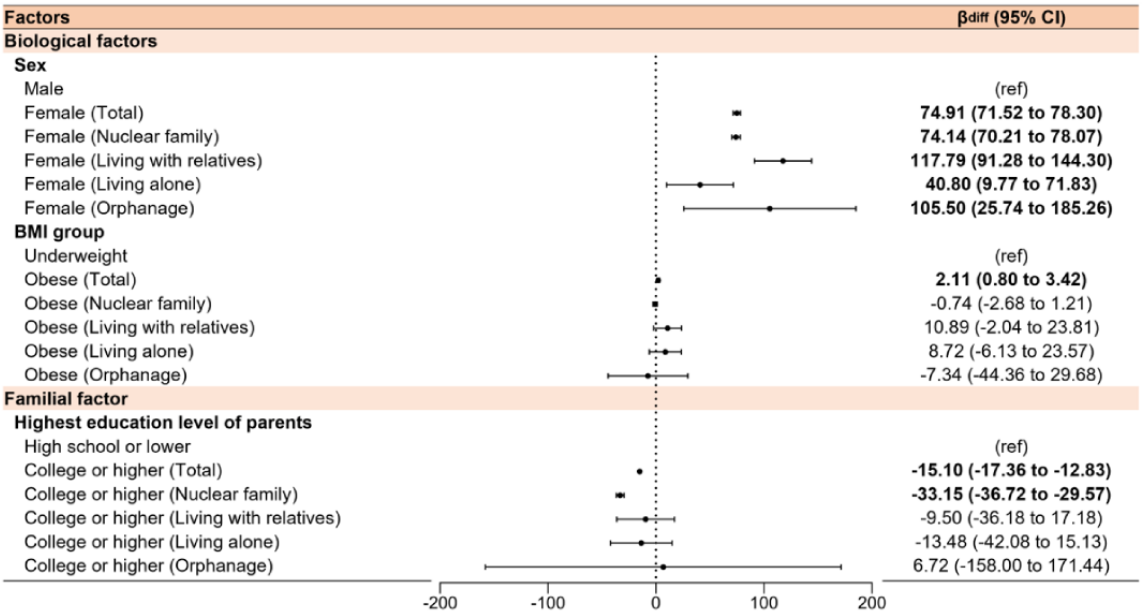
[c]Indicate a significant difference (P<.05).

[d]BMI was divided into four groups according to the 2017 Korean National Growth Charts: underweight (<5%), normal (5% to 84%), overweight (85% to 94%), and obese (≥95%). BMI was calculated as weight in kilograms divided by height in meters squared.

[e]School performance was divided into five groups: low (<20%), middle-low (20% to 39%), middle (40% to 59%), middle-high (60% to 79%), and high (≥80%).

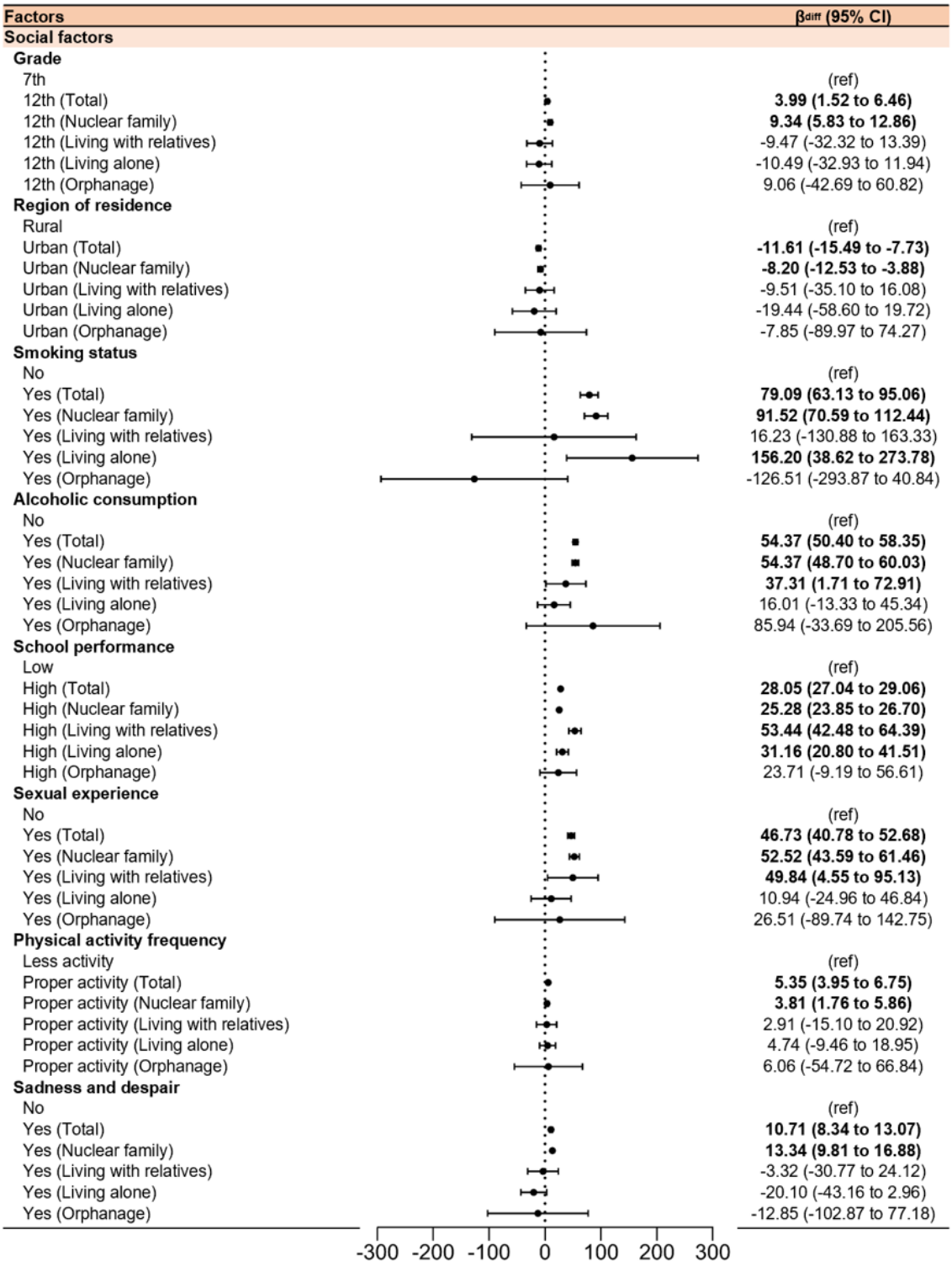
[f]Physical activity frequency was divided into three groups based on the engagement in vigorous aerobic and resistance training more than three days per week: lower activity (neither activity is done for more than three days per week), moderate activity (either one activity), and higher activity (both activities).

Figure 2. β differences of biological and familial factors in weighted average screen time among adolescents before and during the COVID-19 pandemic in KYRBS. KYRBS: Korea Youth Risk Behavior Web-Based Survey.



XSL•FO
RenderX

Figure 3. β differences of social factors in weighted average screen time among adolescents before and during the COVID-19 pandemic in KYRBS. KYRBS: Korea Youth Risk Behavior Web-Based Survey.



Adjusted models presented in Tables S5-S8 in Multimedia Appendix 2 explored associations across family structures. Before the pandemic, men in nuclear families (β=–10.50, 95% CI –11.92 to –9.08) and those living with relatives (β=–15.14, 95% CI –21.58 to –8.71) had lower screen time. In contrast, women living alone (β=14.80, 95% CI 3.57-26.03) showed higher screen time. During the pandemic, screen time significantly increased for women across all family structures: nuclear families (β_{diff}=74.14, 95% CI 70.21-78.07), living with relatives (β_{diff} 117.79, 95% CI 91.28-144.30), living alone (β_{diff} 40.80, 95% CI 9.77-71.83), and in orphanages (β_{diff} 105.50, 95% CI 25.74-185.26; Figure 3).

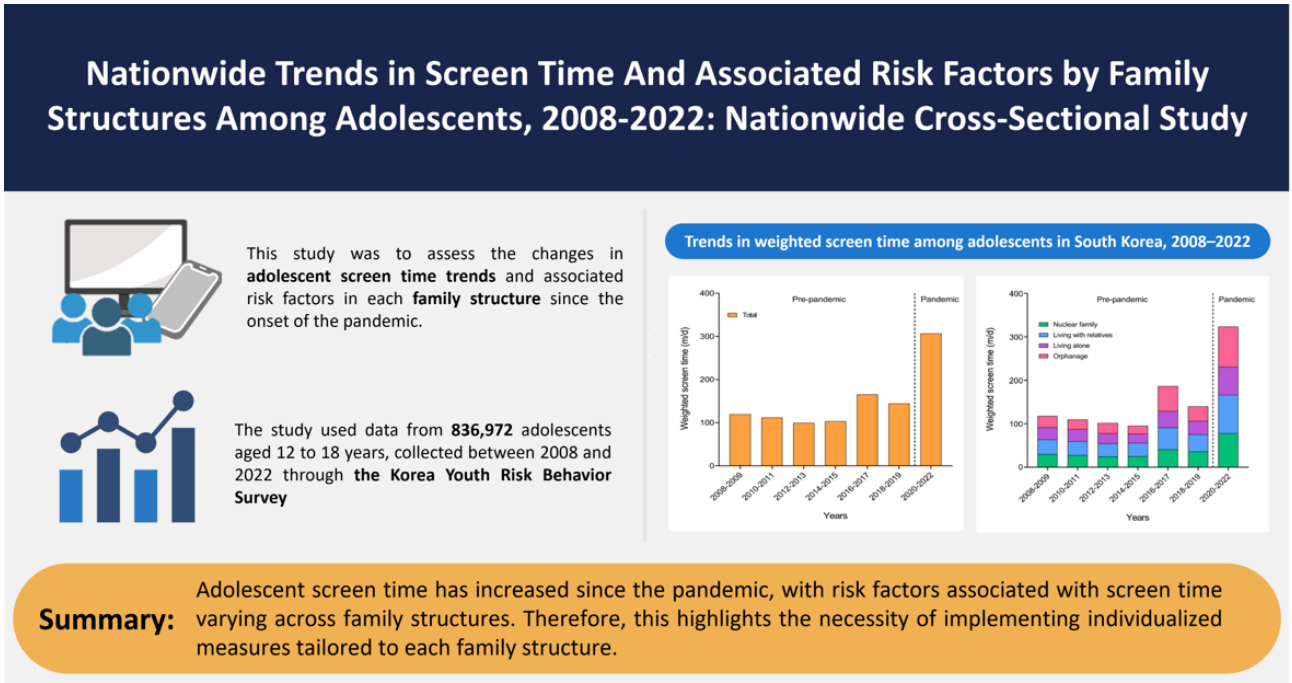
XSL•FO
RenderX

Discussion

Key Findings

This study stands out as the initial comprehensive long-term trend analysis of adolescent screen time and associated factors across various family structures, encompassing the COVID-19 era (Figure 4). There was an overall increase in screen time from 2008 to 2022. Specifically, throughout the pandemic spanning from 2020 to 2022, there was a substantial increase in adolescent screen time compared to the prepandemic. Among various family structures, students within the orphanage group exhibited the most significant increase in screen time since the pandemic onset. All examined factors exhibited statistically significant alterations in their association with adolescent screen time during the pandemic. Particularly, stratified by family structure, it was consistently observed that women had more vulnerability to increased screen time during the pandemic regardless of family structure. Meanwhile, the nuclear family group showed distinctive screen time-associated factors, including grade, region of residence, physical activity frequency, sadness and despair, and highest education level of parents, which were not present in the other family groups.

Figure 4. Graphical conclusion.



Comparison With Previous Studies

Earlier investigations have predominantly focused on the changes in adolescent screen time trends during the COVID-19 pandemic. Studies conducted in countries such as Canada (n=29,027), Germany (n=1711), the United States (n=5412), China (n=1680), and South Korea (n=913,191) consistently reported significant increases in screen time following the onset of the pandemic [14,26-28]. Many of these studies also examined factors influencing screen time, including depression and mental health [10,11]. However, most were limited by relatively small sample sizes, restricted age ranges of participants, or insufficient analysis of trends across specific demographic subgroups, such as those defined by family structure [26-29].

Additionally, while existing studies from Norway (n=4509) and France (n=3720) have explored the association between adolescent screen time and family structure, they often lack diversity in the types of family structures considered and fail to account for the profound shifts introduced by the pandemic [15,16]. These limitations highlight the importance of considering a wider range of family structures and the influence of contextual factors like the pandemic when analyzing screen time trends [27,28].

In contrast, this study addresses these gaps by analyzing a significantly larger sample size (n=836,972) collected over 15 years, allowing for a more robust examination of long-term trends. Furthermore, the stratification of adolescents by diverse family structures—nuclear families, living with relatives, living alone, and orphanages—enables a unique comparative analysis of disparities in screen time trends across these groups. This approach not only provides a more comprehensive understanding of the changes in adolescent screen time but also sheds light on the distinct risk factors associated with different family structures during the pandemic.

Plausible Mechanism

The overall elevation in average screen time of adolescents spanning from 2008 to 2022 could be elucidated by the rise in smartphone penetration among adolescents. The prevalence of smartphones among adolescents in South Korea increased by 15.3%, rising from 79.3% in 2015 to 94.6% in 2022 [30]. This upward trend would have played a role in the elevation of screen time among Korean adolescents over the 15 years. Particularly, a substantial increase in screen time with the onset of the pandemic is likely due to the increased duration of indoor activities. During the pandemic, there was an increase in time spent indoors at home and a decrease in outdoor activities,

XSL•FO
RenderX

including walking, cycling, and commuting [31]. Students were restricted from the absence of the school community covering a range of cultural, sports, social, and emotional experiences. Instead, there was an increase in internet use time for video watching and gaming with the onset of the pandemic, suggesting an influence on the rapid rise in screen time [32]. Fundamentally, the rise in screen time could be attributed to the shift in the pandemic lifestyles of students: decreased outdoor activities and various web-based engagements.

Across all family structures, a marked increase in screen time was evident following the pandemic. However, there was a specific increase in screen time among students belonging to the orphanage group. This suggests a potential influence of the presence of parents or relatives on screen time during the pandemic. The Royal College of Paediatrics and Child Health emphasized the significance of the roles of guardians in managing adolescent screen time. They recommended careful parental consideration when introducing technology to adolescents and setting individualized regulations for their screen time [16]. This detailed assistance, initially inadequate for the orphanage group, became further deficient during the pandemic due to the obstruction of support from teachers in schools. It is likely to have contributed to the substantial screen time adjustment failure observed in the orphanage group.

We analyzed the association between various biological, social, and familial factors and screen time according to the family structure. Among the diverse family structures, the nuclear family group exhibited several unique screen time-associated factors from the onset of the pandemic that were not observed in other family structures. In particular, the region of residence indicated that students residing in urban regions were less susceptible to increased screen time during the pandemic compared to their counterparts in rural regions. This finding is likely linked to parental abilities associated with the region of residence. Parents have been identified as having heightened concerns related to screen-related risks, such as cyberbullying and exposure to inappropriate or harmful content [33]. The perceived difference in screen time between regions is likely because parents residing in urban regions exhibited higher levels of parenting compared to those living in rural regions [34]. Particularly, the increase in indoor time during the pandemic induced an increase in the time spent with parents, indicating an intensified influence of parenting on adolescents. Therefore, it may be that students residing in urban regions were less vulnerable to increased screen time during the pandemic than students residing in rural regions.

Clinical and Policy Implications

This study offers a reminder of the recent rapid increase in screen time among adolescents. Considering the rise in screen time since the pandemic, it is imperative for clinical professionals and policymakers to promptly propose relevant regulation solutions. As a noticeable increase in screen time has been observed among adolescents belonging to the orphanage group following the pandemic, officials should implement more intensive management strategies for this particular group. Heightened management is also required for women adolescents, who exhibited vulnerability to increased screen time across all family structures during the pandemic. We additionally suggest the development of a customized screen time management strategy tailored to the specific family structures of each adolescent. Providing individualized solutions targeting the risk factors associated with each family structure can contribute to a more effective reduction in screen time. Specifically, focusing on the distinct risk factors (grade, region of residence, physical activity frequency, sadness and despair, and highest education level of parents) within the nuclear family group could prompt targeted management strategies for this particular cohort. Given the physical and mental health and socioemotional well-being concerns associated with excessive screen time [10,35,36], appropriate efforts from clinical, social, and political professionals are needed to normalize the rapidly increased screen time among adolescents since the pandemic. To address this, it is essential to expand international health expenditure for pandemic response and preparedness [37]. Thus, securing long-term and sustainable funding is crucial to address the rise in adolescents' digital device use and the associated mental health challenges [38].

In particular, investments in digital education programs, enhanced mental health services, and community support networks are vital to mitigate the physical and mental health impacts of increased screen time among adolescents during the pandemic. Additionally, policy and educational initiatives should focus on leveraging the already identified positive aspects of increased screen time during restricted social interactions, such as its potential for information exchange, education, and communication [39]. Pandemic response funds should be prioritized to address these issues and support adolescents in growing up in a healthy digital environment. Appropriate allocation of funds and policy support will play a critical role in addressing the challenges of increased adolescent screen time and alleviating health disparities exacerbated by the pandemic.

Limitations

This study has several limitations. First, variations in survey questions from 2008 to 2022 affected data comparability. Specifically, while smartphone use time was the focus in 2017 and from 2020 to 2022, other years emphasized internet use time, excluding total screen time such as television viewing. Nevertheless, to the best of our knowledge, this is the first study to analyze long-term data spanning 15 years, providing comprehensive insights into prepandemic and pandemic-era trends in adolescent screen time. Second, the self-response methodology in KYRBS may not fully reflect the characteristics of the actual population. However, KYRBS has been historically acknowledged as a credible dataset [40,41], and the weight-based analysis conducted with over 800,000 survey respondents further enhances the representativeness and validity of the screen time trends observed in this study. Third, this study is limited to Korean adolescents, restricting its representation of global trends. Despite this, the focus on a specific demographic allowed us to examine the nuanced relationships between screen time and family structures in the Korean context, providing insights relevant to similar societies. Finally, the cross-sectional design prevents establishing causal relationships between screen time and associated factors, highlighting the need for longitudinal and comprehensive studies. However, the

XSL•FO
RenderX

study emphasizes the need for longitudinal research and offers a foundational understanding of screen time trends and influencing factors across diverse family structures.

Conclusions

This study examined adolescent screen time trends from 2008 to 2022 across various family structures, revealing a consistent increase over time, with the steepest rise observed during the COVID-19 pandemic. Screen time patterns varied significantly by family structure, with orphaned adolescents experiencing the most pronounced increase, while other groups, such as those in nuclear families, also showed distinct trends influenced by factors like parental education and region of residence. These findings underscore the importance of tailored interventions, including family-type–specific workshops, institutional programs, and digital literacy resources, to address the diverse needs of adolescents. Reducing regional disparities and promoting alternative activities are also critical for mitigating the negative impacts of excessive screen time. Future research on broader structural and cultural influences on digital engagement is necessary to develop comprehensive and sustainable strategies for managing adolescent screen time.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (RS-2024-00460379) and the Ministry of Science and ICT, South Korea, under the Information Technology Research Center) support program (IITP-2024-RS-2024-00438239) supervised by the Institute for Information and Communications Technology Planning and Evaluation. This research was supported by grants from the NRF funded by the Korean government (RS-2024-00465719). The funders had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Data Availability

The datasets generated or analyzed during this study are available from the Korea Disease Control and Prevention Agency repository [42].

Authors' Contributions

DKY had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final version before submission. SK, HJ, YS, and MKS drafted the manuscript and conducted statistical analyses. KL, JP, HL, LS, ED, GF, LB, GFLS, MAT, MR, and DP reviewed and edited the manuscript. SW and DKY supervised the study and are the guarantors for this study. SK, HJ, YS, and MKS contributed equally as the first authors. SW and DKY contributed equally as corresponding authors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The trend in the average screen time (m/d) of adolescents before and during the COVID-19 pandemic, weighted mean (95% CI), in the Korea Youth Risk Behavior Web-Based Survey.
[DOCX File , 35 KB - publichealth_v11i1e57962_app1.docx]

Multimedia Appendix 2
Trends and factors associated with adolescent screen time before and during the COVID-19 pandemic by household type in the Korea Youth Risk Behavior Web-Based Survey.
[DOCX File , 100 KB - publichealth_v11i1e57962_app2.docx]

References

1. Cha Y, Jung W, Seo M, Rahmati M. The emerging pandemic recent: SARS-CoV-2. Life Cycle 2023;3:e2. [doi: 10.54724/lc.2023.e2]
2. Lee SW, Yuh WT, Yang JM, Cho Y, Yoo IK, Koh HY, et al. Nationwide results of COVID-19 contact tracing in South Korea: individual participant data from an epidemiological survey. JMIR Med Inform 2020;8(8):e20992 [FREE Full text] [doi: 10.2196/20992] [Medline: 32784189]
3. Maqsood A, Abbas J, Rehman G, Mubeen R. The paradigm shift for educational system continuance in the advent of COVID-19 pandemic: mental health challenges and reflections. Curr Res Behav Sci 2021;2:100011 [FREE Full text] [doi: 10.1016/j.crbeha.2020.100011] [Medline: 38620741]

XSL•FO
RenderX

4. Rahmat TE, Raza S, Zahid H, Abbas J, Sobri FAM, Sidiki SN. Nexus between integrating technology readiness 2.0 index and students' e-library services adoption amid the COVID-19 challenges: implications based on the theory of planned behavior. J Educ Health Promot 2022;11:50 [FREE Full text] [doi: 10.4103/jehp.jehp_508_21] [Medline: 35372596]
5. Trott M, Driscoll R, Irlado E, Pardhan S. Changes and correlates of screen time in adults and children during the COVID-19 pandemic: a systematic review and meta-analysis. EClinicalMedicine 2022;48:101452 [FREE Full text] [doi: 10.1016/j.eclinm.2022.101452] [Medline: 35615691]
6. Abbas J, Aman J, Nurunnabi M, Bano S. The impact of social media on learning behavior for sustainable education: evidence of students from selected universities in Pakistan. Sustainability 2019;11(6):1683. [doi: 10.3390/su11061683]
7. Abbas J. The impact of coronavirus (SARS-CoV2) epidemic on individuals mental health: the protective measures of Pakistan in managing and sustaining transmissible disease. Psychiatr Danub 2020;32(3-4):472-477 [FREE Full text] [doi: 10.24869/psyd.2020.472] [Medline: 33370755]
8. Abbas J, Al-Sulaiti K, Lorente D, Shah S, Shahzad U. Economic Growth and Environmental Quality in a Post-Pandemic World. United Kingdom: Routledge; 2023.
9. Meng Q, Yan Z, Abbas J, Shankar A, Subramanian M. Human–computer interaction and digital literacy promote educational learning in pre-school children: mediating role of psychological resilience for kids' mental well-being and school readiness. Int J Hum Comput Interact 2023;41(1):16-30. [doi: 10.1080/10447318.2023.2248432]
10. Boers E, Afzali MH, Newton N, Conrod P. Association of screen time and depression in adolescence. JAMA Pediatr 2019;173(9):853-859 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.1759] [Medline: 31305878]
11. Tandon PS, Zhou C, Johnson AM, Gonzalez ES, Kroshus E. Association of children's physical activity and screen time with mental health during the COVID-19 pandemic. JAMA Netw Open 2021;4(10):e2127892 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.27892] [Medline: 34596669]
12. Wise J. Screen time: two hour daily limit would improve children's cognition, study finds. BMJ 2018;362:k4070. [doi: 10.1136/bmj.k4070]
13. Yang Q, Wang H, Wu H, Li W, Zhang Y, Yao Y, et al. Effect of online game policy on smartphone game play time, addiction, and emotion in rural adolescents of China. BMC Psychiatry 2023;23(1):814 [FREE Full text] [doi: 10.1186/s12888-023-05325-3] [Medline: 37940898]
14. Oh J, Kim M, Rhee SY, Rahmati M, Koyanagi A, Smith L, et al. National trends in the prevalence of screen time and its association with biopsychosocial risk factors among Korean adolescents, 2008-2021. J Adolesc Health 2024;74(3):504-513. [doi: 10.1016/j.jadohealth.2023.10.021] [Medline: 38085201]
15. Langøy A, Smith ORF, Wold B, Samdal O, Haug EM. Associations between family structure and young people's physical activity and screen time behaviors. BMC Public Health 2019;19(1):433 [FREE Full text] [doi: 10.1186/s12889-019-6740-2] [Medline: 31023280]
16. Ashton JJ, Beattie RM. Screen time in children and adolescents: is there evidence to guide parents and policy? Lancet Child Adolesc Health 2019;3(5):292-294. [doi: 10.1016/S2352-4642(19)30062-8] [Medline: 30853301]
17. Park S, Yon H, Ban CY, Shin H, Eum S, Lee SW, et al. National trends in alcohol and substance use among adolescents from 2005 to 2021: a Korean serial cross-sectional study of one million adolescents. World J Pediatr 2023;19(11):1071-1081. [doi: 10.1007/s12519-023-00715-9] [Medline: 36977821]
18. Lee JH, Son Y, Park J, Lee H, Choi Y, Lee M, et al. Comparison of national trends in physical activity among adolescents before and during the COVID-19 pandemic: a nationally representative serial study in South Korea. Heliyon 2024;10(21):e40004. [doi: 10.1016/j.heliyon.2024.e40004] [Medline: 39553655]
19. Oh J, Lee M, Lee H, Yang H, Park J, Rahmati M, et al. Hand and oral hygiene practices of South Korean adolescents before and during the COVID-19 pandemic. JAMA Netw Open 2023;6(12):e2349249 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.49249] [Medline: 38147331]
20. Park S, Son Y, Lee H, Lee H, Lee J, Kang J, et al. Sex-specific trends in the prevalence of osteoarthritis and rheumatoid arthritis from 2005 to 2021 in South Korea: nationwide cross-sectional study. JMIR Public Health Surveill 2024;10:e57359 [FREE Full text] [doi: 10.2196/57359] [Medline: 39486025]
21. Hong S, Son Y, Lee M, Lee JH, Park J, Lee H, et al. Association between sociodemographic factors and vaccine acceptance for influenza and SARS-CoV-2 in South Korea: nationwide cross-sectional study. JMIR Public Health Surveill 2024;10:e56989. [doi: 10.2196/56989] [Medline: 39629701]
22. Park S, Yim Y, Lee M, Lee H, Park J, Lee JH, et al. Longitudinal trends in depression, suicidal ideation, and suicide attempts by family structure in South Korean adolescents, 2009-2022: a nationally representative serial study. Asian J Psychiatr 2024;98:104122. [doi: 10.1016/j.ajp.2024.104122] [Medline: 38959550]
23. Yoo IK, Marshall DC, Cho JY, Yoo HW, Lee SW. N-Nitrosodimethylamine-contaminated ranitidine and risk of cancer in South Korea: a nationwide cohort study. Life Cycle 2021;1:e1. [doi: 10.54724/lc.2021.e1]
24. Yoo HW, Jin HY, Yon DK, Effenberger M, Shin YH, Kim SY, et al. Non-alcoholic fatty liver disease and COVID-19 susceptibility and outcomes: a Korean nationwide cohort. J Korean Med Sci 2021;36(41):e291 [FREE Full text] [doi: 10.3346/jkms.2021.36.e291] [Medline: 34697932]
25. Lee SW. Regression analysis for continuous independent variables in medical research: statistical standard and guideline of life cycle committee. Life Cycle 2022;2:e3. [doi: 10.54724/lc.2022.e3]

XSL•FO
RenderX

26. Madigan S, Eirich R, Pador P, McArthur BA, Neville RD. Assessment of changes in child and adolescent screen time during the COVID-19 pandemic: a systematic review and meta-analysis. JAMA Pediatr 2022;176(12):1188-1198 [FREE Full text] [doi: 10.1001/jamapediatrics.2022.4116] [Medline: 36342702]
27. Schmidt SCE, Anedda B, Burchartz A, Eichsteller A, Kolb S, Nigg C, et al. Physical activity and screen time of children and adolescents before and during the COVID-19 lockdown in Germany: a natural experiment. Sci Rep 2020;10(1):21780 [FREE Full text] [doi: 10.1038/s41598-020-78438-4] [Medline: 33311526]
28. Nagata JM, Cortez CA, Cattle CJ, Ganson KT, Iyer P, Bibbins-Domingo K, et al. Screen time use among US adolescents during the COVID-19 pandemic: findings from the Adolescent Brain Cognitive Development (ABCD) study. JAMA Pediatr 2022;176(1):94-96 [FREE Full text] [doi: 10.1001/jamapediatrics.2021.4334] [Medline: 34724543]
29. Xiao S, Yan Z, Zhao L. Physical activity, screen time, and mood disturbance among Chinese adolescents during COVID-19. J Psychosoc Nurs Ment Health Serv 2021;59(4):14-20. [doi: 10.3928/02793695-20201104-04] [Medline: 33180950]
30. Kwon S, Kim R, Lee JT, Kim J, Song S, Kim S, et al. Association of smartphone use with body image distortion and weight loss behaviors in Korean adolescents. JAMA Netw Open 2022;5(5):e2213237 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.13237] [Medline: 35594044]
31. Park S, Kim HJ, Kim S, Rhee SY, Woo HG, Lim H, et al. National trends in physical activity among adults in South Korea before and during the COVID-19 pandemic, 2009-2021. JAMA Netw Open 2023;6(6):e2316930 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.16930] [Medline: 37273204]
32. Major D, Fazekas-Pongor V, Pártos K, Tabák AG, Ungvari ZI, Eörsi D, et al. Effect of school lockdown due to the COVID-19 pandemic on screen time among adolescents in hungary: a longitudinal analysis. Front Public Health 2023;11:1233024 [FREE Full text] [doi: 10.3389/fpubh.2023.1233024] [Medline: 38098820]
33. Kimball HG, Fernandez F, Moskowitz KA, Kang M, Alexander LM, Conway KP, et al. Parent-perceived benefits and harms associated with internet use by adolescent offspring. JAMA Netw Open 2023;6(10):e2339851 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.39851] [Medline: 37883086]
34. Han J, Hao Y, Cui N, Wang Z, Lyu P, Yue L. Parenting and parenting resources among Chinese parents with children under three years of age: rural and urban differences. BMC Prim Care 2023;24(1):38 [FREE Full text] [doi: 10.1186/s12875-023-01993-y] [Medline: 36726105]
35. Mayor S. Added screen time linked to more depression in adolescents, study finds. BMJ 2019;366:l4691. [doi: 10.1136/bmj.l4691] [Medline: 31311792]
36. Stiglic N, Viner RM. Effects of screentime on the health and well-being of children and adolescents: a systematic review of reviews. BMJ Open 2019;9(1):e023191 [FREE Full text] [doi: 10.1136/bmjopen-2018-023191] [Medline: 30606703]
37. Global Burden of Disease 2021 Health Financing Collaborator Network. Global investments in pandemic preparedness and COVID-19: development assistance and domestic spending on health between 1990 and 2026. Lancet Global Health 2023;11(3):e385-e413 [FREE Full text] [doi: 10.1016/S2214-109X(23)00007-4] [Medline: 36706770]
38. Abbas J. Crisis management, transnational healthcare challenges and opportunities: the intersection of COVID-19 pandemic and global mental health. Res Globalization 2021;3:100037 [FREE Full text] [doi: 10.1016/j.resglo.2021.100037]
39. Yu S, Abbas J, Draghici A, Negulescu OH, Ain NU. Social media application as a new paradigm for business communication: the role of COVID-19 knowledge, social distancing, and preventive attitudes. Front Psychol 2022;13:903082 [FREE Full text] [doi: 10.3389/fpsyg.2022.903082] [Medline: 35664180]
40. Kim Y, Choi S, Chun C, Park S, Khang YH, Oh K. Data resource profile: the Korea youth risk behavior web-based survey (KYRBS). Int J Epidemiol 2016;45(4):1076-1076e. [doi: 10.1093/ije/dyw070] [Medline: 27380796]
41. Bae J, Joung H, Kim JY, Kwon KN, Kim YT, Park SW. Test-retest reliability of a questionnaire for the Korea youth risk behavior web-based survey. J Prev Med Public Health 2010;43(5):403-410 [FREE Full text] [doi: 10.3961/jpmph.2010.43.5.403] [Medline: 20959711]
42. Korea Disease Control and Prevention Agency (KDCA). URL: https://www.kdca.go.kr/yhs/ [accessed 2025-02-22]

Abbreviations

KYRBS: Korea Youth Risk Behavior Web-Based Survey

XSL•FO
RenderX

Edited by W Goedel; submitted 01.03.24; peer-reviewed by J Abbas, D Carvalho, YH Kim, S Kim; comments to author 12.01.25; revised version received 17.01.25; accepted 27.01.25; published 10.03.25.

Please cite as:
Kim S, Jo H, Son Y, Shin MK, Lee K, Park J, Lee H, Smith L, Dragioti E, Fond G, Boyer L, López Sánchez GF, Tully MA, Rahmati M, Pizzol D, Woo S, Yon DK
Nationwide Trends in Screen Time and Associated Risk Factors by Family Structures Among Adolescents, 2008-2022: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e57962
URL: https://publichealth.jmir.org/2025/1/e57962
doi:10.2196/57962
PMID:

©Seokjun Kim, Hyesu Jo, Yejun Son, Min Kyung Shin, Kyeongmin Lee, Jaeyu Park, Hayeon Lee, Lee Smith, Elena Dragioti, Guillaume Fond, Laurent Boyer, Guillermo F López Sánchez, Mark A Tully, Masoud Rahmati, Damiano Pizzol, Selin Woo, Dong Keon Yon. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

How Information Exposure Shapes Risk Perceptions and Vaccination Intentions Among Gay, Bisexual, and Other Men Who Have Sex With Men: Cross-Sectional Survey Study

Doug Cheung[1*], ScM; Luyao Xie[1*], PhD; Lijuan Wang[2], MSc; Siyu Chen[1], PhD; Xinge Li[2], MSc; Zheng Zhang[2], MSc; Xinyue Chen[2], MPH; Shen Ge[2], MSc; Fuk-yuen Yu[1], MSSC; Yuan Fang[3], PhD; Zihuang Chen[4], BSc; Zhennan Li[5], BSc; Fenghua Sun[3], PhD; Phoenix Mo[1], PhD; Yingjie Liu[2], MSc; Zixin Wang[1], PhD

[1]Centre for Health Behaviours Research, JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Beijing Chaoyang District Center for Disease Control and Prevention, Beijing, China

[3]Department of Health and Physical Education, the Education University of Hong Kong, Hong Kong, China (Hong Kong)

[4]Danlan Goodness, Beijing, China, Beijing, China

[5]Hongshiliu Goodness, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Zixin Wang, PhD
Centre for Health Behaviours Research
JC School of Public Health and Primary Care
The Chinese University of Hong Kong
Room 508, School of Public Health, Prince of Wales Hospital, 30-32 Ngan Shing Street
Hong Kong
China (Hong Kong)
Phone: 852 2252 8474
Email: wangzx@cuhk.edu.hk

Abstract

Background: Exposure to information about mpox may shape distinct perceptual processes that influence vaccination intent. Understanding how such information impacts perceived risk and vaccination intention is crucial for designing effective risk communication and public health messaging, particularly among populations at high risk such as gay, bisexual, and other men who have sex with men (GBMSM).

Objective: This study examined the specific pathways through which mpox information exposure and associated perceptual processes influence vaccination intent. Differences between GBMSM in Beijing and Hong Kong were also examined to explore potential contextual influences.

Methods: We conducted a cross-sectional survey of mpox-unvaccinated GBMSM in Hong Kong (n=470) and Beijing (n=519) between November 2023 and March 2024. Structural equation modeling was conducted to estimate the direct and indirect effects of information exposure, perceptual processes (eg, perceived control and threat perceptions), and perceived risk on vaccination intent. Multigroup structural equation modeling was used to estimate the effect measure modification by city.

Results: Exposure to positive mpox information significantly enhanced perceived control (β=0.33; *P*=.001) and increased vaccination intention through heightened perceived risk of contracting mpox (β=0.27; *P*<.001) in the following 6 months. The indirect effect of positive information exposure on vaccination intent via perceived control and risk was significant in Hong Kong (*B*=0.01; Wald test: *Z*=3.05 and *P*=.002) but not in Beijing (*B*=−0.01; *P*=.25). Conversely, negative information exposure primarily increased threat perceptions (Hong Kong: *B*=0.33 and *P*=.001; Beijing: *B*=0.93 and *P*<.001) but did not consistently translate to increased perceived risk of contracting mpox (Hong Kong: *B*=−0.10 and *P*=.31; Beijing: *B*=0.20 and *P*=.11). Notable contextual differences emerged between Beijing and Hong Kong—participants in Beijing reported higher levels of information exposure (eg, international mpox statistics; mean score 2.17, SD 0.97 vs 1.79, SD 0.89; *P*<.001) and more nonregular sex partners (mean 2.20, SD 6.73 vs 1.65, SD 3.37; *P*=.02), but the associations among information exposure, perceived risk, and vaccination intent were weaker than in Hong Kong participants.

XSL•FO
RenderX

Conclusions: Positive mpox-related information strongly promotes vaccination intent by enhancing perceived control and amplifying perceived risk, particularly in settings with accessible vaccination programs such as Hong Kong. Conversely, negative information appears less effective in driving vaccination intention. Tailored, stigma-free communication is crucial for improving vaccination uptake, especially in mainland China, where subsidized vaccines are scarce and perceptual pathways linking information exposure to vaccination intent are relatively weak. Enhancing access to vaccines and addressing contextual barriers can further optimize the impact of such interventions.

(JMIR Public Health Surveill 2025;11:e70635) doi:10.2196/70635

KEYWORDS

mpox; gay, bisexual, and other men who have sex with men; GBMSM; multigroup structural equation modeling

Introduction

Background

The 2022 multicountry mpox outbreak spread rapidly via air travel and undetected local transmissions among gay, bisexual, and other men who have sex with men (GBMSM) driven by social stigma, marginalization, and discrimination [1-5]. These factors often compel GBMSM to seek anonymous and multiple same-sex connections in high-risk environments, facilitating "superspreader" events and exacerbating both local and global mpox transmission [2,5-8]. The mpox outbreak exemplifies the critical importance of a One Health approach by revealing the fundamental interconnectedness or existential disconnectedness among human, animal, and environmental health. Of concern, originally a zoonotic disease with animal-human transmission in parts of Africa, mpox has adapted to "human-to-human" transmission networks worldwide, particularly in urban environments and initially mostly affecting GBMSM. This evolution demonstrates how ecological shifts, human mobility, and social determinants—including stigma and health care inequities—collectively shape infectious disease dynamics [2]. In 2023, Asia experienced sustained imported mpox cases, with mainland China documenting 501 cases in August alone [9-11]. The region faces unique challenges due to persistent stigmatization of same-sex behaviors, limited mpox vaccination programs, and the easing of COVID-19 travel restrictions [12,13], raising concerns about a resurgence if vaccine uptake among GBMSM remains low [4,6,14].

While Western countries effectively curtailed the outbreak through culturally sensitive campaigns—such as leveraging gay-identifying physicians and influencers to boost vaccine acceptance [15,16]—Asian contexts have struggled [17]. Despite the US Centers for Disease Control and Prevention recommending the 2-dose JYNNEOS vaccine for high-risk GBMSM [18], costs and access remain significant barriers. For example, JYNNEOS costs up to US $300 per dose in higher-income settings, whereas Hong Kong has offered fully subsidized vaccination since October 2022. However, this program is largely inaccessible to mainland Chinese GBMSM, and mpox-specific vaccines are still under development [19-21]. Even in places with subsidized programs such as Hong Kong and Taiwan, vaccine uptake remains suboptimal (21.7% in Hong Kong [22] and 59.9% in Taiwan [23], much lower than in Western cities such as Paris at 78% [24]) and below the modeled 80% threshold needed to prevent further outbreaks [25].

A growing body of research has examined risk perception and media exposure among GBMSM [22,26-28], but few studies have explored how varying types and sources of mpox-related information shape illness perceptions and vaccination intentions—especially in Chinese contexts. Most notably, the psychological mechanisms connecting information exposure with preventive intentions and actions remain poorly understood in Asia, where negative media portrayals and lesbian, gay, bisexual, transgender, and queer (LGBTQ+) stigma are widespread [29-31]. This constitutes a critical research gap as understanding these mechanisms is essential for designing effective, culturally sensitive risk communication strategies.

The common sense model of illness self-regulation (CSM) provides a valuable framework for examining how different types of mpox-related information influence GBMSM's preventive intentions and behaviors. The extended CSM framework proposes that emotional states such as concerns, depression, and distress are more strongly correlated with threat-related illness perceptions, including the identification of symptom severity, duration, and their impacts on individuals [32]. For example, exposure to stigmatizing news about deaths due to mpox and HIV or AIDS coinfection may evoke negative emotions and elevate GBMSM's perceptions of mpox severity and impact, leading to maladaptive coping responses such as denial, avoidance, or delayed treatment seeking [32,33]. Conversely, accurate and timely information about mpox transmission, symptoms, and prevention methods may contribute to adaptive illness representations, such as illness coherence and controllability, which predict positive coping behaviors such as adherence to preventive measures and timely health care seeking [34,35]. Applying the extended CSM to mpox-related information exposure among GBMSM can provide insights into the psychological mechanisms driving preventive intentions and behaviors. This understanding can inform the development of targeted risk communication strategies that promote adaptive coping and vaccination. These relationships remain unexplored in the context of culturally sensitive mpox risk communication and vaccination campaigns for GBMSM in the Asia-Pacific region.

Behavioral intentions have been theorized (eg, theory of planned behavior) and shown to predict actual health-seeking behaviors [36], such as COVID-19 vaccination among UK adults aged 50 to 64 years [37], human papillomavirus vaccination among Dutch girls [38], and influenza vaccination for patients with diabetes in China [39]. However, Jongen et al [40] found that an increased level of perceived risk for mpox more consistently

XSL•FO
RenderX

predicted and explained the difference between high intention for mpox vaccination and actual vaccination uptake among Dutch GBMSM attending sexual health services in a later survey wave. Similarly, in a cross-sectional survey of Canadian GBMSM, those with a lower perceived susceptibility to mpox were more likely to remain unvaccinated despite being offered vaccination after attending a sexually transmitted infection (STI) clinic [41]. Therefore, mpox illness perceptions may play a stronger role in the relationship between behavioral intent and subsequent uptake for GBMSM if it increases the perceived risk of contracting mpox—this hypothesized relationship is yet to be explored [42].

Mainland China and Hong Kong play pivotal roles in the global resurgence of mpox, largely due to their extensive interconnectivity and prominence as international air travel hubs. These 2 regions present distinct policy and social environments—Beijing, as the capital of mainland China, has limited vaccine availability and a higher level of societal stigma [43], whereas Hong Kong has implemented a fully subsidized vaccine program and has generally more open societal attitudes toward sexual minority groups [44]. Both cities experience significant cross-border interaction and are major population centers yet differ markedly in their public health policies, media environments, and levels of LGBTQ+ acceptance [44]. This contrast allows for a nuanced comparison of how structural factors—such as vaccine access, stigma, and risk communication—interact to shape health behaviors among GBMSM.

Objectives

In 2023, China reported 1712 confirmed mpox cases, with the Beijing municipality (258 cases) and Guangdong province (342 cases) accounting for more than one-third of the country's total cases [9]. Hong Kong, despite having only 68 confirmed cases, is surrounded by the Guangdong province with millions of daily cross-border travelers and observed steady local confirmed cases from January 2024 to August 2024 [45]. Using the extended CSM framework applied through multigroup structural equation modeling (SEM), this study aimed to examine how differential exposure to mpox information and perceptual processes may influence perceived susceptibility and vaccination intentions among GBMSM in these 2 highly interconnected municipal areas.

Given the multidimensional nature of information exposure, illness perceptions, and behavioral intent, this study used SEM because it enables the simultaneous estimation of complex, interrelated pathways and mediating effects, which are essential for understanding how latent constructs—such as threat perception and personal control—link information exposure to risk perception and vaccination intent. Unlike standard regression, SEM allows for the modeling of latent variables using multiple observed indicators, providing a more accurate representation of psychosocial processes. In addition, we used multigroup SEM to directly compare these pathways between GBMSM in Hong Kong and Beijing, allowing us to test for contextual differences in the relationships among key variables.

We hypothesized that regional differences in case numbers, risk communication, vaccine program availability, and cultural values lead to different levels of mpox information exposure, illness perceptions, and subsequent perceived risk and vaccination intentions. Specifically, we propose that the content of mpox information—factual, positive, or negative—may affect vaccination intentions through distinct perceptual processes by shaping threat or control perceptions. We further anticipate that these information attributes will differentially impact perceived risk and vaccination intent and that geographical and social differences between Hong Kong and Beijing may moderate these effects based on local case burdens, information exposure, perceived susceptibility, and vaccine access.

Methods

Study Design and Sample

A cross-sectional survey was conducted with GBMSM in Beijing and Hong Kong from November 2023 to March 2024. Participants had to meet the following inclusion criteria: (1) being Chinese male individuals aged ≥18 years and (2) having engaged in anal intercourse with at least one man in the previous 6 months (geographic information and location of the study sites are presented in Figure 1). Recruitment and data collection methods were similar between Beijing and Hong Kong. In both cities, upon approval of the owners, trained and experienced field-workers approached potential participants in gay bars at different time slots on weekdays and weekends. On sites, the field-workers briefed potential participants about the study and gave them an information sheet. Online recruitment was also conducted in both cities by posting study advertisements on popular gay websites (in Hong Kong only), social media platforms such as Facebook (in Hong Kong only) or Weibo and WeChat (in Beijing only), and partner-seeking mobile apps frequently used by local GBMSM (Blued in Beijing and Grindr in Hong Kong). In Hong Kong, Facebook is the most popular social media platform, and Grindr is the most widely used gay dating app. However, these 2 online platforms are not available in mainland China. Weibo and WeChat are the 2 most commonly used social media platforms in mainland China, with functions very similar to those of Facebook. Blued is the most widely used gay dating app in mainland China, with features and functions similar to those of Grindr. Therefore, Weibo, WeChat, and Blued were chosen to facilitate online recruitment in Beijing. Potential participants who were interested in the study could contact the research team via private messaging, telephone, WhatsApp or WeChat, or email. Recruitment was supplemented through referrals made by peers and community-based organizations. Participants were screened for study eligibility. Participant anonymity was strictly maintained, and verbal informed consent was obtained. Various contact methods were used to arrange telephone interviews, which were conducted in Mandarin in Beijing and in Cantonese in Hong Kong.

XSL•FO
RenderX

Figure 1. Geographical mapping of Beijing and Hong Kong with respect to gay, bisexual, and other men who have sex with men participants recruited from these municipal areas.

Ethical Considerations

This study was reviewed and approved by the Survey and Behavioural Research Ethics Committee of the Chinese University of Hong Kong (reference SBRE-22-0488). The research adhered to established ethical guidelines for human participant research. Verbal informed consent was obtained from all participants before taking part in the study. Participants were provided with a detailed information sheet outlining the purpose of the study, procedures, risks, and benefits before consenting. Consent was specifically obtained for the primary data collection, and no additional consent was required for the analysis as all data were anonymized.

Participant anonymity was strictly maintained throughout the study. Data collected during the phone interviews were deidentified at the point of collection, ensuring that no personally identifiable information was stored or linked to the responses. All data were securely stored in encrypted files accessible only to authorized researchers. Participants were compensated for their time upon completing the 30-minute phone interviews. In Hong Kong, participants received an HK $50 (US $6.40) supermarket coupon, and in Beijing, they were provided with a ¥20 (US $3) e-coupon. No identifiable images or personal data of individual participants were included in the manuscript or supplementary materials. As such, there is no risk of participant identification.

Participants and Data Collection

Recruitment and data collection procedures were standardized for both Beijing and Hong Kong. Trained field-workers, with permission from venue owners, approached potential participants at various times during the week and on weekends in gay bars. The participants were given a brief description of the study and an information sheet. Online recruitment included posting advertisements on popular gay websites (in Hong Kong), social media platforms (Facebook in Hong Kong and Weibo and WeChat in Beijing), and partner-seeking apps (Blued in Beijing and Grindr in Hong Kong). Interested individuals could contact the research team via private messages, phone, WhatsApp or WeChat, or email. Peer referrals and community-based organizations also supported recruitment efforts. Participant anonymity was strictly maintained, and verbal informed consent was obtained. Various contact methods were used to arrange telephone interviews, which were conducted in Mandarin in Beijing and in Cantonese in Hong Kong. To ensure content consistency, the study questionnaire was available in simplified Chinese for Beijing and traditional Chinese for Hong Kong.

Questionnaire Design

The questionnaire was developed by a team of public health researchers, epidemiologists, health psychologists, physicians, and representatives of nongovernmental organizations. It was then pilot-tested with 5 GBMSM, and their feedback was used to refine the final version.

Measures

Background Characteristics

Participants provided information on their sociodemographic characteristics (such as age, educational level, relationship

XSL•FO
RenderX

status, income, and employment status); sexual orientation; and their experiences with HIV and STI testing, HIV pre-exposure prophylaxis (PrEP), and other HIV prevention services used in the previous 12 months. They also reported HIV-related risk behaviors over the previous 6 months, including engaging in condomless anal intercourse with different types of male partners, the number of sexual partners, psychoactive substance use, and the use of drugs during sexual activity. Sexualized drug use was defined as the consumption of substances such as ketamine, methamphetamine, cocaine, cannabis, ecstasy, nonprescription hypnotics, heroin, cough suppressants, amyl nitrite (poppers), γ-hydroxybutyric acid or γ-butyrolactone, foxy, and mephedrone before or during sex.

Mpox-Specific Information Content Exposure

Participants were asked to rate how often they encountered information on 7 common mpox-related topics using a scale from 1 (almost never) to 4 (often). The topics were (1) mpox statistics from other countries (eg, case numbers and mortality rates); (2) mpox statistics specific to Hong Kong (eg, case numbers and mortality rates); (3) the Hong Kong government's response to the mpox outbreak; (4) information on mpox discrimination and stigma toward LGBTQ+ individuals; (5) negative information about patients with mpox, such as severe symptoms and fatalities; (6) positive information about patients with mpox, such as recovery stories; and (7) information regarding mpox prevention and treatment. Details on the aforementioned items have been published elsewhere [22].

Perceived Risk of Contracting Mpox

Participants were asked the following: "Based on your current health and sexual behaviors, how do you rate your risk of contracting mpox within the following year?" (1=none; 5=very high).

Mpox Illness Representations

Mpox illness representations were assessed using the Brief Illness Perception Questionnaire [46], modified by asking participants to imagine a hypothetical scenario of contracting mpox and replacing "illness" with "mpox" in the original format [46]. This approach has been used in previous studies to assess illness representations of those at risk of a disease but who do not have the disease [47,48]. Participants were asked to rank each item on a Likert scale ranging from 0 to 10. These items included consequences ("If you contracted mpox, how much does your illness affect your life?"), timeline ("If you contracted mpox, how long do you think your illness will continue?"), personal control ("If you contracted mpox, how much control do you feel you have over your illness?"), treatment control ("If you contracted mpox, how much do you think existing treatment can help your illness?"), identity ("If you contracted mpox, how much do you experience symptoms from the illness?") and concern ("If you contracted mpox, how concerned are you about your illness?"), coherence ("If you contracted mpox, how well do you feel you understand mpox?"), and emotions ("If you contacted mpox, how much does your illness affect you emotionally?").

Mpox Vaccination Intent and Uptake Behaviors

Participants were asked whether they had received the mpox vaccine (yes or no) and, if not, how likely they were to receive it in the following 6 months (1=very low; 5=very high).

Analysis

Sample characteristics were described using total row proportions, means, and SDs and by location (Beijing and Hong Kong), with differences between locations evaluated using the Pearson chi-square test or Wilcoxon rank sum test (analytic sample size: N=989; Hong Kong: n=470; Beijing: n=519) [49]. Zero-order (Pearson) correlations were computed to evaluate correlations between key variables in this study. Missing data were imputed using the nonparametric random forest algorithm [50]. Linear regression was conducted to examine the bivariate relationship between background variables and mpox vaccination intention in the following 6 months. All significant background variables in univariate regression were adjusted for in the following analyses aided by directed acyclic graphs to avoid omitted confounding effects and conditioning on a collider [49,51].

SEM was conducted to examine the hypothesized model using a 2-step approach. First, a confirmatory factor analysis was conducted to assess the goodness of fit of the measurement model. The construct reliability of the latent variables was evaluated using both the Cronbach α and McDonald ω [52]. SEM was used to test the proposed structural model. The mean and variance–adjusted weighted least squares estimator was used to estimate modeled parameters accounting for ordinally distributed endogenous outcomes [49,53]. Indirect effects were estimated using Bollen-Stine bootstrapping with 5000 repetitions, which is robust against deviation from multivariate normality corresponding to mean and variance–adjusted weighted least squares estimator and ordinal endogenous outcomes [54]. Overall model fit was determined using (1) the scaled chi-square badness-of-fit index (P>.05), (2) a scaled root mean square error of approximation (S-RMSEA) of <0.06, (3) the scaled comparative fit index (S-CFI), and (4) a scaled Tucker-Lewis index (S-TLI) of >0.95 [51,55,56].

Multigroup SEM was further used to examine potential differences in mpox information exposure and their associations with mpox perceptions across cities (moderation by city). We tested for partial scalar invariance in the measurement model and partial path coefficient invariance in the structural model. Invariance was determined by a difference in the S-CFI of <0.01 and a significant log-likelihood ratio test (P<.05) between the baseline models (ie, configural fits) and models with equality constraints on each of the latent factor loadings (weak or metric invariance) and thresholds or intercepts (strong or scalar invariance) [57,58]. The moderation of parameter estimates from the final multigroup structural equation model was examined using a 2-sided Wald test for heterogeneity using the appropriate unstandardized coefficients (B) and variance parameters [59], with P<.05 indicating a significant difference. We conducted sensitivity analyses of the direct and indirect effects of multigroup SEM by replicating the same model restricted to participants from Hong Kong and Beijing separately (Multimedia Appendix 1). Psychometric analyses were

XSL•FO
RenderX

conducted using the *psych* and *lavaan* packages in R (R Foundation for Statistical Computing) [60,61].

Results

Sample Description

Of the eligible participants approached through outreach in gay venues, online recruitment, and peer referral (n=713 from Hong Kong and n=602 from Beijing), 86% (613/713) from Hong Kong and 87% (524/602) from Beijing provided verbal consent and completed telephone interviews. Of the original total sample (N=1137), 0.44% (5/1137) of participants from Beijing and 12.58% (143/1137) of participants from Hong Kong were vaccinated against mpox and were removed from the analytic sample. Table 1 presents the unvaccinated participant characteristics by total sample (N=989) and group (Beijing: 519/989, 52.5%; Hong Kong: 470/989, 47.5%). Among the participants from both groups, most were aged ≥25 years (859/989, 86.9%), tertiary educated (836/989, 84.5%), single (752/989, 76%), and employed full time (774/989, 78.3%); identified as homosexual (823/989, 83.2%); and earned above the median city income (612/989, 61.9%). Mpox vaccination rates were significantly higher in Hong Kong than in Beijing (116/519, 22.4% vs 4/470, 0.9%; *P*<.001); however, the average levels of vaccination intention (mean score 2.31, SD 1.24 in Beijing vs mean score 2.32, SD 1.22 in Hong Kong; *P*=.60) and perceived risk of contracting mpox (mean score 1.84, SD 0.88 in Beijing vs mean score 1.86, SD 0.82 in Hong Kong; *P*=.90) among those who were unvaccinated were not significantly different between Beijing and Hong Kong. GBMSM in Hong Kong were less likely to be married, cohabit with a woman, or identify as bisexual or other. Participants in Beijing reported a higher uptake of sexual health services, including HIV testing (440/517, 85.1% vs 185/470, 39.4%; *P*<.001) and PrEP use (100/517, 19.3% vs 39/470, 8.3%; *P*=.03), in the previous 6 months.

XSL•FO
RenderX

Table 1. Sample characteristics among mpox-unvaccinated gay, bisexual, and other men who have sex with men participants from Hong Kong and Beijing, China (N=989).

Characteristic	Overall	Beijing (n=519)	Hong Kong (n=470)	*P* value[a]
Sociodemographic characteristics, n (%)				
Age group (y)				<.001
18-24	130 (13.1)	64 (12.3)	66 (14)	
25-34	488 (49.3)	276 (53.2)	212 (45.1)	
35-44	256 (25.9)	138 (26.6)	118 (25.1)	
≥45	115 (11.6)	41 (7.9)	74 (15.7)	
Educational level				.03
Secondary or lower	153 (15.5)	68 (13.1)	85 (18.1)	
Tertiary or higher	836 (84.5)	451 (86.9)	385 (81.9)	
Current relationship status				<.001
Currently single	752 (76)	365 (70.3)	387 (82.3)	
Married or cohabiting with a man	199 (20.1)	119 (22.9)	80 (17)	
Married or cohabiting with a woman	38 (3.8)	35 (6.7)	3 (0.6)	
Monthly income level				<.001
Below the city median	339 (34.3)	190 (36.6)	149 (31.7)	
Above the city median	612 (61.9)	294 (56.6)	318 (67.7)	
Missing	38 (3.8)	35 (6.7)	3 (0.6)	
Current employment status				.20
Full time	774 (78.3)	414 (79.8)	360 (76.6)	
Other	215 (21.7)	105 (20.2)	110 (23.4)	
Sexual orientation				<.001
Homosexual	823 (83.2)	403 (77.6)	420 (89.4)	
Bisexual, heterosexual, or uncertain	166 (16.8)	116 (22.4)	50 (10.6)	
Health service use in the previous 6 mo, n (%)				
HIV testing				<.001
No	364 (36.8)	79 (15.2)	285 (60.6)	
Yes	625 (63.2)	440 (84.8)	185 (39.4)	
STI[b] testing				<.001
No	509 (51.5)	187 (36)	322 (68.5)	
Yes	480 (48.5)	332 (64)	148 (31.5)	
HIV pre-exposure prophylaxis use				<.001
No	850 (85.9)	419 (80.7)	431 (91.7)	
Yes	139 (14.1)	100 (19.3)	39 (8.3)	
Sexual behaviors in the previous 6 mo				
Number of regular sex partners, mean (SD)	1.42 (2.27)	1.71 (2.91)	1.10 (1.11)	.003
Condomless anal sex with regular sex partners, n (%)				.02
No	534 (54)	299 (57.6)	235 (50)	
Yes	455 (46)	220 (42.4)	235 (50)	
Number of nonregular sex partners, mean (SD)	1.94 (5.40)	2.20 (6.73)	1.65 (3.37)	.02
Condomless anal sex with nonregular sex partners, n (%)				.02
No	806 (81.5)	409 (78.8)	397 (84.5)	

XSL•FO
RenderX

Characteristic	Overall	Beijing (n=519)	Hong Kong (n=470)	*P* value[a]
Yes	183 (18.5)	110 (21.2)	73 (15.5)	
Sexualized drug use in the previous 6 mo, n (%)				<.001
No	854 (86.3)	424 (81.7)	430 (91.5)	
Yes	135 (13.7)	95 (18.3)	40 (8.5)	
Risk perception and intention to receive mpox vaccine, mean (SD)				
Perceived risk of contracting mpox	1.84 (0.88)	1.83 (0.82)	1.86 (0.93)	.90
Mpox vaccination intent	2.31 (1.24)	2.32 (1.22)	2.29 (1.27)	.60

[a]Pearson chi-square test or Wilcoxon rank sum test for group comparison.

[b]STI: sexually transmitted infection.

Distribution, Construct Reliability, and Validity Among Latent Variables

Table 2 shows the means and SDs of the observed items of the latent factors from the measurement model. The average self-rating of most observed items of the latent factors was higher among participants in Beijing than among those in Hong Kong except for some mpox illness perceptions (Brief Illness Perception Questionnaire) measures. Reliability coefficients indicated adequate internal consistency for most latent factors except for the control perception domain of mpox illness perception. Thus, only 1 personal control item from mpox illness perceptions was used in subsequent structural equation models. Zero-order correlations showed high correlations among factual, positive, and negative mpox information content (Table 3). The threat and personal control perception domains of mpox illness perceptions had low to moderate correlations with other latent variables, indicating discriminant validity that was consistent across the Beijing and Hong Kong subsamples.

XSL•FO
RenderX

Table 2. Descriptive statistics of observed indicators and reliabilities of latent variables among mpox-unvaccinated gay, bisexual, and other men who have sex with men participants from Hong Kong and Beijing, China (N=989).

	Score range	Beijing (n=519)			Hong Kong (n=470)			*P* value[a]
		Score, mean (SD)	Cronbach α	McDonald ω	Score, mean (SD)	Cronbach α	McDonald ω	
Frequency of exposure to information								
Factual content			0.90	0.90		0.70	0.78	
International mpox statistics	1-4	2.17 (0.97)			1.79 (0.89)			<.001
Local mpox statistics	1-4	2.21 (0.94)			1.94 (0.96)			<.001
Government's responses	1-4	2.16 (0.94)			1.83 (0.88)			<.001
Positive content			0.86	0.85		0.72	0.77	
Patients recovered from mpox	1-4	2.35 (1.00)			1.69 (0.87)			<.001
Prevention and treatment of mpox	1-4	2.31 (0.96)			2.10 (0.95)			<.001
Negative content			0.73	0.74		0.73	0.75	
Discrimination and stigma against LGBTQ+[b] individuals	1-4	2.67 (1.06)			2.46 (1.16)			.001
Patients with severe symptoms and deaths	1-4	2.33 (0.99)			2.10 (1.01)			<.001
Mpox illness perceptions (BIPQ[c])								
Threat perceptions			0.94	0.94		0.77	0.78	
Consequences	0-10	7.93 (2.72)			7.35 (2.22)			<.001
Timeline	0-10	7.24 (2.68)			5.42 (2.26)			<.001
Identity	0-10	7.57 (2.57)			7.50 (1.93)			.59
Concern	0-10	8.20 (2.59)			7.77 (2.40)			.004
Emotion	0-10	7.81 (2.69)			7.78 (2.12)			.82
Control perceptions			0.60	0.61		0.33	0.54	
Coherence[d]	0-10	5.53 (3.10)			4.92 (2.21)			<.001
Personal control	0-10	6.72 (2.86)			6.22 (2.21)			.001
Treatment control[d]	0-10	7.23 (2.74)			7.44 (1.86)			.12

[a]*P* value from Kruskal-Wallis test for group comparison.

[b]LGBTQ+: lesbian, gay, bisexual, transgender, and queer.

[c]BIPQ: Brief Illness Perception Questionnaire.

[d]Excluded from analysis due to differential item cross-loadings across groups, contributing to poor reliability and convergent validity.

XSL•FO
RenderX

Table 3. Zero-order (Pearson) correlation matrix among key factors from the measurement model among total gay, bisexual, and other men who have sex with men participants (N=989) and separately for those from Beijing (n=524) and Hong Kong (n=613).

	Factual information	Positive information	Negative information	Threat perceptions	Personal control perceptions	Perceived risk of contracting mpox	Vaccination intent for mpox
Total							
Factual information	1	—[a]	—	—	—	—	—
Positive information	0.85	1	—	—	—	—	—
Negative information	0.84	0.81	1	—	—	—	—
Threat perceptions	0.21	*0.14*[b]	0.31	1	—	—	—
Personal control perceptions	0.36	*0.39*	*0.30*	0.27	1	—	—
Perceived risk of contracting mpox	*0.05*	*0.08*	*0.08*	*0.05*	*0.03*	1	—
Vaccination intent for mpox	0.22	0.19	0.22	0.16	*0.11*	0.27	1
Beijing							
Factual information	1	—	—	—	—	—	—
Positive information	0.85	1	—	—	—	—	—
Negative information	0.84	0.86	1	—	—	—	—
Threat perceptions	0.18	*0.11*	0.31	1	—	—	—
Personal control perceptions	0.31	*0.33*	*0.28*	0.27	1	—	—
Perceived risk of contracting mpox	*0.11*	*0.14*	*0.19*	*0.11*	*−0.003*	1	—
Vaccination intent for mpox	0.29	0.23	0.25	0.16	*0.07*	0.27	1
Hong Kong							
Factual information	1	—	—	—	—	—	—
Positive information	0.83	1	—	—	—	—	—
Negative information	0.86	0.78	1	—	—	—	—
Threat perceptions	0.19	*0.07*	0.31	1	—	—	—
Personal control perceptions	0.43	*0.49*	*0.36*	0.17	1	—	—
Perceived risk of contracting mpox	*0.001*	*0.06*	*−0.02*	*0.01*	*0.13*[c]	1	—

[a]Not applicable.

[b]Coefficients in italics signify qualitatively different variable-by-variable correlations between the total sample and the subsamples by city.

[c]Significant difference of the variable-by-variable correlation coefficient from the null value (coefficient=0).

Table S1 in Multimedia Appendix 2 shows that the bivariate relationships of having tertiary education or higher, having an income level above the city's median income, having been tested for HIV and STIs and using PrEP, having higher numbers of regular and nonregular sex partners in the previous 6 months, and having a higher perceived risk for contracting mpox in the following 12 months were associated with higher levels of intent to vaccinate against mpox in the following 6 months. Figure S1 in Multimedia Appendix 2 shows the directed acyclic graphs of the hypothesized causal relationships to derive a minimal sufficient adjustment set for confounding control consisting of age, educational level, HIV testing, PrEP use, and number of sexual partners [62,63]. Multimedia Appendix 3 shows the Strengthening the Reporting of Observational Studies in Epidemiology statement checklist.

XSL•FO
RenderX

SEM Results

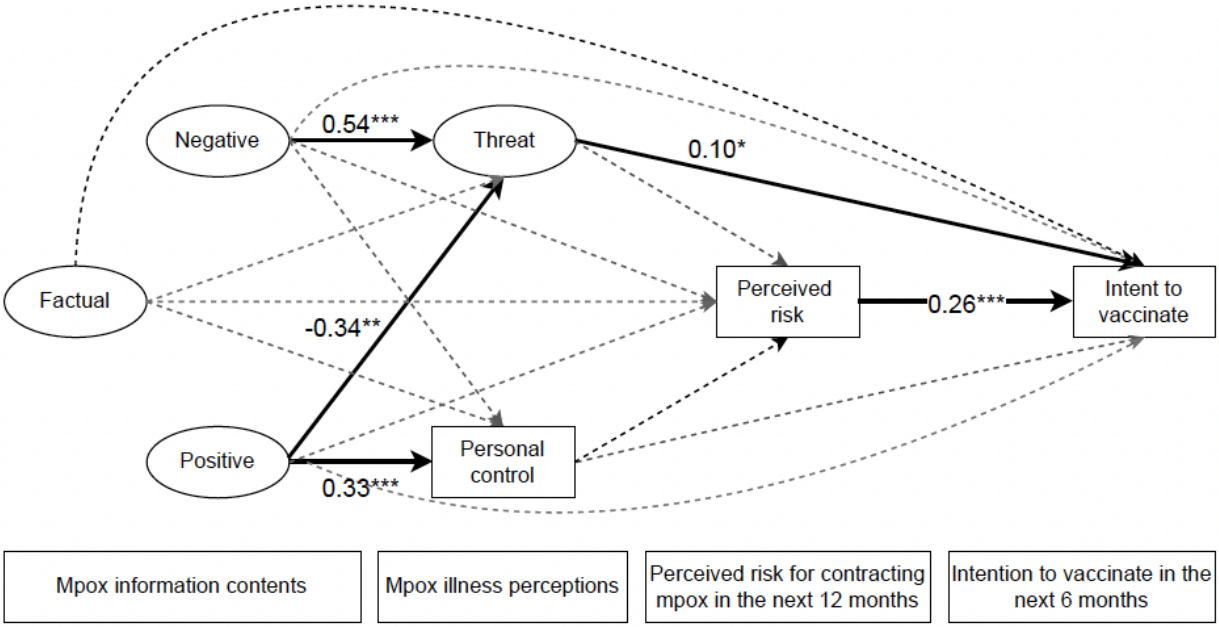
Model Fit Statistics

The measurement model yielded a good fit to the data (χ^2_{124}=161.0; S-CFI=0.989; S-TLI=0.986; S-RMSEA=0.049, 90% CI 0.041-0.057; standardized root mean square residual=0.030; weighted root mean square residual=0.920). Its standardized factor loading ranged from 0.72 to 0.89 (*P*<.001 in all cases). The structural model also yielded a good fit to the data (χ^2_{356}=323.0; S-CFI=0.986; S-TLI=0.992; S-RMSEA=0.030, 90% CI 0.025-0.035; standardized root mean square residual=0.019; weighted root mean square residual=1.083).

Path Coefficient (Direct Effects)

Figure 2 shows the statistically significant path coefficients of the structural equation model combining Hong Kong and Beijing participants. Positive mpox information content (standardized path coefficient β=−0.34; *P*=.001) was negatively associated with threat perceptions, whereas negative mpox information content (β=0.54; *P*<.001) was positively associated with threat perceptions of mpox. Only positive mpox information content (β=0.33; *P*=.001) was positively associated with personal control perceptions of mpox; factual mpox information content (β=0.17; *P*<.001) was marginally associated with personal control perceptions. Neither threat perceptions nor personal control perceptions were associated with perceived risk of contracting mpox. Perceived risk of contracting mpox (β=0.27; *P*<.001) and threat perceptions (β=0.10; *P*=.01) were both positively associated with intention to receive the mpox vaccine in the following 6 months.

Figure 2. Path diagram of structural equation model with statistically significant standardized coefficients for the total mpox-unvaccinated gay, bisexual, and other men who have sex with men participants (N=989). The model was adjusted for educational level, age, HIV testing, and pre-exposure prophylaxis use; number of nonregular partners; and number of regular partners in the previous 6 months. **P*<.05; ***P*<.01; ****P*<.001.



Mediated Pathways (Indirect Effects)

Threat perceptions of mpox negatively mediated (unstandardized path coefficient [B]=−0.043, 95% CI −0.82 to −0.004) the association between positive mpox information content and the intention to receive the mpox vaccine. Threat perceptions of mpox positively mediated (B=0.07, 95% CI 0.012-0.128) the association between negative mpox information content and vaccination intention.

Multigroup Analysis and Invariance Testing

Overview

As indicated in Table 4, the results of the multigroup analysis of the measurement model showed that the configural model (baseline) had an acceptable fit (comparative fit index [CFI]=0.991; root mean square error of approximation [RMSEA]=0.050). The metric invariance model, with all factor loadings constrained to be equal across groups, showed a statistically meaningful decrease in model fit compared to the configural model (ΔCFI=−0.006; ΔRMSEA=0.011). However, partial scalar invariance was achieved by freeing the intercepts of the 2 indicators (ΔCFI=0.000; ΔRMSEA=−0.003). As in the SEM literature, multigroup analyses can be conducted if partial scalar invariance is supported [64-66].

XSL•FO
RenderX

Table 4. Measurement and structural invariance model fit statistics for multigroup structural equation models among mpox-unvaccinated gay, bisexual, and other men who have sex with men participants from Beijing and Hong Kong.

	Chi-square (*df*)	CFI[a]	RMSEA[b]	ΔChi-square per df	*P* value	ΔCFI[c]	ΔRMSEA[d]
Measurement model							
Configural fit (baseline)	212.3 (96)	0.991	0.050	Reference	—[e]	—	—
Metric invariance	295.2 (104)	0.985	0.061	10.36	<.001	−0.006	0.011
Scalar invariance	398.3 (119)	0.978	0.069	8.09	<.001	−0.013	0.019
Partial scalar invariance	228.7 (111)	0.991	0.046	1.10	.15	0.000	−0.003
Full structural equation model							
Configural fit (baseline)	387.7 (338)	0.997	0.017	Reference	—	—	—
Metric invariance	443.5 (346)	0.993	0.024	6.976	<.001	−0.013	0.007
Scalar invariance	512.1 (399)	0.992	0.024	2.039	<.001	−0.044	0.020
Partial scalar invariance	414.5 (373)	0.997	0.015	0.765	.97	−0.007	0.002

[a]CFI: comparative fit index.

[b]ΔCFI threshold of >0.01 indicates noninvariance model misfit.

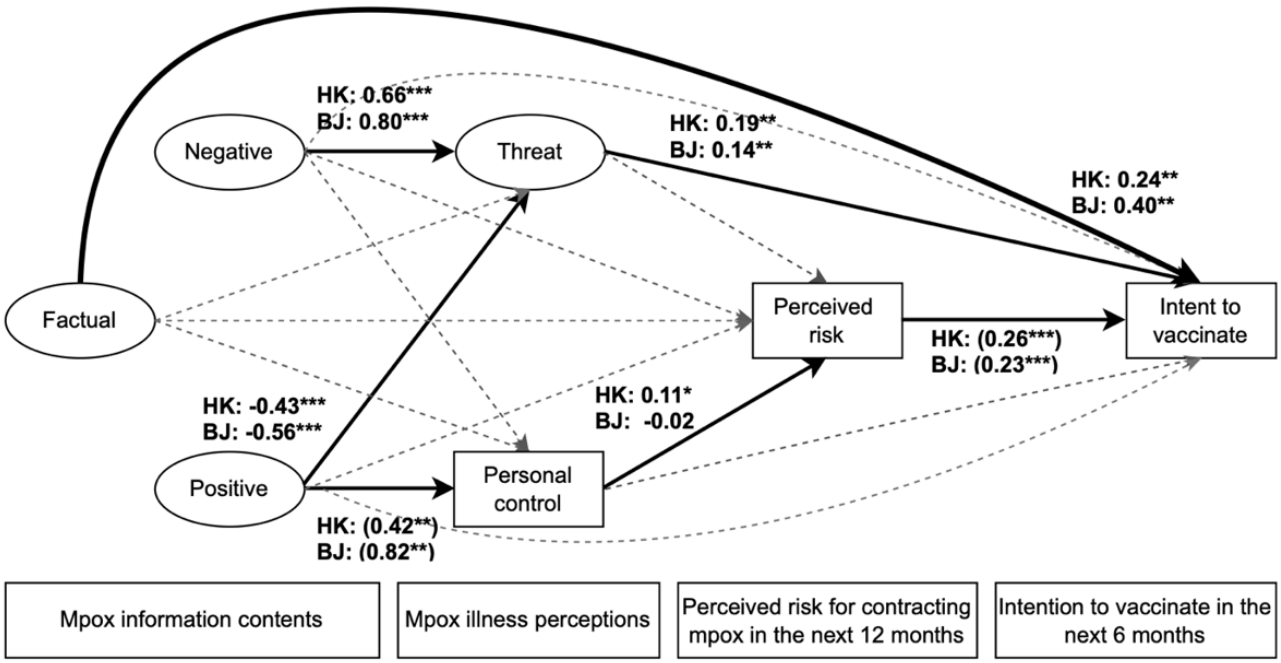
[c]RMSEA: root mean square error of approximation.

[d]ΔRMSEA threshold of >0.01 indicates noninvariance model misfit.

[e]Not applicable.

Figure 3 shows the standardized regression path coefficients of the multigroup structural equation model. To determine whether the city moderated the paths, nested models were formed by comparing equality-constrained parameter models with the unconstrained free parameter model by city using the likelihood ratio test (chi-square≥3.8 per df as the critical value). Bolded paths indicate significant coefficients for either the Hong Kong or Beijing subsample. Estimates in parentheses indicate coefficients estimated with equality-constrained parameters and vice versa.

Figure 3. Path diagram of multigroup structural equation model with statistically significant standardized coefficients for gay, bisexual, and other men who have sex with men participants in Hong Kong (HK; n=470) and Beijing (BJ; n=519). The model was adjusted for education, age, HIV testing, and pre-exposure prophylaxis use; number of nonregular partners; and number of regular partners in the previous 6 months. Path coefficients in parentheses indicate that the coefficient parameters were estimated with equality constraints. **P*<.05; ***P*<.01; ****P*<.001.



XSL•FO
RenderX

Relationships Among Attributes of Mpox Information Exposure, Threat Perceptions, Perceived Risk, and Vaccination Intent

Table 5 shows unstandardized direct and indirect effect coefficients of the relationships among attributes of mpox information exposure, threat perception, perceived risk, and vaccination intent. Exposure to positive mpox information was negatively associated with threat perceptions in Hong Kong (*B*=−0.22, 95% CI –0.38 to –0.06) and Beijing (*B*=−0.65, 95% CI –1.00 to –0.30), with a stronger association in Beijing (Wald test: *Z*=2.34 and *P*=.02). Exposure to negative information was positively associated with threat perceptions in Hong Kong (*B*=0.33, 95% CI 0.15-0.51) and Beijing (*B*=0.93, 95% CI 0.52-1.34), with a stronger association in Beijing (Wald test: *Z*=−3.40 and *P*=.002). Threat perception was positively associated with vaccination intent in Hong Kong (*B*=0.39, 95% CI 0.15-0.63) and Beijing (*B*=0.12, 95% CI 0.01-0.23), with a stronger association in Hong Kong (Wald test: *Z*=2.14 and *P*=.03). Negative information exposure had an insignificant negative association with perceived risk in Hong Kong (*B*=−0.10, 95% CI –0.30 to 0.10) and an insignificant positive association with perceived risk in Beijing (*B*=0.20, 95% CI –0.05 to 0.46), but the difference between cities was significant (Wald test: *Z*=−2.06 and *P*=.04).

Table 5. Unstandardized direct and indirect effect coefficients and Wald test of heterogeneity of effects by group.

	Hong Kong, *B*[a] (SE; 95% CI)	Beijing, *B* (SE; 95% CI)	Wald test[b]	
			Z	*P* value
Direct effects				
Positive[c]→threat[d]	*−0.22 (0.08; −0.38 to −0.06)*[e]	*−0.65 (0.18; −1.00 to −0.30)*	*2.34*	*.02*
Negative[f]→threat	*0.33 (0.09; 0.15 to 0.51)*	*0.93 (0.21; 0.52 to 1.34)*	*−3.40*	*.002*
Positive→risk[g]	0.14 (0.10; −0.06 to 0.34)	0.04 (0.11; −0.18 to 0.26)	0.65	.52
Negative→risk	−0.10 (0.10; −0.30 to 0.10)	0.20 (0.13; −0.05 to 0.46)	*−2.06*	*.04*
Threat→intent[h]	*0.39 (0.12; 0.15 to 0.63)*	*0.12 (0.06; 0.01 to 0.23)*	*2.14*	*.03*
Personal control[i]→risk	*0.05 (0.02; 0.01 to 0.09)*	−0.02 (0.01; −0.04 to 0.01)	*3.05*	*.002*
Personal control→intent	*0.06 (0.03; 0.001 to 0.12)*	−0.02 (0.03; −0.05 to 0.02)	*2.53*	*.01*
Indirect effects[j]				
Positive→threat→intent	*−0.09 (0.04; −0.17 to −0.007)*	−0.08 (0.04; −0.16 to 0.00)	−0.13	.89
Negative→threat→intent	*0.13 (0.05; 0.03 to 0.23)*	*0.11 (0.06; 0.002 to 0.22)*	0.19	.85
Positive→personal control→risk→intent	*0.01 (0.01; 0.002 to 0.03)*	−0.01 (0.00; −0.01 to 0.003)	*2.52*	*.01*
Negative→personal control→risk→intent	0.00 (0.00; −0.01 to 0.003)	0.00 (0.00; −0.002 to 0.005)	−1.20	.23
Positive→personal control→intent	0.05 (0.03; −0.002 to 0.10)	−0.01 (0.01; −0.04 to 0.02)	*2.17*	*.03*
Negative→personal control→intent	−0.01 (0.01; −0.04 to 0.01)	0.004 (0.01; −0.007 to 0.01)	−1.13	.26

[a]Unstandardized path coefficients.

[b]Wald test of heterogeneity of effects across groups.

[c]Positive mpox information latent variable.

[d]Threat perception latent variable composed of items from the Brief Illness Perception Questionnaire (BIPQ).

[e]Values in italics correspond to $P<0.05$.

[f]Negative mpox information latent variable.

[g]Perceived risk of contracting mpox within the following 12 months.

[h]Intent to receive the mpox vaccine in the following 6 months.

[i]Personal control item from the BIPQ.

[j]Bollen-Stine bootstrapped SEs and 95% CIs with 5000 repetitions.

XSL•FO
RenderX

Relationships Among Personal Control Perceptions, Perceived Risk of Contracting Mpox, and Vaccination Intent

Personal control perception was positively associated with perceived mpox risk in Hong Kong (*B*=0.05, 95% CI 0.01-0.09) and insignificantly negatively associated with perceived mpox risk in Beijing (*B*=−0.02, 95% CI –0.04 to 0.01), with a significant city difference (Wald test: *Z*=3.05 and *P*=.002). It was also positively associated with vaccination intent in Hong Kong (*B*=0.06, 95% CI 0.001-0.12) but not in Beijing (*B*=−0.02, 95% CI –0.05 to 0.02), with a significant city difference (Wald test: *Z*=2.53 and *P*=.01).

Indirect Effects

The indirect effect of positive mpox information exposure on vaccination intent through personal control and mpox risk was significant and positive in Hong Kong (*B*=0.01, 95% CI 0.002-0.03) but not in Beijing (*B*=−0.01, 95% CI –0.01 to 0.003), with a significant difference between cities (Wald test: *Z*=2.52 and *P*=.01). The indirect effect through personal control was significant and positive in Hong Kong (*B*=0.05, 95% CI –0.002 to 0.01) but not in Beijing (*B*=−0.01, 95% CI –0.007 to 0.01), with a significant difference between cities (Wald test: *Z*=2.17 and *P*=.03).

Discussion

Principal Findings

Since the 2022 multicountry mpox outbreak, mpox has continued to spread in places with little access to vaccines, whereas its resurgence across countries is imminent [67]. China is home to >12 million GBMSM and has the highest mpox case burden in the Western Pacific region [4,68]. However, only 1% (5/524) of participants from China's mpox epicenter, Beijing, had been vaccinated against mpox. To anticipate the regulatory approval of mpox vaccines in mainland China, this study examined the novel pathways through which exposure to different attributes of mpox information and perceptual processes could influence GBMSM's vaccination decisions, specifically by raising GBMSM's level of perceived risk of contracting mpox. Our findings provide timely implications for GBMSM-tailored risk communication and community partnerships to raise awareness of mpox among GBMSM across borders.

Our study highlights the critical roles of mpox-related information content and disease-related perceptions in shaping perceived risk and subsequent vaccine-related decision-making among GBMSM, extending the findings of Jongen et al [40]. In Hong Kong, increased exposure to positive mpox information was associated with higher personal control and vaccination intent, mediated by increased perceived risk. However, this relationship was not observed in Beijing. Factual mpox information was associated with increased vaccination intent in both cities, as demonstrated by the multigroup structural equation model. Correlation analyses also revealed that perceived risk was positively correlated with personal control perceptions in Hong Kong but not in Beijing. Vaccination intent had a stronger positive correlation with personal control in Hong Kong than in Beijing. These results suggest that, in places where fully government-sponsored mpox vaccines are in place, promoting factual and positive mpox information (eg, positive portrayal of patients with mpox, recovery stories, and health advice) could potentially increase vaccination intent by raising the awareness of their level of perceived risk of contracting mpox, which may also promote GBMSM's personal control of mpox, which may include having accessible vaccination programs in place at GBMSM's disposal and their related promotional campaigns enacting effective coping strategies. This finding is consistent with behavior change theories, including the extended CSM framework underlying the role of control perceptions in motivating vaccine uptake behaviors [32].

Our study shows that exposure to negative mpox information could raise GBMSM's threat perceptions of mpox without necessarily increasing their perceived risk of contracting mpox and subsequent vaccination intent but that it is associated with increased behavioral intention for mpox vaccination. These results suggest that the association between negative mpox information and behavioral intention for mpox vaccinations through threat perceptions may be less predictive of actual vaccination behaviors as compared to positive content through personal control, which in turn increases both risk perceptions and behavioral intention to receive the mpox vaccine. Highly threatening or stigmatizing content may trigger maladaptive coping behaviors such as denial, avoidance, or disengagement [33]. Rather than motivating action, overwhelming threat can result in fatalism or psychological withdrawal, especially in communities with limited resources or structural barriers to prevention (eg, lack of vaccine access in mainland China). This finding may correspond to the high prevalence of sensationalized headlines and exaggerated reporting of severe mpox cases, mortality, and HIV or AIDS coinfection across regions [69]. This negatively charged media coverage that appeals to the public's consumption of polarizing views and affective processes without providing risk information relevant to GBMSM, including unmoderated homophobic user comments on media outlets (homophobic slurs remain unmoderated in one of Hong Kong's online news outlets [70]), could counter mpox control efforts [4]. Among Chinese adults exposed to disgusting graphic mpox information, those with higher disgust endorsed more avoidant preventive measures and perceived higher mpox severity but not susceptibility or perceived risk of contracting mpox [71]. In addition, higher levels of mpox-related stigma attenuate the association between perceived severity and preventive intent [71]. Emotionally provoked viewers, including GBMSM, may dissociate from their own mpox risk and view those at risk as "distant others" (eg, people living with HIV) who should be receiving the vaccine rather than themselves [72]. In addition, exposure to stigmatizing mpox information could galvanize GBMSM's self-stigma and identity concealment, which negatively predicted mpox vaccination intent and uptake among Canadian GBMSM in the study by Le Forestier et al [8].

Despite higher levels of mpox information exposure and illness perceptions, GBMSM from Beijing were different from those from Hong Kong. This suggests that GBMSM from Beijing may not be able to identify relevant mpox risk information even

XSL•FO
RenderX

with more exposure to local cases and case information. In Hong Kong, health authorities collaborate with community partners and media channels to disseminate tailored mpox risk communication and vaccination promotion for GBMSM and remove hate speech toward LGBTQ+ individuals in online news outlets [73,74]. Moreover, our finding that personal control perception more strongly predicted vaccination intention among GBMSM in Hong Kong than in Beijing may be explained by several contextual and theoretical factors. First, Hong Kong's public health infrastructure and government-sponsored vaccination programs are widely regarded as transparent and accessible, which can enhance individuals' trust in health authorities and the perceived efficacy of their own preventive actions. This aligns with the CSM, which posits that, when individuals believe that their actions (such as seeking vaccination) will be supported and effective within their health care environment, perceptions of personal control are more likely to translate to concrete health behaviors [32,33]. In contrast, in Beijing and mainland China, limited access to mpox vaccines, lower visibility of LGBTQ+-inclusive health campaigns, and persistent stigma may undermine individuals' confidence that their actions will be effective or socially supported. In addition, Hong Kong's relatively open media environment and consistent dissemination of affirmative health messages may reinforce the notion that vaccination is both accessible and beneficial, thus strengthening the link between personal control and intent. Conversely, in Beijing, mixed or negative messaging, combined with structural and social barriers, may weaken this relationship [75,76]. Thus, the interplay among health care system trust, policy transparency, and supportive media narratives likely explains why personal control perceptions are more predictive of vaccination intention in Hong Kong.

Conversely, GBMSM from mainland China may lack adequate exposure to relevant environmental risk communication. Leiwen et al [77] found that only one-fourth of mainland Chinese GBMSM correctly identified GBMSM as being at risk of contracting mpox, whereas half believed that only people living with HIV are at risk. Similarly, Zheng et al [78] found that more than half of GBMSM incorrectly identified that there were effective medications for mpox and less than half could identify >3 ways to prevent mpox infections. These findings suggest that tailored and comprehensive risk communication for GBMSM in mainland China could improve their mpox risk self-assessment and promote preventive measures even when vaccines are not within the range of options, for example, reducing the number of sexual partners and using condoms to reduce the risk of mpox transmission through direct genital contact [79]. Such communication could be synergized with positive portrayals of stories of patients with mpox, such as gay-identifying physicians and social media influencers with mpox [15,16], in raising GBMSM's cognitive perceptions of mpox and reducing their emotional distance on perceived susceptibility while empowering them to enact preventive behaviors.

Limitations

The results of this study should be interpreted within their limitations. First, the cross-sectional design cannot ascertain causality, specifically the temporal direction of how previous information affects subsequent intent. It is causally plausible that behavioral intent assessed at 1 point in time is a function of previous information and psychological processes. Longitudinal studies are preferred to clarify causality and the potential for reverse or bidirectional effects. Second, our recruitment strategy—using social media, gay bars, and peer referrals—may have introduced selection bias. This approach likely overrepresented younger, more digitally connected, and socially active GBMSM who are comfortable accessing online platforms or community venues while potentially underrepresenting older, less connected, or more marginalized individuals who may not participate in such spaces. As a result, the findings may not fully reflect the experiences or perspectives of the entire GBMSM population in Beijing and Hong Kong, particularly those less engaged with the LGBTQ+ community or less reachable through digital or venue-based outreach. Third, there is a possibility that unmeasured confounders influenced our study results. Factors such as previous vaccination hesitancy, underlying attitudes toward vaccination, previous experiences with health care systems, or structural barriers to health care access were not directly measured but could have independently affected both information exposure and vaccination intentions. Fourth, although our study included participants from 2 major urban centers—Beijing and Hong Kong—the generalizability of our findings to GBMSM living in smaller cities, rural areas, or other regions of China may be limited. GBMSM in less urbanized areas may have different patterns of information exposure, levels of stigma, health care access, or vaccine availability, which could influence both risk perceptions and preventive behaviors. Finally, our sample of participants from Beijing had a higher level of education, included less individuals who identified as bisexual, and had been more recently HIV tested than participants in another respondent-driven sample of GBMSM from Beijing [80]. This indicates that we were more likely to have overestimated the prevalence of behavioral intent for mpox vaccination. However, our population parameter estimate is still more conservative than those in other studies (>80%) with larger sample sizes [78,81]. On the other hand, participants from Hong Kong had similar characteristics to those of a larger, more representative sample [82]. All data were self-reported and, therefore, were subject to recall and social desirability biases. The use of trained interviewers with structured questionnaires and interview designs in this study may have reduced some reporting bias. However, questions related to sexual behavior could be stigmatized when asked by interviewers, which could have been underreported.

Conclusions

This study highlights the critical role of mpox-related information exposure in influencing behavioral intentions toward vaccination among GBMSM in Beijing and Hong Kong. Positive information exposure, particularly regarding recovery stories and prevention, significantly increased vaccination intent by enhancing personal control perceptions and perceived risk of contracting mpox among GBMSM in Hong Kong, where a fully government-subsidized mpox vaccination program is in place. Conversely, while negative information related to mpox increased threat perceptions and vaccination intentions, it did

XSL•FO
RenderX

not reliably translate to a higher perceived risk, especially in Beijing. These findings suggest several practical implications for public health messaging aimed at increasing mpox vaccine uptake among GBMSM in China. Health agencies should prioritize accurate, positively framed, and stigma-free information—such as recovery stories and practical prevention advice—delivered through channels or key opinion leaders trusted by the target community. Proactive efforts are needed to counter misinformation and stigma, for example, through myth-busting campaigns and the promotion of supportive narratives on both social and traditional media platforms. In addition to the careful moderation of stigmatizing messaging, news media channels should implement stricter standards when framing their reporting on mpox and moderate user comments on their platforms to prevent the spread of misinformation that could harm public health. Partnerships with LGBTQ+ organizations and community leaders can enhance the credibility and reach of messaging, whereas targeted outreach via social media and mobile apps can ensure that information is accessible to those most at risk. Finally, communications should provide clear guidance on accessing vaccination and related services and be culturally sensitive and inclusive to empower GBMSM to assess their risk realistically and take preventive action.

Acknowledgments

The authors would like to thank all participants of this study. The authors did not use generative artificial intelligence in any portion of the manuscript. This study was supported by the AIDS Trust Fund (reference MSS376R) and the 2021 Beijing Key Specialty Program for Major Epidemic Prevention and Control. The content of this paper represents the opinions of the authors' organizations. It does not represent the position of the AIDS Trust Fund. The AIDS Trust Fund is not responsible for any claims, demands, or liabilities whatsoever from or in connection with the use of any information contained in this paper or the participation in the sponsored project. Patients or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Data Availability

The datasets generated or analyzed during this study are not publicly available because they contain sensitive personal behaviors but are available from the corresponding author on reasonable request.

Authors' Contributions

DC, LX, LW, SC, FS, PM, YL, and ZW conceived and designed the study. DC, LX, LW, SC, and ZW developed the methodology. LW, SC, LX, ZZ, SG, XC, FY, YF, ZC, and ZL handled data curation. DC, LX, and ZW conducted formal analyses. LW, SC, LX, ZZ, SG, XC, FY, YF, ZC, ZL, FS, PM, YJ, and ZW were responsible for project administration. DC, LX, and ZW drafted the manuscript. LW, SC, XL, ZZ, SG, XC, YF, ZC, ZL, FS, PM, and YL reviewed and edited the first draft of the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensitivity analyses with unstandardized direct and indirect effect coefficients conducted separately on gay, bisexual, and other men who have sex with men participant subsets from Hong Kong and Beijing.
[DOCX File , 17 KB - publichealth_v11i1e70635_app1.docx]

Multimedia Appendix 2
Bivariate relationship between baseline characteristics and mpox vaccination intention in the following 6 months among mpox-unvaccinated gay, bisexual, and other men who have sex with men participants in Beijing and Hong Kong and directed acyclic graph representing causal relationships among variables to guide the selection of variables for statistical analyses to support causal inference.
[DOCX File , 190 KB - publichealth_v11i1e70635_app2.docx]

Multimedia Appendix 3
Strengthening the Reporting of Observational Studies in Epidemiology statement—checklist of items that should be included in reports of cross-sectional studies.
[DOCX File , 20 KB - publichealth_v11i1e70635_app3.docx]

References

XSL•FO
RenderX

1. Multi-country monkeypox outbreak: situation update. World Health Organization. 2022 Jun 10. URL: https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON392 [accessed 2024-07-17]
2. Ajmal H, Hunter E, Duggan J, Timoney C, Walsh C. Agent-based modelling of mpox infection dynamics: simulating disease transmission and control strategies. J Artif Soc Soc Simul 2024;27(2):6. [doi: 10.18564/jasss.5380]
3. Paredes MI, Ahmed N, Figgins M, Colizza V, Lemey P, McCrone JT, et al. Underdetected dispersal and extensive local transmission drove the 2022 mpox epidemic. Cell 2024 Mar 14;187(6):1374-86.e13 [FREE Full text] [doi: 10.1016/j.cell.2024.02.003] [Medline: 38428425]
4. Zhu Z, Pan X. Preventing stigma against the gay community to curb mpox transmission in the Western Pacific region. J Glob Health 2024 Feb 09;14:03012 [FREE Full text] [doi: 10.7189/jogh.14.03012] [Medline: 38330199]
5. Cheung DH, Samoh N, Jonas K, Lim SH, Kongjareon Y, Guadamuz TE. Partying in Thailand: never testing among men who have sex with men who attend sexualized (drug) parties. Sex Transm Dis 2024 May 01;51(5):352-358 [FREE Full text] [doi: 10.1097/OLQ.0000000000001937] [Medline: 38301623]
6. Lin M, Xin Y, Wang J, Nie P, Yan Q, Wang L, et al. Analysing monkeypox epidemic drivers: policy simulation and multi-index modelling across 39 nations. J Glob Health 2024 Feb 09;14:04037 [FREE Full text] [doi: 10.7189/jogh.14.04037] [Medline: 38333932]
7. Baral S, Logie CH, Grosso A, Wirtz AL, Beyrer C. Modified social ecological model: a tool to guide the assessment of the risks and risk contexts of HIV epidemics. BMC Public Health 2013 May 17;13(1):482 [FREE Full text] [doi: 10.1186/1471-2458-13-482] [Medline: 23679953]
8. Le Forestier JM, Page-Gould E, Chasteen A. Identity concealment may discourage health-seeking behaviors: evidence from sexual-minority men during the 2022 global mpox outbreak. Psychol Sci 2024 Feb 12;35(2):126-136. [doi: 10.1177/09567976231217416] [Medline: 38215021]
9. Ren R, Li C, Bai W, Wang Y, Li D, Ding F, et al. The epidemiological characteristics of mpox cases - China, 2023. China CDC Wkly 2024 Jun 28;6(26):619-623 [FREE Full text] [doi: 10.46234/ccdcw2024.118] [Medline: 38966310]
10. Endo A, Jung SM, Miura F. Mpox emergence in Japan: ongoing risk of establishment in Asia. Lancet 2023 Jun 10;401(10392):1923-1924. [doi: 10.1016/S0140-6736(23)00766-3] [Medline: 37269865]
11. Duarte PM, Adesola RO, Priyadarsini S, Singh R, Shaheen MN, Ogundijo OA, et al. Unveiling the global surge of mpox (monkeypox): a comprehensive review of current evidence. The Microbe 2024 Sep;4:100141. [doi: 10.1016/j.microb.2024.100141]
12. Akhmetzhanov AR, Wu PH. Transmission potential of mpox in mainland China, June-July 2023: estimating reproduction number during the initial phase of the epidemic. PeerJ 2024;12:e16908 [FREE Full text] [doi: 10.7717/peerj.16908] [Medline: 38344294]
13. Asakura TR, Jung SM, Murayama H, Ghaznavi C, Sakamoto H, Teshima A, et al. Projecting international mpox spread in Asia: ongoing global health risk. medRxiv Preprint posted online on April 19, 2024 [FREE Full text] [doi: 10.1101/2024.04.17.24305832]
14. Molla J, Sekkak I, Mundo Ortiz A, Moyles I, Nasri B. Mathematical modeling of mpox: a scoping review. One Health 2023 Jun;16:100540 [FREE Full text] [doi: 10.1016/j.onehlt.2023.100540] [Medline: 37138928]
15. Man recovering from monkeypox describes experience. YouTube. URL: https://www.youtube.com/watch?v=p_3tMj2M0nA [accessed 2024-08-11]
16. Irshad U. Overcoming stigma—contracting mpox as a minoritised trainee. Lancet Infect Dis 2023 Apr;23(4):400. [doi: 10.1016/s1473-3099(23)00064-6]
17. Brand SP, Cavallaro M, Cumming F, Turner C, Florence I, Blomquist P, et al. The role of vaccination and public awareness in forecasts of mpox incidence in the United Kingdom. Nat Commun 2023 Jul 11;14(1):4100 [FREE Full text] [doi: 10.1038/s41467-023-38816-8] [Medline: 37433797]
18. Vaccine for mpox prevention in the United States. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/mpox/hcp/vaccine-considerations/index.html [accessed 2025-05-27]
19. Mpox vaccination programme. Centre for Health Protection. URL: https://www.chp.gov.hk/en/features/106090.html [accessed 2023-11-19]
20. Current CDC vaccine price list. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/vaccines-for-children/php/awardees/current-cdc-vaccine-price-list.html [accessed 2024-11-10]
21. Bavarian Nordic boosted by monkeypox vaccine. Financial Times. URL: https://www.ft.com/content/4b63d5ae-ab71-4db4-85ea-fb67ac633600 [accessed 2024-11-10]
22. Cheung DH, Chen S, Fang Y, Sun F, Zhang Q, Yu FY, et al. Influences of mpox disease perceptions, sources and contents of information exposure on mpox vaccine uptake among gay, bisexual, and other men who have sex with men in Hong Kong, China. Vaccine 2024 Apr 02;42(9):2337-2346. [doi: 10.1016/j.vaccine.2024.02.083] [Medline: 38480102]
23. Huang MF, Chang YP, Lin CW, Yen CF. Factors related to mpox-vaccine uptake among men who have sex with men in Taiwan: roles of information sources and emotional problems. Vaccines (Basel) 2024 Mar 20;12(3):332 [FREE Full text] [doi: 10.3390/vaccines12030332] [Medline: 38543965]

XSL•FO
RenderX

24. Palich R, Jedrzejewski T, Schneider L, Qatib N, Orriere T, Gosset D, et al. High uptake of vaccination against mpox in men who have sex with men (MSM) on HIV pre-exposure prophylaxis (PrEP) in Paris, France. Sex Transm Infect 2023 Dec 28;99(8):552-553 [FREE Full text] [doi: 10.1136/sextrans-2023-055885] [Medline: 37507222]
25. Gan G, Janhavi A, Tong G, Lim JT, Dickens BL. The need for pre-emptive control strategies for mpox in Asia and Oceania. Infect Dis Model 2024 Mar;9(1):214-223 [FREE Full text] [doi: 10.1016/j.idm.2023.12.005] [Medline: 38293686]
26. Ren F, Liu J, Miao J, Xu Y, Zhang R, Fan J, et al. Public awareness, specific knowledge, and worry about mpox (monkeypox): a preliminary community-based study in Shenzhen, China. Front Public Health 2023 Feb 14;11:1077564 [FREE Full text] [doi: 10.3389/fpubh.2023.1077564] [Medline: 36866102]
27. Owens C, Hubach RD. An exploratory study of the mpox media consumption, attitudes, and preferences of sexual and gender minority people assigned male at birth in the United States. LGBT Health 2023 Jul 03;10(5):401-407. [doi: 10.1089/lgbt.2022.0251] [Medline: 36735618]
28. May T, Towler L, Smith LE, Horwood J, Denford S, Rubin GJ, et al. Mpox knowledge, behaviours and barriers to public health measures among gay, bisexual and other men who have sex with men in the UK: a qualitative study to inform public health guidance and messaging. BMC Public Health 2023 Nov 17;23(1):2265 [FREE Full text] [doi: 10.1186/s12889-023-17196-0] [Medline: 37978506]
29. Edinger A, Valdez D, Walsh-Buhi E, Trueblood JS, Lorenzo-Luaces L, Rutter LA, et al. Misinformation and public health messaging in the early stages of the mpox outbreak: mapping the Twitter narrative with deep learning. J Med Internet Res 2023 Jun 06;25:e43841 [FREE Full text] [doi: 10.2196/43841] [Medline: 37163694]
30. Wang G, Ma X. Representations of LGBTQ+ issues in China in its official English-language media: a corpus-assisted critical discourse study. Crit Discourse Stud 2020 Mar 06;18(2):188-206. [doi: 10.1080/17405904.2020.1738251]
31. Manalastas EJ, Ojanen TT, Torre BA, Ratanashevorn R, Hong BC, Kumaresan V, et al. Homonegativity in Southeast Asia: attitudes toward lesbians and gay men in Indonesia, Malaysia, the Philippines, Singapore, Thailand, and Vietnam. Asia Pac Soc Sci Rev 2017 Jun 30;17(1). [doi: 10.59588/2350-8329.1120]
32. Hagger MS, Orbell S. The common sense model of illness self-regulation: a conceptual review and proposed extended model. Health Psychol Rev 2022 Sep;16(3):347-377 [FREE Full text] [doi: 10.1080/17437199.2021.1878050] [Medline: 33461402]
33. Chan RC, Mak WW. Common sense model of mental illness: understanding the impact of cognitive and emotional representations of mental illness on recovery through the mediation of self-stigma. Psychiatry Res 2016 Dec 30;246:16-24. [doi: 10.1016/j.psychres.2016.09.013] [Medline: 27644016]
34. Luo S, She R, Lau MM, Lau JT. Would illness representations of COVID-19 and COVID-19 fears during clinic visits promote or reduce behavioral intention to seek medical consultations for flu symptoms? A random telephone survey in Hong Kong, China. Front Public Health 2022 Jun 10;10:903290 [FREE Full text] [doi: 10.3389/fpubh.2022.903290] [Medline: 35757650]
35. Broadbent E, Wilkes C, Koschwanez H, Weinman J, Norton S, Petrie KJ. A systematic review and meta-analysis of the Brief Illness Perception Questionnaire. Psychol Health 2015 Aug 26;30(11):1361-1385. [doi: 10.1080/08870446.2015.1070851] [Medline: 26181764]
36. Limbu YB, Gautam RK, Zhou W. Predicting vaccination intention against COVID-19 using theory of planned behavior: a systematic review and meta-analysis. Vaccines (Basel) 2022 Nov 26;10(12):2026 [FREE Full text] [doi: 10.3390/vaccines10122026] [Medline: 36560436]
37. Griffin B, Conner M, Norman P. Applying an extended protection motivation theory to predict Covid-19 vaccination intentions and uptake in 50-64 year olds in the UK. Soc Sci Med 2022 Apr;298:114819 [FREE Full text] [doi: 10.1016/j.socscimed.2022.114819] [Medline: 35245755]
38. Pot M, van Keulen HM, Ruiter RA, Eekhout I, Mollema L, Paulussen TW. Motivational and contextual determinants of HPV-vaccination uptake: a longitudinal study among mothers of girls invited for the HPV-vaccination. Prev Med 2017 Jul;100:41-49 [FREE Full text] [doi: 10.1016/j.ypmed.2017.04.005] [Medline: 28389328]
39. Ye L, Fang T, Cui J, Zhu G, Ma R, Sun Y, et al. The intentions to get vaccinated against influenza and actual vaccine uptake among diabetic patients in Ningbo, China: identifying motivators and barriers. Hum Vaccin Immunother 2021 Jan 02;17(1):106-118 [FREE Full text] [doi: 10.1080/21645515.2020.1761201] [Medline: 32460620]
40. Jongen VW, Groot Bruinderink ML, Boyd A, Koole JC, Teker B, Dukers-Muijrers NH, et al. What determines mpox vaccination uptake? Assessing the effect of intent-to-vaccinate versus other determinants among men who have sex with men. Vaccine 2024 Jan 12;42(2):186-193 [FREE Full text] [doi: 10.1016/j.vaccine.2023.12.018] [Medline: 38072753]
41. Gilbert M, Ablona A, Chang HJ, Grennan T, Irvine MA, Sarai Racey C, et al. Uptake of mpox vaccination among transgender people and gay, bisexual and other men who have sex with men among sexually-transmitted infection clinic clients in Vancouver, British Columbia. Vaccine 2023 Apr 06;41(15):2485-2494 [FREE Full text] [doi: 10.1016/j.vaccine.2023.02.075] [Medline: 36894397]
42. Xiao X. Follow the heart or the mind? Examining cognitive and affective attitude on HPV vaccination intention. Atl J Commun 2019 Dec 26;29(2):93-105. [doi: 10.1080/15456870.2019.1708743]

XSL•FO
RenderX

43. Gao S, Zeng Z, Xin Q, Yang M, Feng X, Liu X, et al. Global transboundary transmission path and risk of mpox revealed with Least Cost Path model. Int J Infect Dis 2024 Sep;146:107101 [FREE Full text] [doi: 10.1016/j.ijid.2024.107101] [Medline: 38777082]
44. Kong TS. Transnational queer sociological analysis of sexual identity and civic-political activism in Hong Kong, Taiwan and mainland China. Br J Sociol 2019 Dec 12;70(5):1904-1925. [doi: 10.1111/1468-4446.12697] [Medline: 31402452]
45. CHP investigates confirmed mpox case. The Government of the Hong Kong Special Administrative Region. URL: https://www.info.gov.hk/gia/general/202408/02/P2024080200436.htm?fontSize=1 [accessed 2024-08-11]
46. Weinman J, Petrie KJ. Illness perceptions: a new paradigm for psychosomatics? J Psychosom Res 1997 Feb;42(2):113-116. [doi: 10.1016/s0022-3999(96)00294-2] [Medline: 9076639]
47. Kawuki J, Nambooze J, Chan PS, Chen S, Liang X, Mo PK, et al. Differential practice and associated factors of COVID-19 personal preventive measures among the slum and estate communities of Uganda: a community-based cross-sectional survey. J Glob Health 2023 Sep 15;13:06039 [FREE Full text] [doi: 10.7189/jogh.13.06039] [Medline: 37712388]
48. Nanteer-Oteng E, Kretchy IA, Nanteer DO, Kretchy JP, Osafo J. Hesitancy towards COVID-19 vaccination: the role of personality traits, anti-vaccine attitudes and illness perception. PLOS Glob Public Health 2022 Dec 28;2(12):e0001435 [FREE Full text] [doi: 10.1371/journal.pgph.0001435] [Medline: 36962915]
49. Hoyle RH. Handbook of Structural Equation Modeling. New York, NY: Guilford Publications; 2012.
50. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012 Jan 01;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]
51. Kline RB. Principles and Practice of Structural Equation Modeling Fifth Edition. New York, NY: Guilford Press; 2023.
52. Hayes AF, Coutts JJ. Use omega rather than Cronbach's alpha for estimating reliability. But…. Commun Methods Meas 2020 Feb 11;14(1):1-24. [doi: 10.1080/19312458.2020.1718629]
53. Nevitt J, Hancock G. Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. Struct Equ Modeling 2001 Jul 1;8(3):353-377. [doi: 10.1207/s15328007sem0803_2]
54. Bollen KA, Stine RA. Bootstrapping goodness-of-fit measures in structural equation models. Sociol Methods Res 1992 Nov 01;21(2):205-229. [doi: 10.1177/0049124192021002004]
55. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
56. Hooper D, Coughlan J, Mullen MR. Evaluating model fit: a synthesis of the structural equation modelling literature. Dublin Institute Technol 2008 [FREE Full text] [doi: 10.21427/D79B73]
57. Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Modeling 2002 Apr;9(2):233-255. [doi: 10.1207/s15328007sem0902_5]
58. Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct Equ Modeling 2007 Dec 05;14(3):464-504. [doi: 10.1080/10705510701301834]
59. Kaufman JS, MacLehose RF. Which of these things is not like the others? Cancer 2013 Dec 15;119(24):4216-4222 [FREE Full text] [doi: 10.1002/cncr.28359] [Medline: 24022386]
60. Revelle W. Using the psych package to generate and test structural models. The Personality Project. 2017 Jan 7. URL: https://personality-project.org/r/psych_for_sem.pdf [accessed 2025-05-27]
61. Rosseel Y. lavaan: an R package for structural equation modeling. J Stat Softw 2012;48(2):1-36 [FREE Full text] [doi: 10.18637/jss.v048.i02]
62. Ankan A, Wortel IM, Textor J. Testing graphical causal models using the R package "dagitty". Curr Protoc 2021 Feb 16;1(2):e45. [doi: 10.1002/cpz1.45] [Medline: 33592130]
63. DAGitty v3.1. Dagitty. URL: https://dagitty.net/dags.html?id=REaJCCcR [accessed 2025-05-28]
64. Al-Swidi A, Al Yahya M. Training transfer intention and training effectiveness: assessing the gender differences using multi-group structural equation modelling approach. Int J Organ Anal 2017 Nov 06;25(5):839-860 [FREE Full text] [doi: 10.1108/ijoa-07-2016-1043]
65. Mo PK, Chan VW, Wang X, Lau JT. Gender difference in the association between internet addiction, self-esteem and academic aspirations among adolescents: a structural equation modelling. Comput Educ 2020 Oct;155:103921. [doi: 10.1016/j.compedu.2020.103921]
66. Hirschfeld G, von Brachel R. Multiple-group confirmatory factor analysis in R – a tutorial in measurement invariance with continuous and ordinal indicators. Pract Assess Res Eval 2014 Jul;19(7) [FREE Full text] [doi: 10.7275/qazy-2946]
67. MacIntyre R, Grulich A, Hammoud M. Mpox is on the rise in Australia. Here's what to know about the virus – and who should get vaccinated. UNSW. 2024 Jun 5. URL: https://www.unsw.edu.au/newsroom/news/2024/06/mpox-on-rise-Australia [accessed 2024-08-11]
68. Yang Z, Wei L, Xie W, Chen L, Yang Z, Zhang Y, et al. Estimating changes in population size and behavioral characteristics in men who have sex with men between 2014 and 2019: longitudinal study. JMIR Public Health Surveill 2022 Aug 16;8(8):e34150 [FREE Full text] [doi: 10.2196/34150] [Medline: 35972779]
69. Movahedi Nia Z, Bragazzi N, Asgary A, Orbinski J, Wu J, Kong J. Mpox panic, infodemic, and stigmatization of the two-spirit, lesbian, gay, bisexual, transgender, queer or questioning, intersex, asexual community: geospatial analysis, topic

XSL•FO
RenderX

modeling, and sentiment analysis of a large, multilingual social media database. J Med Internet Res 2023 May 01;25:e45108 [FREE Full text] [doi: 10.2196/45108] [Medline: 37126377]
70. Monkeypox | The Philippines recorded its first case this year, and many Southeast Asian countries are on high alert. Hong Kong News. 2024. URL: https://tinyurl.com/bdn9ffhz [accessed 2025-06-05]
71. Cao X, Zheng N. How does disgust affect mpox prevention? Examining the underlying mechanisms of perceived severity and perceived susceptibility moderated by stigma. Health Commun 2025 Apr 25;40(4):713-724 [FREE Full text] [doi: 10.1080/10410236.2024.2364377] [Medline: 38916051]
72. Devine PG, Tauer JM, Barron KE, Elliot AJ, Vance KM, Harmon-Jones E. Moving beyond attitude change in the study of dissonance-related processes: an update on the role of discomfort. In: Harmon-Jones E, editor. Cognitive Dissonance: Reexamining a Pivotal Theory in Psychology. Washington, DC: American Psychological Association; 2019.
73. Kung KH. Hong Kong's mpox situation and public health measures. Centre for Health Protection. 2023 Dec 7. URL: https://icidportal.ha.org.hk/Home/File?path=/Training%20Calendar/182/Surveillance%20and%20public%20health%20response%20to%20mpox%20in%20Hong%20Kong.pdf [accessed 2025-05-27]
74. Hong Kong Pride Parade. Facebook. URL: https://www.facebook.com/watch/?v=1111865123496834 [accessed 2024-08-11]
75. Musa S, Cilovic-Lagarija S, Kavazovic A, Bosankic-Cmajcanin N, Stefanelli A, Scott NA, et al. COVID-19 risk perception, trust in institutions and negative affect drive positive COVID-19 vaccine intentions. Int J Public Health 2022 Apr 11;67:1604231 [FREE Full text] [doi: 10.3389/ijph.2022.1604231] [Medline: 35479763]
76. Chung GK, Chan YH, Chan SM, Chen JK, Wong H, Chung RY. The impact of trust in government on pandemic management on the compliance with voluntary COVID-19 vaccination policy among adolescents after social unrest in Hong Kong. Front Public Health 2022 Sep 16;10:992895 [FREE Full text] [doi: 10.3389/fpubh.2022.992895] [Medline: 36660556]
77. Fu L, Sun Y, Li Y, Wang B, Yang L, Tian T, et al. Perception of and vaccine readiness towards mpox among men who have sex with men living with HIV in China: a cross-sectional study. Vaccines (Basel) 2023 Feb 23;11(3):528 [FREE Full text] [doi: 10.3390/vaccines11030528] [Medline: 36992114]
78. Zheng M, Qin C, Qian X, Yao Y, Liu J, Yuan Z, et al. Knowledge and vaccination acceptance toward the human monkeypox among men who have sex with men in China. Front Public Health 2022;10:997637 [FREE Full text] [doi: 10.3389/fpubh.2022.997637] [Medline: 36388271]
79. Clinical overview of mpox. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/mpox/hcp/clinical-overview/index.html [accessed 2025-05-27]
80. Sun Y, Lu H, Ye J, Li D, Li G. Awareness and use of HIV pre-exposure prophylaxis and factors associated with awareness among MSM in Beijing, China. Sci Rep 2023 Jan 11;13(1):554 [FREE Full text] [doi: 10.1038/s41598-023-27485-8] [Medline: 36631515]
81. Zheng M, Du M, Yang G, Yao Y, Qian X, Zhi Y, et al. Mpox vaccination hesitancy and its associated factors among men who have sex with men in China: a national observational study. Vaccines (Basel) 2023 Aug 30;11(9):1432 [FREE Full text] [doi: 10.3390/vaccines11091432] [Medline: 37766109]
82. PRiSM - HIV Prevalence and Risk behavioural Survey of Men who have sex with men in Hong Kong 2017. Special Preventive Programme, Centre for Health Protection, Department of Health. 2018 May. URL: https://www.aids.gov.hk/english/surveillance/sur_report/prism2017e.pdf [accessed 2025-05-27]

Abbreviations

CFI: comparative fit index
CSM: common sense model of illness self-regulation
GBMSM: gay, bisexual, and other men who have sex with men
LGBTQ+: lesbian, gay, bisexual, transgender, and queer
PrEP: pre-exposure prophylaxis
RMSEA: root mean square error of approximation
S-CFI: scaled comparative fit index
SEM: structural equation modeling
S-RMSEA: scaled root mean square error of approximation
STI: sexually transmitted infection
S-TLI: scaled Tucker-Lewis index

XSL•FO
RenderX

Edited by A Mavragani; submitted 28.12.24; peer-reviewed by P Magalia Duarte, R Adesola, S Salajegheh Tazerji; comments to author 13.03.25; revised version received 22.04.25; accepted 07.05.25; published 18.06.25.

Please cite as:
Cheung D, Xie L, Wang L, Chen S, Li X, Zhang Z, Chen X, Ge S, Yu FY, Fang Y, Chen Z, Li Z, Sun F, Mo P, Liu Y, Wang Z
How Information Exposure Shapes Risk Perceptions and Vaccination Intentions Among Gay, Bisexual, and Other Men Who Have Sex With Men: Cross-Sectional Survey Study
JMIR Public Health Surveill 2025;11:e70635
URL: https://publichealth.jmir.org/2025/1/e70635
doi:10.2196/70635
PMID:

©Doug Cheung, Luyao Xie, Lijuan Wang, Siyu Chen, Xinge Li, Zheng Zhang, Xinyue Chen, Shen Ge, Fuk-yuen Yu, Yuan Fang, Zihuang Chen, Zhennan Li, Fenghua Sun, Phoenix Mo, Yingjie Liu, Zixin Wang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.06.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Determinants of Health-Promoting Behaviors Among Indonesian Adolescents Living in Child Welfare Institutions: Cross-Sectional Survey Study

Aloysia Ispriantari[1,2], PhD; Hyejung Lee[1,3], PhD; Sue Kim[1,3], PhD; Hyeonkyeong Lee[1,3], PhD; Anna Lee[1,3], PhD; Chang Park[4], PhD

[1]College of Nursing, Yonsei University, Seoul, Republic of Korea

[2]Department of Nursing, Faculty of Health Sciences, Institute of Technology and Health Science RS dr Soepraoen, Malang, Indonesia

[3]Mo-Im Kim Nursing Research Institute, College of Nursing, Yonsei University, Seoul, Republic of Korea

[4]College of Nursing, University of Illinois Chicago, Chicago, United States

Corresponding Author:
Hyejung Lee, PhD
Mo-Im Kim Nursing Research Institute
College of Nursing
Yonsei University
50-1, Yonsei-ro, Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 3345
Email: hlee26@yuhs.ac

Abstract

Background: Adolescents in child welfare institutions often face inadequate facility resources, limited caregiver support, and restricted access to health care services. These obstacles impede their physical, mental, and social development during adolescence, resulting in significant health vulnerabilities. Engaging in health-promoting behaviors (HPBs) can enhance their overall health and quality of life, potentially contributing to improved long-term well-being.

Objective: This study aimed to identify the HPBs of adolescents in child welfare institutions in Indonesia and explore the determinants influencing these behaviors using the health promotion model.

Methods: This study used a cross-sectional survey design. Adolescents living in welfare institutions were recruited from January 14, 2024, to February 3, 2024. After obtaining institutional review board approval, 6 research assistants visited 17 institutions in Malang and collected data from participants who provided their consent using tablet PCs that linked to the questionnaire. The variables studied included HPBs, health literacy, self-esteem, perceived barriers to action, perceived self-efficacy, and social support. Multivariate structural analysis was conducted using SPSS Statistics (version 26.0; IBM Corp) and SPSS Amos (version 26.0; IBM Corp).

Results: A total of 276 adolescents participated in this study. Adolescents' HPBs differed significantly based on age group (P=.03), educational level (P=.04), duration of stay in child welfare institutions (P=.03), and the institutions' accreditation level (P=.02). In the final model, perceived self-efficacy (β=0.538; P<.001) and social support (β=0.256; P<.001) together accounted for 47.9% of the variance in HPBs. Self-esteem was positively correlated with perceived self-efficacy (β=0.184; P<.001) and social support (β=0.303; P<.001) but negatively correlated with perceived barriers to action (β=−0.194; P<.01). Health literacy was also negatively correlated with perceived barriers to action (β=−0.234; P<.001). Self-esteem indirectly affected HPBs through perceived self-efficacy (β=0.099; P<.01) and social support (β=0.078; P<.001).

Conclusions: To improve the HPBs of adolescents living in child welfare institutions, their self-esteem needs to be increased to further enhance their self-efficacy and social support. Careful attention and monitoring of HPBs among these adolescents may lead to better health outcomes and support their transition from child welfare institutions to the broader community.

(JMIR Public Health Surveill 2025;11:e75024) doi:10.2196/75024

XSL•FO
RenderX

KEYWORDS

adolescents; health-promoting behaviors; child welfare institutions; Indonesia; health promotion model; structural equation model

Introduction

Background

Adolescents living in child welfare institutions—commonly referred to as orphanages or childcare institutions—face unique challenges that may adversely affect their overall health and well-being. These adolescents are more prone to experiencing physical, mental, and social difficulties than their peers living in family environments [1]. Furthermore, these challenges can lead to hopelessness and hinder their ability to thrive [2].

Health-promoting behaviors (HPBs) refer to individuals' actions and choices to improve their health and prevent the occurrence of diseases [3]. The World Health Organization emphasizes HPBs as a key strategy for improving health outcomes and enhancing quality of life [4]. During adolescence, engagement in HPBs is linked to improved academic performance, stronger interpersonal relationships, and a lower risk of future chronic diseases [5]. Thus, adopting HPBs during adolescence not only is beneficial for the present but also serves as a crucial investment in their future health and well-being.

Although HPBs have many advantages, barriers to their adoption exist for adolescents in child welfare institutions. Limited facility resources and a lack of caregiver support to fulfill adolescents' physical and psychological needs can lead to low self-esteem and poor decision-making skills [6]. In addition, poor national health insurance coverage and inadequate government registration may lead to low awareness and limited access to health care services, further restricting HPBs [7].

This is especially true in the Indonesian context, which faces its own unique challenges. Because of family economic hardships, adolescents are placed in child welfare institutions with the expectation that they will receive better care and education than they did with their families [8]. However, the vast majority of institutions in Indonesia are private and lack adequate government oversight and support, meaning that they do not always comply with national care standards [9]. Consequently, adolescents living in such conditions become more susceptible to unhealthy behaviors, including smoking, poor sleep, low physical activity, and challenges in maintaining overall well-being [10]. Despite the recognized importance of HPBs during adolescence, research on the factors influencing HPBs among adolescents living in child welfare institutions in Indonesia remains limited.

The health promotion model (HPM) developed by Pender and Murdaugh [3] has frequently been used to examine the relationship between relevant factors and HPBs in the adolescent population. The HPM encompasses 3 interconnected domains: individual characteristics and experiences, behavior-specific cognitions and affect, and behavioral outcomes and HPBs. These serve as the end point, with the aim to achieve positive health outcomes. Individual characteristics such as previous behavior, health literacy, self-esteem, and sociodemographic factors significantly shape subsequent actions, with their impact varying depending on the specific behavior being considered. Behavior-specific cognitions and affect—including perceived benefits and barriers to action, perceived self-efficacy, activity-related affect, and interpersonal and situational influences—are highly motivational. These factors can be modified through targeted interventions, making their measurement essential for evaluating intervention effectiveness.

Previous studies have established an explicit connection between various factors that are considered determinants of HPBs among adolescents [11-14]. Perceived barriers to action were the strongest and most important predictor of HPBs, exhibiting a negative relationship—the lower the perceived barriers to action, the higher the HPBs among adolescents and young adults [11,12]. Perceived self-efficacy also plays a crucial role in determining HPBs, with higher perceived self-efficacy correlating with increased HPBs [13]. Higher perceived self-efficacy is also correlated with lower perceived barriers to action [14].

Social support also contributes to HPBs, with greater social support associated with higher levels of HPBs [15,16]. Moreover, social support can enhance a sense of belonging, which, in turn, increases perceived self-efficacy and further promotes HPBs among adolescents. This is especially critical during adolescence, when individuals begin to understand health-related information and make independent health decisions [17]. Health literacy not only reduces perceived barriers to action but also improves perceived self-efficacy [18,19]. Furthermore, individuals with higher health literacy are more likely to seek and receive greater social support from their networks [20].

Self-esteem as a key individual characteristic is a significant factor. Higher self-esteem reduces perceived barriers to action while simultaneously promoting greater perceived self-efficacy and social support [21-23].

Objectives

This study aimed to identify the HPBs of adolescents in child welfare institutions in Indonesia and explore the determinants influencing these behaviors using the HPM. Given the relationships between the variables reviewed, this study examined health literacy and self-esteem as components of the individual characteristic domain and perceived barriers to action, perceived self-efficacy, and social support as elements of the behavior-specific cognitions domain. The HPBs examined as behavioral outcomes in this study were health responsibility, physical activity, nutrition, positive life perspective, interpersonal relationships, stress management, and spiritual health [24]. We specifically hypothesized that perceived barriers to action would be negatively correlated with HPBs, whereas perceived self-efficacy and social support would be positively correlated with HPBs. The findings of this study will provide insights into the challenges and opportunities for promoting HPBs in this population by identifying their determinants using structural equation modeling (SEM).

XSL•FO
RenderX

Methods

Study Design and Setting

This study used a cross-sectional design. Conducted in January 2024 and February 2024, this study took place in Malang, Indonesia. First, convenience sampling was used to select the child welfare institutions in Malang. A total of 17 child welfare institutions agreed to participate in this study. Inclusion criteria for this study were adolescents aged 13 to 18 years who had been living in child welfare institutions for at least 6 months and were willing to participate. Adolescents with developed mental or physical disabilities were excluded.

The sample size estimation for SEM is flexible. Several rules of thumb are commonly used, such as 10 cases per variable, 5 or 10 observations per estimated parameter, and a minimum sample size of 100 to 200 [25]. A sample size of 200 is often considered the gold standard in SEM and is widely used in studies [26,27]. Thus, this recommendation was applied while accounting for a 20% dropout rate, as reported in previous studies [28,29]. The required sample size for this study was calculated to be a minimum of 240.

Six research assistants collected data from participants. These assistants were third- and fifth-semester undergraduate nursing students who had completed the pediatric nursing theory and clinical practicum courses. The principal investigator provided instructions on the study's purpose and managed the link to the questionnaire and other relevant procedures. During data collection, research assistants visited the institutions and presented the questionnaire to participants using tablets or notebooks, allowing them to complete the survey via the KoboToolbox platform. In addition, when participants needed help understanding any question, research assistants provided individualized explanations. Participants were given sufficient time to provide their answers, and completing the questionnaires took approximately 45 minutes.

Before the answers were submitted online, the research assistants verified that no responses were missing. Of the 280 adolescents who participated, 4 (1.4%) were excluded because they selected identical responses to every question. The final analysis included 276 participants.

Ethical Considerations

The institutional review board of the Institute of Technology, Science, and Health RS Dr Soepraoen in Malang, Indonesia, approved this study (approval KEPK-EC-10/XII/2023). As adolescents are a vulnerable group requiring guardian consent for participation, the heads of the child welfare institutions provided written consent for the study, and all participants were asked to provide assent before data collection. The researchers ensured the privacy and confidentiality of participant data by not recording personal identifiers and by storing data in password-protected web-based storage accessible only to the research team. Participants who completed the questionnaire received a souvenir and a lunch box with an equivalent value of approximately US $1.50.

General Participant Demographics

Demographic information about the participants included age (in years), sex (male or female), educational level (primary school, junior high school, senior high school, or undergraduate), orphan status (maternal orphan, paternal orphan, double orphan, or not orphan), and duration of residence in child welfare institutions (in years). In addition, the accreditation level of the child welfare institutions was recorded. Indonesian child welfare institutions receive accreditation ratings of A (excellent), B (good), or C (satisfactory) or are unaccredited, reflecting varying levels of compliance with national standards for child welfare services. Higher ratings indicate a greater capacity to provide quality care within an institution.

HPB Measurement

In this study, HPBs refer to actions taken by adolescents to maintain and improve their health within the environment of a child welfare institution. The Adolescent Lifestyle Profile–Revised 2 [30] was used to measure HPBs. This instrument consists of 44 items across 7 domains: health responsibility, physical activity, nutrition, positive life perspective, interpersonal relationships, stress management, and spiritual health. The items were rated using a 4-point Likert scale (1=*never*, 2=*sometimes*, 3=*often*, and 4=*always*), with total possible scores ranging from 44 to 176. Higher scores indicate better HPBs among adolescents. The overall Cronbach α was 0.86 in this study.

Health Literacy Measurement

Health literacy refers to adolescents' ability to seek, understand, evaluate, and use health information to make informed decisions about their health. The Health Literacy Assessment Scale for Adolescents [31], consisting of 15 items, was used to measure health literacy. The instrument includes 3 subscales: communication, confusion, and functional health literacy. Items were rated on a 5-point Likert scale (1=*always*; 5=*never*), with total possible scores ranging from 15 to 75. Higher scores indicate greater health literacy. The Cronbach α for the scale was 0.70 in this study.

Self-Esteem Measurement

Self-esteem refers to adolescents' self-perception and evaluation of their own personal worth and abilities. The Rosenberg Self-Esteem Scale [32] adapted for the Indonesian context [33] was used to measure adolescents' self-esteem. This instrument consists of 8 items rated on a 4-point Likert scale (1=*strongly disagree*, 2=*disagree*, 3=*agree*, and 4=*strongly agree*) for items 1, 2, 3, 6, and 8, with reverse scoring applied to items 4, 5, and 7. The scores range from 8 to 32, with higher scores indicating higher self-esteem. In this study, the Cronbach α was 0.68.

Perceived Barrier to Action Measurement

Perceived barriers to action refer to adolescents' perceptions of the barriers or obstacles that prevent them from engaging in HPBs. The Barriers to Health-Promoting Activities Scale [34], consisting of 18 items, was used to measure this perception. Items were rated on a 4-point Likert scale (1=*never*; 4=*routinely*), with total possible scores ranging from 18 to 72.

XSL•FO
RenderX

Higher scores indicate greater perceived barriers to HPBs. In this study, the Cronbach α of the scale was 0.85.

Perceived Self-Efficacy Measurement

Perceived self-efficacy refers to adolescents' belief in their ability to perform and maintain HPBs. The Self-Rated Abilities for Health and Practices Scale–Adolescent Version [35] was used. This instrument includes 28 items across 4 subscales: nutrition, psychological well-being, exercise, and responsible health practices. This instrument was rated on a 5-point Likert scale (1=*cannot do at all*, 2=*a little*, 3=*somewhat*, 4=*mostly*, and 5=*certainly can do*), with total possible scores ranging from 28 to 140. Higher scores indicate greater perceived self-efficacy among adolescents. In this study, the Cronbach α was 0.92.

Social Support Measurement

Social support refers to the emotional, instrumental, appraisal, and informational support that adolescents receive from caregivers, peers, and friends to help them engage in HPBs. The original Child and Adolescent Social Support Scale for Healthy Behaviors [36] is used to assess social support for healthy behaviors in children and adolescents from parents, teachers, classmates, close friends, and people in the school environment. For this study, this instrument was adapted to focus on 3 sources of support: caregivers, peers, and friends residing in child welfare institutions. This instrument consists of 36 items and 2 subscales: frequency and importance of social support. Frequency responses were rated on a 6-point Likert scale (1=*never*; 6=*always*), and importance responses were rated on a 3-point Likert scale (1=*not important*; 3=*very important*), with total possible scores ranging from 72 to 324. Higher scores indicate greater social support. The Cronbach α of the scale was 0.96 in this study.

Translation Process

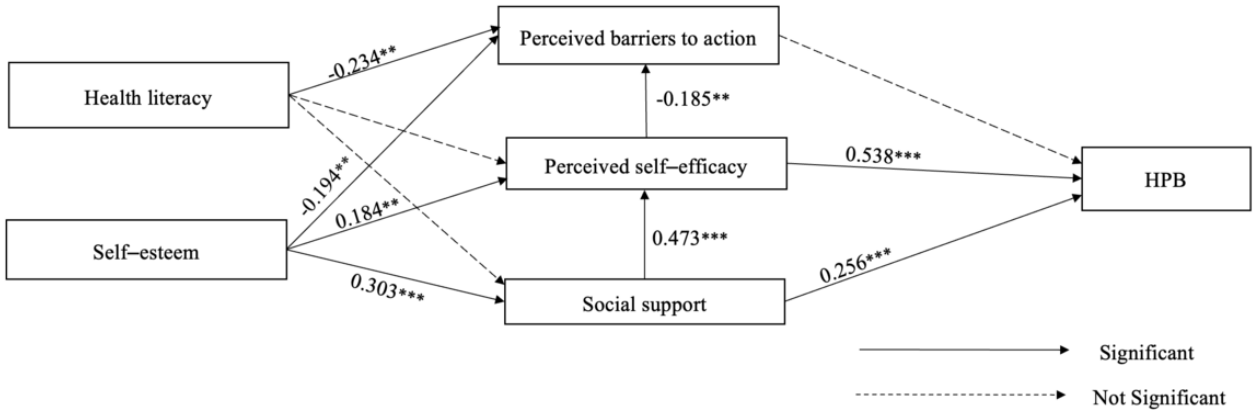
Among the instruments, the following 5 questionnaires—the Adolescent Lifestyle Profile–Revised 2, Health Literacy Assessment Scale for Adolescents, Barriers to Health-Promoting Activities Scale, Self-Rated Abilities for Health and Practices Scale–Adolescent Version, and Child and Adolescent Social Support Scale for Healthy Behaviors—were translated into Indonesian using forward and backward translation techniques before use [37]. To ensure the validity of the translated scales, a translation process was followed step by step. Two faculty members from the pediatric nursing department and one with a background in English literature who had worked in the health field for >5 years took part in the following translation process: (1) 2 translators independently translated the original instrument from English into Indonesian (version A); (2) a third translator, who was unfamiliar with the original instruments, translated version A back into English (version B); (3) all translators reviewed and compared version B with the original version to identify conceptual discrepancies, adapt culturally relevant terms, and develop a refined forward translation (version C); (4) for cognitive debriefing, 5 adolescents were interviewed to assess whether the translated items were easily understood and accurately conveyed the intended meaning of the original questionnaire (version D); and (5) the translators reconvened to review version D for conceptual equivalence and make final proofreading adjustments, resulting in the finalized version used in this study (version E).

Statistical Analysis

SPSS Statistics (version 26.0; IBM Corp) was used to conduct descriptive analyses of variables, internal consistency tests, independent 2-tailed *t* tests, 1-way ANOVAs, Pearson correlations, and post hoc analyses. The SPSS Amos software (version 26.0; IBM Corp) was used to conduct path analyses examining the regression coefficients and the effects (direct, indirect, and total) among variables, as well as evaluating the fit of the structural model.

The determinants of HPBs among adolescents living in child welfare institutions were identified using SEM. In this study, 6 variables included in the model were observed variables with a single indicator. The 2 exogenous variables—health literacy and self-esteem—served as independent variables (predictors) of the other observed variables in the model. The 4 endogenous variables included perceived barriers to action, perceived self-efficacy, social support, and HPBs. HPBs were identified as the outcome of the model, whereas perceived barriers to action, perceived self-efficacy, and social support were identified as mediators between the exogenous variables and HPBs. Figure 1 shows the hypothesized model for this study.

Figure 1. Path coefficient diagram of the final model of health-promoting behaviors (HPBs) among Indonesian adolescents living in child welfare institutions. *P<.05; **P<.01; ***P<.001.



XSL•FO
RenderX

A structural equation model was also used to test the hypotheses. First, the model was identified, followed by an evaluation of its overall fit. The number of distinct sample moments in this model was 21, and the number of distinct parameters to be estimated was 18. Therefore, the *df* were calculated as 21−18=3. As the *df* were positive, a minimum was achieved, allowing for further testing [27]. Four tests were used to assess the goodness of fit: the Tucker-Lewis index (>0.90), comparative fit index (>0.90), standardized root mean square residual (<0.08), and root mean square error of approximation (<0.08) [38]. Bootstrapping analysis with 5000 samples and 95% bias-corrected CIs determined the significance of both indirect and direct effects.

Results

Participant Demographic and HPB Analysis

Of the 276 participants, 149 (54%) were female. Most (155/276, 56.2%) were aged 13 to 15 years, with a mean age of 15.3 (SD 1.7) years. In addition, 43.1% (119/276) were in senior high school, whereas 42.8% (118/276) were in junior high school. Approximately half (141/276, 51.1%) of the participants were not orphans, and 60.9% (168/276) had been living in child welfare institutions for 1 to 3 years, with a mean duration of 3.4 (SD 2.5) years. Table 1 presents participants characteristics and comparison of HPBs.

Table 1. Participant characteristics and comparison of health-promoting behaviors (HPBs) between groups (N=276).

Variable and category	Participants, n (%)	Variable mean (SD)	HPBs—ALP-R2[a] (score range 44-176), mean (SD)	*P* value
Age (y)		15.3 (1.7)		.03
13-15	155 (56.2)		122.0 (15.0)	
16-18	121 (43.8)		125.0 (12.6)	
Sex		—[b]		.40
Male	127 (46)		124.2 (14.3)	
Female	149 (54)		122.5 (13.9)	
Educational level (school)		—		.04
Primary	28 (10.1)		116.8 (13.1)	
Junior high school	118 (42.8)		123.6 (14.7)	
Senior high school	119 (43.1)		125.0 (13.4)	
College	11 (4)		118.3 (12.5)	
Orphan status		—		.73
Not orphan	141 (51.1)		124.2 (14.1)	
Maternal	23 (8.3)		120.9 (13.9)	
Paternal	88 (31.9)		123.8 (14.5)	
Double	24 (8.7)		119.0 (11.5)	
Duration of residence in child welfare institutions (y)		3.4 (2.5)		.03
1-3	168 (60.9)		123.6 (14.0)	
4-5	77 (27.9)		125.2 (14.1)	
>5	31 (11.2)		117.0 (13.3)	
Accreditation level of institutions		—		.02
A	20 (7.2)		117.3 (15.0)	
B	125 (45.3)		125.5 (13.4)	
C	51 (18.5)		118.2 (13.9)	
Unaccredited	80 (29)		124.7 (14.0)	

[a]ALP-R2: Adolescent Lifestyle Profile–Revised 2.

[b]Not applicable.

HPB scores were significantly different by age group (*P*=.03), educational level (*P*=.04), duration of residence in child welfare institutions (*P*=.03), and accreditation level of institutions (*P*=.02). On the basis of the post hoc analysis, adolescents aged 13 to 15 years had lower HPB scores than adolescents aged 16 to 18 years (*P*=.03). Primary school students had lower HPB scores than college students (*P*=.03). HPB scores were higher among those living in child welfare institutions for 1 to 3 years and 4 to 5 years than among those living in them for >5 years (*P*<.05). Adolescents in level B institutions had higher HPB

XSL•FO
RenderX

scores than those in level C institutions (*P*=.02), and adolescents living in unaccredited institutions had higher HPB scores than those living in level C institutions (*P*=.03).

Correlation Between Major Variables

The correlation analysis showed that HPBs had a significant positive correlation with perceived self-efficacy (*r*=0.66; *P*<.001), social support (*r*=0.53; *P*<.001), health literacy (*r*=0.13; *P*=.03), and self-esteem (*r*=0.31; *P*<.001). Conversely, HPBs had a significant negative correlation with perceived barriers to action (*r*=−0.15; *P*=.01). Table 2 presents the results of the correlation analysis between variables.

Table 2. Correlation analysis between variables.

	Health literacy	Self-esteem	Perceived barriers to action	Perceived self-efficacy	Social support
Health literacy					
r	1	0.26	−0.31	0.16	0.02
P value	—[a]	<.001	<.001	<.001	.76
Self-esteem					
r	0.26	1	−0.32	0.35	0.29
P value		—	<.001	<.001	<.001
Perceived barriers to action					
r	−0.31	−0.32	1	−0.29	−0.23
P value	<.001	<.001	—	<.001	<.001
Perceived self-efficacy					
r	0.16	0.35	−0.29	1	0.53
P value	.01	<.001	<.001	—	<.001
Social support					
r	0.02	0.29	−0.23	0.53	1
P value	.76	<.001	<.001	<.001	—
HPBs[b]					
r	0.13	0.31	−0.15	0.66	0.53
P value	.03	<.001	.01	<.001	<.001

[a]Not applicable

[b]HPB: health-promoting behavior.

HPBs and Significant Determinants

The model in this study indicated a good fit (Tucker-Lewis index=0.947; comparative fit index=0.989; standardized root mean square residual=0.000; root mean square error of approximation=0.070). The results of the SEM analysis are presented in Table 3. In the HPB pathway, only perceived self-efficacy (β=0.538; *P*<.001) and social support (β=0.256; *P*<.001) had significant positive correlations with HPBs. These 2 variables explained 47.9% of the variance in HPBs. In the perceived barriers to action pathway, perceived self-efficacy (β=−0.185; *P*<.01), health literacy (β=−0.234; *P*<.001), and self-esteem (β=−0.194; *P*<.01) showed significant negative correlations with perceived barriers to action, collectively explaining 18.8% of its variance. In the perceived self-efficacy pathway, only self-esteem (β=0.184; *P*<.001) and social support (β=0.473; *P*<.001) had significant positive correlations, explaining 32.8% of its variance. In the social support pathway, only self-esteem (β=0.303; *P*<.001) had a significant positive correlation with social support, explaining 8.6% of its variance.

XSL•FO
RenderX

Table 3. Effect coefficients of the health-promoting behavior (HPB) model among adolescents living in child welfare institutions.

Dependent and independent variable	Direct effect (β)	Indirect effect (β)	Total effect (β)	SMC[a]
HPBs				0.479
Perceived barriers to action	0.065	—[b]	0.065	
Perceived self-efficacy	0.538[c]	−0.012	0.526[c]	
Social support	0.256[c]	0.249[c]	0.505[c]	
Perceived barriers to action				0.188
Perceived self-efficacy	−0.185[d]	—	−0.185[d]	
Social support	—	−0.088[d]	−0.088[d]	
Health literacy	−0.234[d]	−0.013	−0.247[c]	
Self-esteem	−0.194[d]	−0.061[d]	−0.254[c]	
Perceived self-efficacy				0.328
Health literacy	0.098	−0.029	0.069	
Self-esteem	0.184[d]	0.143[c]	0.327[d]	
Social support	0.473[c]	—	0.473[c]	
Social support				0.086
Health literacy	−0.061	—	−0.061	
Self-esteem	0.303	—	0.303[c]	

[a]SMC: squared multiple correlation.

[b]Not applicable.

[c]*P*<.001.

[d]*P*<.01.

In addition, there was a significant indirect effect of self-esteem on HPBs through perceived self-efficacy (β=0.099; *P*<.001), social support (β=0.078; *P*<.001), and the combination of social support and perceived self-efficacy (β=0.077; *P*<.001). The final model of HPBs among adolescents living in child welfare institutions is shown in Figure 1.

Discussion

Principal Findings

This study explored the determinants of HPBs among adolescents living in Indonesian child welfare institutions. It primarily found significant differences in adolescents' HPBs based on age group, educational level, duration of stay in child welfare institutions, and the accreditation level of the institutions. Older adolescents and those with higher educational levels demonstrated better HPBs, which may be attributed to greater responsibility and autonomy [39]. Interestingly, adolescents who had lived in child welfare institutions for >5 years had significantly lower HPBs than those who had lived there for shorter periods, such as 1 to 3 years and 4 to 5 years. This may be due to adaptation to regular daily schedules, although prolonged stays without proper knowledge might lead to the normalization of poor HPBs. In addition, adolescents who have just entered these institutions might receive more attention and support from caregivers and peers than long-term residents. These findings provide valuable insights into the necessity and timing of delivering internal education to promote healthy behaviors. In terms of accreditation levels, adolescents living in level B institutions showed significantly higher HPBs than those in level C institutions. However, adolescents living in level A institutions did not demonstrate better HPBs than those living in level B or C institutions. Higher-level institutions are assumed to have better access to services and more structured systems than lower-level ones, potentially offering adolescents more resources to engage in good HPBs. These unexpected findings might be due to the small sample size of level A institutions and the fact that the accreditation process primarily focuses on environmental and human support factors rather than the promotion of health behaviors among the residents. This aspect is suggested as a criterion in the future accreditation process for Indonesian child welfare institutions.

Through the SEM analysis, this study found that perceived self-efficacy and social support together accounted for 47.9% of adolescents' HPBs. HPBs are key to happiness and a good quality of life, and self-efficacy increases HPBs at every stage of life [40]. When individuals have confidence in their ability to perform HPBs, they are more likely to engage in these behaviors [4]. Among adolescents in child welfare institutions, social support appears to play a crucial role in the development of good HPBs, helping bring more structure to their daily activities.

XSL•FO
RenderX

This study challenges existing frameworks by highlighting that perceived barriers to action were not significant to HPBs in this context, providing insights into the unique needs of these adolescents. This finding aligns with findings from Thailand [41]. This could be attributed to the unique environment of child welfare institutions, where structured activities and support systems mitigate the barriers typically faced by adolescents living with their families. The organized nature of these institutions—with regulated schedules and caregiver supervision—can encourage adolescents to engage in HPBs, often leaving health decision-making to caregivers and the institution. This is supported by the fact that most participants in this study (196/276, 71%) lived in accredited child welfare institutions (level A, B, or C), which have already met the minimum national care standards for children.

Interestingly, health literacy did not correlate with perceived self-efficacy in this study. This finding differs from those of studies conducted in Australia, China, Turkey, and Germany, which found a positive correlation between health literacy and adolescent self-efficacy [18]. This discrepancy may be due to adolescents in child welfare institutions having moderate health literacy but lacking confidence in performing HPBs. In general, adolescents are not fully independent at this stage and often lack autonomy in their activities, as well as practical experience in carrying them out [42].

In addition, this study identified the link between self-esteem, social support, and perceived self-efficacy and HPBs, a connection that has emerged in recent literature on this unique population. It may provide a foundation for health professionals to develop interventions that address HPBs by improving self-esteem and enhancing social support for targeted adolescents such as those in this study.

Comparison With Prior Work

No previous studies—to our knowledge—have examined HPBs among adolescents living in child welfare institutions. Much of the existing research has focused on healthy adolescent populations, with studies conducted in Turkey [43], Portugal [44], Iran [45], and India [46]. However, this study explored HPBs among a specific group of adolescents—those living in child welfare institutions—which has been understudied in Indonesia and other countries. This study also highlights several factors under the HPM that influence HPBs, providing a comprehensive understanding of the determinants of these behaviors.

This study found that self-esteem had a significant indirect effect on HPBs through perceived self-efficacy and social support. This supports previous studies indicating that higher self-esteem improves perceived self-efficacy, which is crucial for problem-solving and adopting HPBs [47,48]. Individuals with high self-esteem tend to have greater perceived self-efficacy in problem-solving [49,50].

This study also found that only self-esteem showed a significant positive correlation with social support and self-efficacy and indirectly affected HPBs. Higher self-esteem leads to more active social engagement, thereby increasing perceived social support [51]. Interestingly, self-esteem explained only 8.6% of the variance in social support. Social support is a complex and multidimensional construct [52], and we assessed it only from caregivers, peers, and friends living in child welfare institutions. In addition, only 8.7% (24/276) of the participants were double orphans, indicating that most participants still had family members with whom they maintained contact, which may also influence their social support. The limitation of the social support measurement may have affected the strength of the observed relationship, and future studies should investigate the role of family support among adolescents living in child welfare institutions who are not double orphans.

Furthermore, self-esteem was found to be moderate in this study, which is consistent with previous research indicating that adolescents living in child welfare institutions generally exhibit low to moderate levels of self-esteem [53]. Hence, interventions to improve self-esteem could enhance social support and, consequently, HPBs. In child welfare institutions, fostering a nurturing environment that enhances self-esteem and interpersonal relationships can be used to boost adolescents' perceived self-efficacy, thereby promoting positive HPBs.

This study also confirmed that social support is positively correlated with perceived self-efficacy, which helps reduce perceived barriers to action. These findings align with those of a study from the United States that showed that higher perceived self-efficacy reduces perceived barriers to physical activity—a component of HPBs [54]. Greater social support leads to higher perceived self-efficacy and fewer perceived barriers. Given the limited studies on this topic—especially among adolescents—this study provides valuable insights into the relationships among self-esteem, social support, perceived self-efficacy, and perceived barriers to action.

The key finding of this study was the nuanced relationship between health literacy and HPBs among adolescents in child welfare institutions. Surprisingly, health literacy did not significantly influence perceived self-efficacy or social support, suggesting that the structured environment of these institutions provided consistent information and support, reducing reliance on individual health literacy. In addition, adolescents are still in the process of developing their health-related knowledge, and the lack of health education interventions within child welfare institutions may limit the impact of health literacy on their ability to benefit from social support. However, lower health literacy significantly increased perceived barriers to action, indicating that, even within a supportive environment, understanding health information is crucial for overcoming challenges. The structured environment of the institutions mitigated the negative impact of these perceived barriers to HPBs, suggesting that consistent support and resources can buffer the effects of perceived obstacles. These findings underscore the need for interventions that enhance health literacy, empowering adolescents to overcome barriers while leveraging the supportive nature of structured environments in child welfare institutions to promote healthier behaviors and improve adolescents' overall well-being.

Clinical and Policy Implications

The findings of this study highlight the need for policies specific to the improvement of HPBs in adolescents living in child

XSL•FO
RenderX

welfare institutions in Indonesia. The large number of unrecorded child welfare institutions hinders the provision of nursing interventions to these children and adolescents. Coordination between the Ministry of Social Affairs and Ministry of Health is essential, starting with data collection from child welfare institutions and regular monitoring of the health and welfare of children in these institutions. This will enable health professionals to provide optimal health services and interventions for these adolescents.

While school-based interventions are often considered effective, a review found that these interventions were insufficient for improving HPBs [55]. This is consistent with conditions in Indonesia, where the concept of school nurses or other health professionals at school is not widespread. Only a few private schools have health care professionals who work full time or part time. This may lead to inadequate health interventions in schools. Public health centers are the closest access points for adolescents. Public health centers conduct annual health screenings for students and have a youth health care program that usually provides health services and education for schools at least twice a year. Within these programs, health professionals can also administer HPB assessments and provide education for adolescents in schools.

It has also been suggested to extend interventions beyond the school setting into the home [55]. This is feasible in Indonesia. This study found that social support from caregivers and friends in child welfare institutions was more significant than that from peers. Unfortunately, public health centers do not have specific programs for child welfare institutions. However, they have an Islamic boarding school health care center (*poskestren*) program, which focuses on health services to Islamic boarding schools. Public health centers can include child welfare institutions in this program, conducting health interventions for adolescents, caregivers, and other children in these institutions, thus improving HPBs among adolescents in Indonesia.

Study Limitations

Despite our best efforts, a few limitations exist in this study. First, convenience sampling methods were used in a single city, so generalizing the findings should be done with caution. Second, this study had a cross-sectional design, which requires careful interpretation of causality as bidirectional relationships between variables may exist. Therefore, we encourage future studies to use a longitudinal design to gain a deeper understanding of HPBs among adolescents living in child welfare institutions over time. Third, because half of the participants had at least one parent, social support from parents—which was not assessed—may have influenced the level of social support they received. Therefore, we suggest that future studies assess social support from parents and its potential impact on HPBs. Fourth, exploratory factor analysis was conducted; however, no confirmatory factor analyses were conducted on the translated instruments, which may have impacted the validity of the measures. Finally, although the research assistants were trained to ensure consistent interpretation of the instruments, periodic meetings were not held to calibrate interpretations or address issues that arose during data collection. This may have affected the reliability and accuracy of the data collection process.

Conclusions

On the basis of the HPM and testing using SEM analysis, both perceived self-efficacy and social support directly affected HPBs, whereas self-esteem indirectly affected HPBs among adolescents in child welfare institutions in Indonesia. In addition, health literacy only influenced perceived barriers to action, but these perceived barriers to action did not contribute to HPBs. The key findings of our study contribute to the development of health interventions aimed at improving self-efficacy and social support by enhancing self-esteem, ultimately leading to increased HPBs among adolescents living in child welfare institutions in Indonesia. We hope that these interventions will lead to improved health outcomes and quality of life for these individuals as they transition into adulthood and move from child welfare institutions to the broader community.

Acknowledgments

This study was supported by the Yonsei University College of Nursing and the Lee (SK) Mija Global Leadership Research Scholarship. The authors would also like to thank all the heads and caregivers of the 17 child welfare institutions for granting them access to participants. They also extend their gratitude to the 6 nursing students who assisted with data collection. The authors declare that no generative artificial intelligence was used while writing the manuscript. The funder had no involvement in the study design; data collection, analysis, or interpretation; or writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AI and HL developed the study concept. SK, HL, and AL contributed to the study design, and CP provided statistical expertise for the analysis. All authors contributed to writing the manuscript and approved the final version.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Bhatt KB, Apidechkul T, Srichan P, Bhatt N. Depressive symptoms among orphans and vulnerable adolescents in childcare homes in Nepal: a cross-sectional study. BMC Psychiatry 2020 Sep 25;20(1):466 [FREE Full text] [doi: 10.1186/s12888-020-02863-y] [Medline: 32977787]
2. Öztürk Ş, Ekinci M. The effect of structured education on self-esteem and the suicide probability of male adolescents living in orphanages. Arch Psychiatr Nurs 2018 Aug;32(4):604-609 [FREE Full text] [doi: 10.1016/j.apnu.2018.03.011] [Medline: 30029755]
3. Pender NJ, Murdaugh CL, Parsons MA. Health Promotion in Nursing Practice. London, UK: Pearson Education; 2011.
4. Bakouei F, Jalil Seyedi-Andi S, Bakhtiari A, Khafri S. Health promotion behaviors and its predictors among the college students in Iran. Int Q Community Health Educ 2018 Jun 18;38(4):251-258 [FREE Full text] [doi: 10.1177/0272684x18781780]
5. Anderson E, Durstine JL. Physical activity, exercise, and chronic diseases: a brief review. Sports Med Health Sci 2019 Dec;1(1):3-10 [FREE Full text] [doi: 10.1016/j.smhs.2019.08.006] [Medline: 35782456]
6. Kholifah S. Empowerment of orphanage caregivers with the IHC model to improve adolescent psychological needs. Front Community Serv Empow 2022 Sep 24;1(3) [FREE Full text] [doi: 10.35882/ficse.v1i3.18]
7. Moffa M, Cronk R, Fejfar D, Dancausse S, Padilla LA, Bartram J. A systematic scoping review of hygiene behaviors and environmental health conditions in institutional care settings for orphaned and abandoned children. Sci Total Environ 2019 Mar 25;658:1161-1174 [FREE Full text] [doi: 10.1016/j.scitotenv.2018.12.286] [Medline: 30677980]
8. Jafar N, Nusu AC, Suriah S. The implementation of balanced nutrition using “Piring Makanku” on food consumption of orphanage children in Makassar city. Open Access Maced J Med Sci 2020 Sep 15;8(T2):75-80 [FREE Full text] [doi: 10.3889/oamjms.2020.5202]
9. Roja ML. Kebijakan pendidikan anak terlantar di panti asuhan St. Louis de Monfort kota Kupang. SocioEdu 2020;1(1):1-10 [FREE Full text] [doi: 10.59098/socioedu.v1i1.238]
10. Mansoer WW, Putri RD, Sulaiman N. Attachment to significant figures, resilience, and delinquency among adolescents in orphanages in Jakarta. In: Proceedings of the 2nd International Conference on Intervention and Applied Psychology. 2018 Presented at: ICIAP 2018; September 13-14, 2018; West Java, Indonesia. [doi: 10.2991/iciap-18.2019.14]
11. Ferreira Silva RM, Mendonça CR, Azevedo VD, Raoof Memon A, Noll PR, Noll M. Barriers to high school and university students' physical activity: a systematic review. PLoS One 2022 Apr 4;17(4):e0265913 [FREE Full text] [doi: 10.1371/journal.pone.0265913] [Medline: 35377905]
12. Xiang B, Wong HM, Perfecto AP, McGrath CP. Modelling health belief predictors of oral health and dental anxiety among adolescents based on the Health Belief Model: a cross-sectional study. BMC Public Health 2020 Nov 23;20(1):1755 [FREE Full text] [doi: 10.1186/s12889-020-09784-1] [Medline: 33228621]
13. Wang CW, Chen DR. Associations of sugar-sweetened beverage knowledge, self-efficacy, and perceived benefits and barriers with sugar-sweetened beverage consumption in adolescents: a structural equation modeling approach. Appetite 2022 Jan 01;168:105663 [FREE Full text] [doi: 10.1016/j.appet.2021.105663] [Medline: 34455023]
14. Neumann RJ, Ahrens KF, Kollmann B, Goldbach N, Chmitorz A, Weichert D, et al. The impact of physical fitness on resilience to modern life stress and the mediating role of general self-efficacy. Eur Arch Psychiatry Clin Neurosci 2022 Jun 07;272(4):679-692 [FREE Full text] [doi: 10.1007/s00406-021-01338-9] [Medline: 34622343]
15. Khalid A, Morawska A, Turner KM. Pakistani orphanage caregivers' perspectives regarding their caregiving abilities, personal and orphan children's psychological wellbeing. Child Care Health Dev 2023 Jan;49(1):145-155 [FREE Full text] [doi: 10.1111/cch.13027] [Medline: 35771173]
16. Mahasneh AM. The relationship between subjective well-being and social support among Jordanian university students. Psychol Russ 2022;15(2):53-64 [FREE Full text] [doi: 10.11621/pir.2022.0204] [Medline: 36699712]
17. Ma C, Ren Z, Chen Z, Li C. The association between interactive health literacy and dietary behaviors among Chinese college students: a large-scale cross-sectional study. Front Psychol 2024 May 30;15:1363885 [FREE Full text] [doi: 10.3389/fpsyg.2024.1363885] [Medline: 38873517]
18. Ceylan D, Akan-Çelen FN, Özkan S, Aycan Z. Promoting adolescent health: health literacy, self-efficacy and internet use. Turk J Pediatr 2022;64(1):110-121 [FREE Full text] [doi: 10.24953/turkjped.2021.1264] [Medline: 35286037]
19. Poon PK, Tam KW, Lam T, Luk AK, Chu WC, Cheung P, et al. Poor health literacy associated with stronger perceived barriers to breast cancer screening and overestimated breast cancer risk. Front Oncol 2022;12:1053698 [FREE Full text] [doi: 10.3389/fonc.2022.1053698] [Medline: 36686831]
20. Liu B, Tian L, Yang S, Wang X, Luo J. Effects of multidimensional self-esteems on health promotion behaviors in adolescents. Front Public Health 2022;10:847740 [FREE Full text] [doi: 10.3389/fpubh.2022.847740] [Medline: 35558531]
21. Zhou Y, Fan C, Zhang S. A longitudinal examination of loneliness in left-behind children: the interaction between self-esteem and academic self-efficacy matters. BMC Psychol 2025 Apr 02;13(1):327 [FREE Full text] [doi: 10.1186/s40359-025-02635-x] [Medline: 40176175]
22. Hu S, Cai D. Development of perceived family support and positive mental health in junior high school students: a three-year longitudinal study. J Early Adolesc 2022 May 11;43(2):216-243. [doi: 10.1177/02724316221101534]

XSL•FO
RenderX

23. Gonzalez SZ. Ethnic identity, self-efficacy, academic adjustment, and anxiety symptoms in Latinx undergraduates. The Florida State University. 2023. URL: https://repository.lib.fsu.edu/islandora/object/fsu:928100/datastream/PDF/download [accessed 2025-08-21]
24. Gaete J, Olivares E, Godoy MI, Cárcamo M, Montero-Marín J, Hendricks C, et al. Adolescent Lifestyle Profile-Revised 2: validity and reliability among adolescents in Chile. J Pediatr (Rio J) 2021;97(1):52-60 [FREE Full text] [doi: 10.1016/j.jped.2019.11.005] [Medline: 32084438]
25. Hahs-Vaughn DL. Applied Multivariate Statistical Concepts. Milton Park, UK: Taylor & Francis; 2024.
26. Crockett S. A five-step guide to conducting SEM analysis in counseling research. Counsel Outcome Res Eval 2017 Mar 16;3(1):30-47 [FREE Full text] [doi: 10.1177/2150137811434142]
27. Kline RB. Principles and Practice of Structural Equation Modeling. New York, NY: Guilford Publications; 1998.
28. Cho S, Lee H, Oh EG, Kim GS, Kim YC, Park CG. Health-related quality of life among migrant workers: the impact of health-promoting behaviors. Nurs Health Sci 2020 Jun 31;22(2):318-327. [doi: 10.1111/nhs.12660] [Medline: 31667923]
29. Kim SJ, Yoo IY. Health promotion behavior of Chinese international students in Korea including acculturation factors: a structural equation model. Asian Nurs Res (Korean Soc Nurs Sci) 2016 Mar;10(1):25-31 [FREE Full text] [doi: 10.1016/j.anr.2015.10.008] [Medline: 27021831]
30. Hendricks C, Murdaugh C, Pender N. The Adolescent Lifestyle Profile: development and psychometric characteristics. J Natl Black Nurses Assoc 2006 Dec;17(2):1-5. [Medline: 17410753]
31. Manganello JA, DeVellis RF, Davis TC, Schottler-Thal C. Development of the Health Literacy Assessment Scale for Adolescents (HAS-A). J Commun Healthc 2015 Jun 17;8(3):172-184 [FREE Full text] [doi: 10.1179/1753807615y.0000000016]
32. Rosenberg M. Society and the Adolescent Self-Image. Princeton, NJ: Princeton University Press; 2015.
33. Alwi MA, Razak A. Adaptasi Rosenberg's Self-Esteem di Indonesia. Universitas Negeri Makassar. 2022. URL: https://download.garuda.kemdikbud.go.id/article.php?article=3509418&val=30680&title=Adaptasi%20Rosenberg%27s%20Self-Esteem%20di%20Indonesia [accessed 2025-08-22]
34. Becker H, Stuifbergen AK, Sands D. Development of a scale to measure barriers to health promotion activities among persons with disabilities. Am J Health Promot 1991;5(6):449-454 [FREE Full text] [doi: 10.4278/0890-1171-5.6.449] [Medline: 10171668]
35. Chilton JM, Gosselin KP, Haas BK. Development of the self-rated abilities for health practices-adolescent version: a self-efficacy measure. J Nurs Meas 2018 Apr 01;26(1):134-141. [doi: 10.1891/1061-3749.26.1.134] [Medline: 29724284]
36. Menon V, Demaray MK. Child and adolescent social support scale for healthy behaviors: scale development and assessment of the relation between targeted social support and body size dissatisfaction. Children's Health Care 2013 Jan;42(1):45-66 [FREE Full text] [doi: 10.1080/02739615.2013.753800]
37. Lee WL, Chinna K, Lim Abdullah K, Zainal Abidin I. The forward-backward and dual-panel translation methods are comparable in producing semantic equivalent versions of a heart quality of life questionnaire. Int J Nurs Pract 2019 Feb;25(1):e12715 [FREE Full text] [doi: 10.1111/ijn.12715] [Medline: 30515964]
38. Wang YA, Rhemtulla M. Power analysis for parameter estimation in structural equation modeling: a discussion and tutorial. Adv Meth Pract Psychol Sci 2021 Mar 23;4(1) [FREE Full text] [doi: 10.1177/2515245920918253]
39. Halpern-Felsher B, Baker M, Stitzel S. Decision-making in adolescents and young adults. In: Diefenbach M, Miller-Halegoua S, Bowen D, editors. Handbook of Health Decision Science. New York, NY: Springer; 2016:157-167.
40. Sheeran P, Maki A, Montanaro E, Avishai-Yitshak A, Bryan A, Klein WM, et al. The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: a meta-analysis. Health Psychol 2016 Nov;35(11):1178-1188. [doi: 10.1037/hea0000387] [Medline: 27280365]
41. Sangsawang J, Supavititpatana B, Sriarporn P. Factors related to health promoting behaviors among adolescent mothers. Nurs J CMU 2019;46(4):59-69 [FREE Full text]
42. Tabrizi JS, Doshmangir L, Khoshmaram N, Shakibazadeh E, Abdolahi HM, Khabiri R. Key factors affecting health promoting behaviors among adolescents: a scoping review. BMC Health Serv Res 2024 Jan 11;24(1):58 [FREE Full text] [doi: 10.1186/s12913-023-10510-x] [Medline: 38212786]
43. Ozturk F, Ayaz-Alkaya S. Health literacy and health promotion behaviors of adolescents in Turkey. J Pediatr Nurs 2020;54:e31-e35 [FREE Full text] [doi: 10.1016/j.pedn.2020.04.019] [Medline: 32446664]
44. Tomás CC, Queirós PJ, Ferreira TD. Health-promoting behaviors: psychometric properties of an assessment tool. Texto Contexto Enferm 2015 Mar;24(1):22-29. [doi: 10.1590/0104-07072015000700014]
45. Musavian AS, Pasha A, Rahebi SM, Atrkar Roushan Z, Ghanbari A. Health promoting behaviors among adolescents: a cross-sectional study. Nurs Midwifery Stud 2014 Apr;3(1):e14560 [FREE Full text] [doi: 10.17795/nmsjournal14560] [Medline: 25414892]
46. Swaminathan SS, Boratne AV, Patil R, Sankaran A. Health promoting behavior among school going adolescents in Puducherry-a cross-sectional study. Indian J Community Health 2020 Sep 30;32(3):479-485. [doi: 10.47203/ijch.2020.v32i03.004]

XSL•FO
RenderX

47. Asakereh A, Yousofi N. Reflective thinking, self-efficacy, self-esteem and academic achievement of Iranian EFL students. Int J Educ Psychol 2018;7(1):68-89 [FREE Full text] [doi: 10.17583/ijep.2018.2896]
48. Saunders JF, Frazier LD, Nichols-Lopez KA. Self-esteem, diet self-efficacy, body mass index, and eating disorders: modeling effects in an ethnically diverse sample. Eat Weight Disord 2016 Sep;21(3):459-468 [FREE Full text] [doi: 10.1007/s40519-015-0244-6] [Medline: 26703131]
49. Flynn DM, Chow P. Self-efficacy, self-worth and stress. Education 2017;138(1):83. [doi: 10.4324/9781315523019-3]
50. Mailey EL, Phillips SM, Dlugonski D, Conroy DE. Overcoming barriers to exercise among parents: a social cognitive theory perspective. J Behav Med 2016 Aug;39(4):599-609 [FREE Full text] [doi: 10.1007/s10865-016-9744-8] [Medline: 27108160]
51. Shi J, Wang L, Yao Y, Su N, Zhao X, Zhan C. Family function and self-esteem among Chinese university students with and without grandparenting experience: moderating effect of social support. Front Psychol 2017;8:886 [FREE Full text] [doi: 10.3389/fpsyg.2017.00886] [Medline: 28611720]
52. Acoba EF. Social support and mental health: the mediating role of perceived stress. Front Psychol 2024 Feb 21;15:1330720 [FREE Full text] [doi: 10.3389/fpsyg.2024.1330720] [Medline: 38449744]
53. Mohammadzadeh M, Awang H, Kadir Shahar H, Ismail S. Emotional health and self-esteem among adolescents in Malaysian orphanages. Community Ment Health J 2018 Jan;54(1):117-125 [FREE Full text] [doi: 10.1007/s10597-017-0128-5] [Medline: 28315972]
54. Dishman RK, McIver KL, Dowda M, Saunders RP, Pate RR. Self-efficacy, beliefs, and goals: moderation of declining physical activity during adolescence. Health Psychol 2019 Jun;38(6):483-493 [FREE Full text] [doi: 10.1037/hea0000734] [Medline: 30973746]
55. Nyman J, Tornivuori A, Salanterä S, Barroso T, Parisod H. Systematic review of digital interventions to support refusal self-efficacy in child and adolescent health promotion. Health Promot Int 2022 Oct 01;37(5):daac085 [FREE Full text] [doi: 10.1093/heapro/daac085] [Medline: 36166268]

Abbreviations

HPB: health-promoting behavior
HPM: health promotion model
SEM: structural equation modeling

Edited by A Mavragani, T Sanchez; submitted 27.03.25; peer-reviewed by A Adeoye, S Du, C Onah; comments to author 21.05.25; revised version received 12.07.25; accepted 07.08.25; published 17.09.25.

Please cite as:
Ispriantari A, Lee H, Kim S, Lee H, Lee A, Park C
Determinants of Health-Promoting Behaviors Among Indonesian Adolescents Living in Child Welfare Institutions: Cross-Sectional Survey Study
JMIR Public Health Surveill 2025;11:e75024
URL: https://publichealth.jmir.org/2025/1/e75024
doi:10.2196/75024
PMID:40962318

©Aloysia Ispriantari, Hyejung Lee, Sue Kim, Hyeonkyeong Lee, Anna Lee, Chang Park. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Ultraprocessed Foods Associated With Depressive Symptoms Among Chinese Adolescents: Cross-Sectional Study

Honglv Xu[1], MD, PhD; Jieru Yang[1], MSc; Dongyue Hu[1], MSc; Xiaoxiao Li[1], MSc; Feihui Zhou[1], MD

School of Medicine, Kunming University, kunming, China

Corresponding Author:
Honglv Xu, MD, PhD
School of Medicine
Kunming University
Puxin Road 2
Kunming, 650214
China
Phone: 86 0871 65036573
Fax: 86 0871 65036573
Email: x hl@126.com

Abstract

Background: Ultraprocessed food (UPF) consumption is very common among adolescents. Previous studies have suggested an association between UPF consumption and adolescent mental health problems, but studies on multiethnic adolescents in China are rare.

Objective: This study investigated the association between UPF consumption and depressive symptoms in multiethnic adolescents in Yunnan Province, China.

Methods: A cluster sampling of 8500 middle school students from 11 counties in Yunnan Province was conducted. The Chinese version of the Depression Anxiety Stress Scale-21 was administered to assess depressive symptoms, and a food frequency questionnaire was used to collect UPF consumption data. A generalized linear model was used to analyze the association between UPF consumption and depressive symptoms, and a restricted cubic spline was used to fit the correlation plot. The association analysis was performed after adjusting for demographic and confounding variables that might have influenced depressive symptoms.

Results: The detection rates of depressive symptoms in adolescents were 28.3% (2402/8500) for the sample, 25.5% (1069/4184) for males, and 30.9% (1333/4316) for females. After adjusting for confounding variables, instant food (all participants: β=.22, 95% CI 0.13-0.32; *P*<.001; males: β=.15, 95% CI 0.03-0.27; *P*=.01; females: β=.29, 95% CI 0.15-0.43; *P*<.001), carbonated beverage (β=.20, 95% CI 0.12-0.29; *P*<.001; β=.15, 95% CI 0.04-0.26; *P*=.006; β=.29, 95% CI 0.16-0.43; *P*<.001;), energy beverage (β=.21, 95% CI 0.10-0.32; *P*<.001; β=.22, 95% CI 0.08-0.36; *P*=.002; β=.22, 95% CI 0.04-0.40; *P*=.01), and other beverage (β=.24, 95% CI 0.14-0.34; *P*<.001; β=.21, 95% CI 0.08-0.35; *P*=.002; β=.25, 95% CI 0.10-0.41; *P*=.001) consumption were associated with depressive symptoms for all participants, males, and females, respectively. Females were more sensitive to the association between UPF consumption and depressive symptoms. Moreover, female interaction with instant foods (β=.31, 95% CI 0.18-0.44; *P*<.001), carbonated beverages (β=.27, 95% CI 0.15-0.39; *P*<.001), energy beverages (β=.18, 95% CI 0.01-0.35; *P*=.04), and other sugar-sweetened beverages (β=.17, 95% CI 0.02-0.32; *P*=.03) was associated with depressive symptoms, respectively.

Conclusions: These results indicated that UPF consumption was related to depressive symptoms in Yunnan multiethnic adolescents. Health education and behavioral interventions that reduce UPF consumption could protect adolescents from depressive symptoms.

(JMIR Public Health Surveill 2025;11:e75061) doi:10.2196/75061

KEYWORDS

adolescents; depressive symptoms; eating behavior; mental health; ultraprocessed food

XSL•FO
RenderX

Introduction

Depression is one of the most common mental illnesses in the world. In 2008, the World Health Organization ranked major depression as the third leading cause of disease burden worldwide and projected that the disease would rank first by 2030 [1]. The Global Burden of Disease Study from 1990 to 2019 shows that depression rates are on the rise [2]. Recently, mental health problems among adolescents have become a widespread concern [3] and are a main disease burden for adolescents [4]. Adolescence is a peak period for mental health problems [5]. Depression is the primary cause of disability in adolescents [6]. Depressive symptoms are the most common psychopathological symptoms in adolescents and significantly increase self-injury risk (odds ratio [OR] 2.19, 95% 2.03-2.36) and death by suicide (OR 2.67, 95% CI 1.72-4.14) [3]. A systematic review and meta-analysis demonstrated that the point prevalence of elevated depressive symptoms among adolescents increased from 24% (95% CI 19%-28%) between 2001 and 2010 to 37% (95% CI 32%-42%) between 2011 and 2020 [7]. A German study reported that the prevalence of depressive symptoms in adolescents aged 12-17 years was 8.2% (95% CI 6.5%-9.9%) [8]. In a UK school-based survey, the prevalence of depressive symptoms among 11- to 16-year-olds ranged from 24% to 32% [9]. The prevalence rates of depressive symptoms among adolescents may be higher in developing countries than in developed countries; the rates were 52.3% in India [10] and 26.2% among Malaysian secondary school students [11]. A meta-analysis of Chinese children and adolescents showed a total prevalence of depressive symptoms of 19.85% (95% CI 14.75%-24.96%), and the prevalence rates in the eastern, central, western, and northeastern regions were 17.8%, 23.7%, 22.7%, and 14.5%, respectively [12]. Our previous sample survey of 10- to 20-year-olds from 4 cities in China determined a 27.3% detection rate for depressive symptoms [13]. Current studies claim depressive symptoms are affected by various factors such as academic pressure, growth and development, eating behavior, physical activity, and social media use [3,14,15].

The global consumption of ultraprocessed foods (UPF) has rapidly increased. In 2010, the NOVA food classification system was proposed [16]. After undergoing physical, biological, and chemical processes and containing multiple ingredients and additives, UPF is separated from nature before consumption or preparation as dishes [16]. UPF has the advantages of attractive packaging, ease of consumption, attractive taste, and high palatability, and has gained popularity, especially among adolescents [17]. UPF consumption by adolescents in developed countries is higher; the proportion of energy intake from UPF in American adolescents is higher than that in adults [18]. The UPF energy intake in British middle school students accounts for 77.8% of their total lunch energy intake [19]. UPF consumption among adolescents in developing countries has also attracted widespread attention. According to a Romanian survey, 18- to 23-year-olds generally consume pastries (41.4%), while sweets (12.4%), pastries (11.1%), and sugar-sweetened beverages (SSB; 11.2%) are highly addictive, and males consume more than females [20]. Data from the Global School Health Survey of 9 Southeast Asian countries suggest that 40% of 12- to 15-year-olds consume carbonated beverages [21]. Adolescents aged 14-17 years in Spain consume 7.72 UPF servings per day; the reported consumption rates for processed meats, chocolate products, snacks, chocolate drinks, and soft drinks were 52.14%, 49.29%, 48.21%, 46.07%, and 33.57%, respectively [22]. UPF consumption by adolescents is associated with social isolation, diet globalization, and parental rearing style [23].

Existing research demonstrates that unhealthy eating behaviors and eating patterns of adolescents are potentially associated with mental health issues [24]. UPF consumption is associated with depressive symptoms (OR 1.44, 95% CI 1.14-1.82) [25] and significantly associated with an increased risk of depression (relative risk 1.31, 95% CI 1.21-1.41) [26]. Western dietary patterns (sweetened beverages, processed or junk foods, and foods rich in saturated fatty acids) have been linked to an increased risk of depression among adolescents [27]. SSB, fried foods, processed meats, and baked products are associated with an increased risk of depression. SSB, sweets, and snack consumption are associated with depressive symptoms in adolescents (OR 1.62, 95% CI 1.35-1.95) [28]. Studies in China, the United Kingdom, Brazil, and other countries have found that the consumption of sweets, biscuits, snacks, and fast foods and depressive symptoms in adolescents are significantly positively correlated [29-32]. For instance, in the Brazilian National School Health Survey, boys and girls who consumed ≥6 UPF reported more depressive symptoms and other psychopathologies than those who consumed ≤3 UPF 24 hours before the survey (boys: β=.27, girls: β=.31) [33].

Although several studies have explored the association between UPF consumption and depressive symptoms in adolescents, their findings have been inconsistent. Studies have shown that the consumption of soft drinks and processed meat products is not associated with depressive symptoms in adolescents [30]. There are limited studies on the association between UPF consumption and depressive symptoms of adolescents in China, and no studies on multiethnic adolescents in China have been conducted. Chinese multiethnic adolescents mainly live in the border areas of China and are influenced by the dietary culture of neighboring countries and multiethnic groups; thus, their UPF consumption may be higher. Therefore, this study investigated the association between UPF consumption and depressive symptoms in Chinese multiethnic adolescents.

Methods

Study Design

This is a cross-sectional study. This study is part of a behavior and subhealth study of Yunnan adolescents that aims to explore the association between lifestyle factors and physical and mental subhealth among adolescents in Yunnan, China, through a series of cohort studies [34].

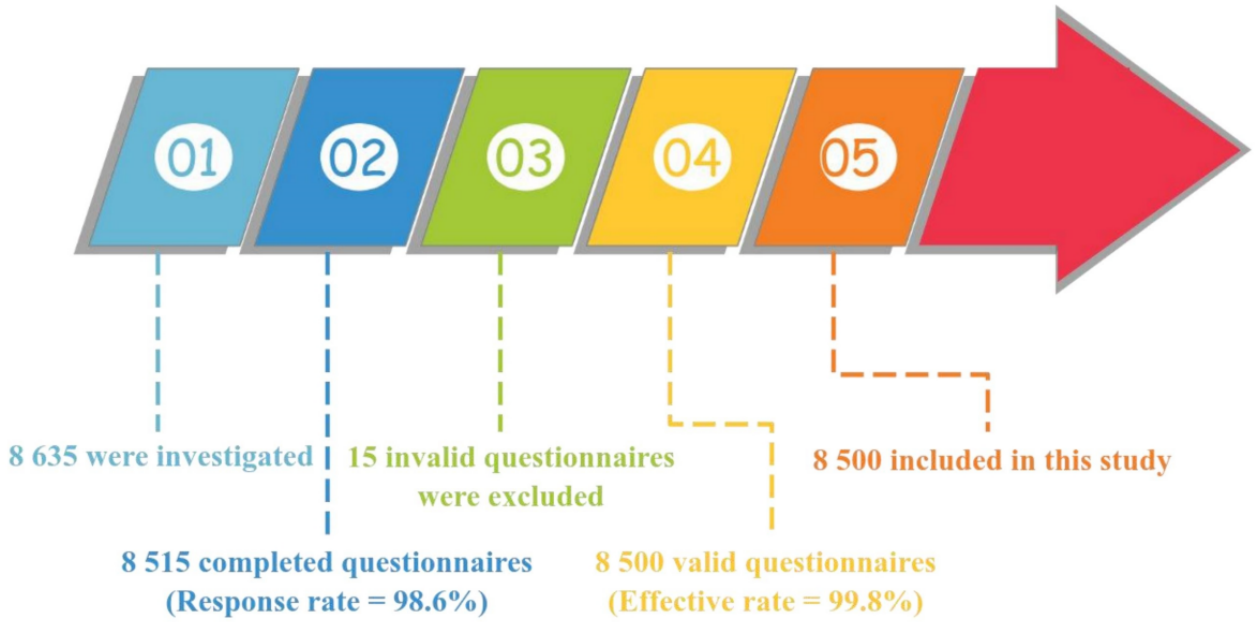
Study Setting and Sample

From October to December 2022, a sample questionnaire survey was conducted with first-year (grade 7) middle school students from 11 counties in Yunnan Province, China. Yunnan Province lies along China's southwestern border with Myanmar, Laos,

XSL•FO
RenderX

and Vietnam. Yunnan has 26 ethnic groups and 25 ethnic minority groups. A total of 8635 multiethnic adolescents were surveyed, and 8515 questionnaires were completed, for a 98.6% response rate. There were 8500 valid questionnaires; the effective response rate was 99.8%. The average age of participants was 12.64 (SD 0.63; range 12-15) years; 4184 (49.2%) were males and 4136 (48.8%) were females; and 6358 (74.8%) were from rural areas and 2142 (25.2%) were from urban areas. There were 3856 (45.3%) Han and 4659 (54.7%) ethnic minority participants. Figure 1 presents the flowchart of participants. Tables 1 and 2 present the distributions of the demographic variables.

Figure 1. Flow chart of participants.



XSL•FO
RenderX

Table 1. Distribution of depressive symptoms in 12-year-old adolescents.

Variables	Normal, n (%)	Depressive symptoms, n (%)	Total, n (%)	Chi-square (*df*)	*P* value
Sex				18.3 (1)	<.001
Male	1280 (47.2)	388 (39.3)	1668 (45.1)		
Female	1431 (52.8)	599 (60.7)	2030 (54.9)		
Ethnicity				14.2 (4)	.007
Han	1143 (42.2)	372 (37.7)	1515 (41)		
Hani	680 (25.1)	255 (25.9)	935 (25.3)		
Yi	460 (17)	168 (17)	628 (17)		
Lisu	76 (2.8)	48 (4.9)	124 (3.4)		
Other	352 (13)	144 (14.6)	496 (13.4)		
The number of close friends				110.2 (2)	<.001
0-2	230 (8.5)	192 (19.5)	422 (11.4)		
3-4	418 (15.4)	198 (20.1)	616 (16.7)		
≥5	2063 (76.1)	597 (60.5)	2660 (71.9)		
Residence				4.6 (1)	.03
Rural	1822 (67.2)	700 (70.9)	2522 (68.2)		
Urban	889 (32.8)	287 (29.1)	1176 (31.8)		
The only child in the family				5.8 (1)	.02
Yes	406 (15)	180 (18.2)	586 (15.8)		
No	2305 (85)	807 (81.8)	3112 (84.2)		
Family type				16.6 (3)	.001
2-parent family	2259 (83.3)	765 (77.5)	3024 (81.8)		
1-parent family	250 (9.2)	121 (12.3)	371 (10)		
Combined family	171 (6.3)	84 (8.5)	255 (6.9)		
Other	31 (1.1)	17 (1.7)	48 (1.3)		
Father’s education level				9.3 (4)	.05
Illiteracy	268 (9.9)	126 (12.8)	394 (10.7)		
Elementary school	567 (20.9)	197 (20)	764 (20.7)		
Secondary school	1295 (47.8)	470 (47.6)	1765 (47.7)		
High school	372 (13.7)	136 (13.8)	508 (13.7)		
University	209 (7.7)	58 (5.9)	267 (7.2)		
Mother’s education level				7.4 (4)	.12
Illiteracy	457 (16.9)	197 (20)	654 (17.7)		
Elementary school	588 (21.7)	221 (22.4)	809 (21.9)		
Secondary school	1148 (42.3)	398 (40.3)	1546 (41.8)		
High school	328 (12.1)	117 (11.9)	445 (12)		
University	190 (7)	54 (5.5)	244 (6.6)		
Father’s occupation				12.7 (5)	.03
Civil servant	1186 (43.7)	459 (46.5)	1645 (44.5)		
Worker	215 (7.9)	63 (6.4)	278 (7.5)		
Staff	569 (21)	211 (21.4)	780 (21.1)		
Merchant	260 (9.6)	64 (6.5)	324 (8.8)		
Farmer	109 (4)	40 (4.1)	149 (4)		

XSL•FO
RenderX

Variables	Normal, n (%)	Depressive symptoms, n (%)	Total, n (%)	Chi-square (*df*)	*P* value
Other	372 (13.7)	150 (15.2)	522 (14.1)		
Mother's occupation				8.8 (5)	.12
Civil servant	1285 (47.4)	503 (51)	1788 (48.4)		
Worker	178 (6.6)	48 (4.9)	226 (6.1)		
Staff	345 (12.7)	127 (12.9)	472 (12.8)		
Merchant	298 (11)	88 (8.9)	386 (10.4)		
Farmer	140 (5.2)	46 (4.7)	186 (5)		
Other	465 (17.2)	175 (17.7)	640 (17.3)		

XSL•FO
RenderX

Table 2. Distribution of depressive symptoms in 13- to 15-year-old adolescents.

Variables	Normal, n (%)	Depressive symptoms, n (%)	Total, n (%)	Chi-square (*df*)	*P* value
Sex				14.7 (1)	<.001
Male	1835 (54.2)	681 (48.1)	2516 (52.4)		
Female	1552 (45.8)	734 (51.9)	2286 (47.6)		
Ethnicity				98.1 (4)	<.001
Han	1766 (52.1)	575 (40.6)	2341 (48.8)		
Hani	440 (13)	210 (14.8)	650 (13.5)		
Yi	550 (16.2)	201 (14.2)	751 (15.6)		
Lisu	269 (7.9)	208 (14.7)	477 (10)		
Other	362 (10.7)	221 (15.6)	583 (12.1)		
The number of close friends				90.3 (2)	<.001
0-2	278 (8.2)	245 (17.3)	523 (10.9)		
3-4	516 (15.2)	229 (16.2)	745 (15.5)		
≥5	2593 (76.6)	941 (66.5)	3534 (73.6)		
Residence				0.3 (1)	.62
Rural	2712 (80)	1124 (79.4)	3836 (79.9)		
Urban	675 (19.9)	291 (20.6)	966 (20.1)		
The only child in the family				2.6 (1)	.11
Yes	455 (13.4)	215 (15.2)	670 (14)		
No	2932 (86.6)	1200 (84.8)	4132 (86)		
Family type				37.9 (3)	<.001
2-parent family	2794 (82.5)	1060 (74.9)	3854 (80.3)		
1-parent family	312 (9.2)	179 (12.7)	491 (10.2)		
Combined family	227 (6.7)	135 (9.5)	362 (7.5)		
Other	54 (1.6)	41 (2.9)	95 (2)		
Father’s education level				15.7 (4)	.004
Illiteracy	448 (13.2)	242 (17.1)	690 (14.4)		
Elementary school	831 (24.5)	361 (25.5)	1192 (24.8)		
Secondary school	1569 (46.3)	597 (42.2)	2166 (45.1)		
High school	387 (11.4)	161 (11.4)	548 (11.4)		
University	152 (4.5)	54 (3.8)	206 (4.3)		
Mother’s education level				24.9 (4)	<.001
Illiteracy	684 (20.2)	362 (25.6)	1046 (21.8)		
Elementary school	869 (25.7)	367 (25.9)	1236 (25.7)		
Secondary school	1422 (42)	504 (35.6)	1926 (40.1)		
High school	291 (8.6)	136 (9.6)	427 (8.9)		
University	121 (3.6)	46 (3.3)	167 (3.5)		
Father’s occupation				8.4 (5)	.14
Civil servant	1813 (53.5)	715 (50.5)	2528 (52.6)		
Worker	159 (4.7)	65 (4.6)	224 (4.7)		
Staff	709 (20.9)	309 (21.8)	1018 (21.2)		
Merchant	243 (7.2)	92 (6.5)	335 (7)		
Farmer	80 (2.4)	40 (2.8)	120 (2.5)		

XSL•FO
RenderX

Variables	Normal, n (%)	Depressive symptoms, n (%)	Total, n (%)	Chi-square (*df*)	*P* value
Other	383 (11.3)	194 (13.7)	577 (12)		
Mother's occupation				11.3 (5)	.045
Civil servant	2047 (60.4)	825 (58.3)	2872 (59.8)		
Worker	137 (4)	46 (3.3)	183 (3.8)		
Staff	403 (11.9)	172 (12.2)	575 (12)		
Merchant	268 (7.9)	104 (7.3)	372 (7.7)		
Farmer	99 (2.9)	39 (2.8)	138 (2.9)		
Other	433 (12.8)	229 (16.2)	662 (13.8)		

Cluster random sampling was used to conduct a questionnaire survey among adolescents in 11 counties of Yunnan Province, and 1-2 middle schools were selected from each county. SPSS software (version 23.0; IBM Corp) was used for random sampling. All students in grade 7 from the sampled middle schools were included in the survey.

The survey included demographic variables, eating behaviors, anxiety symptoms, and confounding factors that may influence depressive symptoms in adolescents, such as family changes, hospital experiences, screen time, smoking, drinking behavior, physical activity, family history of depression, learning stress, and the impact of COVID-19. The selection of these confounding factors was based on references and previous research findings of the research group.

A pilot test was conducted on 131 middle school students before the field survey to assess participant understanding of the questionnaire. The pilot test results showed that middle school students were able to fully understand the content of the questionnaire, and the average time taken to complete the questionnaire was 16 minutes. The adolescents gathered in a classroom to complete the survey. Trained investigators were on-site to answer questions from adolescents face-to-face. Adolescents completed the paper version of the questionnaire independently, which takes approximately 15-20 minutes.

Measures

Covariates

The demographic variables included sex, age, ethnicity, only child status, family residence, self-rated family economic status, family type, number of close friends, parental education, and parental occupation.

Depressive Symptom Assessment

The Children's Depression Inventory (CDI) was used to assess depressive symptoms in adolescents [35]. The CDI is applicable for the self-assessment of depressive symptoms in children and adolescents between the ages of 7 and 17 years; it is based on self-perception within the past 2 weeks. The CDI consists of 27 items, each of which is scored on a 3-point scale of "occasionally," "often," and "always" (0-2 points), respectively. The total CDI score ranges from 0 to 54; higher scores indicate more severe depressive symptoms. In this study, a total score of ≥19 was considered positive for depressive symptoms, and a score <19 was considered negative for depressive symptoms [35]. Research on Chinese children and adolescents aged 7-17 years shows that the CDI has good reliability and validity with a Cronbach α of 0.84 [36]. In this study, Cronbach α for the CDI was 0.853.

UPF Consumption Assessment

The eating frequency questionnaire prepared by our research group, which has good reliability and validity, was used to evaluate the eating behavior of adolescents [37]. According to the NOVA food classification system proposed by Monteiro [16], UPF refers to food that is separated from nature before being consumed or prepared as dishes, undergoes physical, biological, or chemical processes, and contains multiple ingredients and additives, for example, soft drinks, ice cream, packaged sweets, salty snacks, processed meat, and prepared frozen foods. This survey focused on the average weekly UPF consumption of adolescents in the previous month, including staple processed food (eg, fried bread stick, fried dough cake, 200 g per serving), processed meat (eg, ham, sausage, ham sausage, Chinese bacon, canned meat, beef jerky, chicken steak, 100 g per serving), processed eggs (eg, salted duck egg, spiced corned egg, 60 g per serving), preserved fruits (eg, preserved apple, 100 g per serving), instant foods (eg, instant noodles, 100 g per serving), salted snacks (refers to industrial production, packaging, and sales, as well as long-term preserved salty food; eg, spicy strip, salty crackers, 100 g per serving), sweet snacks (refers to industrial production, packaging, and sales, as well as long-term preserved sweet food; eg, chocolate, candy, 50 g per serving), SSB (eg, carbonated beverages, juice beverages, energy beverages, milk beverages [refers to a category of SSB made by blending milk with various additives, such as sugar, sweeteners, and acidifiers, eg, wang zai milk], 500 mL per bottle), other SSB (refers to SSB other than the sweet snacks, carbonated beverages, juice beverages, energy beverages, and milk beverages; eg, vitamin beverages, sports beverages, 500 mL per bottle). We used concise questions to assess UPF consumption frequency. For instance, how many bottles of Coca-Cola (500 mL per serving) did you drink per week in the last 1 month? There were 10 options: 0, 1, 2, 3, 4, 5, 6, 7, 8, and ≥9 times. Adolescents choose their consumption patterns based on their actual situations. The Cronbach α of the dietary frequency questionnaire in this study was 0.935.

Data Analysis

We used EpiData (version 3.0; EpiData Association) to build the database, and SPSS software (version 23.0; IBM Inc) and

XSL•FO
RenderX

R software (version 4.3.1; R Foundation for Statistical Computing) were used for statistical analysis. In the SPSS software, descriptive statistics, a chi-square test, and a generalized linear model (GLM) analysis were performed. Restricted cubic spline (RCS) analysis was performed using R software. A chi-square test was used to compare the difference in depressive symptoms detection rate among middle school students with different demographic characteristics. A GLM and RCS were used to analyze the association between UPF consumption and depressive symptoms. In the model, the UPF consumption frequency (independent variable) and depressive symptom score (dependent variable) were continuous variables. An association analysis was performed after adjusting for demographic and confounding variables that might have influenced depressive symptoms. The association between UPF consumption and depressive symptoms was analyzed using the RCS, and an RCS correlation plot was fitted. A GLM was used to analyze the association between the interaction term of sex, UPF consumption, and depressive symptoms. The test standard was α=.05.

Ethical Considerations

This study was approved by the Ethics Committee of the School of Medicine, Kunming University (approval 20210222), and was carried out in accordance with the Declaration of Helsinki. Ethical guidelines were followed during all the stages of this study. The survey was conducted anonymously with informed consent from the schools, guardians, and students. All participants joined the study voluntarily, and no compensation was provided to them. All data were kept confidential, deidentified, and anonymous.

Results

Distribution of Depressive Symptoms in Adolescents With Different Demographic Variables

The total detection rate of depressive symptoms was 28.3% (2402/8500) among middle school students, 25.5% (1069/4184) among males, and 30.9% (1333/4316) among females; the difference was statistically significant (χ^2_1=29.83; *P*<.001). The detection rate of depressive symptoms in 12-year-olds was 26.7% (987/3698), and that in 13- to 15-year-olds was 29.5% (1415/4802); the difference was statistically significant (χ^2_1=7.95; *P*=.005). The stratified analysis by age (12 years and 13-15 years) showed statistically significant differences in the detection rate of depressive symptoms among 12-year-olds for all demographic variables except parental education level and maternal occupation. Among adolescents aged 13-15 years, there were statistically significant differences in the detection rates of depressive symptoms for other demographic variables except for only child, residence, and parental occupation (*P*<.05).

Association of UPF Consumption With Depressive Symptoms in Adolescents

The consumption rates (frequency ≥1) of staple processed food, processed meat, processed eggs, preserved fruits, instant foods, salted snacks, sweet snacks, carbonated beverages, juice beverages, energy beverages, milk beverages and other SSB for adolescents are presented in Table 3.

Table 3. Consumption of ultraprocessed foods among adolescents.

Variable	Adolescents (N=8500), n (%)	Males (n=4184), n (%)	Females (n=4316), n (%)
Staple processed food	4712 (55.4)	2274 (54.3)	2438 (56.5)
Processed meat	5761 (67.8)	2870 (68.6)	2891 (67.0)
Processed eggs	3878 (45.6)	1943 (46.4)	1935 (44.8)
Preserved fruits	4387 (51.6)	2037 (48.7)	2350 (54.4)
Instant foods	5234 (61.6)	2648 (63.3)	2586 (59.9)
Salted snacks	5702 (67.1)	2759 (65.9)	2943 (68.2)
Sweet snacks	5271 (62.0)	2490 (59.5)	2781 (64.4)
Carbonated beverages	6022 (70.8)	3187 (76.2)	2835 (65.7)
Juice beverages	5612 (66.0)	2687 (64.2)	2925 (67.8)
Energy beverages	3716 (43.7)	2109 (50.4)	1607 (37.2)
Milk beverages	6025 (70.9)	2732 (65.3)	3293 (76.3)
Other SSB[a]	3739 (44.0)	1899 (45.4)	1840 (42.6)

[a]SSB: sugar-sweetened beverages.

Table 4 presents the results of the GLM analysis of the association between UPF consumption and depressive symptoms in adolescents after adjusting for variables (Model 2). Table S1 in Multimedia Appendix 1 shows the analysis results without adjusting for variables (Model 1). After adjusting for confounding variables, instant food (β=.22, 95% CI 0.13-0.32; *P*<.001), carbonated beverage (β=.20, 95% CI 0.12-0.29; *P*<.001), energy beverage (β=.21, 95% CI 0.10-0.32; *P*<.001), and other SSB (β=.24, 95% CI 0.14-0.34; *P*<.001) consumption were associated with depressive symptoms for all participants. The sex-stratified analysis showed that instant food (β=.15, 95% CI 0.03-0.27; *P*=.01), carbonated beverage (β=.15, 95% CI 0.04-0.26; *P*=.006), energy beverage (β=.22, 95% CI 0.08-0.36; *P*=.002), milk beverage (β=–0.14, 95% CI –0.26 to

XSL•FO
RenderX

–0.02; *P*=.03), and other SSB (β=.21, 95% CI 0.08-0.35; *P*=.002) consumption was associated with depressive symptoms in males. In addition, instant food (β=.29, 95% CI 0.15-0.43; *P*<.001), carbonated beverage (β=.29, 95% CI 0.16-0.43; *P*<.001), energy beverage (β=.22, 95% CI 0.04-0.40; *P*=.01), and other SSB (β=.25, 95% CI 0.10-0.41; *P*=.001) consumption were associated with depressive symptoms in females.

XSL•FO
RenderX

Table 4. Association of ultraprocessed food consumption with depressive symptoms in adolescents. Adjusted for age, ethnicity, residence, sex, the only child in the family, parental occupation, family type, parental educational level, the number of close friends, self-perceived academic stress, self-perceived socioeconomic status, family changes, hospital experience, smoking, alcohol consumption, family history of depression, left-behind experience, physical activity, and the impact of COVID-19. The analysis of all participants also adjusted for sex.

Variable	β (95% CI)	SE	Wald chi-square (*df*)	*P* value
All participants				
Staple processed food	–.08 (–0.17 to 0.00)	0.04	3.7 (1)	.06
Processed meat	–.04 (–0.12 to 0.04)	0.04	1.0 (1)	.33
Processed eggs	–.04 (–0.14 to 0.05)	0.05	0.8 (1)	.38
Preserved fruits	–.04 (–0.14 to 0.05)	0.05	0.8 (1)	.37
Instant foods	.22 (0.13 to 0.32)	0.05	23.3 (1)	<.001
Salted snacks	–.03 (–0.10 to 0.05)	0.04	0.5 (1)	.50
Sweet snacks	–.04 (–0.12 to 0.04)	0.04	1.0 (1)	.31
Carbonated beverages	.20 (0.12 to 0.29)	0.04	23.0 (1)	<.001
Juice beverages	–.06 (–0.15 to 0.03)	0.05	1.6 (1)	.21
Energy beverages	.21 (0.10 to 0.32)	0.06	14.7 (1)	<.001
Milk beverages	–.07 (–0.15 to 0.02)	0.04	2.3 (1)	.13
Other SSB[a]	.24 (0.14 to 0.34)	0.05	20.8 (1)	<.001
Males				
Staple processed food	–.11 (–0.23 to 0.00)	0.06	3.8 (1)	.05
Processed meat	–.03 (–0.13 to 0.08)	0.05	0.3 (1)	.59
Processed eggs	.01 (–0.11 to 0.14)	0.06	0.0 (1)	.84
Preserved fruits	–.03 (–0.17 to 0.10)	0.07	0.2 (1)	.62
Instant foods	.15 (0.03 to 0.27)	0.06	6.1 (1)	.01
Salted snacks	.01 (–0.10 to 0.12)	0.05	0.0 (1)	.85
Sweet snacks	.00 (–0.12 to 0.11)	0.06	0.0 (1)	.94
Carbonated beverages	.15 (0.04 to 0.26)	0.05	7.6 (1)	.006
Juice beverages	–.02 (–0.14 to 0.10)	0.06	0.1 (1)	.76
Energy beverages	.22 (0.08 to 0.36)	0.07	9.9 (1)	.002
Milk beverages	–.14 (–0.26 to –0.02)	0.06	4.9 (1)	.03
Other SSB	.21 (0.08 to 0.35)	0.07	9.8 (1)	.002
Females				
Staple processed food	–.04 (–0.17 to 0.09)	0.07	0.4 (1)	.54
Processed meat	–.04 (–0.17 to 0.09)	0.06	0.4 (1)	.55
Processed eggs	–.09 (–0.24 to 0.05)	0.07	1.6 (1)	.20
Preserved fruits	–.06 (–0.21 to 0.08)	0.07	0.7 (1)	.40
Instant foods	.29 (0.15 to 0.43)	0.07	16.6 (1)	<.001
Salted snacks	–.07 (–0.19 to 0.04)	0.06	1.6 (1)	.20
Sweet snacks	–.09 (–0.20 to 0.02)	0.06	2.6 (1)	.11
Carbonated beverages	.29 (0.16 to 0.43)	0.07	18.9 (1)	<.001
Juice beverages	–.12 (–0.26 to 0.02)	0.07	2.8 (1)	.10
Energy beverages	.22 (0.04 to 0.40)	0.09	6.0 (1)	.01
Milk beverages	.00 (–0.13 to 0.12)	0.06	0.0 (1)	.96
Other SSB	.25 (0.10 to 0.41)	0.08	10.1 (1)	.001

[a]SSB: sugar-sweetened beverages.

XSL•FO
RenderX

Tables 5 and 6 present the association between the interaction of sex, UPF consumption, and depressive symptoms in adolescents. The results showed that female×instant foods (β=.31, 95% CI 0.18-0.44; *P*<.001), female×carbonated beverages (β=.27, 95% CI 0.15-0.39; *P*<.001), female×energy beverages (β=.18, 95% CI 0.01-0.35; *P*=.04), and female×other SSB (β=.17, 95% CI 0.02-0.32; *P*=.03) were associated with depressive symptoms, respectively. However, no association was observed between the interaction of these UPFs and male and depressive symptoms.

Table 5. Association between interaction of sex, ultraprocessed food consumption, and depressive symptoms in adolescents (Model 1: unadjusted for variables).

Variable	β (95% CI)	SE	Wald chi-square (*df*)	*P* value
Male×instant foods	.14 (0.01 to 0.26)	0.06	4.7 (1)	.03
Female×instant foods	.44 (0.30 to 0.58)	0.07	38.6 (1)	<.001
Male×carbonated beverages	.09 (–0.02 to 0.19)	0.05	2.5 (1)	.12
Female×carbonated beverages	.31 (0.18 to 0.44)	0.07	21.7 (1)	<.001
Male×energy beverages	.12 (–0.03 to 0.27)	0.08	2.4 (1)	.12
Female×energy beverages	.19 (0.00 to 0.37)	0.09	3.8 (1)	.05
Male×other SSB[a]	.15 (0.01 to 0.30)	0.07	4.4 (1)	.04
Female×other SSB	.18 (0.02 to 0.34)	0.08	4.6 (1)	.03

[b]SSB: sugar-sweetened beverages.

Table 6. Association between interaction of sex, ultraprocessed food consumption, and depressive symptoms in adolescents (Model 2: adjusted for age, ethnicity, residence, the only child in the family, parental occupation, family type, parental educational level, the number of close friends, self-perceived academic stress, self-perceived socioeconomic status, family changes, hospital experience, smoking, alcohol consumption, family history of depression, left-behind experience, physical activity, and the impact of COVID-19).

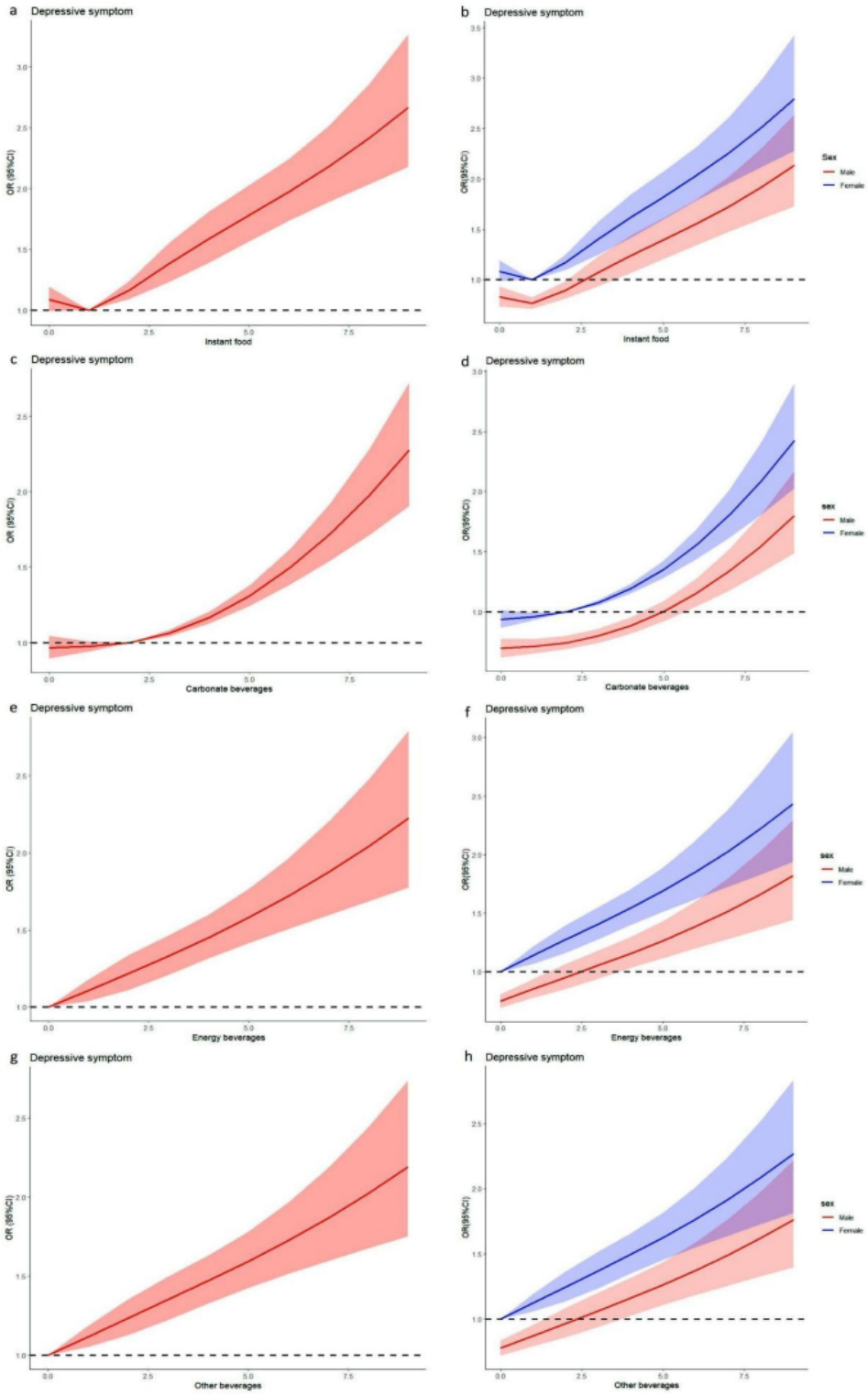
Variable	β (95% CI)	SE	Wald chi-square (*df*)	*P* value
Male×instant foods	.11 (–0.01 to 0.22)	0.06	3.3 (1)	.07
Female×instant foods	.31 (0.18 to 0.44)	0.07	21.6 (1)	<.001
Male×carbonated beverages	.06 (–0.05 to 0.15)	0.05	1.2 (1)	.28
Female×carbonated beverages	.27 (0.15 to 0.39)	0.06	19.4 (1)	<.001
Male×energy beverages	.12 (–0.02 to 0.25)	0.07	2.7 (1)	.10
Female×energy beverages	.18 (0.01 to 0.35)	0.09	4.1 (1)	.04
Male×other SSB[a]	.12 (–0.01 to 0.25)	0.07	3.1 (1)	.08
Female×other SSB	.17 (0.02 to 0.32)	0.08	5.0 (1)	.03

[a]SSB: sugar-sweetened beverages.

Figure 2 displays the results of the RCS analysis for the association between UPF consumption and depressive symptoms in adolescents. Overall, UPF consumption was significantly associated with depressive symptoms in adolescents, and females were more sensitive to UPF consumption (exposure) than males.

XSL•FO
RenderX

Figure 2. Association of ultraprocessed food consumption with depressive symptoms in adolescents. OR: odds ratio.



Discussion

Principal Findings

In this study, the detection rate for depressive symptoms among adolescents of multiple ethnic groups in Yunnan was higher (2402/8500, 28.3%), which is consistent with the results of our previous study on adolescents in 4 cities in China (27.3%) [13]. This detection rate was higher than that in a meta-analysis of middle school students in mainland China (24.3%) and that of middle school students in eastern China (23.3%) [38,39]; however, it was lower than that for Bangladeshi adolescents (36.6%) [40] and Indian adolescents (52.3%) [10]. Our data showed the detection rate of depressive symptoms for adolescents aged 13-15 years (1415/4802, 29.5%) was higher than that for adolescents aged 12 years (987/3698, 26.7%). A school-based survey in the United Kingdom determined that the detection rate of depressive symptoms among 15- to

XSL•FO
RenderX

16-year-olds (32%) was higher than that among 11- to 12-year-olds (24%) [9]. In addition, we found the detection rate of depressive symptoms in females (1333/4316, 30.9%) was significantly higher than that for males (1069/4184, 25.5%), which is consistent with the conclusions of other studies [40]. For example, a Bangladeshi study reported a significantly higher prevalence of depressive symptoms among female adolescents than male adolescents (32.5% vs 25.3%) [40]. The higher prevalence of depressive symptoms among female adolescents is consistent worldwide. Numerous studies have suggested that depressive symptoms increase during adolescence, adolescent females are more susceptible to depressive symptoms, and sex differences appear only during adolescence [41]. A meta-analysis found that sex differences peaked in adolescence (OR 3.02, 95% CI 2.76-3.30 for ages 13-15 years), then declined and remained stable in adulthood [42]. Moreover, the longitudinal study of Australian children also found that between the ages of 4 and 14 years, females were more likely than males to be on a trajectory for increased depressive symptoms [43]. Therefore, schools and families need to pay more attention to depressive symptoms in adolescents, especially female adolescents.

Our study suggests that fast foods, carbonated beverages, energy beverages, and other SSB (except those listed in this study) were positively associated with depressive symptoms in both males and females. Other studies have also determined that UPF consumption is associated with depressive symptoms [25,29,31]. Our survey of Chinese college freshmen demonstrated that carbonated beverage and milk tea consumption were associated with depressive symptoms in males, whereas carbonated beverages, tea beverages, and milk tea consumption were associated with depressive symptoms in females [44]. Higher UPF consumption in Spanish adolescents aged 14-17 years was associated with a higher detection rate of depressive symptoms, and the higher the intake, the more pronounced the depressive symptoms [22]. The consumption of sweets, biscuits, snacks, and fast food among adolescents in England, Wales, and Northern Ireland was significantly positively associated with depressive symptoms [30]. Instant noodles, instant rice, and other instant foods are very popular among Chinese adolescents because they are easy to carry and eat. Our survey indicated that the consumption rate of instant foods such as instant noodles was as high as 61.6% (5234/8500), which has not been previously reported in China. Our previous research found an association between carbonated beverage and processed snack consumption and anxiety symptoms in Yunnan college students [45]. This was the first study to report a correlation between instant food consumption and depressive symptoms among Chinese multiethnic adolescents.

Furthermore, the data from this study showed that milk beverage consumption was negatively associated with depressive symptoms in males. The milk beverages in this study were also SSB; thus, we will further observe and evaluate these associations in a cohort study. The RCS analysis revealed that instant food, carbonated beverages, energy beverages, and other SSB were associated with depressive symptoms in males and females, and the association between UPF consumption and depressive symptoms in females was more sensitive. We further analyzed the interaction between sex and UPF and its association with depressive symptoms. The results indicated that females who interact with instant foods, carbonated beverages, energy beverages, and other SSBs were associated with depressive symptoms, respectively. However, no association was observed between the interaction of males and these 4 types of UPFs and depressive symptoms. The research data further support the view that females who consumed UPF were more likely to experience depressive symptoms. Therefore, when schools develop intervention measures to reduce depression symptoms among adolescents through dietary behavior intervention, they need to take into account sex-specific factors. For female adolescents, it is even more necessary to strengthen health education and behavioral intervention.

Several mechanisms by which UPF consumption is associated with depressive symptoms are possible. First, UPF consumption is associated with sleep dissatisfaction in adolescents [28]. The low sleep quality is a predictor of depressive symptoms in adolescents. Previous studies have suggested that the fourth quartile of UPF (44.3%-81.8% of total calories; prevalence ratio=1.14) was associated with a higher prevalence of poor sleep quality in adolescents [46]. A systematic review and meta-analysis showed that the higher the intake of UPF, the greater the risk of insomnia among adolescents (OR 5, 95% CI 1.21-1.99), and there is a clear dose-response relationship [47]. Second, UPF consumption may lead to disturbances in the intestinal flora and increase stress susceptibility of adolescents through the hypothalamic-pituitary-adrenal axis, which can lead to depressive symptoms occurring [5]. Experimental studies have shown that excessive intake of UPF leads to a reduced abundance of potentially beneficial butyrogenic microbes (*Roseburia*, *Faecalbacterium prausnitzii*, *Lactobacillus*, *Bifidobacterium*, and *Lachnospiraceae*), and the abundance of potentially pathogenic or opportunistic microorganisms (*Alistiopsis and Pasteurella*) increases [48]. Finally, the intake of UPF leads to an increase in chronic inflammatory levels, which causes the occurrence of depressive symptoms [49]. Adolescents with UPF energy consumption ≥30% (tertile 3 of UPF) had a 79% increase in interleukin-8 levels and had higher C-reactive protein levels when compared with adolescents in tertile 1 of UPF [50]. UPF intake may lead to a dysregulation of immune cell homeostasis and inflammation, including decreased levels of anti-inflammatory cytokines and increased levels of proinflammatory cytokines and chemokines [51].

Adolescents have great potential for physical, cognitive, and social development, and adolescence is regarded as the second critical window after infancy [52]. The mental health and growth of adolescents are related to the future of a country and require extensive attention. Adolescent depression is a global concern. Adolescence is a period of rapid emotional and cognitive development and critical life transitions. Notably, the treatment of depression is expensive. A Netherlands study of 11- to 18-year-olds with subclinical depressive symptoms showed that the social costs associated with depressive symptoms amounted to €54 million (US $57 million) per year [53]. Therefore, prevention and early intervention of adolescent depression is a top priority, using strategies that target predisposing factors, causes, and depression symptoms. Modifiable risk factors are

XSL•FO
RenderX

critical for reducing depression prevalence among adolescents. Eating behavior interventions have been recommended as an effective measure to prevent mental illness [31]. Therefore, it may be necessary to expand the monitoring of UPF consumption behaviors and depressive symptoms in adolescents, strengthen health education in schools and families, and promote healthy eating environments in schools to reduce UPF consumption and improve depressive symptoms in adolescents.

Strengths and Limitations

This study has several strengths. First, the sample comprised adolescents from 11 counties in Yunnan Province, China, with good representativeness and extrapolation of research results. Second, this study was the first to report the association between UPF consumption and depressive symptoms in Chinese multiethnic adolescents. Finally, we used a variety of models to analyze the association between UPF consumption and depressive symptoms and obtained stable results.

This study also has a few limitations. First, information bias was inevitable in the investigation, such as recall and social expectation biases. Second, the current assessment of UPF intake relies on frequency data without quantification in terms of grams or caloric contribution, which is limited for an accurate analysis of the association between the precise intake of UPF and depressive symptoms in adolescents. Finally, cross-sectional data cannot determine causal associations, and bidirectional associations may need to be clarified in further cohort studies.

Conclusion

This study found that the consumption of instant foods, carbonated beverages, and energy beverages was related to depressive symptoms among Chinese multiethnic adolescents. It is important to note that the observed association may have a reverse correlation, where depressive symptoms could lead to an increase in the intake of ultrapalmitic fatty acids among adolescents. This needs to be further confirmed in cohort studies. The results highlighted the importance of reducing UPF consumption among adolescents. These findings should be replicated in cohort studies to determine the causal relationship and the direction. The government, schools, and families should take effective measures to reduce UPF consumption among Chinese multiethnic adolescents to support mental health development in adolescence.

Acknowledgments

We thank the schools that took part in the survey and all the participants. This work was supported by the National Natural Science Foundation of China (82160622), the Young Talents Special Project of the “Xingdian Talent Support Plan” in Yunnan Province (XDYC-QNRC-2022-0289), the Yunnan Fundamental Research Projects (202501AT070481), the Scientific Research Foundation of Yunnan Provincial Department of Education (2025Y1087), and the Introduced Talents Scientific Research Project of Kunming University (YJL2103).

Data Availability

Data are available upon reasonable request.

Authors' Contributions

Conceptualization: HX
Data curation: HX, JY, DH, XL, FZ
Formal analysis: HX, JY, DH
Methodology: HX
Writing–original draft: HX
Writing–review & editing: JY, DH, XL, FZ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Results of the generalized linear model analysis of the association between ultraprocessed food consumption and depressive symptoms in adolescents without adjusting for variables.
[DOCX File , 21 KB - publichealth_v11i1e75061_app1.docx]

References

1. Malhi GS, Mann JJ. Depression. Lancet 2018;392(10161):2299-2312. [doi: 10.1016/S0140-6736(18)31948-2] [Medline: 30396512]
2. Liu J, Liu Y, Ma W, Tong Y, Zheng J. Temporal and spatial trend analysis of all-cause depression burden based on global burden of disease (GBD) 2019 study. Sci Rep 2024;14(1):12346 [FREE Full text] [doi: 10.1038/s41598-024-62381-9] [Medline: 38811645]

XSL•FO
RenderX

3. Senior M, Pierce M, Taxiarchi VP, Garg S, Edge D, Newlove-Delgado T, et al. 5-year mental health outcomes for children and adolescents presenting with psychiatric symptoms to general practitioners in England: a retrospective cohort study. Lancet Psychiatry 2024;11(4):274-284 [FREE Full text] [doi: 10.1016/S2215-0366(24)00038-5] [Medline: 38490760]
4. Eray, Tekcan D, Baran Y. More anxious or more shy? Examining the social anxiety levels of adolescents with primary enuresis nocturna: a controlled study. J Pediatr Urol 2019;15(4):343.e1-343.e5. [doi: 10.1016/j.jpurol.2019.04.002] [Medline: 31036479]
5. Nota MHC, Nicolas S, O'Leary OF, Nolan YM. Outrunning a bad diet: interactions between exercise and a Western-style diet for adolescent mental health, metabolism and microbes. Neurosci Biobehav Rev 2023;149:105147. [doi: 10.1016/j.neubiorev.2023.105147] [Medline: 36990371]
6. Reangsing C, Punsuwun S, Schneider JK. Effects of mindfulness interventions on depressive symptoms in adolescents: a meta-analysis. Int J Nurs Stud 2021;115:103848. [doi: 10.1016/j.ijnurstu.2020.103848] [Medline: 33383273]
7. Shorey S, Ng ED, Wong CHJ. Global prevalence of depression and elevated depressive symptoms among adolescents: a systematic review and meta-analysis. Br J Clin Psychol 2022;61(2):287-305. [doi: 10.1111/bjc.12333] [Medline: 34569066]
8. Wartberg L, Kriston L, Thomasius R. Depressive symptoms in adolescents. Dtsch Arztebl Int 2018;115(33-34):549-555 [FREE Full text] [doi: 10.3238/arztebl.2018.0549] [Medline: 30189974]
9. Randhawa SS, Marcenes W, Stansfeld SA, Bernabé E. Depressive symptoms and traumatic dental injuries among adolescents. Dent Traumatol 2020;36(2):192-197. [doi: 10.1111/edt.12524] [Medline: 31667970]
10. Gupta S, Das S, Das M, Banerjee S, Neogi R, Mukherjee S. Prevalence and correlates of depression, anxiety, and stress among high school students in a block of Hooghly district, West Bengal: across-sectional study. J Educ Health Promot 2023;12:345 [FREE Full text] [doi: 10.4103/jehp.jehp_1769_22] [Medline: 38144014]
11. Ang AL, Wahab S, Abd Rahman FN, Hazmi H, Md Yusoff R. Depressive symptoms in adolescents in Kuching, Malaysia: prevalence and associated factors. Pediatr Int 2019;61(4):404-410. [doi: 10.1111/ped.13778] [Medline: 30597707]
12. Rao W, Xu D, Cao X, Wen S, Che W, Ng CH, et al. Prevalence of depressive symptoms in children and adolescents in China: a meta-analysis of observational studies. Psychiatry Res 2019;272:790-796. [doi: 10.1016/j.psychres.2018.12.133] [Medline: 30832200]
13. Xu H, Guo J, Wan Y, Zhang S, Yang R, Xu H, et al. Association between screen time, fast foods, sugar-sweetened beverages and depressive symptoms in Chinese adolescents. Front Psychiatry 2020;11:458 [FREE Full text] [doi: 10.3389/fpsyt.2020.00458] [Medline: 32528328]
14. Steare T, Lewis G, Lange K, Lewis G. The association between academic achievement goals and adolescent depressive symptoms: a prospective cohort study in Australia. Lancet Child Adolesc Health 2024;8(6):413-421 [FREE Full text] [doi: 10.1016/S2352-4642(24)00051-8] [Medline: 38642573]
15. Zahedi H, Djalalinia S, Sadeghi O, Zare Garizi F, Asayesh H, Payab M, et al. Breakfast consumption and mental health: a systematic review and meta-analysis of observational studies. Nutr Neurosci 2022;25(6):1250-1264. [doi: 10.1080/1028415X.2020.1853411] [Medline: 33314992]
16. Monteiro CA, Levy RB, Claro RM, Castro IRRD, Cannon G. A new classification of foods based on the extent and purpose of their processing. Cad Saude Publica 2010;26(11):2039-2049 [FREE Full text] [doi: 10.1590/s0102-311x2010001100005] [Medline: 21180977]
17. De Amicis R, Mambrini SP, Pellizzari M, Foppiani A, Bertoli S, Battezzati A, et al. Ultra-processed foods and obesity and adiposity parameters among children and adolescents: a systematic review. Eur J Nutr 2022;61(5):2297-2311 [FREE Full text] [doi: 10.1007/s00394-022-02873-4] [Medline: 35322333]
18. Zhang Z, Jackson SL, Martinez E, Gillespie C, Yang Q. Association between ultraprocessed food intake and cardiovascular health in US adults: a cross-sectional analysis of the NHANES 2011-2016. Am J Clin Nutr 2021;113(2):428-436 [FREE Full text] [doi: 10.1093/ajcn/nqaa276] [Medline: 33021623]
19. Parnham JC, Chang K, Rauber F, Levy RB, Millett C, Laverty AA, et al. The ultra-processed food content of school meals and packed lunches in the United Kingdom. Nutrients 2022;14(14):2961 [FREE Full text] [doi: 10.3390/nu14142961] [Medline: 35889918]
20. Mititelu M, Oancea C, Neac u SM, Musuc AM, Gheonea TC, Stanciu TI, et al. Evaluation of junk food consumption and the risk related to consumer health among the Romanian population. Nutrients 2023;15(16):3591 [FREE Full text] [doi: 10.3390/nu15163591] [Medline: 37630781]
21. Shawon MSR, Jahan E, Rouf RR, Hossain FB. Psychological distress and unhealthy dietary behaviours among adolescents aged 12-15 years in nine South-East Asian countries: a secondary analysis of the global school-based health survey data. Br J Nutr 2023;129(7):1242-1251. [doi: 10.1017/S0007114522002306] [Medline: 35856273]
22. Reales-Moreno M, Tonini P, Escorihuela RM, Solanas M, Fernández-Barrés S, Romaguera D, et al. Ultra-processed foods and drinks consumption is associated with psychosocial functioning in adolescents. Nutrients 2022;14(22):4831 [FREE Full text] [doi: 10.3390/nu14224831] [Medline: 36432518]
23. Moore JB, Jee S, Kemper BJ, Maddock JE, Li R. Association between social isolation and eating alone on foods consumed by Chinese adolescents. J Healthy Eat Act Living 2021;1(2):74-83 [FREE Full text] [Medline: 37789910]

XSL•FO
RenderX

24. Voráčová J, Badura P, Hamrik Z, Holubčíková J, Sigmund E. Unhealthy eating habits and participation in organized leisure-time activities in Czech adolescents. Eur J Pediatr 2018;177(10):1505-1513. [doi: 10.1007/s00431-018-3206-y] [Medline: 30014303]
25. Lane MM, Gamage E, Travica N, Dissanayaka T, Ashtree DN, Gauci S, et al. Ultra-processed food consumption and mental health: a systematic review and meta-analysis of observational studies. Nutrients 2022;14(13):2568 [FREE Full text] [doi: 10.3390/nu14132568] [Medline: 35807749]
26. Mazloomi SN, Talebi S, Mehrabani S, Bagheri R, Ghavami A, Zarpoosh M, et al. The association of ultra-processed food consumption with adult mental health disorders: a systematic review and dose-response meta-analysis of 260,385 participants. Nutr Neurosci 2023;26(10):913-931. [doi: 10.1080/1028415X.2022.2110188] [Medline: 36094005]
27. Chopra C, Mandalika S, Kinger N. Does diet play a role in the prevention and management of depression among adolescents? A narrative review. Nutr Health 2021;27(2):243-263. [doi: 10.1177/0260106020980532] [Medline: 33416032]
28. Malmir H, Mahdavi FS, Ejtahed H, Kazemian E, Chaharrahi A, Mohammadian Khonsari N, et al. Junk food consumption and psychological distress in children and adolescents: a systematic review and meta-analysis. Nutr Neurosci 2023;26(9):807-827. [doi: 10.1080/1028415X.2022.2094856] [Medline: 35816403]
29. Qian K, Wang D, Sun Y, Ye X, Wang D, Xu H. Association of dietary patterns with depressive symptoms in Chinese adolescents: a cross-sectional study. Front Nutr 2023;10:1180858 [FREE Full text] [doi: 10.3389/fnut.2023.1180858] [Medline: 37497054]
30. El Ansari W, Adetunji H, Oskrochi R. Food and mental health: relationship between food and perceived stress and depressive symptoms among university students in the United Kingdom. Cent Eur J Public Health 2014;22(2):90-97 [FREE Full text] [doi: 10.21101/cejph.a3941] [Medline: 25230537]
31. Coletro HN, Mendonça RDD, Meireles AL, Machado-Coelho GLL, Menezes MCD. Ultra-processed and fresh food consumption and symptoms of anxiety and depression during the COVID-19 pandemic: COVID inconfidentes. Clin Nutr ESPEN 2022;47:206-214 [FREE Full text] [doi: 10.1016/j.clnesp.2021.12.013] [Medline: 35063203]
32. Faisal-Cury A, Leite MA, Escuder MML, Levy RB, Peres MFT. The relationship between ultra-processed food consumption and internalising symptoms among adolescents from São Paulo city, Southeast Brazil. Public Health Nutr 2022;25(9):2498-2506 [FREE Full text] [doi: 10.1017/S1368980021004195] [Medline: 34610854]
33. Mesas AE, González AD, de Andrade SM, Martínez-Vizcaíno V, López-Gil JF, Jiménez-López E. Increased consumption of ultra-processed food is associated with poor mental health in a nationally representative sample of adolescent students in Brazil. Nutrients 2022;14(24):5207 [FREE Full text] [doi: 10.3390/nu14245207] [Medline: 36558369]
34. Xu H, Zhao Y, Tan R, Li M, Yu C, Rui D, et al. Association of consumption of sugar-sweetened beverages with elevated blood pressure among college students in Yunnan Province, China. Public Health Nutr 2024;27(1):e85 [FREE Full text] [doi: 10.1017/S1368980024000569] [Medline: 38418286]
35. Timbremont B, Braet C, Dreessen L. Assessing depression in youth: relation between the children's depression inventory and a structured interview. J Clin Child Adolesc Psychol 2004;33(1):149-157. [doi: 10.1207/S15374424JCCP3301_14] [Medline: 15028549]
36. Wang L, Feng Z, Yang G, Yang Y, Wang K, Dai Q, et al. Depressive symptoms among children and adolescents in western china: an epidemiological survey of prevalence and correlates. Psychiatry Res 2016;246:267-274. [doi: 10.1016/j.psychres.2016.09.050] [Medline: 27741479]
37. Xu H, Zhang M, Yang C, Fu Y, Yang J, Yang W, et al. Fast food and anxiety symptoms of Chinese multi-ethnic adolescents: the moderating role of regulatory emotional self-efficacy. J Affect Disord 2025;374:160-170. [doi: 10.1016/j.jad.2025.01.036] [Medline: 39800067]
38. Tang X, Tang S, Ren Z, Wong DFK. Prevalence of depressive symptoms among adolescents in secondary school in mainland China: a systematic review and meta-analysis. J Affect Disord 2019;245:498-507. [doi: 10.1016/j.jad.2018.11.043] [Medline: 30439677]
39. Tan Z, Zhang Z, Zhang Y. Establishment and validation of depressive symptom predictive model in middle school studentsJ. Chin J Sch Health 2024;45(7):998-1002. [doi: 10.16835/j.cnki.1000-9817.2024219]
40. Anjum A, Hossain S, Sikder T, Uddin ME, Rahim DA. Investigating the prevalence of and factors associated with depressive symptoms among urban and semi-urban school adolescents in Bangladesh: a pilot study. Int Health 2022;14(4):354-362 [FREE Full text] [doi: 10.1093/inthealth/ihz092] [Medline: 31693088]
41. Serio B, Kohler R, Ye F, Lichenstein SD, Yip SW. A multidimensional approach to understanding the emergence of sex differences in internalizing symptoms in adolescence. Dev Cogn Neurosci 2022;58:101182 [FREE Full text] [doi: 10.1016/j.dcn.2022.101182] [Medline: 36495789]
42. Salk RH, Hyde JS, Abramson LY. Gender differences in depression in representative national samples: meta-analyses of diagnoses and symptoms. Psychol Bull 2017;143(8):783-822 [FREE Full text] [doi: 10.1037/bul0000102] [Medline: 28447828]
43. Lewis AJ, Sae-Koew JH, Toumbourou JW, Rowland B. Gender differences in trajectories of depressive symptoms across childhood and adolescence: a multi-group growth mixture model. J Affect Disord 2020;260:463-472. [doi: 10.1016/j.jad.2019.09.027] [Medline: 31539681]

XSL•FO
RenderX

44. Xu H, Yang Z, Liu D, Yu C, Zhao Y, Yang J, et al. Mediating effect of physical sub-health in the association of sugar-sweetened beverages consumption with depressive symptoms in Chinese college students: a structural equation model. J Affect Disord 2023;342:157-165. [doi: 10.1016/j.jad.2023.09.020] [Medline: 37730148]
45. Jiang Y, Su Y, Su Y. Longitudinal association between processed food consumption and anxiety symptoms among college students in Yunnan Province. Chin J Sch Health 2024;45(2):178-182. [doi: 10.16835/j.cnki.1000-9817.2024051]
46. Sousa RDS, Bragança MLBM, Oliveira BRD, Coelho CCNDS, Silva AAMD. Association between the degree of processing of consumed foods and sleep quality in adolescents. Nutrients 2020;12(2):462 [FREE Full text] [doi: 10.3390/nu12020462] [Medline: 32059416]
47. Pourmotabbed A, Awlqadr FH, Mehrabani S, Babaei A, Wong A, Ghoreishy SM, et al. Ultra-processed food intake and risk of insomnia: a systematic review and meta-analysis. Nutrients 2024;16(21):3767 [FREE Full text] [doi: 10.3390/nu16213767] [Medline: 39519600]
48. Lane M, Howland G, West M, Hockey M, Marx W, Loughman A, et al. The effect of ultra-processed very low-energy diets on gut microbiota and metabolic outcomes in individuals with obesity: a systematic literature review. Obes Res Clin Pract 2020;14(3):197-204. [doi: 10.1016/j.orcp.2020.04.006] [Medline: 32546361]
49. Tristan Asensi M, Napoletano A, Sofi F, Dinu M. Low-grade inflammation and ultra-processed foods consumption: a review. Nutrients 2023;15(6):1546 [FREE Full text] [doi: 10.3390/nu15061546] [Medline: 36986276]
50. Martins GMDS, França AKTC, Viola PCDAF, Carvalho CAD, Marques KDS, Santos AMD, et al. Intake of ultra-processed foods is associated with inflammatory markers in Brazilian adolescents. Public Health Nutr 2022;25(3):591-599 [FREE Full text] [doi: 10.1017/S1368980021004523] [Medline: 34726140]
51. Maki KA, Sack MN, Hall KD. Ultra-processed foods: increasing the risk of inflammation and immune dysregulation? Nat Rev Immunol 2024;24(7):453-454. [doi: 10.1038/s41577-024-01049-x] [Medline: 38831164]
52. Chen T, Dong B, Dong Y, Li J, Ma Y, Liu D, et al. Matching actions to needs: shifting policy responses to the changing health needs of Chinese children and adolescents. Lancet 2024;403(10438):1808-1820. [doi: 10.1016/S0140-6736(23)02894-5] [Medline: 38643776]
53. Bodden DHM, van den Heuvel MWH, Engels RCME, Dirksen CD. Societal costs of subclinical depressive symptoms in Dutch adolescents: a cost-of-illness study. J Child Psychol Psychiatry 2022;63(7):771-780 [FREE Full text] [doi: 10.1111/jcpp.13517] [Medline: 34496447]

Abbreviations

CDI: Children's Depression Inventory
GLM: generalized linear model
OR: odds ratio
RCS: restricted cubic spline
SSB: sugar-sweetened beverages
UPF: ultraprocessed food

Edited by A Mavragani, T Sanchez; submitted 30.03.25; peer-reviewed by A Ivaturi, B Najand; comments to author 08.06.25; revised version received 30.07.25; accepted 28.08.25; published 20.10.25.

Please cite as:
Xu H, Yang J, Hu D, Li X, Zhou F
Ultraprocessed Foods Associated With Depressive Symptoms Among Chinese Adolescents: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e75061
URL: https://publichealth.jmir.org/2025/1/e75061
doi:10.2196/75061
PMID:

©Honglv Xu, Jieru Yang, Dongyue Hu, Xiaoxiao Li, Feihui Zhou. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations Between Both HIV and Metabolic Comorbidity and Self-Reported Mpox Among Men Who Have Sex With Men: Multicenter Cross-Sectional Study

Yujie Liu[1], PhD; Xin Ge[1], MD; Yinghuan Zhang[1,2], MD; Yuxuan Wang[1,2], MD; Meihui Zhang[1,2], MD; Jianyu Chen[3], PhD; Gang Xu[2], PhD; Jiechen Zhang[4*], PhD; Ying Wang[1*], PhD; Yong Cai[1,5*], PhD

[1]Public Health Research Center, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]Shanghai Jiao Tong University School of Public Health, Shanghai, China

[3]College of Public Health, Shanghai University of Medicine & Health Sciences, Shanghai, China

[4]Dermatology Department, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[5]Center for Community Health Care, China Hospital Development Institute, Shanghai Jiao Tong University, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Yong Cai, PhD
Public Health Research Center
Tongren Hospital, Shanghai Jiao Tong University School of Medicine
No.1111 Xianxia Road
Shanghai, 200336
China
Phone: 86 021 52039999
Email: caiyong202028@hotmail.com

Abstract

Background: Men who have sex with men (MSM) face a disproportionately high risk of mpox infection, and China has recently experienced a rapid increase in the reported cases. This population also has a high prevalence of HIV, which has been identified as a critical factor in understanding the vulnerability to mpox. In addition, metabolic diseases frequently co-occur with HIV and share immunometabolic pathways, raising concerns that they may interact to confer additional risk of mpox infection.

Objective: This study examines the potential interaction between HIV and metabolic comorbidity in relation to self-reported mpox among MSM in China.

Methods: A cross-sectional study was conducted among MSM aged 18 to 76 years from October 2023 to March 2024 in 6 representative provincial regions of China. Participants completed an anonymous questionnaire on HIV infection, metabolic diseases (hypertension, diabetes mellitus, and hyperlipidemia), and mpox infection. Metabolic comorbidity was defined as the presence of more than one of these conditions. Logistic regression models were used to examine associations, and additive and multiplicative interactions between HIV and metabolic comorbidity were assessed.

Results: Of the 2403 MSM, 56 (2.33%) reported mpox, 199 (8.28%) reported HIV, and 325 (13.52%) reported at least one metabolic comorbidity (hypertension, diabetes, or hyperlipidemia). Both HIV (odds ratio [OR] 4.81, 95% CI 2.29-9.64) and metabolic comorbidity (OR 2.62, 95% CI 1.27-5.14) were associated with higher odds of mpox infection. A dose-response relationship was observed, with the odds of mpox increasing with the number of conditions (per-condition trend: OR 3.03, 95% CI 1.86-4.83). While multiplicative interaction was not statistically significant (interaction term=2.98, 95% CI 0.68-13.70; *P*=.15), additive interaction metrics suggested a possible excess association (relative excess risk due to interaction=10.80, 95% CI 1.21-37.52; attributable proportion due to interaction=0.74, 95% CI 0.07-0.87; synergy index=4.99, 95% CI 1.19-20.86). Compared to the participants without HIV or metabolic comorbidity, those with HIV and metabolic comorbidity had higher odds of mpox infection (OR 14.51, 95% CI 4.83-40.70).

Conclusions: This study suggests that HIV and metabolic comorbidity were each associated with higher odds of self-reported mpox, and exploratory analyses indicated a possible additive interaction. Given the reliance on self-reported diagnoses and the cross-sectional design, the findings should be interpreted with caution due to reporting bias and reverse causation. Further studies

XSL•FO
RenderX

are needed to confirm these associations and better understand the comprehensive health needs of MSM with co-occurring conditions.

(JMIR Public Health Surveill 2025;11:e83450) doi:10.2196/83450

KEYWORDS

HIV; metabolic comorbidity; mpox; men who have sex with men; additive interaction

Introduction

Human mpox is a zoonotic disease caused by the monkeypox virus, originally endemic in Central and West African regions [1]. In May 2022, an outbreak of mpox occurred, characterized by a rapid increase in cases reported in previously nonendemic countries [2]. In response to the global outbreak of mpox, the World Health Organization (WHO) declared a public health emergency of international concern on July 23, 2022 [3]. In May 2023, the WHO announced that mpox no longer constituted a public health emergency of international concern [4]. In contrast to the global trend, the number of reported mpox cases in China has increased rapidly, with cases mainly concentrated in economically developed regions [5]. As of August 2025, China has reported 3531 cases and ranks among the 10 most affected countries [6].

Mpox primarily spreads through person-to-person transmission via close contact, including sexual activity. It has disproportionately affected men who have sex with men (MSM), who account for 85.4% of confirmed cases worldwide [6,7]. In China, 93.8% of mpox cases were identified within the MSM population [8]. Mpox is characterized by a skin rash or mucosal lesions that progress through several stages. These lesions can appear anywhere on the body, commonly affecting the face; palms; soles; and mucosal surfaces such as the mouth, genitals, and anus [7]. Vaccination has been shown to provide protective effects [9]. However, effective control in China remains challenging due to multiple factors, including difficulties in tracing infection sources, managing close contacts, and addressing stigma toward MSM [10,11].

MSM are also a vulnerable group to HIV infection. Worldwide, high prevalence rates of HIV have been consistently reported among MSM [12-14]. In China, consecutive cross-sectional surveys in the Jiangsu province documented prevalence estimates ranging from 8% to 9.8% between 2016 and 2020 [15]. During the global mpox outbreak, a substantial proportion of cases occurred in people living with HIV [16]. Studies have consistently documented overlap between mpox and HIV [17,18], which may reflect multiple, non–mutually exclusive factors such as overlapping sexual networks, differences in testing and health care access, and immune compromise [19]. These findings underscore HIV as a critical factor in understanding vulnerability to mpox within MSM populations.

Beyond this overlap, HIV is strongly associated with a broader range of chronic conditions, particularly metabolic comorbidity [20]. A recent meta-analysis estimated that the prevalence of metabolic syndrome among HIV-positive individuals was 25.3%, with significantly higher odds compared with their HIV-negative counterparts [21]. Even in those with normal body weight, HIV has been linked to poorer metabolic health, including unfavorable blood pressure and lipid and glucose profiles [22]. Such metabolic dysfunction is associated with chronic immune dysregulation and impaired antiviral responses, providing biologically plausible pathways through which metabolic comorbidity may influence susceptibility to viral infections [23].

Building on these observations, the co-occurrence of HIV and metabolic comorbidity may confer excess risk beyond individual effects. MSM are disproportionately exposed to adverse social conditions, and syndemic theory highlights that co-occurring health conditions can interact to generate excessive disease burdens [24]. Evidence from other infectious diseases supports this concern. In the context of COVID-19, coexistence of HIV and metabolic disorders has been shown to exacerbate disease severity [25]. Evidence from tuberculosis research demonstrates that co-occurrence of HIV and diabetes is associated with elevated mortality and measurable biological interaction, indicating synergistic rather than additive effects [26,27]. Although HIV and metabolic conditions have each been associated with mpox individually [18,28], whether their combined presence could pose additional mpox risk remains unclear.

This study aimed to address this knowledge gap by investigating the relationship between HIV, metabolic comorbidity, and mpox among MSM in 6 regions of China. Specifically, our hypothesis was that (1) HIV and metabolic comorbidity would be interrelated and individually associated with mpox infection and (2) HIV and metabolic comorbidity may interact in a way associated with a greater likelihood of mpox infection.

Methods

Participants and Procedure

A nationwide cross-sectional study was conducted from October 2023 to March 2024 in 6 geographically representative provincial regions of China: Shanghai, Guangdong, Xinjiang, Shaanxi, Yunnan, and Liaoning. Given the relatively hidden nature of the MSM population, recruitment was conducted in partnership with nongovernmental organizations (NGOs) that provide services and outreach to MSM.

Before the survey started, NGO staff received standardized training on study protocols, informed consent, and quality control procedures. Relying on their established networks and trust within MSM communities, the trained staff promoted the study and invited participation. Recruitment channels included both offline venues (eg, community events and in-person outreach) and web-based platforms (eg, WeChat groups). Potential participants were informed of the study's objectives,

XSL•FO
RenderX

procedures, and potential benefits before completing the questionnaire.

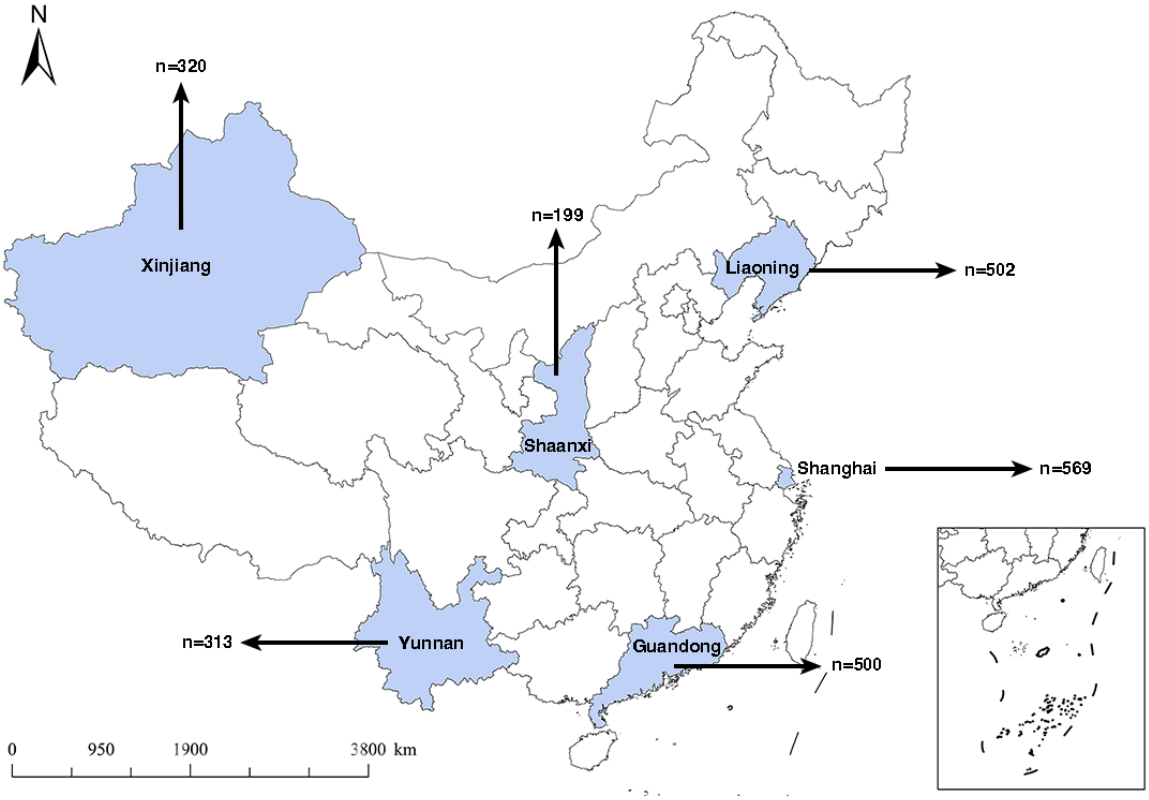
The inclusion criteria were as follows: (1) male and aged ≥18 years, (2) reporting of engagement in sexual behavior with men within the previous 6 months, and (3) regular residence in one of the 6 selected provincial regions (primarily residing locally for most of the previous 6 months). The exclusion criteria were as follows: (1) completion of the questionnaire in <300 seconds, (2) incorrect answers to quality control questions, and (3) IP address indicating a location outside the surveyed provincial regions. All eligible participants provided electronic informed consent and then completed the anonymous questionnaire via the Wenjuanxing web-based survey platform (Changsha Ranxing Information Technology Co).

The sample size was calculated using the prevalence of mpox as the primary outcome. Based on previous literature, the prevalence in MSM is approximately 5.6% [29]. Setting the allowable error (δ) at 0.20, the following sample size formula for cross-sectional surveys was applied:

A sample size of 1619 was obtained. Considering a nonresponse rate of 15%, the final sample size for this survey was determined to be at least 1905 participants.

In this study, a total of 2481 questionnaires were collected, of which 2403 (96.86%) were valid. Figure 1 shows the geographic location and sample size of each surveyed area.

Figure 1. Distribution of the participants.



Measurements

HIV Infection

Participants were asked the following: *What was your recent voluntary HIV testing (VHT) result*? A response of *positive* was considered a self-reported HIV infection.

Metabolic Comorbidity

This study primarily focused on 3 common metabolic diseases: hypertension, diabetes mellitus, and hyperlipidemia. Participants were asked whether they had been diagnosed with any of these conditions. The presence of at least one of these conditions was defined as having a metabolic comorbidity, consistent with a previous study [30]. Participants with metabolic comorbidity were referred to as the metabolically unhealthy group, whereas those without these conditions were referred to as the metabolically healthy group.

Self-Reported Mpox Infection

Participants were asked the following: *Have you been diagnosed with mpox by a doctor*? A response of *yes* was classified as a self-reported diagnosis of mpox.

Covariates

Sociodemographic characteristics included age, educational level (junior high school and lower, senior high school, and college and higher), marital status (single, married, and divorced or widowed), employment status (mental labor, manual labor, student, freelance work, and unemployed), duration of residence in the surveyed area (local, <5 years, and >5 years), and sexual orientation (homosexual and others).

XSL•FO
RenderX

Behavioral and clinical factors included the self-reported number of male sexual partners in the previous 6 months (1 or none, 2, and ≥3), condom use during anal intercourse with men in the previous 6 months (never, rarely, mostly, and always), and current use of antiretroviral therapy (yes or no).

Statistical Analysis

Descriptive analyses were conducted to examine demographic characteristics and the prevalence of HIV and metabolic diseases. Chi-square tests were used to compare the distribution of categorical variables between the mpox infection and noninfection groups, whereas independent-sample 2-tailed *t* tests were used for continuous variables. Logistic regression models were constructed to investigate the bivariate associations among HIV, metabolic comorbidity, and mpox, as well as the associations between the number of health conditions and mpox infection while adjusting for demographic factors.

We examined the interaction between HIV and metabolic comorbidity in relation to mpox infection using logistic regression analyses [31], with 95% CIs estimated using the mover method as implemented in the *interactionR* R package (R Foundation for Statistical Computing) [32]. Interactions on both multiplicative and additive scales were assessed using interaction terms and indicator variables, including relative excess risk due to interaction (RERI), attributable proportion due to interaction (AP), and synergy index (SI) [33]. If the 95% CI of the AP or RERI contained 0 or the 95% CI of the SI contained 1, this indicated the absence of an interaction effect.

Sensitivity analyses were conducted to assess the robustness of our findings. Given the relatively small number of mpox cases, we performed Firth penalized logistic regression using the R package *logistf* to reduce small-sample bias and improve the stability of estimates [34]. In addition, to evaluate the potential impact of unmeasured confounding, we used the R package *EValue* to compute *E* values [35]. The *E* value quantifies the minimum strength of association that an unmeasured confounder would need to have with both the exposure and the outcome, conditional on the measured covariates, to fully explain the observed association.

Ethical Considerations

All procedures performed in studies involving human participants were in accordance with the 1964 Declaration of Helsinki. The study protocol was approved by the Ethics Committee of Shanghai University of Medicine and Health Sciences (2023-MSMMPOX-22-310222197604080237). Written informed consent was obtained from all participants. Participants were informed that participation was voluntary and that they could withdraw from the study at any time without consequences. All data were anonymized before analysis, and no personally identifiable information was collected. Participants received a ¥50 (approximately US $7) gift card as compensation for their time.

Results

Characteristics of the Participants

Table 1 presents the characteristics of the 2403 participants aged 18 to 76 years, with a mean age of 30.59 (SD 8.02) years. Most of the participants were unmarried (2148/2403, 89.38%) and employed (1973/2403, 82.11%), whereas 78.57% (1888/2403) had an educational level of college or higher. A total of 38.95% (936/2403) of the participants were locals, and 24.84% (597/2403) had lived in the surveyed area for over 5 years. The sexual orientation of most participants (1887/2403, 78.53%) was homosexual. A total of 2.33% (56/2403) self-reported an mpox infection, and 8.28% (199/2403) self-reported an HIV infection. The prevalence rates of hypertension, diabetes mellitus, hyperlipidemia, and metabolic comorbidity were 6.45% (155/2403), 2.87% (69/2403), 9.24% (222/2403), and 13.52% (325/2403), respectively. The prevalence was 9.57% (230/2403) for participants with 1 metabolic condition and 3.95% (95/2403) for those with 2 or more conditions.

XSL•FO
RenderX

Table 1. Participant characteristics by self-reported mpox infection (N=2403).

	Total	Mpox infection		*P* value
		No (n=2347)	Yes (n=56)	
Age (y), mean (SD)	30.59 (8.02)	30.66 (8.02)	27.38 (7.42)	.002
Educational level, n (%)				.55
Junior high school or lower	157 (6.53)	152 (6.48)	5 (9)	
Senior high school	358 (14.90)	352 (15.00)	6 (11)	
College or higher	1888 (78.57)	1843 (78.53)	45 (80)	
Marital status, n (%)				<.001
Single	2035 (84.69)	1997 (85.09)	38 (68)	
Married	255 (10.61)	238 (10.14)	17 (30)	
Divorced or widowed	113 (4.70)	112 (4.77)	1 (2)	
Employment status, n (%)				.86
Mental labor	386 (16.06)	376 (16.02)	10 (18)	
Manual labor	774 (32.21)	757 (32.25)	17 (30)	
Student	320 (13.32)	315 (13.42)	5 (9)	
Freelance work	813 (33.83)	792 (33.75)	21 (38)	
Unemployed	110 (4.58)	107 (4.56)	3 (5)	
Duration of residence, n (%)				.47
Local	936 (38.95)	912 (38.86)	24 (43)	
<5 y	870 (36.20)	848 (36.13)	22 (39)	
≥5 y	597 (24.84)	587 (25.01)	10 (18)	
Sexual orientation, n (%)				.33
Homosexual	1887 (78.53)	1846 (78.65)	41 (73)	
Others	516 (21.47)	501 (21.35)	15 (27)	
HIV, n (%)				<.001
No	2204 (91.72)	2164 (92.20)	40 (71)	
Yes	199 (8.28)	183 (7.80)	16 (29)	
Hypertension, n (%)				.003
No	2248 (93.55)	2201 (93.78)	47 (84)	
Yes	155 (6.45)	146 (6.22)	9 (16)	
Diabetes, n (%)				.05
No	2334 (97.13)	2282 (97.23)	52 (93)	
Yes	69 (2.87)	65 (2.77)	4 (7)	
Hyperlipidemia, n (%)				.02
No	2181 (90.76)	2135 (90.97)	46 (82)	
Yes	222 (9.24)	212 (9.03)	10 (18)	
Metabolic status, n (%)				.003
Healthy	2078 (86.48)	2037 (86.79)	41 (73)	
Unhealthy	325 (13.52)	310 (13.21)	15 (27)	
Number of sexual partners, n (%)				.94
1 or none	1206 (50.19)	1179 (50.23)	27 (48)	
2	598 (24.89)	584 (24.88)	14 (25)	
≥3	599 (24.93)	584 (24.88)	15 (27)	

XSL•FO
RenderX

	Total	Mpox infection		*P* value
		No (n=2347)	Yes (n=56)	
Condom use, n (%)				<.001
Never	126 (5.24)	123 (5.24)	3 (5)	
Rarely	248 (10.32)	231 (9.84)	17 (30)	
Mostly	634 (26.38)	607 (25.86)	27 (48)	
Always	1395 (58.05)	1386 (59.05)	9 (16)	
Antiretroviral therapy use, n (%)				.06
No	2356 (98.04)	2303 (98.13)	53 (95)	
Yes	47 (1.96)	44 (1.87)	3 (5)	

Of the sociodemographic factors, age (*P*=.002) and marital status (*P*<.001) were associated with mpox infection. In addition, self-reported HIV (*P*<.001), hypertension (*P*=.003), hyperlipidemia (*P*=.02), and metabolic comorbidity (*P*=.003) were associated with higher odds of self-reported mpox infection.

Bivariate Associations Among HIV, Metabolic Comorbidity, and Mpox Infection

Table S1 in Multimedia Appendix 1 presents the results of the adjusted logistic regression examining bivariate associations among HIV, metabolic comorbidity, and mpox infection. After adjusting for covariates, HIV was positively associated with metabolic comorbidity (odds ratio [OR] 2.49, 95% CI 1.63-3.74). Both HIV (OR 4.81, 95% CI 2.29-9.64) and metabolic comorbidity (OR 2.62, 95% CI 1.27-5.14) were associated with higher odds of mpox infection.

Association Between Number of Health Conditions and Mpox Infection

Table 2 shows the results of the adjusted logistic regression examining the association between number of health conditions (ie, HIV and metabolic comorbidity) and mpox infection. After adjusting for covariates, an increase in the number of conditions was positively associated with mpox infection, suggesting a dose-response pattern (linear trend=3.03, 95% CI 1.86-4.83). Individuals with the simultaneous presence of HIV and metabolic comorbidity were associated with higher odds of mpox infection (OR 10.88, 95% CI 3.63-30.13).

Table 2. Association between number of health conditions (HIV and metabolic comorbidity) and self-reported mpox infection.

	Participants, n/N (%)	OR[a] (95% CI)	*P* value
Number of health conditions			
0	33/1927 (1.71)	Reference	Reference
1	15/428 (3.5)	2.16 (1.08-4.15)	.02
2	8/48 (17)	14.03 (4.64-39.59)	<.001
Linear trend	—[b]	3.03 (1.86-4.83)	<.001
Co-occurring conditions			
No	48/2355 (2.04)	Reference	Reference
Yes	8/48 (17)	10.88 (3.63-30.13)	<.001

[a]OR: odds ratio; adjusting for age, educational level, marital status, employment status, duration of residence, sexual orientation, number of sexual partners, condom use, and antiretroviral therapy use.

[b]Not applicable.

Interaction Between HIV and Metabolic Comorbidity in Relation to Mpox Infection

Table 3 shows the results of the multivariable logistic regression examining interaction between HIV and metabolic comorbidity in relation to mpox infection. No statistically significant interaction was found between HIV and metabolic comorbidity (OR 2.98, 95% CI 0.68-13.70; *P*=.15). The RERI, AP, and SI for mpox infection resulting from HIV+metabolic comorbidity were 10.80 (95% CI 1.21-37.52), 0.74 (95% CI 0.07-0.87), and 4.99 (95% CI 1.19-20.86), respectively. The findings suggest a possible additive interaction between HIV and metabolic comorbidity. When examining each metabolic disease separately, HIV showed an interaction with hypertension in relation to mpox infection, whereas no interaction was observed with diabetes mellitus or hyperlipidemia (Table S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Table 3. Interaction between HIV and metabolic comorbidity in relation to mpox infection.

	OR[a] (95% CI)	*P* value
Interaction model		
Unhealthy versus healthy	1.54 (0.61-3.83)	.36
HIV positive versus HIV negative	3.17 (1.32-7.60)	.009
Interaction term	2.98 (0.68-13.70)	.15
Multiplicative scale	2.98 (0.67-13.21)	—[b]
RERI[c]	10.80 (1.21-37.52)	—
AP[d]	0.74 (0.07-0.87)	—
SI[e]	4.99 (1.19-20.86)	—

[a]OR: odds ratio, adjusting for age, educational level, marital status, employment status, duration of residence, sexual orientation, number of sexual partners, condom use, and antiretroviral therapy use.

[b]Not applicable.

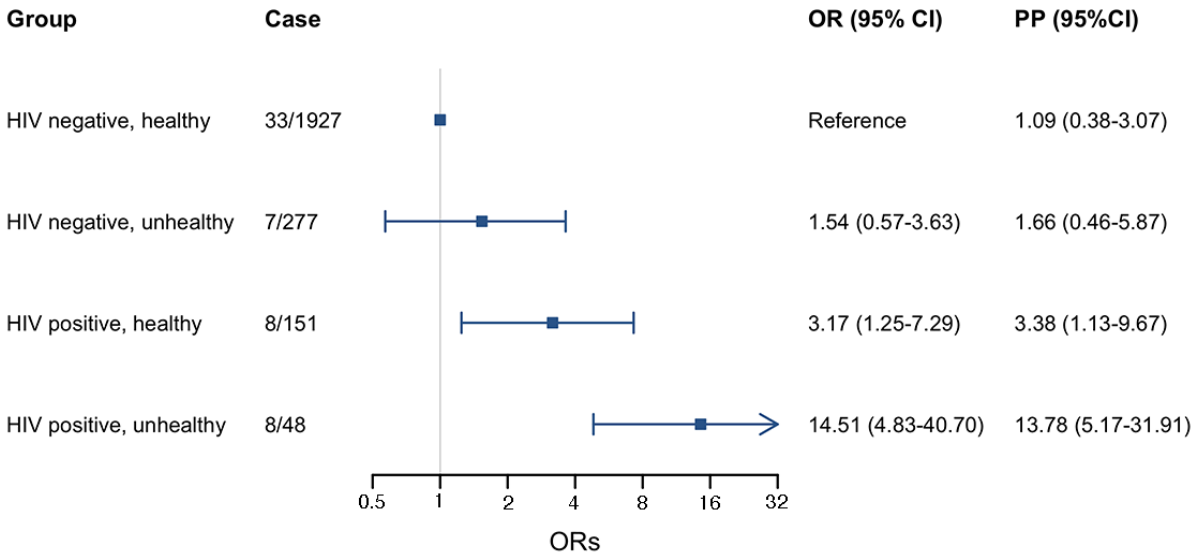
[c]RERI: relative excess risk due to interaction.

[d]AP: attributable proportion due to interaction.

[e]SI: synergy index.

Figure 2 shows the relative odds of mpox infection by HIV infection and metabolic status. Compared to individuals without self-reported HIV and who were metabolically healthy, those with self-reported HIV but who were metabolically healthy (OR 3.17, 95% CI 1.25-7.29) and those with both self-reported HIV and metabolic comorbidity (OR 14.51, 95% CI 4.83-40.70) had higher odds of self-reported mpox infection.

Figure 2. Odds ratios of self-reported mpox infection by HIV infection and metabolic status. OR: odds ratio; PP: predicted probabilities (%).



Sensitivity Analysis

The estimates of the additive interaction between HIV and metabolic comorbidity obtained using the Firth penalized likelihood logistic regression were consistent with those from the main analysis. Specifically, the RERI was 9.56 (95% CI 1.16-30.90), the AP was 0.71 (95% CI 0.08-0.86), and the SI was 4.38 (95% CI 1.20-15.94), supporting a possible additive interaction. Penalized regression models assessing mpox infection by HIV infection and metabolic status produced results that were in line with the primary findings (Table S3 in Multimedia Appendix 1).

In the quantitative bias analysis, the group with self-reported HIV but who was metabolically healthy yielded an *E* value of 5.79 (lower 95% CI bound=1.81), suggesting that an unmeasured confounder would need to be associated with both exposure and outcome by an OR of at least 5.79 to fully account for the observed association. For the group with both self-reported HIV and metabolic comorbidity, the *E* value was 28.51 (lower 95% CI bound=9.13), indicating that only a strong unmeasured confounder could explain the observed association.

XSL•FO
RenderX

Discussion

Principal Findings

This study offers valuable insights into the associations among HIV, metabolic comorbidity, and self-reported mpox. In our sample of MSM in China, we observed a clustering of HIV and metabolic comorbidity, both of which were individually associated with mpox infection. In addition, our exploratory analyses suggested a possible additive interaction between HIV and metabolic comorbidity in relation to mpox infection.

MSM, as a sexual minority population, often face social, economic, and environmental disadvantages, leading to a higher prevalence of adverse health conditions [36]. In this study among MSM, the prevalence rates of HIV and metabolic comorbidity were 8.28% (199/2403) and 13.52% (325/2403), respectively, with a positive correlation observed between the 2 health conditions. Consistent with previous findings that metabolic diseases tend to occur earlier in people living with HIV [20], we found that 24.1% (48/199) of MSM who were infected with HIV simultaneously exhibited metabolic comorbidity. The underlying factors contributing to these comorbidities are multifaceted and include antiretroviral therapy use, unhealthy lifestyle behaviors (ie, smoking and alcohol use), and chronic inflammation despite suppressed HIV replication [37].

Existing studies among MSM have predominantly examined the interplay of psychological, social, and structural factors using the syndemic framework [38-40], whereas relatively few have focused on HIV and co-occurring physical health conditions [41,42]. This study provides exploratory observations on the association between HIV and metabolic comorbidity in relation to mpox infection among MSM. While previous research has documented associations between both HIV and metabolic diseases and mpox infection individually [18,28], our analysis, through calculation of additive interaction measures (RERI, AP, and SI), suggests a previously unrecognized possibility of interaction between HIV and metabolic comorbidity. These findings suggest that MSM with both HIV and metabolic comorbidity may warrant further investigation in future research and consideration in the design of integrated health services.

In this study, we identified 2.33% (56/2403) of cases of mpox infection among MSM in China. This prevalence rate is notably higher than the 0.73% reported in a previous nationwide survey conducted in early August 2023 [43], primarily because a substantial number of confirmed mpox cases in China were reported after that [8]. In addition, despite the clear clustering of HIV with metabolic comorbidity, the prevalence rates of each metabolic condition, namely, hypertension (155/2403, 6.45%), diabetes mellitus (69/2403, 2.87%), and hyperlipidemia (222/2403, 9.24%), were relatively low in our sample. This may explain the nonsignificant interaction observed between HIV and each separate metabolic disease in this study.

Implications

Our findings suggest that HIV and metabolic health issues tend to cluster and interact, and exploratory analyses suggested a possible additive interaction in relation to mpox infection. Understanding the individual and contextual factors underlying these interactions—such as substance use, living conditions, and social networks—will be important for informing future research [41,44]. Although this study was conducted in China, the results may offer insights relevant to other low- and middle-income countries, where HIV and noncommunicable diseases are highly prevalent and health systems often underresourced [45,46]. Strengthening resilient health systems capable of delivering integrated care is crucial for advancing universal health coverage [47]. Previous research has shown that interventions focusing solely on individual conditions such as HIV have limited effectiveness in improving overall health [48,49]. Building on these lessons, global experiences from infectious disease emphasize the need for systematic health system redesign and stronger international collaboration [50]. In this context, MSM with multiple conditions may represent a subgroup warranting attention in future comprehensive health services. However, the design and evaluation of such interventions should be based on stronger empirical evidence from future studies.

Limitations

This study has several limitations. First, HIV status, mpox status, and metabolic comorbidity were self-reported, which may have introduced misclassification and reporting bias. Given that mpox-related stigma among MSM can discourage care-seeking behaviors [51], underreporting of previous diagnoses is plausible. In epidemiological studies, nondifferential misclassification due to underreporting typically reduces sensitivity while maintaining high specificity, which tends to bias effect estimates toward the null [52]. Therefore, our findings are likely conservative and should be interpreted as exploratory. Second, the sample was recruited through NGOs and web-based platforms, potentially overrepresenting MSM who are more engaged with health services and limiting generalizability to hidden or higher-risk subgroups. Third, the cross-sectional design precludes causal inference. As a result, the temporal sequence among HIV, metabolic comorbidity, and self-reported mpox cannot be established, and reverse causation remains possible.

Fourth, the number of mpox cases was relatively small, which limited statistical power and precision, especially for interaction analyses. The wide CIs for the interaction terms reflect this instability. Although we conducted sensitivity analyses using the Firth penalized likelihood logistic regression to mitigate sparse-data bias and found consistent results, the findings should still be interpreted as exploratory. Fifth, our study focused on 3 common metabolic diseases—hypertension, diabetes mellitus, and hyperlipidemia—potentially overlooking other important conditions such as obesity. Finally, while we adjusted for several behavioral and clinical covariates, other relevant factors such as sex work, history of other sexually transmitted diseases, viral load, and vaccination status were not collected, which may have led to residual confounding. However, our quantitative bias analysis indicated that only unmeasured confounders with very strong associations with both exposure and outcome could fully explain the observed results, suggesting that the main findings are relatively robust. Taken together, these limitations highlight

XSL•FO
RenderX

the need for future studies with larger and more representative samples; objective diagnostic measures; and more comprehensive assessments of biomedical, behavioral, and contextual factors.

Conclusions

Our findings suggest a potential clustering and additive interaction of HIV and metabolic comorbidity in relation to self-reported mpox infection among MSM. These exploratory results highlight the importance of considering co-occurring health conditions among MSM and may offer insights for similar populations in other settings with high burdens of HIV and noncommunicable diseases.

Acknowledgments

This study was funded by the key discipline projects of the Shanghai Three-Year Action Plan for Public Health (grant GWVI-11.1-29), Science and Technology Commission of Shanghai Municipality (20JC1410204), and public health research project from the Shanghai Disease Control and Prevention Administration (2024GKM30). The authors would like to thank Shanghai Qingai Health Center (Shanghai), Consultation Center of AIDS Aid and Health Service (Liaoning), Trans China (Yunnan), Xinjiang Rong'ai Health Consulting Co. Ltd (Xinjiang), Xi'an Zhitong AIDS Prevention and Health Center (Shaanxi), Lingnan Partner Community Support Center (Guangdong), and Professor Xiang-Sheng Chen (Institute of Dermatology, Chinese Academy of Medical Science and Peking Union Medical College, Nanjing, China) for their contributions to data collection.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to privacy or ethical restrictions but are available from the corresponding author on reasonable request.

Authors' Contributions

YC and YW contributed to the conceptualization and methodology. YC, JZ, and JC contributed to funding acquisition and data curation. YL, YZ, and YW contributed to the formal analysis. YL and XG contributed to the writing of the original draft. YL, MZ, and GX contributed to the review and editing of the manuscript. All authors contributed to the final approval of the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional tabulated results.
[DOCX File , 18 KB - publichealth_v11i1e83450_app1.docx]

Multimedia Appendix 2
STROBE checklist.
[DOC File , 73 KB - publichealth_v11i1e83450_app2.doc]

References

1. Pourriyahi H, Aryanian Z, Afshar ZM, Goodarzi A. A systematic review and clinical atlas on mucocutaneous presentations of the current monkeypox outbreak: with a comprehensive approach to all dermatologic and nondermatologic aspects of the new and previous monkeypox outbreaks. J Med Virol 2023 Feb;95(2):e28230. [doi: 10.1002/jmv.28230] [Medline: 36254380]
2. Mpox (monkeypox) outbreak 2022. World Health Organization. URL: https://www.who.int/emergency/situations/monkeypox-oubreak-2022 [accessed 2025-10-23]
3. WHO director-general declares the ongoing monkeypox outbreak a public health emergency of international concern 2022. World Health Organization. 2023 Jul 23. URL: https://www.who.int/europe/news/item/23-07-2022-who-director-general-declares-the-ongoing-monkeypox-outbreak-a-public-health-event-of-international-concern [accessed 2025-10-23]
4. Mpox IHR emergency committee 2024. World Health Organization. URL: https://tinyurl.com/4pnj6xjd [accessed 2025-10-23]
5. Ren R, Li C, Bai W, Wang Y, Li D, Ding F, et al. The epidemiological characteristics of mpox cases - China, 2023. China CDC Wkly 2024 Jun 28;6(26):619-623. [doi: 10.46234/ccdcw2024.118] [Medline: 38966310]
6. 2022-23 mpox outbreak: global trends. World Health Organization. URL: https://worldhealthorg.shinyapps.io/mpx_global/ [accessed 2025-10-23]

XSL•FO
RenderX

7. Fact sheets of Mpox. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/monkeypox [accessed 2025-10-23]
8. Mpox outbreak surveillance for November 2023. Chinese Center for Disease Control and Prevention. URL: https://www.chinacdc.cn/jkyj/crb2/yl/hd/gnyq/202409/t20240906_298327.html [accessed 2025-10-23]
9. Zhang XS, Niyomsri S, Mandal S, Mohammed H, Mindlin M, Dugbazah B, et al. Cost-effectiveness of vaccination strategies to control future Mpox outbreaks in england: a modelling study. Lancet Reg Health Eur 2025 Jul 1;55:101364 [FREE Full text] [doi: 10.1016/j.lanepe.2025.101364] [Medline: 40950948]
10. Shang W, Yan J, Zhang H, Liang W, Liu J, Liu M. Mpox containment strategies in China, 2022-2023: from preparedness to legal management. BMJ Glob Health 2025 Aug 26;10(8):e019279 [FREE Full text] [doi: 10.1136/bmjgh-2025-019279] [Medline: 40858338]
11. Liu Y, Zhang J, Cai Y. Mpox-related stigma and healthcare-seeking behavior among men who have sex with men. Glob Health Res Policy 2025 Apr 14;10(1):16 [FREE Full text] [doi: 10.1186/s41256-025-00418-w] [Medline: 40229912]
12. Wang P, Wogayehu A, Fahad A, Menschner C, Bolden B, Rosmarin-DeStefano C, et al. HIV prevalence and associated factors among men who have sex with men (MSM) in New Jersey, U.S., 2017. AIDS Behav 2022 May;26(5):1652-1659. [doi: 10.1007/s10461-021-03516-1] [Medline: 34787759]
13. Ribeiro Banze A, Muleia R, Nuvunga S, Boothe M, Semá Baltazar C. Trends in HIV prevalence and risk factors among men who have sex with men in Mozambique: implications for targeted interventions and public health strategies. BMC Public Health 2024 Apr 27;24(1):1185 [FREE Full text] [doi: 10.1186/s12889-024-18661-0] [Medline: 38678173]
14. Wardhani BD, Grulich AE, Kawi NH, Prasetia Y, Luis H, Wirawan GB, et al. Very high HIV prevalence and incidence among men who have sex with men and transgender women in Indonesia: a retrospective observational cohort study in Bali and Jakarta, 2017-2020. J Int AIDS Soc 2024 Nov;27(11):e26386. [doi: 10.1002/jia2.26386] [Medline: 39448551]
15. Zhou Y, Chen Y, Lu J, Zhang Z, Sun Q, Liu X, et al. Sustained high HIV incidence among men who have sex with men in Jiangsu province, China: based on the limiting-antigen avidity EIA method and six consecutive surveys, 2016-2021. Front Public Health 2023 Nov 20;11:1277570 [FREE Full text] [doi: 10.3389/fpubh.2023.1277570] [Medline: 38054072]
16. Ortiz-Martínez Y, Zambrano-Sanchez G, Rodríguez-Morales AJ. Monkeypox and HIV/AIDS: when the outbreak faces the epidemic. Int J STD AIDS 2022 Sep;33(10):949-950. [doi: 10.1177/09564624221114191] [Medline: 35822876]
17. Curran KG, Eberly K, Russell OO, Snyder RE, Phillips EK, Tang EC, Monkeypox, HIV, and STI Team. HIV and sexually transmitted infections among persons with monkeypox - eight U.S. jurisdictions, May 17-July 22, 2022. MMWR Morb Mortal Wkly Rep 2022 Sep 09;71(36):1141-1147 [FREE Full text] [doi: 10.15585/mmwr.mm7136a1] [Medline: 36074735]
18. Liu BM, Rakhmanina NY, Yang Z, Bukrinsky MI. Mpox (monkeypox) virus and its co-infection with HIV, sexually transmitted infections, or bacterial superinfections: double whammy or a new prime culprit? Viruses 2024 May 15;16(5):784 [FREE Full text] [doi: 10.3390/v16050784] [Medline: 38793665]
19. Ortiz-Saavedra B, Montes-Madariaga ES, Cabanillas-Ramirez C, Alva N, Ricardo-Martínez A, León-Figueroa DA, et al. Epidemiologic situation of HIV and monkeypox coinfection: a systematic review. Vaccines (Basel) 2023 Jan 22;11(2):246 [FREE Full text] [doi: 10.3390/vaccines11020246] [Medline: 36851124]
20. Willig AL, Overton ET. Metabolic consequences of HIV: pathogenic insights. Curr HIV/AIDS Rep 2014 Mar;11(1):35-44. [doi: 10.1007/s11904-013-0191-7] [Medline: 24390642]
21. Trachunthong D, Tipayamongkholgul M, Chumseng S, Darasawang W, Bundhamcharoen K. Burden of metabolic syndrome in the global adult HIV-infected population: a systematic review and meta-analysis. BMC Public Health 2024 Sep 28;24(1):2657 [FREE Full text] [doi: 10.1186/s12889-024-20118-3] [Medline: 39342258]
22. Lake JE, Li X, Palella FJ, Erlandson KM, Wiley D, Kingsley L, et al. Metabolic health across the BMI spectrum in HIV-infected and HIV-uninfected men. AIDS 2018 Jan 02;32(1):49-57 [FREE Full text] [doi: 10.1097/QAD.0000000000001651] [Medline: 28926404]
23. Perakakis N, Harb H, Hale BG, Varga Z, Steenblock C, Kanczkowski W, et al. Mechanisms and clinical relevance of the bidirectional relationship of viral infections with metabolic diseases. Lancet Diabetes Endocrinol 2023 Sep;11(9):675-693. [doi: 10.1016/S2213-8587(23)00154-7] [Medline: 37524103]
24. Ouafik MR, Buret L, Scholtes B. Mapping the current knowledge in syndemic research applied to men who have sex with men: a scoping review. Soc Sci Med 2022 Aug;306:115162. [doi: 10.1016/j.socscimed.2022.115162] [Medline: 35779501]
25. Krishnan A, Woreta TA, Sims OT, Hamilton JP, Potter JJ, Alqahtani SA. Impact of nonalcoholic fatty liver disease on clinical outcomes in patients with COVID-19 among persons living with HIV: a multicenter research network study. J Infect Public Health 2023 May;16(5):673-679 [FREE Full text] [doi: 10.1016/j.jiph.2023.02.008] [Medline: 36931143]
26. Sinha P, Moll AP, Brooks RP, Deng YH, Shenoi SV. Synergism between diabetes and human immunodeficiency virus in increasing the risk of tuberculosis. Int J Tuberc Lung Dis 2018 Jul 01;22(7):793-799 [FREE Full text] [doi: 10.5588/ijtld.17.0936] [Medline: 29914606]
27. Houck K, Chakhaia T, Gorvetzian S, Critchley JA, Schechter MC, Magee MJ. Diabetes mellitus and tuberculosis treatment outcomes: interaction assessment between hyperglycemia and human immunodeficiency virus in the state of Georgia, 2015-2020. Open Forum Infect Dis 2023 May 10;10(6):ofad255 [FREE Full text] [doi: 10.1093/ofid/ofad255] [Medline: 37383249]

XSL•FO
RenderX

28. Ruiz-Pacheco JA, Castillo-Díaz LA, Arreola-Torres R, Fonseca-Coronado S, Gómez-Navarro B. Diabetes mellitus: lessons from COVID-19 for monkeypox infection. Prim Care Diabetes 2023 Apr;17(2):113-118 [FREE Full text] [doi: 10.1016/j.pcd.2023.01.008] [Medline: 36737358]
29. Minhaj FS, Singh V, Cohen SE, Townsend MB, Scott H, Szumowski J, et al. Prevalence of undiagnosed monkeypox virus infections during global mpox outbreak, United States, June-September 2022. Emerg Infect Dis 2023 Nov;29(11):2307-2314 [FREE Full text] [doi: 10.3201/eid2911.230940] [Medline: 37832516]
30. Ul-Haq Z, Mackay DF, Fenwick E, Pell JP. Impact of metabolic comorbidity on the association between body mass index and health-related quality of life: a Scotland-wide cross-sectional study of 5,608 participants. BMC Public Health 2012 Feb 24;12:143 [FREE Full text] [doi: 10.1186/1471-2458-12-143] [Medline: 22364437]
31. Tsai AC, Venkataramani AS. Syndemics and health disparities: a methodological note. AIDS Behav 2016 Feb;20(2):423-430 [FREE Full text] [doi: 10.1007/s10461-015-1260-2] [Medline: 26662266]
32. Alli BY. InteractionR: an R package for full reporting of effect modification and interaction. Softw Impacts 2021 Nov;10:100147. [doi: 10.1016/j.simpa.2021.100147]
33. Andersson T, Alfredsson L, Källberg H, Zdravkovic S, Ahlbom A. Calculating measures of biological interaction. Eur J Epidemiol 2005;20(7):575-579. [doi: 10.1007/s10654-005-7835-x] [Medline: 16119429]
34. Heinze G, Ploner M, Jiricka L, Steiner G. logistf: Firth's bias-reduced logistic regression. The Comprehensive R Archive Network. 2025 Apr 16. URL: https://CRAN.R-project.org/package=logistf [accessed 2025-10-23]
35. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the e-value. Ann Intern Med 2017 Aug 15;167(4):268-274. [doi: 10.7326/M16-2607] [Medline: 28693043]
36. Fredriksen-Goldsen KI, Kim HJ, Barkan SE, Muraco A, Hoy-Ellis CP. Health disparities among lesbian, gay, and bisexual older adults: results from a population-based study. Am J Public Health 2013 Oct;103(10):1802-1809 [FREE Full text] [doi: 10.2105/AJPH.2012.301110] [Medline: 23763391]
37. Petoumenos K, Huang R, Hoy J, Bloch M, Templeton DJ, Baker D, et al. Prevalence of self-reported comorbidities in HIV positive and HIV negative men who have sex with men over 55 years-the Australian Positive and Peers Longevity Evaluation Study (APPLES). PLoS One 2017 Sep 08;12(9):e0184583 [FREE Full text] [doi: 10.1371/journal.pone.0184583] [Medline: 28886173]
38. Qian Y, Detels R, Comulada WS, Hidalgo MA, Lee SJ, Biello KB, et al. Longitudinal analysis of overlapping psychosocial factors predicting incident hospitalization among mixed HIV serostatus men who have sex with men in the multicenter AIDS cohort study. AIDS Behav 2024 Sep;28(9):1-12. [doi: 10.1007/s10461-024-04356-5] [Medline: 38703339]
39. Guimarães MD, McKinnon K, Dourado I, Veras MA, Magno L, Almeida M, et al. A potential syndemic effect associated with symptoms of depression among men who have sex with men. Braz J Psychiatry 2022 Aug 29;44(5):517-521 [FREE Full text] [doi: 10.47626/1516-4446-2022-2454] [Medline: 36423231]
40. Houang ST, Kafka JM, Choi SK, Meanley SP, Muessig KE, Bauermeister JA, et al. Co-occurring epidemic conditions among Southern U.S. Black men who have sex with men in an online eHealth intervention. AIDS Behav 2023 Feb;27(2):641-650. [doi: 10.1007/s10461-022-03799-y] [Medline: 35986818]
41. Byg B, Bazzi AR, Funk D, James B, Potter J. The utility of a syndemic framework in understanding chronic disease management among HIV-infected and type 2 diabetic men who have sex with men. J Community Health 2016 Dec;41(6):1204-1211. [doi: 10.1007/s10900-016-0202-x] [Medline: 27137196]
42. Bhardwaj A, Kohrt BA. Syndemics of HIV with mental illness and other noncommunicable diseases: a research agenda to address the gap between syndemic theory and current research practice. Curr Opin HIV AIDS 2020 Jul;15(4):226-231 [FREE Full text] [doi: 10.1097/COH.0000000000000627] [Medline: 32487814]
43. Du M, Zheng M, Liu J, Yang G, Yao Y, Qian X, et al. The prevalence of mpox and its association with sexual behavior among Chinese men who have sex with men in early August 2023. J Med Virol 2023 Dec;95(12):e29320. [doi: 10.1002/jmv.29320] [Medline: 38102917]
44. Pirrone I, Dieleman M, Reis R, Pell C. Syndemic contexts: findings from a review of research on non-communicable diseases and interviews with experts. Glob Health Action 2021 Jan 01;14(1):1927332 [FREE Full text] [doi: 10.1080/16549716.2021.1927332] [Medline: 34308786]
45. Wong EB, Olivier S, Gunda R, Koole O, Surujdeen A, Gareta D, Vukuzazi Team. Convergence of infectious and non-communicable disease epidemics in rural South Africa: a cross-sectional, population-based multimorbidity study. Lancet Glob Health 2021 Jul;9(7):e967-e976 [FREE Full text] [doi: 10.1016/S2214-109X(21)00176-5] [Medline: 34143995]
46. Ross JL, Rupasinghe D, Chanyachukul T, Crabtree Ramírez B, Murenzi G, Kwobah E, Sentinel Research Network of the International epidemiology Databases to Evaluate AIDS. Comorbidities and HIV-related factors associated with mental health symptoms and unhealthy substance use among older adults living with HIV in low- and middle-income countries: a cross-sectional study. J Int AIDS Soc 2025 Mar;28(3):e26434 [FREE Full text] [doi: 10.1002/jia2.26434] [Medline: 40045453]
47. Debie A, Nigusie A, Gedle D, Khatri RB, Assefa Y. Building a resilient health system for universal health coverage and health security: a systematic review. Glob Health Res Policy 2024 Jan 04;9(1):2 [FREE Full text] [doi: 10.1186/s41256-023-00340-z] [Medline: 38173020]

XSL•FO
RenderX

48. Halkitis PN, Wolitski RJ, Millett GA. A holistic approach to addressing HIV infection disparities in gay, bisexual, and other men who have sex with men. Am Psychol 2013;68(4):261-273. [doi: 10.1037/a0032746] [Medline: 23688093]
49. Herbst JH, Beeker C, Mathew A, McNally T, Passin WF, Kay LS, Task Force on Community Preventive Services. The effectiveness of individual-, group-, and community-level HIV behavioral risk-reduction interventions for adult men who have sex with men: a systematic review. Am J Prev Med 2007 Apr;32(4 Suppl):S38-S67. [doi: 10.1016/j.amepre.2006.12.006] [Medline: 17386336]
50. Zhang XX, Jin YZ, Lu YH, Huang LL, Wu CX, Lv S, et al. Infectious disease control: from health security strengthening to health systems improvement at global level. Glob Health Res Policy 2023 Sep 05;8(1):38 [FREE Full text] [doi: 10.1186/s41256-023-00319-w] [Medline: 37670331]
51. Ge X, Liu Y, Liu S, Tao S, Xu C, Wu B, et al. Can social support improve mpox care-seeking among men who have sex with men through stigma reduction: a chain mediation analysis from China. Glob Health Res Policy 2025 Oct 01;10(1):49 [FREE Full text] [doi: 10.1186/s41256-025-00452-8] [Medline: 41029767]
52. Copeland KT, Checkoway H, McMichael AJ, Holbrook RH. Bias due to misclassification in the estimation of relative risk. Am J Epidemiol 1977 May;105(5):488-495. [doi: 10.1093/oxfordjournals.aje.a112408] [Medline: 871121]

Abbreviations

AP: attributable proportion due to interaction
MSM: men who have sex with men
NGO: nongovernmental organization
OR: odds ratio
RERI: relative excess risk due to interaction
SI: synergy index
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 03.09.25; peer-reviewed by F Wang, S Chen; comments to author 24.09.25; accepted 13.10.25; published 01.12.25.

Please cite as:
Liu Y, Ge X, Zhang Y, Wang Y, Zhang M, Chen J, Xu G, Zhang J, Wang Y, Cai Y
Associations Between Both HIV and Metabolic Comorbidity and Self-Reported Mpox Among Men Who Have Sex With Men: Multicenter Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e83450
URL: https://publichealth.jmir.org/2025/1/e83450
doi:10.2196/83450
PMID:

©Yujie Liu, Xin Ge, Yinghuan Zhang, Yuxuan Wang, Meihui Zhang, Jianyu Chen, Gang Xu, Jiechen Zhang, Ying Wang, Yong Cai. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 01.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association of Frailty Index at 66 Years of Age with Health Care Costs and Utilization Over 10 Years in Korea: Retrospective Cohort Study

Jieun Jang[1], PhD; Anna Kim[2], PhD; Mingee Choi[3], PhD; Ellen P McCarthy[4,5], MPH, PhD; Brianne Olivieri-Mui[6], MPH, PhD; Chan Mi Park[4], MD, MPH; Jae-Hyun Kim[7], PhD; Jaeyong Shin[3,8], MD, MPH, PhD; Dae Hyun Kim[4,5], MD, MPH, SCD

[1]Department of Preventive Medicine, College of Medicine, Dongguk University, Gyeongju, Republic of Korea

[2]School of Economics, Yonsei University, Seoul, Republic of Korea

[3]Department of Preventive Medicine, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

[4]Frailty Research Center, Hinda and Arthur Marcus Institute for Aging Research, Hebrew SeniorLife, Harvard Medical School, Boston, MA, United States

[5]Division of Gerontology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, United States

[6]Department of Health Sciences, Bouvé College of Health Sciences, Northeastern University, Boston, MA, United States

[7]Department of Health Administration, College of Health Science, Dankook University, Cheonan, Republic of Korea

[8]Department of Policy Analysis and Management, College of Human Ecology, Cornell University, Ithaca, NY, United States

Corresponding Author:
Jaeyong Shin, MD, MPH, PhD
Department of Preventive Medicine, Yonsei University College of Medicine, 50 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: The long-term economic impact of frailty measured at the beginning of elderhood is unknown.

Objective: The objective of our study was to examine the association between an individual's frailty index at 66 years of age and their health care costs and utilization over 10 years.

Methods: This retrospective cohort study included 215,887 Koreans who participated in the National Screening Program for Transitional Ages at 66 years of age between 2007 - 2009. Frailty was categorized using a 39-item deficit accumulation frailty index: robust (<0.15), prefrail (0.15 - 0.24), and frail (≥0.25). The primary outcome was total health care cost, while the secondary outcomes were inpatient and outpatient health care costs, inpatient days, and number of outpatient visits. Generalized estimating equations with a gamma distribution and identity link function were used to investigate the association between the frailty index and health care costs and utilization until December 31, 2019.

Results: The study population included 53.3% (n=115,113) women, 32.9% (n=71,082) with prefrailty, and 9.7% (n=21,010) with frailty. The frailty level at 66 years of age was associated with higher cumulative total costs (robust to frail: $19,815 to $28.281; *P*<.001), inpatient costs (US $11,189 to US $16,627; *P*<.001), and outpatient costs (US $8,625 to US $11,654; *P*<.001) over the next 10 years. In the robust group, a one-year increase in age was associated with increased total health care costs (mean change per beneficiary per year: US $206.2; SE: $1.2; *P*<.001), inpatient costs (US $126.8; SE: $1.0; *P*<.001), and outpatient costs (US $74.4; SE: $0.4; *P*<.001). In the frail group, the increase in total health care costs was greater compared to the robust group (difference in mean cost per beneficiary per year: US $120.9; SE: $5.3; *P*<.001), inpatient costs (US $102.8; SE: $5.22; *P*<.001), and outpatient costs (US $15.6; SE: $1.5; *P*<.001). Similar results were observed for health care utilization (*P*<.001). Among the robust group, a one-year increase in age was associated with increased inpatient days (mean change per beneficiary per year: 0.9 d; *P*<.001) and outpatient visits (2.1 visits; *P*<.001). In the frail group, inpatient days increased annually compared to the robust group (difference in the mean inpatient days per beneficiary per year: 1.5 d; *P*<.001), while outpatient visits increased to a lesser extent (difference in the mean outpatient visits per beneficiary per year: −0.2 visits; *P*<.001).

Conclusions: Our study demonstrates the potential utility of assessing frailty at 66 years of age in identifying older adults who are more likely to incur high health care costs and utilize health care services over the subsequent 10 years. The long-term high health care costs and utilization associated with frailty and prefrailty warrants public health strategies to prevent and manage frailty in aging populations.

(JMIR Public Health Surveill 2025;11:e50026) doi:10.2196/50026

XSL•FO
RenderX

KEYWORDS

frailty index; health care costs; health care utilization; elderly; Korea; frailty; aging; utilization; older adults; sociodemographic; cost; prevention

Introduction

Health care systems face the challenge of managing increasing health care costs in aging populations [1,2]. South Korea experiences one of the fastest rates of population aging in the world, with 40% of its population expected to be aged over 65 years by 2050 [3]. The country is already experiencing a deficit in health insurance budget [4]. Proactive identification of individuals who are likely to incur high health care costs and utilization is critical for developing strategies to control health care expenditures in the aging society.

Frailty is defined by a decline in physiological reserves across multiple organ systems, leading to increased susceptibility to poor health outcomes following stressors [5]. Previous studies have found that frailty was associated with increased health care costs and utilization [6-10]. Moreover, individuals with frailty accounted for 43.9% of preventable health care expenditures [11]. One of the established models for measuring frailty is the deficit accumulation frailty model, expressed as a frailty index (FI), which is calculated by dividing the number of deficits present by the total number of deficits considered [12,13]. The FI can range from 0 to 1, with higher scores indicating greater frailty [13]. It can be derived from an existing database that contains information on standardized health assessments.

Since 2007, all Koreans who turn 66 years old have been invited to participate in a comprehensive health evaluation as part of the National Screening Program for Transitional Ages at government-approved clinics, hospitals, and public health facilities [14]. This examination assesses lifestyle, medical history, functional status, and laboratory tests, providing data to quantify a deficit accumulation FI on a national level. We have recently shown that higher FI scores at age 66 were associated with faster development of age-related chronic diseases over the subsequent 10 years among nearly 1 million Koreans [15]. The availability of standardized health assessments through the National Screening Program for Transitional Ages provides a unique opportunity to identify the FI across a large cohort. However, it is uncertain whether the FI at age 66 can predict long-term health care costs and utilization.

We conducted a nationwide cohort study of Koreans who participated in the National Screening Program for Transitional Ages in 2007 - 2009 to examine the association between the FI at age 66 and subsequent health care costs and utilization over 10 years. South Korea offers a unique context for research in frailty due to its nationwide single-payer health insurance system, which ensures both universal coverage, and provides detailed and standardized claims data for the entire population. This system facilitates long-term follow-up of health care costs and utilization, enabling population-level analyses that are rarely feasible in other health care settings. In this study, we hypothesized that a higher FI at age 66 would be associated with a greater increase in health care costs and utilization over 10 years.

Methods

Data Sources

We accessed the National Screening Program for Transitional Ages database (2007 - 2009), which was linked to the Korean National Health Insurance database (2004 - 2019), through the Korean National Health Insurance Corporation research program. The dataset included a 35% (n=435,572) random sample of adults who reached the age of 66 years between 2007-2009. The screening program database includes information on lifestyle indicators, medical history, functional status, and laboratory measurements. The Korean National Health Insurance database includes ICD-10 diagnostic codes, sociodemographic variables, health service claims, health care utilization, and long-term care insurance claims [16].

Study Population

We included individuals with complete sociodemographic information who participated in the screening program (n=222,480) through the 435,572 enrollees of the Korean National Health Insurance aged 66 years between 2007 - 2009. We excluded those with (1) duplicate records (n=1460), (2) insufficient data (<80% of the necessary items) for calculating frailty (n=5119), or (3) death within the month of examination or the month following the examination (n=14). Our final cohort included 215,887 enrollees (Figure S1 in Multimedia Appendix 1). This cohort included both community-dwelling older adults and long-term care residents.

Measurement of Frailty and Other Characteristics

The detailed procedure for constructing the FI has been described previously [15]. Briefly, we followed the standard procedure established by Searle et al [17]. The variables were selected as health deficits if they met the following criteria: (1) associated with health status, (2) prevalence increased with age, (3) did not saturate too early, and (4) covered a range of organ systems. We calculated the FI (range: 0 to 1; higher scores indicated greater frailty) based on 39 health-deficit items in the following health domains: medical history (15 items), biometric or laboratory measures (8 items), physical health (2 items), psychological health (8 items), and disability (6 items) (the definition of each item is provided in Table S1 in Multimedia Appendix 1) [15]. Frailty categories were defined using the previously used cut off points [15,18,19]: robust (<0.15), prefrail (0.15 to <0.25), and frail (≥0.25). We also assessed the presence of chronic conditions using ICD-10 diagnosis codes (ie, 1 inpatient or 2 outpatient diagnoses) from the previous year. Additionally, the following characteristics were obtained from the screening examination: sex, annual income level (quantiles), insurance status (employee insurance, self-employed insurance, or medical aid for low income), residential area (capital,

XSL•FO
RenderX

metropolitan, or rural areas), and examination year (2007, 2008, or 2009).

Outcome Measurements and Follow Up

The primary outcome was total health care costs per beneficiary per year. These health care costs were the sum of reimbursements from the Korean National Health Insurance and beneficiaries' cost-sharing for inpatient and outpatient care. Secondary outcomes were inpatient costs, outpatient costs, inpatient days, and the number of outpatient visits per beneficiary per year. Follow up began on the day after the screening examination and lasted until the following, depending on whichever occurred earlier: date of death, 10 years from the screening examination, or December 31, 2019. To account for variations in inflation, we adjusted the costs to 2007 Korean Won (KRW) using an annual conversion factor (converted into US dollars at an exchange rate of 1 USD=1200 KRW). The extreme cost values were truncated at the 1st and 99th percentiles (ie, replacing values smaller than the 1st percentile with the 1st percentile value, and values larger than the 99th percentile with the 99th percentile).

Statistical Analysis

We used the χ^2 test to compare baseline characteristics by frailty category. The cumulative annual costs (total health care costs, inpatient costs, and outpatient costs) were compared using one-way ANOVA. Additionally, we compared the annual rates of health care utilization (inpatient days and number of outpatient visits) over 10 years, according to the frailty category. To investigate the association between the FI at age 66 years and health care costs and utilization over 10 years, generalized estimating equations were applied with a gamma distribution and the identity link function [20,21]. This was done to model right-skewed and over-dispersed health care cost and utilization data, accounting for repeated annual measures of costs and utilization within each beneficiary. Subgroup analyses were conducted by sex and by health insurance type. We tested the homogeneity of the interaction term between time and frailty categories across sex and health insurance type [22,23]. All models were adjusted for sex, annual income, insurance status, residential area, and examination years.

$$Y_{it}=\beta_0+\beta_1 Prefrail_{it}+\beta_2 frail_{it}+\beta_3 Time_{it}+\beta_4 Time_{it}*Prefrail_{it}+\beta_5 Time_{it}*frail_{it}+\beta_6 X_{it}+\varepsilon_{it}$$

- Y_{iT}: Dependent variable for an individual participant i observed for time T
- $T=t–c$; where time T is defined as year t minus the calendar year in which an individual participant i underwent a medical examination year C^i (2007, 2008, or 2009)
- $Time_t$: Time in years [0 (67 years), 1 (68 years), 2 (69 years),…, 9 (76 years)]
- $Prefrail_{iT}$: Dummy variable (1 for prefrail group, 0 otherwise)
- X_i: Independent variables

The analyses were performed using SAS Enterprise Guide (version 7.15; SAS Institute,) and STATA (version 15; Stata Corporation). A two-sided P-value <.05 was considered statistically significant.

Ethical Considerations

This study was exempted from review by the Institutional Review Board of Ajou University Health System (AJIRB-MED-EXP-20‐127) as the data used in this study were deidentified and secondary analysis was performed using public data. No compensation was provided to the participants.

Results

The study population included 123,795 (57.3%) women, 71,082 individuals with prefrailty (32.9%), and 21,010 individuals with frailty (9.7%) (Table 1). Individuals with greater frailty were more likely to be women (robust vs frail: 44.8% vs 68.6%) and medical aid recipients (robust vs frail: 1.8% vs 7.8%). Over the 10-years follow up, 20,189 (9.4%) individuals died. The survival rate to age 76 years was higher in the robust group than in the prefrail and frail groups (robust: 91.8%; pre-frail: 90.3%; frail: 85.2%) (Table S2 in Multimedia Appendix 1).

Figure 1 shows the cumulative health care costs over 10 years by frailty category in individuals aged 66. Individuals in the frail group incurred the highest cumulative total health care costs (US $28,281), followed by the prefrail (US $23,793), and robust groups (US $19,815; P<.001). Compared to the robust group, the frail group was associated with higher cumulative inpatient costs (US $11,189 vs US $16,627) and outpatient costs (US $8,625 vs US $11,654) for the subsequent 10 years.

XSL•FO
RenderX

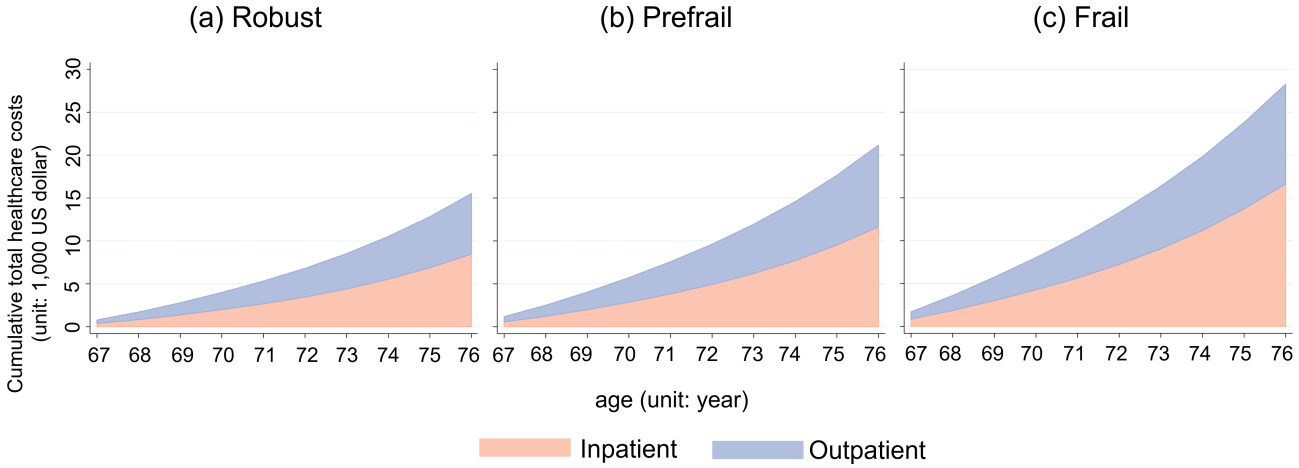
Table . Characteristics of Korean adults who participated in the National Screening Program for Transitional Ages at 66 years of age.

Characteristics	Total (N=215,887), n (%)	Frailty category[a], n (%)			*P* value[b]
		Robust (n=123,795)	Prefrail (n=71,082)	Frail (n=21,010)	
Gender					<.001
Men	100,774 (46.68)	68,390 (55.24)	25,780 (36.27)	6604 (31.43)	
Women	115,113 (53.32)	55,405 (44.76)	45,302 (63.73)	14,406 (68.57)	
Annual income (US $)					<.001
Quartile 1 (lowest)	46,306 (21.45)	26,944 (21.77)	14,525 (20.43)	4837 (23.02)	
Quartile 2	33,002 (15.29)	19,065 (15.40)	10,716 (15.08)	3221 (15.33)	
Quartile 3	55,973 (25.93)	32,026 (25.87)	18,624 (26.20)	5323 (25.34)	
Quartile 4 (highest)	80,606 (37.34)	45,760 (36.96)	27,217 (38.29)	7629 (36.31)	
Insurance status					<.001
Employee insurance	67,302 (31.17)	37,738 (30.48)	22,880 (32.19)	6684 (31.81)	
Self-employed insurance	141,955 (65.75)	83,769 (67.67)	45,492 (64.00)	12,694 (60.42)	
Medical aid for low income	6630 (3.07)	2288 (1.85)	2710 (3.81)	1632 (7.77)	
Residential area					<.001
Capital area	77,728 (36)	44,030 (35.57)	25,886 (36.42)	7812 (37.18)	
Metropolitan area	53,811 (24.93)	30,909 (24.97)	17,487 (24.60)	5415 (25.77)	
Rural area	84,348 (39.07)	48,856 (39.47)	27,709 (38.98)	7783 (37.04)	
Examination year					<.001
2007	53,907 (24.97)	29,875 (24.13)	18,297 (25.74)	5735 (27.30)	
2008	86,427 (40.03)	48,638 (39.29)	29,050 (40.87)	8739 (41.59)	
2009	75,553 (35.00)	45,282 (36.58)	23,735 (33.39)	6536 (31.11)	

[a]Frailty categories were defined as robust (frailty index <0.15, prefrail (0.15 to <0.25), and frail (≥0.25).

[b]P values were calculated using the χ^2 test for categorical variables.

Figure 1. Cumulative growth in total health care costs according to frailty index at age 66 years (A) Robust; (B) Prefrail; (C) Frail categories. The X-axis represents age (years) and the Y-axis represents cumulative total health care costs (US $). The pink area represents inpatient costs and blue area represents outpatient costs.



Over a period of 10 years, the annual growth in total health care costs, inpatient costs, and outpatient costs was greater in the frail group than in the prefrail and robust groups (Figure 2). Multivariable analyses (Table 2; Table S3 in Multimedia Appendix 1) showed that the frail group had higher mean total health care costs (difference: $827.2, SE: $20.0; *P*<.001), inpatient costs ($432.2, SE: $16.3; *P*<.001), and outpatient costs ($395.0, SE: $7.9; *P*<.001) than the robust group at baseline.

XSL•FO
RenderX

In the robust group, a one-year increase in age was associated with increased total health care costs (mean change per beneficiary per year: $206.2; SE: $1.2; *P*<.001), inpatient costs ($126.8; SE: $1.0; *P*<.001), and outpatient costs ($74.4; SE: $0.4; *P*<.001). In the frail group, there were greater increases in total health care costs each year compared to the robust group (difference in the mean cost per beneficiary per year: $120.9; SE: $5.3; *P*<.001), inpatient costs ($102.8; SE: $5.2; *P*<.001), and outpatient costs ($15.6; SE: $1.5; *P*<.001). The prefrail group had higher mean total health care costs, inpatient costs, and outpatient costs than the robust group at baseline, with the annual change per one-year increase in age falling between the robust and frail groups (Table 2).

Figure 2. Trends in annual health care costs per beneficiary per year according to frailty index at age 66 years. The X-axis represents age (years) and the Y-axis is annual health care costs per beneficiary per year (US dollar). The navy diamond line represents frail individuals, the blue triangle line represents pre-frail individuals, and the sky-blue square line represents robust individuals.

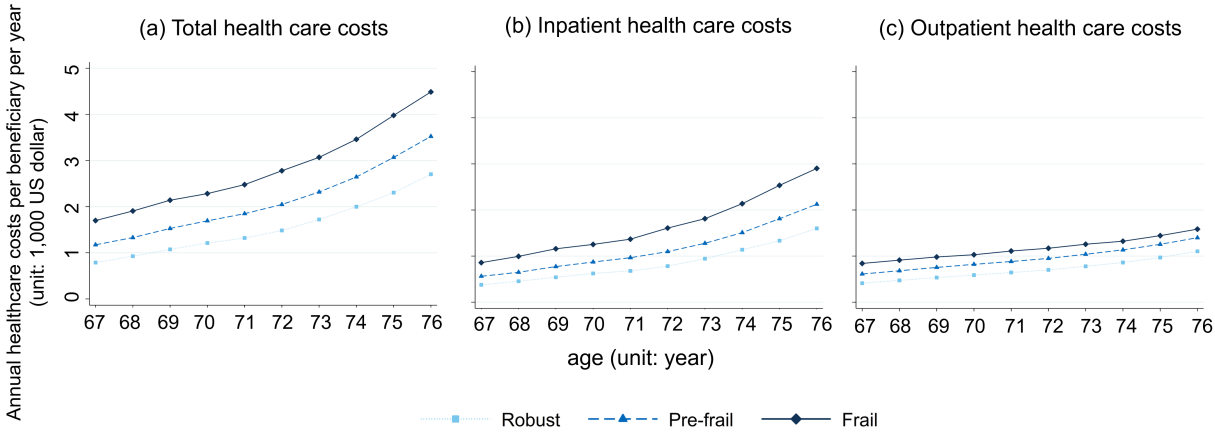


Table . Association between frailty at 66 years of age and annual health care costs and health care utilization over 10 years.[a] results of the generalized estimating equations models that examined the association between frailty at 66 years of age and annual health care costs and health care utilization over 10 years.

Characteristics	Total health care costs		Inpatient costs		Outpatient costs		Number of inpatient days		Number of outpatient visits	
	β (SE)[b]	*P* value[c]	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value	β (SE)	*P* value
Frailty category[d]										
Robust	Ref.[e]		Ref.		Ref.		Ref.		Ref.	
Pre-frail	354 (8.4)	<.001	169.4 (6.4)	<.001	180.6 (3.5)	<.001	1.7 (0)	<.001	16.7 (0)	<.001
Frail	827.2 (20)	<.001	432.2 (16.3)	<.001	395 (7.9)	<.001	5 (0)	<.001	36.0 (0.1)	<.001
Year (Ref: robust)										
Per 1-year increase	206.2 (1.2)	<.001	126.8 (1)	<.001	74.4 (0.4)	<.001	0.9 (0)	<.001	2.1 (0)	<.001
Year * Frailty category										
Year * Prefrail	52 (2.4)	<.001	38.7 (2.1)	<.001	13.1 (0.7)	<.001	0.3 (0)	<.001	0.1 (0)	<.001
Year * Frail	120.9 (5.3)	<.001	102.8 (5.2)	<.001	15.6 (1.5)	<.001	0.9 (0)	<.001	-0.2 (0)	<.001

[a]Generalized estimation equation models were used for the analysis. The models were adjusted for the examination year, sex, annual income, insurance status, and residential area.

[b]β (SE) represents the regression co-efficient (β) and its standard error (SE).

[c]P values indicate the significance level of the comparisons.

[d]Frailty categories were defined as robust (frailty index<0.15), pre-frail (0.15 to <0.25), and frail (≥0.25).

[e]Ref. denotes the reference category used for comparisons.

XSL•FO
RenderX

The examination of health care utilization over 10 years showed that the frail group had a greater increase in inpatient days and the outpatient visits over 10 years than the prefrail or robust groups (Figure S2 and Figure S3 in Multimedia Appendix 1). The frail group had a greater mean number of inpatient days (5 d; SE: 0 d; *P*<.001) and outpatient visits (36 visits; SE: 0.1; *P*<.001) than the robust group at baseline (Table 2). In the robust group, a one-year increase in age was associated with increases in inpatient days (mean change per beneficiary per year: 0.9 d; SE: 0 d; *P*<.001) and outpatient visits (mean change per beneficiary per year: 2.1 visits; SE: 0 visits; *P*<.001). In the frail group, the number of inpatient days increased more each year compared to the robust group (difference in the mean inpatient days per beneficiary per year: 0.9 d; SE: 0 d; *P*<.001), whereas the number of outpatient visits increased to a lesser degree (difference in the mean outpatient visits per beneficiary per year: −0.2 visits; SE: 0 visits; *P*<.001). The prefrail group had a greater mean number of inpatient days and outpatient visits than the robust group at baseline, with more increases in both inpatient days and outpatient visits annually (Table 2).

In subgroup analyses, men with frailty had a greater increase in total health care costs than women with frailty (*P*-for-interaction:<.001) (Table S4 in Multimedia Appendix 1). Frail people with low-income medical aid were more likely to incur higher total health care costs than those with other insurance types and frailty groups; however, the observed difference was not significant (*P*-for-interaction: .74) (Table S5 in Multimedia Appendix 1).

Discussion

Principal Findings

Using a nationwide Korean cohort, we found that the frailty level at 66 years of age was associated with higher health care costs and health care utilization over the subsequent 10 years. The growth in annual health care costs and health care utilization, particularly inpatient days, was greater in individuals with frailty than in prefrail or robust individuals. Furthermore, prefrail individuals, who made up 32.9% of the population, had higher health care costs and utilization than robust individuals. Given the high prevalence of prefrailty and associated long-term health care costs and utilization, our findings suggest the importance of identifying frailty and prefrailty to control health care costs and utilization in aging populations.

Our findings are consistent with previous studies demonstrating that frailty is associated with increased use of health care resources [24-26]. A population-based cohort study showed that the association between frailty onset and increasing self-reported health care costs was prominent in inpatient care and informal nursing care [6]. A meta-analysis of 7 cohorts of community-dwelling older adults found that health care costs of prefrail and frail older adults were higher than robust individuals [27]. In addition, those who were frail faced a greater risk of hospitalization, skilled nursing facility stays [28,29], emergency department visits [30], and institutionalization [31], compared to robust individuals. Our study contributes to the existing literature by examining 10-year trajectories of health care costs and utilization in a nationally representative cohort of Koreans aged 66 years. The choice of this time point in the beginning of elderhood removes the effect of chronologic age and emphasizes the importance of early identification, prevention, and management of frailty and prefrailty. Early identification can provide the opportunity to proactively address the needs of these individuals to lower health care costs and utilization in the future [32,33]. Frailty may be prevented or delayed by interventions such as physical activity, nutrition, and comprehensive geriatric assessment [34]. Therefore, by linking frailty with long-term health care costs and utilization, our findings expand the understanding of frailty's economic impact beyond short-term or disease-specific analyses commonly found in the literature.

Previous research suggested that incorporating frailty into a diagnosis-based model, such as the hierarchical condition category method (which is used to predict Medicare health care costs), may improve the accuracy of cost projections [35]. However, most prediction models for health care costs used in Korea do not account for frailty [36]. Although health care cost prediction models have been developed using the National Screening Program for Transitional Ages database, they have been designed for the entire population rather than specifically for older adults [37]. Our findings suggest that frailty metrics could enhance existing health care cost prediction models, providing a more nuanced understanding of health care needs in aging populations.

In our study, the associations between frailty and health care costs and utilization were stronger in men than women. Li et al [38] and colleagues found that frailty or worsening frailty had a stronger association with increased hospitalization and outpatient costs in men than in women. However, other studies did not find an interaction between frailty and sex on catastrophic health expenditures [39] or health care utilization (outpatient visit, inpatient visit, and inpatient length of stay) [40]. There are several explanations for higher health care costs among frail men than among frail women in our study. Men are generally more likely to develop serious health conditions that can be costly to manage, such as coronary heart disease, cancer, cerebrovascular disease, emphysema, cirrhosis of the liver, kidney disease, and atherosclerosis [41]. In addition, men are more likely to engage in risky health behaviors such as smoking and drinking and may be less likely to seek health care services for health issues, leading to delayed treatment [42]. These findings suggest that taking into account the differences in health risks and needs between frail men and frail women may help to improve health outcomes and reduce costs in older adults.

Limitations

Our study had important limitations. First, we were unable to assess costs and utilization of noncovered services including the costs of outpatient prescription medications. Second, selection bias is possible due to nonparticipation of otherwise eligible individuals in the National Screening Program for Transitional Ages and deaths, which may affect the longitudinal cost trajectories. However, we previously reported no major differences in characteristics between nonparticipants and participants in the program [15]. Third, the association between frailty and health care costs and utilization may be subject to

XSL•FO
RenderX

the choice of frailty definitions. Both the deficit accumulation FI and frailty phenotype have been associated with increased health care costs in previous studies [8,43]. Fourth, causality may not be inferred from our observational data. Finally, our findings may not be generalizable to other countries with different health care systems and financing structures.

Conclusion

Our study demonstrates the potential utility of assessing frailty at 66 years of age to identify older adults who are more likely to incur high health care costs and utilize health care services in the subsequent 10 years. The long-term high health care costs and utilization associated with frailty and prefrailty call for public health approaches to prevent and manage frailty in aging populations.

Acknowledgments

This study used customized data provided by the Korean National Health Insurance Service and was supported by faculty research grant 6-2021-0112 from Yonsei University College of Medicine (JS), grant KSPM-2021-01 from the Korean Society for Preventive Medicine and the Korea Health Technology R&D Project through the Korea Health Industry Development Institute funded by grant HI21C1249 from the Ministry of Health & Welfare, Korea (JJ), and by Mid-Career Investigator Award K24AG073527 from the National Institute on Aging (DHK). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

An earlier version of this study was presented at the Gerontological Society of America Annual Scientific Meeting on November 6, 2022 in Indianapolis, Indiana, and a poster was presented at the American Geriatrics Society Annual Scientific Meeting on May 4, 2022 in Long Beach, California.

Data Availability

The datasets generated and analyzed during this study are not publicly available. We obtained permission to use the datasets for the present study.

Authors' Contributions

Conceptualization: JJ, JS, DHK
Data curation: JJ, JS, DHK, AK
Formal analysis: JJ, DHK, JS, AK
Project administration: JJ, JS, DHK
Resources: JJ, JS, DHK
Supervision: JS, DHK
Writing – original draft: JJ, JS, DHK
Writing – review and editing: JJ, JS, DHK, AK

Conflicts of Interest

JJ reported receiving grant funding from the Korea Health Industry Development Institute during the conduct of the study. JS reported receiving grant funding from Yonsei University College of Medicine and the Korean Society for Preventive Medicine during the conduct of the study and being the chief executive officer and founder of Evertri, a company unrelated to the work that was submitted. DHK reported receiving grant funding from the National Institutes of Health during the conduct of the study and personal fees from Alosa Health and Village MD outside the submitted work. Generative AI was not used in any portion of the manuscript writing. No other disclosures were reported.

Multimedia Appendix 1
Analysis of Frailty index, healthcare utilization trends, and cost variations over time.
[DOCX File, 229 KB - publichealth_v11i1e50026_app1.docx]

References

1. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 8;388(10053):1545-1602. [doi: 10.1016/S0140-6736(16)31678-6] [Medline: 27733282]
2. Osborn R, Doty MM, Moulds D, Sarnak DO, Shah A. Older Americans were sicker and faced more financial barriers to health care than counterparts in other countries. Health Aff (Millwood) 2017 Dec;36(12):2123-2132. [doi: 10.1377/hlthaff.2017.1048]

XSL•FO
RenderX

3. World social report 2023: Leaving no one behind In an ageing world. United Nations. 2023. URL: https://www.un.org/development/desa/dspd/world-social-report/2023-2.html
4. National health insurance statistical yearbook 2021. Health Insurance Review and Assessment service. 2023.
5. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 2;381(9868):752-762. [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
6. Hajek A, Bock JO, Saum KU, et al. Frailty and healthcare costs-longitudinal results of a prospective cohort study. Age Ageing 2018 Mar 1;47(2):233-241. [doi: 10.1093/ageing/afx157] [Medline: 29036424]
7. Hoogendijk EO, Afilalo J, Ensrud KE, Kowal P, Onder G, Fried LP. Frailty: implications for clinical practice and public health. Lancet 2019 Oct 12;394(10206):1365-1375. [doi: 10.1016/S0140-6736(19)31786-6] [Medline: 31609228]
8. Sirven N, Rapp T. The cost of frailty in France. Eur J Health Econ 2017 Mar;18(2):243-253. [doi: 10.1007/s10198-016-0772-7] [Medline: 26914932]
9. Bock JO, König HH, Brenner H, et al. Associations of frailty with health care costs--results of the ESTHER cohort study. BMC Health Serv Res 2016 Apr 14;16(1):128. [doi: 10.1186/s12913-016-1360-3] [Medline: 27074800]
10. Shi SM, Olivieri-Mui B, McCarthy EP, Kim DH. Changes in a frailty index and association with mortality. J Am Geriatr Soc 2021 Apr;69(4):1057-1062. [doi: 10.1111/jgs.17002] [Medline: 33377190]
11. Figueroa JF, Joynt Maddox KE, Beaulieu N, Wild RC, Jha AK. Concentration of potentially preventable spending among high-cost medicare subpopulations: an observational study. Ann Intern Med 2017 Nov 21;167(10):706-713. [doi: 10.7326/M17-0767] [Medline: 29049488]
12. Mitnitski AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. ScientificWorldJournal 2001 Aug 8;1:323-336. [doi: 10.1100/tsw.2001.58] [Medline: 12806071]
13. Kaskirbayeva D, West R, Jaafari H, et al. Progression of frailty as measured by a cumulative deficit index: a systematic review. Ageing Res Rev 2023 Feb;84:101789. [doi: 10.1016/j.arr.2022.101789] [Medline: 36396032]
14. Kim HS, Shin DW, Lee WC, Kim YT, Cho B. National screening program for transitional ages in Korea: a new screening for strengthening primary prevention and follow-up care. J Korean Med Sci 2012 May;27 Suppl(Suppl):S70-S75. [doi: 10.3346/jkms.2012.27.S.S70] [Medline: 22661875]
15. Jang J, Jung H, Shin J, Kim DH. Assessment of frailty index at 66 years of age and association with age-related diseases, disability, and death over 10 years in Korea. JAMA Netw Open 2023 Mar 1;6(3):e2248995. [doi: 10.1001/jamanetworkopen.2022.48995] [Medline: 36862415]
16. Lee J, Lee JS, Park SH, Shin SA, Kim K. Cohort profile: The National Health Insurance Service-National Sample Cohort (NHIS-NSC), South Korea. Int J Epidemiol 2017 Apr 1;46(2):e15-e15. [doi: 10.1093/ije/dyv319] [Medline: 26822938]
17. Searle SD, Mitnitski A, Gahbauer EA, Gill TM, Rockwood K. A standard procedure for creating a frailty index. BMC Geriatr 2008 Sep 30;8(1):1-10. [doi: 10.1186/1471-2318-8-24] [Medline: 18826625]
18. Rockwood K, Mitnitski A. Frailty in relation to the accumulation of deficits. J Gerontol A Biol Sci Med Sci 2007 Jul;62(7):722-727. [doi: 10.1093/gerona/62.7.722] [Medline: 17634318]
19. Shi SM, McCarthy EP, Mitchell S, Kim DH. Changes in predictive performance of a frailty index with availability of clinical domains. J Am Geriatr Soc 2020 Aug;68(8):1771-1777. [doi: 10.1111/jgs.16436] [Medline: 32274807]
20. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. Biometrics 1988 Dec;44(4):1049-1060. [Medline: 3233245]
21. Hanley JA, Negassa A, Edwardes MDD, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. Am J Epidemiol 2003 Feb 15;157(4):364-375. [doi: 10.1093/aje/kwf215] [Medline: 12578807]
22. Fleiss JL. The statistical basis of meta-analysis. Stat Methods Med Res 1993;2(2):121-145. [doi: 10.1177/096228029300200202] [Medline: 8261254]
23. Rothman K. Spreadsheets for the analysis of epidemiologic data. RTI Health Solutions. 2002. URL: https://www.drugepi.org/dope/software#Episheet
24. Petermann-Rocha F, Lyall DM, Gray SR, et al. Associations between physical frailty and dementia incidence: a prospective study from UK Biobank. Lancet Healthy Longev 2020 Nov;1(2):e58-e68. [doi: 10.1016/S2666-7568(20)30007-6] [Medline: 36094146]
25. Veronese N, Cereda E, Stubbs B, et al. Risk of cardiovascular disease morbidity and mortality in frail and pre-frail older adults: results from a meta-analysis and exploratory meta-regression analysis. Ageing Res Rev 2017 May;35:63-73. [doi: 10.1016/j.arr.2017.01.003] [Medline: 28143778]
26. Cheng MH, Chang SF. Frailty as a risk factor for falls among community dwelling people: evidence from a meta-analysis. J Nurs Scholarsh 2017 Sep;49(5):529-536. [doi: 10.1111/jnu.12322] [Medline: 28755453]
27. Chi J, Chen F, Zhang J, et al. Impacts of frailty on health care costs among community-dwelling older adults: A meta-analysis of cohort studies. Arch Gerontol Geriatr 2021;94:104344. [doi: 10.1016/j.archger.2021.104344] [Medline: 33516075]
28. Ensrud KE, Kats AM, Schousboe JT, et al. Frailty phenotype and healthcare costs and utilization in older women. J Am Geriatr Soc 2018 Jul;66(7):1276-1283. [doi: 10.1111/jgs.15381] [Medline: 29684237]
29. Ensrud KE, Kats AM, Schousboe JT, et al. Frailty phenotype and healthcare costs and utilization in older men. J Am Geriatr Soc 2020 Sep;68(9):2034-2042. [doi: 10.1111/jgs.16522]

XSL•FO
RenderX

30. Kojima G. Frailty as a predictor of emergency department utilization among community-dwelling older people: a systematic review and meta-analysis. J Am Med Dir Assoc 2019 Jan;20(1):103-105. [doi: 10.1016/j.jamda.2018.10.004] [Medline: 30470576]
31. Kojima G. Frailty as a predictor of nursing home placement among community-dwelling older adults: a systematic review and meta-analysis. J Geriatr Phys Ther 2018;41(1):42-48. [doi: 10.1519/JPT.0000000000000097] [Medline: 27341327]
32. Holroyd-Leduc J, Resin J, Ashley L, et al. Giving voice to older adults living with frailty and their family caregivers: engagement of older adults living with frailty in research, health care decision making, and in health policy. Res Involv Engagem 2016;2(1):23. [doi: 10.1186/s40900-016-0038-7] [Medline: 29062523]
33. Abbasi M, Rolfson D, Khera AS, Dabravolskaj J, Dent E, Xia L. Identification and management of frailty in the primary care setting. CMAJ 2018 Sep 24;190(38):E1134-E1140. [doi: 10.1503/cmaj.171509] [Medline: 30249759]
34. Walston J, Buta B, Xue QL. Frailty screening and interventions: considerations for clinical practice. Clin Geriatr Med 2018 Feb;34(1):25-38. [doi: 10.1016/j.cger.2017.09.004] [Medline: 29129215]
35. Johnston KJ, Wen H, Joynt Maddox KE. Relationship of a claims-based frailty index to annualized medicare costs: a cohort study. Ann Intern Med 2020 Apr 21;172(8):533-540. [doi: 10.7326/M19-3261] [Medline: 32252070]
36. Lee S, Moon Y. A study of the financial projection of health expenditures of the aged of national health insurance-focused on the healthy ageing of EU. J Crit Soc Pol 2018 Feb 28;58:53-93. [doi: 10.47042/ACSW.2018.02.58.53]
37. Choi Y, An J, Ryu S, Kim J. Development and evaluation of machine learning-based high-cost prediction model using health check-up data by the national health insurance service of Korea. IJERPH 2022;19(20):13672. [doi: 10.3390/ijerph192013672]
38. Li CY, Al Snih S, Chou LN, et al. Frailty transitions predict healthcare use and Medicare payments in older Mexican Americans: a longitudinal cohort study. BMC Geriatr 2020 Jun 1;20(1):189. [doi: 10.1186/s12877-020-01583-y] [Medline: 32487037]
39. Fan L, Hou XY, Liu Y, Chen S, Wang Q, Du W. Catastrophic health expenditure associated with frailty in community-dwelling Chinese older adults: a prospective cohort analysis. Front Public Health 2021;9:718910. [doi: 10.3389/fpubh.2021.718910] [Medline: 34568260]
40. Fan L, Tian Y, Wang J, et al. Frailty predicts increased health care utilization among community-dwelling older adults: a longitudinal study in China. J Am Med Dir Assoc 2021 Sep;22(9):1819-1824. [doi: 10.1016/j.jamda.2021.01.082] [Medline: 33662331]
41. Case A, Paxson C. Sex differences in morbidity and mortality. Demography 2005 May;42(2):189-214. [doi: 10.1353/dem.2005.0011] [Medline: 15986983]
42. Harris CR, Jenkins M. Gender differences in risk assessment: Why do women take fewer risks than men? Judgm decis mak 2006 Jul;1(1):48-63. [doi: 10.1017/S1930297500000346]
43. Comans TA, Peel NM, Hubbard RE, Mulligan AD, Gray LC, Scuffham PA. The increase in healthcare costs associated with frailty in older people discharged to a post-acute transition care program. Age Ageing 2016 Mar;45(2):317-320. [doi: 10.1093/ageing/afv196] [Medline: 26769469]

Abbreviations

KRW: Korean Won

Edited by A Mavragani; submitted 16.06.23; peer-reviewed by LK Chen, P Heyn; revised version received 29.11.24; accepted 29.11.24; published 27.01.25.

Please cite as:
Jang J, Kim A, Choi M, McCarthy EP, Olivieri-Mui B, Park CM, Kim JH, Shin J, Kim DH
Association of Frailty Index at 66 Years of Age with Health Care Costs and Utilization Over 10 Years in Korea: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e50026
URL: https://publichealth.jmir.org/2025/1/e50026
doi:10.2196/50026

© Jieun Jang, Anna Kim, Minjee Choi, Ellen P McCarthy, Brianne Olivieri-Mui, Chan Mi Park, Jae-Hyun Kim, Jaeyong Shin, Dae Hyun Kim. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of the National Essential Public Health Service Package on Blood Pressure Control in Chinese People With Hypertension: Retrospective Population-Based Longitudinal Study

Feiran Wei[1*], MD; You Ge[1,2*], MD; Han Li[2], MPH; Yuan Liu[3], MSc

[1]School of Public Health, Southeast University, Nanjing, China

[2]Department of Infectious Disease, The Second Hospital of Nanjing, Nanjing, China

[3]Institute for Chronic Disease Management, Jining No. 1 People's Hospital, 0802 Huoju, Jining, China

[*]these authors contributed equally

Corresponding Author:
Yuan Liu, MSc
Institute for Chronic Disease Management, Jining No. 1 People's Hospital, 0802 Huoju, Jining, China

Abstract

Background: The National Essential Public Health Service Package (NEPHSP) was launched in 2009 to tackle poor blood pressure control in Chinese people with hypertension; however, it's effect is still unclear.

Objective: In a retrospective population-based longitudinal study, we aimed to evaluate effect of the NEPHSP on blood pressure control.

Methods: A total of 516,777 patients registered in the NEPHSP were included. The blood pressure control data were assessed based on the Residence Health Record System dataset. We longitudinally evaluated the effects of the NEPHSP on blood pressure control by analyzing changes in blood pressure at quarterly follow-ups. Both the degree and trend of the blood pressure changes were analyzed. We conducted stratified analysis to explore effects of the NEPHSP on blood pressure control among subgroups of participants with specific characteristics.

Results: The mean baseline systolic blood pressure (SBP) and diastolic blood pressure (DBP) were 147.12 (SD 19.88) mm Hg and 85.11 (SD 11.79) mm Hg, respectively. The control rates of baseline SBP and DBP were 39.79% (205,630/516,777) and 69.21% (357,685/516,777). Compared to baseline, the mean SBP decreased in each quarter by 5.06 mm Hg (95% CI −5.11 to −5.00; *P*<.001), 6.69 mm Hg (95% CI; −6.74 to −6.63; *P*<.001), 10.30 mm Hg (95% CI −10.34 to −10.23; *P*<.001), and 6.63 mm Hg (95% CI −6.68 to −6.57; *P*<.001). The SBP control rates increased in each quarter to 53.12% (274,493/516,777; β coefficient=0.60, 95% CI 0.59-0.61; *P*<.001), 56.61% (292,537/516,777; β coefficient=0.76, 95% CI 0.75-0.77; *P*<.001), 63.4% (327,648/516,777; β coefficient=1.08, 95% CI 1.07-1.09; *P*<.001), and 55.09% (284,711/516,777; β coefficient=0.69, 95% CI 0.68-0.70; *P*<.001). Compared to baseline, the mean DBP decreased in each quarter by 1.75 mm Hg (95% CI −1.79 to −1.72; *P*<.001), 2.64 mm Hg (95% CI −2.68 to −2.61; *P*<.001), 4.20 mm Hg (95% CI −4.23 to −4.16; *P*<.001), and 2.64 mm Hg (95% CI −2.68 to −2.61; *P*<.001). DBP control rates increased in each quarter to 78.11% (403,641/516,777; β coefficient=0.52, 95% CI 0.51-0.53; *P*<.001), 80.32% (415,062/516,777; β coefficient=0.67, 95% CI 0.66-0.68; *P*<.001), 83.17% (429,829/516,777; β coefficient=0.89, 95% CI 0.88-0.90; *P*<.001), and 79.47% (410,662/516,777; β coefficient=0.61, 95% CI 0.60-0.62; *P*<.001). The older age group had a larger decrease in their mean SBP (β coefficient=0.87, 95% CI 0.85-0.90; *P*<.001) and a larger increase in SBP control rates (β coefficient=0.054, 95% CI 0.051-0.058; *P*<.001). The participants with cardiovascular disease (CVD) had a smaller decrease in their mean SBP (β coefficient=−0.38, 95% CI −0.41 to −0.35; *P*<.001) and smaller increase in SBP control rates (β coefficient=−0.041, 95% CI −0.045 to −0.037; *P*<.001) compared to the blood pressure of participants without CVD.

Conclusions: The NEPHSP was effective in improving blood pressure control of Chinese people with hypertension. Blood pressure control of older individuals and those with CVD need to be intensified.

(JMIR Public Health Surveill 2025;11:e65783) doi:10.2196/65783

KEYWORDS

hypertension; primary care; public health; blood pressure control; cardiovascular disease

XSL•FO
RenderX

Introduction

The Global Burden of Disease Study reported that cardiovascular diseases (CVDs) are the leading noncommunicable disease contributing to the increase in disease-adjusted life years in China [1-3]. Hypertension, the major modifiable risk factor for CVDs, is prevalent in China [4-12]. According to the data of a recent survey conducted nationwide, the prevalence of hypertension was 54.7% among Chinese adults over 18 years old, and reportedly, merely 60.1% of Chinese adults with hypertension accepted antihypertensive treatments [13]. The low hypertension treatment rate has led to a suboptimal control of blood pressure among Chinese adults with hypertension [13]. Merely 7.2% of Chinese adults with hypertension had their blood pressures under control [13].

To tackle the poor hypertension control and growing burden of CVDs, the Chinese government launched the National Essential Public Health Service Package (NEPHSP) in 2009 [14]. The NEPHSP was a set of public health services which were available for all community-dwelling residents [14-18]. The public health services contain 4 modules, which are screening, monitoring, regular follow-up, and individualized interventions for hypertension control. The 4 modules were provided by primary health care professionals [14-18]. NEPHSP registration has been continuously expanded since it was launched in 2009 [14,18]. A total of 109 million Chinese adults diagnosed with hypertension registered with the NEPHSP from 2009 to 2019 [14]. The government investment increased from 15 to 84 RMB (US $2.05 to US $11.46) per person annually [15,18]. Because of the registration expansion and investment increase for the NEPHSP, the proportion of Chinese adult with hypertension who accepted antihypertensive treatment increased by 59.58% from 2009 to 2019 [14,18].

Although the registration expansion and investment increase for the NEPHSP improved the hypertension treatment rate, the NEPHSP lacks quality-oriented evaluations, which leads to an uncertainty whether the NEPHSP is effective for blood pressure control [14-18]. Furthermore, research demonstrating the effectiveness of the NEPHSP for blood pressure control in Chinese adults with hypertension is lacking [14-18]. To fulfill the research gap, this study aimed to explore effect of the NEPHSP on blood pressures control of Chinese adult with hypertension.

Methods

Study Design

This study was designed as longitudinal study to evaluate effect of the NEPHSP intervention on blood pressure control. The longitudinal study observed changes in the health status and clinical indicators of individuals who accepted treatment in a period of time, in order to explore the association between treatment and disease control [19-22]. Longitudinal studies have been commonly used in clinical research to evaluate the effects of treatment [19-22]. In this longitudinal study, we repetitively measured the blood pressure in participants who accepted hypertension care services offered by the NEPHSP in each quarter of 2023. All participants had a recorded blood pressure measurement prior to joining the NEPHSP in 2022. With the baseline outcomes in 2022 as the control, the degree and trend of the blood pressure changes after receiving NEPHSP in 2023 were analyzed longitudinally.

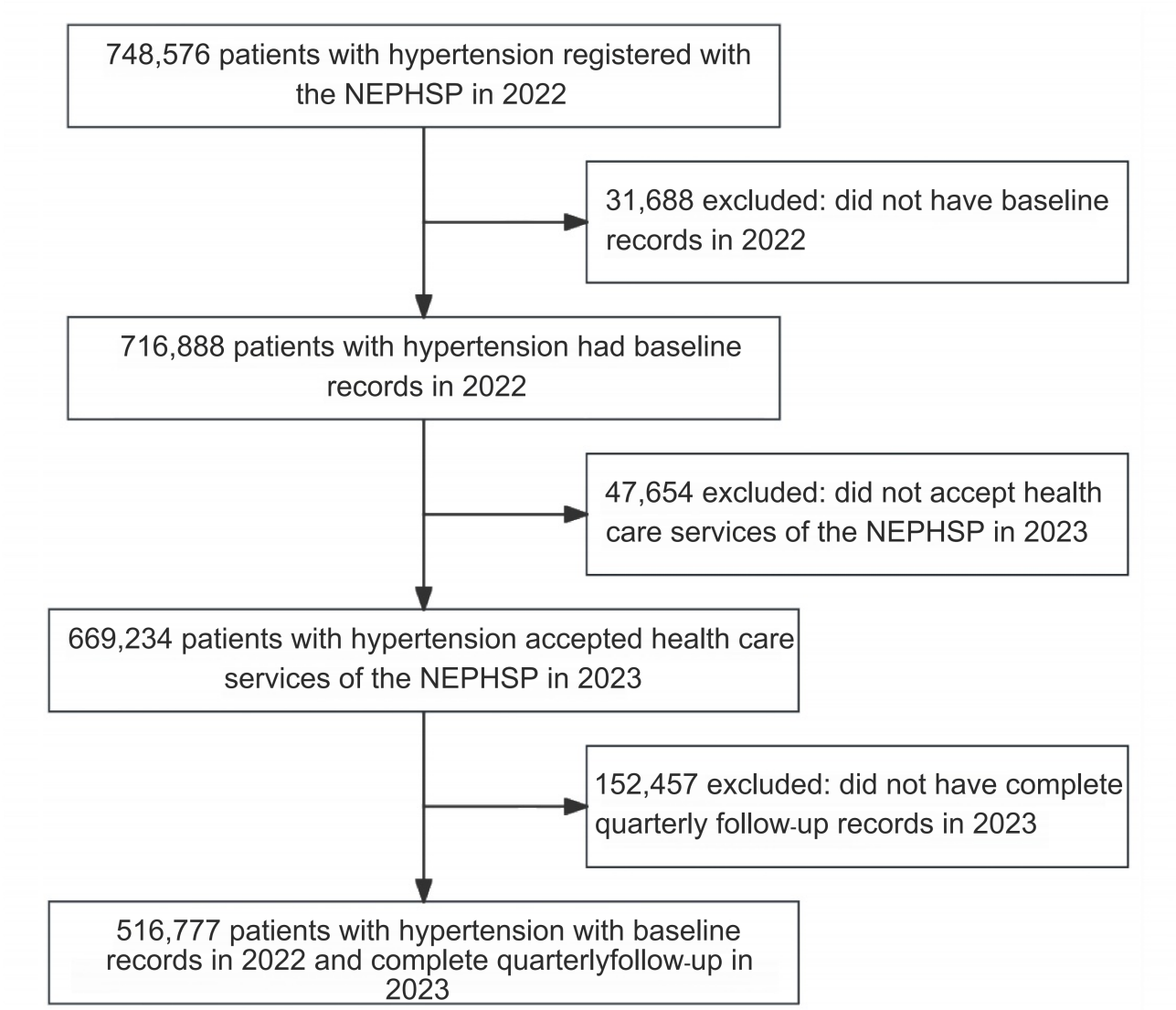
The data in this study were extracted from the Residence Health Record System, which is an electronic health record system built to recording demographic and health information of people registering with the NEPHSP [16,17]. This study used data of the Residence Health Record System built in Jining City, which is a national pilot city of the Chronic Disease Comprehensive Prevention and Control programs and is located in eastern China.

Study Participants

We analyzed data from the NEPHSP participants from 1593 primary care institutions in the 9 districts in Jining City. Eligibility criteria were (1) individuals registering with the NEPHSP as a patient with hypertension, (2) individuals who had baseline records of health examinations in 2022, (3) individuals who accepted health care services offered by the NEPHSP in 2023, and (4) individuals who had records for quarterly follow-ups conducted in 2023. The flowchart of participant inclusion is described in Figure 1.

XSL•FO
RenderX

Figure 1. The flowchart of the participant inclusion. NEPHSP: National Essential Public Health Service Package.

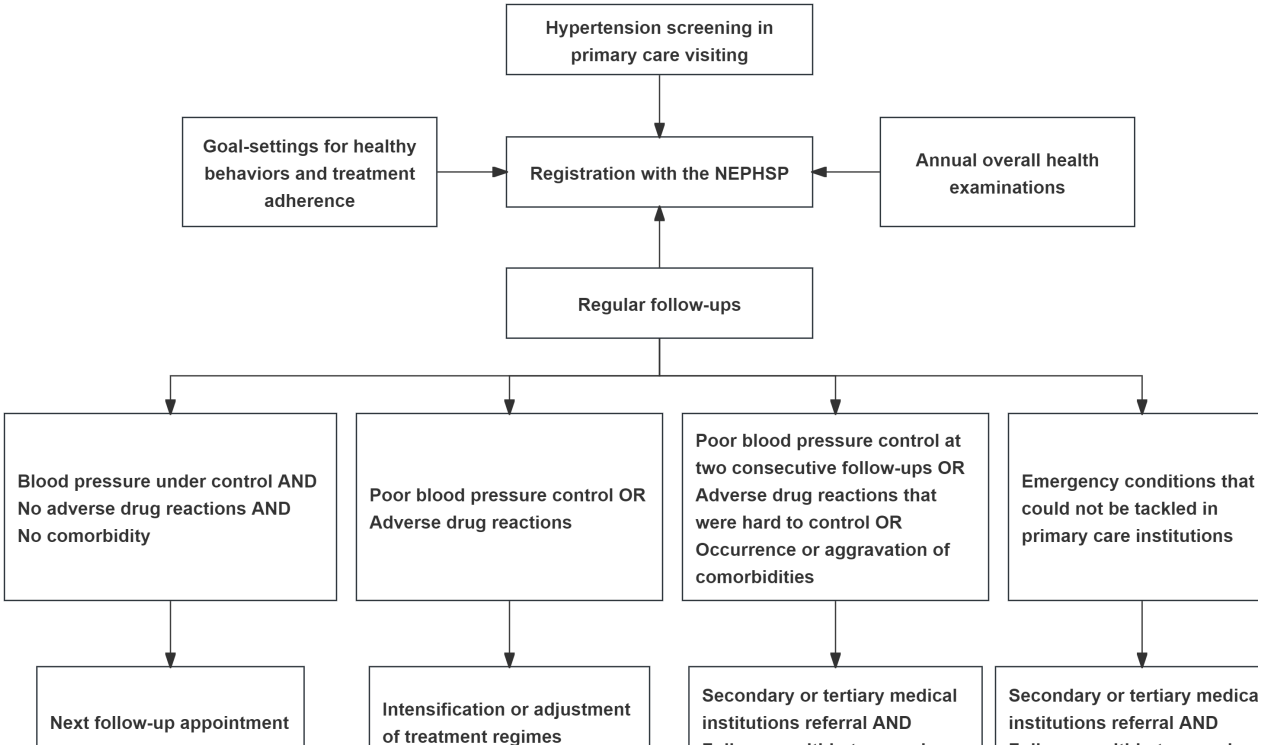


Intervention

The NEPHSP was divided into 4 dimensions: screening, routine follow-ups, individualized interventions, and annual health examinations. The explicit flowchart of the NEPHSP is described in Figure 2.

XSL•FO
RenderX

Figure 2. Flowchart of the NEPHSP. NEPHSP: National Essential Public Health Service Package.



The screening was conducted at primary care institutions when community residents had an appointment with primary health care professionals. The criteria for a hypertension diagnosis were a systolic blood pressure (SBP) over 140 mm Hg or a diastolic blood pressure (DBP) over 90 mm Hg for 3 separate measurements on different days. The residents were invited to register with the NEPHSP if they were diagnosed with hypertension.

The routine follow-up contained 5 facets: (1) assessing emergency conditions, (2) evaluating hypertensive symptoms, (3) measuring blood pressure, (4) evaluating healthy behaviors, and (5) evaluating medication treatment regimens and compliance. Assessing emergency conditions was the first part of the routine follow-up. The participants would be referred to secondary or tertiary care settings if there were emergency conditions that were hard to tackle in primary care institutions. If no emergency conditions were identified, the remaining parts would be conducted. The regular follow-ups were provided quarterly.

The individualized interventions contained 3 components: medication treatment adjustment, referral channel to superior medical institutions, and healthy behavior changes. According to health conditions that were evaluated in the regular follow-ups, primary care professionals would make decisions on medication treatment adjustments for the participants. If the health issues could not be handled in primary care settings, the primary care professionals would transfer the patient to superior medical institutions. The explicit processes and criteria of medication treatment adjustment and referral were divided into 3 dimensions. First, if blood pressure was under control, there were no adverse drug reactions, and there was not an occurrence of comorbidities or an aggravation of comorbidities, the next follow-up appointments were made for the participants. Second, if blood pressures were uncontrolled or adverse drug reactions were identified, the primary health care professionals would adjust the antihypertensive medication regimens. The conditions of the participants would be re-evaluated within 2 weeks. Third, if the patients had refractory uncontrolled blood pressure, adverse drug reactions that could not be tackled in primary care settings, or an occurrence and aggravation of comorbidities, the primary health care professionals would refer the participants to secondary or tertiary care settings. The conditions of the participants would be re-evaluated within 2 weeks.

Changes in health behavior were available for all participants, and goal-setting was used to formulate their health behaviors. The primary health care professionals developed shared goals with the participants. The targets for changes in health behavior included medication treatment compliance, dietary changes and weight control, control of smoking and alcohol drinking, and exercise engagement. The achievements of the health behavior changes were evaluated at the routine follow-ups. Annual health examinations were available for all patients. The health examinations contained overall physical and bioclinical examinations, as well as lifestyle and medication treatment evaluations.

Outcome Variables

We evaluated the impact of the NEPHSP on blood pressure control in Chinese patients with hypertension through changes in SBP and DBP and changes in the control rates of SBP and DBP. The control rates of SBP and DBP were the proportion of study participants whose SBP and DBP values were in the normal range, respectively. The normal range of SBP and DBP in the Chinese hypertension guideline is defined as <140 mm Hg and <90 mm Hg, respectively [23].

XSL•FO
RenderX

We used the mean blood pressure values at baseline in 2022 and at 4 quarterly follow-up visits in 2023 to depict trends in blood pressure for the patients before and after receiving the NEPHSP services. The baseline values of SBP and DBP were averaged by the blood pressure measurements of the health examinations in 2022. The mean follow-up values of SBP and DBP were averaged by blood pressure measurements of the routine follow-ups. We extracted the blood pressure data from the Residence Health Record System, which recorded the SBP and DBP values measured in the annual health examinations and the regular follow-ups of the NEPHSP participants.

Statistical Analysis

Baseline characteristics of the participants were statistically described. The quantitative variables were described as mean and SD, while categorical variables were described as counts and percentage. The generalized estimating equation (GEE) model was used to analyze the repetitively measured blood pressure data that contained a continuous time variable [24]. The GEE model was capable of analyzing numerical and categorical data that were repetitively measured [24]. Both the trend and degree of the blood pressure changes were analyzed. First, the variable of time was set as the continuous variable to analyze the trend of the blood pressure changes in the GEE models. Linear trend tests were conducted to examine whether the changes in the blood pressure values with the continuous time variation were linear. The linear trend test were reflected by *P* for trend [25,26]. If the changes in the blood pressure values for the continuous time variation were linear, we used the coefficients of GEE models to reflect linear trends of blood pressure changes. Second, we also set the variable of time as the categorical variable to analyze the degree of blood pressure changes in the GEE models. The baseline blood pressure values in 2022 were chosen as the reference group. All statistical analyses were carried out by R software version 4.0.4 (R Core Team). The analysis was performed with 2-tailed tests at an α level of .05.

The covariates were first selected based on evidence of previous studies that had proven the factors that could impact blood pressure [27-34]. Subsequently, we also conducted univariate analyses to confirm whether the covariates had an impact on blood pressure. The following participant characteristics were adjusted as covariates: age, sex, BMI, waist circumference, exercise engagement, alcohol drinking and smoking status, fasting plasma glucose, total cholesterol, triglycerides, low density lipoprotein, high density lipoprotein, baseline SBP and DBP values, diagnosis of cardiovascular and renal diseases, and duration since hypertension diagnosis. The results of univariate analyses and corresponding information of each covariate in the GEE models are described in Multimedia Appendix 1 and Multimedia Appendix 2.

Sensitivity and Stratified Analyses

A series of stratified and sensitivity analyses were conducted to identify whether the changes in blood pressure were different by subgroup. The subgroups were predefined as an older age group (≥65 years old) and younger age group (<65 years old), a male group and female group, a CVD diagnosis group and no CVD diagnosis group, and a baseline controlled SBP group and baseline uncontrolled SBP group. We fit GEE models to assess the changes in blood pressure separately for those subgroups. We then tested the interactive effects between the time variables (continuous time variable) and the subgroup indicators using the total sample, to formally evaluate whether the changes in blood pressure differed between the subgroups.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and approved by the Committee on Human Research of the Jining No. 1 People's Hospital (2023 Ethical Approval No. KYLL-204609‐176). The need for informed consent was waived due to the retrospective nature of this study.

Results

Baseline Characteristics

The baseline characteristics are accessible in Table 1. A total of 516,777 participants were involved in this study. The mean age was 68.59 (SD 9.57) years, 58.56% (n=302,109) of the participants were female, and 66.94% (n=345,923) of the participants were over 65 years old. The mean values for baseline SBP and DBP were 157.12 (SD 19.88) mm Hg and 85.11 (SD 11.79) mm Hg, respectively. The mean values of baseline SBP were above the normal criteria [19]. The control rates of baseline SBP and DBP were 39.79% (205,630/516,777) and 69.21% (357,685/516,777), respectively.

XSL•FO
RenderX

Table . Baseline characteristics.

Characteristics	Values
Age (years), mean (SD)	68.39 (9.57)
Sex, n/N (%)	
Male	214,668/516,777 (41.54)
Female	302,109/516,777 (58.56)
DBP[a] (mm Hg), mean (SD)	85.11 (11.79)
DBP control rate, n/N (%)	357685/516,777 (69.21)
Cardiovascular disease diagnosis, n/N (%)	124409/516,777 (24.07)
SBP[b] (mm Hg), mean (SD)	
All participants	147.12 (19.88)
Older participants	149.81 (20.37)
Younger participants	141.67 (17.64)
Participants with cardiovascular disease	149.14 (18.48)
Participants without cardiovascular disease	144.47 (20.02)
SBP control rate, n/N (%)	
All participants	205,630/516,777 (39.79)
Older participants	114,649/345,923 (33.14)
Younger participants	90,981/170,854 (53.25)
Participants with cardiovascular disease	48,552/124,409 (39.02)
Participants without cardiovascular disease	157,078/392,368 (40.03)

[a]DBP: diastolic blood pressure.

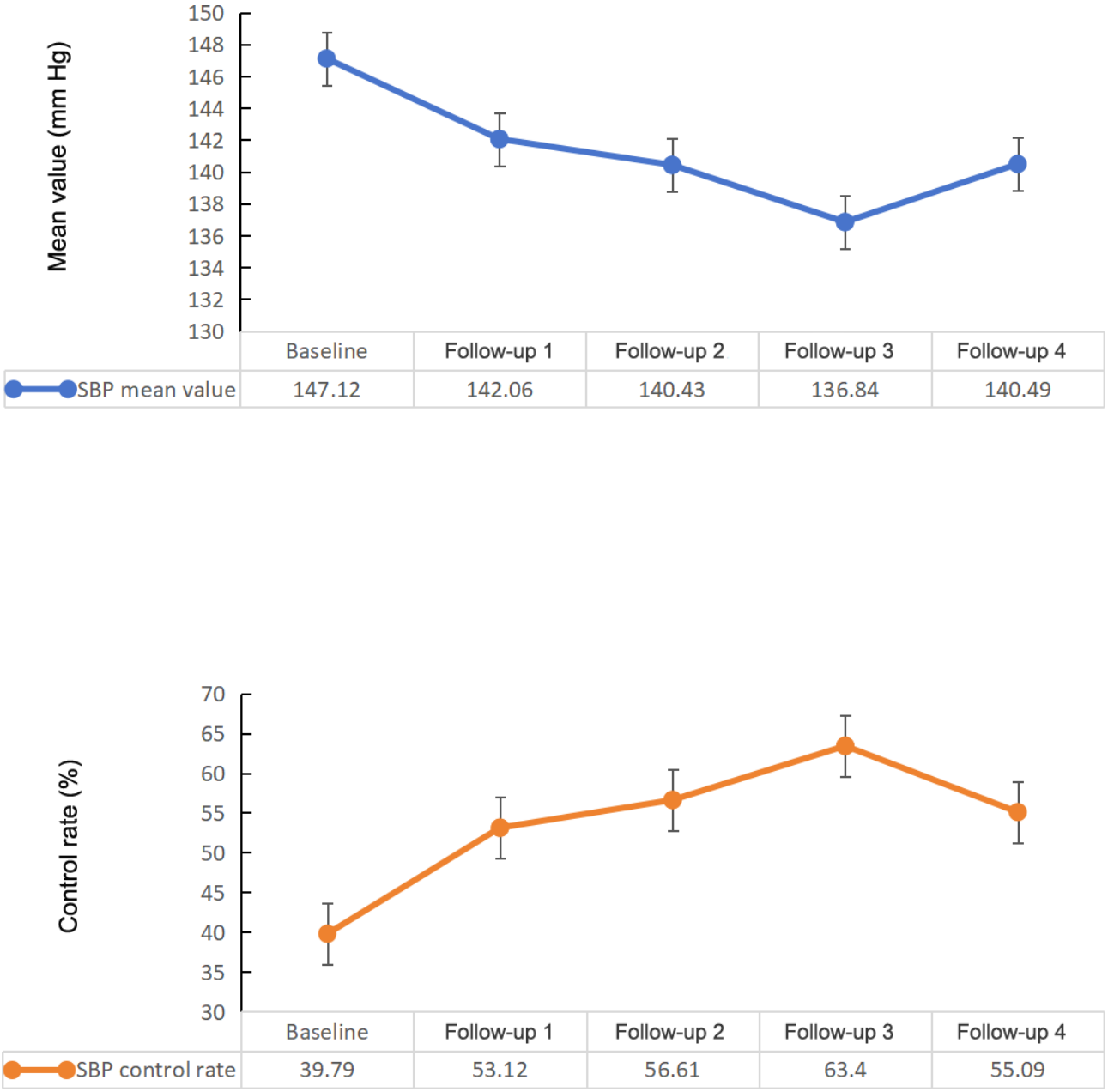
[b]SBP: systolic blood pressure.

Changes in SBP Control

The changes in SBP control are described in Figure 3. The SBP values had a linear change trend with continuous time variation (β coefficient=−1.85, 95% CI −1.86 to −1.93; *P*<.001).

XSL•FO
RenderX

Figure 3. Changes in SBP control. SBP: systolic blood pressure.



The SBP values showed a decreasing trend (Table 2, Figure 3). Compared to baseline, SBP decreased in each quarter by 5.06 mm Hg, 6.69 mm Hg, 10.60 mm Hg, and 6.63 mm Hg (Table 3).

Table . Slopes of the blood pressure changes from baseline to the fourth follow-up.

Variable	β coefficient (95% CI)	*P* value
SBP[a]	−1.85 (−1.86 to −1.93)	<.001
SBP control rate	0.185 (0.183-0.187)	<.001
DBP[b]	−0.77 (−0.78 to −0.76)	<.001
DBP control rate	0.169 (0.166-0.171)	<.001

[a]SBP: systolic blood pressure.

[b]DBP: diastolic blood pressure.

XSL•FO
RenderX

Table . Degree of change in systolic blood pressure (SBP) control.

Appointment	SBP		SBP control rate	
	β coefficient (95% CI)	*P* value	β coefficient (95% CI)	*P* value
Follow-up 1	−5.00 (−5.11 to −5.00)	<.001	0.18 (0.183-0.187)	<.001
Follow-up 2	−6.69 (−6.74 to −6.63)	<.001	0.60 (0.59-0.61)	<.001
Follow-up 3	−10.60 (−10.54 to −10.63)	<.001	0.76 (0.75-0.77)	<.001
Follow-up 4	−6.63 (−6.68 to −6.57)	<.001	1.08 (1.07-1.09)	<.001

SBP control rates for the NEPHSP participants showed an increasing trend (Figure 3, Table 2). Compared to baseline, SBP control rates increased in each quarter to 53.12% (274,493/516,777), 56.61% (292,537/516,777), 63.4% (327,648/516,777), and 55.09% (284,711/516,777; Table 3).

Changes in DBP Control

The changes in DBP control are described in Figure 4. The DBP values had a linear trend with continuous time variation (β coefficient=−1.85, 95% CI −1.86 to −1.93; *P*<.001). The DBP values of the NEPHSP participants showed a decreasing trend (Table 2, Figure 4). Compared to baseline DBP, DBP decreased in each quarter by 1.75 mm Hg, 2.64 mm Hg, 4.20 mm Hg, and 2.64 mm Hg (Table 4).

XSL•FO
RenderX

Figure 4. Changes in DBP control. DBP: diastolic blood pressure.

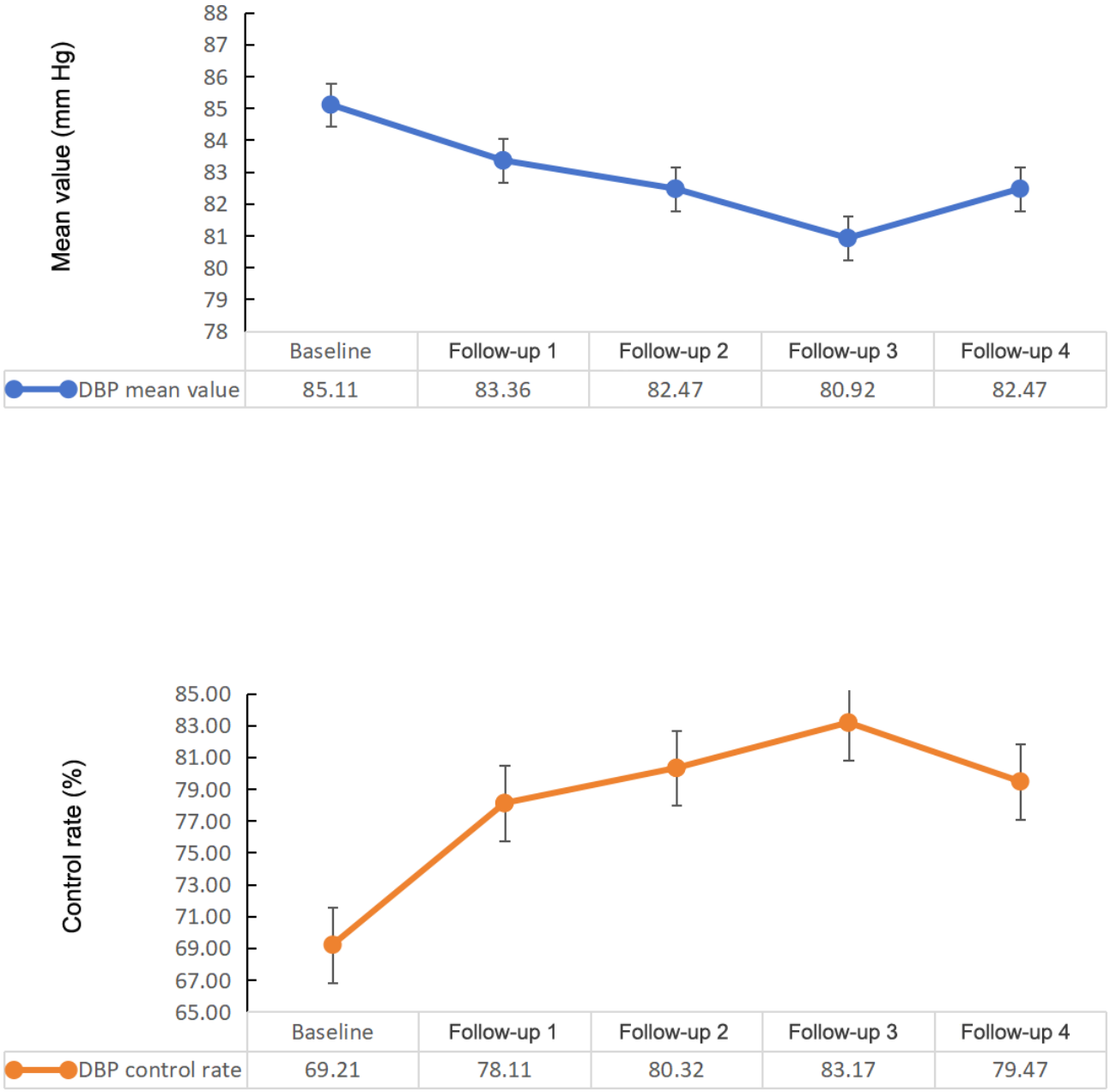


Table . Degree of change in diastolic blood pressure (DBP) control.

Appointment	DBP		DBP control rate	
	β coefficient (95% CI)	*P* value	β coefficient (95% CI)	*P* value
Follow-up 1	−1.75 (−1.79 to −1.72)	<.001	0.62 (0.61-0.63)	<.001
Follow-up 2	−2.64 (−2.68 to −2.61)	<.001	0.67 (0.66-0.68)	<.001
Follow-up 3	−4.20 (−4.63 to −4.16)	.01	0.89 (0.88-0.90)	<.001
Follow-up 4	−2.64 (−2.68 to −2.61)	<.001	0.61 (0.60-0.62)	<.001

DBP control rates for the NEPHSP participants showed an increasing trend (Table 2, Figure 4). Compared to baseline, DBP control rates increased in each quarter to 78.11% (403,641/516,777), 80.32% (415,062/516,777), 83.17% (429,829/516,777), and 79.47% (410,662/516,777; Table 4).

Sensitivity and Stratified Analyses

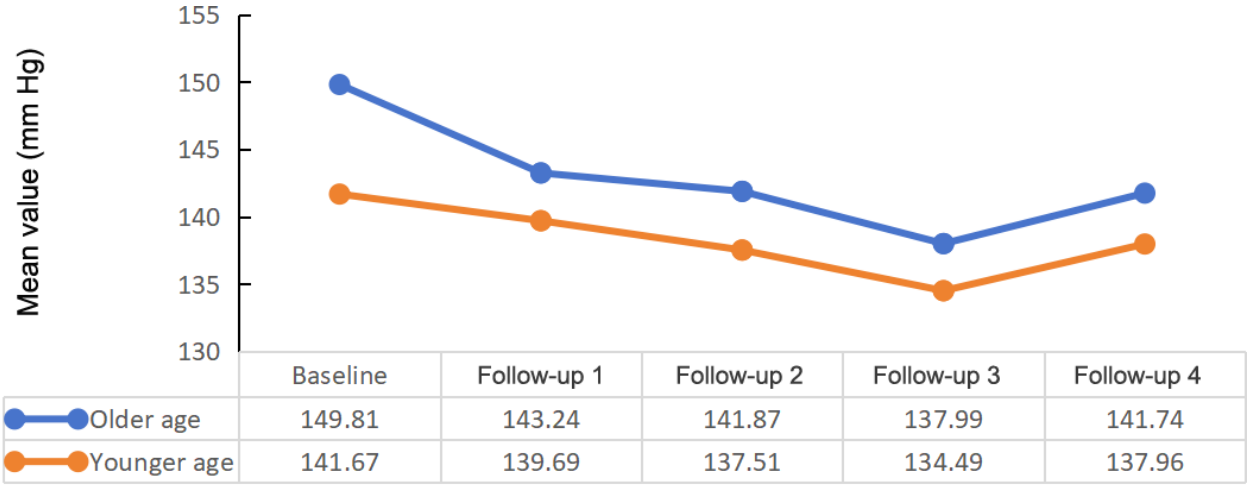
Our sensitivity analysis stratified by the predefined subgroups showed that a decrease in blood pressure and an increase in blood pressure control rates were consistent across subgroups of age, sex, and a diagnosis of baseline CVD. The parameters of the sensitivity analysis are accessible in Multimedia Appendix 3. The changes in SBP were converse between the baseline

XSL•FO
RenderX

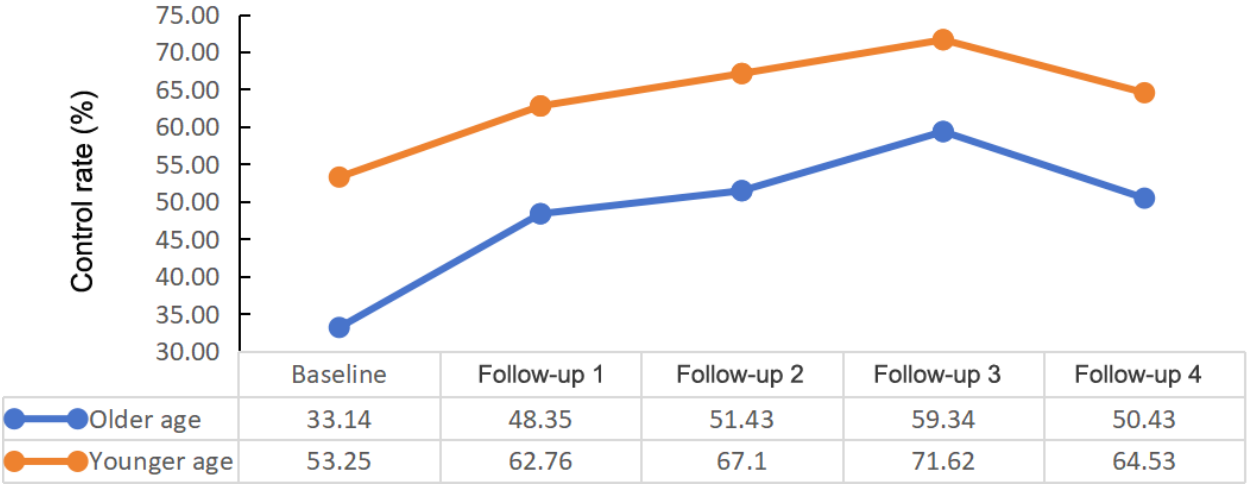
controlled SBP and uncontrolled SBP groups. The baseline controlled SBP group had an increased control of their SBP, while the baseline uncontrolled SBP group had a decreased control of their SBP. The SBP control rates decreased in the baseline controlled SBP group, while the SBP control rates increased in the baseline uncontrolled SBP group.

The changes in SBP control in the older and younger age groups are described in Figure 5. Compared with younger age group, the older age group showed a larger decrease in SBP and a larger increase in SBP control rates. The SBP of the older age group were higher than the younger age group. The SBP control rates of the older age group were significantly lower than the younger age group (P<.001).

Figure 5. Systolic blood pressure control of the age groups.



	Baseline	Follow-up 1	Follow-up 2	Follow-up 3	Follow-up 4
Older age	149.81	143.24	141.87	137.99	141.74
Younger age	141.67	139.69	137.51	134.49	137.96

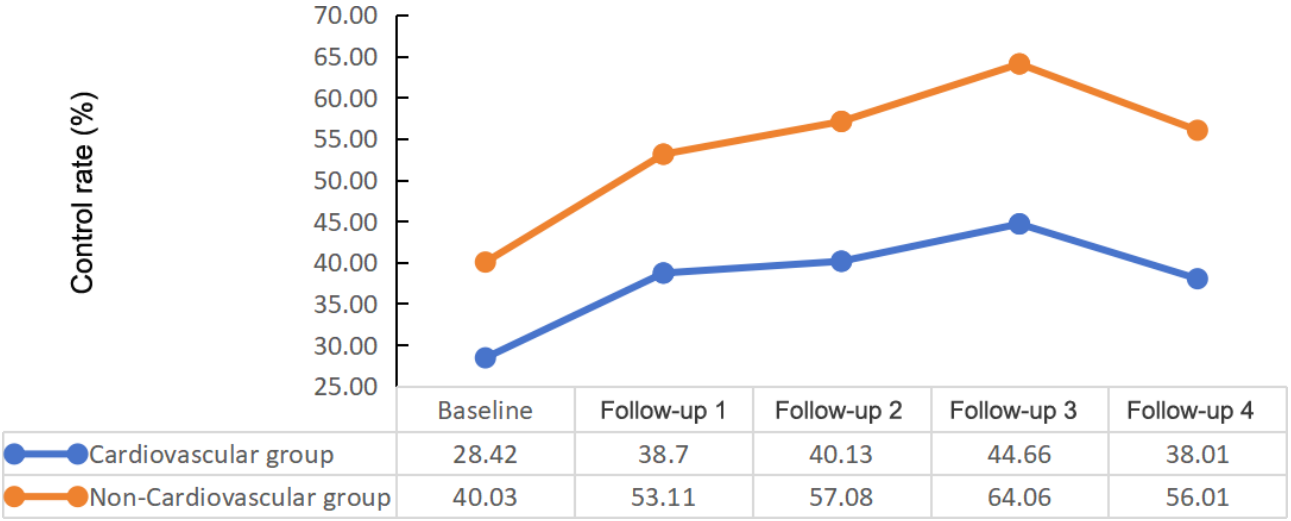


	Baseline	Follow-up 1	Follow-up 2	Follow-up 3	Follow-up 4
Older age	33.14	48.35	51.43	59.34	50.43
Younger age	53.25	62.76	67.1	71.62	64.53

The changes in SBP control of the participants with and without CVD were described in Figure 6. The participants with CVD had a smaller decrease in SBP and smaller increase in SBP control rates than the participants without CVD. The SBP control rates of the participants with CVD were significantly lower than the participants without CVD (P<.001).

XSL•FO
RenderX

Figure 6. Systolic blood pressure control of the cardiovascular disease group.



	Baseline	Follow-up 1	Follow-up 2	Follow-up 3	Follow-up 4
Cardiovascular group	28.42	38.7	40.13	44.66	38.01
Non-Cardiovascular group	40.03	53.11	57.08	64.06	56.01

The stratified analysis by sex demonstrated that the decrease in SBP values and the increase in SBP control rates were greater in the female group than in male group; however, the differences were relatively small.

Discussion

Principal Findings

To our knowledge, this is the first study that longitudinally explored effect of the NEPHSP on blood pressure control. The study findings demonstrated that the NEPHSP was effective for improving SBP and DBP control. The NEPHSP decreased the values of SBP and DBP and increased the control rates of SBP and DBP.

The routine follow-up that monitored blood pressure and the individualized interventions of the NEPHSP might contribute to the improvement of SBP and DBP control. The routine follow-ups generated clinically crucial data that could recognize prioritization of the participant subgroups who had poorly controlled blood pressures [35]. Previous studies also reported that the regular population review that enables the formulation of customized interventions for improving blood pressure control is an effective strategy for hypertension management [35-48]. Individualized interventions were formulated according to the evaluations of routine follow-ups to identify appropriate strategies to improve blood pressure control in participants with poorly controlled blood pressure [35,36,47].

Although the NEPHSP improved the control of blood pressure, the mean values of SBP were still slightly higher than the normal criteria. This might be caused by therapeutic inertia [33]. Therapeutic inertia was defined as a failure of health care professionals to initiate or intensify treatment appropriately during visits [33,49]. Previous research demonstrated that the failure to appropriately initiate or intensify treatment was introduced by the perspective of health care professionals that near-target blood pressures were acceptable and the overconcern about adverse effects caused by treatment intensification [47,48,50,51]. Because of the doctor-related factors leading to therapeutic inertia, participants with near-target blood pressures might not accept intensification of antihypertensive treatments, which could contribute to the slightly higher values of blood pressure [47,48,50,51]. Although therapeutic inertia is associated with poor blood pressure control and cardiovascular events, therapeutic inertia is common in hypertension management in primary care settings [47,48,51]. Given this situation, interventions for handling therapeutic inertia are important for improving hypertension management in primary care settings [33,47,48,51]. However, in terms of the NEPHSP, no component for tackling therapeutic inertia were embedded. The relevant components addressing therapeutic inertia can be considered for inclusion in the NEPHSP [33].

The subgroup analysis showed that the older participants had greater blood pressure improvement compared to the younger participants. This finding might be relevant to the referral mechanism of the NEPHSP. Compared with younger patients with hypertension, older patients with hypertension have more complex health conditions, such as frailty, comorbidities, and polypharmacy [34,52-58]. These complex health conditions can introduce challenges for blood pressure control in primary care settings, which are positively associated with poor blood pressure control [34,52-58]. To improve poor blood pressure control in older patients with complex health conditions in primary care settings, a referral mechanism to superior medical institutions is important [34,46,47,57-59]. In terms of the NEPHSP, the older participants with complex health conditions and poor blood pressure control were recognized via the regular follow-up evaluations. Based on the health condition evaluations of the regular follow-ups, the older participants could accept more referrals to superior medical institutions than the younger participants. Through the referral to superior medical institutions, the older participants could receive more intensive antihypertensive treatments than the treatments in primary care settings [34,46,47,57-59]. Because of the intensive antihypertensive treatments provided in superior medical institutions, the older participants had more blood pressure control improvement than the younger participants [34,46,47,57-59].

Although older participants had a greater improvement in blood pressure control than younger participants, the SBP of the older participants were still higher than the younger participants.

XSL•FO
RenderX

Moreover, the SBP control rates for the older participants were significantly lower than the younger participants. These study findings might be associated with concerns about adverse events caused by blood pressure–lowering pharmacotherapy [33,34,45-58]. However, previous robust research evidence demonstrated that decreasing the blood pressures of older adults to normal range is safe and effective for reducing risks of cardiovascular events [59,60]. Consequently, older people with hypertension should have a similar attitude toward blood pressure control as younger people with hypertension [59,60]. Given this situation, strategies to intensify blood pressure control in older participants engaging in the NEPHSP are necessary [59,60].

The improvement in the blood pressure control for the participants with CVD was less than the participants without CVD, while blood pressure control for people with hypertension and CVD needs to be more intensive than people with hypertension but without CVD [60]. Similar to the older participants, the patients with hypertension with CVD also required more referrals to superior medical institutions to achieve optimal blood pressure control [61-66]. Compared to patients with hypertension without CVD, the patients with CVD demand more intensive blood pressure control to prevent cardiovascular events, and blood pressure control in patients with hypertension and CVD is challenging in primary care settings [61-66]. The main challenges are potential adverse drug effects caused by polypharmacy and emergent cardiovascular events caused by intensive antihypertensive treatments [61-66]. Given this situation, referral to superior medical institutions is necessary for patients with hypertension and CVD to achieve optimal blood pressure control via intensive and safe antihypertensive treatments [61-66].

Although the NEPHSP had the algorithm to evaluate health conditions of the participants with comorbidities and subsequent referral mechanism to superior medical settings, the improvement of blood pressure control in the participants with CVD was still less than the participants without CVD. The lesser improvement in blood pressure control in the participants with CVD could be caused by the health evaluations of the regular follow-ups. The health condition evaluations might fail to accurately assess the health conditions of the participants with CVD, as the evaluations did not have systematic tools such as checklists, which could lead to failure to refer the participants who needed referral [61-66]. Furthermore, no explicit criteria for referral of participants with CVD were predefined, which could lead to confusion for the primary health care professionals when making the decisions on referral [61-66]. Given that the blood pressure control in people with hypertension and CVD needs to be more intensive than people without CVD, strategies for intensifying blood pressure control in the NEPHSP participants with CVD should be considered [61-66].

The subgroup analysis showed that the participants with controlled SBP at baseline had an increase in SBP and a decrease in SBP control rates, which were converse to the SBP results of all study participants. These results could be associated with the health evaluation algorithm of the regular follow-ups of the NEPHSP. According to the health evaluation algorithm, the participants who had controlled blood pressures that were evaluated in the regular follow-ups would not accept specific interventions. Because specific interventions were not provided, the participants with controlled SBP at baseline could have an increase in SBP and a decrease in SBP control rates. Although SBP of the participants was controlled at baseline, the increase of SBP values within the normal range could also increase CVD events [58]. Given this situation, the health evaluation algorithm and interventions of the NEPHSP for the participants with controlled SBP at baseline were suggested to be reformulated [58-61].

Blood lipid and glucose control could impact hypertension management. Previous studies reported that the triglyceride-glucose index (TyG), which is calculated by multiplying fasting triglyceride by fasting glucose, had a positive relationship with blood pressure control [31,32,61,62]. The increase of the TyG could contribute to poor blood pressures control [31,32,61,63]. A linear dose-response relationship between changes in the TyG and the change in blood pressure was identified [31,32,62,63]. To achieve optimal blood pressure control, blood lipid and glucose should be concurrently managed in people with hypertension [31,32,62,63]. However, in terms of the NEPHSP, no treatment targets for blood lipid and glucose were predefined for the study participants, which might cause ignorance toward blood lipid and glucose control. The ignorance of blood lipid and glucose control might cause a high TyG, which could contribute to the near-target blood pressure levels in this study [31,32,62,63]. Given this situation, further studies are required to confirm whether blood lipid and glucose control impact blood pressure control of the NEPHSP participants, in order to clarify the target for improving blood pressure control in this population.

Limitations

This study had several limitations. Although the effects of the NEPHSP on blood pressure control were modest, this study could not confirm the factors impacting blood pressure control of the NEPHSP. The group divisions for subgroup analysis were general and not explicit. Further studies that include a gradient group division to explore factors impacting hypertension control of NEPHSP are needed. Randomization for group division was not applied in the subgroup analysis, which could lead to an uneven distribution of confounding factors. Further studies that use randomization or matched group division are need to confirm the factors impacting hypertension control of the NEPHSP. This study explored baseline characteristics that could impact hypertension control of the NEPHSP, and future studies exploring the impact of the factor trajectories on hypertension control are needed. In addition, this study evaluated NEPHSP in a pilot city, and further nationwide studies are needed.

The blood pressure data in this study were the clinical records obtained from routine blood pressure measurements that were conducted via the Riva-Rocci and Korotkoff techniques [64-66]. The Riva-Rocci and Korotkoff techniques are still considered a cornerstone in evaluation of blood pressure levels in clinical trials [64-66]. However, the techniques have limitations, including that they can provide only a limited number of blood pressure values and can not reflect blood pressure fluctuations over 24 hours [64-66]. As a result, potential influences of blood

XSL•FO
RenderX

pressure fluctuations over 24 hours in our study findings were uncertain [64-66]. Further studies that use advanced devices such as ambulatory blood pressure monitoring to explore impacts of 24-hour blood pressure fluctuations on effects of the NEPHSP on blood pressure control are required. Despite the uncertainty of impacts of 24-hour blood pressure fluctuations, our study findings also provide preliminary evidence for the control of blood pressure by the NEPHSP. First, our study was conducted at population level. The large scale of the study participants might reduce the impact of 24-hour blood pressure fluctuations on the changes in blood pressure levels to some degree. Second, the follow-up time was one year for this study, which could offset the impact of blood pressure fluctuations over 24 hours.

This study was designed as a self-controlled longitudinal study. We compared blood pressure levels before the participants accepted health services from the NEPHSP versus after accepting the health services. The changes in blood pressure levels could preliminarily reflect the effect of the NEPHSP on blood pressure control for the participants with hypertension. However, the data monitoring blood pressure changes were only available for the NEPHSP participants and were inaccessible for people with hypertension who did not register with the NEPHSP. Because of the inaccessibility of blood pressure monitoring data for this population, this study was unable to compare blood pressure control between the NEPHSP participants and people who did not register with the NEPHSP. Given this situation, further cohort and control trials are required to confirm advantages of the NEPHSP.

Conclusion

The NEPHSP was effective for improving blood pressure control of Chinese people with hypertension. However, the SBP of the study participants was near-target, which could imply therapy inertia in the NEPHSP participants. A suggested strategy to tackle potential therapy inertia was to intensify blood pressure control of the NEPHSP. The effects of the NEPHSP on blood pressure control in older participants and participants with CVD were modest. Embedding intensive treatment modules into the NEPHSP was suggested to improve blood pressure control of older participants and participants with CVD. Further studies confirming effects of the NEPHSP on blood pressure control for people with specific characteristics, such as older age and comorbidities, are required.

Acknowledgments

We are grateful to all the staff who participated in this research survey.

Data Availability

The data that support the findings of this study are available on reasonable request from the corresponding author. The data are not publicly available due to their containing information that could compromise the privacy of the participants.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Univariate regression analysis of covariates.
[XLSX File, 11 KB - publichealth_v11i1e65783_app1.xlsx]

Multimedia Appendix 2
Corresponding information of each covariate in the generalized estimating equation regressions.
[XLSX File, 12 KB - publichealth_v11i1e65783_app2.xlsx]

Multimedia Appendix 3
Subgroup analysis.
[XLSX File, 10 KB - publichealth_v11i1e65783_app3.xlsx]

References

1. GBD 2019 Viewpoint Collaborators. Five insights from the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1135-1159. [doi: 10.1016/S0140-6736(20)31404-5] [Medline: 33069324]
2. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
3. Lewington S, Lacey B, Clarke R, et al. The burden of hypertension and associated risk for cardiovascular mortality in China. JAMA Intern Med 2016 Apr;176(4):524-532. [doi: 10.1001/jamainternmed.2016.0190] [Medline: 26975032]

XSL•FO
RenderX

4. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249. [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
5. Roth GA, Mensah GA, Johnson CO, et al. Global burden of chronic respiratory diseases and risk factors, 1990-2019: an update from the Global Burden of Disease Study 2019. EClinMed 2023;59:101956. [doi: 10.1016/j.eclinm.2023.101956] [Medline: 37125402]
6. Yang ZJ, Liu J, Ge JP, et al. Prevalence of cardiovascular disease risk factor in the Chinese population: the 2007-2008 China National Diabetes and Metabolic Disorders Study. Eur Heart J 2012 Jan;33(2):213-220. [doi: 10.1093/eurheartj/ehr205] [Medline: 21719451]
7. Li D, Lv J, Liu F, et al. Hypertension burden and control in mainland China: analysis of nationwide data 2003-2012. Int J Cardiol 2015 Apr 1;184:637-644. [doi: 10.1016/j.ijcard.2015.03.045] [Medline: 25771229]
8. Feng XL, Pang M, Beard J. Health system strengthening and hypertension awareness, treatment and control: data from the China Health and Retirement Longitudinal Study. Bull World Health Organ 2014 Jan 1;92(1):29-41. [doi: 10.2471/BLT.13.124495] [Medline: 24391298]
9. He J, Gu D, Chen J, et al. Premature deaths attributable to blood pressure in China: a prospective cohort study. Lancet 2009 Nov 21;374(9703):1765-1772. [doi: 10.1016/S0140-6736(09)61199-5] [Medline: 19811816]
10. Yang G, Wang Y, Zeng Y. Rapid health transition in China, 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet 2013 Jun 8;381(9882):1987-2015. [doi: 10.1016/S0140-6736(13)61097-1.] [Medline: 23746901]
11. Wang J, Zhang L, Wang F, Liu L, Wang H, China National Survey of Chronic Kidney Disease Working Group. Prevalence, awareness, treatment, and control of hypertension in China: results from a national survey. Am J Hypertens 2014 Nov;27(11):1355-1361. [doi: 10.1093/ajh/hpu053] [Medline: 24698853]
12. Li W, Gu H, Teo KK, et al. Hypertension prevalence, awareness, treatment, and control in 115 rural and urban communities involving 47 000 people from China. J Hypertens 2016 Jan;34(1):39-46. [doi: 10.1097/HJH.0000000000000745] [Medline: 26630211]
13. Lu J, Lu Y, Wang X, et al. Prevalence, awareness, treatment, and control of hypertension in China: data from 17 million adults in a population-based screening study (China PEACE Million Persons Project). Lancet 2017 Dec 9;390(10112):2549-2558. [doi: 10.1016/S0140-6736(17)32478-9] [Medline: 29102084]
14. You L, Zhao J, Chen X. National essential public health services programs over the past decade research report two: progress and achievements of the implementation of National Essential Public Health Services Programs over the past decade. Chin Gen Pract 2022;25:3209-3220. [doi: 10.12114/j.issn.1007-9572.2022.0407]
15. Fang G, Yang D, Wang L, Wang Z, Liang Y, Yang J. Experiences and challenges of implementing universal health coverage with China's National Basic Public Health Service Program: literature review, regression analysis, and insider interviews. JMIR Public Health Surveill 2022;8(7):e31289. [doi: 10.2196/31289] [Medline: 35867386]
16. Li X, Lu J, Hu S, et al. The primary health-care system in China. Lancet 2017 Dec 9;390(10112):2584-2594. [doi: 10.1016/S0140-6736(17)33109-4] [Medline: 29231837]
17. Li X, Krumholz HM, Yip W, et al. Quality of primary health care in China: challenges and recommendations. Lancet 2020 Jun 6;395(10239):1802-1812. [doi: 10.1016/S0140-6736(20)30122-7] [Medline: 32505251]
18. Xiong S, Jiang W, Meng R, et al. Factors associated with the uptake of national essential public health service package for hypertension and type-2 diabetes management in China's primary health care system: a mixed-methods study. Lancet Reg Health West Pac 2023 Feb;31:100664. [doi: 10.1016/j.lanwpc.2022.100664] [Medline: 36879777]
19. Liu J, Colditz GA. Optimal design of longitudinal data analysis using generalized estimating equation models. Biom J 2017 Mar;59(2):315-330. [doi: 10.1002/bimj.201600107] [Medline: 27878852]
20. Newman AB. An overview of the design, implementation, and analyses of longitudinal studies on aging. J Am Geriatr Soc 2010 Oct;58 Suppl 2(Suppl 2):S287-S291. [doi: 10.1111/j.1532-5415.2010.02916.x] [Medline: 21029055]
21. Schneeweiss S, Rassen JA, Brown JS, et al. Graphical depiction of longitudinal study designs in health care databases. Ann Intern Med 2019 Mar 19;170(6):398-406. [doi: 10.7326/M18-3079] [Medline: 30856654]
22. Hund H, Gerth S, Lossnitzer D, Fegeler C. Longitudinal data driven study design. Stud Health Technol Inform 2014;205(373-7):373-377. [Medline: 25160209]
23. Wang JG, Chia YC, Chen CH, et al. What is new in the 2018 Chinese hypertension guideline and the implication for the management of hypertension in Asia? J Clin Hypertens (Greenwich) 2020 Mar;22(3):363-368. [doi: 10.1111/jch.13803] [Medline: 31955513]
24. Huang Y, Pan J. Penalized joint generalized estimating equations for longitudinal binary data. Biom J 2022 Jan;64(1):57-73. [doi: 10.1002/bimj.202000336] [Medline: 34587284]
25. Lee DH, Keum N, Hu FB, et al. Predicted lean body mass, fat mass, and all cause and cause specific mortality in men: prospective US cohort study. BMJ 2018 Jul 3;362:k2575. [doi: 10.1136/bmj.k2575] [Medline: 29970408]
26. Zhu J, Chen C, Lu L, Shikany JM, D'Alton ME, Kahe K. Folate, vitamin B6, and vitamin B12 status in association with metabolic syndrome incidence. JAMA Netw Open 2023 Jan 3;6(1):e2250621. [doi: 10.1001/jamanetworkopen.2022.50621] [Medline: 36630134]

XSL•FO
RenderX

27. Zhang W, He K, Zhao H, et al. Association of body mass index and waist circumference with high blood pressure in older adults. BMC Geriatr 2021 Apr 19;21(1):260. [doi: 10.1186/s12877-021-02154-5] [Medline: 33874889]
28. Bowling CB, Sloane R, Pieper C, et al. Association of sustained blood pressure control with multimorbidity progression among older adults. J Am Geriatr Soc 2020 Sep;68(9):2059-2066. [doi: 10.1111/jgs.16558] [Medline: 32501546]
29. Karmali KN, Lloyd-Jones DM, van der Leeuw J, et al. Blood pressure-lowering treatment strategies based on cardiovascular risk versus blood pressure: a meta-analysis of individual participant data. PLoS Med 2018 Mar;15(3):e1002538. [doi: 10.1371/journal.pmed.1002538] [Medline: 29558462]
30. Czernichow S, Castetbon K, Salanave B, et al. Determinants of blood pressure treatment and control in obese people: evidence from the general population. J Hypertens 2012 Dec;30(12):2338-2344. [doi: 10.1097/HJH.0b013e3283593010] [Medline: 23032140]
31. Ren Y, Cheng L, Song Y, et al. Ideal cardiovascular health index and high-normal blood pressure in elderly people: evidence based on real-world data. Sci Rep 2024 May 3;14(1):10166. [doi: 10.1038/s41598-024-60906-w] [Medline: 38702348]
32. Aljuraiban GS, Alharbi FJ, Aljohi AO, et al. Triglyceride-glucose index (TyG index) in association with blood pressure in adults: a retrospective study. Int J Gen Med 2024;17:3395-3402. [doi: 10.2147/IJGM.S469147] [Medline: 39130487]
33. Masoli JAH, Mensah E, Rajkumar C. Age and ageing cardiovascular collection: blood pressure, coronary heart disease and heart failure. Age Ageing 2022 Aug 2;51(8):afac179. [doi: 10.1093/ageing/afac179] [Medline: 35934320]
34. Wong MCS, Wang HHX, Cheung CSK, et al. Factors associated with multimorbidity and its link with poor blood pressure control among 223,286 hypertensive patients. Int J Cardiol 2014 Nov 15;177(1):202-208. [doi: 10.1016/j.ijcard.2014.09.021] [Medline: 25499379]
35. Glynn LG, Murphy AW, Smith SM, Schroeder K, Fahey T. Interventions used to improve control of blood pressure in patients with hypertension. Cochrane Database Syst Rev 2010 Mar 17;17(3):CD005182. [doi: 10.1002/14651858.CD005182.pub4] [Medline: 20238338]
36. Jaffe MG, Lee GA, Young JD, Sidney S, Go AS. Improved blood pressure control associated with a large-scale hypertension program. JAMA 2013 Aug 21;310(7):699-705. [doi: 10.1001/jama.2013.108769] [Medline: 23989679]
37. Jafar TH, Gandhi M, de Silva HA, et al. A community-based intervention for managing hypertension in rural South Asia. N Engl J Med 2020 Feb 20;382(8):717-726. [doi: 10.1056/NEJMoa1911965] [Medline: 32074419]
38. Schwalm JD, McCready T, Lopez-Jaramillo P, et al. A community-based comprehensive intervention to reduce cardiovascular risk in hypertension (HOPE 4): a cluster-randomised controlled trial. Lancet 2019 Oct 5;394(10205):1231-1242. [doi: 10.1016/S0140-6736(19)31949-X] [Medline: 31488369]
39. Anand TN, Joseph LM, Geetha AV, Prabhakaran D, Jeemon P. Task sharing with non-physician health-care workers for management of blood pressure in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Glob Health 2019 Jun;7(6):e761-e771. [doi: 10.1016/S2214-109X(19)30077-4] [Medline: 31097278]
40. Schlaich MP, Kiuchi MG, Nolde JM, Lugo-Gavidia LM, Carnagarin R. Therapeutic inertia in hypertension management - status quo in primary care. J Hypertens 2021 Jun 1;39(6):1107-1108. [doi: 10.1097/HJH.0000000000002830] [Medline: 33967214]
41. Phillips LS, Branch WT, Cook CB, et al. Clinical inertia. Ann Intern Med 2001 Nov 6;135(9):825-834. [doi: 10.7326/0003-4819-135-9-200111060-00012] [Medline: 11694107]
42. Milman T, Joundi RA, Alotaibi NM, Saposnik G. Clinical inertia in the pharmacological management of hypertension: a systematic review and meta-analysis. Medicine (Baltimore) 2018 Jun;97(25):e11121. [doi: 10.1097/MD.0000000000011121] [Medline: 29924011]
43. Lebeau JP, Cadwallader JS, Vaillant-Roussel H, et al. General practitioners' justifications for therapeutic inertia in cardiovascular prevention: an empirically grounded typology. BMJ Open 2016 May 13;6(5):e010639. [doi: 10.1136/bmjopen-2015-010639] [Medline: 27178974]
44. Kerr EA, Zikmund-Fisher BJ, Klamerus ML, Subramanian U, Hogan MM, Hofer TP. The role of clinical uncertainty in treatment decisions for diabetic patients with uncontrolled blood pressure. Ann Intern Med 2008 May 20;148(10):717-727. [doi: 10.7326/0003-4819-148-10-200805200-00004] [Medline: 18490685]
45. Masoli JAH, Delgado J, Pilling L, Strain D, Melzer D. Blood pressure in frail older adults: associations with cardiovascular outcomes and all-cause mortality. Age Ageing 2020 Aug 24;49(5):807-813. [doi: 10.1093/ageing/afaa028] [Medline: 32133525]
46. Boyd C, Smith CD, Masoudi FA, et al. Decision making for older adults with multiple chronic conditions: executive summary for the American Geriatrics Society Guiding Principles on the Care of Older Adults With Multimorbidity. J Am Geriatr Soc 2019 Apr;67(4):665-673. [doi: 10.1111/jgs.15809] [Medline: 30663782]
47. Benetos A, Petrovic M, Strandberg T. Hypertension management in older and frail older patients. Circ Res 2019 Mar 29;124(7):1045-1060. [doi: 10.1161/CIRCRESAHA.118.313236] [Medline: 30920928]
48. Anker D, Santos-Eggimann B, Zwahlen M, et al. Blood pressure control and complex health conditions in older adults: impact of recent hypertension management guidelines. J Hum Hypertens 2021 Mar;35(3):280-289. [doi: 10.1038/s41371-020-0334-4] [Medline: 32346124]

XSL•FO
RenderX

49. Bowling CB, Lee A, Williamson JD. Blood pressure control among older adults with hypertension: narrative review and introduction of a framework for improving care. Am J Hypertens 2021 Apr 2;34(3):258-266. [doi: 10.1093/ajh/hpab002] [Medline: 33821943]
50. Zhang WL, Cai J. STEP to blood pressure management of elderly hypertension: evidence from Asia. Hypertens Res 2022 Apr;45(4):576-582. [doi: 10.1038/s41440-022-00875-7] [Medline: 35277670]
51. Blood Pressure Lowering Treatment Trialists' Collaboration. Age-stratified and blood-pressure-stratified effects of blood-pressure-lowering pharmacotherapy for the prevention of cardiovascular disease and death: an individual participant-level data meta-analysis. Lancet 2021 Sep 18;398(10305):1053-1064. [doi: 10.1016/S0140-6736(21)01921-8] [Medline: 34461040]
52. Curfman G, Bauchner H, Greenland P. Treatment and control of hypertension in 2020: the need for substantial improvement. JAMA 2020 Sep 22;324(12):1166-1167. [doi: 10.1001/jama.2020.13322] [Medline: 32902571]
53. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA 2005 Aug 10;294(6):716-724. [doi: 10.1001/jama.294.6.716] [Medline: 16091574]
54. Valderas JM, Starfield B, Roland M. Multimorbidity's many challenges: a research priority in the UK. BMJ 2007 Jun 2;334(7604):1128. [doi: 10.1136/bmj.39226.427095.3A] [Medline: 17540918]
55. Kennard L, O'Shaughnessy KM. Treating hypertension in patients with medical comorbidities. BMJ 2016 Feb 16;352:i101. [doi: 10.1136/bmj.i101] [Medline: 26884124]
56. Sarkar C, Dodhia H, Crompton J, et al. Hypertension: a cross-sectional study of the role of multimorbidity in blood pressure control. BMC Fam Pract 2015 Aug 7;16:98. [doi: 10.1186/s12875-015-0313-y] [Medline: 26248616]
57. Paulsen MS, Andersen M, Thomsen JL, et al. Multimorbidity and blood pressure control in 37 651 hypertensive patients from Danish general practice. J Am Heart Assoc 2012 Dec 31;2(1):e004531. [doi: 10.1161/JAHA.112.004531] [Medline: 23525411]
58. Messerli FH, Hofstetter L, Rimoldi SF, Rexhaj E, Bangalore S. Risk factor variability and cardiovascular outcome: JACC review topic of the week. J Am Coll Cardiol 2019 May 28;73(20):2596-2603. [doi: 10.1016/j.jacc.2019.02.063] [Medline: 31118154]
59. Mehlum MH, Liestøl K, Kjeldsen SE, et al. Blood pressure variability and risk of cardiovascular events and death in patients with hypertension and different baseline risks. Eur Heart J 2018 Jun 21;39(24):2243-2251. [doi: 10.1093/eurheartj/ehx760] [Medline: 29365085]
60. Mezue K, Goyal A, Pressman GS, Matthew R, Horrow JC, Rangaswami J. Blood pressure variability predicts adverse events and cardiovascular outcomes in SPRINT. J Clin Hypertens (Greenwich) 2018 Sep;20(9):1247-1252. [doi: 10.1111/jch.13346] [Medline: 29984884]
61. Stevens SL, Wood S, Koshiaris C, et al. Blood pressure variability and cardiovascular disease: systematic review and meta-analysis. BMJ 2016 Aug 9;354:i4098. [doi: 10.1136/bmj.i4098] [Medline: 27511067]
62. Wang Y, Yang W, Jiang X. Association between triglyceride-glucose index and hypertension: a meta-analysis. Front Cardiovasc Med 2021;8:644035. [doi: 10.3389/fcvm.2021.644035] [Medline: 34136539]
63. Xu AR, Jin Q, Shen Z, Zhang J, Fu Q. Association between the risk of hypertension and triglyceride glucose index in Chinese regions: a systematic review and dose-response meta-analysis of a regional update. Front Cardiovasc Med 2023;10:1242035. [doi: 10.3389/fcvm.2023.1242035] [Medline: 37583585]
64. Parati G, Bilo G, Mancia G. Blood pressure measurement in research and in clinical practice: recent evidence. Curr Opin Nephrol Hypertens 2004 May;13(3):343-357. [doi: 10.1097/00041552-200405000-00013] [Medline: 15073495]
65. Sewell K, Halanych JH, Russell LB, et al. Blood pressure measurement biases in clinical settings, Alabama, 2010-2011. Prev Chronic Dis 2016 Jan 7;13:E01. [doi: 10.5888/pcd13.150348] [Medline: 26741995]
66. Atherosclerosis, Hypertension, and Obesity in the Young Committee of the American Heart Association Council on Cardiovascular Disease in the Young, Alpert B, McCrindle B, et al. Recommendations for blood pressure measurement in human and experimental animals; part 1: blood pressure measurement in humans. Hypertension 2006 Jul;48(1):e3. [doi: 10.1161/01.HYP.0000229661.06235.08] [Medline: 16769991]

Abbreviations

CVD: cardiovascular disease
DBP: diastolic blood pressure
GEE: generalized estimating equation
NEPHSP: National Essential Public Health Service Package
SBP: systolic blood pressure
TyG: triglyceride-glucose index

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 26.08.24; peer-reviewed by W Wang, Y Ren; revised version received 17.12.24; accepted 19.12.24; published 06.02.25.

Please cite as:
Wei F, Ge Y, Li H, Liu Y
Impact of the National Essential Public Health Service Package on Blood Pressure Control in Chinese People With Hypertension: Retrospective Population-Based Longitudinal Study
JMIR Public Health Surveill 2025;11:e65783
URL: https://publichealth.jmir.org/2025/1/e65783
doi:10.2196/65783

© Feiran Wei, You Ge, Han Li, Yuan Liu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 6.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Incidence of Deliberate Self-Harm in Hong Kong Before and During the COVID-19 Pandemic: Population-Wide Retrospective Cohort Study

Luke Y F Luk[1,2*], BSc; Christie J Y Ching[3*], MPH; Tsz Fung Yip[1,2], BEng; Sunny C L Chan[3], MStat; Catherine Y T Lam[4]; Elizabeth T C Lam[5]; Serena R B Yue[4], MSc; Hoi Ching Pang[4]; Janet Y H Wong[6], PhD; Carlos K H Wong[2,7,8], PhD; Chak Kwan Tong[9], MBBS; Tafu Yamamoto[10], MBChB; Timothy H Rainer[3,11], MD; Abraham K C Wai[3,11,12], MD; Joshua W K Ho[1,2], PhD

[1]School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[10]Department of Accident & Emergency, Yan Chai Hospital, Hong Kong SAR, China (Hong Kong)

[11]Department of Accident & Emergency, Queen Mary Hospital, Hong Kong SAR, China (Hong Kong)

[12]Department of Accident & Emergency, The University of Hong Kong-Shenzhen Hospital, Shenzhen, China

[2]Laboratory of Data Discovery for Health Limited (D24H), Hong Kong Science Park, Hong Kong SAR, China (Hong Kong)

[3]Department of Emergency Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, G06, G/F, Jockey Club Building for Interdisciplinary Research, 5 Sassoon Road, Pokfulam, Hong Kong SAR, China (Hong Kong)

[4]Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[5]Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[6]School of Nursing & Health Sciences, Hong Kong Metropolitan University, Hong Kong SAR, China (Hong Kong)

[7]Department of Family Medicine and Primary Care, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[8]Department of Pharmacology and Pharmacy, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[9]Intensive Care Unit, Tuen Mun Hospital, Hong Kong SAR, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Abraham K C Wai, MD
Department of Emergency Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, G06, G/F, Jockey Club Building for Interdisciplinary Research, 5 Sassoon Road, Pokfulam, Hong Kong SAR, China (Hong Kong)

Abstract

Background: COVID-19 ended on May 5, 2023, and since then Hong Kong reported increased mental distress, which was speculated to be from the policies implemented during the pandemic. Despite this, longitudinal surveillance of deliberate self-harm (DSH) incidences throughout the pandemic in Hong Kong remained insufficient.

Objective: The objective of this study was to outline the changes in DSH incidences before and during the COVID-19 pandemic in Hong Kong, with respect to sex, age, and co-occurring mental health issues.

Methods: A quasi-experiment was conducted using an interrupted time series design to estimate the impact of the pandemic on DSH-related emergency department (ED) visits. This design enabled the estimation of DSH-related ED visits based on prepandemic data from 2016 to 2019, assuming the pandemic had not occurred, and allowed for a comparison with observed DSH-related ED visits during the pandemic. The descriptive results were reported as the observed monthly DSH-related ED visits and observed incidence ratios during the pandemic. Afterwards, a negative binomial model was fitted to the prepandemic data (2016 - 2019) and adjusted for temporal trends, seasonality, and population variation to estimate the expected monthly DSH-related ED visits and adjusted incidence ratios (aIRs).

Results: Between January 2016 and December 2022, a total of 31,893 DSH episodes were identified. Initial descriptive analysis showed a significant difference in demographic characteristics (sex) and clinical characteristics (death within 28 d, diagnoses of co-occurring mental health issues, public assistance pay code, and triage level). Subsequent interrupted time-series analysis demonstrated significantly increasing trends in comparison with the prepandemic period. As reported in the aIRs among young adult males (aIR in 2020=1.34, *P*=.002; 2021: aIR=1.94, *P*<.001; and 2022: aIR=2.53, *P*<.001), adult males (aIR in 2020=1.58, *P*<.001; 2021: aIR=2.64, *P*<.001; and 2022: aIR=3.13, *P*<.001), adult females (aIR in 2020=1.13, *P*=.01; 2021: aIR=1.52, *P*<.001; and 2022: aIR=1.64, *P*<.001), and older male adults (aIR in 2020=1.53, *P*<.001; 2021: aIR=2.37, *P*<.001; and 2022: aIR=3.01, *P*<.001).

XSL•FO
RenderX

Conclusions: The average annual DSH-related ED visits increased during the pandemic period. Therefore, there is a need to raise awareness for such vulnerable groups in Hong Kong to prepare for postpandemic spillover.

(JMIR Public Health Surveill 2025;11:e57500) doi:10.2196/57500

KEYWORDS

emergency department; COVID-19; deliberate self-harm; mental health; self-harm; self-injury; self-violence; Hong Kong; SARS-Cov-2

Introduction

During the early phase of COVID-19 pandemic, Hong Kong experienced a significant increase in mental health distress within the community [1-3]. Surveys indicated that an alarming 53.9% (299/555) of participants exhibited moderate to very high levels of psychological distress [4]. In response, the Hong Kong Chief Executive allocated HK $300 million (the conversion rate at the latest revision date of the Chief Executive's 2020 policy address [November 25, 2020] was US $0.129 per unit HKD) to raise mental health awareness [5]. An advisory committee subsequently developed a comprehensive, community-based initiative. This initiative included an education and health promotion program called "Shall We Talk," emotional support workshop training for school staff, mental health first aid training for service professionals, and the deployment of suicide bereavement liaison officers. Despite these efforts, there were ongoing concerns about the sustainability of these resources [6].

The association between stress and deliberate self-harm (DSH) was well-documented in the literature [7-11]. However, recent local studies on DSH risk factors, conducted before the pandemic, identified factors such as sex, age, social welfare status, and co-occurring mental health disorders (ie, depression, bipolar disorder, personality disorder, and substance misuse) [12]. The literature from other countries on this topic is limited and sometimes contradictory, with mixed definitions of self-harm and varying conclusions about youth as a risk factor [13-17]. This highlights the need for longitudinal surveillance to better understand the pre-existing risk factors of DSH.

One of the major challenges in emergency departments (EDs) is reducing waiting times without compromising the quality of health care, which includes preventing readmission and reducing health care costs. The Hospital Authority in Hong Kong has pledged to limit the waiting time for emergency (triage category 2) patients to 15 minutes, and for urgent (triage category 3) patients to 30 minutes [18]. However, by May 2022, over 1500 patients were still waiting for admissions in ED [19]. This underscores the necessity for effective health policies to manage DSH-related ED visits.

The aim of this study was to describe the impact of COVID-19 on DSH incidences in Hong Kong through a longitudinal analysis of DSH-related ED visits before and during the pandemic. This study was stratified by sex, age, and co-occurring mental health issues. The findings from this study justify the need for further surveillance to monitor a possible spill-over effect post the COVID-19 pandemic.

Methods

Study Overview and Settings

This territory-wide cohort study analyzed data from January 2020 to December 2022, which was further categorized into pandemic months (January 2020 to December 2022) and the prepandemic months (January 2016 to December 2019).

Data Sources and International Classification of Diseases, Ninth Revision, Clinical Modification Codes

Data were sourced from the Clinical Data Analysis and Reporting System (CDARS), an administrative clinical database managed by the Hospital Authority of the Hong Kong Special Administrative Region of China [20], encompassing ED data from 18 public hospitals.

DSH and mental health issues were identified using both nursing codes for self-inflicted injury and the *ICD-9-CM* (*International Classification of Diseases, Ninth Revision, Clinical Modification*) diagnosed by clinicians (Table S1 in Multimedia Appendix 1).

In addition, yearly and midyearly population counts for sex and age groups were obtained from the government in Hong Kong [21].

Outcome Measures

This study defined DSH according the National Institute for Clinical Excellence guidelines, which describe it as "self-poisoning or injury, irrespective of the apparent purpose of the act" [22]. This guideline was adopted due to the absence of patients' intent in the medical records from CDARS.

Other key definitions included age groups, public assistance pay codes, triage categories, and mental health issues. Age was categorized into teenagers (12 - 17 y), young adults (18 - 24 y), adults (25 - 64 y), and older adults (65 - 84 y). Public assistance pay codes referred to the recipients of the Hospital Authority medical fee waiver schemes (Residential Care Service Voucher, Comprehensive Social Security Assistance, and Old Age Living Allowance) [23]. Triage categories 1 to 3 were defined in Table S2 in Multimedia Appendix 1 [18]. Mental health issues included mood disorders, neurosis, alcohol and drug-associated mental health conditions, and other psychotic illnesses and developmental disorders. These diagnoses were identified by *ICD-9-CM* codes, as defined in Table S1 in Multimedia Appendix 1.

Data Analysis

First, a descriptive analysis of the outcome measures was conducted to identify statistical significance before and during

XSL•FO
RenderX

COVID-19. Fisher exact and Mann-Whitney *U* tests were used to compare DSH incidences, sex, age groups, and co-occurring mental health issues.

Next, a quasi-experiment was conducted using an interrupted time series (ITS) design to estimate the impact of the COVID-19 pandemic on DSH-related ED visits. Quasi-experiments estimate causal relationships between different factors and outcomes when random assignment is not possible. The ITS design enabled the estimation of DSH-related ED visits during the pandemic period (2020 to 2022) based on prepandemic data (2016 to 2019). This allowed for a comparison between the observed DSH-related ED visits and the expected DSH-related ED visits in the absence of the pandemic.

Segmented regression is a method commonly used in ITS, typically involves assuming linear trends and identifying significant break points. Given the complexity of COVID-19's impact on self-harm ED visits, a nonsegmented regression model was chosen, fitted with prepandemic data instead of a segmented-regression model fitted with all data. Self-harm–related ED visits were aggregated into monthly counts to ITS analysis for each group. This analysis was done by fitting prepandemic data to the negative binomial models, with a linear term for trend, and optionally with Fourier terms to account for seasonality, using population terms as an offset. The candidate models can be described by:

$$E[y]=\exp(\beta_0+\beta_{trend}x+\sum_{k=1}^{n}[\beta_{1,k}\sin(2\pi kxt/12)+\beta_{2,k}\cos(2\pi kxt/12)])population$$

Or

$$E[y]=\exp(\beta_0+\beta_{trend}\cdot xt)\cdot population$$

The dependent variable y represents the monthly count of self-harm incidents for a particular subgroup. Where xt is the number of months from the start of the study period, n is the highest degree of Fourier terms, $\beta_0\beta_{trend}$ is the prepandemic slope of incidence rate, $\beta_{1,k}\beta_{2,k}$ are coefficients for seasonality. The inclusion of the Fourier term and its degree n were guided by the Akaike Information Criterion (AIC). The final model is the one with the minimal AIC value, ensuring the best balance between model fit and complexity.

The findings from this study were presented as observed incidence ratio (oIR) and adjusted incidence ratio (aIR). The respective formulas were as follows:

$$oIR=\frac{Observed\ annual\ DSH\text{-}related\ ED\ visits}{Prepandemic\ annual\ DSH\text{-}related\ ED\ visits}$$

$$aIR=\frac{Observed\ annual\ DSH\text{-}related\ ED\ visits\ rate}{Expected\ annual\ DSH\text{-}related\ ED\ visits\ rate}$$

In the formulas, the expected annual DSH-related ED visits were estimated using negative binomial models and accounted for trends, seasonal variations, and population changes. Annual counts were estimated by summing monthly counts, assuming the population remained constant. The aIR compares the observed number of DSH-related ED visits during the pandemic to the expected number based on prepandemic data, providing an estimate of the COVID-19 pandemic's impact on DSH-related ED visits. Time series graphs were then generated to illustrate these trends. All *P* values less than .05 were considered statistically significant. All statistical analyses were performed using RStudio (version 4.1.0; R Foundation for Statistical Computing).

Ethical Considerations

The study protocol received approval from the institutional review board of the Hospital Authority, Hong Kong West Cluster (reference UW 20 - 112; Multimedia Appendix 2). Informed consent was waivered, and all the data were anonymized.

Results

Overview

A total of 31,893 cases of DSH-related ED visits were identified during the study period. Initial descriptive analysis compared demographic characteristics between the prepandemic and pandemic periods (Table 1). The total number of ED visits with DSH increased by 1175 out of 15,359 (7.7%) patients. There were significantly higher DSH-related ED visits in males (*P*<.001), co-occurring mental health issues diagnosis (*P*<.001), triage category 3 (urgent), and 4 (semiurgent) (*P*<.001) during the pandemic. Conversely, there were significantly less deaths within 28 days (*P*<.001) and less patients with public assistance pay code during the pandemic (*P*=.046). These trends were further explored with subgroup analysis using ITS models (Table 2, Figures 1-8). Further subgroup analysis of co-occurring mental health issues diagnoses were explored in Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Demographic characteristics in patients with deliberate self-harm who visited the emergency department before and during the pandemic (total N=31,893).

Characteristic		Time period		*P* value
		Prepandemic (n=15,359)	Pandemic (n=16,534)	
Age (years), median (IQR)		38 (25)	39 (30)	.60[a]
Sex, n (%)				<.001[b]
	Female	8184 (53)	8258 (50)	
	Male	7175 (47)	8276 (50)	
Deaths within 28 days, n (%)				<.001[b]
	No	15,102 (98)	16,352 (99)	
	Yes	257 (2)	182 (1)	
Co-occurring mental health issues diagnosis, n (%)				<.001[b]
	No	5991 (39)	4878 (30)	
	Yes	9368 (61)	11,656 (71)	
Public assistance pay code, n (%)				.046[b]
	No	12,060 (79)	13,134 (79)	
	Yes	3299 (22)	3400 (21)	
Triage, n (%)				<.001[a]
	Category 1 (critical)	1885 (12)	1329 (8)	
	Category 2 (emergent)	5944 (39)	5450 (33)	
	Category 3 (urgent)	6429 (42)	8539 (52)	
	Category 4 (semiurgent)	1055 (7)	1169 (7)	
	Category 5 (nonurgent)	29 (0)	30 (0)	

[a]Mann-Whitney *U* test.

[b]Fisher exact test.

XSL•FO
RenderX

Table . The number of average yearly visits, observed incidence ratio, and adjusted incidence ratio in deliberate self-harm–related emergency department visits between 2016 and 2022.

Patient characteristic	Years									
	2016 - 2019	2020			2021			2022		
	Average yearly visits, n (%)	oIR[a]	aIR[b]	*P* value	oIR	aIR	*P* value	oIR	aIR	*P* value
Female teenager (12 - 17 years old)										
Total	251 (100)	1.71	0.78	.98	3.26	1.10	.27	2.90	0.73	.95
All mental health issues	156 (62)	1.89	0.86	.86	3.99	1.34	.06	3.31	0.82	.81
Alcohol and drug associated mental health conditions	14 (6)	2.69	1.50	.11	4.58	2.07	.05	3.78	1.39	.28
Male teenager (12 - 17 years old)										
Total	77 (100)	1.09	0.92	.70	1.71	1.36	.07	1.65	1.25	.21
All mental health issues	37 (48)	1.47	1.07	.35	2.67	1.73	.04	2.50	1.47	.15
Alcohol and drug associated mental health conditions	6 (8)	2.24	2.46	.01	3.04	3.52	.01	4.16	5.13	.02
Female young adult (18 - 24 years old)										
Total	287 (100)	1.02	1.06	.25	1.38	1.48	.002	1.57	1.71	<.001
All mental health issues	149 (52)	1.21	1.24	.04	1.53	1.60	.005	1.83	1.91	.004
Alcohol and drug associated mental health conditions	47 (16)	1.16	1.64	.008	1.61	2.65	.001	1.91	3.58	.001
Male young adult (18 - 24 years old)										
Total	177 (100)	0.93	1.34	.002	1.15	1.94	<.001	1.30	2.53	<.001
All mental health issues	107 (61)	0.96	1.48	.01	1.37	2.53	<.001	1.61	3.51	<.001
Alcohol and drug associated mental health conditions	66 (37)	1.08	2.04	.001	1.68	4.08	<.001	2.08	6.38	<.001
Female adult (25 - 64 years old)										
Total	1316 (100)	0.94	1.13	.01	1.15	1.52	<.001	1.13	1.64	<.001

XSL•FO
RenderX

Patient characteristic	Years									
	2016 - 2019	2020			2021			2022		
	Average yearly visits, n (%)	oIR[a]	aIR[b]	*P* value	oIR	aIR	*P* value	oIR	aIR	*P* value
All mental health issues	765 (58)	0.96	1.27	<.001	1.26	1.89	<.001	1.23	1.23	<.001
Alcohol and drug associated mental health conditions	320 (24)	1.20	1.88	<.001	1.86	3.55	<.001	1.88	4.35	<.001
Male adult (25 - 64 years old)										
Total	1289 (100)	1.17	1.58	<.001	1.72	2.64	<.001	1.78	3.13	<.001
All mental health issues	915 (71)	1.24	1.83	<.001	2.00	3.49	<.001	2.09	4.28	<.001
Alcohol and drug associated mental health conditions	713 (55)	1.35	2.07	<.001	2.32	4.27	<.001	2.39	5.26	<.001
Female older adult (65 - 84 years old)										
Total	128 (100)	1.25	1.04	.33	1.54	1.19	.15	1.81	1.32	.08
All mental health issues	59 (46)	1.27	1.27	.09	1.64	1.63	.02	2.14	2.13	.002
Alcohol and drug associated mental health conditions	14 (11)	1.85	1.29	.23	3.78	2.27	.07	3.93	2.05	.16
Male older adult (65 - 84 years old)										
Total	177 (100)	1.37	1.53	<.001	2.03	2.37	<.001	2.43	3.01	<.001
All mental health issues	104 (59)	1.60	1.71	<.001	2.57	2.84	<.001	3.19	3.66	<.001
Alcohol and drug associated mental health conditions	70 (40)	1.92	2.11	<.001	3.35	3.83	<.001	4.10	4.93	<.001

[a]oIR: observed incidence ratio.

[b]aIR: adjusted incidence ratio.

XSL•FO
RenderX

Figure 1. Subgroup analysis of female teenagers (12 - 17 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

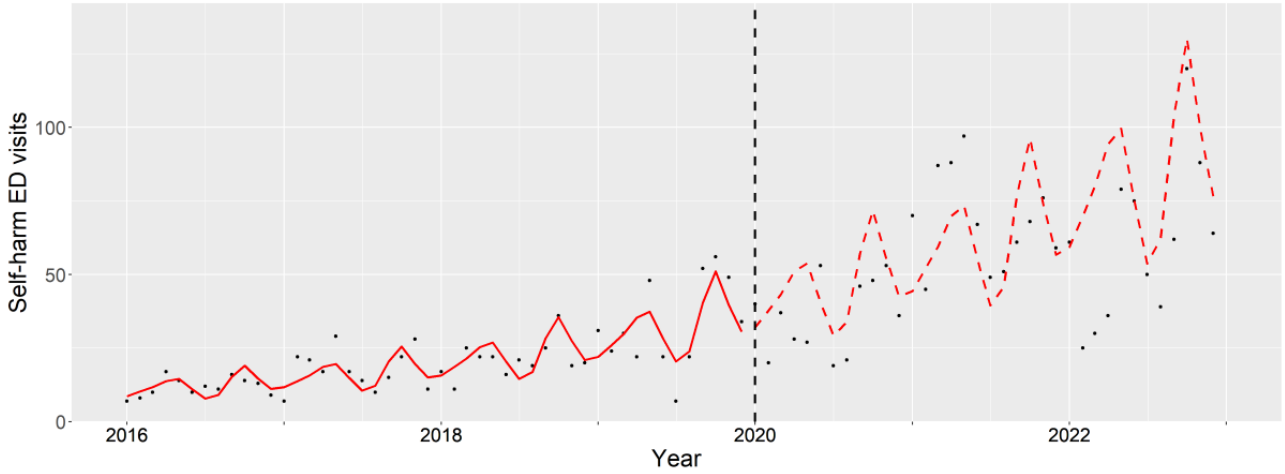


Figure 2. Subgroup analysis of male teenagers (12 - 17 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

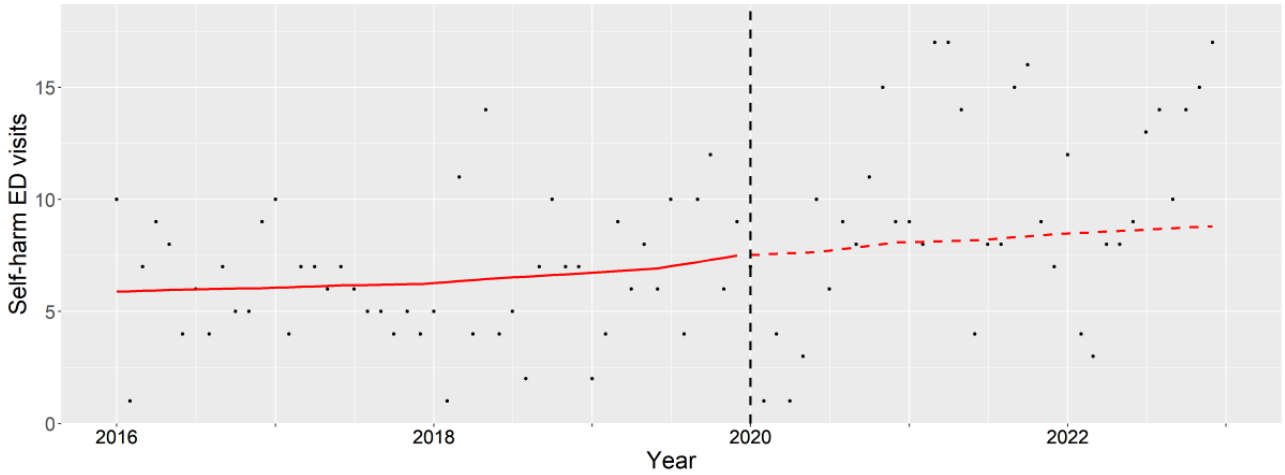
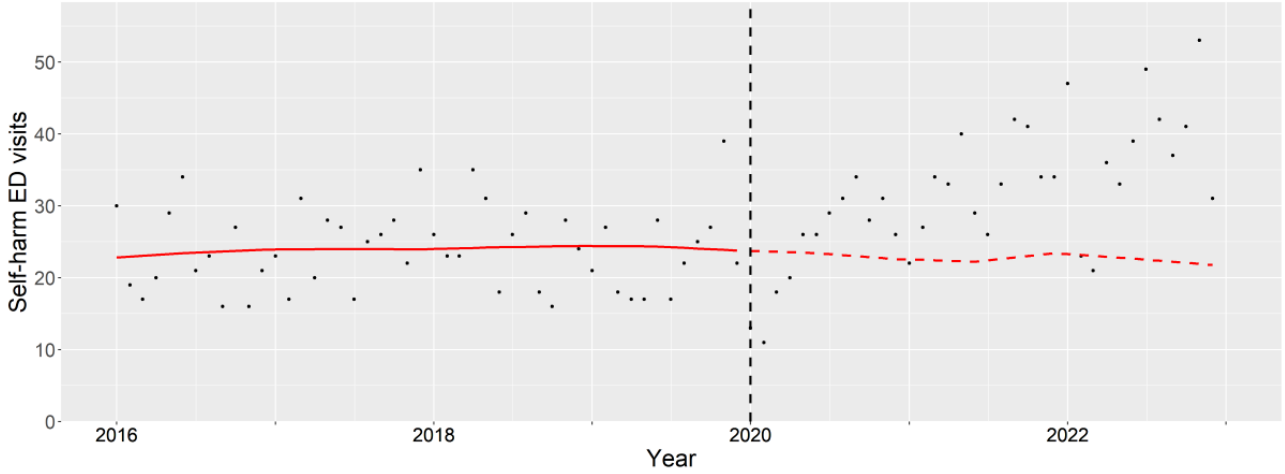


Figure 3. Subgroup analysis of female young adults (18 - 24 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.



XSL•FO
RenderX

Figure 4. Subgroup analysis of male young adults (18 - 24 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

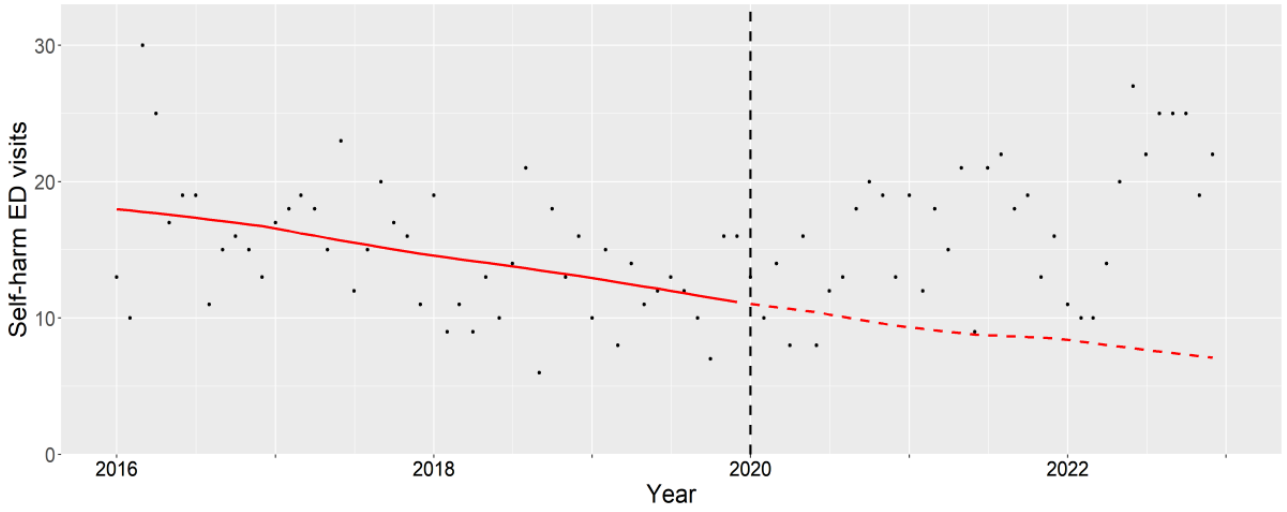


Figure 5. Subgroup analysis of female adults (25 - 64 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

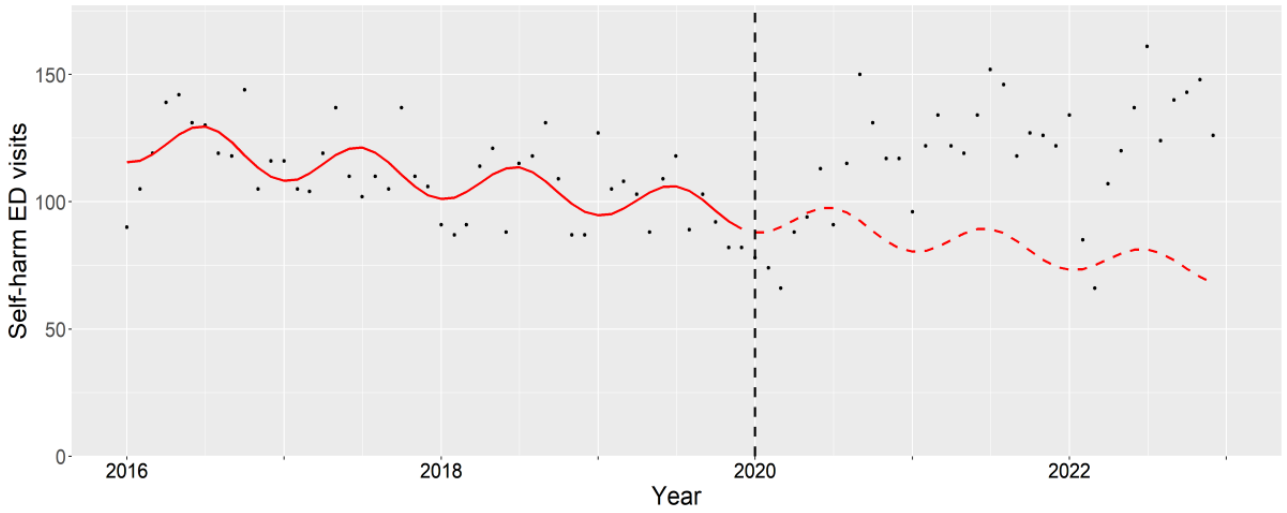
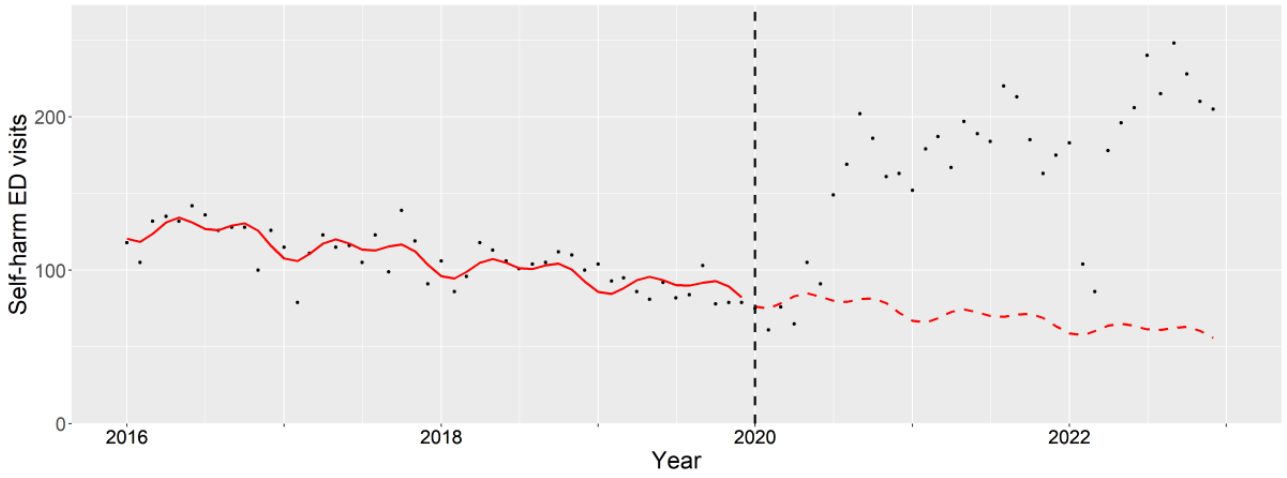


Figure 6. Subgroup analysis of male adults (25 - 64 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020).



XSL•FO
RenderX

Figure 7. Subgroup analysis of female older adults (65 - 84 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.

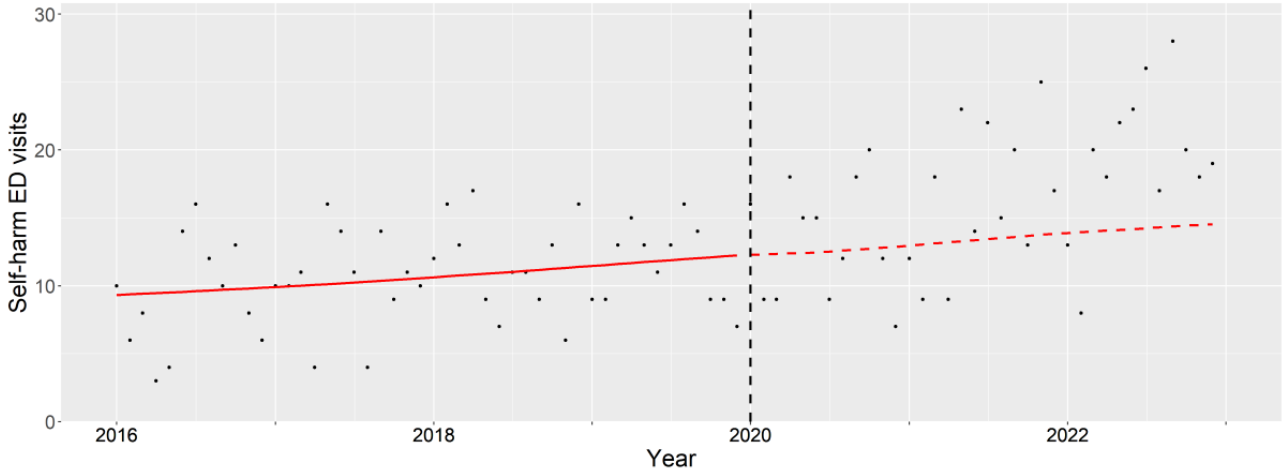
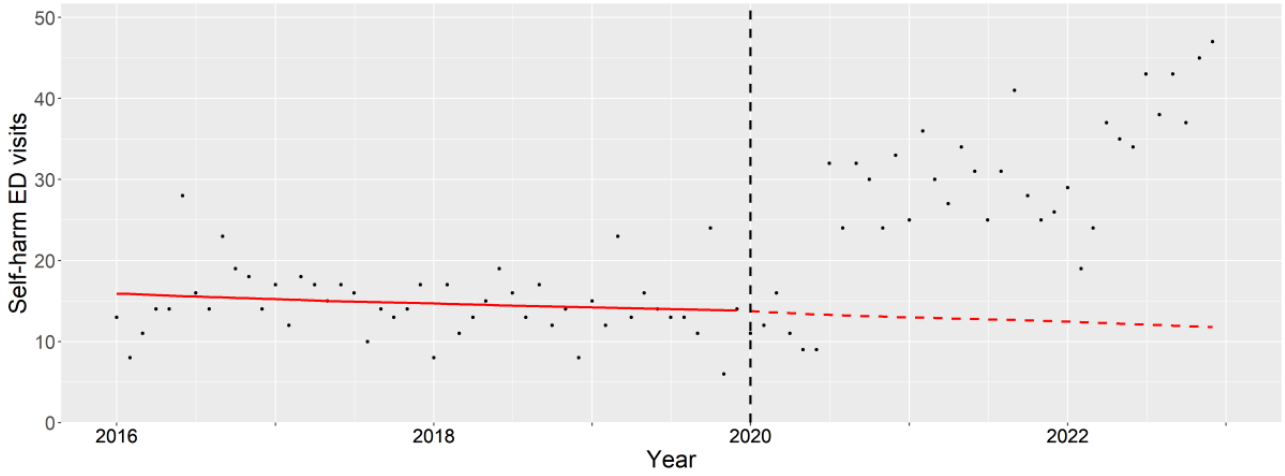


Figure 8. Subgroup analysis of male older adults (65 - 84 years old). The observed (black dots) and expected (red line) number of deliberate self-harm–related emergency department visits between 2016 and 2022. The vertical black dash line indicates the beginning of pandemic period (January 2020). ED: emergency department.



Teenagers

For teenage females, the oIRs and aIR showed no difference (2020: *P*=.98, 2021: *P*=.27, and 2022: *P*=.95), as reflected by the overlap of the expected and observed trends (Figure 1).

For teenage males, the aIRs showed no difference (2020: *P*=.70, 2021: *P*=.07, and 2022: *P*=.21), as highlighted by the overlap of the observed and expected trends during the study period (Figure 2). However, the aIRs among those with alcohol and drug-associated mental health conditions were significantly different and increased during the pandemic (2020: *P*=.01, 2021: *P*=.01, and 2022: *P*=.02).

Young Adults

For female young adults, the observed trend was higher than expected during the pandemic (Figure 3). The highest aIR were observed among those with alcohol and drug-associated mental health conditions, with significant difference (2020: *P*=.008, 2021: *P*=.001, and 2022: *P*=.001).

For male young adults, the aIRs between 2020 and 2022 all significantly increased during the pandemic (2020: *P*=.002, 2021: *P*<.001, and 2022: *P*<.001). Further illustrated when the observed trend was higher than the expected trend during the pandemic (Figure 4). Those with alcohol and drug associated mental health conditions had the highest aIRs with statistical significance (2020: *P*=.001, 2021: *P*<.001, and 2022: *P*<.001).

Adults

The aIRs among adult females significantly increased during the pandemic (2020: *P*=.01, 2021: *P*<.001, and 2022: *P*<.001). This was highlighted when the expected trend decreased while the observed trend increased during the pandemic (Figure 5). The highest aIRs were seen in patients with co-occurring alcohol and drug associated mental health conditions with statistical significance (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001).

For adult males, the aIRs significantly increased during the pandemic (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001). As illustrated when the expected trend decreased while the observed trend increased during the pandemic (Figure 6). Significantly highest aIRs were observed among those with co-occurring alcohol and drug associated mental health condition (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001).

XSL•FO
RenderX

Older Adults

During 2021 and 2022, older female adults with mental health issues exhibited significantly higher aIRs (2021: *P*=.02, 2022: *P*=.002). These rates were elevated compared with the expected levels during 2021 and 2022 but declined during 2020 (Figure 7).

For older male adults, the aIRs significantly increased throughout the pandemic period (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001). Further underscored by the divergence between the decreasing expected rates and the increasing observed rates during the pandemic (Figure 8). The most pronounced aIRs with statistical significance were among found in older male adults with co-occurring alcohol and drug associated mental health conditions (2020: *P*<.001, 2021: *P*<.001, and 2022: *P*<.001).

Discussion

Principal Results

The findings from this study revealed differences in DSH incidences before and during the pandemic, with a notable rise among males and individuals with co-occurring alcohol and drug-related mental health issues.

Previous studies conducted before the pandemic also reported higher rates of DSH among males, which contrasts with the trends observed during our study period [12,24]. This discrepancy may be influenced by cultural factors [12]. For instance, past research has shown that while Chinese males in Hong Kong demonstrate health-seeking behavior, they often hesitate to consult doctors, preferring over-the-counter medications instead [25]. Such tendencies could be exacerbated by the mental distress linked to the isolation policies, as noted in another Hong Kong-based study [26]. However, further research is essential to elucidate the impact of COVID-19 policies on the rising incidences of male DSH in Hong Kong.

In addition, the increased incidences of DSH among individuals with alcohol and drug-related mental health issues observed in this study align with local government reports. The Narcotics Division of the Security Bureau in Hong Kong reported a surge in the use of cocaine, cannabis, ketamine, and heroin combined with methamphetamine during the pandemic [27]. In April 2020, the Centre of Health Protection in Hong Kong also noted a 5.5% increase in alcoholic consumption since the outbreak [28]. This increase in substance use may have been fueled by misinformation circulated on social media, suggesting that alcohol or smoking could protect against COVID-19 [29]. Despite the well-established link between self-harm and substance misuse [30], further research is needed to determine the association between these behaviors and COVID-19 policies.

Several additional patterns emerged from this study. First, there was a decline in DSH-related ED visits in early 2020 and 2022 as well as peaking staying times (Multimedia Appendices 3-6), consistent with the findings from both local and international studies [31-33]. This decline may be attributed to the stigma surrounding mental health issues, which can complicate health-seeking behavior, delay necessary care, and result in more severe DSH presentation requiring readmission to the ED [32,34]. Second, a significant decrease in patients requiring social security suggests that patients performing DSH may come from higher socioeconomic backgrounds. However, socioeconomic profiles varied across epidemic waves [35], indicating the need for further research to validate these findings. Finally, patients performing DSH might have been triaged to less urgent categories due to the changes in ED policies aimed at managing overcrowded cases of infection. These additional findings suggest that patients performing DSH had reduced access to health care services during COVID-19, potentially leading to worsening conditions and long-term repercussions postpandemic.

Given the empirical evidence from this study and existing literature, continued surveillance post the COVID-19 pandemic is essential to determine whether a spillover effect occurred. This need is particularly pressing in light of recent changes in the Hong Kong government's policies such as the announcement of a liquor tax cut to promote trade in 2024 [36]. Ongoing research and surveillance will be crucial in developing targeted interventions, guiding policy-making, and enhancing mental health services. By closely monitoring these trends, we can better address the mental health challenges posed by the pandemic and its aftermath, ultimately improving outcomes for vulnerable populations.

Strengths and Limitations

This study has several strengths and limitations. A major strength was the large sample size, which enhances reliability of the findings. In addition, the trend, seasonal, and population factors were adjusted for in the ITS analysis. However, there are notable limitations as well. The absence of data from outpatient and private clinics may reduce the generalizability of our results. Furthermore, the use of *ICD-9-CM* codes to identify DSH cases could compromise the accuracy of our findings, as later versions of the *ICD* (*International Classification of Diseases*) included more specific codes for DSH, but only *ICD-9-CM* was available on CDARS. Also, the data captured patients who presented to the ED with DSH, but neglected those not admitted through the ED. Another limitation is the lack of consideration for other potential confounding variables, such as the changing social conditions in 2019. In addition, the use of linear interpolation to adjust population counts may introduce bias, as population changes during the pandemic could affect the results.

Conclusion

The results indicate increasing adjusted DSH incidence ratios among males and individuals with alcohol and drug-related mental health conditions. This underscores the need for continued surveillance beyond the COVID-19 pandemic to determine if there is a spill-over effect. Importantly, all stakeholders—including policy makers, psychiatrists, and health care professionals—should collaborate to raise awareness and provide support for these vulnerable groups in Hong Kong. Targeted interventions and policies are essential to address the mental health challenges exacerbated by the pandemic and ensure better outcomes for those affected.

XSL•FO
RenderX

Acknowledgments

This work was supported in part by AIR@InnoHK administered by Innovation and Technology Commission of Hong Kong.

Authors' Contributions

CJYC and LYFL contributed to conception and design of the study. CJYC, LYFL, and AKCW performed the study. CJYC, LYFL, SCLC, TFY, CYTL, ETCL, SRBY, and HCP carried out the data analysis. CJYC, LYFL, TFY, JYHW, CKHW, CKT, TY, THR, AKCW, and JWKH contributed to the revision and read and approved the submitted version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional tables.
[DOCX File, 44 KB - publichealth_v11i1e57500_app1.docx]

Multimedia Appendix 2
Ethical approval.
[DOCX File, 128 KB - publichealth_v11i1e57500_app2.docx]

Multimedia Appendix 3
The average observation room staying time between the years 2019 and 2022. The red highlight corresponds to the dip in the monthly deliberate self-harm–related emergency department visits seen in Figure 1.
[PNG File, 45 KB - publichealth_v11i1e57500_app3.png]

Multimedia Appendix 4
The average total staying time between the years 2019 and 2022. The red highlight corresponds to the dip in the monthly deliberate self-harm–related emergency department visits seen in Figure 1.
[PNG File, 41 KB - publichealth_v11i1e57500_app4.png]

Multimedia Appendix 5
The average cubicle time between the years 2019 and 2022. The red highlight corresponds to the dip in the monthly deliberate self-harm–related emergency department visits seen in Figure 1.
[PNG File, 41 KB - publichealth_v11i1e57500_app5.png]

Multimedia Appendix 6
The average waiting time to triage between the years 2019 and 2022. The red highlight corresponds to the dip in the monthly deliberate self-harm–related emergency department visits seen in Figure 1.
[PNG File, 41 KB - publichealth_v11i1e57500_app6.png]

References

1. Zhao SZ, Wong JYH, Luk TT, Wai AKC, Lam TH, Wang MP. Mental health crisis under COVID-19 pandemic in Hong Kong, China. Int J Infect Dis 2020 Nov;100:431-433. [doi: 10.1016/j.ijid.2020.09.030] [Medline: 32947051]
2. Choi EPH, Hui BPH, Wan EYF. Depression and Anxiety in Hong Kong during COVID-19. Int J Environ Res Public Health 2020 May 25;17(10):3740. [doi: 10.3390/ijerph17103740] [Medline: 32466251]
3. Ni MY, Yao XI, Leung KSM, et al. Depression and post-traumatic stress during major social unrest in Hong Kong: a 10-year prospective cohort study. Lancet 2020 Jan;395(10220):273-284. [doi: 10.1016/S0140-6736(19)33160-5]
4. Chair SY, Chien WT, Liu T, et al. Psychological Distress, Fear and Coping Strategies among Hong Kong People During the COVID-19 Pandemic. Curr Psychol 2023;42(3):2538-2557. [doi: 10.1007/s12144-021-02338-7] [Medline: 34690470]
5. Legislative Council of the Hong Kong Special Administrative Region of the People's Republic of China. Measures to promote mental health during COVID-19. 2021. URL: https://www.legco.gov.hk/research-publications/english/essentials-2021ise16-measures-to-promote-mental-health-during-covid-19.htm#endnote12 [accessed 2025-01-31]
6. The Hong Kong Special Administrative Region of the People's Republic of China. The Chief Executive's 2020 Policy Address. 2020. URL: https://www.policyaddress.gov.hk/2020/eng/p144.html [accessed 2025-01-31]

XSL•FO
RenderX

7. Nock MK, Mendes WB. Physiological arousal, distress tolerance, and social problem-solving deficits among adolescent self-injurers. J Consult Clin Psychol 2008 Feb;76(1):28-38. [doi: 10.1037/0022-006X.76.1.28] [Medline: 18229980]
8. Guerry JD, Prinstein MJ. Longitudinal prediction of adolescent nonsuicidal self-injury: examination of a cognitive vulnerability-stress model. J Clin Child Adolesc Psychol 2010;39(1):77-89. [doi: 10.1080/15374410903401195] [Medline: 20390800]
9. Kaess M, Hille M, Parzer P, Maser-Gluth C, Resch F, Brunner R. Alterations in the neuroendocrinological stress response to acute psychosocial stress in adolescents engaging in nonsuicidal self-injury. Psychoneuroendocrinology 2012 Jan;37(1):157-161. [doi: 10.1016/j.psyneuen.2011.05.009] [Medline: 21676550]
10. Liu RT, Frazier EA, Cataldo AM, Simon VA, Spirito A, Prinstein MJ. Negative life events and non-suicidal self-injury in an adolescent inpatient sample. Arch Suicide Res 2014;18(3):251-258. [doi: 10.1080/13811118.2013.824835] [Medline: 24712970]
11. Burke TA, Hamilton JL, Abramson LY, Alloy LB. Non-suicidal self-injury prospectively predicts interpersonal stressful life events and depressive symptoms among adolescent girls. Psychiatry Res 2015 Aug 30;228(3):416-424. [doi: 10.1016/j.psychres.2015.06.021] [Medline: 26165966]
12. Chai Y, Luo H, Yip PSF. Prevalence and risk factors for repetition of non-fatal self-harm in Hong Kong, 2002-2016: A population-based cohort study. Lancet Reg Health West Pac 2020 Sep;2:100027. [doi: 10.1016/j.lanwpc.2020.100027] [Medline: 34327378]
13. Isumi A, Doi S, Yamaoka Y, Takahashi K, Fujiwara T. Do suicide rates in children and adolescents change during school closure in Japan? The acute effect of the first wave of COVID-19 pandemic on child and adolescent mental health. Child Abuse Negl 2020 Dec;110(Pt 2):104680. [doi: 10.1016/j.chiabu.2020.104680] [Medline: 32847679]
14. O'Connor RC, Wetherall K, Cleare S, et al. Mental health and well-being during the COVID-19 pandemic: longitudinal analyses of adults in the UK COVID-19 Mental Health & Wellbeing study. Br J Psychiatry 2021 Jun;218(6):326-333. [doi: 10.1192/bjp.2020.212] [Medline: 33081860]
15. Robillard CL, Turner BJ, Ames ME, Craig SG. Deliberate self-harm in adolescents during COVID-19: The roles of pandemic-related stress, emotion regulation difficulties, and social distancing. Psychiatry Res 2021 Oct;304:114152. [doi: 10.1016/j.psychres.2021.114152] [Medline: 34371298]
16. Hawton K, Lascelles K, Brand F, et al. Self-harm and the COVID-19 pandemic: A study of factors contributing to self-harm during lockdown restrictions. J Psychiatr Res 2021 May;137:437-443. [doi: 10.1016/j.jpsychires.2021.03.028] [Medline: 33774538]
17. Bismark M, Scurrah K, Pascoe A, Willis K, Jain R, Smallwood N. Thoughts of suicide or self-harm among Australian healthcare workers during the COVID-19 pandemic. Aust N Z J Psychiatry 2022 Dec;56(12):1555-1565. [doi: 10.1177/00048674221075540] [Medline: 35128948]
18. Hospital Authority Hong Kong Special Administrative Region of the People's Republic of China. Accident & Emergency (A&E). URL: https://www.ha.org.hk/visitor/ha_serviceguide_details.asp?Content_ID=10051 [accessed 2025-01-31]
19. Hospital Authority Hong Kong Special Administrative Region of the People's Republic of China. Multi-tiered Triage and Treatment Strategy to battle a tsunami of COVID-19 cases. 2022. URL: https://www3.ha.org.hk/ehaslink/issue119/en/cover_story-2.html [accessed 2025-01-31]
20. Cheng MCY, Tong YH, Kwok TC, et al. Development journey of clinical data analysis and reporting system (CDARS) in hospital authority of Hong Kong. MedInfo. 2010.
21. Census and Statistics Department The Government of the Hong Kong Special Adminstrative Region. Table 1B: Population by Sex and Age [database on the Internet]. URL: https://www.censtatd.gov.hk/en/web_table.html?id=1B [accessed 2025-01-31]
22. Self-harm: assessment, management and preventing recurrence. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/guidance/ng225/chapter/Terms-used-in-this-guideline [accessed 2025-01-31]
23. Hospital Authority Hong Kong Special Administrative Region of the People's Republic of China. Waiving of Medical Charges (For Eligible Persons). URL: https://www.ha.org.hk/visitor/ha_visitor_index.asp?Content_ID=259365&Lang=ENG&Dimension=100 [accessed 2025-01-31]
24. Liu CH, Chen PC, Chen JH, Yeh CC. Changes in self-harm- and violence-related urgent psychiatric consultation in the emergency department during the different stages of the COVID-19 pandemic. BMC Psychiatry 2022 Jun 7;22(1):384. [doi: 10.1186/s12888-022-04029-4] [Medline: 35672738]
25. Rochelle TL. Health and help-seeking behavior among Chinese men in Hong Kong: The influence of culture. Psychol Men Masculinit 2019;20(1):71-81. [doi: 10.1037/men0000146]
26. Fong TCT, Chang K, Ho RTH. Association between quarantine and sleep disturbance in Hong Kong adults: The mediating role of COVID-19 mental impact and distress. Front Psychiatry 2023;14:1127070. [doi: 10.3389/fpsyt.2023.1127070] [Medline: 36926466]
27. Lam RPK. The impact of the COVID-19 pandemic on acute toxicity related to recreational drug abuse presenting to emergency departments. : Narcotics Division, Security Bureau, Hong Kong Special Administrative Region of the People's Republic of China; 2023.

XSL•FO
RenderX

28. The Centre for Health Protection DoH, The Government of Hong Kong Special Administrative Region. Alcohol Drinking, COVID-19 and NCDs. 2020.
29. Luk TT, Zhao S, Weng X, et al. Exposure to health misinformation about COVID-19 and increased tobacco and alcohol use: a population-based survey in Hong Kong. Tob Control 2021 Nov;30(6):696-699. [doi: 10.1136/tobaccocontrol-2020-055960] [Medline: 32855353]
30. Chai Y, Luo H, Wong GHY, et al. Risk of self-harm after the diagnosis of psychiatric disorders in Hong Kong, 2000-10: a nested case-control study. Lancet Psychiatry 2020 Feb;7(2):135-147. [doi: 10.1016/S2215-0366(20)30004-3] [Medline: 31974072]
31. Wai AKC, Wong CKH, Wong JYH, et al. Changes in Emergency Department Visits, Diagnostic Groups, and 28-Day Mortality Associated With the COVID-19 Pandemic: A Territory-Wide, Retrospective, Cohort Study. Ann Emerg Med 2022 Feb;79(2):148-157. [doi: 10.1016/j.annemergmed.2021.09.424] [Medline: 34742591]
32. Lopes S, Soares P, Gama A, et al. Association between perception of COVID-19 risk, confidence in health services and avoidance of emergency department visits: results from a community-based survey in Portugal. BMJ Open 2022 Jul 8;12(7):e058600. [doi: 10.1136/bmjopen-2021-058600] [Medline: 35803630]
33. DelPozo-Banos M, Lee SC, Friedmann Y, et al. Healthcare contacts with self-harm during COVID-19: An e-cohort whole-population-based study using individual-level linked routine electronic health records in Wales, UK, 2016-March 2021. PLoS ONE 2022;17(4):e0266967. [doi: 10.1371/journal.pone.0266967] [Medline: 35476839]
34. Maguire E, Kavalidou K, Bannan N, Doherty AM, Jeffers A. Substance use and self-harm emergency department presentations during COVID19: evidence from a National Clinical Programme for Self-Harm. Ir J Psychol Med 2023 Sep;40(3):424-429. [doi: 10.1017/ipm.2022.36] [Medline: 35915057]
35. Zhang A, Shi W, Tong C, et al. The fine-scale associations between socioeconomic status, density, functionality, and spread of COVID-19 within a high-density city. BMC Infect Dis 2022 Mar 21;22(1):274. [doi: 10.1186/s12879-022-07274-w] [Medline: 35313829]
36. The Hong Kong Special Administrative Region of the People's Republic of China. The Chief Executive's 2024 Policy Address. 2024. URL: https://www.policyaddress.gov.hk/2024/en/p65.html [accessed 2025-01-31]

Abbreviations

AIC: Akaike Information Criterion
aIR: adjusted incidence ratio
CDARS: Clinical Data Analysis and Reporting System
DSH: deliberate self-harm
ED: emergency department
***ICD*:** *International Classification of Diseases*
***ICD-9-CM*:** *International Classification of Diseases, Ninth Revision, Clinical Modification*
ITS: interrupted time series
oIR: observed incidence ratio

Edited by A Nakhostin-Ansari; submitted 19.02.24; peer-reviewed by A Costanza, AH Hoveidaei, MA Khadembashiri; revised version received 16.12.24; accepted 21.12.24; published 10.02.25.

Please cite as:
Luk LYF, Ching CJY, Yip TF, Chan SCL, Lam CYT, Lam ETC, Yue SRB, Pang HC, Wong JYH, Wong CKH, Tong CK, Yamamoto T, Rainer TH, Wai AKC, Ho JWK
Incidence of Deliberate Self-Harm in Hong Kong Before and During the COVID-19 Pandemic: Population-Wide Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e57500
URL: https://publichealth.jmir.org/2025/1/e57500
doi:10.2196/57500

© Luke Y F Luk, Christie J Y Ching, Tsz Fung Yip, Sunny C L Chan, Catherine Y T Lam, Elizabeth T C Lam, Serena R B Yue, Hoi Ching Pang, Janet Y H Wong, Carlos K H Wong, Chak Kwan Tong, Tafu Yamamoto, Timothy H Rainer, Abraham K C Wai, Joshua W K Ho. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic

XSL•FO
RenderX

information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Investigating the Trajectories of Poor Vision in Children and Adolescents in Wuhan, China From 2016 to 2019: Prospective Cohort Study

Lijuan Xu[1*], MM; Hanjia Li[2*], MPH; Fang Li[3*], MPH; Tinghui Zhang[2], BS; Jingyan Yan[2], MPH; Hong Yan[2], PhD; Lu He[1], MM; Bin Yu[2,4], PhD

[1]Physical Examination Center, Renmin Hospital of Wuhan University, Wuhan, China

[2]Department of Epidemiology and Health Statistics, School of Public Health, Wuhan University, 115 Donghu Rd, Wuchang District, Wuhan, China

[3]Wuhan Center for Disease Control and Prevention, Wuhan, China

[4]Population and Health Research Center, Wuhan University, Wuhan, China

[*]these authors contributed equally

Corresponding Author:
Bin Yu, PhD
Department of Epidemiology and Health Statistics, School of Public Health, Wuhan University, 115 Donghu Rd, Wuchang District, Wuhan, China

Abstract

Background: Poor vision is a challenging public health problem among children and adolescents globally and in China. It is well-recognized that early onset of poor vision and progressing to moderate and severe poor vision will increase the risk of irreversible blinding complications. To achieve the national goal of poor vision control and prevention, it is essential to investigate and understand the development of poor vision among children and adolescents in China.

Objective: This study aims to investigate the progression of poor vision among children and adolescents in Wuhan, China, based on a prospective cohort and to provide scientific evidence for the development and implementation of effective poor vision prevention and control programs.

Methods: Data were derived from a 4-year prospective cohort (2016 - 2019) of primary and middle school students (N=108,585) in Wuhan, China. Vision condition was measured using the standard logarithmic visual acuity charts. A group-based trajectory model was used to identify trajectories of poor vision overall and by gender and region.

Results: The mean age of the study subjects was 11.13 (SD 3.33) years, 200,110 (53.91%) were male and the majority (354404, 95.48%) were from urban areas. The prevalence of poor vision was 58.51% in 2016, 58.95% in 2017, 53.83% in 2018, and 54.79% in 2019. Group-based trajectory model identified 3 groups, including normal vision group (NVG) (27.4%), vision decline group (VDG) (17.8%), and moderate poor vision group (MPVG) (54.8%). A higher proportion of girls (57.8%) were in the MPV group compared to boys (50.5%), and the VDG showed greater changes in girls compared to boys. Furthermore, urban students (55.3%) had a higher proportion of MPV compared to rural students (47.5%), while urban students (17.2%) had a smaller proportion in the VDG compared to rural students (24%). Further analyses showed that as age increased, the likelihood of being categorized in the NVG decreased (β=−.417, *P*<.001), while the likelihood of being in the VDG (β=.058, *P*<.001) increased. Compared with boys, girls were more likely to be categorized in the VDG (β=.597, *P*<.001) and MPV group (β=.362, *P*<.001). Rural students were less likely than urban students to be categorized in the VDG (β=−.311, *P*<.001).

Conclusions: The prevalence of poor vision among children and adolescents in Wuhan has remained high over the years, with a slight decrease in recent years. The study identified three groups: normal vision, vision decline, and moderate poor vision. Girls and students from urban areas were more likely to have moderate poor vision, while boys and rural students had a higher proportion of vision decline. These findings provide valuable information for implementing poor vision prevention and control policies in the region.

(JMIR Public Health Surveill 2025;11:e53028) doi:10.2196/53028

KEYWORDS

children and adolescents; poor vision; group-based trajectory model; myopia; gender difference; China; normal vision group; vision decline group; moderate poor vision group; prevalence

XSL•FO
RenderX

Introduction

With the increasing prevalence of poor vision among children and adolescents in recent years, vision problems have been recognized as a major public health challenge globally in the 21st century [1] and in China [2]. Childhood and adolescence are well-accepted as the most vulnerable periods for poor visual development. Visual acuity reaches its full development by approximately 8 years of age, while visual cortex plasticity begins to diminish from the age of 2 years onward [1]. Poor vision can exert significant psychological, educational, and social impacts on children and adolescents, including psychological distress [3], perceived barriers [4], and poor academic achievement [5]. One systematic review and meta-analysis indicated that approximately one-third of children and adolescents were affected by myopia worldwide, and the number was projected to 740 million by 2050 [6] . Another review reported the prevalence of poor vision in children and adolescents as 7.26%, 7.34%, and 2.91% based on uncorrected visual acuity, presenting visual acuity and best corrected visual acuity, respectively, with refractive errors being the most common cause [7].

Compared with other countries, children and adolescents in East Asia reported a higher prevalence of poor vision. For example, data from a literature review indicated that the estimated prevalence of myopia among adolescents in East Asia was 69%, far greater than 16.7% in White adolescents and 14.3% in Hispanic or Latino adolescents [8]. Further, it is well-known that children and adolescents in China experience a significant burden of poor vision. Findings from a systematic review showed that studies from China reported a much higher prevalence of poor vision than those from other countries (13.33% - 41.17% in China vs 1.34% - 29.42% in other countries) [7]. Data from Chinese government indicated that the myopia rate among children and adolescents was 53.6% in 2018,decreased to 50.2% in 2019, and reversed to 52.7% in 2020 [9]. Additionally, data from a review showed that the prevalence of vision impairment in China increased from 19% at age 6 years to 66.9% at age 17 years [10]. To achieve the national goal of controlling and preventing poor vision, it is essential to investigate and understand its development among children and adolescents in China.

It is well-recognized that progression to moderate and severe poor vision increases the risk of irreversible blinding complications, such as retinal detachment, glaucoma, and myopic maculopathy [11]. Previous studies have shown that individuals with early-onset poor vision are at a higher risk of progressing to severe poor vision and vision impairment [12-14]. Therefore, it is valuable to identify the trends of poor vision progression among children and adolescents. However, to the best of our knowledge, most studies among children and adolescents in China have been cross-sectional, limiting the ability to examine the incidence and progression of poor vision [15,16]. Based on a prospective cohort, this study aimed to investigate the progression of poor vision among children and adolescents in China. Findings of the study will provide evidence for the future design and implementation of effective interventions and prevention programs for poor vision.

Methods

Participants and Sampling

Study data were derived from a prospective cohort of students in primary and middle schools in Wuchang District, Wuhan City, China. The cohort was developed based on a screening project that included all primary and middle schools in the Wuchang District. All students who were currently studying in the selected schools, with oral or written informed consent collected from the parents or guardians, were recruited for the study. Students who were absent from school on the test day or had severe eye diseases or conditions that may influence the testing results (eg, infections, surgeries) were excluded. Enrollment proportions were 91.10% in 2016, 91.43% in 2017, 95% in 2018, and 96.88% in 2019. The screening team from the Physical Examination Center, Renmin Hospital of Wuhan University visited each primary or middle school on a scheduled day. Schools prepared a classroom for the testing in advance, and students in that school entered the room one class at a time until all students were tested. The screening team completed the screening test for all students in the primary or middle schools in Wuchang District; the process was repeated every year between 2016 and 2019. Wuhan is one of the largest metropolitan cities in central China, with a population of 15 million and a per capita GDP (gross domestic product) of $25,000. Wuchang District is one of the 13 districts of Wuhan and ranks among the largest, with a population of 1.27 million [17]. We aimed to investigate the vision status among students and provide evidence and recommendations for decision-makers. Data from 2016 to 2019 were used in this study, with 89,872, 91,060, 93,391, and 96,849 students enrolled in each year, respectively.

Ethical Considerations

The original study was a free governmental screening project for which the informed consent was waived from the parents or guardians. This study involved secondary data analysis of screening data and was approved by the Medical Ethics Committee at Wuhan University School of Medicine (WHU-LFMD-IRB2023031). Informed consent for the secondary data analysis was waived, and all data were was deidentified. No compensation was provided to the participants.

Measurements

Vision Status Assessment

Participants' visual acuity was measured by the standard logarithmic visual acuity charts, which has been widely used in poor vision screening in ophthalmology clinics and schools for over 20 years in China [18]. Participants were asked to stand at a distance of 5 m from the visual acuity chart and identify the direction of the "E optotype." Visual acuity was recorded using the 5-point method; this is an original visual acuity recording method in China, which stipulates normal vision as ≥5.0 and no light perception as 0. The participant's visual acuity is the recorded value of visual acuity corresponding to the smallest row for which they can correctly identify more than half the directions of the "E" optotype [18]. All inspectors received professional training.

XSL•FO
RenderX

The screening results were categorized based on the requirements of the technical standard for physical examination for students [19]; normal vision was defined as 5.0 in both left and right eyes. In cases where the visual acuity of left and right eyes differed, the lower value was used for classification. Mild poor vision was defined as a visual acuity of 4.9, moderate poor vision as <4.9 but >4.5, and severe poor vision as ≤4.5.

Demographic Variables

Demographic variables, including age (in years), sex (male or female), and place of origin (urban or rural), were provided by the schools where the children and adolescents were enrolled.

Statistical Analysis

Descriptive analyses (eg, frequency, proportion, mean, standard deviation) were used to summarize the sample characteristics. The group-based trajectory model (GBTM) was used to analyze the longitudinal data, explore the heterogeneity within the study sample, and further identify different developmental trajectories for poor vision. The underlying assumption of the GBTM is that the overall study sample consists of several latent subgroups with distinct developmental trajectories or patterns. Findings from the GBTM could help researchers and decision-makers to have a deeper understanding of the variability within the population and how different individuals progress over time. The GBTM has been widely used in various areas, including behaviors, mental health [20], and medical conditions [21]. It is an application of finite mixture modeling, which can approximate different developmental trajectories with a finite number of groups through maximum likelihood estimation [22]. The model accommodates missing data for certain time points, as demonstrated in previous studies [23,24]. Here, $Y_i=\{y_{i1},y_{i2},...,y_{it}\}$ denotes individual i's longitudinal sequence of the kth measured over T=4 time periods, π_g denotes the probability that individual i belongs to g^{th} group, $P^g(Y_i)$ represents the conditional probability of Y_i for a given membership in group g. The probability of the individual i's longitudinal sequence is expressed by Equation 1.

(1)P(Yi)=∑g=1GπgPgYi

Furthermore, GBTM assumes that the random variables $y_{it}(t=1,2,...,T)$ are independent of each other when within groups. Therefore, we can use Equation 2 to definite $P^g(Y_i)$.

(2)PgYi=∏t=1TPgyit

We tested models with varying numbers of trajectory groups and orders of polynomial function, and then selected the best model based on four criteria [25-30] : (a) a lower absolute value of Bayesian Information Criteria , (b) the average posterior probability (Avepp) for class membership >0.7 for all groups, (c) all groups included at least 5% of the total sample, and (d) a simplified and rational model. Following the establishment of the foundational model, the incorporation of age as a time-dependent covariate, along with initial age, sex, and place of origin as independent covariates, was undertaken to ascertain the impact of these variables on the progression of visual acuity. All analyses were implemented using SAS (version 9.4; SAS Institute) and R software (version 4.2.1; R Foundation for Statistical Computing).

Results

Characteristics of the Study Sample

Results in Table 1 show that among the 89,872 children and adolescents in Wave 1 (2016), 48,431 (53.89%) were male, 83,102 (92.47%) were from urban areas, and the mean age of the sample was 11.24 (SD 3.33) years. Among the 91,060 participants in Wave 2 (2017), 49,056 (53.87%) were male, 85,826 (94.25%) were from urban areas, and the mean age of the sample was 11.16 (SD 3.32) years. Among the 93,391 participants in Wave 3 (2018), 50,350 (53.91%) were male, 90,503 (96.91%) were from urban areas, and the mean age of the sample was 11.09 (SD 3.33) years. In Wave 4 (2019), 52,273 (53.97%) were male, 94,973 (98.06%) were from urban areas, and the mean age of the sample was 11.04 (SD 3.33) years.

XSL•FO
RenderX

Table . Characteristics of the children and adolescents in Wuhan, China, between Wave 1 in 2016 - Wave 4 in 2019.

Variables	Number of participants between Wave I (2016) and Wave IV (2019)			
	Wave 1 (n=89,872)	Wave 2 (n=91,060)	Wave 3 (n=93,391)	Wave 4 (96,849)
Sex, n (%)				
Male	48,431 (53.89)	49,056 (53.87)	50,350 (53.91)	52,273 (53.97)
Female	41,441 (46.11)	42,004 (46.13)	43,041 (46.09)	44,576 (46.03)
Age (years), mean (SD)	11.24 (3.33)	11.16 (3.32)	11.09 (3.33)	11.04 (3.33)
Grade, n (%)				
Grade 1	9161 (10.19)	9350 (10.27)	10,460 (11.2)	10,659 (11.01)
Grade 2	8878 (9.88)	9224 (10.13)	9253 (9.91)	10,649 (11)
Grade 3	8663 (9.64)	8966 (9.85)	9140 (9.79)	9170 (9.47)
Grade 4	7788 (8.67)	8609 (9.45)	8844 (9.47)	9190 (9.49)
Grade 5	7953 (8.85)	7672 (8.43)	8667 (9.28)	8878 (9.17)
Grade 6	7690 (8.56)	7972 (8.75)	7774 (8.32)	8698 (8.98)
Grade 7	7777 (8.65)	7757 (8.52)	7762 (8.31)	7735 (7.99)
Grade 8	7665 (8.53)	7713 (8.47)	7640 (8.18)	7715 (7.97)
Grade 9	7671 (8.54)	7513 (8.25)	7552 (8.09)	7479 (7.72)
Grade 10	5795 (6.45)	5722 (6.28)	5692 (6.1)	6005 (6.2)
Grade 11	5715 (6.36)	5718 (6.28)	5625 (6.02)	5701 (5.89)
Grade 12	5116 (5.69)	4844 (5.32)	4979 (5.33)	4944 (5.11)
Region, n (%)				
Rural	6770 (7.53)	5234 (5.75)	2888 (3.09)	1876 (1.94)
Urban	83,102 (92.47)	85,826 (94.25)	90,503 (96.91)	94,973 (98.06)

Prevalence of Poor Vision

Results in Table 2 show that the prevalence of poor vision in the study sample was 58.51% in 2016, 58.95% in 2017, 53.83% in 2018, and 54.79% in 2019. The prevalence of poor vision from 2016 to 2019 ranged from 12.71% to 12.32% for mild poor vision, from 20.46% to 21.05% for moderate poor vision, and from 25.34% to 21.42% for severe poor vision. Additional information by gender and region is presented in Table 2.

XSL•FO
RenderX

Table . Prevalence of poor vision among children and adolescents in Wuhan, China, from 2016 to 2019.

Vision condition	Number of participants between 2016 and 2019			
	2016 (N=82,477)	2017 (N=90,396)	2018 (N=91,548)	2019 (N=93,068)
Overall, n (%)				
Normal	34,224 (41.49)	37,107 (41.05)	42,266 (46.17)	42,079 (45.21)
Mild poor vision	10,480 (12.71)	10,389 (11.49)	9211 (10.06)	11,467 (12.32)
Moderate poor vision	16,873 (20.46)	18,145 (20.07)	17,622 (19.25)	19,589 (21.05)
Severe poor vision	20,900 (25.34)	24,755 (27.39)	22,449 (24.52)	19,933 (21.42)
Males, n (%)				
Normal	19,737 (44.13)	21,303 (43.77)	24,142 (48.88)	23,969 (47.67)
Mild poor vision	5422 (12.12)	5382 (11.06)	4838 (9.8)	5928 (11.79)
Moderate poor vision	9075 (20.29)	9596 (19.72)	9320 (78.87)	10,402 (20.69)
Severe poor vision	10,491 (23.46)	12,387 (15.45)	11,087 (22.45)	9980 (19.85)
Females, n (%)				
Normal	14,487 (38.37)	15,804 (37.87)	18,124 (42.99)	18,110 (42.32)
Mild poor vision	5058 (13.4)	5007 (12)	4373 (10.37)	5539 (12.94)
Moderate poor vision	7798 (20.66)	8549 (20.49)	8302 (19.69)	9187 (21.47)
Severe poor vision	10,409 (27.57)	12,368 (29.64)	11362 (26.95)	9953 (23.26)
Urban students, n (%)				
Normal	30,985 (40.76)	34,530 (40.54)	40,703 (45.91)	41,251 (45.19)
Mild poor vision	9616 (12.65)	9894 (11.62)	8932 (10.07)	11,252 (12.33)
Moderate poor vision	15,567 (20.48)	17,081 (20.06)	16,997 (19.17)	19,179 (21.01)
Severe poor vision	19,852 (26.11)	23,665 (27.79)	22,034 (24.85)	19,596 (21.47)
Urban males, n (%)				
Normal	17,757 (43.37)	19,727 (43.23)	23,159 (48.54)	23,450 (47.62)
Mild poor	4940 (12.07)	5106 (11.19)	4681 (9.81)	5817 (11.81)
Moderate poor vision	8307 (20.29)	8991 (19.7)	8997 (18.86)	10,162 (20.64)
Severe poor vision	9938 (24.27)	11,811 (25.88)	10,879 (22.8)	9816 (19.93)
Urban females, n (%)				
Normal	13,228 (37.71)	14,803 (37.44)	17,544 (42.84)	17,801 (42.35)
Mild poor vision	4676 (13.33)	4788 (12.11)	4251 (10.38)	5435 (12.93)
Moderate poor vision	7260 (20.7)	8090 (20.46)	8000 (19.54)	9017 (21.45)
Severe poor vision	9914 (28.26)	11854 (29.98)	11,155 (27.24)	9780 (23.27)
Rural students, n (%)				
Normal	3239 (50.16)	2577 (49.31)	1563 (54.23)	828 (46.26)
Mild poor vision	864 (13.38)	495 (9.47)	279 (9.68)	215 (12.01)
Moderate poor vision	1306 (20.23)	1064 (20.36)	625 (21.69)	410 (22.91)
Severe poor vision	1048 (16.23)	1090 (20.86)	415 (14.4)	337 (18.83)
Rural males, n (%)				
Normal	1980 (52.34)	1576 (51.96)	983 (58.82)	519 (50.19)
Mild poor vision	482 (12.74)	276 (9.1)	157 (9.4)	111 (10.74)
Moderate poor vision	768 (20.3)	605 (19.95)	323 (19.33)	240 (23.21)
Severe poor vision	553 (14.62)	576 (18.99)	208 (12.45)	164 (15.86)
Rural, females, n (%)				

XSL•FO
RenderX

Vision condition	Number of participants between 2016 and 2019			
	2016 (N=82,477)	2017 (N=90,396)	2018 (N=91,548)	2019 (N=93,068)
Normal	1259 (47.08)	1001 (45.64)	580 (47.89)	309 (40.87)
Mild poor vision	382 (14.29)	219 (9.99)	122 (10.07)	104 (13.76)
Moderate poor vision	538 (20.12)	459 (20.93)	302 (24.94)	170 (22.49)
Severe poor vision	495 (18.51)	514 (23.44)	207 (17.09)	173 (22.88)

Trajectories of Poor Vision

Figure 1 shows the three distinct trajectories that were identified based on the lowest absolute Bayesian Information Criteria value (−361460.5). Group 1 (n=29752, 27.4%; Avepp=0.76) was characterized by a low degree of poor vision with a reduction in first 2 years and was named the normal vision group (NVG). Group 2 (n=19328, 17.8%; Avepp=0.72) was characterized by a rapid increase in poor vision and was named the vision decline group (VDG). Group 3 (n=59505, 54.8%; Avepp=0.93) was characterized by a stable level of moderate poor vision and was named the moderate poor vision group (MPV).

Figure 1. Trajectories of poor vision among children and adolescents in Wuhan, China, 2016 - 2019. Dotted lines represent the actual value, solid lines represent the predicted value. NVG: normal vision group, VDG: vision decline group, MPV: moderate poor vision group.

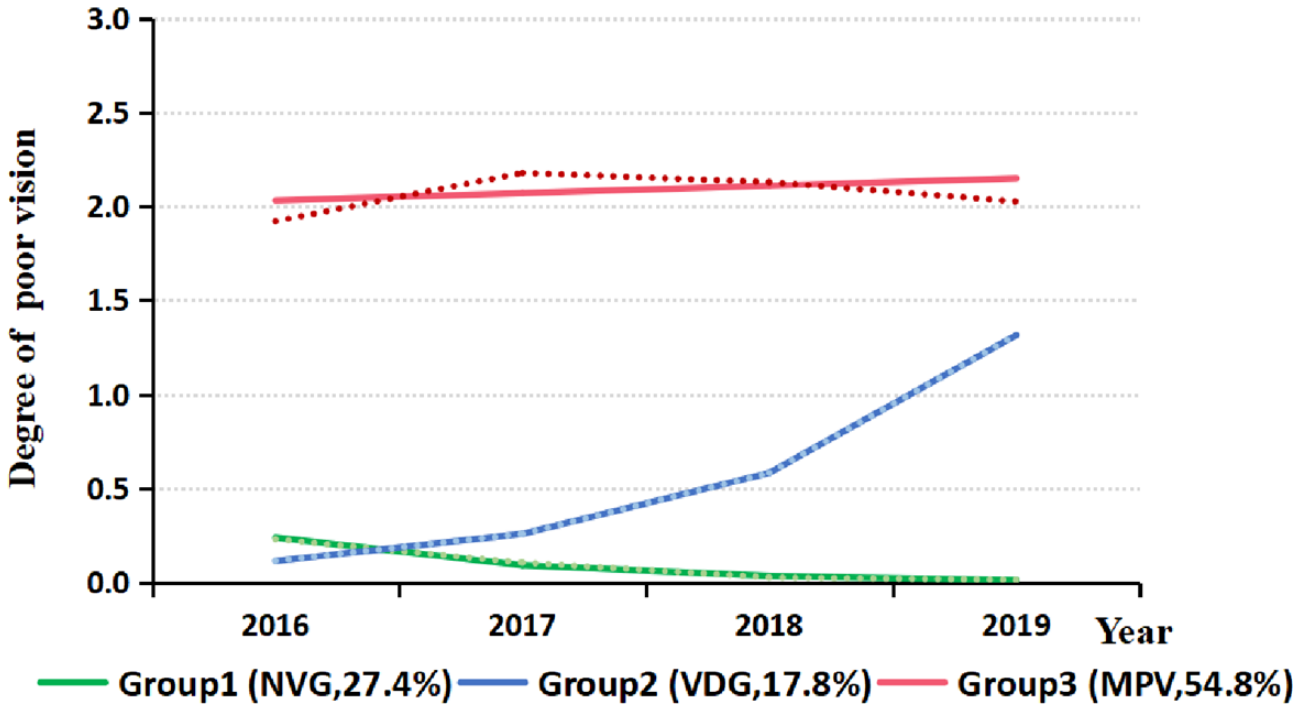


Figure 2 shows the trajectories of poor vision by sex and region, with a total of three trajectories identified. Compared to males, females had a higher proportion in Group 3 (57.8% vs 50.5%), and a lower proportion in Group 2 (17.8% vs 27.9%). Compared to rural areas, students in the urban areas had a higher proportion in Group 3 (55.3% vs 47.5%) and a lower proportion in Group 2 (17.2% vs 24.5%).

XSL•FO
RenderX

Figure 2. Trajectories of poor vision among children and adolescents in Wuhan, China, by sex and region between 2016-2019. Dotted lines represent the actual value, solid lines represent the predicted value. NVG: normal vision group; VDG: vision decline group; MPV: moderate poor vision group.

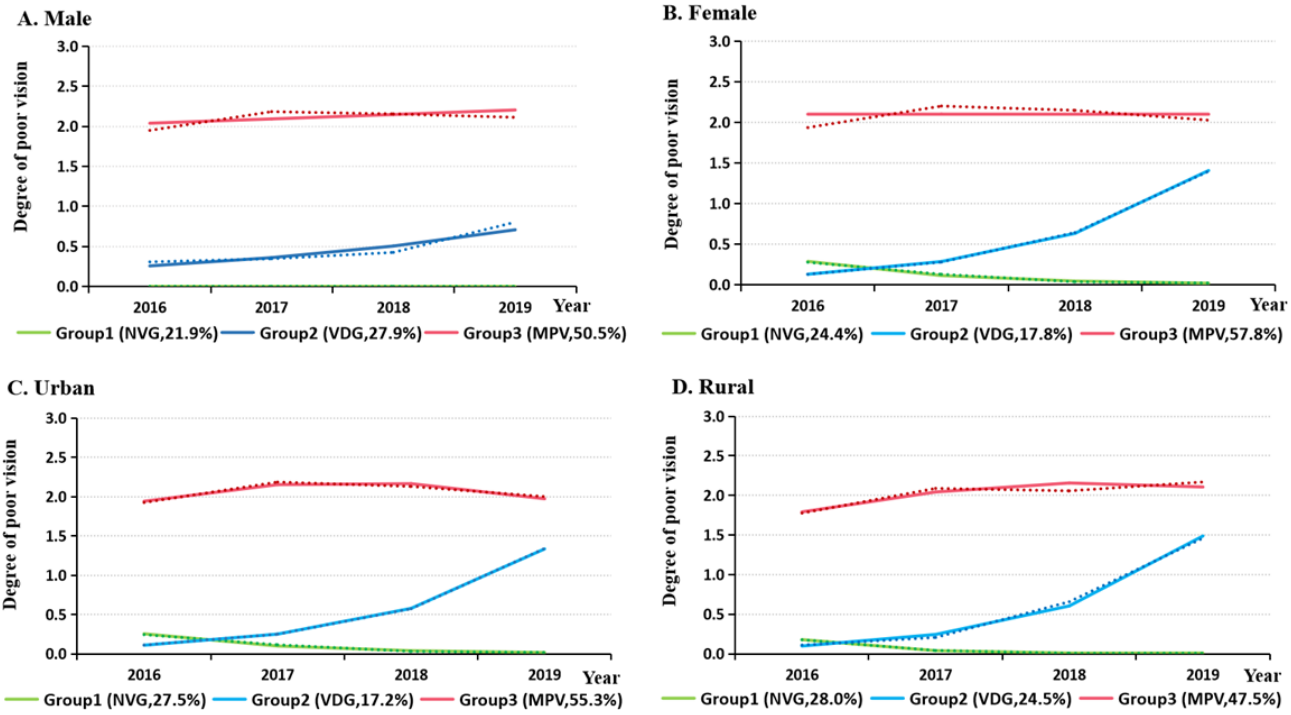
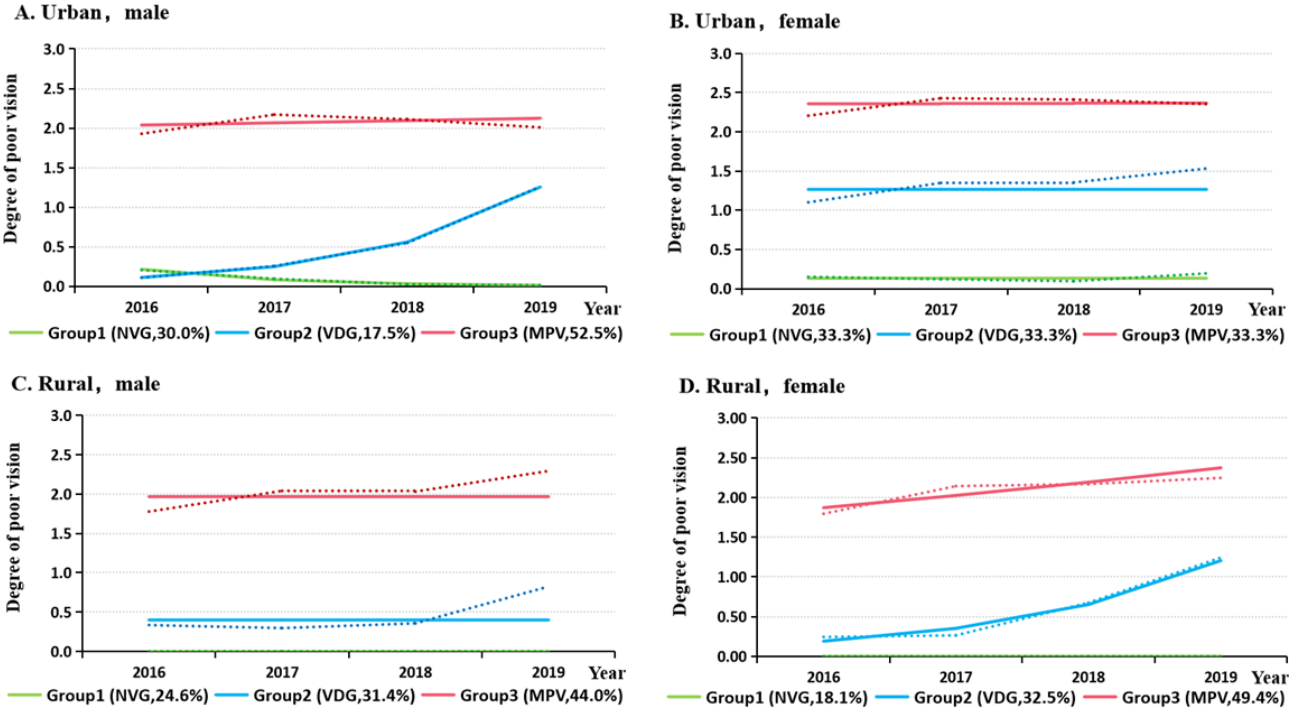


Figure 3 shows the trajectories of poor vision by the interaction of sex and region. The highest proportion in Group 3 was found among urban females (56.8%), while the lowest proportion was found in rural males (43.6%). The highest proportion in Group 2 was found among rural males (32.8%), followed by urban females (27.2%), rural females (25.8%), and urban males (17.5%).

Figure 3. Trajectories of poor vision among children and adolescents in Wuhan, China, 2016 - 2019, across sex and regions. Dotted lines represent the actual value, solid lines represent the predicted value. NVG: normal vision group, VDG: vision decline group, MPV: moderate poor vision group.



Factors Influencing Trajectory Membership and Progression

Data in Table 3 show that with increasing age, students were less likely to be categorized into Group 1 (β=−0.417, *P*<.001), and more likely to be categorized into Group 2 (β=.058,<.001). For group membership (Group 1 was used as reference), compared to boys, girls were more likely to be categorized into Group 2 (β=.597, *P*<.001) and Group 3 (β=.362, *P*<.001). Additionally, compared to urban students, students from rural

XSL•FO
RenderX

area were less likely to be categorized into Group 2 (β=−.311, *P*<.001). Further, with increasing initial enrollment age, students were more likely to be categorized into Group 1 (β=.369, *P*<.001) and Group 3 (β=.096, *P*<.001).

Table . Factors influencing trajectory membership and progression of poor vision among children and adolescents in Wuhan, China between 2016 and 2019.

Group	Variables	β	*t* test (*df*)[a]	*P* value
Trajectories progression				
Group 1[b]	Age	−.417	23.730 (85,820)	<.001
Group 2[c]	Age	.058	58.250 (85,820)	<.001
Group 3[d]	Age	.002	0.208 (85,820)	.84
Group membership (ref: Group 1)				
Group 2	Sex, (ref:Boys)			
	Girls	.597	22.709 (49,894)	<.001
	Region (ref:Urban)			
	Rural	−.311	−6.973 (5,092)	<.001
	Initial age	.369	33.711 (84,929)	<.001
Group 3	Sex (ref:Boys)			
	Females	.362	9.338 (49,994)	<.001
	Region (ref:Urban)			
	Rural	−.007	−0.113 (5,092)	.910
	Initial age	.096	5.391 (84,929)	<.001

[a]2-tailed t test was applied in the analysis.

[b]Group 1: normal vision group (NVG).

[c]Group 2: vision decline group (VDG).

[d]Group 3: moderate poor vision group (MPV).

Discussion

Principal Findings

This study investigated the progression of poor vision by analyzing data from a prospective cohort of primary and middle school students in Wuhan, China. To the best of our knowledge, this is the first study to use longitudinal data to identify the progression pattern of poor vision among children and adolescents. The findings of the study deepen our understanding of the development of poor vision and provide valuable evidence for devising and implementing intervention programs.

Population-based research indicates that poor vision in most school children results from uncorrected refractive errors [31,32]. Visual acuity has been widely used as a proxy measure for refractive error in children in previous studies [33-35] and is useful for estimating the prevalence of myopia in large populations that undergo vision screening when refraction is not feasible [36]. In this study, we found that the prevalence of poor vision among children and adolescents was 58.51% in 2016, 58.95% in 2017, 53.83% in 2018, and 54.79% in 2019, showing a decline in the latter two years. It is well-recognized that Chinese students are under greater academic pressure, spend more time on near-work activities and less time on outdoor activities, which may contribute to the high rates of poor vision [28]. The declines observed in the latter two years may be associated with the nationwide policy change regarding myopia prevention and control [29].

In this study, we identified 3 different development trajectories of poor vision among children and adolescents. Notably, more than half of the participants were categorized into the progression pattern of MPV during 2016 and 2019, amounting to 49,880 primary and middle school students in Wuchang district. If this pattern was applied to Wuhan City, approximately 589,135 primary and middle students would be fall into the MPV category. Meanwhile, our results show that 17.8% of the students enrolled were categorized into the VDG, which was characterized by a progression from normal vision to MPV. This finding suggests that the magnitude of vision loss is increasing every year. A study conducted by Ma et al [10] also showed that the prevalence rate of vision impairment was associated with grade and age in a nonlinear pattern. The underlying mechanisms of the rapid increasing pattern from normal to moderate poor vision may include high academic stress, long time of study, frequent use of electronic devices, lack of outdoor physical activities, among others [28,37-40]. These findings also highlight the urgent need for effective intervention and prevention programs to control poor vision development; potential actions include encouraging students to play outdoor sports, reducing their academic stress, and limiting time spent on electronic devices [41,42]. Further, we found that 27.4% of the participants maintained relatively normal vision

XSL•FO
RenderX

during the study period. Future studies should investigate the characteristics of this group and explore the potential protective factors against poor vision.

This study also suggests that being girls are at a higher risk of poor vision than boys, with the average poor vision rate being 59.61% in girls compared to 53.89% in boys, consistent with previous studies [26,43]. Moreover, a study on nationally representative samples conducted by Li and Mo [39] reported that 64.49% of girls were myopic while 54.44% of the boys. This trend was further confirmed in the trajectory analysis, where a greater proportion of females were classified into the MPV group compared to males. Although fewer girls belonged to the VDG, female children and adolescents experienced a greater decline in vision than males. The findings of this study indicated that girls had a greater risk of being assigned to both the VDG and the MPV than to the NVG; this result may be attributed to the fact that girls may spend less time on outdoor activities and more time on near-work tasks than boys [28,44,45].

Further, study findings suggest that urban students had a higher risk of poor vision than youth from rural areas, with a higher prevalence of poor vision and a greater proportion in the MPV category. Compared to urban students, rural children, and adolescents were less likely to be classified into the VDG than the NVG. Previous studies have similarly demonstrated higher rates of poor vision among urban students than those in rural areas [46]. This phenomenon could result from higher academic pressure, more time spent using electronic devices, and less outdoor activities among urban students [47,48].

Limitations

Although this study included a large population, it had several limitations. First, the results of the standard for logarithmic visual acuity charts were used as the sole criterion to determine the visual acuity of participants, and the cycloplegic refraction was not measured, which may have introduced measurement errors. We used the term poor vision instead of myopia throughout the paper to avoid the overestimation. Second, this study was only performed in a single district in Wuhan, and caution should be exercised when generalizing the findings to other areas. Third, there is a lack of data on demographic variables and risk factors for poor vision, that limited further analysis.

Conclusions

This study identified 3 groups with distinct poor vision progression patterns, including normal vision, visual decline, and MPV. Furthermore, the prevalence of poor vision was higher in girls than boys and higher in urban than rural students. Therefore, particular attention should be paid to girls and urban students for the prevention and control of poor vision among primary and secondary school students in the future.

Acknowledgments

The study was supported by Hubei Provincial Outstanding Young Talent Project for Public Health (PI: Dr. Bin Yu), Fundamental Research Funds for the Cross Innovation Talent Project of Renmin Hospital of Wuhan University (PI: Lijuan Xu), and the Funding for Scientific Research Projects from Wuhan Municipal Health Commission (WY22A07, PI: Dr. Fang Li).

Data Availability

The datasets generated during and/or analyzed during this study are not publicly available due to the surveillance policies from the government but are available from the corresponding author on reasonable request.

Authors' Contributions

Data curation: LX, HL
Statistical analysis: HL, FL, TZ, JY
Supervision: HY
Writing – original draft: LX, HL and FL
Writing – review & editing: BY, LH
Author Lu He is the co-corresponding author of this paper and can be reached at: Dr. Lu He, Physical Examination Center, Renmin Hospital of Wuhan University, Wuhan, Hubei Zhang Road (formerly Ziyang Road) Wuchang District No. 99 Jiefang Road 238, Wuhan, China, 86 13545211012; rm003619@whu.edu.cn.

Conflicts of Interest

None declared.

References

1. The Lancet Child Adolescent Health. Vision for the future. Lancet Child Adolesc Health 2021 Mar;5(3):155. [doi: 10.1016/S2352-4642(21)00029-8] [Medline: 33609489]
2. Pan CW, Wu RK, Li J, Zhong H. Low prevalence of myopia among school children in rural China. BMC Ophthalmol 2018 Jun 11;18(1):140. [doi: 10.1186/s12886-018-0808-0] [Medline: 29890943]

XSL•FO
RenderX

3. Pinquart M, Pfeiffer JP. Psychological well-being in visually impaired and unimpaired individuals: a meta-analysis. Brit J Vis Imp 2011 Jan;29(1):27-45. [doi: 10.1177/0264619610389572]
4. Greguol M, Gobbi E, Carraro A. Physical activity practice among children and adolescents with visual impairment--influence of parental support and perceived barriers. Disabil Rehabil 2015;37(4):327-330. [doi: 10.3109/09638288.2014.918194] [Medline: 24828394]
5. Horvat-Gitsels LA, Cortina-Borja M, Rahi JS. Do adolescents with impaired vision have different intentions and ambitions for their education, career and social outcomes compared to their peers? Findings from the Millennium Cohort Study. Br J Ophthalmol 2023 Dec 18;108(1):159-164. [doi: 10.1136/bjo-2021-320972] [Medline: 36307166]
6. Liang J, Pu Y, Chen J, et al. Global prevalence, trend and projection of myopia in children and adolescents from 1990 to 2050: a comprehensive systematic review and meta-analysis. Br J Ophthalmol 2024 Sep 24:bjo-2024-325427. [doi: 10.1136/bjo-2024-325427] [Medline: 39317432]
7. Yekta A, Hooshmand E, Saatchi M, et al. Global prevalence and causes of visual impairment and blindness in children: a systematic review and meta-analysis. J Curr Ophthalmol 2022;34(1):1-15. [doi: 10.4103/joco.joco_135_21] [Medline: 35620376]
8. Rudnicka AR, Kapetanakis VV, Wathern AK, et al. Global variations and time trends in the prevalence of childhood myopia, a systematic review and quantitative meta-analysis: implications for aetiology and early prevention. Br J Ophthalmol 2016 Jul;100(7):882-890. [doi: 10.1136/bjophthalmol-2015-307724]
9. Transcript of the national health commission's press conference. China Government Network, Propaganda. 2022. URL: http://www.nhc.gov.cn/xcs/s3574/202207/16833fe2630e49338195edd979b6d386.shtml [accessed 2023-06-05]
10. Ma Y, Zhang X, He F, et al. Visual impairment in rural and migrant Chinese school-going children: prevalence, severity, correction and associations. Br J Ophthalmol 2022 Feb;106(2):275-280. [doi: 10.1136/bjophthalmol-2020-317072] [Medline: 33127829]
11. Hu Y, Ding X, Guo X, Chen Y, Zhang J, He M. Association of age at myopia onset with risk of high myopia in adulthood in a 12-year follow-up of a Chinese cohort. JAMA Ophthalmol 2020 Nov 1;138(11):1129-1134. [doi: 10.1001/jamaophthalmol.2020.3451] [Medline: 32940622]
12. Chua SYL, Sabanayagam C, Cheung YB, et al. Age of onset of myopia predicts risk of high myopia in later childhood in myopic Singapore children. Ophthalmic Physiol Opt 2016 Jul;36(4):388-394. [doi: 10.1111/opo.12305] [Medline: 27350183]
13. Williams KM, Hysi PG, Nag A, Yonova-Doing E, Venturini C, Hammond CJ. Age of myopia onset in a British population-based twin cohort. Ophthalmic Physiol Opt 2013 May;33(3):339-345. [doi: 10.1111/opo.12042] [Medline: 23510367]
14. Gwiazda J, Hyman L, Dong LM, et al. Factors associated with high myopia after 7 years of follow-up in the Correction of Myopia Evaluation Trial (COMET) Cohort. Ophthalmic Epidemiol 2007;14(4):230-237. [doi: 10.1080/01658100701486459] [Medline: 17896302]
15. Li SM, Wei S, Atchison DA, et al. Annual Incidences and Progressions of Myopia and High Myopia in Chinese Schoolchildren Based on a 5-Year Cohort Study. Invest Ophthalmol Vis Sci 2022 Jan 3;63(1):8. [doi: 10.1167/iovs.63.1.8] [Medline: 34989760]
16. Wang SK, Guo Y, Liao C, et al. Incidence of and factors associated with myopia and high myopia in Chinese children, based on refraction without cycloplegia. JAMA Ophthalmol 2018 Sep 1;136(9):1017-1024. [doi: 10.1001/jamaophthalmol.2018.2658] [Medline: 29978185]
17. Communiqué of the seventh national population census of Wuhan city [Chinese]. Wuhan Bureau of Statistics, Census Column. 2021. URL: http://tjj.wuhan.gov.cn/ztzl_49/pczl/202109/t20210916_1779157.shtml [accessed 2023-06-05]
18. Standard no.GB/T 11533-2011 [Chinese]. National Standards Full Text Public System. 2011. URL: https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=A9F9E03A346211223DE34421A85CA1C8 [accessed 2023-06-05]
19. Ministry of Health of the PRC. Technical Standard for Physical Examination for Students: GB/T26343 — 2010: Standards Press of China; 2011. URL: https://www.codeofchina.com/standard/GBT26343-2010.html [accessed 2023-06-06]
20. Wei L, Tian J, Guo M, et al. Trajectories of sexual risk behaviors and the associated factors among young men who have sex with men in China. Front Public Health 2022;10:854616. [doi: 10.3389/fpubh.2022.854616] [Medline: 35387185]
21. Beltrame TB, de Oliveira AC, Manera F, Fernandes RC, Castilho BM, Höfelmann DA. Trajectory of blood pressure levels and weight gain in pregnant women: A group-based approach. Nutr Metab Cardiovasc Dis 2023 Oct;33(10):2019-2027. [doi: 10.1016/j.numecd.2023.06.006] [Medline: 37500346]
22. Haviland A, Nagin DS, Rosenbaum PR, Tremblay RE. Combining group-based trajectory modeling and propensity score matching for causal inferences in nonexperimental longitudinal data. Dev Psychol 2008 Mar;44(2):422-436. [doi: 10.1037/0012-1649.44.2.422] [Medline: 18331133]
23. Le Brocque RM, Dow BL, McMahon H, et al. The course of posttraumatic stress in Children: examination of symptom trajectories and predictive factors following admission to pediatric intensive care. Pediatr Crit Care Med 2020 Jul;21(7):e399-e406. [doi: 10.1097/PCC.0000000000002316] [Medline: 32224826]
24. Crothers A, Haeusler GM, Slavin MA, et al. Examining health-related quality of life in pediatric cancer patients with febrile neutropenia: factors predicting poor recovery in children and their parents. EClinicalMedicine 2021 Oct;40:101095. [doi: 10.1016/j.eclinm.2021.101095] [Medline: 34746716]

XSL•FO
RenderX

25. Kiadaliri A, Englund M. Variability in end-of-life healthcare use in patients with osteoarthritis: a population-based matched cohort study. Osteoarthr Cartil 2021 Oct;29(10):1418-1425. [doi: 10.1016/j.joca.2021.07.001]
26. Schuster AK, Krause L, Kuchenbäcker C, et al. Prevalence and time trends in myopia among children and adolescents. Dtsch Arztebl Int 2020 Dec 11;117(50):855-860. [doi: 10.3238/arztebl.2020.0855] [Medline: 33612155]
27. Yotsukura E, Torii H, Inokuchi M, et al. Current prevalence of myopia and association of myopia with environmental factors among schoolchildren in Japan. JAMA Ophthalmol 2019 Nov 1;137(11):1233-1239. [doi: 10.1001/jamaophthalmol.2019.3103] [Medline: 31415060]
28. Wang J, Ying GS, Fu X, et al. Prevalence of myopia and vision impairment in school students in Eastern China. BMC Ophthalmol 2020 Jan 2;20(1):2. [doi: 10.1186/s12886-019-1281-0] [Medline: 31898504]
29. The notice on the issue of implementation plan for comprehensive prevention and control of myopia in children and adolescents. Ministry of Education of the People's Republic of China. 2018. URL: http://www.moe.gov.cn/srcsite/A17/moe_943/s3285/201808/t20180830_346672.html [accessed 2024-11-06]
30. Yang J, Huang A, Chen D. Characteristics of physical developmental trajectories of infants and children aged 0-3 years based on cluster trajectory model and influencing factors. Chinese Journal of Reproductive 2023;34(1):25-34.
31. Fan DSP, Lam DSC, Lam RF, et al. Prevalence, incidence, and progression of myopia of school children in Hong Kong. Invest Ophthalmol Vis Sci 2004 Apr;45(4):1071-1075. [doi: 10.1167/iovs.03-1151] [Medline: 15037570]
32. Goh PP, Abqariyah Y, Pokharel GP, Ellwein LB. Refractive error and visual impairment in school-age children in Gombak District, Malaysia. Ophthalmology 2005 Apr;112(4):678-685. [doi: 10.1016/j.ophtha.2004.10.048] [Medline: 15808262]
33. Tay MT, Au Eong KG, Ng CY, Lim MK. Myopia and educational attainment in 421,116 young Singaporean males. Ann Acad Med Singap 1992 Nov;21(6):785-791. [Medline: 1295418]
34. Au Eong KG, Tay TH, Lim MK. Education and myopia in 110,236 young Singaporean males. Singapore Med J 1993 Dec;34(6):489-492. [Medline: 8153707]
35. Xiang F, He M, Zeng Y, Mai J, Rose KA, Morgan IG. Increases in the prevalence of reduced visual acuity and myopia in Chinese children in Guangzhou over the past 20 years. Eye (Lond) 2013 Dec;27(12):1353-1358. [doi: 10.1038/eye.2013.194] [Medline: 24008929]
36. Sun HP, Li A, Xu Y, Pan CW. Secular trends of reduced visual acuity from 1985 to 2010 and disease burden projection for 2020 and 2030 among primary and secondary school students in China. JAMA Ophthalmol 2015 Mar;133(3):262-268. [doi: 10.1001/jamaophthalmol.2014.4899] [Medline: 25429523]
37. Shen DY, Li XH. Influencing factors of myopia among primary and secondary school students in Shenzhen. Chinese J School Health 2020;41:583-587.
38. Bian H. Investigation on the current situation of myopia among pupils in Baotou City and analysis of related factors. Int Eye Sci 2021;21:1834-1838.
39. Li H, Mo J. Myopia and risk factors among Chinese adolescents. Modern preventive medicine (14):2552-2557 [FREE Full text]
40. Guan H, Yu NN, Wang H, et al. Impact of various types of near work and time spent outdoors at different times of day on visual acuity and refractive error among Chinese school-going children. PLoS One 2019;14(4):e0215827. [doi: 10.1371/journal.pone.0215827] [Medline: 31026279]
41. Morgan IG, French AN, Ashby RS, et al. The epidemics of myopia: aetiology and prevention. Prog Retin Eye Res 2018 Jan;62:134-149. [doi: 10.1016/j.preteyeres.2017.09.004] [Medline: 28951126]
42. Li SM, Ran AR, Kang MT, et al. Effect of text messaging parents of school-aged children on outdoor time to control myopia: a randomized clinical trial. JAMA Pediatr 2022 Nov 1;176(11):1077-1083. [doi: 10.1001/jamapediatrics.2022.3542] [Medline: 36155742]
43. Yin C, Gan Q, Xu P, et al. Dietary patterns and associations with myopia in Chinese children. Nutrients 2023 Apr 18;15(8):1946. [doi: 10.3390/nu15081946]
44. Guo L, Yang J, Mai J, et al. Prevalence and associated factors of myopia among primary and middle school-aged students: a school-based study in Guangzhou. Eye (Lond) 2016 Jun;30(6):796-804. [doi: 10.1038/eye.2016.39] [Medline: 26965016]
45. Yin Y, Qiu C, Qi Y. Myopia in Chinese adolescents: its influencing factors and correlation with physical activities. Comput Math Methods Med 2022 Aug 24;2022:1-10. [doi: 10.1155/2022/4700325]
46. Huang J, Zhang Y, Du K, et al. Prevalence of poor vision and eyeglass wear among primary and high school students in urban and rural western China: a comparative study. Chin J Public 2020;36(6):849-852. [doi: 10.11847/zgggws1122103]
47. Pan CW, Ramamurthy D, Saw SM. Worldwide prevalence and risk factors for myopia. Ophthalmic Physiol Opt 2012 Jan;32(1):3-16. [doi: 10.1111/j.1475-1313.2011.00884.x] [Medline: 22150586]
48. Pan CW, Dirani M, Cheng CY, Wong TY, Saw SM. The age-specific prevalence of myopia in Asia: a meta-analysis. Optom Vis Sci 2015 Mar;92(3):258-266. [doi: 10.1097/OPX.0000000000000516] [Medline: 25611765]

Abbreviations

Avepp: average posterior probability
GBTM: Group-based trajectory model

XSL•FO
RenderX

GDP: gross domestic product
MPV: moderate poor vision
NVG: normal vision group
VDG: vision decline group

Edited by A Mavragani; submitted 23.09.23; peer-reviewed by L Marković, W Wang; revised version received 01.01.25; accepted 02.01.25; published 18.02.25.

Please cite as:
Xu L, Li H, Li F, Zhang T, Yan J, Yan H, He L, Yu B
Investigating the Trajectories of Poor Vision in Children and Adolescents in Wuhan, China From 2016 to 2019: Prospective Cohort Study
JMIR Public Health Surveill 2025;11:e53028
URL: https://publichealth.jmir.org/2025/1/e53028
doi:10.2196/53028

© Lijuan Xu, Hanjia Li, Fang Li, Tinghui Zhang, Jingyan Yan, Hong Yan, Lu He, Bin Yu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Data-Driven Clinical Subgroups for Cervical Cancer Prevention With Machine Learning: Population-Based, External, and Diagnostic Validation Study

Zhen Lu[1], MSc; Binhua Dong[2,3], MSc; Hongning Cai[4], MSc; Tian Tian[5], PhD; Junfeng Wang[6], PhD; Leiwen Fu[7], PhD; Bingyi Wang[8,9], PhD; Weijie Zhang[1], BD; Shaomei Lin[10], BD; Xunyuan Tuo[11], MSc; Juntao Wang[12], MSc; Tianjie Yang[13], MSc; Xinxin Huang[14], MSc; Zheng Zheng[13], MSc; Huifeng Xue[15], MSc; Shuxia Xu[16], BD; Siyang Liu[1], PhD; Pengming Sun[2,3,17], MD, PhD; Huachun Zou[18,19,20], PhD

[1]School of Public Health (Shenzhen), Sun Yat-sen University, Shenzhen, China

[10]Department of Gynecology, Shunde Women's and Children's Hospital of Guangdong Medical University, Foshan, China

[11]Department of Gynecology, Gansu Provincial Maternity and Child-care Hospital, Lanzhou, China

[12]Department of Gynecology, Guiyang Maternal and Child Health Care Hospital, Guiyang, China

[13]Department of Gynecology, Shenzhen Maternity & Child Healthcare Hospital, Shenzhen, China

[14]The Ministry of Health, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[15]Center for Cervical Disease Diagnosis and Treatment, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[16]Department of Pathology, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[17]School of Group Medicine and Public Health, Peking Union Medical College, Beijing, China

[18]School of Public Health, Fudan University, Shanghai, China

[19]Shenzhen Campus, Sun Yat-sen University, Shenzhen, China

[2]Department of Gynecology, Laboratory of Gynecologic Oncology, Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics & Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[20]Fujian Maternity and Child Health Hospital, College of Clinical Medicine for Obstetrics and Gynecology and Pediatrics, Fujian Medical University, Fuzhou, China

[3]Fujian Key Laboratory of Women and Children’s Critical Diseases Research, Fuzhou, China

[4]Department of Gynecology, Maternal and Child Health Hospital of Hubei Province (Women and Children's Hospital of Hubei Province) Wuhan, Wuhan, China

[5]School of Public Health, Xinjiang Medical University, Urumqi, China

[6]Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, Netherlands

[7]Beijing Chest Hospital, Capital Medical University, Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing, China

[8]Institute for HIV/AIDS Control and Prevention, Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China

[9]Department of HIV/AIDS Control and Prevention, Guangdong Provincial Academy of Preventive Medicine, Guangzhou, China

Corresponding Author:
Huachun Zou, PhD
School of Public Health, Fudan University, Shanghai, China

Abstract

Background: Cervical cancer remains a major global health issue. Personalized, data-driven cervical cancer prevention (CCP) strategies tailored to phenotypic profiles may improve prevention and reduce disease burden.

Objective: This study aimed to identify subgroups with differential cervical precancer or cancer risks using machine learning, validate subgroup predictions across datasets, and propose a computational phenomapping strategy to enhance global CCP efforts.

Methods: We explored the data-driven CCP subgroups by applying unsupervised machine learning to a deeply phenotyped, population-based discovery cohort. We extracted CCP-specific risks of cervical intraepithelial neoplasia (CIN) and cervical cancer through weighted logistic regression analyses providing odds ratio (OR) estimates and 95% CIs. We trained a supervised machine learning model and developed pathways to classify individuals before evaluating its diagnostic validity and usability on an external cohort.

XSL•FO
RenderX

Results: This study included 551,934 women (median age, 49 years) in the discovery cohort and 47,130 women (median age, 37 years) in the external cohort. Phenotyping identified 5 CCP subgroups, with CCP4 showing the highest carcinoma prevalence. CCP2–4 had significantly higher risks of CIN2+ (CCP2: OR 2.07 [95% CI: 2.03 - 2.12], CCP3: 3.88 [3.78 - 3.97], and CCP4: 4.47 [4.33 - 4.63]) and CIN3+ (CCP2: 2.10 [2.05 - 2.14], CCP3: 3.92 [3.82 - 4.02], and CCP4: 4.45 [4.31 - 4.61]) compared to CCP1 (P<.001), consistent with the direction of results observed in the external cohort. The proposed triple strategy was validated as clinically relevant, prioritizing high-risk subgroups (CCP3-4) for colposcopies and scaling human papillomavirus screening for CCP1-2.

Conclusions: This study underscores the potential of leveraging machine learning algorithms and large-scale routine electronic health records to enhance CCP strategies. By identifying key determinants of CIN2+/CIN3+ risk and classifying 5 distinct subgroups, our study provides a robust, data-driven foundation for the proposed triple strategy. This approach prioritizes tailored prevention efforts for subgroups with varying risks, offering a novel and scalable tool to complement existing cervical cancer screening guidelines. Future work should focus on independent external and prospective validation to maximize the global impact of this strategy.

(JMIR Public Health Surveill 2025;11:e67840) doi:10.2196/67840

KEYWORDS

cervical cancer; human papillomavirus; screening; machine learning; cervical tumor; cancer; carcinoma; tumor; malignant; ML; phenomapping strategy; logistic regression; regression; population-based; validation study; cancer prevention; validity; usability; algorithm; surveillance; electronic health record; EHR

Introduction

Cervical cancer is the fourth most common cancer among women, with an estimated 660,000 new cases and 350,000 deaths globally in 2022 [1,2]. It is the most common cancer in 25 countries and the leading cause of cancer death in 37 countries. Despite being largely preventable through human papillomavirus (HPV) vaccination [3], the high incidence and persistence of high-risk HPV (hrHPV) remain the primary risk factors for cervical cancer [4] and related diseases [5-7]. In November 2020, the World Health Organization (WHO) launched a global initiative [8] to eliminate cervical cancer as a public health problem, emphasizing a triple intervention strategy: vaccinating at least 90% of girls against HPV by age 15 years, screening 70% of women with a high-performance test by ages 35 and 45 years, and treating at least 90% of detected precancerous lesions and invasive cancers. Yet, globally, an estimated 1.6 billion (67%) of 2.3 billion women aged 20 - 70 years have never been screened for cervical cancer [9], and in China, 5-year screening coverage among women aged 35 - 49 years was only 33%. Establishing high-quality, sustainable, and acceptable cervical cancer prevention (CCP) with broad coverage in resource-limited regions remains a critical challenge [10-12]. While HPV-based screening has demonstrated significant benefits [13-18], leading the WHO to recommend it as the primary method for CCP globally [19], its adoption in low- and middle-income countries (LMICs) is limited due to health inequities, resource constraints, and limited access to affordable, clinically validated HPV tests [9,20]. Addressing these barriers requires personalized CCP strategies.

We hypothesize that data-driven CCP subgroups, identified using machine learning, can provide valuable insights into CCP, enabling personalized strategies to reduce the risk of cervical intraepithelial neoplasia (CIN) and cervical cancer. Personalized CCP strategies depend on the unique phenotypic profile of each individual. The population, with its varying phenotypic diversity, comprises heterogeneous subgroups that reflect multiple underlying behavior patterns and causes of disease. Clinical practice guidelines also recommend that cervical cancer screening should be tailored to an individual's risk profile of HPV infection [21,22]. Therefore, the discovery of CCP subgroups could inform and improve the development of new screening strategies, public health policies, and clinical decision-making, as well as contribute to trial design, despite the complex causal relationships among individual risk factors [4]. To date, no study has identified CCP subgroups. The incomplete understanding of these subgroups across large, diverse populations, coupled with insufficient validation through various methods, has hindered the implementation of the triple intervention strategy. Machine learning-based approaches allow for the phenotyping of entire populations based on individual characteristics, enabling the prediction of phenotypic clusters and disease onset [23,24]. Additionally, by defining a computational phenomap—a mathematical construct of individual phenotypes based on baseline measures [24]—computational methods can assess the heterogeneous risk effects of distinct subpopulations. This approach accounts for the phenotypic diversity within populations and their subsequent histopathological diagnoses.

As such, in this study, we aimed to: (1) test the hypothesis that distinct CCP subgroups exhibit differential risks of cervical precancer or cancer based on their complex phenotypic profiles (development); (2) demonstrate internal validity (within a dataset and across methods), as well as external and diagnostic validity (across datasets; validation); and (3) propose a computational phenomapping strategy with clinical relevance and pathways to improve global access to CCP (impact).

Methods

Data Source

Deidentified data were extracted from electronic health records (EHRs) of the national cervical cancer screening program in China. In summary, our study included eligible women aged

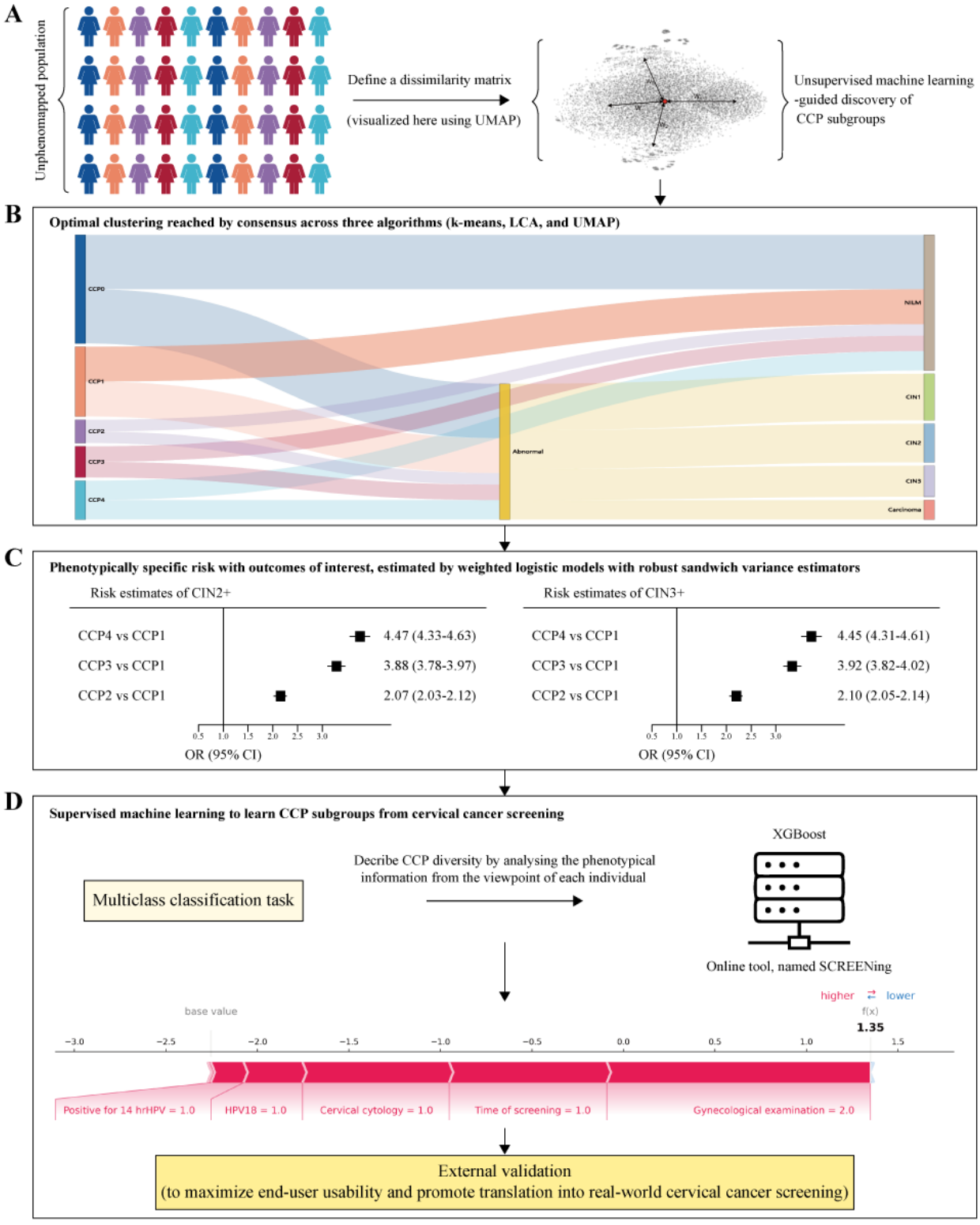
XSL•FO
RenderX

25 - 65 years who participated in the cervical cancer screening. Data from Fujian Province (2014 - 2023) were used to establish a discovery cohort to train the models. Additionally, EHR data from 5 other regions—Shenzhen City, Foshan City, Hubei Province, Gansu Province, and Guizhou Province—were employed as an external cohort to validate the generalizability of the models across diverse populations. Details on the study design, as illustrated in Figure 1, are available in Appendix S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 1. Study design. (**A**) We explored the data-driven cervical cancer prevention (CCP) subgroups by applying unsupervised machine learning to a deeply phenotyped, population-based discovery cohort. In this population-based, external, and diagnostic validation study, we aimed to use different machine learning methods to test our hypothesis (**B**) that individuals within distinct CCP subgroups exhibit differential risks of cervical precancer or cancer based on their complex phenotypic profiles. After identifying CCP subgroups, we extracted CCP-specific risks of outcomes of interest (**C**) conditionally on all predefined and algorithmically selected features through weighted logistic regression analyses providing average odds ratio (OR) estimates. Additionally, we stratified individuals based on their key features to conduct subgroup analyses. Finally, we trained a supervised machine learning model and developed pathways (**D**) to classify individuals using features most consistently linked to CCP subgroups before evaluating its diagnostic validity and usability on an external cohort. Pathologically abnormal diagnoses consisted of CIN1, CIN2, CIN3, and carcinoma. UMAP: uniform manifold approximation and projection; CCP: cervical cancer prevention; LCA: latent class analysis; NILM: negative for intraepithelial lesion or malignancy; CIN1: cervical intraepithelial neoplasia grade 1; CIN2: cervical intraepithelial neoplasia grade 2; CIN3: cervical intraepithelial neoplasia grade 3; CIN2+: cervical intraepithelial neoplasia grade 2 or worse; CIN3+: cervical intraepithelial neoplasia grade 3 or worse; OR: odds ratio; XGBoost: eXtreme gradient boosting; SCREENing: clinical Subgroups for CeRvical cancEr prEvention using computational phenomaps and machine learning; hrHPV: high-risk human papillomavirus; HPV: human papillomavirus.

XSL•FO
RenderX



Ethical Considerations

This study was approved by the Ethics Committee of Fujian Maternal and Child Health Hospital (2023KY141). No additional informed consent was required for this analysis. All participant data were deidentified prior to analysis. No images of identifiable individuals are included in the manuscript or supplementary materials.

Data Preprocessing

In line with the consensus that detecting and treating CIN2 or CIN3 [25,26], key premalignant cervical conditions, can prevent progression to invasive cervical cancer, our primary and secondary outcomes were CIN2+ and CIN3+, respectively. These outcomes were confirmed by histopathology and align with those commonly used in previous studies [26-29]. We selected features based on prior research [22,30-34] and input from clinical, biostatistical, and epidemiological experts. These included demographic characteristics, cervical cancer screening

XSL•FO
RenderX

history, HPV infection status, and medical examination results. Details on data preprocessing are provided in Appendix S1 in Multimedia Appendix 1.

Unsupervised Machine Learning

We calculated phenotypic distances between individuals using Gower's distance [35], a dissimilarity metric suitable for mixed continuous and categorical data. To visualize phenotypic variation in the population, we employed uniform manifold approximation and projection (UMAP) [36] to construct a phenomap. This approach enhances interpretability by presenting the distributions of individuals within a multidimensional phenotypic space, capturing the full range of baseline phenotypes. To identify and categorize the underlying CCP subgroups in which individuals exhibit phenotypic similarity, we employed and compared 3 methods to capture and validate CCP diversity: k-means [37], latent class analysis [38], and UMAP [39]. The final number of CCP subgroups was determined by reaching a consensus across all 3 approaches. Further details are provided in Appendix S1 in Multimedia Appendix 1.

Risk Estimates and an Algorithm to Identify Subgroups

To further explore the association between CCP subgroups and outcomes of interest, we calculated CCP-specific risk estimates (odds ratios, ORs) [40] and 95% CIs using inverse probability weights [41]. Additionally, we trained an eXtreme gradient boosting (XGBoost) algorithm [42] to predict 5 CCP subgroups with differential risks of outcomes, as described in our previous work [23]. Briefly, the model was trained and calibrated using an isotonic regression algorithm, and internally validated in the discovery cohort. The SHapley Additive exPlanations (SHAP) method was employed to identify each feature's relative contribution [23,43] and enhance the model's explainability. Model performance was evaluated using OVR AUROC (one-versus-rest area under the receiver operating characteristics curve, extended for multiple classes), Brier score, and calibration curves as primary metrics. In external validation, the model's diagnostic validity was assessed by comparing CCP-specific ORs between datasets to evaluate cross-cluster and cross-dataset risk differences. Further details are provided in Appendix S1 in Multimedia Appendix 1.

Developing Pathways to Improve Impact

We named our computational phenomapping strategy SCREENing (clinical Subgroups for CeRvical cancEr prEvention using computational pheNomaps and machine learnING). To assess real-world usability and effectiveness, we consulted 11 clinical experts and 3 epidemiologists on its clinical relevance, justification, result interpretability, and potential impact on screening strategies and public health policymaking.

Statistical Analysis

Given the importance of menopause in women, we specifically examined CCP-specific risks across age using spline analyses, with interaction tests to assess whether age modified these risks. Subgroup analyses were also performed by stratifying women based on key features. Details are provided in Appendix S1 in Multimedia Appendix 1. We followed the Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research [44] and the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis statement [45]. Data management was performed using the lulab.utils R package [46]. All analyses [47] were conducted using Python (version 3.11.6), SAS Enterprise Guide (version 7.1), and R (version 4.3.2; R Foundation for Statistical Computing).

Results

Characteristics of Cohorts

The study included 551,934 women (median age 49 years [42,48]; 10% infected with hrHPV, 1.2% with HPV-16, and 0.6% with HPV-18) in the discovery cohort (Table 1) and 47,130 women (median age 37 years [32,45]; 16.6% with hrHPV, 3.9% with HPV-16, and 1.4% with HPV-18) in the external cohort (Table S3 in Multimedia Appendix 1). In the discovery cohort, 9932 were pathologically diagnosed with cervical abnormalities, including 533 with carcinoma. In contrast, the external cohort showed a higher prevalence of cervical abnormalities at 11.1% (5251).

XSL•FO
RenderX

Table . Characteristics by cervical cancer prevention (CCP) subgroups in the discovery cohort. Categorical features are summarized as numbers (percentages), and continuous features as median (Q1, Q3), as appropriate.

Characteristics	CCP0[a] (n=542,002)	CCP1 (n=2242)	CCP2 (n=3770)	CCP3 (n=2278)	CCP4 (n=1642)	Total (n=551,934)
Age (years), median (Q1, Q3)	48.00 (42.00, 54.00)	48.00 (42.00, 54.00)	48.00 (43.00, 55.00)	51.00 (44.00, 57.00)	48.00 (43.00, 54.00)	49.00 (42.00, 54.00)
Race/ethnicity, n (%)						
Han	534,023 (98.53)	2214 (98.75)	3707 (98.33)	2236 (98.16)	1607 (97.87)	543,787 (98.52)
Others	4528 (0.84)	15 (0.67)	30 (0.80)	29 (1.27)	20 (1.22)	4622 (0.84)
Missing	3451 (0.64)	13 (0.58)	33 (0.88)	13 (0.57)	15 (0.91)	3525 (0.64)
History of cervical cancer screening, n (%)						
Missing	234 (0.04)	0 (0.00)	2 (0.05)	0 (0.00)	0 (0.00)	236 (0.04)
No	411,029 (75.84)	0 (0.00)	3768 (99.95)	2278 (100.00)	1642 (100.00)	418,717 (75.86)
Yes	130,739 (24.12)	2242 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)	132,981 (24.09)
Time of previous screening, n (%)						
Missing	785 (0.14)	13 (0.58)	2 (0.05)	0 (0.00)	0 (0.00)	800 (0.14)
No previous screening	411,029 (75.84)	0 (0.00)	3768 (99.95)	2278 (100.00)	1642 (100.00)	418,717 (75.86)
Within 3 years from now	65,519 (12.09)	1091 (48.66)	0 (0.00)	0 (0.00)	0 (0.00)	66,610 (12.07)
More than 3 years ago	64,669 (11.93)	1138 (50.76)	0 (0.00)	0 (0.00)	0 (0.00)	65,807 (11.92)
Gynecological examination, n (%)						
Missing	2180 (0.40)	1 (0.04)	1 (0.03)	0 (0.00)	0 (0.00)	2182 (0.40)
Normal	417,921 (77.11)	1508 (67.26)	3763 (99.81)	1568 (68.83)	0 (0.00)	424,760 (76.96)
Abnormal	121,901 (22.49)	733 (32.69)	6 (0.16)	710 (31.17)	1642 (100.00)	124,992 (22.65)
Positive for high-risk HPV [b,c], n (%)						
No	496,234 (91.56)	99 (4.42)	206 (5.46)	27 (1.19)	67 (4.08)	496,633 (89.98)
Yes	45,768 (8.44)	2143 (95.58)	3564 (94.54)	2251 (98.81)	1575 (95.92)	55,301 (10.02)
Positive for low-risk HPV [d], n (%)						
No	534,313 (98.58)	2035 (90.77)	3664 (97.19)	1813 (79.59)	1602 (97.56)	543,427 (98.46)
Yes	7689 (1.42)	207 (9.23)	106 (2.81)	465 (20.41)	40 (2.44)	8507 (1.54)
Positive for possible high-risk HPV [e], n (%)						
No	537,999 (99.26)	2130 (95.00)	3686 (97.77)	1987 (87.23)	1617 (98.48)	547,419 (99.18)
Yes	4003 (0.74)	112 (5.00)	84 (2.23)	291 (12.77)	25 (1.52)	4515 (0.82)
Number of HPV infections	0.00 (0.00, 0.00)	1.00 (1.00, 2.00)	1.00 (1.00, 1.00)	2.00 (2.00, 3.00)	1.00 (1.00, 1.00)	0.00 (0.00, 0.00)
Positive for HPV-16, n (%)						
No	538,375 (99.33)	1684 (75.11)	2745 (72.81)	1494 (65.58)	1199 (73.02)	545,497 (98.83)
Yes	3627 (0.67)	558 (24.89)	1025 (27.19)	784 (34.42)	443 (26.98)	6437 (1.17)
Positive for HPV-18, n (%)						
No	539,716 (99.58)	2002 (89.30)	3442 (91.30)	1887 (82.84)	1508 (91.84)	548,555 (99.39)
Yes	2286 (0.42)	240 (10.70)	328 (8.70)	391 (17.16)	134 (8.16)	3379 (0.61)
Positive for HPV-31, n (%)						
No	540,143 (99.66)	2138 (95.36)	3662 (97.14)	2063 (90.56)	1588 (96.71)	549,594 (99.58)

XSL•FO
RenderX

Characteristics	CCP0[a] (n=542,002)	CCP1 (n=2242)	CCP2 (n=3770)	CCP3 (n=2278)	CCP4 (n=1642)	Total (n=551,934)
Yes	1859 (0.34)	104 (4.64)	108 (2.86)	215 (9.44)	54 (3.29)	2340 (0.42)
Positive for HPV-33, n (%)						
No	539,290 (99.50)	2096 (93.49)	3608 (95.70)	1978 (86.83)	1562 (95.13)	548,534 (99.38)
Yes	2712 (0.50)	146 (6.51)	162 (4.30)	300 (13.17)	80 (4.87)	3400 (0.62)
Positive for HPV-35, n (%)						
No	540,820 (99.78)	2197 (97.99)	3708 (98.36)	2160 (94.82)	1622 (98.78)	550,507 (99.74)
Yes	1182 (0.22)	45 (2.01)	62 (1.64)	118 (5.18)	20 (1.22)	1427 (0.26)
Positive for HPV-39, n (%)						
No	537,675 (99.20)	2092 (93.31)	3638 (96.50)	2021 (88.72)	1600 (97.44)	547,026 (99.11)
Yes	4327 (0.80)	150 (6.69)	132 (3.50)	257 (11.28)	42 (2.56)	4908 (0.89)
Positive for HPV-45, n (%)						
No	540,998 (99.81)	2205 (98.35)	3741 (99.23)	2204 (96.75)	1632 (99.39)	550,780 (99.79)
Yes	1004 (0.19)	37 (1.65)	29 (0.77)	74 (3.25)	10 (0.61)	1154 (0.21)
Positive for HPV-51, n (%)						
No	537,654 (99.20)	2053 (91.57)	3590 (95.23)	1954 (85.78)	1565 (95.31)	546,816 (99.07)
Yes	4348 (0.80)	189 (8.43)	180 (4.77)	324 (14.22)	77 (4.69)	5118 (0.93)
Positive for HPV-52, n (%)						
No	525,904 (97.03)	1622 (72.35)	3019 (80.08)	1396 (61.28)	1276 (77.71)	533,217 (96.61)
Yes	16,098 (2.97)	620 (27.65)	751 (19.92)	882 (38.72)	366 (22.29)	18,717 (3.39)
Positive for HPV-56, n (%)						
No	539,439 (99.53)	2133 (95.14)	3682 (97.67)	2076 (91.13)	1610 (98.05)	548,940 (99.46)
Yes	2563 (0.47)	109 (4.86)	88 (2.33)	202 (8.87)	32 (1.95)	2994 (0.54)
Positive for HPV-58, n (%)						
No	534,764 (98.66)	1877 (83.72)	3297 (87.45)	1691 (74.23)	1413 (86.05)	543,042 (98.39)
Yes	7238 (1.34)	365 (16.28)	473 (12.55)	587 (25.77)	229 (13.95)	8892 (1.61)
Positive for HPV-59, n (%)						
No	540,034 (99.64)	2180 (97.23)	3710 (98.41)	2124 (93.24)	1620 (98.66)	549,668 (99.59)
Yes	1968 (0.36)	62 (2.77)	60 (1.59)	154 (6.76)	22 (1.34)	2266 (0.41)
Positive for HPV-66, n (%)						
No	540,640 (99.75)	2186 (97.50)	3720 (98.67)	2180 (95.70)	1614 (98.29)	550,340 (99.71)
Yes	1362 (0.25)	56 (2.50)	50 (1.33)	98 (4.30)	28 (1.71)	1594 (0.29)
Positive for HPV-68, n (%)						
No	538,250 (99.31)	2124 (94.74)	3654 (96.92)	2018 (88.59)	1604 (97.69)	547,650 (99.22)
Yes	3752 (0.69)	118 (5.26)	116 (3.08)	260 (11.41)	38 (2.31)	4284 (0.78)
Positive for HPV-11, n (%)						
No	541,435 (99.90)	2218 (98.93)	3764 (99.84)	2235 (98.11)	1631 (99.33)	551,283 (99.88)
Yes	567 (0.10)	24 (1.07)	6 (0.16)	43 (1.89)	11 (0.67)	651 (0.12)
Positive for HPV-42, n (%)						
No	540,587 (99.74)	2199 (98.08)	3745 (99.34)	2175 (95.48)	1639 (99.82)	550,345 (99.71)
Yes	1415 (0.26)	43 (1.92)	25 (0.66)	103 (4.52)	3 (0.18)	1589 (0.29)
Positive for HPV-43, n (%)						
No	541,315 (99.87)	2212 (98.66)	3759 (99.71)	2210 (97.01)	1641 (99.94)	551,137 (99.86)

XSL•FO
RenderX

Characteristics	CCP0[a] (n=542,002)	CCP1 (n=2242)	CCP2 (n=3770)	CCP3 (n=2278)	CCP4 (n=1642)	Total (n=551,934)
Yes	687 (0.13)	30 (1.34)	11 (0.29)	68 (2.99)	1 (0.06)	797 (0.14)
Positive for HPV-44, n (%)						
No	540,505 (99.72)	2212 (98.66)	3753 (99.55)	2201 (96.62)	1636 (99.63)	550,307 (99.71)
Yes	1497 (0.28)	30 (1.34)	17 (0.45)	77 (3.38)	6 (0.37)	1627 (0.29)
Positive for HPV-6, n (%)						
No	540,748 (99.77)	2207 (98.44)	3748 (99.42)	2199 (96.53)	1635 (99.57)	550,537 (99.75)
Yes	1254 (0.23)	35 (1.56)	22 (0.58)	79 (3.47)	7 (0.43)	1397 (0.25)
Positive for HPV-81, n (%)						
No	539,354 (99.51)	2178 (97.15)	3747 (99.39)	2136 (93.77)	1632 (99.39)	549,047 (99.48)
Yes	2648 (0.49)	64 (2.85)	23 (0.61)	142 (6.23)	10 (0.61)	2887 (0.52)
Positive for HPV-53, n (%)						
No	538,445 (99.34)	2139 (95.41)	3689 (97.85)	2015 (88.45)	1617 (98.48)	547,905 (99.27)
Yes	3557 (0.66)	103 (4.59)	81 (2.15)	263 (11.55)	25 (1.52)	4029 (0.73)
Cervical cytology examination, n (%)						
NILM[f]	42,746 (7.89)	489 (21.81)	721 (19.12)	438 (19.23)	345 (21.01)	44,739 (8.11)
No examination due to negative for high-risk HPV	487,235 (89.90)	3 (0.13)	7 (0.19)	0 (0.00)	0 (0.00)	487,245 (88.28)
ASC-US[g]	7577 (1.40)	808 (36.04)	1185 (31.43)	721 (31.65)	531 (32.34)	10822 (1.96)
LSIL[h]	2245 (0.41)	488 (21.77)	772 (20.48)	546 (23.97)	286 (17.42)	4337 (0.79)
AGC[i]	99 (0.02)	18 (0.80)	29 (0.77)	13 (0.57)	30 (1.83)	189 (0.03)
Missing but positive for high-risk HPV	779 (0.14)	99 (4.42)	231 (6.13)	159 (6.98)	81 (4.93)	1349 (0.24)
ASC-H[j]	902 (0.17)	147 (6.56)	363 (9.63)	163 (7.16)	145 (8.83)	1720 (0.31)
AGC-FN[k]	7 (0.00)	4 (0.18)	9 (0.24)	3 (0.13)	6 (0.37)	29 (0.01)
HSIL[l]	411 (0.08)	185 (8.25)	443 (11.75)	233 (10.23)	215 (13.09)	1487 (0.27)
Carcinoma	1 (0.00)	1 (0.04)	10 (0.27)	2 (0.09)	3 (0.18)	17 (0.00)
Histopathological diagnosis, n (%)						
NILM	542,002 (100.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	542,002 (98.20)
CIN1[m]	0 (0.00)	1542 (68.78)	2456 (65.15)	1445 (63.43)	948 (57.73)	6391 (1.16)
CIN2[n]	0 (0.00)	22 (0.98)	29 (0.77)	15 (0.66)	16 (0.97)	82 (0.01)
CIN2/3[o]	0 (0.00)	551 (24.58)	1113 (29.52)	716 (31.43)	485 (29.54)	2865 (0.52)
CIN3[p]	0 (0.00)	16 (0.71)	19 (0.50)	13 (0.57)	13 (0.79)	61 (0.01)

XSL•FO
RenderX

Characteristics	CCP0[a] (n=542,002)	CCP1 (n=2242)	CCP2 (n=3770)	CCP3 (n=2278)	CCP4 (n=1642)	Total (n=551,934)
Carcinoma	0 (0.00)	111 (4.95)	153 (4.06)	89 (3.91)	180 (10.96)	533 (0.10)

[a]CCP: cervical cancer prevention.

[b]HPV: human papillomavirus.

[c]high-risk HPV: HPV-16/18/31/33/35/39/45/51/52/56/58/59/66/68.

[d]low-risk HPV: HPV-11/40/42/43/44/6/61/72/81.

[e]possible high-risk HPV: HPV-53/70/73/82/83.

[f]NILM: negative for intraepithelial lesion or malignancy.

[g]ASC-US: atypical squamous cells of undetermined significance.

[h]LSIL: low-grade squamous intraepithelial lesion.

[i]AGC: atypical glandular cells.

[j]ASC-H: atypical squamous cells, cannot exclude high-grade squamous intraepithelial lesion.

[k]AGC-FN: AGC-favor neoplastic.

[l]HSIL: high-grade squamous intraepithelial lesion.

[m]CIN1: cervical intraepithelial neoplasia grade 1.

[n]CIN2: cervical intraepithelial neoplasia grade 2.

[o]CIN2/3 is reflective of CIN2 or CIN3, i.e., HSIL; carcinoma consists of AIS (adenocarcinoma in situ) and cancer.

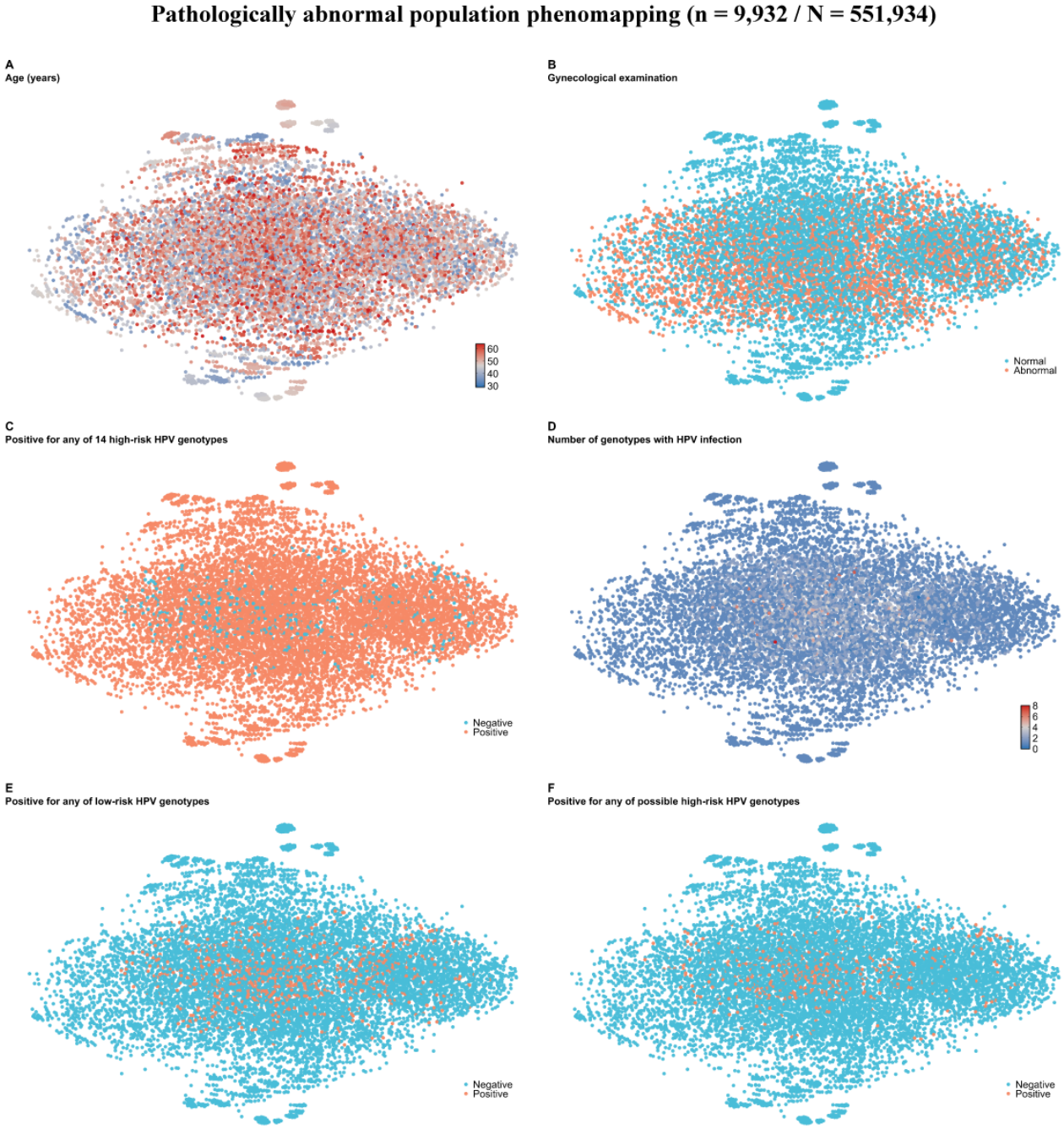
[p]CIN3: cervical intraepithelial neoplasia grade 3.

Data-Driven Subgroups

The discovery cohort was extensively phenotyped based on 31 features (Figure 2). Visual assessment of the risk phenomaps revealed that nearly all 6 features were heterogeneously distributed in the phenomic space to varying degrees, indicating distinct phenotypic neighborhoods. In the discovery cohort, the CCP subgroups identified across 3 algorithms and all metrics were stable (Figure 3A and Table S4 in Multimedia Appendix 1), with the optimal number of clusters being 4 in the subpopulation of 9932 women with cervical abnormalities. The distribution of cervical abnormalities across subgroups is presented in Figure 3B. Additionally, women with normal screening results were treated as another CCP subgroup (Table 1). Following a detailed analysis of each subgroup's features, we labeled the 5 identified subgroups as follows: (0) healthy, (1) early onset, (2) screening-targeted, (3) late onset, and (4) carcinoma-specific. CCP1 had the highest prevalence of CIN1, while CCP4 had the highest prevalence of carcinoma (Figure 3C). Most features were able to discriminate well between the subgroups. In the external cohort, the subgroups were consistent (Table S3 in Multimedia Appendix 1), though CCP1 was not identified due to missing information on previous screenings. The distribution of subgroups was similar across the discovery and external cohorts, with CCP0 being the most common subgroup and CCP4 having the highest prevalence of carcinoma.

XSL•FO
RenderX

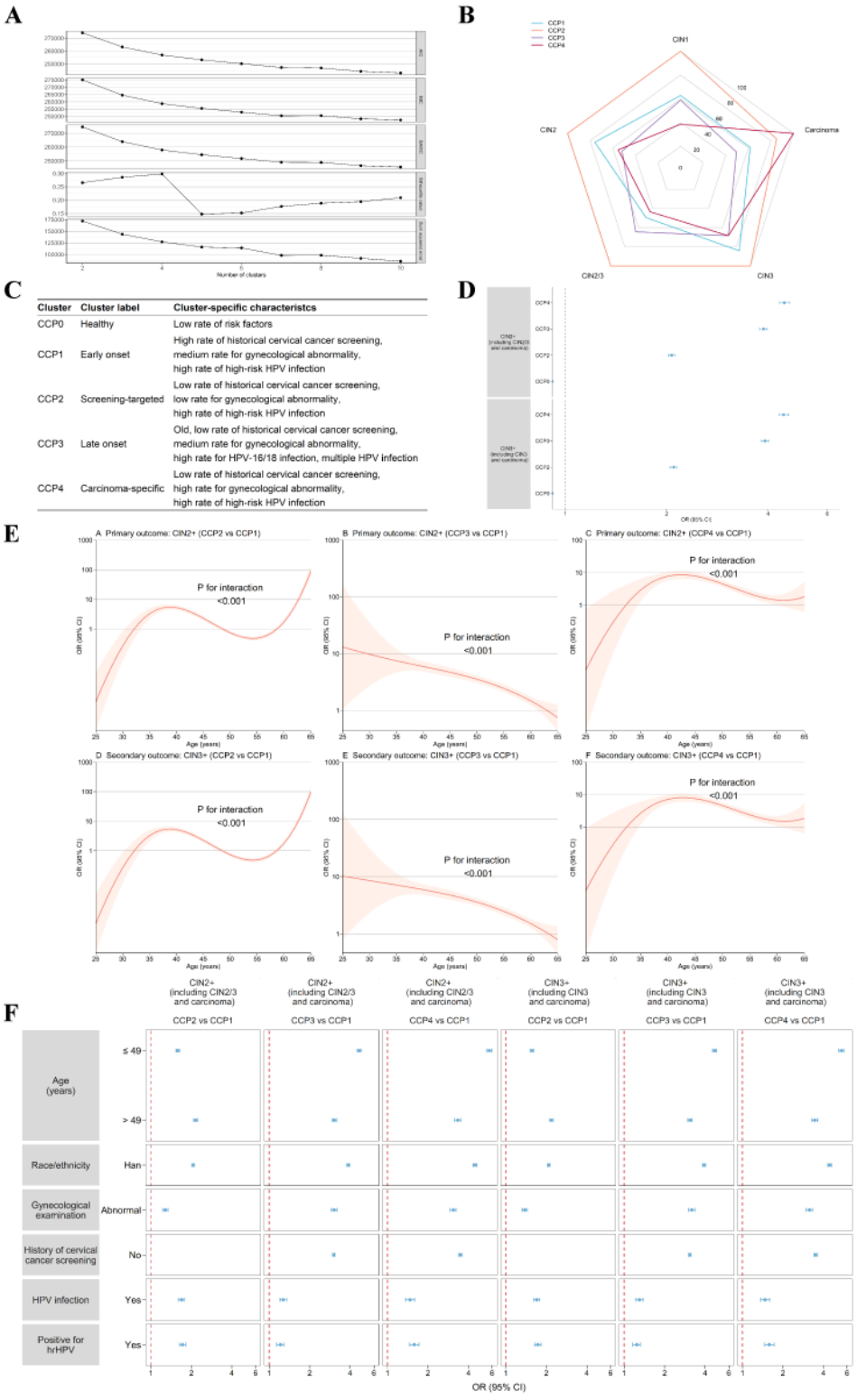
Figure 2. Manifold representations of the phenotypic architecture of 9932 individuals pathologically diagnosed with cervical abnormality from the discovery cohort (N=551,934). A total of 9932 individuals are embedded in the phenotypic space based on dissimilarity metrics (Gower's distance) derived from 31 included phenotypic features; thus, phenotypically similar women tend to be topologically closer. In the subfigures, each dot represents an individual, with coloring based on the value of features. Since the dimensionality reduction is nonlinear, axes have been omitted, and only the comparisons between distances are meaningful. Pathologically abnormal diagnoses consisted of CIN1, CIN2, CIN3, and carcinoma. 14 high-risk HPV: HPV-16/18/31/33/35/39/45/51/52/56/58/59/66/68; low-risk HPV: HPV-11/40/42/43/44/6/61/72/81; possible high-risk HPV: HPV- 53/70/73/82/83.



XSL•FO
RenderX

Figure 3. Determination of the 5 CCP subgroups with differential risks of CIN2+/CIN3+. (**A**) Elbow plot to determine the optimal number of clusters. Across 3 algorithms (k-means, LCA, and UMAP), identified CCP subgroups were stable, with the optimal number of clusters being 4 in 9932 women pathologically diagnosed with cervical abnormality. Women with normal screening test findings were considered as one independent CCP subgroup free of cervical cancer (Table 1). (**B**) Relative prevalence of pathologically abnormal diagnoses across CCP subgroups. For each pathologically abnormal diagnosis, the highest prevalence was designated as 100%, and the prevalence in each of the other CCP subgroups was relative to that prevalence (0 - 100). (**C**) Characteristics of CCP subgroups. (**D**) CCP-specific risk estimates of CIN2+/CIN3+ compared with the CCP1 subgroup. Adjusted ORs (dots) and 95% CIs (error bars) are presented. The dashed line marks an OR of 1.00; lower limits of 95% CIs with values greater than 1.00 indicate significantly increased risk. (**E**) The CCP-specific risks of CIN2+/CIN3+ across age compared with the CCP1 subgroup. Age was transformed into a restricted cubic spline function for the analyses. *P* value was based on 2-sided Chi-squared test on the interaction between age and CCP subgroups. A *P* value of <.05 suggests that age modifies the association between CCP subgroups and CIN2+/CIN3+. Light-colored bands represent 95% CIs. (**F**) Subgroup analyses of CCP-specific risk estimates of CIN2+/CIN3+ compared with the CCP1 subgroup. Adjusted ORs (dots) and 95% CIs (error bars) are presented. The dashed line marks an OR of 1.00; lower limits of 95% CIs with values greater than 1.00 indicate significantly increased risk. Empty subfigures suggest insufficient samples for the analysis. AIC: Akaike information criterion; BIC: Bayesian information criterion; SABIC: sample size-adjusted BIC; CCP: cervical cancer prevention; CIN1: cervical intraepithelial neoplasia grade 1; CIN2: cervical intraepithelial neoplasia grade 2; CIN3: cervical intraepithelial neoplasia grade 3; CIN2+: cervical intraepithelial neoplasia grade 2 or worse; CIN3+: cervical intraepithelial neoplasia grade 3 or worse; OR: odds ratio; hrHPV: high-risk human papillomavirus.

XSL•FO
RenderX



Risk Estimates and Diagnostic Validity

In the diagnostic validity analysis of the discovery cohort (Figure 3D), women in CCP2-4 exhibited a significantly increased risk of both CIN2+ (CCP2: OR 2.07, 95% CI [2.03 - 2.12]; CCP3: 3.88 [3.78 - 3.97]; CCP4: 4.47 [4.33 - 4.63]) and CIN3+ (2.10 [2.05 - 2.14]; 3.92 [3.82 - 4.02]; 4.45 [4.31 - 4.61]) compared to CCP1. Risk analysis across age groups showed that the risks for CIN2+/CIN3+ were evident in the age ranges of 34 - 47 and 60 - 65 in CCP2, 26 - 61 in CCP3, and 35 - 60 in CCP4, respectively (Figure 3E). Interaction analyses between age and subgroups revealed that the risks of CIN2+/CIN3+ increased with age in the 34 - 39 and 60 - 65 age ranges in CCP2, in the 35-43/44 age range in CCP4. Conversely, the risks decreased with age in CCP2 within the 40 - 47 age range and in CCP4 within the 45 - 60 age range (all *P* for interaction <.001). Subgroup analyses indicated that the risks for CIN2+/CIN3+ were present across nearly all subgroups, categorized by age, race, gynecological examination, screening history, hrHPV

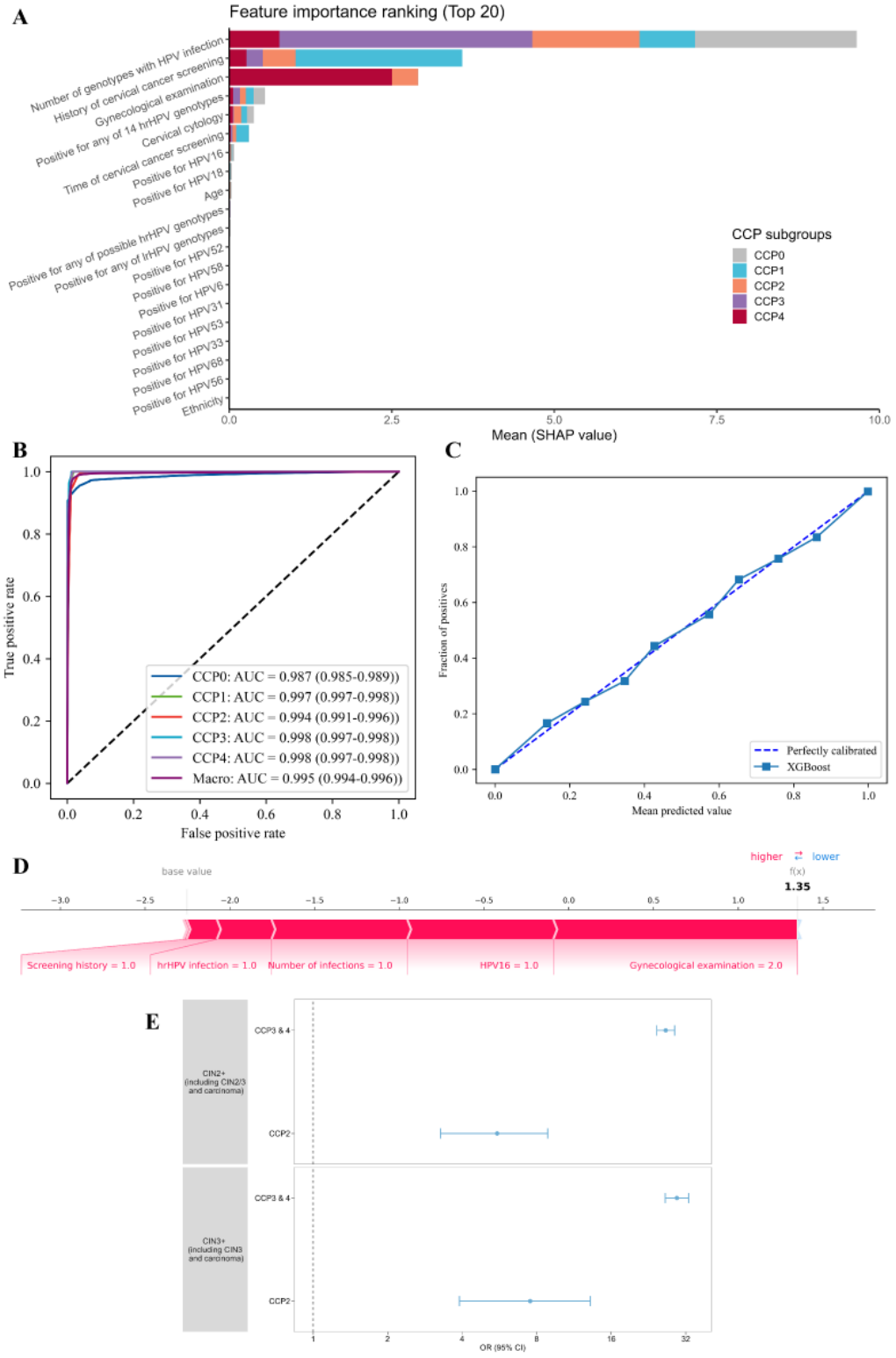
XSL•FO
RenderX

infection, and number of infections (Figure 3F and Table S5 in Multimedia Appendix 1).

SHAP analysis identified the top 10 key features for prediction, including the number of infections, screening history, gynecological examination, hrHPV infection, cervical cytology, time since previous screening, HPV-16/18 infection, age, and possible hrHPV infection (Figure 4A). The final XGBoost model demonstrated excellent discrimination (OVR AUROC 0.995 [0.994 - 0.996]), with CCP0 showing the lowest AUROC of 0.987 [0.985 - 0.989], and strong calibration (Brier score 0.021 [0.020 - 0.022]) (Figure 4B–C and Table S6 in Multimedia Appendix 1). For the external diagnostic validity of the identified subgroups, women in CCP2-4 showed differential and increased risk for both CIN2+ (CCP2: 5.54 [3.27 - 8.86]; CCP3&4: 26.56 [24.44 - 28.88]) and CIN3+ (7.53 [3.90 - 13.18]; 29.47 [26.46 - 32.86]) compared to CCP0. The cross-cluster risk differences were consistent across the discovery and external cohorts. To illustrate the model's explainability, Figure 4D presents the SHAP plot for a woman in the CCP4 subgroup, who had no previous screening and tested positive for HPV-16 and CIN1 based on cervical cytology. Given this information, our model predicted an elevated risk of CCP4. Additionally, a screenshot of our browser-accessible tool based on the model developed in this study, to clarify how the model can be easily accessed via a browser in clinical settings, is shown in Figure S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 4. Feature importance, discrimination, calibration, explainability, and external diagnostic validity. (**A**) The top 20 features for prediction of CCP subgroups with differential risks of CIN2+/CIN3+ are shown. The y-axis represents the features included in the model development (in descending order of importance), and the x-axis indicates the mean of SHAP values. The 10 most important features were selected to train the final parsimonious model. (**B**) ROC curves in the internal validation from the discovery cohort. The dark purple line represents the macro-average of all 5 CCP subgroups. (**C**) Calibration curve of the alignment between predicted and observed CCP subgroups for the final model. The solid line corresponds to the calibration curve, with the dashed line corresponding to a reference for perfect calibration (ie, perfect alignment between the predicted and observed probabilities). (**D**) Explanation of the SCREENing tool (ie, the inference process of the final model with a woman in the CCP4 subgroup). (**E**) In external validation, to assess cross-cluster and cross-dataset risk differences, the model was measured and compared by diagnostic validity (CCP-specific ORs) between datasets. CCP: cervical cancer prevention; hrHPV: high-risk human papillomavirus; lrHPV: low-risk human papillomavirus; SHAP: SHapley Additive exPlanations; AUC: area under receiver operating characteristics curve; XGBoost: eXtreme gradient boosting.



XSL•FO
RenderX

Developed Pathways

Based on our analyses, we proposed a triple SCREENing strategy to prioritize CCP subgroups with varying CIN2+/CIN3+ risks: (1) top priority for colposcopy referrals to women in CCP3-4 in resource-constrained settings, (2) higher priority given for scaling up organized, population-based HPV screening programs (with adequate follow-up) for women in CCP1-2, and (3) lowest priority for CCP0 women, with large-scale screening limited to resource-rich settings. To assess its impact, we evaluated the real-world usability and effectiveness of SCREENing. Sample clinicians reported that the identified subgroups and included features were clinically relevant for risk-based cervical cancer management, particularly for colposcopy referrals in CCP3-4. They also found the strategy transparent, interpretable, and generalizable across clinical settings, with the browser-accessible model feasible and effective for indicating risk probabilities during consultations. Sample epidemiologists highlighted the strategy's potential to advance screening practices, public health policies, and trial design by enabling phenomapping, estimating subgroup-specific risk profiles, and establishing high-performance, cost-effective CCP for resource-limited regions. They noted its capability for testing effectiveness and cost-effectiveness in well-designed, prospective studies.

Discussion

This study represents one of the largest EHR analyses to date, employing diverse machine learning methods and robust validation approaches to classify subgroups and predict CIN2+/CIN3+ risk. By leveraging comprehensive and interconnected phenotypic features, our study identified 5 CCP subgroups with varying risks. CCP2–4 had significantly higher risks of CIN2+ (CCP2: OR 2.07 [95% CI: 2.03 - 2.12], CCP3: 3.88 [3.78 - 3.97], CCP4: 4.47 [4.33 - 4.63]) and CIN3+ (CCP2: 2.10 [2.05 - 2.14], CCP3: 3.92 [3.82 - 4.02], CCP4: 4.45 [4.31 - 4.61]) compared to CCP1 (P<.001), consistent with the direction of results observed in the external cohort. Our findings offer a robust foundation for the proposed triple SCREENing strategy. This tailored approach prioritizes high-risk subgroups, providing actionable insights for cervical cancer prevention, particularly in LMICs.

To this end, various approaches, including single-cell transcriptomic analysis [49], sequence and phylogenetic analysis [50,51], and cluster analysis [48,52-54], have been employed to define genetic and cellular heterogeneity [55] in cervical cancer. These studies suggest that distinct subpopulations exhibit heterogeneous risk effects, linking individual features to varying absolute cervical cancer risk. Our 5 CCP subgroups align with findings from 2 major population-based clustering studies [56,57]. One study used a Poisson regression–based CEM clustering algorithm [56] to identify clusters of Indian states with similar cervical cancer incidence patterns. However, due to its focus on approximating missing data on sexual behavior, HPV prevalence, or cervical cancer incidence [56], the clusters were not ideal for screening purposes. That study did not provide sufficient details to estimate cluster-specific risks, and its reliance on features not typically available in routine screening limited real-world applicability. Additionally, strict model assumptions and the inclusion of relatively few features further constrained its generalizability. Another study performed hierarchical clustering of HPV-related methylation sites to identify subgroups of patients with cervical cancer [57]. While informative for prognosis and clinical management, the 3 clusters were not designed for cervical cancer screening. Instead, they are more suitable for guiding prognosis assessment, refining risk stratification, and optimizing treatment strategies in clinical practice.

In contrast, our study, specifically designed for cervical cancer screening, identified complex interactions among factors such as the number of infections, screening history, gynecological examination, hrHPV and HPV-16/18 infections, cervical cytology, time since previous screening, age, and possible hrHPV infection as key determinants of the 5 CCP subgroups with varying CIN2+/CIN3+ risks. Compared to previous studies [56,57], our design and identified subgroups are more representative of real-world cervical cancer screening. Leveraging the comprehensive and interconnected phenotypic features from EHRs, our algorithm extracted CCP-specific risk estimates, providing a robust foundation for the proposed triple SCREENing strategy. This strategy prioritizes subgroups with differential CIN2+/CIN3+ risks, offering a tailored approach to CCP. We recommend implementing SCREENing as a supplemental tool to existing guidelines [58-60], while accounting for the unique priorities and constraints of LMICs. SCREENing enables clinicians to perform effective, risk-based screening, followed by adequate diagnosis, surveillance, and management, while empowering policy makers to optimize public health policies and resource allocation. This approach has the potential to mitigate resource shortages in LMICs, reduce delays in diagnosis and treatment, and enhance screening efficiency by focusing efforts on high-risk populations, ultimately maximizing population-level benefits. Additionally, our analysis highlights the modifying effect of age on CCP-specific CIN2+/CIN3+ risks, emphasizing the need for greater attention to menopause and age [61] in cervical cancer screening strategies.

Methodologically, our study advances the external validation of machine learning algorithms for identifying subgroups and predicting cervical precancer or cancer risk using large-scale routine EHRs—a rarity in previous studies, which have often been limited to small samples. Our robust, structured framework of internal, external, and diagnostic validation enhances the acceptability and generalizability of unsupervised and supervised machine learning approaches in routine CCP and is adaptable to other diseases. The identified subgroups demonstrated good accuracy and diagnostic validity for CIN2+/CIN3+ both within and across datasets, though performance was lower in scenarios with missing key features. Differences in diagnostic validity across datasets may reflect variations in HPV infection patterns and previous screening histories, which influence the distribution of risk factors. Our findings of 5 subgroups and the proposed strategy are novel, offering a framework for assessing validity in screening, follow-up surveillance, and treatment for cervical cancer. This study signals the potential for more effective and targeted approaches to CCP.

XSL•FO
RenderX

This study represents one of the largest EHR analyses to date, employing multiple machine learning methods, datasets, and validation approaches to classify subgroups and predict CIN2+/CIN3+ risk. However, there are several limitations. First, a major limitation of this study is that our models were not externally validated beyond China. HPV infection patterns and the epidemiology of cervical abnormalities differ across countries and regions, which may impact the generalizability of our models. Given this variability, and although we used data from 5 distinct regions within China, we strongly recommend further validation of our models' performance before applying them in settings not included in this study. We acknowledge that externally validating the model in additional multicenter studies worldwide is crucial for assessing its transferability and applicability across different clinical settings. Additionally, the tool may face challenges in fully adapting to various real-world scenarios when implemented outside the controlled validation environment of this study. To address these concerns, we plan to conduct independent external and prospective validations, as well as pilot implementation across diverse clinical settings in future studies. These efforts will evaluate the model's effectiveness using a wider range of data, particularly in the context of real-world cervical cancer prevention across regions beyond those included in this study. We also encourage independent researchers to validate our model in their own settings, where feasible. While this study marks a significant step in the development of the SCREENing tool, we acknowledge that further research and validation are necessary to establish its effectiveness in real-world applications.

Second, we acknowledge that the use of retrospective EHR data may introduce biases and may not fully capture the real-time challenges of clinical practice. These challenges include issues such as missing data, data errors, and the heterogeneity of EHR systems across different settings. Although we made efforts to control for the quality of the EHR data, it remains inherently difficult to fully address these concerns without significant infrastructure changes, as well as ongoing monitoring and data validation. Therefore, we explicitly caution that the findings of this study should be interpreted with caution. In light of this limitation, future studies should aim to bridge the gap between retrospective analyses and the practical challenges of data collection in clinical settings. Third, while this study focused on routine EHR data, we acknowledge that incorporating multi-omics data—such as genomic, proteomic, and imaging data—could further enhance the model's robustness and performance, providing valuable new insights into cervical abnormalities. We propose this as an area for future research to better understand the progression of cervical cancer and to offer novel perspectives on controlling and eliminating cervical cancer. Fourth, due to the limitations of routinely collected EHR content, we were unable to include several key sexual behavior features, such as the number of sexual partners and oral contraceptive use [62], in our analyses. These factors are known to significantly influence HPV infection and cervical cancer risk. While the model in this study demonstrated high performance, the absence of such data constrained our ability to further analyze the impact of sexual behaviors on findings, such as subgroup characterization, which could provide valuable insights. We recommend that future studies incorporate these data to validate our findings and offer a more comprehensive understanding of this field.

In conclusion, this study underscores the potential of leveraging machine learning algorithms and large-scale routine EHRs to enhance cervical cancer prevention strategies. By identifying key determinants of CIN2+/CIN3+ risk and classifying 5 distinct subgroups, our study provides a robust, data-driven foundation for the proposed triple SCREENing strategy. This approach prioritizes tailored prevention efforts for subgroups with varying risks, offering a novel and scalable tool to complement existing cervical cancer screening guidelines. Future work should focus on independent external and prospective validation to maximize the global impact of this strategy.

Acknowledgments

We thank all the contributors for their efforts in this study. We also thank all funding sources for their funding. This research was funded by Fujian Province's Third Batch of Flexible Introduction of High-Level Medical Talent Teams (TD202307), Merck Investigator Studies Program [100073, 59484], National Natural Science Foundation of China (82271658), Major Scientific Research Program for Young and Middle-aged Health Professionals of Fujian Province, China (2021ZQNZD011), Fujian Province Central Government-Guided Local Science and Technology Development Project (2023L3019), Fujian Provincial Health Technology Project (2024GGA061), and Fujian Provincial Science and Technology Innovation Joint Fund (2021Y9171).

Data Availability

The data that support the findings of this study are available from the corresponding authors upon reasonable request.

Authors' Contributions

ZL conceived and designed the study in consultation with HZ. ZL analyzed the data and interpreted the results. ZL drafted the manuscript with all authors (BD, HC, TT, JW, LF, BW, WZ, TY, XT, JW, SL, XH, ZZ, HX, SX, SL, PS, and HZ) critically reviewing the manuscript. HZ and SL provided administrative and material support, as well as supervision and mentorship.
Each author contributed important intellectual content during drafting or revision of the manuscript and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved.

XSL•FO
RenderX

All authors had access to all the data in this study and approved the final version of the manuscript. The corresponding authors attests that all the listed authors meet the authorship criteria and that no others meeting the criteria have been omitted. We confirm that no generative AI tools, such as ChatGPT or other language models, were used in the writing or preparation of this manuscript.
S Liu, PS, and HZ are co-corresponding authors.
The complete address for the three co-corresponding authors are as follows:
S Liu: West Building 1 Medical Garden 4th floor, Gongchang Road 66, Sun Yat-sen University (Shenzhen Campus) Shenzhen, Guangming District, 518107, China; liusy99@mail.sysu.edu.cn (SL); +86 135 4329 6396.
PS: 18 Daoshan Road, Fuzhou 350001, Fujian, P.R. China (Phone: +86-591-87558732; Fax: +86-591-87551247; email: fmsun1975@fjmu.edu.cn).
HZ: Room 435, Bld #8, 130 Dongan Road, Xuhui District, Shanghai (200032), China; zouhuachun@fudan.edu.cn (HZ); +86 21 5423 7707.
ZL, BD, HC, TT, and Junfeng Wang are co-first authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data Supplement.
[DOCX File, 380 KB - publichealth_v11i1e67840_app1.docx]

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Chen YC, Chen YY, Su SY, et al. Projected time for the elimination of cervical cancer under various intervention scenarios: age-period-cohort macrosimulation study. JMIR Public Health Surveill 2024 Apr 18;10:e46360. [doi: 10.2196/46360] [Medline: 38635315]
3. Lama Y, Hu D, Jamison A, Quinn SC, Broniatowski DA. Characterizing trends in human papillomavirus vaccine discourse on Reddit (2007-2015): an observational study. JMIR Public Health Surveill 2019 Mar 27;5(1):e12480. [doi: 10.2196/12480] [Medline: 30916662]
4. Walboomers JMM, Jacobs MV, Manos MM, et al. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J Pathol 1999 Sep;189(1):12-19. [doi: 10.1002/(SICI)1096-9896(199909)189:1<12::AID-PATH431>3.0.CO;2-F]
5. Lu Z, Zhou X, Fu L, et al. Prognostic prediction models for oropharyngeal squamous cell carcinoma (OPSCC): a protocol for systematic review, critical appraisal and meta-analysis. BMJ Open 2023 Oct 12;13(10):e073375. [doi: 10.1136/bmjopen-2023-073375] [Medline: 37827742]
6. Yang SP, Lin XY, Hu M, Cai CF. The prognostic and predictive effects of human papillomavirus status in hypopharyngeal carcinoma: population-based study. JMIR Public Health Surveill 2022 Dec 16;8(12):e40185. [doi: 10.2196/40185] [Medline: 36525304]
7. Yang L, Lu Z, Bian J, Li F, Zou H. Association between chronic viral infection-related hospitalization and risk of cardiovascular disease: a population-based cohort study. J Med Virol 2024 Jan;96(1):e29350. [doi: 10.1002/jmv.29350] [Medline: 38180233]
8. WHO. Global strategy to accelerate the elimination of cervical cancer as a public health problem. 2020. URL: https://www.who.int/publications/i/item/9789240014107
9. Bruni L, Serrano B, Roura E, et al. Cervical cancer screening programmes and age-specific coverage estimates for 202 countries and territories worldwide: a review and synthetic analysis. Lancet Glob Health 2022 Aug;10(8):e1115-e1127. [doi: 10.1016/S2214-109X(22)00241-8] [Medline: 35839811]
10. Gravitt PE, Silver MI, Hussey HM, et al. Achieving equity in cervical cancer screening in low- and middle-income countries (LMICs): strengthening health systems using a systems thinking approach. Prev Med 2021 Mar;144:106322. [doi: 10.1016/j.ypmed.2020.106322]
11. Arbyn M, Simon M, Peeters E, et al. 2020 list of human papillomavirus assays suitable for primary cervical cancer screening. Clin Microbiol Infect 2021 Aug;27(8):1083-1095. [doi: 10.1016/j.cmi.2021.04.031]
12. Poljak M, Oštrbenk Valenčak A, Gimpelj Domjanič G, Xu L, Arbyn M. Commercially available molecular tests for human papillomaviruses: a global overview. Clin Microbiol Infect 2020 Sep;26(9):1144-1150. [doi: 10.1016/j.cmi.2020.03.033] [Medline: 32247892]
13. Arbyn M, Verdoodt F, Snijders PJF, et al. Accuracy of human papillomavirus testing on self-collected versus clinician-collected samples: a meta-analysis. Lancet Oncol 2014 Feb;15(2):172-183. [doi: 10.1016/S1470-2045(13)70570-9] [Medline: 24433684]

XSL•FO
RenderX

14. Mezei AK, Armstrong HL, Pedersen HN, et al. Cost-effectiveness of cervical cancer screening methods in low- and middle-income countries: a systematic review. Int J Cancer 2017 Aug 1;141(3):437-446. [doi: 10.1002/ijc.30695] [Medline: 28297074]
15. Camara H, Nosi S, Munnull G, et al. Women's acceptability of a self-collect HPV same-day screen-and-treat program in a high burden setting in the Pacific. BMC Health Serv Res 2022 Dec 12;22(1):1514. [doi: 10.1186/s12913-022-08842-1] [Medline: 36510192]
16. Ronco G, Dillner J, Elfström KM, et al. Efficacy of HPV-based screening for prevention of invasive cervical cancer: follow-up of four European randomised controlled trials. The Lancet 2014 Feb;383(9916):524-532. [doi: 10.1016/S0140-6736(13)62218-7]
17. Arbyn M, Ronco G, Anttila A, et al. Evidence regarding human papillomavirus testing in secondary prevention of cervical cancer. Vaccine (Auckl) 2012 Nov 20;30 Suppl 5:F88-F99. [doi: 10.1016/j.vaccine.2012.06.095] [Medline: 23199969]
18. Bouvard V, Wentzensen N, Mackie A, et al. The IARC perspective on cervical cancer screening. N Engl J Med 2021 Nov 11;385(20):1908-1918. [doi: 10.1056/NEJMsr2030640] [Medline: 34758259]
19. WHO. Guidelines for screening and treatment of precancerous lesions for cervical cancer prevention. 2021. URL: https://www.who.int/publications/i/item/9789240030824
20. Mungo C, Guliam A, Chinula L, et al. Comparison of the ScreenFire and Xpert HPV assays for the detection of human papillomavirus and cervical precancer among women living with HIV in Malawi. medRxiv 2024 Feb 23:2024.02.21.24303142. [doi: 10.1101/2024.02.21.24303142] [Medline: 38617305]
21. WHO. Comprehensive cervical cancer control: a guide to essential practice. 2014.
22. Cheung LC, Egemen D, Chen X, et al. 2019 ASCCP risk-based management consensus guidelines: methods for risk estimation, recommended management, and validation. J Low Genit Tract Dis 2020 Apr;24(2):90-101. [doi: 10.1097/LGT.0000000000000528] [Medline: 32243306]
23. Zou H, Lu Z, Weng W, et al. Diagnosis of neurosyphilis in HIV-negative patients with syphilis: development, validation, and clinical utility of a suite of machine learning models. EClinicalMedicine 2023 Aug;62:102080. [doi: 10.1016/j.eclinm.2023.102080] [Medline: 37533423]
24. Oikonomou EK, Spatz ES, Suchard MA, Khera R. Individualising intensive systolic blood pressure reduction in hypertension using computational trial phenomaps and machine learning: a post-hoc analysis of randomised clinical trials. Lancet Digit Health 2022 Nov;4(11):e796-e805. [doi: 10.1016/S2589-7500(22)00170-4] [Medline: 36307193]
25. Melnikow J, Henderson JT, Burda BU, Senger CA, Durbin S, Weyrich MS. Screening for cervical cancer with high-risk human papillomavirus testing: updated evidence report and systematic review for the US Preventive Services Task Force. J Am Med Assoc 2018 Aug 21;320(7):687-705. [doi: 10.1001/jama.2018.10400] [Medline: 30140883]
26. Xiao T, Wang C, Yang M, et al. Use of virus genotypes in machine learning diagnostic prediction models for cervical cancer in women with high-risk human papillomavirus infection. JAMA Netw Open 2023 Aug 1;6(8):e2326890. [doi: 10.1001/jamanetworkopen.2023.26890] [Medline: 37531108]
27. van der Waal D, Bekkers RLM, Dick S, et al. Risk prediction of cervical abnormalities: the value of sociodemographic and lifestyle factors in addition to HPV status. Prev Med 2020 Jan;130:105927. [doi: 10.1016/j.ypmed.2019.105927]
28. Rothberg MB, Hu B, Lipold L, et al. A risk prediction model to allow personalized screening for cervical cancer. Cancer Causes Control 2018 Mar;29(3):297-304. [doi: 10.1007/s10552-018-1013-4]
29. Lee CH, Peng CY, Li RN, et al. Risk evaluation for the development of cervical intraepithelial neoplasia: development and validation of risk-scoring schemes. Int J Cancer 2015 Jan 15;136(2):340-349. [doi: 10.1002/ijc.28982] [Medline: 24841989]
30. Bouvard V, Baan R, Straif K, et al. A review of human carcinogens--Part B: biological agents. Lancet Oncol 2009 Apr;10(4):321-322. [doi: 10.1016/s1470-2045(09)70096-8] [Medline: 19350698]
31. Schreiberhuber L, Barrett JE, Wang J, et al. Cervical cancer screening using DNA methylation triage in a real-world population. Nat Med 2024 Aug;30(8):2251-2257. [doi: 10.1038/s41591-024-03014-6] [Medline: 38834848]
32. Marcus JZ, Cason P, Downs LS, Einstein MH, Flowers L. The ASCCP Cervical Cancer Screening Task Force Endorsement and Opinion on the American Cancer Society updated cervical cancer screening guidelines. J Low Genit Tract Dis 2021 Jul 1;25(3):187-191. [doi: 10.1097/LGT.0000000000000614] [Medline: 34138787]
33. Moy LM, Zhao FH, Li LY, et al. Human papillomavirus testing and cervical cytology in primary screening for cervical cancer among women in rural China: comparison of sensitivity, specificity, and frequency of referral. Int J Cancer 2010 Aug 1;127(3):646-656. [doi: 10.1002/ijc.25071] [Medline: 19960441]
34. Saslow D, Solomon D, Lawson HW, et al. American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology screening guidelines for the prevention and early detection of cervical cancer. Am J Clin Pathol 2012 Apr;137(4):516-542. [doi: 10.1309/AJCPTGD94EVRSJCG] [Medline: 22431528]
35. Gower JC. A general coefficient of similarity and some of its properties. Biometrics 1971 Dec;27(4):857. [doi: 10.2307/2528823]
36. McInnes L, Healy J, Saul N, Großberger L. UMAP: uniform manifold approximation and projection. arXiv. Preprint posted online on 2018arXiv:180203426.
37. Reese JT, Blau H, Casiraghi E, et al. Generalisable long COVID subtypes: findings from the NIH N3C and RECOVER programmes. eBioMedicine 2023 Jan;87:104413. [doi: 10.1016/j.ebiom.2022.104413] [Medline: 36563487]

XSL•FO
RenderX

38. Zhou W, He MM, Wang F, Xu RH, Wang F, Zhao Q. Latent class analysis-derived classification improves the cancer-specific death stratification of molecular subtyping in colorectal cancer. NPJ Precis Oncol 2023 Jun 23;7(1):60. [doi: 10.1038/s41698-023-00412-w] [Medline: 37353681]
39. Faghri F, Brunn F, Dadu A, et al. Identifying and predicting amyotrophic lateral sclerosis clinical subgroups: a population-based machine-learning study. Lancet Digit Health 2022 May;4(5):e359-e369. [doi: 10.1016/S2589-7500(21)00274-0] [Medline: 35341712]
40. Lu Z, Fu L, Yang L, et al. Hesitancy to undergo SARS-CoV-2 rapid antigen testing in China: nationwide cross-sectional study. JMIR Public Health Surveill 2023 May 1;9:e43555. [doi: 10.2196/43555] [Medline: 36888911]
41. Xu E, Xie Y, Al-Aly Z. Long-term neurologic outcomes of COVID-19. Nat Med 2022 Nov;28(11):2406-2415. [doi: 10.1038/s41591-022-02001-z]
42. Chen T, Guestrin C, editors. XGBoost: a scalable tree boosting system. 2016 Presented at: KDD '16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
43. Lundberg SM, Erion G, Chen H, et al. From local explanations to global understanding with explainable AI for trees. Nat Mach Intell 2020 Jan;2(1):56-67. [doi: 10.1038/s42256-019-0138-9] [Medline: 32607472]
44. Luo W, Phung D, Tran T, et al. Guidelines for developing and reporting machine learning predictive models in biomedical research: a multidisciplinary view. J Med Internet Res 2016 Dec 16;18(12):e323. [doi: 10.2196/jmir.5870] [Medline: 27986644]
45. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD statement. Ann Intern Med 2015 Jan 6;162(1):55-63. [doi: 10.7326/M14-0697]
46. Lu Z. lulab.utils: Supporting Functions Maintained by Zhen Lu. 2024. URL: https://CRAN.R-project.org/package=lulab.utils
47. Lu Z. SCREENing. URL: https://github.com/Leslie-Lu/SCREENing
48. Rosty C, Sheffer M, Tsafrir D, et al. Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. Oncogene 2005 Oct 27;24(47):7094-7104. [doi: 10.1038/sj.onc.1208854] [Medline: 16007141]
49. Li Y, Wang C, Ma A, et al. Identification of HPV oncogene and host cell differentiation associated cellular heterogeneity in cervical cancer via single‐cell transcriptomic analysis. J Med Virol 2023 Aug;95(8). [doi: 10.1002/jmv.29060]
50. Sait K, Turki R, Abuzenadah AM, Jiffiri OH, Bohmaidah A, Sohrab SS. Genetic diversity and phylogenetic analysis of HPV 16 & 18 variants isolated from cervical specimens of women in Saudi Arabia. Saudi J Biol Sci 2019 Feb;26(2):317-324. [doi: 10.1016/j.sjbs.2018.05.005] [Medline: 31485171]
51. Balanda M, Fernández J, Vergara N, et al. Genetic variability of human papillomavirus type 66 L1 gene among women presenting for cervical cancer screening in Chile. Med Microbiol Immunol 2019 Dec;208(6):757-771. [doi: 10.1007/s00430-019-00621-w] [Medline: 31165237]
52. Wan F, Miao X, Quraishi I, Kennedy V, Creek KE, Pirisi L. Gene expression changes during HPV-mediated carcinogenesis: a comparison between an in vitro cell model and cervical cancer. Int J Cancer 2008 Jul 1;123(1):32-40. [doi: 10.1002/ijc.23463] [Medline: 18398830]
53. Vázquez-Ortíz G, Ciudad CJ, Piña P, et al. Gene identification by cDNA arrays in HPV-positive cervical cancer. Arch Med Res 2005;36(5):448-458. [doi: 10.1016/j.arcmed.2005.04.016] [Medline: 16099320]
54. Widschwendter A, Gattringer C, Ivarsson L, et al. Analysis of aberrant DNA methylation and human papillomavirus DNA in cervicovaginal specimens to detect invasive cervical cancer and its precursors. Clin Cancer Res 2004 May 15;10(10):3396-3400. [doi: 10.1158/1078-0432.CCR-03-0143] [Medline: 15161694]
55. Shiri Aghbash P, Shirvaliloo M, Khalo Abass Kasho A, Alinezhad F, Nauwynck H, Bannazadeh Baghi H. Cluster of differentiation frequency on antigen presenting-cells: the next step to cervical cancer prognosis? Int Immunopharmacol 2022 Jul;108:108896. [doi: 10.1016/j.intimp.2022.108896] [Medline: 35640377]
56. Man I, Georges D, Bonjour M, Baussano I. Approximating missing epidemiological data for cervical cancer through footprinting: a case study in India. Elife 2023 May 25;12:e81752. [doi: 10.7554/eLife.81752] [Medline: 37227260]
57. Yang S, Wu Y, Wang S, et al. HPV-related methylation-based reclassification and risk stratification of cervical cancer. Mol Oncol 2020 Sep;14(9):2124-2141. [doi: 10.1002/1878-0261.12709] [Medline: 32408396]
58. Li M, Wei L, Sui L, et al. Guidelines for cervical cancer screening in China. Gynecol Obstet Clin Med 2023 Dec;3(4):189-194. [doi: 10.1016/j.gocm.2023.10.005]
59. Sawaya GF, Smith-McCune K, Kuppermann M. Cervical cancer screening: more choices in 2019. J Am Med Assoc 2019 May 28;321(20):2018-2019. [doi: 10.1001/jama.2019.4595] [Medline: 31135834]
60. Wang W, Arcà E, Sinha A, Hartl K, Houwing N, Kothari S. Cervical cancer screening guidelines and screening practices in 11 countries: a systematic literature review. Prev Med Rep 2022 Aug;28:101813. [doi: 10.1016/j.pmedr.2022.101813] [Medline: 35637896]
61. Hickey M, Basu P, Sassarini J, et al. Managing menopause after cancer. The Lancet 2024 Mar;403(10430):984-996. [doi: 10.1016/S0140-6736(23)02802-7]

XSL•FO
RenderX

62. Gierisch JM, Coeytaux RR, Urrutia RP, et al. Oral contraceptive use and risk of breast, cervical, colorectal, and endometrial cancers: a systematic review. Cancer Epidemiol Biomarkers Prev 2013 Nov 1;22(11):1931-1943. [doi: 10.1158/1055-9965.EPI-13-0298]

Abbreviations

CCP: cervical cancer prevention
CIN: cervical intraepithelial neoplasia
EHR: electronic health record
HPV: human papillomavirus
hrHPV: high-risk HPV
LMIC: low- and middle-income country
OR: odds ratio
OVR AUROC: one-versus-rest area under the receiver operating characteristics curve
SCREENing: clinical Subgroups for CeRvical cancEr prEvention using computational pheNomaps and machine learnING
SHAP: SHapley Additive exPlanations
UMAP: uniform manifold approximation and projection
WHO: World Health Organization
XGBoost: eXtreme gradient boosting

Edited by A Mavragani; submitted 22.10.24; peer-reviewed by G Guerrero-Contreras, ZP Liu; revised version received 29.01.25; accepted 29.01.25; published 19.03.25.

Please cite as:
Lu Z, Dong B, Cai H, Tian T, Wang J, Fu L, Wang B, Zhang W, Lin S, Tuo X, Wang J, Yang T, Huang X, Zheng Z, Xue H, Xu S, Liu S, Sun P, Zou H
Identifying Data-Driven Clinical Subgroups for Cervical Cancer Prevention With Machine Learning: Population-Based, External, and Diagnostic Validation Study
JMIR Public Health Surveill 2025;11:e67840
URL: https://publichealth.jmir.org/2025/1/e67840
doi:10.2196/67840

© Zhen Lu, Binhua Dong, Hongning Cai, Tian Tian, Junfeng Wang, Leiwen Fu, Bingyi Wang, Weijie Zhang, Shaomei Lin, Xunyuan Tuo, Juntao Wang, Tianjie Yang, Xinxin Huang, Zheng Zheng, Huifeng Xue, Shuxia Xu, Siyang Liu, Pengming Sun, Huachun Zou. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Survival Tree Analysis of Interactions Among Factors Associated With Colorectal Cancer Risk in Patients With Type 2 Diabetes: Retrospective Cohort Study

Sarah Tsz Yui Yau, MSc, MPH; Chi Tim Hung, MBBS; Eman Yee Man Leung, PhD; Albert Lee, MPH, MD, LLM; Eng Kiong Yeoh, MBBS

JC School of Public Health and Primary Care, The Chinese University of Hong Kong, 4/F, School of Public Health Building, Prince of Wales Hospital, Shatin, Hong Kong, China (Hong Kong)

Corresponding Author:
Eman Yee Man Leung, PhD
JC School of Public Health and Primary Care, The Chinese University of Hong Kong, 4/F, School of Public Health Building, Prince of Wales Hospital, Shatin, Hong Kong, China (Hong Kong)

Abstract

Background: Colorectal cancer (CRC) and diabetes share many common lifestyle risk factors, such as obesity. However, it remains largely unknown how different factors interact to influence the risk of CRC development among patients with diabetes.

Objective: This study aimed to identify the interaction patterns among factors associated with the risk of CRC incidence among patients with diabetes.

Methods: This is a retrospective cohort study conducted using electronic health records from Hong Kong. Patients who were diagnosed with type 2 diabetes and received care in general outpatient clinics between 2010 and 2019 without cancer history were included and followed up until December 2019. A conditional inference survival tree was applied to examine the interaction patterns among factors associated with the risk of CRC.

Results: A total of 386,325 patients were included. During a median follow-up of 6.2 years (IQR 3.3-8.0), 4199 patients developed CRC. Patients were first partitioned into 4 age groups by increased levels of CRC risk (≤54 vs 55 to 61 vs 62 to 73 vs >73 years). Among patients aged more than 54 years, male sex was the dominant risk factor for CRC within each age stratum and the associations lessened with age. Abdominal obesity (waist-to-hip ratio >0.95) and longer duration of diabetes (median 12, IQR 7-18 vs median 4, IQR 1-11 years) were identified as key risk factor for CRC among men aged between 62 and 73 years and women aged more than 73 years, respectively.

Conclusions: This study suggests the interaction patterns among age, sex, waist-to-hip ratio, and duration of diabetes on the risk of CRC incidence among patients with diabetes. Findings of the study may help identify target groups for public health intervention strategies.

(JMIR Public Health Surveill 2025;11:e62756) doi:10.2196/62756

KEYWORDS

colorectal cancer; risk factor; interaction; type 2 diabetes; survival analysis; decision tree; recursive partitioning; segmentation; risk stratification; public health

Introduction

Colorectal cancer (CRC) is the third most frequently diagnosed cancer worldwide [1]. Previous studies have shown that patients with diabetes are 1.3 times likely to develop CRC when compared with those without diabetes [2]. CRC and type 2 diabetes mellitus share many common lifestyle risk factors, such as obesity [3,4], smoking [5,6], heavy alcohol use [7,8], physical inactivity [9,10], Western dietary pattern [11-13], and processed meat [12-14].

While CRC and diabetes share many overlapping risk factors, little is known about how different risk factors interact in the presence of a pre-existing diabetes condition which may influence the risk of CRC development. For example, while older age and male sex are risk factors for many cancers including CRC, it remains less certain how the association differs by age and sex [15], and how to optimally separate patients with diabetes into subgroups by age and sex according to their risk levels. A previous meta-analysis [16] showed that men are 1.83 times likely to have CRC when compared with women. However, most CRC screening guidelines for average-risk individuals are age-based (mostly starting from 50 y [17], or 45 y at the earliest [18]). It remains largely unknown how the risk levels differentiate by age and sex within the population with diabetes. Several studies have incorporated

XSL•FO
RenderX

age-sex [19,20] and diabetes [19] into their proposed screening strategies. Furthermore, while adiposity is a common risk factor for diabetes and CRC [3,21], most studies focused on general obesity [3,22]. It is less clear whether general obesity may adequately represent abdominal obesity in predicting CRC risk [22-24], and whether adiposity becomes a predominant risk factor in certain subgroups of patients. A previous meta-analysis [22] found that when comparing highest with lowest categories of waist-to-hip ratio, men and women in the highest categories of waist-to-hip ratio are 1.47 and 1.3 times likely to have CRC respectively when compared with their same-sex counterparts in the lowest category. Furthermore, it is less understood whether duration of diabetes [15] or antidiabetic drug treatment [21] may influence the risk of CRC development, given the interactions between gut microbiota with obesity [25,26], insulin resistance [25,27], diabetes [25], antidiabetic drugs [25], diet [25,26], immunity [25,26], and CRC [26]. Previous studies [15,28-30] demonstrated mixed findings on the associations between duration of diabetes and CRC risk.

Previous research suggests type 2 diabetes as a heterogeneous disease [31-35], and various approaches such as clustering techniques [31-34] have been applied to explore the subtypes. For example, Ahlqvist et al [31] has used 6 variables (including age, BMI, and HbA_{1c}) to stratify patients with adult-onset diabetes into 5 distinct clusters, namely “severe autoimmune diabetes”, “severe insulin-deficient diabetes”, “severe insulin-resistant diabetes”, “mild obesity-related diabetes”, and “mild age-related diabetes”, where the latter 4 clusters may represent different subtypes within type 2 diabetes. Identifying subtypes of type 2 diabetes may help individualize most appropriate treatment to help improve clinical outcomes [31,33]. Furthermore, the interplay between genetic and environmental factors [36] may collectively contribute to the risk of diabetes and CRC. A previous meta-analysis on gene-environment interaction studies found insufficient evidence for the interaction effects between genetic and environmental factors [37]. However, a recent large-scale gene-environment interaction study [38] suggested that common genetic variants related to insulin signaling and immune function may potentially modify the association between diabetes and CRC risk. In addition, some medical conditions may share common pathophysiology with diabetes. For example, obesity, immune-mediated inflammatory bowel diseases, and metabolic disorders [39] are characterized by systematic low-grade inflammation, which may in turn promote CRC development.

Given the potential complex interplay among interrelated factors for diabetes and CRC, in this study, tree-structured survival analysis is applied to examine the potential interaction patterns among a set of covariates on CRC risk. Conditional inference survival tree [40] is a tree-structured (or recursive partitioning) algorithm embedded with statistical theory of conditional inference. Tree-structured algorithms are able to account for multicollinearity among the set of covariates, and explore the potential interactions among more than 2 covariates without an exhaustive search of all possible combinations. Compared with other tree-structured algorithms, conditional inference survival tree has the advantages of (1) incorporating a theoretical framework, (2) avoiding overfitting, (3) minimizing selection bias toward covariates with many possible values, and (4) not requiring explicit pruning [40].

This study aims to examine whether there are interaction patterns among factors associated with the risk of CRC incidence among patients with pre-existing type 2 diabetes using survival tree analysis.

Methods

Study Design and Study Population

This is a retrospective cohort study performed using territory-wide electronic health records of Hong Kong. The Hospital Authority (HA) is statutory body managing 43 public hospitals, 49 specialist outpatient clinics, and 74 general outpatient clinics over the territory. The HA maintains a centralized data repository that stores information on patients’ demographics, disease diagnoses, prescription records, laboratory measurements, as well as inpatient and outpatient visits. Since the HA provides approximately 90% specialist and inpatient care [41], most of the cancer cases were captured in the records system. Disease diagnoses were coded with the *International Classification of Diseases, Ninth Revision* (*ICD-9*) or *International Classification of Diseases, Tenth Revision* (*ICD-10*) or the *International Classification of Primary Care, Second Edition* (*ICPC-2*). Data were accessed through HA Data Collaboration Lab.

Patients

Patients who (1) were diagnosed with diabetes and (2) received a first diabetes complication screening assessment at any of the general outpatient clinics between 2010 and 2019 were initially included. Patients who (1) were diagnosed with non–type 2 diabetes, (2) had a missing date of diabetes diagnosis, (3) were diagnosed with diabetes below the age of 18 years, or (4) had a history of malignancy were excluded. Index date was defined as date of the first assessment. Patients were followed up until a cancer diagnosis, death, or December 31, 2019, whichever occurred earlier. Since the diagnosis of one cancer may influence the diagnosis of another cancer [42], those who received a cancer diagnosis at sites other than colon or rectum during follow-up were excluded. In addition, to minimize reverse causality, those who had less than 6 months of follow-up [42] were also excluded.

The diagnosis of type 2 diabetes was based on *ICPC-2* code (T90), and defined as a clinical diagnosis by clinicians, with 2 abnormal test results of plasma glucose and presentation of clinical symptoms. Type 2 diabetes is pharmaceutically treated by glucose lowering agents (most commonly used drugs are metformin, sulphonylurea, and insulin, since these drugs are less expensive options) [41]. HbA_{1c} is the primary measure used to monitor blood glucose control.

Outcome

The outcome of interest was diagnosis of CRC (*ICD-9*: 153 - 154; *ICD-10*: C18-21) during follow-up. Patients who presented with CRC symptoms (such as rectal bleeding or blood in stool) were recommended to take a fecal occult blood test. If the results were positive, patients would be referred for a

XSL•FO
RenderX

diagnostic colonoscopy procedure. For CRC diagnosis, gastroenterologists would perform a colonoscopy procedure to remove polyps and take biopsy samples for further examination by pathologists to provide a diagnosis. CRC are generally classified into 4 consensus molecular subtypes (CMSs), namely CMS1 (microsatellite instability immune), CMS2 (canonical), CMS3 (metabolic), and CMS4 (mesenchymal), with distinct characteristics and developmental pathways [43,44]. Possible treatment options include surgery, radiotherapy, chemotherapy, and targeted therapy.

Covariates

Information on input data was ascertained during the first assessment. Candidate split variables included demographics (age and sex), duration of diabetes, medical history (ischemic heart disease, cerebrovascular disease, heart failure, hypertension, chronic kidney disease, liver cirrhosis, chronic obstructive pulmonary disease, pneumonia, and family history of diabetes), medication use (antidiabetic drugs, aspirin, nonsteroidal anti-inflammatory drugs, anticoagulants, antiplatelets, antihypertensive drugs, and statins), lifestyle behaviors (alcohol use and smoking), anthropometric measurements (waist-to-hip ratio and BMI), and laboratory measurements (HbA_{1c}, fasting glucose, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, triglycerides, and serum creatinine). Duration of diabetes was the time difference between diabetes diagnosis and the first assessment. Medication use was defined as whether patients had been prescribed a drug at the time of the assessment. Antidiabetic drugs included were metformin, sulfonylurea, insulin, and dipeptidyl peptidase-4 inhibitors. Laboratory measurements were taken from most recent results to the assessment date.

Data Analysis

Conditional inference survival tree [40] was applied to examine the interaction patterns among factors associated with the risk of CRC incidence. At each split, a global null hypothesis of independence between any of the candidate split variables and the risk of CRC was first tested at a prespecified α level. If rejected, a set of partial null hypotheses of independence between each covariate and the risk of CRC were then tested, and the covariate with strongest association or smallest Bonferroni-corrected *P* value would be selected as split variable. The partitioning procedures were recursively conducted until the global null hypothesis cannot be rejected. For continuous variables, the cutoff point was the optimal value to maximize the between-node differences in survival probability. The statistical significance threshold α level and maximum depth of the survival tree were set at .01 and 4, respectively. Each path from the root node to a terminal node represents an interaction pattern [45]. The effects of a split variable are conditional on split variables chosen at its ancestor nodes. Patients were partitioned into mutually exclusive subgroups at terminal nodes with most homogenous within-group survival outcomes. The CRC-free survival of subgroups of patients with diabetes was examined using Kaplan–Meier method. Model performance was evaluated using area under the curve as metric. In post hoc analyses, the associations between identified important factors and the risk of CRC were estimated using Cox regression and reported in adjusted hazard ratio (aHR) with 95% CI. Data analyses were performed using R software (version 4.2.3; R Foundation for Statistical Computing).

Ethical Considerations

Ethics approval for secondary data analysis was provided by the Chinese University of Hong Kong – Survey and Behavioural Research Ethics Committee (SBRE-22-0386). Patient consent was waived since individuals were not identifiable in this study.

Results

Overview

Figure 1 shows the flowchart of patient selection. Of the 386,325 patients included, 4199 patients developed CRC during a median follow-up of 6.2 (IQR 3.3-8.0) years. The incidence rates among female and male patients aged more than 54 years were 1.90 and 2.94 per 1000 person-years, respectively. On the other hand, the incidence rate among patients aged ≤54 years across both sexes was 0.54 per 1000 person-years. The survival tree first partitioned patients into 4 age groups (≤54 vs 55 to 61 vs 62 to 73 vs >73 years) by increasing levels of CRC risk. Among patients aged more than 54 years, male sex was identified as most dominant risk factor for CRC within each age stratum. Waist-to-hip ratio and sulfonylurea use (characterized by long duration of diabetes) emerged as important factor in differentiating the risk of CRC among men aged between 62 and 73 years and women aged more than 73 years, respectively (Figure 2 and Tables 1 and 2).

XSL•FO
RenderX

Figure 1. Flowchart of patient selection.

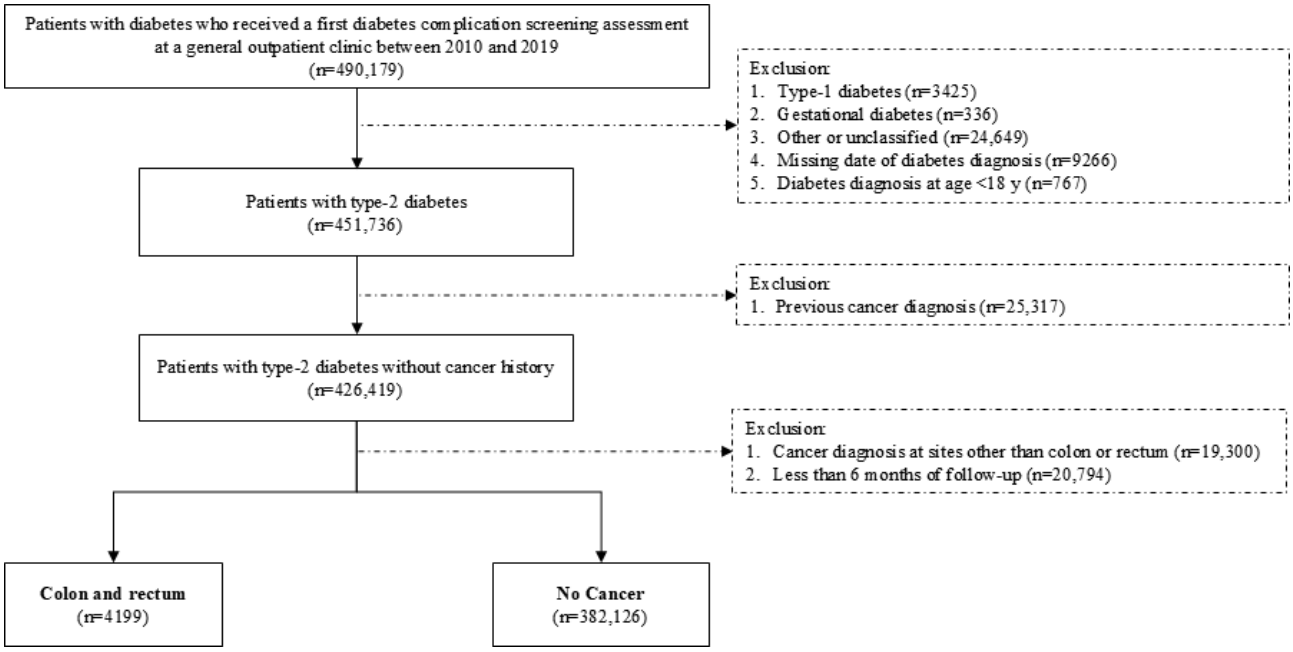
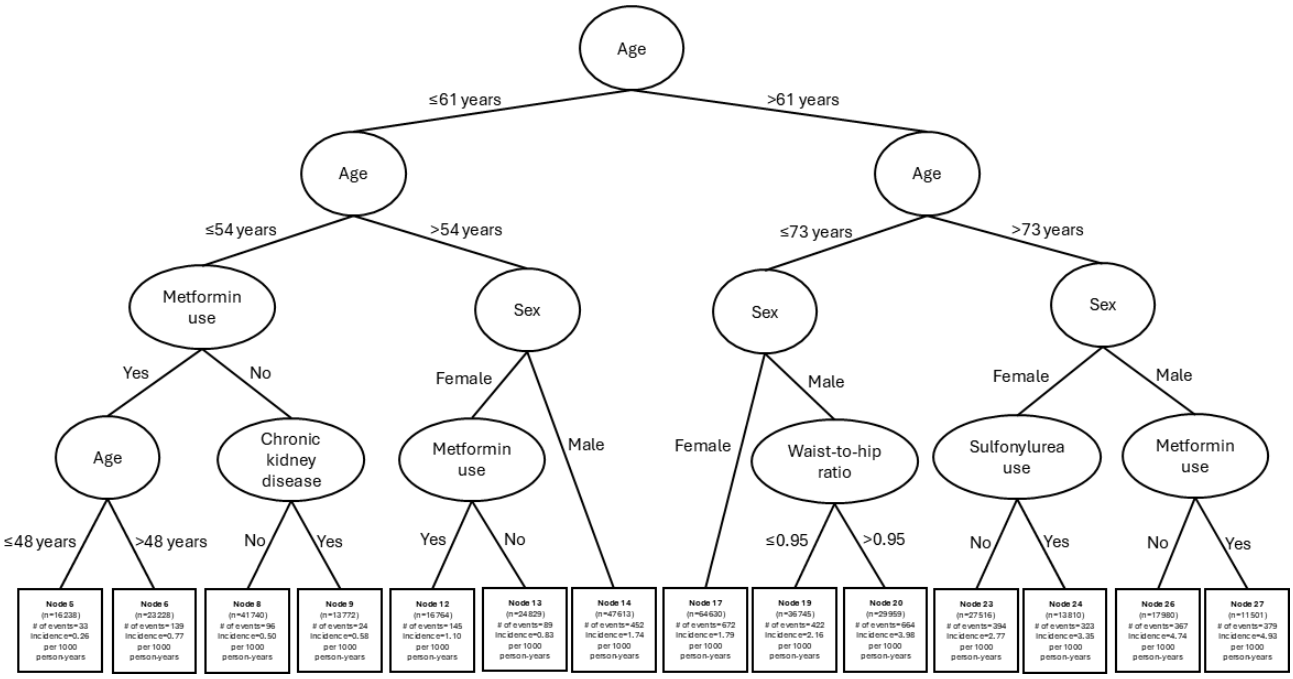


Figure 2. Survival tree diagram illustrating the interaction patterns among key factors on colorectal cancer incidence among patients with diabetes.



XSL•FO
RenderX

Table . Characteristics of different subgroups of patients with diabetes.

Characteristics	Node (5,6,8,9)	Node (12,13)	Node14	Node17	Node19	Node20	Node23	Node24	Node (26,27)
	(n=94,978)	(n=41,593)	(n=47,613)	(n=64,630)	(n=36,745)	(n=29,959)	(n=27,516)	(n=13,810)	(n=29,481)
Colorectal cancer cases during follow (up, n (%)	292 (0.31)	234 (0.56)	452 (0.95)	672 (1.04)	422 (1.15)	664 (2.22)	394 (1.43)	323 (2.34)	746 (2.53)
Demographics									
Male, n (%)	53,228 (56.04)	0 (0)	47,613 (100)	0 (0)	36,745 (100)	29,959 (100)	0 (0)	0 (0)	29,481 (100)
Age at assessment (y), mean (SD)	47.68 (6.04)	58.1 (2.00)	58.09 (2.00)	66.95 (3.40)	66.67 (3.36)	66.91 (3.38)	79.85 (4.58)	79.52 (4.35)	79.03 (4.21)
Duration of diabetes (y), median (IQR)	2 (0 - 5)	2 (1-7)	3 (1-8)	4 (1-10)	4 (1-10)	3 (1-9)	4 (1-11)	12 (7-18)	6 (1-12)
Medical history									
Ischemic heart disease, n (%)	3096 (3.26)	1026 (2.47)	3979 (8.36)	3625 (5.61)	4135 (11.25)	3351 (11.19)	2908 (10.57)	1500 (10.86)	4519 (15.33)
Cerebrovascular disease, n (%)	2353 (2.48)	1378 (3.31)	2501 (5.25)	3579 (5.54)	3062 (8.33)	2496 (8.33)	3086 (11.22)	1396 (10.11)	4108 (13.93)
Heart failure, n (%)	827 (0.87)	326 (0.78)	720 (1.51)	1005 (1.56)	753 (2.05)	708 (2.36)	1373 (4.99)	665 (4.82)	1470 (4.99)
Hypertension, n (%)	69,712 (73.40)	33,998 (81.74)	39,564 (83.09)	58,252 (90.13)	31,876 (86.75)	27,702 (92.47)	26,271 (95.48)	13,405 (97.07)	27,904 (94.65)
Chronic kidney disease, n (%)	13,772 (14.50)	4563 (10.97)	7806 (16.39)	8112 (12.55)	6529 (17.77)	5213 (17.40)	6748 (24.52)	0 (0)	5971 (20.25)
Liver cirrhosis, n (%)	2376 (2.50)	732 (1.76)	1123 (2.36)	1159 (1.79)	653 (1.78)	626 (2.09)	373 (1.36)	193 (1.40)	430 (1.46)
Chronic obstructive pulmonary disease, n (%)	75 (0.08)	33 (0.08)	206 (0.43)	139 (0.22)	453 (1.23)	475 (1.59)	251 (0.91)	97 (0.70)	1087 (3.69)
Pneumonia, n (%)	1695 (1.78)	582 (1.40)	1158 (2.43)	1402 (2.17)	1260 (3.43)	1084 (3.62)	1656 (6.02)	761 (5.51)	2511 (8.52)
Family history of diabetes, n (%)	59,255 (62.39)	23,293 (56.00)	25,487 (53.53)	28,911 (44.73)	15,127 (41.17)	11,406 (38.07)	7260 (26.38)	4412 (31.95)	7062 (23.95)
Lifestyle behaviors									
Current drinker or ex-drinker, n (%)	32,141 (33.84)	5011 (12.05)	23,558 (49.48)	6130 (9.48)	16,429 (44.71)	14,456 (48.25)	2082 (7.57)	1001 (7.25)	11,157 (37.84)

XSL•FO
RenderX

Characteristics	Node (5,6,8,9)	Node (12,13)	Node14	Node17	Node19	Node20	Node23	Node24	Node (26,27)
	(n=94,978)	(n=41,593)	(n=47,613)	(n=64,630)	(n=36,745)	(n=29,959)	(n=27,516)	(n=13,810)	(n=29,481)
Current smoker or ex-smoker, n (%)	31,727 (33.40)	1951 (4.69)	24,514 (51.49)	2596 (4.02)	19,260 (52.42)	17,033 (56.85)	1781 (6.47)	934 (6.76)	16,182 (54.89)
Anthropometric measurements									
Waist-to-hip ratio, mean (SD)	0.93 (0.06)	0.92 (0.06)	0.95 (0.05)	0.93 (0.06)	0.92 (0.03)	1 (0.04)	0.94 (0.07)	0.94 (0.07)	0.96 (0.06)
Body mass index (kg/m^2), mean (SD)	27.26 (4.82)	26.16 (4.33)	26.16 (3.82)	25.81 (4.10)	24.52 (3.23)	26.87 (3.42)	25.31 (3.87)	24.95 (3.73)	24.95 (3.29)
Laboratory measurements									
HbA_{1c} (%), mean (SD)	7.62 (1.72)	7.38 (1.40)	7.47 (1.56)	7.24 (1.24)	7.27 (1.42)	7.34 (1.37)	6.96 (1.09)	7.32 (1.15)	7.13 (1.24)
Fasting glucose (mmol/L), mean (SD)	8 (2.60)	7.65 (2.27)	7.78 (2.37)	7.45 (2.01)	7.51 (2.16)	7.55 (2.13)	7.15 (1.79)	7.39 (2.05)	7.23 (1.90)
Low-density lipoprotein cholesterol (mmol/L), mean (SD)	2.77 (0.83)	2.84 (0.86)	2.66 (0.81)	2.71 (0.84)	2.55 (0.79)	2.57 (0.79)	2.59 (0.83)	2.64 (0.79)	2.47 (0.76)
High-density lipoprotein cholesterol (mmol/L), mean (SD)	1.21 (0.31)	1.36 (0.34)	1.19 (0.30)	1.36 (0.34)	1.25 (0.33)	1.15 (0.28)	1.39 (0.36)	1.32 (0.35)	1.23 (0.33)
Triglycerides (mmol/L), mean (SD)	1.81 (1.56)	1.62 (1.06)	1.65 (1.26)	1.59 (0.94)	1.39 (0.96)	1.62 (1.10)	1.54 (0.83)	1.53 (0.87)	1.34 (0.81)
Serum creatinine (µmol/L), mean (SD)	75.2 (42.11)	65.7 (29.00)	88.52 (45.86)	71.79 (32.63)	93.11 (44.86)	95.17 (43.13)	82.34 (37.03)	86.73 (34.08)	106.12 (44.41)

XSL•FO
RenderX

Table . Medication use of different subgroups of patients with diabetes.

	Node (5,6,8,9)	Node (12,13)	Node14	Node17	Node19	Node20	Node23	Node24	Node (26,27)
	(n=94,978)	(n=41,593)	(n=47,613)	(n=64,630)	(n=36,745)	(n=29,959)	(n=27,516)	(n=13,810)	(n=29,481)
Antidiabetic drugs, n (%)									
Metformin	39,466 (41.55)	16,764 (40.30)	18,438 (38.72)	26,945 (41.69)	13,084 (35.61)	12,524 (41.80)	8271 (30.06)	10,160 (73.57)	11,501 (39.01)
Sulfonylurea	23,445 (24.68)	10,078 (24.23)	12,184 (25.59)	17,740 (27.45)	9315 (25.35)	8905 (29.72)	0 (0)	13,810 (100)	9608 (32.59)
Insulin	6460 (6.80)	2154 (5.18)	2983 (6.27)	3703 (5.73)	2362 (6.43)	1991 (6.65)	1801 (6.55)	541 (3.92)	2053 (6.96)
Dipeptidyl peptidase-4 inhibitors	3497 (3.68)	1292 (3.11)	2006 (4.21)	2383 (3.69)	1671 (4.55)	1134 (3.79)	1059 (3.85)	481 (3.48)	1302 (4.42)
Sodium-glucose co-transporter-2 inhibitors	371 (0.39)	77 (0.19)	215 (0.45)	115 (0.18)	154 (0.42)	46 (0.15)	30 (0.11)	0 (0)	35 (0.12)
Glucagon-like peptide-1 receptor agonists	123 (0.13)	18 (0.04)	26 (0.05)	11 (0.02)	4 (0.01)	11 (0.04)	2 (0.01)	0 (0)	2 (0.01)
Glucosidase inhibitors	308 (0.32)	155 (0.37)	181 (0.38)	301 (0.47)	128 (0.35)	145 (0.48)	45 (0.16)	146 (1.06)	165 (0.56)
Meglitinide	35 (0.04)	20 (0.05)	15 (0.03)	27 (0.04)	10 (0.03)	9 (0.03)	9 (0.03)	7 (0.05)	15 (0.05)
Glitazone	463 (0.49)	144 (0.35)	153 (0.32)	245 (0.38)	91 (0.25)	131 (0.44)	19 (0.07)	90 (0.65)	91 (0.31)
Any of the above	51,272 (53.98)	20,504 (49.30)	24,369 (51.18)	33,517 (51.86)	18,012 (49.02)	15,998 (53.40)	10,366 (37.67)	13,810 (100)	16,343 (55.44)
Aspirin, n (%)	374 (0.39)	4591 (11.04)	9924 (20.84)	12,660 (19.59)	10,418 (28.35)	8628 (28.80)	9160 (33.29)	4323 (31.30)	11,498 (39.00)
Nonsteroidal anti-inflammatory drugs, n (%)	49,797 (52.43)	26,009 (62.53)	24,009 (50.43)	38,648 (59.80)	17,489 (47.60)	15,470 (51.64)	15,693 (57.03)	7137 (51.68)	13,781 (46.75)
Anticoagulants, n (%)	2816 (2.96)	910 (2.19)	2636 (5.54)	2513 (3.89)	2704 (7.36)	2021 (6.75)	2024 (7.36)	759 (5.50)	2651 (8.99)
Antiplatelets, n (%)	3381 (3.56)	1429 (3.44)	4104 (8.62)	3854 (5.96)	470 (1.28)	2826 (9.43)	3687 (13.40)	0 (0)	3728 (12.65)
Antihypertensive drugs, n (%)	48,987 (51.58)	26,534 (63.79)	30,941 (64.98)	48,850 (75.58)	25,567 (69.58)	24,265 (80.99)	23,146 (84.12)	12,537 (90.78)	24,676 (83.70)
Statins, n (%)	37,282 (39.25)	20,845 (50.12)	24,015 (50.44)	36,367 (56.27)	19,802 (53.89)	16,520 (55.14)	15,871 (57.68)	5939 (43.01)	15,378 (52.16)

Age and Sex

The optimal cutoffs identified to differentiate the increasing levels of CRC risk with age were 54, 61, and 73 years (Figure 2 and Figure 3) . Among patients aged more than 54 years, male sex was the most important risk factor for CRC within each age group (Figure 2). Men had an elevated risk of developing CRC when compared with women within each age stratum (for 55 to 61 years: aHR 1.81, 95% CI 1.54 - 2.12; for 62 to 73 years: aHR 1.70, 95% CI 1.54 - 1.87; and for >73 years: aHR 1.64, 95% CI 1.48 - 1.82), after controlling for age and duration of diabetes. On the other hand, among the youngest age group (≤54 years), no split variable was identified to have individually significant effects on the risk of CRC after adjusting for age, sex, and duration of diabetes (Table 3).

XSL•FO
RenderX

Figure 3. Kaplan–Meier curves for colorectal cancer-free survival across distinct subgroups of patients with diabetes.

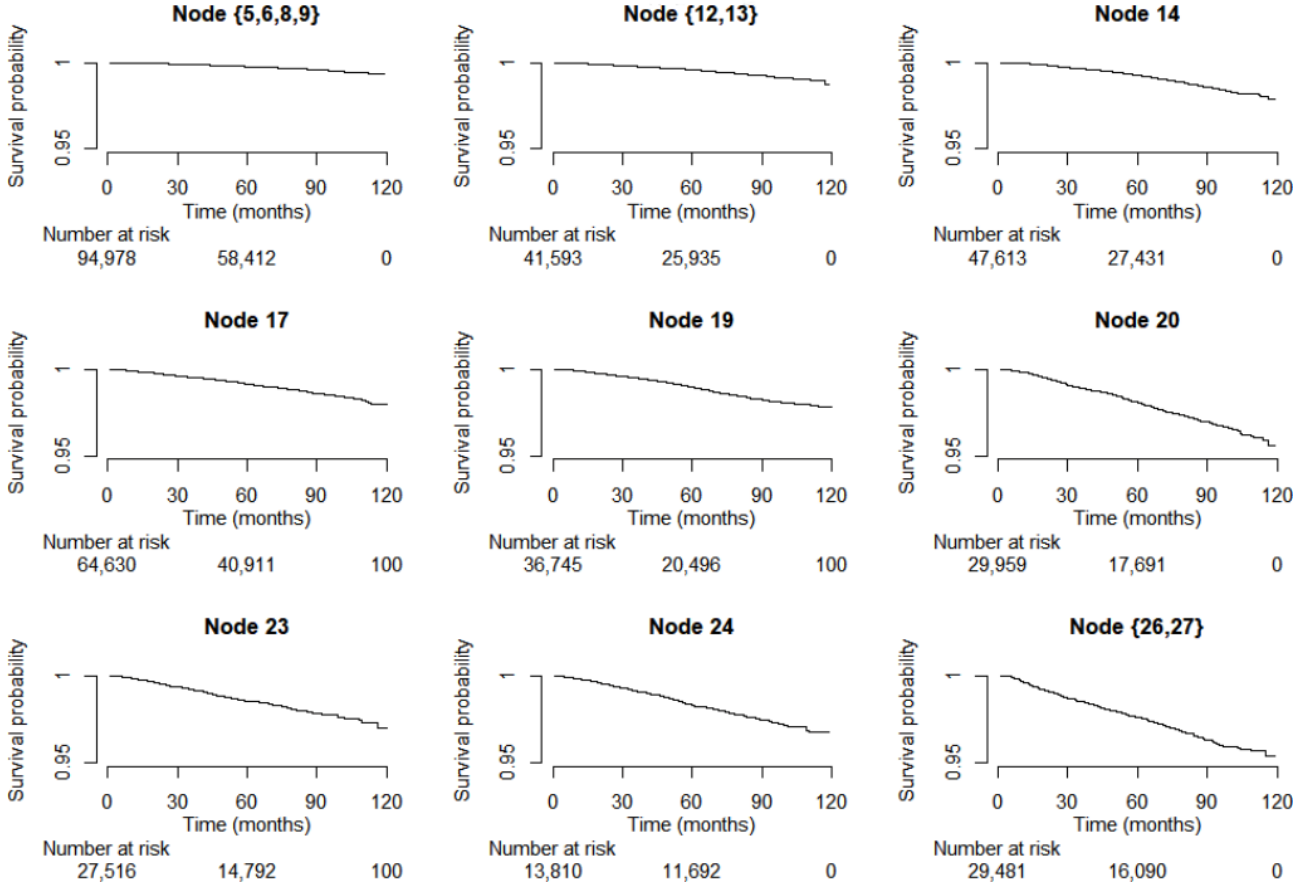


Table . Comparisons in adjusted hazard ratios of split variables between comparison nodes at the third level (comparison 1 was adjusted for age, sex, and duration of diabetes; and comparisons 2 to 4 were controlled for age and duration of diabetes).

Node	Characteristics	Comparison 1, metformin use in the youngest aHR[a] (95% CI)	Comparison 2, sex in middle-young aHR (95% CI)	Comparison 3, sex in middle-old aHR (95% CI)	Comparison 4, sex in the oldest aHR (95% CI)
Node (5,6)	Metformin users aged ≤54 years	1 (reference)	—[b]	—	—
Node (8,9)	Metformin nonusers aged ≤54 years	1.15 (0.90 - 1.47)	—	—	—
Node (12,13)	Female aged between 55 and 61 years	—	1 (reference)	—	—
Node 14	Male aged between 55 and 61 years	—	1.81 (1.54 - 2.12)	—	—
Node 17	Female aged between 62 and 73 years	—	—	1 (reference)	—
Node (19,20)	Male aged between 62 and 73 years	—	—	1.70 (1.54 - 1.87)	—
Node (23,24)	Female aged more than 73 years	—	—	—	1 (reference)
Node (26,27)	Male aged more than 73 years	—	—	—	1.64 (1.48 - 1.82)

[a]aHR: adjusted hazard ratio.

[b]Not available.

XSL•FO
RenderX

Age, Sex, and Waist-to-Hip Ratio

Among male patients aged between 62 and 73 years, elevated waist-to-hip ratio (>0.95) appeared as dominant risk factor for CRC (Figure 2). Those with a ratio of 0.95 above had an increased risk of developing CRC (aHR 1.82, 95% CI 1.61 - 2.05) when compared with those with a ratio ≤0.95, when controlling for age and duration of diabetes (Table 4). The corresponding mean waist-to-hip ratios of these 2 subgroups were 1.00 (SD 0.04) and 0.92 (SD 0.03).

XSL•FO
RenderX

Table . Comparisons in adjusted hazard ratios of split variables between sibling nodes at the bottom level (comparisons 1 to 2 were made with adjustment of age, sex, and duration of diabetes; comparisons 3, 4, and 6 were adjusted for age and duration of diabetes; and comparison 5 was adjusted for age, duration of diabetes, HbA_{1c} and fasting glucose).

Node	Characteristics	Comparison 1, age in the youngest, aHR[a] (95% CI)	Comparison 2, chronic kidney disease in the youngest, aHR (95% CI)	Comparison 3, metformin use in the middle-young female aHR (95% CI)	Comparison 4, waist-to-hip ratio in the middle-old male aHR (95% CI)	Comparison 5, sulfonylurea use in the oldest female aHR (95% CI)	Comparison 6, metformin use in the oldest male aHR (95% CI)
Node 5	Metformin users aged ≤48 years	1 (reference)	—[b]	—	—	—	—
Node 6	Metformin users aged between 49 and 54 years	1.45 (0.77 - 2.71)	—	—	—	—	—
Node 8	Metformin nonusers aged ≤54 years without chronic kidney disease	—	1 (reference)	—	—	—	—
Node 9	Metformin nonusers aged ≤54 years with chronic kidney disease	—	1.43 (0.89 - 2.29)	—	—	—	—
Node 12	Female metformin users aged between 55 and 61 years	—	—	1.06 (0.80 - 1.40)	—	—	—
Node 13	Female metformin nonusers aged between 55 and 61 years	—	—	1 (reference)	—	—	—
Node 19	Male aged between 62 and 73 years with waist-to-hip ratio ≤0.95	—	—	—	1 (reference)	—	—
Node 20	Male aged between 62 and 73 years with waist-to-hip ratio >0.95	—	—	—	1.82 (1.61 - 2.05)	—	—
Node 23	Female sulfonylurea nonusers aged >73 years	—	—	—	—	1 (reference)	—
Node 24	Female sulfonylurea users aged>73 years	—	—	—	—	1.19 (1.02 - 1.39)	—
Node 26	Male metformin nonusers aged >73 years	—	—	—	—	—	1 (reference)
Node 27	Male metformin users aged >73 years	—	—	—	—	—	1.03 (0.89 - 1.19)

[a]aHR: adjusted hazard ratio.

[b]Not available.

Age, Sex, and Duration of Diabetes

Among female patients aged more than 73 years, sulfonylurea use emerged as key factor in differentiating the risk of CRC (Figure 2). Those who had been prescribed sulfonylurea were characterized by long duration of diabetes when compared with those who were not prescribed with sulfonylurea (median 12, IQR 7-18 vs median 4, IQR 1-11 years). Among the oldest

XSL•FO
RenderX

female group, sulfonylurea users had a higher risk of developing CRC (aHR 1.19, 95% CI 1.02 - 1.39) when compared with sulfonylurea nonusers, controlling for age, duration of diabetes, HbA_{1c}, and fasting glucose (Table 4).

Model Performance

The areas under the curve of the tree model at 2, 5, and 7 years were 0.713, 0.696, and 0.685, respectively.

Discussion

Principal Findings

This study adopted a supervised learning algorithm approach, namely conditional inference survival tree embedded with statistical theory, to identifying key factors and their optimal cutoffs to differentiate the risk of CRC and separate patients with type 2 diabetes into mutually exclusive subgroups of most homogenous CRC risk. Age, sex, waist-to-hip ratio, and sulfonylurea use exhibited interaction patterns on the risk of CRC incidence among the study diabetes cohort. Despite older age, male sex, and obesity being established individual risk factors for CRC, the optimal cutoffs for age among the study cohort were identified to partition patients according to their CRC risk levels using a data-driven approach. Furthermore, while male sex emerged as most important risk factor among patients aged more than 54 years, abdominal obesity appeared as dominant risk factor among men aged 62 to 73 years. In addition, sulfonylurea use (as characterized by long duration of diabetes) was identified as key factor in differentiating CRC risk among oldest women aged more than 73 years.

This study found that male sex emerged as dominant risk factor for CRC among patients aged more than 54 years and the associations weakened with age, implying the potentially differential associations between male sex and CRC risk with age. A previous meta-analysis [16] found that the associations between male sex and CRC risk grew markedly from age 50 years but dropped gradually after reaching 60 years. A similar trend was observed in this study, where male sex was found to be an important risk factor from 55 years onwards, and the associations declined with age. In a more recent study performed among young adults who received colonoscopies in South Korea [46], male sex was only found to be a significant risk factor for overall colorectal neoplasia among patients aged 30 to 39 years, but not among those aged between 20 and 29 years. However, male sex was not associated with advanced colorectal neoplasia within each age group [46]. In another study conducted among patients who underwent colonoscopies in Korea [47], male sex was found to be a risk factor for colorectal adenoma among the overall cohort, but not younger patients aged less than 50 years. Furthermore, epigenomic evidence supports that the second period of abrupt changes in immunity occurs earlier in men than in women (early 60s vs late 60s), and the changes are more pronounced in men [48]. The discrepancy (5 to 6 years) in second period of accelerated aging coincides with life expectancy of men and women [48]. This change in immunity could partially explain the dominance of male sex on CRC risk starting from middle adulthood, and the effects decline over time as both sexes undergo rapid changes.

Furthermore, the survival tree model selected abdominal obesity (waist-to-hip ratio >0.95) over general obesity (BMI) as a dominant risk factor among middle-old men aged between 62 and 73 years. In a previous study conducted among patients who received colonoscopies in Korea [49], abdominal obesity was only found to be significantly associated with the risk of colorectal neoplasia in men but not women. Several reviews also summarized that obesity has a stronger effect on the risk of CRC among men than women [50-52], and weight gain later in life is a key risk factor for CRC among men [50,51]. Physiologically, men are predisposed to accumulating visceral fat than women [53], and this could account for the stronger link between obesity and CRC in men. The dominance of abdominal obesity as key risk factor among middle-old men could be due to decline in testosterone with age. Previous research has shown that lower testosterone in men is linked to abdominal obesity and metabolic abnormalities [51,53]. Furthermore, individuals tend to accumulate visceral fat and lose skeletal muscle with age, rendering general obesity being not sufficiently representative of abdominal obesity among older individuals [54].

In addition, oldest women aged more than 73 years who took sulfonylurea and had a longer duration of diabetes (median 12, IQR 7-18 years) appeared to have a greater risk of CRC risk than oldest women who did not take sulfonylurea and had a shorter duration of diabetes (median 4, IQR 1-11 years) in this study. It is unclear whether the elevated risk is due to longer duration of diabetes or sulfonylurea use. In the existing literature, the association between sulfonylurea use and cancer risk remains controversial. Previous studies generally reported a positive association [21] while some did not find any association [55,56]. In a study performed in Korea [57], sulfonylurea use was only found to be associated with an increased risk of CRC among patients aged 65 years or older, but not among those below the age of 65 years. In this study, sulfonylurea use only appeared to be a key factor associated with an increased risk of CRC among women aged more than 73 years. Given sulfonylurea users being characterized by long duration of diabetes, it is possible that changing gut microbiota over the course of diabetes progression [25,27] may eventually influence carcinogenesis of the colon and rectum [26].

This study identified the optimal cutoffs for age to partition patients with type 2 diabetes in the study cohort by their risk levels of CRC using a data-driven approach. While male sex is known to be a risk factor for many cancers, including CRC, the association between male sex and CRC appears to be more apparent among older patients with diabetes (aged more than 54 years). Nevertheless, the links become less strong with age. Furthermore, abdominal obesity becomes a dominant risk factor over general obesity among male patients with diabetes aged between 62 and 73 years. In addition, women aged more than 73 years with a long duration of diabetes could be potentially at greater risk of CRC than their counterparts with a shorter duration of diabetes. Findings of the study may offer valuable insights into identifying profiles of potential target groups for future public health interventions.

Strengths of this study include incorporation of a diabetes cohort with a relatively large sample size, regular follow-up, and

XSL•FO
RenderX

routine surveillance, availability of information from electronic health records, and adoption of a supervised learning algorithm such that selection of key factors is directly driven by outcome. However, several limitations may exist. First, information on some potential confounders such as dietary exposure, family history of CRC, insulin, or C-peptide was not available. Second, dosage and duration of medication use was not examined. Third, different types of sulfonylureas were not differentially evaluated in this study. Finally, the dominance of key factors and optimal cutoffs may vary across different populations. Future research is warranted to verify the findings. Examples include validation of optimal age-sex segmentation for CRC risk in a prospective cohort, incorporation of a comparison group without diabetes, exploration of the role of adiposity (general and abdominal) in CRC risk over the life course across sexes in a prospective cohort with a long-term follow-up, and new-user design in exploring drug effects (such as commonly used antidiabetic drugs and newer drugs including semaglutide) on CRC risk. Reproducible results of identified subtypes of type 2 diabetes associated with differential CRC risk may help guide public health intervention strategies and optimize clinical outcomes [33].

Conclusions

This study suggested the potential interaction patterns among age, sex, abdominal obesity, and duration of diabetes on the risk of CRC among patients with diabetes. While older age, male sex, and obesity are well-established risk factors for CRC, the combinations of risk factors may differentially contribute to different levels of CRC risk among patients with diabetes. Findings of the study may potentially help identify target groups for public health intervention strategies.

Acknowledgments

This study was supported by Hong Kong SAR government - Strategic Public Policy Research Funding Scheme (S2019.A4.015.19S).

Authors' Contributions

Conceptualization and methodology was contributed by STYY, EYML, and EKY. Data curation, formal analysis, and writing – original draft preparation was contributed by STYY. Writing – review and editing was contributed by CTH, EYML, AL, and EKY. Supervision was handled by CTH and EKY. Funding acquisition was handled by AL. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
2. Larsson SC, Orsini N, Wolk A. Diabetes mellitus and risk of colorectal cancer: a meta-analysis. J Natl Cancer Inst 2005 Nov 16;97(22):1679-1687. [doi: 10.1093/jnci/dji375] [Medline: 16288121]
3. Lauby-Secretan B, Scoccianti C, Loomis D, et al. Body fatness and cancer--viewpoint of the IARC working group. N Engl J Med 2016 Aug 25;375(8):794-798. [doi: 10.1056/NEJMsr1606602] [Medline: 27557308]
4. Jayedi A, Soltani S, Motlagh SZT, et al. Anthropometric and adiposity indicators and risk of type 2 diabetes: systematic review and dose-response meta-analysis of cohort studies. BMJ 2022 Jan 18;376:e067516. [doi: 10.1136/bmj-2021-067516] [Medline: 35042741]
5. Botteri E, Iodice S, Bagnardi V, Raimondi S, Lowenfels AB, Maisonneuve P. Smoking and colorectal cancer: a meta-analysis. JAMA 2008 Dec 17;300(23):2765-2778. [doi: 10.1001/jama.2008.839] [Medline: 19088354]
6. Pan A, Wang Y, Talaei M, Hu FB, Wu T. Relation of active, passive, and quitting smoking with incident type 2 diabetes: a systematic review and meta-analysis. Lancet Diabetes Endocrinol 2015 Dec;3(12):958-967. [doi: 10.1016/S2213-8587(15)00316-2] [Medline: 26388413]
7. McNabb S, Harrison TA, Albanes D, et al. Meta-analysis of 16 studies of the association of alcohol with colorectal cancer. Int J Cancer 2020 Feb 1;146(3):861-873. [doi: 10.1002/ijc.32377] [Medline: 31037736]
8. Knott C, Bell S, Britton A. Alcohol consumption and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis of more than 1.9 million individuals from 38 observational studies. Diabetes Care 2015 Sep;38(9):1804-1812. [doi: 10.2337/dc15-0710] [Medline: 26294775]
9. Wolin KY, Yan Y, Colditz GA. Physical activity and risk of colon adenoma: a meta-analysis. Br J Cancer 2011 Mar 1;104(5):882-885. [doi: 10.1038/sj.bjc.6606045] [Medline: 21304525]
10. Aune D, Norat T, Leitzmann M, Tonstad S, Vatten LJ. Physical activity and the risk of type 2 diabetes: a systematic review and dose-response meta-analysis. Eur J Epidemiol 2015 Jul;30(7):529-542. [doi: 10.1007/s10654-015-0056-z] [Medline: 26092138]

XSL•FO
RenderX

11. Garcia-Larsen V, Morton V, Norat T, et al. Dietary patterns derived from principal component analysis (PCA) and risk of colorectal cancer: a systematic review and meta-analysis. Eur J Clin Nutr 2019 Mar;73(3):366-386. [doi: 10.1038/s41430-018-0234-7] [Medline: 30050075]
12. Fung TT, Schulze M, Manson JE, Willett WC, Hu FB. Dietary patterns, meat intake, and the risk of type 2 diabetes in women. Arch Intern Med 2004 Nov 8;164(20):2235-2240. [doi: 10.1001/archinte.164.20.2235] [Medline: 15534160]
13. van Dam RM, Rimm EB, Willett WC, Stampfer MJ, Hu FB. Dietary patterns and risk for type 2 diabetes mellitus in U.S. men. Ann Intern Med 2002 Feb 5;136(3):201-209. [doi: 10.7326/0003-4819-136-3-200202050-00008] [Medline: 11827496]
14. Farvid MS, Sidahmed E, Spence ND, Mante Angua K, Rosner BA, Barnett JB. Consumption of red meat and processed meat and cancer incidence: a systematic review and meta-analysis of prospective studies. Eur J Epidemiol 2021 Sep;36(9):937-951. [doi: 10.1007/s10654-021-00741-9] [Medline: 34455534]
15. Ma Y, Yang W, Song M, et al. Type 2 diabetes and risk of colorectal cancer in two large U.S. prospective cohorts. Br J Cancer 2018 Nov;119(11):1436-1442. [doi: 10.1038/s41416-018-0314-4] [Medline: 30401889]
16. Nguyen SP, Bent S, Chen YH, Terdiman JP. Gender as a risk factor for advanced neoplasia and colorectal cancer: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2009 Jun;7(6):676-681. [doi: 10.1016/j.cgh.2009.01.008] [Medline: 19514116]
17. Bénard F, Barkun AN, Martel M, von Renteln D. Systematic review of colorectal cancer screening guidelines for average-risk adults: summarizing the current global recommendations. World J Gastroenterol 2018 Jan 7;24(1):124-138. [doi: 10.3748/wjg.v24.i1.124] [Medline: 29358889]
18. US Preventive Services Task Force, Davidson KW, Barry MJ, et al. Screening for colorectal cancer: US Preventive Services Task Force recommendation statement. JAMA 2021 May 18;325(19):1965-1977. [doi: 10.1001/jama.2021.6238] [Medline: 34003218]
19. Ali Khan U, Fallah M, Sundquist K, Sundquist J, Brenner H, Kharazmi E. Risk of colorectal cancer in patients with diabetes mellitus: a Swedish nationwide cohort study. PLoS Med 2020 Nov;17(11):e1003431. [doi: 10.1371/journal.pmed.1003431] [Medline: 33186354]
20. Chen X, Heisser T, Cardoso R, Hoffmeister M, Brenner H. Personalized initial screening age for colorectal cancer in individuals at average risk. JAMA Netw Open 2023 Oct 2;6(10):e2339670. [doi: 10.1001/jamanetworkopen.2023.39670] [Medline: 37878311]
21. Giovannucci E, Harlan DM, Archer MC, et al. Diabetes and cancer: a consensus report. Diabetes Care 2010 Jul;33(7):1674-1685. [doi: 10.2337/dc10-0666] [Medline: 20587728]
22. Dong Y, Zhou J, Zhu Y, et al. Abdominal obesity and colorectal cancer risk: systematic review and meta-analysis of prospective studies. Biosci Rep 2017 Dec 22;37(6):BSR20170945. [doi: 10.1042/BSR20170945] [Medline: 29026008]
23. Safizadeh F, Mandic M, Schöttker B, Hoffmeister M, Brenner H. Central obesity may account for most of the colorectal cancer risk linked to obesity: evidence from the UK Biobank prospective cohort. Int J Obes (Lond) 2024 Nov 19. [doi: 10.1038/s41366-024-01680-7] [Medline: 39562688]
24. Safizadeh F, Mandic M, Hoffmeister M, Brenner H. Colorectal cancer and central obesity. JAMA Netw Open 2025 Jan 2;8(1):e2454753. [doi: 10.1001/jamanetworkopen.2024.54753] [Medline: 39820694]
25. Crudele L, Gadaleta RM, Cariello M, Moschetta A. Gut microbiota in the pathogenesis and therapeutic approaches of diabetes. EBioMedicine 2023 Nov;97:104821. [doi: 10.1016/j.ebiom.2023.104821] [Medline: 37804567]
26. Wong SH, Yu J. Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. Nat Rev Gastroenterol Hepatol 2019 Nov;16(11):690-704. [doi: 10.1038/s41575-019-0209-8] [Medline: 31554963]
27. Takeuchi T, Kubota T, Nakanishi Y, et al. Gut microbial carbohydrate metabolism contributes to insulin resistance. Nature New Biol 2023 Sep;621(7978):389-395. [doi: 10.1038/s41586-023-06466-x] [Medline: 37648852]
28. Campbell PT, Deka A, Jacobs EJ, et al. Prospective study reveals associations between colorectal cancer and type 2 diabetes mellitus or insulin use in men. Gastroenterology 2010 Oct;139(4):1138-1146. [doi: 10.1053/j.gastro.2010.06.072] [Medline: 20633560]
29. Yang YX, Hennessy S, Lewis JD. Type 2 diabetes mellitus and the risk of colorectal cancer. Clin Gastroenterol Hepatol 2005 Jun;3(6):587-594. [doi: 10.1016/s1542-3565(05)00152-7] [Medline: 15952101]
30. Hu FB, Manson JE, Liu S, et al. Prospective study of adult onset diabetes mellitus (type 2) and risk of colorectal cancer in women. JNCI Journal of the National Cancer Institute 1999 Mar 17;91(6):542-547. [doi: 10.1093/jnci/91.6.542]
31. Ahlqvist E, Storm P, Käräjämäki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. Lancet Diabetes Endocrinol 2018 May;6(5):361-369. [doi: 10.1016/S2213-8587(18)30051-2] [Medline: 29503172]
32. Zaharia OP, Strassburger K, Strom A, et al. Risk of diabetes-associated diseases in subgroups of patients with recent-onset diabetes: a 5-year follow-up study. Lancet Diabetes Endocrinol 2019 Sep;7(9):684-694. [doi: 10.1016/S2213-8587(19)30187-1] [Medline: 31345776]
33. Misra S, Wagner R, Ozkan B, et al. Precision subclassification of type 2 diabetes: a systematic review. Commun Med (Lond) 2023 Oct 5;3(1):138. [doi: 10.1038/s43856-023-00360-3] [Medline: 37798471]

XSL•FO
RenderX

34. Sarría-Santamera A, Orazumbekova B, Maulenkul T, Gaipov A, Atageldiyeva K. The identification of diabetes mellitus subtypes applying cluster analysis techniques: a systematic review. Int J Environ Res Public Health 2020 Dec 18;17(24):9523. [doi: 10.3390/ijerph17249523] [Medline: 33353219]
35. Kononenko IV, Smirnova OM, Mayorov AY, Shestakova MV. Classification of diabetes. World Health Organization 2019. What's new? Diabetes mellitus 2020;23(4):329-339. [doi: 10.14341/DM12405]
36. Virolainen SJ, VonHandorf A, Viel K, Weirauch MT, Kottyan LC. Gene-environment interactions and their impact on human health. Genes Immun 2023 Feb;24(1):1-11. [doi: 10.1038/s41435-022-00192-6] [Medline: 36585519]
37. Yang T, Li X, Montazeri Z, et al. Gene-environment interactions and colorectal cancer risk: an umbrella review of systematic reviews and meta-analyses of observational studies. Int J Cancer 2019 Nov 1;145(9):2315-2329. [doi: 10.1002/ijc.32057] [Medline: 30536881]
38. Dimou N, Kim AE, Flanagan O, et al. Probing the diabetes and colorectal cancer relationship using gene - environment interaction analyses. Br J Cancer 2023 Aug;129(3):511-520. [doi: 10.1038/s41416-023-02312-z] [Medline: 37365285]
39. Adolph TE, Meyer M, Jukic A, Tilg H. Heavy arch: from inflammatory bowel diseases to metabolic disorders. Gut 2024 Jul 11;73(8):1376-1387. [doi: 10.1136/gutjnl-2024-331914] [Medline: 38777571]
40. Hothorn T, Hornik K, Zeileis A. Unbiased recursive partitioning: a conditional inference framework. J Comput Graph Stat 2006 Sep;15(3):651-674. [doi: 10.1198/106186006X133933]
41. Lau IT. A clinical practice guideline to guide a system approach to diabetes care in Hong Kong. Diabetes Metab J 2017 Apr;41(2):81-88. [doi: 10.4093/dmj.2017.41.2.81] [Medline: 28447435]
42. Johnson JA, Carstensen B, Witte D, et al. Diabetes and cancer (1): evaluating the temporal relationship between type 2 diabetes and cancer incidence. Diabetologia 2012 Jun;55(6):1607-1618. [doi: 10.1007/s00125-012-2525-1] [Medline: 22476947]
43. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015 Nov;21(11):1350-1356. [doi: 10.1038/nm.3967] [Medline: 26457759]
44. Fessler E, Medema JP. Colorectal cancer subtypes: developmental origin and microenvironmental regulation. Trends Cancer 2016 Sep;2(9):505-518. [doi: 10.1016/j.trecan.2016.07.008] [Medline: 28741479]
45. Ramezankhani A, Tohidi M, Azizi F, Hadaegh F. Application of survival tree analysis for exploration of potential interactions between predictors of incident chronic kidney disease: a 15-year follow-up study. J Transl Med 2017 Nov 28;15(1):240. [doi: 10.1186/s12967-017-1346-x] [Medline: 29183386]
46. Kim NH, Jung YS, Yang HJ, et al. Prevalence of and risk factors for colorectal neoplasia in asymptomatic young adults (20-39 years old). Clin Gastroenterol Hepatol 2019 Jan;17(1):115-122. [doi: 10.1016/j.cgh.2018.07.011] [Medline: 30025922]
47. Lee SE, Jo HB, Kwack WG, Jeong YJ, Yoon YJ, Kang HW. Characteristics of and risk factors for colorectal neoplasms in young adults in a screening population. World J Gastroenterol 2016 Mar 14;22(10):2981-2992. [doi: 10.3748/wjg.v22.i10.2981] [Medline: 26973394]
48. Márquez EJ, Chung CH, Marches R, et al. Sexual-dimorphism in human immune system aging. Nat Commun 2020 Feb 6;11(1):751. [doi: 10.1038/s41467-020-14396-9] [Medline: 32029736]
49. Kim NH, Jung YS, Park JH, Park DI, Sohn CI. Abdominal obesity is more predictive of advanced colorectal neoplasia risk than overall obesity in men: a cross-sectional study. J Clin Gastroenterol 2019 Aug;53(7):e284-e290. [doi: 10.1097/MCG.0000000000001086] [Medline: 29939870]
50. Baraibar I, Ros J, Saoudi N, et al. Sex and gender perspectives in colorectal cancer. ESMO Open 2023 Apr;8(2):101204. [doi: 10.1016/j.esmoop.2023.101204] [Medline: 37018873]
51. Kim H, Giovannucci EL. Sex differences in the association of obesity and colorectal cancer risk. Cancer Causes Control 2017 Jan;28(1):1-4. [doi: 10.1007/s10552-016-0831-5]
52. Ning Y, Wang L, Giovannucci EL. A quantitative analysis of body mass index and colorectal cancer: findings from 56 observational studies. Obes Rev 2010 Jan;11(1):19-30. [doi: 10.1111/j.1467-789X.2009.00613.x] [Medline: 19538439]
53. Tchernof A, Després JP. Pathophysiology of human visceral obesity: an update. Physiol Rev 2013 Jan;93(1):359-404. [doi: 10.1152/physrev.00033.2011] [Medline: 23303913]
54. Barzilai N, Huffman DM, Muzumdar RH, Bartke A. The critical role of metabolic pathways in aging. Diabetes 2012 Jun;61(6):1315-1322. [doi: 10.2337/db11-1300] [Medline: 22618766]
55. Soranna D, Scotti L, Zambon A, et al. Cancer risk associated with use of metformin and sulfonylurea in type 2 diabetes: a meta-analysis. Oncologist 2012;17(6):813-822. [doi: 10.1634/theoncologist.2011-0462] [Medline: 22643536]
56. Zhao H, Liu Z, Zhuo L, et al. Sulfonylurea and cancer risk among patients with type 2 diabetes: a population-based cohort study. Front Endocrinol (Lausanne) 2022;13:874344. [doi: 10.3389/fendo.2022.874344] [Medline: 35846337]
57. Shin CM, Kim N, Han K, et al. Anti-diabetic medications and the risk for colorectal cancer: a population-based nested case-control study. Cancer Epidemiol 2020 Feb;64:101658. [doi: 10.1016/j.canep.2019.101658] [Medline: 31887708]

Abbreviations

aHR: adjusted hazard ratio
CMS: consensus molecular subtype

XSL•FO
RenderX

CRC: colorectal cancer
HA: Hospital Authority
***ICD-10*:** *Interactional Classification of Diseases, Tenth Revision*
***ICD-9*:** *Interactional Classification of Diseases, Ninth Revision*
***ICPC-2*:** *International Classification of Primary Care, Second Edition*

Edited by AS Bhagavathula; submitted 31.05.24; peer-reviewed by C Baxter, WQ Chen; revised version received 24.01.25; accepted 18.02.25; published 29.04.25.

Please cite as:
Yau STY, Hung CT, Leung EYM, Lee A, Yeoh EK
Survival Tree Analysis of Interactions Among Factors Associated With Colorectal Cancer Risk in Patients With Type 2 Diabetes: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e62756
URL: https://publichealth.jmir.org/2025/1/e62756
doi:10.2196/62756

© Sarah Tsz Yui Yau, Chi Tim Hung, Eman Yee Man Leung, Albert Lee, Eng Kiong Yeoh. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Area-Level Indices and Health Care Use in a Pediatric Brain and Central Nervous System Tumor Cohort: Observational Study

Yvette H Tran[1*], BS, MS; Seho Park[2*], PhD; Scott L Coven[3,4*], DO, MPH; Eneida A Mendonca[5,6*], MD, PhD

[1]Department of Health Policy and Management, Richard M Fairbanks School of Public Health, Indiana University Indianapolis, 1050 Wishard Boulevard, Indianapolis, IN, United States

[2]Department of Industrial and Data Engineering, Hongik University, Seoul, Republic of Korea

[3]Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN, United States

[4]Riley Children's Hospital, Indianapolis, IN, United States

[5]Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[6]University of Cincinnati College of Medicine, Cincinnati, OH, United States

[*]all authors contributed equally

Corresponding Author:
Yvette H Tran, BS, MS
Department of Health Policy and Management, Richard M Fairbanks School of Public Health, Indiana University Indianapolis, 1050 Wishard Boulevard, Indianapolis, IN, United States

Abstract

Background: While survival among pediatric patients with cancer has advanced, disparities persist. Public health tools such as the Area Deprivation Index, the Child Opportunity Index (COI), and the Social Vulnerability Index (SVI) are potential proxies for social determinants of health and could help researchers, public health practitioners, and clinicians identify neighborhoods or populations most likely to experience adverse outcomes. However, evidence regarding their relationship with health care use, especially in the pediatric population with cancer, remains mixed.

Objective: We sought to evaluate the relationship between emergency department (ED) visits and hospitalizations with these area-level indices in our study population.

Methods: We conducted a cross-sectional study of pediatric patients with brain and central nervous system tumors in a single Midwestern state who were diagnosed between 2010 and 2020. We fitted zero-inflated Poisson models for counts of ED and inpatient visits to determine if any of these use measures were associated with our 3 area-level indices. Finally, we mapped index quintiles onto neighborhoods to visualize and compare how each index differentially ranks neighborhoods.

Results: Our study cohort consisted of 524 patients; 78.6% (n=412) of them had no recorded ED visit, and 39.7% (n=208) had no record of hospitalization. Moderate (coefficient=0.306; *P*=.01) and high (coefficient=0.315; *P*=.01) deprivation were associated with more ED visits. Both low child opportunity (coefficient=0.497; *P*<.001) and very high child opportunity (coefficient=0.328; *P*=.01) were associated with more ED visits. All quintiles of SVI were associated with ED visits, but the relationship was not dose-dependent. Low and very high deprivation were associated with hospitalizations, but COI and SVI were not. Additionally, by overlaying index quintiles onto census tracts and census block groups, we showed that most patients who had an ED visit lived in disadvantaged neighborhoods based on Area Deprivation Index rankings, but not necessarily COI or SVI rankings.

Conclusions: Although indices provide useful context about the environment in which our patient population resides in, we found little evidence that neighborhood conditions as measured by these indices consistently or reliably relate to health care use.

(JMIR Public Health Surveill 2025;11:e66834) doi:10.2196/66834

KEYWORDS

child opportunity; neighborhood deprivation; area deprivation; social vulnerability; health care use; social determinants of health; pediatrics; health disparities

Introduction

Extant research has attributed social determinants of health (SDOH) to differential health outcomes and health risks, sometimes more so than clinical care [1]. In pediatric populations with cancer, factors such as race, social deprivation, socioeconomic status, and insurance coverage are significantly associated with inequities in access to clinical trials and survival outcomes [2,3]. Even if pediatric patients with cancer do survive, they remain at higher risk than the general population for all-cause and health-related mortality [4]. Thus, mitigating these nonclinical factors may improve the life course of this vulnerable

XSL•FO
RenderX

population. However, to progress toward this goal, clinicians, public health professionals, health care organizations, payers, and policy makers must gather SDOH data. SDOH information can be collected in a variety of ways, including through aggregating individual responses to social risk assessments, using national databases (eg, Agency for Healthcare Research and Quality's SDOH Database), or relying on existing composite indices (or neighborhood indices) that rank or score geographic areas (eg, census tracts, census block groups, counties, or zip codes) on multiple sociodemographic and socioeconomic variables.

Using neighborhood indices as an approach to approximate SDOH affords several advantages. They offer convenience by operationalizing SDOH in the form of easy-to-interpret numeric rankings or scores and can be integrated into electronic health records or clinical data warehouses as clinical decision support tools [5,6]. The addition of neighborhood indices may also improve the performance risk prediction models when compared to models that only use patient- or individual-level characteristics [7-9]. Moreover, they may be valuable to investigators because their data are often publicly available and can be easily linked to patient data by matching geocodes—thus facilitating work in identifying nonclinical predictors of health outcomes, risk stratifying patient and patient populations by SDOH, and building prediction models. Finally, indices could serve as viable complements to screenings or substitutes when it is not feasible to assess patient populations' social needs, either due to workflow and time constraints, lack of training, or patients' uneasiness with revealing sensitive [10].

Researchers, clinicians, policy makers, and payors have access to numerous publicly available area-level indices. Three indices that have been linked to measures of health care use include the Area Deprivation Index (ADI), the Child Opportunity Index (COI), and the Social Vulnerability Index (SVI). The ADI was originally created by the US Health Resources and Services Administration and contains education, employment, housing, and poverty indicators [11,12]. Data used to build the ADI primarily come from the Census and the American Community Survey, and both sources have been demonstrated to underestimate foreign-born and noncitizen populations [13]. Some indices, such as the COI, are pediatric-specific. Investigators at Brandeis University and the Kirwan Institute at the Ohio State University collaborated to create and validate the COI, which ranks neighborhoods on health, socioeconomic, and education, and environmental indicators related to child health outcomes [14,15]. Not all indices were intended for research and clinical practice. For example, the Centers for Disease Control and Prevention and the Agency for Toxic Substances and Disease Registry originally created the SVI to help communities plan for and manage emergency responses to natural disasters and infectious disease outbreaks and include measures of socioeconomic status, household composition, minority status and language, housing, and transportation [16,17]. However, in recent years, investigators have used SVI scores to operationalize SDOH to determine associations between SDOH and health care access, surgical outcomes, and cancer screening [18-20].

Indices may be useful for informing clinical care, policy, and research; however, many publicly available indices exist without clear guidance or a gold standard for when or which index should be used. In many cases, investigators choose indices based on familiarity or available geographic information [21]. Our work seeks to examine the concordance of 3 indices by testing these indices on a single-state pediatric central nervous system (CNS) tumor cohort. We selected this cohort because brain and spinal cord tumors are the leading disease-related causes of death among children and teens and despite significant gains in survival outcomes, this population remains at elevated risk for poor health and higher health care use than the general pediatric population [22].

In this study, we raised and answered the following questions: do neighborhood indices correlate with emergency department (ED) use and hospitalizations, and is there discordance between these neighborhood indices? We hypothesize that neighborhood indices are correlated with downstream outcomes such as a higher likelihood of ED visits and hospitalizations, but effect sizes will be small. We also hypothesize that indices are discordant due to varying weighting techniques and input data used to construct the indices. Therefore, we hypothesize that indices also cannot be used interchangeably. Our study aims include determining whether indices are associated with health care use (ED visits and inpatient encounters).

Methods

Study Setting, Data, and Population

This cross-sectional analysis uses patient-level data from the Indiana Network for Patient Care research database, which includes most health care organizations throughout Indiana. We used data from this repository to identify pediatric patients with brain and CNS tumors. We then cross-referenced the patients with relevant diagnosis codes to state cancer registry data to exclude patients who might have had health care encounters in Indiana but were not Indiana residents or did not receive a plurality of care in Indiana. Our inclusion criteria were any patient diagnosed before the age of 21 years and any patient with an encounter between January 1, 2010, through December 31, 2020. We bounded the cutoff age at 21 years rather than 18 years to reflect the common practice in pediatric oncologists to treat patients up to 21 years and the *Bright Futures* guidelines from the American Academy of Pediatrics. We excluded any patients' missing sex, date of birth, or addresses. We then geocoded patients' addresses and assigned each patient a census block group or census tract. For patients with multiple addresses in their electronic health records, we geocoded the modal address. Data for the ADI, COI, and SVI were publicly available and downloadable as either text files or comma-delimited files. These indices were then linked to the patient cohort data by matching on census tracts for the COI and the SVI and census block groups for the ADI. The ADI is only validated at the census block group level; thus, we did not aggregate census block groups up to the census tract level.

Measures

We evaluated 2 visit types: inpatient encounters (hospitalizations) and ED. For each patient in our cohort, we

XSL•FO
RenderX

computed the number of ED encounters and hospitalizations. Our main analyses were count data for each visit type. We also created 2 binary measures for whether a patient had any ED visit or any hospitalization during our study period. Finally for each visit type, we computed a categorical measure (0 visits, at least 1 visit, and 2 or more visits).

Our main exposure variables were index quintiles, which we derived from raw scores for each index. The ADI and COI score neighborhoods numerically, from 1 to 100 based on deprivation or opportunity. For the ADI, higher scores mean more severe deprivation or disadvantage, while for the COI, higher scores mean greater opportunity. The SVI scores neighborhoods (census tracts) on a scale of 0 to 1, with values closer to 1 indicating higher vulnerability. Other controls and covariates included age at diagnosis, diagnosis category, sex, race as a binary indicator of "White" or "non-White," ethnicity (Hispanic or Latino vs not Hispanic or Latino), urbanicity (a binary indicator of urban or nonurban based on rural-urban codes), and distance (in kilometers) to the state's children's hospital where the patients were diagnosed or had at least 1 encounter.

Statistical Analysis

Descriptive analyses were first used to summarize the patient cohort's characteristics. For categorical variables, we presented frequencies and proportions. For continuous variables, we presented means and SDs. Next, we fitted a series of zero-inflated Poisson regression models for each of our outcome variables. In these models, we regressed visit type (ED visit counts or inpatient visit counts) on index quintiles and controlled for age at diagnosis, patient sex, diagnosis category, race, ethnicity, urbanicity, and visit year. We selected distance to the pediatric cancer center as our "inflate" variable based on the assumption that the farther away the patient is from the site of cancer treatment, the more likely the patient would rely on ED for some care. With regard to hospitalizations, patients closer to the pediatric cancer center might have their health better managed—reducing the possibility of being hospitalized. Alternatively, patients living far away from the pediatric cancer center, especially those in remote or rural areas, might use less care, therefore contributing excess zeros in our data. This approach addresses our overdispersion problem by accounting for an excess 0 ED visits and an excess number of patients with no hospitalizations in our data. We considered findings statistically significant at the α=.05 level. Statistical analyses were conducted in Stata SE (version 18; StataCorp LLC).

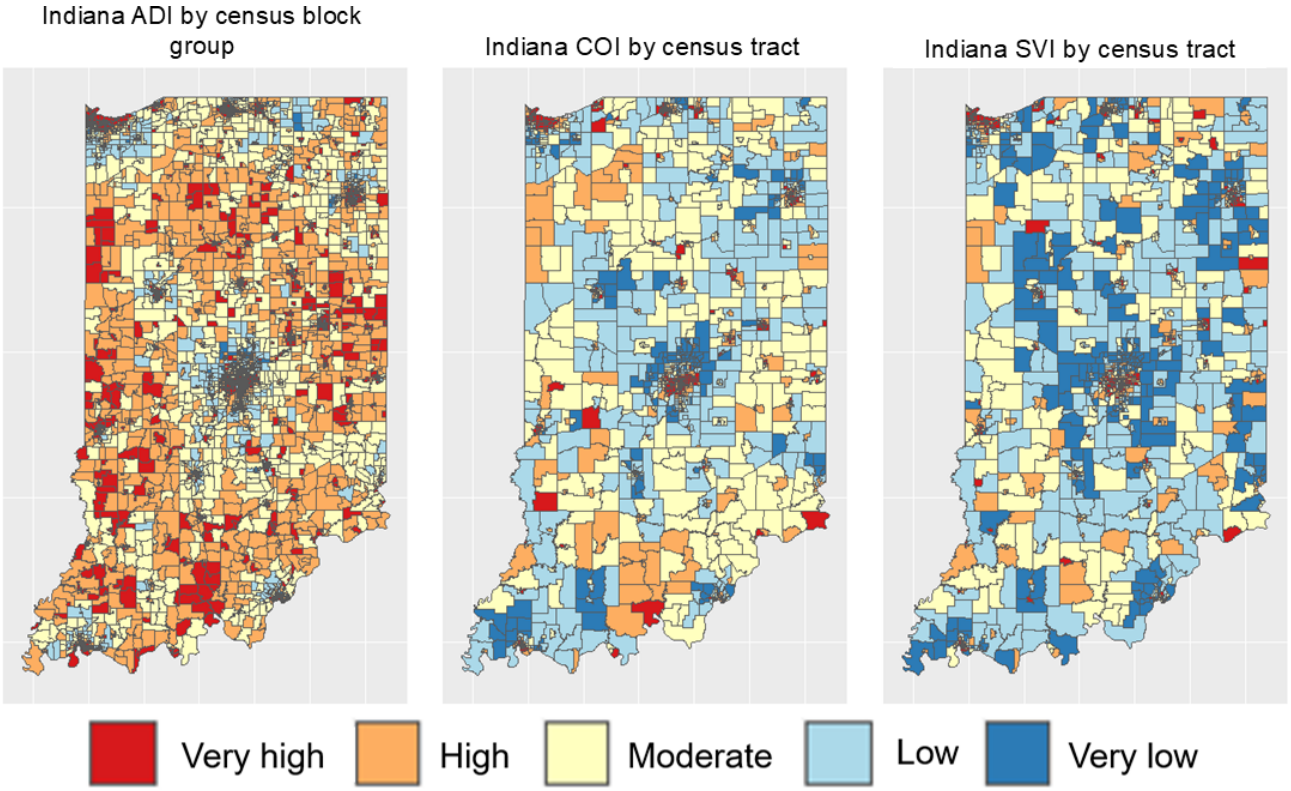
Supplemental Analyses

To test alternative specifications, we also fitted multivariable logit and ordered logit models for each visit type in additional analyses, which we report in Multimedia Appendix 1. In logit models, we regressed visit type as a binary indicator of "0" for never having a visit or "1" for having at least 1 visit on the index quintile and controlled for age at diagnosis, patient sex, diagnosis category, race, ethnicity, urbanicity, visit year, and distance to the pediatric cancer center. In our ordered logit models, we regressed visit type on the same set of explanatory variables and controls. However, our outcome variables were categorical variables with 3 levels ("never having a visit," "1 visit," and "more than 1 visit").

For each index, we overlaid index quintiles onto census tracts or census block groups to help readers visualize how indices classify or categorize neighborhoods in Indiana (Figure 1). All dataset assembly and visualizations were conducted using R (version 4.0.4; R Foundation for Statistical Computing) in the RStudio environment.

XSL•FO
RenderX

Figure 1. Map of index quintiles by census tracts or census block groups. ADI: Area Deprivation Index; COI: Child Opportunity Index; SVI: Social Vulnerability Index.



Ethical Considerations

This study received expedited approval from the Indiana University Institutional Review Board (IRB #13628) on December 3, 2021, and satisfied the waiver of authorization criteria in accordance with 45 CFR 164.512(i)(2)(ii). Because we analyzed deidentified secondary data, study participants did not receive compensation. All data were stored and analyzed in Health Insurance Portability and Accountability Act–compliant environments.

Results

Sample Characteristics

After applying our inclusion and exclusion criteria, our study cohort contained 524 patients. Table 1 presents sample characteristics. Of 524 eligible patients, 78.6% (n=412) had 0 ED visits, and 39.7% (n=208) had no hospitalizations. The average age at diagnosis was 7.9 (SD 5.4) years. Our patient cohort contained a higher proportion of male patients (n=297, 56.7%) to female patients (n=227, 43.3%). Additionally, the majority of our patient cohort was non-Hispanic White, urban-dwelling, and residing approximately 130 km from the local pediatric cancer center.

XSL•FO
RenderX

Table . Descriptive statistics.

	Values (N=524)
Age at diagnosis (years), mean (SD)	7.9 (5.4)
Sex, n (%)	
Female	227 (43.3)
Male	297 (56.7)
Race, n (%)	
Non-White	68 (12.9)
White	456 (87)
Ethnicity, n (%)	
Hispanic or Latino	36 (6.9)
Not Hispanic or Latino	488 (93.1)
Diagnosis category, n (%)	
Blastoma	68 (12.9)
Cytoma	140 (26.7)
Glioma	99 (18.9)
Other brain tumors or malignancies	174 (33.2)
Unspecified brain tumor	43 (8.2)
Urban, n (%)	
No	166 (31.7)
Yes	358 (68.3)
Area Deprivation Index quintile[a], n (%)	
Very low deprivation	106 (20.5)
Low deprivation	102 (19.7)
Moderate deprivation	105 (20.3)
High deprivation	104 (20.1)
Very high deprivation	101 (19.5)
Child Opportunity Index quintile[b], n (%)	
Very low opportunity	120 (22.9)
Low opportunity	134 (25.6)
Moderate opportunity	127 (24.2)
High opportunity	47 (8.9)
Very high opportunity	117 (22.3)
Social Vulnerability Index quintile[c], n (%)	
Very low vulnerability	107 (20.4)
Low vulnerability	111 (21.2)
Moderate vulnerability	95 (18.1)
High vulnerability	94 (17.9)
Very high vulnerability	117 (22.3)
Distance to cancer center (km)	130.1 (114.1)
Emergency visits, n (%)	
None	412 (78.6)
At least 1	80 (21.4)

XSL•FO
RenderX

	Values (N=524)
Inpatient visits, n (%)	
None	208 (39.7)
At least 1	316 (60.3)

[a]Area Deprivation Index rankings were converted into quintiles in order of lowest to highest deprivation.

[b]Child Opportunity Index rankings were converted into quintiles in order of lowest to highest opportunity.

[c]Social Vulnerability Index rankings were converted into quintiles in order of lowest to highest vulnerability.

ED Visits

In comparing zero-inflated models for ED visits, we observed similarities and differences. Across all 3 models presented in Table 2, children who were diagnosed at an older age and male patients were less likely to have ED visits. Furthermore, urban-dwelling patients were more likely to have received care in an ED. However, we observed discordance among indices. Moderate (coefficient=0.306; *P*=.01) and high (coefficient=0.315; *P*=.01) deprivation were associated with more ED visits compared to very low deprivation. In contrast, both low child opportunity (coefficient=0.497; *P*<.001) and very high child opportunity (coefficient=0.328; *P*=.01) were associated with more ED visits compared to very low child opportunity. With regard to the SVI, all levels were associated with increased ED visits compared to very low social vulnerability.

XSL•FO
RenderX

Table . Zero-inflated Poisson regression output, emergency department visits.

	ADI[a]		COI[b]		SVI[c]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Age at diagnosis (years)	−0.027 (−0.04 to −0.01)	<.01	−0.029 (−0.04 to −0.01)	<.01	−0.027 (−0.04 to −0.01)	<.001
Male	−0.279 (−0.43 to −0.13)	<.001	−0.272 (−0.42 to −0.122)	<.001	−0.281 (−0.43 to −0.13)	<.001
Diagnosis category						
Blastoma	__[d]	—	—	—	—	—
Cytoma	−0.284 (−0.52 to −0.05)	.02	−0.333 (−0.56 to −0.10)	.004	−0.322 (−0.56 to −0.09)	.007
Glioma	−0.277 (−0.55 to −0.01)	.05	−0.235 (−0.50 to 0.03)	.08	−0.241 (−0.51 to 0.03)	.08
Other	−0.194 (−0.41 to 0.03)	.09	−0.191 (−0.41 to 0.03)	.09	−0.208 (−0.43 to 0.01)	.07
Unspecified	−0.486 (−0.81 to −0.16)	.001	−0.460 (−0.79 to −0.13)	.006	−0.494 (−0.82 to −0.17)	.001
Race (White)	−0.195 (−0.42 to 0.03)	.08	−0.102 (−0.33 to 0.13)	.38	−0.068 (−0.29 to 0.15)	.55
Not Hispanic or Latino	−0.074 (−0.44 to 0.29)	.69	0.049 (−0.31 to 0.40)	.79	−0.014 (−0.35 to 0.38)	.94
Urban	0.221 (0.03 to 0.41)	.03	0.275 (0.07 to 0.48)	.01	0.214 (−0.02 to 0.41)	.03
Index quintile						
1=Very low	—	—	—	—	—	—
2=Low	−0.045 (−0.31 to 0.22)	.74	0.497 (0.27 to 0.72)	<.001	0.31 (0.06 to 0.56)	.01
3= Moderate	0.306 (0.07 to 0.54)	.01	0.210 (−0.03 to 0.45)	.08	0.47 (0.20 to 0.73)	<.001
4=High	0.315 (0.07 to 0.56)	.01	−0.190 (−0.50 to 0.12)	.22	0.61 (0.36 to 0.86)	<.001
5=Very high	0.202 (−0.06 to 0.47)	.13	0.328 (0.08 to 0.58)	.01	0.61 (0.35 to 0.87)	<.001
Visit year	−0.003 (−0.03 to 0.02)	.83	0.001 (−0.02 to 0.03)	.94	−0.001 (−0.03 to 0.02)	.94
Inflate						
Distance to pediatric cancer center (km)	0.007 (0.00-0.00)	<.001	0.007 (0.01 to 0.01)	<.001	0.01 (0.01 to 0.01)	<.001

[a]Area Deprivation Index (ADI) rankings were converted into quintiles in order of lowest to highest deprivation.

[b]Child Opportunity Index (COI) rankings were converted into quintiles in order of lowest to highest opportunity.

[c]Social Vulnerability Index (SVI) rankings were converted into quintiles in order of lowest to highest vulnerability.

[d]Not available.

Hospitalizations

In comparing zero-inflated model output among indices, we also find that indices do not produce similar estimates. Furthermore, we also do not observe a dose-dependent relationship between hospitalization and level of deprivation, opportunity, or vulnerability. Both low deprivation (coefficient=0.273; *P*=.003) and very high deprivation (coefficient=0.207; *P*=.03) appear to be associated with more hospitalization episodes. Meanwhile, models using the COI or the SVI produced null findings (Table 3). Across all zero-inflated models with hospitalization counts as the outcome variable, urbanicity was associated with fewer hospitalization episodes.

XSL•FO
RenderX

Table . Zero-inflated Poisson regression output, hospitalizations.

	ADI[a]		COI[b]		SVI[c]	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Age at diagnosis (years)	−0.007 (−0.02 to 0.00)	.21	−0.004 (−0.01 to 0.01)	.42	−0.006 (−0.02 to 0.00)	.27
Male	−0.021 (−0.14 to 0.09)	.72	−0.040 (−0.16 to 0.08)	.50	−0.031 (−0.15 to 0.08)	.59
Diagnosis category						
Blastoma	__[d]	—	—	—	—	—
Cytoma	−1.311 (−1.48 to −1.14)	<.001	−1.285 (−1.45 to −1.12)	<.001	−1.296 (−1.46 to −1.13)	<.001
Glioma	−1.166 (−1.37 to −0.96)	<.001	−1.161 (−1.37 to −0.96)	<.001	−1.168 (−1.37 to −0.96)	<.001
Other	−0.465 (−0.60 to −0.33)	<.001	−0.435 (−0.57 to −0.30)	<.001	−0.453 (−0.59 to −0.32)	<.001
Unspecified	−0.727 (−0.99 to −0.47)	<.001	−0.716 (−0.98 to −0.45)	<.001	−0.721 (−0.98 to −0.46)	<.001
Race (White)	0.014 (−0.18 to 0.21)	.89	0.008 (−0.19 to 0.20)	.94	−0.010 (−0.20 to 0.18)	.92
Not Hispanic or Latino	−0.091 (−0.29 to 0.10)	.35	−0.083 (−0.28 to 0.12)	.42	−0.111 (−0.31 to 0.09)	.29
Urban	−0.209 (−0.33 to −0.08)	<.001	−0.178 (−0.31 to −0.04)	.01	−0.226 (−0.35 to −0.10)	<.001
Index quintile						
1=Very low	—	—	—	—	—	—
2=Low	0.273 (0.09 to 0.46)	<.001	0.083 (−0.08 to 0.24)	.32	0.035 (−0.14 to 0.21)	.69
3=Moderate	0.122 (−0.05 to 0.30)	.17	0.084 (−0.08 to 0.25)	.33	−0.321 (−0.21 to 0.14)	.72
4=High	0.125 (−0.05 to 0.30)	.17	0.004 (−0.22 to 0.23)	.97	−0.104 (−0.29 to 0.08)	.27
5=Very high	0.207 (−0.02 to 0.39)	.03	0.033 (−0.16 to 0.23)	.74	−0.018 (−0.21 to 0.17)	.85
Visit year	−0.034 (−0.05 to −0.02)	<.001	−0.037 (−0.05 to −0.02)	<.001	−0.034 (−0.05 to −0.02)	<.001
Inflate						
Distance to pediatric cancer center (km)	−0.000 (−0.00 to 0.00)	.59	−0.000 (−0.00 to 0.00)	.49	−0.000 (−0.00 to 0.00)	.48

[a]Area Deprivation Index (ADI) rankings were converted into quintiles in order of lowest to highest deprivation.

[b]Child Opportunity Index (COI) rankings were converted into quintiles in order of lowest to highest opportunity.

[c]Social Vulnerability Index (SVI) rankings were converted into quintiles in order of lowest to highest vulnerability.

[d]Not available.

Supplemental Analyses

In our logit models, we did not find any relationship between index quintile and hospitalization (Tables S1-S3 in Multimedia Appendix 1). Moderate (adjusted odds ratio [aOR] 1.55, 95% CI 1.14-2.09; *P*=.005) and high deprivation (aOR 1.47, 95% CI 1.07-2.02; *P*=.02) were associated with increased odds of ED use compared to the reference category (very low area deprivation; Table S4 in Multimedia Appendix 1). Similarly, with the COI, we did not find a dose-dependent relationship. Low opportunity (aOR 1.46, 95% CI 1.09-1.96; *P*=.01) and moderate opportunity (aOR 1.72, 95% CI 1.29-2.30; *P*<.001) were associated with increased odds of ED visits compared to very low opportunity (Table S5 in Multimedia Appendix 1). When we regressed the binary indicator of any ED visit on the SVI quintile, we did observe that moderate to very high social vulnerability was significantly associated with higher odds of ED use but not low social vulnerability (Table S6 in Multimedia

XSL•FO
RenderX

Appendix 1). In ordered logit models, we also failed to uncover any concordance or consistency among indices in predicting ED use or hospitalization. Across the ordered logit models, we did consistently find that the farther patients are from the pediatric cancer center, the lower the odds of ED and inpatient visits (Tables S7-S12 in Multimedia Appendix 1).

Maps

For the most part, all 3 indices covered all neighborhoods in Indiana, except for the SVI, which lacked data for 1 census tract. After overlaying index quintiles onto neighborhoods, we found some differences and similarities in how indices rank or score neighborhoods. Although it was not fully possible to compare the ADI to the COI or SVI since the ADI is validated at the census block group level and the COI and SVI are validated at the census tract level, all 3 indices ranked neighborhoods in Central Indiana (near Indianapolis) less favorably due to higher deprivation, lower opportunity, or higher social vulnerability (Figure 1 and Table 4). Of all 3 indices, the ADI, which produces neighborhood estimates at a more granular geographic unit, ranked neighborhoods as highly or very highly deprived at 68% of the time compared to 42% for the COI and 37% for the SVI. Additionally, the ADI was more likely to classify neighborhoods' near state borders as highly or very highly deprived.

Table . Frequency table for index categories.

	ADI[a]	COI[b]	SVI[c],
Very low, n (%)	89 (1.7)	342 (22.7)	281 (18.7)
Low, n (%)	464 (8.9)	283 (18.8)	337 (22.4)
Moderate, n (%)	1128 (21.7)	379 (25.2)	328 (21.8)
High, n (%)	1612 (31)	345 (22.9)	310 (20.6)
Very high, n (%)	1909 (36.7)	156 (10.4)	248 (16.5)
Total census block group or census tracts, n	5202	1505	1504

[a]ADI: Area Deprivation Index.

[b]COI: Child Opportunity Index.

[c]SVI: Social Vulnerability Index.

Discussion

Principal Findings

In several of our models, we did find statistically significant associations between indices and ED visits or hospitalizations, but these findings were often isolated to specific quintiles of the indices. In other words, we did not observe evidence of dose-dependent or meaningful associations between these indices and ED or inpatient visits. Additionally, we also observed discordance between the indices in both our regression models and our maps. We expected to observe discordance due to the difference in input variables, data sources, and weighting decisions made by those who created and validated these indices. By demonstrating discordance in our findings, we contribute further evidence that indices are not interchangeable—that selection of index has implications for conclusions. Furthermore, while scholars in the pediatric cancer domain have incorporated indices into epidemiological and predictive modeling, our study is one of the first to compare multiple indices. Still, more research is needed to help establish a gold standard and an informative guide for researchers to reference in selecting which index to use, and many scholars have made significant progress toward qualitatively comparing the indices based on input variables, geographic nits, and study contexts [21,23].

Despite the lack of meaningful associations between indices and health care use measures, we found consistent associations between urbanicity and health care use as well as age and health care use. In our zero-inflated models, patients in urban settings had greater ED use and lower hospital use, which contradicts a corpus of studies showing that rural ED use often outpaces urban ED use [24,25]. Most hospitalizations are not predictable or planned events, and we lacked sufficient data to explain why we might see lower hospitalizations among urban-dwelling patients. However, it is possible that because physician or provider supply tends to be higher in urban areas, patients may have greater access to outpatient care and thus having health conditions and comorbidities better managed than nonurban-dwelling patients. The Indiana Network for Patient Care is one of the largest state health databases in the United States, so we believe that our data capture the majority (if not all) of a patient's encounters; however, it is possible that we may have missing use data from smaller, nonintegrated or independent EDs that do not contribute or share data with the Indiana Health Information Exchange. Furthermore, we may also have failed to capture out-of-state encounters among patients who live closer to the border of other Midwestern states (eg, Ohio, Illinois, or Kentucky). Consistent across our models, children diagnosed with a brain or CNS tumor later in childhood had lower health care use than children diagnosed at a younger age. We believe this might be because children undergoing cancer treatment have less developed organs, which may make them more vulnerable to adverse side effects of treatment [26]. This finding also contradicts some studies, showing that young adults or adolescents tend to have more ED visits because their treatment plan closely resembles treatment plans for adult patients with cancer—often resulting in a higher risk of toxicity from treatment [27]. Without additional context into what additional comorbidities or chronic conditions the patients might

XSL•FO
RenderX

have and their treatment plans, we cannot make any strong conclusions about why age matters.

Limitations

This study is not without several limitations. First, our findings are not generalizable to other patient populations, given that our cohort consisted of only pediatric patients with brain and CNS cancer found in one Midwest state's health information exchange. Consequently, our patient cohort was relatively small, containing 524 patients and over 30,000 encounters. Second, a plurality of our patient cohort was of non-Hispanic or Latino ethnicity and of White race, thus limiting our ability to detect racial and ethnic disparities. Third, we had limited information about patients' health status and health care use behaviors and could not stratify or adjust for the severity of cancer, the presence of comorbidities, or if any use was preventable. Fourth, comparing indices was complicated by the fact that indices were validated at different geographical levels. For example, the COI and SVI were validated at census tract levels, while the ADI was validated at the census block group level. Census block groups are more granular than census tracts and may be more sensitive or pick up more variation than indices validated at larger geographic units such as census tracts or counties. Finally, this study qualitatively and descriptively compared indices and did not use more complex quantitative approaches such as factor analyses.

Conclusions

This study contributes evidence to understanding how neighborhood indices can be used and when they might not have much utility. We found that even though the indices occasionally produced similar results in regression models, there was also much discordance among indices. More research is needed to understand how indices are used in clinical settings and whether the information provided by these indices is relevant to clinicians, patients, and patient outcomes. Although our work has limited generalizability outside of pediatric populations with cancer and our analyses should be tested on larger or more diverse patient cohorts, our work suggests that indices may not be useful for clinicians or other providers in helping them screen out or screen for patients who may need social services. Other methods or approaches such as household material hardship screening may yield more accurate patient-level information, with the tradeoff being low response rates to surveys, submission of incomplete surveys, or staff time needed to administer and analyze survey results [28]. Future work will consider qualitative and mixed method approaches to identify clinicians using these indices and to seek insight into how such clinicians interpret and make decisions with such information.

Acknowledgments

The authors thank the Regenstrief Data Services Team for extracting patient data for this study and Jay Colbert at the Polis Center at Indiana University Indianapolis for geocoding support. YHT reports funding support from the National Library of Medicine (grant T15LM012502).

Authors' Contributions

YHT, SP, SLC, and EAM contributed to the conception and design of the study. YHT carried out the analyses with assistance from SP in the interpretation of findings. YHT drafted the initial manuscript, and EAM, SLC, and SP contributed to the writing, revision, and approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental analyses conducted to check for robustness and sensitivity of findings.
[DOCX File, 382 KB - publichealth_v11i1e66834_app1.docx]

References

1. Hood CM, Gennuso KP, Swain GR, Catlin BB. County health rankings: relationships between determinant factors and health outcomes. Am J Prev Med 2016 Feb;50(2):129-135. [doi: 10.1016/j.amepre.2015.08.024] [Medline: 26526164]
2. Graetz DE. The most vulnerable pediatric cancer patients. Lancet Reg Health Am 2023;21:100480. [doi: 10.1016/j.lana.2023.100480]
3. Tran YH, Coven SL, Park S, Mendonca EA. Social determinants of health and pediatric cancer survival: a systematic review. Pediatr Blood Cancer 2022 May;69(5):e29546. [doi: 10.1002/pbc.29546] [Medline: 35107854]
4. Ehrhardt MJ, Liu Q, Dixon SB, et al. Association of modifiable health conditions and social determinants of health with late mortality in survivors of childhood cancer. JAMA Netw Open 2023 Feb 1;6(2):e2255395. [doi: 10.1001/jamanetworkopen.2022.55395] [Medline: 36763361]

XSL•FO
RenderX

5. Gardner BJ, Pedersen JG, Campbell ME, McClay JC. Incorporating a location-based socioeconomic index into a de-identified i2b2 clinical data warehouse. J Am Med Inform Assoc 2019 Apr 1;26(4):286-293. [doi: 10.1093/jamia/ocy172] [Medline: 30715327]
6. Miliard M. Cerner unveils new interoperability tools, as CEO Brent Shafer says “innovation is accelerating”. Healthcare IT News. 2020. URL: https://www.healthcareitnews.com/news/cerner-unveils-new-interoperability-tools-ceo-brent-shafer-says-innovation-accelerating [accessed 2024-09-24]
7. Mehaffey JH, Hawkins RB, Charles EJ, et al. Socioeconomic “Distressed Communities Index” improves surgical risk-adjustment. Ann Surg 2020 Mar;271(3):470-474. [doi: 10.1097/SLA.0000000000002997] [Medline: 30741732]
8. Phelos HM, Kass NM, Deeb AP, Brown JB. Social determinants of health and patient-level mortality prediction after trauma. J Trauma Acute Care Surg 2022 Feb 1;92(2):287-295. [doi: 10.1097/TA.0000000000003454] [Medline: 34739000]
9. Vest JR, Kasthurirathne SN, Ge W, Gutta J, Ben-Assuli O, Halverson PK. Choice of measurement approach for area-level social determinants of health and risk prediction model performance. Inform Health Soc Care 2022 Jan 2;47(1):80-91. [doi: 10.1080/17538157.2021.1929999] [Medline: 34106026]
10. Phillips J, Richard A, Mayer KM, Shilkaitis M, Fogg LF, Vondracek H. Integrating the social determinants of health into nursing practice: nurses’ perspectives. J Nurs Scholarsh 2020 Sep;52(5):497-505. [doi: 10.1111/jnu.12584] [Medline: 32654364]
11. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible—the neighborhood atlas. N Engl J Med 2018 Jun 28;378(26):2456-2458. [doi: 10.1056/NEJMp1802313] [Medline: 29949490]
12. Singh GK. Area deprivation and widening inequalities in US mortality, 1969-1998. Am J Public Health 2003 Jul;93(7):1137-1143. [doi: 10.2105/ajph.93.7.1137] [Medline: 12835199]
13. Rothbaum J, Eggleston J, Bee A, Klee M, Mendez-Smith B. Addressing nonresponse bias in the american community survey during the pandemic using administrative data (ACS21-RER-05). US Census Bureau. 2021. URL: https://www.census.gov/library/working-papers/2021/acs/2021_Rothbaum_01.html [accessed 2025-04-24]
14. Acevedo-Garcia D, McArdle N, Hardy EF, et al. The Child Opportunity Index: improving collaboration between community development and public health. Health Aff (Millwood) 2014 Nov;33(11):1948-1957. [doi: 10.1377/hlthaff.2014.0679]
15. Acevedo-Garcia D, Noelke C, McArdle N, et al. Racial and ethnic inequities in children’s neighborhoods: evidence from the new Child Opportunity Index 2.0. Health Aff (Millwood) 2020 Oct 1;39(10):1693-1701. [doi: 10.1377/hlthaff.2020.00735]
16. At a glance: CDC/ATSDR Social Vulnerability Index. Agency for Toxic Substances and Disease Registry. 2022. URL: https://www.atsdr.cdc.gov/place-health/php/svi/index.html [accessed 2022-09-14]
17. Flanagan BE, Gregory EW, Hallisey EJ, Heitgerd JL, Lewis B. A Social Vulnerability Index for disaster management. J Homel Secur Emerg Manag 2011 Jan 5;8(1). [doi: 10.2202/1547-7355.1792]
18. Al Rifai M, Jain V, Khan SU, et al. State-level Social Vulnerability Index and healthcare access: the behavioral risk factor surveillance system survey. Am J Prev Med 2022 Sep;63(3):403-409. [doi: 10.1016/j.amepre.2022.03.008] [Medline: 35504796]
19. Paro A, Hyer JM, Diaz A, Tsilimigras DI, Pawlik TM. Profiles in social vulnerability: the association of social determinants of health with postoperative surgical outcomes. Surgery 2021 Dec;170(6):1777-1784. [doi: 10.1016/j.surg.2021.06.001] [Medline: 34183179]
20. Schraw JM, Peckham-Gregory EC, Rabin KR, Scheurer ME, Lupo PJ, Oluyomi A. Area deprivation is associated with poorer overall survival in children with acute lymphoblastic leukemia. Pediatr Blood Cancer 2020 Sep;67(9):e28525. [doi: 10.1002/pbc.28525] [Medline: 32573920]
21. Bucholz EM, McCormick AD, Ronai C. The Child Opportunity Index and other social determinants of health neighborhood indexes: important considerations for choosing the correct index. JAMA Pediatr 2024 Dec 1;178(12):1370-1372. [doi: 10.1001/jamapediatrics.2024.3146] [Medline: 39401022]
22. Subramanian S, Ahmad T. Childhood Brain Tumors: StatPearls; 2022.
23. Trinidad S, Brokamp C, Mor Huertas A, et al. Use of area-based socioeconomic deprivation indices: a scoping review and qualitative analysis. Health Aff (Millwood) 2022 Dec 1;41(12):1804-1811. [doi: 10.1377/hlthaff.2022.00482]
24. Greenwood-Ericksen MB, Kocher K. Trends in emergency department use by rural and urban populations in the United States. JAMA Netw Open 2019 Apr 5;2(4):e191919. [doi: 10.1001/jamanetworkopen.2019.1919] [Medline: 30977849]
25. Muelleman RL, Sullivan AF, Espinola JA, Ginde AA, Wadman MC, Camargo Jr CA. Distribution of emergency departments according to annual visit volume and urban–rural status: implications for access and staffing. Acad Emerg Med 2010 Dec;17(12):1390-1397. [doi: 10.1111/j.1553-2712.2010.00924.x]
26. Tonorezos ES, Cohn RJ, Glaser AW, et al. Long-term care for people treated for cancer during childhood and adolescence. Lancet 2022 Apr 16;399(10334):1561-1572. [doi: 10.1016/S0140-6736(22)00460-3] [Medline: 35430023]
27. Kaul S, Russell H, Livingston JA, Kirchhoff AC, Jupiter D. Emergency department visits by adolescent and young adult cancer patients compared with pediatric cancer patients in the United States. J Adolesc Young Adult Oncol 2018 Oct;7(5):553-564. [doi: 10.1089/jayao.2018.0026] [Medline: 29924663]
28. Coven SL, Tran YH, Park S, Mendonça EA. DEI-05. Assessing household material hardship in children with central nervous system (CNS) tumors. Neuro-oncology 2024 Jun 18;26(Suppl 4):iv43-iv40. [doi: 10.1093/neuonc/noae064.152]

XSL•FO
RenderX

Abbreviations

ADI: Area Deprivation Index
aOR: adjusted odds ratio
CNS: central nervous system
COI: Child Opportunity Index
ED: emergency department
SDOH: social determinants of health
SVI: Social Vulnerability Index

Edited by A Mavragani; submitted 01.10.24; peer-reviewed by A Gourishankar, D Poddighe; revised version received 15.02.25; accepted 16.02.25; published 02.05.25.

Please cite as:
Tran YH, Park S, Coven SL, Mendonca EA
Area-Level Indices and Health Care Use in a Pediatric Brain and Central Nervous System Tumor Cohort: Observational Study
JMIR Public Health Surveill 2025;11:e66834
URL: https://publichealth.jmir.org/2025/1/e66834
doi:10.2196/66834

© Yvette H Tran, Seho Park, Scott L Coven, Eneida A Mendonca. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 2.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Peer Relationships Are a Direct Cause of the Adolescent Mental Health Crisis: Interpretable Machine Learning Analysis of 2 Large Cohort Studies

Heiner Stuke[1,2], MD; Robert Schlack[3], PhD; Michael Erhart[4], PhD; Anne Kaman[4], PhD; Ulrike Ravens-Sieberer[4], PhD; Christopher Irrgang[1], PhD

[1]Centre for Artificial Intelligence in Public Health Research of the Robert Koch-Institute, Nordufer 20, Berlin, Germany

[2]Department of Psychiatry and Neurosciences at the Charité Campus Mitte, Charité University Hospital Berlin, Berlin, Germany

[3]Department B Epidemiology and Health Monitoring, Robert Koch-Institute, Berlin, Germany

[4]Department of Child and Adolescent Psychiatry and Psychotherapy and Psychosomatics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Corresponding Author:
Heiner Stuke, MD
Centre for Artificial Intelligence in Public Health Research of the Robert Koch-Institute, Nordufer 20, Berlin, Germany

Abstract

Background: Converging evidence indicates an adolescent mental health crisis in Western societies that has developed and exacerbated over the past decade. The proposed driving factors of this trend include more screen time, physical inactivity, and social isolation, but their causal influence on mental health is insufficiently understood.

Objective: The objective of this study is to test whether and based on which predictor variables the development of mental health in adolescents in the last decade can be predicted and to better understand the causal chain of factors at work.

Methods: We implemented an interpretable machine learning pipeline based on gradient boosting regression with repeated cross-validation to assess the development of mental health throughout adolescence in members of 2 longitudinal cohort studies, the British Millenium cohort (MC; n=8599) and the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) cohort (n=1212). In total, 144 (MC) and 102 (KiGGS) predictors assessed at the age of around 13.8 years (MC) and 11.6 years (KiGGS) were used to assess mental health at the ages of around 16.7 years (MC) and 16.4 years (KiGGS). Based on these predictive models, we used permutation-based feature importance analyses to identify relevant predictors and predictor domains. Moreover, we performed partial dependence analyses in a causal inference framework to determine the direct effects of physical inactivity, screen time, and peer problems on the development of mental health.

Results: The average cross-validated Pearson correlation coefficient (r) between predicted and true mental health in late adolescence was 0.614 (MC) and 0.466 (KiGGS). Feature importance analyses indicated a strong impact of preexisting mental health and weaker impacts of sex (female as a risk factor), physical health (chronic disease as a risk factor), lifestyle, and socioeconomic and family factors (eg, low parental education, income, and mental health as risk factors). Causal inference analyses suggested a strong direct effect of peer relationships, but only a small direct effect of physical inactivity and a very small direct effect of screen time.

Conclusions: Mental health development during adolescence can be assessed by a combination of variables from early adolescence. Peer problems represent an important direct cause of mental health development, and their deterioration may contribute to the current mental health crisis.

(JMIR Public Health Surveill 2025;11:e60125) doi:10.2196/60125

KEYWORDS

mental health; depression; epidemiology; machine learning; prediction; risk factors; adolescence; adolescent; adolescent health; peer; health crisis; Western countries; longitudinal study; cohort study; screen time; physical inactivity; social isolation; public health; assessments; analogous analyses; Pearson correlation

Introduction

Adolescence is a critical period of life when people explore, experiment, and are exposed to particular risk factors such as substance use, violence, bullying, and academic problems. It is when the course is set for health in adulthood [1]. The current debate about an adolescent and young adult mental health crisis in Western countries [2,3] is therefore all the more important, especially because it is substantiated by a number of recent

XSL•FO
RenderX

findings: a study comparing 2 UK birth cohorts found that while no difference was found during childhood, there were markedly more emotional problems during adolescence in individuals born 2000 - 2002 compared to individuals born 1991 - 1992 [4]. Consistently, a recent comparison of 2 methodologically standardized and comparable Dutch population surveys showed almost a doubling of affective disorders between the years around 2008 and around 2021, with the increase being particularly pronounced among young adults [5]. In the United States, after a years-long decline, rising depression among adolescents has been reported since about 2010 [6], accompanied by an increase in suicide rates in this age group [7]. While prevalence rates for externalizing mental health problems among children and adolescents in Germany were declining before the COVID-19 pandemic, emotional problems also increased here between 2003 - 2006 and 2014 - 2017, especially among girls [8]. These alarming results are in line with findings from other countries, although there is heterogeneity and effect sizes appear slightly smaller in meta-analytic summaries [9]. While reduced stigma and an increased readiness to report mental health problems might partly contribute to these trends, it is remarkable that rising prevalence rates are also reported in studies that used the same standardized and well-validated clinical interview over different time periods [5]. Moreover, investigations relating time trends of self-reported mental disorders to time trends of stigmatization of mental disorders in the same regions did not find evidence that reduced stigma explains the increase in self-reported mental health problems [10].

Summarizing the evidence suggests that since about 2010, there seems to be a marked cross-national increase in mental health problems of adolescents, particularly concerning internalizing or emotional problems and affective disorders. The causes for this increase are insufficiently understood [2]. Proposed risk factors cover rising screen time and social media consumption [11], worsened peer relationships with loneliness and social isolation [12], and physical inactivity [13]. However, the predictive strength and the causal interplay of these factors are still unclear and debated.

To address this research gap, we used multivariable machine learning models to assess the development of mental health from early to late adolescence in members of longitudinal cohort studies. This allows us to model nonlinear relationships between numerous predictors in early adolescence and the future development of mental health, as well as to test the quality of the resulting predictions by means of cross-validation. We investigated 2 large independent population-based cohort studies of adolescents, the UK Millennium Cohort (MC) study [14] and the German Health Interview and Examination Survey for Children and Adolescents (KiGGS) [15]. This approach allows a robust comparison of results between 2 independent cohorts from different countries.

The analyzed cohort studies were conducted in countries affected by the mental health crisis, and the investigated time period covers exactly the "critical period" of the crisis development from 2010 onwards. Thus, the developed models can also be used to add further evidence to the debate on driving factors of this deterioration. More specifically, it can be examined which simulated changes in risk factors that have been implicated in the ongoing mental health crisis actually led to worse predicted mental health. However, assessing beyond pure predictive relationships the causal impact of these risk factors on the development of mental health involves stronger assumptions and challenges interpretation [16]. In this work, we draw on our multivariable models in conjunction with recent research on the causal interpretation of predictive models [17] and on the appropriate selection of covariates [18]. Based on this approach, we assess the independent effect of proposed driving factors for the adolescent mental health crisis (increased screen time, less physical activity, and worsened peer relationships).

We aimed to answer 3 related research questions. (1) How well can mental health in late adolescence be assessed based on assessments in early adolescence? (2) Which assessments in early adolescence are important for these assessments? (3) What are the independent impacts of physical activity, screen time, and peer relationships on unfavorable mental health trajectories of adolescents in the years after 2010?

Methods

Analyzed Cohort Studies

For reciprocal validation, analogous analyses were conducted in 2 European cohort studies of child and adolescent health. The first one was the MC study investigates the development of children born in 2000-2002 in the United Kingdom [19]. For this study, we assessed the participants' mental health at the age of around 17 (hereafter referred to as "follow-up acquisition") based on assessments at the age of around 14 ("baseline acquisition"). The second one was the KiGGS, a population-based study for German children and adolescents with 3 longitudinal assessment waves [15]. In KiGGS, we used participants' assessments at the age of around 12 years ("Baseline acquisition") to assess their mental health at the age of around 16 years ("Follow-up acquisition"). A participant flowchart for both cohorts is shown in Figure S1 in Multimedia Appendix 1, and the final sample is described in Table 1.

XSL•FO
RenderX

Table . Descriptive statistics summary of cohort participants.

	Millennium cohort	KiGGS[a] cohort
Number of eligible participants (m/f), n (n/n)	8599 (4268/4331)	1212 (599/613)
Period of baseline acquisition	January 2015 to March 2016	June 2009 to June 2012
Period of follow-up acquisition	January 2018 to March 2019	September 2014 to August 2017
Age at baseline (years), mean (SD)	13.76 (0.46)	11.62 (0.62)
Age at follow-up (years), mean (SD)	16.73 (0.46)	16.37 (0.65)
Total number of predictors at baseline, n	144	102

[a]KiGGS: German Health Interview and Examination Survey for Children and Adolescents.

Ethical Considerations

All MC waves were approved by medical research ethics committees, specifically by the National Research Ethics Service Research Ethics Committee London—Central for the baseline acquisition (approval 13/LO/1786) and by the National Research Ethics Service Research Ethics Committee North East—York for the follow-up acquisition (approval 17/NE/0341). The ethics committee of the Charité – Universitätsmedizin Berlin has approved the KiGGS baseline study (approval 101/2000) and KiGGS Wave 1 (approval EA2/058/09), and the ethics committee of the Hannover Medical School has approved KiGGS Wave 2 (approval 2275 - 2014).

Outcome, Predictors, and Sample Weights

Consistent with previous studies [4] and because of its availability in both cohorts at both examination times, we used the emotional symptoms sum score of the parents-reported Strength and Difficulties Questionnaire (SDQ) as the outcome measure in both cohorts. The SDQ is among the most established assessment tools for mental health in children and adolescents [20]. Its emotional symptoms subscale (SDQ-E) consists of 5 items on a 0 - 2 Likert scale, which are summed to give a total score ranging from 0 to 10. It has demonstrated a high accuracy in detecting depressive and anxiety disorders [21].

Based on consultations with experts in adolescent mental health (RS, URS), the study assessments were systematically screened and a broad set of predictors of future mental health was included and assigned to the domains of mental health and well-being (eg, questions on depressive and anxiety symptoms, history of self-harm [22]), physical health (chronic diseases, BMI [23]), psychological and cognitive testing (school grades, cognitive assessments, moral attitudes), socioeconomic and family factors (family climate, socioeconomic status, and mental health problems of the parents [24,25]), lifestyle and peer relationships (substance use, media consumption, activities [11,26,27]), and others (sex [28]). A complete list of selected predictors, domain assignments, and the percentage of missing values can be found in Tables S1 and S2 in Multimedia Appendix 1. The analysis of predictor domains in addition to individual predictors aimed to simplify the interpretation of the feature importance analyses and to stabilize the results against the inclusion or exclusion of individual predictors. Hence, in contrast to data-driven dimension reduction techniques, the focus here was not on an explanation of the highest possible variance of the predictors by the factors but on their conceptual interpretability. However, post-hoc tests showed that these content-based domain assignments significantly better captured the latent factor structure of the data as compared to random assignments (Text S1 in Multimedia Appendix 1).

Both investigated cohort studies aimed for representativeness for the total population of the surveyed country. However, because individuals with certain sociodemographic criteria were underrepresented in the sampled cohort compared with the total population, both cohorts provide weight variables to give individuals from underrepresented groups more weight in the analyses. To further increase the generalizability of our results to the overall country population, we used these weights (GOVWT2 variable for the MC, wQSpop variable for KiGGS) in all univariable and multivariable analyses.

Univariable Comparisons

To obtain unadjusted associations between individual predictors and outcomes, we calculated linear models with each predictor as the only independent variable, based on data completed using K-nearest neighbors imputation (KNNImputer class of the scikit-learn toolbox for Python [29] with default parameters) and weighted with population sample weights. Exploratory tests on the quality of the values imputed in this way revealed highly significant correlations between imputed and actual values, which would not be achievable with univariate imputation methods such as mean or median imputation (Text S2 in Multimedia Appendix 1). Here, numerical predictors were z-standardized (so that the resulting coefficient values are identical to Pearson correlation coefficients). These analyses make it possible to assess correlative (uncorrected) relationships between the individual predictors of interest and mental health in later adolescence.

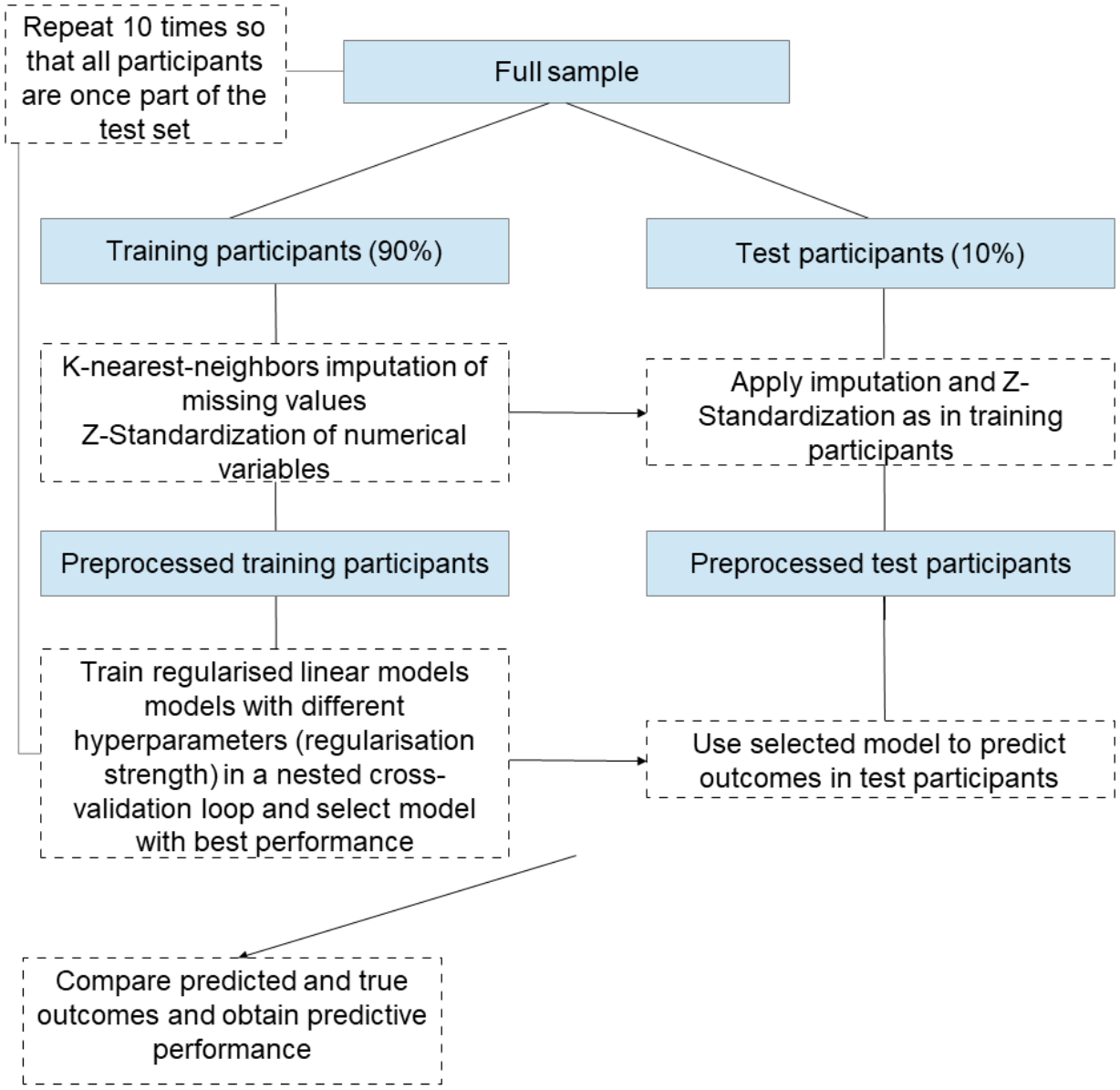
Multivariable Machine Learning Pipelines

Separately for both cohorts, we computed multivariable models that assess mental health at follow-up with all predictors combined. We used 10-fold cross-validation for a regularized linear and a nonlinear pipeline within the scikit-learn environment (Figures 1 and 2 for details). To correct for chance effects, the analyses were repeated 10 times with different random number generator seeds, and the results were averaged (so that the validation altogether amounts to a 10 times repeated 10-fold cross-validation). Nonlinear regression can improve the accuracy of predictions by modeling nonlinear relationships between predictors and outcomes, for example, in the form of interactions (ie, the effect of social media consumption on future

XSL•FO
RenderX

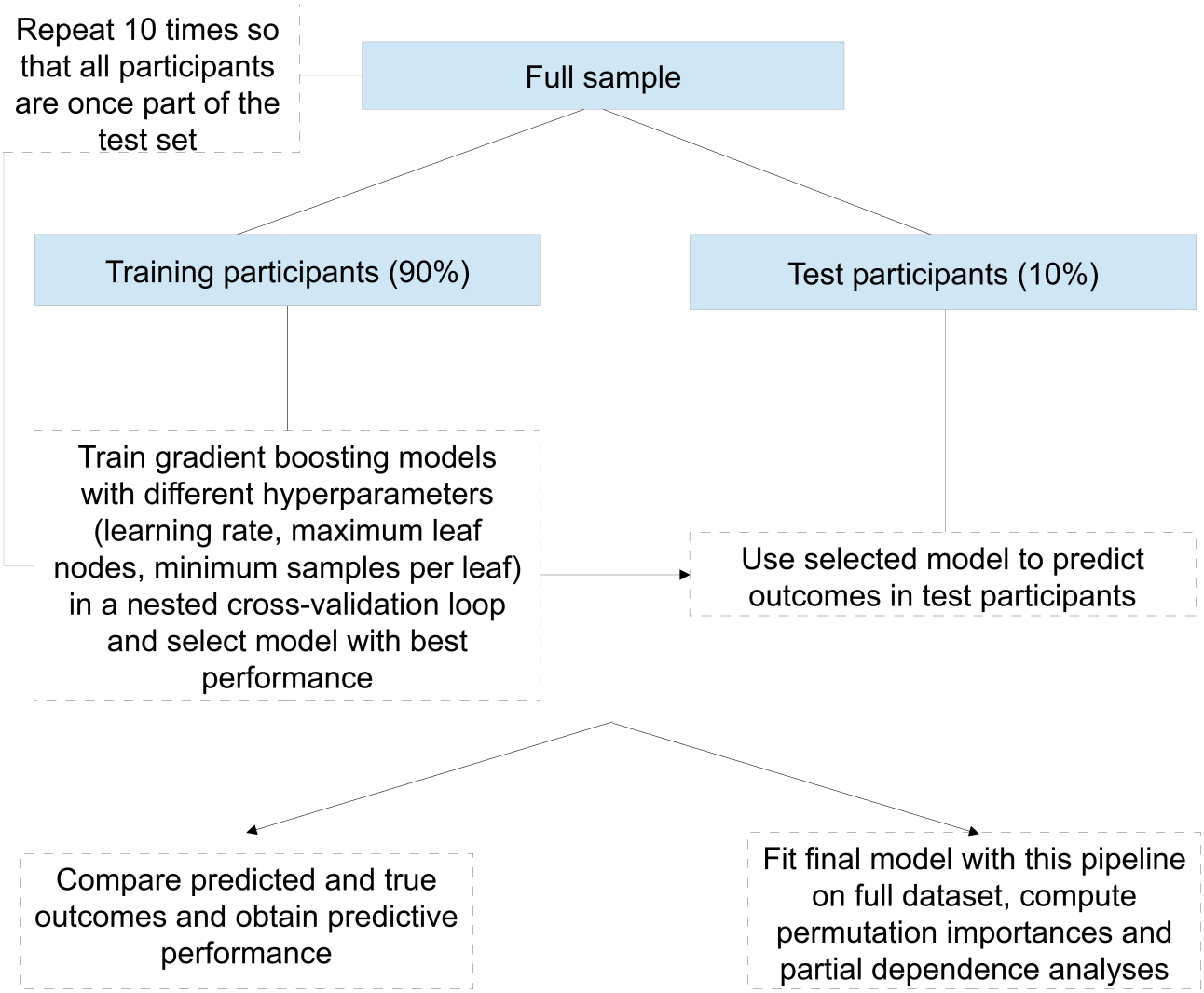
mental health might depend on gender) or nonlinear effects of single predictors (ie, very low or very high BMI values might both predict worse mental health). We used gradient boosting models (GBM [30]), which are currently considered the most accurate predictive tools for tabular datasets [31]. Specifically, we applied histogram-based GBMs as implemented in the HistGradientBoosting classes of scikit-learn for Python, which can automatically handle categorical and missing data. Hence, no imputation and encoding of categorical variables was performed as preprocessing in the nonlinear pipeline. The predictive performance was assessed with Pearson correlation coefficients and the mean squared error between predicted and true SDQ-E scores at follow-up in test set participants.

Figure 1. The linear pipeline comprised one-hot encoding of categorical variables, K-nearest neighbor imputation of missing values (as implemented in the scikit-learn KNNImputer class with default parameters), Z-standardization of numerical variables, and fitting of a linear model to predict the outcomes. To improve the predictive performance through mitigation of overfitting, we performed L2 regularization of the model's regression coefficients (Ridge regression), where the optimal regularization strength was determined in a nested 5-fold cross-validation loop in each training set.



XSL•FO
RenderX

Figure 2. The nonlinear pipeline comprised fitting of a gradient boosting regression model to predict the outcomes, where the optimum value of crucial hyperparameters was determined in a nested 5-fold cross-validation loop in each training set. Iterations (adding of new trees to the predictor) were stopped when the predictor's performance did not improve over the last 10 iterations. Due to the model's native ability to handle missing values and categorical variables, no such preprocessing (imputation and encoding) was necessary.



Predictor Importance Estimation

To obtain a measure of the importance of individual predictors in the multivariable models, we calculated permutation-based feature importances as implemented in the PermutationImportance class of the Alibi explain toolbox for Python [32]. This technique assesses the decline in a model's performance when the associations between predictor and outcome are removed by randomly interchanging the values of 1 feature. The magnitude of decline in the model's performance afforded by this permutation quantifies the extent to which the model relies on that specific predictor [33]. Here, we report the percentage change in performance when the predictor is intact compared to when it is perturbed.

Since the importance of individual predictors in multivariable models with many predictors is difficult to interpret (ie, the importance of a predictor decreases when a correlated predictor is added), we, moreover, applied a variant of this algorithm, which calculated the importance of a priori defined domains of predictors (Tables S1 and S2 in Multimedia Appendix 1) domains of predictors. The procedure here was identical to that for individual predictors (the predictors belonging to a domain are permuted together, and the associated decline in predictive performance is determined). To correct for potential chance effects, we fitted 10 GBMs with different random generators and averaged the predictor importances over these 10 models.

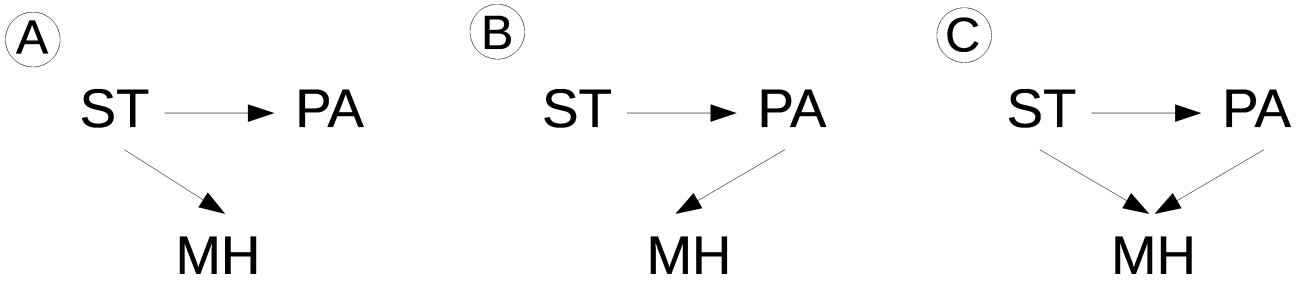
Causal Effects of Predictor Changes

We assessed the causal effect of factors whose change is discussed as the cause of the observed adolescent mental health crisis (physical inactivity, screen time, peer problems). In contrast to linear models, nonlinear regressors such as GBMs do not have simple coefficients that capture the adjusted influence of a predictor on the outcome. Hence, we artificially changed the baseline values of the relevant predictors for all participants to estimate the adjusted effects of these factors on mental health development during adolescence. All other predictors were kept unchanged, and we assessed mental health with these changed predictor values. This approach is similar to partial dependence plots [34] and can yield causal estimates when "backdoor paths" [35] between changed predictor and outcome are closed through appropriate inclusion of covariate predictors in the model [17]. However, ensuring that all

XSL•FO
RenderX

backdoor paths are closed would require complete knowledge of the real causal graph, which is hardly possible in practice. Therefore, heuristics were developed for the best possible selection of predictor covariates to isolate causal effects, which can be applied with incomplete knowledge of the causal graph [18]. In line with this approach, we included available assessments of causes of mental health and of the factors investigated for causal effects on mental health into the predictive models. It should be noted that since this approach might lead to the inclusion of variables on the causal path from predictor to outcome (ie, effect mediators), the total causal effect (ie, the direct effect plus the effect indirectly mediated by mediators) of the predictors cannot be determined in this way [17,18]. Thus, the influence of the calculated predictors calculated is not an answer to the question, "How would mental health change in late adolescence if this predictor had a different value in early adolescence and the values of all other variables affected by this predictor also changed accordingly?" but rather to the question, "How would mental health change in late adolescence if this predictor had a different value in early adolescence but all other variables (including those potentially affected by the changed predictor) remained the same?" (Figure 3). This means that if an effect is found for a predictor in such an analysis, it represents a direct (ie, not conveyed via other mediators included in the model) cause of the further development of mental health.

Figure 3. Exemplary possible causal relationships between screen time (ST), physical activity (PA), and mental health (MH). (**A**) Screen time could reduce physical activity and worsen mental health. Screen time thus generates a statistical relationship between physical activity and mental health, although a causal relationship is missing in this example. In this case, a partial dependence analysis would falsely show an improvement in mental health due to an increase in physical activity if screen time is not included in the model. The inclusion of screen time in the model closes this "backdoor path," and conditional on screen time, physical activity and mental health become independent. A partial dependence analysis in this case (ie, with inclusion of screen time in the model) would correctly show an improvement in mental health through a reduction in screen time but not through an increase in physical activity. (**B**) High screen time could reduce physical activity and (only) thereby worsen mental health. In this case, without including physical activity in the model, there would be a statistical relationship between screen time and mental health. Conditioned on physical activity (by including this predictor in the model), screen time and mental health become statistically independent. If physical activity is not included in the model, a partial dependence analysis would show an improvement in mental health afforded by a reduction in screen time. However, if physical activity is included, a partial dependence analysis would show no effect of a reduction in screen time, but only of an increase in physical activity. (**C**) High screen time could have a direct negative effect on mental health and an additional indirect (mediated) effect by reducing physical activity. Without including physical activity in the model, in this case both the direct and the indirect (mediated by physical activity) effect on mental health would be "attributed" to screen time, and a partial dependence analysis would accordingly show the total effect (direct plus indirect effect) of a reduction in screen time. Including physical activity in the model would "attribute" the effect of screen time mediated by physical activity to physical activity. In this case, a partial dependence analysis would only show the direct effect of screen time on mental health.



Specifically, we tested the effect of screen time in both cohorts by setting all questions relating to time spent with media consumption to the minimum value (lowest), the 1/3 percentile (low), the 2/3 percentile (high), and the maximum value (highest). In both cohorts, these were questions about the hours spent using TV, computers, cell phones, and social media per day (details on all items in Tables S1 and S2 in Multimedia Appendix 1). Similarly, we set the values for questions on physical inactivity in both cohorts to minimum, 1/3 percentile, 2/3 percentile, and maximum and tested the impact on predicted mental health in late adolescence. These were questions on the number of days spent last week doing moderate to vigorous physical activity and the total minutes in moderate to vigorous physical activity as objectively recorded by the accelerometer in the MC and questions about the total hours of sports per week and the number of days physically active in the last 7 days in the KiGGS cohort (Tables S1 and S2 in Multimedia Appendix 1 for details). For peer problems, we changed parent-reported peer problems (as assessed in the SDQ with items like "The child has at least one good friend" or "The child is generally liked by other children"), the assessment of happiness with friendships, the feeling that "there is no one I feel close to," and that "there is someone I trust whom I would turn to if I had problems" in the MC. Similarly, we changed the parent- and self-reported peer problems in the SDQ and the social support scale sum score value in the KiGGS cohort. Again, details on the items used to measure peer problems can be found in Tables S1 and S2 in Multimedia Appendix 1.

Similar to the feature importance estimation, we fitted 10 GBMs with different random generators and averaged the predicted mental health under all predictor changes over these 10 models to compensate for random calculations during model training.

Data and Code Availability and Adherence to Reporting Guidelines

The analysis code used to generate the study results is freely available at the GitHub page of the authors[36]. For privacy reasons, the data cannot be made public, but data access can be requested at the institutes implementing the cohort studies (MC: Center for Longitudinal Studies of the University College London; KiGGS cohort: Department of Epidemiology and Health Monitoring of the Robert Koch Institute). The reporting of the model development and validation followed the Transparent Reporting of a multivariable prediction model for

XSL•FO
RenderX

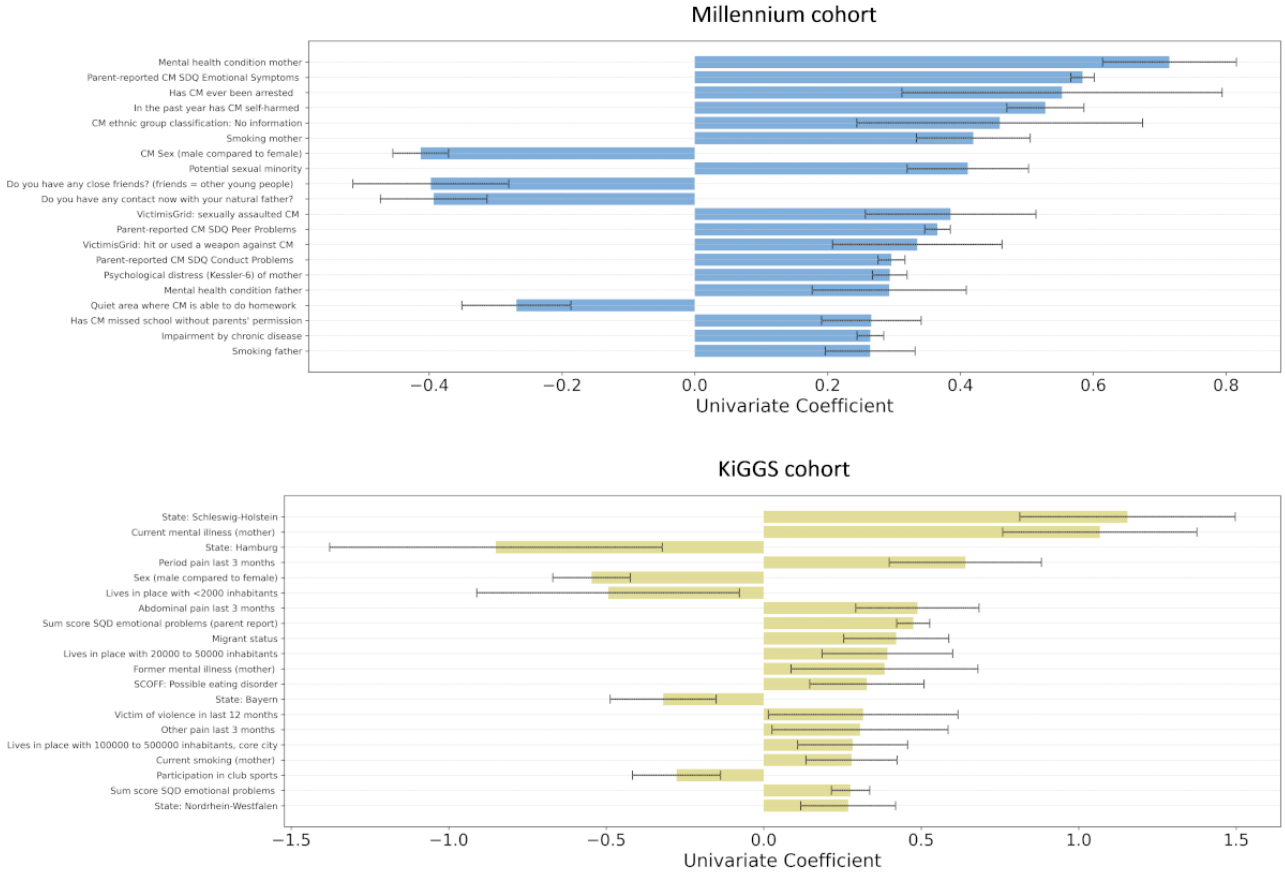
Individual Prognosis Or Diagnosis (TRIPOD [37]) guidelines (Table S4 in Multimedia Appendix 1).

Results

Univariable Relationships Between Baseline Predictors and Future Mental Health

The strongest statistically significant unadjusted linear relationships between baseline predictors and mental health at follow-up are shown in Figure 4 and coefficients of all predictors can be found in Tables S1 and S2 in Multimedia Appendix 1. Previous mental health problems, mental health problems of the parents, conduct problems, physical health problems and pain, low socioeconomic status, female sex, victimization, and potential exposure to marginalization and discrimination were related to worse future mental health in both cohorts.

Figure 4. Univariable (unadjusted) linear relationships and 95% confidence intervals between predictors assessed at baseline and mental health (Strength and Difficulties Questionnaire—Emotional [SDQ-E]) at follow-up for the Millennium Cohort (upper row) and the KiGGS cohort (lower row). For reasons of conciseness, only the 20 most important predictors with statistical significance are shown and sorted in descending order of absolute coefficients. A complete overview of all predictors, including their univariable coefficients and their assignment to predictor domains, can be found in Tables S1 and S2 in Multimedia Appendix 1. For binary predictors, a "yes"-response was generally coded as 1, meaning that a negative value indicates less future mental health problems when the question was answered with "yes." Ordinal-scaled items were ordered so that higher values indicated stronger agreement with a question.



Predictive Performance of Multivariable Models

The average cross-validated predictive performance of the multivariable linear and nonlinear models is summarized in Table 2 (Pearson correlation coefficients between predicted and true SDQ-E scores as well as mean squared error). In both cohorts, the nonlinear model performed better on average and was hence used for further analyses.

XSL•FO
RenderX

Table . Predictive performance (linear correlation between predicted and true SDQ-E[a] scores in test set participants) for different models and both cohorts (mean value and SD of 10 repeats of 10-fold cross-validation).

Model	Millennium cohort		KiGGS[b] cohort	
	r (SD over cross-validation repeats)	MSE[c] (SD over cross-validation repeats)	*r* (SD over cross-validation repeats)	MSE (SD over cross-validation repeats)
Linear	0.614 (.002)	3.203 (.005)	0.384 (.02)	2.869 (.02)
Nonlinear (gradient boosting)	0.614 (.002)	3.194 (.009)	0.466 (.03)	2.754 (.04)

[a]SDQ-E: Strength and Difficulties Questionnaire—Emotional.

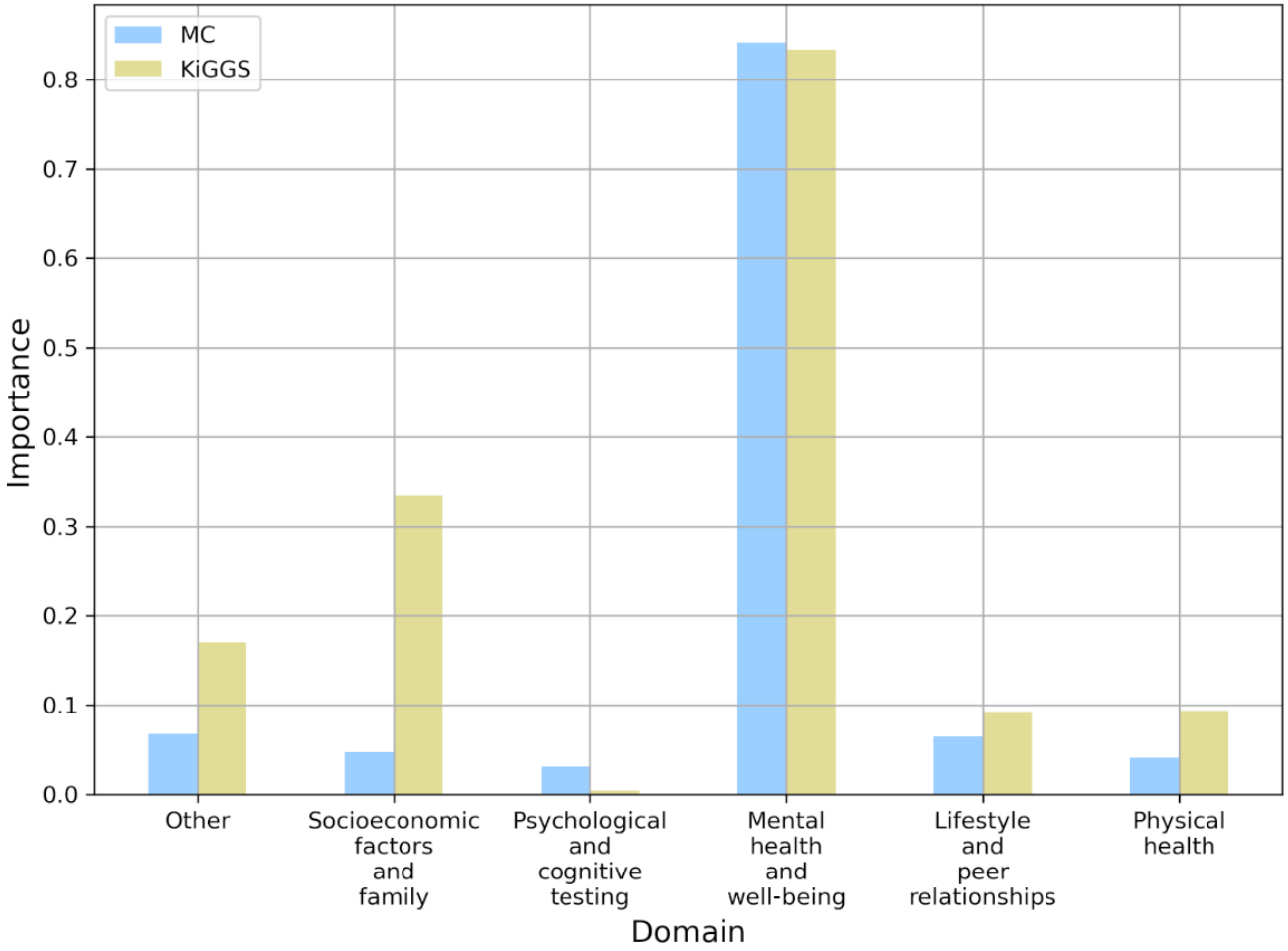
[b]KiGGS: German Health Interview and Examination Survey for Children and Adolescents.

[c]MSE: mean squared error.

Predictor Importance

The permutation-based importance of defined predictor domains is shown in Figure 5. Mental health assessments at baseline were the most important domain in both cohorts, but other predictors (participants sex, physical health such as chronic diseases, sociodemographic factors such as parental education and mental health, and lifestyle factors such as physical activity) also contributed to the predictions. The importance of each individual predictor as well as their assignments to predictor domains can be found in Tables S1 and S2 in Multimedia Appendix 1.

Figure 5. Permutation-based feature importance of predictor domains (percentage improvement by keeping the predictor domain intact compared to perturbing it).



Direct Effects of Screen Time, Physical Activity, and Peer Relationships

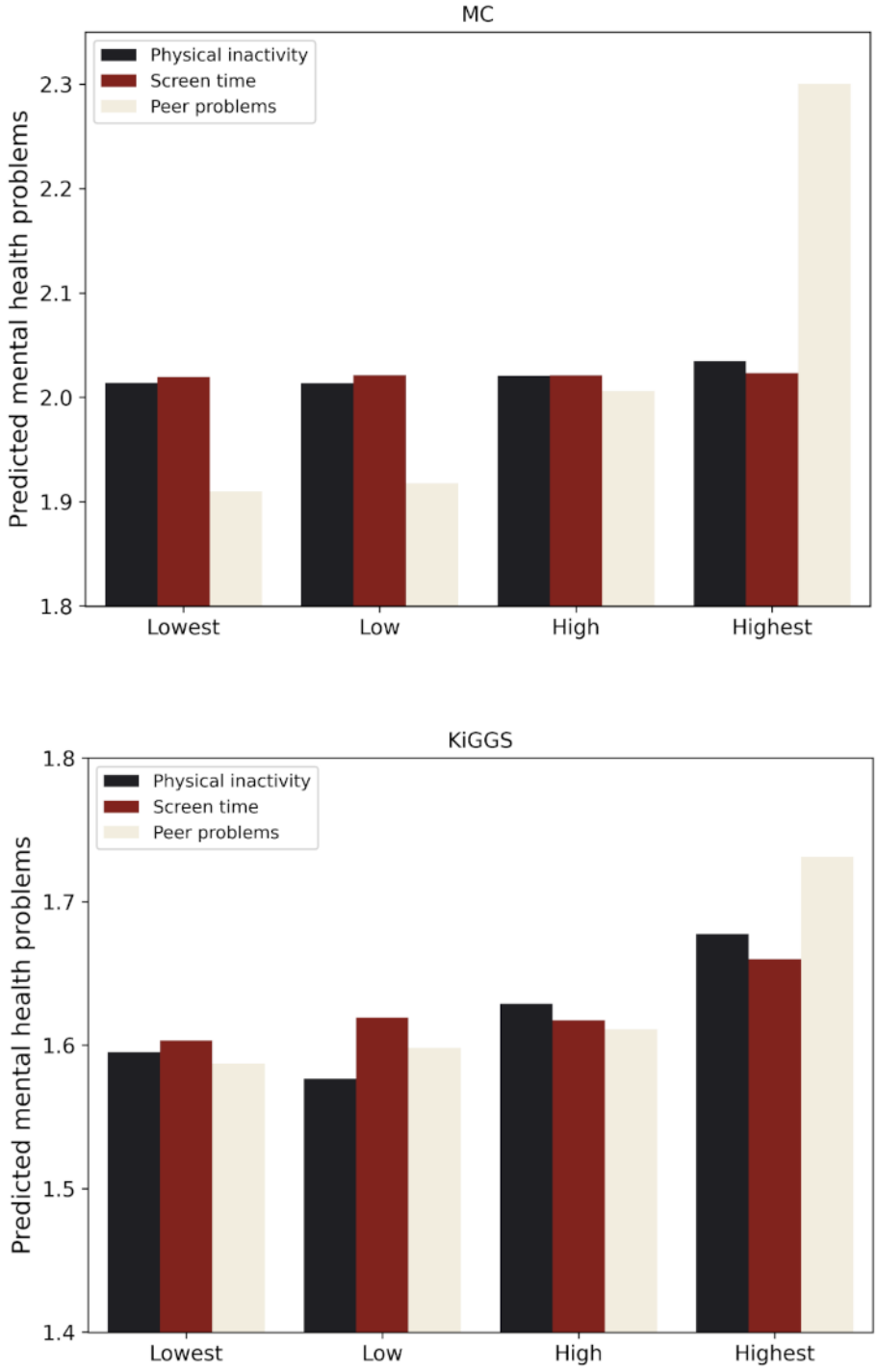
The direct causal effects of physical inactivity, screen time, and peer problems in early adolescence on mental health in late adolescence are summarized in Figure 6. It can be seen that in both cohorts, predicted mental health problems in late adolescence increased with growing peer problems. Both cohorts also showed more predicted mental health problems for high levels of physical inactivity, although this effect was small in the Millennium Cohort. Increased screen time did not change predictions in the MC, while in KiGGS, predictions with the

XSL•FO
RenderX

highest screen time levels showed slightly more mental health problems compared to lower levels. These results show that, independent of changes in other predictors, a deterioration in peer relationships in particular leads to a worse development of mental health in both cohorts.

XSL•FO
RenderX

Figure 6. Direct effect of setting the value of predictors pertaining to screen time, physical activity, and peer relationships in early adolescence to minimum (lowest), 1/3 percentile (low), 2/3 percentile (high), and maximum (highest) on predicted mental health in late adolescence (keeping all other predictors constant). Screen time variables comprised all questions relating to time spent with media consumption (TV, computer, cell phone, and social media). Physical inactivity variables comprised days last week spent doing moderate to vigorous physical activity and total minutes in moderate to vigorous physical activity as objectively recorded by the accelerometer in the MC and total hours of sports per week and number of days physically active in the last 7 days in the KiGGS cohort. Peer problems variables comprised parent-reported peer problems (in the SDQ), the assessment of happiness with friendships, the feeling that "there is no one I feel close to," and that "there is someone I trust whom I would turn to if I had problems" in the MC and parent- and self-reported peer problems and the social support scale sum score value in the KiGGS cohort. In Tables S1 and S2 in Multimedia Appendix 1, the corresponding variables are color-coded in accordance with the figure.



XSL•FO
RenderX

Discussion

In this study, we used interpretable machine learning pipelines on 2 large European cohort studies for 2 related goals: First, we tested how well and based on which assessments the development of mental health during adolescence can be assessed. Second, we assessed the causal effects that high screen time, low physical activity, and worsened peer relationships had on an unfavorable development of mental health. Because the analyzed time period lies in the critical years after 2010, we thereby added evidence to the discussion on reasons for the mental health crisis unfolding during that time period.

Using cross-validation, we showed that multivariable models had medium to high predictive accuracy with Pearson correlations between predicted and actual mental health in late adolescence between 0.38 and 0.61. Although direct comparisons are complicated by differences in the study design and outcome measures, this predictive accuracy is in the upper range of the few studies published to date on the prediction of mental health problems in adolescents using comprehensive multivariable models [38,39]. Predictive performance was improved with nonlinear GBMs, which are known to perform well for medium-sized tabular datasets [31]. However, these figures also make it clear that the calculations of the models should be interpreted as assessments and not as absolutely accurate predictions. The relative superiority of the nonlinear model compared with the linear model suggests that at least some relationships between the investigated predictors and the development of mental health are not optimally modeled in a linear and additive fashion. This is also reflected, for example, in the disproportionately strong increase in predicted mental health problems with very severe peer problems.

We used univariable and multivariable feature importance analyses to identify predictive factors. Here, we were able to replicate and extend previously described risk and resilience factors. As expected and consistently demonstrated in previous studies [40,41], mental health problems in early adolescence were the strongest predictors of mental health problems in late adolescence. This result captures a certain stability of mental health problems during adolescence, but the contribution of other factors to the predictions shows that the development of adolescent mental health depends on multiple determinants that go beyond a mere “autoregressive” continuation of baseline mental health problems. Among family factors, mental health problems of the parents were strongly related to mental health problems in their adolescent children. For preventive strategies, it is important to understand the mechanisms of this intergenerational transmission of risk and to develop targeted interventions based on this understanding [42]. Indicators of physical health problems (chronic health conditions, self-reported pain, and impairments in general health) were also associated with future mental health problems. This finding substantiates current evidence that physical health problems in early life predict a worsening of mental health in adolescence and that children with chronic conditions might hence be particularly in need of support [23]. Low socioeconomic status (ie, income and education) of the parents predicted an unfavorable development of their children’s mental health during adolescence, which corroborates previous findings with different cohorts [43,44]. In both cohorts, we moreover found that female sex predicted worse mental health in late adolescence also when statistically correcting for baseline mental health symptoms in young adolescence. This finding is in accordance with a well-established gender gap in depressive symptoms, which starts to manifest during puberty, increases until adulthood, and has multifaceted reasons [28].

We leveraged recent advances in covariate selection and causal interpretation of predictive models [17,18] to approach the causal impacts of factors implicated in the deterioration of adolescent mental health in the past decade (less physical activity, more screen time, and worse peer relationships). This analysis revealed improved predicted mental health in late adolescence when more physical activity was assumed in early adolescence. However, the direct effect was small, particularly in the MC. We found that, except small effects of very much screen time in the KiGGS cohort, reducing screen time alone did not directly entail meaningful improvements in the predicted development of mental health in both cohorts. Conversely, good peer relationships predicted better mental health trajectories in both cohorts even when all other factors were held constant.

What do these results suggest in the context of current debates on the causes of the youth mental health crisis?

First, according to these analyses, adolescents would have had a more favorable mental health trajectory during the period of the observed deterioration in adolescent mental health if they had better peer relationships at ages 12-14 years, even if a wide range of other determinants of mental health had remained unchanged. It is noteworthy in this context that an overall deterioration in peer relationships of adolescents was observed across countries during the same period as the deterioration of mental health [45] and that proposed influencing factors such as increased screen time were shown to potentially have a detrimental impact on peer relationships [46]. It thus seems plausible that increased peer problems, themselves the result of a complex change in behavioral patterns of adolescents, represent an important direct cause of the deterioration of mental health. Given the growing evidence for the effectiveness of general prevention programs (such as antibullying school interventions) and individual therapeutic strategies (such as social skills training or self-esteem reflections) to improve peer relationships [47], it appears that a combination of these interventions would be a promising public health approach to the adolescent mental health crisis.

Second, adolescents would not have altered mental health trajectories if they had less screen time at ages 12-14 years when the other determinants of mental health had remained unchanged. This finding ties well with a recent panel network analysis showing that when multiple determinants of mental health were taken into account, social media use only had a very small influence on the development of mental health [48]. There is an intense debate around relationships between screen time, particularly social media use, and mental health in adolescents and young adults. The debate is fueled by a remarkable coincidence between the mass spread of social media and the deterioration of adolescent mental health around 2010 [49].

XSL•FO
RenderX

However, longitudinal studies could not prove a clear association between social media use and adolescent mental health [50,51]. It seems plausible that heavy social media use overall is part of a complex change in adolescent behaviors that might, potentially mediated through increased peer problems, contribute to mental health problems [11]. As illustrated in Figure 3, the absence of a direct effect in our study would be compatible with a model in which screen time exerts a negative effect on mental health via other downstream factors (eg, factors impacted by screen time and themselves impacting mental health, such as time spent with friends, body dissatisfaction, and sleep problems). Such modifiable factors on the causal pathway between exposures such as screen time and mental health outcomes could be targets for intervention strategies. Characterizing them more precisely would be an important undertaking for future research, which could use causal inference in combination with higher frequency surveys [18].

Third, adolescents would have a slightly improved mental health trajectory if they had been more physically active at ages 12-14 years, even when the other determinants of mental health had remained unchanged, especially those with very low levels of physical activity. In line with this finding, exercise interventions were shown to have small beneficial effects on adolescent mental health [52], and a recent review of reviews concluded that there is only "partial evidence" for a causal impact of physical activity on adolescent mental health [53]. Interestingly, the authors of this review discuss the possibility that the association between physical activity and adolescent mental health "may be linear, curvilinear, or contain a threshold, after which no further gains in mental health are made" [53], which would be consistent with our findings that increases in physical activity only have a beneficial direct effect in the case of low baseline levels. However, the small effect strengths, together with meta-studies that used aggregate data to demonstrate that the level of physical activity in adolescents has not relevantly decreased in Western countries between 2001 and 2016 [54], render it unlikely that the observed substantial worsening of adolescent mental in the past decade is due to decreased physical activity.

Some limitations of our study are relevant when interpreting the study results. First, causal inference from observational data is challenging, and results can depend on assumptions that are hard to test empirically. If the research questions cannot be investigated in a randomized controlled manner and one has to resort to observational data (as in our case), the best possible selection of covariates must be undertaken to approach causal interpretations. We applied principles developed for this purpose in our study [18]. Nevertheless, it should be kept in mind that the evidence for causal relationships in observational studies is generally lower compared to randomized controlled trials [55]. The results should therefore be interpreted in conjunction with other research work (eg, cohort studies, clinical observations, and analysis of aggregated time series) in order to gain a comprehensive picture of the factors contributing to the deterioration of young people's mental health. Moreover, the inability to reliably distinguish confounders from mediators only allows for the calculation of direct effects, and it is possible that the factors under investigation have additional effects through the modification of mediator variables. To characterize such mediating relationships more precisely, future research could explore multiple-wave longitudinal studies in conjunction with causal inference [18]. Second, although the mutual validation of results in 2 independent cohorts is a strength, both cohorts differed in some design features, which reduces the comparability. In particular, the MC is age-homogenous, whilst in the age-heterogeneous KiGGS cohort, participants with an age similar to the MC participants were selected, leading to a smaller case number. Due to a longer interval between baseline and follow-up examination, participants in the KiGGS cohort were younger at baseline. This could contribute to observed discrepancies, for example, to the result that the direct effect of physical activity was more pronounced in the KiGGS cohort. In addition, predictors were only partially overlapping between the 2 studies. Despite these differences, it is noteworthy that key predictors of future mental health, such as the parents' mental health, the participants' baseline mental health, sex, peer problems, and physical activity, were found in both cohorts and there was a qualitative agreement on the estimated direct causal effects. The robustness of these results would be further strengthened by an analysis of additional cohorts from other countries. Third, not all proposed driving factors of the observed adolescent mental health deterioration could be investigated with our study. For instance, factors affecting society as a whole, such as growing social inequality and pressure to perform in education, have also been suggested [56,57] but could not be analyzed with our research design, which leveraged interindividual differences in risk factors to assess their impact on mental health development. Relatedly, the influence of new stressors such as the COVID pandemic, the war in Ukraine, and the increasingly noticeable effects of climate change could not be examined due to the time period under study. However, once appropriate data are available, our ML pipeline could be readily adapted in future work to predict the risk for mental health problems in the face of societal crises and identify its determining factors.

In summary, we showed that the development of adolescent mental health can be assessed by a combination of variables from early adolescence. Our analyses suggest a strong influence of peer relationships, and a deterioration of these through different reasons may contribute as a causal factor to the observed adolescent mental health crisis. Interventions to improve social relationships with peers could be explored as countermeasures against further deterioration of adolescent mental health.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Additional content.
[DOCX File, 98 KB - publichealth_v11i1e60125_app1.docx]

References

1. Sawyer SM, Afifi RA, Bearinger LH, et al. Adolescence: a foundation for future health. Lancet 2012 Apr 28;379(9826):1630-1640. [doi: 10.1016/S0140-6736(12)60072-5] [Medline: 22538178]
2. Gunnell D, Kidger J, Elvidge H. Adolescent mental health in crisis. Br Med J 2018 Jun 19;361:k2608. [doi: 10.1136/bmj.k2608] [Medline: 29921659]
3. Youth mental health crisis management. Nat Mental Health 2023 Aug 7;1(8):525-526. [doi: 10.1038/s44220-023-00112-1]
4. Armitage JM, Kwong ASF, Tseliou F, et al. Cross-cohort change in parent-reported emotional problem trajectories across childhood and adolescence in the UK. Lancet Psychiatry 2023 Jul;10(7):509-517. [doi: 10.1016/S2215-0366(23)00175-X] [Medline: 37244272]
5. Ten Have M, Tuithof M, van Dorsselaer S, Schouten F, Luik AI, de Graaf R. Prevalence and trends of common mental disorders from 2007-2009 to 2019-2022: results from the Netherlands Mental Health Survey and Incidence Studies (NEMESIS), including comparison of prevalence rates before vs. during the COVID-19 pandemic. World Psychiatry 2023 Jun;22(2):275-285. [doi: 10.1002/wps.21087] [Medline: 37159351]
6. Keyes KM, Gary D, O'Malley PM, Hamilton A, Schulenberg J. Recent increases in depressive symptoms among US adolescents: trends from 1991 to 2018. Soc Psychiatry Psychiatr Epidemiol 2019 Aug;54(8):987-996. [doi: 10.1007/s00127-019-01697-8] [Medline: 30929042]
7. Ruch DA, Sheftall AH, Schlagbaum P, Rausch J, Campo JV, Bridge JA. Trends in suicide among youth aged 10 to 19 years in the United States, 1975 to 2016. JAMA Netw Open 2019 May 3;2(5):e193886. [doi: 10.1001/jamanetworkopen.2019.3886] [Medline: 31099867]
8. Baumgarten F, Junker S, Schlack R. Prevalence and time trends of self-reported mental health problems among children and adolescents between 11 and 17 years in the KiGGS study. Z Kinder Jugendpsychiatr Psychother 2023 Jul;51(4):311-320. [doi: 10.1024/1422-4917/a000936] [Medline: 37417965]
9. Shorey S, Ng ED, Wong CHJ. Global prevalence of depression and elevated depressive symptoms among adolescents: a systematic review and meta-analysis. Br J Clin Psychol 2022 Jun;61(2):287-305. [doi: 10.1111/bjc.12333] [Medline: 34569066]
10. Gagné T, Henderson C, McMunn A. Is the self-reporting of mental health problems sensitive to public stigma towards mental illness? A comparison of time trends across English regions (2009-19). Soc Psychiatry Psychiatr Epidemiol 2023 Apr;58(4):671-680. [doi: 10.1007/s00127-022-02388-7] [Medline: 36473961]
11. Abi-Jaoude E, Naylor KT, Pignatiello A. Smartphones, social media use and youth mental health. CMAJ 2020 Feb 10;192(6):E136-E141. [doi: 10.1503/cmaj.190434] [Medline: 32041697]
12. Twenge JM, Haidt J, Blake AB, McAllister C, Lemon H, Le Roy A. Worldwide increases in adolescent loneliness. J Adolesc 2021 Dec;93:257-269. [doi: 10.1016/j.adolescence.2021.06.006] [Medline: 34294429]
13. Kandola A, Lewis G, Osborn DPJ, Stubbs B, Hayes JF. Depressive symptoms and objectively measured physical activity and sedentary behaviour throughout adolescence: a prospective cohort study. Lancet Psychiatry 2020 Mar;7(3):262-271. [doi: 10.1016/S2215-0366(20)30034-1] [Medline: 32059797]
14. Joshi H, Fitzsimons E. The Millennium Cohort Study: the making of a multi-purpose resource for social science and policy. Longit Life Course Stud 2016 Oct 26;7(4). [doi: 10.14301/llcs.v7i4.410]
15. Mauz E, Lange M, Houben R, et al. Cohort profile: KiGGS cohort longitudinal study on the health of children, adolescents and young adults in Germany. Int J Epidemiol 2020 Apr 1;49(2):375-375k. [doi: 10.1093/ije/dyz231] [Medline: 31794018]
16. Prosperi M, Guo Y, Sperrin M, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare. Nat Mach Intell 2020 Jul 13;2(7):369-375. [doi: 10.1038/s42256-020-0197-y]
17. Zhao Q, Hastie T. Causal interpretations of black-box models. J Bus Econ Stat 2019;2019. [doi: 10.1080/07350015.2019.1624293] [Medline: 33132490]
18. VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol 2019 Mar;34(3):211-219. [doi: 10.1007/s10654-019-00494-6] [Medline: 30840181]
19. Connelly R, Platt L. Cohort profile: UK Millennium Cohort Study (MCS). Int J Epidemiol 2014 Dec;43(6):1719-1725. [doi: 10.1093/ije/dyu001] [Medline: 24550246]
20. Vostanis P. Strengths and Difficulties Questionnaire: research and clinical applications. Curr Opin Psychiatry 2006 Jul;19(4):367-372. [doi: 10.1097/01.yco.0000228755.72366.05] [Medline: 16721165]
21. Armitage JM, Tseliou F, Riglin L, et al. Validation of the Strengths and Difficulties Questionnaire (SDQ) emotional subscale in assessing depression and anxiety across development. PLoS ONE 2023;18(7):e0288882. [doi: 10.1371/journal.pone.0288882] [Medline: 37467238]
22. Patalay P, Fitzsimons E. Development and predictors of mental ill-health and wellbeing from childhood to adolescence. Soc Psychiatry Psychiatr Epidemiol 2018 Dec;53(12):1311-1323. [doi: 10.1007/s00127-018-1604-0] [Medline: 30259056]

XSL•FO
RenderX

23. Määttä H, Honkanen M, Hurtig T, Taanila A, Ebeling H, Koivumaa-Honkanen H. Childhood chronic condition and subsequent self-reported internalizing and externalizing problems in adolescence: a birth cohort study. Eur J Pediatr 2022 Sep;181(9):3377-3387. [doi: 10.1007/s00431-022-04505-9] [Medline: 35796794]
24. Ridley M, Rao G, Schilbach F, Patel V. Poverty, depression, and anxiety: Causal evidence and mechanisms. Science 2020 Dec 11;370(6522):eaay0214. [doi: 10.1126/science.aay0214] [Medline: 33303583]
25. Reupert A, Maybery D. What do we know about families where parents have A mental illness? A systematic review. Child Youth Serv 2016 Apr 2;37(2):98-111. [doi: 10.1080/0145935X.2016.1104037]
26. van Harmelen AL, Kievit RA, Ioannidis K, et al. Adolescent friendships predict later resilient functioning across psychosocial domains in a healthy community cohort. Psychol Med 2017 Oct;47(13):2312-2322. [doi: 10.1017/S0033291717000836] [Medline: 28397612]
27. Brownlie E, Beitchman JH, Chaim G, Wolfe DA, Rush B, Henderson J. Early adolescent substance use and mental health problems and service utilisation in a school-based sample. Can J Psychiatry 2019 Feb;64(2):116-125. [doi: 10.1177/0706743718784935] [Medline: 29929386]
28. Kuehner C. Why is depression more common among women than among men? Lancet Psychiatry 2017 Feb;4(2):146-158. [doi: 10.1016/S2215-0366(16)30263-2] [Medline: 27856392]
29. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825-2830.
30. Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal 2002 Feb;38(4):367-378. [doi: 10.1016/S0167-9473(01)00065-2]
31. Shwartz-Ziv R, Armon A. Tabular data: deep learning is not all you need. Inf Fusion 2022 May;81:84-90. [doi: 10.1016/j.inffus.2021.11.011]
32. Klaise J, Looveren A, Vacanti G, Coca A. Alibi explain: algorithms for explaining machine learning models. J Mach Learn Res 2021;22(1):8194-8200.
33. Breiman L. Random forests. Mach Learn 2001;45(1):5-32. [doi: 10.1023/A:1010933404324]
34. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Statist 2001 Oct;29(5):1189-1232. [doi: 10.1214/aos/1013203451]
35. Pearl J. [Bayesian Analysis in Expert Systems]: Comment: graphical models, causality and intervention. Statist Sci 1993;8(3):266-269. [doi: 10.1214/ss/1177010894]
36. Climate and Societal Analytics from the Centre for Artificial Intelligence in Public Health Research, Robert Koch Institute. GitHub. URL: https://github.com/ClimSocAna [accessed 2025-05-06]
37. Collins GS, Reitsma JB, Altman DG, Moons KGM, Group T. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. The TRIPOD Group. Circulation 2015 Jan 13;131(2):211-219. [doi: 10.1161/CIRCULATIONAHA.114.014508] [Medline: 25561516]
38. Tate AE, McCabe RC, Larsson H, Lundström S, Lichtenstein P, Kuja-Halkola R. Predicting mental health problems in adolescence using machine learning techniques. PLoS One 2020;15(4):e0230389. [doi: 10.1371/journal.pone.0230389] [Medline: 32251439]
39. Koning NR, Büchner FL, Leeuwenburgh NA, et al. Identification of child mental health problems by combining electronic health record information from different primary healthcare professionals: a population-based cohort study. BMJ Open 2022 Jan 12;12(1):e049151. [doi: 10.1136/bmjopen-2021-049151] [Medline: 35022168]
40. Johnson D, Dupuis G, Piche J, Clayborne Z, Colman I. Adult mental health outcomes of adolescent depression: a systematic review. Depress Anxiety 2018 Aug;35(8):700-716. [doi: 10.1002/da.22777] [Medline: 29878410]
41. Jones PB. Adult mental health disorders and their age at onset. Br J Psychiatry Suppl 2013 Jan;54(s54):s5-10. [doi: 10.1192/bjp.bp.112.119164] [Medline: 23288502]
42. Leijdesdorff S, van Doesum K, Popma A, Klaassen R, van Amelsvoort T. Prevalence of psychopathology in children of parents with mental illness and/or addiction: an up to date narrative review. Curr Opin Psychiatry 2017 Jul;30(4):312-317. [doi: 10.1097/YCO.0000000000000341] [Medline: 28441171]
43. Joinson C, Kounali D, Lewis G. Family socioeconomic position in early life and onset of depressive symptoms and depression: a prospective cohort study. Soc Psychiatry Psychiatr Epidemiol 2017 Jan;52(1):95-103. [doi: 10.1007/s00127-016-1308-2] [Medline: 27837235]
44. Hakulinen C, Mok PLH, Horsdal HT, et al. Parental income as a marker for socioeconomic position during childhood and later risk of developing a secondary care-diagnosed mental disorder examined across the full diagnostic spectrum: a national cohort study. BMC Med 2020 Nov 16;18(1):323. [doi: 10.1186/s12916-020-01794-5] [Medline: 33190641]
45. Buecker S, Mund M, Chwastek S, Sostmann M, Luhmann M. Is loneliness in emerging adults increasing over time? A preregistered cross-temporal meta-analysis and systematic review. Psychol Bull 2021 Aug;147(8):787-805. [doi: 10.1037/bul0000332] [Medline: 34898234]
46. Lawrence D, Hunter SC, Cunneen R, et al. Reciprocal relationships between trajectories of loneliness and screen media use during adolescence. J Child Fam Stud 2022 May;31(5):1306-1317. [doi: 10.1007/s10826-021-02066-3]

XSL•FO
RenderX

47. Pollak I, Mitic M, Birchwood J, et al. A systematic review of intervention programs promoting peer relationships among children and adolescents: methods and targets used in effective programs. Adolescent Res Rev 2023 Sep;8(3):297-321. [doi: 10.1007/s40894-022-00195-4]
48. Panayiotou M, Black L, Carmichael-Murphy P, Qualter P, Humphrey N. Time spent on social media among the least influential factors in adolescent mental health: preliminary results from a panel network analysis. Nat Mental Health 2023 May 8;1(5):316-326. [doi: 10.1038/s44220-023-00063-7]
49. Twenge JM, Joiner TE, Rogers ML, Martin GN. Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time. Clin Psychol Sci 2018 Jan;6(1):3-17. [doi: 10.1177/2167702617723376]
50. Course-Choi J, Hammond L. Social media use and adolescent well-being: a narrative review of longitudinal studies. Cyberpsychol Behav Soc Netw 2021 Apr;24(4):223-236. [doi: 10.1089/cyber.2020.0020] [Medline: 33103917]
51. Odgers CL, Jensen MR. Annual research review: adolescent mental health in the digital age: facts, fears, and future directions. J Child Psychol Psychiatry 2020 Mar;61(3):336-348. [doi: 10.1111/jcpp.13190] [Medline: 31951670]
52. Hale GE, Colquhoun L, Lancastle D, Lewis N, Tyson PJ. Review: physical activity interventions for the mental health and well-being of adolescents - a systematic review. Child Adolesc Ment Health 2021 Nov;26(4):357-368. [doi: 10.1111/camh.12485] [Medline: 34105239]
53. Biddle SJH, Ciaccioni S, Thomas G, Vergeer I. Physical activity and mental health in children and adolescents: an updated review of reviews and an analysis of causality. Psychol Sport Exerc 2019 May;42:146-155. [doi: 10.1016/j.psychsport.2018.08.011]
54. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. Lancet Child Adolesc Health 2020 Jan;4(1):23-35. [doi: 10.1016/S2352-4642(19)30323-2] [Medline: 31761562]
55. O'Neil M, Berkman N, Hartling L, et al. Observational evidence and strength of evidence domains: case examples. Syst Rev 2014 Apr 23;3:35. [doi: 10.1186/2046-4053-3-35] [Medline: 24758494]
56. Högberg B, Strandh M, Hagquist C. Gender and secular trends in adolescent mental health over 24 years – the role of school-related stress. Soc Sci Med 2020 Apr;250:112890. [doi: 10.1016/j.socscimed.2020.112890]
57. Collishaw S, Sellers R. Trends in child and adolescent mental health prevalence, outcomes, and inequalities. In: Taylor E, Verhulst F, Wong JCM, Yoshida K, editors. Mental Health and Illness of Children and Adolescents: Springer Singapore; 2020:63-73. [doi: 10.1007/978-981-10-2348-4_9]

Abbreviations

GBM: gradient boosting model
KiGGS: German Health Interview and Examination Survey for Children and Adolescents
MC: British Millenium cohort
SDQ: Strength and Difficulties Questionnaire
SDQ-E: Strength and Difficulties Questionnaire—Emotional
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis

Edited by Z Wu; submitted 02.05.24; peer-reviewed by F Yang, Y Huang; revised version received 18.03.25; accepted 22.03.25; published 12.05.25.

Please cite as:
Stuke H, Schlack R, Erhart M, Kaman A, Ravens-Sieberer U, Irrgang C
Peer Relationships Are a Direct Cause of the Adolescent Mental Health Crisis: Interpretable Machine Learning Analysis of 2 Large Cohort Studies
JMIR Public Health Surveill 2025;11:e60125
URL: https://publichealth.jmir.org/2025/1/e60125
doi:10.2196/60125

© Heiner Stuke, Robert Schlack, Michael Erhart, Anne Kaman, Ulrike Ravens-Sieberer, Christopher Irrgang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Validation and Refinement of Scores to Predict Stroke Risk: Prospective Cohort Study

Hua Meng[1,2*], PhD; Zhuo Liu[3*], MA; Dongfeng Pan[4], BM; Xinya Su[3], PhD; Wenwen Lu[5], MA; Xingtian Wang[6], MA; Yuhui Geng[3], MA; Xiaojuan Ma[3], MA; Peifeng Liang[7], PhD

[1]Hubei Provincial Clinical Research Center for Alzheimer's Disease, Tianyou Hospital, School of Medicine, Wuhan University of Science and Technology, Wuhan, China

[2]Brain Science and Advanced Technology Institute, Wuhan University of Science and Technology, Wuhan, China

[3]School of Public Health, Ningxia Medical University, Yinchuan, China

[4]Department of Emergency Medicine, People's Hospital of Ningxia Hui Autonomous Region, Ningxia Medical University, Yinchuan, China

[5]Futian Center for Chronic Disease Control, Shenzhen, China

[6]Medical Record Statistics Department, General hospital of Ningxia Medical University, Yinchuan, China

[7]Department of Medical Affair, People's Hospital of Ningxia Hui Autonomous Region, Ningxia Medical Univeristy, 301 Zhengyuan North Street, Yinchuan, China

[*]these authors contributed equally

Corresponding Author:
Peifeng Liang, PhD
Department of Medical Affair, People's Hospital of Ningxia Hui Autonomous Region, Ningxia Medical Univeristy, 301 Zhengyuan North Street, Yinchuan, China

Abstract

Background: In China, the "8+2" stroke risk score has been widely used to identify individuals at high risk of stroke, despite insufficient evidence confirming its predictive ability for stroke events.

Objective: We aimed to validate the risk score's ability to predict the risk of stroke within a 10-year timeframe in community cohort populations and to optimize the scoring method to improve its predictive accuracy.

Methods: By reviewing previous literature to obtain the parameters for constructing the logistic regression model and the Rothman-Keller model, the risk threshold points of the models were determined using a sample of 100,000 participants. For this population-based cohort study, 22,259 community residents were recruited in 2013 from one urban and rural monitoring site in Ningxia, China. The occurrence of stroke was established by a combination of self-reporting and review of hospitalization electronic records (the *International Statistical Classification of Diseases and Related Health Problems 10th Revision*: I60-63). A logistic regression model and a Rothman-Keller model were used to refine the 8-factor stroke risk score to predict the 10-year stroke risk. The performance of the model was assessed by the area under the receiver operating characteristic curve and net reclassification improvement.

Results: The threshold points for low and medium risk in the logistic regression model and the Rothman-Keller model are risk scores of 0.062 and 0.002, respectively. The threshold points for medium and high risk are risk scores of 0.165 and 0.005, respectively. A total of 11,692 community residents aged 40 years or older who met the inclusion criteria completed the 10-year follow-up. According to the "8+2" stroke risk score, the stroke incidence in the low-risk (n=8908), medium-risk (n=1074), and high-risk groups (n=1710) was 4.5%, 14.7%, and 12.3%, respectively. The logistic regression model and the Rothman-Keller model demonstrated significant differences in area under the receiver operating characteristic curve values when compared to the "8+2" stroke risk score (Z=2.60, *P*=.001; Z=3.47, *P*=.009, respectively). However, no significant difference was observed between the logistic regression model and the Rothman-Keller model (Z=0.688, *P*=.49). Relative to the risk score, the absolute net reclassification improvement of the Rothman-Keller model was 0.051 (*P*=.01) and of the logistic regression model was 0.010 (*P*=.62).

Conclusions: Our study confirmed that the "8+2" stroke risk score does not effectively predict stroke events. But the Rothman-Keller model may enhance the ability to identify individuals at high risk for stroke. Future research should incorporate more specific biomarkers and multimodal imaging features to develop more accurate risk prediction models.

(JMIR Public Health Surveill 2025;11:e72497) doi:10.2196/72497

KEYWORDS

stroke; Rothman-Keller model; logistics model; risk prediction; "8+2" stroke risk score

XSL•FO
RenderX

Introduction

The Global Burden of Disease (GBD) 2021 study found that the number of new cases of stroke increased by 70.2% from 1990 to 2021 [1], highlighting a relatively insufficient emphasis on prevention, particularly in low-income countries. The number of patients with stroke in China is currently the highest in the world. Stroke is the leading cause of death and disability among adults in the country [2]. China has the highest risk of stroke globally, with an overall lifetime risk of 39.9% [3]. Consequently, China has implemented a series of measures to address the increasing burden of stroke, advocating for an integrated approach to stroke prevention, treatment, management, and rehabilitation [4].

The National Ministry of Health of China launched a significant national project called the "China National Stroke Screening Survey (CNSSS)" in 2009 to tackle the challenge posed by stroke. The China Stroke Prevention Project Committee (CSPPC) was established in April 2011. Volunteers aged 40 years and older were recruited through structured face-to-face questionnaires, and the "8+2" risk scorecard is used to screen participants and to identify high-risk groups [5]. The "8" refers to 8 risk factors (hypertension, heart disease, smoking, dyslipidemia, diabetes, physical inactivity, overweight, and family history of stroke [FHS]), and the "2" refers to transient ischemic attacks (TIAs) and previous strokes. According to the judging criteria, respondents are categorized into low-risk, medium-risk, and high-risk groups. In 2020, a total of 268,000 individuals in the high-risk group for stroke were identified across more than 240 project areas across the country [6].

Risk assessment is an effective tool for identifying prevention priorities [7]. The Framingham Stroke Risk Profile is recognized as one of the earliest and most widely used simple stroke risk assessment tools. However, validation studies in domestic populations have found that it tends to overestimate the actual stroke incidence to some extent [8]. Since its publication, the pooled cohort risk assessment equations have also been controversial, as some external validation studies suggest that this risk assessment model may overestimate the risk of atherosclerotic cardiovascular disease [9]. A study compared the performance of the Framingham cardiovascular risk equation, the pooled cohort equations, and the China-Population Attributable Risk equations in predicting the 5-year risk of atherosclerotic cardiovascular disease, including ischemic stroke. In the Uyghur and Kazakh populations, all 3 risk assessment equations consistently underestimated the risk [10]. Furthermore, although the "8+2" risk score tool has been widely used, its predictive ability remains unclear.

Therefore, we aimed to validate the "8+2" stroke risk score for predicting the 10-year risk of stroke in community cohort populations, and to optimize the scoring method to improve predictive accuracy.

Methods

Data Source

This study was a cohort study. This study followed the Transparent reporting of multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines.

The cohort was part of the China Stroke High-risk Population Screening and Intervention Program (CSHPSIP), an ongoing nationwide population-based program [5,11]. The study participants were recruited from the screening site in the Jinfeng District of Yinchuan City, Ningxia Hui Autonomous Region. A total of 22,259 community residents were enrolled in 2013 from one urban and rural monitoring site, and the outcome ascertainment was completed in 2023. The inclusion criteria for the screening participants are individuals who are aged 40 or older, permanent residents (those who have lived in the area for 6 months or more), and those who voluntarily participate by signing an informed consent form [12]. Patients younger than 40 years, those with a history of stroke or TIA, or individuals recruited after 2014 were excluded from the study cohort. The data quality control process is detailed in the Multimedia Appendix 1.

Risk Factors Measurement

Based on the "Stroke Screening and Prevention Technical Specifications" promulgated by the National Health and Family Planning Commission's Stroke Screening and Prevention Engineering Committee, the following risk factors were assessed: hypertension, heart disease, smoking, dyslipidemia, diabetes, physical inactivity, overweight, and FHS. The detailed criteria for each risk factor are shown in Table S1 in Multimedia Appendix 1 which is also available on the China Stroke and Cardiovascular Disease website.

The criteria for classifying individuals into high-, medium-, and low-risk stroke groups were as follows: the high-risk group was defined as having 3 or more risk factors; the medium-risk group was characterized by 3 or less risk factors along with a history of chronic diseases (such as hypertension, diabetes, and heart disease); and the low-risk group was defined as having 3 or less risk factors without any history of chronic diseases (Table S1 in Multimedia Appendix 1) [13]. For high-risk individuals, follow-up visits are conducted by primary health care institutions at 6 months and 12 months after the initial assessment. For moderate-risk populations, primary health care institutions conduct a single follow-up visit at 12 months to evaluate and address their associated risk factors.

Outcome

We recorded stroke as an endpoint event by searching electronic hospitalization records in Ningxia in June 2023. Stroke was identified using the diagnostic code I60-63 from the *International Statistical Classification of Diseases and Related Health Problems 10th Revision* (*ICD-10*).

Logistic Regression Model

Logistic regression was used to determine the odds ratios (ORs) of every risk factor for incident stroke [14]. The basic equation for regression with multiple independent variables is:

XSL•FO
RenderX

$$Y=\ln(P1-P)=\alpha+\beta1X1++\beta2X2+ +\beta iXi$$

Y is the estimated continuous outcome; α is the intercept. This is considered a constant value; β is the beta coefficients; and X_i is each risk factor.

Rothman-Keller Model

The Rothman-Keller model, initially developed by Kenneth J. Rothman and David B. Keller in the early 1970s, was designed to assess the combined impact of tobacco and alcohol consumption on the risk of oral and pharyngeal cancers [15]. This model provides a statistical framework that enables researchers to quantify both the independent and joint contributions of various risk factors to disease risk. It has been adapted and applied to a wide range of health conditions and diseases, including early-onset colorectal cancer and mild cognitive impairment [16,17]. Its flexibility in considering both additive and multiplicative effects of risk factors makes it a valuable tool for public health research and individual risk prediction.

The Rothman-Keller model uses the binomial distribution function method for risk classification. It calculates the benchmark proportion of incidence and risk scores based on the population exposure rate and OR of each risk factor. In addition, it estimates an individual's relative risk of developing a disease by calculating their combined risk scores. The parameters of the Rothman-Keller model are calculated as follows:

1. Baseline morbidity ratio ( ): $\rho=1\sum i=1nORi\times Pi=1-PAR\%$
   P_i: the exposure rate of individuals exposed to a risk factor in the whole population; OR_i: the odds ratios of exposure to a risk factor; and *PAR*%: population attributable risk percentage.
2. Risk score ( ): $S=\rho\times ORi$
3. Total risk score ( ):
   Pi: risk factor scores for S≥1, q_i: risk factor scores for S<1.
4. Individual risk prediction score: Individual risk of stroke=the incidence of stroke × . This expected risk of stroke is a relative value because it is measured against the overall incidence rate in the population. It can help us understand whether the individual's likelihood of developing a stroke is higher or lower than the average level of the population.

The population exposure rate of every risk factor was derived from the literature [18]. The OR of exposure to a risk factor was sourced from the logistic regression model. Data for 100,000 participants were randomly generated by the binomial distribution functions of risk factors collected from the literature to identify nodes of high, medium, and low risk in models. The exposure rate of a risk factor in the study population was P_0. We generate 100,000 random P values of 0: 1, $P<P_0$ was recorded as 1 (ie, exposure), and $P>P_0$ was recorded as 0 (ie, nonexposure). Each risk factor simulates a column of data, summarizing the exposure of 100,000 community residents to each risk factor.

Statistical Analysis

Missing Data Interpolation

Among 11,692 participants, 1 individual did not have information on hypertension, and 2556 individuals lacked information on blood lipid levels. In accordance with previous studies [12,19,20], the incomplete data for hypertension and dyslipidemia were imputed simultaneously by multiple imputations (n=25) using the R package MICE (Stef van Buuren) [21]. Based on the Akaike information criterion value, 2 of the interpolation datasets were selected and the same analysis was performed on the selected interpolation dataset to identify results that were likely to be robust. The detailed data report of the other interpolation set is presented in the Tables S2-S5 in Multimedia Appendix 1.

Model Evaluation

First, we assessed the discrimination for the Rothman-Keller model and the logistic regression model using the area under the receiver operating characteristic curve (AUC). Our primary objective was to determine whether the predictive capability of the Rothman-Keller model surpassed that of the conventional logistic regression model, particularly in a multiclassification situation. When the increase in AUC is not statistically significant, its interpretation can become challenging [22-24]. Therefore, in addition to the AUC, we incorporated the absolute net reclassification improvement (NRI) to evaluate the relative performance of the 2 models. If the absolute NRI is greater than 0 or less than or equal to 1, it indicates a positive improvement, showing that the predictive ability of the new index has improved compared to the old index for stroke events. Conversely, if absolute NRI is less than 0 or greater than or equal to −1, it signifies a negative change, suggesting an improvement in the predictive ability of the new model for no stroke events. If the absolute NRI is equal to 0, it means that the new model shows no improvement. In our study, we analyzed the reclassification and absolute NRI for individuals who experienced a stroke event and those who did not. For individuals who have had a stroke, being reclassified into a higher-risk group was deemed an improvement in classification, whereas being reclassified into a lower-risk group was considered a failure.

A P value less than .05 was considered to indicate statistical significance. The statistical analyses were performed using R 4.2 software (R Core Team).

Ethical Considerations

The ethics review committee of The People's Hospital of Ningxia Hui Autonomous Region approved this study (approval number: 2020-KY-053). Patients provided informed consent for using the data. Data were deidentified. No compensation was provided.

Results

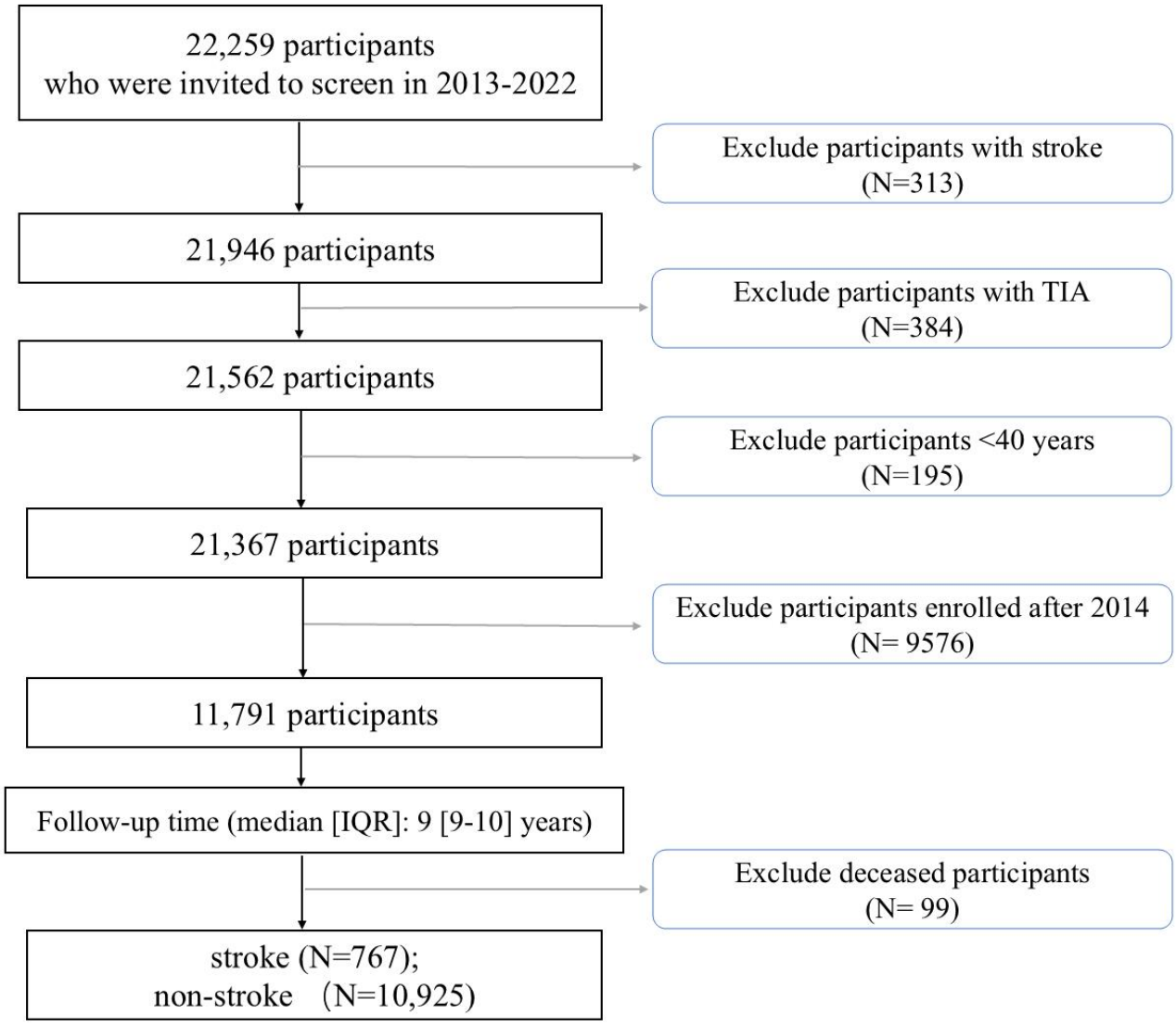
Characteristics of the Study Cohort

A total of 22,259 community residents were recruited in 2013. After excluding individuals with a history of stroke (n=313), and TIA (n=384), as well as participants younger than 40 years

XSL•FO
RenderX

old (n=195) and those recruited after 2014 (n=9576), 11,791 participants were included in the follow-up cohort. After a 10-year follow-up period, 99 participants were lost to follow-up. Finally, a total of 11,692 eligible participants were included in the final analysis (Figure 1). A total of 767 participants (6.6%) had a stroke by the end of the follow-up period. Based on the "8+2" stroke risk score, the 10-year stroke incidence among the 3 stroke risk groups of the community residents was as follows: low-risk group 4.47% (n=8908, 398 stroke cases); medium-risk group 14.71% (n=1074, 158 stroke cases); and high-risk group 12.34% (n=1710, 211 stroke cases) (Table 1). Kaplan-Meier survival curves are shown in Figure S1 in Multimedia Appendix 1.

Figure 1. Study participants screening process.



XSL•FO
RenderX

Table . Characteristics of participants based on different risk levels according to the "8+2" stroke risk score at baseline in Ningxia (N=11,692).

Characteristics	All participants, N (%)	High-risk participants, n (%)	Medium-risk participants, n (%)	Low-risk participants, n (%)
Total	11,692 (100)	1710 (14.6)	1074 (9.2)	8908 (76.2)
Age (years)				
40 - 49	5457 (46.7)	468 (27.4)	227 (21.1)	4762 (53.5)
50 - 59	3298 (28.2)	600 (35.1)	318 (29.6)	2380 (26.7)
60 - 69	2050 (17.5)	500 (29.2)	336 (31.3)	1214 (13.6)
≥70	887 (7.6)	142 (8.3)	193 (18.0)	552 (6.2)
Sex				
Female	5797 (49.6)	919 (53.7)	608 (56.6)	4270 (47.9)
Male	5895 (50.4)	791 (46.3)	466 (43.4)	4638 (52.1)
District				
Urban	6057 (51.8)	1212 (70.9)	189 (17.6)	4656 (52.3)
Rural	5635 (48.2)	498 (29.1)	885 (82.4)	4252 (47.7)
Family history of stroke				
Yes	361 (3.1)	289 (16.9)	15 (1.4)	57 (0.6)
No	11,331 (96.9)	1421 (83.1)	1059 (98.6)	8851 (99.4)
Heart disease				
Yes	461 (3.9)	265 (15.5)	196 (18.2)	0 (0)
No	11,231 (96.1)	1445 (84.5)	878 (81.8)	8908 (100)
Hypertension				
Yes	1684 (14.4)	903 (52.8)	781 (72.7)	0 (0)
No	10,008 (85.6)	807 (47.2)	293 (27.3)	8908 (100)
Dyslipidemia				
Yes	1639 (14.0)	1343 (78.5)	113 (10.5)	183 (2.1)
No	10,053 (86.0)	367 (21.5)	961 (89.5)	8725 (97.9)
Diabetes				
Yes	395 (3.4)	190 (11.1)	205 (19.1)	0 (0)
No	11,297 (96.6)	1520 (88.9)	869 (80.9)	8908 (100)
Smoking				
Yes	1355 (11.6)	439 (25.7)	92 (8.6)	824 (9.3)
No	10,337 (88.4)	1271 (74.3)	982 (91.4)	8084 (90.7)
Overweight				
Yes	2362 (20.2)	1100 (64.3)	267 (24.9)	995 (11.2)
No	9330 (79.8)	610 (35.7)	807 (75.1)	7913 (88.8)
Physical inactivity				
Yes	2983 (25.5)	1204 (70.4)	198 (18.4)	1581 (17.7)
No	8709 (74.5)	506 (29.6)	876 (81.6)	7327 (82.3)

Model Construction

The baseline incidence ratio (ρ) of 8 factors was obtained to calculate the population attributable risk percentage (PAR%) through a previous study [18]. The OR values of these factors were assessed using a logistic regression model. The parameters of the logistic model and the Rothman-Keller model are shown in Table 2.

XSL•FO
RenderX

Table . Parameters of risk exposure factors in the logistic model and Rothman-Keller model.

Risk factor	P_i[a]	OR_i[b] (95% CI)	β_i[c]	PAR (%)[d]	ρ[e]	S[f]
Hypertension						
Yes	0.580	3.00 (2.49 - 3.59)	1.097	50.5	0.495	1.485
No	0.420	1			0.495	0.495
Diabetes						
Yes	0.297	2.21 (1.67 - 2.90)	.793	17.4	0.826	1.825
No	0.703	1			0.826	0.826
Dyslipidemia						
Yes	0.297	0.94 (0.76 - 1.17)	–.060	19.6	0.804	0.756
No	0.703	1			0.804	0.804
Heart diseases						
Yes	0.691	1.24 (0.89 - 1.69)	.211	50.4	0.496	0.615
No	0.309	1			0.496	0.496
Smoking						
Yes	0.213	0.88 (0.69 - 1.10)	–.133	8.2	0.918	0.808
No	0.787	1			0.918	0.918
Overweight						
Yes	0.054	1.51 (1.27 - 1.79)	.411	2.1	0.979	1.478
No	0.946	1			0.979	0.979
Physical inactivity						
Yes	0.515	1.01 (0.85 - 1.20)	.010	34.9	0.651	0.658
No	0.485	1			0.651	0.651
Family history of stroke						
Yes	0.085	1.40 (1.01 - 1.92)	.338	5.1	0.949	1.329
No	0.915	1			0.949	0.949

[a]P_i: the exposure rate of individuals exposed to a risk factor in the whole population

[b]OR_i: the odds ratios of exposure to a risk factor;

[c] β_i is the beta coefficient

[d] PAR%: population attributed risk percentage

[e]ρ: baseline morbidity ratio

[f]S: risk score

After ranking the stroke incidence risk based on a dataset of 100,000 random entries, 2 nodes were selected for subdividing the risk groups into low-risk, medium-risk, and high-risk categories using the logistic regression model: node a (ID=25844, risk prediction score=0.06168760) and node b (ID=77778, risk prediction score=0.16451650). In addition, for the Rothman-Keller model, 2 other nodes were chosen for the same purpose: node A (ID=25426, risk prediction score=0.0021993391) and node B (ID=64553, risk prediction score =0.0047898060) (Figure S2 in Multimedia Appendix 1).

Evaluation of the Model

The sensitivity, specificity, and AUC of the “8+2” stroke risk score, the Rothman-Keller model, and the logistic regression model are presented in Table 3. The logistic regression model and the Rothman-Keller model demonstrated significant differences in AUC values compared to the “8+2” stroke risk score (Z=2.60, $P<.05$; Z=3.47, $P<.05$, respectively). However, no difference was observed between the logistic regression model and the Rothman-Keller model (Z=0.688, $P>.05$). The comparison of receiver operating characteristic curve is shown in Figure S3 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . The discrimination of "8+2" stroke risk score, logistic and Rothman-Keller model.

	Sensitivity	Specificity	AUC[a] (95% CI)	Z value (*P*) [b]	Z value (*P*)[c]
"8+2" stroke risk score	0.48	0.79	0.627 (0.619 - 0.636)		
Logistic model	0.41	0.85	0.649 (0.641 - 0.658)	3.47 (*P*=.001)	
Rothman-Keller model	0.52	0.74	0.646 (0.637 - 0.654)	2.60 (*P*=.009)	0.688 (*P*=.492)

[a]AUC: area under the curve

[b]:Versus "8+2" stroke risk score;

[c]:Versus logistic model

From the NRI, we found that the majority of participants remained at the same level of risk for developing a stroke as predicted by the "8+2" stroke risk score, the logistic regression model, and the Rothman-Keller model (ie, along the diagonal from the lower left to the upper right). However, some community residents were reclassified as having a different level of risk for developing stroke (Figure S4 in Multimedia Appendix 1). According to the "8+2" stroke risk score, the NRI for reclassification of stroke events by the Rothman-Keller model was 7.8%, while the NRI for nonstroke events was −2.7%. The absolute NRI was then estimated to be 0.051 (*P*=.01), calculated using the sum of the net estimated for individuals who developed a stroke and those who did not. The NRI for the logistic regression model was 1.7% for stroke events and −0.7% for nonstroke events, resulting in an absolute NRI of 0.010 (*P*=.62) .

Discussion

Principal Findings

This cohort study followed 11,692 individuals aged 40 years and older for a duration of 10 years. According to the "8+2" stroke risk score, the stroke incidence in the low-risk (n=8908), medium-risk (n=1074), and high-risk groups (n=1710) was 4.5%, 14.7%, and 12.3%, respectively. We developed a logistic regression model and a Rothman-Keller model to validate and optimize the risk score. Through a comparative analysis of the performance of the 3 models, we found that the Rothman-Keller model exhibited the best performance.

We verified the efficacy of the model using an actual database. There was no significant difference in the AUC values between the logistic regression model and the Rothman-Keller model. To evaluate model performance more accurately, we used the NRI for a more in-depth analysis. The NRI assesses the effects of low-, medium-, and high-risk reclassification for both stroke and nonstroke events, resulting in a net reclassification that provides a more accurate estimate than that obtained with other approaches [23]. Positive values of the stroke NRI indicate that the model effectively identifies patients with stroke, enabling physicians to initiate targeted detection or treatment to prevent stroke events. In contrast, a decrease in the NRI for nonstroke events suggests that community residents with a low or medium risk as determined by the "8+2" stroke risk score may actually be at a higher risk of having a stroke. Based on the overall NRI, we conclude that the Rothman-Keller model enhances the reclassification of both stroke and nonstroke events [25]. This finding aligns with the results of previous studies. Researchers used the Rothman-Keller model to predict the likelihood of mild cognitive impairment in older Chinese individuals. Upon validation with actual population data, it was found that the model had appropriate accuracy and performed well in terms of predictive efficacy [16,26]. The model can be adjusted and optimized based on new research data and epidemiological changes, thereby maintaining its predictive power in a timely manner. Its methodology can be applied to risk assessments for other populations and chronic diseases, demonstrating significant universality.

There are many model studies aimed at predicting stroke risk [27-30]. For instance, SCORE2 is a risk assessment tool developed using extensive data from a large number of European populations. It is designed to evaluate the risk of cardiovascular disease in both men and women across 4 distinct risk areas in Europe within a 10-year period [31]. While this tool is widely applied, it has limitations in terms of ethnicity and geography. Significant prediction errors may occur when applying it to populations with considerable differences [9,32]. A study estimated the 10-year risk of stroke based on a cohort analysis and found that factors such as age, systolic blood pressure, diastolic blood pressure, FHS, atrial fibrillation, diabetes, and others can significantly predict the incidence of stroke [27]. Compared with other models, this study developed a Rothman-Keller model based on questionnaire information to identify new risk nodes through simulation datasets, which provided a basis for stroke prevention. Furthermore, the model's predictive power and accuracy were verified using real-world data. Logistic regression analysis indicated that smoking, heart disease, dyslipidemia, and physical inactivity were not related to stroke, which may be attributed to variations in demographic and stroke subtypes differences. This result was consistent with the results of previous Mendelian randomization studies [33-37].

The limitations of this study are as follows: first, the dataset used for verification only included participants from only one region. Studies in other provinces are necessary to evaluate the efficacy of our model. When research results are extrapolated to other populations with significant differences, it may be essential to consider the exposure rates of risk factors and OR or RR values for those populations in order to update and optimize the model. This process enhances the predictive accuracy and applicability of the Rothman-Keller model. Second, we were unable to perform a subgroup analysis on the various types of stroke. The primary objective of the program is to identify and intervene with high-risk populations to prevent the occurrence of stroke or reduce its risk, rather than focusing on the subtypes of ischemic stroke and hemorrhagic stroke.

XSL•FO
RenderX

Furthermore, given the limitations of medical resources and the acceptability of screening, the program may prioritize the implementation of more accessible and universal preventive measures. These measures include controlling blood pressure, quitting smoking, and increasing physical activity, all of which are effective in preventing both ischemic stroke and hemorrhagic stroke [38,39]. Third, during the decade, participants may have received lifestyle interventions, pharmaceutical treatments, and early clinical treatment that influenced the incidence of stroke events. However, in the medium- and high-risk groups for stroke, these patients represent a certain proportion. Our model primarily assesses the variations in the initial screening judgments, and the outcome events remain consistent across different models, making it unlikely to influence the study results. Finally, the low sensitivity observed in this study may be attributed to the lack of several important stroke prediction factors from the risk scoring scale, thereby limiting its predictive ability. In 2021, the Guidelines for Stroke Prevention and Treatment in China recommended including homocysteine testing in routine screening and conducting carotid artery examinations for high-risk populations when conditions permit. Our research findings further support this recommendation. Although this study has some limitations, it also presents several advantages. The diagnostic criteria for risk factors in this project are based on relevant guidelines and standards established by the China Health Commission. Staff members undergo standardized training and assessment, and only those who pass the assessment are qualified to conduct screening tasks. Therefore, the identification of risk factors in this study had high accuracy and credibility. We used the binomial distribution of risk factors to construct a random dataset of 100,000 community residents, which allowed us to determine the high-, medium-, and low-risk boundary values of the models. The variables in the model were easily obtained and predicted estimates could be derived through straightforward calculations. We included a substantial number of community residences from the CSHPSIP over a decade for external validation, which exhibited good discrimination and calibration.

Conclusions

In conclusion, the Rothman-Keller model may improve the predictive efficacy of stroke screening models. In the future, verification will need to be carried out in a wider population and combined with more risk factors. The Rothman-Keller model for assessing individualized stroke risk, combined with interactive information platforms for health education, is beneficial for decreasing the incidence of stroke among high-risk groups.

Acknowledgments

This work was supported by the Ningxia Natural Science Foundation (grant number 2023AAC03445) and Key R&D Project of Ningxia Hui Autonomous Region (grant number 2021BEG03099).

Data Availability

The datasets used and/or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LPF and PDF conceived the study. PDF, SXY, LWW, WXT, LZ, GYH, and MXJ designed and supervised the study. MH, SXY, LWW, WXT, LZ, GYH, and MXJ participated in data collection. MH, SXY, and LPF performed the whole data integration and analysis. MH and LZ wrote the first draft of the manuscript. LPF and MH improved the research and edited the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary data to this article can be found in Multimedia Appendix 1.
[DOCX File, 4311 KB - publichealth_v11i1e72497_app1.docx]

References

1. Feigin VL, Abate MD, Abate YH, et al. Global, regional, and national burden of stroke and its risk factors, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Neurol 2024 Oct;23(10):973-1003. [doi: 10.1016/S1474-4422(24)00369-7]
2. Wang LD, Peng B, Zhang HQ, et al. Brief report on stroke prevention and treatment in China, 2020. Chin J Cerebrovasc Dis 2022;19(2):136-144. [doi: 10.3969/j.issn.1672-5921.2022.02.011]
3. GBD 2016 Lifetime Risk of Stroke Collaborators, Feigin VL, Nguyen G, et al. Global, Regional, and Country-Specific Lifetime Risks of Stroke, 1990 and 2016. N Engl J Med 2018 Dec 20;379(25):2429-2437. [doi: 10.1056/NEJMoa1804492] [Medline: 30575491]
4. Tu WJ, Wang LD, Special Writing Group of China Stroke Surveillance Report. China stroke surveillance report 2021. Mil Med Res 2023 Jul 19;10(1):33. [doi: 10.1186/s40779-023-00463-x] [Medline: 37468952]

XSL•FO
RenderX

5. Chao BH, Yan F, Hua Y, et al. Stroke prevention and control system in China: CSPPC-Stroke Program. Int J Stroke 2021 Apr;16(3):265-272. [doi: 10.1177/1747493020913557] [Medline: 32223541]
6. Prevention RoS, Group TiCW. Brief report on stroke prevention and treatment in China, 2021. Chinese Journal of Cerebrovascular Diseases 2023;20(11):783-793. [doi: 10.3969/j.issn.1672-5921.2023.11.009]
7. Chinese Medical Association Branch of Neurology, Group of Cerebrovascular diseases, Branch of Neurology, Chinese Medical Association. Expert consensus on the use of the Chinese Ischemic Stroke Risk Assessment Scale. Chin J Neurol 2016;49(7):519-525. [doi: 10.3760/cma.j.issn.1006-7876.2016.07.003]
8. Zhang Y, Fang X, Guan S, et al. Validation of 10-year stroke prediction scores in a community-based cohort of Chinese older adults. Front Neurol 2020;11(986):33192957. [doi: 10.3389/fneur.2020.00986]
9. Liu X, Shen P, Zhang D, et al. Evaluation of atherosclerotic cardiovascular risk prediction models in China: results from the CHERRY study. JACC Asia 2022 Feb;2(1):33-43. [doi: 10.1016/j.jacasi.2021.10.007] [Medline: 36340248]
10. Jiang Y, Ma R, Guo H, et al. External validation of three atherosclerotic cardiovascular disease risk equations in rural areas of Xinjiang, China. BMC Public Health 2020 Dec;20(1):32993590. [doi: 10.1186/s12889-020-09579-4]
11. Tu W, Yan F, Chao B, Ji X, Wang L. Status of hyperhomocysteinemia in China: results from the China Stroke High-risk Population Screening Program, 2018. Front Med 2021 Dec;15(6):903-912. [doi: 10.1007/s11684-021-0871-4] [Medline: 34893949]
12. Guan T, Ma J, Li M, et al. Rapid transitions in the epidemiology of stroke and its risk factors in China from 2002 to 2013. Neurology (ECronicon) 2017 Jul 4;89(1):53-61. [doi: 10.1212/WNL.0000000000004056] [Medline: 28566547]
13. Guo J, Bai Y, Ding M, et al. Analysis of carotid ultrasound screening of high-risk groups of stroke based on big data technology. J Healthc Eng 2022;2022(6363691):6363691. [doi: 10.1155/2022/6363691] [Medline: 35126935]
14. Stoltzfus JC. Logistic regression: a brief primer. Acad Emerg Med 2011 Oct;18(10):1099-1104. [doi: 10.1111/j.1553-2712.2011.01185.x] [Medline: 21996075]
15. Rothman K, Keller A. The effect of joint exposure to alcohol and tobacco on risk of cancer of the mouth and pharynx. J Chronic Dis 1972 Dec;25(12):711-716. [doi: 10.1016/0021-9681(72)90006-9] [Medline: 4648515]
16. Wang B, Shen T, Mao L, Xie L, Fang QL, Wang XP. Establishment of a risk prediction model for mild cognitive impairment among elderly Chinese. J Nutr Health Aging 2020;24(3):255-261. [doi: 10.1007/s12603-020-1335-2] [Medline: 32115605]
17. Gu J, Li Y, Yu J, et al. A risk scoring system to predict the individual incidence of early-onset colorectal cancer. BMC Cancer 2022 Jan 29;22(1):122. [doi: 10.1186/s12885-022-09238-4] [Medline: 35093005]
18. Dong S, Fang J, Li Y, Ma M, Hong Y, He L. The population attributable risk and clustering of stroke risk factors in different economical regions of China. Medicine (Baltimore) 2020 Apr;99(16):e19689. [doi: 10.1097/MD.0000000000019689]
19. Luik A, Radzewitz A, Kieser M, et al. Cryoballoon versus open irrigated radiofrequency ablation in patients with paroxysmal atrial fibrillation: the prospective, randomized, controlled, noninferiority FreezeAF study. Circulation 2015 Oct 6;132(14):1311-1319. [doi: 10.1161/CIRCULATIONAHA.115.016871] [Medline: 26283655]
20. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009 Jun 29;338:b2393. [doi: 10.1136/bmj.b2393] [Medline: 19564179]
21. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate Imputation by Chained Equations in R. J Stat Softw 2011 Dec;45(3):1-67. [doi: 10.18637/jss.v045.i03]
22. Pencina MJ, D'Agostino RB, D'Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008 Jan 30;27(2):157-172. [doi: 10.1002/sim.2929] [Medline: 17569110]
23. Wilson PWF, Pencina M, Jacques P, Selhub J, D'Agostino R Sr, O'Donnell CJ. C-reactive protein and reclassification of cardiovascular risk in the Framingham Heart Study. Circ Cardiovasc Qual Outcomes 2008 Nov;1(2):92-97. [doi: 10.1161/CIRCOUTCOMES.108.831198] [Medline: 20031795]
24. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. JAMA 2017 Oct 10;318(14):1377-1384. [doi: 10.1001/jama.2017.12126] [Medline: 29049590]
25. Leening MJG, Vedder MM, Witteman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Ann Intern Med 2014 Jan 21;160(2):122-131. [doi: 10.7326/M13-1522] [Medline: 24592497]
26. Wang Q, Zhou S, Zhang J, et al. Risk assessment and stratification of mild cognitive impairment among the Chinese elderly: attention to modifiable risk factors. J Epidemiol Community Health 2023 Aug;77(8):521-526. [doi: 10.1136/jech-2022-219952] [Medline: 37321832]
27. Chien KL, Su TC, Hsu HC, et al. Constructing the prediction model for the risk of stroke in a Chinese population: report from a cohort study in Taiwan. Stroke 2010 Sep;41(9):1858-1864. [doi: 10.1161/STROKEAHA.110.586222] [Medline: 20671251]
28. Arafa A, Kokubo Y, Sheerah HA, et al. Developing a stroke risk prediction model using cardiovascular risk factors: the Suita study. Cerebrovasc Dis 2022;51(3):323-330. [doi: 10.1159/000520100] [Medline: 34844243]
29. Yang S, Han Y, Yu C, et al. Development of a model to predict 10-year risk of ischemic and hemorrhagic stroke and ischemic heart disease using the China Kadoorie Biobank. Neurology (ECronicon) 2022 Jun 7;98(23):e2307-e2317. [doi: 10.1212/WNL.0000000000200139] [Medline: 35410902]

XSL•FO
RenderX

30. Hong C, Pencina MJ, Wojdyla DM, et al. Predictive accuracy of stroke risk prediction models across Black and White race, sex, and age groups. JAMA 2023 Jan 24;329(4):306-317. [doi: 10.1001/jama.2022.24683] [Medline: 36692561]
31. Hageman S, Pennells L, Ojeda F, et al. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. Eur Heart J 2021 Jul 1;42(25):2439-2454. [doi: 10.1093/eurheartj/ehab309]
32. Amarenco P. Five-year risk of stroke after TIA or minor ischemic stroke. N Engl J Med 2018 Oct 18;379(16):1579-1581. [doi: 10.1056/NEJMc1808913]
33. Larsson SC, Burgess S, Michaëlsson K. Smoking and stroke: a mendelian randomization study. Ann Neurol 2019 Sep;86(3):468-471. [doi: 10.1002/ana.25534] [Medline: 31237718]
34. Li Q, Yan S, Li Y, Kang H, Zhu H, Lv C. Mendelian randomization study of heart failure and stroke subtypes. Front Cardiovasc Med 2022;9(844733):35463787. [doi: 10.3389/fcvm.2022.844733]
35. Beheshti S, Madsen CM, Varbo A, Benn M, Nordestgaard BG. Relationship of familial hypercholesterolemia and high low-density lipoprotein cholesterol to ischemic stroke: Copenhagen General Population Study. Circulation 2018 Aug 7;138(6):578-589. [doi: 10.1161/CIRCULATIONAHA.118.033470] [Medline: 29593013]
36. Valdes-Marquez E, Parish S, Clarke R, et al. Relative effects of LDL-C on ischemic stroke and coronary disease: a Mendelian randomization study. Neurology (ECronicon) 2019 Mar 12;92(11):e1176-e1187. [doi: 10.1212/WNL.0000000000007091] [Medline: 30787162]
37. Bahls M, Leitzmann MF, Karch A, et al. Physical activity, sedentary behavior and risk of coronary artery disease, myocardial infarction and ischemic stroke: a two-sample Mendelian randomization study. Clin Res Cardiol 2021 Oct;110(10):1564-1573. [doi: 10.1007/s00392-021-01846-7] [Medline: 33774696]
38. Gu H, Shao S, Liu J, et al. Age- and sex-associated impacts of body mass index on stroke type risk: a 27-year prospective cohort study in a low-income population in China. Front Neurol 2019;10(456):456. [doi: 10.3389/fneur.2019.00456] [Medline: 31118920]
39. Wang J, Wen X, Li W, Li X, Wang Y, Lu W. Risk factors for stroke in the Chinese population: a systematic review and meta-analysis. J Stroke Cerebrovasc Dis 2017 Mar;26(3):509-517. [doi: 10.1016/j.jstrokecerebrovasdis.2016.12.002] [Medline: 28041900]

Abbreviations

AUC: area under the receiver operating characteristic curve
CNSSS: China National Stroke Screening Survey
CSHPSIP: China Stroke High-risk Population Screening and Intervention Program
CSPPC: China Stroke Prevention Project Committee
FHS: family history of stroke
GBD: Global Burden of Disease
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems 10th Revision*
NRI: net reclassification improvement
TIA: transient ischemic attack
TRIPOD: Transparent reporting of multivariable prediction model for Individual Prognosis or Diagnosis

Edited by A Mavragani, T Sanchez; submitted 11.02.25; peer-reviewed by IL Rissanen, X Zhou; revised version received 27.06.25; accepted 04.07.25; published 21.08.25.

Please cite as:
Meng H, Liu Z, Pan D, Su X, Lu W, Wang X, Geng Y, Ma X, Liang P
Validation and Refinement of Scores to Predict Stroke Risk: Prospective Cohort Study
JMIR Public Health Surveill 2025;11:e72497
URL: https://publichealth.jmir.org/2025/1/e72497
doi:10.2196/72497

©Hua Meng, Zhuo Liu, Dongfeng Pan, Xinya Su, Wenwen Lu, Xingtian Wang, Yuhui Geng, Xiaojuan Ma, Peifeng Liang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Catch-Up of Routine School-Based Immunizations Since the COVID-19 Pandemic: 4-Year Observational Cohort Study

Kelvin Nartey[1*], MPH, PhD; Laura Reifferscheid[1*], MSc, RN; Marilou Kiely[2], PhD, RN; Shannon E MacDonald[1], PhD, RN; Canadian Immunization Research Network

[1]Faculty of Nursing, University of Alberta, Edmonton Clinic Health Academy, 11405-87 Avenue, Edmonton, AB, Canada

[2]Department of Social and Preventive Medicine, Faculty of Medicine, Université Laval, Quebec City, QC, Canada

[*]these authors contributed equally

Corresponding Author:
Shannon E MacDonald, PhD, RN
Faculty of Nursing, University of Alberta, Edmonton Clinic Health Academy, 11405-87 Avenue, Edmonton, AB, Canada

Abstract

Background: School-based immunization programs (SBIPs) have been widely used to efficiently deliver vaccines to children and adolescents, achieving higher and more equitable coverage than community-based approaches, though implementation varies across regions. However, the COVID-19 pandemic disrupted SBIPs, leading to significant declines in vaccine coverage and highlighting the need for long-term assessments to determine the effectiveness of recovery efforts, especially as current research has largely overlooked SBIPs or focused narrowly on select vaccines.

Objective: This study aimed to assess the impact of the COVID-19 pandemic and ongoing catch-up immunization efforts on vaccine coverage and administration volume among school-aged children in the Canadian province of Alberta. We also sought to determine the magnitude of any remaining vaccine deficits and estimate the duration of additional catch-up efforts required to address those deficits.

Methods: In this retrospective cohort study, we used population-wide administrative health data to calculate monthly cumulative vaccine coverage among Grade 1, 6, and 9 cohorts, comparing four pandemic cohorts (2019 - 2020, 2020 - 2021, 2021 - 2022, and 2022 - 2023 grade-years) to a prepandemic cohort for each grade. We also calculated vaccine administration rates as monthly count of vaccines delivered, stratified by routine (ie, received during the school year as scheduled) and catch-up (ie, received after the regularly scheduled school year) doses. For each vaccine, we calculated vaccine coverage deficits remaining at the end of the study period and estimated time to clear those deficits based on observed catch-up uptake rates.

Results: All Grade 6 pandemic cohorts had lower human papillomavirus (HPV) and hepatitis B vaccine coverage at the end of the school year (2.6 - 60 percentage points lower and 1.2 - 41.7 percentage points lower, respectively), though coverage had increased to near prepandemic levels after approximately 3 years of follow-up. Coverage among the Grade 1 and 9 cohorts remained below prepandemic levels at study end; the largest deficits were noted in the 2022 - 2023 Grade 1 cohort (7.87 percentage points lower for measles-containing vaccines, 4.64 percentage points lower for pertussis-containing vaccines) and the 2020 - 2021 cohort for Grade 9 (14.0 percentage points lower for meningococcal conjugate ACYW-135 vaccine, 12.9 percentage points lower for pertussis-containing vaccines). The additional catch-up time required to reach prepandemic coverage levels for all cohorts was estimated to be less than 1 year for Grade 6 cohorts, 5.6 years for Grade 1 cohorts, and 4.8 years for Grade 9 cohorts.

Conclusions: Partial recovery of vaccine coverage was found for cohorts impacted by interruptions to school-based programming during the COVID-19 pandemic. Further efforts to alleviate coverage deficits among Grade 1 cohorts may be addressed in routine Grade 6 programs, while targeted efforts in postsecondary institutions may be required to address delays in Grade 9 immunizations.

(JMIR Public Health Surveill 2025;11:e79669) doi:10.2196/79669

KEYWORDS

vaccination; measles-mumps-rubella; human papillomavirus; hepatitis B; meningococcal conjugate ACYW-135; Tdap

Introduction

School settings have been used in varying capacities around the world to administer vaccines to children and adolescents. Worldwide, the first grade is most frequently targeted for school-based immunization programs (SBIPs), though delivery to preadolescents and adolescents is increasing [1]. SBIPs are attractive, largely due to the ability to allow rapid delivery to a large portion of the target population [2], and have been shown to achieve higher and more equitable vaccine coverage than community-based delivery alone [1,3,4]. SBIPs have been used to support vaccine delivery across Canada for decades, though

XSL•FO
RenderX

there is regional variation in the grades and vaccines offered, as well as how these settings are used for catch-up immunization and implementation of mandatory immunization policies [5].

Despite previous successes, school closures and social distancing measures implemented during the COVID-19 pandemic resulted in significant decreases in vaccine coverage among those routinely immunized through SBIPs [6]. These critical drops in vaccine coverage raised concerns about possible resurgences of vaccine-preventable diseases [7] and left immunization programs with the added challenge of identifying and vaccinating those who missed the recommended doses. In addition, increased parental vaccine hesitancy in some areas of the world [8] has led to concerns that decreases in vaccine coverage may persist or become more pronounced [9].

Research exploring pandemic-related disruptions to routine childhood immunization programs has focused mainly on children aged 5 years and younger [10,11] or adolescent coverage outside of SBIPs [12-15], with minimal attention paid to SBIPs to date [16]. Those evaluating coverage among children eligible for immunization through SBIPs have limited their analysis to select vaccines (human papillomavirus [HPV] vaccine only [17]; HPV and meningococcal conjugate ACYW-135 [MenC-ACYW] only [6]), providing a limited picture of the full extent and impact of catch-up activities. In addition, current evidence points to immunization deficits that have persisted past the initial months of pandemic-related health care and school disruptions [14,15,18]. Assessment of vaccine coverage over the longer term is needed to determine whether recovery efforts have been effective or additional strategies are needed.

The aim of this study was to assess the impact of the COVID-19 pandemic and ongoing catch-up immunization efforts on vaccine coverage and administration volume among school-aged children in the Canadian province of Alberta, using more than four years of observation since the pandemic. We also sought to determine the magnitude of any remaining vaccine deficits and estimate the duration of additional catch-up efforts required to address those deficits.

Methods

Study Design and Setting

This retrospective observational cohort study was conducted in Alberta, which has a population of 4.8 million people. Almost 99% of Albertans are registered with the universal, publicly funded Alberta Health Care Insurance Plan (AHCIP), which provides access to all vaccines included in the provincial routine immunization schedule free of charge. With the exception of travel and seasonal vaccines (for example, influenza and COVID-19), all vaccines for school-aged children are provided exclusively by public health nurses through community-based public health centers and SBIPs, which includes screening and vaccinating children in Grade 1, Grade 6, and Grade 9. The Grade 1 program acts as a catch-up opportunity for those who missed preschool immunizations, while the Grades 6 and 9 programs are the main touchpoints for offering scheduled adolescent vaccines [19]. While schools are used to offer and provide vaccines, no vaccines are mandatory for school attendance in Alberta. The province's routine immunization schedule is provided in Multimedia Appendix 1. This study is reported in accordance with the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) cohort reporting guidelines [20].

Data Sources and Population

The data set was created by linking the Provincial School Immunization Record (PSIR), which contains individual-level records of student enrollment for each year, including grade, age, and sex, with the Immunization and Adverse Reaction to Immunization (Imm/ARI) repository, which includes a record of all publicly-funded vaccines provided in the province. We excluded those who lived in the city of Lloydminster (as vaccines are delivered by a neighboring province), and those who died or left the province before the study end date. To focus on institutions with SBIPs, we excluded schools that offer online or distance education only, as well as postsecondary, continuing education, summer, or evening or weekend schools. Schools on First Nations reserves were also excluded, as immunization data is not consistently submitted to the Imm/ARI repository. To align with the standard ages of assessment in the province, we also restricted cohorts based on individual age as of September 01 of the school year, including only those 5 - 7 years of age for Grade 1, 9 - 11 years of age for Grade 5, 10 - 12 years of age for Grade 6, and 13 - 15 years of age for Grade 9.

We examined 5 annual cohorts for each grade targeted for SBIPs, including 4 pandemic cohorts (2019 - 2020, 2020 - 2021, 2021 - 2022, and 2022 - 2023) and 1 prepandemic cohort. For Grades 1 and 9, the prepandemic cohort was defined as the cohort eligible for immunization during the 2018 - 2019 school year. For Grade 6, we used the 2017 - 2018 Grade 5 group as the prepandemic cohort, as the target grade for vaccine delivery changed from Grade 5 to Grade 6 in 2018 - 2019, resulting in no SBIP for Grade 5 or 6 in 2018 - 2019.

For each grade, we focused on the vaccines routinely delivered through SBIPs, including a second dose of measles-mumps-rubella (MMR) or measles-mumps-rubella—varicella (MMRV), and fifth dose of diphtheria, tetanus, acellular pertussis, polio (DTap-IPV) in Grade 1; the 2-dose HPV and hepatitis B series in Grade 6; and, the single dose of MenC-ACYW and tetanus, diphtheria, and acellular pertussis (Tdap) series in Grade 9. We included all pertussis-containing vaccines in the DTaP-IPV (Grade 1) and Tdap (Grade 9) analyses.

Data Analysis

For each vaccine, stratified by grade-year cohort, we calculated cumulative vaccine coverage, that is, the proportion of the grade-year cohort who had received the vaccine dose of interest, calculated for each month of observation. We determined cumulative coverage for each cohort from the end of their respective school year (eg, July 2020 for the 2019 - 2020 cohorts) to the end of the study period (July 2024), aiming to capture the impact of both the intensive and routine catch-up activities. As the Grade 1 immunization program is a routine catch-up opportunity (ie, children would have had the

XSL•FO
RenderX

opportunity to receive the vaccines of interest before entering their Grade 1 year), we included monthly cumulative coverage from the start of each cohort's Grade 1 year (eg, September 2019 for the 2019 - 2020 cohorts). For the Grade 6 HPV and hepatitis B series, we calculated both complete (2 dose) and 1+ dose coverage, given recent recommendations that 1 dose of HPV vaccine may confer immunity. During the prepandemic (2017 - 2018) school year, a 3-dose program for both HPV and hepatitis B was in place, which subsequently switched to a 2-dose program; thus, complete coverage was defined as receipt of 3 doses for the 2017 - 2018 cohort, and at least 2 doses for all subsequent cohorts. To estimate pandemic impacts, we calculated three measures of comparison for each vaccine of interest, using the prepandemic cohort as the counterfactual for pandemic cohorts: (1) cumulative coverage difference, (2) vaccine administration volume, and (3) cumulative immunization deficits.

To determine cumulative coverage differences, cumulative monthly coverage for each pandemic cohort was compared to the analogous month for the prepandemic cohort. For example, the July 2021 coverage for the 2019 - 2020 pandemic cohort represents coverage at 12 months from end of grade year and thus was compared to the coverage for the prepandemic cohort at 12 months from the end of grade year. We calculated 95% CI for binomial proportions, for both coverage and difference in coverage.

Vaccine administration, that is, a count of the number of vaccine doses delivered, was assessed for each grade-year cohort for each month of observation, collapsed into "school year" (from September to June) and "summer" (July and August) for each study year. We calculated administration for 1 vaccine for each grade: for Grade 6 and 9, we chose the vaccine with the highest administration volume (HPV and MenC-ACYW, respectively); for Grade 1, we chose to assess MMR or MMRV administration, given currently ongoing measles outbreaks globally. Administration counts were stratified by "routine" (ie, received during the school year as scheduled) and "catch-up" (ie, received after the regularly scheduled school year) immunization. We used dose count rather than calculating a delivery rate, which would correct for differences in cohort sizes, as we were interested in the magnitude and persistence of changes to vaccine administration volume associated with catch-up efforts.

We calculated the cumulative deficits for each vaccine at the end of the study period by comparing the observed number of individuals vaccinated for all pandemic cohorts to an expected number of individuals, estimated based on vaccine coverage for the prepandemic cohort at the end of the study period. For each pandemic cohort, the expected number of individuals vaccinated at the end of the study period was calculated by multiplying the observed cohort vaccine coverage at the end of the study period by the number of individuals in the pandemic cohort. We then summed the expected numbers of individuals vaccinated for all pandemic cohorts and calculated the difference from the observed number vaccinated at the end of the study period. Finally, we compared this calculated difference to observed maximum and average annual catch-up vaccine uptake, that is, number of individuals vaccinated [21], to estimate time to clear any accumulated deficits in coverage. All analyses were completed using SAS (version 9.4; SAS Institute Inc), and Microsoft Excel (version 1808).

Ethical Considerations

This study received approval from the Health Research Ethics Board (HREB) at the University of Alberta (Pro00134011). Informed consent to participate was waived by the HREB. All study data were deidentified.

Results

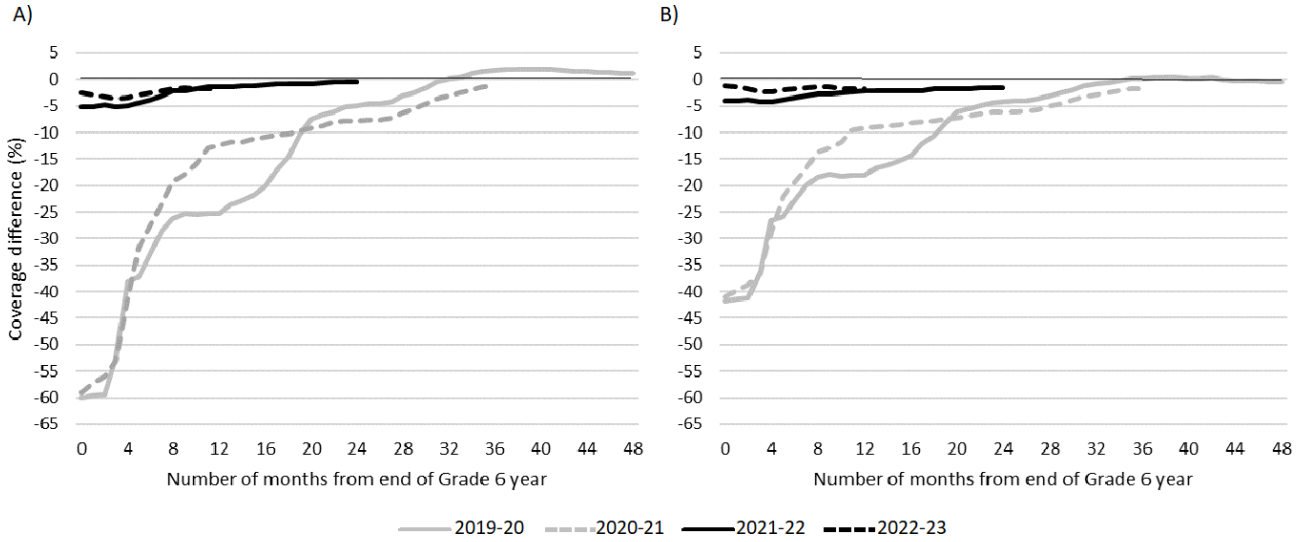
A total of 799,961 children were included in our study cohorts; 270,603 were in the Grade 1 cohorts, 273,163 in Grade 6, and 256,195 in Grade 9. The total number in each grade-year cohort is provided in Multimedia Appendix 1.

Cumulative Coverage Difference

Cumulative vaccine coverage for the Grade 6 pandemic cohorts was lower than prepandemic levels at the end of each school year, although they reached at or near prepandemic coverage by the end of the observation period (see Figure 1). The 2019 - 2020 and 2020 - 2021 cohorts finished their school years with the largest deficits: 60 and 59 percentage points lower for HPV (see Figure 1A and Table S2 in Multimedia Appendix 1); and 41.7 and 41.0 percentage points lower for hepatitis B (see Figure 1B and Table S3 in Multimedia Appendix 1). For the 2019 - 2020 cohort, coverage levels were at or above prepandemic levels by February (HPV) and May (hepatitis B) of 2023. Another cohort, the 2021 - 2022 cohort, reached prepandemic levels for HPV by the end of the study period. For all other cohorts, coverage levels were less than 2 percentage points lower than prepandemic levels by the end of the observation period for both vaccines. For comparison, coverage for the prepandemic cohort ranged from 65.6% (end of Grade 6 year) to 75.4% (at 48 months of follow-up) for HPV (see Table S2 in Multimedia Appendix 1), and 73.7% (end of Grade 6 year) to 83% (at 48 months of follow-up) for hepatitis B (see Table S3 in Multimedia Appendix 1). HPV vaccine coverage remained below hepatitis B coverage levels throughout the study period, for all cohorts. Cumulative 1+ dose coverage for the most impacted cohort (2020 - 21) was approximately 15 percentage points lower than prepandemic levels at the end of the school year for both vaccines, recovering to within 2 percentage points of prepandemic levels within 31 (HPV) and 36 (hepatitis B) months of observation (see Figure S1 in Multimedia Appendix 1).

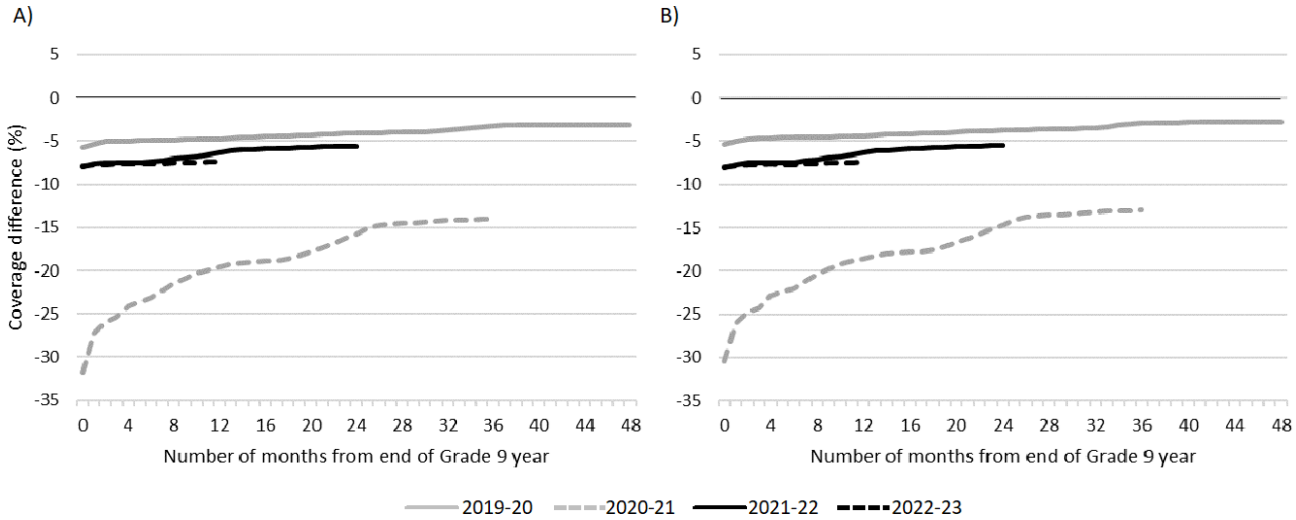
XSL•FO
RenderX

Figure 1. Differences in monthly complete (2 dose) coverage compared to prepandemic school year (2017 - 2018) for each Grade 6 cohort for: (A) human papillomavirus (HPV) and (B) hepatitis B. Differences were calculated by comparing the cumulative monthly coverage for each pandemic cohort to the cumulative coverage in the analogous month for the prepandemic cohort.



Cumulative coverage for all Grade 9 pandemic cohorts remained below prepandemic coverage throughout the study period (see Figure 2). The largest deficits in coverage were observed for the 2020 - 2021 Grade 9 cohort, with coverage for both vaccines more than 30 percentage points lower than prepandemic at the end of the school year, with deficits persisting through to the end of observation: 14 percentage points (95% CI 13.5 - 14.5) lower for MenC-ACYW coverage (see Figure 2A and Table S4 in Multimedia Appendix 1) and 12.9 percentage points (12.4 - 13.4) lower for Tdap (see Figure 2B and Table S5 in Multimedia Appendix 1). Coverage for the prepandemic cohort ranged from 84.9% (at end of Grade 9 year) to 86.6% (after 48 months of follow-up) for MenC-ACYW (see Table S4 in Multimedia Appendix 1), and from 85.6% (at end of Grade 9 year) to 88.2% (at 48 months of follow-up) for Tdap (see Table S5 in Multimedia Appendix 1).

Figure 2. Differences in monthly complete coverage compared to prepandemic school year (2018 - 19) for each Grade 9 cohort for: (A) Meningococcal conjugate ACYW-135 (MenC-ACYW; 1 dose), and (B) Tetanus, diphtheria, and acellular pertussis (Tdap; 1 dose). Differences were calculated by comparing the cumulative monthly coverage for each pandemic cohort to the cumulative coverage in the analogous month for the prepandemic cohort.

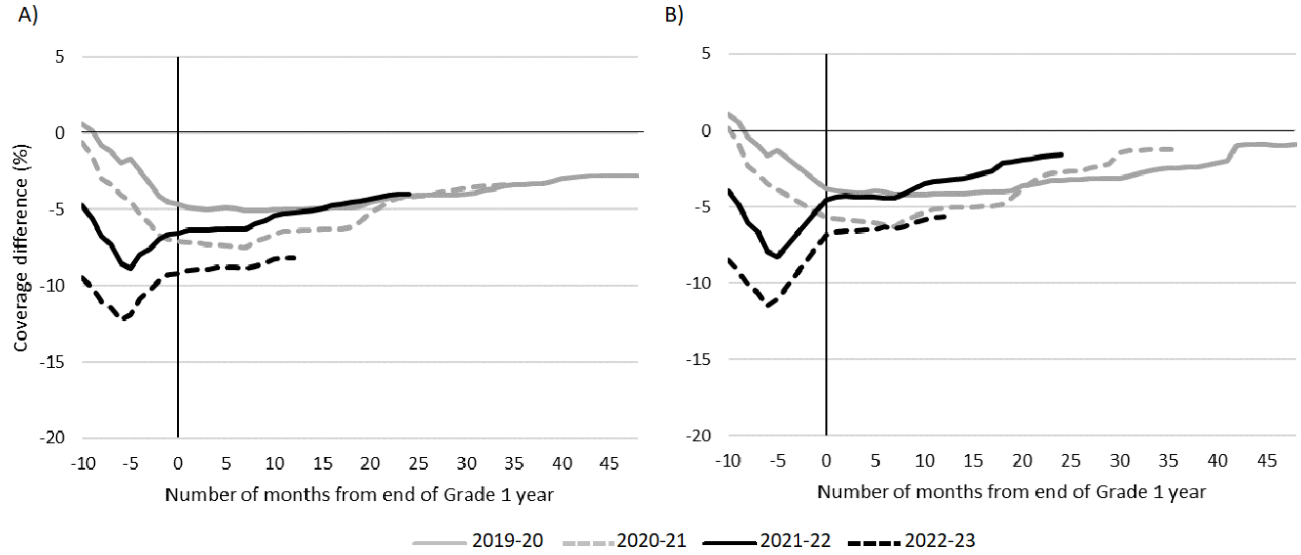


For all Grade 1 pandemic cohorts, cumulative coverage remained below prepandemic levels throughout the observation period, for both MMR or MMRV and DTaP-IPV (see Figure 3). For both vaccines, coverage at the start of the Grade 1 year was similar to prepandemic levels for the 2019 - 2020 and 2020 - 2021 cohorts, but significantly lower for the 2021 - 2022 and 2022 - 2023 cohorts. During the school year, coverage for the 2019 - 2020 and 2020 - 2021 cohorts continued to decline relative to prepandemic levels. In contrast, the 2021 - 2022 and 2022 - 2023 cohorts ended their Grade 1 years with coverage deficits similar to the start of the school year. For all cohorts, coverage deficits gradually decreased during the remaining period of observation. MMR or MMRV coverage for all cohorts remained significantly below prepandemic levels at the end of the study period, ranging from 3.06 percentage points lower (2.69 - 3.44) for the 2019 - 2020 cohort, to 7.87 percentage points lower (7.46 - 8.28) for the 2022 - 2023 cohort (see Figure 3A and Table S7 in Multimedia Appendix 1). Smaller deficits were observed for DTaP-IPV coverage, ranging from 0.5 percentage points lower (0.11 - 0.92) in the 2019 - 2020 cohort to 4.64 percentage points lower (4.20 - 5.08) in the 2022 - 2023 cohort (see Figure 3B and Table S6 in Multimedia Appendix

XSL•FO
RenderX

1). Coverage for the prepandemic cohort increased during the period of comparison: from 76.6% (76.3% - 77%; September 2018) to 90.5% (90.2% - 90.7%; July 2023) for MMR or MMRV (see Table S7 in Multimedia Appendix 1); from 74.3% (74% - 74.7%; September 2018) to 87.1% (86.8% - 87.4%; July 2023) for DTap-IPV (see Table S6 in Multimedia Appendix 1). Compared to DTaP-IPV, generally larger deficits in MMR or MMRV coverage were observed in all pandemic Grade 1 cohorts; MMR or MMRV coverage remained higher than DTaP-IPV in the 2019 - 2020, 2020 - 2021, and 2021 - 2022 cohorts, but was lower in the 2022 - 2023 cohort.

Figure 3. Differences in monthly complete coverage compared to prepandemic school year (2018 - 19) for each Grade 1 cohort for: (A) Measles-mumps-rubella (MMR) or measles-mumps-rubella—varicella (MMRV; 2 doses), and (B) Diphtheria, tetanus, acellular pertussis, and polio (DTap-IPV; 5 doses). Differences were calculated by comparing the cumulative monthly coverage for each pandemic cohort to the cumulative coverage in the analogous month for the prepandemic cohort.

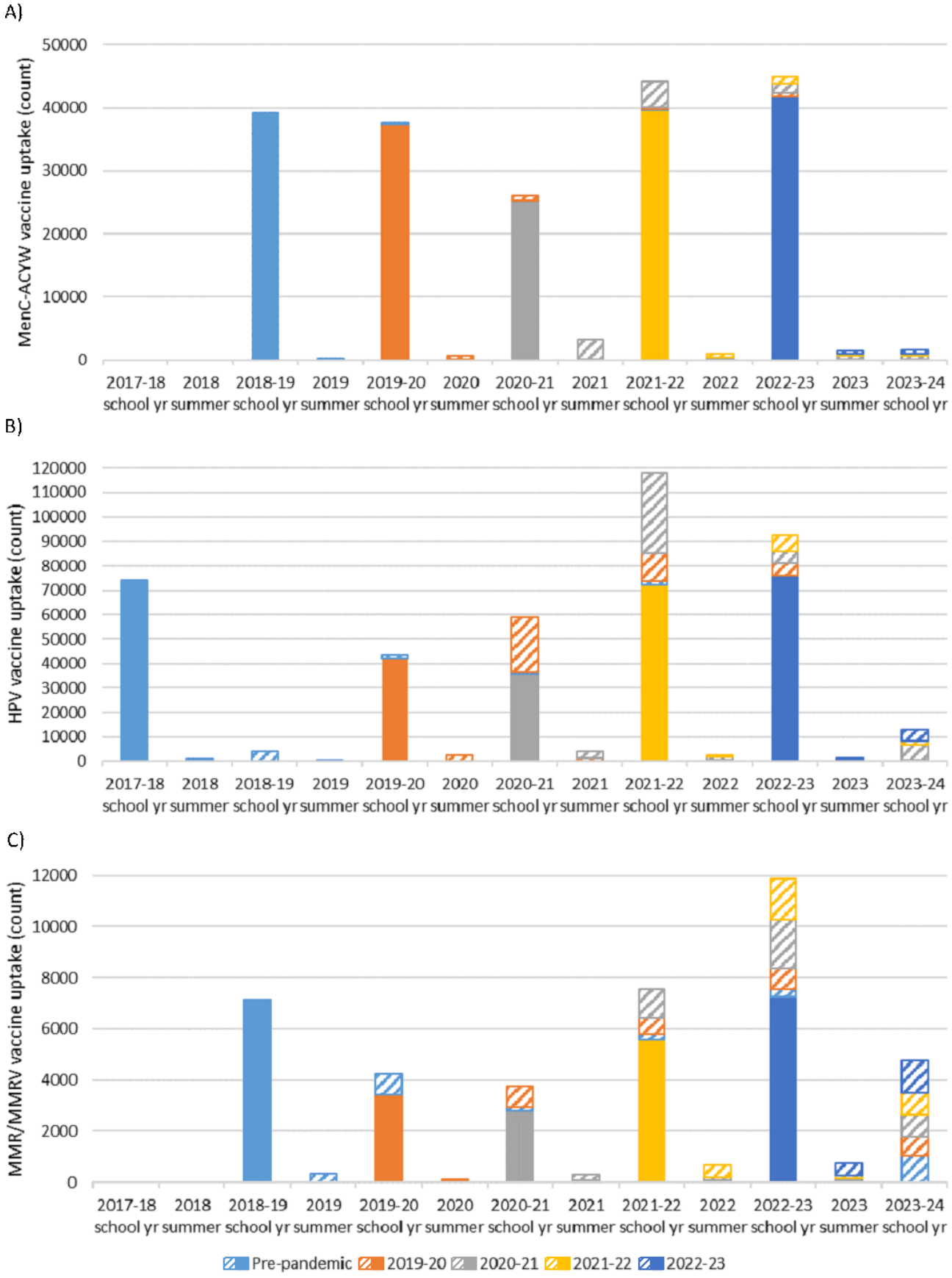


Vaccine Administration

Drops in vaccine administration (ie, count of the number of vaccine doses delivered) were observed during the 2019 - 2020 and 2020 - 2021 school years, with volume increasing to above prepandemic levels during the 2021 - 2022 and 2022 - 2023 school years (see Figure 4). For Grade 1, the largest volume of catch-up immunization occurred during the 2023-2024 year (2023 summer and 2023 - 2024 school year [from July 2023 to June 2024], total of 5521 MMR or MMRV doses delivered). For both Grade 6 and Grade 9, the largest volume of catch-up immunization was completed during the 2021-2022 year (from July 2021 to June 2022), with totals of 50,290 (HPV) and 7780 (MenC-ACYW) doses delivered, respectively. Immunization of the prepandemic cohort continued throughout the study period for all vaccines, with vaccine delivery observed during each summer and school year included. While generally higher vaccine administration volumes were noted during the school years, we are unable to identify whether those doses were delivered through SBIPs or through other public health vaccination efforts.

XSL•FO
RenderX

Figure 4. Vaccine administration volume for: (A) Meningococcal conjugate ACYW-135 (MenC-ACYW), **(B)** human papillomavirus (HPV), and (C) measles-mumps-rubella (MMR) or measles-mumps-rubella—varicella (MMRV) by grade-year cohort. Solid bars represent routine vaccine administration, and hatched bars represent catch-up vaccine administration.



XSL•FO
RenderX

Cumulative Immunization Deficits

The cumulative deficits of unimmunized or incompletely immunized individuals for each vaccine, aggregated for all years, are presented in Table 1. The estimated time to clear these deficits is also presented, including the minimum time required (based on maximum effort observed in the previous years) and average time required (based on the average effort observed in the previous years). Deficits were highest for the Grade 1 and Grade 9 cohorts, with correspondingly greater time estimated to clear these deficits.

Table . Cumulative vaccine deficits among pandemic cohorts at the end of the study period (July 2024), relative to maximum vaccine coverage for prepandemic grade cohorts, and estimated time to clear deficit.

Grade and vaccine	Number of individuals vaccinated		Cumulative deficit	Estimated remaining catch-up time (years)	
	Expected[a]	Observed		Minimum[b]	Average[c]
Grade 1					
DTaP-IPV[d]	198,332	183,132	−15,200	3.1	5.6
MMR[e] or MMRV[f]	199,198	184,409	−14,789	3.0	5.4
Grade 6					
HPV[g]	167,633	161,700	−5933	0.14	0.37
Hepatitis B	186,029	176,768	−9261	0.22	0.78
Grade 9					
MenC-ACYW[h]	180,871	164,436	−16,435	2.1	4.8
Tdap[i]	185,057	167,623	−17,434	2.2	4.7

[a]Estimated number of individuals vaccinated in pandemic cohorts to achieve coverage equivalent to maximum coverage observed in pre-pandemic cohort.

[b]Calculated using maximum annual catch-up uptake observed for each grade.

[c]Calculated using average annual catch-up uptake observed for each grade.

[d]DTaP-IPV: diphtheria, tetanus, acellular pertussis, and polio.

[e]MMR: measles-mumps-rubella.

[f]MMRV: measles-mumps-rubella—varicella.

[g]HPV: human papillomavirus.

[h]MenC-ACYW: meningococcal conjugate ACYW-135.

[i]Tdap: tetanus, diphtheria, and acellular pertussis.

Discussion

Principal Findings

We explored the cumulative impact of immunization program interruptions and subsequent catch-up efforts that have occurred since the COVID-19 pandemic on vaccine coverage and administration among school-aged children. Our results indicate that catch-up efforts resulted in significant recovery of vaccine coverage levels; however, important gaps remain. HPV and hepatitis B vaccine coverage levels had reached or were near prepandemic levels in all pandemic Grade 6 cohorts by the end of follow-up, though recovery took approximately 36 months for cohorts with the most significant relative deficits in coverage (2019 - 20 and 2020 - 21 grade-year). Our findings regarding persistent coverage deficits among Grade 1 and Grade 9 cohorts, more than 3 school years after pandemic-related school closures, are particularly concerning, given the recent elevation in pertussis [22] and measles [23] cases recorded in Canada. Both programs face unique challenges, with vaccine delivery to Grade 1 cohorts deprioritized early in the pandemic, and limited access to Grade 9 cohorts as they age out of routine SBIPs. While the routine SBIPs provide additional opportunities to catch up Grade 1 cohorts, efforts outside of secondary schools will be required to reach Grade 9 students who remain undervaccinated.

We found significant, though largely temporary, drops in HPV and hepatitis B vaccine coverage among Grade 6 cohorts. These results indicate that potential changes in immunization attitudes had little to no impact on uptake for these vaccines. This is particularly important to note for the HPV vaccine, which has historically been subject to large drops in uptake due to changes in public sentiment [24]. Similar to our findings, a largely temporary drop in HPV and hepatitis B vaccine coverage was observed in the province of Quebec, which also has a well-established SBIP [17]. However, that study observed a return to near prepandemic 1+ dose coverage levels after only one year of catch-up efforts for the most impacted cohort (2020 - 2021 school year), after experiencing drops in coverage larger than reported here, highlighting the importance of contextualizing findings within regional strategies used for intensive catch-up programs. While results from Alberta indicate that catch-up efforts were successful at returning coverage to prepandemic levels, coverage of both HPV and hepatitis B remains significantly below national vaccine coverage goals of 90% [25].

XSL•FO
RenderX

We found catch-up immunization efforts for Grade 1 cohorts were delayed relative to the Grade 6 and 9 cohorts, confirming that primary vaccine series were prioritized in catch-up efforts [26]. Concerningly, the most recent (2022 - 2023) Grade 1 cohort in our study showed the largest coverage deficit compared to prepandemic levels. This cohort would have been scheduled to receive their preschool vaccines during the initial stages of the pandemic, when access to immunization services was limited and infant vaccines were prioritized [27]. Other studies also noted relatively larger declines in immunization for this age group compared to infants [13,15,28,29]. While the Grade 1 SBIP coverage improvements for this cohort were similar to the SBIP in the prepandemic cohort, deficit levels from the preschool program remain. It will be important to understand whether the limited catch-up during the Grade 1 program is due to individual (eg, increased vaccine refusal, decreased school attendance) or programmatic (eg, staff capacity, school access) factors.

Our results indicate that catch-up efforts impacted coverage for all cohorts, across all vaccines, throughout the study period, demonstrating the importance of maximizing opportunities for immunization. While intensive catch-up in different grades was prioritized at different times, these efforts represent notable increases in vaccine delivery workload. In addition, these numbers do not capture unseen efforts related to mitigating potentially increased student stress [30], and heightened school collaboration required to implement catch-up immunization efforts. It is unclear whether it is feasible for providers to sustain this level of catch-up efforts for the estimated additional time required to address outstanding deficits. Finally, while ongoing catch-up efforts can simplify vaccine access and prompt vaccine uptake, they also have the potential to contribute to vaccine fatigue. Vaccine fatigue, described as individual "inertia or inaction towards vaccine information or instruction due to perceived burden and burnout" [31], has the potential to lead to ongoing decreases in vaccine uptake. More information on the impact of repeat offering is needed to determine whether those subsequently vaccinated simply missed previous opportunities, or their willingness to be vaccinated changed due to programmatic (eg, removing COVID-19 immunization from schools) or contextual (eg, infectious disease outbreaks) factors. While ongoing monitoring of vaccine-specific coverage deficits can help appropriately direct program efforts, optimizing the use of schools for catch-up programs will require careful consideration of the impact of these efforts on providers, children and their parents or caregivers, and schools.

Strengths and Limitations

This study provides a comprehensive summary of how vaccine coverage among children targeted through routine SBIPs was impacted by the COVID-19 pandemic and subsequent catch-up immunization efforts. Rather than evaluating a 1-time measure of vaccine coverage at a set age, we followed each cohort after grade completion to study end (July 2024). As a result, total cumulative vaccine coverage for our prepandemic comparison cohorts reflects vaccine activity that occurred before (ie, routine) and during the pandemic (ie, catch-up). Thus, the observed catch-up vaccination for this cohort may also reflect pandemic-related decreases in vaccination activity, causing an underestimation of pandemic impacts by the end of our study period.

Our work benefited from a population-based vaccine repository with information on all vaccines provided in the province, as well as complete denominator data. However, it is possible that vaccines received outside of the province were not completely captured in the vaccine repository, which could lead to underestimations in coverage. We also did not capture prepandemic trends that may have inflated or deflated coverage differences noted for the pandemic cohorts. Our HPV and hepatitis B prepandemic comparison year reflected coverage for the 3-dose series offered to Grade 5 students, which we compared to cohorts exposed to the 2-dose series offered in Grade 6. The decrease in series doses could have contributed to lower coverage deficits noted in the pandemic cohorts. We were also unable to identify whether catch-up vaccines were delivered through school programs or through other efforts (eg, outreach programs and community clinic visits). Future work should consider the impact of catch-up efforts on inequities in vaccine coverage.

Conclusions

We found coverage deficits that occurred during and after the COVID-19 pandemic persisted for at least 3 years after the pandemic-related school and health facility closures were initiated in March 2020. While we noted recovery to prepandemic vaccine coverage levels for vaccines scheduled in Grade 6, deficits remain in those targeted for delivery in Grades 1 and 9. Existing Grade 6 and 9 routine programs may provide additional opportunities to address immunization gaps among younger cohorts; strengthening these should be a priority. However, reaching those who missed Grade 9 vaccines will require targeted efforts in high school and post-secondary educational institutions. A better understanding of the approaches used for successful catch-up, in the context of routine delivery approaches, is required to ensure intensive and routine catch-up programs are optimized.

Acknowledgments

The authors would like to thank Laura Ekekhor and Bharath Sunil Nair for their assistance with compiling data tables. This study was supported by the Public Health Agency of Canada and the Canadian Institutes of Health Research, through the Canadian Immunization Research Network (CIRN) and the Government of Alberta (grant 015486).

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are not publicly available as the steward of the data is the Government of Alberta, which maintains the data for the purpose of health system administration, but a formal data request may be submitted to the Government of Alberta.

Authors' Contributions

Conceptualization: KN (lead), SEM (supporting), LR (supporting), MK (supporting)
Data curation: KN
Formal analysis: KN (lead), LR (equal)
Funding acquisition: SEM
Methodology: KN (lead), LR (supporting), SEM (supporting)
Project administration: SEM (lead), LR (supporting)
Resources: SEM
Supervision: SEM
Visualization: LR (lead), KN (supporting), SEM (supporting)
Writing – original draft: LR
Writing – review & editing: KN (lead), SEM (equal), MK (supporting)

Conflicts of Interest

MK is an active member of the Quebec Immunization Committee. None declared by the other authors.

Multimedia Appendix 1
Supplementary figures and tables.
[PDF File, 816 KB - publichealth_v11i1e79669_app1.pdf]

References

1. Vandelaer J, Olaniran M. Using a school-based approach to deliver immunization—global update. Vaccine (Auckl) 2015 Jan 29;33(5):719-725. [doi: 10.1016/j.vaccine.2014.11.037] [Medline: 25523525]
2. Brotherton JML, Murray SL, Hall MA, et al. Human papillomavirus vaccine coverage among female Australian adolescents: success of the school-based approach. Med J Aust 2013 Nov 4;199(9):614-617. [doi: 10.5694/mja13.10272] [Medline: 24182228]
3. Musto R, Siever JE, Johnston JC, Seidel J, Rose MS, McNeil DA. Social equity in human papillomavirus vaccination: a natural experiment in Calgary Canada. BMC Public Health 2013 Jul 9;13:640. [doi: 10.1186/1471-2458-13-640] [Medline: 23837819]
4. Wang J, Ploner A, Sparén P, et al. Mode of HPV vaccination delivery and equity in vaccine uptake: a nationwide cohort study. Prev Med 2019 Mar;120:26-33. [doi: 10.1016/j.ypmed.2018.12.014] [Medline: 30593796]
5. Adedzi KA, Dubé E. School-based immunization programs: an effective strategy for achieving high vaccination rates? Canadian Public Health Association. 2020. URL: https://www.canvax.ca/sites/default/files/2021-03/CPHA_School-BasedImmunizationPrograms_2020_EN.pdf [accessed 2025-08-26]
6. Sell H, Raj Paudel Y, Voaklander D, MacDonald SE. School immunization coverage in adolescents during the COVID-19 pandemic: a retrospective cohort study. Vaccine (Auckl) 2023 Feb 10;41(7):1333-1341. [doi: 10.1016/j.vaccine.2023.01.011] [Medline: 36642632]
7. Mulholland K, Kretsinger K, Wondwossen L, Crowcroft N. Action needed now to prevent further increases in measles and measles deaths in the coming years. The Lancet 2020 Dec;396(10265):1782-1784. [doi: 10.1016/S0140-6736(20)32394-1]
8. The state of the world's children 2023: for every child, vaccination. UNICEF. 2023. URL: https://www.unicef.org/reports/state-worlds-children-2023#SOWC [accessed 2025-08-26]
9. Larson A, Skolnik A, Bhatti A, Mitrovich R. Addressing an urgent global public health need: Strategies to recover routine vaccination during the COVID-19 pandemic. Hum Vaccin Immunother 2022 Jan 31;18(1). [doi: 10.1080/21645515.2021.1975453] [Medline: 1975453]
10. DeSilva MB, Haapala J, Vazquez-Benitez G, et al. COVID-19 and completion of select routine childhood vaccinations. Pediatrics 2025 Apr 1;155(4):e2024068244. [doi: 10.1542/peds.2024-068244] [Medline: 40132643]
11. Yunusa A, Cabral C, Anderson E. The impact of the Covid-19 pandemic on the uptake of routine maternal and infant vaccines globally: A systematic review. PLOS Glob Public Health 2022;2(10):e0000628. [doi: 10.1371/journal.pgph.0000628] [Medline: 36962643]
12. Cunniff L, Alyanak E, Fix A, et al. The impact of the COVID-19 pandemic on vaccination uptake in the United States and strategies to recover and improve vaccination rates: A review. Hum Vaccin Immunother 2023 Aug 1;19(2):2246502. [doi: 10.1080/21645515.2023.2246502] [Medline: 37671468]

XSL•FO
RenderX

13. Eiden AL, DiFranzo A, Bhatti A, et al. Changes in vaccine administration trends across the life-course during the COVID-19 pandemic in the United States: a claims database study. Expert Rev Vaccines 2023;22(1):481-494. [doi: 10.1080/14760584.2023.2217257] [Medline: 37218717]
14. Ji C, Piché-Renaud PP, Apajee J, et al. Impact of the COVID-19 pandemic on routine immunization coverage in children under 2 years old in Ontario, Canada: A retrospective cohort study. Vaccine (Auckl) 2022 Mar 15;40(12):1790-1798. [doi: 10.1016/j.vaccine.2022.02.008] [Medline: 35164987]
15. Jiter N, Dykstra A, Loper O, Kuramoto S, Schultz J. Tri-State evaluation of the effects of the COVID-19 pandemic on routine vaccine uptake in Iowa, Minnesota, and North Dakota, 2017-2021. J Community Health 2024 Dec;49(6):977-992. [doi: 10.1007/s10900-024-01355-4] [Medline: 38581625]
16. Gallant AJ, Steenbeek A, Halperin SA, Parsons Leigh J, Curran JA. Identifying and addressing the impacts of the COVID-19 pandemic on school-based immunisation programmes in the Canadian Maritimes: a mixed methods study protocol. BMJ Open 2023 Jun 26;13(6):e073172. [doi: 10.1136/bmjopen-2023-073172] [Medline: 37369397]
17. Dionne M, Sauvageau C, Kiely M, et al. School-based vaccination program against HPV and Hepatitis B: A longitudinal analysis of vaccine coverage between 2015 and 2021 in Quebec. Vaccine (Auckl) 2024 Jan 1;42(1):17-23. [doi: 10.1016/j.vaccine.2023.11.055] [Medline: 38044245]
18. Jeevakanthan A, Roubos S, Hong C, et al. Routine vaccination coverage at ages 2 and 7, before, during, and after the COVID-19 pandemic: Results from the STARVAX surveillance system. Can J Public Health 2025 Apr;116(2):284-289. [doi: 10.17269/s41997-024-00956-9] [Medline: 39461928]
19. Immunization and routine immunization schedule. Government of Alberta. 2025. URL: https://www.alberta.ca/immunization-routine-schedule [accessed 2025-03-31]
20. Checklist for cohort studies STROBE. Strengthening the reporting of observational studies in epidemiology (STROBE). 2007. URL: https://www.strobe-statement.org/checklists [accessed 2025-01-23]
21. MacDonald SE, Russell ML, Liu XC, et al. Are we speaking the same language? An argument for the consistent use of terminology and definitions for childhood vaccination indicators. Hum Vaccin Immunother 2019;15(3):740-747. [doi: 10.1080/21645515.2018.1546526] [Medline: 30457475]
22. Previl S. Whooping cough cases are surging in canada. What's going on? Global News. 2024. URL: https://globalnews.ca/news/10730538/whooping-cough-surge-canada-causes-vaccine-rates [accessed 2025-04-28]
23. Public Health Agency of Canada. Measles and rubella weekly monitoring report — canada.ca. 2025 URL: https://health-infobase.canada.ca/measles-rubella [accessed 2025-04-28]
24. Hansen PR, Schmidtblaicher M, Brewer NT. Resilience of HPV vaccine uptake in Denmark: Decline and recovery. Vaccine (Auckl) 2020 Feb 11;38(7):1842-1848. [doi: 10.1016/j.vaccine.2019.12.019] [Medline: 31918860]
25. Vaccination coverage goals and vaccine preventable disease reduction targets by 2025. Public Health Agency of Canada. 2022. URL: https://www.canada.ca/en/public-health/services/immunization-vaccine-priorities/national-immunization-strategy/vaccination-coverage-goals-vaccine-preventable-diseases-reduction-targets-2025.html [accessed 2022-09-04]
26. Sell H, Assi A, Driedger SM, et al. Continuity of routine immunization programs in Canada during the COVID-19 pandemic. Vaccine (Auckl) 2021 Sep 15;39(39):5532-5537. [doi: 10.1016/j.vaccine.2021.08.044] [Medline: 34426028]
27. MacDonald SE, Paudel YR, Kiely M, et al. Impact of the COVID-19 pandemic on vaccine coverage for early childhood vaccines in Alberta, Canada: a population-based retrospective cohort study. BMJ Open 2022 Jan 25;12(1):e055968. [doi: 10.1136/bmjopen-2021-055968] [Medline: 35078849]
28. Iusitini L, Pacheco G, Schober T. Assessing the impact of the COVID-19 pandemic on childhood vaccine uptake with administrative data. SSM Popul Health 2024 Jun;26:101657. [doi: 10.1016/j.ssmph.2024.101657] [Medline: 38596363]
29. Kujawski SA, Yao L, Wang HE, Carias C, Chen YT. Impact of the COVID-19 pandemic on pediatric and adolescent vaccinations and well child visits in the United States: A database analysis. Vaccine (Auckl) 2022 Jan 31;40(5):706-713. [doi: 10.1016/j.vaccine.2021.12.064] [Medline: 35012776]
30. MacDonald NE, Comeau JL, Dubé È, Bucci LM. COVID-19 and missed routine immunizations: designing for effective catch-up in Canada. Can J Public Health 2020 Aug;111(4):469-472. [doi: 10.17269/s41997-020-00385-4] [Medline: 32761546]
31. Su Z, Cheshmehzangi A, McDonnell D, da Veiga CP, Xiang YT. Mind the "vaccine fatigue". Front Immunol 2022;13:839433. [doi: 10.3389/fimmu.2022.839433] [Medline: 35359948]

Abbreviations

DTap-IPV: diphtheria, tetanus, acellular pertussis, and polio
HPV: human papillomavirus
MenC-ACYW: meningococcal conjugate ACYW-135
MMR: measles-mumps-rubella
MMRV: measles-mumps-rubella—varicella
SBIP: school-based immunization program
Tdap: tetanus, diphtheria, and acellular pertussis

XSL•FO
RenderX

Edited by A Mavragani; submitted 26.06.25; peer-reviewed by K Li, T Schober; revised version received 05.08.25; accepted 06.08.25; published 23.09.25.

Please cite as:
Nartey K, Reifferscheid L, Kiely M, MacDonald SE, Canadian Immunization Research Network
Catch-Up of Routine School-Based Immunizations Since the COVID-19 Pandemic: 4-Year Observational Cohort Study
JMIR Public Health Surveill 2025;11:e79669
URL: https://publichealth.jmir.org/2025/1/e79669
doi:10.2196/79669

© Kelvin Nartey, Laura Reifferscheid, Marilou Kiely, Shannon E MacDonald, Canadian Immunization Research Network. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Association Between Posting WeChat Moments and the Risk of Depressive Symptoms Among Middle-Aged and Older Chinese Adults: Prospective National Cohort Study

Wei Wang[1*], MMed; Hui Wang[2*], BM; Xinru Hu[3], BSc; Qian Yu[4], MSc; Fangyi Chen[3], BM; Xirui Qiu[5], BM; Xiaoxiao Wang[6], PhD

[1]GCP Center, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), Nanjing, China

[2]Public Health Service Center of Nanjing Jiangbei New District, Nanjing, China

[3]Department of Public Health, School of Medicine and Holistic Integrative Medicine, Nanjing University of Chinese Medicine, Nanjing, China

[4]Department of Science and Technology, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), Nanjing, China

[5]Department of Respiratory and Critical Medicine, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), Nanjing, China

[6]Clinical Research Institute, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), 155 Hanzhong Road, Nanjing, China

[*]these authors contributed equally

Corresponding Author:
Xiaoxiao Wang, PhD
Clinical Research Institute, Affiliated Hospital of Nanjing University of Chinese Medicine (Jiangsu Province Hospital of Chinese Medicine), 155 Hanzhong Road, Nanjing, China

Abstract

Background: The association between social media usage and the risk of depressive symptoms has attracted increasing attention. WeChat is a popular social media software in China. The impact of using WeChat and posting WeChat moments on the risk of developing depressive symptoms among community-based middle-aged and older adults in China is unknown.

Objective: The objective was to assess the association between using WeChat and posting WeChat moments and the risk of depressive symptoms among middle-aged and older adults in China.

Methods: A prospective national cohort study was designed based on the data obtained from the fourth and fifth waves of the China Health and Retirement Longitudinal Study (CHARLS). The strength of association between using WeChat and posting WeChat moments and the risk of depressive symptoms was estimated by modified Poisson regressions. Depressive symptoms were determined using the 10-item Center for Epidemiologic Studies Depression Scale. Meanwhile, the heterogeneity of the associations was explored through multiple subgroup analyses. Moreover, multiple sensitivity analyses were performed to verify the robustness of the associations between the exposures and depressive symptoms.

Results: A total of 9670 eligible participants were included in the cohort study, and the incidence rate of depressive symptoms was 19.08% (1845/9670, 95% CI 19.07% - 19.09%) from the fourth to fifth waves of the CHARLS. Using WeChat (adjusted relative risk [aRR] 0.691, 95% CI 0.582 - 0.520) and posting WeChat moments (aRR 0.673, 95% CI 0.552 - 0.821) reduced the risk of depressive symptoms among middle-aged and older Chinese adults. The association between the exposures and depressive symptoms was robust, proved through multiple sensitivity analyses (all P<.05). However, the associations were heterogeneous in certain subgroup catagories, such as solitude, duration of sleep at night, nap after lunch, physical activity, and having multiple chronic conditions.

Conclusions: Using WeChat and especially posting WeChat moments can mitigate the risk of depressive symptoms among community-based middle-aged and older Chinese adults. However, there is likely a need for a longer follow-up period to explore the impact of the exposures on the risk of long-term depressive outcomes.

(JMIR Public Health Surveill 2025;11:e62730) doi:10.2196/62730

KEYWORDS

cohort study; depression; depressive symptoms; mental health; middle-aged adults; modified Poisson regression; older adults; WeChat

XSL•FO
RenderX

Introduction

Depression is a common emotional disorder [1-4] resulting in limited psychosocial functioning and a decreased quality of life [5] and is also an independent risk factor of all-cause mortality and other chronic conditions [6-8]. A meta-analysis [3] reported that the prevalence of depressive symptoms among older adults in mainland China was 20% (95% CI 17.5% - 22.8%), and Chen et al [9] reported an unweighted incidence rate of depressive symptoms of 34.77% based on the China Health and Retirement Longitudinal Study (CHARLS). These findings highlight the need for an increased attention to depressive symptoms in China. Therefore, understanding modifiable factors associated with depressive symptoms is essential for developing effective public health strategies to prevent them. Multiple well-known factors associated with depressive symptoms include genetic factors, psychosocial stress, changes in the hypothalamic-pituitary-adrenal axis, inflammation, and others [5,10,11].

In terms of psychosocial factors, there has been a growing body of studies on the association or correlation between social media usage (SMU) and depression in recent years. Studies conducted by Lin et al [12], Primack et al [13] and Perlis et al [14] demonstrated that SMU, particularly on Facebook, TikTok (known as Douyin in China) and Snapchat, was associated with an increased risk of depression among US adults. However, Cotten et al [15] analyzed the data from 6 waves of the National Health and Aging Trends Study and suggested that SMU may not be related to depression in older adults. Based on the above studies, the association between SMU and depression could be attributed to problematic SMU [16,17], while moderate and reasonable use of social media may not cause harm to mental health.

WeChat (known as Weixin in China) is a popular social media software in China, which had 1.04 billion monthly active users according to Tencent [18]. The primary functions of WeChat include texting, voice and video calling, and browsing and posting moments that encompass both positive and negative events happening among friends or oneself, as well as facilitating financial management. From the perspective of these functionalities, WeChat distinguishes itself from platforms like Facebook or TikTok, which focus on interests and preferences that can foster addiction or problematic SMU.

It is hypothesized that using WeChat and posting WeChat moments would not result in an increased risk of depressive symptoms among community-based middle-aged and older Chinese adults, but rather potentially mitigate the risk of developing depressive symptoms. In fact, two cross-sectional studies using data from the fourth wave of the CHARLS and the fifth wave of the China Family Panel Study reported an association between WeChat usage and a decreased risk of depressive symptoms among middle-aged and older Chinese adults [19,20]. However, the two studies were unable to determine a temporal relationship between WeChat usage and depressive symptoms, thus precluding any causal inferences regarding the relationship between the observed exposures and depressive symptoms. Additionally, findings from the study conducted by Qu et al [19] indicated that participants who posted WeChat moments did not exhibit a decreased risk of depressive symptoms, which contradicted the findings of Zhang and Liang [20].

In response to these conflicting findings, a national, prospective cohort study was designed, leveraging data from both the fourth and fifth waves of the CHARLS. The aim of this study was to elucidate the relationship between the use of WeChat—with particular emphasis on posting WeChat moments—and the risk of depressive symptoms.

Methods

Study Design and Study Population

The study was characterized as a national, prospective cohort study, deriving its data from the fourth and fifth waves of the CHARLS. Initiated in 2011, the CHARLS represents a longitudinal national survey aimed at collecting extensive data pertinent to community-based middle-aged and older adults in China. The scope of the CHARLS encompasses high-quality, representative panel data covering a broad spectrum of topics. These topics include demographic variables, socioeconomic status, familial relationships, health status, health care usage, and others. More information about the CHARLS is accessible at their official website [21].

The intersection of the following 3 datasets from the CHARLS were compiled for the raw dataset: demographic backgrounds from the fourth wave, health status and functioning from the fourth wave, and the 10-item Center for Epidemiologic Studies Depression Scale (CESD-10) from the fifth wave. Participants from the raw dataset that (1) were aged <45 years, (2) were missing scores for the CESD-10 in the fourth wave, (3) had a CESD-10 score ≥10 in the fourth wave, or (4) were missing scores for the CESD-10 in the fifth wave were excluded.

Assessment of Exposures

In the fourth wave of the CHARLS, participants were asked, “Do you use WeChat?” with possible responses being yes or no. Participants who answered yes were categorized as WeChat users; otherwise, they were categorized as non-WeChat users. Subsequently, the participants using WeChat were further asked, “Do you post WeChat moments?” with possible responses being yes or no. Participants who answered yes were categorized as WeChat moments users; otherwise, they were categorized as non-WeChat moments users.

Therefore, according to whether or not they used WeChat, the participants were stratified into 2 groups: WeChat users and non-WeChat users. Furthermore, according to whether or not the participant posted WeChat moments, they were stratified into 3 groups: non-WeChat users, non-WeChat moments users and WeChat moments users. In this study, we first focused on the comparison between the incidence of depressive symptoms among the WeChat users and non-WeChat users, then that among non-WeChat users, non-WeChat moments users and WeChat moments users.

XSL•FO
RenderX

Assessment of the Outcome

The outcome of this study was depressive symptoms, which were measured using the CESD-10 scores from the fifth wave of the CHARLS as a dichotomous variable (yes or no). Each item of the CESD-10 was rated on a 4-point Likert scale ranging from 0 to 3, corresponding to the frequency of the symptoms (0=rarely or not at all; 1=some or a little of the time; 2=occasionally or a moderate amount of time; 3=most or all of the time). The aggregate score, which was the sum of the responses to all items, thus ranged from 0 to 30. A cumulative score of 10 or above was used as a threshold for identifying participants with depressive symptoms, whereas scores below this threshold indicated the absence of such symptoms. Additionally, for sensitivity analysis, alternative cutoff points of 11 and 12 on the CESD-10 were also employed to further ascertain the presence of depressive symptoms.

Assessment of Covariates

The following characteristics obtained from the fourth wave of the CHARLS were considered as covariates: age, sex (male or female), race (Han or minority), faith (yes or no), type of community (village or city/town), marital status (married and cohabiting, married and separated, or others), education (illiterate, not finished primary school/home school/primary school, junior high school, or senior high school and above), smoking habits (currently, ever, or never), alcohol consumption (yes or no), duration of sleep at night (<6 hours, ≥6 hours to <8 hours, or ≥8 hours), nap after lunch (yes or no), and physical activity (none, mild, moderate, or vigorous). Additionally, the social activity score was a covariate [22], which was the frequency of 11 kinds of activities rated as never (score=0), not regularly (score=1), almost every week (score=2), or almost daily (score=3). These activities were assembled to a sum score based on the frequency level, and the total scores for social activities could range from 0-33 points.

The basic activities of daily living (BADL) score was also a covariate [23], which included items for dressing, bathing, feeding, moving from a bed to a chair, using the toilet, and maintaining continence. All items had 4 potential answers, including (1) do it without difficulty, (2) do it but with difficulty, (3) do it with difficulty and need help, or (4) cannot do it, which were rated as 0, 1, 2, or 3 points, respectively. BADL scores were the sum of points for all items. Another covariate was life satisfaction. This question had 5 alternative answers: completely satisfied, very satisfied, somewhat satisfied, not very satisfied, or not at all satisfied. The responses were then stratified into 3 categories: completely/very satisfied, somewhat satisfied, or not very/not at all satisfied. Last, the number of chronic conditions (5 categories: no chronic conditions, 1 chronic condition, 2 chronic conditions, 3 chronic conditions, and ≥4 chronic conditions) and the number of parts with body pain were covariates.

Statistical Methods

Descriptive analyses and differential comparisons were conducted on the characteristics of the eligible participants. The quantitative variables with a normal distribution were described using mean and SD, and a 2-tailed *t* test or one-way ANOVA was used for comparing the difference between two groups or among three groups, respectively. Quantitative variables with skewed distributions were described using the median and IQR, and a Wilcoxon rank sum test or Kruskal-Wallis H test was used for the comparison. Disordered multinomial variables were described by the number of cases and percentage, and a χ^2 test or Fisher exact test was used for the comparison. Ordinal multinomial variables were described by the frequency and percentage, and a Wilcoxon rank sum test or Kruskal-Wallis H test was used for comparison.

For all eligible participants, univariable modified Poisson regressions were used to analyze the association between using WeChat and posting moments with the risk of depressive symptoms. Next, variables with $P<.05$ from between-group comparisons, along with adjusted variables in previous studies [19,20], were included as confounders in the multivariable modified Poisson regressions to analyze the association between the exposures and depressive symptoms. Meanwhile, the generalized variance-inflation factor of covariates in the multivariable regressions were calculated to test for collinearity (Table S1 in Multimedia Appendix 1).

For subgroup analyses, all eligible participants were divided into different subgroups according to sex (male or female), age (<60 years old or ≥60 years old), type of community (village or city/town), solitude (living with a heterosexual partner or not), education (primary school and below or junior high school and above), alcohol consumption (yes or no), duration of sleep at night (<6 hours, 6 - 8 hours, or ≥8 hours), nap after lunch (yes or no), physical activity (vigorous, moderate, mild, or none), and chronic conditions (≥2 chronic conditions or not). In the same way, multivariable modified Poisson regressions with the confounders, except the grouping variable, were performed to estimate the strength of association between the exposures and depressive symptoms.

There were missing values in variables for smoking habits, alcohol consumption, duration of sleep at night, nap after lunch, physical activity, social activity score, BADL score, and the number of body parts with pain, of which the proportions of missing values are shown in the Table S2 in Multimedia Appendix 1. Hot deck imputation was used to impute the missing values, and all statistical analyses were based on that data. Smoking habits, social activity score, BADL score, and the number of body parts with pain were not considered as confounders in the main analysis, due to the proportion of missing values being >40%. However, the above variables were adjusted for in the sensitivity analyses.

Multiple sensitivity analyses were conducted to validate the robustness of the association between the exposures and the outcome. First, the outcome of depressive symptoms was redefined as a CEDS-10 score of ≥11 or 12 in the multivariable modified Poisson regressions. Second, eligible participants diagnosed with emotional, nervous, or psychiatric problems in the fourth wave of the CHARLS were excluded. Third, to improve the comparability of the exposed groups with the control group, eligible participants aged >83 years in the non-WeChat users group were excluded since the oldest user in the WeChat users group was 83 years old (Table S3 in

XSL•FO
RenderX

Multimedia Appendix 1). Fourth, additional confounders (smoking habits, BADL score, social activity score, and the number of body parts with pain) were adjusted for in the main analysis. Last, multivariable binary logistic regressions were performed to estimate the association between exposures and depressive symptoms.

All statistical analyses were completed by R 4.3.3 (R Core Team) with R packages *VIM*, *rqlm*, *car* and *forestploter* and the Mengte Cloud version 1.0 statistics platform (Wuxi Mengte Yi Shu Tong Co, Ltd). A 2-sided hypothesis test was used, and the significance level was set at α=.05.

Ethical Considerations

The ethical approval for the CHARLS was approved by the Ethics Review Committee of Peking University (IRB00001052-11015). Prior to participation, all study participants were required to sign an informed consent form, thereby confirming their voluntary engagement in the survey process. The study data were anonymous. China's Health Commission, Ministry of Science and Technology, and other departments jointly issued the Ethical Review Measures for Life Science and Medical Research involving Human Beings in February 2023, in which Article 32 mentioned 4 situations of exemption from ethical review. The secondary analysis of this public database CHARLS satisfied the first 2 of these: the data were free and publicly available and were available through a web-based application. Importantly, the data obtained from the platform were anonymous, and the identity information of all respondents was unknown. In addition, the study did not use biological samples from the survey subjects. Regarding compensation and informed consent, because this study was a secondary analysis, along with the subjects being anonymous, the respondents did not receive secondary compensation or sign a new informed consent form.

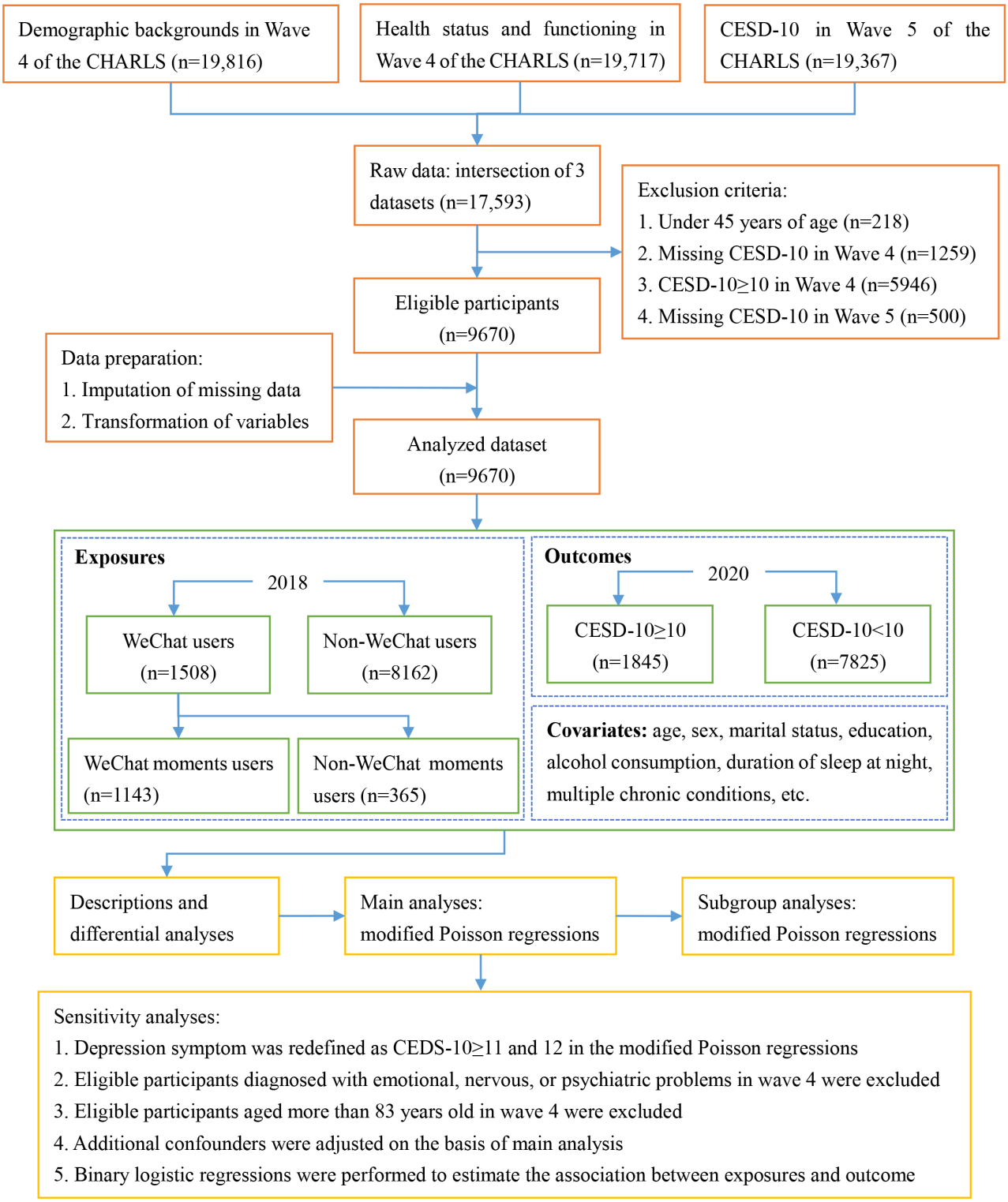
Results

The Characteristics of Eligible Participants

The process for screening participants is shown in Figure 1. A total of 9670 eligible participants were included in the final analysis. The mean age of these participants was 61.01 (SD 9.35) years. Of all eligible participants, 47.55% (n=4598) were female. A majority (n=6733, 69.63%) resided in villages, and 82.99% (n=8025) reported being married or cohabiting. Alcohol consumption was reported by 38.73% (n=3745) of participants, and 62.79% (n=6072) indicated they took a nap after lunch. A total of 92.84% (n=8978) engaged in physical activity, and 74.96% (n=7249) had been diagnosed with chronic conditions. The median scores of the CESD-10 were 4 (IQR 2-7) in the fourth wave and 6 (IQR 3-9) in the fifth wave. The incidence rate of depressive symptoms was 19.08% (95% CI 19.07% - 19.09%) from the fourth to fifth waves of the CHARLS (Table 1). Participants who used WeChat, regardless of whether they posted on WeChat moments, tended to be younger, male, non-religious, urban residents, or have a higher level of education. They were also more likely to be married or cohabiting, be current or former smokers, be alcohol consumers, sleep less than 8 hours at night, take afternoon naps, engage in physical activity, have high social activity scores, have low BADL scores, have fewer body parts in pain, and have no chronic conditions (Table 1 and S4 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 1. The flowchart for screening eligible participants, data governance, and statistical analysis. CESD-10: 10-item Center for Epidemiologic Studies Depression Scale; CHARLS: China Health and Retirement Longitudinal Study.



XSL•FO
RenderX

Table . The baseline characteristics and incidence of depressive symptoms of eligible participants in the cohort study based on the China Health and Retirement Longitudinal Study.

Characteristics	Total (n=9670)	Non-WeChat users (n=8162)	WeChat users (n=1508)		*P* value
			Non-WeChat moments users (n=365)	WeChat moments users (n=1143)	
Age (years), mean (SD)	61.01 (9.35)	62.05 (9.32)	55.68 (7.71)	55.28 (7.14)	<.001
Sex, n (%)					<.001
Male	5072 (52.45)	4204 (51.51)	204 (55.89)	664 (58.09)	
Female	4598 (47.55)	3958 (48.49)	161 (44.11)	479 (41.91)	
Race, n (%)					.48
Han	8990 (92.97)	7578 (92.84)	344 (94.25)	1068 (93.44)	
Minority	680 (7.03)	584 (7.16)	21 (5.75)	75 (6.56)	
Faith, n (%)					.02
Yes	946 (9.78)	821 (10.06)	40 (10.96)	85 (7.44)	
No	8724 (90.22)	7341 (89.94)	325 (89.04)	1058 (92.56)	
Type of community, n (%)					<.001
Village	6733 (69.63)	6085 (74.55)	176 (48.22)	472 (41.29)	
City/Town	2937 (30.37)	2077 (25.45)	189 (51.78)	671 (58.71)	
Education, n (%)					<.001
Illiterate	1721 (17.80)	1704 (20.88)	12 (3.29)	5 (0.44)	
Not finish primary school/home school/primary school	4040 (41.78)	3713 (45.49)	100 (27.40)	227 (19.86)	
Junior high school	2453 (25.37)	1903 (23.32)	133 (36.44)	417 (36.48)	
Senior high school and above	1456 (15.06)	842 (10.32)	120 (32.88)	494 (43.22)	
Marital status, n (%)					<.001
Married and cohabiting	8025 (82.99)	6727 (82.42)	307 (84.11)	991 (86.70)	
Married and separated	574 (5.94)	454 (5.56)	32 (8.77)	88 (7.70)	
Others	1071 (11.08)	981 (12.02)	26 (7.12)	64 (5.60)	
Smoking habits, n (%)					.006
Currently	333 (3.44)	271 (3.32)	8 (2.19)	54 (4.72)	
Ever	394 (4.07)	319 (3.91)	13 (3.56)	62 (5.42)	
Never	8943 (92.48)	7572 (92.77)	344 (94.25)	1027 (89.85)	
Alcohol consumption, n (%)					<.001
No	5925 (61.27)	5237 (64.16)	194 (53.15)	494 (43.22)	
Yes	3745 (38.73)	2925 (35.84)	171 (46.85)	649 (56.78)	
Duration of sleep at night, n (%)					<.001
<6 hours	2573 (26.61)	2229 (27.31)	88 (24.11)	256 (22.40)	
≥6 hours to <8 hours	4282 (44.28)	3456 (42.34)	190 (52.05)	636 (55.64)	
≥8 hours	2815 (29.11)	2477 (30.35)	87 (23.84)	251 (21.96)	
Nap after lunch, n (%)					<.001
No	3598 (37.21)	3129 (38.34)	131 (35.89)	338 (29.57)	
Yes	6072 (62.79)	5033 (61.66)	234 (64.11)	805 (70.43)	

XSL•FO
RenderX

Characteristics	Total (n=9670)	Non-WeChat users (n=8162)	WeChat users (n=1508)		*P* value
			Non-WeChat moments users (n=365)	WeChat moments users (n=1143)	
Physical activity, n (%)					<.001
No	692 (7.16)	657 (8.05)	13 (3.56)	22 (1.92)	
Mild	2741 (28.35)	2321 (28.44)	96 (26.30)	324 (28.35)	
Moderate	3072 (31.77)	2462 (30.16)	143 (39.18)	467 (40.86)	
Vigorous	3165 (32.73)	2722 (33.35)	113 (30.96)	330 (28.87)	
Social activity score, median (IQR)	3 (2-5)	3 (2-4)	5 (3-6)	5 (3-8)	<.001
BADL[a] score, mean (SD)	6.32 (1.02)	6.34 (1.04)	6.23 (0.71)	6.25 (0.90)	.006
Life satisfaction, n (%)					<.001
Completely/very	4010 (41.47)	3508 (42.98)	135 (36.99)	367 (32.11)	
Somewhat	5262 (54.42)	4319 (52.92)	211 (57.81)	732 (64.04)	
Not very/not at all	398 (4.12)	335 (4.10)	19 (5.21)	44 (3.85)	
Number of chronic conditions, n (%)					<.001
0	2421 (25.04)	1960 (24.01)	113 (30.96)	348 (30.45)	
1	2574 (26.62)	2179 (26.70)	98 (26.85)	297 (25.98)	
2	1971 (20.38)	1699 (20.82)	79 (21.64)	193 (16.89)	
3	1243 (12.85)	1051 (12.88)	45 (12.33)	147 (12.86)	
≥4	1461 (15.11)	1273 (15.60)	30 (8.22)	158 (13.82)	
Number of body parts with pain, median (IQR)	3 (2-6)	3 (2-6)	3 (1-5)	3 (1-5)	.007
CESD-10[b] in 2020, median (IQR)	6 (3-9)	6 (3-10)	4 (2-7)	4 (1-7)	<.001
Depressive symptoms, n (%)					<.001
No	7825 (80.92)	6457 (79.11)	326 (89.32)	1042 (91.16)	
Yes	1845 (19.08)	1705 (20.89)	39 (10.68)	101 (8.84)	

[a]BADL: basic activities of daily living.

[b]CESD-10: 10-item Center for Epidemiologic Studies Depression Scale.

Association Between Exposures and Depressive Symptoms

As shown in Figure 2 and Table S5 in Multimedia Appendix 1, among all eligible participants, the risk of depressive symptoms in the WeChat users group was 0.556 times lower than that in the non-WeChat users group, as indicated by the crude relative risk (cRR) of 0.444 (95% CI 0.378 - 0.523; *P*<.001) using a univariable modified Poisson regression. Furthermore, after adjusting for potential confounders in the multivariable modified Poisson regression models, the risk of depressive symptoms in the WeChat users group was 0.309 times lower than that in the non-WeChat users group (adjusted relative risk [aRR] 0.691, 95% CI 0.582 - 0.820; *P*<.001). These results suggested a significant association between WeChat usage and a decreased risk of depressive symptoms, implying that WeChat may serve as a protective factor against depressive symptoms in community-based middle-aged and older Chinese adults.

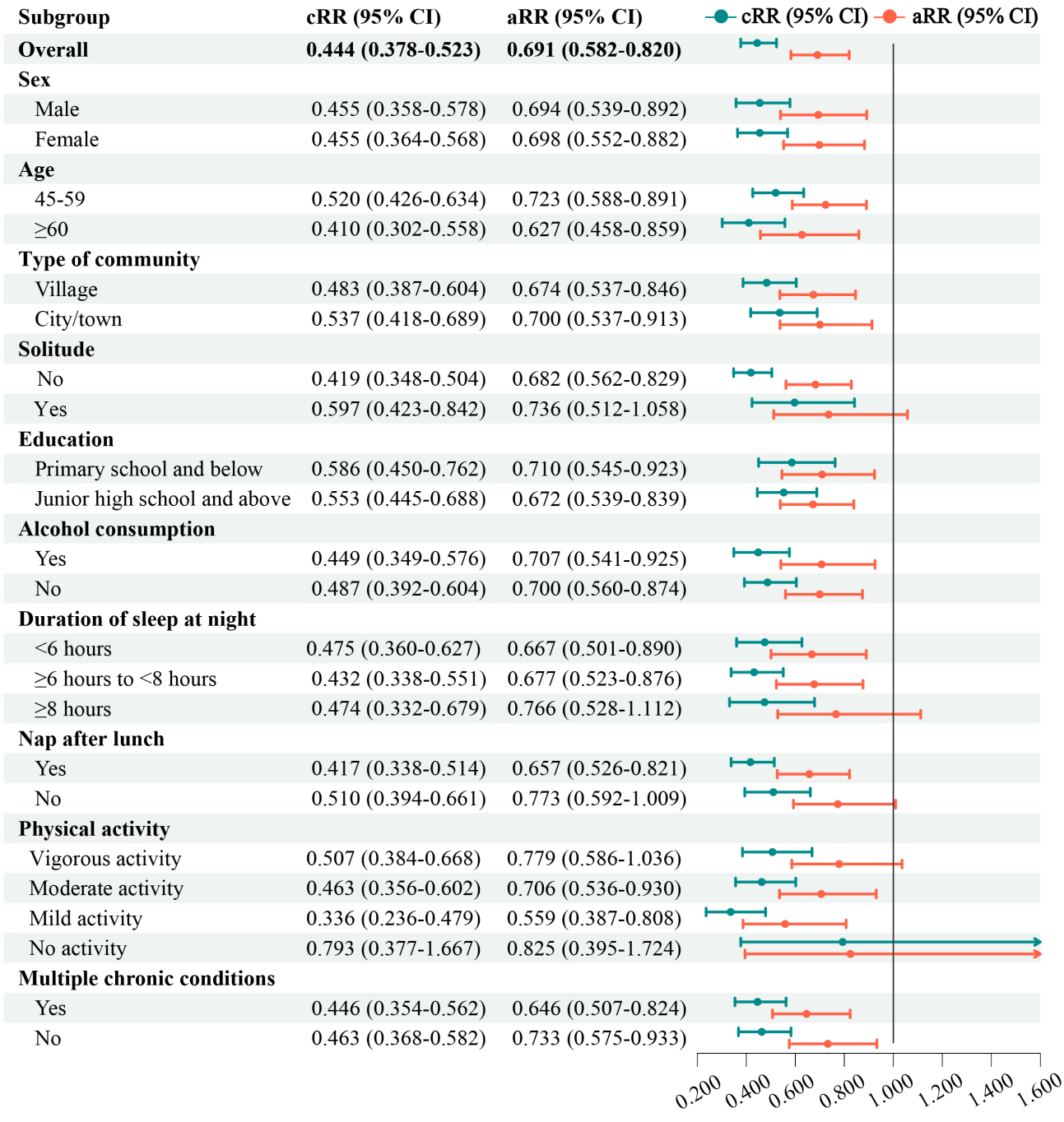
Subgroup analyses showed that WeChat usage did not reduce the risk of depressive symptoms among those who were solitary, had a sleep duration of ≥8 hours at night, did not have nap after lunch, engaged in vigorous physical activity, or did not engage in any physical activity. However, in other subgroups, the result indicated that WeChat usage was associated with a lower risk of depressive symptoms.

As depicted in Figure 3 and Table S6 in Multimedia Appendix 1, compared to non-WeChat users, both the non-WeChat moments users (aRR 0.739, 95% CI 0.550 - 0.992; *P*=.044) and WeChat moments users (aRR 0.673, 95% CI 0.552 - 0.821; *P*<.001) exhibited a lower risk of depressive symptoms. In the subgroups of females (aRR 0.655, 95% CI 0.432 - 0.996) and those without multiple chronic conditions (aRR 0.589, 95% CI

XSL•FO
RenderX

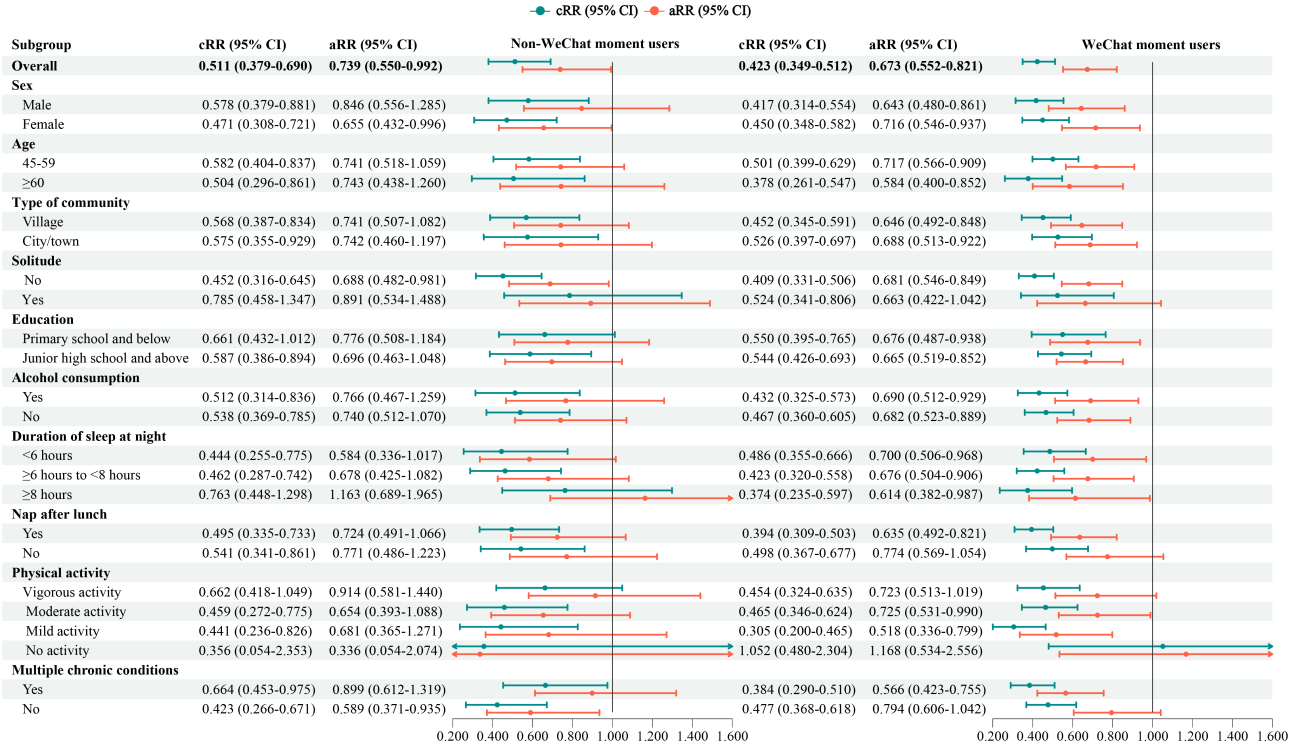
0.371–0.935), non-WeChat moments users had a lower risk of depressive symptoms compared to non-WeChat users. Additionally, with the exception of individuals who did not live with a spouse or opposite sex, did not nap after lunch, engaged in vigorous physical activity, did not engage in any physical activity, or did not have multiple chronic conditions, results from the subgroup analyses suggested that posting WeChat moments was associated with a reduced risk of depressive symptoms.

Figure 2. The association between using WeChat and risk of depressive symptoms among middle-aged and older Chinese adults in the cohort study based on the CHARLS. aRR: adjusted relative risk; cRR: crude relative risk.



XSL•FO
RenderX

Figure 3. The association between using WeChat with or without posting WeChat moments and risk of depressive symptoms among middle-aged and older Chinese adults in the cohort study based on CHARLS. aRR: adjusted relative risk; cRR: crude relative risk.



Sensitivity Analyses

The results of sensitivity analyses are presented in Table 2. (1) When depressive symptoms were redefined as CESD-10 scores of ≥11 or 12, the multivariable Poisson regressions indicated that WeChat usage reduced the risk of depressive symptoms in middle-aged and older individuals, regardless of whether they posted WeChat moments. (2) The multivariable Poisson regression was performed again after excluding participants who had emotional, nervous, or psychiatric problems in the fourth wave of the CHARLS, and results were consistent with the main analysis. (3) Given that none of the WeChat users were over 83 years old, another multivariable Poisson regression was performed after excluding non-WeChat users over 83 years old, which yielded results similar to the main analysis. (4) Additional adjustments were made for smoking habits, social activity score, BADL score, and the number of body parts with pain. The outcome of depressive symptoms remained consistent with those of the main analysis. (5) Odds ratios from multivariable binary logistic regressions were used to estimate the associations, and the results demonstrated the robustness of the main analysis.

XSL•FO
RenderX

Table . Multiple sensitivity analyses to determine the association between exposures and depressive symptoms that were identified by the main analysis.

Sensitivity analysis and variables	β[a] (SE)	aRR[b]/aOR[c] (95% CI)	*P* value
Depressive symptoms was defined as CESD-10[d] of ≥11 in the modified Poisson regressions			
Use WeChat (vs no)			
Yes	–0.370 (0.087)	0.691 (0.582 - 0.820)	<.001
Use WeChat and post moments (vs do not use WeChat)			
Use WeChat and not post moments	–0.303 (0.151)	0.738 (0.549 - 0.992)	.04
Use WeChat and post moments	–0.396 (0.101)	0.673 (0.552 - 0.821)	<.001
Depressive symptoms was defined as CESD-10 of ≥12 in the modified Poisson regressions			
Use WeChat (vs no)			
Yes	–0.402 (0.100)	0.669 (0.550 - 0.813)	<.001
Use WeChat and post moments (do not use WeChat)			
Use WeChat and not post moments	–0.267 (0.168)	0.766 (0.551 - 1.064)	.11
Use WeChat and post moments	–0.457 (0.117)	0.633 (0.504 - 0.796)	<.001
Participants diagnosed with emotional, nervous, or psychiatric problems in Wave 4 were excluded			
Use WeChat (vs no)			
Yes	–0.399 (0.089)	0.671 (0.564 - 0.799)	<.001
Use WeChat and post moments (do not use WeChat)			
Use WeChat and not post moments	–0.325 (0.153)	0.723 (0.536 - 0.975)	.03
Use WeChat and post moments	–0.428 (0.103)	0.652 (0.532 - 0.799)	<.001
Participants aged more than 83 years in Wave 4 were excluded			
Use WeChat (vs no)			
Yes	–0.363 (0.088)	0.696 (0.586 - 0.826)	<.001
Use WeChat and post moments (vs do not use WeChat)			
Use WeChat and not post moments	–0.298 (0.151)	0.742 (0.552 - 0.998)	.048
Use WeChat and post moments	–0.388 (0.101)	0.678 (0.556 - 0.827)	<.001
Smoking habits, social activity score, BADL[e] score, and the number of body parts with pain were adjusted for the main analysis			
Use WeChat (vs no)			
Yes	–0.376 (0.090)	0.687 (0.576 - 0.819)	<.001
Use WeChat and post moments (vs do not use WeChat)			
Use WeChat and not post moments	–0.306 (0.152)	0.736 (0.546 - 0.992)	.045
Use WeChat and post moments	–0.403 (0.103)	0.668 (0.546 - 0.818)	<.001
Multivariable logistic regressions for analyzing the association between exposures and depressive symptoms			
Use WeChat (vs no)			
Yes	–0.431 (0.102)	0.650 (0.532 - 0.794)	<.001
Use WeChat and post moments (vs do not use WeChat)			
Use WeChat and not post moments	–0.361 (0.178)	0.697 (0.492 - 0.988)	.04
Use WeChat and post moments	–0.457 (0.117)	0.633 (0.504 - 0.797)	<.001

[a]β: regression coefficient.

[b]aRR: adjusted relative risk.

[c]aOR: adjusted odds ratio.

[d]CESD-10: 10-item Center for Epidemiologic Studies Depression Scale.

[e]BADL: basic activities of daily living.

XSL•FO
RenderX

Discussion

Principal Findings

This prospective national cohort study confirmed that using WeChat and posting WeChat moments was a protective factor for depressive symptoms among community based middle-aged and older adults in China. In terms of the absolute magnitude of the aRR, participants who post WeChat moments had a lower risk of depressive symptoms than those who do not post WeChat moments. The results of multiple sensitivity analysis also illustrated the robustness of the association between the exposures and depressive symptoms. In addition, the associations were heterogeneous across several subgroups. For example, irrespective of posting WeChat moments, the association between WeChat usage and the risk of depression was not significant among participants whose duration of sleep at night was more than or equal to 8 hours, while the association was statistically significant in the other two subgroups for the duration of sleep. Also, only in the subgroups where participants engaged in mild and moderate physical activity did using WeChat and posting WeChat moments mitigate the risk of depressive symptoms.

First, this study demonstrated that WeChat usage mitigated the risk of future depressive symptoms, which was consistent with the conclusions of 3 existing cross-sectional studies carried out by Qu et al [19], Wang et al [24], and Zhang and Liang [20]. Second, the study further suggested that posting WeChat moments was significantly associated with a decreased risk of depressive symptoms, and that was consistent with the finding of the cross-sectional study conducted by Zhang and Liang [20], but it was inconsistent with the finding of the cross-sectional study conducted by Qu et al [19]. The results from this prospective cohort study are more convincing because it meets a necessary and inarguable criterion for causal inference: the temporality of cause and effect [25]. Last, multiple subgroup analyses were performed to explore the heterogeneity of the associations for the different characteristics of the participants, which were not conducted in previous studies about WeChat usage and depression.

Many studies have shown that loneliness, social isolation, lack of family and social support, and lack of social activity are risk factors for anxiety and depression in adults [26-30]. On the one hand, WeChat is an instant message platform where WeChat users can communicate with family and friends to relieve the feeling of loneliness, especially for older adults whose children or relatives are not around them [31]. Importantly, a video call on WeChat allows users to see each other, which is more convenient for emotional expression and perception and users can obtain emotional support. WeChat users not only can maintain and improve existing social relationships through WeChat, but they can also expand their social circle to make more like-minded friends and engage in more social activity [32].

WeChat moments is another common and primary function of WeChat, which is an emotional space composed of words, pictures, and short videos posted by users. Users of WeChat moments can post what is happening around them or express their feelings by text or video, browse the content posted or reproduced by their WeChat friends, and interact with their friends through the act of liking and commenting on posts and short videos. Similarly, users can obtain emotional support and emotional communication, as well as release negative emotions by posting WeChat moments. Therefore, the results of Nan et al [33] were supported, which found that internet usage, including WeChat usage, could directly or indirectly reduce the risk of depression through improving interpersonal relationships. Since solitude, a longer sleep duration, and no physical activity are all risk factors for depressive symptoms, it is difficult to offset the negative impact of these factors by using WeChat or posting WeChat moments. Therefore, the heterogeneity of the associations among certain subgroups was observed. Importantly, the heterogeneity of the associations shown in the subgroup analysis suggested that a one-size-fits-all model should be used with caution, and recommendations for individualization or stratified treatment are necessary, especially when research is used to inform policy and clinicians' decisions [34-36].

Strengths and Limitations

There are several strengths of this study. First, this study is a prospective national cohort study, of which the sample sizes are large and the participants are representative. Second, modified Poisson regressions were used to precisely estimate the strength of the associations, which is a statistical approach more appropriate for this scenario given the high incidence of depressive symptoms. Third, multiple subgroups analyses were performed to explore the heterogeneity of the associations between the exposures and depressive symptoms among participants with different characteristics. Last, multiple sensitivity analyses were performed to validate the robustness of the associations between the exposures and depressive symptoms based on the main analysis.

Meanwhile, there are several limitations of this study. First, as the questions about WeChat usage and WeChat moments usage were only collected in the 2018 and 2020 surveys in the CHARLS database, the data for only these two surveys were used in this study. The corresponding follow-up period in this study was short, and a longer follow-up is needed to further confirm our conclusion. Second, the use of subgroups and more exposure groups (WeChat users were further divided into WeChat moments users and non-WeChat moments users) reduced the sample size and statistical power of the regression analyses. In particular, fewer participants used WeChat but did not post moments, which may have resulted in wide CIs for the associations and heterogeneity of the associations that are not easily explained. Third, some studies have suggested that excessive SMU or problematic SMU may increase the risk of depression [16,37,38]. However, the CHARLS database lacks relevant information on the frequency or duration of WeChat and WeChat moments usage, so this study cannot analyze the association between the frequency of WeChat use and depressive symptoms. Fourth, limited by the information collected in the database, other factors behind the use of WeChat and WeChat moments that may influence depressive symptoms were not considered in this study, such as the personality of participants (ie, outgoing or introverted). These factors can relate to whether participants post moments and affect the risk of depressive

XSL•FO
RenderX

symptoms. Last, WeChat is only one of many social media software, and the participants may also use other social media software in addition to WeChat. Also, limited by the relevant information provided by CHARLS, this study was unable to investigate the effects of other social media and the joint use of them on mental health.

Conclusion

In summary, the prevention of depression in community-based middle-aged and older adults calls for the joint efforts of individuals, families, and society. On the one hand, middle-aged and older adults are recommended to moderately use social media and the internet to learn more about new things or keep in touch with family and friends [39]. At the same time, middle-aged and older adults are recommended to participate in offline social activities and physical activity, if their physical capacity allows. On the other hand, it is suggested that children should actively contact their older family members, both electronically and offline, to provide them with emotional and material support. In addition, especially for older adults, it is recommended that the government and community organizations provide more community support [40], such as group recreational activities, which can also be important for the prevention of depression. In addition, the next waves of the CHARLS data will be obtained to determine the impact of using WeChat and posting WeChat moments on the long-term risk of depressive symptoms.

Acknowledgments

This work was supported by Preventive Medicine Scientific Research Project of Jiangsu Province in 2023 (Ym2023103), Medical Research Program of Health Commission of Jiangsu Province (M2021001), Natural Science Foundation of Nanjing University of Traditional Chinese Medicine (XZR2021005), and Academic Promotion Special Project of Department Director in 2022 (Y2022ZR14). All funding organizations had no role in the study design, implementation, analysis, or interpretation of the data. All authors thank the CHARLS team of Peking University for its excellent contribution and all participants surveyed for their efforts.

Data Availability

The data used in the cohort study can be accessed through the official website of the CHARLS [21].

Authors' Contributions

WW and HW are credited as cofirst authors for their significant contributions. XW is recognized as the corresponding author. WW, HW, and XW jointly worked on the study's concept and design. WW and HW also took on the analysis and interpretation of data. The drafting of the manuscript was a joint effort by WW and HW, with guidance from XW and with additional support from XH, QY, and FC who concentrated on assessing the results, especially focusing on the methods and analytics. All authors participated in critically revising the manuscript for important intellectual content, had full access to all the study's data, reviewed and revised the manuscript drafts, and gave their final approval for the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional data tables.
[DOC File, 252 KB - publichealth_v11i1e62730_app1.doc]

References

1. James SL, Abate D, Abate KH. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1789-1858. [doi: 10.1016/S0140-6736(18)32279-7] [Medline: 30496104]
2. Charlson FJ, Baxter AJ, Cheng HG, Shidhaye R, Whiteford HA. The burden of mental, neurological, and substance use disorders in China and India: a systematic analysis of community representative epidemiological studies. Lancet 2016 Jul 23;388(10042):376-389. [doi: 10.1016/S0140-6736(16)30590-6] [Medline: 27209143]
3. Tang T, Jiang J, Tang X. Prevalence of depressive symptoms among older adults in mainland China: a systematic review and meta-analysis. J Affect Disord 2021 Oct 1;293:379-390. [doi: 10.1016/j.jad.2021.06.050] [Medline: 34246000]
4. Gu L, Xie J, Long J, et al. Epidemiology of major depressive disorder in mainland china: a systematic review. PLoS One 2013;8(6):e65356. [doi: 10.1371/journal.pone.0065356] [Medline: 23785419]
5. Malhi GS, Mann JJ. Depression. Lancet 2018 Nov 24;392(10161):2299-2312. [doi: 10.1016/S0140-6736(18)31948-2] [Medline: 30396512]
6. Bennett S, Thomas AJ. Depression and dementia: cause, consequence or coincidence? Maturitas 2014 Oct;79(2):184-190. [doi: 10.1016/j.maturitas.2014.05.009] [Medline: 24931304]

XSL•FO
RenderX

7. Meng R, Yu C, Liu N, et al. Association of depression with all-cause and cardiovascular disease mortality among adults in China. JAMA Netw Open 2020 Feb 5;3(2):e1921043. [doi: 10.1001/jamanetworkopen.2019.21043] [Medline: 32049295]
8. Wang S, Mao S, Xiang D, Fang C. Association between depression and the subsequent risk of Parkinson's disease: a meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry 2018 Aug 30;86:186-192. [doi: 10.1016/j.pnpbp.2018.05.025] [Medline: 29859854]
9. Chen R, Chen Q, Lu G, et al. Sleep duration and depressive symptoms in Chinese middle-aged and older adults: the moderating effects of grip strength. J Affect Disord 2023 Oct 15;339:348-354. [doi: 10.1016/j.jad.2023.07.059] [Medline: 37451435]
10. Hasler G. Pathophysiology of depression: do we have any solid evidence of interest to clinicians? World Psychiatry 2010 Oct;9(3):155-161. [doi: 10.1002/j.2051-5545.2010.tb00298.x] [Medline: 20975857]
11. Mendelson T, Tandon SD. Prevention of depression in childhood and adolescence. Child Adolesc Psychiatr Clin N Am 2016 Apr;25(2):201-218. [doi: 10.1016/j.chc.2015.11.005] [Medline: 26980124]
12. Lin LY, Sidani JE, Shensa A, et al. Association between social media use and depression among U.S. young adults. Depress Anx 2016 Apr;33(4):323-331. [doi: 10.1002/da.22466] [Medline: 26783723]
13. Primack BA, Shensa A, Sidani JE, Escobar-Viera CG, Fine MJ. Temporal associations between social media use and depression. Am J Prev Med 2021 Feb;60(2):179-188. [doi: 10.1016/j.amepre.2020.09.014] [Medline: 33309454]
14. Perlis RH, Green J, Simonson M, et al. Association between social media use and self-reported symptoms of depression in US adults. JAMA Netw Open 2021 Nov 1;4(11):e2136113. [doi: 10.1001/jamanetworkopen.2021.36113] [Medline: 34812844]
15. Cotten SR, Ghaiumy Anaraky R, Schuster AM. Social media use may not be as bad as some suggest: implication for older adults. Innov Aging 2023;7(3):igad022. [doi: 10.1093/geroni/igad022] [Medline: 37179658]
16. Shannon H, Bush K, Villeneuve PJ, Hellemans KG, Guimond S. Problematic social media use in adolescents and young adults: systematic review and meta-analysis. JMIR Ment Health 2022 Apr 14;9(4):e33450. [doi: 10.2196/33450] [Medline: 35436240]
17. Lopes LS, Valentini JP, Monteiro TH, et al. Problematic social media use and its relationship with depression or anxiety: a systematic review. Cyberpsychol Behav Soc Netw 2022 Nov;25(11):691-702. [doi: 10.1089/cyber.2021.0300] [Medline: 36219756]
18. Tencent announces 2018 first quarter results. Tencent. 2018. URL: https://static.www.tencent.com/storage/uploads/2019/11/09/3e68a59c9f2568245fd0532efa0d7f78.pdf [accessed 2018-05-16]
19. Qu X, Houser SH, Zhang J, Wen J, Zhang W. Association between using social media WeChat and depressive symptoms among middle-aged and older people: findings from a national survey. BMC Geriatr 2022 Apr 22;22(1):351. [doi: 10.1186/s12877-022-03054-y] [Medline: 35448984]
20. Zhang C, Liang X. Association between WeChat use and mental health among middle-aged and older adults: a secondary data analysis of the 2020 China Family Panel Studies database. BMJ Open 2023 Oct 16;13(10):e073553. [doi: 10.1136/bmjopen-2023-073553] [Medline: 37844989]
21. China Health and Retirement Longitudinal Study. National Development Institute, Peking University. URL: http://charls.pku.edu.cn/ [accessed 2024-10-28]
22. Li H, Li C, Wang A, et al. Associations between social and intellectual activities with cognitive trajectories in Chinese middle-aged and older adults: a nationally representative cohort study. Alzheimers Res Ther 2020 Sep 25;12(1):115. [doi: 10.1186/s13195-020-00691-6] [Medline: 32977839]
23. Liu H, Ma Y, Lin L, Sun Z, Li Z, Jiang X. Association between activities of daily living and depressive symptoms among older adults in China: evidence from the CHARLS. Front Public Health 2023;11:1249208. [doi: 10.3389/fpubh.2023.1249208] [Medline: 38035294]
24. Wang G, Duan J, Kan Q, Zhou Y, Cheng Z, Tang S. The correlation analysis of WeChat usage and depression among the middle-aged and elderly in China: the mediating role of social participation. BMC Public Health 2023 Mar 10;23(1):462. [doi: 10.1186/s12889-023-15349-9] [Medline: 36899336]
25. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. Am J Public Health 2005;95 Suppl 1:S144-S150. [doi: 10.2105/AJPH.2004.059204] [Medline: 16030331]
26. Lee SL, Pearce E, Ajnakina O, et al. The association between loneliness and depressive symptoms among adults aged 50 years and older: a 12-year population-based cohort study. Lancet Psychiatry 2021 Jan;8(1):48-57. [doi: 10.1016/S2215-0366(20)30383-7] [Medline: 33181096]
27. Zhang Y, Kuang J, Xin Z, et al. Loneliness, social isolation, depression and anxiety among the elderly in Shanghai: findings from a longitudinal study. Arch Gerontol Geriatr 2023 Jul;110:104980. [doi: 10.1016/j.archger.2023.104980] [Medline: 36863166]
28. Elmer T, Geschwind N, Peeters F, Wichers M, Bringmann L. Getting stuck in social isolation: solitude inertia and depressive symptoms. J Abnorm Psychol 2020 Oct;129(7):713-723. [doi: 10.1037/abn0000588] [Medline: 32672987]
29. Ibarra-Rovillard MS, Kuiper NA. Social support and social negativity findings in depression: perceived responsiveness to basic psychological needs. Clin Psychol Rev 2011 Apr;31(3):342-352. [doi: 10.1016/j.cpr.2011.01.005] [Medline: 21382539]

XSL•FO
RenderX

30. Roh HW, Hong CH, Lee Y, et al. Participation in physical, social, and religious activity and risk of depression in the elderly: a community-based three-year longitudinal study in Korea. PLoS One 2015;10(7):e0132838. [doi: 10.1371/journal.pone.0132838] [Medline: 26172441]
31. Song L, Ge Y, Zhang X. The relationship between WeChat use by Chinese urban older adults living alone and their subjective well-being: the mediation role of intergenerational support and social activity. Psychol Res Behav Manag 2021;14:1543-1554. [doi: 10.2147/PRBM.S330827] [Medline: 34629911]
32. Guo L. WeChat as a semipublic alternative sphere: exploring the use of WeChat among Chinese older adults. Int J Commun 2017 Jan 27;11:408-428 [FREE Full text]
33. Nan Y, Xie Y, Hu Y. Internet use and depression among Chinese older adults: the mediating effect of interpersonal relationship. Front Public Health 2023;11:1102773. [doi: 10.3389/fpubh.2023.1102773] [Medline: 36935716]
34. Cao Q, Buskens E, Hillege HL, Jaarsma T, Postma M, Postmus D. Stratified treatment recommendation or one-size-fits-all? A health economic insight based on graphical exploration. Eur J Health Econ 2019 Apr;20(3):475-482. [doi: 10.1007/s10198-018-1013-z] [Medline: 30374630]
35. Wang W, Yu Q, Hu XR, Wang XX. Heterogeneous analysis on association between exposure and outcome in cohort studies: based on propensity score matching and subpopulation treatment effect pattern plot. Chin J Epidemiol 2024 May 14;45(5):748-754. [doi: 10.3760/cma.j.cn112338-20231121-00303]
36. Goligher EC, Lawler PR, Jensen TP, et al. Heterogeneous treatment effects of therapeutic-dose heparin in patients hospitalized for COVID-19. JAMA 2023 Apr 4;329(13):1066-1077. [doi: 10.1001/jama.2023.3651] [Medline: 36942550]
37. Valkenburg PM. Social media use and well-being: what we know and what we need to know. Curr Opin Psychol 2022 Jun;45:101294. [doi: 10.1016/j.copsyc.2021.12.006] [Medline: 35016087]
38. Majeed M, Irshad M, Fatima T, Khan J, Hassan MM. Relationship between problematic social media usage and employee depression: a moderated mediation model of mindfulness and fear of COVID-19. Front Psychol 2020;11:557987. [doi: 10.3389/fpsyg.2020.557987] [Medline: 33391075]
39. Chiu CJ, Hu YH, Lin DC, Chang FY, Chang CS, Lai CF. The attitudes, impact, and learning needs of older adults using apps on touchscreen mobile devices: results from a pilot study. Comput Hum Behav 2016 Oct;63:189-197. [doi: 10.1016/j.chb.2016.05.020]
40. Zhang C, Tang D, Wang Y, Jiang S, Liu X. Community support and promoting cognitive function for the elderly. Front Psychol 2022;13:942474. [doi: 10.3389/fpsyg.2022.942474] [Medline: 36148108]

Abbreviations

aRR: adjusted relative risk
BADL: basic activities of daily living
CESD-10: 10-item Center for Epidemiologic Studies Depression Scale
CHARLS: China Health and Retirement Longitudinal Study
cRR: crude relative risk
SMU: social media usage

Edited by A Mavragani; submitted 04.06.24; peer-reviewed by S Sheth, S Munusamy, W Zheng; revised version received 15.11.24; accepted 17.11.24; published 13.01.25.

Please cite as:
Wang W, Wang H, Hu X, Yu Q, Chen F, Qiu X, Wang X
The Association Between Posting WeChat Moments and the Risk of Depressive Symptoms Among Middle-Aged and Older Chinese Adults: Prospective National Cohort Study
JMIR Public Health Surveill 2025;11:e62730
URL: https://publichealth.jmir.org/2025/1/e62730
doi:10.2196/62730

© Wei Wang, Hui Wang, Xinru Hu, Qian Yu, Fangyi Chen, Xirui Qiu, Xiaoxiao Wang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Community-Built Environment's Effect on Physical Activity and Depressive Symptoms Trajectories Among Middle-aged and Older Adults: Chinese National Longitudinal Study

Kaili Zhang[1]; Bowen Huang[2]; Prasanna Divigalpitiya[3]

[1]Graduate School of Human-Environment Studies, Kyushu University, Fukuoka, Japan

[2]Zigong Academy of Urban Planning and Design, Zigong, Sichuan Province, China

[3]Faculty of Human-Environment Studies, Kyushu University, Fukuoka, Japan

Corresponding Author:
Bowen Huang
Zigong Academy of Urban Planning and Design
Number 19, Donghuan Road
Bancang Street, High-Tech Zone
Zigong, Sichuan Province, 643031
China
Phone: 86 15882073304
Email: jpby5223248@gmail.com

Abstract

Background: The effects of physical activity (PA) across different domains and intensities on depressive symptoms remain inconclusive. Incorporating the community-built environment (CBE) into longitudinal analyses of PA's impact on depressive symptoms is crucial.

Objective: This study aims to examine the effects of PA at different intensities—low-intensity PA (eg, walking activities) and moderate-to-vigorous-intensity PA (eg, activities requiring substantial effort and causing faster breathing or shortness of breath)—across leisure-time and occupational domains on depressive symptom trajectories among middle-aged and older adults. Additionally, it investigated how CBEs influence depressive symptoms and PA trajectories.

Methods: This longitudinal study included 6865 middle-aged and older adults from the China Health and Retirement Longitudinal Survey. A CBE variable system was developed using a community questionnaire to assess attributes of the physical built environment. Depressive symptoms were measured using the Center for Epidemiologic Studies Depression Scale. Latent growth curve modeling was applied to analyze 3 waves of the cohort data (2015, 2018, and 2020) to explore the differential effects of PA on depressive symptoms and the role of the CBE.

Results: In the 2015 and 2018 waves, higher low-intensity leisure-time physical activity (LTPA) was associated with lower depressive symptoms (β=–.025, *P*=.01 and β=–.027, *P*=.005, respectively). Across all waves, moderate-to-vigorous-intensity LTPA showed no significant predictive effects (*P*=.21 in 2015, *P*=.57 in 2018, and *P*=.85 in 2020, respectively). However, higher occupational physical activity (OPA), particularly at moderate-to-vigorous intensities, was consistently associated with higher depressive symptoms. Parallel process latent growth curve modeling revealed that the initial level of total LTPA negatively predicted the initial level of depressive symptoms (β=–.076, *P*=.01). OPA exhibited dual effects, positively predicting the initial level of depressive symptoms (β=.108, *P*<.001) but negatively predicting their upward trajectory (β=–.136, *P*=.009). Among CBE variables, better infrastructure conditions (β=–.082, *P*<.001) and greater accessibility to public facilities (β=–.036, *P*=.045) negatively predicted the initial level of depressive symptoms. However, greater accessibility to public facilities positively predicted the upward trajectory of depressive symptoms (β=.083, *P*=.04). Better infrastructure conditions (β=.100, *P*=.002) and greater accessibility to public transport (β=.060, *P*=.01) positively predicted the initial level of total LTPA. Meanwhile, better infrastructure conditions (β=–.281, *P*<.001) and greater accessibility to public facilities (β=–.073, *P*<.001) negatively predicted the initial level of total OPA. Better infrastructure conditions positively predicted the declining trajectory of total OPA (β=.100, *P*=.004).

XSL•FO
RenderX

Conclusions: This study underscores the importance of considering the differential effects of PA across domains and intensities on depressive symptoms in public policies and guidelines. Given the influence of the environment on PA and depressive symptoms, targeted community measures should be implemented.

(JMIR Public Health Surveill 2025;11:e64564) doi:10.2196/64564

KEYWORDS

community-built environment; physical activity; depressive symptom trajectories; middle-aged and older adults; latent growth curve modeling; longitudinal study

Introduction

In communities, the physical built environment and health behaviors, such as physical activity (PA), can significantly influence residents' mental health [1]. Overall, well-designed community-built environments (CBEs)—including factors such as walkability, appropriate density, and safety—are beneficial in reducing depressive symptoms. By contrast, unfavorable built environments, such as poor-quality housing and insufficient green spaces, may exacerbate depressive symptoms [2]. Exercise and PA have been shown to be as effective as some antidepressant medications in improving depressive symptoms [3].

With the acceleration of the global aging process [4], depression among middle-aged and older adults has become a significant public health issue worldwide. The World Health Organization (WHO) estimates that the prevalence of depression in the global aging population ranges from 10% to 20% [5]. In China, the pooled prevalence of depressive symptoms among seniors aged 60 years and older is approximately 22.7% [6]. Depression poses a serious threat to the mental health of middle-aged and older adults and can also lead to physical illness and an increased risk of death [7].

In all studies examining factors influencing depressive symptoms, PA has become a focal point. PA can be easily incorporated into the daily context of a community, whether in a living or work environment, making it a universally accessible and straightforward intervention. Its practicality and ease of implementation have made PA a key topic in discussions about improving mental health [8]. The beneficial effects of PA on alleviating depressive symptoms are well-documented. Physically inactive individuals are more likely to experience depressive symptoms compared with those who engage in regular PA [9]. Recent studies have shown that this effect is not influenced by age [10]. Middle-aged and older adults are particularly susceptible to depressive disorders due to factors such as the gradual decline in physical function, an increase in chronic illnesses, and the potential for significant life changes, including widowhood, retirement, and living alone after middle age [11]. Older adults who engage in insufficient PA are at a higher risk of depression compared with those who maintain adequate levels of PA [12].

Many studies exploring the relationship between PA and depressive symptoms have primarily focused on exercise-based PA [13]. However, PA can also occur in occupational settings or outside structured exercise programs. PA shows different health outcomes depending on the context, such as work versus leisure settings [14]. Recently, researchers have begun to examine the "PA paradox"—the concept that leisure-time physical activity (LTPA), such as sports and recreational activities, has positive effects on health, while PA in occupational settings may have negative consequences for both physical and mental health [15]. LTPA, regardless of intensity, is associated with a lower risk of depressive symptoms. By contrast, occupational PA (OPA) or transportation-related PA does not show a similar reduction in the risk of depressive symptoms [16]. Workers who engage in work-related physical activities tend to report higher levels of depressive symptoms [17], and individuals with the highest levels of OPA are more likely to experience depressive symptoms [18]. Globally, a significant proportion of people continue to engage in manual labor, particularly in many developing countries. Therefore, it is important to distinguish between different domains or motivations when examining the effect of PA on depressive symptoms, as this distinction may lead to entirely different results. However, only a small number of studies have made such distinctions to examine the effects of various PA domains on depressive symptoms [19].

There are still several limitations and gaps in the existing research on the effects of PA on depressive symptoms. First, while the beneficial effects of exercise on depressive symptoms are well-established, the complex impacts of different domains (such as exercise and work) or PA intensity remain inconclusive [20]. There is still a lack of basic evidence, particularly for middle-aged and older adults, and further research is urgently needed to provide clearer insights. Second, most studies on the effect of PA on depressive symptoms offer only a cross-sectional perspective, with few utilizing cohort data to explore the longitudinal impact. By focusing on individual trajectories and patterns over time, longitudinal studies can provide valuable insights into the developmental continuum, help determine the stability of the association between PA and depressive symptoms, and analyze the dynamics of these interactions [21].

Furthermore, the CBE is an important factor influencing mental health, including depressive symptoms and mental health inequalities [22]. Depressive symptoms are associated with the built environment, and different physical attributes of the community may impact them through various behavioral or social pathways [23]. Many characteristics of the CBE may expose residents to environmental stressors—such as poor housing quality, uncomfortable population densities, low safety, and poor accessibility to amenities—that can negatively affect their mental health [24]. As for built environment measures, the most common ones are related to the "5D" elements: density, diversity, design, destination accessibility, and distance to transit

XSL•FO
RenderX

[25]. These elements are often evaluated individually or together.

Despite the growing body of literature on the effects of CBE on depressive symptoms, there is still limited evidence concerning middle-aged and older adults. Few studies have highlighted that improving community walkability, particularly the accessibility and safety of public transportation, is crucial for middle-aged individuals with depression [2]. Higher levels of residential greenery have been significantly associated with lower odds of depressive symptoms among older adults living in the community [26]. Additionally, high population density and proximity to destinations contribute to better mental health, while walkable environments are linked to reduced depression, particularly among Asian women [27]. Conversely, the availability of local retail stores has been associated with an increased likelihood of depressive symptoms in older men [28]. Thus, inconsistent results may arise in different contexts or regions. Therefore, it is crucial to consider the local community context when assessing the effects of CBE on depressive symptoms and PA. Furthermore, among longitudinal studies exploring the relationship between PA and depressive symptoms, few have taken CBE into account [29]. Therefore, providing insights into how CBE influences the trajectories of depressive symptoms and PA is invaluable.

To address these gaps, this study focused on middle-aged and older adults aged 45 years and older in China, using data from 3 waves of repeated measurements to identify the trajectory of depressive symptoms, various types of PA, and the built environment of communities. The study aimed to answer the following 2 research questions: (1) What are the longitudinal effects of LTPA and OPA on depressive symptoms among middle-aged and older adults? (2) How does the CBE influence the trajectory of depressive symptoms and the 2 domains of PA in this population?

Methods

Data Source and Study Sample

The data for this study were obtained from the latest 3 waves of tracking data (2015, 2018, and 2020) from the China Health and Retirement Longitudinal Survey (CHARLS), a large microsurvey database. CHARLS covers 450 communities across 28 provinces in China, using a multistage stratified probability proportional to size sampling method to conduct the survey. A detailed description can be found in Appendix S1 in Multimedia Appendix 1 (see also [30-32]). As part of the International Survey of Health in Aging, CHARLS has been extensively utilized in interdisciplinary studies focusing on aging and health issues [33].

Figure S1 in Multimedia Appendix 1 illustrates the sample screening process. For each wave of data, this study excluded samples with missing PA data (missing PA data were randomly absent: participants were randomly selected to answer PA-related questions, and the absence of available records indicated that they did not respond to PA-related questions) and missing depressive symptom scores (individuals who left more than 2 questions unanswered on the scale were considered invalid for the depressive symptom measure. Among the respondents excluded from the analysis, approximately 80% had missing data for all 10 items on the scale in each wave). Additionally, samples lost due to attrition or participant death were excluded. After matching data across the 3 waves, participants younger than 45 years were also excluded. This resulted in a final sample of 6865 respondents, with 20,595 observation points. For samples with missing values for sex and education level, data from adjacent waves were used.

Ethics Approval

The original data collection was approved by the Biomedical Ethics Review Committee of Peking University (IRB00001052-11015). All data used in this study were fully anonymized, and participants provided informed consent. The data for this study were registered and obtained through the official CHARLS website [34].

Depressive Symptoms

Depressive symptoms were assessed using the Center for Epidemiologic Studies Depression Scale (CES-D), as provided by CHARLS. Respondents answered 10 questions reflecting feelings and behaviors associated with depression [35]. The total score, which ranged from 0 to 30, was calculated by summing the responses to these 10 questions, with reverse scoring applied to the 2 positive questions. Higher CES-D scores indicated more severe depressive symptoms. The Cronbach α coefficients for the CES-D scale in this study were 0.794, 0.800, and 0.796 in 2015, 2018, and 2020, respectively. In the descriptive statistical analysis, a CES-D score of ≥10 was classified as indicative of depression [36]. Additionally, we calculated the Reliable Change Index (RCI) for the CES-D scores across waves to assess whether changes in participants' depressive symptom scores over multiple waves were statistically significant (|RCI|>1.96) [37].

Physical Activity Level

The CHARLS used items similar to those in the International Physical Activity Questionnaire (short form) to assess respondents' levels of PA. It collected data on the duration and weekly frequency of each type of PA: low intensity (walking activities), moderate intensity (activities that cause faster breathing than usual, such as cycling, playing table tennis, and practicing yoga), and high intensity (activities that cause shortness of breath, such as fast running and playing basketball) [38]. This study used the metabolic equivalent (MET) method to calculate PA levels. To determine an individual's weekly MET for each PA intensity, we multiplied the corresponding duration, frequency, and coefficient for each activity. Following previous studies [30,31], we performed data organization and calculations for the 3 intensity levels of PA, including data cleaning, outlier elimination, and data truncation. Details regarding the CHARLS Questionnaire on PA, as well as the data processing and calculation methods used in this study, are provided in Appendix S1 in Multimedia Appendix 1. We scaled all PA data by a factor of 100 to meet the data requirements for statistical analysis before inputting them into the model calculation. As a result, the actual unit of MET reported in the models was 100 MET minutes per week.

XSL•FO
RenderX

The CHARLS database collected information on the purpose of each PA (4 categories: work, exercise, recreation, and others), allowing for the examination of the differential effects of various PA domains on depressive symptoms. Activities labeled as "other" in CHARLS were not explicitly defined and included a mix of multiple purposes, so they were excluded from the calculations for this study. PA data were categorized into 2 domains: LTPA and OPA. Within each domain, PA was further classified by intensity: low intensity (walking activities) and moderate-to-vigorous intensity (activities that require significant effort and lead to faster breathing or shortness of breath). The corresponding MET values were calculated and aggregated.

Construction of the Community-Built Environment Variables System

In this study, a set of variables was developed to assess the physical built environments of communities based on the community questionnaire data set provided by CHARLS. This data set was released in a single issue as part of the national baseline survey, updated in December 2013, with no subsequent updates in later waves. Based on the concepts of 3D, 5D, and other typical indicator systems and their extensions [25,39], and considering data availability and the characteristics of middle-aged and older adults, this study selected 7 core variables to evaluate the CBE attributes of each community: residential density, density of public facilities, diversity of sports venues, diversity of senior care facilities, accessibility of public facilities, accessibility of public transport, and infrastructure conditions. These variables are outlined in Table 1.

Table 1. Community-built environment variable system constructed from the community questionnaire of the China Health and Retirement Longitudinal Survey.

Built environment element	Variable	Calculation method: unit (weight)
Density	• Residential density • Density of public facilities	• Community population/total area of the community: pieces/km2 (N/A[a]) • The total number of basic public facilities in the community/total area of the community: pieces/km2 (N/A)
Diversity	• Diversity of sports venues • Diversity of senior care facilities	• Whether or not the community has any of the 8 sports venues: a "yes" answer will be scored as 1 point for that venue, a "no" answer will be scored as 0 points for that venue, and the total values will be added up. The maximum score is 8: points (N/A) • Whether or not the community has any of the 7 senior care facilities; diversity levels were calculated as described above.
Destination accessibility	• Accessibility of public facilities	• The average distance from the community office to the most commonly used type of each facility (if this facility is located within the community, the answer is 0): km (N/A).
Distance to transit	• Accessibility of public transport	• The actual distance from the community office to the most commonly used bus stop: km (43.86%). • The total number of actual bus lines accessible to the community: pieces (56.14%).
Infrastructure	• Infrastructure conditions	• Pavement quality of the main roads, where pathway/dirt/unpaved road=1 point, sandstone road=2 points, and paved road=3 points: points (13.62%) • The total number of days roads are impassable: days (12.31%) • The proportion of households that use purified tap water: percentage (18.30%) • Availability of sewer system in the community. yes=1 point and no=0 points: points (17.56%) • The level of waste disposal, where moved away by truck=5 points, buried in the village=4 points, burn away=3 points, put into nearby river=2 points, and do not manage=1 point: points (17.13%) • The proportion of households that use electricity: percentage (17.07%) • The condition of community toilets, where inside toilet with water=5, inside toilet without water=4, outside toilet with water=3, outside public toilet without water=2, and open air=1: points (4.00%)

[a]N/A: not applicable.

Among the variables, the accessibility of public transport and infrastructure conditions were assessed by calculating multiple subvariables. The CRITIC weighting method was used to determine the weight of each subvariable. Detailed information on the content of the facilities and the source of the question numbers can be found in Table S1 in Multimedia Appendix 1. Furthermore, due to limitations in the survey data, this study could not include variables such as green spaces and road network connectivity, despite their recognized importance [40,41]. All CBE variables were normalized using the min-max normalization method, resulting in a processed data range of 0-1, before being included in the model calculations. Further details can be found in Appendix S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Covariates

Three types of variables were used as control variables: demographic and socioeconomic status, health behavior and status, and housing characteristics. These variables were selected due to their association with mental health levels or their potential impact on depressive symptoms [42]. Baseline data were chosen as covariates for inclusion in the model calculations. Specifically, demographic and socioeconomic status included age, sex (male or female), education level (categorized into 4 groups: illiteracy; primary and lower; middle, high, and vocational school; and 3-year college/bachelor's degree and higher), marital status (married or other), cohabiting status (with or without a cohabiting partner), and personal annual income. Personal annual income encompasses salaries and transfer income, such as pensions, benefits, and subsidies.

Health behavior and status included smoking (current smoker or nonsmoker), drinking (categorized into 3 groups: drinking more than once a month, drinking less than once a month, and nondrinker), the number of chronic diseases (total of 10, as outlined in Appendix S2 and Table S2 in Multimedia Appendix 1), and disability status (disabled or nondisabled). Housing characteristics were assessed based on the presence of various amenities: elevator, barrier-free facilities, toilet flushing, electricity, running water, bathing facility, gas/natural gas, heating, broadband, air purifier, and tidiness, for a total of 11 items. One point was assigned for each present characteristic or amenity, and the total score was calculated. Higher scores indicate better housing quality.

Statistical Analysis

This study utilized latent growth curve modeling (LGCM) to analyze the trajectories and longitudinal associations among the variables. LGCM, a variant of structural equation modeling, is a statistical method commonly used in longitudinal studies with latent variables [43]. It uses the intercept to represent the initial level and the slope to capture the rate of change. First, an unconditional linear LGCM was applied to each of the 3 measures of PA and depressive symptoms to examine their trajectories in middle-aged and older adults. In the second step, all control variables were incorporated into a linear growth model to construct a conditional linear LGCM for depressive symptoms. PA was included as a time-varying covariate to examine the effect of each wave of PA on depressive symptoms. To test whether these effects remained stable over time (2015-2018-2020), equality constraints were applied to the effects at each wave, and likelihood ratio tests (LRTs) were conducted to assess the statistical equivalence of these effects. In the third step, an LGCM was constructed for both depressive symptoms and PA, with 7 categories of built environment (CBE) as time-invariant covariates. A parallel process LGCM was used to examine the dynamic associations among CBE, PA, and depressive symptoms. The study also incorporated the adjusted individual weights provided by CHARLS for each wave of data in all models to minimize representational bias due to response differences. The acceptable thresholds for model fit indices were as follows: Comparative Fit Index>0.90, Tucker-Lewis Index>0.90, root mean square error of approximation<0.08, and standardized root mean square residual<0.08 [44]. Additionally, the Harman single-factor test was conducted to examine the data for common method bias. The total variance explained by the first common factor was 19.76% (<40%), indicating that the study was not significantly affected by common method bias [45]. Data cleaning was performed using Stata 18.0 (StataCorp), all LGCM analyses were conducted with Mplus 8.10 (Muthén and Muthén), and the Harman single-factor test was executed using IBM SPSS 26.

Results

Characteristics of the Sample at Baseline

A total of 6865 participants were included in the study sample, with the detailed statistical characteristics presented in Table S2 in Multimedia Appendix 1. The male-to-female ratio in the sample was 0.88, indicating a nearly balanced distribution. The respondents' education levels were relatively low: of the 6865 participants, 4652 (67.76%) did not receive education beyond secondary school, reflecting the social context of the middle-aged and older adult generation in China. The marriage rate in the sample was very high (6836/6865, 99.58%), and a majority of respondents (5767/6865, 84%) lived with a cohabiting partner. The mean annual personal income of participants was 15,712.63 Chinese yuan (1 yuan=US $0.14). Among the respondents, the majority were nonsmokers (4986/6865, 72.63%) and nondrinkers (4400/6865, 64.09%). On average, participants reported having 2.10 chronic diseases, and the average score for housing characteristics was 4.06. Additionally, 872 (12.70%) participants reported having a disability.

Developmental Trajectories of Physical Activity and Depressive Symptoms

The statistics for depressive symptoms and various types of PA among middle-aged and older adults across the 3 data waves are presented in Table 2. The proportion of participants with depression (CES-D scores≥10) ranged from approximately 30% to 40% over the study period (2178/6865, 31.73%, in 2015; 2469/6865, 35.97%, in 2018; and 2666/6865, 38.83%, in 2020). Additionally, in each wave, participants consistently reported higher levels of OPA than LTPA. The amount of moderate-to-vigorous-intensity OPA was greater than that of low-intensity OPA, while low-intensity LTPA was more prevalent than moderate-to-vigorous-intensity LTPA.

Unconditional LGCM was used to examine the developmental trends of each variable. The model fit indices, along with the mean intercepts and slopes for the 3 models, are presented in Table 3. Overall, the models demonstrated a good fit. Depressive symptoms and LTPA levels showed a linear upward trend across the 3 measurement waves (with positive and significant slopes), while OPA levels displayed a linear downward trend. Furthermore, based on the RCI statistics for changes in depressive symptoms (Appendix S2 and Table S3 in Multimedia Appendix 1), despite the overall increasing trend in depressive symptoms, most participants did not show significant changes across waves (–1.96≤RCI≤1.96). Specifically, 5506 of 6865 (80.20%) individuals from 2015 to 2018, 5483 of 6865 (79.87%) individuals from 2018 to 2020, and 5391 of 6865 (78.53%)

XSL•FO
RenderX

individuals from 2015 to 2020 showed no significant changes (–1.96≤RCI≤1.96). Over the longer period from 2015 to 2020, 948 of 6865 (13.81%) participants experienced a significant increase (RCI>1.96) in depressive symptoms, while 526 of 6865 (7.66%) participants experienced a significant decrease (RCI<–1.96).

Table 2. Statistical characteristics of depressive symptoms and physical activity among middle-aged and older adults across three waves of data.[a]

Variable	Wave 2015	Wave 2018	Wave 2020
Depressive symptoms, mean (SD)	7.81 (6.27)	8.46 (6.32)	8.85 (6.46)
Number of depression, n/N (%)	2178/6865 (31.73)	2469/6865 (35.97)	2666/6865 (38.83)
Number of nondepression, n/N (%)	4687/6865 (68.27)	4396/6865 (64.03)	4199/6865 (61.17)
Total LTPA[b], mean (SD)	974.38 (2130.84)	1248.28 (2412.47)	1590.66 (2748.37)
Low-intensity LTPA, mean (SD)	593.47 (1198.14)	792.45 (1356.73)	902.70 (1356.9)
Moderate-to-vigorous-intensity LTPA, mean (SD)	380.91 (1514.96)	455.83 (1719.31)	687.96 (2140.36)
Total OPA[c], mean (SD)	5529.50 (6988.45)	4834.27 (6593.79)	3745.39 (5744.90)
Low-intensity OPA, mean (SD)	1234.15 (1907.24)	1190.91 (1925.46)	719.76 (1570.59)
Moderate-to-vigorous-intensity OPA, mean (SD)	4295.35 (5985.89)	3643.36 (5600.78)	3025.63 (5064.63)

[a]A score of the Center for Epidemiologic Studies Depression Scale ≥10 was defined as depression. The unit of physical activity in this table is metabolic equivalent minutes/week.

[b]LTPA: leisure-time physical activity.

[c]OPA: occupational physical activity.

Table 3. Estimates and fit index of unconditional LGCMs.[a]

Estimates	LGCM[b] for depressive symptoms		LGCM for total LTPA[c]		LGCM for total OPA[d]	
	Estimate	SE	Estimate	SE	Estimate	SE
Means						
Intercept	7.855[e]	0.073	9.661[e]	0.246	56.000[e]	0.803
Slope	0.516[e]	0.038	3.048[e]	0.190	–9.087[e]	0.439
Covariance (intercept with slope)	–0.319	0.486	–18.383[f]	8.409	–46.701[e]	5.570
Fit index						
Chi-square (*df*)	4.180 (1)	N/A[g]	1.182 (1)	N/A	3.442 (1)	N/A
Root mean square error of approximation (90% CI)	0.022 (0.004-0.045)	N/A	0.005 (0.000-0.033)	N/A	0.031 (0.013-0.053)	N/A
Comparative Fit Index	0.999	N/A	1.000	N/A	0.998	N/A
Tucker-Lewis Index	0.998	N/A	0.999	N/A	0.994	N/A
Standardized root mean square residual	0.006	N/A	0.004	N/A	0.008	N/A

[a]The unit of physical activity in this table is 100 metabolic equivalent minutes/week.

[b]LGCM: latent growth curve modeling.

[c]LTPA: leisure-time physical activity.

[d]OPA: occupational physical activity.

[e]*P*<.001.

[f]*P*<.05.

[g]N/A: not applicable.

Effects of Leisure-Time and Occupational Physical Activity on Depressive Symptoms

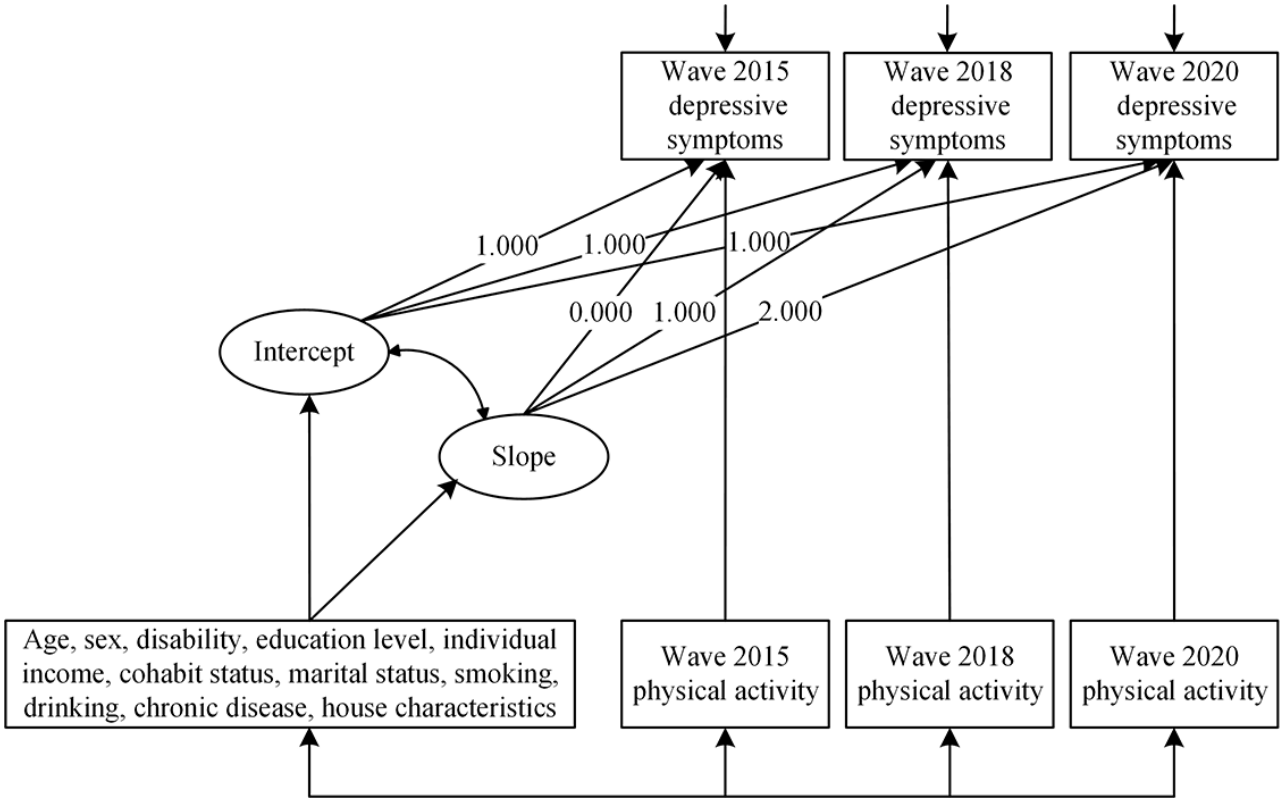
Based on the time-varying covariate LGCM specified in Figure 1, the direct effects of all types of PA on depressive symptoms were calculated for each wave, and equality constraints were tested. The results are presented in Table 4. The model fit for all pathways was satisfactory (Appendix S2 and Table S4 in Multimedia Appendix 1). In both 2015 and 2018, total LTPA (β=–.023, *P*=.02 and β=–.025, *P*=.01, respectively) and low-intensity LTPA (β=–.025, *P*=.01 and β=–.027, *P*=.005, respectively) negatively predicted depressive symptoms.

XSL•FO
RenderX

However, this effect was not statistically significant in 2020 (*P*=.51). No significant differences were observed in the pathways across multiple waves, as the LRT results were not significant (*P*=.31 in 2015, *P*=.65 in 2018, and *P*=.14 in 2020, respectively). Moderate-to-vigorous-intensity LTPA did not show a significant predictive effect on depressive symptoms in any wave (*P*=.21 in 2015, *P*=.57 in 2018, and *P*=.85 in 2020, respectively).

By contrast, in waves 2015, 2018, and 2020, both total OPA (β=.041, *P*<.001; β=.022, *P*=.02; and β=.027, *P*=.03, respectively) and moderate-to-vigorous-intensity OPA (β=.036, *P*<.001; β=.020, *P*=.04; and β=.019, *P*=.049, respectively) positively predicted depressive symptoms. The positive predictive effect of low-intensity OPA on depressive symptoms was significant only in wave 2015 (β=.030, *P*=.003). Furthermore, the LRT results indicated that the effect of low-intensity OPA on depressive symptoms differed significantly between waves 2015 and 2018 (*P*=.048) and between waves 2015 and 2020 (*P*=.02). Across the 3 waves of data, the effects of both total OPA (*P*=.27 in 2015; *P*=.98 in 2018; and *P*=.33 in 2020, respectively) and moderate-to-vigorous-intensity OPA on depressive symptoms remained stable with no significant differences observed (*P*=.27 in 2015, *P*=.87 in 2018, and *P*=.41 in 2020, respectively). Given the representativeness, statistical significance, and stability of the effects across waves, this study proceeded with the inclusion of total LTPA and total OPA for further analysis.

Figure 1. Graphical representation of the time-varying covariate latent growth curve modeling with the depressive symptoms and physical activity as time-varying variables. The 11 covariates as time-invariant covariates were plotted in a box for mapping simplicity. They were calculated concurrently in each latent growth curve model analysis.



XSL•FO
RenderX

Table 4. Standardized estimate, SE (in parentheses), and significance of the time-varying covariate latent growth curve modeling.[a]

Variables	Wave 2015 depressive symptoms	Wave 2018 depressive symptoms	Wave 2020 depressive symptoms	LRT[b] 2015 vs 2018	LRT 2018 vs 2020	LRT 2015 vs 2020
2015 Total LTPA[c]	*–0.023 (0.010)*[d]	N/A[e]	N/A	*P*=.31	N/A	N/A
2018 Total LTPA	N/A	*–0.025 (0.010)*[d]	N/A	N/A	*P*=.65	N/A
2020 Total LTPA	N/A	N/A	–0.015 (0.022)	N/A	N/A	*P*=.14
2015 Low-intensity LTPA	*–0.025 (0.010)*[d]	N/A	N/A	*P*=.93	N/A	N/A
2018 Low-intensity LTPA	N/A	*–0.027 (0.010)*[f]	N/A	N/A	*P*=.19	N/A
2020 Low-intensity LTPA	N/A	N/A	*–0.010 (0.010)*	N/A	N/A	*P*=.20
2015 Moderate-to-vigorous-intensity LTPA	–0.012 (0.010)	N/A	N/A	*P*=.18	N/A	N/A
2018 Moderate-to-vigorous-intensity LTPA	N/A	0.005 (0.009)	N/A	N/A	*P*=.57	N/A
2020 Moderate-to-vigorous-intensity LTPA	N/A	N/A	–0.002 (0.009)	N/A	N/A	*P*=.35
2015 Total OPA[g]	*0.041 (0.010)*[h]	N/A	N/A	*P*=.27	N/A	N/A
2018 Total OPA	N/A	*0.022 (0.010)*[d]	N/A	N/A	*P*=.98	N/A
2020 Total OPA	N/A	N/A	*0.027 (0.012)*[d]	N/A	N/A	*P*=.33
2015 Low-intensity OPA	*0.030 (0.010)*[f]	N/A	N/A	*P=.048*[d]	N/A	N/A
2018 Low-intensity OPA	N/A	0.012 (0.010)	N/A	N/A	*P*=.22	N/A
2020 Low-intensity OPA	N/A	N/A	–0.005 (0.010)	N/A	N/A	*P=.02*[d]
2015 Moderate-to-vigorous-intensity OPA	*0.036 (0.010)*[h]	N/A	N/A	*P*=.27	N/A	N/A
2018 Moderate-to-vigorous-intensity OPA	N/A	*0.020 (0.010)*[d]	N/A	N/A	*P*=.87	N/A
2020 Moderate-to-vigorous-intensity OPA	N/A	N/A	*0.019 (0.010)*[d]	N/A	N/A	*P*=.41

[a]Italicized values indicate significant results.

[b]LRT: likelihood ratio test.

[c]LTPA: leisure-time physical activity.

[d]*P*<.05.

[e]N/A: not applicable.

[f]*P*<.01.

[g]OPA: occupational physical activity.

[h]*P*<.001.

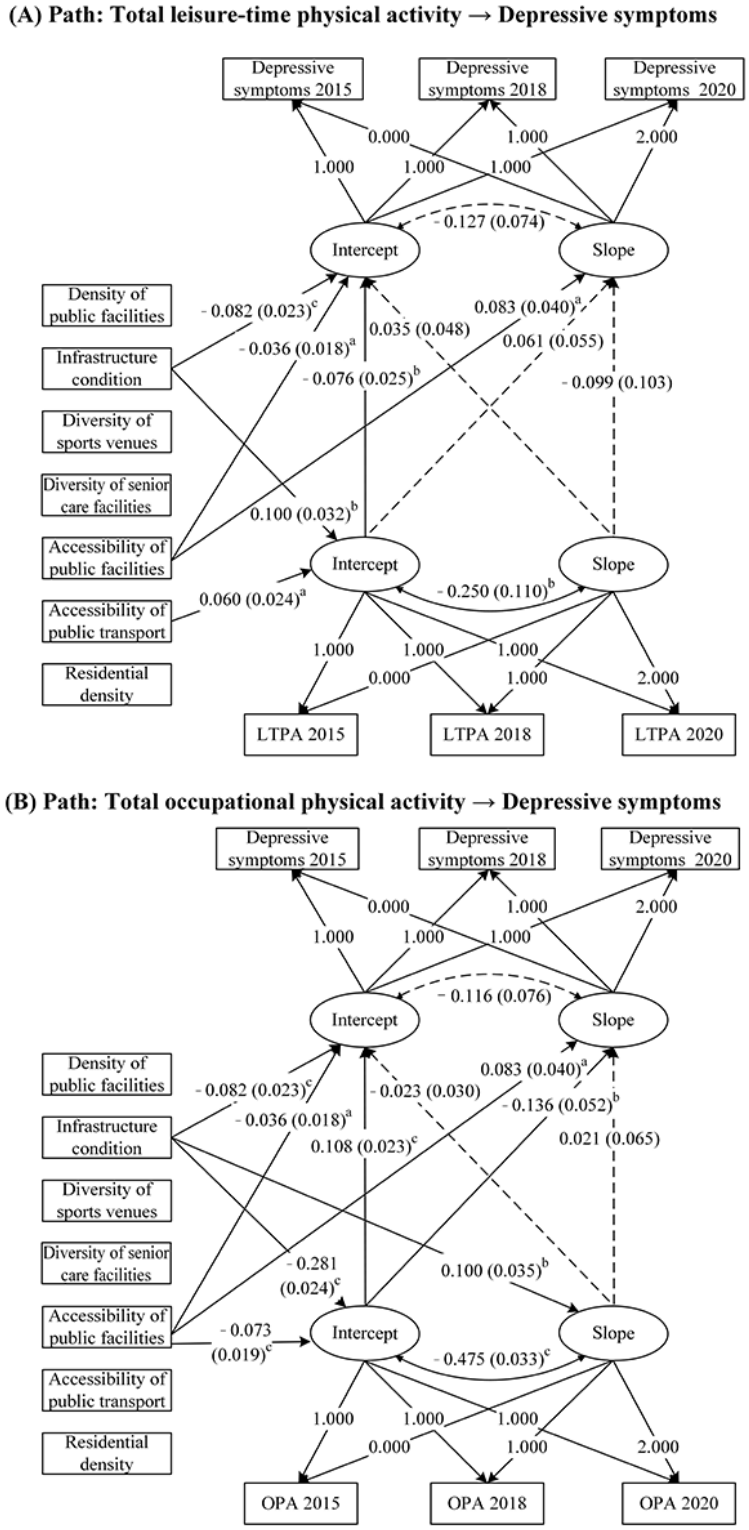
Longitudinal Associations Among Community-Built Environment, Physical Activity, and Depressive Symptoms

This study further used a parallel process LGCM to analyze the dynamic associations between CBE, the 2 domains of PA, and depressive symptoms (Figure 2). The models demonstrate a good fit (Appendix S2 and Table S4 in Multimedia Appendix 1). First, the intercept (initial level) of total LTPA negatively predicted the initial level of depressive symptoms (β=–.076, *P*=.003). The initial level of total OPA positively predicted the initial level of depressive symptoms (β=.108, *P*<.001) and negatively predicted the rate of increase in depressive symptoms (β=–.136, *P*=.009).

Regarding the effect of CBE on the trajectory of depressive symptoms, infrastructure conditions and accessibility of public facilities negatively predicted the initial level of depressive symptoms (β=–.082, *P*<.001 and β=–.036, *P*=.045, respectively). However, accessibility to public facilities positively predicted the rate of increase in depressive symptoms (β=.083, *P*=.04). Additionally, infrastructure conditions (β=.100, *P*=.002) and accessibility of public transport (β=.060, *P*=.01) positively predicted the intercept of total LTPA. By contrast, infrastructure conditions (β=–.281, *P*<.001) and accessibility of public facilities (β=–.073, *P*<.001) negatively predicted the initial level of total OPA. Infrastructure conditions also positively predicted the declining trend of total OPA over time (β=.100, *P*=.004).

XSL•FO
RenderX

Figure 2. Parallel process latent growth curve modeling to estimate the effect of physical activity on depressive symptoms, with the community-built environment serving as a time-invariant covariate. Dashed lines indicate nonsignificant paths. The nonsignificant paths for the 7 community-built environment variables are not displayed. All control variables are evaluated simultaneously. [a]*P*<.05; [b]*P*<.01; [c]*P*<.001. LTPA: leisure-time physical activity; OPA: occupational physical activity.



Discussion

Longitudinal Effect of Physical Activity on Depressive Symptoms Among Middle-Aged and Older Adults

First, the analysis based on the linear unconditional LGCM revealed an upward trend in depressive symptoms among middle-aged and older adults from 2015 to 2020. This finding aligns with previous studies using CHARLS data [33], suggesting that depressive symptoms may be influenced by the aging process. However, further analysis using the RCI indicated that while depressive symptoms showed an upward trend, this change appeared to occur primarily at the group level. At the

XSL•FO
RenderX

individual level, fluctuations in depressive symptoms were relatively small for most participants, with only a minority experiencing significant increases. For the majority, changes did not reach statistically significant levels (Table S3 in Multimedia Appendix 1). Additionally, the average levels of depressive symptoms across the 3 waves were slightly below the CES-D clinical screening threshold (Table 2) [36], indicating that most participants exhibited mild or subclinical depressive symptoms. Therefore, the core focus of this study is to understand and explore the potential effects of CBE and PA on changes in depressive symptoms. These changes should not be overly simplified as the onset of clinical depression or an increase in its severity.

Existing studies examining the effects of PA on depressive symptoms often focus on a single domain or intensity level. However, given that PA across different intensities and domains may have varying effects on depressive symptoms [46], our study contributes to the literature by further categorizing PA into low- and moderate-to-high-intensity levels within the domains of LTPA and OPA. These results highlight the importance of the differences observed. By analyzing data from 3 waves of repeated measurements, we found that LTPA, particularly low-intensity LTPA, consistently served as a negative predictor of depressive symptoms among middle-aged and older adults in China. This suggests that greater participation in low-intensity LTPA was associated with fewer depressive symptoms in this population (Table 4). Our findings are consistent with previous studies that emphasize the benefits of moderate and light exercise over vigorous exercise for middle-aged and older adults [47]. Engaging in lower-intensity activities, such as walking, may also help prevent depression [48], with lower LTPA levels being associated with fewer depressive symptoms [20].

However, in all waves of the analysis, we did not observe a significant association between moderate-to-vigorous-intensity LTPA and depressive symptoms. This finding is inconsistent with some existing literature. For example, certain cross-sectional studies have suggested that both low- and moderate-to-vigorous-intensity PA positively impact the mental health of middle-aged and older adults [49]. Furthermore, both small and large amounts of LTPA are associated with a lower prevalence of depression, with comparable effects [19]. These discrepancies may arise from differences in how PA intensity is categorized, the participant selection criteria, or regional characteristics across the studies.

This study also clearly observed that after controlling for socioeconomic factors such as income levels, higher moderate-to-vigorous-intensity OPA was associated with higher depressive symptoms. This negative predictive effect remained significant and consistent across all waves of data (Table 4). This result aligns with previous studies. Research from Brazil and South Korea reported that participants with higher levels of OPA were more likely to experience depressive symptoms [20], and that the higher the level of work-related PA, the greater the likelihood of depressive symptoms [50]. Similarly, OPA appears to be associated with an increased risk of depression [51]. However, these results contradict those of other studies. For instance, research from the United States suggested that OPA does not significantly predict depression regardless of gender [52], while a study from Ghana found that depression decreased as work-related PA increased [53]. As our study controlled for socioeconomic characteristics, these differences suggest that the effect of OPA on depressive symptoms may vary depending on the level of economic development or overall social progress in a given region.

Collectively, these findings support the important perspective that the effects of PA on depressive symptoms vary depending on the domain and intensity of the activity. This study adds to the existing research on the PA paradox and PA parameters [54]. Overall, our findings validate and expand upon the PA paradox in the mental health domain, showing that LTPA, particularly low-intensity exercise in LTPA, is associated with lower depressive symptoms, whereas moderate-to-vigorous-intensity OPA is linked to higher depressive symptoms. These 2 distinct forms of PA exert opposite effects on depressive symptoms in middle-aged and older adults [51]. Leisure exercise and occupational labor, as fundamentally different drivers of PA, likely result in varying degrees of subjective willingness and well-being, which may have divergent long-term effects on mental health [55]. Particularly in many developing countries, moderate-to-vigorous-intensity OPA is often linked to limited individual job skills and is usually accompanied by poorer working conditions and lower levels of welfare protection.

When examining the dynamic effects of PA on depressive symptoms, this study further revealed the heterogeneous characteristics of different domains of PA. Using parallel-processed LGCM, we found that the initial level of total LTPA in middle-aged and older adults negatively predicted the initial level of depressive symptoms. This indicated that individuals with higher initial LTPA levels tended to have lower initial levels of depressive symptoms (Figure 2). This finding is consistent with the results of Heesch et al [56], who reported that low LTPA and walking levels were associated with reduced anxiety and depression scores. One possible explanation is that individuals with depressive symptoms are less likely to engage in PA. Additionally, confounding factors may have influenced this association.

This study also identified, for the first time, that OPA is not only significantly positively associated with the initial level of depressive symptoms, but also significantly predicts the trajectory of change in depressive symptoms. Specifically, middle-aged and older adults with higher initial OPA levels tended to exhibit higher initial levels of depressive symptoms. However, higher OPA levels appeared to suppress the upward trajectory of depressive symptoms over time (Figure 2). This finding highlights the complexity and duality of OPA's influence on the depressive symptom trajectory. On the one hand, high-intensity or repetitive OPA may increase physical burdens, contributing to higher initial depressive symptoms. On the other hand, the "steeling effect" resulting from prolonged OPA may enhance individuals' resilience to negative experiences and improve physiological adaptability, thereby mitigating the increase in depressive symptoms over time [57,58]. Moreover, OPA may exert long-term positive effects on mental health through mechanisms such as improving physical metabolism,

XSL•FO
RenderX

maintaining a regular daily routine, and reducing sedentary behavior. These benefits may partially offset the negative effects.

Effects of Community-Built Environment on Trajectories of Depressive Symptoms and Physical Activity

To the best of our knowledge, few studies have attempted to develop indicator systems or variable sets for international microsurvey databases to assess a community's physical environment. Jones-Smith and Popkin [59] proposed a scale for assessing community contextual characteristics but did not pay enough attention to physical built environments. Taking the CHARLS community questionnaire as an example, this study developed CBE indicators for international microsurvey databases, which can provide references for other studies. Specifically, incorporating the variables of the diversity of senior care facilities and sports venues directly addresses the unique needs of adults aged 45 years or older. Assessing the level of access to senior care and exercise facilities within the community is particularly important for this age group. The density of public facilities and accessibility of public facilities reflect the richness of community services and the convenience of facility distribution, providing valuable insights into middle-aged and older adults' ease of access to essential daily services. These factors are closely linked to the overall livability and quality of life within the community. Additionally, the infrastructure conditions variable covered several fundamental aspects critical to the community environment, such as road quality, waste disposal systems, sewer availability, and waste management practices. By comprehensively accounting for these infrastructure elements, this system effectively represents the community environmental realities in China.

Our study identified infrastructure conditions and accessibility to public facilities as the 2 most critical CBE factors influencing the trajectory of depressive symptoms among middle-aged and older adults (Figure 2). First, these factors significantly predicted the initial level of depressive symptoms; middle-aged and older adults residing in communities with better infrastructure conditions and higher accessibility to public facilities exhibited significantly lower initial levels of depressive symptoms. This finding aligns with those reported in the existing literature. Fan et al [60] highlighted the crucial role of community factors such as infrastructure and senior centers in mitigating depressive symptoms among middle-aged and older adults. Similarly, Li et al [61] found that a lack of infrastructure in villages was positively associated with the prevalence of depression in older adults in rural China. Moreover, the availability of recreational facilities in urban communities was associated with a lower incidence of depressive symptoms [62]. Perceived proximity to community facilities is an important factor influencing depressive symptoms among low-income older adults. Improving accessibility to such facilities may help alleviate depressive symptoms [63]. From a mechanistic perspective, high-quality infrastructure provides a safer and more comfortable living environment and reduces the sources of psychological stress. Concurrently, convenient public facilities promote social participation and PA, offering environmental support for mental well-being.

Through a long-term analysis of depressive symptom trajectories, our study not only confirmed the effect of accessibility to public facilities on the initial level of depressive symptoms but also revealed its predictive role in the upward trajectory of depressive symptoms. Contrary to expectations, greater accessibility to public facilities was found to significantly accelerate the upward trajectory of depressive symptoms among middle-aged and older adults in China (Figure 2). This can be attributed to 2 reasons. First, higher accessibility may be associated with greater community activity and more complex social interactions, which can introduce additional environmental stressors, such as noise, social conflicts, or other pressures that negatively affect mental health [64]. Second, the inadequate service quality of public facilities and insufficient consideration of the needs of vulnerable populations could also play a key role. Middle-aged and older adults may have higher expectations of accessible public facilities; however, if these facilities fail to meet their needs—due to issues such as overcrowding, usage restrictions, or poor design—it may lead to disappointment or frustration [65]. In developing countries, public facilities are often designed with a focus on quantity and coverage, whereas service capacity and the needs of disadvantaged groups are frequently overlooked. This limitation may exacerbate mental health challenges in middle-aged and older adults by introducing potential risk factors.

Our study further demonstrates the significant effect of CBE on the initial level of total LTPA, emphasizing the importance of infrastructure conditions and accessibility to public transport. The results indicated that middle-aged and older adults living in communities with better infrastructure or higher public transport accessibility had significantly higher initial LTPA levels. These findings provide new insights into how the built environment influences health behaviors among middle-aged and older adults. Additionally, middle-aged and older adults in communities with better infrastructure and higher public facility accessibility engaged in significantly lower initial OPA. Furthermore, infrastructure conditions positively predicted a declining trajectory in OPA levels over time (Figure 2). As residential communities are not direct venues for occupational activity, a plausible explanation is that infrastructure conditions and public facility accessibility reflect the overall quality, location advantages, and amenities provided within the community. In China, residents of higher-tier communities are typically part of the nonlabor force population, while manual laborers tend to reside in communities with poorer infrastructure and fewer amenities [66]. This community-level differentiation significantly impacts the intensity and type of occupational activity, reflecting the observed community segregation and spatial stratification within Chinese society [67]. However, due to the lack of data on community-level economic variables, this study could not fully account for the potential influence of these factors on the results. Further research is needed to validate these associations and explore the underlying mechanisms.

XSL•FO
RenderX

Implications

Our findings could offer valuable insights for public health management practices and community environmental planning, with relevance to other developing countries or regions. When formulating community PA guidelines, it is crucial to consider the distinct effects of different types of PA. LTPA should be integrated into community mental health intervention programs, with a focus on encouraging middle-aged and older adults to engage in activities such as walking. However, relevant authorities should also guide middle-aged and older adults in reducing the intensity of OPA, optimizing workplace conditions and workflows, and providing mental health support. This recommendation is particularly critical for developing countries, where physical labor constitutes a significant portion of the workforce.

From the perspective of CBE, improving community infrastructure and enhancing the accessibility of public facilities are key strategies for preventing and addressing mental health issues in middle-aged and older adults, while also promoting PA. This includes, but is not limited to, upgrading road quality, improving water and sewage systems, and enhancing electricity services. These measures can lead to better depressive symptom outcomes and encourage LTPA in middle-aged and older populations. Furthermore, improving public transportation accessibility has the potential to promote LTPA in this demographic. However, we recommend that governments carefully consider the potential environmental stressors associated with enhanced public facilities and focus on improving their service capacity and quality. Special attention should be given to addressing the needs of vulnerable populations to mitigate negative effects and more effectively promote community mental health.

Limitations and Future Research

First, we excluded participants with missing PA or depressive symptom scores. While this approach is common, it may introduce bias, as it is challenging to determine whether the missing values are associated with participants who have lower PA levels or higher depressive symptom scores. Second, although the findings offer critical empirical support for the PA paradox, the study is constrained by its design and data availability, with limited consideration of factors such as workplace environment, working conditions, and job autonomy. These factors may mediate or confound the relationship between PA and mental health. Future research should incorporate these variables to comprehensively elucidate the complex mechanisms linking OPA and mental health. Third, the wave 2020 data used in this study were collected during the COVID-19 pandemic, which may have influenced the results. Fourth, due to space constraints and the study's objectives, insufficient attention was given to subgroup analyses or population heterogeneity, which future research could explore in greater detail. Finally, the CBE data in the CHARLS were published for only 1 wave without subsequent follow-up, limiting this study's ability to examine changes in the role of CBE in the relationship between PA and depressive symptom trajectories. Future research should incorporate dynamic geographical information to capture the effects of CBE changes on mental health or PA trajectories.

Conclusions

Communities play an integral role in the daily lives of residents. This study offers valuable insights from developing countries by highlighting the effects of PA on depressive symptom trajectories and the role of CBE factors as physical determinants. The findings reveal that the impact of PA on depressive symptoms varies by domain and intensity. Overall, higher levels of LTPA were associated with lower depressive symptom levels among middle-aged and older adults, whereas OPA, particularly at moderate-to-vigorous intensities, had negative effects on depressive symptoms. Additionally, OPA exhibits a dual effect: while a higher initial level of total OPA was associated with a higher initial level of depressive symptoms, it also contributed to slowing the upward trajectory of depressive symptoms over time. Among the CBE factors, infrastructure conditions and accessibility to public facilities emerged as the most critical determinants influencing depressive symptom trajectories. Better infrastructure and greater accessibility to public facilities were linked to lower initial levels of depressive symptoms; however, greater accessibility to public facilities might exacerbate the upward trajectory of depressive symptoms. Furthermore, infrastructure conditions and accessibility to public transport played positive roles in promoting LTPA.

These findings provide novel perspectives on the PA paradox in mental health research, offering valuable guidance for developing more precise public health strategies. Such strategies should aim to promote PA types that effectively reduce depressive symptoms while minimizing potential adverse effects. By identifying the specific impacts of CBE factors, this study underscores that optimizing the built environment is not merely a physical enhancement but also a vital public health approach to improving residents' mental health. We advocate for enhancing community environments to encourage beneficial PA and alleviate depressive symptoms among middle-aged and older adults.

Acknowledgments

This work was supported by JST SPRING (grant JPMJSP2136). We sincerely thank the handling editor and all the reviewers for their time and effort, as their feedback and comments greatly contributed to improving this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Additional analyses.
[DOCX File , 258 KB - publichealth_v11i1e64564_app1.docx]

References

1. Sokale IO, Conway SH, Douphrate DI. Built environment and its association with depression among older adults: a systematic review. TOPHJ 2022 Mar 18;15(1):e187494452202030 [FREE Full text] [doi: 10.2174/18749445-v15-e2202030]
2. Koohsari M, Yasunaga A, McCormack G, Shibata A, Ishii K, Nakaya T, et al. Depression among middle-aged adults in Japan: the role of the built environment design. Landscape and Urban Planning 2023 Mar;231:104651. [doi: 10.1016/j.landurbplan.2022.104651]
3. Dinas PC, Koutedakis Y, Flouris AD. Effects of exercise and physical activity on depression. Ir J Med Sci 2011 Jun;180(2):319-325. [doi: 10.1007/s11845-010-0633-9] [Medline: 21076975]
4. Dogra S, Dunstan DW, Sugiyama T, Stathi A, Gardiner PA, Owen N. Active aging and public health: evidence, implications, and opportunities. Annu Rev Public Health 2022 Apr 05;43:439-459 [FREE Full text] [doi: 10.1146/annurev-publhealth-052620-091107] [Medline: 34910580]
5. Brundtland GH. From the World Health Organization. Mental health: new understanding, new hope. JAMA 2001 Nov 21;286(19):2391. [doi: 10.1001/jama.286.19.2391] [Medline: 11712923]
6. Zhang L, Xu Y, Nie H, Zhang Y, Wu Y. The prevalence of depressive symptoms among the older in China: a meta-analysis. Int J Geriatr Psychiatry 2012 Sep;27(9):900-906. [doi: 10.1002/gps.2821] [Medline: 22252938]
7. van den Bosch M, Meyer-Lindenberg A. Environmental exposures and depression: biological mechanisms and epidemiological evidence. Annu Rev Public Health 2019 Apr 01;40:239-259. [doi: 10.1146/annurev-publhealth-040218-044106] [Medline: 30633709]
8. Soini E, Rosenström T, Määttänen I, Jokela M. Physical activity and specific symptoms of depression: a pooled analysis of six cohort studies. J Affect Disord 2024 Mar 01;348:44-53 [FREE Full text] [doi: 10.1016/j.jad.2023.12.039] [Medline: 38128736]
9. De Mello MT, Lemos VDA, Antunes HKM, Bittencourt L, Santos-Silva R, Tufik S. Relationship between physical activity and depression and anxiety symptoms: a population study. J Affect Disord 2013 Jul;149(1-3):241-246. [doi: 10.1016/j.jad.2013.01.035] [Medline: 23489405]
10. Wassink-Vossen S, Collard RM, Penninx BW, Hiles SA, Oude Voshaar RC, Naarding P. The reciprocal relationship between physical activity and depression: does age matter? Eur Psychiatry 2018 Jun;51:9-15. [doi: 10.1016/j.eurpsy.2017.12.029] [Medline: 29510298]
11. Abdoli N, Salari N, Darvishi N, Jafarpour S, Solaymani M, Mohammadi M, et al. The global prevalence of major depressive disorder (MDD) among the elderly: a systematic review and meta-analysis. Neurosci Biobehav Rev 2022 Jan;132:1067-1073. [doi: 10.1016/j.neubiorev.2021.10.041] [Medline: 34742925]
12. Kong JY, Hong H, Kang H. Relationship between physical activity and depressive symptoms in older Korean adults: moderation analysis of muscular strength. BMC Geriatr 2022 Nov 21;22(1):884 [FREE Full text] [doi: 10.1186/s12877-022-03610-6] [Medline: 36411413]
13. Kim JH. Regular physical exercise and its association with depression: a population-based study short title: exercise and depression. Psychiatry Res 2022 Mar;309:114406. [doi: 10.1016/j.psychres.2022.114406] [Medline: 35074644]
14. Clays E, Lidegaard M, De Bacquer D, Van Herck K, De Backer G, Kittel F, et al. The combined relationship of occupational and leisure-time physical activity with all-cause mortality among men, accounting for physical fitness. Am J Epidemiol 2014 Mar 01;179(5):559-566. [doi: 10.1093/aje/kwt294] [Medline: 24305575]
15. Temporelli PL. Is physical activity always good for you? The physical activity paradox. Eur Heart J Suppl 2021 Oct;23(Suppl E):E168-E171 [FREE Full text] [doi: 10.1093/eurheartj/suab115] [Medline: 34650379]
16. He F, Li Y, Hu Z, Zhang H. Association of domain-specific physical activity with depressive symptoms: a population-based study. Eur Psychiatry 2022 Dec 12;66(1):e5 [FREE Full text] [doi: 10.1192/j.eurpsy.2022.2350] [Medline: 36503700]
17. Joo MJ, Jang YS, Jang YS, Park EC. Association between work-related physical activity and depressive symptoms in Korean workers: data from the Korea national health and nutrition examination survey 2014, 2016, 2018, and 2020. BMC Public Health 2023 Sep 08;23(1):1752 [FREE Full text] [doi: 10.1186/s12889-023-16631-6] [Medline: 37684616]
18. Oliveira AJ, Pancoti BM, Oliveira BRR, de Figueiredo JA, Paravidino V. Associations of occupational physical activity and active transport with depressive symptoms in a middle-income country. Public Health 2024 Jul;232:195-200. [doi: 10.1016/j.puhe.2024.04.033] [Medline: 38805868]
19. Schuch FB, Werneck AO, Vancampfort D, Stubbs B, Teychene M, Lotufo PA, et al. Cross-sectional associations of leisure and transport related physical activity with depression and anxiety. J Psychiatr Res 2021 Aug;140:228-234. [doi: 10.1016/j.jpsychires.2021.05.053] [Medline: 34118641]
20. Lopes MVV, Matias TS, da Costa BGG, Schuch FB, Chaput JP, Samara Silva K. The relationship between physical activity and depressive symptoms is domain-specific, age-dependent, and non-linear: an analysis of the Brazilian national health survey. J Psychiatr Res 2023 Mar;159:205-212. [doi: 10.1016/j.jpsychires.2023.01.041] [Medline: 36739848]
21. Lu S, Liu Y, Guo Y, Ho HC, Song Y, Cheng W, et al. Neighbourhood physical environment, intrinsic capacity, and 4-year late-life functional ability trajectories of low-income Chinese older population: A longitudinal study with the parallel process

XSL•FO
RenderX

of latent growth curve modelling. EClinicalMedicine 2021 Jun;36:100927 [FREE Full text] [doi: 10.1016/j.eclinm.2021.100927] [Medline: 34189445]
22. Nie P, Clark AE, D'Ambrosio C, Ding L. Income-related health inequality in urban China (1991-2015): the role of homeownership and housing conditions. Health Place 2022 Jan;73:102743. [doi: 10.1016/j.healthplace.2022.102743] [Medline: 35045352]
23. Xiao J, Zhao J, Luo Z, Liu F, Greenwood D. The impact of built environment on mental health: a COVID-19 lockdown perspective. Health Place 2022 Sep;77:102889 [FREE Full text] [doi: 10.1016/j.healthplace.2022.102889] [Medline: 36027740]
24. Lederbogen F, Kirsch P, Haddad L, Streit F, Tost H, Schuch P, et al. City living and urban upbringing affect neural social stress processing in humans. Nature 2011 Jun 22;474(7352):498-501. [doi: 10.1038/nature10190] [Medline: 21697947]
25. Ewing R, Cervero R. Travel and the built environment: a synthesis. Transportation Research Record: Journal of the Transportation Research Board 2001 Jan 01;1780(1):87-114. [doi: 10.3141/1780-10]
26. Wang P, Wang M, Shan J, Liu X, Jing Y, Zhu H, et al. Association between residential greenness and depression symptoms in Chinese community-dwelling older adults. Environ Res 2024 Feb 15;243:117869 [FREE Full text] [doi: 10.1016/j.envres.2023.117869] [Medline: 38070849]
27. Koohsari M, McCormack G, Nakaya T, Shibata A, Ishii K, Yasunaga A, et al. Urban design and Japanese older adults' depressive symptoms. Cities 2019 Apr;87:166-173. [doi: 10.1016/j.cities.2018.09.020]
28. Saarloos D, Alfonso H, Giles-Corti B, Middleton N, Almeida OP. The built environment and depression in later life: the health in men study. Am J Geriatr Psychiatry 2011 May;19(5):461-470. [doi: 10.1097/JGP.0b013e3181e9b9bf] [Medline: 20808136]
29. Gubbels JS, Kremers SPJ, Droomers M, Hoefnagels C, Stronks K, Hosman C, et al. The impact of greenery on physical activity and mental health of adolescent and adult residents of deprived neighborhoods: a longitudinal study. Health Place 2016 Jul;40:153-160. [doi: 10.1016/j.healthplace.2016.06.002] [Medline: 27322564]
30. Li X, Zhang W, Zhang W, Tao K, Ni W, Wang K, et al. Level of physical activity among middle-aged and older Chinese people: evidence from the China Health and Retirement Longitudinal Study. BMC Public Health 2020 Nov 10;20(1):1682 [FREE Full text] [doi: 10.1186/s12889-020-09671-9] [Medline: 33172439]
31. Fan M, Lyu J, He P. Chinese guidelines for data processing and analysis concerning the International Physical Activity Questionnaire. Zhonghua Liu Xing Bing Xue Za Zhi 2014 Aug;35(8):961-964. [Medline: 25376692]
32. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)-short and long forms. IPAQ. 2005. URL: http://www.ipaq.ki.se/scoring [accessed 2024-12-23]
33. Gao Y, Jia Z, Zhao L, Han S. The effect of activity participation in middle-aged and older people on the trajectory of depression in later life: national cohort study. JMIR Public Health Surveill 2023 Mar 23;9:e44682 [FREE Full text] [doi: 10.2196/44682] [Medline: 36951932]
34. China Health and Retirement Longitudinal Study (CHARLS) website. China Health and Retirement Longitudinal Study (CHARLS). URL: https://charls.pku.edu.cn/ [accessed 2024-12-12]
35. Park SH, Yu HY. How useful is the center for epidemiologic studies depression scale in screening for depression in adults? An updated systematic review and meta-analysis. Psychiatry Res 2021 Aug;302:114037. [doi: 10.1016/j.psychres.2021.114037] [Medline: 34098160]
36. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10(2):77-84. [Medline: 8037935]
37. Jacobson NS, Truax P. Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. J Consult Clin Psychol 1991 Feb;59(1):12-19. [doi: 10.1037//0022-006x.59.1.12] [Medline: 2002127]
38. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
39. Cervero R, Kockelman K. Travel demand and the 3Ds: density, diversity, and design. Transportation Research Part D: Transport and Environment 1997 Sep;2(3):199-219. [doi: 10.1016/S1361-9209(97)00009-6]
40. Cardinali M, Beenackers MA, van Timmeren A, Pottgiesser U. The relation between proximity to and characteristics of green spaces to physical activity and health: a multi-dimensional sensitivity analysis in four European cities. Environ Res 2024 Jan 15;241:117605 [FREE Full text] [doi: 10.1016/j.envres.2023.117605] [Medline: 37956752]
41. Cui X, Wen J, Yang H, Helbich M, Dijst M, Roberts H, et al. Pathways from street network design to symptoms of depression among emerging adults in China. Travel Behaviour and Society 2023 Oct;33:100644. [doi: 10.1016/j.tbs.2023.100644]
42. Zhang W, Wang T, Wang A. Impact of physical activity intensity on longitudinal trajectories of cognitive function and depressive symptoms in middle-aged and older Chinese adults: eight-year prospective study. J Affect Disord 2022 Oct 15;315:64-69. [doi: 10.1016/j.jad.2022.07.012] [Medline: 35839942]
43. Verbeke G, Fieuws S, Molenberghs G, Davidian M. The analysis of multivariate longitudinal data: a review. Stat Methods Med Res 2014 Feb;23(1):42-59. [doi: 10.1177/0962280212445834] [Medline: 22523185]

XSL•FO
RenderX

44. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
45. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
46. Meng Y, Luo Y, Qin S, Xu C, Yue J, Nie M, et al. The effects of leisure time physical activity on depression among older women depend on intensity and frequency. J Affect Disord 2021 Dec 01;295:822-830. [doi: 10.1016/j.jad.2021.08.142] [Medline: 34706452]
47. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
48. Mammen G, Faulkner G. Physical activity and the prevention of depression: a systematic review of prospective studies. Am J Prev Med 2013 Nov;45(5):649-657. [doi: 10.1016/j.amepre.2013.08.001] [Medline: 24139780]
49. Loprinzi PD. Objectively measured light and moderate-to-vigorous physical activity is associated with lower depression levels among older US adults. Aging Ment Health 2013;17(7):801-805. [doi: 10.1080/13607863.2013.801066] [Medline: 23731057]
50. Ryu J, Jeong A, Min JH, Lee DH, Lee J, Song IH, et al. The relationship between domain-specific physical activity and depressive symptoms in Korean adults: analysis of the Korea National Health and Nutrition Examination Survey. J Affect Disord 2022 Apr 01;302:428-434. [doi: 10.1016/j.jad.2022.01.097] [Medline: 35090946]
51. Werneck AO, Stubbs B, Szwarcwald CL, Silva DR. Independent relationships between different domains of physical activity and depressive symptoms among 60,202 Brazilian adults. Gen Hosp Psychiatry 2020;64:26-32. [doi: 10.1016/j.genhosppsych.2020.01.007] [Medline: 32086172]
52. Lin L, Halgin RP, Well AD, Ockene I. The relationship between depression and occupational, household, and leisure-time physical activity. Journal of Clinical Sport Psychology 2008;2(2):95-107. [doi: 10.1123/jcsp.2.2.95]
53. Nestor A, Henry Kofi M. The association between work-related physical activity and depression. Journal of Physical Activity Research 2017;2(1):1-6 [FREE Full text] [doi: 10.12691/jpar-2-1-1]
54. Coenen P, Huysmans MA, Holtermann A, Krause N, van Mechelen W, Straker LM, et al. Towards a better understanding of the 'physical activity paradox': the need for a research agenda. Br J Sports Med 2020 Sep;54(17):1055-1057. [doi: 10.1136/bjsports-2019-101343] [Medline: 32265218]
55. Luo M, Gupta N, Holtermann A, Stamatakis E, Ding D. Revisiting the 'physical activity paradox' in a Chinese context: occupational physical activity and mortality in 142,302 urban working adults from the China Kadoorie Biobank study. Lancet Reg Health West Pac 2022 Jun;23:100457. [doi: 10.1016/j.lanwpc.2022.100457] [Medline: 35602414]
56. Heesch KC, Burton NW, Brown WJ. Concurrent and prospective associations between physical activity, walking and mental health in older women. J Epidemiol Community Health 2011 Sep;65(9):807-813. [doi: 10.1136/jech.2009.103077] [Medline: 20515892]
57. Kok AAL, Twisk JWR, Blom F, Beekman ATF, Huisman M. Steeling or sensitizing? A longitudinal examination of how ongoing accumulation of negative life events affects depressive symptoms in older adults. J Gerontol B Psychol Sci Soc Sci 2021 Nov 15;76(10):2041-2053 [FREE Full text] [doi: 10.1093/geronb/gbab114] [Medline: 34171092]
58. Rutter M. Resilience in the face of adversity. Protective factors and resistance to psychiatric disorder. Br J Psychiatry 1985 Dec;147:598-611. [doi: 10.1192/bjp.147.6.598] [Medline: 3830321]
59. Jones-Smith JC, Popkin BM. Understanding community context and adult health changes in China: development of an urbanicity scale. Soc Sci Med 2010 Oct;71(8):1436-1446 [FREE Full text] [doi: 10.1016/j.socscimed.2010.07.027] [Medline: 20810197]
60. Fan VS, Mahadevan R, Leung J. Effect of income inequality, community infrastructure and individual stressors on adult depression. Health Promot Int 2021 Mar 12;36(1):46-57. [doi: 10.1093/heapro/daaa036] [Medline: 32277828]
61. Li LW, Liu J, Zhang Z, Xu H. Late-life depression in rural China: do village infrastructure and availability of community resources matter? Int J Geriatr Psychiatry 2015 Jul;30(7):729-736. [doi: 10.1002/gps.4217] [Medline: 25333218]
62. Deng Y, Paul DR. The relationships between depressive symptoms, functional health status, physical activity, and the availability of recreational facilities: a rural-urban comparison in middle-aged and older Chinese adults. Int J Behav Med 2018 Jun;25(3):322-330. [doi: 10.1007/s12529-018-9714-3] [Medline: 29498014]
63. Chen YY, Wong GHY, Lum TY, Lou VWQ, Ho AHY, Luo H, et al. Neighborhood support network, perceived proximity to community facilities and depressive symptoms among low socioeconomic status Chinese elders. Aging Ment Health 2016;20(4):423-431. [doi: 10.1080/13607863.2015.1018867] [Medline: 25775108]
64. Pineda VS. What is inclusive and accessible public space? jps 2022 Jun 26;7(2):5-8. [doi: 10.32891/jps.v7i2.1607]
65. Huang W, Lin G. The relationship between urban green space and social health of individuals: a scoping review. Urban Forestry & Urban Greening 2023 Jul;85:127969. [doi: 10.1016/j.ufug.2023.127969]
66. Simon A, John K, Lina S, Qingjie X. Contrasting paradigms: segmentation and competitiveness in the formation of the Chinese labour market. In: Unemployment, Inequality and Poverty in Urban China. London, UK: Routledge; Mar 10, 2006:228-251.

XSL•FO
RenderX

67. Zhou S, Deng L, Kwan M, Yan R. Social and spatial differentiation of high and low income groups' out-of-home activities in Guangzhou, China. Cities 2015 Jun;45:81-90. [doi: 10.1016/j.cities.2015.03.002]

Abbreviations

CBE: community-built environment
CES D: Center for Epidemiologic Studies Depression Scale
CHARLS: China Health and Retirement Longitudinal Survey
LGCM: latent growth curve modeling
LRT: likelihood ratio test
LTPA: leisure-time physical activity
MET: metabolic equivalent
OPA: occupational physical activity
PA: physical activity
RCI: Reliable Change Index
WHO: World Health Organization

Edited by A Mavragani; submitted 20.07.24; peer-reviewed by R Liang, Y Liu, S Fujihara, E Luo; comments to author 08.11.24; revised version received 08.12.24; accepted 11.12.24; published 13.01.25.

Please cite as:
Zhang K, Huang B, Divigalpitiya P
Identifying Community-Built Environment's Effect on Physical Activity and Depressive Symptoms Trajectories Among Middle-aged and Older Adults: Chinese National Longitudinal Study
JMIR Public Health Surveill 2025;11:e64564
URL: https://publichealth.jmir.org/2025/1/e64564
doi:10.2196/64564
PMID:39804686

©Kaili Zhang, Bowen Huang, Prasanna Divigalpitiya. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.01.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Factors Affecting Nonresponse Among Female Participants in the Korea Nurses' Health Study: Longitudinal Cohort Survey Study

Young Taek Kim[1], PhD; Chiyoung Cha[2], PhD; Gumhee Baek[2], MS; Bohye Kim[3], PhD; Bo Mi Song[4], PhD; Joong-Yeon Lim[4], PhD; Hyun-Young Park[5], PhD; Juh Hyun Shin[6], PhD

[1]Division for Gender-based Violence Research, Korean Women's Development Institute, Seoul, Republic of Korea

[2]College of Nursing, Ewha Research Institute of Nursing Science, Ewha Womans University, Seoul, Republic of Korea

[3]Nursing Workforce Support Center, Korean Nursing Association, Seoul, Republic of Korea

[4]Division of Population Health Research, Department of Precision Medicine, National Institute of Health, Cheongju, Republic of Korea

[5]National Institute of Health, Cheongju, Republic of Korea

[6]School of Nursing, George Washington University, Washington, DC, United States

Corresponding Author:
Chiyoung Cha, PhD
College of Nursing
Ewha Research Institute of Nursing Science
Ewha Womans University
#202 Helen Building
52 Ewhayeodae-gil, Seodaemun-gu
Seoul, 03760
Republic of Korea
Phone: 82 232772883
Fax: 82 232772850
Email: chiyoung@ewha.ac.kr

Abstract

Background: The major drawback of a cohort study design is the loss to follow-up, which increases selection bias and threatens external validity, particularly in online surveys. It is important to identify factors beyond population or demographics that influence nonresponse rates in cohort studies.

Objective: This study aimed to examine the nonresponse rate and associated factors over a 10-year follow-up period among female participants in the Korea Nurses' Health Study using data from the initial and subsequent surveys.

Methods: The Korea Nurses' Health Study recruited 20,613 female nurses in 2013 using simple random sampling. The participants were followed up 10 times through 2022. We identified the demographic, work-related, survey-related, and psychological characteristics of nonresponding nurses during the 10-year follow-up and compared them with those who continued to participate. Descriptive statistics, chi-square tests, and multivariate logistic regression models were used for the analysis.

Results: The nonresponse rate of the follow-up surveys from the 2nd to the 11th survey varied between 25.5% (5258/20,613; second survey) and 61.2% (12,620/20,613; sixth survey). The influence of age, education, and the usability of survey websites on nonresponse lasted up to the 11th survey. Nurses in their 20s were less likely to respond to the follow-up surveys than those in their 30s. Those who had an associate degree and neutral feelings about the usability of the survey websites were less likely to respond to the follow-up surveys than those who were satisfied with the initial survey. The influence of geographical region, hospital size, and psychological factors—including stress, fatigue, and sleep disturbance—was evident from the second to the sixth survey.

Conclusions: When designing and recruiting female nurse participants for community-based cohort studies, researchers should consider the factors that influence nonresponse and adopt tailored strategies based on demographic characteristics. In addition, improving the usability of survey websites is recommended to reduce nonresponses at follow-up in cohort studies involving female participants.

International Registered Report Identifier (IRRID): RR2-DOI: https://doi.org/10.4178/epih.e2024048

(JMIR Public Health Surveill 2025;11:e68038) doi:10.2196/68038

XSL•FO
RenderX

KEYWORDS

cohort studies; female; follow-up studies; logistic models; longitudinal studies; nonresponse; nurses

Introduction

Background

A cohort study was conducted in 1935 in the United States to determine the incidence of tuberculosis [1]. Despite their costly and time-consuming nature, cohort studies have been recognized as valuable longitudinal research methods. Cohort studies have been widely conducted to identify associations between known and unknown risk factors for specific diseases [2]. One of the major drawbacks of the cohort study design is the loss of participants at follow-up. With awareness of the differences in health disparities and determinants of health between men and women, cohort studies have been conducted to identify salient factors related to health and illness among women across several countries [3-5]. In 2013, the Korea Nurses' Health Study, a national cohort study of 20,613 female nurses aged between 20 and 45 years, was initiated to determine the sociodemographic, occupational, and lifestyle factors affecting the health and disease status of Korean women [6]. Follow-up surveys were conducted annually for more than 10 years, allowing the recognition of long-term patterns formed from early to middle adulthood. Participants accessed the Korea Nurses' Health Study through an online platform, with invitations delivered via phone, SMS text messages, or email.

Nonresponse is a significant issue in longitudinal studies, with response rates varying widely across different populations. In general, a response rate of 50% is considered acceptable [7]. Studies of nurses in Western countries have often reported response rates below 30% [3,8-10]. Similarly, health surveys in the Netherlands reported nonresponse rates of 53% [11], while the Thai Nurse Cohort Study reported nonresponse rates of 39.8% [12]. In Korea, the Nurses' Health Study, a prospective cohort study of women's health that has been ongoing for more than 10 years, has a nonresponse rate of 50% [13]. In this study, we examined the nonresponse rate and its associated factors in up to 10 years of follow-up surveys.

Literature Review

Nonresponse generally describes cases in which individuals do not participate in follow-up data collection [14] and the loss of participants over time, particularly in later data collection surveys, owing to factors such as loss of contact, incapacity, or refusal to continue participating [15]. Nonresponse can also occur when participants are lost to follow-up, choose to withdraw from the study, do not respond to follow-up surveys, or die during the study period [16]. This issue is particularly pronounced in online surveys, where a high rate of nonresponse increases the likelihood of nonrandom missing data [17]. Factors such as indifference to the study, rapidly changing health care environments, and illness might also contribute to nonresponse [3,17]. Previous cohort studies have examined factors related to nonresponse, with a focus on demographic factors, such as age [18], gender [19], education level [19-21], socioeconomic status [19], and health status [22]. However, participation in a survey can also be influenced by psychological factors, including the presence of stressful life events [23], depression, and anxiety disorders [21], along with participants' perceptions of the survey, which have rarely been studied. This is particularly significant, as most previous cohort studies identifying factors associated with nonresponse have focused on individuals with health issues [21,24] or older adult populations [25,26]. Moreover, survey satisfaction and the time required to complete the surveys are factors that can influence future participation in cohort studies. Participants who were satisfied with the survey process were more likely to participate in a follow-up survey [23]. Conversely, long or burdensome surveys may increase nonresponse in subsequent surveys [8]. Thus, a comprehensive understanding of psychological and survey-related factors is crucial to address nonresponse rates in cohort studies.

Methods

Design and Participants

The Korea Nurses' Health Study, initiated in 2013 with 20,613 participants, was originally developed based on the Nurses' Health Study 3 conducted in the United States [10], which is the oldest cohort study on women's health involving nurses as participants. For the Korea Nurses' Health Study, registered nurses aged 20 to 45 years working at hospitals across Korea at the time of data collection were recruited through simple random sampling. According to the Korean Nursing Association, 157,569 women of childbearing age were identified as nurses. The target sample size was determined based on population parameters and was calculated to be 20,000 female nurses, considering a 95% CI with a 1% permissible error [6]. After recruiting 20,613 registered nurses in the first survey in 2013, follow-up surveys were subsequently conducted: second survey (March 2014 to April 2019), third survey (November 2014 to April 2019), fourth survey (September 2015 to April 2019), fifth survey (June 2016 to April 2019), sixth survey (September 2017 to April 2019), seventh survey (December 2018 to April 2019), eighth survey (October 2019 to April 2020), ninth survey (October 2020 to April 2021), 10th survey (October 2021 to August 2022), and 11th survey (March 2022 to April 2023).

Participants in the initial survey were invited to participate in the follow-up surveys. This approach ensured high response rates without affecting the consistency of response and nonresponse rate calculations because the timing of survey availability did not influence participant responses [13]. While deaths during the study period were not specifically investigated, the number of nurses eligible to respond changed over time due to refusals to continue follow-up contact. Participants who left the nursing field during the study period were included in follow-up surveys to capture changes in health outcomes and ensure that the study reflected the broader scope of women's health. In addition, the participants were notified through SMS text messages and email reminders to encourage their participation. Surveys were accessible via both web and mobile platforms, ensuring convenience for participants.

XSL•FO
RenderX

Measures

On the basis of the literature, we included demographic and work-related characteristics, psychological variables, and survey-related characteristics from the initial survey to examine the factors associated with nonresponse in the follow-up surveys. The participants self-reported these characteristics throughout the survey.

Demographic and Work-Related Characteristics

The demographic characteristics included age and geographical region, categorized as Seoul (the capital), other metropolitan cities, and smaller cities or rural areas. Work-related characteristics included hospital size (number of beds); job position (manager or head nurse and charge or staff nurse); and work unit (special unit, general ward, or delivery room).

Psychological Variables

Four scales were used to measure the psychological variables: depression, stress, fatigue, and sleep. Depressive symptoms were measured using a nine-item, 4-point Likert scale from the Patient Health Questionnaire [27]. Possible scores range from 0 to 27, with higher scores indicating more severe depressive symptoms. On the basis of a previous study, a cutoff score of 10 was used for no to mild and moderate to severe symptoms [27]. Its sensitivity and specificity were 88% [27]. In this study, the Cronbach α was 0.892. Stress was measured using the Perceived Stress Scale (PSS)-4 [28], which consists of four 5-point Likert-type items. The possible scores range from 0 to 16, with higher scores indicating higher stress levels. The cutoff scores were determined based on the literature, and the top tertile of the scores (≥11) was categorized as *high stress*, with the other tertiles considered as *low stress* [29]. The PSS-4 has been confirmed to be reliable and valid compared to the original 14-item PSS [28]. Cronbach α was 0.691. Fatigue was measured using the Chalder Fatigue Scale [30], which comprises 11 items rated on a 4-point Likert scale. The possible scores range from 0 to 33, with higher scores indicating higher levels of fatigue. Scores between 0 and 16 were categorized as low, those between 17 and 21 as moderate, and those above 22 as high fatigue [30]. In this study, Cronbach α was 0.918. Sleep disturbance was measured using the Jenkins Sleep Questionnaire [31], which evaluates sleep problems during the preceding 4 weeks. This scale comprises 4 items rated on a 6-point Likert-type scale. Scores range from 0 to 20, with higher scores indicating greater sleep disturbance. On the basis of the cutoff point of 12 or more, which indicates severe sleep problems, participants were classified into groups with little sleep disturbance and those with high frequency of sleep disturbance (<12 and ≥12) [31]. In this study, the Cronbach α was 0.875. The collinearity of the 4 psychological variables was tested using the Pearson correlation coefficient (r<0.4 for all 4 psychological variables).

Survey-Related Characteristics

Survey-related characteristics included 3 items: time spent on survey completion, usability of the survey website, and overall impression of the survey website. Time spent completing the survey was categorized as <15 minutes, 15 to 30 minutes, and ≥30 minutes. Usability and overall impressions of the survey website were assessed using end-of-survey questions, such as "What is your overall impression of this survey website?" and "What was your experience using the survey website?" The responses were categorized as unsatisfactory, neutral, or satisfactory.

Data Analysis

Descriptive statistics were used to examine nonresponse rates and factors. Chi-square tests were used to compare the respondent and nonrespondent groups using initial survey data. A fully adjusted logistic regression model with available variables from the initial survey was used to identify factors associated with nonresponse. Missing data were excluded from analyses involving these variables. SPSS software (version 29; IBM Corp) was used for statistical analysis.

Ethical Considerations

For the initial survey in 2013 and the follow-up surveys until 2015, the institutional review board of the National Institute of Health in Korea reviewed and approved the study (#2013-03CON-03-P, 2014-05EXP-01-1C-A, 2014-08EXP-01-2C-A, and 2015-05-EXP-04-3C-A). For the follow-up surveys conducted in 2016, the Ewha Womans University Institutional Review Board reviewed and approved the study (#117-4, EWURB-17-3.0-20170901, ewha-201904-0012-01, ewha-201904-0012-09, ewha-201904-0012-12, and ewha-201904-0012-15). The Korea Nurses' Health Study survey instrument was meticulously designed by integrating questions from various validated sources to ensure comprehensive coverage of relevant topics. The quality of the survey was ensured through a rigorous pilot testing phase with nurses and expert groups to verify the clarity, relevance, and logical sequence of the questions. This process helped identify and rectify potential issues, such as redundancy and omission of important inquiries. No essential items from the original tools were omitted during the integration of questions from multiple sources. In cases where redundancy was identified, questions were asked only once to streamline the survey without compromising its psychometric properties. The questionnaire construction and refinement of the Korea Nurses' Health Study have been described in greater detail elsewhere [6]. The participants provided informed electronic consent before participating in the baseline and follow-up surveys. The participants were informed of the research purpose, anonymity, and confidentiality. All data were fully deidentified before analysis, with access restricted to authorized investigators through secure servers in compliance with Korea Nurses' Health Study data protection policies. Individual-level information was neither disclosed nor reported. Beverage vouchers were provided as an incentive.

Results

Overview

The nonresponse rates of the participants in the Korea Nurses' Health Study are shown in Table 1. Participants from the baseline survey (N=20,613) were contacted for the follow-up surveys. The nonresponse rate varied between 25.5% (5258/20,613; second survey) and 61.2% (12,620/20,613; sixth

XSL•FO
RenderX

survey), with the sixth survey having the highest nonresponse rate.

Table 1. Nonresponse rate of the Korea Nurses' Health Study participants (first survey: N=20,613)

Survey number	Nonresponding participants, n (%)
2	5258 (25.5)
3	7765 (37.7)
4	9963 (48.3)
5	9086 (44.1)
6	12,620 (61.2)
7	11,955 (58.0)
8	10,360 (50.3)
9	9957 (48.3)
10	10,503 (51)
11	10,357 (47.9)

Comparison of Respondent and Nonrespondent Groups

Differences between the respondent and nonrespondent groups from the second to the eleventh survey were compared based on participant characteristics from the first survey (Multimedia Appendix 1). In the initial survey, most participants were in their 20s (12,059/20,613, 58.5%) with a mean age of 29.4 (SD 5.9) years. All participants worked in hospitals, and 45.2% (9316/20,613) held a bachelor degree [13]. Age, education, stress, and perceived usability of the survey at baseline were found to be statistically different across the follow-up surveys, up to the 11th survey. Overall impressions of the survey website at baseline differed significantly in all follow-up surveys except the second. Nurses dissatisfied with the usability of the survey responded more to the second to fourth surveys than satisfied nurses, and nurses dissatisfied with the overall survey responded more to the third and fourth surveys. Geographical region and fatigue level in the baseline survey differed significantly between the respondent and nonrespondent groups up to the seventh survey. Hospital size at baseline was significantly associated with the response status in the fifth survey. The working units significantly differed until the third survey was conducted.

Factors Affecting Nonresponse

A fully adjusted logistic regression analysis was performed to identify the factors associated with nonresponse in the follow-up surveys (Tables 2 and 3). Age, education level, and perceived usability of the survey website at baseline were significant factors that showed a consistent association with nonresponse throughout the 11th survey. In the third survey, nurses in their 20s were less likely to respond compared to those in their 30s (odds ratio [OR] 0.93, 95% CI 0.87-1.00); this association remained significant in all surveys through the 11th, except the second. In the second survey, nurses with an associate degree were less likely to respond than those with a bachelor degree (OR 0.84, 95% CI 0.79-0.91) or a master degree or higher (OR 0.75, 95% CI 0.64-0.88); this pattern remained significant through the 11th survey. Nurses who reported neutral feelings about the usability of the survey website were more likely not to respond in the second survey compared to those who perceived the survey website to be useful in the second survey (OR 0.90, 95% CI 0.81-0.99), and this association was consistently significant until the 11th survey.

The association between geographical region and hospital size lasted until the fifth and sixth surveys. In the second survey, nurses working in Seoul were more likely to respond compared to those working in metropolitan areas (OR 1.16, 95% CI 1.06-1.27) and other regions (OR 1.17, 95% CI 1.08-1.27). This association remained consistently significant until the sixth survey was conducted. Nurses who worked in hospitals with more than 300 beds were less likely to respond to the second survey compared to their counterparts (OR 1.40, 95% CI 1.29-1.52); this association was consistently significant until the fifth survey.

Age at baseline was a significant factor for nonresponse throughout the follow-up surveys, except for the second survey. Nurses in their 30s were less likely to be nonrespondents than those in their 20s in the third survey (OR 0.93, 95% CI 0.868-0.999), and this was consistently significant up to the 11th survey (OR 0.82, 95% CI 0.764-0.875).

XSL•FO
RenderX

Table 2. Fully adjusted logistic regression of second to sixth surveys for nonresponse in the Korea Nurses' Health Study participants[a].

Variables	Second survey, OR (95% CI)	Third survey, OR (95% CI)	Fourth survey, OR (95% CI)	Fifth survey, OR (95% CI)	Sixth survey, OR (95% CI)
Age (y; reference: 20 to 29)					
30 to 39	1.02 (0.94-1.10)	0.93[b] (0.87-1.00)	0.87[c] (0.81-0.93)	0.83[d] (0.78-0.90)	0.77[c] (0.72-0.82)
≥40	0.90 (0.77-1.04)	0.92 (0.80-1.06)	1.13 (1.00-1.30)	1.00 (0.87-1.14)	0.87 (0.76-1.00)
Geographical region (reference: Seoul)					
Metropolitan cities	1.16[d] (1.06-1.27)	1.24[c] (1.15-1.34)	1.36[c] (1.26-1.47)	1.26[c] (1.17-1.36)	1.17[c] (1.08-1.27)
Others	1.17[c] (1.08-1.27)	1.22[c] (1.13-1.31)	1.20[c] (1.11-1.29)	1.17[c] (1.09-1.26)	1.13[d] (1.04-1.21)
Education (reference: associate degree)[c]					
Bachelor degree	0.84 (0.79-0.91)	0.85 (0.80-0.91)	0.84 (0.79-0.90)	0.82 (0.77-0.87)	0.79 (0.74-0.84)
Master degree or higher	0.75 (0.64-0.88)	0.72 (0.63-0.83)	0.76 (0.67-0.87)	0.75 (0.66-0.86)	0.77 (0.67-0.88)
Hospital size (reference: ≥300 beds)					
<300 beds	1.40[c] (1.29-1.52)	1.36[c] (1.26-1.46)	1.24[c] (1.16-1.34)	1.15[c] (1.07-1.24)	1.04 (0.96-1.12)
Job role (reference: charge or staff nurse)					
Manager or head nurse	1.04 (0.87-1.25)	0.99 (0.84-1.16)	1.09 (0.93-1.27)	1.05 (0.90-1.23)	1.13 (0.96-1.32)
Work unit type (reference: general ward)					
Special unit	1.01 (0.94-1.09)	0.97 (0.91-1.04)	0.98 (0.92-1.05)	1.03 (0.96-1.10)	1.03 (0.97-1.11)
Delivery room or others	1.11[b] (1.01-1.23)	1.07 (0.98-1.17)	1.04 (0.95-1.13)	1.06 (0.97-1.15)	1.05 (0.96-1.14)
Depressive symptoms (reference: <10)					
≥10	1.00 (0.92-1.09)	0.98 (0.91-1.06)	0.98 (0.91-1.06)	0.94 (0.87-1.02)	0.91[b] (0.84-0.99)
Stress (reference: <11)					
≥11	1.25[b] (1.05-1.48)	1.16 (0.99-1.35)	1.12 (0.96-1.30)	1.17[b] (1.00-1.37)	1.19[b] (1.01-1.41)
Fatigue (reference: approximately 17 to 21)					
<17	0.89[d] (0.82-0.96)	0.92[b] (0.86-0.99)	0.90[d] (0.84-0.97)	0.95 (0.89-1.03)	0.95 (0.89-1.03)
≥22	1.04 (0.95-1.14)	1.08 (1.00-1.18)	1.05 (0.97-1.14)	1.10[b] (1.01-1.19)	1.12[c] (1.03-1.21)
Sleep disturbance (reference: <12)					
≥12	0.88[d] (0.80-0.96)	0.94 (0.87-1.02)	0.93 (0.86-1.01)	0.97 (0.89-1.05)	0.98 (0.90-1.07)
Survey response time (min; reference: <15)					
15 to 30	0.99 (0.92-1.06)	0.97 (0.91-1.03)	0.97 (0.91-1.04)	0.98 (0.92-1.04)	1.02 (0.95-1.09)
≥30	0.96 (0.83-1.12)	0.95 (0.83-1.08)	0.98 (0.86-1.11)	1.02 (0.90-1.16)	1.04 (0.91-1.19)
Usability of the survey website (reference: neutral)					
Satisfied	0.90[b] (0.81-0.99)	0.83[c] (0.76-0.90)	0.85[c] (0.79-0.93)	0.82[c] (0.75-0.89)	0.85[c] (0.78-0.93)
Unsatisfied	0.84 (0.65-1.10)	0.82 (0.65-1.04)	0.89 (0.71-1.11)	1.02 (0.82-1.28)	0.95 (0.75-1.19)
Comprehensive feelings on the survey (reference: neutral)					
Satisfied	1.02 (0.92-1.12)	1.04 (0.96-1.14)	1.01 (0.93-1.10)	0.97 (0.89-1.05)	0.92 (0.84-1.00)
Unsatisfied	1.07 (0.86-1.34)	1.13 (0.93-1.38)	1.04 (0.86-1.26)	1.214[b] (1.00-1.47)	1.25[b] (1.02-1.54)

[a]Nagelkerke R^2 values: 0.015 (second survey), 0.018 (third survey), 0.018 (fourth survey), 0.018 (fifth survey), 0.019 (sixth survey). Model fit *P* values were all <.001 across surveys.

[b]*P*<.05.

[c]*P*<.001.

[d]*P*<.01.

XSL•FO
RenderX

Table 3. Fully adjusted logistic regression of 7th to 11th surveys for nonresponse in the Korea Nurses' Health Study participants[a].

Variables	7th survey, OR (95% CI)	8th survey, OR (95% CI)	9th survey, OR (95% CI)	10th survey, OR (95% CI)	11th survey, OR (95% CI)
Age (y; reference: 20 to 29)					
30 to 39	0.84[b] (0.78-0.90)	0.88[b] (0.82-0.94)	0.82[b] (0.76-0.88)	0.93[c] (0.87-1.00)	0.82[b] (0.76-0.88)
≥40	1.02 (0.89-1.16)	1.04 (0.91-1.19)	0.85[b] (0.74-0.97)	1.02 (0.90-1.17)	0.97 (0.85-1.10)
Geographical region (reference: Seoul)					
Metropolitan cities	1.09[c] (1.01-1.17)	1.02 (0.95-1.10)	1.05 (0.97-1.13)	0.97 (0.90-1.05)	0.98 (0.91-1.06)
Others	1.03 (0.96-1.11)	0.98 (0.92-1.06)	1.03 (0.96-1.11)	1.00 (0.93-1.07)	0.99 (0.92-1.06)
Education (reference: associate degree)					
Bachelor degree[b]	0.81 (0.76-0.87)	0.84 (0.79-0.90)	0.85 (0.80-0.91)	0.85 (0.80-0.90)	0.82 (0.77-0.87)
Master degree or higher	0.80[d] (0.70-0.91)	0.87[c] (0.76-0.99)	0.84[c] (0.74-0.96)	0.79[b] (0.69-0.90)	0.82[d] (0.72-0.94)
Hospital size (reference: ≥300 beds)					
<300 beds	1.05 (0.97-1.13)	1.02 (0.94-1.09)	1.07 (1.00-1.15)	0.99 (0.92-1.06)	1.01 (0.94-1.09)
Job role (reference: charge or staff nurse)					
Manager or head nurse	1.11 (0.95-1.31)	1.04 (0.89-1.22)	1.03 (0.88-1.21)	0.96 (0.82-1.13)	1.03 (0.88-1.21)
Work unit type (reference: general ward)					
Special unit	1.06 (0.99-1.13)	1.03 (0.97-1.10)	1.05 (0.99-1.12)	1.04 (0.98-1.11)	1.07[c] (1.01-1.15)
Delivery room or others	0.94 (0.87-1.03)	1.00 (0.92-1.09)	0.97 (0.89-1.05)	0.96 (0.88-1.04)	0.98 (0.90-1.06)
Depressive symptoms (reference: <10)					
≥10	0.93 (0.86-1.00)	0.96 (0.89-1.04)	0.98 (0.91-1.06)	0.93 (0.86-1.00)	0.96 (0.89-1.04)
Stress (reference: <11)					
≥11	1.20[c] (1.02-1.41)	1.20[c] (1.03-1.40)	1.09 (0.94-1.28)	1.16 (0.99-1.35)	1.10 (0.95-1.28)
Fatigue (reference: approximately 17 to 21)					
<17	0.96 (0.89-1.03)	1.01 (0.94-1.08)	1.00 (0.93-1.07)	1.04 (0.97-1.12)	0.96 (0.90-1.04)
≥22	1.05 (0.97-1.14)	1.05 (0.97-1.13)	1.03 (0.96-1.12)	1.07 (0.94-1.11)	1.00 (0.92-1.08)
Sleep disturbance (reference: <12)					
≥12	0.96 (0.89-1.05)	0.98 (0.91-1.06)	1.01 (0.93-1.09)	1.03 (0.95-1.12)	1.02 (0.94-1.11)
Survey response time (min; reference: <15)					
15 to 30	1.03 (0.97-1.10)	1.02 (0.95-1.08)	0.99 (0.93-1.06)	0.99 (0.93-1.05)	0.98 (0.92-1.05)
≥30	0.98 (0.86-1.11)	1.02 (0.90-1.16)	0.90 (0.79-1.02)	0.99 (0.87-1.13)	1.01 (0.88-1.14)
Usability of the survey website (reference: neutral)					
Satisfied	0.83[b] (0.76-0.90)	0.84[b] (0.77-0.92)	0.86[b] (0.79-0.93)	0.89[d] (0.82-0.97)	0.89[d] (0.82-0.97)
Unsatisfied	1.15 (0.91-1.45)	1.16 (0.93-1.46)	1.14 (0.91-1.42)	1.30[c] (1.04-1.63)	1.06 (0.85-1.33)
Comprehensive feelings on the survey (reference: neutral)					
Satisfied	0.93 (0.86-1.02)	0.96 (0.88-1.04)	0.92 (0.85-1.01)	0.93 (0.86-1.02)	0.93 (0.85-1.01)
Unsatisfied	1.26[c] (1.03-1.54)	1.21 (1.00-1.47)	1.17 (0.97-1.42)	1.13 (0.93-1.37)	1.15 (0.95-1.40)

[a]Nagelkerke R^2 values: 0.015 (7th survey), 0.009 (8th survey), 0.012 (9th survey), 0.008 (10th survey), 0.011 (11th survey). Model fit *P* values were all <.001 across surveys.

[b]*P*<.001.

[c]*P*<.05.

[d]*P*<.01.

XSL•FO
RenderX

Geographical region was significantly related to nonresponse until the seventh survey, and hospital size was significant until the fifth survey. In the second survey, nurses working in metropolitan cities (OR 1.16, 95% CI 1.06-1.27) and other areas (OR 1.17, 95% CI 1.08-1.27) were more likely to be nonrespondents than those working in Seoul. This was consistently significant up to the sixth survey in other areas (OR 1.13, 95% CI 1.04-1.21) and up to the seventh survey in metropolitan cities (OR 1.09, 95% CI 1.01-1.17). Nurses working in hospitals with fewer than 300 beds were more likely to be nonrespondents than their counterparts in the second survey (OR 1.40, 95% CI 1.29-1.52), which was consistently significant up to the fifth survey (OR 1.15, 95% CI 1.07-1.24).

Psychological factors, including fatigue, stress, and sleep disturbance, were significant non–response-related factors in some follow-up surveys, but did not show consistency across follow-up surveys.

Finally, it has been reported that the demographic characteristics of participants, such as gender and ethnicity, may influence their responses to cohort studies [17]. However, the Korea Nurses' Health Study participants comprised only female Korean nurses, leaving little room to explain the influence of gender and ethnicity on nonresponse in community-based cohort studies and future studies.

Discussion

Principal Findings

In this study, we reported the nonresponse rate and associated factors for more than 10 follow-up surveys. Our findings are crucial for informing future research on the factors that should be considered when designing and recruiting cohort studies, especially for women and nurses. Cohort study results should be interpreted carefully, considering random nonresponse factors. Thus, researchers must report nonresponses that occur during the study period when synthesizing studies using quality evaluation to provide a high level of evidence [32].

As of March 2025, the ongoing Korea Nurses' Health Study has conducted more than 10 follow-up surveys, with nonresponse rates ranging from 25% (5258/20,613) to 61% (12,620/20,613) and averaging around 50% (10,306/20,613). Unexpectedly, nonresponse peaked at the sixth survey from September 2017 to September 2019, although there was no change in the recruitment method. This may be attributed to reduced promotional activities following a change in the managing research institution in 2018. After the change in the research institute, we started the seventh survey and focused on promotional activities to increase response rates.

In this study, the age and education level of the nurses and the perceived usability of the survey website from the baseline survey were nonresponse factors throughout the 11th follow-up survey. Fewer nurses in their 20s at the baseline survey responded to the follow-up surveys than those in their 30s. This may be related to career stability and professional confidence, which often increase with experience [33-35]. Conversely, younger nurses may face greater challenges when transitioning from student to professional nurses [36,37], considering that the mean age of newly graduated nurses is 24.7 years in Korea [38], which is much younger than the 29 years in the United States [39]. Notably, nurses aged more than 40 years did not show significantly different nonresponse rates from those in their 20s, indicating that factors other than age, such as career development, leadership roles in the organization, and life-work balance, may influence survey response at follow-up. These findings suggest the need for supportive strategies tailored to different career stages to enhance participation in longitudinal studies.

Consistent with previous studies [25,40-43], educational level was associated with nonresponse. Nurses with associate degrees were less likely to respond to follow-up surveys than those with higher education levels. Targeting more participants with less education may be a strategy for retaining this group in follow-up surveys of cohort studies. Contrary to our findings, 1 longitudinal study that investigated nonresponse to follow-up surveys among an aging population did not report education as a factor related to nonresponse [26]. However, the influence of education may not be applicable to older adults in cohort studies.

In addition, nurses working in hospitals with more than 300 beds and those located in Seoul were more likely to respond to the follow-up surveys through the fifth and sixth rounds, respectively, compared to their counterparts. Most large hospitals are located in Seoul, and university-affiliated hospitals and large corporation-operated hospitals may have more professional and systematic organizational support for nurses [44,45], which might have influenced their value in health research. However, because nurses may change their jobs or relocate, baseline information on location and workplace may not retain its predictive value in later follow-ups.

Although the association between psychological factors, such as depression, stress, fatigue, sleep disturbance, and nonresponse in the follow-up surveys, was not consistent in this study, there is some evidence of psychological factors influencing the follow-up surveys in the literature. Studies have reported that experiencing stressful life events has been shown to result in lower subsequent follow-up responses [40], and respondents reported better subjective health and healthier lifestyle behaviors than nonrespondents [40,46]. Therefore, further systematic research on this topic is required.

Interestingly, the time taken to complete the survey was not a significant factor in nonresponse; rather, the perceived usability of the survey website was related to nonresponse. This suggests that participants were willing to return to the survey regardless of the length of time spent; however, technical issues in completing the surveys might be a barrier to participation in future research. The presentation of survey websites is an important but often overlooked part of data collection, as it can influence not only the response rate but also the measurement error [47]. Researchers who use websites to deliver surveys in cohort studies should carefully consider the wording and display of questions, including question ordering, to increase the response rates [48].

Strategies for participant recruitment and retention in longitudinal research should be specifically developed to minimize the dropout risk. Current suggestions for designing

XSL•FO
RenderX

cohort studies focus on recruiting and maintaining retention while considering the education, age, and psychometric characteristics of registered nurses. Further research on ethnic minority groups [49] is required worldwide. The factors associated with the nonresponse rates in this study should also be considered when designing cohort studies involving women. As nonresponse rates differed according to demographic factors such as age, education, and geographical region, tailored strategies should be developed to lower the nonresponse rate of participants with specific demographic characteristics. Examples of these strategies include posting flyers in certain regions, such as small cities, retrieving participants, and advertising the research through social networking services to gain access to those who are younger and who do not live or work in metropolitan areas.

Limitations

The data were measured only at baseline, which limited the ability to assess the influence of geographical location changes over time, such as participants moving from Seoul to metropolitan or urban areas in subsequent survey waves. Furthermore, the statistical analysis of confounding variables may yield further information regarding nonresponse. We could not identify participants who died during the follow-up period. However, participants who had left the nursing profession were included because this study focused on women's health rather than their professional roles. In addition, 3 to 3.7% of participants could not be reached due to lost contact information.

Conclusions

The nonresponse factors identified by the Korea Nurses' Health Study during the 10-year follow-up survey were age, education, and comprehensive feelings about the survey. We suggest that future research examine the extent of nonresponse bias and its potential impact on the observed associations to further enhance the validity and reliability of the findings. The implementation of weighting methods to account for nonresponse may also be considered in subsequent analyses. Customized strategies that reflect these factors should be developed and tested in future studies to increase follow-up survey response rates.

Acknowledgments

The authors would like to thank all the nurses who completed the survey and made this study possible. This study was funded by the Korea Disease Control and Prevention Agency (grants 2016ER630500, 2016ER630501, and 2022ER0602-01).

Authors' Contributions

YTK conducted statistical analyses and drafted the manuscript. JHS contributed to the methods, results, and discussion sections. CC drafted the manuscript and handled the project administration, supervision, validation, and review. GB contributed to the results and discussion sections. BK was responsible for writing the introduction and methods sections. BMS, JYL, and HYP were involved in the project administration, supervision, validation, and review. All the authors wrote, reviewed, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Differences in baseline characteristics from second to sixth and from 7th to 11th surveys between respondent and nonrespondent groups among Korea Nurses' Health Study participants (N=20,613).
[DOCX File , 82 KB - publichealth_v11i1e68038_app1.docx]

References

1. Doll R. Cohort studies: history of the method I. prospective cohort studies. Soz Präventivmed 2001 Mar;46(2):75-86. [doi: 10.1007/bf01299724]
2. Omair A. Selecting the appropriate study design: case–control and cohort study designs. J Health Spec 2016 Aug 09;4(1):37. [doi: 10.4103/1658-600x.173842]
3. Hundrup YA, Simonsen MK, Jørgensen T, Obel EB. Cohort profile: the Danish nurse cohort. Int J Epidemiol 2012 Oct 18;41(5):1241-1247. [doi: 10.1093/ije/dyr042] [Medline: 21421694]
4. Colditz GA. The nurses' health study: a cohort of US women followed since 1976. J Am Med Womens Assoc (1972) 1995 Apr;50(2):40-44. [Medline: 7722205]
5. Hayashi K, Ideno Y, Nagai K, Lee JS, Yasui T, Kurabayashi T, et al. Complaints of reduced cognitive functioning during perimenopause: a cross-sectional analysis of the Japan Nurses' Health Study. Womens Midlife Health 2022 Jun 05;8(1):6 [FREE Full text] [doi: 10.1186/s40695-022-00076-9] [Medline: 35659793]
6. Kim O, Ahn Y, Lee HY, Jang HJ, Kim S, Lee JE, et al. The Korea Nurses' Health Study: a prospective cohort study. J Womens Health (Larchmt) 2017 Aug;26(8):892-899 [FREE Full text] [doi: 10.1089/jwh.2016.6048] [Medline: 28771383]
7. Babbie ER. Survey Research Methods. Belmont, CA: Wadsworth Publishing Company; 1973.

XSL•FO
RenderX

8. Kaplowitz MD, Hadlock TD, Levine R. A comparison of web and mail survey response rates. Public Opinion Q 2004 Apr 22;68(1):94-101. [doi: 10.1093/poq/nfh006]
9. Lawson CC, Johnson CY, Chavarro JE, Lividoti Hibert EN, Whelan EA, Rocheleau CM, et al. Work schedule and physically demanding work in relation to menstrual function: the Nurses' Health Study 3. Scand J Work Environ Health 2015 Mar 30;41(2):194-203 [FREE Full text] [doi: 10.5271/sjweh.3482] [Medline: 25634477]
10. Bao Y, Bertoia ML, Lenart EB, Stampfer MJ, Willett WC, Speizer FE, et al. Origin, methods, and evolution of the three nurses' health studies. Am J Public Health 2016 Sep;106(9):1573-1581. [doi: 10.2105/ajph.2016.303338]
11. Mölenberg FJ, de Vries C, Burdorf A, van Lenthe FJ. A framework for exploring non-response patterns over time in health surveys. BMC Med Res Methodol 2021 Feb 18;21(1):37 [FREE Full text] [doi: 10.1186/s12874-021-01221-0] [Medline: 33602123]
12. Sawaengdee K, Tangcharoensathien V, Theerawit T, Thungjaroenkul P, Thinkhamrop W, Prathumkam P, et al. Thai nurse cohort study: cohort profiles and key findings. BMC Nurs 2016 Feb 17;15(1):10 [FREE Full text] [doi: 10.1186/s12912-016-0131-0] [Medline: 26893589]
13. Cha C, Jung H, Kim S, Lee JE, Ko KP, Cho E, et al. Korea Nurses' Health Study and the health of reproductive-aged women: a cohort profile. Epidemiol Health 2024 Apr 30;46:e2024048 [FREE Full text] [doi: 10.4178/epih.e2024048] [Medline: 38697860]
14. Bu F. Non-response and attrition in longitudinal studies. J Epidemiol Community Health 2022 Oct 28;76(12):971. [doi: 10.1136/jech-2022-219861] [Medline: 36307201]
15. Brilleman SL, Pachana NA, Dobson AJ. The impact of attrition on the representativeness of cohort studies of older people. BMC Med Res Methodol 2010 Aug 05;10(1):71 [FREE Full text] [doi: 10.1186/1471-2288-10-71] [Medline: 20687909]
16. Costello H, Hayes GM, Francis PT. Factors affecting withdrawal and donation attrition in the brains for dementia research cohort. Int J Geriatr Psychiatry 2018 Dec 12;33(12):1709-1716. [doi: 10.1002/gps.4975] [Medline: 30209830]
17. Yu T, Chen J, Gu NY, Hay JW, Gong CL. Predicting panel attrition in longitudinal HRQoL surveys during the COVID-19 pandemic in the US. Health Qual Life Outcomes 2022 Jul 06;20(1):104 [FREE Full text] [doi: 10.1186/s12955-022-02015-8] [Medline: 35794553]
18. Stewart AC, Cossar R, Walker S, Wilkinson AL, Quinn B, Dietze P, et al. Strategies to maximise study retention and limit attrition bias in a prospective cohort study of men reporting a history of injecting drug use released from prison: the prison and transition health study. BMC Med Res Methodol 2021 Sep 12;21(1):185 [FREE Full text] [doi: 10.1186/s12874-021-01380-0] [Medline: 34511067]
19. Matthews FE, Chatfield M, Brayne C, Medical Research Council Cognitive Function and Ageing Study. An investigation of whether factors associated with short-term attrition change or persist over ten years: data from the Medical Research Council Cognitive Function and Ageing Study (MRC CFAS). BMC Public Health 2006 Jul 18;6(1):185. [doi: 10.1186/1471-2458-6-185] [Medline: 16848886]
20. Gustavson K, von Soest T, Karevold E, Røysamb E. Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study. BMC Public Health 2012 Oct 29;12:918. [doi: 10.1186/1471-2458-12-918] [Medline: 23107281]
21. Lamers F, Hoogendoorn AW, Smit JH, van Dyck R, Zitman FG, Nolen WA, et al. Sociodemographic and psychiatric determinants of attrition in the Netherlands Study of Depression and Anxiety (NESDA). Compr Psychiatry 2012 Jan;53(1):63-70 [FREE Full text] [doi: 10.1016/j.comppsych.2011.01.011] [Medline: 21397218]
22. Chatfield MD, Brayne CE, Matthews FE. A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. J Clin Epidemiol 2005 Jan;58(1):13-19. [doi: 10.1016/j.jclinepi.2004.05.006] [Medline: 15649666]
23. Gibbs L, Molyneaux R, Whiteley S, Block K, Harms L, Bryant RA, et al. Distress and satisfaction with research participation: impact on retention in longitudinal disaster research. Int J Disaster Risk Reduct 2018 Mar;27:68-74. [doi: 10.1016/j.ijdrr.2017.09.038]
24. Zook PM, Jordan C, Adams B, Visness CM, Walter M, Pollenz K, et al. Retention strategies and predictors of attrition in an urban pediatric asthma study. Clinical Trials 2010 Jun 22;7(4):400-410. [doi: 10.1177/1740774510373798]
25. Cacioppo JT, Cacioppo S. The population-based longitudinal Chicago Health, Aging, and Social Relations Study (CHASRS): study description and predictors of attrition in older adults. Arch Sci Psychol 2018 Feb 26;6(1):21-31 [FREE Full text] [doi: 10.1037/arc0000036] [Medline: 30221101]
26. Jacobsen E, Ran X, Liu A, Chang CC, Ganguli M. Predictors of attrition in a longitudinal population-based study of aging. Int Psychogeriatr 2021 Aug;33(8):767-778. [doi: 10.1017/s1041610220000447]
27. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
28. Cohen S. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The Social Psychology of Health. Newbury Park, CA: SAGE Publications; 1988:31-67.
29. Rubin LH, Meyer VJ, J Conant R, Sundermann EE, Wu M, Weber KM, et al. Prefrontal cortical volume loss is associated with stress-related deficits in verbal learning and memory in HIV-infected women. Neurobiol Dis 2016 Aug;92(Pt B):166-174 [FREE Full text] [doi: 10.1016/j.nbd.2015.09.010] [Medline: 26408051]

XSL•FO
RenderX

30. Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, Wright D, et al. Development of a fatigue scale. J Psychosom Res 1993 Feb;37(2):147-153. [doi: 10.1016/0022-3999(93)90081-p] [Medline: 8463991]
31. Jenkins CD, Stanton BA, Niemcryk SJ, Rose RM. A scale for the estimation of sleep problems in clinical research. J Clin Epidemiol 1988 Jan;41(4):313-321. [doi: 10.1016/0895-4356(88)90138-2] [Medline: 3351539]
32. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions version 6.5. Cochrane Collaboration. 2024. URL: https://www.cochrane.org/authors/handbooks-and-manuals/handbook [accessed 2025-08-20]
33. Yang SH. The effects of GRIT and self-leadership on nursing performance. Soc Converg Knowl Transact 2023;11(2):93-108 [FREE Full text] [doi: 10.22716/sckt.2023.11.2.018]
34. Salminen H, Miettinen M. The role of perceived development opportunities on affective organizational commitment of older and younger nurses. Int Stud Manage Organ 2019 Feb 20;49(1):63-78. [doi: 10.1080/00208825.2019.1565094]
35. Walters G, Hoffart N, Kring D, Whitley T, Horne L, Almotairy M. Work readiness of newly licensed RNs. J Nurs Adm 2022 Sep 01;52(9):469-473. [doi: 10.1097/NNA.0000000000001184] [Medline: 35973193]
36. Lee Y, Kim JL, Kim SH, Chae J. Effect of an age-stratified working environment and hospital characteristics on nurse turnover. HIRA Res 2022 May 31;2(1):106-119. [doi: 10.52937/hira.22.2.1.106]
37. Seo HJ, Kim GY, Chang SJ. Effects of hospital characteristics on employment rate, working period and retirement of ward nurses in Korea: a retrospective cohort study based on HIRAS data. J Korean Acad Nurs 2020 Dec;50(6):837-847. [doi: 10.4040/jkan.20190] [Medline: 33441530]
38. Min A, Baek W, Choi S. What are the key factors influencing newly graduated nurses' preference for choosing their workplace? a best-worst scaling approach. J Nurs Scholarsh 2023 Nov 08;55(6):1204-1214. [doi: 10.1111/jnu.12920] [Medline: 37291734]
39. National Council of State Boards of Nursing homepage. National Council of State Boards of Nursing. URL: https://www.ncsbn.org/ [accessed 2024-07-04]
40. Kolaja CA, Belding JN, Boparai SK, Castañeda SF, Geronimo-Hara TR, Powell TM, et al. Survey response over 15 years of follow-up in the Millennium Cohort Study. BMC Med Res Methodol 2023 Sep 09;23(1):205 [FREE Full text] [doi: 10.1186/s12874-023-02018-z] [Medline: 37689640]
41. Burke SL, Hu T, Naseh M, Fava NM, O'Driscoll J, Alvarez D, et al. Factors influencing attrition in 35 Alzheimer's Disease Centers across the USA: a longitudinal examination of the National Alzheimer's Coordinating Center's Uniform Data Set. Aging Clin Exp Res 2019 Sep 10;31(9):1283-1297 [FREE Full text] [doi: 10.1007/s40520-018-1087-6] [Medline: 30535620]
42. Kuh D, Wong A, Shah I, Moore A, Popham M, Curran P, et al. The MRC National Survey of Health and Development reaches age 70: maintaining participation at older ages in a birth cohort study. Eur J Epidemiol 2016 Nov 19;31(11):1135-1147 [FREE Full text] [doi: 10.1007/s10654-016-0217-8] [Medline: 27995394]
43. Park JH, Hwang HY. Predictors of turnover intention among nurses in small and medium-sized hospitals. J Korean Acad Nurs Adm 2017;23(5):471-482. [doi: 10.11111/jkana.2017.23.5.471]
44. Cho SH, Lee JY, Hong KJ, Huh I. The relationship between average length of stay and nurse staffing in general hospitals from 1996 to 2016. J Korean Acad Nurs Adm 2020;26(5):521-532. [doi: 10.11111/jkana.2020.26.5.521]
45. Van Loon AJ, Tijhuis M, Picavet HS, Surtees PG, Ormel J. Survey non-response in the Netherlands: effects on prevalence estimates and associations. Ann Epidemiol 2003 Feb;13(2):105-110. [doi: 10.1016/s1047-2797(02)00257-0] [Medline: 12559669]
46. Couper MP. Web Surveys: a review of issues and approaches. Public Opin Q 2000;64(4):464-494. [doi: 10.1086/318641]
47. Fan W, Yan Z. Factors affecting response rates of the web survey: a systematic review. Comput Hum Behav 2010 Mar;26(2):132-139. [doi: 10.1016/j.chb.2009.10.015]
48. Nápoles AM, Chadiha LA, Resource Centers for Minority Aging Research. Advancing the science of recruitment and retention of ethnically diverse populations. Gerontologist 2011 Jun 11;51 Suppl 1(Suppl 1):S142-S146 [FREE Full text] [doi: 10.1093/geront/gnr019] [Medline: 21565815]
49. Sutton E, Cowan A, Haverty C, Jones B, Moe A, Russell D. Main street, USA enrollment: community hospital recruitment strategies for diverse and equitable selection of pregnant participants in a prospective cohort study ID. Obstet Gynecol 2023;141(5S):39S-40S. [doi: 10.1097/01.AOG.0000930232.53045.07]

Abbreviations

OR: odds ratio
PSS: Perceived Stress Scale

XSL•FO
RenderX

Edited by T Sanchez, A Mavragani; submitted 26.10.24; peer-reviewed by C Kolaja, L Zhu; comments to author 18.11.24; revised version received 07.03.25; accepted 17.08.25; published 20.10.25.

Please cite as:
Kim YT, Cha C, Baek G, Kim B, Song BM, Lim JY, Park HY, Shin JH
Factors Affecting Nonresponse Among Female Participants in the Korea Nurses' Health Study: Longitudinal Cohort Survey Study
JMIR Public Health Surveill 2025;11:e68038
URL: https://publichealth.jmir.org/2025/1/e68038
doi:10.2196/68038
PMID:

©Young Taek Kim, Chiyoung Cha, Gumhee Baek, Bohye Kim, Bo Mi Song, Joong-Yeon Lim, Hyun-Young Park, Juh Hyun Shin. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Health Impact of Suspected Interpersonal Violence Against Adults: Retrospective Cohort Study

Rui Barbosa[1], MD; Rita Lopes[2], MD; Tiago Taveira-Gomes[2,3,4,5,6], MD, PhD; Carla Ponte[7], MD; Teresa Magalhães[1,2,3,8], MD, PhD

[1]Department of Public Health and Forensic Sciences, and Medical Education, Faculty of Medicine, University of Porto, Porto, Portugal

[2]MTG Research and Development Lab, Porto, Portugal

[3]Center for Health Technology and Services Research (CINTESIS@RISE), Porto, Portugal

[4]Department of Community Medicine, Information and Decision in Health, Faculty of Medicine, University of Porto, Porto, Portugal

[5]Faculty of Health Sciences, University Fernando Pessoa (FCS-UFP), Porto, Portugal

[6]SIGIL Scientific Enterprises, Dubai, United Arab Emirates

[7]USF Porta do Sol - ULS Matosinhos, Senhora da Hora, Portugal

[8]TOXRUN–Toxicology Research Unit, University Institute of Health Sciences, Advanced Polytechnic and University Cooperative (CESPU), Gandra, Portugal

Corresponding Author:
Rita Lopes, MD
MTG Research and Development Lab
Rua Professor Joaquim Bastos 102 2A
Porto, 4200-604
Portugal
Phone: 351 911729305
Email: rita.lopes@mtg.pt

Abstract

Background: Interpersonal violence (IV) has an extensive and profound impact on health, representing a public health concern. Different health outcomes have been identified based on the characteristics of the survivor and the abuser, their relationship, the type of violence perpetrated, and the cumulative effect of multiple violent experiences.

Objective: The main objective of this study was to estimate the likelihood of negative health outcomes, such as substance abuse, mental health, and somatic disorders, occurring in adults presenting a clinical suspicion of IV.

Methods: We performed a retrospective, observational cohort study, using secondary data from electronic health records of adult patients of the Local Health Unit of Matosinhos (ULSM). The control cohort included all patients aged between 18 and 59 years followed at ULSM between January 1, 2008, and May 9, 2024, while the violence cohort included patients within the same age range who were suspected survivors of IV within the same time period. Exposure was defined by the presence of one or more of the IV text expressions or codes in the patient's electronic health record. Data regarding violence suspicion, comorbidities, and health outcomes were obtained by conducting a text search on clinical notes, as well as *ICD-9* (*International Classification of Diseases, Ninth Revision*) and *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) and International Classification of Primary Care 2 codes. The follow-up period was 10 years. To estimate the risk of developing health outcomes, we constructed a cohort model using a Cox proportional hazards model adjusted at baseline for age and sex.

Results: The control cohort included 154,145 patients, and the violence cohort included 36,835 patients. Suspected survivors of IV had a higher hazard ratio of developing alcohol abuse (3.05, 95% CI 2.87-3.24), drug abuse (6.03, 95% CI 4.5-8.06), suicidal ideation (5.50, 95% CI 4.78-6.32), major psychiatric disorder (4.46, 95% CI 4.38-4.53), chronic pain (3.70, 95% CI 3.60-3.81), sleep disorders (3.57, 95% CI 3.46-3.68), use of antidepressants (3.57, 95% CI 3.51-3.64), use of anxiolytics (2.98, 95% CI 2.93-3.03), and eating disorder (2.72, 95% CI 2.06-3.59). Regarding somatic health conditions, suspected violence exposure was linked to a higher hazard ratio for metabolic dysfunction–associated steatotic liver disease (6.91, 95% CI 6.37-7.50), chronic immune inflammatory disorder (3.68, 95% CI 3.93-3.44), asthma (2.64, 95% CI 2.53-2.74), chronic kidney disease (2.49, 95% CI 2.39-2.59), hypercholesterolemia (2.16, 95% CI 2.12-2.20), and early heart disease (2.01, 95% CI 1.93-2.10).

XSL•FO
RenderX

Conclusions: Exposure to violence was linked to a higher likelihood of developing adverse events related to substance abuse, mental, and somatic health. Our findings lead to a deeper understanding of the complex burden of violence on health, uncovering new relationships between IV and health outcomes while validating those already explored.

(JMIR Public Health Surveill 2025;11:e65191) doi:10.2196/65191

KEYWORDS

interpersonal violence; health outcomes; health risk behaviors; mental health; somatic health conditions; health care diagnosis

Introduction

The World Health Organization (WHO) defines violence as "the intentional use of physical force or power, threatened or current, against oneself, another person, or against a group or community, that either results in or has a high likelihood of resulting in injury, death, psychological harm, maldevelopment, or deprivation." Thus, it includes active and passive behaviors, mostly of a repetitive nature. It is present across every stage of life, regardless of sex, and in different sociocultural contexts and forms [1].

In Portugal, data from the National Institute of Statistics from 2023 shows that 20.1% of people aged between 18 and 74 years have been survivors of physical or sexual violence in adulthood. Women are primarily affected in the context of intimacy, while men are more often survivors of violence from nonpartners [2].

Interpersonal violence (IV) is no longer regarded only as a criminal justice matter. Its impact on health is extensive and profound, and it is recognized as a significant public health problem, which requires the attention of health care providers [1,3]. Overall, IV has short-, medium-, and long-term effects. In fact, they are not limited to immediate physical trauma but include physiological changes, psychological consequences, and coping strategies that may increase the risk of negative health outcomes, years or decades following the primary exposure [4-7]. There are several mechanisms behind these transformations, some happening in the first years of life [5,8-11]: (1) changes in the brain's white and grey matter, its volume and functional connective; (2) altered neurotransmitter metabolism; (3) changes in the neuroendocrine stress response; (4) chronic inflammation; (5) impaired glucose metabolism; (6) changes in the microbiome in response to stress; and (7) epigenetic mechanisms through DNA methylation. Many of these alterations seem related to the hypothalamic-pituitary-adrenal (HPA) axis, which integrates and coordinates the central and peripheral aspects of the mammalian stress response. Thus, violence exposure acts as a trigger and places survivors at high risk for the development of physical and mental health illnesses, as well as latent and genetically programmed diseases [5,7].

Several studies describe different health outcomes according to the survivor's and abuser's personal characteristics and their relationship, as well as the type of violence perpetrated. Intimate partner violence is shown to be related, for example, with mental health disorders (where genetics have a significant preponderance [11]), substance consumption, and somatic diseases, as obesity, metabolic syndrome, type 2 diabetes (T2D; due to increased cortisol levels and suppressed insulin levels [10]), hypertension (as a result of chronic activation of HPA axis, and autonomic nervous system [9]) among others [6,7,9-13]. Bullying and cyberbullying seem to be associated with depression, anxiety, and somatization symptoms [6]. Sexual abuse is found to be mainly related to depression, anxiety, suicide attempts, substance abuse, sleep disorders, gastrointestinal disorders, or nonepileptiform seizures [5,6,14]. Economic abuse is considered associated with increased depression, anxiety, suicidal ideation, and posttraumatic stress disorder [15].

Furthermore, it must be outlined that survivors of IV frequently experience multiple forms of it [6,16]. Experiencing or being exposed to one type of violence increases the risk of becoming a survivor of other types of violence or inflicting harm on others [17,18]. The association of different types of IV or its reiteration may result in different and potentially more severe negative impacts on health consequences [18,19]. This way, to understand the full burden of IV health outcomes, we should look to the cumulative effect of the multiple experiences of violence during the lifespan [18].

Ultimately, there are still associations to be unveiled and better understood that may revolutionize the approach to survivors. The better we understand the possible long-term outcomes of IV, the more capable doctors are to minimize the burden of violence on survivors' development and well-being.

The main objective of this study was to estimate the likelihood of negative health outcomes, such as substance abuse, mental health, and somatic disorders, occurring in adults presenting a clinical suspicion of IV. The study also aimed to establish the incidence of individuals with suspected exposure to violence followed in the Local Health Unit of Matosinhos (Unidade Local de Saúde de Matosinhos [ULSM]) and to characterize this population both demographically and according to prior clinical history.

Methods

Study Design and Setting

A real-world, retrospective, observational, and analytical cohort study was performed using data from electronic health records (EHR) of adult users of ULSM.

ULSM is a health unit that provides primary, secondary, and tertiary health care, being composed of 1 hospital and 14 primary care centers, serving a population of around 176.774 inhabitants.

The study's inclusion criteria for the violence cohort were: (1) patients aged between 18 and 59 years; (2) sought the ULSM service between January 1, 2008, and May 9, 2024; (3) having text expressions (according to predefined keywords, described

XSL•FO
RenderX

in Table 1) related to IV documented in the clinical notes or codes or at least one *ICD-9* (*International Classification of Diseases, Ninth Revision*) or *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) or an International Classification of Primary Care 2 (ICPC-2) IV-related code occurring after the age of 18 years (Table 1); and (4) having at least one appointment at ULSM in the 365 days before the index date. This last criterion was established to ensure the capturing of a population regularly followed at ULSM and to exclude sporadic visitors, such as tourists, thus increasing accuracy in the characterization of the local population.

The control cohort consisted of an incident population from ULSM, between 18 and 59 years old, who sought the ULSM service between January 1, 2008, and May 9, 2024, and had at least one appointment at ULSM in the previous 365 days before the index date.

No exclusion criteria were used in this investigation other than not meeting the inclusion criteria. Formal sample size calculations were not conducted, as the sample consisted of all the patients eligible according to the inclusion criteria.

The two study cohorts were not mutually exclusive; therefore, the same patient could enter the two cohorts at different times during the study period.

The index data corresponded to the date when each patient met all the cohort criteria. The patients' medical history, also described as prior outcomes, was evaluated using all records available in the preindex period. The follow-up period was 10 years (3650 days).

XSL•FO
RenderX

Table 1. Codes and text expressions used to define interpersonal violence suspicion in the study[a].

Classification	Code
ICD-9[b]	• 995.80 Adult maltreatment, unspecified • 995.81 Adult physical abuse • 995.82 Adult emotional or psychological abuse • 995.83 Adult sexual abuse • 995.84 Adult neglect (nutritional) • 995.85 Other adult abuse and neglect

ICD-10[c]	• T74 Adult and child abuse, neglect, and other maltreatment, confirmed
	  • T74.01 Adult neglect or abandonment, confirmed
	  • T74.11 Adult physical abuse, confirmed
	  • T74.21 Adult sexual abuse, confirmed
	  • T74.31 Adult psychological abuse, confirmed
	  • T74.51 Adult forced sexual exploitation, confirmed
	  • T74.61 Adult forced labor exploitation, confirmed
	  • T74.91 Unspecified adult maltreatment, confirmed
	  • T74.A1 Adult financial abuse, confirmed
	• T76 Adult and child abuse, neglect, and other maltreatment, suspected
	  • T76.01 Adult neglect or abandonment, suspected
	  • T76.11 Adult physical abuse, suspected
	  • T76.21 Adult sexual abuse, suspected
	  • T76.31 Adult psychological abuse, suspected
	  • T76.51 Adult forced sexual exploitation, suspected
	  • T76.61 Adult forced labor exploitation, suspected
	  • T76.91 Unspecified adult maltreatment, suspected
	  • T76.A1 Adult financial abuse, suspected

ICPC-2[d]	• Z12 Relational problem with partner • Z13 Partner behavioral problem • Z25 Violent act or event
Text expressions	• [Aa]bandon • [Aa]gred • [Aa]gress • [Aa]mea[çc]a • [Aa]pert[ãa]o • [Aa]pert[õo]es • [Bb]alead • [Bb]at[ei] • [Bb]ofet • [Cc]hapad • [Dd]estru • [Dd]ispar • [Ee]mpurr • [Es]sfaquead • [Ee]sgana • [Ee]stalad • [Ee]strang • [Ff]acad • [Hh]umilh • [Ii]nsult • [Mm]urr • [Oo]brigada a • [Pp]ancad • [Pp]aulad • [Pp]r[aá]tica[s]* sex • [Pp]edrad • [Pp]erseg • [Pp]ontap • [Ss]apatad • [Ss]oco • [Ss]ufoc

[a]Patients aged between 18 and 59 years followed in the Local Health Unit of Matosinhos should have at least one of the codes or text expressions to qualify for the violence cohort.

XSL•FO
RenderX

[b]ICD-9: International Classification of Diseases, Ninth Revision.

[c]ICD-10: International Statistical Classification of Diseases, Tenth Revision.

[d]ICPC-2: International Classification of Primary Care, 2nd edition.

Exposure and Outcome Definitions

Exposure was defined by the presence of one or more of the IV text-expressions or *ICD-9*, *ICD-10*, or ICPC-2 codes in the patient's EHRs described in Table 1.

Study Variables

The variable categories included in the study were (1) sex, (2) age, (3) health risk behaviors, (4) mental health disorders, and (5) somatic health conditions.

The health risk behaviors category included the following variables: (1) alcohol abuse and (2) drug abuse.

Mental health disorders encompassed (1) major psychiatric disorders, (2) posttraumatic stress disorder, (3) suicidal ideation, (4) chronic pain, (5) headache, (6) sleep disorders, (7) memory disorders, (8) eating disorders, and (9) psychotropic drug use (sedatives, antidepressants, anxiolytics, and antipsychotics).

Within the somatic health conditions category, the included variables were (1) T2D, (2) hypercholesterolemia, (3) metabolic dysfunction–associated steatotic liver disease (MASLD), (4) hypertension, (5) chronic kidney disease, (6) early heart disease, (7) asthma, (8) chronic inflammatory disorder, (9) cancer, and (10) unspecified illness.

All variables described above were used for cohort characterization, using all information available in the preindex period, within study dates.

The events used to estimate hazard ratios (HRs) included all of the health outcomes included in the categories "health risk behaviors," "mental health disorders," or "somatic health conditions" occurring up to 3650 days after the index date.

Statistical Analysis

Analysis was conducted through local execution of analytical programs developed using VERO technology that were compiled for ULSM target data infrastructure built upon Apache Spark Framework version 3.2.1 and source data harmonized according to the OMOP CDM 5.4 standard. The package implements a complete data engineering pipeline to transform the source data into the final dataset, as well as the execution of statistical analysis and the generation of results reports. Continuous variables were described using the average and SD, while categorical variables were expressed by absolute and relative frequencies. Age was analyzed as a continuous and a categorical variable. Patients were stratified into 10-year age categories in order to better characterize the age distribution of the study's population. To estimate the risk of developing health outcomes, we constructed a cohort model using a Cox proportional hazards model adjusted at baseline for age and sex. HRs were calculated with 95% CIs. Time at risk was calculated from the index date until the development of the outcome of interest. Patients were censored in case of leaving the database. Missing data were not imputed for any variables. Analysis was performed using only the available data.

Ethical Considerations

The Health Ethics Committee and Data Protection Officer of ULSM authorized the use of the database for this study, under approval codes 14/CES/JAS of 02-11-2022 and 04/CLPSI/2022 of 03-02-2022. Data processing and analysis were performed using analytical programs developed for this purpose that were sent for execution at ULSM servers. Researchers did not have access to patient-level data. Only aggregated data were shared with the researchers. No data extraction took place in the context of this study. Informed consent was waived by the institution, invoking public health and research interests, following the General Data Protection Regulation. No type of compensation was given to participants.

Results

Baseline Characteristics of Incident Population

The baseline characteristics of the control and suspected IV cohorts are described in Table 2.

The population on both cohorts was mostly female (83,636/154,145, 54.3% in the control and 23,521/36,835, 63.9% in the suspected IV cohort), average age was 32.7 (SD 23.6) in the control and 42.6 (SD 19.6) in the IV cohort.

Regarding patients' clinical history, there was a higher prevalence of both drug and alcohol abuse in the suspected IV cohort. The same was true for the major psychiatric disorders, chronic pain, sleep disorders, and the use of psychotropic drugs.

Regarding somatic health conditions, there was a higher prevalence of all studied conditions in the suspected IV cohort when compared to the control, with the biggest differences concerning cancer (1723/154,145, 1.1% in the control; 5704/36,835, 15.5% in the suspected IV cohort), hypertension (7131/154,145, 4.6% vs 13,343/36,835, 36.2%), and hypercholesterolemia (8401/154,145, 5.5% vs 13,819/36,835, 37.5%).

XSL•FO
RenderX

Table 2. Baseline characteristics of control and suspected violence cohorts[a].

	Control group (n=154,145)	Suspected IV[b] group (n=36,835)
Sex, n (%)		
Female	83,636 (54.3)	23,521 (63.9)
Male	70,509 (45.7)	13,314 (36.1)
Age (years)		
Mean (SD)	32.7 (23.6)	42.6 (19.6)
Range, n (%)		
10-19	31,275 (20.3)	1746 (4.7)
20-29	36,792 (23.9)	5906 (16)
30-39	31,892 (20.7)	8334 (22.6)
40-49	29,015 (18.8)	9739 (26.4)
50-59	25,171 (16.3)	11,110 (30.2)
Cohort components, n (%)		
Violence codes	108 (0.1)	1909 (5.2)
Violence registries	4665 (3)	34,985 (95)
Prior outcomes		
Health risk behaviors , n (%)		
Alcohol abuse	180 (0.1)	1001 (2.7)
Drug abuse	9 (0.1)	43 (0.1)
Mental health disorders , n (%)		
Major psychiatric disorder	4765 (3.1)	16,378 (44.5)
Posttraumatic stress disorder	22 (0.1)	69 (0.2)
Suicidal ideation	41 (0.1)	211 (0.6)
Chronic pain	493 (0.3)	4178 (11.3)
Headache	1640 (1.1)	2793 (7.6)
Sleep disorders	706 (0.5)	3884 (10.5)
Memory disorders	92 (0.1)	368 (1)
Eating disorders	71 (0.1)	79 (0.2)
Psychotropic drug use		
Sedatives	4399 (2.9)	9365 (25.4)
Antidepressants	3320 (2.2)	15,065 (40.9)
Anxiolytics	7987 (5.2)	22,304 (60.6)
Antipsychotics	1170 (0.8)	4154 (11.3)
Somatic health conditions , n (%)		
Type 2 diabetes mellitus	2983 (1.9)	1803 (4.9)
Hypercholesterolemia	8401 (5.5)	13,819 (37.5)
Metabolic dysfunction–associated steatotic liver disease	20 (0.1)	254 (0.7)
Hypertension	7131 (4.6)	13,343 (36.2)
Chronic kidney disease	398 (0.3)	653 (1.8)
Early heart disease	2538 (1.6)	1324 (3.6)
Asthma	2603 (1.7)	2393 (6.5)
Chronic immune inflammatory disorder	213 (0.1)	668 (1.8)

XSL•FO
RenderX

	Control group (n=154,145)	Suspected IV[b] group (n=36,835)
Cancer	1723 (1.1)	5704 (15.5)
Unspecified illness	1472 (1)	2986 (8.1)

[a]All patients were followed at the Local Health Unit of Matosinhos between January 1, 2008, and May 9, 2024, and aged between 18 and 59 years.
[b]IV: interpersonal violence.

Health Outcomes 10 Years After the Episode of Suspected Violence

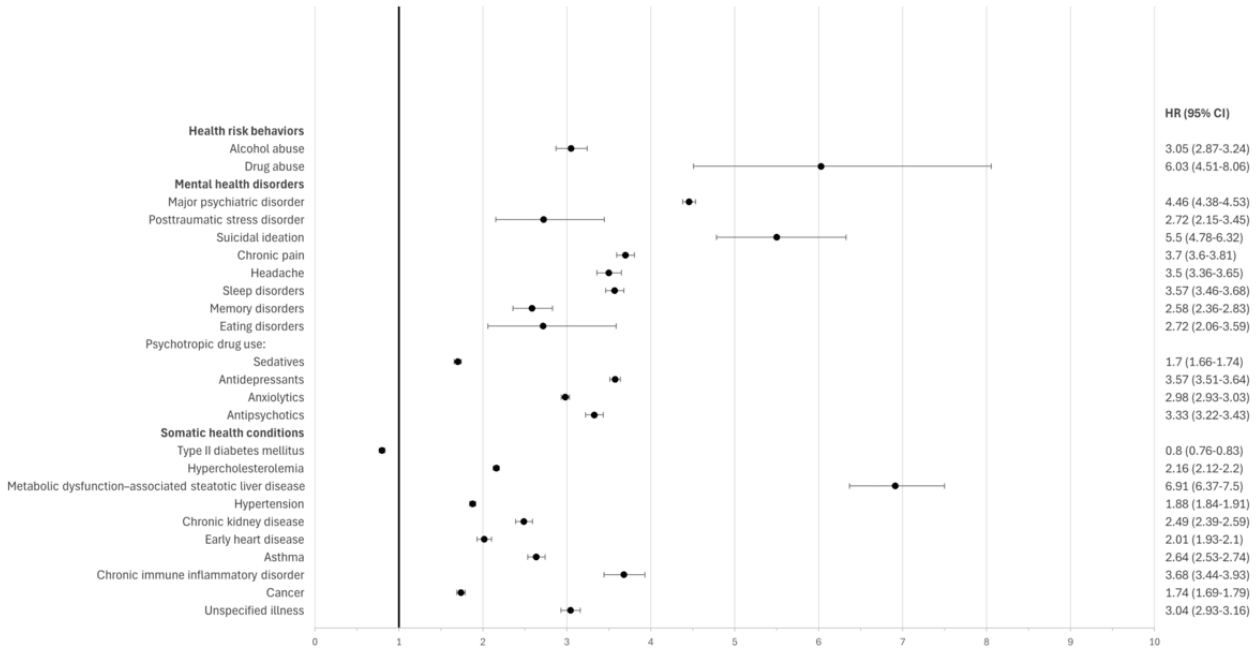
The results obtained from the 3650-day follow-up after the index date are represented in Figure 1 and Table 3.

For every outcome studied, the follow-up period for adult survivors of IV was consistently shorter compared to other adults. The shortest follow-up periods were observed for the use of anxiolytics (mean 1064, SD 1151), major psychiatric disorder (mean 1227, SD 1257), and hypercholesterolemia (mean 1241, SD 1171).

Health outcomes did come up earlier in survivors of IV, compared to adults who were not survivors of IV. The shortest times to event occurred in early heart disease (mean 43, SD 127), T2D (mean 123, SD 294), and hypertension (mean 381, SD 532).

The violence cohort had a higher HR of developing the outcomes studied, except for T2D (0.80, 95% CI 0.76-0.83). The outcomes that stood out, as they presented the highest HR values, were fatty liver disease (6.91, 95% CI 6.37-7.50), drug abuse (6.03, 95% CI 4.5-8.06), suicidal ideation (5.50, 95% CI 4.78-6.32), and major psychiatric disorder (4.46, 95% CI 4.38-4.53).

Figure 1. Health outcomes up to 3650 days postindex in the suspected violence cohort. All patients were followed at the Local Health Unit of Matosinhos between January 1, 2008, and May 9, 2024, and were aged between 18 and 59 years.



XSL•FO
RenderX

Table 3. Health outcomes after a suspected experience of IV in adults up to 10 years of follow-up[a].

	Ev/100PY[b]	Follow-up (days), mean (SD)	Time to event, mean (SD)	HR[c] (95% CI)
Health risk behaviors				
Alcohol abuse	0.53	2244 (1285)	1055 (941)	3.05 (2.87-3.24)
Drug abuse	0.03	2294 (1273)	1412 (1046)	6.03 (4.05-8.06)
Mental health disorders				
Major psychiatric disorder	16.34	1227 (1257)	642 (836)	4.46 (4.38-4.53)
Posttraumatic stress disorder	0.04	2293 (1274)	948 (931)	2.72 (2.15-3.45)
Suicidal ideation	0.11	2285 (1275)	1055 (990)	5.50 (4.78-6.32)
Chronic pain	3.25	2007 (1291)	1295 (1014)	3.70 (3.60-3.81)
Headache	1.25	2177 (1283)	1378 (989)	3.50 (3.36-3.65)
Sleep disorders	2.75	2048 (1292)	1274 (994)	3.57 (3.46-3.68)
Memory disorders	0.26	2273 (1275)	1554 (1062)	2.58 (2.36-2.83)
Eating disorders	0.02	2294 (1273)	981 (945)	2.72 (2.06-3.59)
Psychotropic drug use				
Sedatives	3.72	1924 (1318)	976 (970)	1.70 (1.66-1.74)
Antidepressants	13.36	1331 (1288)	672 (862)	3.57 (3.51-3.64)
Anxiolytics	20.95	1064 (1151)	646 (808)	2.98 (2.93-3.03)
Antipsychotics	2.56	2056 (1337)	900 (1017)	3.33 (3.22-3.43)
Somatic health conditions				
Type 2 diabetes mellitus	0.84	2174 (1328)	123 (294)	0.80 (0.76-0.83)
Hypercholesterolemia	16.00	1241 (1171)	808 (814)	2.16 (2.12-2.20)
Metabolic dysfunction–associated steatotic liver disease	0.45	2256 (1288)	996 (1092)	6.91 (6.37-7.50)
Hypertension	11.11	1342 (1325)	381 (532)	1.88 (1.84-1.91)
Chronic kidney disease	1.42	2174 (1298)	1176 (1121)	2.49 (2.39-2.59)
Early heart disease	0.60	2201 (1320)	43(127)	2.01 (1.93-2.10)
Asthma	1.20	2168 (1311)	844 (933)	2.64 (2.53-2.74)
Chronic immune inflammatory disorder	0.49	2245 (1288)	961 (1015)	3.68 (3.44-3.93)
Cancer	2.79	2026 (1324)	915 (997)	1.74 (1.69-1.79)
Unspecified illness	1.59	2133 (1291)	1242 (1003)	3.04 (2.93-3.16)

[a]All patients were followed at the Local Health Unit of Matosinhos between January 1, 2008, and May 9, 2024, and aged between 18 and 59 years.

[b]Ev/100PY: event rate per person-year.

[c]HR: hazard ratio.

Discussion

Principal Findings

The analysis of EHR data from adult users of ULSM revealed that patients who were suspected survivors of IV were more likely to develop negative outcomes when compared to the control cohort. This was true for the different health outcomes regarding health risk behaviors, mental health, and somatic conditions studied, except for T2D. In addition to reinforcing existing literature on the subject, we advanced a potential link between the experience of violence and the onset of MASLD.

We found that adult suspected survivors of IV were more likely to develop drug and alcohol abuse, respectively. This goes according to Rivara et al [6], Stubbs and Szoeke [12], Kyle [20], and Colaprico et al [21], which describe a relation between different types of IV and alcohol and substance use. Health risk behaviors, as substance abuse and alcohol abuse, may work as coping mechanisms to violence exposure [6,20,21], but it is important to note that this relationship is bidirectional. Substance abuse and worsened mental health may result in a higher risk of being a survivor of violence or inflicting harm on others [20].

According to WHO, “a mental disorder is characterized by a clinically significant disturbance in an individual’s cognition,

XSL•FO
RenderX

emotional regulation, or behavior" [22]. In this study, adult suspected survivors of IV were more likely to develop each of the mental disorders studied. This goes according to other studies, such as Rivara et al [6], which summarizes the relation between intimate partner violence and depression, posttraumatic stress disorder, suicidal ideation, and chronic mental illness. Kyle [20] and Stubbs and Szoeke [12] describe similar findings, adding anxiety disorders and eating disorders, respectively. In this study, suspected survivors of IV were more likely to develop memory disorders, which is consistent with findings described by Campbell et al [23].

Sexual abuse is also related to depression, anxiety, sleep disorders, suicide, eating disorders, and posttraumatic stress disorders, according to Rivara et al [6]. Bullying and cyberbullying in the youth and children population seem to be linked to depression, anxiety, self-esteem, posttraumatic stress disorder, sleep disturbances, and headache [6]. Colaprico et al [21] described a relation between workplace bullying in adults and disorders as anxiety, depression, suicidal tendencies, and headache.

Another finding of this study was the higher likelihood of the use of psychotropic drugs. This finding seems to be directly related to the type of mental health disorders developed and their comorbidities, as these drugs may be part of a treatment strategy to respond to them.

Literature describes correlations between IV in adults and various health issues, including cardiovascular diseases, hypertension, elevated cortisol levels, cancer, sexually transmitted infections, a higher risk of developing cancer, gastrointestinal disorders, such as irritable bowel syndrome, and asthma [5-7,12].

Lin et al [24] found that children exposed to adverse childhood experiences had higher risks of dyslipidemia, asthma, liver disease, kidney disease, and arthritis, similar to our findings in the adult population.

We found a potential link between adult IV and MASLD. This possible correlation with violence has not yet been described in the literature for the adult population. However, a recent cross-sectional study found a moderate risk in men for specific forms of child maltreatment [25]. Recent studies have focused on the relation between NALFD and mental health disorders. Shea et al [26] described in their systematic review a high prevalence of depression, anxiety, and stress among adults with MASLD. Kang et al [27] concluded that higher perceived stress was independently associated with an increased prevalence of MASLD. Soto-Angona et al [28] have also come to the conclusion that MASLD is part of a complex system of mental and noncommunicable somatic disorders, with a shared underlying pathogenesis. This is based on common lifestyle and environmental risk factors, mediated by dysregulation of inflammation, oxidative stress pathways, and mitochondrial function.

This study found an inverse correlation between the experience of IV and T2D. Some studies have found a link between intimate partner violence and T2D [7,10,12]. Studies performed on adverse childhood experiences also correlate these experiences to the higher prevalence of T2D in childhood [29,30] This way, the lower HR of developing T2D in the IV cohort may be explained by a higher prevalence of the disease in children exposed to violence—a higher baseline prevalence could lead to lower incidence in adulthood as a bigger percentage of the cohort could have developed the disease prior to index date [30]. In this study, 4.9% of the IV cohort had T2D at the index date compared to only 1.9% of the control cohort.

Survivors of violence may be more susceptible to losing follow-up due to instability in their lives, which can result in missed appointments, relocation, or increased mortality, for example.

As previously said, the association between IV and negative health outcomes may be partially explained by the impact of chronic stress on various systems that regulate body homeostasis, particularly the HPA axis [5,8,16,31,32]. Hyperactivation of the HPA axis leads to neuroendocrine and immune disruption and altered systemic cortisol production [5,8,31]. Survivors of violence are at a higher risk of resurvivorization across their lifespan, exposing them to chronic and cumulative stress. This snowball effect may result in the negative outcomes explored in this study [16,32]. In fact, violence may even start before birth (intrauterine violence), in cases of violence against pregnant women, perpetuating through every phase of life, having a cumulative effect with tremendous consequences [16].

Even though it seems clear that violence is deleterious, we are still not able to comprehend the full extent of its consequences. Studies tend to address IV within each subtype of violence or a specific subgroup of the population, overlooking the intricate relation between different forms of violence and recurrent experiences of it. This limits our capacity to draw conclusions and compare patterns across diverse populations. A lot of studies have been done on children, trying to understand the direct impact of violence on health during childhood, but also its long-term consequences through their adult life [4,8,29-31,33-35]. There is also a great amount of literature regarding intimate partner violence, particularly among women [7,12,15,20,23,36-38]. However, there is limited research on other subtypes of IV. The body of research regarding men's health and violence is less robust than that for women [18], creating barriers in understanding the true burden in this sex and whether there are significant differences in response to violence between sexes.

Since multiple experiences of violence across a person's lifespan are common, we will only understand the full impact of IV by considering the co-occurrence and interconnections of the multiple experiences of it. Violence must be conceptualized from a dose-response perspective, considering the accumulation of exposure over time and across multiple domains [19]. Scott-Storey et al [32] propose an approach aligned with this. They developed a cumulative lifetime violence severity scale reflecting dimensions of type, focus (target or perpetrator), timing (childhood or adulthood), context, frequency, and degree of distress. This tool was designed only for men. Similar tools could be developed for both sexes, adjusting to each population,

XSL•FO
RenderX

allowing a profound understanding of violence's negative impact on health.

In this study, we chose to have a global approach to IV in adults, as we could not develop a dose-response perspective or differentiate each type of violence in each patient, due to the poor codification on this matter and the broad vocabulary used to describe it in EHR.

Our results underscore the significant burden of IV, which affects multiple facets of survivors' lives. Health consequences have their own repercussions that substantially diminish survivors' quality of life, affecting other spheres of life, such as job loss and career abandonment, or increased health care costs [1,15]. Violence is therefore conceptualized as a structural determinant of health. Its presence negatively impacts patients' health outcomes and self-development. Moreover, distinct patterns of cumulative lifetime exposure to violence result in varying degrees of symptom severity [13].

Violence screening represents a cost-effective prevention strategy that should be standard practice [39]. Several barriers prevent an effective detection of the cases: (1) survivors may struggle with disclosure due to shame, fear of being judged, of not being believed [40], or of violence escalation; (2) health care practitioners may face challenges such as insufficient time and skills to identify cases, discomfort, confidentiality concerns, or their own experiences with IV [39,40]; and (3) clinical-level factors may include inadequate information, lack of screening protocols, and insufficient support [39,40]. These constraints may lead to a misdiagnosis of cases. Furthermore, due to medical privacy and doubts associated with the case complexity, some situations might only be recorded in confidential registers, which may contribute to an apparent underidentification of IV cases, making it difficult to understand the full extent of the problem.

Since almost all clinicians across their practice will encounter patients who were survivors of IV, they should be trained with the necessary skills to identify and report these situations and respond appropriately, reducing their harmful consequences and preventing reoccurrence [37,39,40]. Public institutions must have a direct role in the prevention of violence, addressing underlying causes, and creating appropriate measures and programs that can give the needed support to survivors and their families.

A deeper understanding of the long-term consequences of IV enables an improvement in the approach to survivors in medical practice. By identifying the negative health outcomes that may arise from exposure to violence, clinicians can develop a proactive and targeted follow-up of survivors. An early detection of these pathologies and an understanding of the complex and multifaceted needs of survivors contribute to a more personalized, preventive, and effective medical care.

Limitations

Omission bias is possible due to certain aspects of this study. Physicians might not document or code all suspected IV cases they detect in the EHR. Additionally, they may choose to record these cases only in confidential health registers.

Given that the identification of suspected cases of IV was done through data from the EHR of adult users, we implemented comprehensive selection criteria and used a dual approach that integrates all relevant ICPC-2, *ICD-9*, and *ICD-10* codes, as well as all relevant clinical notes at every intervention point for any health care professional.

While we tried to include pertinent and broad terminology used by physicians to describe violence scenarios in clinical notes, it is not possible to account for the all variety of possible expressions that may have been used to describe a context of violence; as such, there may be a potential subestimation of violence-exposed patients in the respective cohort.

Furthermore, the variability in coding and terminology used to describe violence between countries and institutions may compromise the comparability and reproducibility of our findings when applied to different clinical settings.

Another potential limitation of this study is the reporting bias that may be associated with patients who are suspected of IV, as health care professionals may have a stronger tendency to report health care outcomes, especially those connected with mental health, regarding these patients.

In addition, given the fact that there is no methodical registration of variables as socioeconomic status, race or ethnicity, and educational level, it was not possible to adjust the data for these possible confounders; therefore, confounding bias may be present.

By including all the patients who met the inclusion criteria, without any additional exclusion criteria, we minimized selection bias and enhanced the detection of patients exposed to violence. Our comprehensive approach, encompassing all the eligible population at ULSM during the study period, resulted in a large and diverse sample, which strengthens the validity and generalizability of our findings.

A validation study of these data should be carried out. We intend to extend the study to other regions of Portugal and collaborate with research groups in other countries, which would allow us to assemble larger cohorts, conduct robust statistical analyses, and gain a deeper understanding of violence to improve prevention, intervention, and treatment. The use of real-world data is recommended as it provides the possibility of analyzing larger patient populations over longer timelines [41]. We also intend to compare the effects of IV between sexes and to develop study designs that incorporate the cumulative impact of violence across multiple domains and stages of life, while characterizing the studied populations with further socioeconomic demographic data such as profession, employment situation, level of education, and family structure.

Conclusions

This research has demonstrated that adults suspected to be survivors of IV are more likely to develop negative health outcomes than the general population. This was true for health risk behaviors, as well as for mental health and somatic conditions. Notably, this study introduces a novel perspective on the matter, suggesting that violence is linked with the development of MASLD.

XSL•FO
RenderX

Furthermore, adults suspected to be survivors of IV developed health outcomes earlier and had a shorter period of follow-up, compared to the adults with no exposure to violence. These findings highlight the complex impact of violence on health, underscoring the urgent need to address this endemic problem in society. Health providers play a crucial role in the early detection of cases, follow-up, and treatment of survivors of IV. Future multicentric research (including international research groups) should focus on validating these findings, trying to have a dose-response perspective. Such efforts will enable more robust statistical analyses and a deeper understanding of violence's consequences, ultimately contributing to improved prevention, detection, and treatment strategies.

Acknowledgments

The authors would like to acknowledge the Local Healthcare Unit of Matosinhos for granting the permissions that enabled this study.

Data Availability

Data sharing is not applicable to this article as no datasets were generated during this study. Patient-level data used in this study are not publicly available, as the Zero Trust model was used. Researchers accessed only aggregated and anonymized results, ensuring full compliance with data protection, ethical, and privacy standards. This article incorporates all the available statistical results.

Conflicts of Interest

None declared.

References

1. Krug EG. World Report on Violence and Health. Geneva, Switzerland: World Health Organization; 2002.
2. Inquérito sobre segurança no espaço público e privado [Survey on safety in public and private spaces]. Instituto Nacional de Estatística. 2023. URL: https://ec.europa.eu/eurostat/web/products-statistical-reports/w/ks-ft-22-005 [accessed 2025-09-25]
3. Ziaei S, Hammarström A. The relationship between interpersonal violence in adulthood and mental health: a longitudinal study based on the Northern Swedish Cohort. BMC Public Health 2023;23(1):637 [FREE Full text] [doi: 10.1186/s12889-023-15525-x] [Medline: 37013550]
4. Strathearn L, Giannotti M, Mills R, Kisely S, Najman J, Abajobir A. Long-term cognitive, psychological, and health outcomes associated with child abuse and neglect. Pediatrics 2020;146(4):e20200438 [FREE Full text] [doi: 10.1542/peds.2020-0438] [Medline: 32943535]
5. Keeshin BR, Cronholm PF, Strawn JR. Physiologic changes associated with violence and abuse exposure: an examination of related medical conditions. Trauma Violence Abuse 2012;13(1):41-56. [doi: 10.1177/1524838011426152] [Medline: 22186168]
6. Rivara F, Adhia A, Lyons V, Massey A, Mills B, Morgan E, et al. The effects of violence on health. Health Aff (Millwood) 2019;38(10):1622-1629. [doi: 10.1377/hlthaff.2019.00480] [Medline: 31589529]
7. Clemente-Teixeira M, Magalhães T, Barrocas J, Dinis-Oliveira RJ, Taveira-Gomes T. Health outcomes in women victims of intimate partner violence: a 20-year real-world study. Int J Environ Res Public Health 2022;19(24):17035 [FREE Full text] [doi: 10.3390/ijerph192417035] [Medline: 36554916]
8. Neves I, Dinis-Oliveira RJ, Magalhães T. Epigenomic mediation after adverse childhood experiences: a systematic review and meta-analysis. Forensic Sci Res 2021;6(2):103-114 [FREE Full text] [doi: 10.1080/20961790.2019.1641954] [Medline: 34377567]
9. Nguyen KA, Abrahams N, Jewkes R, Mhlongo S, Seedat S, Myers B, et al. The associations of intimate partner violence and non-partner sexual violence with hypertension in South African women. Int J Environ Res Public Health 2022;19(7):4026. [doi: 10.3390/ijerph19074026] [Medline: 35409715]
10. Sanderson M, Cook M, Brown LL, Mallett V, Coker AL. Lifetime interpersonal violence or abuse and diabetes rates by sex and race. Am J Prev Med 2023;65(5):783-791 [FREE Full text] [doi: 10.1016/j.amepre.2023.06.007] [Medline: 37302511]
11. Daníelsdóttir HB, Aspelund T, Shen Q, Halldorsdottir T, Jakobsdóttir J, Song H, et al. Adverse childhood experiences and adult mental health outcomes. JAMA Psychiatry 2024;81(6):586-594. [doi: 10.1001/jamapsychiatry.2024.0039] [Medline: 38446452]
12. Stubbs A, Szoeke C. The effect of intimate partner violence on the physical health and health-related behaviors of women: a systematic review of the literature. Trauma Violence Abuse 2022;23(4):1157-1172. [doi: 10.1177/1524838020985541] [Medline: 33541243]
13. Scott-Storey K, O'Donnell S, Perrin N, Wuest J. Cumulative lifetime violence, gender, social determinants of health and mental health in Canadian men: a latent class analysis. J Fam Violence 2024;39(4):665-680 [FREE Full text] [doi: 10.1007/s10896-023-00502-0] [Medline: 38644980]

XSL•FO
RenderX

14. Csoboth CT, Birkás E, Purebl G. Living in fear of experiencing physical and sexual abuse is associated with severe depressive symptomatology among young women. J Womens Health (Larchmt) 2005;14(5):441-448. [doi: 10.1089/jwh.2005.14.441] [Medline: 15989417]
15. Johnson L, Chen Y, Stylianou A, Arnold A. Examining the impact of economic abuse on survivors of intimate partner violence: a scoping review. BMC Public Health 2022;22(1):1014 [FREE Full text] [doi: 10.1186/s12889-022-13297-4] [Medline: 35590302]
16. Graça H, Frazão SL, Magalhães T, Vieira-Pinto P, Gomes J, Taveira-Gomes T. Health conditions in older adults suspected of being maltreated: a 20-year real-world study. J Clin Med 2023;12(16):5247 [FREE Full text] [doi: 10.3390/jcm12165247] [Medline: 37629290]
17. Wilkins N, Myers L, Kuehl T, Bauman A, Hertz M. Connecting the dots: state health department approaches to addressing shared risk and protective factors across multiple forms of violence. J Public Health Manage Pract 2018:S32-S41 [FREE Full text] [doi: 10.1097/PHH.0000000000000669] [Medline: 29189502]
18. Scott-Storey K, O'Donnell S, Busolo D, DiTommaso E, Malcolm J, Taylor P, et al. Cumulative lifetime violence severity, social determinants and anxiety in a national sample of Canadian men. BMC Psychiatry 2022;22(1):265 [FREE Full text] [doi: 10.1186/s12888-022-03865-8] [Medline: 35421946]
19. Margolin G, Vickerman KA, Oliver PH, Gordis EB. Violence exposure in multiple interpersonal domains: cumulative and differential effects. J Adolesc Health 2010;47(2):198-205 [FREE Full text] [doi: 10.1016/j.jadohealth.2010.01.020] [Medline: 20638013]
20. Kyle J. Intimate partner violence. Med Clin North Am 2023;107(2):385-395. [doi: 10.1016/j.mcna.2022.10.012] [Medline: 36759104]
21. Colaprico C, Grima D, Shaholli D, Imperiale I, La Torre G. Workplace bullying in Italy: a systematic review and meta-analysis. Med Lav 2023;114(6):e2023049 [FREE Full text] [doi: 10.23749/mdl.v114i6.14673] [Medline: 38060211]
22. Mental disorders. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/mental-disorders [accessed 2024-07-01]
23. Campbell JC, Anderson JC, McFadgion A, Gill J, Zink E, Patch M, et al. The effects of intimate partner violence and probable traumatic brain injury on central nervous system symptoms. J Womens Health (Larchmt) 2018;27(6):761-767 [FREE Full text] [doi: 10.1089/jwh.2016.6311] [Medline: 29083256]
24. Lin L, Wang HH, Lu C, Chen W, Guo VY. Adverse childhood experiences and subsequent chronic diseases among middle-aged or older adults in China and associations with demographic and socioeconomic characteristics. JAMA Netw Open 2021;4(10):e2130143 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.30143] [Medline: 34694390]
25. Siewert-Markus U, Ittermann T, Klinger-König J, Grabe HJ, Stracke S, Völzke H, et al. Childhood maltreatment and risk of metabolic dysfunction-associated steatotic liver disease—evidence of sex-specific associations in the general population. J Psychosom Res 2024;183:111829 [FREE Full text] [doi: 10.1016/j.jpsychores.2024.111829] [Medline: 38896985]
26. Shea S, Lionis C, Kite C, Lagojda L, Uthman O, Dallaway A, et al. Non-alcoholic fatty liver disease and coexisting depression, anxiety and/or stress in adults: a systematic review and meta-analysis. Front Endocrinol (Lausanne) 2024;15:1357664 [FREE Full text] [doi: 10.3389/fendo.2024.1357664] [Medline: 38689730]
27. Kang D, Zhao D, Ryu S, Guallar E, Cho J, Lazo M, et al. Perceived stress and non-alcoholic fatty liver disease in apparently healthy men and women. Sci Rep 2020;10(1):38 [FREE Full text] [doi: 10.1038/s41598-019-57036-z] [Medline: 31913326]
28. Soto-Angona, Anmella G, Valdés-Florido MJ, De Uribe-Viloria N, Carvalho AF, Penninx BWJH, et al. Non-alcoholic fatty liver disease (NAFLD) as a neglected metabolic companion of psychiatric disorders: common pathways and future approaches. BMC Med 2020;18(1):261 [FREE Full text] [doi: 10.1186/s12916-020-01713-8] [Medline: 32998725]
29. Hughes K, Bellis MA, Hardcastle KA, Sethi D, Butchart A, Mikton C, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. Lancet Public Health 2017;2(8):e356-e366 [FREE Full text] [doi: 10.1016/S2468-2667(17)30118-4] [Medline: 29253477]
30. Martins B. Adverse childhood experiences and health outcomes: a 20-year real-world study. 2023. URL: https://repositorio-aberto.up.pt/handle/10216/150662 [accessed 2024-06-28]
31. Novais M, Henriques T, Vidal-Alves M, Magalhães T. When problems only get bigger: the impact of adverse childhood experience on adult health. Front Psychol 2021;12:693420 [FREE Full text] [doi: 10.3389/fpsyg.2021.693420] [Medline: 34335410]
32. Scott-Storey K, O'Donnell S, Wuest J, MacIntosh J, Merritt-Gray M. Cumulative lifetime violence severity scale: development and initial testing among men. BMC Public Health 2020;20(1):418 [FREE Full text] [doi: 10.1186/s12889-020-08551-6] [Medline: 32228553]
33. Oberth C, Goulter N, McMahon RJ. The comparative and cumulative impact of different forms of violence exposure during childhood and adolescence on long-term adult outcomes. Dev Psychopathol 2022;34(4):1313-1328. [doi: 10.1017/S0954579420002254] [Medline: 33779534]
34. Wiens K, Gillis J, Nicolau I, Wade TJ. Capturing risk associated with childhood adversity: independent, cumulative, and multiplicative effects of physical abuse, sexual abuse, and family violence on mental disorders and suicidality. Perm J 2020;24 [FREE Full text] [doi: 10.7812/TPP/19.079] [Medline: 32240085]

XSL•FO
RenderX

35. Webster EM. The impact of adverse childhood experiences on health and development in young children. Global Pediatr Health 2022;9:2333794X221078708 [FREE Full text] [doi: 10.1177/2333794X221078708] [Medline: 35237713]
36. Dicola D, Spaar E. Intimate partner violence. Am Fam Physician. 2016. URL: https://www.aafp.org/pubs/afp/issues/2016/1015/p646.html [accessed 2024-06-23]
37. Miller CJ, Adjognon OL, Brady JE, Dichter ME, Iverson KM. Screening for intimate partner violence in healthcare settings: an implementation-oriented systematic review. Implement Res Pract 2021;2. [doi: 10.1177/26334895211039894] [Medline: 36712586]
38. Kifle ME, Aychiluhm SB, Anbesu EW. Global prevalence of intimate partner violence during the COVID-19 pandemic among women: systematic review and meta-analysis. BMC Womens Health 2024;24(1):127 [FREE Full text] [doi: 10.1186/s12905-023-02845-8] [Medline: 38368323]
39. Jiang L, Sutherland MA, Hutchinson MK, Si B. A multi-center structural equation modeling approach to investigate interpersonal violence screening for public health promotion. Front Public Health 2021;9:637222 [FREE Full text] [doi: 10.3389/fpubh.2021.637222] [Medline: 34178912]
40. Hegarty K, McKibbin G, Hameed M, Koziol-McLain J, Feder G, Tarzia L, et al. Health practitioners' readiness to address domestic violence and abuse: a qualitative meta-synthesis. PLoS One 2020;15(6):e0234067 [FREE Full text] [doi: 10.1371/journal.pone.0234067] [Medline: 32544160]
41. Magalhães T, Dinis-Oliveira RJ, Taveira-Gomes T. Digital health and big data analytics: implications of real-world evidence for clinicians and policymakers. Int J Environ Res Public Health 2022;19(14):8364 [FREE Full text] [doi: 10.3390/ijerph19148364] [Medline: 35886214]

Abbreviations

EHR: electronic health record
HPA: hypothalamic-pituitary-adrenal
HR: hazard ratios
ICD-9: International Classification of Diseases, Ninth Revision
ICD-10: International Statistical Classification of Diseases, Tenth Revision
ICPC-2: International Classification of Primary Care, 2nd edition
IV: interpersonal violence
MASLD: metabolic dysfunction–associated steatotic liver disease
T2D: type 2 diabetes mellitus
ULSM: Unidade Local de Saúde de Matosinhos
WHO: World Health Organization

Edited by A Mavragani; submitted 13.08.24; peer-reviewed by E Finegood, AA Villa-Rueda; comments to author 12.03.25; revised version received 02.06.25; accepted 04.07.25; published 30.10.25.

Please cite as:
Barbosa R, Lopes R, Taveira-Gomes T, Ponte C, Magalhães T
Health Impact of Suspected Interpersonal Violence Against Adults: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e65191
URL: https://publichealth.jmir.org/2025/1/e65191
doi:10.2196/65191
PMID:

©Rui Barbosa, Rita Lopes, Tiago Taveira-Gomes, Carla Ponte, Teresa Magalhães. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Spatial Clusters of Condyloma Acuminata and the Regional Risk Factors in South Korea: Bayesian Spatial Regression Analysis

Joonsu Jang, MSci; Seyul Park, BA; Byung Chul Chun, MPH, MD, PhD

Department of Preventive Medicine, College of Medicine, Korea University, 73 Goryeodae-ro, Seong-buk Gu, Seoul, Republic of Korea

Corresponding Author:
Byung Chul Chun, MPH, MD, PhD
Department of Preventive Medicine, College of Medicine, Korea University, 73 Goryeodae-ro, Seong-buk Gu, Seoul, Republic of Korea

Abstract

Background: Condyloma acuminata (CA), the clinical manifestation of infection with low-risk human papillomaviruses 6 and 11, is a common sexually transmitted infection (STI) with recurrent lesions and notable psychosocial and health system burden. Recent evidence indicates a substantial global burden, with prevalence ranging from 0.5% to 33.1% and incidence ranging from 24 to 2940 per 100,000 person-years, varying by age, sex, time, and geography. In South Korea, national insurance data show sustained increases in patients receiving care for CA during 2010 to 2019. Beyond individual behaviors, spatial proximity and contextual factors can produce clustered STI risk. However, the municipal-level spatial distribution of CA in Korea and its contextual correlates remain understudied.

Objective: This study aimed to identify high-risk geographic clusters of CA in South Korea and determine the regional factors associated with its incidence rates.

Methods: We conducted an ecological analysis using 2019 municipal-level data from the National Health Insurance Service of Korea. Spatial autocorrelation of CA incidence rates was evaluated using Moran's I, and clustering was assessed with Getis-Ord Gi* to detect high-risk clusters. We then analyzed potential regional determinants using two Bayesian spatial regression models: the intrinsic conditional autoregressive model and the Besag–York–Mollié model. Key municipal-level variables included health behaviors, socioeconomic indicators, health care access, adult entertainment venue density, and risk of sexual violence. Results are reported as adjusted relative risks (aRRs) with 95% credible intervals (CrIs).

Results: A total of 52,009 CA cases were identified in 2019, 70.03% (36,421/52,009) of which were in men. We found significant positive spatial autocorrelation in CA incidence rates (Moran's I>0, *P*<.001), indicating nonrandom spatial clustering. The Getis-Ord Gi* analysis revealed several high-incidence clusters (*hotspots*) in metropolitan and southeastern regions of South Korea. In the Bayesian spatial models, higher CA incidence rates were associated with a greater share of the municipal budget spent on social welfare (aRR 1.005, 95% CrI 1.001 - 1.009), a higher percentage of single-person households (aRR 1.034, 95% CrI 1.025 - 1.043), and more adult entertainment establishments per 10,000 people (aRR 1.006, 95% CrI 1.001 - 1.012).

Conclusions: We identified significant geographic hotspots of CA and several community-level risk factors driving these patterns in South Korea. These findings highlight the importance of spatial surveillance and targeted public health interventions in high-risk areas. Adapting STI prevention programs to address local social determinants may help reduce the spread of CA in the identified hotspots.

(JMIR Public Health Surveill 2025;11:e76751) doi:10.2196/76751

KEYWORDS

sexually transmitted infections; condyloma acuminata; cluster analysis; spatial epidemiology; social determinants

Introduction

Background

Condyloma acuminata (CA), the clinical manifestation of infection with low-risk human papillomaviruses 6 and 11, remains a persistent public health concern due to its high incidence, recurrent course, psychosocial burden, and nontrivial health system costs [1]. Although generally asymptomatic, CA lesions can itch, bleed, and cause irritation or discomfort in the anogenital region [2]. Recent studies indicate that CA remains a substantial global burden, with prevalence ranging from 0.5% to 33.1% and incidence rates ranging from 24 to 2940 per 100,000 person-years, varying by age, sex, time, and geography [3]. In South Korea, population-based analyses using the Korean National Health Insurance Service (KNHIS) data showed sustained increases in patients receiving care for CA during 2010 - 2019 [4]. Earlier Health Insurance Review and Assessment Service in South Korea data (2007 - 2015) also

XSL•FO
RenderX

documented rising prevalence and economic burden at the national level [5].

Several individual-level risk factors for CA have been identified in prior studies. These include early age at first sexual intercourse, engaging in sexual activity with multiple partners, unprotected intercourse [6], a history of other sexually transmitted infections (STIs) [7], and smoking tobacco [8]. However, most of these studies focused on individual behaviors and did not examine how regional or community-level factors contribute to the geographical heterogeneity of CA incidence rates.

Beyond person-to-person transmission, spatial proximity effects can produce geographically clustered STI patterns through several ecological mechanisms. Evidence from large urban settings demonstrated that exposure to high-STI neighborhoods increases infection risk in both geographically proximate and socially connected areas through spillover [9]. Also, municipalities that share contextual determinants such as urbanicity, socioeconomic deprivation, and access to sexual health services tend to exhibit correlated STI risks over space and time. Recent spatial studies of STIs demonstrate these mechanisms across settings, including spatiotemporal clustering of HIV in China with strong associations to urbanization and health-system indicators, and persistent hot spots of curable STIs in subnational analyses [10]. Korean data also showed spatial clustering for STI, underscoring that geographically structured ecological determinants can shape STI risk beyond individual behavior [11]. These mechanisms justify spatial analyses to identify clustered risk and its contextual determinants.

Despite the increasing CA burden, evidence on the municipal-level spatial distribution of CA in South Korea remains limited. Specifically, no study has quantified municipal-level spatial clustering of CA in South Korea and examined how local contextual factors relate to CA incidence rate. Filling this gap is essential for geographically targeted prevention and service planning.

Objectives

We hypothesized that (1) CA incidence rates would cluster spatially (hot spots) and (2) area-level determinants would partially explain this clustering. To address these, we aimed to quantify municipal-level clustering of CA in South Korea and identify which regional characteristics are associated with higher CA incidence rates to inform geographically targeted interventions.

Methods

Study Design and Setting

We conducted a cross-sectional ecological study at the municipal level in South Korea for the year 2019. Reporting followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist for cross-sectional studies (Checklist 1). The unit of analysis was the Si-Gun-Gu (municipalities, including cities, counties, and districts), of which there are 250 nationwide. By focusing on area-level data across all municipalities, this design allowed us to assess geographic variations and identify municipal-level risk factors associated with CA incidence rates in a single year.

Data Sources and Variables

This study analyzed nationwide health insurance claims generated from the KNHIS. The KNHIS insures approximately 97% of residents, with the remainder covered by the Medical Aid program [12]. The KNHIS claims constitute almost complete national dataset for epidemiologic research in Korea. The claims files include dates of service, diagnosis and procedure codes, provider type and location, and basic beneficiary characteristics needed for population-level analyses; personally identifiable information is not available to investigators [13]. Because this was a secondary analysis of collected data, there was no clinical recruitment or patient enrollment.

The primary outcome was the incidence rates of CA in each municipality during 2019. CA cases were identified using the *International Classification of Diseases, 10th Revision, Clinical Modification* code A63.0 recorded in the principal or additional diagnosis fields. To define incident cases for 2019, first A63.0-coded claim was selected in 2019 as the index event. To avoid counting follow-up care from the same clinical episode as new incidents, a 30-day reattendance window was applied: any subsequent A63.0-coded claims occurring ≤30 days after the index claim were considered part of the same episode and not counted as additional incident events. The incident events were assigned to the municipality recorded in the claims and aggregated across the 250 municipalities of South Korea for 2019. Annual municipal incidence rates per 100,000 people were computed using midyear resident-registry denominators from Statistics Korea [14].

We assembled a broad set of municipal characteristics as potential risk factors for CA, drawing on multiple data sources. These variables encompassed several domains including health behaviors, welfare-related variables, health conditions, socioeconomic variables, health care accessibility, adult entertainment establishments, and the risk of sexual violence (Table S1 in Multimedia Appendix 1). Additionally, two demographic covariates were included to adjust for population structure: the sex ratio and the median age of the population in each municipality. These adjustments account for differences in the age and gender distribution that might influence CA incidence rates.

All municipal-level variables were obtained from authoritative national data sources for the year 2019 to ensure consistency with the outcome data. Health behavior, health condition, and welfare and socioeconomic indicators were extracted from the 2019 Korean Community Health Survey and the Korean Statistical Information Service databases [14,15]. The number of adult entertainment establishments in each area was gathered from the government's open data portal [16], which provides official public data on local businesses and facilities. The regional risk of sexual violence was derived from a report published by the Korea Institute of Criminology and Justice [17]. By integrating these diverse data sources, we compiled a comprehensive dataset of regional variables hypothesized to be associated with CA incidence rates.

XSL•FO
RenderX

Spatial Autocorrelation

The global Moran's I statistics for the incidence rates of CA across the 250 municipalities were calculated to examine whether the geographic distribution of CA incidence rates exhibited spatial autocorrelation [18]. Moran's I is a measure of global spatial autocorrelation that ranges from –1 to +1, where values close to +1 indicate strong clustering of similar values and values close to –1 indicate spatial dispersion of values. A value near 0 would indicate that CA incidence rates are spatially random. A spatial weights matrix was constructed based on the K-nearest neighboring municipalities for each area (K=4), using great-circle distances between municipality centroids to define neighbors. The K-nearest neighbors approach ensures that every municipality has a consistent number of neighboring areas considered, even for those with no direct geographic borders. Given this weights matrix, Moran's I was computed for the CA incidence rates per 100,000 population, and a permutation test with 999 Monte Carlo simulations was conducted to determine the significance of the observed Moran's I value. The null hypothesis for this test is that CA incidence rates are distributed randomly in space; a significant Moran's I (P<.05) would indicate that spatial autocorrelation is present, justifying the use of spatial modeling techniques in subsequent analyses [19].

Cluster Detection (Hotspot Analysis)

In addition to global autocorrelation, we explored local clusters of high or low CA incidence rates using the Getis-Ord Gi* statistic [20]. The Getis-Ord Gi* is a local spatial statistic that identifies *hotspots* and *coldspots*, meaning areas where the incidence rates are significantly higher or lower than would be expected by chance, given the values in neighboring areas. For each municipality, Gi* considers the value of CA incidence rates in that area and those of its surrounding neighbors (using K=4) and computes a z score. A high positive z score with a significant P value (<.05) indicates a clustering of high incidence rates (a hotspot), whereas a high negative z score indicates a clustering of low incidence rates (a coldspot). Municipalities with a Gi* statistic exceeding the 95% CI bounds were mapped as significant clusters of elevated or reduced CA incidence rates. This hotspot analysis complements the global Moran's I by identifying where spatial clusters occur, providing insight into specific high-risk regions.

Variable Selection and Nonspatial Regression Analysis

Prior to building spatial models, we conducted a nonspatial regression analysis to identify important predictors of CA incidence rates and to ensure that our model would account for overdispersion in the count data [21]. Initially, a Poisson regression framework was considered because the number of CA cases in each municipality is a count outcome. However, we tested for overdispersion in the count data by comparing the variance to the mean and using an overdispersion test statistic [22]. The overdispersion test indicated that the variance of CA cases across municipalities was significantly greater than the mean (P<.05). We therefore chose a negative binomial regression model to account for the overdispersion.

Using the negative binomial framework, bivariate regression analyses were performed for each candidate independent variable. In these models, the outcome was the count of CA cases in the municipality, and the log of the municipality's population was included as an offset term. Each candidate variable was entered individually to assess its crude association with CA incidence rates. A liberal significance criterion of P≤.25 was used to screen for potential predictors [23]. In other words, any variable that showed an association with CA incidence rates at P≤.25 in the bivariate analysis was retained for consideration in the multivariable model. This cutoff is more inclusive than the conventional .05 level and was chosen to avoid prematurely excluding variables that might become significant in a multivariable context [23].

Next, a multivariable negative binomial regression model was used, including all candidate variables that passed the screening threshold. Variance inflation factors were computed for the covariates to check for multicollinearity, with a variance inflation factor >5 used as an indicator of high multicollinearity that could distort the model estimates [24]. If any pair of variables were highly collinear, we planned to remove or combine one of them, favoring the variable with more direct relevance to CA or the one that yielded the better model fit. The final nonspatial model was selected through a backward elimination procedure. Starting with the full set of selected covariates, we iteratively removed the least significant variable one at a time and compared model fit using the Akaike Information Criterion. The final model provided a baseline set of risk factors for CA incidence rates without accounting for spatial effects.

Bayesian Spatial Modeling

We incorporated spatial dependence into our regression modeling using a Bayesian approach because the exploratory spatial analysis revealed evidence of nonrandom spatial clustering in CA incidence rates. A Bayesian hierarchical spatial modeling was adopted because the outcome comprises municipal counts exhibiting overdispersion and residual spatial autocorrelation. Traditional nonhierarchical spatial regression often struggles to accommodate these characteristics in small-area count settings due to the absence of area-level random effects to capture residual dependence [25]. The Bayesian framework enables us to specify a negative binomial likelihood with a population offset, incorporate area-level spatial random effects that borrow strength from neighboring municipalities, and quantify uncertainty with posterior intervals.

Two types of Bayesian spatial regression models were implemented to account for residual spatial autocorrelation in the incidence rates of CA across municipalities. These were the intrinsic conditional autoregressive (ICAR) model and the Besag–York–Mollié (BYM) model [26]. Both models are commonly used in disease mapping to improve estimates by *borrowing strength* from neighboring areas and to control for spatially structured confounding [27]. The ICAR model introduces a spatially structured random effect that assumes nearby areas have similar risk, whereas the BYM model includes two random effect components—one spatially structured and one unstructured—allowing it to capture both localized spatial clustering and independent heterogeneity at the municipal level.

XSL•FO
RenderX

The municipal incidence rates of CA were modeled using a negative binomial likelihood. We let Y_i denote the counts of CA in municipality i (i=1,…,250).

(1)Yi~ NB(r, pi), pi=rr+λi, E(Yi)=λi, Var(Yi)=λi+λi2r

Here, r>0 is the size (overdispersion) parameter and p_i is the success probability. The mean λi is linked to covariates with a population offset:

(2)log λi=α+log Ni+∑kβkXi,k+ηi

where α is the intercept, N_i is the municipal population (offset), $X_{i,k}$ are municipal-level covariates with coefficients βk, and ηi is a spatial random effect. To account for residual spatial autocorrelation, we first specified ηi under an ICAR prior, which imposes spatial smoothing such that neighboring municipalities have similar risk:

(3)ηi=si, s=(s1, …, s250)T~ ICAR (W, σs2)

The spatial weights matrix W was built using K=4 K-nearest neighbors computed from municipal centroids in projected coordinates. We then considered the BYM extension, which augments the ICAR component with an independent, spatially unstructured term to capture residual heterogeneity:

(4)ηi=si+ui,ui~N(0,σu2)

Both ICAR and BYM models are widely used in disease mapping to borrow strength across adjacent areas and to mitigate spatially structured confounding [28,29]. We fitted both ICAR and BYM specifications and compared performance using deviance information criterion (DIC) values with practical and theoretical considerations. For regression outputs, the posterior median of the exponentiated fixed effects was reported as an adjusted relative risk (aRR) with a 95% credible interval (CrI). For clarity, in the Bayesian framework, a 95% CrI denotes the range containing 95% of the posterior probability mass for the parameter, given the model and prior. Covariate effects were considered significant when the 95% CrI for the aRR does not include 1.00.

The posterior marginal distribution of the model parameters was approximated using the integrated nested Laplace approximation [30]. A flat prior distribution was applied to the intercept α and the regression parameters β corresponding to each explanatory variable and covariate. For the variance associated with spatially structured and spatially unstructured terms, a weakly informative prior known as the Penalized Complexity prior was applied [31].

We conducted a sensitivity analysis to examine the robustness of the findings to the choice of spatial neighborhood structure. The primary analysis defined spatial adjacency using each municipality's four nearest neighbors (K=4). For sensitivity testing, the Bayesian spatial model was rerun under two alternative neighbor definitions: first using three nearest neighbors (K=3) and then using five nearest neighbors (K=5) for each municipality. All the results were compared across the different neighbor specifications: spatial autocorrelation statistics, identified hotspots, and key regression coefficients. The sensitivity analysis confirmed that the overall spatial patterns and the identified significant predictors of CA incidence rates remained consistent when using K=3 or K=5, suggesting that our findings are robust to reasonable variations in the spatial neighborhood definition.

All statistical analyses were performed using R version 4.3.1 (R Foundation for Statistical Computing). Spatial autocorrelation and clustering statistics (global Moran's I and Getis-Ord Gi*) were computed with the *spdep* package [32]; all maps were produced with *tmap* [33]; the nonspatial negative binomial baseline model was fitted using *MASS* [34]; overdispersion was evaluated with *AER* [22]; and Bayesian spatial models were estimated with *R-INLA* [30]. The shape file was obtained from the National Spatial Geographic Information Service [35].

Ethical Considerations

This study was conducted using aggregated, deidentified health data and publicly available municipal statistics. The Institutional Review Board of Korea University reviewed and approved this study (IRB approval number: KUIRB-2021-0382-01).

Results

Descriptive Statistics

In 2019, a total of 52,009 cases of CA were recorded across the 250 municipalities, with approximately 70.03% (36,421/52,009) of cases in men. Across municipalities, the mean number of CA cases was 207.3 (SD 179.5) per municipality, and the mean CA incidence rate was 95.8 per 100,000 people (Table 1). The burden of CA varied widely between regions: the number of cases ranged from as low as 3 to as high as 886 in different municipalities, and incidence rates ranged from 1.46 to 184.21 per 100,000 people. This represents more than a 100-fold difference in incidence rates across communities, highlighting substantial geographic disparities in CA occurrence. The coefficient of variation for CA cases was 0.87, reflecting this high relative dispersion. By comparison, the potential regional risk factors examined showed coefficients of variation ranging from 0.01 up to 0.85, indicating that none of these variables exhibited as much regional variability as the incidence rates of CA.

XSL•FO
RenderX

Table . Descriptive statistics for condyloma acuminata (CA; *ICD-10-CM*[a] A63.0) and potential regional risk factors across 250 municipalities examined in this nationwide cross-sectional ecological study of South Korea (2019).[b]

		Mean (SD)	Minimum	Median	Maximum	CV[c]
Cases of CA, n		207.26 (179.48)	3	159.50	886	0.87
Incidence rates of CA (per 100,000 people)		95.83 (27.08)	1.46	95.04	184.21	0.28
Health behavior						
	Exercise and physical activities (%)	25.36 (6.88)	5.50	24.75	56.90	0.27
	Current smoking (%)	20.27 (3.09)	11.60	20.25	28.80	0.15
	Alcohol consumption (%)	59.28 (4.36)	44.80	59.90	71.00	0.07
	High-risk drinking (%)	14.16 (2.98)	5.80	14.10	24.80	0.21
Welfare-related variables						
	Social welfare facilities per 100,000 people	19.12 (11.83)	2.30	16.65	73.70	0.62
	Share of municipal budget on social welfare (%)	34.70 (14.71)	11.40	32.90	66.60	0.42
	Subjective health perception (%)	42.64 (6.98)	29.70	41.40	68.30	0.16
	EQ-5D index	0.96 (0.01)	0.93	0.96	0.99	0.01
	Healthy living practice (%)	29.02 (9.48)	9.60	28.50	54.70	0.33
Health conditions						
	Prevalence of depression (%)	3.04 (1.38)	0.40	2.90	7.20	0.45
	Prevalence of obesity (%)	34.38 (3.72)	24.30	34.25	44.80	0.11
	Prevalence of diabetes (%)	8.04 (1.32)	5.00	8.00	11.80	0.16
	Prevalence of hypertension (%)	19.44 (2.28)	14.90	19.30	27.40	0.12
Socioeconomic variables						
	Low educational attainment (%)	51.39 (11.65)	14.84	52.32	73.13	0.23
	Single-person household (%)	23.90 (5.50)	12.99	23.80	39.05	0.23
	Divorce per 1000 people	2.18 (0.42)	1.20	2.20	4.10	0.19
	Financial autonomy (%)	25.16 (14.01)	7.05	20.93	68.90	0.56
Health care accessibility						
	Doctors per 1000 people	2.78 (2.20)	1.00	2.30	19.60	0.79
	Unmet medical facilities (%)	6.48 (3.40)	0.90	6.00	19.10	0.52
Adult entertainment and sexual violence						

XSL•FO
RenderX

		Mean (SD)	Minimum	Median	Maximum	CV[c]
	Adult entertainment establishments per 10,000 people	7.70 (6.53)	0.11	6.32	64.29	0.85
	Sexual violence risk	99.95 (18.22)	68.53	96.22	203.78	0.18
Covariates						
	Median age (y)	47.05 (6.03)	36.80	45.55	61.00	0.13
	Sex ratio (%)	101.53 (6.72)	87.60	100.40	134.70	0.07

[a]*ICD-10-CM*: *International Classification of Diseases, 10th Revision, Clinical Modification.*

[b]The table summarizes municipal-level determinants alongside annual CA cases and incidence rates per 100,000 people.

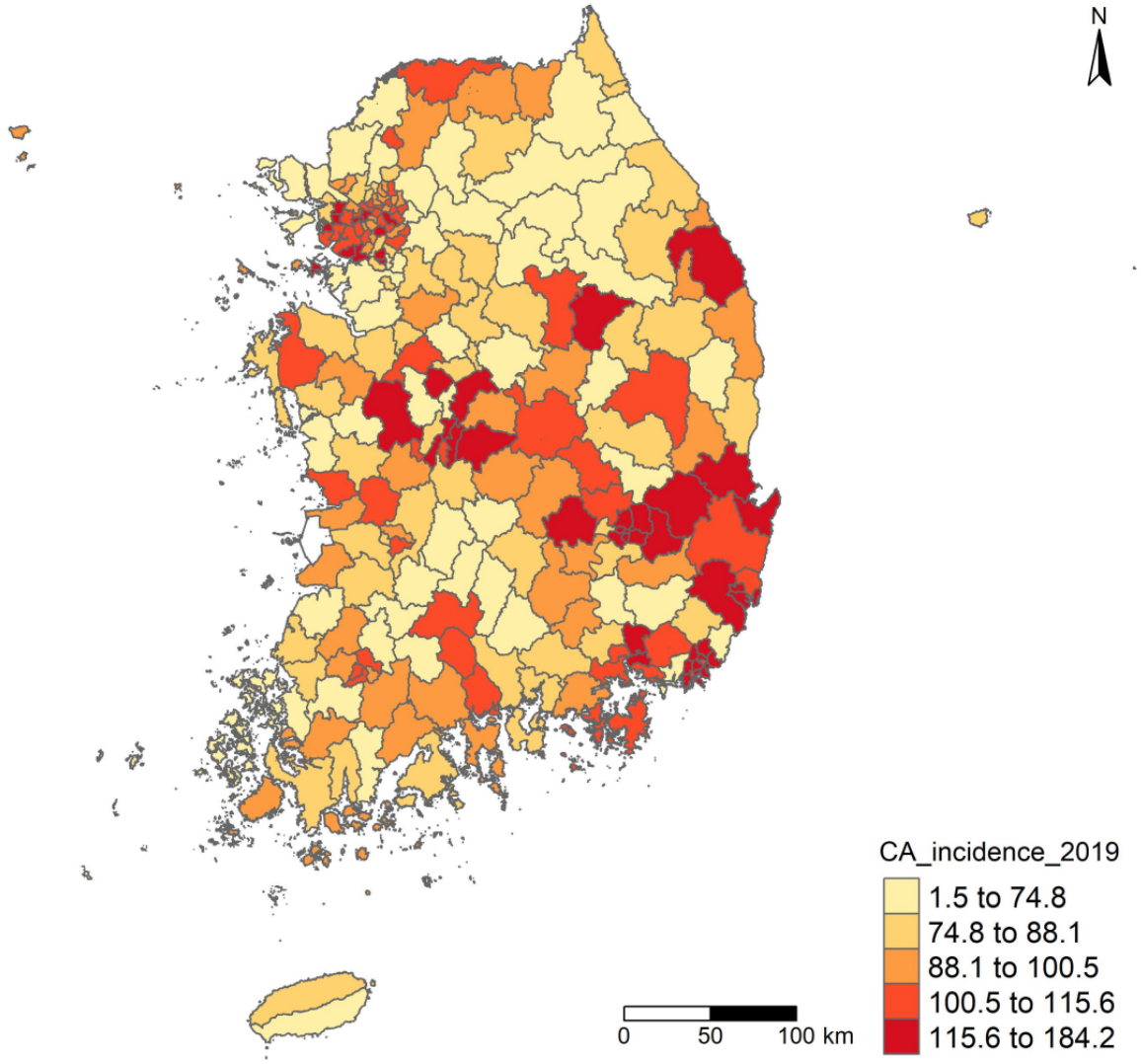
[c]CV: coefficient of variance.

Spatial Clustering

The choropleth map of CA incidence rates (Figure 1; Multimedia Appendix 2) indicated that municipalities in certain metropolitan and southeastern parts of South Korea showed higher incidence rates compared to those in the northeastern and southwestern regions, where rates were lower. Consistent with these visual patterns, the Global Moran's I analysis confirmed significant positive spatial autocorrelation. Specifically, the incidence rates of CA were spatially clustered across municipalities (Moran's I for incidence rates=0.335, $P<.001$), as were the case counts (Moran's I=0.395, $P<.001$). Local cluster detection using the Getis-Ord Gi* statistic further identified specific areas of high and low incidence rates. The Gi* analysis identified several significant hotspots of increased CA incidence rates in the urbanized metropolitan area and the southeastern region (Figure 2).

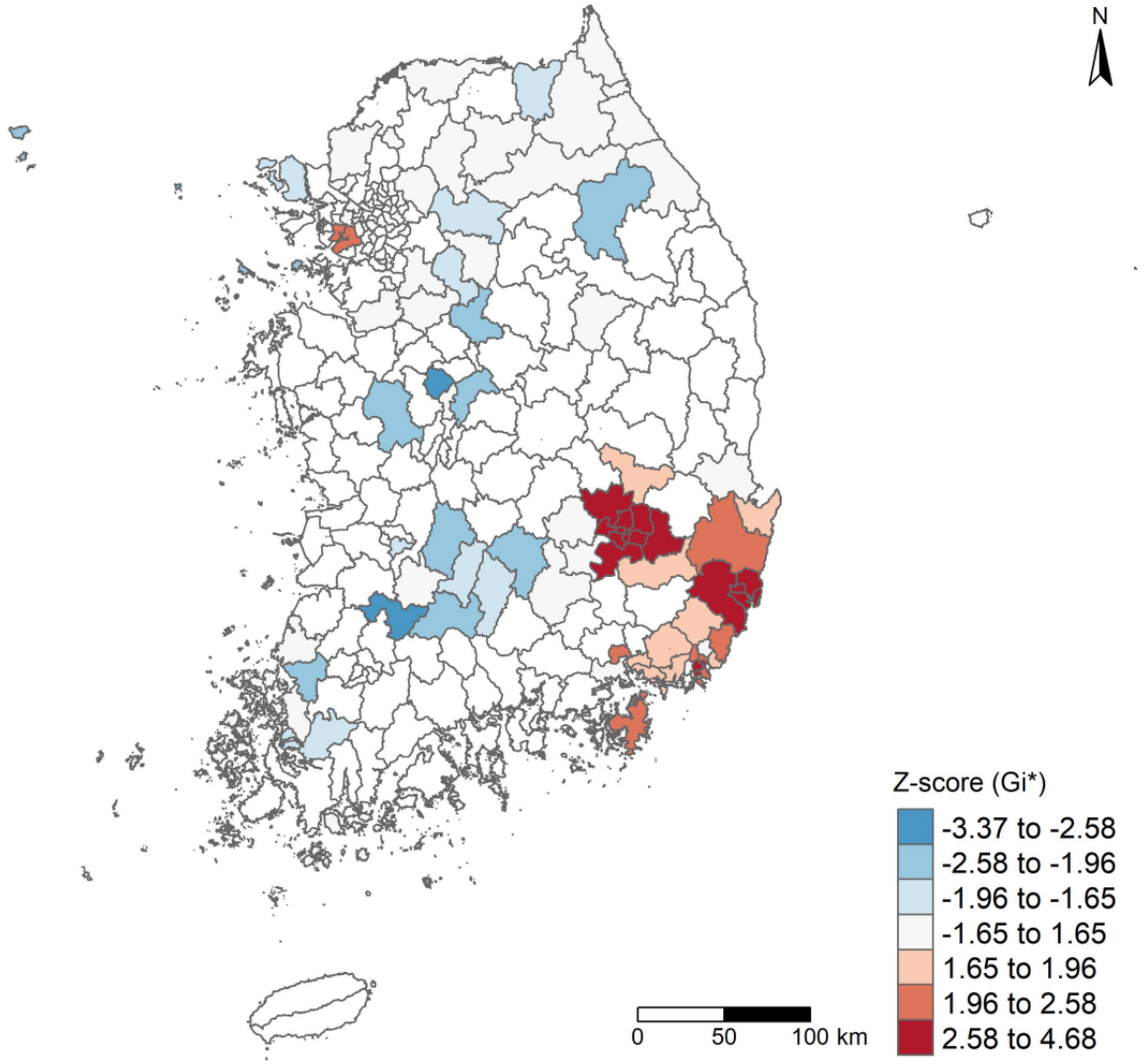
XSL•FO
RenderX

Figure 1. Geographic distribution of condyloma acuminata (CA; *International Classification of Diseases, 10th Revision, Clinical Modification* A63.0) incidence rates per 100,000, by municipality (n=250), in this nationwide cross-sectional ecological study of South Korea (2019). Incident CA cases were aggregated to the municipality of residence; denominators are midyear resident registry populations. The choropleth shades each municipality by CA incidence rates. Lighter to darker tones indicate lower to higher incidence rates.



XSL•FO
RenderX

Figure 2. Hotspot and coldspot clustering of condyloma acuminata (CA; *International Classification of Diseases, 10th Revision, Clinical Modification* A63.0) incidence rates per 100,000 based on the Getis–Ord Gi* analysis at the municipality (n=250) level for South Korea (2019). Spatial weights were defined by K-nearest neighborhood contiguity (K=4). Municipalities are colored by Gi* *z* score: statistically significant hotspots are shown in red tones (clusters of high incidence rates), and statistically significant coldspots are shown in blue tones (clusters of low incidence rates); nonsignificant areas are shown in white. Significance thresholds are $|z|\geq1.96$ ($P<.05$) and $|z|\geq2.58$ ($P<.01$).



Regression Modeling

After the preliminary nonspatial analysis and variable selection procedure, five candidate variables were selected for the spatial regression modeling: social welfare facilities per 100,000 people, share of municipal budget on social welfare (%), percentage of single-person households, prevalence of hypertension (%), and number of adult entertainment establishments per 10,000 people. In addition, two demographic covariates, median age and sex ratio, were included in the models to adjust for population structure. Model fit was similar for the ICAR and BYM specifications (ICAR: DIC 2422.38; BYM: DIC 2421.36), with ΔDIC=1.02. While the DIC difference was minimal, we selected the BYM model for theoretical reasons. The BYM model accounts for both structured spatial dependence and unstructured spatial heterogeneity, providing a more comprehensive representation of spatial effects. The BYM model decomposed the total spatial variance into structured (precision=101.48, 95% CI 34.84 - 250.05) and unstructured components (precision=17,485.45, 95% CI 202.07 - 119,072.66), indicating the presence of both spatial clustering and municipality-specific variation. Therefore, we present the results from the BYM model as the final spatial regression model (Table 2). The ICAR model yielded nearly identical fixed effect estimates, with magnitudes of fixed effect, directions, and statistical significance remaining consistent across both models, demonstrating that our findings are robust to spatial model specification.

The Bayesian spatial regression identified three key regional factors that were significantly associated with higher CA incidence rates after adjusting for covariates and spatial effects. First, the share of the municipal budget devoted to social welfare showed a positive association: for each 1% increase in welfare budget share, the aRR of CA increased to 1.004 (95% CrI

XSL•FO
RenderX

1.001 - 1.008), suggesting a small but significant increase in risk. Second, the proportion of single-person households in a municipality was strongly associated with CA incidence rates: each 1% increase in single-person households corresponded to the aRR of 1.033 (95% CrI 1.024 - 1.041). Third, the density of adult entertainment establishments was linked to higher CA incidence rates: for each additional adult entertainment establishment per 10,000 population, the aRR was 1.006 (95% CrI 1.001 - 1.012).

Sensitivity analyses confirmed the robustness of the spatial regression findings. We refit the Bayesian models using alternative spatial neighborhood definitions. The model fit remained best when using K=4 neighbors, as evidenced by the lowest DIC at K=4 compared to K=3 or K=5. Importantly, the direction and magnitude of the associations for the identified risk factors were consistent across these different spatial weights (Table S2 in Multimedia Appendix 1).

Table . Association between municipal-level determinants and incidence rates of condyloma acuminata (*ICD-10-CM*[a]A63.0) in this nationwide cross-sectional ecological study of South Korea (2019).

	Nonspatial model (multiple negative binomial regression)[b]	Spatial models[c]	
		ICAR[d]	BYM[e]
Social welfare facilities per 100,000 people, aRR[f] (95% CrI[g])	0.998 (0.994 - 1.001)	0.998 (0.995 - 1.002)	0.998 (0.995 - 1.002)
Share of municipal budget on social welfare (%), aRR (95% CrI)	1.006 (1.003 - 1.010)	1.005 (1.001 - 1.008)	1.004 (1.001 - 1.008)
Single-person household (%), aRR (95% CrI)	1.030 (1.022 - 1.039)	1.032 (1.024 - 1.041)	1.033 (1.024 - 1.041)
Prevalence of hypertension (%), aRR (95% CrI)	0.987 (0.973 - 1.002)	0.997 (0.981 - 1.013)	0.998 (0.983 - 1.014)
Adult entertainment establishments per 10,000 people, aRR (95% CrI)	1.008 (1.003 - 1.013)	1.006 (1.001 - 1.012)	1.006 (1.001 - 1.012)
Median age (y), aRR (95% CrI)	1.019 (1.011 - 1.027)	1.014 (1.006 - 1.023)	1.014 (1.005 - 1.022)
Sex ratio (%), aRR (95% CrI)	1.008 (1.002 - 1.014)	1.005 (0.999 - 1.011)	1.004 (0.998 - 1.010)
Dispersion parameter (*r*), posterior median (95% CrI)	22.518 (17.825 - 27.974)	27.799 (20.853 - 36.094)	29.022 (21.620 - 38.022)
Precision of spatially structured component (τs), posterior median (95% CrI)	—[h]	160.414 (42.191 - 478.918)	101.485 (34.842 - 250.054)
Precision of spatially unstructured component (τu), posterior median (95% CrI)	—	—	17,485.451 (202.075 - 119,072.658)
DIC[i]	2440.852	2422.379	2421.362

[a]*ICD-10-CM*: *International Classification of Diseases, 10th Revision, Clinical Modification.*

[b]Nonspatial generalized regression model.

[c]Two Bayesian spatial models with municipal adjacency based on K-nearest neighborhood contiguity (K=4).

[d]ICAR: intrinsic conditional auto-regressive model.

[e]BYM: Besag–York–Mollié model.

[f]aRR: adjusted relative risk (fixed effect).

[g]CrI: credible interval.

[h]Not applicable.

[i]DIC: deviance information criterion.

Discussion

Principal Findings

This study revealed geographic disparities in the incidence rates of CA across South Korea. A total of 52,009 CA cases were analyzed, and the municipal-level incidence rates ranged from as low as 1.46 to 184.21 per 100,000 people—over a 100-fold difference between the lowest and highest areas (Table 1). Such variation indicates substantial spatial heterogeneity in CA burden. Indeed, we found significant positive global spatial autocorrelation (Moran's I=0.33, *P*<.001), confirming that municipalities with high CA incidence rates tended to cluster near each other. The choropleth map of CA incidence rates (Figure 1) and the Getis-Ord Gi* hotspot analysis (Figure 2) consistently identified distinct high-risk clusters in the urbanized metropolitan regions and the southeastern part of the country. These hotspots emphasize that the transmission of CA is not

XSL•FO
RenderX

distributed evenly but rather concentrated in certain communities at an increased risk. In addition to mapping clusters, our analysis identified key regional factors associated with the spatial distribution of CA. Using Bayesian spatial regression modeling, we found three significant municipal-level risk factors that remained positively associated with higher CA incidence rates after adjusting for demographic covariates and spatial autocorrelation (Table 2). First, municipalities that spent a higher share of their budget on social welfare programs had a higher relative risk of CA. Second, areas with a higher proportion of single-person households showed increased CA incidence rates. Third, municipalities with more adult entertainment establishments per 10,000 people had higher CA rates. The Bayesian model that included both spatially structured and unstructured random effects—the BYM model—provided a better fit than the purely spatial ICAR model, indicating that accounting for local heterogeneity in addition to broad spatial trends improved the explanation of CA incidence rates patterns.

Our findings of spatial clustering of CA incidence rates are consistent with patterns observed in other settings. The spatial autocorrelation observed in CA incidence rates does not result from direct person-to-person transmission between neighboring individuals but rather from the geographic clustering of municipal-level risk factors that create similar high-risk environments in spatially proximate areas. Previous geographic studies have also identified localized hotspots for STIs. For example, mapping of STIs in Portugal found distinct high-rate clusters in certain regions [36], and a cross-sectional analysis reported that per-capita STI incidence rates tended to increase systematically with urban population size [37]. Likewise, cross-setting analyses report nonrandom spatial clustering of STIs that aligns with contextual determinants such as urbanicity, socioeconomic conditions, and access to sexual health services [9,38]. In our study, municipalities with higher proportions of single-person households and greater density of adult entertainment establishments exhibited a higher incidence of CA, suggesting that local sexual market structures and alcohol-related social environments may influence opportunities for risky behavior [39]. Furthermore, the highest CA incidence rates were found in densely populated urban and southeastern areas, which is consistent with other studies and suggests that large urban centers often serve as hubs for STI transmission. This aligns with the concept of core transmission areas, where densely interconnected sexual networks facilitate wider spread of infection [40]. Our results support this concept, as the identified metropolitan clusters may act as reservoirs that support wider disease spread. A similar spatial concentration of STIs has been demonstrated in the United States, where county-level analyses using hotspot methods have identified significant clusters of increased STI prevalence [41]. Both our study and prior research suggest that geographic context, especially urbanicity and regional connectivity, plays a central role in STI dynamics.

The positive association we observed between the prevalence of single-person households and CA incidence rates is supported by epidemiological data on sexual behavior. Single individuals are more likely to have multiple sexual partners compared to married or cohabiting individuals [42], and having multiple partners is a well-established risk factor for acquiring CA and other STIs [43]. A Turkish study similarly noted that unmarried status was linked to higher risk of human papillomavirus infection [44]. Thus, areas with many one-person households may inherently have higher levels of casual or transient sexual partnering in the population, contributing to greater CA transmission. Our findings tentatively extend this individual-level evidence to the community level, indicating that municipalities with a higher proportion of single-person households also tend to have higher incidence rates of CA.

The finding that municipalities with more adult entertainment establishments have higher CA incidence rates is also consistent with previous research. Such establishments may foster networks of high-risk sexual contacts. Studies in the United States have documented that individuals working in or around the adult entertainment industry have a higher STI prevalence and engage in riskier sexual behaviors than the general population [45,46]. These studies, combined with our findings, point to adult entertainment establishments as important sites of STI transmission. They likely attract or employ core groups with higher numbers of sexual encounters, thereby amplifying the spread of infection in the surrounding community. Our findings reinforce the need for targeted public health intervention in such areas to reduce local transmission of CA.

In contrast, the positive association we found between higher social welfare expenditure and CA incidence rates appears to be a novel observation with no direct precedent in the literature. This counterintuitive finding should be interpreted with caution. One plausible explanation is reverse causation or confounding: municipalities facing a high burden of CA may allocate a higher share of their budget to social services and public health, including STI management and prevention programs. In other words, increased welfare spending may be a response to the existing public health needs rather than a causal driver of infection. Alternatively, this association may reflect underlying social deprivation regions with need for welfare support could simultaneously have higher STI rates, as suggested by broader public health studies linking low socioeconomic status to increased STI prevalence [47]. CA is a condition requiring long-term management with a tendency for recurrence [48]. Communities heavily impacted by recurrent CA cases may incur higher health care and social support costs, prompting larger welfare budgets. Since no prior study, to our knowledge, has examined this factor in relation to CA, our result highlights a new avenue for research. Further studies are needed to disentangle whether higher social welfare spending is a marker of underlying social determinants that facilitate CA spread or whether it reflects an intensified local response to control an ongoing epidemic [49].

Limitations

This study has limitations. First, the analysis was conducted at the municipal level using an ecological design. All data, including CA cases and risk factors, were aggregated by municipality; individual-level exposures and behaviors could not be assessed. Consequently, the associations we identified do not necessarily imply causation at the individual level, and they may be susceptible to the ecological fallacy. Second, the

XSL•FO
RenderX

cross-sectional nature of our study limits our ability to infer temporal trends or causality. We captured a one-year snapshot of CA incidence rates and regional characteristics, but the situation may evolve over time. Third, our reliance on the NHIS records for CA cases means the data are subject to any inconsistencies in diagnosis or reporting. Genital warts carry social stigma [50], which could lead to underdiagnosis or underreporting in certain areas. For instance, individuals might avoid seeking treatment for an STI due to shame, or clinicians might use nonspecific diagnostic codes to spare patient discomfort. Such variability in health care–seeking behavior and diagnostic practices across municipalities could introduce measurement bias in the recorded incidence rates. Fourth, while we included a wide range of regional covariates, it is possible that unmeasured confounders also influence the spatial distribution of CA. Finally, the spatial resolution of our analysis was at the municipality level; intra-municipal variation or clustering at a finer neighborhood scale could not be examined with the available data. Despite these limitations, this study provides a foundation for addressing CA incidence rates through evidence-based strategies, focusing on identified hotspots and regional risk factors.

Conclusions

In summary, our nationwide spatial analysis of CA in South Korea revealed significant geographic clustering, with hotspots concentrated in the metropolitan and southeastern regions, and demonstrated that area-level factors, particularly a higher prevalence of single-person households and greater density of adult entertainment establishments, are associated with increased incidence rates of CA. These findings highlight the value of integrating spatial epidemiology into routine surveillance to identify communities at greatest risk and to guide the efficient allocation of targeted interventions. Moreover, the observed association with social welfare expenditure suggests that policymakers should consider how recurrent STI burdens strain local resources and ensure that funding for prevention and care is targeted to the most affected areas. Future research should explore the causal pathways linking these regional characteristics to CA transmission and evaluate the impact of geographically targeted public health interventions over time.

Acknowledgments

This work was supported by the Korea Disease Control and Prevention Agency (grant number 2021-07-010). The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to restrictions imposed by the Korean National Health Insurance Service (KNHIS). The KNHIS data can only be accessed and analyzed within a restricted secure network environment, and only aggregated results can be exported from this system. Researchers interested in accessing similar data should apply directly to the KNHIS Data Sharing Service [51], following their established protocols for data access approval. The corresponding author can provide the study's variable list and guidance on the KNHIS application process upon reasonable request.

Authors' Contributions

JJ and BCC conceptualized and designed the study. JJ and SP obtained data. JJ performed the analysis and BCC validated the results. JJ and BCC drafted the manuscript. JJ, SP, and BCC critically revised the intellectual content of the manuscript. All the authors contributed to the reading, provided feedback, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Description and source of study variables and Bayesian spatial regression model results.
[DOCX File, 27 KB - publichealth_v11i1e76751_app1.docx]

Multimedia Appendix 2
Regional distribution of condyloma acuminata (CA) cases in South Korea in 2019. Bright yellow shading represents a low number of CA cases, whereas red shading denotes a high number of CA cases.
[PNG File, 570 KB - publichealth_v11i1e76751_app2.png]

Checklist 1
STROBE checklist.
[PDF File, 193 KB - publichealth_v11i1e76751_app3.pdf]

XSL•FO
RenderX

References

1. Yanofsky VR, Patel RV, Goldenberg G. Genital warts: a comprehensive review. J Clin Aesthet Dermatol 2012 Jun;5(6):25-36. [Medline: 22768354]
2. Grennan D. Genital warts. JAMA 2019 Feb 5;321(5):520. [doi: 10.1001/jama.2018.20181] [Medline: 30721297]
3. You X, Valente S, Machalek DA, et al. Epidemiologic, clinical, and economic burden of anogenital warts among adults: a global systematic review. Int J Infect Dis 2025 Aug 27:108029. [doi: 10.1016/j.ijid.2025.108029] [Medline: 40882763]
4. Kim EJ, Lee JC, Lyu DH, et al. Trends of genital wart in Korea according to treatment method classification: big data analysis of health care in 2010-2019. Investig Clin Urol 2023 Jan;64(1):56-65. [doi: 10.4111/icu.20220318] [Medline: 36629066]
5. Park YJ, Kim JM, Lee BR, Kim TH, Lee EG. Annual prevalence and economic burden of genital warts in Korea: Health Insurance Review and Assessment (HIRA) service data from 2007 to 2015. Epidemiol Infect 2018 Jan;146(2):177-186. [doi: 10.1017/S0950268817002813] [Medline: 29235433]
6. Del Pino M, Vorsters A, Joura EA, et al. Risk factors for human papillomavirus infection and disease: a targeted literature summary. J Med Virol 2024 Feb;96(2):e29420. [doi: 10.1002/jmv.29420] [Medline: 38377121]
7. Kjaer SK, Tran TN, Sparen P, et al. The burden of genital warts: a study of nearly 70,000 women from the general female population in the 4 Nordic countries. J Infect Dis 2007 Nov 15;196(10):1447-1454. [doi: 10.1086/522863] [Medline: 18008222]
8. Kaderli R, Schnüriger B, Brügger LE. The impact of smoking on HPV infection and the development of anogenital warts. Int J Colorectal Dis 2014 Aug;29(8):899-908. [doi: 10.1007/s00384-014-1922-y] [Medline: 24935346]
9. Newmyer L, Evans M, Graif C. Socially connected neighborhoods and the spread of sexually transmitted infections. Demography 2022 Aug 1;59(4):1299-1323. [doi: 10.1215/00703370-10054898] [Medline: 35838157]
10. Xie Z, Chen B, Duan Z. Spatiotemporal analysis of HIV/AIDS incidence in China from 2009 to 2019 and its association with socioeconomic factors: geospatial study. JMIR Public Health Surveill 2024 Jun 7;10:e56229. [doi: 10.2196/56229] [Medline: 38848123]
11. Jang J, Park S, Chun BC. How place shapes genital herpes simplex distribution in South Korea: a Bayesian spatial analysis using National Health Insurance Service data. BMC Public Health 2025 Sep 2;25(1):3023. [doi: 10.1186/s12889-025-24171-4] [Medline: 40898057]
12. Lim SJ, Jang SI. Leveraging National Health Insurance Service Data for public health research in Korea: structure, applications, and future directions. J Korean Med Sci 2025 Mar 3;40(8):e111. [doi: 10.3346/jkms.2025.40.e111] [Medline: 40034096]
13. Cheol Seong S, Kim YY, Khang YH, et al. Data resource profile: the National Health Information database of the National Health Insurance Service in South Korea. Int J Epidemiol 2017 Jun 1;46(3):799-800. [doi: 10.1093/ije/dyw253] [Medline: 27794523]
14. Statistics Korea [Webpage in Korean]. Korean Statistical Information Service. URL: https://kosis.kr/index/index.do [accessed 2025-05-01]
15. Korean Community Health Survey [Webpage in Korean]. Korea Disease Control and Prevention Agency (KDCA). URL: https://chs.kdca.go.kr/chs/index.do [accessed 2024-12-01]
16. Public data portal [Webpage in Korean]. Ministry of the Interior and Safety. URL: https://www.data.go.kr [accessed 2024-12-01]
17. Park J, Kang Y, Kim D, Jeong J. The Development of Crime Risk Assessment Tool and Its Application in South Korea (III): With Focus on Sexual Violence and School Violence: Korean Institute of Criminology; 2014.
18. Besag J, Diggle PJ. Simple Monte Carlo tests for spatial pattern. Appl Stat 1977;26(3):327. [doi: 10.2307/2346974]
19. Kim J, Yoo D, Hong K, Chun BC. Health behaviors and the risk of COVID-19 incidence: a Bayesian hierarchical spatial analysis. J Infect Public Health 2023 Feb;16(2):190-195. [doi: 10.1016/j.jiph.2022.12.013] [Medline: 36580692]
20. Mitchell A. The ESRI Guide to GIS Analysis: Spatial Measurements & Statistics: ESRI; 2005.
21. Wetherill GB, Brown DW. Statistical Process Control: Theory and Practice, 3rd edition: Chapman & Hall; 1991. [doi: 10.1007/978-1-4899-2949-5]
22. Kleiber C, Zeileis A. Applied Econometrics With R: Springer; 2008. [doi: 10.1007/978-0-387-77318-6]
23. Hosmer DWJ, Lemeshow S, Sturdivant RX. Model-building strategies and methods for logistic regression. In: Applied Logistic Regression, 3rd edition: John Wiley & Sons, Inc; 2013:89-151. [doi: 10.1002/9781118548387]
24. Akinwande MO, Dikko HG, Samson A. Variance inflation factor: as a condition for the inclusion of suppressor variable(s) in regression analysis. Open J Stat 2015;05(7):754-767. [doi: 10.4236/ojs.2015.57075]
25. Griffith DA. Assessing spatial dependence in count data: Winsorized and spatial filter specification alternatives to the auto‐Poisson model. Geogr Anal 2006 Apr;38(2):160-179. [doi: 10.1111/j.0016-7363.2006.00681.x]
26. Besag J, York J, Mollie A. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math 1991 Mar;43(1):1-20. [doi: 10.1007/BF00116466]
27. Haining RP, Li G. Modelling Spatial and Spatial-Temporal Data: A Bayesian Approach: CRC Press; 2020.
28. Best N, Richardson S, Thomson A. A comparison of Bayesian spatial models for disease mapping. Stat Methods Med Res 2005 Feb;14(1):35-59. [doi: 10.1191/0962280205sm388oa] [Medline: 15690999]

XSL•FO
RenderX

29. Wakefield J. Disease mapping and spatial regression with count data. Biostatistics 2007 Apr;8(2):158-183. [doi: 10.1093/biostatistics/kxl008] [Medline: 16809429]
30. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J Royal Stat Soc Ser B 2009 Apr 1;71(2):319-392. [doi: 10.1111/j.1467-9868.2008.00700.x]
31. Simpson D, Rue H, Riebler A, Martins TG, Sørbye SH. Penalising model component complexity: a principled, practical approach to constructing priors. Statist Sci 2017;32(1):1-28. [doi: 10.1214/16-STS576]
32. Bivand R. R packages for analyzing spatial data: a comparative case study with areal data. Geogr Anal 2022 Jul;54(3):488-518. [doi: 10.1111/gean.12319]
33. Tennekes M. tmap: thematic maps in R. J Stat Softw 2018;84:1-39. [doi: 10.18637/jss.v084.i06]
34. Venables WN, Ripley BD. Modern Applied Statistics With S, 4th edition: Springer; 2002.
35. National Spatial Geographic Information Service [Webpage in Korean]. Ministry of Data and Statistics. URL: https://sgis.kostat.go.kr/view/index [accessed 2025-03-24]
36. Jardim Santos C, Gomes B, Ribeiro AI. Mapping geographical patterns and high rate areas for sexually transmitted infections in Portugal: a retrospective study based on the National Epidemiological Surveillance System. Sex Transm Dis 2020 Apr;47(4):261-268. [doi: 10.1097/OLQ.0000000000001122] [Medline: 31876867]
37. Patterson-Lomba O, Goldstein E, Gómez-Liévano A, Castillo-Chavez C, Towers S. Per capita incidence of sexually transmitted infections increases systematically with urban population size: a cross-sectional study. Sex Transm Infect 2015 Dec;91(8):610-614. [doi: 10.1136/sextrans-2014-051932] [Medline: 25921021]
38. Shu JT, Jiang T, Li CY, et al. Epidemiological characteristics and socioeconomic factors of sexually transmitted infections in China during 2002-2021. BMC Public Health 2025 Jul 2;25(1):2194. [doi: 10.1186/s12889-025-22515-8] [Medline: 40604694]
39. Theall KP, Scribner R, Cohen D, et al. The neighborhood alcohol environment and alcohol-related morbidity. Alcohol Alcohol 2009;44(5):491-499. [doi: 10.1093/alcalc/agp042] [Medline: 19671569]
40. Jennings JM, Polk S, Fichtenberg C, Chung SE, Ellen JM. Social place as a location of potential core transmitters-implications for the targeted control of sexually transmitted disease transmission in urban areas. Ann Epidemiol 2015 Nov;25(11):861-867. [doi: 10.1016/j.annepidem.2015.07.008] [Medline: 26371418]
41. Stopka TJ, Brinkley-Rubinstein L, Johnson K, et al. HIV clustering in Mississippi: spatial epidemiological study to inform implementation science in the deep south. JMIR Public Health Surveill 2018 Apr 3;4(2):e35. [doi: 10.2196/publichealth.8773] [Medline: 29615383]
42. Lindberg LD, Singh S. Sexual behavior of single adult American women. Perspect Sexual Reproductive 2008 Mar;40(1):27-33. [doi: 10.1363/4002708]
43. Habel LA, Van Den Eeden SK, Sherman KJ, McKnight B, Stergachis A, Daling JR. Risk factors for incident and recurrent condylomata acuminata among women. A population-based study. Sex Transm Dis 1998 Jul;25(6):285-292. [doi: 10.1097/00007435-199807000-00003] [Medline: 9662761]
44. Tamer E, Çakmak SK, İlhan MN, Artüz F. Demographic characteristics and risk factors in Turkish patients with anogenital warts. J Infect Public Health 2016;9(5):661-666. [doi: 10.1016/j.jiph.2015.12.009] [Medline: 26776703]
45. Coyne KM, Banks A, Heggie C, et al. Sexual health of adults working in pornographic films. Int J STD AIDS 2009 Jul;20(7):508-509. [doi: 10.1258/ijsa.2008.008434] [Medline: 19541897]
46. Hill SC, King G, Smith A. Condom use and prevalence of sexually transmitted infection among performers in the adult entertainment industry. Int J STD AIDS 2009 Nov;20(11):809-810. [doi: 10.1258/ijsa.2009.009381] [Medline: 19854885]
47. Rowley J, Vander Hoorn S, Korenromp E, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. Bull World Health Organ 2019 Aug 1;97(8):548-562P. [doi: 10.2471/BLT.18.228486] [Medline: 31384073]
48. Lacey CJN, Woodhall SC, Wikstrom A, Ross J. 2012 European guideline for the management of anogenital warts. J Eur Acad Dermatol Venereol 2013 Mar;27(3):e263-e270. [doi: 10.1111/j.1468-3083.2012.04493.x] [Medline: 22409368]
49. Goyal A, Murray JM. Roadmap to control HBV and HDV epidemics in China. J Theor Biol 2017 Jun 21;423:41-52. [doi: 10.1016/j.jtbi.2017.04.007] [Medline: 28442239]
50. Jeynes C, Chung MC, Challenor R. Shame on you'--the psychosocial impact of genital warts. Int J STD AIDS 2009 Aug;20(8):557-560. [doi: 10.1258/ijsa.2008.008412] [Medline: 19625588]
51. National Health Insurance Service [Webpage in Korean]. URL: https://nhiss.nhis.or.kr/ [accessed 2025-03-24]

Abbreviations

aRR: adjusted relative risk
BYM: Besag–York–Mollié
CA: condyloma acuminata
CrI: credible interval
DIC: deviance information criterion
ICAR: intrinsic conditional autoregressive

XSL•FO
RenderX

KNHIS: Korean National Health Insurance Service
STI: sexually transmitted infection
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 30.04.25; peer-reviewed by B Ngobe, CH Chan, F Pabon-Rodriguez; revised version received 22.09.25; accepted 09.10.25; published 03.11.25.

Please cite as:
Jang J, Park S, Chun BC
Spatial Clusters of Condyloma Acuminata and the Regional Risk Factors in South Korea: Bayesian Spatial Regression Analysis
JMIR Public Health Surveill 2025;11:e76751
URL: https://publichealth.jmir.org/2025/1/e76751
doi:10.2196/76751

© Joonsu Jang, Seyul Park, Byung Chul Chun. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 3.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Hypertension Prevalence Among People Lifted Out of Poverty in China in 2018-2023: Retrospective Spatiotemporal Analysis

Ying Zhang[1,2*], MPH; Dong Xia[2*], PhD; Zhiyu Lv[3], PhD; Chennan Wu[2], PhD; Jiapeng Chen[3], PhD; Lulu Zhang[1,2], PhD

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai, China

[2]Department of Military Health Management, College of Health Service, Naval Medical University, No. 41, Zhengtong Road, Wujiaochang Street, Yangpu District, Shanghai, China

[3]China Population and Development Research Center, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Lulu Zhang, PhD
School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai, China

Abstract

Background: Hypertension is a significant global public health concern, with particular concern in China due to its widespread prevalence. The spatial distribution of hypertension varies significantly, revealing important regional disparities that may impact public health strategies and interventions.

Objective: This study aimed to investigate the temporal trends and spatial distribution characteristics of hypertension prevalence among individuals lifted out of poverty in China, covering the period from 2018 to 2023.

Methods: We used data from the National Health Poverty Alleviation Dynamic Management System to analyze hypertension prevalence among people lifted out of poverty from 2018 to 2023. Long-term trends were assessed using the Joinpoint regression model. Spatial distribution characteristics were examined through global and local spatial autocorrelation, cluster and outlier analyses, and trend surface analyses. These methods provided insights into the spatial aggregation and variability of hypertension prevalence across different regions.

Results: From 2018 to 2023, the prevalence of hypertension among people lifted out of poverty in China increased from 2.71% to 7.29%. The highest rates were observed in the northeast, with Jilin Province ranking first for 3 consecutive years (2020 - 2023), reaching 27.19% in 2023. Linxi County in Inner Mongolia had prevalence rates exceeding 40% for 5 years (2019 - 2023), peaking at 47.99% in 2022. Among the 22 provinces containing poverty-stricken counties, 9 showed significant annual increases, with Guangxi having the highest annual percentage change at 27.9018% (95% CI 7.4095% - 52.3038%). Spatial analysis identified high-high clusters in northern provinces such as Hebei, Jilin, and Inner Mongolia, and low-low clusters in southwestern provinces such as Yunnan and Guizhou. Trend surface analysis revealed a distinct spatial gradient, with the northeast highest and the southwest lowest.

Conclusions: The study revealed a generally increasing trend in hypertension prevalence among people lifted out of poverty in China from 2018 to 2023. The highest prevalence rates were concentrated in northeastern poverty-alleviated counties, while southwestern counties exhibited the lowest prevalence rates.

(JMIR Public Health Surveill 2025;11:e66501) doi:10.2196/66501

KEYWORDS

hypertension; time series analysis; spatial analysis; trend surface analysis; public health; high blood pressure; blood pressure; China; Chinese; JoinPoint regression; temporal trends; spatial distribution; spatiotemporal analysis; global health; prevalence; disparities; poverty; non-communicable disease; epidemic; epidemiology

Introduction

Hypertension is a global public health issue that poses a significant threat to human health and contributes substantially to the disease burden worldwide [1-3]. According to a global pooled analysis, the prevalence of hypertension has steadily increased over the past three decades, especially in low- and middle-income countries, including China [4]. In 2019, China had one of the highest numbers of people with hypertension globally, with substantial disparities across sex, region, and socioeconomic groups [5]. According to a nationwide survey of 298,438 individuals across 31 provinces conducted from 2020 to 2022, the prevalence among residents aged 18 years or older was 31.6%, with males (36.8%) having a higher prevalence than females (26.3%), and rural areas (33.7%) having a higher

XSL•FO
RenderX

prevalence than urban areas (29.1%) [6]. Although prevention and control programs have expanded, the control rate of hypertension remains unsatisfactory in key populations [7,8]. While there is no cure for hypertension, medication can typically maintain blood pressure at a safe level in most patients, effectively reducing the prevalence and mortality rates of cardiovascular disease, stroke, and kidney disease [9,10]. However, structural disparities in health care access persist, particularly in under-resourced regions [11].

People lifted out of poverty are those who, through policy support and their own efforts, have succeeded in escaping poverty and reaching the minimum standard of living set by the state. According to China's current poverty standards, those lifted out of poverty must meet the following conditions: an average annual per capita household income of more than 2300 yuan (at constant 2010 prices) and the realization of the principle of "two no worries, three guarantees," which refers to not worrying about food or clothing, as well as guarantees for compulsory education, basic medical care, and housing security. Among this population, as an important target of precision poverty alleviation, the "established card households" are those poor families that have been included in the national poverty alleviation information system and dynamically managed throughout the poverty alleviation process. By the end of 2020, all 832 poverty-alleviated counties had met the criteria, resulting in a total of 98.99 million rural poor people being lifted out of poverty according to the current standards. This historic achievement significantly reduced absolute poverty in China and yielded profound social benefits. However, the health status of people in these counties still lags behind the national average. To address this, policies such as chronic disease subsidies, family doctor programs, and enhanced medical insurance have been implemented to promote equitable health care [12-14]. Nonetheless, research shows that individuals in poverty-alleviated counties remain disproportionately affected by chronic diseases, including hypertension [15].

Many scholars have studied the knowledge, treatment, and control of hypertension [16,17], but only a few have focused on hypertension prevalence among people lifted out of poverty. Despite the remarkable success of China's efforts to combat poverty, these individuals still face certain vulnerabilities, especially in terms of their generally poor health status. Due to long-term poverty, they have experienced a relative lack of health management and medical resources, which has led to a higher prevalence of some chronic diseases. In this context, our research focused on analyzing hypertension prevalence among people lifted out of poverty, specifically examining its characteristics across temporal and spatial dimensions. This study aimed to provide a scientific basis for developing more targeted health interventions to further improve the quality of life and health of those who have been lifted out of poverty.

Methods

Data Sources

Hypertension prevalence data for people lifted out of poverty in the 832 poverty-alleviated counties nationwide were obtained from the National Health Poverty Alleviation Dynamic Management System. This system provides comprehensive information on people lifted out of poverty, including details on literacy levels, health status, family types, and ethnicity. Specific data include the number of patients, the number of people receiving assistance, the number of times assistance was provided, and reimbursement ratios. National county-level vector maps were obtained from China's National Basic Geographic Information System.

Descriptive Statistics

We described the temporal trends and regional distribution characteristics of hypertension prevalence among people lifted out of poverty from 2018 to 2023. Using the data on the number of patients and the number of people registered in each county, we analyzed the overall temporal trends in hypertension prevalence among this population nationwide. The 832 poverty-alleviated counties were spread across 22 provinces, and we conducted a statistical analysis of hypertension prevalence in the population lifted out of poverty in each province. In addition to the statistical analysis, we also analyzed the spatial distribution characteristics of hypertension prevalence in each county. Hypertension prevalence was calculated using the following formula:

$$Pi=NiTi*100\%$$

$$Ppi=NpiTpi*100\%$$

$$Pci=NciTci*100\%$$

Where i ($2018\leq i\leq 2023$) denotes the six different years, *Pi* denotes the national prevalence of hypertension among people lifted out of poverty in each year, *Ni* denotes the number of people with hypertension among this population nationwide each year, and *Ti* denotes the total number of people lifted out of poverty nationwide each year. p($1\leq p\leq 22$) denotes the 22 provinces, *Ppi* denotes hypertension prevalence among people lifted out of poverty in each province each year, and *Npi* denotes the number of people with hypertension among people lifted out of poverty in each province each year. *Tpi* denotes the total number of people lifted out of poverty in each province each year. c($1\leq c\leq 832$) denotes the 832 counties, *Pci* denotes hypertension prevalence among people lifted out of poverty in each county each year, *Nci* denotes the number of patients with hypertension among those lifted out of poverty in each county each year, and *Tci* denotes the total number of people lifted out of poverty in each county each year.

Time Series Analysis

We used the Joinpoint regression model [18] to analyze temporal trends in hypertension prevalence among people lifted out of poverty across 22 provinces from 2018 to 2023. The model detects statistically significant trend change points (inflection points) through segmented regression and calculates the annual percentage change (APC) for each period to assess how the trend is changing. In contrast to the autoregressive integrated moving average (ARIMA) model and exponential smoothing, Joinpoint allows data to show different trends over time.

The Joinpoint regression model relies on several key assumptions. First, it assumes that the error terms are independent and identically distributed. To address potential autocorrelation in the data, we examined the residuals for any

XSL•FO
RenderX

signs of serial correlation and applied appropriate adjustments if necessary. Second, the model assumes homoscedasticity, meaning that the variance of the error terms is constant over time. In cases where heteroscedasticity was detected, we used weighted least squares to account for varying error variances. Finally, the model assumes that the trend changes are continuous at the joinpoints, ensuring smooth transitions between different linear segments. For our study, the log-transformed independent variable for prevalence was time (year), and the dependent variable was prevalence. We used the observation series (i_1, Ppi_1), ...(i_n, Ppi_n), ($i_1 \leq ... \leq i_n$), where n is the study period, and the regression equation for the log-linear model is:

$$E[Ppi|i]=e\beta 0+\beta 1i+\delta 1(i-\tau 1)++\ldots+\delta k(i-\tau k)+$$

where e is the base of the natural logarithm, k denotes the number of turning points, β_0 is the intercept parameter, β_1 is the regression coefficient for the trend, and δk denotes the regression coefficient for segment k. When $(i-\tau_k)>0$, then $(i-\tau_k)^+)=i-\tau_k$; otherwise, $(i-\tau_k)^+)=0$.

We used the APC as well as the 95% CI to describe the change in each line segment, the change between inflection points, and the change over the entire period. The formula used is as follows:

$$APC=(e\beta 1-1)\times 100\%$$

Spatial Analysis

In our study, we used both global and local spatial autocorrelation analyses to explore the spatial distribution characteristics of hypertension prevalence among people lifted out of poverty. For global spatial autocorrelation, we applied Moran I [19] to assess the spatial correlation between data points across the study area. Moran I ranges from −1 to 1. Positive values indicate positive autocorrelation (spatial clustering); negative values indicate negative autocorrelation (spatial dispersion); and a value of 0 indicates no spatial autocorrelation. The formula for calculating Moran I is as follows:

$$I=cW\cdot\sum k=1c\ \sum j=1c\ wkj(Pck-Pc-)(Pcj-Pc-)\sum k=1c\ (Pck-Pc-)2$$

where c is the sample size, that is, the number of counties in the study area, W is the sum of the spatial weight matrix, that is, $W=\sum k=1c\ \sum j=1c\ wkj$, Pc_k, and Pc_j are the prevalence rates of hypertension in the kth and jth counties, respectively, and Pc- is the mean of the prevalence rates of hypertension in counties that have escaped poverty. w_{kj} is an element in the spatial weight matrix, which denotes the spatial relationship between kth and jth counties (which can usually be a neighborhood weight or a distance weight).

In this study, we used the method of clustering and outlier analysis (Anselin Local Moran I) to examine the local spatial autocorrelation, which can effectively identify local autocorrelation patterns in spatial data and help reveal intrinsic patterns and anomalous features of spatial phenomena. Based on the calculated Local Moran I and the results of the significance tests, four types of spatial clusters and outliers, namely high-high, low-low, high-low, and low-high, were identified. The formula for clustering and outlier analysis is as follows:

$$Ik=(Pck-Pc-)m2\sum j=1c\ wkj(Pcj-Pc-)$$
$$m2=1c\sum k\ (Pck-Pc-)2$$

where I_k is the local Moran I value for the kth county and $m2$ is the variance of the prevalence rate across the 832 poverty-alleviated counties.

In spatial analysis, z values and P values are two key indicators used to assess statistical significance. The z value represents the standardized value of Moran I and measures the deviation of the observed value from the expected value. A larger (positive) z value indicates that clusters of higher values are more significant, while a smaller (negative) z value indicates that clusters of lower values are more significant. The P values indicate the probability of the observed outcome occurring by chance, reflecting the statistical significance of the result.

Trend Surface Analysis

Trend surface analysis [20] is a spatial data analysis technique used to describe the overall trend in spatial data by fitting a mathematical model. We used a second-order polynomial trend surface model to analyze the spatial distribution of hypertension prevalence among people lifted out of poverty. The coefficient of determination (R^2) of the model was 0.569, and the root-mean-square error was 3.541, indicating that the model could reveal the spatial trend better, but there is still room for improvement. Comparing with first-order and higher-order polynomials, the first-order model was unable to capture nonlinear variations, and the higher-order model was prone to overfitting and did not significantly improve the goodness-of-fit. We used the geographic coordinates (longitude and latitude) of each province as the independent variables (x-axis and y-axis) and the prevalence rate as the dependent variable (z-axis). In this setup, the x-axis increases from left to right, representing longitude from low to high (west to east); the y-axis increases from bottom to top, representing latitude from low to high (south to north); and the z-axis increases from low to high, indicating higher prevalence rates. In the 3D spatial trend surface analysis graphs, the xy plane indicates the distribution of people lifted out of poverty across different latitude and longitude locations. The xz plane illustrates the trend of hypertension prevalence relative to longitude, and the yz plane indicates the trend of hypertension prevalence relative to latitude.

Ethical Considerations

The original data were collected through a national platform established and managed by the National Health Commission of China. In accordance with national regulations, data collection activities conducted under the Commission's authority are subject to institutional review board (IRB) approval. This study involved a secondary analysis of deidentified, aggregate county-level data, with no individual identifiers or biospecimens, and therefore did not require additional IRB review. According to International and National Ethical Guidelines (eg, 45 CFR 46.104[d] [4] and Article 23 of China's Ethical Review Guidelines), such secondary analyses of anonymized data are exempt from further IRB approval. The authors have obtained

XSL•FO
RenderX

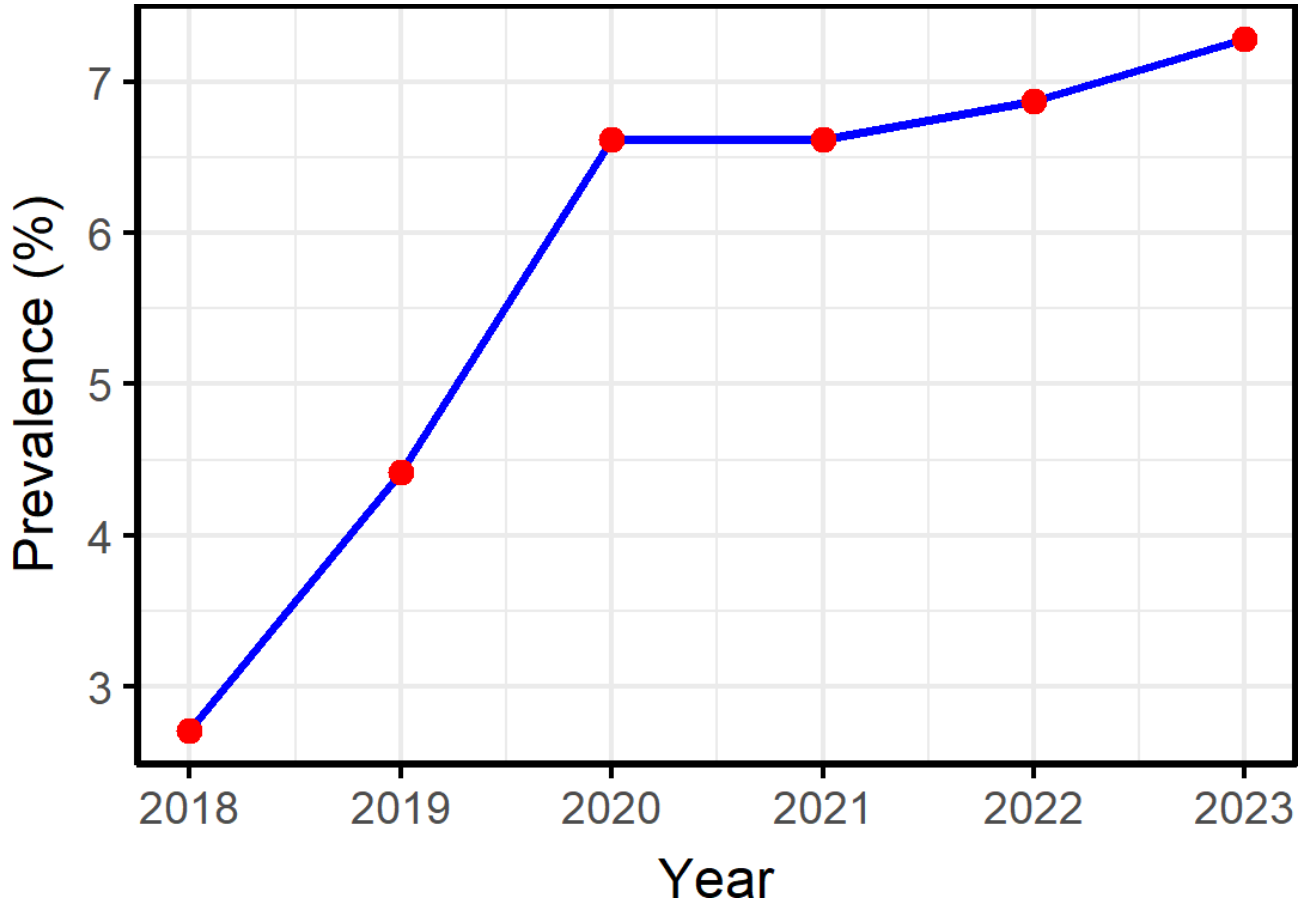
formal permission to access and use the data for academic research purposes.

Results

Epidemiological Characteristics

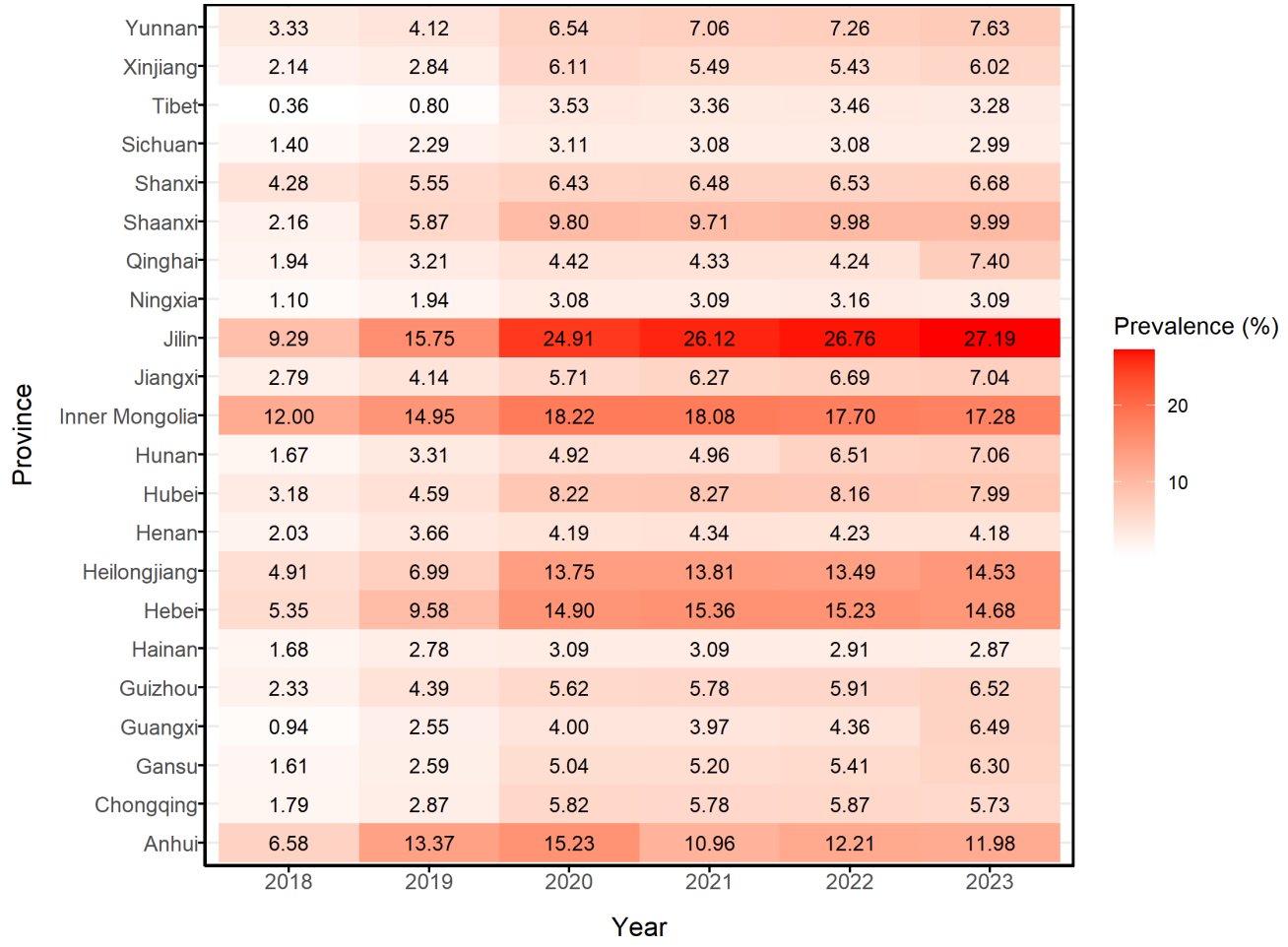
Between 2018 and 2023, hypertension prevalence among people lifted out of poverty showed an overall increasing trend. As shown in Figure 1, the most significant rise occurred between 2018 and 2020, with the prevalence increasing from approximately 2.71% in 2018 to approximately 6.62% in 2020. There was a slight decrease in prevalence from 2020 to 2021, but the overall trend remained relatively stable. From 2021 onward, the prevalence rate began to rise again gradually, reaching a 6-year peak of approximately 7.29% in 2023. We analyzed the annual trend of hypertension prevalence among people lifted out of poverty across the 22 provinces containing poverty-alleviated counties. As shown in Figure 2, hypertension prevalence varied across provinces from 2018 to 2023. Four provinces—Jilin, Heilongjiang, Inner Mongolia Autonomous Region, and Hebei—exhibited the highest prevalence rates. Notably, Jilin Province had the highest prevalence across all 6 years, peaking at approximately 27.19% in 2023. In addition, we analyzed hypertension prevalence in each of the 832 poverty-alleviated counties. As shown in Figure 3, most counties with high hypertension prevalence among people lifted out of poverty were located in the northeast, followed by the northwest and southwest regions, with consistent regional distribution over the 6 years. Analysis revealed that in 2023, the 17 counties with the highest prevalence were distributed across four regions: Jilin Province (n=5), Heilongjiang Province (n=2), Inner Mongolia Autonomous Region (n=2), and Hebei Province (n=8), as illustrated in Figure 4. Notably, Linxi County in the Inner Mongolia Autonomous Region had a prevalence rate exceeding 40% from 2019 to 2023, peaking at 47.99% in 2022. In parallel, Longjing City in Jilin Province showed a continuous increase in prevalence, with an average annual increase of about 3%.

Figure 1. Temporal trends in hypertension prevalence among people lifted out of poverty in China (2018 - 2023).



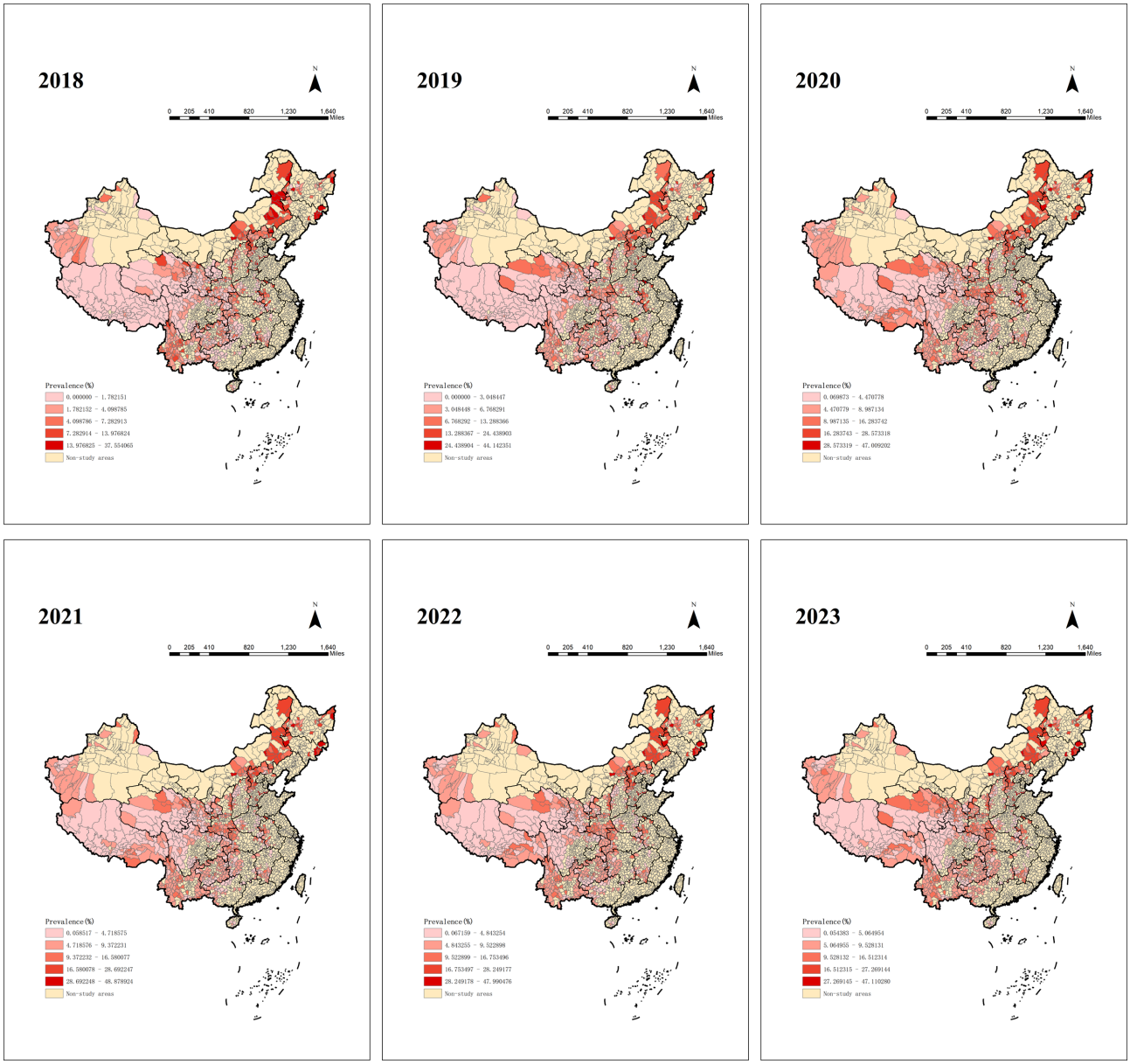
XSL•FO
RenderX

Figure 2. Heatmap of hypertension prevalence among the people lifted out of poverty in 22 provinces of China (2018 - 2023).



XSL•FO
RenderX

Figure 3. County-level hypertension prevalence among poverty-alleviated populations in China (2018 - 2023).



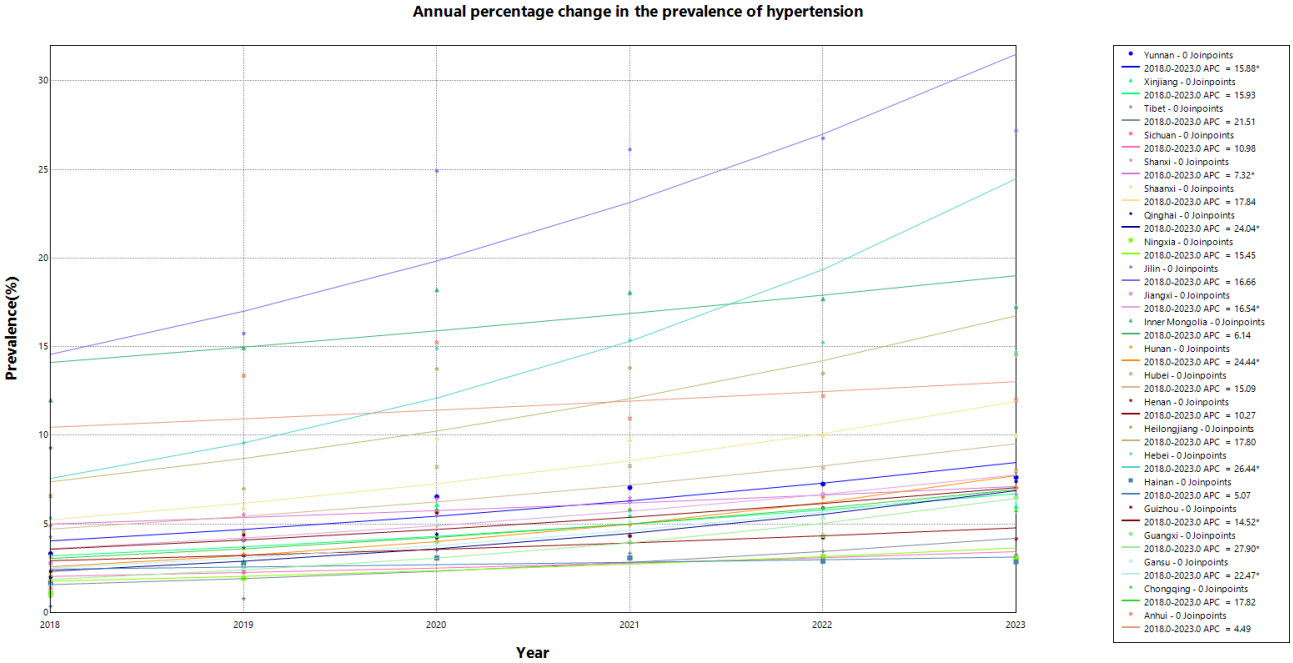
XSL•FO
RenderX

Figure 4. Geographic distribution of the 17 counties with the highest hypertension prevalence among people lifted out of poverty in China (2023).

Temporal Distribution

We analyzed the time trend of hypertension prevalence among people lifted out of poverty across 22 provinces, as shown in Figure 5. Statistically significant APC values were found in only 9 provinces: Yunnan, Shanxi, Qinghai, Jiangxi, Hunan, Hebei, Guizhou, Guangxi, and Gansu. In these provinces, hypertension prevalence showed a consistent year-by-year increase. Guangxi Province had the highest APC value at 27.9018% (95% CI 7.4095% - 52.3038%), followed by Hebei Province with an APC value of 26.4411% (95% CI 0.5839% - 58.9456%), and Hunan Province with the third highest APC value of 24.4372% (95% CI 9.1285% - 41.8933%), as shown in Table 1.

Figure 5. Annual changes in hypertension prevalence among people lifted out of poverty in 22 provinces in China (2018-2023).



XSL•FO
RenderX

Table . Annual trends in hypertension prevalence among people lifted out of poverty in 22 provinces in China (2018-2023).

Cohort	APC[a] (95% CI)	*t* test (*df*)	*P* value
Yunnan	15.8795[b] (3.5653 to 29.6579)	3.6422 (5)	.02
Xinjiang	15.9303 (−5.1830 to 41.7450)	2.0414 (5)	.11
Tibet	21.5128 (−20.1508 to 84.9156)	1.2884 (5)	.27
Sichuan	10.9760 (−3.9215 to 28.1833)	2.0059 (5)	.12
Shanxi	7.3244[b] (0.2287 to 14.9223)	2.8692 (5)	.05
Shaanxi	17.8446 (−7.3675 to 49.9187)	1.8938 (5)	.13
Qinghai	24.0424[b] (9.1032 to 41.0272)	4.6614 (5)	.01
Ningxia	15.4480 (−3.4627 to 38.0631)	2.2295 (5)	.09
Jilin	16.6599 (−1.0574 to 37.5497)	2.5973 (5)	.06
Jiangxi	16.5443[b] (5.2907 to 29.0006)	4.1861 (5)	.01
Inner Mongolia	6.1368 (−2.2056 to 15.1910)	2.0200 (5)	.11
Hunan	24.4372[b] (9.1285 to 41.8933)	4.6240 (5)	.01
Hubei	15.0930 (−4.3224 to 38.4482)	2.1124 (5)	.10
Henan	10.2661 (−4.3873 to 27.1653)	1.9029 (5)	.13
Heilongjiang	17.7968 (−2.8046 to 42.7649)	2.3656 (5)	.08
Hebei	26.4411[b] (0.5839 to 58.9456)	2.8471 (5)	.05
Hainan	5.0657 (−6.8447 to 18.4990)	1.1403 (5)	.32
Guizhou	14.5227[b] (0.2244 to 30.8609)	2.8231 (5)	.05
Guangxi	27.9018[b] (7.4095 to 52.3038)	3.9130 (5)	.02
Gansu	22.4657[b] (2.7094 to 46.0222)	3.1983 (5)	.03
Chongqing	17.8208 (−5.5987 to 47.0505)	2.0546 (5)	.11
Anhui	4.4920 (−13.2676 to 25.8881)	0.6549 (5)	.55

[a]APC: annual percentage change.

[b]Indicates that the annual percent change is significantly different from zero at the α=.05 level.

Spatial Distribution

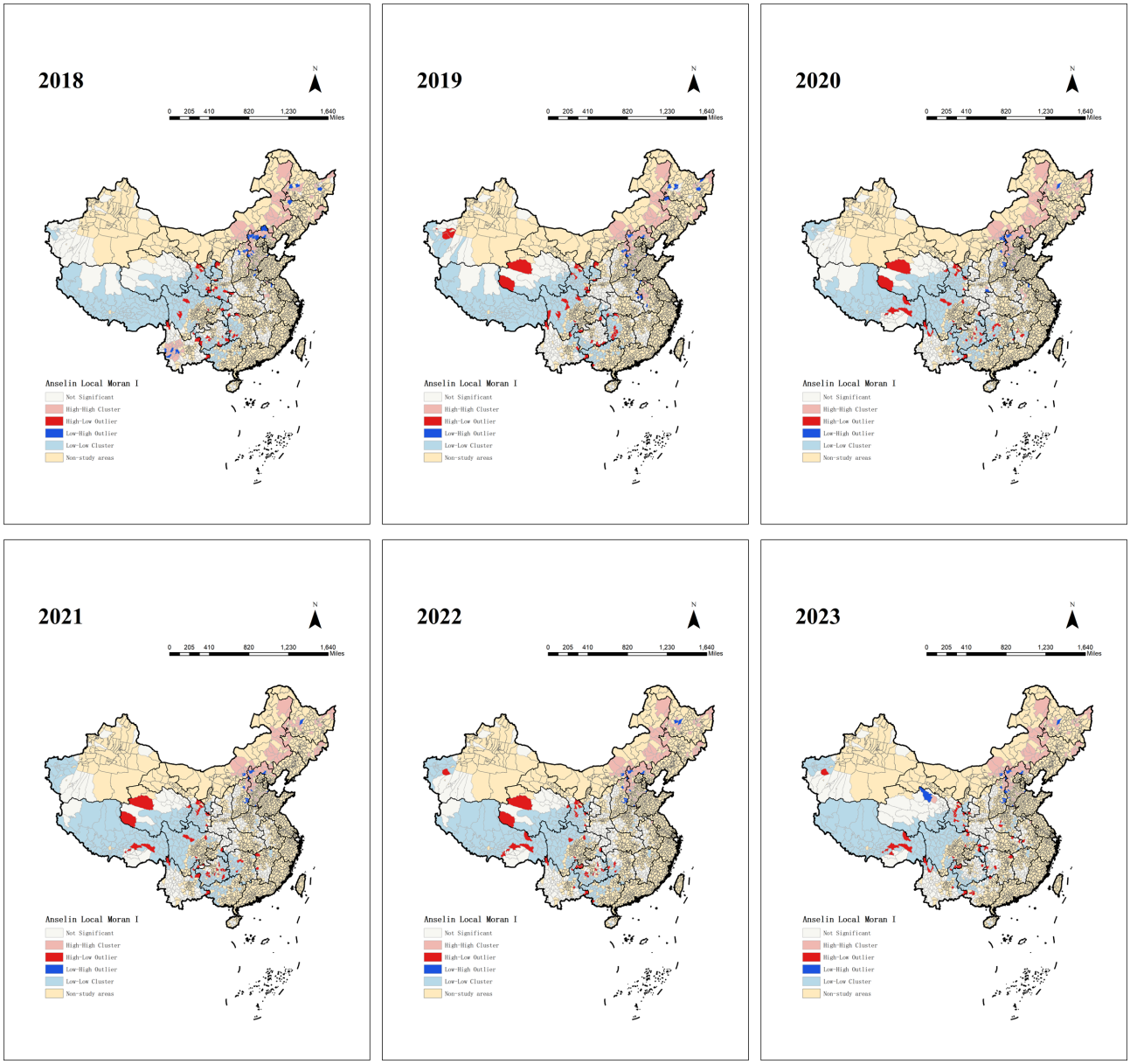
As shown in Table 2, the global spatial autocorrelation analysis revealed that from 2018 to 2023, both Moran I and *z* values were positive. This indicates a significant positive spatial autocorrelation in the distribution of hypertension prevalence among people lifted out of poverty at the county level, reflecting notable spatial clustering. Except for 2020, the variance in the remaining years was 0.000073. This small variance suggests that Moran I showed minimal volatility, indicating stable spatial autocorrelation and limited spatial change trends. As shown in Figure 6, we performed clustering and outlier analysis on hypertension prevalence among people lifted out of poverty in the 832 poverty-alleviated counties. The results showed that the spatial distribution of hypertension prevalence from 2018 to 2023 presents four types of clustering. The high-high clustered areas were mainly distributed in Hebei, Shanxi, Inner Mongolia Autonomous Region, Jilin, Heilongjiang, and Anhui Provinces, where hypertension prevalence was significantly higher than that of the neighboring regions and showed persistent clustering, which may reflect a heavier disease burden in those areas. The identification of high-high aggregation areas is crucial for developing targeted disease prevention and control strategies; therefore, we further identified a county-level list of high-high aggregation areas with high prevalence of hypertension in 2018 - 2023 (refer to Multimedia Appendix 1 for details). The low-low aggregation areas were concentrated in Hunan, Guangxi Zhuang Autonomous Region, Chongqing, Sichuan, Guizhou, Yunnan, Tibet Autonomous Region, Shaanxi, Gansu, Qinghai, Ningxia Hui Autonomous Region, and Xinjiang Uygur Autonomous Region, where hypertension prevalence was significantly lower than that of neighboring regions. High-low aggregation areas were mainly distributed in Hunan, Sichuan, Yunnan, Guizhou, and Gansu Provinces, where hypertension prevalence was relatively high, while neighboring regions had lower rates, possibly reflecting localized disease hotspots. Low-high clusters were mainly located in Hebei, Shanxi, and Heilongjiang Provinces, where hypertension prevalence was lower in these regions but higher in neighboring areas, possibly representing a transitional area of disease spread.

XSL•FO
RenderX

Table . Global spatial autocorrelation analysis of hypertension prevalence among people lifted out of poverty in China (2018-2023).

Year	Moran I	Expected I	Variance	Z value	*P* value	Result
2018	0.256240	−0.001215	0.000073	30.189126	<.001	Clustered
2019	0.272386	−0.001215	0.000073	31.981011	<.001	Clustered
2020	0.319228	−0.001215	0.000074	37.365603	<.001	Clustered
2021	0.329789	−0.001215	0.000073	38.626703	<.001	Clustered
2022	0.315589	−0.001215	0.000073	36.961519	<.001	Clustered
2023	0.282911	−0.001215	0.000073	33.155462	<.001	Clustered

Figure 6. Local spatial clustering of hypertension prevalence among poverty-alleviated counties in China (2018-2023).



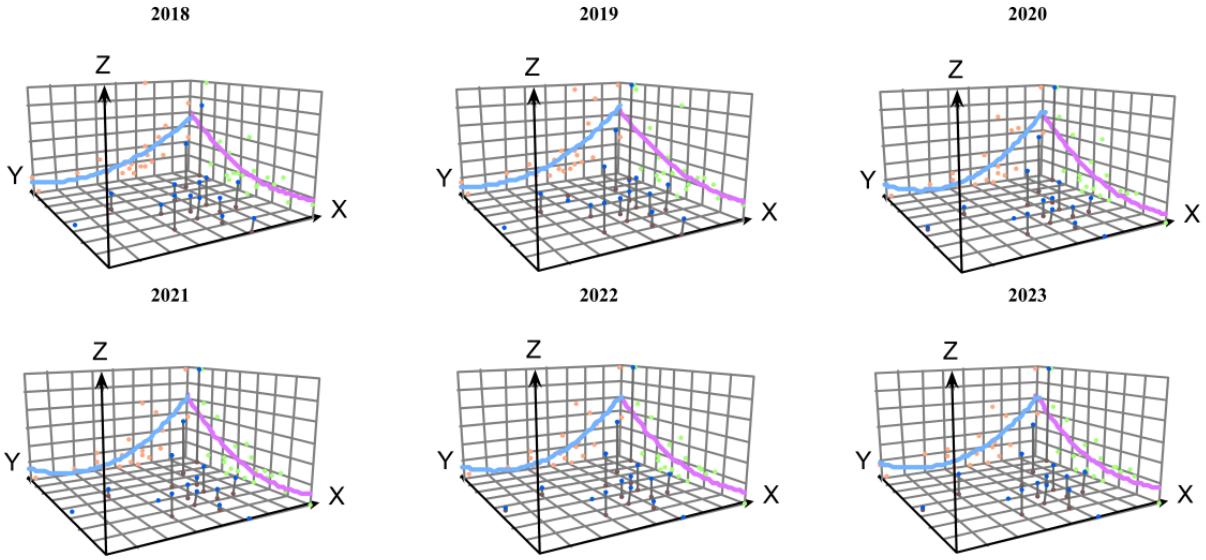
Trend Surface Analysis

As shown in Figure 7, there was a clear spatial heterogeneity in hypertension prevalence among people lifted out of poverty at the provincial level. From 2018 to 2023, hypertension prevalence initially decreased with increasing longitude but then increased again at a faster rate. This indicates that the eastern part of the country generally had a higher prevalence than the western part, with the prevalence being lowest in the central and western regions. In 2018, 2021, and 2023, hypertension prevalence among people lifted out of poverty first decreased and then increased with increasing latitude. In contrast, in the remaining years, the prevalence increased with latitude. Overall, the northern regions exhibited significantly higher prevalence rates than the southern, with the southern and south-central provinces having the lowest prevalence. Therefore,

XSL•FO
RenderX

we conclude that the northeastern region generally has the highest prevalence of hypertension among people lifted out of poverty.

Figure 7. Trend surface analysis of hypertension prevalence among people lifted out of poverty in 22 provinces in China (2018 - 2023).



Discussion

Principal Findings

From 1990 to 2019, the number of people with hypertension doubled globally [4]. In our study, hypertension prevalence among people lifted out of poverty in China increased consistently from 2018 to 2023, with a rise of approximately 4.59% in 2023 compared to the level in 2018. The APC values for all provinces were also positive, indicating a general upward trend. This increase in hypertension prevalence in China is likely attributable to rising life expectancy and lifestyle changes [21]. There are four potential factors contributing to the overall increase in hypertension prevalence. First, the Scientific Research Report on Dietary Guidelines for Chinese Residents, issued in 2021, highlights changes in dietary structure. Specifically, there has been an increase in the consumption of high-calorie, high-fat, high-sugar, and high-salt foods. In addition, the rates of overweight and obesity in rural areas have risen more significantly than in urban areas. Obesity is a major risk factor for hypertension [22]. Second, inadequate sleep and poor sleep quality are associated with an increased risk of hypertension [23]. Insufficient sleep can lead to cognitive decline and poor mental health, while high-pressure environments, mental overstimulation, and anxiety can trigger hypertension [24,25]. Third, population aging significantly contributes to the rising prevalence of hypertension. As Chinese society ages, the proportion of older individuals is increasing, and this age group is particularly susceptible to chronic diseases, such as hypertension [26]. Fourth, the increased prevalence of hypertension is not only closely related to environmental factors and lifestyle but is also significantly influenced by hereditary factors. Studies have shown that individuals with a family history of hypertension have a significantly increased risk of developing the condition. This is because certain genetic variants and genes can affect blood pressure regulatory mechanisms, making individuals with genetic susceptibility more prone to developing hypertension [27,28].

According to the spatial autocorrelation analysis, the high-high aggregation area is mainly distributed in the three northeastern provinces and northern China. The low-high aggregation area is distributed at the junction of northern China and northeastern China, and the high-low aggregation and low-low aggregation areas are mainly located in the west. Trend surface analysis shows that hypertension prevalence is higher in the east than in the west, and in the north than in the south. As early as 1994, a study found that hypertension prevalence was higher in northern China than in southern regions [29].

The spatial heterogeneity in hypertension prevalence among people lifted out of poverty may be explained by the following three factors: First, some studies have shown that high-altitude areas may experience abnormal blood pressure regulation due to hypoxia, which can increase the risk of hypertension [30-32]. The lower prevalence of hypertension in Tibet and Qinghai Provinces may be because residents have been exposed to low oxygen levels from birth and have developed adaptations to better cope with high-altitude conditions [33]. In addition, the diversity of physiological changes may result in different adaptations among high-altitude populations. For instance, research has shown that Tibetan populations exhibit better adaptation than non-Tibetan populations due to higher capillary density and greater oxygen delivery [34]. Experimental studies have also indicated that in Tibetan populations, hypertension prevalence increases by 1.2% for every 100-meter increase in altitude, a phenomenon not observed in all high-altitude populations [35]. In non-Tibetan populations, there is no significant correlation between increased altitude and changes in blood pressure [36]. Second, temperature changes can also impact blood pressure. A study conducted in northwest China found a significant negative correlation between outdoor temperature and blood pressure, with blood pressure

XSL•FO
RenderX

demonstrating notable seasonal variations [37]. This seasonal variation in blood pressure is likely due to changes in ambient temperature, as another study showed that the seasonal variation disappeared after adjusting for temperature [38]. Larger differences in blood pressure between colder and warmer months may increase hypertension prevalence, potentially elevating cardiovascular risk, especially among individuals at high cardiovascular risk and older individuals [39]. Third, dietary differences between the northern and southern regions of China also play a significant role. Regular consumption of high-sodium diets, energy-dense foods, fats, refined carbohydrates, added sugars, and low amounts of fruits and vegetables increases the risk of hypertension and cardiovascular disease. According to the Chinese Guidelines for Adults with Hypertension, issued in 2023, the diet in northeastern China is characterized by stews with a heavy taste, while in central China, food processing methods, such as pickling and curing, are popular, also resulting in heavily seasoned dishes. Meanwhile, southwestern China is known for its preference for spicy food. These guidelines emphasize the need to control oil and salt use across all regions. The Dietary Approaches to Stop Hypertension (DASH) diet was introduced in the United States in 1997 as a dietary pattern to prevent and manage hypertension. This diet focuses on reducing fats and oils while ensuring adequate intake of vegetables, fruits, and low-fat dairy products to maintain optimal levels of potassium, magnesium, and calcium, all of which significantly reduce blood pressure [40-42]. In addition, the Chinese heart-healthy diet and its Sichuan cuisine version, tailored for the Chinese population, have shown significant positive effects in reducing the risk of developing hypertension [43,44].

High-high aggregation areas should be used as key intervention areas to strengthen screening and early intervention, optimize medical resource allocation, and promote low-salt and low-fat diets and moderate exercise through health education. Although the current burden of disease in low-low aggregation areas is low, continuous monitoring is needed. High-low aggregation areas may reflect localized disease hotspots, requiring special attention to the distribution of health care resources, increased investment in primary care, and the development of personalized prevention and control strategies. Low-high aggregation areas may be transitional areas for the spread of the disease, and monitoring and regional joint prevention and control should be strengthened to prevent the spread of hypertension from areas of high prevalence to areas of low prevalence.

Uncontrolled hypertension can result in severe health issues, including stroke, cardiovascular diseases, and kidney diseases, leading to unnecessary human burden and premature death. Despite significant investments by the state in poverty alleviation and rural revitalization [45], the health service system remains inadequate. The prevention and control of chronic diseases, such as hypertension, still face considerable challenges due to limited medical resources and a lack of widespread health education [46]. Both national and local governments should enhance interregional collaborative efforts to address the regional differences and clustering patterns in hypertension prevalence [47]. Establishing cross-regional collaboration mechanisms for hypertension prevention and control, information sharing, resource complementation, and joint efforts can enhance the overall effectiveness of these initiatives. In addition, increased investment in medical resources in impoverished areas has optimized the allocation of these resources. The capacity to prevent and control hypertension in these regions has been strengthened through the development of primary health care institutions, training of primary health care personnel, and the promotion of telemedicine. Nurses can make significant contributions to national hypertension control efforts by leveraging evidence-based practices for the prevention, diagnosis, treatment, and management of hypertension [48]. Furthermore, strengthening health education and public awareness initiatives is crucial for promoting healthy lifestyles and dietary habits. Raising public awareness about hypertension and its prevention involves disseminating knowledge through health promotion activities and educational interventions. Providing targeted education to patients with hypertension can encourage the adoption of healthy behaviors and improve overall health outcomes [49].

Limitations

There are some limitations to this study. First, using the total recorded population in each county that has escaped poverty as the county's total population may introduce statistical errors, as this figure may include mobile populations. Second, the study did not account for seasonal changes, which affect both temperature and dietary patterns and can influence hypertension prevalence. Finally, while this study analyzed the spatial and temporal prevalence of hypertension, further research is needed to examine the underlying factors affecting hypertension prevalence to develop more targeted and effective prevention guidelines.

Conclusions

In this study, we determined the temporal trends and spatial distribution of hypertension prevalence among Chinese individuals lifted out of poverty from 2018 to 2023. Our findings reveal an overall increasing trend in hypertension prevalence year by year. High-prevalence areas were notably concentrated in the northeastern poverty-alleviated counties, while low-prevalence areas were found primarily in the southwestern counties.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (grant 72474220).

XSL•FO
RenderX

Data Availability

The data used in this study were obtained from the National Health Poverty Alleviation Dynamic Management System. Access to this database requires prior application and approval. Due to data use agreements, the datasets are not publicly available. However, the datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

YZ, DX, and LZ contributed to the conceptualization, formal analysis, and methodology; drafted the original manuscript; and reviewed and edited the final version. YZ, DX, ZL, and JC were responsible for data curation, software development, and formal analysis. DX, CW, and LZ contributed to the conceptualization, supervision, and project administration.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of counties with high concentrations of hypertension prevalence (2018-2023).
[DOCX File, 25 KB - publichealth_v11i1e66501_app1.docx]

References

1. Lu J, Lu Y, Wang X, et al. Prevalence, awareness, treatment, and control of hypertension in China: data from 1·7 million adults in a population-based screening study (China PEACE Million Persons Project). Lancet 2017 Dec;390(10112):2549-2558. [doi: 10.1016/S0140-6736(17)32478-9]
2. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on clinical practice guidelines. J Am Coll Cardiol 2018 May 15;71(19):e127-e248. [Medline: 29146535]
3. Hu T, Yang X, Du Y, et al. Trends in the global, regional, and national burden of cardiovascular diseases attributed to high systolic blood pressure from 1990 to 2021 and projections to 2045: a systematic analysis based on GBD 2021 data. BMC Cardiovasc Disord 2025;25(1):390. [doi: 10.1186/s12872-025-04807-4]
4. Cislaghi C, Negri E, La Vecchia C, Levi F. The application of trend surface models to the analysis of time factors in Swiss cancer mortality. Soz Praventivmed 1988;33(7):359-373. [doi: 10.1007/BF02084305] [Medline: 3223104]
5. Cheng H, Gu Y, Ma X, Tang H, Liu X. Urban-rural disparities in hypertension prevalence, awareness, treatment, and control among Chinese middle-aged and older adults from 2011 to 2015: a repeated cross-sectional study. BMC Cardiovasc Disord 2022 Jul 17;22(1):319. [doi: 10.1186/s12872-022-02769-5] [Medline: 35843959]
6. Liu M, He X, Yang X, Wang Z. Interpretation of Report on Cardiovascular Health and Diseases in China 2023. Chin J Cardiovasc Sci 2024;40(8):599-616. [doi: 10.13201/j.issn.1001-1439.2024.08.002]
7. Brettler JW, Arcila GPG, Aumala T, et al. Drivers and scorecards to improve hypertension control in primary care practice: recommendations from the HEARTS in the Americas Innovation Group. Lancet Reg Health Am 2022 May;9:ne. [doi: 10.1016/j.lana.2022.100223] [Medline: 35711685]
8. Jing R, Eggleston K, Lai X, Fang H. Family physician services and blood pressure control in China: A population-based retrospective cohort study. J Health Serv Res Policy 2023 Jul;28(3):163-170. [doi: 10.1177/13558196221149929] [Medline: 36630298]
9. Ma G, Luo A, Shen Z, Duan Y, Shi S, Zhong Z. The status of medication literacy and associated factors of hypertensive patients in China: a cross-sectional study. Intern Emerg Med 2020 Apr;15(3):409-419. [doi: 10.1007/s11739-019-02187-0] [Medline: 31650433]
10. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1736-1788. [doi: 10.1016/S0140-6736(18)32203-7] [Medline: 30496103]
11. Mouton CP, Hayden M, Southerland JH. Cardiovascular health disparities in underserved populations. Prim Care 2017 Mar;44(1):e37-e71. [doi: 10.1016/j.pop.2016.09.019] [Medline: 28164826]
12. Notice of the General Office of the State Council on Printing and Distributing the Medium and Long-Term Plan for the Prevention and Control of Chronic Diseases in China (2017-2025). General Office of the State Council. 2017. URL: https://www.gov.cn/zhengce/content/2017-02/14/content_5167886.htm [accessed 2025-06-20]
13. National health and family planning commission. Notice on Issuing Guiding Opinions on Promoting Family Doctor Contract Services. 2016. URL: https://www.gov.cn/xinwen/2016-06/06/content_5079984.htm [accessed 2025-06-20]
14. Peng Z, Zhu L. The impacts of health insurance on financial strain for people with chronic diseases. BMC Public Health 2021 May 29;21(1):1012. [doi: 10.1186/s12889-021-11075-2] [Medline: 34051775]

XSL•FO
RenderX

15. Wei J, Miao Y, Zhang J, Wu J. Spatial heterogeneity of blood pressure control and its influencing factors in elderly patients with essential hypertension: a small-scale spatial analysis. Health Place 2025 Mar;92:103428. [doi: 10.1016/j.healthplace.2025.103428] [Medline: 39983658]
16. Wang W, Zhang M, Xu CD, et al. Hypertension prevalence, awareness, treatment, and control and their associated socioeconomic factors in China: a spatial analysis of a National Representative Survey. Biomed Environ Sci 2021 Dec 20;34(12):937-951. [doi: 10.3967/bes2021.130] [Medline: 34981717]
17. Zhang M, Shi Y, Zhou B, et al. Prevalence, awareness, treatment, and control of hypertension in China, 2004-18: findings from six rounds of a national survey. BMJ 2023 Jan 11;380:e071952. [doi: 10.1136/bmj-2022-071952] [Medline: 36631148]
18. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000;19. [doi: 10.1002/(sici)1097-0258(20000215)19:3<335::aid-sim336>3.0.co;2-z] [Medline: 10649300]
19. Zhang Q, Ding H, Gao S, et al. Spatiotemporal changes in pulmonary tuberculosis incidence in a low-epidemic area of China in 2005-2020: retrospective spatiotemporal analysis. JMIR Public Health Surveill 2023 Mar 8;9:e42425. [doi: 10.2196/42425] [Medline: 36884278]
20. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021 Sep 11;398(10304):957-980. [doi: 10.1016/S0140-6736(21)01330-1] [Medline: 34450083]
21. Wang JG, Zhang W, Li Y, Liu L. Hypertension in China: epidemiology and treatment initiatives. Nat Rev Cardiol 2023 Aug;20(8):531-545. [doi: 10.1038/s41569-022-00829-z] [Medline: 36631532]
22. Seravalle G, Grassi G. Obesity and hypertension. Pharmacol Res 2017 Aug;122:1-7. [doi: 10.1016/j.phrs.2017.05.013] [Medline: 28532816]
23. Bock JM, Vungarala S, Covassin N, Somers VK. Sleep duration and hypertension: epidemiological evidence and underlying mechanisms. Am J Hypertens 2022 Jan 5;35(1):3-11. [doi: 10.1093/ajh/hpab146] [Medline: 34536276]
24. Boone JL. Stress and hypertension. Prim Care 1991 Sep;18(3):623-649. [doi: 10.1016/S0095-4543(21)00350-X] [Medline: 1946791]
25. Li Y, Sahakian BJ, Kang J, et al. The brain structure and genetic mechanisms underlying the nonlinear association between sleep duration, cognition and mental health. Nat Aging 2022 May;2(5):425-437. [doi: 10.1038/s43587-022-00210-2] [Medline: 37118065]
26. Cheng W, Du Y, Zhang Q, et al. Age-related changes in the risk of high blood pressure. Front Cardiovasc Med 2022;9:939103. [doi: 10.3389/fcvm.2022.939103] [Medline: 36187016]
27. Li AL, Fang X, Zhang YY, Peng Q, Yin XH. Familial aggregation and heritability of hypertension in Han population in Shanghai China: a case-control study. Clin Hypertens 2019;25:17. [doi: 10.1186/s40885-019-0122-z] [Medline: 31428454]
28. Tanira MOM, Al Balushi KA. Genetic variations related to hypertension: a review. J Hum Hypertens 2005 Jan;19(1):7-19. [doi: 10.1038/sj.jhh.1001780] [Medline: 15361889]
29. Huang Z, Wu X, Stamler J, et al. A north-south comparison of blood pressure and factors related to blood pressure in the People's Republic of China: a report from the PRC-USA Collaborative Study of Cardiovascular Epidemiology. J Hypertens 1994 Sep;12(9):1103-1112. [Medline: 7852756]
30. Müller J, Titz A, Schneider SR, et al. The effect of high altitude (2500 m) on incremental cycling exercise in patients with pulmonary arterial hypertension and chronic thromboembolic pulmonary hypertension: a randomised controlled cross-over trial. Eur Respir J 2024;63:2301001. [doi: 10.1183/13993003.01001-2023]
31. Zeng Y, Zhakeer G, Li B, et al. A novel clinical prediction scoring system of high-altitude pulmonary hypertension. Front Cardiovasc Med 2023;10:1290895. [doi: 10.3389/fcvm.2023.1290895] [Medline: 38259305]
32. Li S, Lyu Q, Shi Q, Bai Y, Ren X, Ma J. Intermittent short-duration reoxygenation relieves high-altitude pulmonary hypertension via NOX4/H2O2/PPAR-γ axis. Clin Sci (Lond) 2024 Feb 7;138(3):103-115. [doi: 10.1042/CS20231508] [Medline: 38237016]
33. Zhang Y, Yang Y, Wu X, et al. The association between altitude and the prevalence of hypertension among permanent highlanders. Hypertens Res 2022 Nov;45(11):1754-1762. [doi: 10.1038/s41440-022-00985-2]
34. Beall CM. Two routes to functional adaptation: Tibetan and Andean high-altitude natives. Proc Natl Acad Sci U S A 2007 May 15;104 Suppl 1(Suppl 1):8655-8660. [doi: 10.1073/pnas.0701985104] [Medline: 17494744]
35. Zhang X, Zhang Z, Ye R, Meng Q, Chen X. Prevalence of hypertension and its relationship with altitude in highland areas: a systematic review and meta-analysis. Hypertens Res 2022 Aug;45(8):1225-1239. [doi: 10.1038/s41440-022-00955-8] [Medline: 35705740]
36. Narvaez-Guerra O, Herrera-Enriquez K, Medina-Lezama J, Chirinos JA. Systemic hypertension at high altitude. Hypertension 2018 Sep;72(3):567-578. [doi: 10.1161/HYPERTENSIONAHA.118.11140] [Medline: 30354760]
37. Zheng S, Wang MZ, Cheng ZY, et al. Effects of outdoor temperature on blood pressure in a prospective cohort of Northwest China. Biomed Environ Sci 2021 Feb 20;34(2):89-100. [doi: 10.3967/bes2021.014] [Medline: 33685567]
38. Su D, Du H, Zhang X, et al. Season and outdoor temperature in relation to detection and control of hypertension in a large rural Chinese population. Int J Epidemiol 2014 Dec;43(6):1835-1845. [doi: 10.1093/ije/dyu158] [Medline: 25135908]

XSL•FO
RenderX

39. Wang H, Chen L, Shen D, et al. Association between frequency of spicy food consumption and hypertension: a cross-sectional study in Zhejiang Province, China. Nutr Metab (Lond) 2021 Jul 6;18(1):70. [doi: 10.1186/s12986-021-00588-7] [Medline: 34229688]
40. Ozemek C, Laddu DR, Arena R, Lavie CJ. The role of diet for prevention and management of hypertension. Curr Opin Cardiol 2018 Jul;33(4):388-393. [doi: 10.1097/HCO.0000000000000532] [Medline: 29771736]
41. Filippou CD, Tsioufis CP, Thomopoulos CG, et al. Dietary approaches to stop hypertension (DASH) diet and blood pressure reduction in adults with and without hypertension: a systematic review and meta-analysis of randomized controlled trials. Adv Nutr 2020 Sep 1;11(5):1150-1160. [doi: 10.1093/advances/nmaa041] [Medline: 32330233]
42. Theodoridis X, Chourdakis M, Chrysoula L, et al. Adherence to the DASH diet and risk of hypertension: a systematic review and meta-analysis. Nutrients 2023 Jul 24;15(14):3261. [doi: 10.3390/nu15143261] [Medline: 37513679]
43. Xie W, Wang Y, Sun J, et al. Protocol of a multicenter, single-blind, randomised, parallel controlled feeding trial evaluating the effect of a Chinese Healthy Heart (CHH) diet in lowering blood pressure and other cardiovascular risk factors. BMJ Open 2020 Aug 20;10(8):e036394. [doi: 10.1136/bmjopen-2019-036394] [Medline: 32819944]
44. Su D, Chen H, Guo Y, et al. Effects of the Chinese heart-healthy diet (Sichuan cuisine version) on the 10-year CVD risk and vascular age: a randomised controlled feeding trial. Br J Nutr 2024 Mar 28;131(6):997-1006. [doi: 10.1017/S0007114523002416] [Medline: 37926909]
45. Wang C, Chen X, Hu J, Shahid M. Poverty alleviation and rural revitalization: perspective of fiscal spending and data evidence from 81 Chinese counties. Heliyon 2023 Jul;9(7):e17451. [doi: 10.1016/j.heliyon.2023.e17451] [Medline: 37483731]
46. Wang W, Chen K, Xiao W, Du J, Qiao H. Determinants of health poverty vulnerability in rural areas of Western China in the post-poverty relief era: an analysis based on the Anderson behavioral model. BMC Public Health 2024 Feb 14;24(1):459. [doi: 10.1186/s12889-024-18035-6] [Medline: 38355428]
47. Duan L, Zhang L, Zhang X, Lu S. Trends in output of hypertension management and associated factors in primary care facilities: a latent trajectory analysis in China from 2009 to 2017. BMC Prim Care 2023 Sep 6;24(1):178. [doi: 10.1186/s12875-023-02139-w] [Medline: 37674136]
48. Hannan JA, Commodore-Mensah Y, Tokieda N, et al. Improving hypertension control and cardiovascular health: an urgent call to action for nursing. Worldviews Evid Based Nurs 2022 Feb;19(1):6-15. [doi: 10.1111/wvn.12560] [Medline: 35137506]
49. Chen Y, Li X, Jing G, et al. Health education interventions for older adults with hypertension: a systematic review and meta-analysis. Public Health Nurs 2020 May;37(3):461-469. [doi: 10.1111/phn.12698] [Medline: 31943315]

Abbreviations

APC: annual percentage change
ARIMA: autoregressive integrated moving average
DASH: Dietary Approaches to Stop Hypertension
IRB: institutional review board

Edited by A Mavragani; submitted 17.09.24; peer-reviewed by Y Wang, Y Jia; revised version received 04.06.25; accepted 06.06.25; published 19.08.25.

Please cite as:
Zhang Y, Xia D, Lv Z, Wu C, Chen J, Zhang L
Hypertension Prevalence Among People Lifted Out of Poverty in China in 2018-2023: Retrospective Spatiotemporal Analysis
JMIR Public Health Surveill 2025;11:e66501
URL: https://publichealth.jmir.org/2025/1/e66501
doi:10.2196/66501

© Ying Zhang, Dong Xia, Zhiyu Lv, Chennan Wu, Jiapeng Chen, Lulu Zhang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Distribution and Risk Factors of Scrub Typhus in South Korea, From 2013 to 2019: Bayesian Spatiotemporal Analysis

Jeehyun Kim[1], PhD; Penelope Vounatsou[2], PhD; Byung Chul Chun[1], MD, PhD

[1]Department of Preventive Medicine, College of Medicine, Korea University, 73 Goryeodae-ro, Seoungbuk-gu, Seoul, Republic of Korea
[2]Swiss Tropical and Public Health Institute, University of Basel, Basel, Switzerland

Corresponding Author:
Byung Chul Chun, MD, PhD
Department of Preventive Medicine, College of Medicine, Korea University, 73 Goryeodae-ro, Seoungbuk-gu, Seoul, Republic of Korea

Abstract

Background: Scrub typhus (ST), also known as tsutsugamushi disease, is a common febrile vector-borne illness in South Korea, transmitted by trombiculid mites infected with *Orientia tsutsugamushi*, with rodents serving as the main hosts. Although vector-borne diseases like ST require both a One Health approach and a spatiotemporal perspective to fully understand their complex dynamics, previous studies have often lacked integrated analyses that simultaneously address disease dynamics, vectors, and environmental shifts.

Objective: We aimed to explore spatiotemporal trends, high-risk areas, and risk factors of ST by simultaneously incorporating host and environmental information.

Methods: ST cases were extracted from the 2013 - 2019 Korea National Health Insurance Service data at 250 municipal levels and by epidemiological weeks (*International Classification of Diseases, Tenth Revision, Clinical Modification code*: A75.3). Data on potential risk factors, including the maximum probability of rodent presence, area of dry field farming, forest coverage, woman farmer population, and financial independence, were obtained from publicly available sources. In particular, the maximum rodent presence probability was estimated using a maximum entropy model incorporating ecological and climate variables. Spatial autocorrelation was assessed using Global Moran *I* statistics with 999 Monte Carlo permutations. Spatial and temporal clusters were identified using Getis-Ord Gi* and hot and cold spot trend analyses. Bayesian hurdle models with a spatiotemporal interaction term, accounting for zero-inflated Poisson distribution, were used to identify associations between ST incidence and regional factors. Stratification analyses by gender and age group (0 - 39, 40 - 59, 60 - 79, and ≥80 years) were performed.

Results: Between 2013 and 2019, 95,601 ST patients were reported. ST incidence had positive spatial autocorrelation (*I*=0.600; *P*=.01), with spatial expansion from southwestern to northeastern regions. Spatiotemporal models demonstrated better fit compared with spatial and temporal models, as indicated by lower Watanabe-Akaike information criterion (WAIC) values. Municipalities with higher rodent suitability (β coefficient=0.618; 95% credible interval [CrI] 0.425 - 0.812) and lower financial independence from central government (β coefficient=−0.304; 95% CrI −0.445 to −0.163) had higher likelihoods of increased ST incidence, even after adjusting for spatiotemporal autocorrelation. However, risk factors varied by age group: among individuals aged 40 years or older, ST incidence was positively associated with rodent suitability, while patients in the 0 - 39 years age group showed no association with rodent suitability (β coefficient=0.028; 95% CrI −0.072 to 0.126), and ST incidence was negatively associated with the women farmer population (β coefficient=-0.115; 95% Crl=−0.223 to −0.006).

Conclusions: This is the first study to investigate ST in South Korea using a spatiotemporal framework grounded in a holistic One Health perspective. We elucidated the critical role of spatiotemporal dynamics in ST distribution, highlighting rodent suitability and economic independence as key drivers of disease distribution. Our findings lay the groundwork for evidence-based, region-specific intervention strategies and may inform targeted public health strategies in South Korea and other settings with similar ecological conditions.

(JMIR Public Health Surveill 2025;11:e68437) doi:10.2196/68437

KEYWORDS

Orientia tsutsugamushi; vector-borne diseases; Rodentia; spatiotemporal analysis; One Health

XSL•FO
RenderX

Introduction

Scrub typhus (ST), also known as tsutsugamushi disease, is a vector-borne illness transmitted by trombiculid mites infected with *Orientia tsutsugamushi* [1]. It is a common cause of acute febrile illness across the Asia-Pacific region, including South Korea [1]. Clinical manifestations range from acute fever, headache, rash, and skin eschars at mite bite sites to severe conditions such as acute respiratory distress syndrome, kidney failure, and multiorgan failure, particularly in the absence of timely antibiotic treatment [1,2]. Despite advances in diagnostics and treatment for decades, no effective vaccine exists [2]. The disease primarily occurs during the chigger mite breeding season, typically with an autumn peak in South Korea [3], a seasonal pattern also observed in leptospirosis and hemorrhagic fever with renal syndrome, with rodents as the primary host [4,5].

Analyzing ST requires a holistic approach that considers individual attributes, vectors, and environmental factors, aligned with the One Health perspective [6]. The One Health perspective can improve our understanding of the disease and help identify risk factors by addressing various factors simultaneously, enabling more effective prevention and interventions in person, time, and place [7,8]. In South Korea, *Apodemus agrarius*, the most common rodent species, has been identified as an important reservoir of zoonotic pathogens, including *Orienta tsutsugamushi*, highlighting the need to integrate host surveillance within a One Health framework to support comprehensive public health strategies [9]. However, previous studies have limitations in simultaneously addressing various factors. Most surveys focused primarily on individual experiences [10,11], and ecological studies have addressed environmental factors without considering the distribution of vectors and rodents [12-15]. Previous studies exploring the association between vector distribution and ST infections have overlooked the information environment and human behavior [16], such that they had limitations in suggesting robust risk factors.

Studies on vector-borne diseases have mandated integrated spatiotemporal analyses that encompass disease dynamics, vectors, and environmental shifts. Spatiotemporal methods can help identify high-risk areas [17,18], understand disease transmission [19], monitor disease trends [19], and assess the impact of environmental factors [17,19]. Although previous studies considered both the spatial and temporal axes simultaneously, the studies had limitations in pursuing the One Health approach [12,13,20].

Therefore, we aimed to investigate the spatiotemporal trends, high-risk areas, and risk factors of ST by integrating hosts and environmental information, and spatiotemporal dynamics. These findings are expected to provide a robust foundation for evidence-based policies to alleviate the disease burden, with particular emphasis on rodent distribution as a central determinant of ST risk.

Methods

Data

Scrub Typhus Data

Data on ST cases from 2013 to 2019 were extracted from the National Health Insurance Service (NHIS) of South Korea. We used the NHIS claims data, which cover the entire Korean population under the country's universal health insurance system. Unlike data from the national infectious disease surveillance system, which may be affected by periodic changes in reporting criteria, NHIS data are stable and structured, making them suitable for long-term trend analysis [21]. Therefore, although surveillance reporting criteria were revised in 2019, their impact on NHIS data were limited, and we retained the study period to 2013 - 2019 to ensure sufficient temporal coverage.

ST was defined according to the *International Classification of Diseases, Tenth Revision*, Clinical Modification code (ICD-10-CM code) of A75.3. The data were extracted in 250 municipal-level administrative boundaries on the spatial axis and were extracted in epidemiological week, commonly referred to as Morbidity and Mortality Weekly Report weeks [22], and calendar month on temporal axis. In South Korea, administrative divisions are organized into 4 levels: National, State (Sido), Municipal (Si [cities], Gun [counties], Gu [districts]), Town (Eup-Myun-Dong). The data included information on residential address coded by municipality, gender (men and women), and age group, which were extracted in 20-year interval category (0 - 39 years, 40 - 59 years, 60 - 79 years, and ≥80 years of age) because of privacy-related data extraction regulations of the data provider. Patients with complete data were included in the analyses.

As the standard analysis unit, municipal-level administrative boundaries in 2019 and epidemiological weeks were used to ensure comparability across years. For spatial analysis, a municipal-level shapefile was obtained from the publicly accessible Statistical Geographic Information Service [23]. Population data to calculate incidence per 100,000 population for descriptive analysis was extracted from Statistics Korea [24].

Potential Risk Factors

The maximum probability of rodent presence by municipality was selected as a potential risk factor of ST to consider reservoir distribution, aligned with the One Health perspective [6]. Based on previous studies, information on the maximum probability of the presence of rodents was made using the maximum entropy model (MaxEnt) [25], with point data of rodent existence, human population density [26], and various environmental raster data: 2014 - 2018 average enhanced vegetation index (EVI) [27], land surface temperature (LST) data [27], and a digital elevation model (DEM) [27]. Rodent existence point data were extracted from the 2014 - 2018 National Ecosystem Survey data from the Ministry of Environment [28] regardless of the rodent species and deleted the redundant points. The raster data of human population density in 1×1 km resolution was extracted from "The spatial distribution of population density in 2018,

XSL•FO
RenderX

South Korea" in WorldPop hub [29]. The raster data of EVI and LST were extracted from moderate-resolution imaging spectroradiometer (MODIS) remote sensing data using *MODIStsp* package [30] in R software (version 4.2.1 R Foundation for Statistical Computing). Monthly EVI raster at 1×1 km resolution and daily daytime and nighttime LST at 1×1 km were averaged. The DEM raster at 90×90 m was extracted from the National Spatial Data Infrastructure Portal of the Ministry of Land, Infrastructure, and Transport [31]. The study area of MaxEnt was set outside of a 10 km buffer from rodent existence points for the model to study various environmental areas beyond the point where rodents were discovered. The results of the MaxEnt model showed a better prediction than the random model, with an area under the curve of 0.690. The maximum rodent suitability by municipality was then extracted from the results of the MaxEnt model.

Besides maximum rodent suitability, area of dry field farming [10,11,32,33], forest coverage [32], woman farmer population [11,33], and financial independence as a socio-economic status of the municipality [34] were included based on previous studies and relevance to ST transmission (Table 1). The woman farmer population was defined as the number of females formally engaged in farming activities per municipality and year and used as an ecological indicator of agricultural exposure of women. This variable was selected based on the higher burden of ST among women in Korea and previous findings suggesting strong associations between agricultural activity and ST incidence in women [11,33]. Information on each variable was extracted in year and municipality level. In addition, the population of women and men farmers was extracted from Korean statistical information service at the year and state level to compare the gender ratio in farming population by year and state as supplementary information. Although other potential predictors such as meteorological factors were considered, relevant environmental conditions were already incorporated into rodent suitability construction. Variables were selected considering epidemiological plausibility, municipal-level data availability, and efforts to reduce collinearity among predictors. The shapefile was obtained from the National Spatial Data Infrastructure Portal [31].

Table . Description and data sources of outcome and potential risk factors for scrub typhus incidence in South Korea, 2013-2019.

Variables	Description	Sources
Scrub typhus		
Cases (n)	Cumulative scrub typhus cases by municipality and epidemiological week	NHIS[a]
Potential risk factors		
Maximum rodent suitability	Maximum possibility of rodent presence made by maximum entropy model using point data of rodent existence and various environmental raster data, per municipality (range: 0.0 - 1.0)	NESD[b], WorldPop hub, MODIS[c], NSDIP[d]
Financial independence (%)	Total earned income over total expense, per municipality and year	KOSIS[e]
Forest area (m^2)	Land that forms forests and fields, per municipality and year	KOSIS[e]
Dry field farming area (m^2)	Land that mainly grow plants that do not apply water, per municipality and year	KOSIS[e]
Woman farmer population (n)	Population of women farmers, per municipality and year	KOSIS[e]
Offset		
Population (n)	Human population based on the resident registry in, per municipality	KOSIS[e]

[a]NHIS: National Health Insurance Service.

[b]NESD: National ecosystem survey data.

[c]MODIS: Moderate resolution imaging spectroradiometer.

[d]NSDIP: National Spatial Data Infrastructure Portal.

[e]KOSIS: Korean statistical information service.

Statistical Analysis

Descriptive Analysis

The frequency and incidence of ST were analyzed with respect to calendar year, gender, and age group, similar to the approach used in a previous study [21]. Descriptive statistics, including mean, SD, minimum, median, first quartile, third quartile, maximum, and coefficient of variation, were used to summarize the baseline information for the variables. As the dataset overlaps with that in the previous study [21], relevant findings were referenced. However, this study introduces spatiotemporal analysis methods and adopts more suitable age groupings, offering a refined understanding of ST dynamics over time and space.

XSL•FO
RenderX

To examine the spatiotemporal distribution of ST incidence, choropleth maps depicting ST incidence per 100,000 population with quintile intervals were created, stratified by calendar year, gender, and age group. Since maps only visualize data and cannot provide definitive conclusions due to various sources of uncertainty, Global Moran's *I* statistics were used with 999 Monte Carlo permutations ($P<.05$) to quantify the annual spatial autocorrelation in ST incidence. The Moran's *I* index ranges from −1 to 1, with a positive value signifying a positive spatial autocorrelation. This implies that regions with high variable values are likely to be close to each other. The significant results of Global Moran's *I* statistics could provide verification to consider spatial autocorrelation for further analysis. Subgroup analyses were conducted based on gender and age group.

Furthermore, a Getis-Ord Gi* analysis stratified by calendar year, gender, and age group was performed to identify municipal-level clusters within each stratum. Municipalities with high positive *z* scores, resulting from the Getis-Ord Gi* analysis, indicated "hot spots," signifying spatial clustering of high values. Conversely, "cold spots" were characterized by low negative *z* scores, indicating spatial clustering of low values surrounded by other low-value municipalities. The results of Getis-Ord Gi* analysis provide information on locations with either high or low value of ST incidence that are statistically significantly clustered.

The hot and cold spot trend (HCT) analysis was performed, which is the visualization of a space-time cube in 2 dimensions. The HCT depicts increasing and decreasing *z* scores resulting from Mann–Kendall statistics, classified as "up trend," "down trend," or "no significant trend" with confidence levels of 90%, 95%, and 99%, respectively. Thus, while Getis-Ord Gi* analysis tests clustering on spatial axis, HCT analysis examines statistical trends across both spatial and temporal axes. The results could test whether the ST clusters expand in spatial and temporal perspectives. The three-dimensional space-time cube (x, y, and t) was generated using the "Create Space Time Cube From Defined Locations" tool, with 250 municipal polygons as spatial units and one calendar month as temporal units, resulting in 21,000 bins (250 spatial units×84 temporal units). Missing values in each bin were replaced with zeros. Spatial analyses, including mapping, assessment of spatial autocorrelation, and cluster identification, were performed using R version 4.2.1 [35], and HCT analysis was conducted using ArcGIS Pro version 3.3.1 (Esri) [36].

Modeling

Modeling was performed to explore the association between ST incidence and each potential risk factor after adjusting the effect of space and time trend. For modeling, continuous predictors, including rodent suitability, were standardized due to varying units. Age-standardized ST incidence per 100,000 population was summarized using population data from Population and Housing Census in 2015 [37]. Univariable linear regression was conducted for each variable using age-standardized ST incidence as dependent variable. Variables with a *P* value ≥.25 [38] were identified for exclusion in the multivariable regression (Table S1 in Multimedia Appendix 1). Multicollinearity was assessed using a variance inflation factors threshold of 10 in the multivariable regression (Table S2 in Multimedia Appendix 1).

The primary analytical method used was the Bayesian hurdle Poisson spatiotemporal analysis using the Integrated Nested Laplace Approximation (INLA) approach [39,40], a computationally efficient alternative to the frequently used Markov Chain Monte Carlo simulations [40]. INLA delivers precise and fast results for latent Gaussian models [40] and is widely applied in spatiotemporal models to handle model complexity and vast database dimensions [41,42].

The zero-inflated Poisson distribution was managed using the hurdle Poisson model due to the excessive zeroes in the ST incidence data [39,43]. A hurdle model outperforms a zero-inflated model when covariates have the possibility of zero deflation when using excess zero data [43]. Thus, a hurdle model was constructed with 2 likelihoods (Bernoulli and truncated Poisson) in the INLA [39,44]. The Bernoulli distribution, implemented as "binomial" in the R-INLA, modeled the probability of any case occurrence [40,44], while the zero-truncated Poisson distribution, specified as "poisson," captured the incidence when cases occurred [44]. The age-standardized incidence of ST per municipality, ii=1,…,250 and epiweek, tt=1,…,365 can be represented as Yit, using a hurdle Poisson model as follows:

$$P(Yit=y|p,\mu)=\{p \quad yit=0,(1-p)e-\mu\mu y/y!1-e-\mu \quad yit>0$$

where p denotes the probability that a participant belongs to the zero component. First, to examine the necessity of spatiotemporal interaction term compared to each spatial and temporal term, three base models were constructed as follows:

$$\text{Spatial model 0 } \{logit(pit)=\beta 0p+uip+vip\log(\mu it)=\beta 0\mu+uiμ+viμ$$

$$\text{Temporal model 0 } \{logit(pit)=\beta 0p+\gamma tp+\varphi tp\log(\mu it)=\beta 0\mu+\gamma\mu+\varphi t\mu$$

$$\text{Spatiotemporal model0}\{g(\mu)=\beta 0+uip+vip+\gamma tp+\varphi tp+\delta itpg(\mu)=\beta 0\mu+uiμ+viμ+\gamma\mu+\varphi t\mu+\delta it\mu$$

where μit denotes the exponential value of ST incidence for municipality i and epiweek t, β0 denotes intercept [45]. The model contains ui as an intrinsic conditional auto-regressive (iCAR) term, vi as a spatially unstructured term, γt as a temporally structured effect using an autoregressive model of order 1, and φt that has been specified as φt~Normal0,1τφ [46]. The term δit is an interaction term of spatial and temporal structured effect [46]. The spatial weight matrix, which determines the iCAR term, represents a neighboring relationship among the 250 municipalities. The neighbor matrix was constructed using k-nearest neighborhood, with k as 3. Priors followed the default options of R-INLA.

In addition, to assess the effect of potential risk factors other than maximum rodent suitability on ST incidence, a term for these factors was added to each model. When βj denotes regression parameters of each potential risk factor excluding the maximum possibility of rodent presence, the model was constructed as follows:

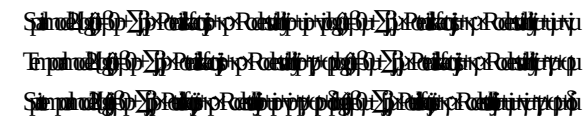
$$\text{Spatial model1}\{g(\mu)=\beta 0p+\sum\beta jxjp+uip+vipg(\mu)=\beta 0\mu+\sum\beta jxj\mu+uiμ+viμ$$

$$\text{Temporal model1}\{g(\mu)=\beta 0p+\sum\beta jxjp+\gamma tp+\varphi tpg(\mu)=\beta 0\mu+\sum\beta jxj\mu+\gamma\mu+\varphi t\mu$$

$$\text{Spatiotemporal model1}\{g(\mu)=\beta 0p+\sum\beta jxjp+uip+vip+\gamma tp+\varphi tp+\delta itpg(\mu)=\beta 0\mu+\sum\beta jxj\mu+uiμ+viμ+\gamma\mu+\varphi t\mu+\delta it\mu$$

XSL•FO
RenderX

Finally, to assess the risk of rodent suitability on ST incidence after accounting for other potential risk factors, and vice versa, a term of rodent suitability was added to each model, where κ denotes the coefficient for rodent suitability, as follows:



To compare the fit of the models, the Watanabe-Akaike information criterion (WAIC) was used. WAIC offers the advantage of averaging over the posterior distribution instead of relying on a single-point estimate, making it particularly valuable for models with hierarchical or mixed structures where point estimates may not be meaningful, even in cases involving singular models [47]. In general, a model with smaller WAIC value was better supported by data [47]. In this study, final model selection considered both WAIC values and the scientific relevance of covariates. Specifically, we prioritized models that allowed evaluation of our primary hypothesis regarding habitat suitability as a risk factor for scrub typhus incidence. Posterior means and 95% credible intervals (CrI) were reported, and model fit was evaluated spatially and temporally by plotting the fitted and predicted values of ST incidence.

Subgroup analyses by sex and age groups (0 - 39, 40 - 59, 60 - 79, and ≥80 years of age) were conducted using ST case counts as the outcome and the log-transformed population as an offset. Information on woman farmer population was excluded when performing subgroup analysis according to gender, to avoid the possibility of overcorrection. Modeling was performed using R software (version 4.2.1; R Foundation for Statistical Computing) [35] (Multimedia Appendix 2).

Sensitivity Analysis

Sensitivity analysis was performed using a different spatial weight matrix constructed through a k-nearest neighbor approach, with k set at 5, to assess the robustness of the results. In addition, another sensitivity analysis was conducted using the number of ST cases as the outcome and log-transformed population as an offset.

Ethical Considerations

This study received an exemption from the Institutional Review Board (IRB) of Korea University (IRB exemption number: KUIRB-2021-0237-02) as the data used in this study were fully anonymized and contained no personally identifiable information. The requirement for obtaining informed consent was waived as the study was observational in design. As the analysis used secondary data, no compensation was provided.

Results

Descriptive Analysis

A total of 95,627 ST cases were extracted between December 30, 2012, and December 31, 2019 from the NHIS data at the municipal level. For the descriptive analysis, 95,601 patients with ST were considered, and for the modeling, 95,582 patients with cases that corresponded to the epiweek-based study period (from the first epiweek of 2013 to the 52nd epiweek of 2019) were included (Figure 1).

XSL•FO
RenderX

Figure 1. Flowchart of scrub typhus cases diagnosed with International Classification of Diseases, 10th Revision-Clinical Modification code A75.3, extracted from the National Health Insurance Service data in South Korea from 2013 to 2019. Epiweek: epidemiological week; NHIS: National Health Insurance Service in South Korea.

XSL•FO
RenderX

Between 2013 - 2019, ST cases (18,822 in 2013; 7770 in 2019) and incidence per 100,000 population (36.8 in 2013; 15.0 in 2019) decreased (Table S3 in Multimedia Appendix 1, [21]). Although a marginal increase was observed in 2016 (cases: 17,086; incidence: 33.1), the cases decreased by around 58.7% between 2013 - 2019, with a 59.02% decrease in incidence over the same period. The ST incidence per 100,000 population was consistently higher in women than in men between 2013 - 2019 (men: 22.0, women: 30.9). The ≥80 years of age group had the highest incidence (94.6), followed by the 60 - 79 (79.4), 40 - 59 (24.1), and 0 - 39 age group (4.96; Table S4 in Multimedia Appendix 1) [21].

Baseline information for each variable is provided in Table 2. The average number of ST cases in 2013 - 2019 at the municipal level and epiweek was 1.0 (SD 3.5) and the coefficient of variance (CV) was 3.4, suggesting regional and temporal variations. The CV of potential risk factors ranged from 0.2 to 1.1, indicating lower regional and temporal variation than ST incidence per 100,000 population (CV=4.2).

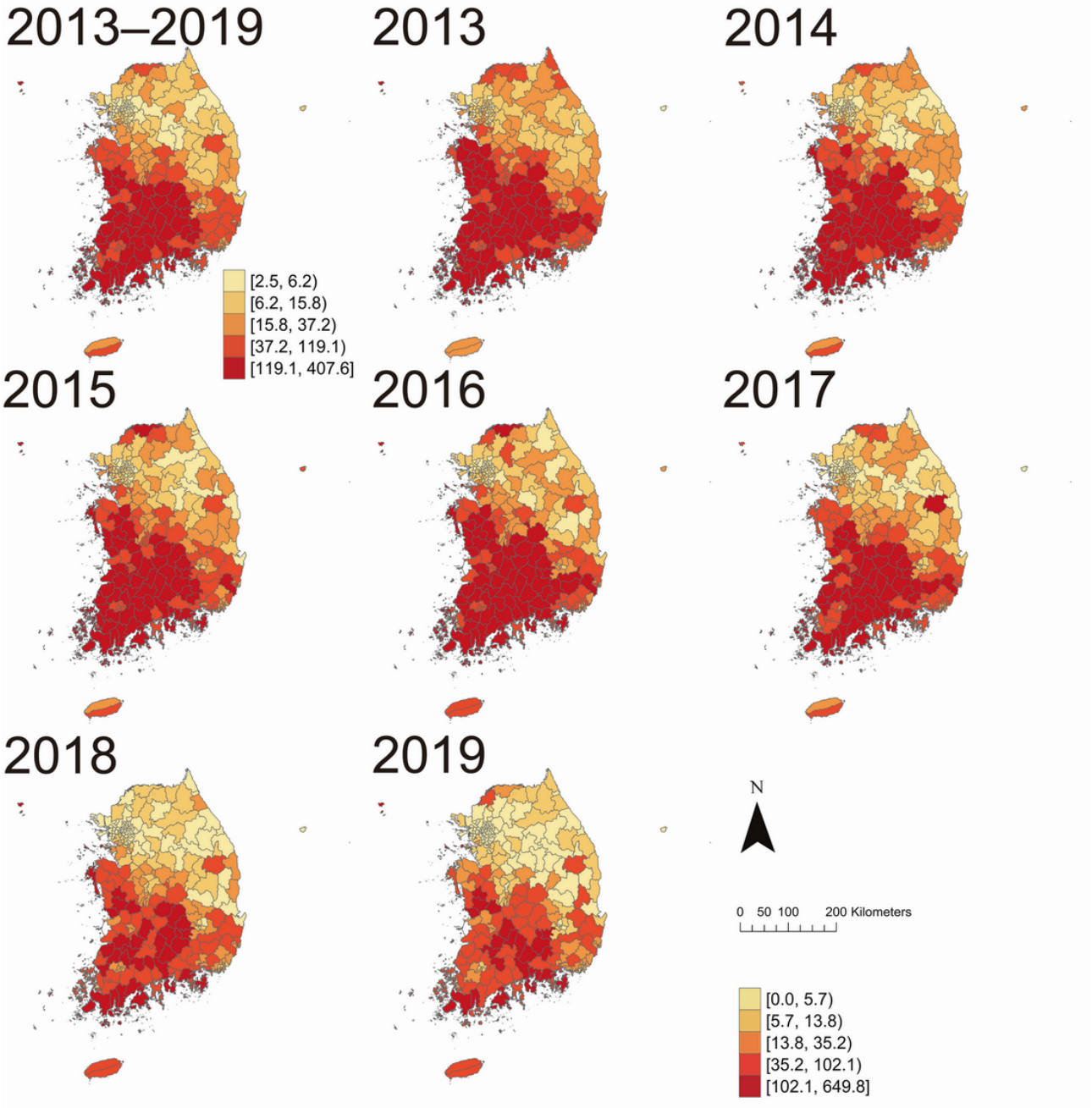
Table . Summary statistics of scrub typhus incidence and potential risk factors by municipality and epidemiological week in South Korea, 2013-2019.

Variables	Mean (SD)	Minimum	Median (Q1-Q3)	Maximum	Coefficient of variance
Scrub typhus					
Cases (n)	1.0 (3.5)	0.0	0.0 (0.0-1.0)	81.0	3.4
Incidence per 100,000 population	1.0 (4.1)	0.0	0.0 (0.0-0.3)	105.7	4.2
Potential risk factors					
Maximum rodent suitability	0.8 (0.2)	0.1	0.8 (0.7-0.9)	1.0	0.2
Financial independence (%)	28.6 (14.9)	7.3	24.2 (16.3-38.9)	75.9	0.5
Forest area (m^2)	255,633,212.9 (287,531,335.5)	0.0	178,314,785.4 (18,240,087.4-405,361,824.9)	1,532,013,213.1	1.1
Dry field farming area (m^2)	30,624,701.2 (32,395,574.9)	0.0	23,566,365.5 (2,706,313.0-52,257,727.8)	210,272,858.0	1.1
Woman farmer population (n)	5,153.8 (4,617.4)	147.0	4286.0 (1068.0-7939.0)	32,932.0	0.9
Offset					
Population (n)	206,370.1 (162,965.9)	9617.0	175,277.0 (59,061.0-314,110.0)	863,720.0	0.8

According to the regional distribution map of ST incidence per 100,000 population in 2013 - 2019, the southwestern parts of South Korea had higher ST incidence (119.1 - 407.6 incidence per 100,000 population) than northeastern parts (Figure 2). Regardless of the year, the ST incidence was the highest in the southwestern region of South Korea. This trend was consistent, regardless of gender (Figure S1 in Multimedia Appendix 1) or age group (Figure S2 in Multimedia Appendix 1). However, over time, the number of municipalities with the highest ST incidence quintile (102.1 - 649.8 incidence per 100,000 population) decreased.

XSL•FO
RenderX

Figure 2. Annual incidence of scrub typhus per 100,000 population in 250 municipalities across South Korea from 2013 to 2019. Each panel represents a specific year, with the top-left panel showing the cumulative incidence from 2013 to 2019. Colors represent incidence rate ranges, where darker red indicates higher incidence and light yellow indicates lower incidence.

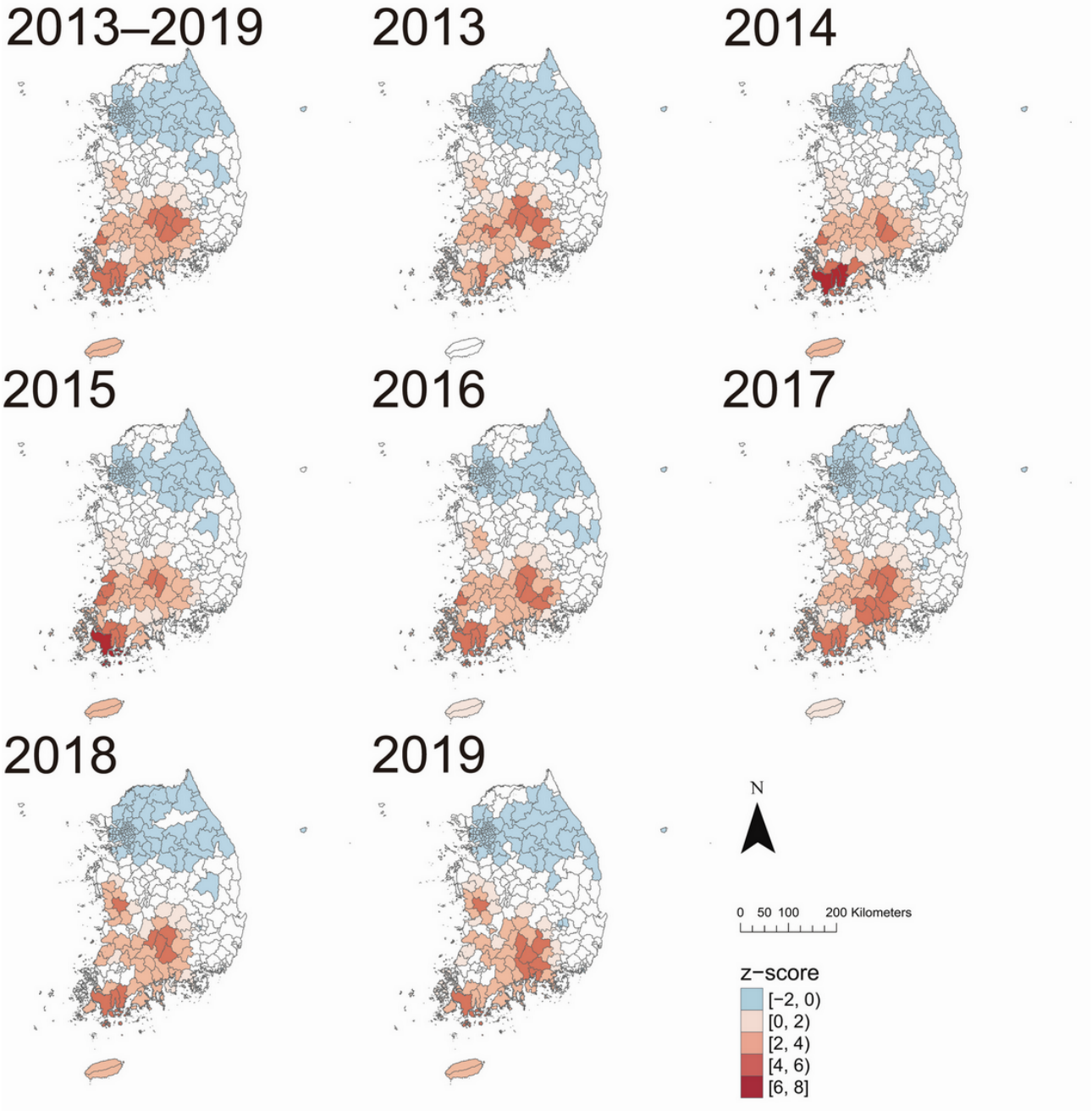


Among all the study participants, the results of the Global Moran's *I* statistics indicated that ST incidence had a positive spatial autocorrelation (*I*=0.600; *P*=.01, Table S5 in Multimedia Appendix 1). Positive spatial autocorrelation was consistently reported in 2013 - 2019, with 2015 having the strongest index value (*I*=0.594; *P*=.01). Positive spatial autocorrelation was consistently reported regardless of gender (men, *I*=0.600, *P*=.01; women, *I*=0.590, *P*=.01) and age group (age 0 - 39 years, *I*=0.564, *P*=.01; age 40 - 59 years, *I*=0.630, *P*=.01; age 60 - 79 years, *I*=0.649, *P*=.01; age≥80 years, *I*=0.622, *P*=.01).

The Getis-Ord Gi* analysis on ST incidence per 100,000 population consistently detected hot and cold spots in the southwestern and northeastern parts of South Korea, respectively, between 2013 - 2019 (Figure 3). The trend persisted, regardless of gender (Figure S1 in Multimedia Appendix 1) and age group (Figure S2 in Multimedia Appendix 1).

XSL•FO
RenderX

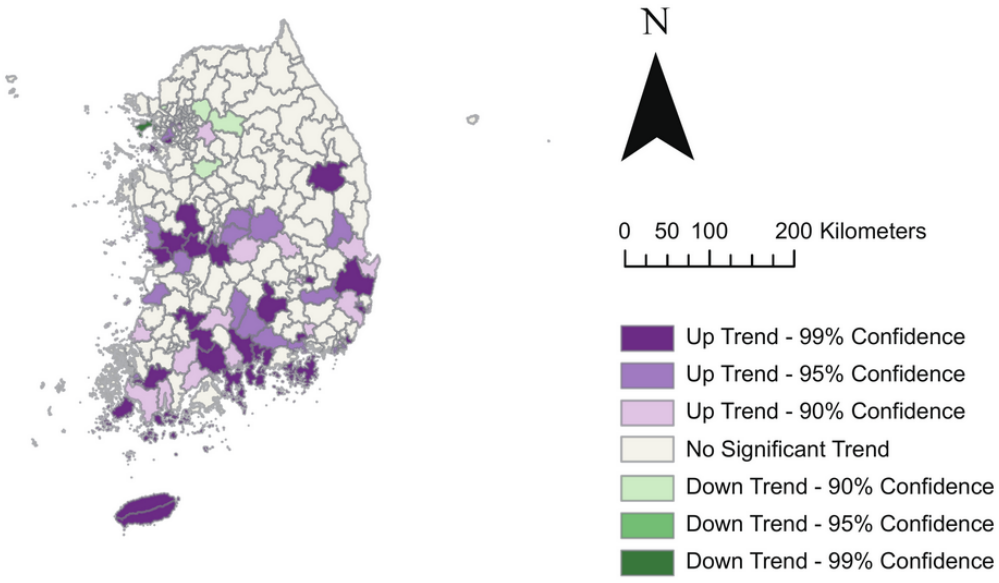
Figure 3. Spatial clusters of scrub typhus incidence per 100,000 population identified using Getis-Ord Gi* analysis across 250 municipalities in South Korea, annually from 2013 to 2019 and for the overall period. Red shades indicate hot spots (high positive *z* score), and blue shades indicate cold spots (low negative *z* score). Color intensity reflects the magnitude of the *z* score resulted from Getis-Ord Gi* analysis.



The HCT results for the total population indicated increasing trends in municipalities across the southern and mid regions of South Korea Figure 4. Municipalities showing increasing trends were also observed in the northwestern part, specifically around the capital region. This pattern was consistent for both genders (Figure S1 in Multimedia Appendix 1), with the southern island of Jeju-do consistently exhibiting an upward trend, regardless of gender and age group. However, the trend varied according to the age group (Figure S2 in Multimedia Appendix 1). Specifically, in the 0 - 39 years age group, municipalities with increasing trends were scattered throughout the country, but the upward trend was more clearly observed in the capital region compared with other age groups. For the 40 - 59, 60 - 79, and ≥80 years age groups, municipalities with increasing trends were more concentrated in the southern and mid regions.

XSL•FO
RenderX

Figure 4. Hot and cold spot trends of scrub typhus incidence per 100,000 population across 250 municipalities in South Korea, from 2013 to 2019, visualized using a 2-dimensional representation of a space-time cube. Purple shades represent up trends (hot spot trends), and green shades indicate down trends (cold spot trends), each categorized by 90%, 95%, and 99% confidence levels. White areas indicate no significant trend.



Modeling

The results of multivariable Bayesian models on ST incidence are presented in Table 3 and Table S6 in Multimedia Appendix 1. Overall, spatiotemporal models demonstrated better fit compared with spatial and temporal models, with lower WAIC values. Among the spatiotemporal models, model 0 had the lowest WAIC value (75,014.8), followed by model 1 (75,019.1), and model 2 (75,037.2). However, given its ability to incorporate relevant covariates and test the primary hypothesis, spatiotemporal model 2 was retained as the model of interest, despite having a slightly higher WAIC than model 0. Sensitivity analysis results also showed that the spatiotemporal model had lower WAIC than spatial and temporal models (Table S7 in Multimedia Appendix 1). By comparing the maps (Figure S3 in Multimedia Appendix 1) and time-series plot (Figure S4 in Multimedia Appendix 1) between the observed and predicted ST incidences using spatiotemporal model 2, the distributions had similarities.

Table . Results of multivariable Bayesian models for age-standardized scrub typhus incidence in South Korea, 2013 to 2019, presented as β coefficients (95% credible intervals).

Model fit metric and variables	Spatial model 2[a]	Temporal model 2[b]	Spatiotemporal model 2[c]
WAIC[d]	211,792.6	115,177.7	75,037.2
Maximum rodent suitability, β coefficient (95% CrI[e])	0.320 (0.075-0.564)	0.493 (0.477-0.509)	0.618 (0.425-0.812)
Financial independence (%), β coefficient (95% CrI)	−0.486 (−0.553 to −0.419)	−0.722 (−0.736 to −0.709)	−0.304 (−0.445 to −0.163)
Forest area (m^2), β coefficient (95% CrI)	0.604 (0.289-0.930)	0.086 (0.074-0.099)	0.171 (−0.076 to 0.421)
Dry field farming area (m^2), β coefficient (95% CrI)	−0.719 (−1.009 to −0.423)	−0.184 (−0.199 to −0.169)	0.040 (−0.228 to 0.311)
Woman farmer population (n), β coefficient (95% CrI)	1.582 (1.525-1.640)	0.272 (0.258-0.287)	0.081 (−0.104 to −0.266)

[a]Model including spatial terms and potential risk factors, including maximum rodent suitability.

[b]Model including temporal terms and potential risk factors, including maximum rodent suitability.

[c]Model including spatial and temporal terms, interaction term of spatial and temporal structured terms, and potential risk factors, including maximum rodent suitability.

[d]WAIC: Watanabe-Akaike information criterion.

[e]CrI: credible interval.

XSL•FO
RenderX

According to the results of spatiotemporal model 2, maximum rodent suitability (β coefficient=0.618, 95% credible interval [CrI] 0.425 - 0.812) was positively associated with ST incidence, while financial independence was negatively associated with ST incidence (β coefficient=−0.304, 95% CrI −0.445 to −0.163).

The subgroup analysis results by gender followed the trends of main analyses, except for forest area having a positive association (Table S8 in Multimedia Appendix 1). The results of the Bayesian spatiotemporal model differed by age group. Table 4 presented the results of the spatiotemporal models by age group. Regardless of age group, financial independence had negative and forest area had positive association with ST incidence. However, the incidence of the 40 - 59, 60 - 79, and ≥80 years of age group had positive association with maximum rodent suitability, while the incidence of 0 - 39 years of age did not have a significant association with rodent suitability (β coefficient=0.028, 95% CrI −0.072 to 0.126). Only the incidence of the 0 - 39 years of age group had a negative association with the women farmer population (β coefficient=−0.115, 95% CrI −0.223 to −0.006). The sensitivity analysis results followed the trends of main analyses, while forest area had significant association (Tables S7 and S9 in Multimedia Appendix 1) and women farmer population did not have significant association with ST incidence in the 0 - 39 year age group like main analysis results (Table S9 in Multimedia Appendix 1).

Table . Results of multivariable Bayesian spatiotemporal models for scrub typhus incidence stratified by age group in South Korea, 2013 to 2019.

Model fit metric and variables	Age group (years)			
	0 - 39	40-59	60 - 79	≥80
Participants, n (%)	8438 (8.8)	28,930 (30.3)	48,129 (50.3)	10,104 (10.6)
WAIC[a]	15,647.3	39,939.7	54,357.4	17,180.7
Maximum rodent suitability, β coefficient (95% CrI[b])	0.028 (−0.072 - 0.126)	0.149 (0.052 - 0.246)	0.223 (0.127 - 0.319)	0.099 (0.017 - 0.180)
Financial independence (%), β coefficient (95% CrI)	−0.350 (−0.431 to −0.268)	−0.262 (−0.329 to −0.194)	−0.206 (−0.269 to −0.143)	−0.136 (−0.199 to −0.072)
Forest area (m^2), β coefficient (95% CrI)	0.472 (0.332 - 0.613)	0.308 (0.180 - 0.436)	0.177 (0.052 - 0.302)	0.174 (0.080 - 0.268)
Dry field farming area (m^2), β coefficient (95% CrI)	0.011 (−0.139 to 0.161)	−0.037 (−0.174 to 0.101)	−0.034 (−0.168 to 0.103)	−0.012 (−0.115 to 0.091)
Woman farmer population (n), β coefficient (95% CrI)	−0.115 (−0.223 to −0.006)	0.036 (−0.046 to 0.118)	0.066 (−0.011 to 0.143)	−0.077 (−0.164 to 0.009)

[a]WAIC: Watanabe-Akaike information criterion.

[b]CrI: credible interval.

Discussion

Principal Findings

This study examined the demographic, spatial, and temporal distributions of the incidence of ST. Notably, ST incidence consistently exhibited spatial hotspots in the southwestern region of South Korea, regardless of year, sex, or age. This pattern aligns with previous findings and can be attributed to the ecological characteristics of the southwestern region, which is characterized by fertile lands conducive for rodent and mite habitats [48].

Moreover, HCT analysis revealed a consistent increasing trend in the southern and mid regions, despite a declining trend in incidence. A similar spatial expansion of ST cases from the southwestern to northeastern regions was noted between 2001 - 2006 [48], which was likely attributed to the effects of climate change, particularly rising temperatures that enhanced vector survival, expanded their geographic range, and accelerated pathogen growth [49]. However, variations in ST trends were observed when age groups were considered. The younger age group (0 - 39 years) exhibited scattered municipalities with increasing trends across the country, whereas older age groups (40 - 59, 60 - 79, and ≥80 years) displayed more concentrated increases in the southern and mid regions. This suggests that age-specific factors play a role in ST transmission dynamics and necessitate tailored prevention strategies based on specific demographic segments.

The current analysis further revealed the importance of incorporating both spatial and temporal dimensions in understanding the epidemiology of ST. Based on the WAIC, models with spatiotemporal interaction term (Spatiotemporal model 0, WAIC=75,014.8; Spatiotemporal model 1, WAIC=75,019.1; Spatiotemporal model 2, WAIC=75,037.2, Table 3, Table S7 in Multimedia Appendix 1) demonstrated a better fit than those with only spatial or temporal terms, regardless of including potential risk factors. This pattern was consistent in a sensitivity analysis using the number of ST cases as the outcome and log-transformed population as an offset (Table S7 in Multimedia Appendix 1), further confirming the benefit of modeling spatiotemporal interaction. This underscores the need for a comprehensive approach that considers both the spatial and temporal dimensions when investigating ST patterns. In particular, the ST data showed distinct temporal patterns in

XSL•FO
RenderX

each area that were inconsistent with those in other regions, and the spatial patterns exhibited considerable variability (Figure S5 in Multimedia Appendix 1). This supports the necessity of including a space-time perspective in the analysis. However, as shown in another sensitivity analysis using a more complex spatial structure, convergence issues occurred in the INLA-based model (Table S9 in Multimedia Appendix 1). This highlights the choice of spatial weight matrix that can affect model convergence in complex Bayesian settings.

We revealed a significant association between ST incidence and specific factors. Our findings quantitatively investigated that municipalities with high rodent suitability were more likely to experience higher ST incidence, supporting previous studies [50-53]. High rodent density, particularly in areas with abundant hosts, contributes significantly to the transmission of *Orientia tsutsugamushi* [52]. In addition, climate variability and rodent abundance jointly influence scrub typhus transmission [53]. The nonsignificant but positive point coefficient observed in the 0‐39 years age group might be due to the smaller number of ST cases and reduced exposure to rodents at home or work. However, as rodent suitability is determined by environmental potential for rodent habitation than actual numbers, future studies should focus on monitoring rodent distribution across the nation and disseminating this information to the public. Additionally, research exploring the relationship between ST cases and the distribution of natural hosts based on this data is needed. Managing human-rodent interaction, as proposed by Sharma et al [54], could be a key strategy for reducing scrub typhus transmission, particularly through rodent habitat control and improved hygiene practices.

Furthermore, this study highlighted the impact of municipality financial status on ST incidence. Municipalities with lower financial independence were more likely to have higher ST incidence, which may reflect the characteristics of rural areas where agricultural and forest environments are common [55,56]. Although these regions may receive targeted support from central government programs, their ecological and occupational conditions likely contribute to the elevated risk.

Following the results of previous studies [8,57], our findings indicated that municipalities with more forest coverage had higher ST incidence in age- and gender–stratified models. Occupational exposure to forest environments, particularly among forestry workers, significantly elevates risk [57]. In addition, Zangpo et al [8] identified activities like cardamom harvesting and living near forests as key risk factors, highlighting the importance of forest proximity in ST transmission. Therefore, implementing targeted preventive strategies, including regular monitoring and enhanced protective education for forestry workers and residents near forested areas, is essential to mitigate future ST cases.

The association between the population of women farmers and ST incidence was mainly observed in the 0‐39 years age group, where women farmer population appeared to have a protective effect on ST incidence at the municipal level. This reflects contrasting trends of increasing ST incidence among men and the decline in the population of women farmers within this age group. From 2003 to 2019, ST incidence became more male-dominant in the 0‐39 years age group, reducing the gender disparity in ST incidence [21]. This age group exhibited a marked increase in ST incidence in the surrounding capital area (Figure S2 in Multimedia Appendix 1), where farming is less prevalent and the proportion of women farmers sharply declined (Figure S6 in Multimedia Appendix 1). Since women farmers have been the main target of ST prevention efforts due to their risk from agricultural work habits during harvest seasons [1,58,59], these findings may indicate a gap in preventive outreach to other at-risk populations, such as younger individuals engaging in nonfarming outdoor activities [60,61]. Considering the changing demographics in both farming populations and ST cases, updated mitigation strategies, such as redefining target populations in nontraditional rural areas, are necessary.

Although previous studies have suggested that dry field farming is associated with outdoor exposure and ST risk [10,33], no statistically significant association was observed in our models. This reflects regional variation in land use, ecological conditions, or population-level exposure, which warrants further investigation.

Strengths and Limitations

To the best of our knowledge, this study is the first to investigate the spatiotemporal dynamics and risk factors of ST in South Korea in a One Health perspective. The considerable diversity observed in the municipal-level epidemic dynamics of ST underscores the necessity for a regional comprehension of the ST infection pattern. This understanding is crucial for the effective implementation of prevention and control strategies tailored to the burden of ST and its determinants at the municipal level. Furthermore, incorporating specific determinants of ST infection into these localized strategies or developing targeted group interventions is warranted. However, this study has some limitations. This study was an ecological study conducted at the municipal level, so the results should not be applied at the individual level. While ecological fallacy and unmeasured confounders may exist, a key advantage of our study is that the results can be directly translated into municipal-level ST interventions. Although our multivariable model enabled simultaneous assessment of multiple risk factors, it may not fully account for variable-specific causal pathways, including distinct confounders and effect modification. Nevertheless, the approach remains valuable insights into overall patterns and supports future hypothesis-driven research. Moreover, the NHIS data provided only municipal residential codes and lacked information on the bite site. As a result, spatial patterns, particularly among younger or more mobile populations, should be interpreted with caution. Further studies that incorporate bite-site information and target the entire nation are required. Although leisure activities such as hiking or camping may contribute to ST risk in younger populations [60], relevant municipal-level indicators were not available. Moreover, ST incidence remains the highest in older adults (Figure 4), despite increasing trends in recreational activity, which suggests that limited explanatory power of leisure exposure. Some popular recreational areas, including mountainous regions and forested areas like Gangwon-do, also showed low incidence (Figures 2 and 3), possibly due to ecological factors such as climate and latitude affecting vector habits [3,8]. In addition, direct

XSL•FO
RenderX

information on mite distribution across the country is unavailable. While vector surveillance programs such as Vector-net could serve as alternatives, their data were not publicly accessible. Therefore, partial compensation has been attempted by using information on rodent suitability, considering that rodents serve as natural hosts of mites and rodent density has been shown to be related to chigger mite distribution [62]. The construction of maximum rodent suitability also incorporated environmental factors, aligning with a One Health approach. Despite these limitations, these findings are expected to provide a foundation for addressing ST through evidence-based strategies based on examined hot spots and risk factors, grounded in a holistic approach that encompasses both One Health and spatiotemporal aspects.

Conclusions

This study provides a detailed analysis of the spatiotemporal dynamics and risk factors of ST in South Korea from 2013 to 2019, emphasizing the importance of incorporating spatiotemporal aspects. The findings indicated positive spatial autocorrelation, with ST expanding from the southwest to the northeast regions. Municipalities with higher rodent suitability, greater forest coverage, and lower financial independence were more likely to experience increased ST incidence, with risk factors varying by age group. This study underscores the value of a holistic, spatiotemporal approach, highlighting the roles of rodent suitability, forest area, and socioeconomic factors in shaping ST dynamics. These findings may inform targeted public health intervention in Korea and provide broader insights applicable to other regions where similar ecological conditions and socioeconomic contexts exist. These insights provide a foundation for developing targeted public health interventions grounded in a One Health perspective to more effectively manage ST in South Korea and similar global settings.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to licensing restrictions by the Korean National Health Insurance Service (NHIS) but are available from the corresponding author on reasonable request and with approval from the NHIS.

Authors' Contributions

JK: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing-original draft, Writing-review and editing. PV: Writing-review and editing. BCC: Conceptualization, Validation, Writing-review and editing, Supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional descriptive and spatiotemporal analyses of scrub typhus incidence, clusters, and risk factors in South Korea, 2013–2019.
[DOCX File, 3729 KB - publichealth_v11i1e68437_app1.docx]

Multimedia Appendix 2
R code for the Bayesian hurdle Poisson spatiotemporal analysis.
[DOCX File, 23 KB - publichealth_v11i1e68437_app2.docx]

Checklist 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.
[DOCX File, 35 KB - publichealth_v11i1e68437_app3.docx]

References

1. Xu G, Walker DH, Jupiter D, Melby PC, Arcari CM. A review of the global epidemiology of scrub typhus. PLoS Negl Trop Dis 2017 Nov;11(11):e0006062. [doi: 10.1371/journal.pntd.0006062] [Medline: 29099844]
2. Jindal N, Abhishek K, Kumar A, et al. A comparative analysis of Orientia Tsutsugamushi (Scrub Typhus): essential information on causes, symptoms, and care. Clin Microbiol Newsl 2025 Mar;50:34-40. [doi: 10.1016/j.clinmicnews.2025.01.004]
3. Kim SH, Jang JY. Correlations between climate change-related infectious diseases and meteorological factors in Korea. J Prev Med Public Health 2010 Sep;43(5):436-444. [doi: 10.3961/jpmph.2010.43.5.436] [Medline: 20959714]
4. Kim MJ. Historical review of leptospirosis in the Korea (1945 - 2015). Infect Chemother 2019 Sep;51(3):315-329. [doi: 10.3947/ic.2019.51.3.315] [Medline: 31583866]
5. Lee SH, Chung BH, Lee WC, Choi IS. Epidemiology of hemorrhagic fever with renal syndrome in Korea, 2001-2010. J Korean Med Sci 2013;28(10):1552. [doi: 10.3346/jkms.2013.28.10.1552]

XSL•FO
RenderX

6. Dantas-Torres F, Chomel BB, Otranto D. Ticks and tick-borne diseases: a One Health perspective. Trends Parasitol 2012 Oct;28(10):437-446. [doi: 10.1016/j.pt.2012.07.003] [Medline: 22902521]
7. Elliott I, Pearson I, Dahal P, Thomas NV, Roberts T, Newton PN. Scrub typhus ecology: a systematic review of Orientia in vectors and hosts. Parasit Vectors 2019 Nov 4;12(1):513. [doi: 10.1186/s13071-019-3751-x] [Medline: 31685019]
8. Zangpo T, Phuentshok Y, Dorji K, et al. Environmental, occupational, and demographic risk factors for clinical scrub typhus, Bhutan. Emerg Infect Dis 2023 May;29(5):909-918. [doi: 10.3201/eid2905.221430] [Medline: 37081000]
9. Kim B, Lee YJ, Choi I, Kang YM, Kwak D, Seo MG. Prevalence and zoonotic potential of pathogens in micromammals (rodents and insectivores) in the Republic of Korea. Acta Trop 2025 Jun;266:107649. [doi: 10.1016/j.actatropica.2025.107649] [Medline: 40355036]
10. Kim DM, Kim KY, Nam HS, Kweon SS, Park MY, Ryu SY. Risk-factors for human infection with Orientia tsutsugamushi: a case-control study in Korea. Clin Microbiol Infect 2008 Feb;14(2):174-177. [doi: 10.1111/j.1469-0691.2007.01901.x] [Medline: 18070125]
11. Ma CJ, Oh GJ, Kang GU, et al. Differences in agricultural activities related to incidence of scrub typhus between Korea and Japan. Epidemiol Health 2017;39:e2017051. [doi: 10.4178/epih.e2017051] [Medline: 29121711]
12. Kang D, Choi J. Bayesian zero-inflated spatio-temporal modelling of scrub typhus data in Korea, 2010-2014. Geospat Health 2018 Nov 9;13(2):215-223. [doi: 10.4081/gh.2018.665] [Medline: 30451461]
13. Kwak J, Kim S, Kim G, Singh VP, Hong S, Kim HS. Scrub typhus incidence modeling with meteorological factors in South Korea. Int J Environ Res Public Health 2015 Jun 29;12(7):7254-7273. [doi: 10.3390/ijerph120707254] [Medline: 26132479]
14. Ni Z, Li S, Xi R, et al. Meteorological factors and normalized difference vegetation index drivers of scrub typhus incidence in Shandong Province based on a 16-year time-frequency analysis. BMC Public Health 2025 Feb 24;25(1):752. [doi: 10.1186/s12889-025-21987-y] [Medline: 39994615]
15. Lin F, Pan K, Li J, et al. Assessment of influential factors and risk area prediction of scrub typhus in Gannan region, China, based on ecological niche modelling and geographical detector. Acta Trop 2025 Aug;268:107726. [doi: 10.1016/j.actatropica.2025.107726] [Medline: 40609850]
16. Seo MG, Song BG, Kim TK, et al. Nationwide incidence of chigger mite populations and molecular detection of Orientia tsutsugamushi in the Republic of Korea, 2020. Microorganisms 2021 Jul 22;9(8):1563. [doi: 10.3390/microorganisms9081563] [Medline: 34442642]
17. Jin HS, Chu C, Han DY. Spatial distribution analysis of scrub typhus in Korea. Osong Public Health Res Perspect 2013 Feb;4(1):4-15. [doi: 10.1016/j.phrp.2012.12.007] [Medline: 24159523]
18. Kumar M, Valiaveetil KA, Chandran J, et al. Scrub typhus: a spatial and temporal analysis from South India. J Trop Pediatr 2022 Jun 6;68(4):fmac054. [doi: 10.1093/tropej/fmac054] [Medline: 35903921]
19. Roberts T, Parker DM, Bulterys PL, et al. A spatio-temporal analysis of scrub typhus and murine typhus in Laos; implications from changing landscapes and climate. PLoS Negl Trop Dis 2021 Aug;15(8):e0009685. [doi: 10.1371/journal.pntd.0009685] [Medline: 34432800]
20. Kim S, Kim Y. Hierarchical Bayesian modeling of spatio-temporal patterns of scrub typhus incidence for 2009–2013 in South Korea. Appl Geogr 2018 Nov;100:1-11. [doi: 10.1016/j.apgeog.2018.08.008]
21. Kim J, Vounatsou P, Chun BC. Changes in seasonality and sex ratio of scrub typhus: a case study of South Korea from 2003 to 2019 based on wavelet transform analysis. BMC Infect Dis 2024 Sep 28;24(1):1066. [doi: 10.1186/s12879-024-09858-0] [Medline: 39342094]
22. MMWR weeks. Centers for Disease Control and Prevention. 2013. URL: https://stacks.cdc.gov/view/cdc/22305 [accessed 2023-04-18]
23. Statistics Korea. Statistical Geographic Information Service [Article in Korean]. URL: https://sgis.kostat.go.kr [accessed 2023-07-10]
24. Statistics Korea. Korean Statistical Information Service. URL: https://kosis.kr [accessed 2023-04-20]
25. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. Ecol Modell 2006 Jan;190(3-4):231-259. [doi: 10.1016/j.ecolmodel.2005.03.026]
26. Ramsden E. The urban animal: population density and social pathology in rodents and humans. Bull World Health Organ 2009 Feb;87(2):82. [doi: 10.2471/blt.09.062836] [Medline: 19274352]
27. Lu L, Sun Z, Qimuge E, et al. Using remote sensing data and species–environmental matching model to predict the potential distribution of grassland rodents in the Northern China. Remote Sens (Basel) 2022;14(9):2168. [doi: 10.3390/rs14092168]
28. National Ecosystem Survey data. Environment Bigdata Platform. URL: https://www.bigdata-environment.kr [accessed 2023-07-31]
29. The spatial distribution of population density in 2018, South Korea. WorldPop Hub. URL: https://hub.worldpop.org/geodata/summary?id=42017 [accessed 2023-08-01]
30. Busetto L, Ranghetti L. MODIStsp: an R package for automatic preprocessing of MODIS Land Products time series. Comput Geosci 2016 Dec;97:40-48 [FREE Full text] [doi: 10.1016/j.cageo.2016.08.020]
31. National Spatial Data Infrastructure Portal. Ministry of Land, Infrastructure, and Transport. URL: https:// [accessed 2023-08-01]

XSL•FO
RenderX

32. Lee SU. Epidemiologic characteristics of scrub typhus on Jeju Island. Epidemiol Health 2017;39:e2017039. [doi: 10.4178/epih.e2017039] [Medline: 28823118]
33. Seo CW. Analysis of factors related to regional occurrence distribution of scrub typhus: 2012~2016. Korean J Clin Lab Sci 2019 Dec 31;51(4):420-427. [doi: 10.15324/kjcls.2019.51.4.420]
34. Stefanoff P, Rosinska M, Samuels S, White DJ, Morse DL, Randolph SE. A national case-control study identifies human socio-economic status and activities as risk factors for tick-borne encephalitis in Poland. PLoS One 2012;7(9):e45511. [doi: 10.1371/journal.pone.0045511] [Medline: 23029063]
35. The R Project for Statistical Computing. R Foundation for Statistical Computing. 2024. URL: https://www.R-project.org
36. ArcGIS pro. 3.3.1 ed. Esri. 2024. URL: https://pro.arcgis.com/en/pro-app/3.3/get-started/get-started.htm [accessed 2025-08-09]
37. 2015 Population and Housing Census. Statstics Korea.: Daejeon, Republic of Korea: Statistics Korea; 2016. URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1502&conn_path=I2 [accessed 2025-04-16]
38. George SL. Identification and assessment of prognostic factors. Semin Oncol 1988 Oct;15(5):462-471. [Medline: 3051402]
39. Gómez-Rubio V. Bayesian Inference with INLA: CRC Press; 2020. [doi: 10.1201/9781315175584]
40. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested laplace approximations. Journal of the Royal Statistical Society Series B 2009 Apr 1;71(2):319-392. [doi: 10.1111/j.1467-9868.2008.00700.x]
41. Khana D, Rossen LM, Hedegaard H, Warner M. A Bayesian spatial and temporal modeling approach to mapping geographic variation in mortality rates for subnational areas with R-INLA. J Data Sci 2018 Jan;16(1):147-182. [Medline: 29520299]
42. Musenge E, Chirwa TF, Kahn K, Vounatsou P. Bayesian analysis of zero inflated spatiotemporal HIV/TB child mortality data through the INLA and SPDE approaches: applied to data observed between 1992 and 2010 in rural North East South Africa. Int J Appl Earth Obs Geoinf 2013 Jun;22(100):86-98. [doi: 10.1016/j.jag.2012.04.001] [Medline: 24489526]
43. Feng CX. A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. J Stat Distrib Appl 2021;8(1):8. [doi: 10.1186/s40488-021-00121-4] [Medline: 34760432]
44. Asmarian N, Ayatollahi SMT, Sharafi Z, Zare N. Bayesian spatial joint model for disease mapping of zero-inflated data with R-INLA: a simulation study and an application to male breast cancer in Iran. Int J Environ Res Public Health 2019 Nov 13;16(22):4460. [doi: 10.3390/ijerph16224460] [Medline: 31766251]
45. Nazia N, Butt ZA, Bedard ML, Tang WC, Sehar H, Law J. Methods used in the spatial and spatiotemporal analysis of COVID-19 epidemiology: a systematic review. Int J Environ Res Public Health 2022 Jul 6;19(14):8267. [doi: 10.3390/ijerph19148267] [Medline: 35886114]
46. Blangiardo M, Cameletti M. Spatial and Spatio‐temporal Bayesian Models with R‐INLA: John Wiley & Sons; 2015. [doi: 10.1002/9781118950203]
47. Gelman A, Hwang J, Vehtari A. Understanding predictive information criteria for Bayesian models. Stat Comput 2014 Nov;24(6):997-1016. [doi: 10.1007/s11222-013-9416-2]
48. W-s K, E-h S, Lee H, Hwang T, Kim H, Lee N, et al. Time-spatial distribution of scrub typhus and its environmental ecology. J Korea Geo Soc 2007;42(6):863-878 [FREE Full text]
49. Charron DF. Potential impacts of global warming and climate change on the epidemiology of zoonotic diseases in Canada. Can J Public Health 2002;93(5):334-335. [doi: 10.1007/BF03404563] [Medline: 12353451]
50. Candasamy S, Ayyanar E, Paily K, Karthikeyan PA, Sundararajan A, Purushothaman J. Abundance & distribution of trombiculid mites & Orientia tsutsugamushi, the vectors & pathogen of scrub typhus in rodents & shrews collected from Puducherry & Tamil Nadu, India. Indian J Med Res 2016 Dec;144(6):893-900. [doi: 10.4103/ijmr.IJMR_1390_15] [Medline: 28474626]
51. Lee IY, Kim HC, Lee YS, et al. Geographical distribution and relative abundance of vectors of scrub typhus in the Republic of Korea. Korean J Parasitol 2009 Dec;47(4):381-386. [doi: 10.3347/kjp.2009.47.4.381] [Medline: 19967086]
52. Lin PR, Tsai HP, Weng MH, et al. Field assessment of Orientia tsutsugamushi infection in small mammals and its association with the occurrence of human scrub typhus in Taiwan. Acta Trop 2014 Mar;131:117-123. [doi: 10.1016/j.actatropica.2013.11.029] [Medline: 24361181]
53. Wei Y, Huang Y, Li X, et al. Climate variability, animal reservoir and transmission of scrub typhus in Southern China. PLoS Negl Trop Dis 2017 Mar;11(3):e0005447. [doi: 10.1371/journal.pntd.0005447] [Medline: 28273079]
54. Sharma PK, Ramakrishnan R, Hutin YJF, et al. Scrub typhus in Darjeeling, India: opportunities for simple, practical prevention measures. Trans R Soc Trop Med Hyg 2009 Nov;103(11):1153-1158. [doi: 10.1016/j.trstmh.2009.02.006] [Medline: 19286238]
55. Luo Y, Zhang L, Lv H, et al. How meteorological factors impacting on scrub typhus incidences in the main epidemic areas of 10 provinces, China, 2006–2018. Front Public Health 2022 Oct 19;10:992555. [doi: 10.3389/fpubh.2022.992555]
56. Devamani CS, Schmidt WP, Ariyoshi K, Anitha A, Kalaimani S, Prakash JAJ. Risk factors for scrub typhus, murine typhus, and spotted fever seropositivity in urban areas, rural plains, and peri-forest hill villages in South India: a cross-sectional study. Am J Trop Med Hyg 2020 Jul;103(1):238-248. [doi: 10.4269/ajtmh.19-0642] [Medline: 32458785]
57. Park JH, Gill B, Acharya D, Yoo SJ, Lee K, Lee J. Seroprevalence and factors associated with scrub typhus infection among forestry workers in national park offices in South Korea. IJERPH 2021 Mar 18;18(6):3131. [doi: 10.3390/ijerph18063131]

XSL•FO
RenderX

58. Kweon SS, Choi JS, Lim HS, et al. Rapid increase of scrub typhus, South Korea, 2001-2006. Emerg Infect Dis 2009 Jul;15(7):1127-1129. [doi: 10.3201/eid1507.080399] [Medline: 19624938]
59. Division of Zoonotic and Vector Borne Disease Control. Guidelines for Managing Tick and Rodent-Borne Infectious Diseases in 2022 [Article in Korean]: Korea Disease Control and Prevention Agency (KDCA); 2022.
60. Kim S, Kim JS, Lee H. Epidemiological characteristics of scrub typhus in Korea, 2009. Osong Public Health Res Perspect 2010 Dec;1(1):55-60. [doi: 10.1016/j.phrp.2010.12.012] [Medline: 24159441]
61. Velay A, Baquer F, Brunet J, et al. Infectious risks associated with outdoor sports activities. Infect Dis Now 2024 Jun;54(4S):104862. [doi: 10.1016/j.idnow.2024.104862] [Medline: 38845331]
62. Kuo CC, Lee PL, Chen CH, Wang HC. Surveillance of potential hosts and vectors of scrub typhus in Taiwan. Parasit Vectors 2015 Dec 1;8:611. [doi: 10.1186/s13071-015-1221-7] [Medline: 26626287]

Abbreviations

CrI: credible interval
CV: coefficient of variance
DEM: digital elevation model
EVI: enhanced vegetation index
HCT: hot and cold spot trend
iCAR: intrinsic conditional auto-regressive
***ICD-10*-CM code:** *International Classification of Diseases, Tenth Revision*, Clinical Modification code
INLA: Integrated Nested Laplace Approximation
LST: land surface temperature
MaxEnt: maximum entropy model
MODIS: moderate resolution imaging spectroradiometer
NHIS: National Health Insurance Service
ST: scrub typhus
WAIC: Watanabe-Akaike information criterion

Edited by A Mavragani; submitted 25.12.24; peer-reviewed by G Xu, K Lee, KD Min; revised version received 06.07.25; accepted 07.07.25; published 10.09.25.

Please cite as:
Kim J, Vounatsou P, Chun BC
Distribution and Risk Factors of Scrub Typhus in South Korea, From 2013 to 2019: Bayesian Spatiotemporal Analysis
JMIR Public Health Surveill 2025;11:e68437
URL: https://publichealth.jmir.org/2025/1/e68437
doi:10.2196/68437

©Jeehyun Kim, Penelope Vounatsou, Byung Chul Chun. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying the Factors Associated With Spatial Clustering of Incident HIV Infection Cases in High-Prevalence Regions: Quantitative Geospatial Study

Qiyu Zhu[1,2*], MPH; Chunnong Jike[3*], MD; Chengdong Xu[4*], PhD; Shu Liang[5*], PhD; Gang Yu[3], MD; Dan Yuan[5], MSc; Hong Mai[6], MD; Yiping Li[5], MD; Lin Xiao[3], MD; Ju Wang[3], MSc; Hong Yang[5], MSc; Fengshun Yuan[5], PhD; Jing Hong[3], MD; Muga Mao[3], MD; Maogang Shen[3], MD; Jing Liu[1], MSc; Lin He[1], PhD; Yuehua Wang[1], PhD; Huanyi Cheng[1], BA; Peng Guan[2], PhD; Yan Jiang[1], PhD; Mengjie Han[1], MPH; Cong Jin[1], PhD; Zhongfu Liu[1], PhD

[1]National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

[2]China Medical University, Shenyang, China

[3]Liangshan Prefecture Center for Disease Control and Prevention, Xichang, China

[4]Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

[5]Sichuan Provincial Center for Disease Control and Prevention, Chengdu, China

[6]The First People's Hospital of Yi Autonomous Prefecture of Liangshan, Xichang, China

[*]these authors contributed equally

Corresponding Author:
Cong Jin, PhD
National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases
National Center for AIDS/STD Control and Prevention
Chinese Center for Disease Control and Prevention
155 Changbai Road
Changping District
Beijing, 102206
China
Phone: 86 1058900922
Email: jinc@chinaaids.cn

Abstract

Background: Incident HIV infection is a critical indicator of an ongoing epidemic, particularly in high-burden regions such as Liangshan Yi Autonomous Prefecture in China, where HIV prevalence exceeds 1% in 4 key counties (Butuo, Zhaojue, Meigu, and Yuexi). Identifying spatial clusters and drivers of recent infections is essential for implementing targeted interventions. Despite advancements in geospatial analyses of HIV prevalence, studies identifying drivers of incident HIV clustering remain limited, especially in low-resource settings.

Objective: This study aims to identify spatial clusters of recent HIV infections and investigate potential driving factors in 4 key counties of the Liangshan Yi Autonomous Prefecture to inform targeted intervention strategies.

Methods: From November 2017 to June 2018, we identified 246 (4.42%) recent HIV infection cases from 5555 newly diagnosed cases through expanded testing of the whole population in 4 key counties of Liangshan Yi Autonomous Prefecture. Recent infection cases were confirmed using limiting antigen avidity enzyme immunoassays or documented seroconversion within 6 months. The spatial distribution of incident HIV infection cases was analyzed using kernel density. Potential drivers, including population density, HIV prevalence, elevation, nighttime light index, urban proximity, and antiretroviral therapy (ART) coverage, were analyzed. The spatial lag regression model was used to identify factors associated with clustering of recent infection cases. The Geodetector *q*-statistic was used to quantify nonlinear interactive effects among these factors.

Results: Significant spatial autocorrelation was observed in the distribution of recent HIV cases (Moran *I*=0.11; *P*<.01). Six spatial clusters were identified, and all were located near urban centers or major roads. Furthermore, 5 factors were identified by the spatial lag regression model as being significantly correlated with the clustering of recent HIV infection cases, including population density (β=0.59; *P*<.001), HIV prevalence (β=0.02; *P*<.001), distance to local urban area (β=–3.10; *P*=.01), SD of

XSL•FO
RenderX

elevation (β=–0.15; *P*=.02), and ART coverage rate (β=183.80; *P*<.01). Geodetector analysis revealed strong interactive effects among these 5 factors, with population density and HIV prevalence exhibiting the largest interactive effect (*q*=0.69).

Conclusions: This study reveals that besides HIV prevalence, urbanization-related factors (population density and proximity to urban area) and transportation accessibility drive incident HIV clustering in Liangshan Yi Autonomous Prefecture. Paradoxically, higher ART coverage was associated with increased transmission, suggesting the need for integrated prevention strategies beyond ART expansion. Furthermore, the township-level geospatial approach provides a valuable model for pinpointing transmission hot spots and tailoring interventions in high-burden regions globally.

(JMIR Public Health Surveill 2025;11:e75291) doi:10.2196/75291

KEYWORDS

HIV; disease transmission; spatial analysis; geographic mapping; public health

Introduction

Background

Incident HIV infection is a direct indicator of an ongoing epidemic. Collecting detailed information on new infection cases has been advocated by the World Health Organization and the Joint United Nations Programme on HIV/AIDS, especially in high-burden areas, to inform trends of HIV transmission. Several laboratory-based assays, such as limiting antigen avidity enzyme immunoassay, have been developed to identify recent HIV infection from newly diagnosed cases and have been proven valid with good credibility [1,2]. The applications of assays for recent HIV infection include providing national estimates of incidence, monitoring HIV incidence in key or sentinel populations, assessing the impact of interventions, and identifying clusters and populations of current epidemics [3-7].

An emerging application of recent HIV infection testing is its integration with geospatial analysis to detect geographic clusters of new transmission, thereby supporting targeted interventions. In Malawi, 3 hot spots of recent infection cases were identified with facility-based HIV testing data [8]. The cluster also indicated service delivery gaps and required an immediate public health response [9]. In several provinces of China, recent infection testing was performed among newly diagnosed HIV cases, and spatial autocorrelation and spatial scan statistics were used to detect hot spots [10-13]. Moreover, in a border prefecture of China, a Bayesian spatiotemporal model was used to estimate the dynamic spatial relative risk of recent infections [14]. Most existing studies focused on describing the spatial distribution, estimating the relative risks, and detecting the hot spots of recent infection cases. However, studies specifically analyzing the potential driving factors of the spatial clustering of incident HIV infection cases remain limited, particularly in low- and middle-income settings and regions with high HIV burden.

Liangshan Yi Autonomous Prefecture, located in Sichuan province in the southwest of China, is one region that has been severely affected by HIV [15,16]. The HIV infection rate among residents in Liangshan Yi Autonomous Prefecture was 0.84% in 2019, with the HIV infection rate in 4 key counties of Butuo, Zhaojue, Meigu, and Yuexi in Liangshan Yi Autonomous Prefecture exceeding 1% [17]. The key to controlling the HIV epidemic in this region is to reveal critical driving factors affecting HIV transmission and implement targeted interventions.

This Study

In this study, we tried to focus on 246 recent HIV infection cases identified from 5555 newly diagnosed HIV cases through an expanded HIV testing from November 2017 to June 2018 in the general population in the 4 key counties. Furthermore, we retrospectively conducted multiple geospatial analyses to identify clusters and high-risk areas of recent HIV infection, aiming at revealing potential factors driving the clustering of incident HIV infection cases in this region and consequently informing targeted interventions to effectively curb the local epidemic.

Methods

Study Scheme

The newly diagnosed HIV cases in this study were identified through expanded HIV testing in residents of the 4 key counties of Butuo, Zhaojue, Yuexi, and Meigu in Liangshan Yi Autonomous Prefecture from November 2017 to June 2018; the testing was incorporated into a physical examination provided as a public health service covering more than 95% of the residents. Participating individuals were informed of HIV testing and could choose to reject the test without impacting other parts of the physical examination service. The individuals signed a physical examination form to approve all testing. For individuals who agreed to HIV testing, a rapid test for detecting HIV specific IgM and IgG (Wondfo) antibodies was performed. Individuals with positive screening results were asked to undergo a confirmatory immunoblot test (MP Biomedicals), and the individuals with positive immunoblot test results were registered in the national HIV database.

Identification of Recent HIV Infection Cases

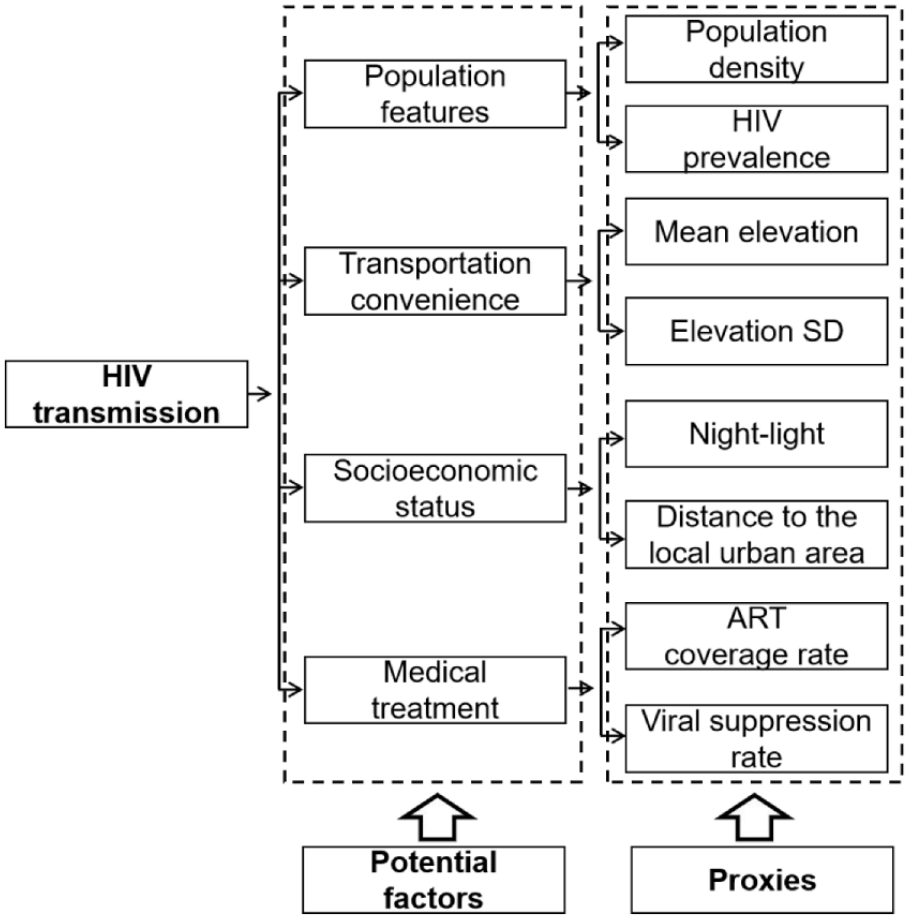
Among newly diagnosed HIV cases, recent HIV infection cases were identified in 2 ways. Most (196/246, 79.7%) of the recent HIV infection cases were identified by a limiting antigen avidity enzyme immunoassay (Beijing Kinghawk Pharmaceutical Co, Ltd). According to the manufacturer's instructions, the mean duration of recent infection of this assay was 130 days. Recent HIV infection cases were also identified through HIV testing history of newly diagnosed HIV cases as having a known negative HIV testing result within 6 months before a positive HIV testing result during the survey period.

XSL•FO
RenderX

Sources of Data

In this study, we analyzed potential factors related to the spatial clustering of recent infection cases. On the basis of previous studies [18,19] and local contextual relevance, we categorized these factors into four aspects: (1) population features, (2) transportation convenience, (3) socioeconomic status, and (4) medical treatment (Figure 1).

Figure 1. Potential driving factors and their proxy factors of HIV transmission. Population features, transportation convenience, socioeconomic status, and medical treatment are 4 potential driving factors of HIV transmission. Population density, HIV prevalence, mean elevation, elevation SD, night-light, distance to local urban center, antiretroviral therapy (ART) coverage rate, and viral suppression rate serve as proxy variables of potential driving factors.



The first aspect was population features. Population density (persons/km²) and HIV prevalence were included to reflect the level of population aggregation and the underlying burden of infection, respectively. Information for calculating population density in the township was obtained from the local statistics bureau. HIV prevalence data were provided by the prefecture center for disease control and prevention.

The second aspect was transportation convenience. Previous studies have shown that transportation accessibility can influence HIV transmission and outcomes [20,21]. Given the mountainous terrain in Liangshan Yi Autonomous Prefecture, variations in elevation may lead to disparities in transportation infrastructure and accessibility. Geographic information on elevation was obtained from the National Catalogue Service for Geographic Information. The mean and SD of elevation were calculated using ArcGIS software (Esri) as proxy variables of transportation convenience.

The third aspect was socioeconomic status. Due to data unavailability, we included the nighttime light index and distance to the local urban area as proxies for socioeconomic status. Nighttime light data were sourced from a platform developed by the Chinese Academy of Sciences, and township-level distances to local urban centers were calculated using ArcGIS software.

The fourth aspect was medical treatment. HIV antiretroviral therapy (ART) coverage and viral suppression are 2 important indicators of HIV treatment. These indicators reflect the availability and effectiveness of HIV treatment services. Data on ART coverage and viral suppression rates were obtained from the local ART quality control center.

Furthermore, demographic and location information of recent HIV infection cases were collected from the national HIV/ AIDS database. Recent HIV infection cases were mapped using village midpoint coordinates extracted from resident address information.

Geographic and Statistical Analyses

All geographic maps were created using ArcGIS (version 10.6). General Moran *I* index was calculated using ArcGIS to determine the spatial autocorrelation of township-level

XSL•FO
RenderX

aggregated number of recent infection cases. Kernel density estimation was performed with ArcGIS to construct a smooth surface for the geographic distribution of recent HIV infection cases. Ordinary least squares regression (OLSR) and spatial lag regression (SLR) were performed using GeoDa (version 1.12.1.139) software to determine the linear correlation between the kernel density of recent HIV infection cases and potential factors. Geodetector *q*-statistic was used to investigate the interactive effect (linear or nonlinear) of a single or every 2 of the proxies on the kernel density of recent HIV infection cases. In general Moran *I* index, OLSR, SLR, and Geodetector *q*-statistic, the statistical unit was township, which means that township-level data were used for analysis. Statistical analyses were performed using SPSS Statistics (version 23.0; IBM Corp) statistical analysis software package. Frequencies of variables were compared using the Fisher exact chi-square test. Considering heterosexual transmission has become the main route of HIV infection, the demographic characteristics between different genders were compared.

Kernel Density Estimation

Kernel density estimation (equation 1) is a process of inferring the probability density at each location, that is, constructing the probability density surface, based on the locations of residents (the points), and it can visualize the density of recent cases of infection. This method has been widely used to construct a smooth surface for the geographic distribution of HIV cases [22,23]. In this study, kernel density estimation was performed using geographic information of recent cases of HIV infection within the study area. The analysis was conducted in ArcGIS software with default parameters, and the study area boundary was used to define the extent of the analysis. The kernel density was calculated as follows [24,25]:

For $dist_i$<*radius*

where $dist_i$ specifies the distance between the point of the recent HIV infection case and location (x, y), *n* represents the total number of recent HIV infection cases, *radius* is the kernel bandwidth, and pop_i is the number of recent HIV infection cases in point . To enable township-level analysis, the average kernel density value within each township was further calculated using ArcGIS.

General Moran I Index

The general Moran *I* index (equation 2) was calculated using ArcGIS software to determine the spatial autocorrelation of township-level aggregated numbers and kernel density of recent infection cases. The general Moran *I* index was calculated as follows [26]:

where *N* represents the number of townships in the study area (*N*=151); w_{ij} denotes the spatial weight between townships *i* and *j*; x_i and x_j are the values of kernel density in the township *i* and *j*, respectively; and specifies the mean kernel density in the study area. The value of the general Moran *I* index ranged between –1 and 1 (positive autocorrelation: Moran *I*>0; negative autocorrelation: Moran *I*<0).

Linear Regression Model

We used both the OLSR model (equation 3) and the SLR model (equation 4) to identify factors associated with the clustering of recent infection cases. These 2 models assumed linear relationships between the dependent variable and explanatory variables. The SLR model further incorporated spatial dependence into the regression model with a spatially lagged dependent variable, which accounted for the fact that the dependent variable (*y*) may be spatially autocorrelated [27]. The fitness of these 2 models was compared to determine which model was most appropriate for analysis. OLSR and SLR were denoted as follows:

$$y = X\beta + \varepsilon \quad \textbf{(3)}$$

$$y = \rho Wy + X\beta + \mu \quad \textbf{(4)}$$

In both equations, *y* is a vector of kernel density of recent infection cases in each township (N=151), *X* is a vector of explanatory variables, and β is the regression coefficient of the explanatory variable. In the OLSR model (equation 3), ε represents the random error. In the SLR model (equation 4), ρ is the coefficient of the spatial lag parameter and represents the level of autocorrelation in the lag term, W_y is a vector of spatial weights, and μ specifies the nonspatial random error. Queen contiguity was used to assign spatial weights to townships (nonzero weights for townships that share a common edge or vertex and zero weights to the others).

Geodetector

Geodetector (Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences), a statistical tool, can measure the spatial stratified heterogeneity and quantify the influencing power of a single or every 2 of the proxies included. Geodetector assumes that explanatory variable and dependent variable would be coupled (linear or nonlinear) in strata if they are correlated, and the nonlinear assumption is its advantage. Thus, we used Geodetector to quantify the influencing power of a single or every 2 proxies with the *q*-statistic value. The *q*-statistic value was expressed as follows [28]:

where *h* (*h*=1,2,...*L*) is the stratification of a single factor (*X*) or every 2 interacting factors, N_h and *N* represent the number of townships in the stratum *h* (h=1,2,...,*L*) and the entire area separately (the stratification was calculated by natural break), and and denote the variances of the kernel density in the stratum *h* and the whole study area, respectively. The *q*-statistic value ranged between 0 and 1, and a larger value indicated higher influencing power.

Ethical Considerations

This study, which involved analyzing the data on recent HIV infection cases identified through expanded HIV testing in Liangshan Yi Autonomous Prefecture, was approved by the institutional review board of the National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control

XSL•FO
RenderX

and Prevention. This expanded HIV testing was incorporated into the local physical examination. All participants signed informed consent forms at the time of their examinations. The physical examination was free, and no financial compensation was provided to participants. No personally identifiable information was present in the data analyzed for this study.

Results

Demographic Characteristics of Recent HIV Infection Cases

A total of 246 (4.4%) individuals who were recently infected with HIV were identified from 5555 newly diagnosed cases in 4 key counties in Liangshan Yi Autonomous Prefecture. A total of 217 (88.2%) of them provided complete demographic information and were included in the analysis of demographic characteristics. Among the 217 individuals who were newly infected (Table 1), 92 (42.4%) were male and 125 (57.6%) were female. In both male and female, individuals aged between 30 and 44 years, those who were married, and those who had not attended school occupied the largest proportion. A significant difference was seen in the transmission modes between different genders (P<.001). A large proportion of individuals who were recently infected acquired HIV through heterosexual transmission, and this number was 87.2% (109/125) and 62% (57/92) in female individuals and male individuals, respectively.

Table 1. Demographic characteristics of individuals who were recently infected with HIV in 4 key counties of Liangshan Yi Autonomous Prefecture, China, from November 2017 to June 2018.

Demographic characteristic	Total (N=217), n (%)	Male (n=92), n (%)	Female (n=125), n (%)	*P* value[a]
Age (y)				.33
1.5-14	15 (6.9)	5 (5.4)	10 (8)	
15-29	69 (31.8)	24 (26.1)	45 (36)	
30-44	103 (47.5)	49 (53.3)	54 (43.2)	
≥45	30 (13.8)	14 (15.2)	16 (12.8)	
Marital status				.45
Single	56 (25.8)	25 (27.2)	31 (24.8)	
Married	135 (62.2)	53 (57.6)	82 (65.6)	
Divorced or widowed	23 (10.6)	13 (14.1)	10 (8)	
Unknown	3 (1.4)	1 (1.1)	2 (1.6)	
Educational attainment				.15
No schooling	143 (65.9)	54 (58.7)	89 (71.2)	
Primary school	66 (30.4)	34 (37)	32 (25.6)	
Middle school and above	8 (3.7)	4 (4.3)	4 (3.2)	
Infection routes				<.001
Heterosexual contact	166 (76.5)	57 (62)	109 (87.2)	
Drug injection	36 (16.6)	30 (32.6)	6 (4.8)	
Others	15 (6.9)	5 (5.4)	10 (8)	

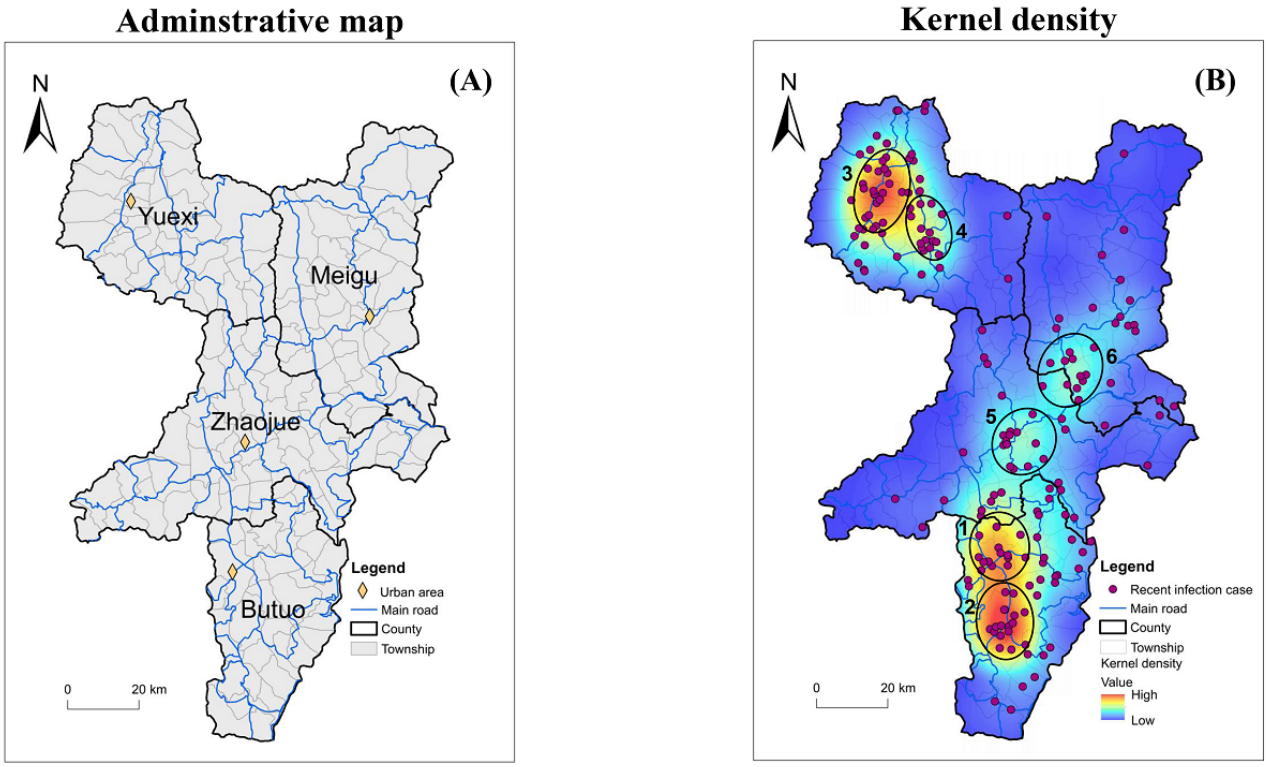
[a]Statistical significance was assessed between male and female individuals by the Fisher exact test.

Spatial Distribution and Clusters of Recent HIV Infection Cases

The general Moran I value of township-level kernel density of recent HIV infection cases was 0.11 (P<.01), which indicated significant positive spatial autocorrelation. To investigate spatial distribution features of recent HIV infection cases in the 4 neighboring counties (Figure 2A), the kernel density of recent HIV cases was calculated and mapped. We found 6 main clusters of recent HIV cases, and all 6 clusters were located at the intersection of main traffic roads (Figure 2B). Except for cluster 6 at the border of Zhaojue and Meigu, which is a transportation portal between Zhaojue and Meigu, other clusters were located around urban areas in the county. Our spatial distribution analyses also indicated spatial variation of recent HIV infection.

XSL•FO
RenderX

Figure 2. Spatial distribution and high-risk areas of recent HIV infection cases in 4 key counties of Liangshan Yi Autonomous Prefecture, China, from November 2017 to June 2018. (A) Administrative map of the 4 key counties. Blue lines indicate main traffic roads, and yellow diamonds indicate the local urban areas of each county. (B) The spatial distribution of recent HIV infection cases is shown using a kernel density map. The color-code shows the density of recent HIV infection cases per square map unit. The black circles show clusters of recent HIV infection cases.



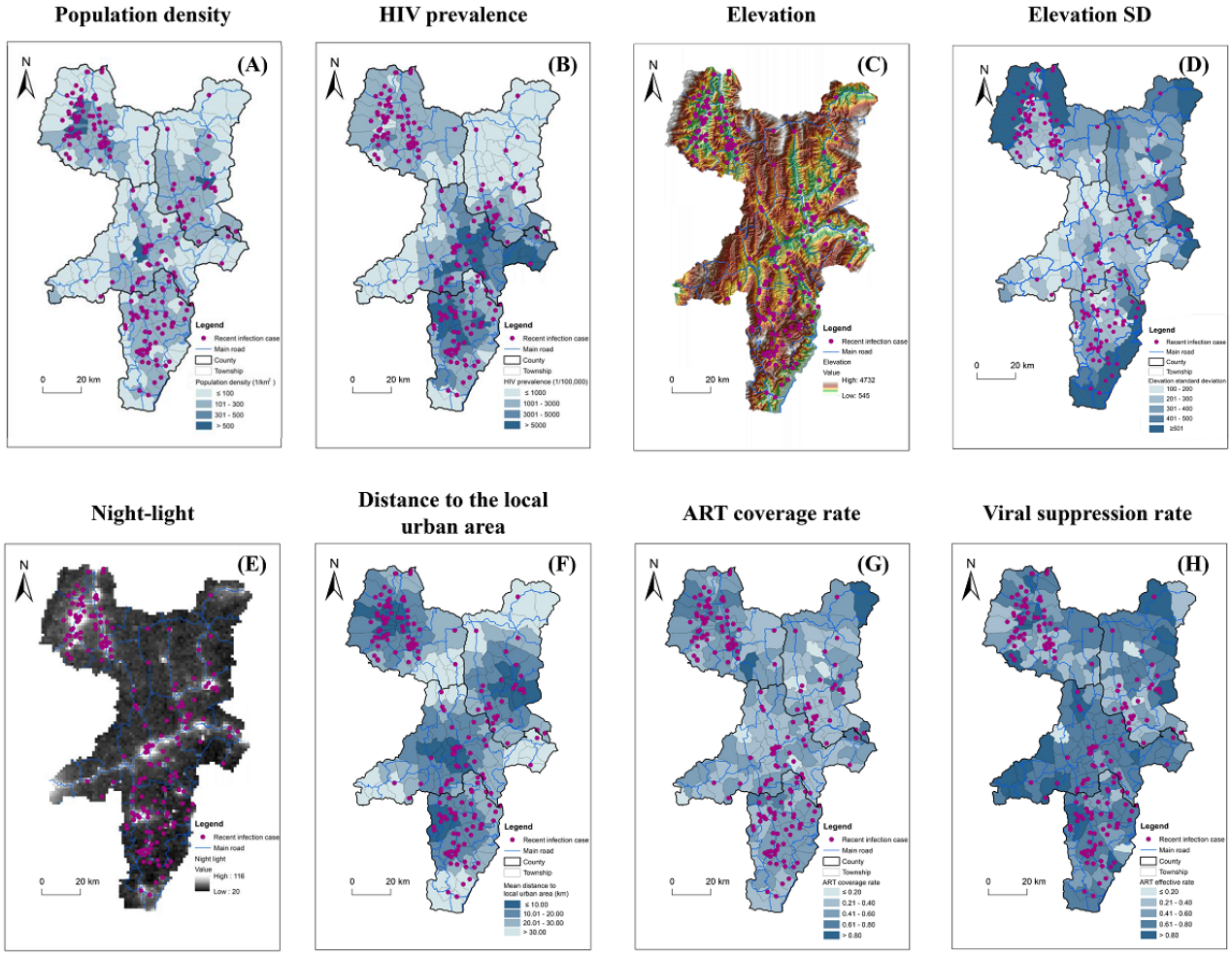
Factors Associated With the Clustering of Recent HIV Infection Cases

The general Moran *I* value of township-level kernel density of recent infection cases was 0.76 ($P<.001$), which indicated significant positive spatial autocorrelation. Therefore, we further used the spatial regression analysis, SLR, which accounts for the fact that the dependent variable (*y*) may be spatially autocorrelated, to perform quantitative analysis and identify factors significantly associated with clustering of recent HIV infection cases and compare it with a nonspatial model, OLSR.

We investigated the factors that might impact the clustering of recent HIV infection cases while considering the unique geographic, demographic, and HIV epidemic features in the area. Population density and HIV prevalence rate of the township were incorporated into the analysis, and we found that all clusters were located in areas with high population density and high HIV prevalence (Figures 3A and 3B). Because the 4 key counties are located in a mountainous area, the altitude was taken into account. The topographical map showed that, except for cluster 2 in Butuo and cluster 4 in Yuexi, all other clusters were located in areas at a lower altitude (Figure 3C). Elevation SD indicates the steepness of local terrain, which can impact the convenience of transportation to a certain extent (Figure 3D). Nighttime light index was used to represent the general economic status at the township level. Because the 4 key counties are lower-income areas, most areas are dark. We found that, except for cluster 2 in Butuo, all other clusters were located in areas with bright night-light, indicating comparatively better economic status (Figure 3E). The distance to the local urban areas of the counties was also mapped, and it showed that all clusters were located within 20 km of the urban areas (Figure 3F). The impact of ART coverage rate and viral suppression rate of individuals infected with HIV before the survey was also examined. The results showed that all clusters were mainly located in areas with comparatively higher ART coverage rates (Figure 3G). Clear patterns related to the distribution relationship between the clusters and the viral suppression rate were not noted (Figure 3H).

XSL•FO
RenderX

Figure 3. Maps of 8 potential factors contributing to the clustering of recent HIV infection cases in 4 key counties of the Liangshan Yi Autonomous Prefecture, China, from November 2017 to June 2018. (A) Map of population density at the township level. The color intensity shows the number of individuals per square km. (B) Map of HIV prevalence at the township level. The color intensity shows the level of HIV prevalence. (C) Elevation map. The color-code shows elevation. (D) Elevation SD map. The color-code shows elevation SD. (E) Night-light map. The color-code shows the nighttime light index. (F) Map of distance to local urban center at the township level. The color intensity shows the distance range. (G) Map of antiretroviral therapy (ART) coverage rate at the township level. The color intensity shows the level of ART coverage rate. (H) Map of the viral suppression rate at the township level. The color intensity shows the level of the viral suppression rate.



In addition to visually mapping potential factors, we used a nonspatial OLSR model and an SLR model to perform quantitative analysis and identify proxies significantly associated with clustering of recent HIV infection cases. Geographic kernel density of recent HIV infection cases was used to represent clustering intensity, and the correlation between kernel density of recent infection cases and proxies, including population density, HIV prevalence, mean elevation, elevation SD, night-light, distance to local urban area, ART coverage rate, and viral suppression rate, was analyzed.

In comparison to the nonspatial OLSR analysis, the SLR model fit better, and there was a significant spatial autocorrelation in the township-level kernel density of recent infection cases (ρ=0.72; P<.001; Table 2). In the SLR model, 85% of the spatial variations in kernel density of recent infection cases can be explained by the proxies included (higher than 62% in the nonspatial OLSR model). In the SLR model, at the township level, population density, HIV prevalence, and ART coverage rate had a positive correlation with kernel density of recent infection cases, while elevation SD and distance to local urban areas had a negative correlation.

XSL•FO
RenderX

Table 2. Results from nonspatial and spatial linear models, namely ordinary least squares regression (OLSR) and spatial lag regression (SLR), and proxy factors associated with clustering of recent HIV infection cases in 4 key counties of Liangshan Yi Autonomous Prefecture, China, from November 2017 to June 2018.

Proxy factors	OLSR[a]		SLR[b]	
	β coefficient	*P* value	β coefficient	*P* value
Population density	0.91	<.001	0.59	<.001
HIV prevalence	0.05	<.001	0.02	<.001
Mean elevation	0.03	.58	–0.03	.45
Elevation SD	0.09	.40	–0.15	.02
Night-light	–7.61	.23	–2.93	.45
Distance to the local county urban area	–11.11	<.001	–3.10	.01
ART[c] coverage rate	491.52	<.001	183.80	<.01
Viral suppression rate	11.73	.88	19.73	.68

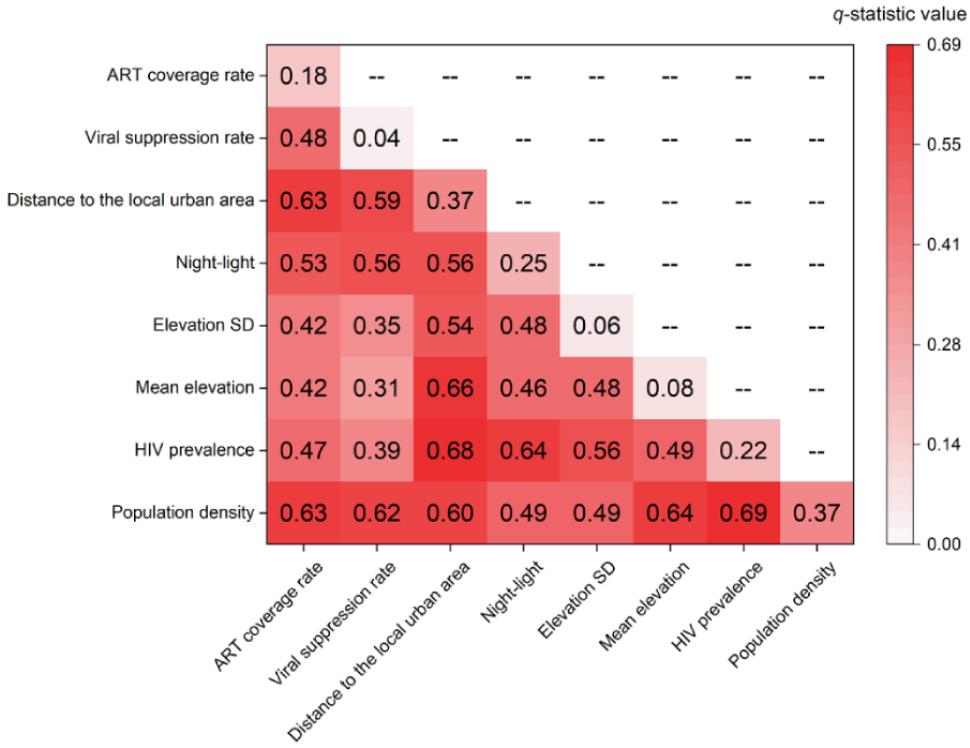
[a]R^2=0.62; *P*<.001; Akaike information criterion=2022.78.

[b]R^2=0.85; ρ=0.72 (represents the level of autocorrelation in the lag term); Akaike information criterion=1911.81; *P*<.001.

[c]ART: antiretroviral therapy.

To further elucidate the nonlinear influencing power of a single and every two proxies on the clustering of recent HIV infection cases, we performed Geodetector analysis. Our results revealed that population density and HIV prevalence showed the highest interactive influence on clustering of recent HIV infection cases, reaching a *q*-statistic value of 0.69 (Figure 4). Furthermore, for every 2 of the 5 proxies, including HIV prevalence, population density, distance to local urban area, SD of elevation, and ART coverage rate, a strong interactive effect was observed through Geodetector analysis (Figure 4). Regarding the influencing power of the single proxy factor, population density and distance to the local urban area had the highest *q*-statistic value of 0.37 (Figure 4).

Figure 4. Interactive effects of 8 potential factors on clustering of recent HIV infection cases in 4 key counties of Liangshan Yi Autonomous Prefecture, China, from November 2017 to June 2018. The color intensity at the intersection of 2 factors shows the extent to which the 2 factors interactively contribute to clustering of recent HIV infection cases. The regions are color-coded from pale red to dark red, and dark red indicates a high q-statistic value. ART: antiretroviral therapy.



XSL•FO
RenderX

Discussion

Principal Findings

Many potential factors could impact local HIV transmission, such as local HIV prevalence, socioeconomic status, transportation convenience, access to HIV-related health care, local culture, etc [19,29-32]. On the basis of the data available at the township level, we categorized the potential driving factors into 4 aspects, which were population features, transportation convenience, socioeconomic status, and access to HIV care. For every potential driving factor that we proposed to study, we tried to use the most appropriate proxy variables for the analyses. Due to a lack of data at the township level for other potential factors, for example, residents' demographic information, access to education, harm reduction, gross domestic product, urbanization situation, and preexposure prophylaxis awareness, these were not included in this study. We also did not include some potential factors that are difficult to quantify, such as attitudes toward casual sex behavior in the Yi ethnic group [19]. However, it is compelling that the factors that we analyzed could explain the clustering of recent HIV infection cases for a high proportion of 85% in the SLR analysis (Table 2). Therefore, these potential factors included in our regression analyses showed a strong explanatory effect on the clustering of recent HIV infection cases.

Our novel geospatial analyses focusing on incident HIV infection allowed us to capture the underlying drivers of the ongoing HIV epidemic. In our study, the geographic distribution of recent HIV infection cases at the township-level has a significant spatial autocorrelation, and the SLR model fitted better than the nonspatial OLSR model, which did not consider spatial features of the area. The SLR model identified 5 factors—population density, HIV prevalence, elevation SD, the distance to the local urban area, and ART coverage rate. In the nonlinear Geodetector analysis, population density and the distance to local urban areas independently contributed to the highest *q*-statistic value of 0.37, and they were the 2 main parameters used to define urbanization [33]. Compared to the nonspatial OLSR model, the SLR model further identified significant factors such as elevation SD, which indicates the steepness of the road, and was used to represent the convenience of transportation. Therefore, our results suggest that areas with higher levels of urbanization and more convenient transportation in the 4 key counties are more inclined to have spatial clustering of incident HIV infection cases. A previous study using spatial clustering analysis of newly diagnosed cases in Mississippi, United States, also showed that urban residence was an important risk factor for HIV acquisition [30] and postulated that HIV spread more rapidly in highly populated urban areas with dense sex work networks [34]. Another study in Lilongwe, Malawi, indicated that due to the proximity to transportation and commercial networks, urbanization may increase the rates of HIV transmission and put women at a higher risk than men [32]. According to a quantitative study in Africa, a 10% increase in urbanization could be associated with a 3.6% increase in adult HIV prevalence [35]. However, while all previous studies used HIV prevalence data for their analysis, our study used HIV incident infection data to reveal the link between urbanization and HIV transmission, an indicator that could serve as a more direct reflection of an ongoing epidemic. Our findings suggest that local governments in areas undergoing urban development should allocate resources toward education on healthy sex behavior to the public and provide related preventive measures to reduce the potential transmission risk of HIV infection.

In our study, the SLR also revealed HIV prevalence and ART coverage rate as significant factors. Meanwhile, the Geodetector analysis showed that both factors had a high interactive effect with population density and distance to the local urban area. Thus, it suggests that these 2 factors may contribute to HIV transmission in Liangshan in a more complex, nonlinear way. It is not surprising that HIV prevalence had a positive relation with HIV transmission. However, it is intriguing that our study found a positive relation between ART coverage and HIV transmission because it has been acknowledged by many studies that expanded ART coverage is associated with decreased HIV transmission [36]. It has been debated whether the preventive effects of ART at an individual level, as shown in the HIV Prevention Trials Network 052 trial in heterosexual couples with HIV serodiscordance, can be replicated at a population level and whether outcomes from mathematical models can be replicated in the real world [37-39]. Early infection, consistent risk behaviors, adherence to therapy, and drug resistance were factors suggested to compromise the preventive effect of expanded antiretroviral coverage [38,39]. In the 4 key counties, there could be higher ART coverage rate in urban areas with better health care services; however, including the context of an extensive sexual network in urban areas and considering the general environment of insufficient detection of HIV infection and poor drug adherence in this area [40], relying solely on expansion of ART may not result in effective reduction in the transmission and incidence of new infections. In addition, the level of pretreatment drug resistance has been moderate in these 4 counties [41]. Therefore, combination prevention strategies, including universal testing, expansion of treatment, increased drug adherence, and other combined targeted measures, are needed to effectively avert the epidemic in the Liangshan area.

Limitations

There are several limitations in our study. The analysis of incident infection cases is generally impacted by the individuals who are newly diagnosed, identified through the local testing program. In this study, the expanded HIV testing covered more than 95% of the residents; therefore, this impact was minimized. However, approximately 5% of the residents were not included in this study due to out-migration for work and population movement, which might still lead to a certain level of bias. It is also noteworthy that the HIV recency assay has a false recent rate (FRR) such that some cases of long-term infection could be misclassified as recent infection. In this study, we excluded HIV cases that received ART and determined the FRR as 1.48% in the 4 key counties. Considering the low FRR, our main findings should not be affected. Furthermore, in modeling the spatial distribution of recent HIV infection cases, kernel density was calculated to reflect the clustering of recent HIV infection cases to address the limited number of recent infections. While kernel density is not a standard epidemiological indicator and its use may affect model interpretation, it serves as a useful

XSL•FO
RenderX

preliminary tool for identifying potential clustering patterns. Future studies with larger datasets and more stable incidence rates are suggested to refine this approach.

Conclusions

Geospatial studies explicitly focused on new HIV infection cases are rare [11]. To the best of our knowledge, our study is the first to use spatial regression analysis to reveal factors associated with clustering of incident HIV infection cases. After revealing 6 clusters of incident HIV infection, we further identified population density, HIV prevalence, elevation SD, distance to local urban area, and ART coverage rate as 5 factors impacting current HIV transmission in the 4 key counties that were severely impacted by the HIV epidemic. Our novel geospatial study at the township level provides a new tool for zooming into a local area and examining unique factors in detail. We hope our analytical tools will be useful in identifying high-risk regions and populations in other areas of high HIV prevalence and thus inform more targeted public health responses, which can be used to avert epidemics.

Acknowledgments

This study was supported by the project of the National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases, Beijing Natural Science Foundation (grant L244069), and National Natural Science Foundation of China (grant numbers 72474226 and 71974199).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

C Jin proposed the research question and designed the study. C Jin and QZ wrote the manuscript. QZ, CX, and SL performed the epidemiology and geographic analyses. C Jike, GY, DY, YL, LX, JW, HY, FY, JH, MM, MS, YW, HC, HM, JL, and LH contributed to laboratory experiments and epidemiology data collection. ZL, MH, YJ, and PG provided expert knowledge. All authors have read and contributed to the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Murphy G, Parry JV. Assays for the detection of recent infections with human immunodeficiency virus type 1. Euro Surveill 2008 Sep 04;13(36):18966 [FREE Full text] [Medline: 18775293]
2. Suligoi B, Regine V, Raimondo M, Rodella A, Terlenghi L, Caruso A, et al. HIV avidity index performance using a modified fourth-generation immunoassay to detect recent HIV infections. Clin Chem Lab Med 2017 Oct 26;55(12):2010-2019. [doi: 10.1515/cclm-2016-1192] [Medline: 28672745]
3. When and how to use assays for recent infection to estimate HIV incidence at a population level. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/44612/9789241501675_eng.pdf [accessed 2025-02-10]
4. Guidelines for second generation HIV surveillance: an update: know your epidemic. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/85511/9789241505826_eng.pdf [accessed 2025-02-10]
5. Currie DW, West CA, Patel HK, Favaloro J, Asiimwe M, Ndagije F, et al. Risk factors for recent HIV infections among adults in 14 countries in Africa identified by population-based HIV impact assessment surveys, 2015-2019. Emerg Infect Dis 2023 Nov;29(11):2325-2334 [FREE Full text] [doi: 10.3201/eid2911.230703] [Medline: 37877591]
6. Srithanaviboonchai K, Yingyong T, Tasaneeyapan T, Suparak S, Jantaramanee S, Roudreo B, et al. Establishment, implementation, initial outcomes, and lessons learned from recent HIV infection surveillance using a rapid test for recent infection among persons newly diagnosed with HIV in Thailand: implementation study. JMIR Public Health Surveill 2024 Dec 26;10:e65124 [FREE Full text] [doi: 10.2196/65124] [Medline: 39724047]
7. Msukwa MT, MacLachlan EW, Gugsa ST, Theu J, Namakhoma I, Bangara F, et al. Characterising persons diagnosed with HIV as either recent or long-term using a cross-sectional analysis of recent infection surveillance data collected in Malawi from September 2019 to March 2020. BMJ Open 2022 Sep 23;12(9):e064707 [FREE Full text] [doi: 10.1136/bmjopen-2022-064707] [Medline: 36153024]
8. Telford CT, Tessema Z, Msukwa M, Arons MM, Theu J, Bangara FF, et al. Geospatial transmission hotspots of recent HIV infection - Malawi, October 2019-March 2020. MMWR Morb Mortal Wkly Rep 2022 Mar 04;71(9):329-334 [FREE Full text] [doi: 10.15585/mmwr.mm7109a1] [Medline: 35239633]
9. Stephens R, Mfungwe H, Chalira D, Luhanga M, Theu J, Thawi R, et al. Notes from the field: public health response to surveillance for recent HIV infections - Malawi, May 2024. MMWR Morb Mortal Wkly Rep 2025 Jun 05;74(20):355-356 [FREE Full text] [doi: 10.15585/mmwr.mm7420a4] [Medline: 40471870]

XSL•FO
RenderX

10. Yang H, Chen M, Hu Y, Xu M, Li Y, Liu L, et al. An assessment of trends in HIV-1 prevalence and incidence and spatio-temporal analyses of HIV-1 recent infection among MSM during the surveillance period between 2018 and 2022 in Sichuan, China. HIV AIDS (Auckl) 2024;16:83-93 [FREE Full text] [doi: 10.2147/HIV.S448096] [Medline: 38464995]
11. Chen M, Ma Y, Chen H, Dai J, Luo H, Yang C, et al. Demographic characteristics and spatial clusters of recent HIV-1 infections among newly diagnosed HIV-1 cases in Yunnan, China, 2015. BMC Public Health 2019 Nov 11;19(1):1507 [FREE Full text] [doi: 10.1186/s12889-019-7557-8] [Medline: 31711447]
12. Yang H, Li Y, He F, Yuan F, Liu L, Li L, et al. Demographic characteristics and hot-spot areas of recent infections among new HIV diagnoses in Sichuan, China, between 2018 and 2020. Infect Drug Resist 2023;16:779-789 [FREE Full text] [doi: 10.2147/IDR.S394828] [Medline: 36779044]
13. Zhang X, Dong L, Zhou T, Chen H, Han Y, Jin X, et al. Analysis of HIV-1 recent infections and hotspot areas in a bordering area in Yunnan Province, China. BMC Public Health 2025 Mar 24;25(1):1122 [FREE Full text] [doi: 10.1186/s12889-025-21877-3] [Medline: 40128682]
14. Yang J, Niu MZ, Zhu QY, Wang YK, Wang JB, Duan X. Spatiotemporal distribution analysis of recent HIV infections in Dehong Prefecture, Yunnan Province, from 2017 to 2022. Chin J AIDS Std 2024;30(11):1136-1140.
15. Liu L, Luan R, Yang W, Zhang L, Zhang J, Nan L, et al. Projecting dynamic trends for HIV/AIDS in a highly endemic area of China: estimation models for Liangshan Prefecture, Sichuan Province. Curr HIV Res 2009 Jul 01;7(4):390-397 [FREE Full text] [doi: 10.2174/157016209788680642] [Medline: 19601774]
16. Yang S, Yang C, Liao Q, Zhai W, Yu G, Xiao L, et al. Analysis of HIV prevalence among pregnant women in Liangshan Prefecture, China, from 2009 to 2015. PLoS One 2017 Sep 7;12(9):e0183418 [FREE Full text] [doi: 10.1371/journal.pone.0183418] [Medline: 28880917]
17. Liu L, Liu L, Yu G, Wang Z, Hu Y, Yuan F. Application of EPP/Spectrum model in the estimation of AIDS epidemic in Liangshan. Chin Prev Med 2020;21(8):57-61 [FREE Full text]
18. Zhu C, Zhou Y, Song X, Jiang Q. Review on the influence factors of AIDS epidemic and control in Liangshan Prefecture Sichuan Province of China. Fudan Univ J Med Sci 2015;42(5):675.
19. Yang S, Luo M, Zhang S, Yao Y, Wang Q, Liao Q. Overview on culture and customs related to AIDS epidemic and prevention among Yi people in Liangshan. Chin J AIDS STD 2017;23(3):271-272 [FREE Full text]
20. Harrison SE, Hung P, Green K, Miller SJ, Paton M, Ahuja D, et al. Does travel time matter?: predictors of transportation vulnerability and access to HIV care among people living with HIV in South Carolina. BMC Public Health 2025 Mar 08;25(1):926 [FREE Full text] [doi: 10.1186/s12889-025-22090-y] [Medline: 40057737]
21. Palk L, Okano JT, Dullie L, Blower S. Travel time to health-care facilities, mode of transportation, and HIV elimination in Malawi: a geospatial modelling analysis. Lancet Glob Health 2020 Dec;8(12):e1555-e1564 [FREE Full text] [doi: 10.1016/S2214-109X(20)30351-X] [Medline: 33220218]
22. Larmarange J, Bendaud V. HIV estimates at second subnational level from national population-based surveys. AIDS 2014 Nov;28 Suppl 4(4):S469-S476 [FREE Full text] [doi: 10.1097/QAD.0000000000000480] [Medline: 25406750]
23. Palk L, Blower S. Geographic variation in sexual behavior can explain geospatial heterogeneity in the severity of the HIV epidemic in Malawi. BMC Med 2018 Mar 09;16(1):22 [FREE Full text] [doi: 10.1186/s12916-018-1006-x] [Medline: 29422096]
24. Silverman BW. Density Estimation for Statistics and Data Analysis. New York, NY: Routledge; 2018.
25. How Kernel density works. ArcMap. URL: https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/how-kernel-density-works.htm [accessed 2025-02-10]
26. Moran PA. NOTES on continuous stochastic phenomena. Biometrika 1950;37(1-2):17-23. [doi: 10.1093/biomet/37.1-2.17]
27. Anselin L. Spatial Econometrics: Methods and Models. Cham, Switzerland: Springer; 2013.
28. Wang JF, Xu CD. Geodetector: principle and prospective. Acta Geographica Sinica 2017;72(1):116-134 [FREE Full text]
29. Hargreaves JR, Bonell CP, Morison LA, Kim JC, Phetla G, Porter JD, et al. Explaining continued high HIV prevalence in South Africa: socioeconomic factors, HIV incidence and sexual behaviour change among a rural cohort, 2001-2004. AIDS 2007 Nov;21 Suppl 7:S39-S48. [doi: 10.1097/01.aids.0000300534.97601.d6] [Medline: 18040163]
30. Stopka TJ, Brinkley-Rubinstein L, Johnson K, Chan PA, Hutcheson M, Crosby R, et al. HIV clustering in Mississippi: spatial epidemiological study to inform implementation science in the deep south. JMIR Public Health Surveill 2018 Apr 03;4(2):e35 [FREE Full text] [doi: 10.2196/publichealth.8773] [Medline: 29615383]
31. Zulu LC, Kalipeni E, Johannes E. Analyzing spatial clustering and the spatiotemporal nature and trends of HIV/AIDS prevalence using GIS: the case of Malawi, 1994-2010. BMC Infect Dis 2014 May 23;14(1):285 [FREE Full text] [doi: 10.1186/1471-2334-14-285] [Medline: 24886573]
32. Amuquandoh A, Escamilla V, Mofolo I, Rosenberg NE. Exploring the spatial relationship between primary road distance to antenatal clinics and HIV prevalence in pregnant females of Lilongwe, Malawi. Int J STD AIDS 2019 Jun 19;30(7):639-646. [doi: 10.1177/0956462419830232] [Medline: 30890119]
33. Alirol E, Getaz L, Stoll B, Chappuis F, Loutan L. Urbanisation and infectious diseases in a globalised world. Lancet Infect Dis 2011 Feb;11(2):131-141. [doi: 10.1016/s1473-3099(10)70223-1]
34. Rothenberg RB, Potterat JJ, Woodhouse DE, Muth SQ, Darrow WW, Klovdahl AS. Social network dynamics and HIV transmission. AIDS 1998 Aug 20;12(12):1529-1536. [doi: 10.1097/00002030-199812000-00016] [Medline: 9727575]

XSL•FO
RenderX

35. Dyson T. HIV/AIDS and urbanization. Popul Dev Rev 2004 Apr 20;29(3):427-442. [doi: 10.1111/j.1728-4457.2003.00427.x]
36. Makhema J, Wirth KE, Pretorius Holme M, Gaolathe T, Mmalane M, Kadima E, et al. Universal testing, expanded treatment, and incidence of HIV infection in Botswana. N Engl J Med 2019 Jul 18;381(3):230-242. [doi: 10.1056/nejmoa1812281]
37. Cohen MS, Chen YQ, McCauley M, Gamble T, Hosseinipour MC, Kumarasamy N, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011 Aug 11;365(6):493-505. [doi: 10.1056/nejmoa1105243]
38. Cohen MS, Dye C, Fraser C, Miller WC, Powers KA, Williams BG. HIV treatment as prevention: debate and commentary--will early infection compromise treatment-as-prevention strategies? PLoS Med 2012 Jul 10;9(7):e1001232 [FREE Full text] [doi: 10.1371/journal.pmed.1001232] [Medline: 22802728]
39. Wu J, Norris J, Jia Y, Wang N. HIV treatment as prevention: contradictory perspectives from dynamic mathematical models. ScientificWorldJournal 2014;2014:760734 [FREE Full text] [doi: 10.1155/2014/760734] [Medline: 25580461]
40. Zhang M, Zhong D, Ma Q, Gao X. Research of influence factors and countermeasures of medication compliance of the AIDS patients in national area. Med J West China 2017;29(08):1165-1168 [FREE Full text]
41. Liu L, Dong A, Liao L, Feng Y, Shao Y, Liang S, et al. Survey of pretreatment HIV drug resistance and genetic transmission network analysis among HIV patients in a high drug-use area of Southwest China. Curr HIV Res 2019;17(6):441-451 [FREE Full text] [doi: 10.2174/1570162X17666191128101426] [Medline: 31778107]

Abbreviations

ART: antiretroviral therapy
FRR: false recent rate
OLSR: ordinary least squares regression
SLR: spatial lag regression

Edited by A Mavragani, T Sanchez; submitted 01.04.25; peer-reviewed by J Zhao, O Lawal; comments to author 08.05.25; revised version received 23.06.25; accepted 04.08.25; published 16.09.25.

Please cite as:
Zhu Q, Jike C, Xu C, Liang S, Yu G, Yuan D, Mai H, Li Y, Xiao L, Wang J, Yang H, Yuan F, Hong J, Mao M, Shen M, Liu J, He L, Wang Y, Cheng H, Guan P, Jiang Y, Han M, Jin C, Liu Z
Identifying the Factors Associated With Spatial Clustering of Incident HIV Infection Cases in High-Prevalence Regions: Quantitative Geospatial Study
JMIR Public Health Surveill 2025;11:e75291
URL: https://publichealth.jmir.org/2025/1/e75291
doi:10.2196/75291
PMID:

©Qiyu Zhu, Chunnong Jike, Chengdong Xu, Shu Liang, Gang Yu, Dan Yuan, Hong Mai, Yiping Li, Lin Xiao, Ju Wang, Hong Yang, Fengshun Yuan, Jing Hong, Muga Mao, Maogang Shen, Jing Liu, Lin He, Yuehua Wang, Huanyi Cheng, Peng Guan, Yan Jiang, Mengjie Han, Cong Jin, Zhongfu Liu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Long-Term Ambient Air Pollution Exposure and the Risk of Cardiovascular and Cerebrovascular Diseases in Rural Chinese Populations: 10-Year Follow-Up of a Multicenter Prospective Cohort Study

Yaqi Zhao[1,2*], BS; Xuefang Cao[1,2*], PhD; Jiang Du[1,2*], PhD; Aiwei He[3], BD; Jun Liang[4], MD; Weitao Duan[5], BS; Yuanzhi Di[1,2], BS; Yijun He[1,2], MD; Boxuan Feng[1,2], BS; Linyu Shen[1,2], MD; Juanjuan Huang[1,2], MD; Zihan Li[1,2], MD; Jianguo Liang[1,2], MD; Hongzhi Li[6], MD; Zisen Liu[5], BS; Fang Liu[3], BD; Shumin Yang[3], MPH; Zuhui Xu[4], MPH; Bin Zhang[5], BS; Jiaoxia Yan[5], BS; Yanchun Liang[7], BD; Rong Liu[8], BD; Fei Shen[6], MD; Qi Jin[1,2], PhD; Henan Xin[1,2], PhD; Lei Gao[1,2], PhD

[1]NHC Key Laboratory of Systems Biology of Pathogens, National Institute of Pathogen Biology, and Center for Tuberculosis Research, Chinese Academy of Medical Sciences & Peking Union Medical College, No.16 Tianrong Street, Beijing, China

[2]Key Laboratory of Pathogen Infection Prevention and Control (Ministry of Education), National Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[3]Tuberculosis Prevention and Control Department, Gansu Provincial Center for Disease Control and Prevention, Lanzhou, China

[4]Tuberculosis Prevention and Control Department, Hunan Provincial Institute of Tuberculosis Prevention and Control, Changsha, Human, China

[5]Tuberculosis Prevention and Control Department, Zhongmu County Center for Disease Control and Prevention, Zhongmu, China

[6]Zhengzhou Sixth People's Hospital, Zhengzhou, Henan, China

[7]Tuberculosis Prevention and Control Department, Longxi County Center for Disease Control and Prevention, Longxi, China

[8]Tuberculosis Prevention and Control Department, Xiangtan County Center for Disease Control and Prevention, Xiangtan, China

[*]these authors contributed equally

Corresponding Author:
Lei Gao, PhD
NHC Key Laboratory of Systems Biology of Pathogens, National Institute of Pathogen Biology, and Center for Tuberculosis Research, Chinese Academy of Medical Sciences & Peking Union Medical College, No.16 Tianrong Street, Beijing, China

Abstract

Background: Long-term follow-up studies investigating the relationship between ambient air pollution and cardiovascular and cerebrovascular diseases (CVD) in rural Chinese populations remain limited.

Objective: This study aimed to investigate the impact of prolonged exposure to particulate matter with aerodynamic diameter ≤2.5 μm ($PM_{2.5}$) on CVD in rural areas of China.

Methods: On the basis of a multicenter population-based prospective study, adult rural residents (aged ≥15 y) from 3 study sites (ie, Xiangtan, Hunan Province; Longxi, Gansu Province; and Zhongmu, Henan Province) with different $PM_{2.5}$ exposure levels were tracked for the incidence of CVD events between 2013 and 2023. The relationship was assessed by applying the Cox proportional hazards model and a trend test.

Results: A total of 15,502 participants were included in the study. During the 10-year follow-up period, for every 1 μg/m^3 increase in $PM_{2.5}$, the risks of CVD, ischemic stroke, coronary heart disease, acute coronary syndrome, and intracerebral hemorrhage increased by 5% (hazard ratio [HR] 1.05, 95% CI 1.04 - 1.06), 7% (HR 1.07, 95% CI 1.06 - 1.08), 8% (HR 1.08, 95% CI 1.07 - 1.09), 9% (HR 1.09, 95% CI 1.06 - 1.11), and 10% (HR 1.10, 95% CI 1.07 - 1.14), respectively. Furthermore, the risk in the high exposure group (Q4) was found to be significantly higher than that in the low exposure group (Q1; *P* for trend <.001). The subgroup analysis indicated that the risk of CVD was higher among older people compared to individuals aged <60 years, and the interaction effect was statistically significant (interaction *P* value=.03).

Conclusions: Our results indicate that long-term exposure to $PM_{2.5}$ significantly increases the risk of CVD in rural areas of China and shows regional differences. This finding may deepen our understanding of the potential public health risks associated with $PM_{2.5}$ exposure and emphasize the crucial role of environmental governance in promoting public health outcomes.

(JMIR Public Health Surveill 2025;11:e81218) doi:10.2196/81218

XSL•FO
RenderX

KEYWORDS

cardiovascular and cerebrovascular diseases; particulate matter; air pollution; prospective study; Cox proportional hazards model

Introduction

Background

Ambient air pollution, considered the most significant environmental risk factor for global mortality and morbidity, particularly particulate matter with aerodynamic diameter ≤2.5 μm ($PM_{2.5}$) with an aerodynamic diameter of less than 2.5 micrometers, has emerged as a significant global public health concern [1]. The Global Burden of Disease report indicates that in 2019, air pollution was responsible for the deaths of 9 million people worldwide, with 61.9% of these fatalities due to increased mortality from cardiovascular and cerebrovascular diseases (CVD) [2]. Furthermore, there are significant variations in air pollution standards worldwide [3]. The limits in China and India are considerably higher than those in high-income nations (at 35 and 40 μg/m^3, respectively). Even in countries with lower levels of air pollution, there is evidence indicating a correlation between air pollution and cardiovascular health [4]. Early large-scale cohort studies reported the impact of long-term exposure to air pollution on CVD events in Europe and the United States. The environmental $PM_{2.5}$ exposure levels were generally <35 μg/m^3 [5,6]. A study of the European Study of Cohorts for Air Pollution Effects project [7] revealed that for every 5 μg/m^3 increase in $PM_{2.5}$ concentration, there was a 13% rise in the risk of coronary events. A cohort study conducted in the United States with 1,934,453 older participants found that for every 1 μg/m^3 increase in $PM_{2.5}$, the risk of transient cerebral ischemia and heart failure episodes increased by 3.5% and 1.9%, respectively [8]. This phenomenon may be attributed to the capacity of smaller particles to penetrate deeper lung regions and enter the bloodstream, thereby inducing systemic inflammation, oxidative stress, and endothelial dysfunction [9]. Consequently, this exposure increases the susceptibility of the CVD systems to the detrimental effects of these particles [10]. However, in many countries, particularly in low- and middle-income countries, the long-term effects of air pollution on health have not been thoroughly investigated [3]. In China, particularly in rural areas, there are numerous instances of biomass burning and coal-fired power plants [4]. The $PM_{2.5}$ exposure levels in these rural areas are significantly higher than those in urban areas, far exceeding the World Health Organization's Air Quality Guidelines [3]. However, epidemiological studies that link long-term air pollution in rural China to specific CVD are limited in scope.

Objectives

This research aims to conduct a multicenter, population-based prospective study to explore the long-term association between $PM_{2.5}$ exposure and the risks of CVD among rural residents from different regions. The findings of this study may contribute to a deeper understanding of the impact of air pollution on health and provide a more robust scientific basis for the development of public health policies in China.

Methods

Study Design and Participants

A population-based multicenter cohort study was initiated in 2013 and was a 10-year follow-up survey of registered residents at 3 study sites using a closed cohort design [11] (ie, Xiangtan, Hunan Province; Longxi, Gansu Province; and Zhongmu, Henan Province) between October 1, 2023, and January 31, 2024, to track the occurrence of CVD. This cohort study was jointly organized by the Institute of Virology at the Chinese Academy of Medical Sciences and the Chinese Center for Disease Control and Prevention. It was divided into 2 phases and spanned a duration of 10 years, from 2013 to 2023. During the initial phase, we gathered the participants' sociodemographic information and disease history. In the second phase, which took place in 2014, 2015, 2018, and 2023, we monitored their disease onset status and corroborated and refined the data through the local chronic disease management system (CDMS). The inclusion criteria for the research participants were as follows: birth date before June 1, 1998 (age ≥15 y); possession of a household registration or residence permit for the village; continuous residence at the research sites for 6 months or longer in the past year; the ability to complete the investigations during the research period; and the provision of voluntary written informed consent. The exclusion criteria were individuals without a residence address at the time of the baseline survey, those lost to follow-up, and those who became pregnant. All eligible current residents living in the 3 selected research sites were included in the 10-year follow-up survey.

Ethical Considerations

The study protocol was approved by the ethics committees of the Institute of Pathogen Biology, Chinese Academy of Medical Sciences, Beijing, China (approval number IPB-2023‐35). Written informed consent was obtained from all study participants, and all participants have the right to withdraw at any time. The study provides breakfast and transportation subsidies to the participants.

Procedure

In this study, measures were implemented to ensure the quality and comparability of data across the 3 research sites. These included standardizing the research protocol, providing uniform training for researchers, and applying consistent disease diagnostic criteria.

Sociodemographic data for each research participant were systematically collected using standardized questionnaires administered by trained interviewers. The data included age, gender, educational attainment, smoking status, alcohol consumption status, weight, height, marital status, per capita household income for 2013 (calculated by dividing the total household income by the number of family members) [12], and a history of hypertension. Household income per capita was categorized based on the national mean level in 2010 (6000 RMB [US $887]) [13]. BMI was categorized as underweight

XSL•FO
RenderX

(<18.5 kg/m^2), normal weight (≥18.5 kg/m^2 to <24.0 kg/m^2), and overweight (≥24.0 kg/m^2) [14].

Assessment of Exposure to Air Pollution

The $PM_{2.5}$ air pollution dataset originates from the near real-time tracking dataset of atmospheric components in China, known as "*Tracking China Air Pollution*" [15,1617]. The $PM_{2.5}$ prediction model establishes a 2-tier machine learning framework. In the first tier, the model uses a resampled training data set and the random forest algorithm to predict high-pollution events. In the second-level model, the second random forest model is established using the residuals between the $PM_{2.5}$ concentration simulated by the Community Multiscale Air Quality model and the observed $PM_{2.5}$ concentration. In the 2-level model, a decision tree–based method is used to establish the association between missing data and other parameters, thereby compensating for the absence of satellite data. The $PM_{2.5}$ prediction at a 1 km resolution integrates high-resolution satellite remote sensing aerosol optical depth data and environmental spatial data, such as road networks, to invert the $PM_{2.5}$ concentration at a 1 km resolution, which is fully covered daily. This model effectively captures changes in $PM_{2.5}$ concentration across various spatiotemporal scales and exhibits high accuracy. The cross-validation determination coefficient (CV-R^2) ranges from 0.86 to 0.90, indicating strong predictive ability (R^2 ranges between 0.80 and 0.84). To evaluate the model's ability to detect variations in $PM_{2.5}$ levels in rural areas with a limited number of monitoring stations and on a local scale, national monitoring station data were used for model evaluation. The results indicate that the high-prediction model performs comparably in out-of-bag evaluation, test data evaluation, and yearly cross-validation evaluation, demonstrating extremely high accuracy and robustness. The long-term $PM_{2.5}$ exposure levels are assigned based on the geographical coordinates of the research participants' permanent residences. The residential address information was collected through questionnaires during the baseline survey. We used the sf package in R software (version 4.4.3; R Foundation for Statistical Computing) to match these addresses with China's township-level administrative centers' database of latitude and longitude and assigned the geographical center's latitude and longitude coordinates of each village to each research participant. All coordinates are part of the WGS84 coordinate system. We then overlaid these coordinates with high-resolution $PM_{2.5}$ raster data, extracted the $PM_{2.5}$ concentration value corresponding to each coordinate point for a specific year and date, and used this as the individual's long-term exposure level. The total $PM_{2.5}$ concentrations for each participant during 4 different exposure windows were estimated, including moving averages for 1-year, 3-year, 5-year, and 10-year periods before the measurement date.

Research Outcome

Statistical Analysis

The composite outcome of this study is the incidence rate of major adverse cardiovascular events. Major adverse cardiovascular events was defined as a composite end point, encompassing the first occurrence of any of the following events: (1) ischemic stroke (IS), (2) coronary heart disease (CHD), (3) acute coronary syndrome (ACS), (4) intracerebral hemorrhage (ICH), and (5) any other form of CVD. In addition to obtaining the relevant disease information for the primary diseases diagnosed by qualified medical institutions (collected through baseline surveys and 10 y follow-up questionnaires), the local CDMS will export data based on ID numbers and match it with our study participants. This CDMS has been developed using the National Basic Public Health Service Management System, which was launched in 2009, and it ensures 100% coverage of all grassroots medical institutions nationwide. If a diagnosis certificate cannot be provided and it is not recorded in the system, it is not considered to have the disease. To assess the impact of $PM_{2.5}$ on specific diseases more specifically, we also included the following outcomes as end points for separate analysis:

- IS: defined as the first hospitalization or emergency event due to ischemic stroke, with the primary diagnosis code being *International Classification of Diseases, Tenth Revision* (*ICD-10*), I63.
- CHD: defined as the first hospitalization or emergency event due to CHD, with the primary diagnosis code being I20-I25. This includes angina pectoris and chronic ischemic heart disease.
- ACS: defined as a severe subtype of CHD, it is characterized by the first hospitalization or emergency event due to ACS, with the primary diagnosis code being I21 (acute myocardial infarction) or I20.0 (unstable angina pectoris).
- ICH: defined as the initial hospitalization or emergency event resulting from ICH. The primary diagnostic code is *ICD-10* I61.
- Any other form of CVD: hospitalization events for other CVD not encompassed by the aforementioned categories (*ICD-10* I00-I99) will be categorized as "other" outcomes and subjected to analysis.

To ensure the accuracy of the outcome definition, we require that each event be coded as the primary diagnosis. For all suspected events, we conduct a secondary confirmation by reviewing the medical records and require that their clinical manifestations and imaging examination results conform to internationally recognized diagnostic standards.

The data were independently entered into the EpiData software (version 3.1; EpiData Consortium) by 2 trained data entry professionals. Any discrepancies between the 2 datasets were resolved through cross-referencing with the original records to ensure data integrity and accuracy. Statistical analyses were performed using R (version 4.4.3; R Foundation for Statistical Computing). Quantitative variables are presented as medians, whereas qualitative variables are summarized as counts (percentages). BMI was calculated as weight (kilogram) divided by height squared (square meter). Fisher exact test and Pearson χ^2 test were used to compare the distribution of categorical variables across groups. The Cox proportional hazards model was used to investigate the association between long-term exposure to $PM_{2.5}$ and the risk of CVD incidence over 1, 3, 5,

XSL•FO
RenderX

and 10 years of follow-up, with continuous adjustments for selected covariates. The selection of covariates was based on prior knowledge and known or potential risk factors for CVD. Specifically, these include demographic variables (eg, age, gender, and BMI), socioeconomic factors (eg, educational attainment and household income), behavioral factors (eg, smoking status and alcohol consumption), and clinical history (eg, a history of hypertension). All variables were collected at the baseline of the study and were included as covariates in the multivariate Cox model during the univariate analysis (P<.05). The time variable in the model was defined as study time (follow-up time) [18]. For every 1 μg/m^3 increase in $PM_{2.5}$ concentration, the hazard ratio (HR) and the corresponding 95% CI were estimated. Subgroup analyses were conducted based on demographic and baseline disease risk factors to determine HRs for specific stratifications. Two-sided P values less than .05 were deemed statistically significant. On the basis of the tertiles of $PM_{2.5}$ concentration, the population was divided into a low exposure group (Q1), a medium exposure group (Q2 and Q3), and a high exposure group (Q4). The quartile groups were used as ordinal variables, assigned values of 1, 2, 3, and 4, for linear trend tests. The P value for trend was calculated to evaluate the overall dose-response trend, and restricted cubic spline analysis was used to examine the curve shape. The interaction P value was determined using the likelihood ratio test to assess the modifying effect of $PM_{2.5}$ on disease associations across various subgroups.

Results

Characteristics of the Study Participants Included in the Analysis in 2013

The information on the 3 study sites and the study participants included in the analysis is detailed in Table 1. Among the 16,636 eligible participants, 15,745 actually participated in the 10-year follow-up survey, yielding a response rate of 94.64%(15,745/16,636). After excluding 57 participants who lacked geographical location information and 186 participants with incomplete data, 15,502 participants were included in the final analysis (Table 1). The total follow-up period spanned 136,310 person-years. Approximately half of the participants were male (7293/15,502), and 22.76% (3528/15,502) of the participants were aged ≥60 years. The age distribution significantly differed among the 3 study sites (P<.001), with a higher proportion of individuals aged ≥60 years in Xiangtan compared to Zhongmu and Longxi (Table 1). Of the participants, 25.96% (4024/15,502) reported current smoking, and 18.35% (2845/15,502) reported alcohol consumption in the past year (Table 1).

XSL•FO
RenderX

Table . Characteristics of the study population at 2013 baseline survey.

Characteristics[a]	Total	Xiangtan	Longxi	Zhongmu	*P* value for χ^2 test
Total, n (%)	15,502	5279 (34.05)	4971 (32.07)	5252 (33.88)	—[b]
Sex, n (%)					<.001
Female	8209 (52.95)	2770 (52.47)	2863 (57.59)	2576 (49.05)	
Male	7293 (47.05)	2509 (47.53)	2108 (42.41)	2676 (50.95)	
Age (y), n (%)					<.001
<60	11,974 (77.24)	3565 (67.53)	4077 (82.02)	4332 (82.48)	
≥60	3528 (22.76)	1714 (32.47)	894 (17.98)	920 (17.52)	
Education, n (%)					<.001
No schooling	3141 (20.26)	517 (9.79)	1241 (24.96)	1383 (26.33)	
Primary school or higher	12,361 (79.74)	4762 (90.21)	3730 (75.04)	3869 (73.67)	
Marital history, n (%)					<.001
Unmarried	3651 (23.55)	784 (14.85)	1468 (29.53)	1399 (26.64)	
Married	11,098 (71.59)	4173 (79.05)	3298 (66.34)	3627 (69.06)	
Divorced	59 (0.38)	28 (0.53)	17 (0.34)	14 (0.27)	
Widowed	694 (4.48)	294 (5.57)	188 (3.79)	212 (4.03)	
Income (RMB), n (%)[c]					<.001
<6000 (US $887)	11,420 (73.67)	3456 (65.47)	3791 (76.26)	4173 (79.46)	
≥6000 (US $887)	4082 (26.33)	1823 (34.53)	1180 (23.74)	1079 (20.54)	
BMI (kg/m^2), n (%)					<.001
<18.5	2818 (18.18)	746 (14.13)	1180 (23.74)	892 (16.98)	
≥18.5 to <24	7155 (46.16)	2725 (51.62)	2367 (47.62)	2063 (39.28)	
≥24	5529 (35.66)	1808 (34.25)	1424 (28.64)	2297 (43.74)	
Smoking history, n (%)					<.001
Never	11,478 (74.04)	3574 (67.70)	4008 (80.63)	3896 (74.18)	
Ever (current and former)	4024 (25.96)	1705 (32.30)	963 (19.37)	1356 (25.82)	
Current drinking status, n (%)					<.001
No	12,657 (81.65)	4385 (83.06)	4415 (88.82)	3857 (73.44)	
Yes	2845 (18.35)	894 (16.94)	556 (11.18)	1395 (26.56)	
History of hypertension, n (%)					<.001
No	14,412 (92.97)	4675 (88.56)	3706 (74.55)	5031 (95.79)	
Yes	1090 (7.03)	604 (11.44)	265 (25.45)	221 (4.21)	

[a]Data might not sum to the total because of missing data.

[b]Not available.

[c]Stratification according to total family income/number of people in the household.

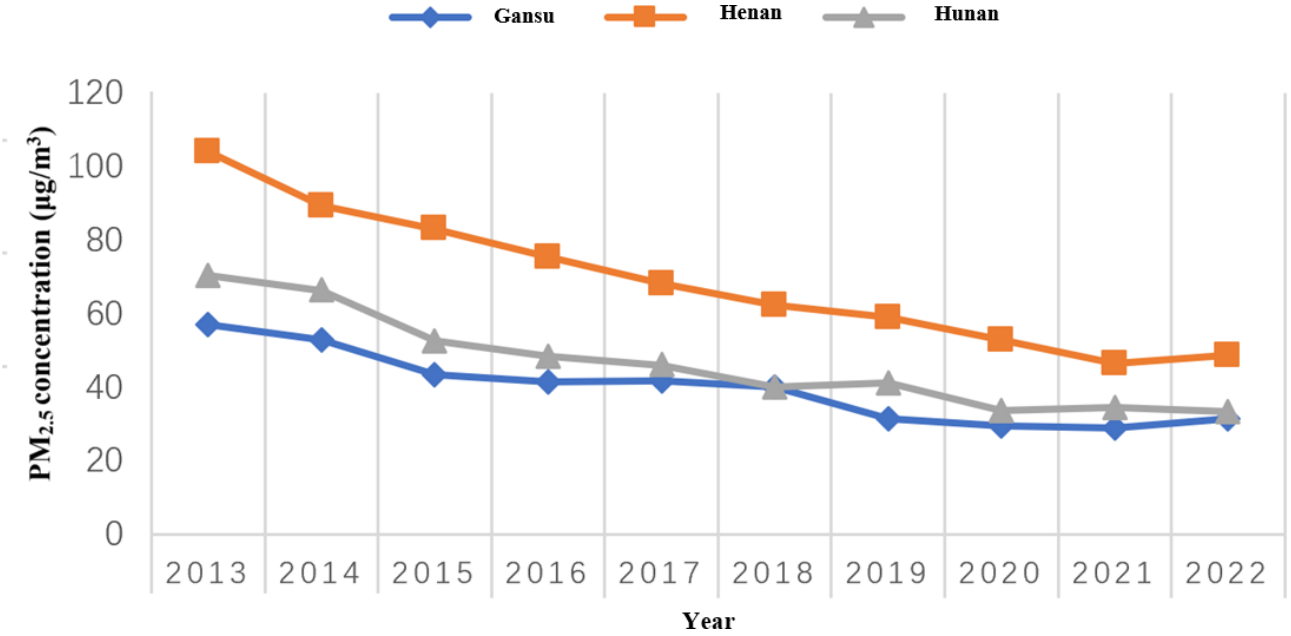
Geographical Distribution of $PM_{2.5}$ Concentration and Regional Differences in the Risk of CVD

The participants were evenly distributed across the 3 sites. From 2013 to 2023, the annual average $PM_{2.5}$ concentrations at these 3 sites exhibited a decreasing trend over time; however, they remained at relatively high levels, ranging from 35.80 to 94.91 $\mu g/m^3$. The median 10-year average $PM_{2.5}$ exposure was 51.50 $\mu g/m^3$. The lowest concentration was found in Gansu Province, whereas the highest was in Henan Province (Figure 1). Over the 10-year period, a total of 628 cases occurred, accounting for 4.05% (628/15,502). Among them, the incidence rates were relatively high among men, individuals aged >60 years, those with an overweight BMI, and those with hypertension (Table

XSL•FO
RenderX

2). For every 1 μg/m^3 increase in $PM_{2.5}$, the risk of CVD increases by 5% (HR 1.05, 95% CI 1.04 - 1.06).

Figure 1. Annual average particulate matter levels at three sites from 2013 to 2022.



XSL•FO
RenderX

Table . Analysis of risk factors for cardiovascular and cerebrovascular diseases.

Characteristics	Cases/total, n/N (%)	*P* value for univariate regression	HR[a,b] (95% CI)
Total	628/15,502 (4.05)	—[c]	—
$PM_{2.5}$[d] (μg/m^3)	—	<.001	1.05 (1.04 - 1.06)
Sex		.006	
Female	303/8209 (3.69)		Reference
Male	325/7293 (4.46)		1.25 (0.97 - 1.61)
Age (y)		<.001	
<60	304/11,974 (2.54)		Reference
≥60	324/3528 (9.18)		3.57 (2.97 - 4.28)
Education		<.001	
Primary school or higher	425/12,361 (3.44)		Reference
No schooling	203/3141 (6.46)		1.19 (0.98 - 1.44)
Marital status		.03	
Married	566/11,098 (5.10)		Reference
Unmarried	7/3651 (0.19)		0.07 (0.03 - 0.16)
Divorced	1/59 (1.69)		0.41 (0.06 - 2.87)
Widowed	54/694 (7.78)		1.20 (0.89 - 1.60)
Income[e] (RMB)		.07	
<6000 (US $887)	473/11,420 (4.14)		1.03 (0.84 - 1.26)
≥6000 (US $887)	155/4082 (3.80)		Reference
BMI (kg/m^2)		<.001	
<18.5	19/2818 (0.67)		0.76 (0.47 - 1.23)
≥18.5 to <24	258/7155 (3.61)		Reference
≥24	351/5529 (6.35)		1.35 (1.14 - 1.59)
Smoking status		<.001	
Never	382/11,478 (3.33)		Reference
Ever (current and former)	246/4024 (6.11)		1.50 (1.17 - 1.90)
Current drinking status		<.001	
No	475/12,657 (3.75)		Reference
Yes	153/2845 (5.38)		1.32 (1.06 - 1.64)
History of hypertension		<.001	
No	505/14,412 (3.50)		Reference
Yes	123/1090 (11.28)		2.27 (1.84 - 2.80)

[a]HR: hazard ratio.

[b]HR analysis included 15,502 participants. HR indicates the increase in incidence risk for every 1 μg/m^3 increase in $PM_{2.5}$ concentration.

[c]Not available.

[d]$PM_{2.5}$: particulate matter with aerodynamic diameter ≤2.5 μm.

[e]Stratification according to total family income/number of people in the household.

On the basis of the estimated $PM_{2.5}$ values from satellite remote sensing within a 1 km radius of the research subjects' residences, the average $PM_{2.5}$ concentration for each year from 2013 to 2022 was calculated, and the changes in $PM_{2.5}$ levels over time for the 3 locations were plotted.

Long-term exposure to varying concentrations of $PM_{2.5}$ over different periods was associated with a slight increase in the incidence of CVD, and this association was statistically significant. Overall, the risk of disease increased significantly with 1-year and 3-year $PM_{2.5}$ exposure, while it slightly

XSL•FO
RenderX

decreased with 5-year and 10-year moving average exposure (Table 3). After adjusting for all potential covariates, during the 1-year exposure window, the overall risk of disease increased by 10% for every 1 μg/m^3 increase in $PM_{2.5}$ (HR 1.10, 95% CI 1.09 - 1.11). Specifically, the risk of IS, CHD, ACS, and ICH increased by 13% (HR 1.13, 95% CI 1.11 - 1.15), 6% (HR 1.06, 95% CI 1.03 - 1.09), 16% (HR 1.16, 95% CI 1.13 - 1.19), and 18% (HR 1.18, 95% CI 1.13 - 1.24), respectively. These associations were statistically significant (Table 3).

Table . The association between $PM_{2.5}$[a] exposure windows and the risk of CVD[b,c].

Diseases	1 y		3 y		5 y		10 y	
	HR[d] (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value	HR (95% CI)	*P* value
CVD	1.10 (1.09 - 1.11)	<.001	1.09 (1.08 - 1.10)	<.001	1.07 (1.06 - 1.08)	<.001	1.05 (1.04 - 1.05)	<.001
Ischemic stroke	1.13 (1.11 - 1.15)	<.001	1.12 (1.10 - 1.14)	<.001	1.10 (1.09 - 1.12)	<.001	1.07 (1.06 - 1.08)	<.001
Coronary heart disease	1.06 (1.03 - 1.09)	.004	1.07 (1.05 - 1.09)	.02	1.07 (1.05 - 1.09)	.009	1.08 (1.07 - 1.09)	<.001
Acute coronary syndrome	1.16 (1.13 - 1.19)	<.001	1.15 (1.12 - 1.18)	<.001	1.12 (1.10 - 1.15)	<.001	1.09 (1.06 - 1.11)	<.001
Intracerebral hemorrhage	1.18 (1.13 - 1.24)	<.001	1.17 (1.12 - 1.22)	<.001	1.14 (1.10 - 1.19)	<.001	1.10 (1.07 - 1.14)	<.001
Other	1.05 (1.01 - 1.09)	.007	1.04 (1.01 - 1.08)	.009	1.04 (1.01 - 1.07)	.009	1.03 (1.01 - 1.05)	.01

[a]$PM_{2.5}$: particulate matter with aerodynamic diameter ≤2.5 μm.

[b]CVD: cardiovascular and cerebrovascular diseases.

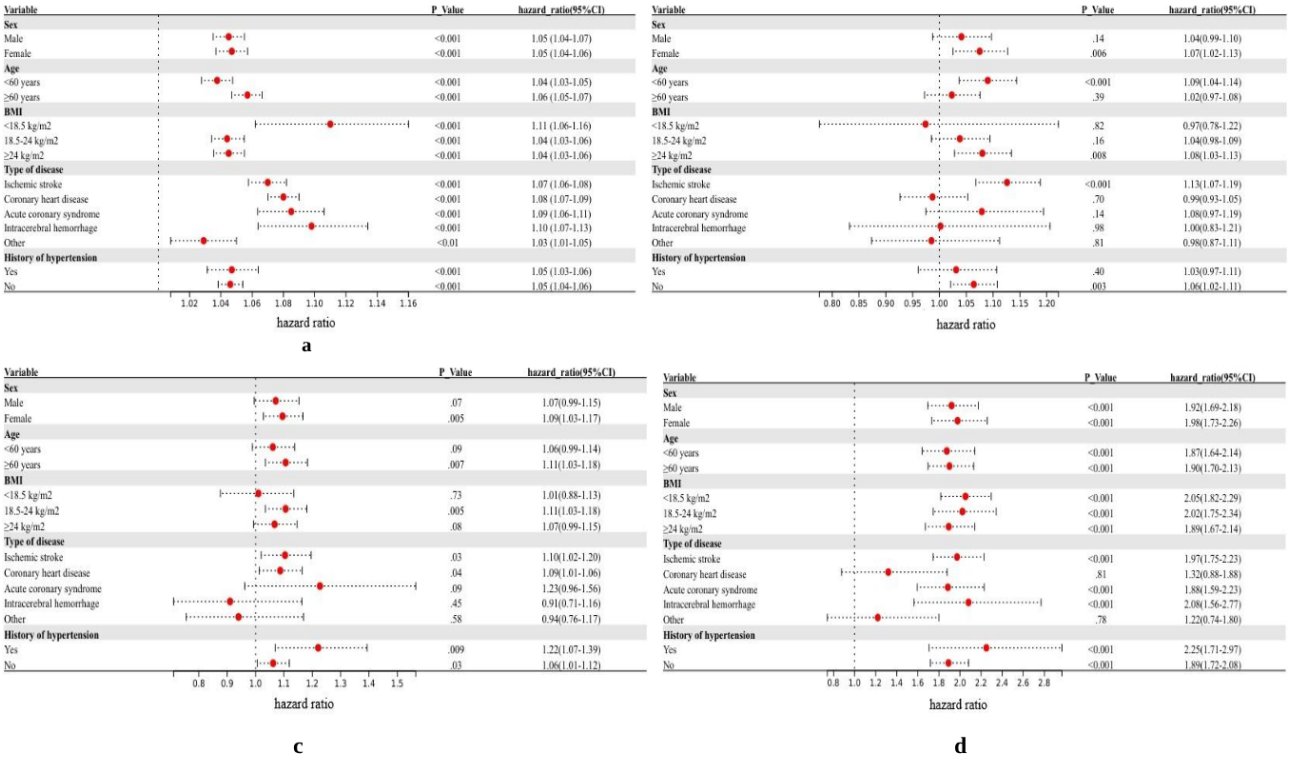
[c]Note: Adjusted HRs (95% CIs) of the occurrence risk are presented by per 1 μg/m^3 increment in moving average ambient $PM_{2.5}$ concentrations in different durations. The Cox model is adjusted for baseline age, gender, income, education level, BMI, smoking and drinking status, marital status, and history of hypertension.

[d]HR: hazard ratio.

Stratified analysis by different regions indicated that at the Zhongmu station, on average, the $PM_{2.5}$ concentration was the highest (73.24 μg/m^3), and its impact on the risk of CVD was higher than that at the Xiangtan and Longxi stations, with an adjusted HR of 1.89 (95% CI 1.73 - 2.06; Figure 2). Considering distinct genders, ages, BMI levels, and hypertension histories, the disease risk in the Zhongmu area is notably higher than in other regions. Additionally, at the Xiangtan, Longxi, and Zhongmu research stations, $PM_{2.5}$ exposure was positively correlated with the risk of IS, with adjusted HR of 1.13 (95% CI 1.07 - 1.19), 1.10 (95% CI 1.02 - 1.20), and 1.97 (95% CI 1.75 - 2.23), respectively (Figure 2). Additionally, the risk of ACS in Xiangtan, Longxi, and Zhongmu was 1.08 (95% CI 0.97 - 1.19), 1.23 (95% CI 0.96 - 1.56), and 1.88 (95% CI 1.59 - 2.23), respectively. For ICH, the risks were 1.00 (95% CI 0.83 - 1.21), 0.91 (95% CI 0.71 - 1.16), and 2.08 (95% CI 1.56 - 2.77), respectively.

XSL•FO
RenderX

Figure 2. The relationship between particulate matter 2.5 ($PM_{2.5}$) and the risk of cardiovascular and cerebrovascular diseases in different sites. (A) Total, (B) Xiangtan, (C) Longxi, and (D) Zhongmou. The Cox model is adjusted for baseline age, gender, income, education level, BMI, smoking and drinking status, marital status, and history of hypertension.



Variation of CVD Risk Across $PM_{2.5}$ Concentration Levels

In this study, a fully adjusted model with a 10-year exposure window was used to perform stratified analyses on the association between $PM_{2.5}$ exposure and the risk of CVD across different exposure groups. The research findings indicate that, in comparison to Q1, the risk levels in Q2, Q3, and Q4 showed a gradual increase. The corresponding HRs were 1.05 (95% CI 0.88 - 1.26), 1.06 (95% CI 1.01 - 1.11), and 1.96 (95% CI 1.67 - 2.29), respectively. Furthermore, the trend test yielded statistically significant results (*P* for trend <.001), suggesting the presence of an overall dose-response relationship (Multimedia Appendix 1). The risk was notably higher in Q4 than in the younger population (aged <60 y). It is notable that the exposure concentration among the older population in different regions was higher than that of the younger population (Multimedia Appendix 1), and the interaction effect based on age was statistically significant (interaction *P* value=.03; Table 4).

XSL•FO
RenderX

Table . Stratified analysis of the association between particulate matter with aerodynamic diameter ≤2.5 μm ($PM_{2.5}$) and the risk of cardiovascular and cerebrovascular diseases.[a]

Characteristics[b]	Quartiles 1 (35.81-44.58μg/m^3)	Quartiles 2 (44.58-51.50 μg/m^3)	Quartiles 3 (51.50-68.77 μg/m^3)	Quartiles 4 (68.77-94.91 μg/m^3)	Interaction *P* value[c]	*P* value for trend[d]
Total	Reference	1.05 (0.88 - 1.26)	1.06 (1.01 - 1.11)	1.96 (1.67 - 2.29)	—[e]	<.001
Age (y)					.03	
<60	Reference	1.37 (0.97 - 1.94)	2.15 (1.49 - 3.08)	2.34 (1.70 - 3.22)		<.001
≥60	Reference	1.26 (0.94 - 1.69)	1.39 (0.91 - 2.12)	3.23 (2.42 - 4.30)		<.001
Sex					.96	
Male	Reference	1.28 (0.93 - 1.76)	1.74 (1.20 - 2.52)	2.63 (1.94 - 3.55)		<.001
Female	Reference	1.38 (1.01 - 1.88)	1.81 (1.23 - 2.68)	2.65 (1.95 - 3.60)		<.001
BMI (kg/m^2)					.02	
<18.5	Reference	0.85 (0.67 - 1.03)	—[f]	3.07 (2.04 - 4.12)		<.001
≥18.5 to <24	Reference	1.26 (0.92 - 1.72)	1.31 (0.81 - 2.11)	2.62 (1.91 - 3.61)		<.001
≥24	Reference	1.41 (1.01 - 1.96)	2.26 (1.60 - 3.21)	2.76 (2.04 - 3.75)		<.001
Education					<.001	
Primary school or higher	Reference	1.23 (0.96 - 1.58)	1.90 (1.39 - 2.59)	1.82 (1.39 - 2.38)		<.001
No schooling	Reference	1.65 (1.00 - 2.71)	1.99 (1.16 - 3.43)	5.65 (3.75 - 8.50)		<.001
Income[g] (RMB)					.10	
<6000 (US $887)	Reference	1.33 (1.01 - 1.74)	1.70 (1.24 - 2.33)	2.89 (2.26 - 3.69)		<.001
≥6000 (US $887)	Reference	1.31 (0.87 - 1.95)	2.18 (1.30 - 3.66)	1.93 (1.22 - 3.05)		<.01
Smoking status					.85	
Never	Reference	1.38 (1.04 - 1.84)	1.89 (1.34 - 2.66)	2.63 (2.00 - 3.46)		<.001
Ever (current and former)	Reference	1.25 (0.87 - 1.79)	1.69 (1.10 - 2.61)	2.73 (1.94 - 3.85)		<.001
Current drinking status					.11	
No	Reference	1.29 (1.01 - 1.65)	1.46 (1.05 - 2.02)	2.64 (2.09 - 3.35)		<.001
Yes	Reference	1.64 (0.92 - 2.92)	3.01 (1.72 - 5.28)	3.10 (1.85 - 5.20)		<.001
History of hypertension					.66	
Yes	Reference	1.21 (1.05 - 1.39)	0.90 (0.77 - 1.05)	1.60 (1.35 - 1.89)		<.001
No	Reference	1.32 (1.02 - 1.71)	1.89 (1.41 - 2.53)	2.65 (2.09 - 3.37)		<.001

[a]The Cox model is adjusted for baseline age, gender, income, education level, BMI, smoking and drinking status, marital status, and history of hypertension.

[b]Data might not sum to the total because of missing data.

[c]Interaction *P* value from likelihood ratio tests.

[d]*P* for trend calculated by treating $PM_{2.5}$ as an ordinal variable (Q1-Q4).

[e]Not applicable.

[f]There were no results in the subgroup analysis, which was due to the model not converging because of the small number of patients.

[g]Stratification according to total family income/number of people in the household.

Discussion

Principal Findings

The results of this multicenter, population-based, prospective study conducted in rural areas of China indicate a significant association between long-term exposure to $PM_{2.5}$ and an increased risk of CVD. In particular, the correlation between long-term exposure to $PM_{2.5}$ (over periods of 1 and 3 y) and disease risk was more pronounced, even after accounting for known risk factors. Notably, there were regional variations and dose-response trends in the risk of disease occurrence at different levels of $PM_{2.5}$ exposure.

XSL•FO
RenderX

The epidemiological evidence of the association between $PM_{2.5}$ and CVD revealed in this study significantly corroborated the findings of multicenter cohort studies conducted worldwide [19]. At the same time, it also highlighted the particularity of rural exposure. A cross-national prospective cohort study [20] has revealed that long-term exposure to outdoor $PM_{2.5}$ is associated with an increased risk of CVD in adults aged between 35 and 70 years. For every 10 $\mu g/m^3$ increase in $PM_{2.5}$ exposure, the risk of CVD death, CVD events, myocardial infarction, and stroke increased by 3%, 5%, 3%, and 8%, respectively [21]. Furthermore, the study revealed that the risk of developing CVD was higher in rural areas compared to urban areas. In our current research, similar results were also observed. In particular, the correlation between 1 to 3 years of $PM_{2.5}$ exposure and disease risk was more pronounced, with regional variations. This may be related to the following reasons: First, short-term exposure could rapidly trigger CVD events through acute inflammation and oxidative stress mechanisms, whereas long-term exposure could exacerbate the occurrence risk of atherosclerosis and CVD through cumulative effects [20]. A meta-analysis conducted by Jeroen de Bont et al [1] revealed that short-term exposure to $PM_{2.5}$ is associated with an increased risk of hypertension, stroke, and myocardial infarction. Long-term exposure to $PM_{2.5}$ was largely related to an increased risk of atherosclerosis, myocardial infarction, hypertension, stroke, and stroke mortality [22]. Second, during the study period, some participants may have relocated to new residences, which could lead to misclassification bias in the exposure assessment, potentially reducing the risk ratio. Finally, during the follow-up period, study participants may have died from causes other than CVD, particularly in long-term follow-ups, which could weaken the estimation of CVD risk. Furthermore, our study demonstrated that the effects of $PM_{2.5}$ on specific CVD differed significantly both in the overall population and in regional stratified analyses. This regional heterogeneity may arise from the complex interplay of multiple mechanisms. The variations in $PM_{2.5}$ concentrations and chemical compositions across different regions contribute to the observed disparities. The most polluted areas, such as Henan, were characterized by high population densities and an energy structure that heavily relied on coal combustion, leading to persistent high levels of air pollution from coal burning [23]. The resulting sulfate and nitrate particles made up over 40% of the $PM_{2.5}$ mass concentration. Additionally, the high-salt dietary habits prevalent in Henan, combined with prolonged exposure to $PM_{2.5}$, may have increased the risk of hypertension and vascular damage [24]. Collectively, these factors contributed to an increased incidence of CVD. In our study, the results of trend tests and restricted cubic spline analysis indicated a dose-response relationship between $PM_{2.5}$ exposure and overall CVD risk. However, this relationship was not strictly linear. These findings suggest that the impact of $PM_{2.5}$ on CVD may exhibit a "*threshold effect*" [25], whereby its toxic effects become markedly pronounced upon exceeding a certain threshold [26]. Evidence suggests that at lower concentrations, cellular autophagy can remove damaged organelles and maintain cellular homeostasis. This compensatory mechanism may partially mitigate the toxic effects of $PM_{2.5}$ on cells, thereby reducing the associated risk [22]. However, when $PM_{2.5}$ concentrations were excessively high, the oxidative stress induced by $PM_{2.5}$ increased in a dose-dependent manner, leading to a significant rise in ROS and cytochrome C expression in vascular endothelial cells. This cascade activates caspase-3, ultimately resulting in DNA fragmentation and cell apoptosis [27]. Furthermore, exposure to high concentrations of $PM_{2.5}$ can trigger multiple programmed cell death pathways in vascular endothelial cells, disrupt tight junction proteins, impair endothelial cell integrity, and consequently cause further damage to cardiovascular tissues, significantly elevating the risk of CVD [28].

Various biological mechanisms have been proposed to explain the association between $PM_{2.5}$ exposure and CVD events, including increased systemic inflammation and oxidative stress, accelerated atherosclerosis, and changes in cardiac autonomic nerve function [29-32]. Exposure to particulate matter is associated with an increased risk of heart disease, primarily through the initiation and promotion of atherosclerosis progression, which underlies the majority of CVD [33]. Exposure to $PM_{2.5}$ has been shown to increase the levels of ROS. The subsequent accumulation of ROS exacerbates oxidative stress, leading to cellular and molecular damage, including DNA, proteins, and lipids [34]. Additionally, exposure to $PM_{2.5}$ promoted the secretion of inflammatory cytokines, leading to endothelial cell activation and a series of pathological changes in the vascular endothelium, thereby fostering the development of CVD [35].

The older population is considered a vulnerable group, susceptible to a range of factors, including immune aging, comorbidities, and environmental influences. Consequently, research focused on this high-risk demographic is of paramount importance. The risk of CVD was elevated in the high-exposure group, whereas it was marginally reduced in the medium and low exposure groups relative to younger individuals. At moderate to low exposure levels, the older population's cumulative physiological compensatory capacity can partially counteract the damage caused by pollution, temporarily maintaining the homeostasis of their internal environment [32]. As previously mentioned, prolonged exposure to high levels results in a substantial rise in the production of reactive oxygen species [34]. This increase surpasses the antioxidant capacity of older people, which is already compromised, thereby exacerbating oxidative damage [36]. Although most studies and expert consensus have reached the conclusion that $PM_{2.5}$ can increase the risk of CVD, it is worth noting that some studies have failed to find a relationship between $PM_{2.5}$ and the risk of CVD [37,38]. These findings underscore the intricate and uncertain dynamic relationship between air pollution and the health outcomes of the older population, emphasizing the need for further research to accurately assess the impact of $PM_{2.5}$ on disease incidence risk and mortality within this demographic.

Limitations

First, the estimated exposure concentrations for each participant were derived from the baseline survey conducted in 2013. During the follow-up period, participants who moved away

XSL•FO
RenderX

were excluded from the analysis, as we lacked comprehensive migration histories for these individuals. This may result in selection bias. Migration patterns in rural areas were typically linked to younger age, higher socioeconomic status, and superior health conditions. This could diminish the gradient between exposure levels and disease risk, potentially leading to an underestimation of the true health impacts of $PM_{2.5}$. Future cohort studies should implement more comprehensive tracking systems to gather exposure data on migrants, thereby minimizing such selection bias as much as possible. Second, we did not evaluate the indoor $PM_{2.5}$ concentration. Ideally, individual exposure should encompass both indoor and outdoor components. Due to the absence of indoor exposure data, we used the outdoor concentration as a proxy variable, which may result in exposure misclassification. In rural areas of China, the use of solid fuels, such as coal and biomass, may exhibit a spatial correlation with outdoor $PM_{2.5}$ concentrations. On the other hand, factors such as rural ventilation practices and house structure can influence the relationship between indoor and outdoor concentrations, thereby increasing the complexity of exposure assessment. Considering these factors, unmeasured indoor exposure was more likely to diminish rather than amplify the risk ratio we reported. Future research can supplement the collection of indoor $PM_{2.5}$ data and develop individual exposure models that integrate indoor and outdoor monitoring to more accurately analyze the impact of indoor pollution sources on individual exposure. Third, we did not adjust for any other key air pollutants in the model. Both NO_2 and SO_2 were strong respiratory irritants that severely damage the respiratory and cardiovascular systems [39]. The interaction with $PM_{2.5}$ may result in a synergistic effect. Owing to the lack of sufficient precise data on air pollution exposure at that time, it may have affected our interpretation of the specific effects of $PM_{2.5}$. However, there was a moderate to high spatial correlation among air pollutants, and including all of them in a multivariate model may lead to collinearity issues [40]. In the analysis of multipollutant models, the effect of $PM_{2.5}$ was typically the most robust [31]. Simultaneously, $PM_{2.5}$ possesses direct cardiovascular toxicity, and its role as a key risk factor is independent. Future research could collect more comprehensive pollutant data at individual exposure levels to more accurately assess the health effects of $PM_{2.5}$. Finally, although we adjusted for several potential confounding factors in the multivariate model, it was impossible to completely eliminate confounding bias due to the presence of unknown and unmeasured residual confounding factors.

Despite these limitations, our research was enhanced by using data from a large-scale, population-based prospective cohort study that featured a 10-year follow-up period and high-quality outcome assessments. Short-term fluctuations in $PM_{2.5}$ can have detrimental health effects, yet long-term exposure may have more significant clinical health implications on CVD morbidity and mortality, as individuals are usually exposed to higher levels of air pollution over an extended period [31]. One significant advantage of this study is that it focuses on rural areas in China, covering a wide range of $PM_{2.5}$ concentrations (ranging from 35.81 to 94.91 $\mu g/m^3$), and conducts risk assessments for different exposure windows of specific CVD, thereby laying a solid scientific foundation for evaluating the exposure-response relationship. The generalizability of the results of this study to other rural populations requires careful consideration. The prospective design, large sample size, and detailed assessment of exposure and confounding factors have enhanced the reliability of the conclusions. Therefore, it is likely that they are applicable to rural populations facing similar environments. However, when extending the results to rural populations with significantly different pollution sources and lifestyles, caution is necessary. Future research to validate our model in other rural environments will help confirm the external validity of these associations.

Conclusions

Our research findings indicate that long-term exposure to $PM_{2.5}$ was significantly associated with an elevated risk of CVD among rural populations, and this association exhibited regional variations. In regions with high levels of $PM_{2.5}$ pollution, comprehensive measures and strategies aimed at reducing air pollution and enhancing public awareness of self-protection should be implemented to mitigate the associated disease risks. This discovery could enhance our understanding of the potential public health risks associated with $PM_{2.5}$ exposure and underscore the important role of environmental governance in promoting public health outcomes. Furthermore, future research efforts should concentrate on clarifying the impacts of $PM_{2.5}$ exposure on the health of various population groups and the underlying mechanisms, thereby contributing to the development of comprehensive intervention measures to mitigate the negative effects of air pollution on public health.

Acknowledgments

The corresponding author would like to express his gratitude to the staff of the disease control departments in various locations of the research site for their assistance in the data collection and quality control processes during the course of the project.

Funding

This study was supported by the CAMS Innovation Fund for Medical Sciences (2021-I2M-1-037), the National Natural Science Foundation of China (82473697 and 82373647), and the National Key Research and Development Program of China (2024YFC2311201).

XSL•FO
RenderX

Data Availability

The datasets used and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LG designed the study. LG, XC, and HX coordinated the study implementation and management. YZ, HX, JD, YD, and YH were responsible for laboratory testing. YH, LS, JH, YD, ZL, JL, and YZ contributed to the field investigation and quality control. WD, AH, JL, ZL, FL, SY, ZX, BZ, JY, RL, FS, and YL contributed to data collection. YZ, LG, and XC did data analyses and wrote the report. YZ, HX, YD, HL, JD, QJ, LG, and XC participated in the data interpretation. YZ, HX, XC, and LG verified the data. All authors contributed to the review and revision and have seen and approved the final version of the manuscript.
The second corresponding author for this paper is HX, Institute of Pathogen Biology, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 16 Tianrong Street, Beijing 102629, China (email: xinhenan@ipbcams.ac.cn; phone number: +86 18519103773).

Conflicts of Interest

None declared.

Multimedia Appendix 1
The nonlinear relationship between particulate matter with aerodynamic diameter ≤2.5 μm and cardiovascular and cerebrovascular diseases and the exposure levels of the older adult population.
[DOCX File, 62 KB - publichealth_v11i1e81218_app1.docx]

Checklist 1
STROBE cohort checklist.
[PDF File, 85 KB - publichealth_v11i1e81218_app2.pdf]

References

1. de Bont J, Jaganathan S, Dahlquist M, Persson Å, Stafoggia M, Ljungman P. Ambient air pollution and cardiovascular diseases: an umbrella review of systematic reviews and meta-analyses. J Intern Med 2022 Jun;291(6):779-800. [doi: 10.1111/joim.13467] [Medline: 35138681]
2. Zhang Y, Wang D, Xu R, et al. Moderating effect of green space on relationship between atmospheric particulate matter and cardiovascular and cerebrovascular disease mortality in Ningxia, China. Environ Res 2025 Apr;270:120931. [doi: 10.1016/j.envres.2025.120931]
3. 2019 world air quality report: region & city PM2.5 ranking. IQAir. 2019. URL: https://www.iqair.com/dl/pdf-reports/2019-World-Air-Report-V8-20200318.pdf [accessed 2025-11-20]
4. Wolf K, Hoffmann B, Andersen ZJ, et al. Long-term exposure to low-level ambient air pollution and incidence of stroke and coronary heart disease: a pooled analysis of six European cohorts within the ELAPSE project. Lancet Planet Health 2021 Sep;5(9):e620-e632. [doi: 10.1016/S2542-5196(21)00195-9] [Medline: 34508683]
5. Khoshakhlagh AH, Mohammadzadeh M, Gruszecka-Kosowska A, Oikonomou E. Burden of cardiovascular disease attributed to air pollution: a systematic review. Global Health 2024 May 3;20(1):37. [doi: 10.1186/s12992-024-01040-0] [Medline: 38702798]
6. Henning RJ. Particulate matter air pollution is a significant risk factor for cardiovascular disease. Curr Probl Cardiol 2024 Jan;49(1 Pt B):102094. [doi: 10.1016/j.cpcardiol.2023.102094] [Medline: 37734693]
7. Cesaroni G, Forastiere F, Stafoggia M, et al. Long term exposure to ambient air pollution and incidence of acute coronary events: prospective cohort study and meta-analysis in 11 European cohorts from the ESCAPE Project. BMJ 2014 Jan 21;348:f7412. [doi: 10.1136/bmj.f7412] [Medline: 24452269]
8. Nethery RC, Josey K, Gandhi P, et al. Air pollution and cardiovascular and thromboembolic events in older adults with high-risk conditions. Am J Epidemiol 2023 Aug 4;192(8):1358-1370. [doi: 10.1093/aje/kwad089] [Medline: 37070398]
9. Bhatnagar A. Cardiovascular effects of particulate air pollution. Annu Rev Med 2022 Jan 27;73:393-406. [doi: 10.1146/annurev-med-042220-011549] [Medline: 34644154]
10. Mallah MA, Soomro T, Ali M, et al. Cigarette smoking and air pollution exposure and their effects on cardiovascular diseases. Front Public Health 2023;11:967047. [doi: 10.3389/fpubh.2023.967047] [Medline: 38045957]
11. Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD]
12. Gao L, Lu W, Bai L, et al. Latent tuberculosis infection in rural China: baseline results of a population-based, multicentre, prospective cohort study. Lancet Infect Dis 2015 Mar;15(3):310-319. [doi: 10.1016/S1473-3099(14)71085-0] [Medline: 25681063]

XSL•FO
RenderX

13. China statistical yearbook—2010. National Bureau of Statistics. URL: https://www.stats.gov.cn/sj/ndsj/2010/indexeh.htm [accessed 2025-11-20]
14. Criteria of weight for adults (WS/T428-2013). Chinese Standard GB/T, GBT, GB. 2013. URL: https://www.chinesestandard.net/PDF/English.aspx/WST428-2013 [accessed 2025-11-20]
15. Ji JS, Zhu A, Bai C, et al. Residential greenness and mortality in oldest-old women and men in China: a longitudinal cohort study. Lancet Planet Health 2019 Jan;3(1):e17-e25. [doi: 10.1016/S2542-5196(18)30264-X] [Medline: 30654864]
16. Yao Y, Liu L, Guo G, Zeng Y, Ji JS. Interaction of Sirtuin 1 (SIRT1) candidate longevity gene and particulate matter (PM2.5) on all-cause mortality: a longitudinal cohort study in China. Environ Health 2021 Mar 14;20(1):25. [doi: 10.1186/s12940-021-00718-x] [Medline: 33715628]
17. Tracking Air Pollution. URL: http://tapdata.org.cn/ [accessed 2025-11-19]
18. Lao XQ, Guo C, Chang LY, et al. Long-term exposure to ambient fine particulate matter (PM2.5) and incident type 2 diabetes: a longitudinal cohort study. Diabetologia 2019 May;62(5):759-769. [doi: 10.1007/s00125-019-4825-1] [Medline: 30706081]
19. Krittanawong C, Qadeer YK, Hayes RB, et al. PM2.5 and cardiovascular health risks. Curr Probl Cardiol 2023 Jun;48(6):101670. [doi: 10.1016/j.cpcardiol.2023.101670] [Medline: 36828043]
20. Yusuf S, Joseph P, Rangarajan S, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet 2020 Mar;395(10226):795-808. [doi: 10.1016/S0140-6736(19)32008-2]
21. Hystad P, Larkin A, Rangarajan S, et al. Associations of outdoor fine particulate air pollution and cardiovascular disease in 157 436 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet Planet Health 2020 Jun;4(6):e235-e245. [doi: 10.1016/S2542-5196(20)30103-0] [Medline: 32559440]
22. Basith S, Manavalan B, Shin TH, et al. The impact of fine particulate matter 2.5 on the cardiovascular system: a review of the invisible killer. Nanomaterials (Basel) 2022 Aug 2;12(15):15. [doi: 10.3390/nano12152656] [Medline: 35957086]
23. Hailiang W, Jifei Z, Huayou L, Lele L, Tian FT, Haofei L. Evaluation of underground water resources utilization in coal mining area sinkholes in Henan province based on AHP-set pair analysis method. Yellow River 2024;46(10):72-77. [doi: 10.3969/j.issn.1000-1379.2024.10.013]
24. Wang W, Li J, Liu Y, et al. Spatiotemporal trends and ecological determinants of cardiovascular mortality among 2844 counties in mainland China, 2006-2020: a Bayesian modeling study of national mortality registries. BMC Med 2022 Nov 30;20(1):467. [doi: 10.1186/s12916-022-02613-9] [Medline: 36451190]
25. Yin P, Brauer M, Cohen A, et al. Long-term fine particulate matter exposure and nonaccidental and cause-specific mortality in a large national cohort of Chinese men. Environ Health Perspect 2017 Nov 7;125(11):117002. [doi: 10.1289/EHP1673] [Medline: 29116930]
26. Pope CA III, Coleman N, Pond ZA, Burnett RT. Fine particulate air pollution and human mortality: 25+ years of cohort studies. Environ Res 2020 Apr;183:108924. [doi: 10.1016/j.envres.2019.108924] [Medline: 31831155]
27. Wang Y, Zhong Y, Liao J, Wang G. PM2.5-related cell death patterns. Int J Med Sci 2021;18(4):1024-1029. [doi: 10.7150/ijms.46421]
28. Han X, Zhuang Y. PM2.5 induces autophagy-mediated cell apoptosis via PI3K/AKT/mTOR signaling pathway in mice bronchial epithelium cells. Exp Ther Med 2021 Jan;21(1):1. [doi: 10.3892/etm.2020.9433] [Medline: 33235610]
29. Pope CA 3rd, Burnett RT, Thurston GD, et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. Circulation 2004 Jan 6;109(1):71-77. [doi: 10.1161/01.CIR.0000108927.80044.7F] [Medline: 14676145]
30. Suwa T, Hogg JC, Quinlan KB, Ohgami A, Vincent R, van Eeden SF. Particulate air pollution induces progression of atherosclerosis. J Am Coll Cardiol 2002 Mar 20;39(6):935-942. [doi: 10.1016/s0735-1097(02)01715-1] [Medline: 11897432]
31. Brook RD, Rajagopalan S, Pope CA 3rd, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation 2010 Jun 1;121(21):2331-2378. [doi: 10.1161/CIR.0b013e3181dbece1] [Medline: 20458016]
32. Fiordelisi A, Piscitelli P, Trimarco B, Coscioni E, Iaccarino G, Sorriento D. The mechanisms of air pollution and particulate matter in cardiovascular diseases. Heart Fail Rev 2017 May;22(3):337-347. [doi: 10.1007/s10741-017-9606-7] [Medline: 28303426]
33. Yu Y, Li H, Gao F. Research progress on the relationship between PM2.5 and cardiovascular diseases and their prevention and treatment. J Integr Tradit Chin West Med Cardiovasc Cerebrovasc Dis 2024;22(20):3718-3723 [FREE Full text]
34. Liu K, Hua S, Song L. PM2.5 exposure and asthma development: the key role of oxidative stress. Oxid Med Cell Longev 2022;2022(1):3618806. [doi: 10.1155/2022/3618806] [Medline: 35419163]
35. van Ierssel SH, Jorens PG, Van Craenenbroeck EM, Conraads VM. The endothelium, a protagonist in the pathophysiology of critical illness: focus on cellular markers. Biomed Res Int 2014;2014:985813. [doi: 10.1155/2014/985813] [Medline: 24800259]
36. Xu MM, Jia YP, Li GX, et al. Relationship between ambient fine particles and ventricular repolarization changes and heart rate variability of elderly people with heart disease in Beijing, China. Biomed Environ Sci 2013 Aug;26(8):629-637. [doi: 10.3967/0895-3988.2013.08.001] [Medline: 23981548]

XSL•FO
RenderX

37. Willocks LJ, Bhaskar A, Ramsay CN, et al. Cardiovascular disease and air pollution in Scotland: no association or insufficient data and study design? BMC Public Health 2012 Mar 22;12:227. [doi: 10.1186/1471-2458-12-227] [Medline: 22440092]
38. Wang M, Beelen R, Stafoggia M, et al. Long-term exposure to elemental constituents of particulate matter and cardiovascular mortality in 19 European cohorts: results from the ESCAPE and TRANSPHORM projects. Environ Int 2014 May;66:97-106. [doi: 10.1016/j.envint.2014.01.026] [Medline: 24561271]
39. Sifan T, Tonglei G, Chen C, et al. The impact of atmospheric gaseous pollutants on the incidence of pulmonary tuberculosis in latent tuberculosis infection patients in rural areas. Chin J Tuberc Control 2024;46(6):654-663. [doi: 10.19982/j.issn.1000-6621.20240061]
40. Billionnet C, Sherrill D, Annesi-Maesano I, GERIE study. Estimating the health effects of exposure to multi-pollutant mixture. Ann Epidemiol 2012 Feb;22(2):126-141. [doi: 10.1016/j.annepidem.2011.11.004] [Medline: 22226033]

Edited by A Mavragani, T Sanchez; submitted 24.07.25; peer-reviewed by T Guo, Y Wu; revised version received 14.10.25; accepted 05.11.25; published 28.11.25.

Please cite as:
Zhao Y, Cao X, Du J, He A, Liang J, Duan W, Di Y, He Y, Feng B, Shen L, Huang J, Li Z, Liang J, Li H, Liu Z, Liu F, Yang S, Xu Z, Zhang B, Yan J, Liang Y, Liu R, Shen F, Jin Q, Xin H, Gao L
Long-Term Ambient Air Pollution Exposure and the Risk of Cardiovascular and Cerebrovascular Diseases in Rural Chinese Populations: 10-Year Follow-Up of a Multicenter Prospective Cohort Study
JMIR Public Health Surveill 2025;11:e81218
URL: https://publichealth.jmir.org/2025/1/e81218
doi:10.2196/81218

© Yaqi Zhao, Xuefang Cao, Jiang Du, Aiwei He, Jun Liang, Weitao Duan, Yuanzhi Di, Yijun He, Boxuan Feng, Linyu Shen, Juanjuan Huang, Zihan Li, Jianguo Liang, Hongzhi Li, Zisen Liu, Fang Liu, Shumin Yang, Zuhui Xu, Bin Zhang, Jiaoxia Yan, Yanchun Liang, Rong Liu, Fei Shen, Qi Jin, Henan Xin, Lei Gao. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Data Parameters From Participatory Surveillance Systems in Human, Animal, and Environmental Health From Around the Globe: Descriptive Analysis

Carrie McNeil[1], MPH, DVM; Nomita Divi[1], BSc, MSPH; Charles Thomas Bargeron IV[2], MS; Andrea Capobianco Dondona[3], MSc, DVM; Kacey C Ernst[4], MPH, PhD; Angela S Gupta[5], MA; Olukayode Fasominu[6], MPH, MD; Lucy Keatts[7], BVetMed, MVSc, MRCVS; Terra Kelly[8], DVM, Dipl. ACZM, PhD; Onicio B Leal Neto[9], PhD; May O Lwin[10], MBA, PhD; Mvuyo Makhasi[11], BSc, MSc; Eric Beda Mutagahywa[12], MSc; Diego Montecino-Latorre[7], BVSc, MPVM, PhD; Sarah Olson[7], PhD; Pranav S Pandit[13], BVSc & AH, MPVM, PhD; Daniela Paolotti[14], PhD; Matt C Parker[15], MPH; Muhammad Haiman Samad[16], MPH; Kara Sewalk[17], MPH; Anita Sheldenkar[16], MS; Lertrak Srikitjakarn[18], DrMedVet; Channé Suy Lan[19], MCA; Michael Wilkes[20], MD, PhD; Terdsak Yano[21], MS, DVM; Mark Smolinski[1], MPH, MD

[1]Ending Pandemics, San Francisco, CA, United States

[2]Center for Invasive Species & Ecosystem Health, University of Georgia, Tifton, GA, United States

[3]Food and Agriculture Organization of United Nations, Rome, Italy

[4]Department of Epidemiology and Biostatistics, College of Public Health, University of Arizona, Tucson, AZ, United States

[5]College of Extension, University of Minnesota, Rochester, MN, United States

[6]Volte Health Systems Limited, Abuja, Nigeria

[7]Health Program, Wildlife Conservation Society, Bronx, NY, United States

[8]EpiEcos LLC, Flagstaff, AZ, United States

[9]College of Public Health, Department of Epidemiology and Biostatistics, University of Arizona, Tuscon, AZ, United States

[10]Wee Kim Wee School of Communication & Information, Nanyang Technological University, Singapore, Singapore

[11]Centre for Respiratory Diseases and Meningitis, National Health Laboratory Service, National Institute for Communicable Diseases, Johannesburg, South Africa

[12]SACIDS Foundation for One Health, Morogoro, United Republic of Tanzania

[13]Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, Davis, CA, United States

[14]Institute for Scientific Interchange, Turin, Italy

[15]Open Dream, Bangkok, Thailand

[16]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

[17]Boston Children's Hospital, Boston, MA, United States

[18]PODD Centre, Chiang Mai University, Chiang Mai, Thailand

[19]Kawsang Co, LTD, Phnom Penh, Cambodia

[20]School of Medicine, University of California, Davis, Davis, CA, United States

[21]Faculty of Veterinary Medicine, Chiang Mai University, Chiang Mai University, Thailand

Corresponding Author:
Carrie McNeil, MPH, DVM
Ending Pandemics
870 Market Street, Suite 528
San Francisco, CA, 94102
United States
Phone: 1 9169522185
Email: carrie.s.mcneil@gmail.com

Abstract

Background: Emerging pathogens and zoonotic spillover highlight the need for One Health surveillance to detect outbreaks as early as possible. Participatory surveillance empowers communities to collect data at the source on the health of animals,

XSL•FO
RenderX

people, and the environment. Technological advances increase the use and scope of these systems. This initiative sought to collate information from active participatory surveillance systems to better understand parameters collected across the One Health spectrum.

Objective: This study aims to develop a compendium of One Health data parameters by examining participatory surveillance systems active in 2023. The expected outcomes of the compendium were to pinpoint specific parameters related to human, animal, and environmental health collected globally by participatory surveillance systems and to detail how each parameter is collected. The compendium was designed to help understand which parameters are currently collected and serve as a reference for future systems and for data standardization initiatives.

Methods: Contacts associated with the 60 systems identified through the One Health Participatory Surveillance System Map were invited by email to provide specific data parameters, methodologies used for data collection, and parameter-specific considerations. Information was received from 38 (63%) active systems. Data were compiled into a searchable spreadsheet-based compendium organized into 5 sections: general, livestock, wildlife, environmental, and human parameters. An advisory group comprising experts in One Health participatory surveillance reviewed the collected parameters, refined the compendium structure, and contributed to the descriptive analysis.

Results: A comprehensive compendium of data parameters from a diverse array of single-sector and multisector participatory surveillance systems was collated and reviewed. The compendium includes parameters from 38 systems used in Africa (n=3, 8%), Asia (n=9, 24%), Europe (n=12, 32%), Australia (n=3, 8%), and the Americas (n=12, 32%). Almost one-third of the systems (n=11, 29%) collect data across multiple sectors. Many (n=17, 45%) focus solely on human health. Variations in data collection techniques were observed for commonly used parameters, such as demographics and clinical signs or symptoms. Most human health systems collected parameters from a cohort of users tracking their own health over time, whereas many wildlife and environmental systems incorporated event-based parameters.

Conclusions: Several participatory surveillance systems have already adopted a One Health approach, enhancing traditional surveillance by identifying shared health threats among animals, people, and the environment. The compendium reveals substantial variation in how parameters are collected, underscoring the need for further work in system interoperability and data standards to allow for timely data sharing across systems during outbreaks. Parameters collated from across the One Health spectrum represent a valuable resource for informing the development of future systems and identifying opportunities to expand existing systems for multisector surveillance.

(JMIR Public Health Surveill 2025;11:e55356) doi:10.2196/55356

KEYWORDS

participatory surveillance; One Health; citizen science; community-based surveillance; digital disease detection; environmental health; wildlife health; livestock health; human health; data standards

Introduction

Background

In a world where anthropogenic landscape change drives zoonotic spillover, climate change exacerbates vulnerabilities in food security, and diseases further threaten biodiversity, effective and timely One Health surveillance mechanisms are imperative. The One Health United Nations Joint Action Plan (2022-2026) [1], Africa's Centers for Disease Control Event-Based Surveillance (EBS) guidance [2], and World Health Organization's "Defining Collaborative Surveillance" [3] underscore the critical need to integrate data across the animal, human, plant, and environmental health sectors. This paper examines the range of data parameters collected across the One Health spectrum through participatory surveillance methodologies.

One Health Surveillance

The One Health approach recognizes the interdependencies of human, animal, and ecosystem health, aiming "to sustainably balance and optimize the health of people, animals, and ecosystems" [4]. Numerous national and subnational systems have undertaken initiatives for data sharing to develop One Health strategies, with support from organizations such as the United Nations, World Bank, and United States Agency for International Development (USAID) [5,6]. However, challenges remain in the implementation of One Health initiatives due to years of siloed professional training, budgets, data system development, data structures, and policies [7]. Rabies and influenza viruses have served as key drivers of multisectoral, collaborative surveillance and offer valuable models for advancing One Health approaches to address multiple pathogens. The need for the inclusion of plant health has more recently come to the forefront due to its critical role in food safety, food security, and ecosystem health [8].

Historically, traditional disease surveillance systems have largely been led by government entities and are often siloed, with a primary focus on human and livestock health. There is growing awareness of the need to harness more wildlife and environmental health event information to provide early warning of broader health threats facing both humans and animals. The Africa's Centers for Disease Control EBS guidance emphasizes the integration of both formal and informal data sources, including hotlines and community-based EBS, to enhance outbreak detection [2]. To strengthen participatory surveillance systems, greater efforts are needed to capture data on putative drivers of pathogen spillover and disease emergence, including landscape change, environmental and climatic variables, and

XSL•FO
RenderX

activities at the human-animal-environmental interface that bring animals and humans into contact with each other [9]. While previous literature has examined One Health tools and individual surveillance systems, no study to date has comprehensively documented the range of parameters collected in One Health participatory surveillance systems [7,10-12].

Participatory Surveillance

Participatory surveillance involves a bidirectional process of receiving and transmitting data for actionable outcomes through the direct engagement of the target population [13]. This approach is increasingly used across multiple sectors to enhance early detection of and rapid response to emerging infections within a One Health framework. By engaging a wide variety of users including the general public, health care workers, rangers, farmers, and outdoor enthusiasts, participatory surveillance systems can provide early warning of potential outbreaks through crowdsourcing self-reported data in near real-time [14]. This active surveillance strategy empowers communities to act as the "eyes and ears" for detecting health threats at the interfaces where animals, humans, and environmental factors converge, creating opportunities for disease emergence.

Participatory surveillance systems play a crucial role in providing evidence-based feedback to users, enabling them to take timely actions to mitigate potential threats and facilitating connections with local resources for prevention and control. In the context of One Health, participatory surveillance extends beyond a human-centered lens to encompass plants, wildlife, livestock, and a multitude of environmental factors—habitats for disease vectors and water and air quality. The organizational structure of system users can vary widely, ranging from the general public to members of sector-specific associations or community alliances. For more structured community-liaison models, feedback is directed to the user group, empowering them to take action on behalf of the communities they serve.

Participatory surveillance has emerged as a powerful tool for enhancing traditional epidemiologic surveillance while fostering community engagement [15]. From reporting disease incidence to promoting mitigation behaviors, participatory surveillance has showcased its versatility across various infectious disease contexts, including influenza, cholera, COVID-19, and Zika [16-22]. One of the key benefits of participatory surveillance is its capacity to address and surmount many of the limitations inherent to traditional surveillance systems. In resource-constrained settings, conventional surveillance efforts often rely on overburdened health care and diagnostic systems and struggle to access difficult-to-reach communities, resulting in reporting delays or gaps. By empowering communities, participatory surveillance enables a proactive role in health monitoring, facilitating both prompt and comprehensive reporting of disease cases [23,24].

Ending Pandemics, a US nonprofit organization working to detect, verify, and contain outbreaks faster (and the supporter of the study presented here) developed the International Workshops on Participatory Surveillance (IWOPS), four of which have been held since 2012 in the United States, the Netherlands, Australia, and Cambodia [25]. IWOPS convened the creators, implementors, and key advocates to advance participatory surveillance approaches and strengthen partnerships among these systems. In 2016, the IWOPS community established a minimum dataset among systems collecting self-reported information to monitor the annual risk of influenza so that a common set of data standards could be developed. As a result, deidentified and disaggregated data are shared on a public platform, Global Flu View, for the purpose of improving situational awareness [26].

While Global Flu View represents one example of what is possible when participatory surveillance systems can "speak" to each other, data standards across the entire spectrum of parameters encompassing One Health are lacking. The One Health Participatory Surveillance Data Parameters Compendium (Multimedia Appendix 1) is a snapshot into the world of participatory surveillance for human, animal, and environmental health. The compendium approach was used as it not only illustrates the breadth and depth of what is being used in currently operating participatory surveillance systems across the globe, but it can serve to inform the needs of those contemplating the development of a participatory surveillance system or expansion of an existing system. As the compendium provides information on how data parameters are collected, it is poised to guide efforts to incorporate additional health sectors, adapt data parameters to align with other systems, and foster collaborations and data sharing. In the future, it could play a pivotal role in advancing data sharing, through informing the development of international data standards.

Methods

Development of Advisory Group

A multisector advisory group was assembled to guide the development of the compendium, support data collection, and provide cross-sectoral reviews encompassing animal, environmental, and public health sector perspectives. Selection criteria for advisory group members included expertise in participatory surveillance systems and integrating data from participatory systems into traditional surveillance systems. Subteams of experts in the areas of human, wildlife, livestock, and environmental health (including plants and vectors) were established. The experts represented a diverse range of public and government organizations, universities or research centers, nongovernmental organizations, and private sector entities. Geographically, they were affiliated with organizations in Asia (n=7), Africa (n=3), Europe (n=2), North America (n=10), and South America (n=2).

Ethical Considerations

The study was reviewed by the University of Arizona Institutional Review Board and deemed not to involve human participants as defined by US Department of Health and Human Services and Food and Drug Administration regulations research (IRB protocol# 00002653). Within the body of the email, individuals were provided a disclosure statement. All participants were notified that they would receive a copy of the compendium, and their system would be listed in the acknowledgments. Data included in the compendium included terms and categories of parameters collected by systems and

XSL•FO
RenderX

did not include any identifiable information. Participants were not compensated, and participation was voluntary.

Data Collection

An invitation letter was emailed to a closed sample of representatives of 60 systems, using contact information from the One Health Participatory Surveillance map and 2022 landscape analysis [13]. Follow-up emails were sent twice, as needed, by members of the compendium advisory group to encourage responses from system representatives who did not reply initially. Emails described the goal for developing the One Health participatory surveillance compendium and requested respondents provide data parameters from their systems, methods for collection, and any special considerations for specific parameters using an attached template. The template had been used for parameter collection for IWOPS IV [27]. System representatives were offered the option to provide the requested information in other formats, such as a link to web-based systems allowing authors to independently collect relevant questions and parameters.

Data collected were in English. Parameters had already been collected for IWOPS IV in 2022 from: AfyaData, Cambodia 115 Hotline, EDDMapS or Wild Spotter, Forest First Pest Detector, Kidenga, MoBuzz, Outbreaks Near Me, One Health App Philippines, Participatory One Health Digital Disease Detection (PODD), SickSense/Open Dream, SURPRISE, World Animal Health Information System, SMART for Health, and WildAlert (formally known as the Wildlife Morbidity and Mortality Event Alert System) [27]. For these systems, emails were sent to confirm the parameters collected and to allow for revisions.

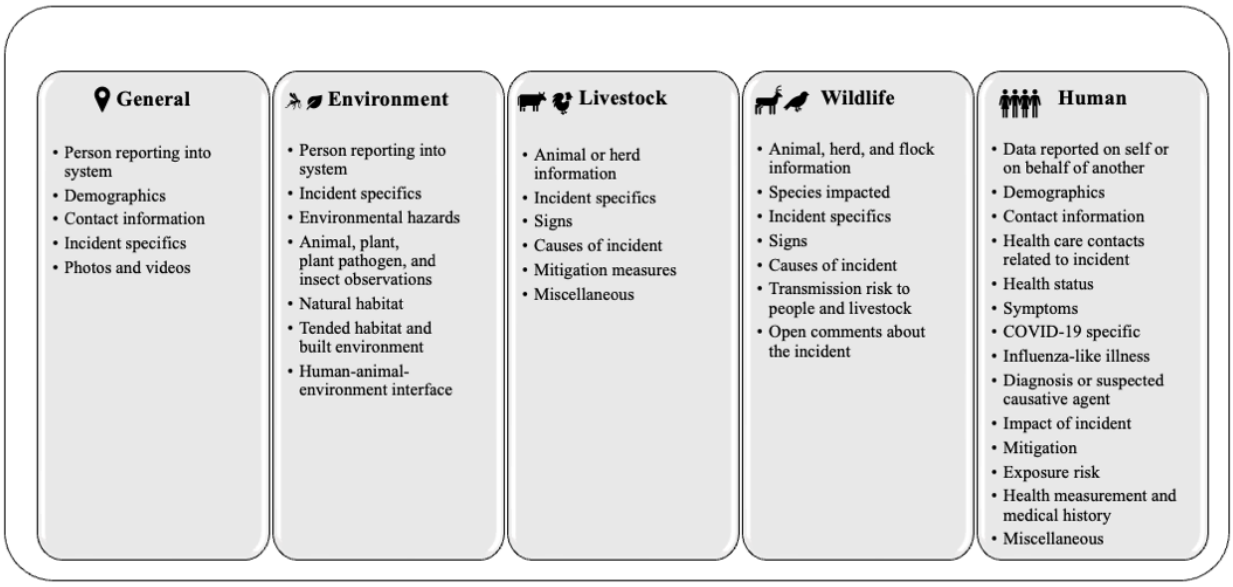
Out of the 60 systems contacted, 41 systems (68%) responded, and 3 systems (5%) responded they were no longer active and were not included in this analysis. Completed responses were received from 36/60 (60%) active systems; data were collected directly from publicly accessible apps or websites from 2/60 (3%) systems. 38 (63%) responses include the 9 country-level programs of InfluenzaNet. Seven of the 9 InfluenzaNet programs submitted 1 set of parameters and 2 (GrippeWeb and Hälsorapport) provided parameters individually.

Data Analysis

The collected parameters were organized into 5 sections in the compendium: general, livestock, wildlife, environmental (including plants and vectors), and human health. Subcategories within each section were reviewed and refined through group discussions with the advisory group subteams specializing in livestock, wildlife, human, and environmental health. The focus of these discussions was to ensure usability, clarity, and accuracy. The human health advisory group subteam determined that parameters would be limited to those they determined directly relevant to infectious disease surveillance. The compendium focuses on parameters and does not describe structure, such as decision trees.

Advisory group representatives reviewed the parameters and overall organization of the compendium and shared perspectives on key aspects to emphasize in the discussion and conclusion of this paper. Descriptive analyses were performed to examine the categories of parameters collected, the number and types of systems represented across sectors, and advisory group feedback on the parameters. Figure 1 outlines the compendium's format, and the categories of data parameters collected.

Figure 1. Categories of data parameters included in the 2023 One Health Participatory Surveillance Compendium.



Results

Participatory Surveillance Systems Represented in the Compendium

Data parameters were collected in English from across 38 participatory surveillance systems in Africa (n=3, 8%), Asia (n=9, 24%), Europe (n=12, 32%), Australia (n=3, 8%), and the Americas (n=12, 32%); only SMART for Health data were collected in several continents (Table 1). Approximately one-third of the systems (n=11, 29%) collect data from more than 1 sector, while the majority focus on a single sector. Many programs (n=17, 45%) are dedicated solely to human health, with fewer focusing exclusively on environmental or animal

XSL•FO
RenderX

health (n=10, 26%). Most users contributing data to these systems are members of the general public (n=30, 79%), though other user groups include trained volunteers (n=6, 16%), wildlife rangers (n=2, 5%), farmers (n=2, 5%), wildlife rehabilitators and biologists (n=2, 5%), informal health care workers (n=1, 3%), and students (n=1, 3%). For example, SURPRISE is a system in which health care workers report on their own health, (n=1, 3%). Systems requesting weekly data reporting (n=16, 42%) primarily focus on human health, though some offer users the flexibility to determine the reporting frequency. Most of the remaining systems (n=20, 53%) allow users to set their reporting frequency.

XSL•FO
RenderX

Table 1. Description of global One Health participatory surveillance systems identified as active in 2023 including sectors included, location, user type, reporting frequency, and technology.

Systems	Sector	Countries	Type of user	Frequency of reporting	User technology
AfyaData	Animal, human, and environment	Tanzania, Mozambique	General public, trained volunteer, farmer, ranger, researchers, and students	Determined by user	SMS and mobile app
Arizona Department of Fish and Game Hotline	Animal and environment	United States	General public	Determined by user	Hotline and email or website
Cambodia 115 Hotline	Animal and human	Cambodia	General public and health officer	Weekly, determined by user	Hotline
California West Nile Bird Reporting	Animal	United States	General public	Determined by user	Hotline and email or website
Colab	Animal, human, and environment	Brazil	General public	Determined by user	Mobile app, hotline, and email or website
CoughWatch SA	Human	South Africa	General public	Weekly	Email or website
DisApp	Environment	India	General public	Determined by user	Mobile app
EDDMapS/Wild Spotter	Animal and environment	United States and Canada	General public and trained volunteer	Determined by user	Mobile app and email or website
eWHIS[a]	Animal and environment	Australia	Trained volunteer	Monthly, determined by user	Hotline and email or website
FluTracking Australia	Human	Australia	General public	Weekly	Email or website
FluTracking HongKong	Human	Hong Kong	General public	Weekly	Email or website
Forest Pest First Detector	Environment	United States	General public	Determined by user	Mobile app and email or website
Guardians of Health	Human	Brazil	General public	Daily	Mobile app
Garden Wildlife Health Project	Animal	Great Britain	General public	Determined by user	Email or website
iMammalia	Animal	Europe	General public	Daily	Mobile app and email or website
InfluenzaNet programs					
FluSurvey.net	Human	United Kingdom	General public	Weekly	Email or website
Grippenet.ch	Human	Switzerland	General public	Weekly	Email or website
Grippenet.fr	Human	France	General public	Weekly	Email or website
GrippeWeb	Human	Germany	General public	Weekly	Email or website
Hälsorapport, Health Report Sweden	Human	Sweden	General public	Weekly	Email or website
Infectieradar.be	Human	Belgium	General public	Weekly	Email or website
Infectieradar.nl	Human	Netherlands	General public	Weekly	Email or website
Influmeter.dk	Human	Denmark	General public	Weekly	Email or website
InfluWeb	Human	Italy	General public	Weekly	Email or website
I-TICK	Environment	United States	General public	Determined by user	Mobile app
Kidenga	Environment and human	United States and Puerto Rico	General public	Weekly	Mobile app
MoBuzz+	Environment and human	Sri Lanka	General public, health authorities	Determined by user	Mobile app and email or website
MosApp	Environment	India	Formal, informal health care workers	Determined by user	Mobile app
Outbreaks Near Me	Human	United States, Canada, Mexico	General public	Weekly, Determined by User	SMS and email or website

XSL•FO
RenderX

Systems	Sector	Countries	Type of user	Frequency of reporting	User technology
PODD[b]	Animal, human, and environment	Thailand	General public, trained volunteer, farmer, and local government	Fortnightly	Mobile app
SMART for Health	Animal and environment	Laos, Cambodia, Vietnam, and Guatemala	Rangers	Determined by user	Mobile app
SickSense or Sabaidee	Human	Thailand	General public	Daily, Weekly, Determined by user	Mobile app
SURPRISE+	Human	Switzerland	Health care workers	Adaptive according to epidemiology	SMS and email or website
Tamil Nadu Population Health Registry	Human	India	Trained volunteer	Daily	Mobile app and paper
Tick Tag Go	Environment	United States	General public	Determined by user	Mobile app and email or website
Weedspotters Australia	Environment	Australia	Trained volunteer	Determined by user	Mobile app
Wild Health Net Ebola	Animal and environment	Republic of Congo	Wildlife biologists	Determined by user	SMS, mobile app, and email or website
WildAlert	Animal	United States	Wildlife rehabilitators, state wildlife agencies, and university researchers	Weekly, determined by user	Email or website

[a]eWHIS: electronic wildlife health information system.

[b]PODD: participatory one health digital disease detection.

General Parameters

General parameters include those not specific to any sector—such as user demographics and contact information, incident location, and options for photo or video uploads. Systems with regular users often collect user demographics and contact information during the initial login, eliminating the need for repeated entry on subsequent logins. Substantial variation exists on how demographic data are collected—particularly as related to gender, ethnicity, education, and profession. In contrast, contact information, such as email and phone numbers, are collected uniformly across most systems. Incident location is commonly collected using automated geolocation. Additionally, some systems include mechanisms for photo or video upload to provide supplementary information.

Environmental Health Parameters

Overview

The range of parameters collected related to environmental health is as broad as the field itself—encompassing activities, such as tracking logging, monitoring water contamination, and detecting invasive species and disease vectors. Over half of the systems (8/15, 53%) collecting environmental health data also include a focus on animal and human health. Kidenga collects data on environmental factors including the risk of mosquito exposure alongside human health symptoms. SMART for Health captures key environmental and epidemiological features as well as information on wildlife and livestock. Participatory surveillance systems vary in scale, with some designed for local, relatively small geographic areas and others operating on global platforms, such as iNaturalist. Several vector-focused systems reported that weather information, such as rainfall and temperature, can be added retrospectively by linking it to date, time, and location. Natural habitat data in these systems includes landscape descriptions, while details of tended habitats include descriptions of yards, gardens, and ponds. The built environment is also considered, with systems collecting data on housing structures, water and sanitation, and energy sources, such as fuel and electricity.

Plants

Three systems collect data on plants, including the Early Detection and Distribution Mapping System, which also gathers data on animals. Collected parameters include date or season, high-resolution digital images, information on the reporter, and detailed location information. Date, time, and GPS coordinates are often extracted directly from digital image metadata and autofilled in the report. Other details include ecosystem type, whether the area is tended or untended, seasonality, and additional species, such as host or insect pests.

Species identification is commonly performed by the primary reporter, with photos allowing for additional verification. Species identification is done in 1 reporting field or across multiple fields, capturing details or images of features such as flowers, seeds, leaves, leaf arrangements, and the entire plant. All the plant-focused systems record data on invasive plants and some databases also include information on legal, noxious weed status. High-quality, high-resolution digital images and tightly calibrated geographic locations are noted to be particularly important in these systems.

XSL•FO
RenderX

Vectors

Almost half (7/15, 46.7%) of the environmental systems are explicitly designed to monitor vectors. Notably, the Early Detection and Distribution Mapping System amalgamates vector, animal, and plant surveillance. Among the 7 vector surveillance systems that primarily focus on the environmental sector, 3 exhibit a multisectoral approach. Kidenga and MoBuzz+ bridge the environmental and human domains.

Vector exposure is collected in several ways across systems including incident-specific data with details on the date and precise location of the vector bite incident, specific vector bite information, and the detection of vectors in the vicinity. Often, observations include the perceived number of vectors and their morphology.

Data encompasses information on any formal diagnoses of vector-borne diseases and community prevalence of disease. Vector identification data incorporates visual documentation using photos uploaded by the user. Photos enable experts receiving the data to visually identify and assess vector size, color, and, where possible, species. Vector habitat data includes breeding sites and the detection of larvae, which use the number of measurement containers at a given location and whether they contain larvae. A number of systems integrate the history of recent outdoor activities to better understand the potential exposure of vectors. Vector mitigation data collects information on the use of pesticides, as well as any treatments for controlling or countering vector exposure, and appropriate clothing.

Animal Health Parameters

Wildlife Health

Nine of the participatory surveillance systems collect data on wild animals. SMART for Health includes parameters to record the health of free-ranging livestock that may be present in protected natural habitats. Most of these systems were developed with the goal of syndromic wildlife disease surveillance except for 2 systems— one that was designed to collect data on geographic distribution and abundance of wild mammalian species and another that focuses on early detection of invasive wildlife species [13]. Among these 9 systems, 4 (44%) also collect information on a range of environmental variables. Many variables focus on risk factors or environmental drivers of adverse health events/disease outbreaks in animals as well as humans such as environmental disturbance, or interactions at the wildlife-livestock-human interface. In addition, WildHealthNet—Ebola integrates data on both wildlife and human health parameters specifically for Ebola Virus Disease surveillance.

Common across most of these systems are parameters specific to wildlife incidents, including species or taxa groups of wild animals involved in the event, date, and numbers and location of affected animals. For location information, a few of the systems have the capability to capture specific GPS coordinates as the user digitally enters data or to generate GPS coordinates from an address through geocoding. Typically, systems also include data on the condition of the animal(s) involved in the incident, including abnormal findings in sick, injured, or dead wild animals and a presenting syndrome or clinical classification of these animals. Depending on the system, data are reported at the level of the individual animal or species.

Wildlife health parameters are collected in a variety of formats including dropdown menus, check boxes, and fields with free text. Categorizing free text data for surveillance purposes through natural language processing models has been applied by WildAlert to categorize data for surveillance purposes. Collection of data related to risk of exposure to humans—such as handling of sick wildlife or bats roosting near a home was also included.

Two of the systems also collect data on definitive diagnosis. For systems that integrate clinical wildlife data, this data, if available, is included for both individual animals as well as events. Other parameters among these systems include circumstances surrounding the incident and the potential causes of illness, injury, or death at both individual and species levels. Furthermore, one system also allows the user to enter data on the estimated population size and number of animals at risk for wildlife incidents enabling estimations of disease occurrence among a population.

Livestock Health

Only 4 surveillance systems collect data on livestock, and all of them are multisectoral. These systems together cover a wide variety of domestic species including livestock, poultry, dogs, and cats. These systems collect data using a range of technologies including calls and mobile apps. They focus on common parameters such as species, age, sex, and location of affected animals, including the magnitude of morbidity and mortality at individual and group (herd or flock) levels. Clinical signs of illness were either reported by body systems (checkboxes) or as free text. AfyaData collects health information on the owner as well as on the animals. Only 2 systems include parameters related to companion animals.

Human Health Parameters

Data parameters on human health are collected by 23/38 (61%) systems, with the majority (17/23, 74%) collecting only human health data. Of the 6 collecting multisector data, 3 systems collect data parameters across all One Health sectors, that is, human, animal, and environmental health.

The data parameters for human health cover virtually every organ system within the human body. Across all the systems monitoring human health, fever-related terms are the most common set of data parameters for reporting symptoms of illness. Additional symptoms indicative of influenza-like illness (ILI), including headache, sore throat, and cough are also among the most common symptom parameters collected. Gastrointestinal symptoms are the next most common set of data parameters across the systems collecting human health data.

In 2020, COVID-19-specific symptoms were added to several of the systems, many of which already collected ILI symptoms to track influenza, but now added “loss of taste or smell,” for example. COVID-19 and influenza diagnostic testing and vaccination history are common among the systems monitoring human health only.

XSL•FO
RenderX

Outbreaks Near Me asks the user to report being "healthy." Only when the user reports not being "healthy" do they receive queries on symptoms, events, or other options available within a specific system. Most systems focus on symptoms of illness and the onset of the timing of the symptoms. Several systems ask about health-seeking behavior and the type of health care system visited. Kidenga asks about the health status of others in the household and others inquire about the number of people in the household.

Absenteeism from work or school or changes in daily routine due to illness is captured by many of the systems. Common parameters also include various risk factors, such as exposure to other people with similar symptoms, exposure to pets, work exposure, and recent travel history. Furthermore, a few of the systems collect data on preexisting medical conditions as well as lifestyle parameters such as exercise, smoking, and alcohol consumption.

Discussion

Overview

The One Health Participatory Surveillance Data Parameters Compendium showcases the breadth and depth of how systems are using community engagement to collect multisector data to better inform shared threats. This resource highlights opportunities to expand One Health surveillance and identifies parameters currently collected across multisector systems. It encompasses a broad array of data modes and mechanisms used for data collection, addressing challenges with optimizing user time for data entry and improving accuracy. Additionally, the compendium reveals sector-specific differences in how and when data are collected.

Principal Findings

Opportunities and Challenges for System Augmentation or Development

Only 3 systems—PODD, AfyaData, and Colab—include data from all 3 sectors. For existing single-sector surveillance systems, the compendium offers a framework to identify key parameters for expanding into this type of multisector, One Health surveillance system. For example, wildlife-focused systems could incorporate data on the surrounding environment or on the health of humans interacting with wildlife. Similarly, human health–focused systems in urban areas could expand to include data parameters on the built environment, proximal wildlife, and the health of companion animals. In rural areas, human health systems could include parameters on the health of livestock and wildlife within the shared environment. Given the increasing incidence of emerging infections, data on drivers of disease emergence—such as environmental and weather changes—are valuable contributions to participatory surveillance systems [7]. With the interconnectedness of human, animal, and environmental health, the expansion of environmental systems to include these components is a logical progression.

Most participatory systems are community-driven and developed; thus, parameters are added to a given system to address specific community needs and concerns. Therefore, piloting terms and questions within a target population is essential to ensure wording and translation are culturally appropriate and that the requested parameters are accurately captured. Ensuring integration with governmental surveillance efforts across multiple ministries and departments is also critical when expanding beyond one sector to support more sustainable and successfully coordinated One Health approaches.

Expanding a surveillance system to collect parameters across sectors requires consideration of the time users are willing to invest. Users can only reasonably be expected to spend a limited amount of time answering questions or selecting from menu options. Many systems address this by incorporating automatic data inputs, such as timestamps or geolocation, to save time and improve accuracy. Decision-tree frameworks are also used to streamline data collection and gather targeted information; for example, a "yes" response to a question about "cough" could lead to a follow-up inquiry about "bloody cough." Wildlife and vector programs frequently rely on images to minimize the number of questions and to ensure accurate species identification.

Advancements in technology allow for automated data collection, such as integrating ambient temperature measurements, which can enhance user experience, save time, and improve data quality. Variations in education and literacy levels within communities must be considered. Working closely with the community during system development can help address these challenges and foster user engagement.

Existing siloes among One Health sectors can limit the effective use of data and hinder cross-sectoral collaboration. Existing system infrastructure as well as sensitivities around data privacy can create barriers that make it difficult to collect and share parameters across sectors. Current systems also operate at varying spatial scales. For example, human health data is often protected by privacy regulations, making it unethical to include precise geocoordinates. On the other hand, some environmental parameters, such as mosquito habitats or locations of dead wildlife, require specific geocoded data to enable required actions. Data privacy is critical and can be addressed through user-specific access layers. Integrated dashboards and similar tools can enable health authorities from multiple sectors to visually identify disease hotspots, pinpoint areas with elevated disease transmission risk, and prioritize mitigation efforts. For instance, in Mo-Buzz+, public health inspectors can review reports on communal dengue cases, identify mosquito breeding sites, and detect community practices that contribute to mosquito breeding. To protect privacy, this type of information could be aggregated to a neighborhood level before being shared with public users, preserving anonymity while providing actional insights.

For livestock and plant health, additional sensitivities regarding the location of incidents may arise due to potential economic impacts. Systems address these concerns in various ways, such as obscuring publicly available information and anonymizing reporter details. Any system handling sensitive information must proactively address these concerns to ensure confidentiality and minimize potential risks.

XSL•FO
RenderX

System Structure and Timing Impact Types of Information Gathered

The types of information gathered by surveillance systems vary based on their structure. Systems designed to collect longitudinal data often include more "general" parameters during the initial registration process, reducing the need to recollect this information during subsequent logins. This approach saves the user time by focusing additional questions only on cases where the user reports "unhealthy" conditions for that week. Collecting longitudinal data from a known population allows these systems to calculate both numerators and denominators for analyzing trends and spatial-temporal patterns. Visualizing reports over time and space strengthens disease tracking and anomaly detection. This approach is most common for human health systems, as well as Garden Wildlife Health and WildAlert. In SMART for Health, users (wildlife rangers) track animal populations over time, allowing for the identification of both population and disease trends.

Some systems that collect data on wildlife and environmental factors rely on users to report opportunistically when they encounter a specific animal or setting. As these encounters may be the only opportunity to collect data on the health event, extensive data may be needed. To make the process user-friendly, these systems often incorporate photographs and automated downloads of data such as geolocation and limit the collection of extensive user and animal demographic details. While one-time reporting for plant or environmental health incidents is highly valuable for enabling prompt action, it rarely supports longitudinal surveillance or trend detection. Identification of ways to expand follow-up reporting by users may be appropriate, particularly in the context of suspected outbreaks. Systems collecting data on human-animal interactions, such as bat exposures, are especially valuable for surveillance initiatives aiming to improve understanding of the underlying factors driving pathogen spillover and disease emergence.

Dropdown menus offer the advantage of standardizing information, which facilitates data recording in the field, promotes consistent data analysis across units, and supports data integration across systems. In contrast, free text data fields capture a wide array of information from the user or when the user may not have the technical background or expertise to properly categorize a case.

Language and cultural barriers within communities can result in disparities in data capture. Addressing these challenges requires strategies such as raising community awareness and education to enhance disease literacy and improve access to health care. Using tailored engagement approaches for specific population segments and fostering collaboration with local stakeholders is essential for improving surveillance and response outcomes.

Data Standards

Improving interoperability and data sharing among participatory surveillance systems can allow for early incident detection. Improving interoperability (the ability of different systems to exchange, interpret, and use data accurately) and data sharing on a regional or global scale will require the development of data standards. Data standards facilitate international data exchange and better collaboration. A data standard is an agreed-upon approach, to allow for consistent measurement, qualification or exchange of an object, process, or unit of information. Data standards include agreements on several topic areas including representation, format, definition, structure, use, documentation, and management of data [28,29]. They enable transparency and understanding, and the use of standards promotes common, clear meanings for data [29]. Several systems in this study collect similar categories of parameters but use different questions or methods to gather the information. This variability is often necessary to account for community-specific differences, such as how questions about ethnic groups are phrased. Ideally, while user interfaces may differ to accommodate these variations, the backend systems should be designed to integrate the collected information.

Sector-Specific Findings

Human Health

It is not surprising to see fever as the most prevalent symptom among systems designed to detect early infection in humans. Many have noted that fever monitored on an ongoing basis with temporal and geospatial visualization could be an early indicator of a communicable disease threat [30,31]. Since respiratory diseases have the highest probability of causing local outbreaks, regional epidemics, and potentially a pandemic, many systems monitor for ILI and this likely accounts for the frequency of ILI-terms among the systems. Several systems collect parameters to complement symptom data reported—such as health care–seeking behavior to help track the severity of illness or travel to help understand exposure risks.

Environmental Health

The inclusion of environmental parameters in multisector systems allows systems to understand the complete context in which an incident occurs. Identification of potential vector-breeding locations, water contamination, and habitat change may allow for intervention before the spread of disease into local animal and human populations. Environmental parameters also capture how animals and people may be impacting the health of the environment.

Like wildlife health, environmental health programs often rely on event-based reporting. For example, a person enters information about vector exposure, seeing a contaminated location, or spotting a presumed invasive species. Automated features, such as automated collection of climate parameters, geolocation, and image uploads, allow the system to accurately capture a situation without solely relying on an observer's subjective reporting. To help standardize data, systems use tools such as iNaturalist, which may include predefined choices, and have internal experts reviewing photographs to correctly identify vector, animal, and plant species. The collection of incident-specific details is foundational for understanding the spatiotemporal distribution of vectors and their potential role in disease transmission.

Despite the comprehensive parameters used, the underreporting of vector-borne disease cases in humans remains a persistent

XSL•FO
RenderX

concern. This phenomenon is especially prevalent in regions with limited access to health care services or where vector-borne diseases are poorly understood [30,31].

Livestock and Wildlife Health

Relatively fewer participatory systems exist for livestock compared to systems dedicated to other sectors. Participatory epidemiology is often used to address specific outbreaks in livestock health [32]; ongoing participatory surveillance systems may not be as common due to heavy reliance on governmental surveillance programs for outbreak detection. In the compendium, livestock parameters are from systems also collecting data on wildlife, people, or the environment, reflecting the interconnectedness of livestock among each of the other sectors. Given the daily proximity of pets to owners and the potential for bidirectional transmission, the expansion of participatory surveillance to inform companion animal outbreak detection is needed.

Similar to environmental health systems, wildlife health systems are often event-based—increasing situational awareness of disease occurrence and informing on common disease threats in real time. Information stemming from these systems are commonly underused as resource for monitoring long-term trends in population health and mostly are untapped as resource by sectors outside of wildlife health.

Across wildlife participatory surveillance systems, data accuracy remains a significant challenge, as users often face difficulties accurately identifying wildlife species and describing abnormalities in sick, injured, or dead wild animals. Additionally, the integration of disparate data sources and achieving interoperability between systems is limited due to the diverse types and formats of data collected. The lack of harmonized data in wildlife health surveillance has been a long-standing issue; however, systems designed for large, multilingual, and strategically diverse user bases hold promise for reversing this trend.

The increasing use of machine learning methods, such as natural language processing, offers new opportunities for categorizing data, and enabling comparisons across multiple surveillance streams and systems [33,34]. However, these approaches require compatible language inputs or effective translation mechanisms. The compendium highlights the value of collecting and analyzing prediagnostic data during incidents, which facilitates the rapid detection of clusters of animals presenting with similar abnormalities that might signify a disease outbreak or other adverse health event in wildlife.

Comparison With Prior Work

This work builds upon the landscape analysis of participatory surveillance programs and IWOPS workshops reviewing data parameters and working toward data standards [13,25]. Prior publications have highlighted frameworks and steps for building One Health surveillance systems, including methods for better cross-sectoral stakeholder engagement and ways to overcome implementation barriers [7,10,11]. Identified barriers to the implementation of One Health surveillance systems include incompatible vocabularies, professional divisions, isolated datasets, and the lack of data standards. When developing new One Health participatory surveillance systems, this compendium provides a concrete starting point with real-world examples to facilitate dialogue on data elements and standards for sharing across siloed sectors. A critical analysis by Behravesh et al [12] of landscaping globally available One Health tools uncovered 50 tools that advance One Health. The results highlight that most of these tools addressed assessment and few addressed planning and prioritization. This compendium provides a practical One Health tool to facilitate and support planning, prioritization discussions, and One Health participatory surveillance system design. Prior studies have noted the value of participatory surveillance in war-affected camps and the use of One Health surveillance to capture the impact of the environment on refugee health [35,36]. Interactions between humans and animals are often ignored with regard to the health of refugees, but animal populations are often closely tied to refugee communities. This compendium can be applied to improve One Health disease surveillance for displaced populations in high-risk regions.

Limitations

Data parameters included in the compendium reflect a point in time and are not fully representative of the entire field of participatory surveillance. The parameters were collected in English, through an English-language survey, which may have limited participation from non-English systems. Furthermore, selection bias could be present, as systems were identified based on the participatory surveillance map. Additional validation of data parameters was not conducted because many systems are not publicly available; the parameters included in the compendium are based on the information directly provided by the systems.

Most systems use a decision-tree framework to structure their questions. The compendium does not include this format, focusing on parameters over the structural design. To partially address this limitation, the “special considerations” column in the compendium provides comments on how specific parameters relate to others. Future iterations of the compendium should explore methods to illustrate system structures, providing context for how parameters are used within their frameworks.

Given the compendium’s focus on outbreak detection, parameters related to chronic disease were not included. Future exploration on how mental health, chronic disease, toxic exposures, and other non-outbreak–related parameters are collected would be informative for participatory surveillance systems aiming to capture a broader set of health parameters.

Conclusions

Participatory surveillance enhances traditional means of health surveillance by empowering communities to identify threats at their source. The compendium highlights how existing participatory surveillance systems are already incorporating a One Health approach. Expanding opportunities for multisector data collection through these participatory systems aligns with recent guidance from the United Nations Quadripartite [3,37]. Additional research is required to inform international data standards for One Health participatory surveillance. Pilot studies could provide useful mechanisms to capture various

XSL•FO
RenderX

multisectoral data collection and sharing techniques. Furthermore, ethical issues of data privacy are always essential to consider when capturing personal information. Participatory surveillance provides critical surveillance data that can be deidentified and collated at a specific geographic level, which facilitates anonymity of the responses. The compendium offers a reference to strengthen community capabilities for developing multisector surveillance systems to inform early detection of health incidents threatening themselves, their families, and the environment. Decision makers and practitioners can use the compendium as a tool to build a new participatory surveillance system or augment their existing systems to include a One Health approach. The compendium highlights the need to work toward system interoperability and data standards to enable community-level surveillance to have national, regional, or global implications in outbreak detection and response.

Acknowledgments

The authors would like to acknowledge support from Marlo Libel for his assistance on this project, and Elizabeth Aaron for copyediting the compendium and assisting in data collection. The University of Arizona and Dr Kacey Ernst for providing human participants' review of this project. The authors appreciate the following systems sharing data parameters to ensure that they had a comprehensive compendium as possible: AfyaData, AZDGF, Cambodia 115 Hotline, California West Nile Bird Reporting, Colab, CoughWatch SA, DisApp, EDDMapS/Wild Spotter, eWHIS, FluTracking Australia, FluTracking HongKong, Forest Pest First Detector, GrippeWeb, Guardians of Health, Garden Wildlife Health Project, Health Report (Hälsorapport) Sweden, iMammalia, InfluenzaNet (including individual country-level programs in Italy, France, Germany, United Kingdom, the Netherlands, Belgium, Switzerland, Sweden, and Denmark), I-Tick, Kidenga, MoBuzz+, Outbreaks Near Me, and PODD. Generative artificial intelligence was not used.

Data Availability

All data generated or analyzed during this study are included in this published article [and its supplementary information files—Multimedia Appendix 1].

Conflicts of Interest

None declared.

Multimedia Appendix 1
One Health Participatory Surveillance Compendium.
[XLSX File (Microsoft Excel File), 672 KB - publichealth_v11i1e55356_app1.xlsx]

References

1. Food and Agriculture Organization of the United Nations, United Nations Environment Programme, World Health Organization, World Organisation for Animal Health. The One Health Joint Plan of Action (2022–2026). Working together for the health of humans, animals, plants and the environment. World Organisation for Animal Health. 2022. URL: https://bulletin.woah.org/?panorama=02-1-4-2023-1_oh-jpa [accessed 2024-12-21]
2. Mercy K, Balajee A, Numbere T, Ngere P, Simwaba D, Kebede Y. Africa CDC's blueprint to enhance early warning surveillance: accelerating implementation of event-based surveillance in Africa. J Public Health Afr 2023;14(8):2827 [FREE Full text] [doi: 10.4081/jphia.2023.2827] [Medline: 37753431]
3. Defining collaborative surveillance: a core concept for strengthening the global architecture for health emergency preparedness, response, and resiliency (HEPR). World Health Organization. URL: https://www.who.int/publications/i/item/9789240074064 [accessed 2024-12-06]
4. One Health High-Level Expert Panel (OHHLEP), Adisasmito WB, Almuhairi S, Behravesh CB, Bilivogui P, Bukachi SA, et al. One Health: a new definition for a sustainable and healthy future. PLoS Pathog 2022;18(6):e1010537 [FREE Full text] [doi: 10.1371/journal.ppat.1010537] [Medline: 35737670]
5. Bordier M, Uea-Anuwong T, Binot A, Hendrikx P, Goutard FL. Characteristics of One Health surveillance systems: a systematic literature review. Prev Vet Med 2020;181:104560. [doi: 10.1016/j.prevetmed.2018.10.005] [Medline: 30528937]
6. Kelly TR, Machalaba C, Karesh WB, Crook PZ, Gilardi K, Nziza J, PREDICT Consortium, et al. Implementing One Health approaches to confront emerging and re-emerging zoonotic disease threats: lessons from PREDICT. One Health Outlook 2020;2:1 [FREE Full text] [doi: 10.1186/s42522-019-0007-9] [Medline: 33824944]
7. One Health High-Level Expert Panel (OHHLEP), Hayman DTS, Adisasmito WB, Almuhairi S, Behravesh CB, Bilivogui P, et al. Developing One Health surveillance systems. One Health 2023;17:100617 [FREE Full text] [doi: 10.1016/j.onehlt.2023.100617] [Medline: 38024258]
8. Rizzo DM, Lichtveld M, Mazet JAK, Togami E, Miller SA. Plant health and its effects on food safety and security in a One Health framework: four case studies. One Health Outlook 2021;3:6 [FREE Full text] [doi: 10.1186/s42522-021-00038-7] [Medline: 33829143]

XSL•FO
RenderX

9. Gruetzmacher K, Karesh WB, Amuasi JH, Arshad A, Farlow A, Gabrysch S, et al. The Berlin principles on One Health—bridging global health and conservation. Sci Total Environ 2021;764:142919 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.142919] [Medline: 33097250]
10. Bordier M, Goutard FL, Antoine-Moussiaux N, Pham-Duc P, Lailler R, Binot A. Engaging stakeholders in the design of one health surveillance systems: a participatory approach. Front Vet Sci 2021;8:646458 [FREE Full text] [doi: 10.3389/fvets.2021.646458] [Medline: 34109232]
11. Uchtmann N, Herrmann JA, Hahn EC, Beasley VR. Barriers to, efforts in, and optimization of integrated One Health surveillance: a review and synthesis. Ecohealth 2015;12(2):368-384. [doi: 10.1007/s10393-015-1022-7] [Medline: 25894955]
12. Behravesh CB, Charron DF, Liew A, Becerra NC, Machalaba C, Hayman DT, et al. An integrated inventory of One Health tools: mapping and analysis of globally available tools to advance One Health. CABI One Health 2024;3:1. [doi: 10.1079/cabionehealth.2024.0017]
13. McNeil C, Verlander S, Divi N, Smolinski M. The landscape of participatory surveillance systems across the One Health spectrum: systematic review. JMIR Public Health Surveill 2022;8(8):e38551 [FREE Full text] [doi: 10.2196/38551] [Medline: 35930345]
14. Pruvot M, Denstedt E, Latinne A, Porco A, Montecino-Latorre D, Khammavong K, et al. WildHealthNet: Supporting the development of sustainable wildlife health surveillance networks in Southeast Asia. Sci Total Environ 2023;863:160748 [FREE Full text] [doi: 10.1016/j.scitotenv.2022.160748] [Medline: 36513230]
15. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. Emerg Themes Epidemiol 2014;11:7 [FREE Full text] [doi: 10.1186/1742-7622-11-7] [Medline: 24991229]
16. Suy Lan C, Sok S, Chheang K, Lan DM, Soung V, Divi N, et al. Cambodia national health hotline—participatory surveillance for early detection and response to disease outbreaks. Lancet Reg Health West Pac 2022;29:100584 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100584] [Medline: 36605884]
17. Leal-Neto O, Egger T, Schlegel M, Flury D, Sumer J, Albrich W, et al. Digital SARS-CoV-2 detection among hospital employees: participatory surveillance study. JMIR Public Health Surveill 2021;7(11):e33576 [FREE Full text] [doi: 10.2196/33576] [Medline: 34727046]
18. Pagliari C, Vijaykumar S. Digital participatory surveillance and the Zika crisis: opportunities and caveats. PLoS Negl Trop Dis 2016;10(6):e0004795 [FREE Full text] [doi: 10.1371/journal.pntd.0004795] [Medline: 27294787]
19. Leal Neto O, Haenni S, Phuka J, Ozella L, Paolotti D, Cattuto C, et al. Combining wearable devices and mobile surveys to study child and youth development in Malawi: implementation study of a multimodal approach. JMIR Public Health Surveill 2021;7(3):e23154 [FREE Full text] [doi: 10.2196/23154] [Medline: 33536159]
20. Mahmud AS, Chowdhury S, Sojib KH, Chowdhury A, Quader MT, Paul S, et al. Participatory syndromic surveillance as a tool for tracking COVID-19 in Bangladesh. Epidemics 2021;35:100462 [FREE Full text] [doi: 10.1016/j.epidem.2021.100462] [Medline: 33887643]
21. Guenin MJ, de Nys HM, Peyre M, Loire E, Thongyuan S, Diallo A, et al. A participatory epidemiological and One Health approach to explore the community's capacity to detect emerging zoonoses and surveillance network opportunities in the forest region of Guinea. PLoS Negl Trop Dis 2022;16(7):e0010462 [FREE Full text] [doi: 10.1371/journal.pntd.0010462] [Medline: 35816491]
22. Garg S, Bhatnagar N, Gangadharan N. A case for participatory disease surveillance of the COVID-19 pandemic in India. JMIR Public Health Surveill 2020;6(2):e18795 [FREE Full text] [doi: 10.2196/18795] [Medline: 32287038]
23. Worsley-Tonks KEL, Bender JB, Deem SL, Ferguson AW, Fèvre EM, Martins DJ, et al. Strengthening global health security by improving disease surveillance in remote rural areas of low-income and middle-income countries. Lancet Glob Health 2022;10(4):e579-e584 [FREE Full text] [doi: 10.1016/S2214-109X(22)00031-6] [Medline: 35303467]
24. Wittwer S, Paolotti D, Lichand G, Leal Neto O. Participatory surveillance for COVID-19 trend detection in Brazil: cross-sectional study. JMIR Public Health Surveill 2023;9:e44517 [FREE Full text] [doi: 10.2196/44517] [Medline: 36888908]
25. Smolinski MS, Crawley AW, Olsen JM, Jayaraman T, Libel M. Participatory disease surveillance: engaging communities directly in reporting, monitoring, and responding to health threats. JMIR Public Health Surveill 2017;3(4):e62 [FREE Full text] [doi: 10.2196/publichealth.7540] [Medline: 29021131]
26. Leal Neto O, Paolotti D, Dalton C, Carlson S, Susumpow P, Parker M, et al. Enabling multicentric participatory disease surveillance for global health enhancement: viewpoint on global flu view. JMIR Public Health Surveill 2023;9:e46644 [FREE Full text] [doi: 10.2196/46644] [Medline: 37490846]
27. IWOPS program. Ending Pandemics. 2022. URL: https://endingpandemics.org/iwops-program [accessed 2024-12-06]
28. Data standards. National Library of Medicine. URL: https://www.nnlm.gov/guides/data-glossary/data-standards [accessed 2024-12-06]
29. Learn about data standards. United States Environmental Protection Agency. URL: https://www.epa.gov/data/learn-about-data-standards [accessed 2024-12-06]
30. Multisectoral approach to the prevention and control of vector-borne diseases: a conceptual framework. World Health Organization. 2020. URL: https://iris.who.int/bitstream/handle/10665/331861/9789240004788-eng.pdf [accessed 2024-12-06]

XSL•FO
RenderX

31. Mubayi A, Castillo-Chavez C, Chowell G, Kribs-Zaleta C, Ali Siddiqui N, Kumar N, et al. Transmission dynamics and underreporting of Kala-azar in the Indian state of Bihar. J Theor Biol 2010;262(1):177-185. [doi: 10.1016/j.jtbi.2009.09.012] [Medline: 19769990]
32. Jost CC, Mariner JC, Roeder PL, Sawitri E, Macgregor-Skinner GJ. Participatory epidemiology in disease surveillance and research. Rev Sci Tech 2007;26(3):537-549 [FREE Full text] [Medline: 18293603]
33. Kelly TR, Pandit PS, Carion N, Dombrowski DF, Rogers KH, McMillin SC, et al. Early detection of wildlife morbidity and mortality through an event-based surveillance system. Proc Biol Sci 2021;288(1954):20210974 [FREE Full text] [doi: 10.1098/rspb.2021.0974] [Medline: 34256001]
34. Mackey T, Purushothaman V, Li J, Shah N, Nali M, Bardier C, et al. Machine learning to detect self-reporting of symptoms, testing access, and recovery associated with COVID-19 on Twitter: retrospective big data infoveillance study. JMIR Public Health Surveill 2020;6(2):e19509 [FREE Full text] [doi: 10.2196/19509] [Medline: 32490846]
35. Sutradhar I, Zaman MH. One Health approaches to improve refugee health. Lancet Glob Health 2021;9(12):e1646-e1647 [FREE Full text] [doi: 10.1016/S2214-109X(21)00416-2] [Medline: 34798016]
36. Wickramage K, Nellapalli P. Community participatory methods in disease surveillance and public health in war-affected camps, and its potential contribution to peace building. Int Electron J Health Educ 2008:95-108 [FREE Full text]
37. Quadripartite One Health intelligence scoping study: actions to develop an effective global One Health intelligence system. Food and Agriculture Organization of the United Nations.: World Health Organization; 2022. URL: https://www.fao.org/3/cc1533en/cc1533en.pdf [accessed 2024-12-06]

Abbreviations

EBS: event-based surveillance
ILI: influenza-like illness
IWOPS: International Workshops on Participatory Surveillance
PODD: Participatory One Health Digital Disease Detection
USAID: United States Agency for International Development

Edited by A Mavragani; submitted 11.12.23; peer-reviewed by D Balate, H Kinzer, C Argyropoulos, F Fathmawati; comments to author 23.09.24; revised version received 07.12.24; accepted 18.12.24; published 26.03.25.

Please cite as:
McNeil C, Divi N, Bargeron IV CT, Capobianco Dondona A, Ernst KC, Gupta AS, Fasominu O, Keatts L, Kelly T, Leal Neto OB, Lwin MO, Makhasi M, Mutagahywa EB, Montecino-Latorre D, Olson S, Pandit PS, Paolotti D, Parker MC, Samad MH, Sewalk K, Sheldenkar A, Srikitjakarn L, Suy Lan C, Wilkes M, Yano T, Smolinski M
Data Parameters From Participatory Surveillance Systems in Human, Animal, and Environmental Health From Around the Globe: Descriptive Analysis
JMIR Public Health Surveill 2025;11:e55356
URL: https://publichealth.jmir.org/2025/1/e55356
doi:10.2196/55356
PMID:

©Carrie McNeil, Nomita Divi, Charles Thomas Bargeron IV, Andrea Capobianco Dondona, Kacey C Ernst, Angela S Gupta, Olukayode Fasominu, Lucy Keatts, Terra Kelly, Onicio B Leal Neto, May O Lwin, Mvuyo Makhasi, Eric Beda Mutagahywa, Diego Montecino-Latorre, Sarah Olson, Pranav S Pandit, Daniela Paolotti, Matt C Parker, Muhammad Haiman Samad, Kara Sewalk, Anita Sheldenkar, Lertrak Srikitjakarn, Channé Suy Lan, Michael Wilkes, Terdsak Yano, Mark Smolinski. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Quality of Life Among Family Caregivers of Individuals With Rare Diseases: Web-Based Population Study on the Validity and Reliability of the Polish World Health Organization Quality of Life-BREF Questionnaire

Piotr Jabkowski[1*], PhD; Jan Domaradzki[2*], PhD; Dariusz Walkowiak[3*], PhD

[1]Faculty of Sociology, Adam Mickiewicz University, Poznan, Poland

[2]Department of Social Sciences and Humanities, Poznan University of Medical Sciences, Rokietnicka 7, Poznań, Poland

[3]Department of Organization and Management in Health Care, Poznan University of Medical Sciences, Poznan, Poland

[*]all authors contributed equally

Corresponding Author:
Jan Domaradzki, PhD
Department of Social Sciences and Humanities, Poznan University of Medical Sciences, Rokietnicka 7, Poznań, Poland

Abstract

Background: Caring for individuals with rare diseases (RDs) presents unique challenges that can significantly impact caregivers' quality of life (QoL). The World Health Organization Quality of Life-BREF (WHOQOL-BREF) is a widely used tool for assessing QoL across different populations.

Objective: This study examines the QoL of caregivers of individuals with RDs and evaluates the psychometric properties of the WHOQOL-BREF in this population.

Methods: A self-administered, anonymous, computer-assisted web-based survey was conducted among family caregivers of individuals with RDs in Poland between March and August 2024. Due to the lack of a national registry of patients with RDs, participants were recruited through convenience sampling via associations, foundations, and organizations of patients with RDs. Eligibility criteria included being 18 years and older of age, speaking Polish, being a caregiver of a person with a confirmed RD diagnosis, and providing informed consent. The survey included sociodemographic questions and the Polish version of the WHOQOL-BREF, which assesses QoL across 4 domains: physical health, psychological health, social relationships, and environment. Internal consistency was assessed using Cronbach α, and confirmatory factor analysis was conducted to examine the instrument's structural validity.

Results: A total of 942 caregivers of individuals with various RDs participated in the study. Confirmatory factor analysis supported the 4-domain structure, with further improvement in a finally modified WHOQOL-BREF model (χ^2_{243}=1043.0; P<.001; Comparative Fit Index=0.919; Tucker-Lewis Index=0.907; root-mean-square error of approximation=0.059). Internal consistency was satisfactory, with Cronbach α values ranging from 0.70 (social relationships) to 0.84 (psychological health). Mean domain scores on a scale of 0 - 100 were 50.2 (SE 0.59; physical health), 54.9 (SE 0.59; psychological health), 51.3 (SE 0.67; social relationships), and 52.1 (SE 0.56; environment), with minimal floor and ceiling effects (≤1.4%) across domains. Younger female caregivers reported significantly lower psychological health (eg, mean 43.6, SE 0.59 vs mean 59.9, SE 10.0 for younger male caregivers) and social relationships (mean 39.3, SE 3.34 vs mean 55.0, SE 4.75) well-being compared to other groups. Exactly 151 (16%) of caregivers rated their overall QoL as poor or very poor, and 289 (30.7%) were dissatisfied or very dissatisfied with their health, with female caregivers reporting more dissatisfaction (n=263, 32.4%) than male caregivers (n=26, 20%). Overall, our findings demonstrate the robust psychometric properties of the WHOQOL-BREF among caregivers of people with RDs and provide domain-specific normative data to guide future research and interventions.

Conclusions: The WHOQOL-BREF is a reliable and valid instrument for assessing QoL among caregivers of individuals with RDs, though the social relationship domain may require further refinement. Caregivers experience varying QoL outcomes depending on demographic factors, highlighting the need for targeted support interventions. Future research should explore cultural adaptations and potential supplementary modules to enhance the instrument's applicability in caregiver populations.

(JMIR Public Health Surveill 2025;11:e72590) doi:10.2196/72590

KEYWORDS

quality of life; rare diseases; caregivers; WHOQOL-BREF; World Health Organization Quality of Life-BREF; validation

XSL•FO
RenderX

Introduction

Quality of life (QoL) is a multidimensional construct that reflects an individual's overall well-being, encompassing physical health, psychological health, social relationships, and environmental factors. In 1991, the World Health Organization (WHO) initiated a project to develop an internationally applicable measure of QoL, leading to the creation of the World Health Organization Quality of Life (WHOQOL) instrument. The WHO defined QoL as an individual's perception of their position in life within the context of their cultural and value systems as well as their personal goals, expectations, standards, and concerns. This definition highlights the subjective nature of QoL, emphasizing that it is shaped by personal experiences and influenced by broader social and environmental contexts. Importantly, QoL assessment provides valuable insights into health outcomes, treatment effectiveness, and overall well-being across diverse populations [1-3].

A subset of QoL research, health-related quality of life (HRQoL), has gained increasing attention in clinical and public health research. HRQoL focuses specifically on the impact of health conditions on an individual's well-being, playing a critical role in health care decision-making, policy development, and patient-centered care [4-6]. Over the past 3 decades, numerous HRQoL assessment tools have been developed, including both generic instruments applicable across populations and disease-specific measures tailored to particular conditions. Despite the proliferation of these tools, a significant limitation remains—the lack of internationally validated instruments capable of capturing HRQoL across diverse cultural and socioeconomic settings [7,8].

To address this challenge, the WHO developed the WHOQOL-100 consisting of 100 questions, an instrument designed to assess subjective well-being while ensuring cross-cultural validity. This instrument was developed through extensive research conducted across 15 field centers worldwide, incorporating perspectives from patients, caregivers, health care professionals, and the general population. The psychometric properties of the WHOQOL-100, including its validity and reliability, were rigorously tested in multiple cultural and linguistic contexts [1]. Recognizing the need for a shorter, more practical instrument, the WHO subsequently introduced the World Health Organization Quality of Life-BREF (WHOQOL-BREF) in 1998, a 26-item version that assesses 4 key domains of HRQoL: physical health, psychological health, social relationships, and environment. The WHOQOL-BREF has since been widely adopted in epidemiological studies, clinical trials, and public health research, demonstrating strong psychometric properties and cross-cultural applicability [2,4,7,9].

Although the WHOQOL-BREF has been validated in numerous populations, its application in certain groups remains underexplored. In Poland, the WHOQOL-BREF was validated in 2006 using a sample of healthy individuals and patients with various diseases and health conditions [6]. However, this validation process overlooked caregivers, an essential yet often neglected population in HRQoL research. Caregivers play a vital role in the health care system, providing continuous support to individuals with chronic illnesses and disabilities. Despite their crucial contributions, caregivers often experience significant physical, emotional, and financial burdens, which in turn affect their own QoL [10-13].

Caregivers of persons with RDs face particularly severe challenges. RDs are often characterized by delayed diagnoses, limited treatment options, and a lack of specialized health care resources, placing an immense strain on families [10,13-17]. Studies suggest that caregivers of persons with RDs report lower QoL compared to those caring for individuals with more common chronic conditions, mainly due to increased caregiving demands, emotional distress, social isolation, and financial hardship. The unpredictable nature of many RDs, coupled with the complexity of care required, exacerbates stress and negatively affects the physical and mental health of caregivers [5,13,18].

The burden of caregiving is multifaceted, affecting various domains of QoL. Physically, caregivers frequently experience chronic fatigue, sleep disturbances, headaches, and weakened immune function due to the demands of continuous care. Psychologically, they are at increased risk for anxiety, depression, burnout, and emotional exhaustion, with some studies indicating higher rates of suicidal ideation among caregivers [17-20]. Socially, caregivers often face reduced participation in social activities, strained relationships, and isolation, as caregiving responsibilities limit their ability to engage with friends and family. Financially, many caregivers are forced to reduce work hours or leave employment altogether, exacerbating economic difficulties and increasing reliance on social support systems. Given these profound challenges, assessing QoL in caregivers of persons with RDs is essential for understanding their specific needs and designing effective interventions to support them [14,16,19,21-23].

Despite the growing recognition of caregiver burden as a public health concern, research validating HRQoL assessment tools specifically for this population remains limited. The WHOQOL-BREF, although widely used across various patient populations, has not been systematically evaluated for its appropriateness in assessing the QoL of caregivers, particularly in the context of rare diseases (RDs). While WHOQOL-BREF is often used in studies on caregivers of individuals with RDs, the resulting QoL domain scores are typically interpreted abstractly—detached from the scale itself—or, less commonly, compared against national normative data [17,24,25]. A validated, standardized tool tailored to caregivers would offer valuable insights for clinical practice, inform health policy, and support the development of targeted interventions, ensuring that the specific needs of caregivers are effectively addressed [2,13,25,26].

To bridge this research gap, this study aims to validate the WHOQOL-BREF for caregivers of persons with RDs and establish population norms specific to this group. By generating normative data, this study seeks to enhance the clinical utility of the WHOQOL-BREF in caregiver research and contribute to the development of targeted support programs and health care policies. A comprehensive understanding of caregiver QoL

XSL•FO
RenderX

is crucial for improving the well-being of both caregivers and the care recipients they support, ultimately leading to better health outcomes for families affected by RDs.

Methods

Study Design

A cross-sectional study was conducted in 2024 in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1) [27].

This study used data from a self-administered, anonymous, computer-assisted web-based survey examining the association between caregiving for a person with an RD and caregivers' perceived QoL. The survey targeted family caregivers of individuals with RDs. Given the absence of a Polish registry of patients with RD and the unknown exact number of affected individuals, a convenience sampling approach was used. Several associations, foundations, and organizations of patients with RD assisted in participant recruitment by sharing the survey link on their websites and Facebook pages.

At the same time, it should be stressed that convenience sampling used in this study introduces selection bias, as participants may not represent the broader population, limiting the generalizability of findings. However, given the specific nature of the study population—namely, the small size of the groups of particular RDs, the absence of a national registry of patients with RD in Poland, and the wide geographic dispersion of caregivers—convenience sampling appeared to be the most practical and, indeed, the only feasible method. However, it is important to acknowledge several limitations associated with this approach. Most importantly, participants recruited via patient organizations may not reflect the full diversity of the broader population. This type of sampling could affect the study outcomes, potentially skewing results based on the demographics of the sample, such as age, sex, socioeconomic status, or disease severity. Consequently, this method introduces selection bias and limits the reliability and generalizability of the findings. Thus, although this strategy was appropriate under the circumstances, future research should aim to use more representative sampling techniques—such as random or stratified sampling—that may better capture the heterogeneity of caregivers of individuals with RDs and allow for a more accurate reflection of the realities of the population as a whole. Nonetheless, again, it should be stressed that due to the unique challenges of studying this population of caregivers of individuals with RDs, the use of other sampling methods is very difficult, if at all feasible.

Setting and Recruitment

The survey targeted family caregivers of individuals diagnosed with RDs. Due to the absence of a national registry of patients with RD in Poland and the unknown exact number of patients with RD, the study used convenience sampling. Recruitment was supported by various patient organizations, foundations, and advocacy groups via their websites and Facebook profiles.

At the first stage of recruitment, the study coordinator contacted several associations, foundations, and organizations of patients with RD to determine whether they were interested in participating in the study. After receiving their consent, a letter with detailed information about the purpose and method of the study, together with an invitation, was published on the organization's Facebook pages and sent out via an internal mailing to members, by which all affiliated caregivers who met the inclusion criteria were invited to participate in the study.

Eligible participants were required to be at least 18 years of age, fluent in Polish, and a parent or family caregiver of a person with a confirmed RD diagnosis. Additionally, participants needed access to the internet, the ability to use electronic devices to complete the survey, and had to provide written informed consent before participation.

The questionnaire consisted of sociodemographic questions concerning both caregivers and persons with RD along with the abbreviated Polish version of the WHOQOL-BREF [6]. This 24-item instrument assesses QoL across 4 domains: physical health, psychological health, social relationships, and environment. Two additional items do not constitute any of the 4 dimensions but are set up to measure the overall QoL and satisfaction with health.

Data Collection Procedure

The survey was conducted between March and August 2024. Initially, the research coordinator contacted multiple associations, foundations, and organizations of patients with RDs to obtain permission for distribution. Upon approval, an invitation letter containing a link to the web-based questionnaire was shared via these platforms. Participants who provided written informed consent completed the survey electronically. To enhance response rates, 3 follow-up reminders were sent in January, March, and June. The survey required approximately 20 - 25 minutes to complete.

Data Analysis

All statistical analyses were conducted using R Statistical Software (version 4.3.1; R Foundation for Statistical Computing). The study used a quantitative research design to analyze the psychometric properties of the WHOQOL-BREF questionnaire for the Polish population of caregivers of people with RDs. Descriptive statistics were calculated to summarize the demographic characteristics of the sample and the overall QoL scores in 4 domains: physical health, psychological health, social relationships, and environment. Reliability analysis was performed using Cronbach α to assess the internal consistency of the questionnaire. Confirmatory factor analysis (CFA) was used to test the hypothesized factor structure of the WHOQOL-BREF by examining the fit of the data to the model. Goodness-of-fit indices, including the Comparative Fit Index (CFI) and root mean-square-error of approximation (RMSEA), were reported to assess model adequacy.

Ethical Considerations

This study was conducted according to the Declaration of Helsinki and was approved by the Poznan University of Medical Sciences Bioethics Committee (KB 228/24, March 13, 2024). All participants were informed about the objectives and methodology of the study in an invitation letter. Informed

XSL•FO
RenderX

consent was obtained electronically through the web-based questionnaire by selecting the "I agree" option. Participants were assured that their participation was voluntary, anonymous, and confidential and that they had the right to withdraw at any time or omit any question without consequence. To ensure privacy, no personal identifying information or images were collected. Participants were also informed that, although some questions may be emotionally sensitive, they could pause or discontinue the survey at any time. Since participation did not involve any direct intervention or additional burden, no financial compensation was offered.

Results

Overview

In this section, we present a validation of the WHOQOL-BREF scale and provide insights into its empirical structure, psychometric properties, and latent score characteristics. We have implemented CFA by comparing a unidimensional baseline model with the established 4D model of the WHOQOL-BREF. Statistical fit indices such as CFI and RMSEA are provided to assess model adequacy, along with modifications to improve dimensional alignment. This section then highlights the robustness of the 4D structure, and the Structural Validity section focuses on the reliability of the WHOQOL-BREF scale, which is examined through the lens of internal consistency across the 4 domains. A descriptive analysis of the scores derived from the 4 dimensions of the WHOQOL-BREF is also presented. Finally, norms for the QoL domains in a Polish population of caregivers of persons with RDs are provided. The norms for the WHOQOL-BREF are presented as mean values, SE of the mean, 25th percentile, median, 75th percentile, minimum, maximum, and percentage of floor and ceiling effect. These norms were estimated by sex and age group.

Study Participants

Table 1 presents characteristics of the sample by sex and age category of caregivers.

Table 2 outlines the demographic and diagnostic characteristics of 942 patients with RDs. Most patients were male (n=586, 62.2%), with a wide age range (0.2 - 78 years), a mean age of 12 (SE 0.32) years, and a median of 10 (IQR 6.0-15.0) years. The most frequently reported conditions included Duchenne muscular dystrophy or Becker muscular dystrophy (n=211, 22.4%), 22q11.2 deletion syndrome (n=131, 13.9%), and Angelman syndrome (n=120, 12.7%). A total of 191 (20.3%) patients were diagnosed with one of 144 other rare or ultrarare conditions. Time to diagnosis varied considerably: while 165 (17.5%) were diagnosed within 3 months, 226 (24%) received a diagnosis within 1 - 2 years, and 130 (13.9%) waited more than 5 years. A small group (n=18, 1.9%) was still in the asymptomatic phase of the disease. These data highlight the diagnostic challenges and heterogeneity of RD populations.

Table . General characteristics of family caregivers of individuals with rare disease.

Age (years)	Male, n (%)	Female, n (%)	Total, n (%)
16 - 29	2 (1.5)	32 (3.9)	34 (3.6)
30 - 39	44 (33.8)	337 (41.5)	381 (40.4)
40 - 49	64 (49.2)	350 (43.1)	414 (43.9)
50 and more	20 (15.4)	93 (11.5)	113 (12.0)
Total	130 (13.8)	812 (86.2)	942 (100)

XSL•FO
RenderX

Table . Characteristics of individuals with rare disease.

Characteristics	Total (N=942), n (%)
Sex	
Female	356 (37.8)
Male	586 (62.2)
Age (years)	
Mean (SE)	12 (0.32)
Range	0.2 - 78
Median (IQR)	10 (6.0-15.0)
Disease	
22q11.2 deletion syndrome	131 (13.9)
Angelman syndrome	120 (12.7)
DMD[a] or BMD[b]	211 (22.4)
Dravet syndrome	106 (11.3)
Ehlers-Danlos syndrome	19 (2)
Glycogen storage disease	10 (1.1)
Kabuki syndrome	11 (1.2)
Neurofibromatosis	33 (3.5)
Niemann-Pick disease	10 (1.1)
PANS[c] or PANDAS[d]	14 (1.5)
Prader-Willi syndrome	10 (1.1)
Tuberous sclerosis complex	14 (1.5)
Turner syndrome	14 (1.5)
Williams syndrome	48 (5.1)
144 other diseases	191 (20.3)
Time spent before the correct diagnosis was made	
0 - 3 months	165 (17.5)
4 - 6 months	106 (11.3)
7 - 12 months	140 (14.9)
1 - 2 years	226 (24.0)
3 - 5 years	157 (16.7)
6 - 10 years	82 (8.7)
11 - 15 years	29 (3.1)
16 - 20 years	10 (1.1)
20+ years	9 (1.0)
Asymptomatic phase of the disease	18 (1.9)

[a]DMD: Duchenne muscular dystrophy.

[b]BMD: Becker muscular dystrophy.

[c]PANS: pediatric acute-onset neuropsychiatric syndrome.

[d]PANDAS: pediatric autoimmune neuropsychiatric disorders associated with streptococcal infection.

Confirmatory Factor Analysis

The CFA results illustrate a comparative assessment of 3 models for the WHOQOL-BREF: the baseline unidimensional model, the original 4D WHO model, and a modified 4D model. The fit indices provide a comprehensive evaluation of the models' performance and structural validity of the WHOQOL-BREF 4D scale (Table 3).

XSL•FO
RenderX

The baseline unidimensional model yielded unsatisfactory fit indices, suggesting a poor representation of the scale's underlying structure. The chi-square statistic (χ^2_{252}=1602.8; P<.001) was significant, and the χ^2/df ratio (6.360) exceeded the acceptable threshold of 5. The CFI (0.862) and Tucker-Lewis Index (TLI=0.849) indicated a suboptimal model fit, while the RMSEA (0.075; 90% CI 0.072 - 0.079) was marginally above the acceptable upper limit. Conversely, the standardized root-mean-square residual (SRMR=0.050) remained within acceptable limits, indicating an adequate residual-based fit. In practical terms, these results suggest that treating the WHOQOL-BREF as a single-factor measure does not fully capture the complexities of caregivers' QoL, as evidenced by the modest CFI or TLI and moderately elevated RMSEA.

Table . Comparison of fit indices for different WHOQOL-BREF[a] latent models for family caregivers of individuals with rare disease.

Model	Chi-square (*df*)	χ^2/df	CFI[b]	TLI[c]	SRMR[d]	RMSEA[e]	RMSEA−95% CI	RMSEA+95% CI
Baseline model	1602.8 (252)	6.4	0.862	0.849	0.050	0.075	0.072	0.079
Original WHOQOL-BREF	1343.9 (246)	5.5	0.888	0.875	0.047	0.069	0.065	0.072
Modified WHOQOL-BREF	1043.0 (243)	4.3	0.919	0.907	0.041	0.059	0.055	0.063

[a]WHOQOL-BREF: World Health Organization Quality of Life-BREF.

[b]CFI: Comparative Fit Index.

[c]TLI: Tucker-Lewis Index.

[d]SRMR: standardized root-mean-square residual.

[e]RMSEA: root-mean-square error of approximation.

In contrast, the initial 4D model demonstrated substantial improvements across all fit indices (χ^2_{246}=1343.9; P<.001). The χ^2/df ratio decreased to 5.463, and the CFI (0.888) and TLI (0.875) approached recommended levels. The RMSEA (0.069; 90% CI 0.065 - 0.072) and SRMR (0.047) further suggested enhanced fit. Adjustments were made based on modification indices to optimize model fit; that is, correlated error terms were introduced between questions 3 and 4 (modification index equals 164.9), questions 12 and 23 (modification index equals 64.8), and questions 19 and 20 (modification index equals 56.4). While these modifications enhanced statistical compatibility, they also entailed conceptual implications. Specifically, it can be posited that each of these item pairs is likely to tap into closely related dimensions of caregiver well-being (eg, physical or psychological strain or environmental or social considerations). Consequently, the modification indices suggest that some questions, though intended to measure slightly different constructs, may elicit overlapping responses from caregivers who face similar daily challenges across multiple domains of life. Rather than perceiving these errors as mere statistical artifacts, it is more fruitful to interpret them as potential indicators of content overlap or a shared contextual factor (eg, stress or time pressure) that influences the way participants interpret these items. For researchers using the WHOQOL-BREF in caregiver populations, acknowledging and examining these correlated errors can foster a more nuanced understanding of how caregiving experiences merge otherwise distinct QoL facets. Future studies could investigate whether such overlaps are unique to certain subgroups or care contexts or if item wording could be refined to differentiate better between seemingly related aspects of QoL.

The modified 4D model demonstrated a significant enhancement (χ^2_{243}=1043.0; P<.001), exhibiting a reduced χ^2/df ratio of 4.292. The CFI (0.919) and TLI (0.907) exceeded the conventional threshold for adequate fit, while the RMSEA (0.059; 90% CI 0.055 - 0.063) and SRMR (0.041) substantiated the model's robustness. The findings suggest that the targeted modifications serve to enhance the correlation between items within their respective dimensions, thereby producing a more robust and precise measurement structure with fit indices that approach commonly accepted benchmarks.

Structural Validity

This section presents a reliability analysis of the items within the WHOQOL-BREF across its 4 dimensions. The analysis uses Cronbach α to evaluate internal consistency, with findings indicating satisfactory to high reliability across domains. Specifically, α values ranged from 0.69 for the social relationships domain to 0.84 for the psychological health domain, demonstrating varying but overall strong consistency (Table 4).

Item-level analyses provide further insight into the impact of individual items on the overall reliability of each domain. Most items contribute positively to reliability, with suggestions for potential refinement in some areas.

XSL•FO
RenderX

Table . Internal consistency scores of the 4 WHOQOL-BREF domains obtained for Polish family caregivers of individuals with rare diseases.

Domain	Items, n	Cronbach α
Physical health	7	0.818
Psychological health	6	0.843
Social relationships	3	0.700
Environment	8	0.807

[a]WHOQOL-BREF: World Health Organization Quality of Life-BREF.

Normalized Scores for Domains of WHOQOL-BREF

This section presents a descriptive analysis of the normalized scores derived from the 4D-modified WHOQOL-BREF scale. Following the recommendations of the WHO, responses to each item were coded from 1 to 5 and transformed to a scale from 0 (lowest QoL) to 100 (highest QoL). Each item was weighted by standardized factor loadings derived from a CFA model before calculating the overall domain score (Table 5).

The factor loadings range from 0.363 to 0.801, indicating varying degrees of item contribution to their respective latent constructs. The physical health domain exhibits robust item representation, with loadings such as question 18 (0.801) and question 16 (0.794) suggesting substantial explanatory power for the physical health construct. The psychological health domain also exhibits high loadings, particularly for question 26 (0.747) and question 19 (0.688), suggesting that these items are key indicators of psychological health. Items in the environmental domain display more variability, with question 12 having the lowest loading (0.456) and question 8 the highest (0.763). The social relationships domain demonstrates moderate item loadings, with question 22 (0.619) being the most influential item and question 20 (0.477) contributing the least.

XSL•FO
RenderX

Table . Standardized factor loadings for the items comprising the 4 WHOQOL-BREF domains for Polish family caregivers of individuals with rare diseases.

WHOQOL-BREF domains and items		Standardized factor loadings
Physical health		
	Question 3	0.456
	Question 4	0.363
	Question 10	0.794
	Question 15	0.673
	Question 16	0.562
	Question 17	0.801
	Question 18	0.658
Psychological health		
	Question 5	0.747
	Question 6	0.689
	Question 7	0.600
	Question 11	0.631
	Question 19	0.795
	Question 26	0.688
Social relationships		
	Question 20	0.776
	Question 21	0.603
	Question 22	0.619
Environment		
	Question 8	0.763
	Question 9	0.677
	Question 12	0.622
	Question 13	0.564
	Question 14	0.675
	Question 23	0.477
	Question 24	0.395
	Question 25	0.380

[a]WHOQOL-BREF: World Health Organization Quality of Life-BREF.

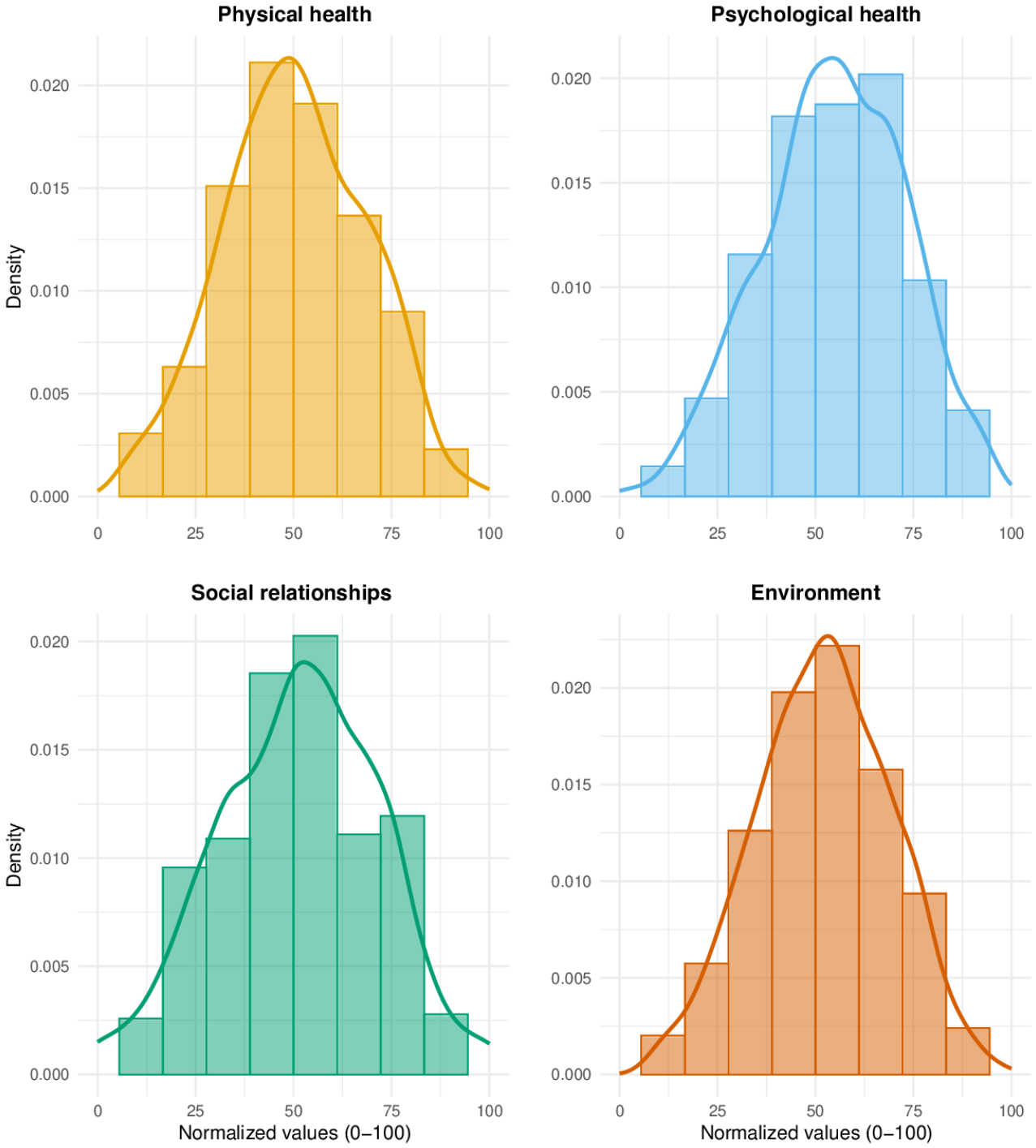
Figure 1 presents histograms and density plots for the 4 dimensions of the WHOQOL-BREF scale: physical health, psychological health, social relationships, and environment domains. Each plot illustrates the distribution of normalized scores (ranging from 0 to 100) for the respective dimension. The overlaid density curves provide a smoothed visualization of the score distribution, highlighting its approximate normality across all domains, with slight deviations observed, particularly in the psychological health and social relationships domains. The detailed descriptive statistics for the distribution of scores within each domain of the WHOQOL-BREF are delineated in Table 6.

The psychological health domain demonstrated the highest mean score (mean 54.9, SE 0.59), indicating relatively better perceived psychological well-being among respondents. Conversely, the physical health domain exhibited the lowest mean score (mean 50.2, SE 0.59). The environment and social relationships domains showed mean scores of 52.1 (SE 0.56) and 51.3 (SE 0.67), respectively. The SEs of the mean were relatively consistent, ranging from 0.56 to 0.67, indicating precision in the central tendency estimates.

Concerning the normality of the data, the skewness and kurtosis values are within acceptable ranges, thereby supporting the visual interpretation from the density plots. However, the Shapiro-Wilk test reveals statistically significant deviations from normality for the physical health, psychological health, and social relationships dimensions ($P<.01$), while the environment domain demonstrates borderline nonsignificance ($P=.05$), suggesting a near-normal distribution.

XSL•FO
RenderX

Figure 1. Histogram and density plots for domains of the WHOQOL-BREF (World Health Organization Quality of Life-BREF) for Polish family caregivers of individuals with rare disease.



XSL•FO
RenderX

Table . Descriptive characteristics of the latent scores for WHOQOL-BREF domains for Polish family caregivers of individuals with rare disease.

Statistic	Environment	Physical health	Psychological health	Social relationships
Mean (SE)	52.1 (0.56)	50.2 (0.59)	54.9 (0.59)	51.3 (0.67)
Range	6.8-97.8	4.8-100.0	0.0-100.0	0.0-100.0
Skewness	−0.060	−0.023	−0.155	−0.137
Kurtosis	2.624	2.551	2.629	2.681
Shapiro-Wilk statistic	0.997	0.995	0.994	0.989
Shapiro-Wilk *P* value	.05	.006	<.001	<.001
Median (IQR)	52.2 (40.4-64.5)	50.0 (37.8-63.9)	54.6 (42.6-67.2)	51.9 (34.9-67.3)
Floor effect (%)	0	0	0.1	1.4
Ceiling effect (%)	0	0.1	0	1.3

[a]WHOQOL-BREF: World Health Organization Quality of Life-BREF.

The analysis of floor and ceiling effects reveals minimal clustering at the extremes, with the social relationships domain exhibiting a 1.4% floor effect and a 1.3% ceiling effect, the highest among all domains. These negligible effects ensure the scale's capacity to capture a wide range of responses without response saturation.

Normative Data for Domains of WHOQOL-BREF

Prior to the presentation of normative data for the domains of the WHOQOL-BREF, an examination of the overall assessment of the QoL and satisfaction with health among Polish caregivers of individuals with RDs is warranted. The data presented in Table 7 reveal notable differences by sex.

Concerning the overall evaluation of QoL, the predominant response among male and female caregivers was "good," with 42.3% (n=55) of male and 39.5% (n=321) of female caregivers indicating this as their response. The second most prevalent response was "neither poor nor good," with 31.5% (n=41) of male and 39.5% (n=321) of female caregivers choosing this option. However, a higher proportion of female caregivers reported a "poor" QoL (n=108, 13.3%) compared to male caregivers (n=19, 14.6%). The percentage of caregivers who rated their QoL as "very poor" was relatively low overall, at 2.5% (n=24), with male caregivers (n=7, 5.4%) more likely to report this than female caregivers (n=17, 2.1%). Conversely, only a small percentage of caregivers reported a "very good" QoL (male: n=8, 6.2% and female: n=45, 5.5%). These findings suggest that perception of the overall QoL is more likely to be situated toward the middle or lower end of the scale for both sexes. Satisfaction with health, however, revealed more pronounced differences between male and female caregivers. While 45.4% (n=59) of male caregivers reported being "satisfied" with their health, this proportion was significantly lower for female caregivers, at 29.2% (n=237). Instead, more female caregivers expressed dissatisfaction (n=237, 29.2% compared to n=22, 16.9% of male caregivers) or were "neither satisfied nor dissatisfied" (female: n=293, 36.1% vs male: n=41, 31.5%). A small proportion of caregivers described themselves as "very satisfied" with their health (male: n=4, 3.1% and female: n=19, 2.3%), while 3.2% (n=30) of caregivers overall reported being "very dissatisfied." These results suggest that male caregivers generally report a higher QoL and greater health satisfaction than female caregivers. The elevated incidence of dissatisfaction and neutral responses among female caregivers may indicate the supplementary burdens they encounter in their caregiving roles or inherent disparities in perceived health status and well-being.

XSL•FO
RenderX

Table . Overall evaluation of the quality of life and satisfaction with the health of family caregivers of individuals with rare diseases by sex.

	Male, n (%)	Female, n (%)	Total, n (%)
How would you rate your quality of life?			
Very poor	7 (5.4)	17 (2.1)	24 (2.5)
Poor	19 (14.6)	108 (13.3)	127 (13.5)
Neither poor nor good	41 (31.5)	321 (39.5)	362 (38.4)
Good	55 (42.3)	321 (39.5)	376 (39.9)
Very good	8 (6.2)	45 (5.5)	53 (5.6)
How satisfied are you with your health?			
Very dissatisfied	4 (3.1)	26 (3.2)	30 (3.2)
Dissatisfied	22 (16.9)	237 (29.2)	259 (27.5)
Neither satisfied nor dissatisfied	41 (31.5)	293 (36.1)	334 (35.5)
Satisfied	59 (45.4)	237 (29.2)	296 (31.4)
Very satisfied	4 (3.1)	19 (2.3)	23 (2.4)

Normative data for the 4 dimensions of the WHOQOL-BREF—physical health, psychological health, social relationships, and environment—segmented by sex and age group are presented in Table 8.

Male caregivers demonstrated higher mean scores in physical health than female caregivers across all age groups. The highest mean score was observed in the youngest male group (16 - 29 years: mean 59.9, SE 10.04), gradually decreasing with age, reaching its lowest value in the 50+ year group (mean 50.4, SE 4.72). Among female caregivers, the 16 - to 29-year age group exhibited the lowest mean score (mean 43.6, SE 2.72), while the 40 - to 49-year age group demonstrated the highest (mean 51.0, SE 1.00). The absence of floor or ceiling effects in this domain indicates that the scale effectively captured a comprehensive range of physical health outcomes without concentrating on the extremes. The observed decline in physical health scores with age is consistent with patterns reported in other populations, suggesting a cumulative impact of caregiving and aging.

Psychological health scores demonstrated relative stability across age groups for both male and female caregivers. Male caregivers exhibited scores ranging from 57.4 to 58.4, while female caregivers displayed scores from 49.5 (16 - 29 years) to 56.2 (40 - 49 years). Notwithstanding the overall consistency, female caregivers in the 16 - to 29-year age group reported significantly lower psychological well-being. No ceiling effects were observed, and only minimal floor effects were present (1.1%) among female caregivers aged 50+ years, indicating that extreme values were rare.

XSL•FO
RenderX

Table . WHOQOL-BREF norms for the Polish population of caregivers of individuals with rare diseases by sex and age.

	Male				Female			
	16 - 29 years	30 - 39 years	40 - 49 years	50+ years	16 - 29 years	30 - 39 years	40 - 49 years	50+ years
Physical health								
Mean (SE)	59.9 (10.04)	58.1 (2.92)	54.0 (2.29)	50.4 (4.72)	43.6 (2.72)	49.5 (0.91)	51.0 (1.00)	45.6 (1.79)
Range	49.9-69.9	21.2-91.5	7.2-92.8	14.5-85.0	17.9-73.7	6.9-92.4	4.8-100.0	7.4-81.7
Median (IQR)	59.9 (54.9-64.9)	60.8 (43.8-72.9)	54.2 (42.9-68.3)	51.6 (33.7-67.4)	44.6 (34.3-52.6)	48.9 (38.7-60.8)	50.0 (38.1-65.6)	47.4 (33.4-57.4)
Floor effect (%)	0	0	0	0	0	0	0	0
Ceiling effect (%)	0	0	0	0	0	0	0.3	0
Psychological health								
Mean (SE)	57.8 (12.33)	58.2 (3.26)	57.4 (2.00)	58.4 (3.98)	49.5 (2.77)	53.6 (0.99)	56.2 (0.96)	52.8 (1.87)
Range	45.5-70.2	7.2-95.5	15.9-91.4	17.9-90.7	21.2-87.7	3.6-95.9	11.7-100.0	0.0-90.7
Median (IQR)	57.8 (51.7-64.0)	63.1 (43.5-71.9)	58.3 (46.1-70.2)	58.0 (48.9-72.3)	49.5 (37.6-60.1)	53.8 (41.4-66.9)	57.7 (45.2-70.4)	54.0 (41.9-66.4)
Floor effect (%)	0	0	0	0	0	0	0	1.1
Ceiling effect (%)	0	0	0	0	0	0	0	0
Social relationships								
Mean (SE)	55.0 (4.75)	52.4 (3.21)	50.8 (2.67)	52.1 (4.82)	39.3 (3.34)	51.2 (1.17)	52.5 (1.05)	50.9 (1.91)
Range	50.2-59.7	7.7-92.3	0.0-100.0	17.3-100.0	0.0-75.0	0.0-100.0	0.0-100.0	0.0-84.5
Median (IQR)	55.0 (52.6-57.3)	57.5 (40.3-67.3)	52.3 (40.9-65.8)	57.4 (33.5-67.3)	41.3 (25.0-50.1)	50.0 (34.7-67.5)	52.2 (40.1-67.5)	52.2 (34.9-67.3)
Floor effect (%)	0	0	4.7	0	6.2	1.2	0.9	1.1
Ceiling effect (%)	0	0	1.6	5	0	1.8	1.1	0
Environment								
Mean (SE)	52.8 (2.02)	57.5 (2.39)	55.3 (1.84)	51.5 (4.09)	48.8 (2.73)	51.1 (0.92)	53.0 (0.93)	48.5 (1.91)
Range	50.8-54.8	24.5-87.2	15.8-83.4	23.0-84.7	14.2-73.3	6.8-97.8	9.5-94.2	9.1-87.8
Median (IQR)	52.8 (51.8-53.8)	59.5 (47.2-68.8)	54.1 (46.3-65.4)	50.2 (34.6-65.4)	48.0 (38.2-61.8)	51.3 (39.8-62.5)	52.0 (41.5-65.5)	50.2 (36.0-62.9)
Floor effect (%)	0	0	0	0	0	0	0	0
Ceiling effect (%)	0	0	0	0	0	0	0	0

[a]WHOQOL-BREF: World Health Organization Quality of Life-BREF.

The social relationships domain displayed greater variability, particularly among female caregivers. The lowest mean score was found in the youngest female group (16 - 29 years: mean 39.3, SE 3.34), suggesting social challenges in this group. In contrast, male caregivers demonstrated more stability, with scores ranging from 50.8 to 55.0 across all age groups. Significant floor effects were observed among female caregivers aged 16 - 29 years (6.2%) and male caregivers aged 40 - 49 years (4.7%), suggesting that many respondents experienced severe difficulties in social relationships. Ceiling effects were generally minimal, with the highest occurrence (5%) observed among male caregivers aged 50 years and older.

The scores in the environment domain demonstrated greater consistency across age groups. The mean score was highest among male caregivers aged 30 - 39 years (mean 57.5, SE 2.39) and lowest among female caregivers aged 50+ years (mean 48.5, SE 1.91). The absence of floor or ceiling effects indicated that responses were distributed uniformly across the scale. This dimension reflects a stable perception of environmental support and resources among caregivers, irrespective of age or sex.

XSL•FO
RenderX

The study's results indicate significant disparities in HRQoL within the sex and age categories of caregivers. Specifically, younger female caregivers exhibited lower scores in various domains of WHOQOL-BREF. These disparities can be attributed to the unique challenges associated with caregiving roles, personal health, and social support systems, particularly for specific demographic groups.

Discussion

Principal Findings

The WHOQOL-BREF is a widely used instrument for assessing QoL across diverse populations [4,28]. Its psychometric properties have been extensively studied, confirming its reliability and validity. However, certain aspects warrant further discussion, including its applicability across cultural contexts, domain-specific reliability concerns, and its use in both general and disease-specific populations [8,9,26,29-32].

Our study demonstrated that the WHOQOL-BREF is an appropriate tool for assessing QoL among caregivers of individuals with RDs. The instrument exhibited good internal consistency, with Cronbach α values generally exceeding 0.70. Internal consistency analysis showed values ranging from 0.69 for the social relationships domain to 0.84 for psychological health, indicating satisfactory to high reliability across domains. The item-level analysis identified specific areas where refinements could enhance reliability. Nevertheless, previous studies have reported lower reliability in the social relationships domain, highlighting a potential need for further refinement [6,8,33].

The construct validity of the WHOQOL-BREF is well established, with CFAs generally supporting its 4-domain structure. The instrument also demonstrates strong convergent and discriminant validity, correlating well with other health measures and effectively distinguishing between groups with differing health conditions [34-37]. Consistent with prior research, our findings confirm that the WHOQOL-BREF generally maintains good internal consistency, with Cronbach α values frequently exceeding 0.70. For instance, one study reported an α of 0.898 for the total score [33].

However, lower α values have often been observed in the social relationship domain, occasionally falling below 0.70. A Thai study reported a Cronbach α of 0.62 for this domain, while the other domains ranged from 0.81 to 0.85 [38]. Similarly, a Greek study found an α of 0.65 for social relationships, with other domains ranging from 0.66 to 0.87 [36]. Several factors may account for this pattern. The social relationships domain comprises only 3 items, which naturally limits internal consistency compared to domains with more items. Additionally, the items assess complex and sensitive aspects of life—such as personal relationships and sexual activity—which may be interpreted differently across individuals, particularly among caregivers of persons with RDs. Cultural, sex-based, and situational factors (eg, emotional exhaustion, social withdrawal, and limited time for interpersonal interaction) may further influence caregivers' responses, contributing to variability in this domain.

Factor analysis has been widely used to evaluate the underlying structure of the WHOQOL-BREF [6,39]. CFA results indicate that a unidimensional model yields poor fit indices, whereas the 4-domain model significantly improves model fit. Further refinements, including correlated error terms, have led to an optimized model with improved goodness-of-fit measures (χ^2/df, CFI, TLI, RMSEA, and SRMR). These findings confirm the structural validity of the WHOQOL-BREF and reinforce its intended dimensional structure [8,32]. Although CFA generally supports the 4-domain model (physical health, psychological health, social relationships, and environment domains), some studies have questioned this structure, reporting that certain items do not clearly discriminate between domains or correlate more strongly with an unintended domain [32,34,36,38]. Exploratory factor analysis in a Norwegian study also resulted in a 4-component solution, partially supporting the established domain structure [8]. Some studies suggest that the physical health domain contributes most significantly to overall QoL [33].

The WHOQOL-BREF has demonstrated strong psychometric properties across various countries and cultures. Research supports its discriminant validity, content validity, internal consistency, and test-retest reliability across different cultural settings, suggesting its suitability for cross-cultural comparisons [4,8,36,38,40]. The instrument has been translated and adapted for multiple languages, including Polish [6], Korean [33], Spanish [35,40], Portuguese [41], Thai [38], Bahasa Indonesia [42], Mongolian [43], Persian [39], Danish [32], Norwegian [8], and French [9]. The translation process generally adheres to rigorous methodologies, such as forward and backward translation and cultural adaptation procedures. Some adaptations have been proposed to enhance cultural relevance, including the incorporation of additional items related to work satisfaction, social activities, and nutrition [36]. While these modifications may improve local applicability, they could also impact comparability across studies [40].

In studies conducted within the general population, the WHOQOL-BREF has been used to establish normative QoL scores, facilitating demographic comparisons [7,9,44]. Factors such as age, sex, education level, and health status have been identified as influential [4,7,9,32,42,44,45]. For instance, some studies suggest that female caregivers report higher social QoL [26,46], while male caregivers tend to report better psychological QoL [7,9,46]. Physical health scores typically decline with age, reflecting the natural progression of age-related health issues [3,4,46,47].

In clinical settings, the WHOQOL-BREF has been validated or used for various disease groups, including cancer, chronic conditions, mental disorders, HIV/AIDS, Huntington disease, Prader-Willi syndrome, hemophilia, neurofibromatosis, rheumatoid arthritis, and spinal cord injury [4,30,31,34,36,48-54]. It serves as a valuable tool for evaluating the impact of illness on QoL and monitoring treatment effectiveness [4,28]. However, some validation studies have raised concerns regarding the internal consistency of the social relationships domain [8,29]. The instrument's generic nature allows for comparisons across different medical conditions,

XSL•FO
RenderX

ensuring broad applicability in health research. Factorial invariance testing has confirmed that the WHOQOL-BREF consistently measures QoL across different patient groups [4,8,34].

Despite its strengths, certain limitations of the WHOQOL-BREF should be acknowledged. The social relationships domain has exhibited weaker reliability in some studies, suggesting a need for refinement or supplementation with additional items. While the environment domain is a notable strength of the instrument, some researchers have recommended incorporating more sensitive response options to improve measurement precision [6,28,29,46].

The WHOQOL-BREF's reliance on self-report data enhances respondent acceptability but also introduces potential biases, such as social desirability and response tendencies. Adherence to standardized scoring guidelines is essential for ensuring accurate interpretation, and researchers should consider variations in response patterns across different populations [6,29,32].

In our study, QoL domain scores were normalized to a 0 - 100 scale. Factor loadings ranged from 0.363 to 0.801, indicating varying contributions of items to their respective constructs. The psychological health domain demonstrated the highest mean score (mean 54.9, SE 0.59), while the physical health domain had the lowest (mean 50.2, SE 0.59). Normality tests showed slight deviations, particularly in the psychological health and social relationships domains, though skewness and kurtosis values remained within acceptable limits. Minimal floor and ceiling effects ensured adequate response distribution. QoL norms were established for a Polish population of caregivers for individuals with RDs, stratified by sex and age group. Male caregivers generally reported higher physical and psychological health scores than female caregivers. Younger female caregivers (16 - 29 years) exhibited the lowest scores in several domains, suggesting potential vulnerabilities in social and psychological well-being. These findings highlight disparities related to caregiving burden, sex, and age-related factors.

Limitations

It is important to consider the limitations of this study before interpreting its results. First, although 942 caregivers replied and completed the survey, the number of caregivers of persons with RDs in Poland is significantly higher. It should be noted, however, that although it can be difficult to extend the results to the total population of caregivers of persons with RDS in Poland, to date, there is no register of individuals with rare conditions in the country. Second, since convenience sampling was used in this survey and the study sample was self-selected, it could further compromise the sample's representativeness and the results' generalizability. Third, since caregivers of persons with different RDs took part in the study, there is a possible bias related to challenges related to different conditions, symptoms, and prognoses. However, since most RDs are small in numbers, it would be difficult, if possible at all, to validate the questionnaire for a single RD. Moreover, despite some differences in caregivers' experiences, numerous previous studies demonstrated that most parents of persons with RDs face similar challenges and express similar needs. Consequently, we believe that the results may still apply to the entire population of family caregivers of persons with RDs in Poland. Especially because the sample is quite diverse, encompassing both mothers and fathers as well as various age groups. Fourth, the self-administering nature of the WHOQOL-BREF is part of its design. However, although the administration may elicit varying degrees of true responses, this bias is consistent across all questionnaires and has no bearing on the accuracy of the psychometric analyses, which are primarily comparative in nature. Fifth, we gathered our data before the Rare Disease Plan for 2024 - 2025 was formally approved by the Council of Ministers in August 2024. Sixth, a comparable sample of caregivers of healthy individuals or those experiencing more common chronic diseases could have been beneficial to the study. Seventh, because the study was conducted as a web-based survey, there is a risk that it did not reach caregivers who lived in distant areas without internet access, who were hesitant to use technological devices and social media, or who did not belong to digital associations, foundations, and organizations of patients with RDs or support groups that aided in collecting the data. Consequently, while there is no direct generalizability of the results shown here to other cross-national samples, it is challenging to evaluate bias and the causes of nonresponses when using web-based survey data. Eighth, a potential bias is also present because only 130 fathers responded to the survey. However, it should be explained that while the sample was obtained through convenience sampling, it is well-documented that in the population of family caregivers, there is a significant underrepresentation of male caregivers. Thus, while these sex differences in caregiving could affect how respondents report their QoL, particularly in areas like psychological health and social relationships due to the small number of male respondents, sex-based factor analysis was not possible. Future research should ensure more balanced sex representation and explore sex differences in WHOQOL-BREF performance. Moreover, the 2006 Polish validation of WHOQOL-BREF did not account for sex, age differences, or caregivers of individuals with RDs, highlighting the need for validation in specific populations and subgroup adaptations. Another limitation of this study is the inclusion of caregivers of individuals with 158 different RDs, which prevented detailed clinical characterization of care recipients. However, only some syndromes, that is, Dravet, Angelman, Williams, 22q11DS, and Duchenne or Becker muscular dystrophy, were represented by multiple caregivers; most conditions had only a few or even single respondents. This reflects the nature of rare, ultrarare, and hyperrare disorders, which affect very small patient populations and make it difficult to validate any tool. Future studies should consider focusing on more homogeneous caregiver groups. Next, since no other QoL instrument has ever been validated in the population of family caregivers of persons with RDs in Poland, this lack of external validation constitutes a limitation of our study and highlights the need for future research dedicated specifically to this group. Finally, this study is also limited by the WHOQOL-BREF's potential insufficiency in capturing social aspects of caregiving; future research should consider using complementary tools like the Zarit Burden Interview.

XSL•FO
RenderX

Conclusions

The WHOQOL-BREF remains a reliable and valid instrument for assessing QoL across diverse populations. Its 4D structure is robust, and normative data provide valuable benchmarks for assessing QoL in specific groups, such as caregivers. Future research should focus on refining the social relationships domain and addressing minor fit inconsistencies to enhance the scale's applicability further. An important question arises regarding whether the social relationships domain, due to inherent differences between societies, patient groups, and caregivers of various conditions, can be conceptualized as a universal construct. While the WHOQOL-BREF remains a valuable tool for assessing QoL in diverse populations, careful consideration of item profiles and potential supplementary modules for specific age groups may further optimize its utility.

Acknowledgments

The authors thank all the caregivers who participated in the study. No generative artificial intelligence tool was used in the writing of this manuscript, calculation, or visualizing the study results. All text, calculation, table, and figures were prepared by the authors themselves.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to reasons of sensitivity but are available from the corresponding author on reasonable request.

Authors' Contributions

JD and DW conceptualized the study. JD, PJ, and DW developed the methodology. JD collected the data. DW curated the data and contributed to validation. PJ performed the formal analysis, software development, validation, and visualization. JD, PJ, and DW had full access to the data, wrote the original draft, and approved the final manuscript. JD administrated and supervised the study. All authors contributed equally to this work.

Conflicts of Interest

None declared.

Checklist 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.
[DOCX File, 34 KB - publichealth_v11i1e72590_app1.docx]

References

1. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995 Nov;41(10):1403-1409. [doi: 10.1016/0277-9536(95)00112-k] [Medline: 8560308]
2. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. Soc Sci Med 1998 Jun;46(12):1569-1585. [doi: 10.1016/s0277-9536(98)00009-4] [Medline: 9672396]
3. Saxena S, Carlson D, Billington R, WHOQOL Group. World Health Organisation Quality Of Life. The WHO Quality of Life assessment instrument (WHOQOL-Bref): the importance of its items for cross-cultural research. Qual Life Res 2001;10(8):711-721. [doi: 10.1023/a:1013867826835] [Medline: 11871592]
4. Hawthorne G, Herrman H, Murphy B. Interpreting the WHOQOL-Brèf: preliminary population norms and effect sizes. Soc Indic Res 2006 May;77(1):37-59. [doi: 10.1007/s11205-005-5552-1]
5. ten Hoedt AE, Maurice-Stam H, Boelen CCA, et al. Parenting a child with phenylketonuria or galactosemia: implications for health-related quality of life. J Inherit Metab Dis 2011 Apr;34(2):391-398. [doi: 10.1007/s10545-010-9267-3] [Medline: 21290186]
6. Jaracz K, Kalfoss M, Górna K, Baczyk G. Quality of life in Polish respondents: psychometric properties of the Polish WHOQOL-Bref. Scand J Caring Sci 2006 Sep;20(3):251-260. [doi: 10.1111/j.1471-6712.2006.00401.x] [Medline: 16922978]
7. West EC, Williams LJ, Stuart AL, Pasco JA. Quality of life in south-eastern Australia: normative values for the WHOQOL-BREF in a population-based sample of adults. BMJ Open 2023 Dec 10;13(12):e073556. [doi: 10.1136/bmjopen-2023-073556] [Medline: 38072488]
8. Hanestad BR, Rustøen T, Knudsen O, Lerdal A, Wahl AK. Psychometric properties of the WHOQOL-BREF questionnaire for the Norwegian general population. J Nurs Meas 2004;12(2):147-159. [doi: 10.1891/jnum.2004.12.2.147] [Medline: 16092712]
9. Baumann C, Erpelding ML, Régat S, Collin JF, Briançon S. The WHOQOL-BREF questionnaire: French adult population norms for the physical health, psychological health and social relationship dimensions. Rev Epidemiol Sante Publique 2010 Feb;58(1):33-39. [doi: 10.1016/j.respe.2009.10.009] [Medline: 20096515]

XSL•FO
RenderX

10. Angelis A, Tordrup D, Kanavos P. Socio-economic burden of rare diseases: a systematic review of cost of illness evidence. Health Policy 2015 Jul;119(7):964-979. [doi: 10.1016/j.healthpol.2014.12.016] [Medline: 25661982]
11. Walkowiak D, Domaradzki J. Parental experiences and needs of caring for a child with 22q11.2 deletion syndrome. Orphanet J Rare Dis 2023 Dec 4;18(1):379. [doi: 10.1186/s13023-023-02980-3] [Medline: 38049864]
12. Domaradzki J, Walkowiak D. Quality of life and caregiving burden associated with parenting a person with Duchenne/Becker muscular dystrophy in Poland. Orphanet J Rare Dis 2024 Nov 30;19(1):450. [doi: 10.1186/s13023-024-03481-7] [Medline: 39616352]
13. Michalík J. Rare diseases and the quality of life of caregivers. Soc Welf 2012 Jun 29;2(1):8-17. [doi: 10.15388/SW.2012.28210]
14. Walkowiak D, Domaradzki J, Mozrzymas R, Kałużny Ł, Walkowiak J. Navigating the unique challenges of caregiving for children with rare diseases: are the care experiences of all caregivers the same? A focus on life-limiting rare diseases. J Clin Med 2024 Aug 1;13(15):4510. [doi: 10.3390/jcm13154510] [Medline: 39124776]
15. Domaradzki J, Walkowiak D. Emotional experiences of family caregivers of children with Dravet syndrome. Epilepsy Behav 2023 May;142:109193. [doi: 10.1016/j.yebeh.2023.109193] [Medline: 37028149]
16. Chung CCY, Ng NYT, Ng YNC, et al. Socio-economic costs of rare diseases and the risk of financial hardship: a cross-sectional study. Lancet Reg Health West Pac 2023 May;34:100711. [doi: 10.1016/j.lanwpc.2023.100711] [Medline: 37283971]
17. Ouattara A, Resseguier N, Cano A, et al. Individual and family determinants for quality of life in parents of children with inborn errors of metabolism requiring a restricted diet: a multilevel analysis approach. J Pediatr 2023 Mar;254:39-47. [doi: 10.1016/j.jpeds.2022.08.060] [Medline: 36265570]
18. Boettcher J, Denecke J, Barkmann C, Wiegand-Grefe S. Quality of life and mental health in mothers and fathers caring for children and adolescents with rare diseases requiring long-term mechanical ventilation. Int J Environ Res Public Health 2020 Dec 2;17(23):33276595. [doi: 10.3390/ijerph17238975] [Medline: 33276595]
19. Bamber MD, Mahony H, Spratling R. Mothers of children with special health care needs: exploring caregiver burden, quality of life, and resiliency. J Pediatr Health Care 2023;37(6):643-651. [doi: 10.1016/j.pedhc.2023.06.003] [Medline: 37516944]
20. Trickett J, Heald M, Oliver C. Sleep in children with Angelman syndrome: parental concerns and priorities. Res Dev Disabil 2017 Oct;69:105-115. [doi: 10.1016/j.ridd.2017.07.017] [Medline: 28844022]
21. Iskrov G, Stefanov R, Ferrelli RM. Health systems for rare diseases: financial sustainability. Ann Ist Super Sanita 2019;55(3):270-275. [doi: 10.4415/ANN_19_03_13] [Medline: 31553322]
22. Filosto M, Mancuso M, Federico A, Padovani A. Unravelling the mysteries: a reflection on rare neurological diseases. Neurol Sci 2024 Mar;45(3):833-835. [doi: 10.1007/s10072-024-07374-2] [Medline: 38302818]
23. Feudtner C, Nye RT, Boyden JY, et al. Association between children with life-threatening conditions and their parents' and siblings' mental and physical health. JAMA Netw Open 2021 Dec 1;4(12):e2137250. [doi: 10.1001/jamanetworkopen.2021.37250] [Medline: 34928360]
24. Guarany NR, Vanz AP, Wilke M, Instituto Genética Para Todos (IGPT), et al. Mucopolysaccharidosis: caregiver quality of life. J Inborn Errors Metab Screen Latin American Society Inborn Errors and Neonatal Screening (SLEIMPN) 2015;3:e150001. [doi: 10.1177/2326409815613804]
25. Silva IMW, Gil-da-Silva-Lopes VL. Comprehensive insights into health services accessibility and quality of life of families with individuals with 22q11.2 deletion syndrome in Brazil. Orphanet J Rare Dis 2024 Jul 6;19(1):255. [doi: 10.1186/s13023-024-03273-z] [Medline: 38971792]
26. von Steinbüchel N, Lischetzke T, Gurny M, Eid M. Assessing quality of life in older people: psychometric properties of the WHOQOL-BREF. Eur J Ageing 2006 Jun;3(2):116-122. [doi: 10.1007/s10433-006-0024-2] [Medline: 28794757]
27. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD] [Medline: 17947786]
28. Development of the World Health Organization WHOQOL-BREF quality of life assessment. The WHOQOL Group. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/s0033291798006667] [Medline: 9626712]
29. Young CA, Mills R, Al-Chalabi A, et al. Measuring quality of life in ALS/MND: validation of the WHOQOL-BREF. Amyotroph Lateral Scler Frontotemporal Degener 2020 Jun 27;21(5-6):364-372. [doi: 10.1080/21678421.2020.1752244] [Medline: 32597226]
30. Pietrzykowski MO, Vranceanu AM, Macklin EA, Mace RA. Minimal clinically important difference in the World Health Organization Quality of Life Brief (WHOQOL-BREF) for adults with neurofibromatosis. Qual Life Res 2024 May;33(5):1233-1240. [doi: 10.1007/s11136-023-03596-7] [Medline: 38214851]
31. Alilović M, Peros-Golubicić T, Radosević-Vidacek B, et al. WHOQOL-BREF questionnaire as a measure of quality of life in sarcoidosis. Coll Antropol 2013 Sep;37(3):701-706. [Medline: 24308206]
32. Noerholm V, Groenvold M, Watt T, Bjorner JB, Rasmussen NA, Bech P. Quality of life in the Danish general population—normative data and validity of WHOQOL-BREF using Rasch and item response theory models. Qual Life Res 2004 Mar;13(2):531-540. [doi: 10.1023/B:QURE.0000018485.05372.d6] [Medline: 15085925]

XSL•FO
RenderX

33. Min SK, Kim KI, Lee CI, Jung YC, Suh SY, Kim DK. Development of the Korean versions of WHO Quality of Life scale and WHOQOL-BREF. Qual Life Res 2002 Sep;11(6):593-600. [doi: 10.1023/a:1016351406336] [Medline: 12206580]
34. Yao G, Wu CH. Factorial invariance of the WHOQOL-BREF among disease groups. Qual Life Res 2005 Oct;14(8):1881-1888. [doi: 10.1007/s11136-005-3867-7] [Medline: 16155775]
35. Lucas-Carrasco R. The WHO Quality of Life (WHOQOL) questionnaire: Spanish development and validation studies. Qual Life Res 2012 Feb;21(1):161-165. [doi: 10.1007/s11136-011-9926-3] [Medline: 21611868]
36. Ginieri-Coccossis M, Triantafillou E, Tomaras V, Soldatos C, Mavreas V, Christodoulou G. Psychometric properties of WHOQOL-BREF in clinical and health Greek populations: incorporating new culture-relevant items. Psychiatriki 2012;23(2):130-142. [Medline: 22796911]
37. Krägeloh CU, Kersten P, Rex Billington D, et al. Validation of the WHOQOL-BREF quality of life questionnaire for general use in New Zealand: confirmatory factor analysis and Rasch analysis. Qual Life Res 2013 Aug;22(6):1451-1457. [doi: 10.1007/s11136-012-0265-9] [Medline: 22983781]
38. Kangwanrattanakul K, Krägeloh CU. Psychometric evaluation of the WHOQOL-BREF and its shorter versions for general Thai population: confirmatory factor analysis and Rasch analysis. Qual Life Res 2024 Feb;33(2):335-348. [doi: 10.1007/s11136-023-03521-y] [Medline: 37906345]
39. Yousefy AR, Ghassemi GR, Sarrafzadegan N, Mallik S, Baghaei AM, Rabiei K. Psychometric properties of the WHOQOL-BREF in an Iranian adult sample. Community Ment Health J 2010 Apr;46(2):139-147. [doi: 10.1007/s10597-009-9282-8] [Medline: 20063062]
40. Benitez-Borrego S, Guàrdia-Olmos J, Urzúa-Morales A. Factorial structural analysis of the Spanish version of WHOQOL-BREF: an exploratory structural equation model study. Qual Life Res 2014 Oct;23(8):2205-2212. [doi: 10.1007/s11136-014-0663-2] [Medline: 24609387]
41. Silva WD, Bonafé FSS, Marôco J, Maloa BFS, Campos J. Psychometric properties of the World Health Organization Quality of Life Instrument-Abbreviated version in Portuguese-speaking adults from three different countries. Trends Psychiatry Psychother 2018;40(2):104-113. [doi: 10.1590/2237-6089-2017-0058]
42. Purba FD, Hunfeld JAM, Iskandarsyah A, et al. Quality of life of the Indonesian general population: test-retest reliability and population norms of the EQ-5D-5L and WHOQOL-BREF. PLoS One 2018;13(5):e0197098. [doi: 10.1371/journal.pone.0197098] [Medline: 29750806]
43. Bat-Erdene E, Hiramoto T, Tumurbaatar E, et al. Quality of life in the general population of Mongolia: normative data on WHOQOL-BREF. PLoS One 2023;18(9):e0291427. [doi: 10.1371/journal.pone.0291427] [Medline: 37773966]
44. Cruz LN, Polanczyk CA, Camey SA, Hoffmann JF, Fleck MP. Quality of life in Brazil: normative values for the WHOQOL-bref in a southern general population sample. Qual Life Res 2011 Sep;20(7):1123-1129. [doi: 10.1007/s11136-011-9845-3] [Medline: 21279448]
45. Perera HN, Izadikhah Z, O'Connor P, McIlveen P. Resolving dimensionality problems with WHOQOL-BREF item responses. Assessment 2018 Dec;25(8):1014-1025. [doi: 10.1177/1073191116678925] [Medline: 27872348]
46. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
47. Baumann M, Couffignal S, Le Bihan E, Chau N. Life satisfaction two-years after stroke onset: the effects of gender, sex occupational status, memory function and quality of life among stroke patients (Newsqol) and their family caregivers (Whoqol-bref) in Luxembourg. BMC Neurol 2012 Sep 25;12(1):105. [doi: 10.1186/1471-2377-12-105] [Medline: 23009364]
48. Taylor WJ, Myers J, Simpson RT, McPherson KM, Weatherall M. Quality of life of people with rheumatoid arthritis as measured by the World Health Organization Quality of Life Instrument, short form (WHOQOL-BREF): score distributions and psychometric properties. Arthritis Rheum 2004 Jun 15;51(3):350-357. [doi: 10.1002/art.20398] [Medline: 15188318]
49. Manee F, Ateya Y, Rassafiani M. A comparison of the quality of life of Arab mothers of children with and without chronic disabilities. Phys Occup Ther Pediatr 2016 Aug;36(3):260-271. [doi: 10.3109/01942638.2015.1076558] [Medline: 26566139]
50. Blay SL, Marchesoni MSM. Association among physical, psychiatric and socioeconomic conditions and WHOQOL-Bref scores. Cad Saude Publica 2011 Apr;27(4):677-686. [doi: 10.1590/s0102-311x2011000400007] [Medline: 21603751]
51. Junca E, Pino M, Santamaría-García H, Baez S. Brain, cognitive, and physical disability correlates of decreased quality of life in patients with Huntington's disease. Qual Life Res 2023 Jan;32(1):171-182. [doi: 10.1007/s11136-022-03220-0] [Medline: 35978062]
52. Trindade GC, Viggiano LGDL, Brant ER, et al. Evaluation of quality of life in hemophilia patients using the WHOQOL-bref and Haemo-A-Qol questionnaires. Hematol Transfus Cell Ther 2019;41(4):335-341. [doi: 10.1016/j.htct.2019.03.010] [Medline: 31409581]
53. Chiu VY, Tsai LP, Wei JT, Tzeng IS, Wu HC. Motor performance in Prader-Willi syndrome patients and its potential influence on caregiver's quality of life. PeerJ 2017;5:e4097. [doi: 10.7717/peerj.4097] [Medline: 29255649]
54. Ihara H, Ogata H, Sayama M, et al. QOL in caregivers of Japanese patients with Prader-Willi syndrome with reference to age and genotype. Am J Med Genet A 2014 Sep;164A(9):2226-2231. [doi: 10.1002/ajmg.a.36634] [Medline: 24953026]

XSL•FO
RenderX

Abbreviations

CFA: confirmatory factor analysis
CFI: Comparative Fit Index
HRQoL: health-related quality of life
QoL: quality of life
RD: rare disease
RMSEA: root-mean-square error of approximation
SRMR: standardized root-mean-square residual
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TLI: Tucker-Lewis Index
WHO: World Health Organization
WHOQOL: World Health Organization Quality of Life
WHOQOL-BREF: World Health Organization Quality of Life-BREF

Edited by A Mavragani; submitted 12.02.25; peer-reviewed by A Adeoye, M Canete-Lairla; revised version received 16.04.25; accepted 08.05.25; published 18.07.25.

Please cite as:
Jabkowski P, Domaradzki J, Walkowiak D
Quality of Life Among Family Caregivers of Individuals With Rare Diseases: Web-Based Population Study on the Validity and Reliability of the Polish World Health Organization Quality of Life-BREF Questionnaire
JMIR Public Health Surveill 2025;11:e72590
URL: https://publichealth.jmir.org/2025/1/e72590
doi:10.2196/72590

© Piotr Jabkowski, Jan Domaradzki, Dariusz Walkowiak. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Estimating the Population Size of People Who Inject Drugs in 3 Cities in Zambia: Capture-Recapture, Successive Sampling, and Bayesian Consensus Estimation Methods

Lauren Parmley[1,2]; Giles Reid[3]; Joyce J Neal[4]; Brave Hanunka[5]; Leigh Tally[5]; Lophina Chilukutu[6]; Tepa Nkumbula[6]; Chipili Mulemfwe[6]; Lazarous Chelu[6]; Ray Handema[7]; John Mwale[8]; Kennedy Mutale[9]; Lloyd Mulenga[10]; Anne F McIntyre[11]; Neena M Philip[3]; Hannah Chung[3]; Maria Lahuerta[1,12]

[1]Strategic Information Unit, ICAP at Columbia University, New York, NY, United States

[10]Ministry of Health, Lusaka, Zambia

[11]Division of Global HIV & TB, Global Health Center, U.S. Centers for Disease Control and Prevention, Atlanta, GA, United States

[12]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, United States

[2]Bilateral Health Office, U.S. Agency for International Development/Southern Africa, 100 Totius Street, Pretoria, South Africa

[3]Survey Unit, ICAP at Columbia University, New York, NY, United States

[4]Division of Global HIV & TB, Global Health Center, U.S. Centers for Disease Control and Prevention, Phnom Penh, Cambodia

[5]Division of Global HIV & TB, Global Health Center, U.S. Centers for Disease Control and Prevention, Lusaka, Zambia

[6]ICAP at Columbia University, Lusaka, Zambia

[7]Department of Biomedical Sciences, Tropical Diseases Research Centre, Ndola, Zambia

[8]National HIV/AIDS/STI/TB Council, Lusaka, Zambia

[9]People Who Inject Drugs Sector, Key Populations Consortium, Lusaka, Zambia

Corresponding Author:
Lauren Parmley
Strategic Information Unit, ICAP at Columbia University, New York, NY, United States

Abstract

Background: Accurate population size estimates (PSE) of key populations—those disproportionately affected by HIV—are critical to forecast need and inform HIV prevention and treatment programs, though they can be difficult to ascertain due to low visibility of these groups. In Zambia, reliable estimates on the number of people who inject drugs are limited, inhibiting public health response.

Objective: We sought to estimate the population size of people who inject drugs in 3 large cities in Zambia, assess how PSEs vary across different estimation methods, and explore the strengths and limitations of each approach.

Methods: We applied 2-source capture-recapture (2S-CRC), 3-source capture-recapture (3S-CRC), and successive sampling population size estimation (SS-PSE) methods in Lusaka, Livingstone, and Ndola, Zambia. 3S-CRC methods included location-based 2S-CRC in combination with a respondent-driven sampling (RDS) survey. Data were collected from November 2021 to February 2022 and analyzed using a Bayesian nonparametric latent class model. SS-PSEs were produced using the RDS recruitment and network sizes. Kruskal tests and general linear models were used to examine sociodemographic and behavioral factors associated with being captured in 2S-CRC among RDS participants. Final city population estimates, incorporating 3S-CRC and SS-PSE with imputed visibility estimates, were generated using a Bayesian consensus estimator.

Results: Bayesian consensus PSEs ranged between 0.5% and 1.8% of the adult male population and were below 1% of the total adult population in each city. Consensus estimates were highest in Lusaka (3700, 95% credible interval [CRI] 1500 - 7500), followed by Ndola (2200, 95% CRI 1600 - 2900) and Livingstone (1200, 95% CRI 900 - 1,900). There was variability in estimates by method, with SS-PSE with imputed visibility generally providing the lowest estimates across cities, excluding Lusaka. Across methods, PSEs and uncertainty bounds (95% confidence interval [CI] or CRI depending on method) ranged from 1510 (95% CRI 1030 - 2070) to 4350 (95% CI 1410 - 18,890) in Lusaka, 360 (95% CI 290 - 530) to 2620 (95% CRI 1510 - 4680) in Livingstone, and 760 (95% CI 390 - 3060) to 4030 (95% CRI 960 - 5480) in Ndola. In all cities, fewer recaptures occurred in capture 3 (RDS) than with location sampling via 2S-CRC. Though results varied across cities, RDS participants captured through 2S-CRC differed from those captured solely through RDS in sociodemographic and behavioral risk factors, including housing, education, injection or needle sharing frequency, time since last injection, receipt of drug treatment, and experience with a peer educator in at least one city.

XSL•FO
RenderX

Conclusions: This study used rigorous methods to produce PSEs in Zambia, and is the first to produce these for major geographies in the country. Through RDS, 3S-CRC reached people who inject drugs with distinct characteristics that were less accessible via location-based sampling (2S-CRC), yielding a PSE that may better reflect the population and informing the Bayesian consensus estimate. Findings from this study can guide program planning and future surveillance activities.

(JMIR Public Health Surveill 2025;11:e66551) doi:10.2196/66551

KEYWORDS

people who inject drugs; population size estimation; Zambia; PWID; HIV

Introduction

In Eastern and Southern Africa, 23% of new HIV infections are estimated to occur among key populations and their partners, roughly 1% of which are estimated to be among people who inject drugs based on available data [1,2]. Though population-level HIV estimates for people who inject drugs are limited, existing data indicate that people who inject drugs experience an HIV incidence rate 6.5 times higher than that of the general population in the region, with the proportion of new HIV infections and the incidence rate ratio slightly increasing from 2010 for this group [2]. Moreover, people who inject drugs endure other harms due to injection drug use, including infections from other blood-borne viruses such as hepatitis C virus (HCV) and hepatitis B virus (HBV), injury, and overdose [3].

Accurate population size estimates (PSEs) of key populations are critical to forecast need and inform HIV and HBV/HCV prevention, care, and treatment program planning, yet they can be difficult to ascertain due to the limited visibility of these groups. Difficulties in sampling key populations, including people who inject drugs, due to discrimination and criminalization, as well as nondisclosure of key populations in health care settings, can hinder PSEs using empiric methods. As a result, the availability and quality of PSEs for people who inject drugs vary. Only half of the countries in sub-Saharan Africa have injection drug use estimates—estimates of varied quality [3]. Although 1.4 million people have been estimated to inject drugs in the region, numbers may be as high as 3.1 million [3].

In Zambia, where punitive laws against drugs and injection equipment impede harm reduction services for people who inject drugs [4,5], estimates on the number of people who inject drugs are largely absent, limiting country-level planning and service delivery coverage estimates. While a formative qualitative assessment of people who use drugs was conducted in 4 cities in Zambia between 2013 and 2015 [6] and a size estimation study with people who inject drugs was conducted in 3 districts in Zambia in 2021 [7], prior methods used were of poor quality, included broad eligibility definitions related to HIV risk, or omitted major geographies in the country.

We sought to estimate the population size of people who inject drugs in 3 large cities in Zambia and document how PSEs vary across estimation methods for these groups. Multiple methods were used in the absence of a gold standard method for PSEs. We present findings, strengths, and limitations of each estimation method, including analyses that highlight method attributes, as well as our approach and results in generating a Bayesian consensus estimate for each city.

Methods

Study Setting

Data were collected from November 2021 to February 2022. Survey cities were selected to include 3 of the most populous cities in Zambia, which have a high number of people who use drugs according to PSEs from a prior study [6]. Each city has a different typology, with Livingstone serving as a major tourist destination and border town, the capital Lusaka serving as the economic and industrial hub of the country, and Ndola serving as a gateway to the Copperbelt province and mining region. Small area estimation techniques suggest the HIV prevalence among the general population in Livingstone, Lusaka, and Ndola is 14.9%, 16.5%, and 14.1%, respectively [8], higher than the national adult HIV prevalence of 9.8% [9].

2-Source Capture-Recapture and 3-Source Capture-Recapture: Captures 1 and 2

In each city, we implemented multiple source capture-recapture where people who inject drugs were “sampled” through 3 captures. This included the distribution of unique objects at places or “hotspots” where people who inject drugs congregate at 2 points in time (captures 1 and 2), followed by a biobehavioral survey (BBS; capture 3). Details and assumptions on capture-recapture methods have been described previously, including that capture-recapture relies on assumptions specific to population closure, capture independence, homogeneous capture probabilities, and clearly defined captures [10-12]. For captures 1 and 2, implemented a week apart, volunteers had a target sample size to distribute up to twice the number of unique objects as the target BBS sample size per capture in each city (target sample size and sample size calculations for the BBS are described further below). Distribution locations and appropriate distribution objects (black and white bracelets for captures 1 and 2, respectively) were identified through a formative assessment and through support from organizations who work with people who inject drugs. Volunteers were assigned to distribute objects at these locations as well as additional locations in the volunteer’s neighborhood or preferred hangout place. Distribution occurred over 3 days to limit the possibility of someone receiving objects from different volunteers. Each day of distribution, volunteers (13-27 per city and capture) received paper-based distribution logs to be completed during distribution, along with their allotted number of objects to distribute in their allocated daily catchment areas. Different volunteers were used for each capture to reduce

XSL•FO
RenderX

dependence between captures, and any participating volunteer was excluded from the seed selection for capture 3. During each of the first 2 captures, volunteers approached both known and potential members of the people who inject drugs community and verified whether individuals had injected drugs within the past 3 months and were 16 years of age or older. After confirming eligibility (same as BBS eligibility below), volunteers inquired whether individuals had previously received the unique object and briefed them about the unique object distribution process and the upcoming survey. If individuals had not previously received the object in that capture, they were offered an object. People who inject drugs who accepted objects in capture 1 were encouraged to keep their object for 2 weeks and told that they may be asked about it in the future. In capture 2, volunteers verified if individuals had received the unique object distributed in capture 1, though they were not required to have it on them as long as they confirmed receiving it. Eligibility, object acceptance or refusal, prior receipt of object, date, time, and location of distribution were recorded on the log and later entered via tablet questionnaire. Any objects that were not distributed were returned to the survey team.

3-Source Capture-Recapture: Capture 3 (BBS)

Capture 3 comprised a BBS with survey methods aligned to the Biobehavioural Survey Guidelines for Populations at Risk for HIV [13]. Participants were recruited using respondent-driven sampling (RDS), a chain referral approach used to recruit populations for whom no sampling frame exists [14]. Individuals were eligible to participate in the BBS if they met the following criteria: self-reported drug injection for nonmedical purposes in the past 3 months; aged 16 years or older; lived in the surveyed city for the past 3 months; spoke English, Bemba, Kaonde, Lozi, Nyanja, or Tonga; willing to provide verbal informed consent; and in possession of a valid survey coupon (for coupon recipients). Seeds, well-connected people who inject drugs who met the eligibility criteria, were identified through a formative assessment and through support from organizations serving people who inject drugs. Seeds were purposively selected to ensure diversity across age, sex, education, area of residence, language, nationality, socioeconomic status, and key populations organization affiliation, and were instructed to recruit 3 of their peers via coupon. Peers who were eligible and participated in the BBS were provided 3 coupons and encouraged to recruit 3 of their peers, and so on.

After providing consent, tablet-based questionnaires were administered via interviewer and covered topics such as demographics, sexual and injection risk behaviors, knowledge, attitudes, and practices related to HIV, stigma and discrimination, and social and structural networks. Data validation and completion checks were preprogrammed to minimize data entry errors. To inform 3-source capture-recapture (3S-CRC), participants were asked about object acceptance or refusal and prior receipt of objects for both captures 1 and 2. Where participants did not physically possess the objects during the interview, they were asked to identify the objects from several images. Interviews lasted approximately 1 hour, after which participants who provided consent were tested for HIV, syphilis, HCV, and HBV.

Using the sample size calculator for survey-based viral load suppression in the Biobehavioural Survey Guidelines for Populations at Risk for HIV as is best practice for BBS [13] and assuming a design effect of 2 and a nonresponse rate of 5%, we calculated a sample size of 195 HIV-positive participants would produce a 2-sided 95% confidence interval (CI) ranging from 50% to 70% when viral load suppression among HIV-positive participants was 60%. To achieve a sample size of 195 HIV-positive participants per survey city, assuming an HIV prevalence of 25%, we estimated a sample size of 780 participants per survey city. After consultation with key population stakeholders in Zambia, it was determined that the target sample size of 780 people who inject drugs would not be achievable in each city, given the expected size of the population and the willingness of people who inject drugs to participate in a study. This sample was subsequently allocated across the 3 cities based on population size and recruitment feasibility. Each city's target sample size was then increased slightly to improve the precision of city estimates and better assure statistical equilibrium and convergence (Lusaka: n=350, Livingstone: n=235, and Ndola: n=260).

Successive Sampling Population Size Estimation

Successive sampling population size estimation (SS-PSE) methods [15] with and without imputed visibility were also used to estimate the population size of people who inject drugs in the survey cities. SS-PSE models the total number of persons in the target population using RDS data from the BBS and, as a result, relies on the inherent assumption that the social network of the population of interest is fully connected. It also assumes that the probability of recruitment for an individual is proportional to their network size, and that the network size is known accurately. SS-PSE methods use the self-reported individual network size question asked in the survey, combined with a prior estimate for the population size, and a Bayesian model to estimate the size of the target population [16]. SS-PSE with imputed visibility is an extension that introduces a measurement error model for network size, which helps account for misreporting and bias in self-reported network size [17].

Bayesian Consensus Estimation

For published reports, technical and stakeholder consultation guided the statistical estimation of a single consensus estimate, which combined estimates from the separate PSE methods.

Consultation took place in 3 parts. First, the project technical staff reviewed and compiled prior information, then computed the PSE using each method. The results were discussed, and an initial decision was made to include only the SS-PSE with imputed visibility and 3S-CRC estimates, based on data quality and plausibility. To elicit feedback from the wider project and key population interest groups, a meeting was convened with the Zambia Key Populations Consortium, a network of key population members. Priors and confidence in each method were finalized for input into the Bayesian consensus estimates calculation (using the "Triangulator" app, further described below).

XSL•FO
RenderX

Statistical Analysis

Data analysis for counts, proportions, and PSEs was conducted using R (version 4.0.5; R Foundation for Statistical Computing). PSEs and credible intervals (CRIs) for 2-source capture-recapture (2S-CRC) and 3S-CRC were derived using a Bayesian nonparametric latent-class model in the *shinyrecap* package [18]. Size estimation was based on the number of participants overlapping (or recaptured) across capture events. The Bayesian model assumes that there are several unobserved subgroups of the population which each have different capture probabilities for each capture event. Using RDS data, bivariate analysis was used to compare sociodemographic factors and drug injection practices by capture history to understand potential differences between those captured through 2S-CRC or the BBS alone.

SS-PSE was computed with the Gile estimator in the *sspse* package (version 0.6) using RDS recruitment histories and self-reported network sizes [19]; measurement errors were estimated with and without imputed visibility. There was very limited information on people who inject drugs in the study cities before our study, so prior populations for SS-PSE were based primarily on the percentage of the population who were estimated to be injection drug users based on a review of previous studies in sub-Saharan Africa [20]. Other evaluations of the people who inject drugs population in Zambia suggested that almost all people who inject drugs were male, so census estimates of the male aged 15 - 64 years population for each city were used to compute priors based on an estimate of 0.6% of this population being people who inject drugs [7]. This resulted in initial estimates of 5000 for Lusaka, 1200 for Ndola, and 400 for Livingstone. The Livingstone prior estimate was increased to 800, because the study contacted more than 800 distinct people who inject drugs during the unique object distribution and the RDS recruitment. To account for the large prior uncertainty while constraining estimates to be plausible proportions of the general population in each city, we set the lower quartile for the prior at half of these estimates, and the upper quartile as twice these estimates. The prior estimates were reviewed and agreed with key population stakeholders and experts from the collaborating Zambia organizations before being input into the SS-PSE calculation.

Results from the RDS survey were produced using R (version 4.0.5, *RDS* package, version 0.9 - 3) with bootstrap variance estimation of 95% CIs [21]. Using Taylor series variance estimation, SAS (version 9.4; SAS Institute Inc) was used for indicator validation and sensitivity analyses. Complete case analysis was used when data were missing, as missing data were minimal.

Final statistical computation of the consensus estimate was carried out using a Bayesian approach, which weights the available estimates and their variance by a "design confidence" weight based on judgment of the methodological reliability of each estimate [22]. Design confidence scores of 75 and 65 were assigned to the 3S-CRC estimates and SS-PSE with imputed visibility estimates, respectively. These confidence factors were set conservatively in collaboration with this study's technical group to account for potentially unmet assumptions and to recognize the substantial divergence of the estimates from different methods from each other and our original priors. Priors for the consensus estimate were set for the lower and upper limits of the people who inject drugs population in each city. Lower bounds used the number of unique individuals captured in each city, and upper bounds were set at approximately 10% of the total city population to avoid unwanted downward influence on the estimates.

Ethical Considerations

This activity was reviewed and approved by the United States Centers for Disease Control and Prevention, Columbia University Medical Center Institutional Review Board, Tropical Diseases Research Centre Ethics Review Committee, and the Zambia Ministry of Health National Health Research Authority. Permission and administrative approval from the Zambia Ministry of Health and the National Health Research Authority were obtained before data collection. Letters of support from the Ministry of Home Affairs (Police) and the Drug Enforcement Commission were obtained to ensure that researchers and people who inject drugs were not prosecuted during the survey's period. Consent was neither required nor sought from people who inject drugs in captures 1 or 2, as aggregate data collected were not classified as human subjects research. To reduce the potential risk of breach of confidentiality, BBS participants provided verbal informed consent. BBS data collection sites were selected to ensure privacy protection, and all survey staff were trained in Good Clinical Practices and signed a confidentiality agreement. Participants received a maximum of K430 (approximately US $27) in cash for compensation of transportation costs, time, and peer referrals in the BBS. Nonidentifying codes were used to link participant questionnaire and biomarker data. Data included in this analysis were deidentified.

Results

2S-CRC: Captures 1 and 2

In Livingstone, 426 and 432 individuals were approached in captures 1 and 2, respectively. In capture 1, a total of 415 (97%) individuals were eligible, and of those, 20 had previously received an object and been sampled in the capture (Table 1). Of those who had not previously been captured and were offered an object, 97% accepted, resulting in a total of 383 people who inject drugs sampled in capture 1. In capture 2, a total of 421 (97%) met study eligibility criteria, and of those, 7 had previously received an object and been sampled in capture 2. Overall, 411 people who inject drugs accepted an object and were sampled in capture 2. Of those sampled in capture 2, a total of 99 (24%) self-reported to having received an object in the first capture and been sampled in capture 1.

XSL•FO
RenderX

Table . Eligibility and object acceptance among people who inject drugs approached across 3-source capture-recapture methods by city, Zambia, 2021.

People who inject drugs approached	City								
	Livingstone			Lusaka			Ndola		
	Capture 1 (n=426), n (%)	Capture 2 (n=432), n (%)	Capture 3 (n=249), n (%)	Capture 1 (n=700), n (%)	Capture 2 (n=763), n (%)	Capture 3 (n=479), n (%)	Capture 1 (n=449), n (%)	Capture 2 (n=473), n (%)	Capture 3 (n=303), n (%)
Self-reported to be people who inject drugs[a]									
Yes	415 (97)	421 (97)	235 (94)	617 (88)	697 (91)	349 (73)	442 (98)	450 (95)	259 (85)
No	11 (3)	11 (3)	14 (6)	83 (12)	66 (9)	130 (27)	7 (2)	23 (5)	44 (15)
Already in capture[b]									
Yes	20 (5)	7 (2)	0 (0)	29 (5)	73 (10)	0	1 (0)	3 (1)	0 (0)
No	395 (95)	414 (98)	235 (100)	588 (95)	624 (90)	349 (100)	441 (100)	447 (99)	259 (100)
Object acceptance[c]									
Accepted	383 (97)	411 (99)	N/A[d]	548 (93)	597 (96)	N/A	428 (97)	432 (93)	N/A
Refused	12 (3)	3 (1)	N/A	40 (7)	27 (4)	N/A	13 (3)	15 (3)	N/A

[a]For capture 3, there were additional eligibility criteria, as noted above in the Methods section.

[b]Among those who self-reported to be people who inject drugs.

[c]Among those not captured in the current capture.

[d]N/A: not applicable.

In Lusaka, 700 and 763 individuals were approached in captures 1 and 2, respectively. In capture 1, a total of 617 (88%) individuals were eligible, and of those, 29 had previously received an object and been sampled in the capture. Of those who had not previously been captured and were offered an object, 93% accepted, resulting in a total of 548 people who inject drugs sampled in capture 1. In capture 2, a total of 697 (91%) met study eligibility criteria, and of those, 73 had previously received an object and been sampled in capture 2. Overall, 597 people who inject drugs accepted an object and were sampled in capture 2. Of those sampled in capture 2, a total of 209 (35%) self-reported to having received an object in the first capture and been sampled in capture 1.

In Ndola, 449 and 473 individuals were approached in captures 1 and 2, respectively. In capture 1, a total of 442 (98%) individuals were eligible, and of those, 1 had previously received an object and had been sampled in the same capture. Of those who had not previously been captured and were offered an object, 97% accepted, resulting in a total of 428 people who inject drugs sampled in capture 1. In capture 2, a total of 450 (95%) met study eligibility criteria, and of those, 3 had previously received an object and been sampled in capture 2. Overall, 432 people who inject drugs accepted an object and were sampled in capture 2. Of those sampled in capture 2, a total of 44 (10%) self-reported having received an object in the first capture and been sampled in capture 1.

PSEs generated using 2S-CRC were highest in Ndola (4030, 95% CRI 960 - 5480), followed by Lusaka (1510, 95% CRI 1030 - 2070) and Livingstone (1550, 95% CRI 870 - 1930; Table 2).

XSL•FO
RenderX

Table . Population size estimates of people who inject drugs by city and estimation method, Zambia, 2021.

PSE[a] method[e]	City		
	Livingstone PSE (95% CI/CRI[b])	Lusaka PSE (95% CI/CRI)	Ndola PSE (95% CI/CRI)
Capture-recapture			
2-source capture-recapture	1550 (870-1930)	1510 (1030-2070)	4030 (960-5480)
3-source capture-recapture	2620 (1510-4680)	2320 (1690-10,740)	3350 (2820-3930)
SS-PSE[c,d]			
SS-PSE	450 (340 - 810)	4350 (1410-18,890)	2010 (690-7400)
SS-PSE with imputed visibility	360 (290 - 530)	2040 (660-30,410)	760 (390 - 3060)

[a]PSE: population size estimation.

[b]CI: confidence interval; CRI: credible interval. Capture-recapture methods report CRIs while SS-PSE report CIs.

[c]SS-PSE: successive sampling population size estimation.

[d]SS-PSE represents medians.

[e] Estimates and CIs/CRIs are rounded to the nearest tens.

Across cities, the most common recruitment locations for captures 1 and 2 were bars, taverns, clubs, or shebeens (informal drinking establishments), as well as streets or parks in Lusaka and Livingstone. More than half of the people who inject drugs in Ndola were recruited from bars, taverns, clubs, or shebeens in both captures and recruitment from this location type increased nearly 2-fold between captures 1 and 2 in Lusaka and Livingstone. Across all cities, many people who inject drugs were recruited from "other" locations, which included free-text responses of homes (capture 1: 25%; capture 2: 5%) and markets or shops (capture 1: 11%; capture 2: 8%).

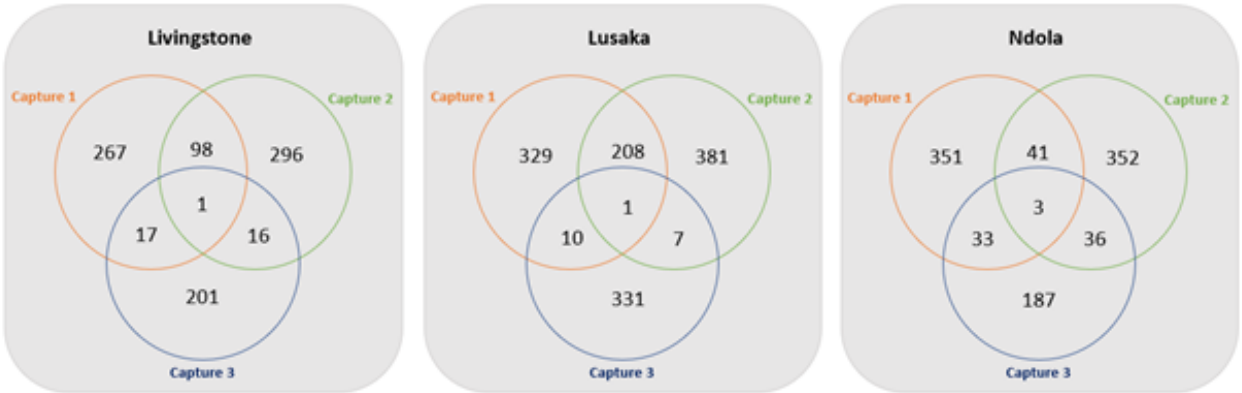
3S-CRC: Capture 3 (BBS)

Lusaka recruited the largest number of purposive seeds (n=12) to participate in the BBS, followed by Ndola (n=8), and Livingstone (n=7). In Livingstone, 249 individuals were screened, 235 were eligible, and 235 were enrolled in the BBS (Table 1). In Lusaka, 479 individuals were screened, 349 were eligible, and 349 were enrolled. In Ndola, 303 individuals were screened, 259 were eligible, and 259 were enrolled. Coupon return rates were highest in Ndola (54%), followed by Lusaka (52%) and Livingstone (48%).

The target population was predominantly male, with an estimated 83% (95% CI 73.9 - 92.9) of people who inject drugs in Livingstone, 96% (95% CI 94.9 - 98.0) of people who inject drugs in Lusaka, and 67% (95% CI 58.2 - 75.6) of people who inject drugs in Ndola assigned male sex at birth. Median age (IQR) was 22 (20-28), 25 (22-29), and 27 (23-32) years for men in Livingstone, Lusaka, and Ndola, respectively, and 29 (23-34), 22 (21-24), and 29 (24-35) years for women in Livingstone, Lusaka, and Ndola, respectively. Most people who inject drugs completed secondary school education or higher (Livingstone: 56%, 95% CI 50.1 - 63.0; Lusaka: 62%, 95% CI 56.7 - 67.4; Ndola: 87%, 95% CI 82.6 - 91.4) and were currently unemployed (Livingstone: 58%, 95% CI 51.4 - 66.0; Lusaka: 78%, 95% CI 72.8 - 83.4; Ndola: 45%, 95% CI 26.6 - 36.0). All but 7 of those recruited across all 3 cities were born in Zambia.

Figure 1 provides overlaps or recaptures across all captures by city, which were used as inputs to generate PSEs using 3S-CRC. In all cities, fewer overlaps occurred with 3S-CRC than with 2S-CRC. PSEs generated using 3S-CRC were highest in Ndola (3350, 95% CRI 2820 - 3930), followed by Livingstone (2620, 95% CRI 1510 - 4680) and Lusaka (2320, 95% CRI 1690 - 10,740; Table 2).

Figure 1. Venn diagram of total captures and recaptures of people who inject drugs from 3-source capture-recapture methods by city, Zambia, 2021.



XSL•FO
RenderX

Bivariate analysis demonstrated differences in sociodemographic factors and drug injection practices among people who inject drugs by capture history (Table 3). In Livingstone, people who inject drugs participating in the BBS who were not captured via location sampling (2S-CRC) had a higher prevalence of homelessness, levels of education completed, and receipt of drug treatment, and differences in the frequency of injection in the last 6 months and of ever "blue toothing"—injecting oneself with a friend's blood to get high—compared to BBS participants who were previously captured. In Lusaka, people who inject drugs participating in the BBS and who were not captured via location sampling had a higher prevalence of living in a home compared to other housing types, more frequent injection in the last 6 months, never using a shared needle, and had differences in levels of education completed, and time of last injection compared to BBS participants who were previously captured. In Ndola, people who inject drugs participating in the BBS who were not captured via location sampling had a higher prevalence of recent injection and never using a shared needle and a lower prevalence of ever speaking to a peer educator or outreach worker about HIV, levels of education completed, and ever blue toothing, and had differences in housing type compared to BBS participants who were previously captured.

XSL•FO
RenderX

Table . Bivariate analysis of sociodemographic factors and drug injection behaviors among people who inject drugs in the respondent-driven sampling biobehavioral survey by capture history and city, Zambia, 2021.

	City								
	Livingstone			Lusaka			Ndola		
	Previously captured[a] (n=34), % (95% CI[b])	Not previously captured (n=201), % (95% CI)	*P* value[c]	Previously captured[a] (n=18), % (95% CI)	Not previously captured (n=331), % (95% CI)	*P* value[c]	Previously captured[a] (n=72), % (95% CI)	Not previously captured (n=187), % (95% CI)	*P* value[c]
Age (years), median (IQR)	22 (18 - 29)	23 (20 - 30)	.01	30 (27 - 32)	24 (21 - 28)	.13	25 (22 - 32)	28 (24 - 34)	.11
Sex			.57			.44			.42
Male	78.5 (59.0 - 97.9)	83.1 (74.9 - 91.4)		93.7 (87.6 - 99.8)	96.6 (94.5 - 98.7)		62.1 (49.2 - 76.0)	68.9 (59.1 - 78.6)	
Female	21.5 (2.1 - 41.0)	16.9 (8.6 - 25.1)		6.3 (0.2 - 12.4)	3.4 (1.3 - 5.5)		37.9 (24.0 - 50.8)	31.1 (21.4 - 40.9)	
Housing type			<.001			<.001			<.001
House	100.0 (89.7 - 100.0)	90.6 (76.8 - 100.0)		19.6 (0.0 - 49.3)	49.7 (43.1 - 56.3)		84.2 (70.5 - 97.9)	83.8 (77.8 - 89.7)	
Apartment	0.0 (0.0 - 10.3)	1.3 (0.4 - 2.3)		20.3 (0.0 - 47.6)	25.0 (19.3 - 30.6)		2.4 (0.0 - 5.3)	2.8 (0.2 - 5.5)	
Dormitory	0.0 (0.0 - 10.3)	0.0 (0.0 - 1.8)		26.2 (0.0 - 58.5)	8.7 (5.1 - 12.4)		11.0 (0.0 - 24.4)	13.4 (7.9 - 18.9)	
Community center	0.0 (0.0 - 10.3)	0.0 (0.0 - 1.8)		0.0 (0.0 - 18.5)	1.9 (0.5 - 3.4)		0.0 (0.0 - 5.0)	0.0 (0.0 - 2.0)	
Street or homeless	0.0 (0.0 - 10.3)	3.1 (1.3 - 5.0)		26.9 (0.0 - 54.9)	10.7 (6.5 - 14.9)		1.9 (1.0 - 2.9)	0.0 (0.0 - 2.0)	
Highest level of education completed			<.001			<.001			<.001
None	4.1 (0.0 - 9.6)	7.4 (0.0 - 22.6)		19.1 (0.0 - 43.0)	5.9 (0.1 - 11.8)		0.0 (0.0 - 5.0)	0.0 (0.0 - 2.0)	
Primary	46.4 (25.9 - 66.9)	35.1 (24.2 - 45.9)		24.1 (0.0 - 56.6)	31.8 (20.9 - 42.6)		4.0 (0.0 - 8.9)	15.8 (9.3 - 22.4)	
Secondary	43.0 (24.7 - 61.2)	51.8 (39.0 - 64.6)		56.9 (23.2 - 90.6)	56.3 (44.5 - 68.0)		81.9 (70.9 - 92.8)	74.9 (65.1 - 84.7)	
Tertiary	6.6 (0.0 - 13.5)	4.2 (1.8 - 6.5)		0.0 (0.0 - 18.5)	5.6 (0.0 - 14.2)		10.7 (3.3 - 18.0)	5.5 (0.0 - 13.0)	
Vocational	0.0 (0.0 - 10.3)	1.5 (0.4 - 2.6)		0.0 (0.0 - 18.5)	0.5 (0.0 - 1.1)		3.5 (1.4 - 5.5)	3.8 (0.2 - 7.4)	
Network size[d]			.17			.96			.97
Less than or equal to median	76.3 (65.6 - 86.9)	85.1 (81.8 - 88.3)		74.0 (48.1 - 98.8)	73.8 (69.4 - 78.2)		78.1 (70.2 - 85.5)	78.3 (73.3 - 83.1)	
Greater than median	23.7 (13.1 - 34.4)	14.9 (11.7 - 18.2)		26.0 (1.2 - 51.9)	26.2 (21.8 - 30.6)		21.9 (14.5 - 29.8)	21.7 (16.9 - 26.7)	
Last injected drugs			.27			<.001			.048
Today or yesterday	39.1 (17.9 - 60.3)	50.5 (37.9 - 63.1)		46.6 (6.6 - 86.6)	47.8 (35.9 - 59.7)		6.6 (0.7 - 12.6)	20.1 (14.0 - 26.2)	
In the past week	41.6 (23.6 - 59.6)	40.7 (27.6 - 53.8)		51.2 (8.8 - 93.7)	45.5 (33.5 - 57.6)		57.7 (43.1 - 71.9)	43.9 (35.5 - 52.4)	
In the past month	15.0 (2.4 - 27.6)	8.3 (3.6 - 12.9)		2.2 (0.0 - 6.1)	5.5 (1.6 - 9.4)		21.0 (8.9 - 33.3)	15.9 (10.3 - 21.6)	

XSL•FO
RenderX

	City								
	Livingstone			Lusaka			Ndola		
	Previously captured[a] (n=34), % (95% CI[b])	Not previously captured (n=201), % (95% CI)	*P* value[c]	Previously captured[a] (n=18), % (95% CI)	Not previously captured (n=331), % (95% CI)	*P* value[c]	Previously captured[a] (n=72), % (95% CI)	Not previously captured (n=187), % (95% CI)	*P* value[c]
In the past 3 months	4.3 (0.0 - 9.9)	0.5 (0.0 - 1.4)		0.0 (0.0 - 18.5)	1.2 (0.0 - 6.7)		14.6 (4.2 - 25.2)	20.0 (12.7 - 27.3)	
Frequency of injection in the last 6 months			<.001			<.001			.98
Less than once a month	0.0 (0.0 - 10.3)	1.0 (0.0 - 2.4)		0.0 (0.0 - 18.5)	2.6 (0.0 - 8.2)		15.6 (0.0 - 36.4)	11.8 (5.2 - 18.5)	
1 to 4 times a month	14.4 (4.7 - 24.1)	13.6 (1.6 - 25.7)		30.1 (0.0 - 71.8)	28.9 (18.0 - 39.9)		61.6 (40.4 - 82.7)	61.8 (53.5 - 70.1)	
2 to 7 times a week/once a day	63.0 (46.6 - 79.5)	51.4 (38.9 - 63.9)		58.0 (25.9 - 90.1)	35.9 (23.5 - 48.2)		17.0 (1.6 - 32.5)	19.7 (13.8 - 25.5)	
2 to 3 times a day	18.0 (5.2 - 30.7)	33.3 (24.1 - 42.4)		11.9 (0.0 - 26.9)	31.8 (23.4 - 40.2)		2.9 (0.0 - 6.3)	3.1 (1.3 - 5.1)	
5 or more times a day	4.7 (2.2 - 7.1)	0.7 (0.1 - 1.3)		0.0 (0.0 - 18.5)	0.8 (0.4 - 1.1)		3.0 (0.0 - 8.2)	3.5 (1.1 - 6.0)	
Used heroin in the last 6 months			.38			.99			.24
Yes	90.0 (82.4 - 97.6)	94.5 (90.2 - 98.8)		95.9 (92.3 - 99.5)	95.9 (93.8 - 98.0)		35.4 (20.7 - 49.7)	26.5 (18.1 - 34.8)	
No	10.0 (2.4 - 17.6)	5.5 (1.2 - 9.8)		4.1 (0.5 - 7.7)	4.1 (2.0 - 6.2)		64.6 (50.3 - 79.3)	73.5 (65.2 - 81.9)	
Locations where self-injected drugs were used in the last 6 months[e]									
At my house	27.2 (12.7 - 42.2)	21.2 (15.7 - 26.7)	.51	31.2 (8.0 - 54.9)	17.9 (12.6 - 23.0)	.25	29.2 (17.3 - 41.7)	24.7 (16.4 - 32.9)	.56
Someone's house	12.2 (0.0 - 24.7)	12.2 (8.3 - 16.1)	.99	6.7 (0.0 - 31.6)	17.3 (12.5 - 22.2)	.32	53.6 (39.4 - 67.8)	58.7 (50.2 - 67.2)	.54
Street or park	16.5 (3.9 - 29.5)	34.4 (27.7 - 41.1)	.06	32.6 (0.0 - 65.9)	11.2 (6.5 - 15.7)	.07	26.3 (14.3 - 38.4)	18.6 (11.6 - 25.7)	.25
Latrines or public toilets	0.0 (0.0 - 10.3)	0.8 (0.4 - 1.3)	—[f]	6.8 (0.0 - 30.3)	4.7 (2.1 - 7.2)	.71	0.0 (0.0 - 5.0)	1.3 (0.0 - 7.8)	—
At a bar or club	0.0 (0.0 - 10.3)	0.6 (0.3 - 1.0)	—	0.0 (0.0 - 18.5)	3.2 (0.0 - 7.0)	—	16.7 (6.4 - 26.9)	6.1 (1.9 - 10.2)	.04
At the drug dealer's place	18.4 (5.3 - 31.8)	29.0 (23.0 - 35.1)	.25	49.9 (21.9 - 76.3)	52.5 (46.0 - 59.1)	.90	10.1 (1.2 - 18.8)	16.3 (10.1 - 22.5)	.30
At a shooting gallery	0.0 (0.0 - 10.3)	0.0 (0.0 - 1.8)	—	28.1 (4.5 - 51.4)	22.4 (17.3 - 27.6)	.67	1.4 (0.0 - 3.3)	0.0 (0.0 - 2.0)	—
Other	44.4 (29.6 - 58.9)	39.7 (32.2 - 47.4)	.67	11.2 (0.0 - 25.3)	25.5 (20.1 - 30.9)	.18	5.9 (1.2 - 10.5)	11.3 (6.8 - 15.8)	.20
Frequency of using a new, sterile needle in the last 6 months			.12			<.001			.20
Never	5.0 (1.0 - 9.0)	0.9 (0.5 - 1.2)		0.0 (0.0 - 18.5)	2.8 (0.8 - 4.8)		11.8 (2.3 - 21.4)	8.2 (4.3 - 12.0)	
Rarely	29.4 (13.1 - 45.8)	29.9 (17.0 - 42.8)		11.7 (0.0 - 43.4)	15.6 (11.3 - 19.9)		35.0 (22.3 - 48.2)	27.5 (19.7 - 35.4)	

XSL•FO
RenderX

	City								
	Livingstone			Lusaka			Ndola		
	Previously captured[a] (n=34), % (95% CI[b])	Not previously captured (n=201), % (95% CI)	*P* value[c]	Previously captured[a] (n=18), % (95% CI)	Not previously captured (n=331), % (95% CI)	*P* value[c]	Previously captured[a] (n=72), % (95% CI)	Not previously captured (n=187), % (95% CI)	*P* value[c]
Half of the time	16.1 (1.4 - 30.9)	18.7 (5.8 - 31.7)		11.3 (0.0 - 36.7)	14.3 (10.3 - 18.2)		5.1 (1.3 - 9.1)	10.7 (5.9 - 15.6)	
Most of the time	5.9 (0.0 - 13.7)	19.7 (12.8 - 26.7)		18.6 (3.3 - 33.9)	15.4 (10.7 - 20.2)		25.5 (14.5 - 36.1)	22.6 (15.7 - 29.4)	
Always	43.5 (22.7 - 64.3)	30.8 (19.9 - 41.7)		58.4 (27.1 - 89.7)	51.9 (45.6 - 58.2)		22.6 (11.6 - 33.3)	31.0 (23.3 - 38.6)	
Frequency of using needles that someone else had already injected with in the last 6 months			.29			<.001			<.001
Never	62.1 (43.6 - 80.6)	61.1 (48.4 - 73.7)		69.4 (37.0 - 100.0)	87.0 (81.9 - 92.2)		39.7 (17.9 - 61.6)	56.3 (48.1 - 64.5)	
Rarely	17.3 (0.0 - 34.8)	15.2 (9.8 - 20.7)		8.5 (0.0 - 35.0)	4.5 (2.4 - 6.5)		19.3 (0.0 - 39.0)	14.0 (8.4 - 19.6)	
Half of the time	2.4 (0.0 - 5.2)	13.2 (0.3 - 26.1)		22.1 (0.0 - 48.8)	6.3 (2.3 - 10.3)		12.1 (4.7 - 19.4)	7.6 (3.5 - 11.7)	
Most of the time	12.0 (5.7 - 18.2)	9.1 (4.0 - 14.2)		0.0 (0.0 - 18.5)	1.2 (0.0 - 3.0)		23.0 (8.6 - 37.4)	17.6 (11.6 - 23.7)	
Always	6.2 (0.0 - 13.1)	1.4 (0.0 - 3.5)		0.0 (0.0 - 18.5)	1.0 (0.0 - 2.2)		5.9 (1.8 - 10.0)	4.4 (1.5 - 7.4)	
Ever injected self with a friend's blood to get high (blue toothing)			.97			.56			.003
Yes	21.2 (6.7 - 35.6)	20.9 (14.8 - 27.0)		10.3 (0.0 - 40.5)	5.6 (2.2 - 9.0)		24.5 (12.8 - 36.4)	7.3 (2.8 - 11.8)	
No	78.8 (64.4 - 93.3)	79.1 (73.0 - 85.2)		89.7 (59.5 - 100.0)	94.4 (91.0 - 97.8)		75.5 (63.6 - 87.2)	92.7 (88.2 - 97.2)	
Ever been arrested due to injection drug use			.29			.20			.68
Yes	78.7 (67.5 - 89.9)	88.6 (83.7 - 93.7)		37.8 (9.4 - 66.3)	63.3 (57.2 - 69.3)		85.8 (68.8 - 100.0)	88.4 (83.1 - 93.7)	
No	11.4 (3.8 - 19.0)	3.7 (0.2 - 6.9)		12.6 (0.0 - 35.7)	11.2 (7.4 - 14.9)		5.3 (0.0 - 18.8)	2.5 (0.7 - 4.3)	
Ever talked to a peer educator or outreach worker about HIV			.66			.67			<.001
Yes	64.8 (48.5 - 81.3)	60.3 (52.5 - 67.9)		63.5 (39.1 - 88.6)	57.8 (51.3 - 64.3)		65.2 (52.1 - 78.7)	55.8 (47.7 - 63.8)	
No	35.2 (18.7 - 51.5)	39.7 (32.1 - 47.5)		36.5 (11.4 - 60.9)	42.2 (35.7 - 48.7)		34.8 (21.3 - 47.9)	44.2 (36.2 - 52.3)	
Ever tested for HIV			.18			.98			
Yes	97.2 (92.0 - 100.0)	89.2 (84.1 - 94.3)		79.5 (55.3 - 100.0)	79.7 (74.5 - 84.8)		91.0 (77.5 - 100.0)	85.8 (79.8 - 91.8)	.40
No	2.8 (0.0 - 8.0)	10.8 (5.7 - 15.9)		20.5 (0.0 - 44.7)	20.3 (15.2 - 25.5)		9.0 (0.0 - 22.5)	14.2 (8.2 - 20.2)	
Ever received drug treatment			<.001			.72			.91
Yes	0.0 (0.0 - 10.3)	5.1 (1.7 - 8.5)		7.0 (0.0 - 31.1)	9.7 (6.2 - 13.3)		5.8 (3.0 - 8.7)	6.3 (2.2 - 10.4)	

XSL•FO
RenderX

	City								
	Livingstone			Lusaka			Ndola		
	Previously captured[a] (n=34), % (95% CI[b])	Not previously captured (n=201), % (95% CI)	*P* value[c]	Previously captured[a] (n=18), % (95% CI)	Not previously captured (n=331), % (95% CI)	*P* value[c]	Previously captured[a] (n=72), % (95% CI)	Not previously captured (n=187), % (95% CI)	*P* value[c]
No	100.0 (89.7 - 100.0)	94.9 (91.5 - 98.3)		93.0 (68.9 - 100.0)	90.3 (86.7 - 93.8)		94.2 (91.3 - 97.0)	93.7 (89.6 - 97.8)	

[a]Previously captured represents people who inject drugs who participated in the biobehavioral survey (capture 3) and reported receiving an object in captures 1 or 2.

[b]CI: confidence interval.

[c]*P* values computed using Kruskal test for difference in medians, and a survey-weighted quasi-binomial general linear model for categorical variables.

[d]Median network size was 5, 4.5, and 3 in Livingstone, Lusaka, and Ndola, respectively.

[e]Responses not mutually exclusive.

[f]Indicates that *P* values were not calculated due to insufficient data or no variation between groups.

SS-PSE

The SS-PSE was highest in Lusaka (4350, 95% CI 1410 - 18,890), followed by Ndola (2010, 95% CI 690 - 7400) and Livingstone (450, 95% CI 340 - 810; Table 2). For all cities, SS-PSE was higher without imputed visibility than with it. SS-PSE with imputed visibility was 360 (95% CI 290 - 530) in Livingstone, 2040 (95% CI 660 - 30,410) in Lusaka, and 760 (95% CI 390 - 3060) in Ndola.

Bayesian Consensus PSE

Integrating 3S-CRC and SS-PSE with imputed visibility estimates, the final Bayesian consensus PSE ranged between 0.5% to 1.8% of the adult male population and all below 1% of the total adult population in each city (Table 4). The consensus PSE was highest in Lusaka (3700, 95% CRI 1500 - 7500), followed by Ndola (2200, 95% CRI 1600 - 2900), and Livingstone (1200, 95% CRI 900 - 1900).

Table . Bayesian consensus population size estimates for people who inject drugs by city, Zambia, 2021.

City	Estimate (median)	95% CRI[a]	% of males aged 15 - 64 years district population[b,c]	% of total people aged 15 - 64 years district population[b,c]
Livingstone	1200	900 - 1900	1.8 (1.4 - 2.9)	0.9 (0.7 - 1.5)
Lusaka	3700	1500 - 7500	0.5 (0.2 - 0.9)	0.2 (0.1 - 0.5)
Ndola	2200	1600 - 2900	1.1 (0.8 - 1.5)	0.6 (0.4 - 0.7)

[a]CRI: credible interval.

[b]District population estimates were used because estimates for smaller catchment areas were not available. In addition to the estimated percentage of the total district population, the estimated percentage of males in the district population is provided because site samples have a high proportion of men. Estimates and CRIs are rounded to the nearest 50. Further, 95% CRI indicate the interval where the true population parameter falls with 95% probability, given the observed data.

[c]District population [23].

Discussion

Principal Findings

Our study found higher estimates of people who inject drugs than prior estimates of people who inject drugs in other parts of Zambia [7], yet ours were relatively consistent with estimates for sub-Saharan Africa where prevalence of injection drug use is estimated to be 0.3% (0.1% - 0.6%) of the adult population and 0.5% (0.1% - 1.0%) among adult males [3]. The prior PSE study in Zambia estimated a much lower number than our Bayesian consensus estimates and provided no uncertainty ranges; a total of 907 people who inject drugs across all 3 districts were estimated after applying correction factors for census enumeration, and 846 people who inject drugs were estimated using wisdom of the crowd methods [7]. The methods used in the prior PSE study in Zambia had several limitations. It applied census methods to enumerate the number of people who inject drugs at hotspots and used wisdom of the crowd methods with purposively recruited "gatekeepers" in Solwezi, Kitwe, and Kabwe [7]. Both methods likely underestimated the number of people who inject drugs, given biases in location sampling.

3S-CRC, used in our study, has increasingly been used to estimate the population size of key populations [24-30] and has several strengths. Estimation using 3 or more captures can account for heterogeneity among captures, including subpopulations with different behaviors or social visibility. In our study, there were more recaptures between location-based captures (2S-CRC) than with the RDS survey (3S-CRC), suggesting that the RDS survey was able to reach people who

XSL•FO
RenderX

were less likely to be found in unique object distribution locations. Future research should explore heterogeneity within networks for people who inject drugs and how differences in recruitment methods influence both recruitment probabilities and the resulting PSE.

BBS participants who had been captured in capture 1 or 2 had differences in housing type and level of education compared to BBS participants not previously captured in all cities, though variations of results by city were seen. These 2 groups had differences in the frequency of sharing needles and frequency of injection, time of last injection, experiences with peer educators or outreach workers, and receipt of drug treatment in one or more cities. Notably, heterogeneity in results by city suggest biases in recruitment methods may be more localized. For example, in Ndola, BBS participants not previously captured via location sampling had lower engagement with a peer educator or outreach worker than those who were previously captured, while other cities had no differences in engagement between these groups.

It is important to acknowledge the assumptions of the PSE approaches used in this study. SS-PSE relies on prior knowledge or assumptions about the target population and assumes a fully connected social network among the target population; given this was the first BBS among people who inject drugs in Zambia, inputs used to generate PSEs through SS-PSE relied on information from prior studies in other settings and may not accurately reflect the context for these geographies in Zambia. Similar assumptions about a connected social network apply to the RDS recruitment process, including the validity of the sample as the third capture-recapture group. Capture-recapture methods ideally assume a closed population, independent capture samples, and equal probability of capturing individuals in the target population, though the use of 3 or more captures and appropriate estimators can allow for some relaxation of these assumptions.

Efforts to minimize the violation of these assumptions were taken in this study. To minimize in- or out-migration of people who inject drugs, the first 2 captures occurred over a short duration, and all captures occurred no more than a week apart. People who inject drugs were sampled from a different set of locations in captures 1 and 2 to help ensure independence and seeds for RDS were selected independently from unique object distributors.

Limitations

We were not able to produce separate PSEs for people who inject drugs by sex, as data on sex were not collected in captures 1 or 2. While we had planned to produce PSEs using service multiplier methods, it was not possible due to limited service providers and a lack of people who inject drugs–specific program data in Zambia. We used wisdom of the crowd methods via the RDS survey, where BBS participants were asked to provide their best estimate (or “most accurate” estimate), as well as their estimate of the highest and lowest possible number of people who inject drugs, aged 16 years or older, living in the survey city. Ultimately, given concerns about the quality and weakness of the method, PSEs generated using wisdom of the crowd were omitted from the Bayesian consensus estimate and excluded from this paper. Some activities in this study were implemented during restrictions associated with the COVID-19 epidemic, which may have limited access to subpopulations of people who inject drugs, particularly those who may have been reached via location sampling methods.

Conclusions

Using rigorous methods, this study provides PSEs for people who inject drugs in Zambia and is the first to produce PSEs for major geographies in the country, including Lusaka, Livingstone, and Ndola. We use multiple methods to derive PSEs; present findings, strengths, and limitations of each method; and document our approach to generating Bayesian consensus estimates. Findings from this study highlight heterogeneity among people who inject drugs in Zambia, including that some people who inject drugs may not be accessible via location sampling, which can guide programming and surveillance strategies to reach these groups. Transparency around underlying data and methodological approaches used to generate consensus estimates, as provided here, can support the field and enhance modeling activities, including extrapolation of size estimation. Critically, the consensus PSE generated in this study can inform HIV and HBV/HCV prevention and treatment program planning, including target setting and coverage estimation, and guide future surveillance activities for key populations in the country.

Acknowledgments

We are grateful to the participants and study team and appreciate the important contributions to the study from the Zambia Survey Advisory Group, including representatives from US Centers for Disease Control and Prevention (CDC), the National HIV/AIDS/sexually transmitted infections (STI)/tuberculosis (TB) Council (NAC), Tropical Diseases Research Centre (TDRC), ICAP, and the Zambia Key Populations Consortium. We thank all civil society organizations participating in the Zambia Key Populations Consortium for their critical support in this study's implementation. From TDRC, we specifically acknowledge Kelvin Kapungu, Shepherd Khondowe, and Innocent Bwalya for their support with data collection and site supervision. This research has been supported by the President's Emergency Plan for AIDS Relief (PEPFAR) through the CDC under the terms of cooperative agreement “Supporting sustainable surveillance systems among key populations (KP) and support the Government of Zambia to improve HIV-related services for KP” (CDC cooperative agreement 1NU2GGH002056).

XSL•FO
RenderX

Disclaimer

LP was not at United States Agency for International Development (USAID) when the research for this paper was conducted. The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the funding agencies.

Data Availability

Datasets generated or analyzed during this study are not publicly available due to privacy and ethical restrictions. All data are the property of the Ministry of Health and NAC and are available upon reasonable request.

Authors' Contributions

LP, JJN, BH, LT, L Chilukutu, TN, CM, L Chelu, RH, JM, KM, and ML contributed to designing this study, developing data collection instruments, and implementing study procedures. L Chilukutu, TN, CM, L Chelu, and RH managed data collection. LP and GR led data cleaning and analysis and drafted this paper. JJN, AFM, BH, and LT were sponsor investigators and did not interact with human participants or have access to identifiable data or specimens for research purposes. All authors critically reviewed this paper, approved this final paper, and had final responsibility for the decision to submit for publication.

Conflicts of Interest

None declared.

References

1. The urgency of now: AIDS at a crossroads. UNAIDS. 2024. URL: https://www.unaids.org/sites/default/files/media_asset/2024-unaids-global-aids-update_en.pdf [accessed 2025-05-28]
2. Korenromp EL, Sabin K, Stover J, et al. New HIV infections among key populations and their partners in 2010 and 2022, by world region: a multisources estimation. J Acquir Immune Defic Syndr 2024 Jan 1;95(1S):e34-e45. [doi: 10.1097/QAI.0000000000003340] [Medline: 38180737]
3. Degenhardt L, Webb P, Colledge-Frisby S, et al. Epidemiology of injecting drug use, prevalence of injecting-related harm, and exposure to behavioural and environmental risks among people who inject drugs: a systematic review. Lancet Glob Health 2023 May;11(5):e659-e672. [doi: 10.1016/S2214-109X(23)00057-8] [Medline: 36996857]
4. Findings from a formative assessment among people who inject drugs in selected towns in Zambia, 2020. : ICAP at Columbia University; 2021.
5. Ministry of Legal Affairs, Government of the Republic of Zambia. Chapter 96. In: The Narcotic Drugs and Psychotropic Substances Act: Government of the Republic of Zambia; 1993. URL: https://www.zambialaws.com/consolidated-statutes/principal-legislation/218-chapter-96narcotic-drugs-and-psychotropic-substances-act [accessed 2025-06-25]
6. Formative assessment of HIV risk and size estimation using census and enumeration methods among key populations most at risk of HIV in Zambia. : Lusaka: Population Council, National HIV/AIDS/STI/TB Council, Tropical Diseases Research Centre, and University of California, San Francisco; 2016.
7. Size estimation of people who inject drugs (PWID) and service mapping in Kabwe, Kitwe and Solwezi. : MAINline; 2021.
8. Mweemba C, Hangoma P, Fwemba I, Mutale W, Masiye F. Estimating district HIV prevalence in Zambia using small-area estimation methods (SAE). Popul Health Metr 2022 Feb 19;20(1):8. [doi: 10.1186/s12963-022-00286-3] [Medline: 35183216]
9. Zambia country fact sheet 2023. UNAIDS. 2023. URL: https://www.unaids.org/en/regionscountries/countries/zambia [accessed 2025-05-28]
10. Hook EB, Regal RR. Capture-recapture methods in epidemiology: methods and limitations. Epidemiol Rev 1995;17(2):243-264. [doi: 10.1093/oxfordjournals.epirev.a036192] [Medline: 8654510]
11. Hook EB, Regal RR. Recommendations for presentation and evaluation of capture-recapture estimates in epidemiology. J Clin Epidemiol 1999 Oct;52(10):917-926. [doi: 10.1016/s0895-4356(99)00060-8] [Medline: 10513754]
12. Nanan DJ, White F. Capture-recapture: reconnaissance of a demographic technique in epidemiology. Chronic Dis Can 1997;18(4):144-148. [Medline: 9445362]
13. WHO, CDC, UNAIDS, FHI360. Biobehavioral survey guidelines for populations at risk for HIV. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/978-92-4-151301-2 [accessed 2025-05-28]
14. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. Soc Probl 1997 May;44(2):174-199. [doi: 10.2307/3096941]
15. Handcock MS, Gile KJ, Mar CM. Estimating hidden population size using respondent-driven sampling data. Electron J Stat 2014;8(1):1491-1521. [doi: 10.1214/14-EJS923] [Medline: 26180577]
16. Wesson PD, Handcock MS, McFarland W, Raymond HF. If you are not counted, you don’t count: estimating the number of African-American men who have sex with men in San Francisco. Int J Epidemiol 2015 Oct;44(suppl_1):i39-i39. [doi: 10.1093/ije/dyv097.132]
17. McLaughlin KR, Handcock MS, Johnston LG, et al. Inference for the visibility distribution for respondent-driven sampling. Presented at: Joint Statistical Meetings; Aug 8-13, 2015; Seattle, WA p. 2259-2267.

XSL•FO
RenderX

18. Fellows IE. Shinyrecap: shiny user interface for multiple source capture recapture models. GitHub. URL: https://fellstat.github.io/shinyrecap/ [accessed 2025-05-28]
19. Handcock MS, Gile KJ, Kim B, McLaughlin KR. Package 'sspse': estimating hidden population size using respondent driven sampling data. The Comprehensive R Archive Network. URL: https://cran.r-project.org/web/packages/sspse/sspse.pdf [accessed 2025-05-28]
20. Degenhardt L, Peacock A, Colledge S, et al. Global prevalence of injecting drug use and sociodemographic characteristics and prevalence of HIV, HBV, and HCV in people who inject drugs: a multistage systematic review. Lancet Glob Health 2017 Dec;5(12):e1192-e1207. [doi: 10.1016/S2214-109X(17)30375-3] [Medline: 29074409]
21. Handcock MS, Gile KJ, Fellows IEF, Neely WW. RDS: respondent-driven sampling. 2012. URL: https://CRAN.R-project.org/package=RDS [accessed 2025-05-28]
22. Fellows IE. Derive consensus estimates of a population quantity from multiple empirical estimates using a hierarchical bayesian model. GitHub. URL: https://github.com/fellstat/triangulator
23. Adjusted population projections for HIV-related program planning based on 2010 Zambia Census. : ZamStat; 2021.
24. Doshi RH, Apodaca K, Ogwal M, et al. Estimating the size of key populations in Kampala, Uganda: 3-source capture-recapture study. JMIR Public Health Surveill 2019 Aug 12;5(3):e12118. [doi: 10.2196/12118] [Medline: 31407673]
25. Guure C, Dery S, Afagbedzi S, et al. National and subnational size estimation of female sex workers in Ghana 2020: comparing 3-source capture-recapture with other approaches. PLoS One 2021;16(9):e0256949. [doi: 10.1371/journal.pone.0256949] [Medline: 34550978]
26. McIntyre AF, Mitchell A, Stafford KA, et al. Key population size estimation to guide HIV epidemic responses in Nigeria: bayesian analysis of 3-source capture-recapture data. JMIR Public Health Surveill 2022 Oct 26;8(10):e34555. [doi: 10.2196/34555] [Medline: 36287587]
27. Musengimana G, Tuyishime E, Remera E, et al. Female sex workers population size estimation in Rwanda using a three-source capture-recapture method. Epidemiol Infect 2021 Mar 18;149:e84. [doi: 10.1017/S0950268821000595] [Medline: 33734058]
28. Tuyishime E, Kayitesi C, Musengimana G, et al. Population size estimation of men who have sex with men in Rwanda: three-source capture-recapture method. JMIR Public Health Surveill 2023 Mar 27;9:e43114. [doi: 10.2196/43114] [Medline: 36972131]
29. Nguyen LT, Patel S, Nguyen NT, et al. Population size estimation of female sex workers in Hai Phong, Vietnam: use of three source capture-recapture method. J Epidemiol Glob Health 2021 Jun;11(2):194-199. [doi: 10.2991/jegh.k.210312.001] [Medline: 33876600]
30. Leclerc P, Vandal AC, Fall A, et al. Estimating the size of the population of persons who inject drugs in the island of Montréal, Canada, using a six-source capture-recapture model. Drug Alcohol Depend 2014 Sep 1;142:174-180. [doi: 10.1016/j.drugalcdep.2014.06.022] [Medline: 25008106]

Abbreviations

2S-CRC: 2-source capture-recapture
3S-CRC: 3-source capture-recapture
BBS: biobehavioral survey
CI: confidence interval
CRI: credible interval
HBV: hepatitis B virus
HCV: hepatitis C virus
PSE: population size estimate
RDS: respondent-driven sampling
SS-PSE: successive sampling population size estimation

Edited by A Mavragani; submitted 18.09.24; peer-reviewed by S Gutreuter, W Hladik; revised version received 02.04.25; accepted 04.04.25; published 30.07.25.

Please cite as:
Parmley L, Reid G, Neal JJ, Hanunka B, Tally L, Chilukutu L, Nkumbula T, Mulemfwe C, Chelu L, Handema R, Mwale J, Mutale K, Mulenga L, McIntyre AF, Philip NM, Chung H, Lahuerta M
Estimating the Population Size of People Who Inject Drugs in 3 Cities in Zambia: Capture-Recapture, Successive Sampling, and Bayesian Consensus Estimation Methods
JMIR Public Health Surveill 2025;11:e66551
URL: https://publichealth.jmir.org/2025/1/e66551
doi:10.2196/66551

XSL•FO
RenderX

© Lauren Parmley, Giles Reid, Joyce J Neal, Brave Hanunka, Leigh Tally, Lophina Chilukutu, Tepa Nkumbula, Chipili Mulemfwe, Lazarous Chelu, Ray Handema, John Mwale, Kennedy Mutale, Lloyd Mulenga, Anne F McIntyre, Neena M. Philip, Hannah Chung, Maria Lahuerta. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 30.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Small-Area Lung Cancer Incidence and Mortality: Cross-Sectional Population-Based Study Using Hospital Discharge and Death Registration Data

Yu He[1*], PhD; Xinxin Xia[2*], MPH; Qing Wang[3*], PhD; Yaoyun Zhang[2,4], MBBS; Ying Meng[2,4], MA; Xiaokang Ji[3], MSc; Qingbo Zhao[3], BM; Yongchao Wang[3], BM; Yifu Zhao[3], BM; Fuzhong Xue[3], PhD; Jin Xu[2], PhD

[1]Department of Journal Management, Chinese Preventive Medicine Association, Beijing, China

[2]China Center for Health Development Studies, Peking University, 38 Xueyuan Road, Haidian District, Beijing, China

[3]School of Public Health, Shandong University, Jinan, China

[4]School of Public Health, Peking University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Jin Xu, PhD
China Center for Health Development Studies, Peking University, 38 Xueyuan Road, Haidian District, Beijing, China

Abstract

Background: Despite rapid development, cancer registries in low- and middle-income countries, such as China, have the persistent problems with up to 6-year delay and a lack of reported details about small areas.

Objective: This study aimed to develop an approach to provide more up-to-date localized cancer surveillance using linked administrative data. We used lung cancer as an example.

Methods: Based on data of hospitalization record front pages (HRFPs) between 2013 and 2022 from all the secondary and tertiary hospitals in Shandong Province, China, we identified incident cases of lung cancer in 2022 with 2013‐2021 being the washout period. Deaths from lung cancer were ascertained for 2022 using linked HRFPs and death registration data. We estimated age-standardized incidence and mortality rates (ASIR and ASMR) of lung cancer in 2022 using Segi world standard population, age-specific incidence and mortality rates by sex, and county-level ASIR and ASMR to illustrate regional disparity. We grouped the counties by municipalities and calculated the Theil indices for within-municipality inequality and between-municipality inequality.

Results: The HRFPs captured 79,672 incident cases of lung cancer in Shandong in 2022 (45,527 males, 34,145 females). The ASIR of lung cancer in Shandong was 42.46 per 100,000 in both sexes (49.19/100,000 in males vs 36.67/100,000 in females). A total of 40,626 lung cancer-specific deaths were ascertained (28,185 men and 12,441 women). The ASMR was 19.76/100,000 in both sexes, 26.29/100,000 and 11.38/100,000 in males and females, respectively. The IQR of county-level ASIR and ASMR were 17.13/100,000 and 10.41/100,000, respectively. The inequality was primarily due to within-municipality disparities, with within-municipality Theil T indices higher than between-municipality Theil T indices (0.0572 vs 0.0033 for ASIR, 0.0824 vs 0.0011 for ASMR).

Conclusions: The cancer surveillance approach based on linked administrative data could provide up-to-date small-area estimates of cancer burden, when cancer registry data are not yet reported and for areas not covered by cancer registries. It could reveal disparity of cancer epidemiology, which provides leads for further investigation into the underlying causes and potential solutions for equity improvement.

(JMIR Public Health Surveill 2025;11:e74062) doi:10.2196/74062

KEYWORDS

cancer incidence; cancer mortality; small-area analysis; administrative health data; epidemiology

Introduction

With an estimated 20 million new cancer cases and 9.7 million cancer deaths in 2022 worldwide, cancer ranks as the third contributor to the global burden of disease [1]. The role of cancer registries as a key source of information for surveillance of cancer burden to reveal the need for global cancer care has been well recognized [2]. At the national and subnational level, timely and locally accurate information about the incidence and mortality of cancer enables the identification of high-risk areas, monitoring of the effectiveness of health interventions, and development of the targeted solutions to address a broad range

XSL•FO
RenderX

of determinants of health, potentially diverse across areas [3,4]. Despite the rapid development of cancer registries in low- and middle-income countries like China, challenges remain regarding delayed reporting [5,6], issues with data accessibility [7-9], and population coverage [10].

With the growing availability of high-quality data based on hospitalization records and other forms of administrative data [11], it is both important and feasible to attempt alternative or complementary approaches to cancer surveillance beyond cancer registries. Indeed, medical insurance claims data have been demonstrated as a reliable source for cancer surveillance [12,13]. Similarly, researchers have also proposed electronic health records, hospital record data in particular, for cancer surveillance as a supplement to or in the absence of a cancer registry [14,15]. Studies from Canada, Spain, and Australia examined the performance of incident hospitalization record in estimating cancer incidence against cancer registries and found high concordance between the two, especially for particular cancer types such as lung cancer and colorectal cancer [16-18]. With the increasing availability of administrative health care data, their use may offer a more detailed and timely surveillance of cancer epidemiology that can inform more effective resource allocation and policy making.

We selected lung cancer as the example in this study due to the heavy burden of this disease. Among all the types of cancer, it has long been the leading cause of both new cancer cases (age-standardized incidence rate [ASIR] in 2022: 40.78/10,000) and cancer deaths (age-standardized mortality rate [ASMR] in 2022: 26.66/10,000) in China [6]. In Shandong Province in particular, lung cancer has been the top cause of cancer death since 2004 [19].

In this study, we sought to develop a timely approach to estimating small-area lung cancer incidence and mortality rates for one of the most populous provinces in China and its counties by linking population-based hospital discharge records and death registration. A small area, depending on the purpose and data availability of the study, may be a neighborhood, a census tract, or a county [20]. In this study, the geographic scale of a small area refers to a county. We also mapped the county-level lung cancer incidence and mortality rates to evaluate the degree of intercounty disparity of lung cancer incidence and mortality.

Methods

Study Setting

Shandong Province is located on the eastern coast of northern China, covering an area of 157,100 km^2. It had a population of 101.6 million people, or about 7.2% of China's total population in 2022 [21]. The province's gross domestic product per capita in the same year was 86,143 Chinese yuan (or US$12,800), equivalent to the level of an upper-middle–income economy by the World Bank categorization [22-24]. Qualified cancer registry data in Shandong covered approximately 33.26 million registered residents, accounting for about 33.24% of the province's total population [25]. Within the province, there are 17 municipalities which in turn consist of 137 counties (including districts and county-level cities). The eastern coastal areas in Shandong are economically more developed, especially Qingdao Municipality and its neighboring municipalities (Weihai, Yantai, Weifang, and Rizhao), compared with other areas in Shandong [26].

Data Sources

This study used the hospitalization record front pages (HRFPs) data from all secondary and tertiary hospitals in Shandong and the death registry covering the whole population in the province. The HRFP covered a wide range of information, including patient demographics, residence, health-related, and hospitalization-related information, among others. Both datasets were stored in the Cheeloo Lifespan Electronic Health Research Data-library (Cheeloo LEAD), which is an electronic health data platform linked via encrypted individual identity numbers [27] . We extracted the following data from HRFP from the Cheeloo LEAD between January 1, 2013, and December 31, 2022: sex, year of birth, date of admission, county of residence, and diagnosis upon admission recorded in *International Classification of Diseases, Tenth Revision (ICD-10)* code [28]. We further extracted records for people who died between January 1, 2022, and December 31, 2022, from the death registers, covering sex, year of birth, county of residence, and cause of death (also based on *ICD-10* code). These data were complemented by county-level and provincial population data from the latest census reported by Shandong Provincial Bureau of Statistics [29].

This article followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline [30] (see Checklist 1).

Identification of Incident Lung Cancer Cases and Deaths

We identified lung cancer–related hospitalization using *ICD-10* codes from C34.0 to C34.9. To avoid classifying prevalent cases as incident cases, we excluded cancer patients with hospitalization records due to lung cancer in previous years (ie, 2013 - 2021). In other words, we adopted a washout time window from 1 year up to 9 years [2]. While a longer washout period appeared to generate more precise results, the benefit of an extra washout year beyond 1 year appeared nonsubstantial (see Figure S1 in Multimedia Appendix 1). For patients with records of multiple hospitalizations for lung cancer, an incident lung cancer case was defined based on the first definite lung cancer–related hospitalization.

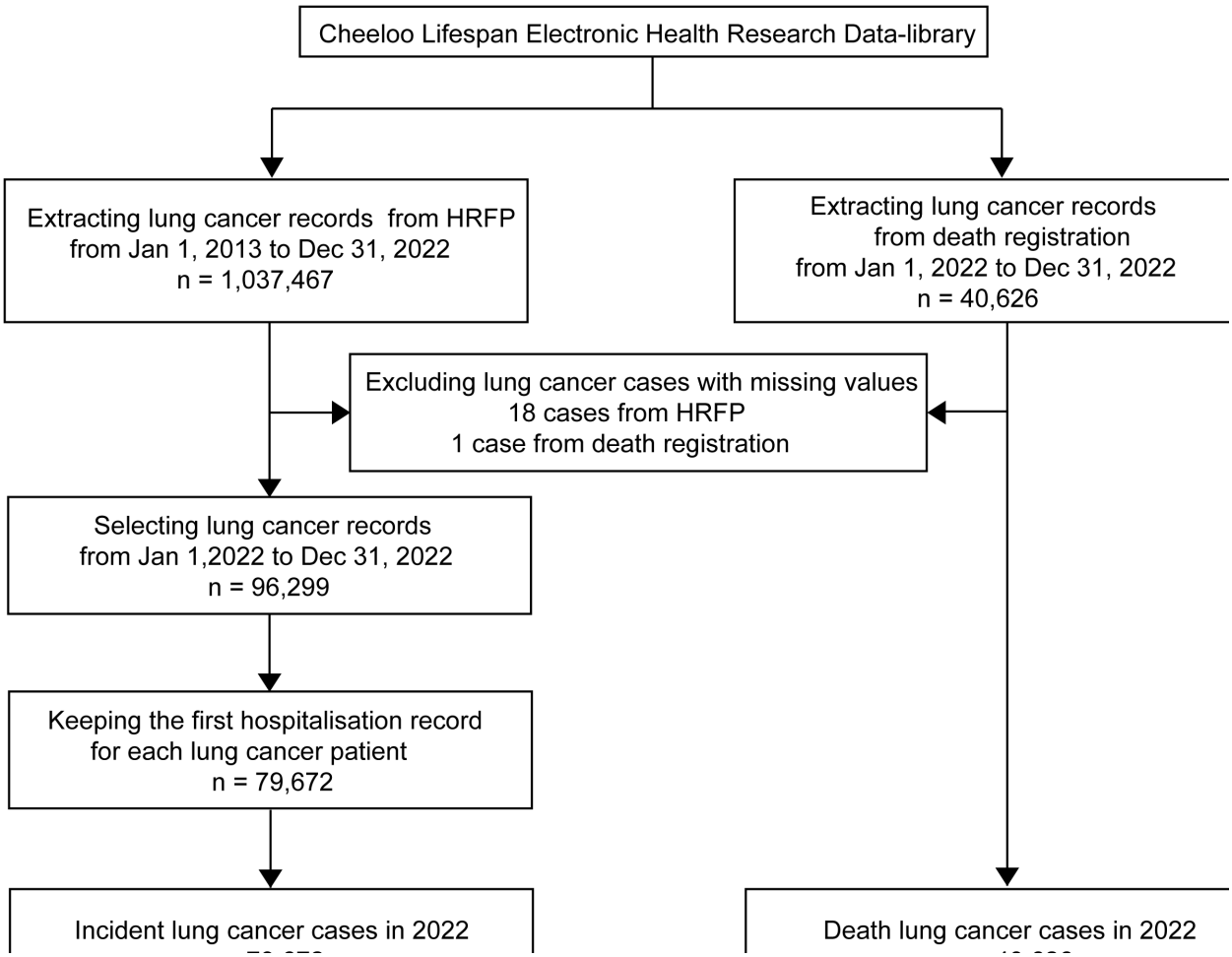
We used the corresponding date of admission as the date of diagnosis [31]. As the number of cases with missing information for each of the variables listed in the previous section was below 10, we removed the cases with missing values (a total of 18 cases were excluded).

We ascertained lung cancer–specific death cases based on lung cancer–related records in the death registers from January 1, 2022, to December 31, 2022, with the corresponding registration date as the date of death. *ICD-10* codes were used to determine the primary cause of death. We excluded 1 death case with missing values.

XSL•FO
RenderX

The procedure of identifying incident lung cancer cases and deaths from lung cancer in Shandong in 2022 is shown in Figure 1.

Figure 1. Identification of incident and death lung cancer cases in Shandong, China, 2022. Hospitalization record front pages were identified as lung cancer with diagnoses in *International Classification of Diseases, Tenth Revision* codes from *C34.0* to *C34.9*. Records from death registration were identified as lung cancer-related death with cause of death in *ICD-10* codes from *C34*.0 to *C34.9*.



Statistical Analysis

We evaluated crude, age-specific, and ASIR of lung cancer in Shandong in 2022 by gender with the corresponding resident population as the denominator (structure of the resident population in Shandong in 2022 by gender is shown in Figure S2 in Multimedia Appendix 1), using the formulas below.

Crude incidence rate=New lung cancer cases in the year of interestTotal population in the same year×100,000

Age-specific incidence rate=New cases of a disease in a given age groupTotal population in the same age group×100,000

Age-standardised incidence rate=∑(Age-specific incidence rate×Standard population in corresponding age group)∑Standard population×100,000

Age-standardization was based on Segi world standard population. Crude, age-specific, and ASMR were calculated following a similar method.

Lung cancer incidence and mortality were also calculated at the municipal and county levels. We calculated the difference between the third quartile (Q3) and the first quartile (Q1) as the IQR to quantify geographical disparity in age-standardized incidence and mortality rates of lung cancer at the county and municipality levels. In addition, we used the Theil index to estimate the overall inequality in these rates [32,33], because it could be disaggregated. We grouped the counties based on municipalities where they were located and calculated the Theil indices (Theil T and Theil L) for these groupings, estimating both within-municipality inequality (disparities between counties within each municipality) and between-municipality inequality (disparities between municipalities). A higher Theil index indicates greater inequality [34]. Details about calculating Theil T and Theil L can be found in xMethods Part S2 in Multimedia Appendix 1.

We also evaluated the internal validity of this cancer surveillance method by examining the stability of incidence rates over the years from 2019 to 2022, referring to a criterion on data quality by the International Agency for Research on Cancer [35]. The same formulas for estimating the ASIR and ASMR of lung

XSL•FO
RenderX

cancer in the year 2022 were adopted for the years 2019 to 2022. Based on the previously reported lung cancer incidence time trends [36], we assumed a threshold of an annual growth rate of 10% as indicating the stability of the data over time. The external validity was assessed by comparing the estimates derived by the method in this study with the cancer statistics reported by the national cancer registry in China for the same time period.

All statistical analyses in this study were performed using R software (version 3.4.1; R Foundation for Statistical Computing), with maps also generated using R software.

Ethical Considerations

This study received ethical approval from the Ethics Committee for Public Health of Shandong University (registration number LL20241105). As the study used preexisting, deidentified secondary data, the Ethics Committee waived the requirement for informed consent. Before analysis, all personally identifiable information was removed and replaced with anonymous identifiers. The privacy and confidentiality of all research subjects were strictly maintained. No compensation was provided to the participants.

Results

Number of Incident Cases and Incidence Rate of Lung Cancer at the Provincial Level

In 2022, there were an estimated 79,672 incident cases of lung cancer in Shandong, including 45,527 males and 34,145 females. The estimated crude incidence of lung cancer in 2022 was 78.40 per 100,000 people. Crude lung cancer incidence rate in males (88.35 per 100,000) was higher than in females (68.15 per 100,000). The ASIR of lung cancer using Segi population was 42.46 per 100,000. The ASIR was 49.19 per 100,000 in males and 36.67 per 100,000 in females, resulting in a male-to-female ratio of 1.34 (Figure 2A).

The incidence of lung cancer rose rapidly after age 50 years, peaked around 75, and gradually dropped in people older than 75 in both men and women. Only in age groups between 35 and 50 was the incidence of lung cancer higher in females than in males.

From 2019 to 2022, the annual growth rate in the lung cancer ASIR in Shandong was within 10%, indicating the reliability of the lung cancer incidence estimates over time (Table S1 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 2. Age- and sex-specific incidence and mortality of lung cancer in Shandong, China, 2022, estimated using hospital discharge and death registration data. (A) Age- and sex-specific incidence rate of lung cancer. (B) Age- and sex-specific mortality rate of lung cancer.

XSL•FO
RenderX

Number of Deaths and Mortality Rate From Lung Cancer at the Provincial Level

A total of 40,626 lung cancer-specific deaths were ascertained in Shandong in 2022, among which 28,185 were men and 12,441 were women. The estimated crude mortality of lung cancer in 2022 was 39.98 per 100,000 people (54.69 per 100,000 in males, and 24.83 per 100,000 in females). The ASMR of lung cancer using the Segi population was 19.76 per 100,000, which was higher in males than in females (26.29 per 100,000 vs 11.38 per 100,000; male-to-female ratio: 2.56).

Age-specific mortality from lung cancer started increasing after age 50 years and peaked in the age group 75 - 80 years in both sexes (Figure 2B). Except for age groups younger than 45 years, in which both men and women experienced a mortality close to zero, the mortality was consistently higher in males compared with females.

Geographical Disparity of Incidence and Mortality at the County Level

The IQR was substantial for county-level ASIR at 17.13 per 100,000 people (Table 1). The Theil index for the ASIR and ASMR of lung cancer was 0.0605 and 0.0835 for Theil T, and 0.0586 and 0.0787 for Theil L, respectively (Table 2). The overall inequality was primarily due to within-municipality disparities, with the within-municipality Theil index accounting for 0.0572 for ASIR and 0.0824 for ASMR in Theil T, and 0.0261 for ASIR and 0.0295 for ASMR in Theil L. The minimal county-level ASIR was observed in Luozhuang of Linyi Municipality (southern Shandong) at 9.50 per 100,000 people, while the maximum was observed in Shinan of Qingdao Municipality (eastern Shandong) at 158.16 per 100,000 people (Figure 3A; see Figure S3A, S4A, and S5A in Multimedia Appendix 1 for maps for crude county-level incidence rate, crude municipal-level incidence rate, and municipal-level ASIR, respectively. Meanwhile, the IQR of municipal-level ASMR was smaller (10.41 per 100,000 people). The minimal county-level ASMR was observed in Penglai of Yantai Municipality (northeastern Shandong) at 1.41 per 100,000 people, while the maximum was observed in Zhanhua of Binzhou Municipality (northern Shandong) at 38.62 per 100,000 (Figure 3B; see Figure S3B, S4B, and S5B in Multimedia Appendix 1 for maps for crude county-level mortality rate, municipal-level mortality rate, and municipal-level ASMR, respectively).

Table . Quartiles and IQRs for age-standardized incidence and mortality rates of lung cancer (per 100,000 people) in Shandong, China, 2022.

	ASIR[a], median (IQR)	ASMR[b], median (IQR)	ASIR, (minimum–maximum)	ASMR, (minimum–maximum)
County-level	38.70 (31.91–49.04)	19.63 (14.71–25.12)	9.50–158.16	1.41–38.62
Municipal-level	41.40 (35.93–49.36)	20.01 (16.86–23.28)	31.01–59.80	13.53–26.17

[a]ASIR: age-standardized incidence rate.

[b]ASMR: age-standardized mortality rate.

Table . Decomposition of intercounty inequality of age-standardized incidence and mortality rates of lung cancer (per 100,000 people) in Shandong, China, 2022.

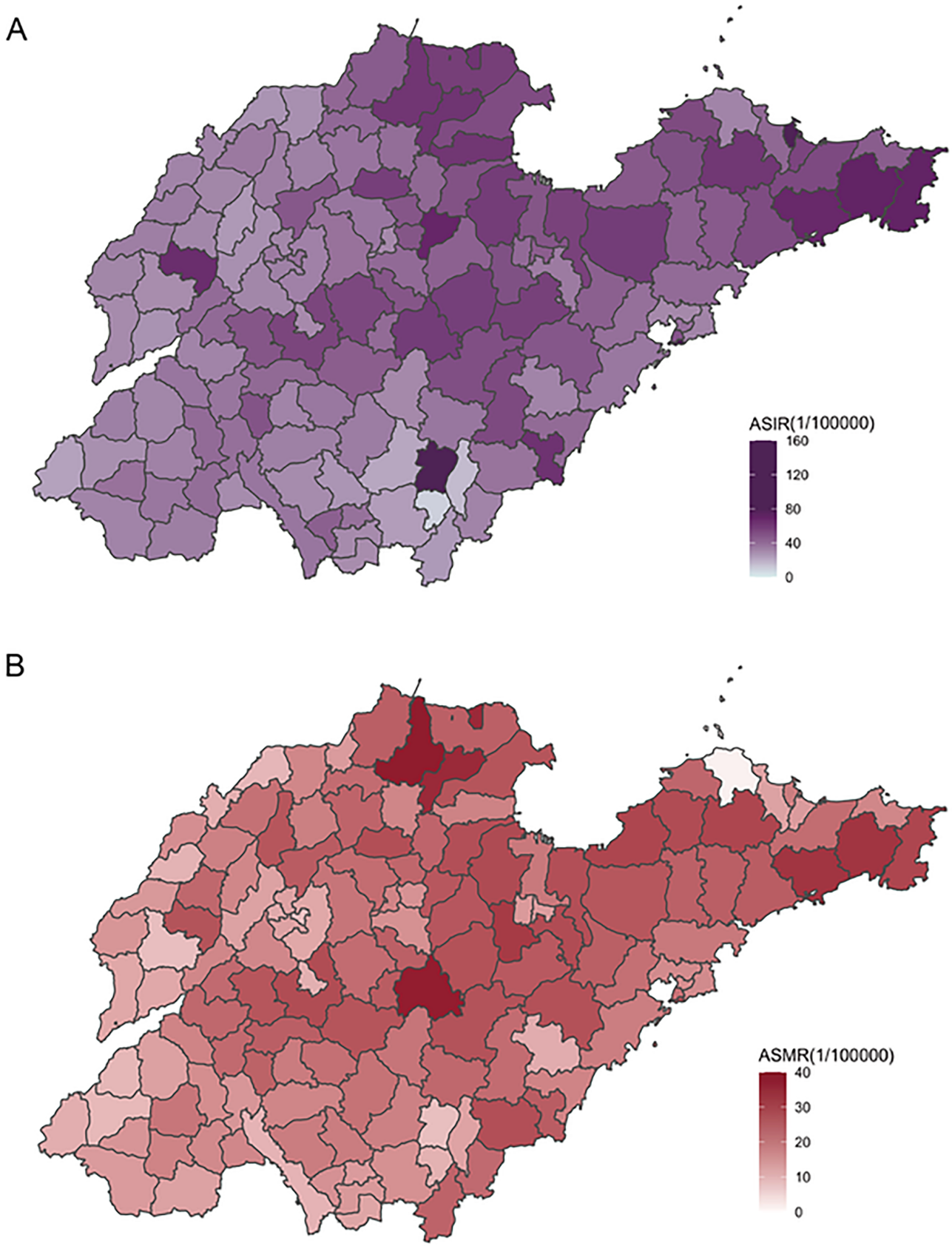
Component	ASIR[a]	ASMR[b]
Overall intercounty inequality		
Theil T	0.0605	0.0835
Theil L	0.0586	0.0787
Between-municipality inequality		
Theil T	0.0033	0.0011
Theil L	0.0325	0.0492
Within-municipality inequality		
Theil T	0.0572	0.0824
Theil L	0.0261	0.0295

[a]ASIR: age-standardized incidence rate.

[b]ASMR: age-standardized mortality rate.

XSL•FO
RenderX

Figure 3. Geographical distribution of (**A**) Age-standardized incidence and (**B**) Mortality rate of lung cancer among county-level divisions in Shandong, China, 2022. ASIR; age-standardized incidence rate; ASMR: age-standardized mortality rate.



XSL•FO
RenderX

Discussion

Principal Findings

This study established a novel approach to cancer surveillance by linking data from hospital discharge records and death registries. Using this approach, we estimated that the age-standardized lung cancer incidence and mortality rates were 42.46 per 100,000 and 19.76 per 100,000, respectively, in Shandong, China, in 2022. The IQR of county-level ASIR was 17.13 per 100,000, while the IQR of county-level ASMR was 10.41 per 100,000. The analyses based on Theil indices suggested that the regional disparity of ASIR and ASMR of lung cancer was primarily due to inequality between counties within each municipality. Hotspots of high ASIR and ASMR counties were also identified. This study demonstrates the potential of using routinely collected administrative hospital records in estimating local cancer incidence and death, offering a reliable approach for up-to-date and small-area cancer surveillance in the future.

For internal validation, we assessed the ASIR and ASMR from 2019 to 2022 and found the rates stable during the period (ASIR: average growth rate 2.242% in males, 6.114% in females; ASMR: average growth rate 2.944% in males, 3.289% in females; see Table S1 in Multimedia Appendix 1). Regarding external validation, we found the lung cancer ASIR and ASMR in Shandong in 2022, we estimated (42.46 per 100,000 and 19.76 per 100,000), comparable to those from the national prediction for the same year by the National Cancer Center based on cancer registry data 2010 - 2018 (40.78 per 100,000 and 26.66 per 100,000, respectively) [6]. Similar to the national level estimates, the lung cancer ASIR and ASMR we found were higher in males than in females, with a male-to-female ASIR ratio of 1.34 (1.71 in the national cancer report) and ASMR ratio of 2.56 (2.69 in the national cancer report) [6]. As mentioned earlier, the estimated annual growth rates from 2019 to 2022 for ASIR using this approach were 10% and thus considered stable [37]. Therefore, our findings are in line with previous studies that demonstrated the comparability of hospital discharge data and cancer registry data [16-18].

The analysis in this study revealed within-municipality disparities in lung cancer estimates. High ASIRs and AMSRs were found in counties to the north and east of Shandong Province (see Figure 3), where the economic development level is relatively higher compared with other parts of the province. This association with regional socio-economic status is similar to that in a previous study on global lung cancer burden, which reported positive correlations between ASIR and ASMR and human development index [38]. A possible explanation might be the regional difference in tobacco smoking prevalence, as a survey in 2016 - 2017 in Shandong found a higher prevalence of smoking in the eastern part compared with the middle area [39]. Air quality and access to diagnostic services might also play a role [40]. Yet there is a lack of evidence on the correlation of those factors with lung cancer incidence and mortality on the geographic scale of small areas in China, which may be explored in the future, for example, in Shandong via linking the data in this study with data sources on the distribution of potential factors such as tobacco smoking exposure, air pollution, etc.

Our approach to rapid and local cancer surveillance by linking two administrative datasets has some advantages that warrant its consideration as a complement to cancer registries. First, it allows the rapid estimates of cancer burden. In this study, we used the records up to 2022 to calculate the incidence and mortality in 2022. It achieved much higher timeliness compared with the 6-year delay of cancer registry data [41]. Compared with medical claims data [2], for which a 6-month delay was expected owing to the lag between discharge and reimbursement, our approach using hospital discharge records could be captured in real-time without delay. Indeed, with more up-to-date data, even more timely estimation can be achieved. Second, the hospital discharge records were generated and reviewed by health care professionals during hospitalization and heavily scrutinized by the health administration, which ensured high data quality as demonstrated in previous studies [42-44]. Third, linked records from all hospitals in an area of interest achieved an almost universal coverage of local residents, while the cancer registries in Shandong only covered 33.24% of its population [31]. It is true that if a patient visited hospitals outside Shandong Province, the corresponding record would not be captured in this study. Yet, due to geographical access and higher reimbursement, patients with cancer were generally diagnosed at health care facilities not far from their residences. Therefore, the impact of multiple hospital visits on the identification of incident cancer cases was considered not leading to substantive underestimation of the results in this study. In areas where the cancer registry sites are yet to be established, the approach using routinely collected hospital discharge records may provide reasonably comprehensive cancer statistics. This also offers a potentially powerful solution for generating high-resolution cancer statistics using existing data in China and other low- and middle-income countries.

While administrative data harbor great potential in disease surveillance, among other health-related topics, there are multiple barriers to use administrative data of sufficient volume for research [45]. First of all, the health-related records may not be digitized and structured into a dataset available for data analysis and linkage, especially in low-resource settings. In addition, data linkage algorithms, data accuracy, and coding standardization make synthesizing data across different institutions/regions a challenging project [46,47].

Limitations

This study has some limitations that warrant consideration when generalizing the findings. First, people who only sought cancer care outside Shandong would be missed. This issue would be magnified in an area lacking cancer-specialty resources. However, given that Shandong is a large middle-income province in China with sufficient cancer treatment capacity, it is highly unlikely that a substantial proportion of residents with cancer would rely purely on inpatient care outside the province. Second, outpatient cancer care was not included in this study. However, for an illness as severe as cancer, diagnostic tests, investigations, and confirmed diagnosis can rarely be completely provided in ambulatory settings, and the Chinese Lung Cancer

XSL•FO
RenderX

Clinical Guidelines recommend hospitalization for surgery in almost all cases [43]. Hence, this limitation is unlikely to lead to a serious underestimation of the cancer incidence.

Implications

Our study has some potential implications. First, we have demonstrated the feasibility of producing timely and small-area cancer incidence and mortality rates for a large population in a low- and middle-income setting. Considering the economic disparity within the province, our approach to cancer surveillance would be potentially suitable for a wide range of countries and areas [48]. Second, the increasing availability and standardization of administrative data means that administrative data should be increasingly made use of in generating more timely and spatially precise estimation of cancer incidence and mortality [49-51]. Third, the obvious disparity within the province (particularly within municipalities) suggests the importance of local level surveillance, policy making, and resource allocation that can reveal and address such geographical inequality. Fourth, linking claims data with hospital records and death registry where possible in the future could further improve the accuracy of estimation. It is also recommended that a comparison between the estimation of cancer incidence and mortality using our approach and cancer registries be conducted to further validate this approach based on linked administrative health data.

Conclusions

This study developed and validated a timely and spatially precise cancer surveillance method for a province of 100 million population in China using administrative data. The ASIR of lung cancer was 42.46 per 100,000, with a between-county IQR of 17.13 per 100,000. Meanwhile, the ASMR of lung cancer was 19.76 per 100,000, with a between-county IQR of 10.41 per 100,000. Besides providing new insights for cancer prevention and control in China, our study demonstrates the power of a novel approach linking hospital discharge records with death registries in enabling timely and geographically sensitive cancer surveillance. The fact that such data were routinely generated in this low- and middle-income setting with large population implies wide applicability as a complement to cancer registries.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (71804004) and Key Program (82330108), Henan Science and Technology Major Program (241100310300). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. Generative artificial intelligence was not used in any portion of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available but may be made available to bona fide researchers upon reasonable request to the corresponding author, Fuzhong Xue (xuefzh@sdu.edu.cn), conditional on obtaining all necessary ethics and institutional approvals.

Authors' Contributions

YH, JX, and FX conceptualized the study. YH, XX, QW, and JX developed the methodology. YH, XX, YZ, YM, XJ, QZ, YW, and YZ performed data curation and formal analyses. QW, FX, and JX were responsible for project administration. YH and XX drafted the manuscript. QW, YZ, YM, XJ, QZ, YW, YZ, and FX reviewed and edited the manuscript. YH, XX, and JX revised the manuscript. All authors have read and agreed to the publication of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary materials.
[DOCX File, 1611 KB - publichealth_v11i1e74062_app1.docx]

Checklist 1
STROBE checklist.
[DOCX File, 22 KB - publichealth_v11i1e74062_app2.docx]

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]

XSL•FO
RenderX

2. The Lancet. Cancer registries: the bedrock of global cancer care. Lancet 2025 Feb 1;405(10476):353. [doi: 10.1016/S0140-6736(25)00189-8] [Medline: 39892893]
3. Fogleman AJ, Mueller GS, Jenkins WD. Does where you live play an important role in cancer incidence in the U.S.? Am J Cancer Res 2015;5(7):2314-2319 [FREE Full text] [Medline: 26328263]
4. Chambers D, Simpson L, Hill-Briggs F, et al. Proceedings of the 8th Annual Conference on the Science of Dissemination and Implementation: Washington, DC, USA. 14-15 December 2015. Implement Sci 2016 Aug 1;11 Suppl 2(Suppl 2):100. [doi: 10.1186/s13012-016-0452-0] [Medline: 27490260]
5. Brenner H, Heisser T, Hoffmeister M. Delayed cancer registration and estimation of screening colonoscopy effects. JAMA Netw Open 2024 Oct 1;7(10):e2435669. [doi: 10.1001/jamanetworkopen.2024.35669] [Medline: 39352704]
6. Han B, Zheng R, Zeng H, et al. Cancer incidence and mortality in China, 2022. J Natl Cancer Cent 2024 Mar;4(1):47-53. [doi: 10.1016/j.jncc.2024.01.006] [Medline: 39036382]
7. Hughes N, Feltbower R, Stark D. The hidden barriers to using cancer registration data. Lancet Oncol 2024 Nov;25(11):1390-1391. [doi: 10.1016/S1470-2045(24)00530-8] [Medline: 39395436]
8. Ye X, Wei J, Li Z, et al. Design and implementation of a mobile system for lung cancer patient follow-up in China and initial report of the ongoing patient registry. Oncotarget 2017 Jan 17;8(3):5487-5497. [doi: 10.18632/oncotarget.13720] [Medline: 27911868]
9. Jena S, Sahoo KC, Samantaray K, Satpathy N, Epari V. Operational feasibility of hospital-based cancer registries in low- and middle-income countries: a systematic review. Cureus 2023 Jul;15(7):e42126. [doi: 10.7759/cureus.42126] [Medline: 37602029]
10. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet 2018 Mar 17;391(10125):1023-1075. [doi: 10.1016/S0140-6736(17)33326-3] [Medline: 29395269]
11. Zhang L, Wang H, Li Q, Zhao MH, Zhan QM. Big data and medical research in China. BMJ 2018 Feb 5;360:j5910. [doi: 10.1136/bmj.j5910] [Medline: 29437562]
12. Bronson MR, Kapadia NS, Austin AM, et al. Leveraging linkage of cohort studies with administrative claims data to identify individuals with cancer. Med Care 2018 Dec;56(12):e83-e89. [doi: 10.1097/MLR.0000000000000875] [Medline: 29334524]
13. Goldsbury DE, Armstrong K, Simonella L, Armstrong BK, O'Connell DL. Using administrative health data to describe colorectal and lung cancer care in New South Wales, Australia: a validation study. BMC Health Serv Res 2012 Nov 9;12(1):387. [doi: 10.1186/1472-6963-12-387] [Medline: 23140341]
14. Conderino S, Bendik S, Richards TB, et al. The use of electronic health records to inform cancer surveillance efforts: a scoping review and test of indicators for public health surveillance of cancer prevention and control. BMC Med Inform Decis Mak 2022 Apr 6;22(1):91. [doi: 10.1186/s12911-022-01831-8] [Medline: 35387655]
15. Penberthy L, McClish D, Pugh A, Smith W, Manning C, Retchin S. Using hospital discharge files to enhance cancer surveillance. Am J Epidemiol 2003 Jul 1;158(1):27-34. [doi: 10.1093/aje/kwg108] [Medline: 12835284]
16. Brackley ME, Penning MJ, Lesperance ML. In the absence of cancer registry data, is it sensible to assess incidence using hospital separation records? Int J Equity Health 2006 Oct 6;5(1):12. [doi: 10.1186/1475-9276-5-12] [Medline: 17026764]
17. Bernal-Delgado EE, Martos C, Martínez N, et al. Is hospital discharge administrative data an appropriate source of information for cancer registries purposes? Some insights from four Spanish registries. BMC Health Serv Res 2010 Jan 8;10(1):9. [doi: 10.1186/1472-6963-10-9] [Medline: 20059773]
18. Goldsbury D, Weber M, Yap S, Banks E, O'Connell DL, Canfell K. Identifying incident colorectal and lung cancer cases in health service utilisation databases in Australia: a validation study. BMC Med Inform Decis Mak 2017 Feb 27;17(1):23. [doi: 10.1186/s12911-017-0417-5] [Medline: 28241763]
19. Fu Z, Jiang F, Lu Z, et al. Changes in cancer mortality in Shandong Province, China: a large population based study. J Natl Cancer Cent 2025 Apr;5(2):132-139. [doi: 10.1016/j.jncc.2024.08.002] [Medline: 40265097]
20. Carstairs V. Small area analysis and health service research. J Public Health (Oxf) 1981 May;3(2):131-139. [doi: 10.1093/oxfordjournals.pubmed.a043414]
21. Shandong Provincial Bureau of Statistics. Shandong Statistical Yearbook 2022 [Book in Chinese]: China Statistics Press; 2022.
22. National Bureau of Statistics of China. China Statistical Yearbook 2022 [Book in Chinese]: China Statistics Press; 2022.
23. Yearly average currency exchange rates. United States Department of the Treasury. URL: https://www.irs.gov/individuals/international-taxpayers/yearly-average-currency-exchange-rates [accessed 2024-12-09]
24. New World Bank country classifications by income level: 2022-2023. World Bank. URL: https://blogs.worldbank.org/en/opendata/new-world-bank-country-classifications-income-level-2022-2023 [accessed 2024-12-09]
25. Fan J, Zhentao F, Zilong L, et al. Analysis of gastric cancer incidence and mortality trends in Shandong Province from 2012 to 2022 and forecast from 2023 to 2030 [Article in Chinese]. Chin J Oncol 2024;46(9):871-877. [doi: 10.3760/cma.j.cn112152-20231227-00387]
26. Shandong Province National Economic and Social Development Statistical Communiqué 2022 [Article in Chinese]: Shandong Provincial Bureau of Statistics; 2022.

XSL•FO
RenderX

27. National Institute of Health and Medical big data [Article in Chinese]. National Institute of Health Data Science of China. URL: http://www.mhdata.sdu.edu.cn [accessed 2024-12-09]
28. International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10). World Health Organization; 2019. URL: https://www.who.int/classifications/icd/en/ [accessed 2024-12-09]
29. Shandong Statistical Yearbook 2020 [Book in Chinese]: Shandong Provincial Bureau of Statistics; 2020.
30. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
31. European Network of Cancer Registries. ENCR Recommendations: Basis of Diagnosis. European Network of Cancer Registries. 2022. URL: https://encr.eu/sites/default/files/Recommendations/ENCR%20Recommendation%20BoD_Oct2022_EN_052025.pdf [accessed 2024-12-09]
32. Theil H. Economics and Information Theory: North-Holland Publishing Company; 1967.
33. Harper S, Lynch J, Meersman SC, Breen N, Davis WW, Reichman ME. An overview of methods for monitoring social disparities in cancer with an example using trends in lung cancer incidence by area-socioeconomic position and race-ethnicity, 1992-2004. Am J Epidemiol 2008 Apr 15;167(8):889-899. [doi: 10.1093/aje/kwn016] [Medline: 18344513]
34. Malakar K, Mishra T. Application of Gini, Theil and concentration indices for assessing water use inequality. IJSE 2017 Oct 9;44(10):1335-1347. [doi: 10.1108/IJSE-01-2016-0017]
35. Bray F, Ferlay J, Laversanne M, et al. Cancer incidence in five continents: inclusion criteria, highlights from volume X and the global status of cancer registration. Int J Cancer 2015 Nov 1;137(9):2060-2071. [doi: 10.1002/ijc.29670] [Medline: 26135522]
36. Gao Y, Zhang M, Liu G, Lin H, Huang H, Xie H. Historical data analysis and future prediction of lung cancer in Zhejiang province, China. Sci Rep 2025 Jul 1;15(1):21813. [doi: 10.1038/s41598-025-07200-5] [Medline: 40594686]
37. Forman D, Bray F, Brewster DH, Gombe MC, Kohler B, Piñeros M. Cancer Incidence in Five Continents Volume X: International Agency for Research on Cancer; 2014.
38. Sharma R. Mapping of global, regional and national incidence, mortality and mortality-to-incidence ratio of lung cancer in 2020 and 2050. Int J Clin Oncol 2022 Apr;27(4):665-675. [doi: 10.1007/s10147-021-02108-2] [Medline: 35020103]
39. Chen RY, Li YF, Long W, et al. Survey on tobacco use and associated factors in population in Shandong province, 2016-2017. Zhonghua Liu Xing Bing Xue Za Zhi 2021 Jul 10;42(7):1200-1204. [doi: 10.3760/cma.j.cn112338-20200903-01123] [Medline: 34814531]
40. Guo H, Chang Z, Wu J, Li W. Air pollution and lung cancer incidence in China: Who are faced with a greater effect? Environ Int 2019 Nov;132:105077. [doi: 10.1016/j.envint.2019.105077] [Medline: 31415963]
41. Zheng RS, Chen R, Han BF, et al. Cancer incidence and mortality in China, 2022. Zhonghua Zhong Liu Za Zhi 2024 Mar 23;46(3):221-231. [doi: 10.3760/cma.j.cn112152-20240119-00035] [Medline: 38468501]
42. Diop M, Strumpf EC, Datta GD. Measuring colorectal cancer incidence: the performance of an algorithm using administrative health data. BMC Med Res Methodol 2018 May 8;18(1):38. [doi: 10.1186/s12874-018-0494-x] [Medline: 29739338]
43. Chinese Medical Association. Chinese Medical Association guidelines for clinical diagnosis and treatment of lung cancer (edition 2024). Zhonghua Zhong Liu Za Zhi 2024;104(34):935-964. [doi: 10.3760/cma.j.cn112152-20240510-00189]
44. Weegar R, Sundström K. Using machine learning for predicting cervical cancer from Swedish electronic health records by mining hierarchical representations. PLOS ONE 2020;15(8):e0237911. [doi: 10.1371/journal.pone.0237911] [Medline: 32822401]
45. Boulanger V, MacLaurin A, Quach C. Barriers and facilitators for using administrative data for surveillance purpose: a narrative overview. J Hosp Infect 2025 Jan;155:25-36. [doi: 10.1016/j.jhin.2024.09.027] [Medline: 39454834]
46. Jutte DP, Roos LL, Brownell MD. Administrative record linkage as a tool for public health research. Annu Rev Public Health 2011;32(1):91-108. [doi: 10.1146/annurev-publhealth-031210-100700] [Medline: 21219160]
47. van Mourik MSM, van Duijn PJ, Moons KGM, Bonten MJM, Lee GM. Accuracy of administrative data for surveillance of healthcare-associated infections: a systematic review. BMJ Open 2015 Aug 27;5(8):e008424. [doi: 10.1136/bmjopen-2015-008424] [Medline: 26316651]
48. Hanna TP, Kangolle ACT. Cancer control in developing countries: using health data and health services research to measure and improve access, quality and efficiency. BMC Int Health Hum Rights 2010 Oct 13;10(1):24. [doi: 10.1186/1472-698X-10-24] [Medline: 20942937]
49. Abraha I, Montedori A, Serraino D, et al. Accuracy of administrative databases in detecting primary breast cancer diagnoses: a systematic review. BMJ Open 2018 Jul 23;8(7):e019264. [doi: 10.1136/bmjopen-2017-019264] [Medline: 30037859]
50. Shah SM, Khan RA. Secondary use of electronic health record: opportunities and challenges. IEEE Access 2022;8:136947-136965. [doi: 10.1109/ACCESS.2020.3011099]
51. Prins H, Kruisinga FH, Büller HA, Zwetsloot-Schonk JHM. Availability and usability of data for medical practice assessment. Int J Qual Health Care 2002 Apr;14(2):127-137. [doi: 10.1093/oxfordjournals.intqhc.a002599] [Medline: 11954682]

XSL•FO
RenderX

Abbreviations

ASIR: age-standardized incidence rate
ASMR: age-standardized mortality rate
Cheeloo LEAD: Cheeloo Lifespan Electronic Health Research Data-library
HRFP: hospitalization record front page
***ICD-10*:** *International Classification of Diseases, Tenth Revision*
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 17.03.25; peer-reviewed by M Jia, SK Mirinezhad, S Zhou; revised version received 16.07.25; accepted 18.07.25; published 29.09.25.

Please cite as:
He Y, Xia X, Wang Q, Zhang Y, Meng Y, Ji X, Zhao Q, Wang Y, Zhao Y, Xue F, Xu J
Small-Area Lung Cancer Incidence and Mortality: Cross-Sectional Population-Based Study Using Hospital Discharge and Death Registration Data
JMIR Public Health Surveill 2025;11:e74062
URL: https://publichealth.jmir.org/2025/1/e74062
doi:10.2196/74062

© Yu He, Xinxin Xia, Qing Wang, Yaoyun Zhang, Ying Meng, Xiaokang Ji, Qingbo Zhao, Yongchao Wang, Yifu Zhao, Fuzhong Xue, Jin Xu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correcting for the Inflated Adult Population Denominator in an English Nationwide Health Care Cohort: Database Analysis Study

Sudhir Venkatesan[1], PhD; Mark Joy[2], PhD; Gavin Jamie[2], MBChB; Debasish Kar[2,3], MSc, MD; Robert Williams[2], BSc; Xuejuan Fan[2], PhD; Wilhelmine Meeraus[4], PhD; Ruby S M Tsang[5], PhD; Kathryn S Taylor[2], PhD; Sylvia Taylor[4], PhD; F D Richard Hobbs[2], MBChB; Sneha N Anand[2], PhD; Chris Robertson[6], PhD; Simon de Lusignan[2], PhD, MBChB

[1]BPM Evidence Statistics, BioPharmaceuticals Medical, AstraZeneca, Cambridge, United Kingdom

[2]Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, United Kingdom

[3]Faculty of Health, University of Plymouth, Plymouth, United Kingdom

[4]Vaccines and Immune Therapies, BioPharmaceuticals R&D, AstraZeneca, Cambridge, United Kingdom

[5]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

[6]Department of Mathematics and Statistics, University of Strathclyde, Glasgow, United Kingdom

Corresponding Author:
Simon de Lusignan, PhD, MBChB
Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, United Kingdom

Abstract

Background: Electronic health care databases are widely used for epidemiological studies. However, they may contain inactive records of individuals no longer participating in the health care system. These inactive records create a methodological challenge as they systematically appear as unexposed with no recorded outcomes. Given the widespread health care system engagement during the COVID-19 pandemic, the English National Health Service (NHS), which hosts a national pandemic planning and research dataset with linkage to COVID-19 vaccination and emergency care data, makes it an ideal setting to identify the extent of overrepresentation due to inactive health care records and assess ways to mitigate them.

Objective: The objective of this study is to report any differences between the general practitioner–registered adult population size based on health care records compared to census estimates for England and to apply methodology that could be used to correct for such differences.

Methods: We compared the number of adult patients within the General Practice Extraction Service Data for Pandemic Planning and Research (GDPPR) with a valid general practitioner registration as of 1st October 2021, with estimates published by the Office for National Statistics (ONS) for the English population. We used an approach adapted from a weighting method to correct for non-response bias in surveys and down-weighted individuals with no evidence of recent activity in their records.

Results: There were 61,194,033 registered NHS patients (in the GDPPR) compared with 56,550,138 in the ONS census-based population. De-duplication on NHS number reduced the population to 57,876,641, including 46,835,968 adults, with the biggest overrepresented group aged 30 - 45 years. Of the 46,835,986, 1,121,954 (2.4%) individuals had their initial weights down-weighted due to non-engagement with the health care system since January 2019. The down-weighting removed most of the differences between NHS and ONS populations.

Conclusions: There are notable differences in the adult population size as per GDPPR when compared to census estimates. While the overall population size in the GDPPR data was seen to be inflated when compared to ONS census estimates, this was differential with respect to sociodemographic variables. A weighting-based approach can be applied to correct for the inflated denominator. Not correcting for it in large health care datasets, including the English NHS data, could introduce selection bias in epidemiological studies.

(JMIR Public Health Surveill 2025;11:e64788) doi:10.2196/64788

KEYWORDS

general practitioners; pandemic planning; COVID-19 vaccines; census; selection bias; emergency medical services; national; service; NHS; weighting; record; size; population; registration; register; Health Informatics

XSL•FO
RenderX

Introduction

Electronic health care databases are widely used resources for conducting epidemiological studies, offering advantages of potentially large sample sizes, reduced costs, and rapid analysis compared to prospective designs. These databases contain longitudinal patient records from routine clinical practice, enabling researchers to evaluate epidemiological associations in real-world settings across diverse populations. However, the secondary use of data collected primarily for administrative or clinical purposes introduces unique methodological challenges that must be addressed to ensure valid inference [1].

An important challenge in database studies is the presence of inactive records—electronic records of individuals who are no longer active participants in the health care system but remain in the database without formal disenrollment. Unlike traditional loss to follow-up in prospective studies where researchers are aware of missing data, inactive records in electronic databases create a particular form of selection bias because they continue to contribute person-time to analyses despite no longer being under observation. While studies have recognized the issue of inactive records potentially inflating the denominator for database [2-4], the precise extent of the issue has not been fully assessed. The COVID-19 pandemic in England offers a valuable opportunity to study denominator inflation, particularly given the high levels of health care engagement observed during this period [5]. England's National Health Service (NHS) provides comprehensive state-funded health care—including primary, community, and specialist services—to approximately 56 million people, creating an ideal setting for examining this methodological issue.

NHS care includes the provision of medications and the administration of nearly all vaccinations. All NHS patients have a unique personal identifier, an NHS number, which allows records to be linked across NHS services. Additionally, primary care is a registration-based system, with each patient registered with a single general practitioner (GP) [6,7]. Furthermore, NHS computerized medical records (CMRs) have been in use in primary care since the 1980s [8]. Additionally, since 2011, there has been automatic transfer of CMR between practices, even when they use different brands of computer systems [9].

During the COVID-19 pandemic, England was one of the first health systems to use COVID-19 vaccines widely, with high levels of COVID-19 primary series vaccine uptake in adults [10]. The NHS has made a large nationwide cohort available for research through NHS England's trusted research environment (TRE) [11]. This cohort included the NHS primary care data collections, known as the General Practice Extraction Service Data for Pandemic Planning and Research (GDPPR), which was linked at the individual level to COVID-19 vaccination, COVID-19 virology test results, hospital, and death certificate data.

The discrepancy in population size between GDPPR and Office for National Statistics (ONS) census estimates has been pointed out before [4]. The phenomenon, in NHS data, where an individual may have multiple unique health care records because of not notifying the health care system when they moved to a different GP practice [4,12,13] is known by different terms in the literature, including over-registration, "ghost" patient records, and inactive records. We subsequently refer to such records as 'inactive records' in this paper. We aimed to assess the extent of any potential discrepancy in population sizes between the NHS England GDPPR dataset and the ONS census estimates (eg, due to over-registration) and then develop and implement a method to minimize the impact of this discrepancy for epidemiological studies.

Methods

Data Source

We used de-identified routinely collected electronic health care record data from primary and secondary care in England accessed via the NHS Digital TRE. Access to the NHS Digital TRE was obtained through a Data Access Request Service (DARS) application. The linked data used for this analysis included GP registration data from GDPPR [10], COVID-19 Second Generation Surveillance System from Pillar 1 and Pillar 2 for COVID-19 PCR testing data, COVID-19 vaccination status data, and Hospital Episode Statistics. Linked and pseudonymized patient-level data were provided to us by NHS Digital within the secure TRE. Details on the data flow, end-to-end linkage process, and data privacy can be found on the NHS Digital website [14]. We used data as of October 1, 2021, for our analyses. In this linked dataset, patient age and sex were obtained from the GDPPR and Personal Demographics Service. The other variables used in this analysis (COVID-19 status, vaccination status, and health care engagement) were defined as yes or no binary flags based on the information from the relevant data source (Second Generation Surveillance System for COVID-19 status, Hospital Episode Statistics for hospital admission).

We assessed the total number of unique, pseudonymized patient identifiers with a valid GP registration as of October 1, 2021, and compared it to estimates published by the ONS for the English population over a comparable time period [15,16]. We refer to this overall GP registration estimate as the "population denominator," as this will constitute the source population that study-specific cohorts will be constructed from in epidemiological studies that will use these CMR data. We then implemented a method for the denominator correction based on the approach used in a Scottish vaccine pharmacovigilance study [17,18] which is, in turn, based on a weighting method to correct for non-response bias in surveys [19]. We adapted this method to apply to the NHS England data.

Step 1

We obtained data from NHS England from GP practices in England as of 1 October 2021. Duplicate records and records missing age and sex were excluded. We used age, sex, and NHS region as the sociodemographic variables for the population weighting. These variables were chosen on the basis that these are some of the key determinants of the differential representation of population groups in the NHS England data [12].

XSL•FO
RenderX

Step 2

We restricted these analyses to those aged 18 years and over for consistency with the UK COVID-19 vaccine rollout policy and also due to difficulties with validating children's GP registration data.

Step 3

We obtained population proportions from the ONS on the projected mid-2020 English population stratified for each year of age (≥18 y) for males and females by the following NHS regions: East of England, London, Midlands, North East and Yorkshire, North West, South East, and South West.

Step 4

Each individual record in the NHS England analytic dataset was then assigned a weight to match the target distributions obtained from Step 4 such that the weights add up to the target population size (ie, England's mid-2020 population). These weights were derived using age, sex, and NHS region. For each individual, k, within the NHS England population, the initial study weight, w, was calculated as:

$$w_{ijrk} = N_{ijr} / n_{ijr}$$

where N is the number of individuals in the ONS census population in age group i with sex j and region r, and n is the corresponding figure for the NHS Digital population.

Step 5

Patient records with recent health care system engagement, defined as *any* of the following since 1st January 2019: GP consultation, hospitalization, laboratory test results, COVID-19 testing, and COVID-19 vaccination were reweighted to 1, as such an encounter would confirm that the individual health care record is active.

Step 6

Weights for those patient records without a recent health care system engagement were updated such that the total revised weights (in the overall NHS England data) sum to the ONS estimate of the population of England age ≥18 years. For each individual, k, in age group i with sex j and region r, the revised weight, w*, was calculated as follows:

$$w^*_{ijrk} = S_w - S_h S_w - S_{w_nh}$$

where S_w is the sum of the original weights (and therefore equals the total ONS population size), S_h is the sum of the updated weights for individuals with a recent health care engagement (as defined in Step 5), and S_{w_nh} is the sum of the original weights (assigned in Step 4) of individuals without a recent health care engagement.

A validation of the revised weights was performed by comparing the patient profile (in terms of age, sex, and region) of the unvaccinated population in the NHS England data to the unvaccinated patient profile information provided by the ONS [20].

Finally, we also conducted an illustrative bias analysis to demonstrate the impact of the bias introduced by not accounting for inactive records in a hypothetical vaccine effectiveness (VE) study.

Ethical Considerations

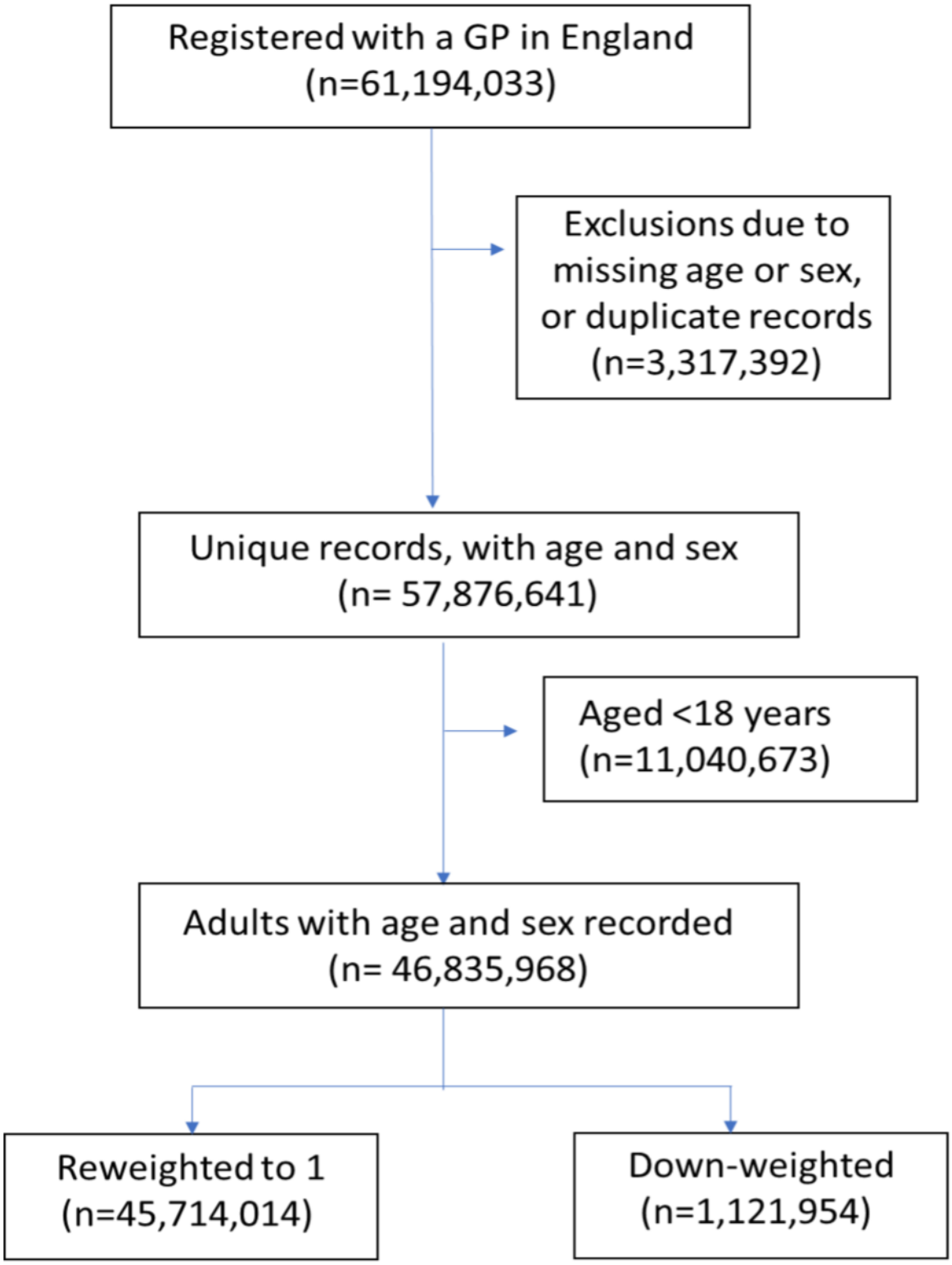
We undertook this work as a part of the RAVEN study, which investigated the effectiveness of the AZD1222 vaccine using CMR from England (Real-world Oxford/AstraZeneca Vaccine Effectiveness Study in England; NCT05047822) [21].

Results

There are 61,194,033 individuals registered at GP practices in England as of 1 October 2021, according to NHS England [22]. However, the ONS estimated the population of England to be 56,550,138 in mid-2020 (including 44,456,850 records for individuals aged ≥18 y) [23]. After de-duplication (ie, removing duplicate records with the same pseudonymized patient identifier) and removing records with missing age or sex, the NHS England data had a total of 57,876,641 unique NHS numbers (including 46,835,968 records for individuals aged ≥18 y) (Figure 1). Age-stratified (categories of age) counts of the NHS England population when compared to ONS estimates revealed that individuals aged 30 - 45 years were the most overrepresented groups, and those between 65 and 75 years were among the most underrepresented (Figure 2).

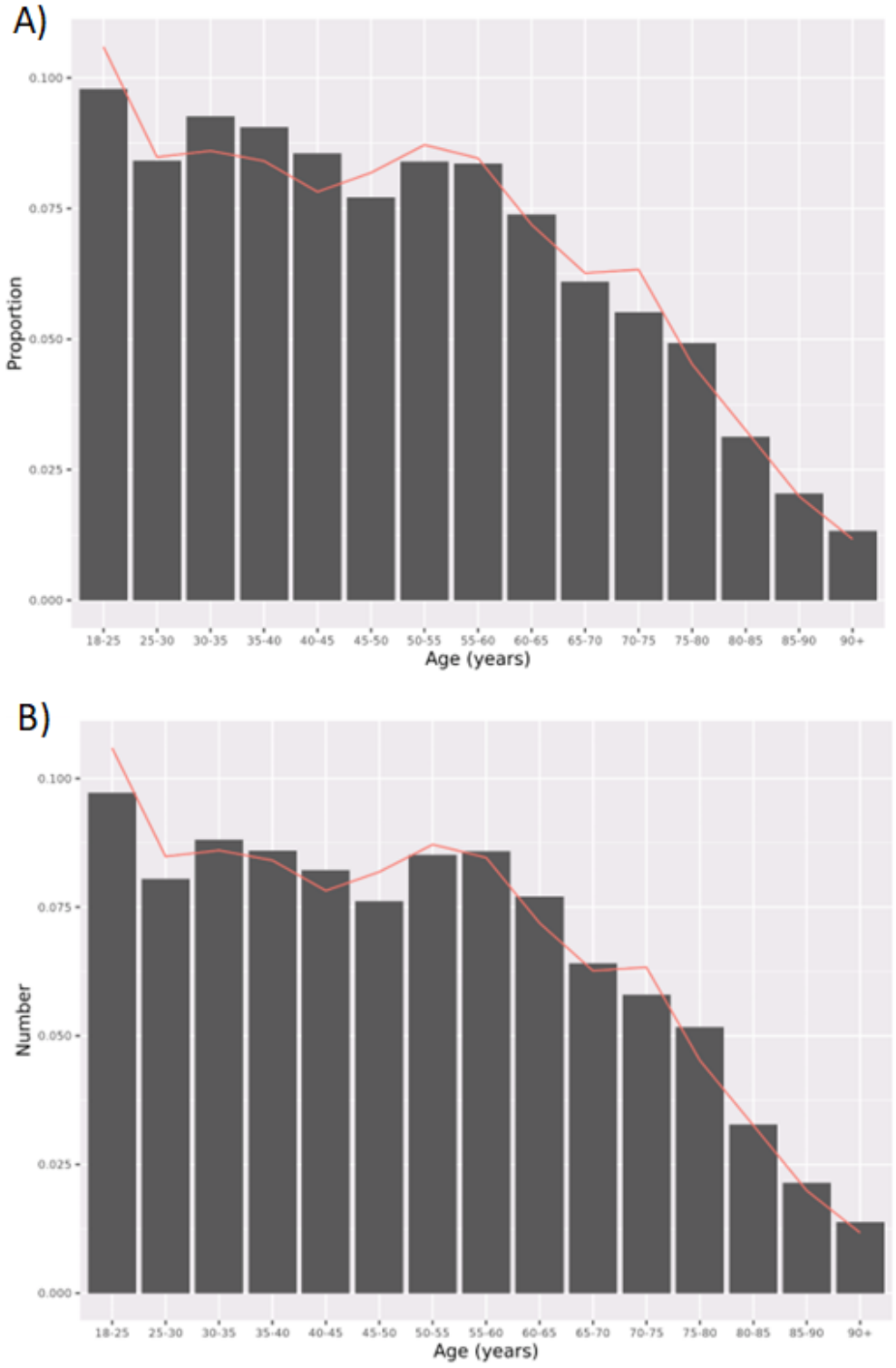
XSL•FO
RenderX

Figure 1. Study flow diagram showing exclusions at each stage.



XSL•FO
RenderX

Figure 2. Age distribution (by age group) of the NHS England data (columns) compared to ONS Census data (line) (A) before and (B) after applying weights for inflated denominator correction.



A total of 97.6% of the NHS England adult population had their initial weights revised to 1 on the basis that they had had some form of engagement with the health care system. Of 46,835,968 individuals, 1,121,954 (2.4%) had their initial weights down-weighted due to non-engagement with the health care system since January 2019. The median (IQR) age of the overall NHS population (≥18 years) was 48 years (18-63), whereas the median (IQR) of the down-weighted population was 36 years (18-44). Of the 46,835,968 NHS England records (population aged ≥18 years), 82% had a record of at least 1 COVID-19 vaccination, 62% had a record of a COVID-19 test, 97% had a recorded GP visit, 28% had been admitted to hospital at some point, and 24% had a record of a lab result (any lab test for any condition) since the 1st of January 2019.

Once the final weights were derived, the magnitude of the overrepresentation of the 30 - 45 years age category was

XSL•FO
RenderX

decreased, although some degree of overrepresentation in these age groups still persisted (Figure 2).

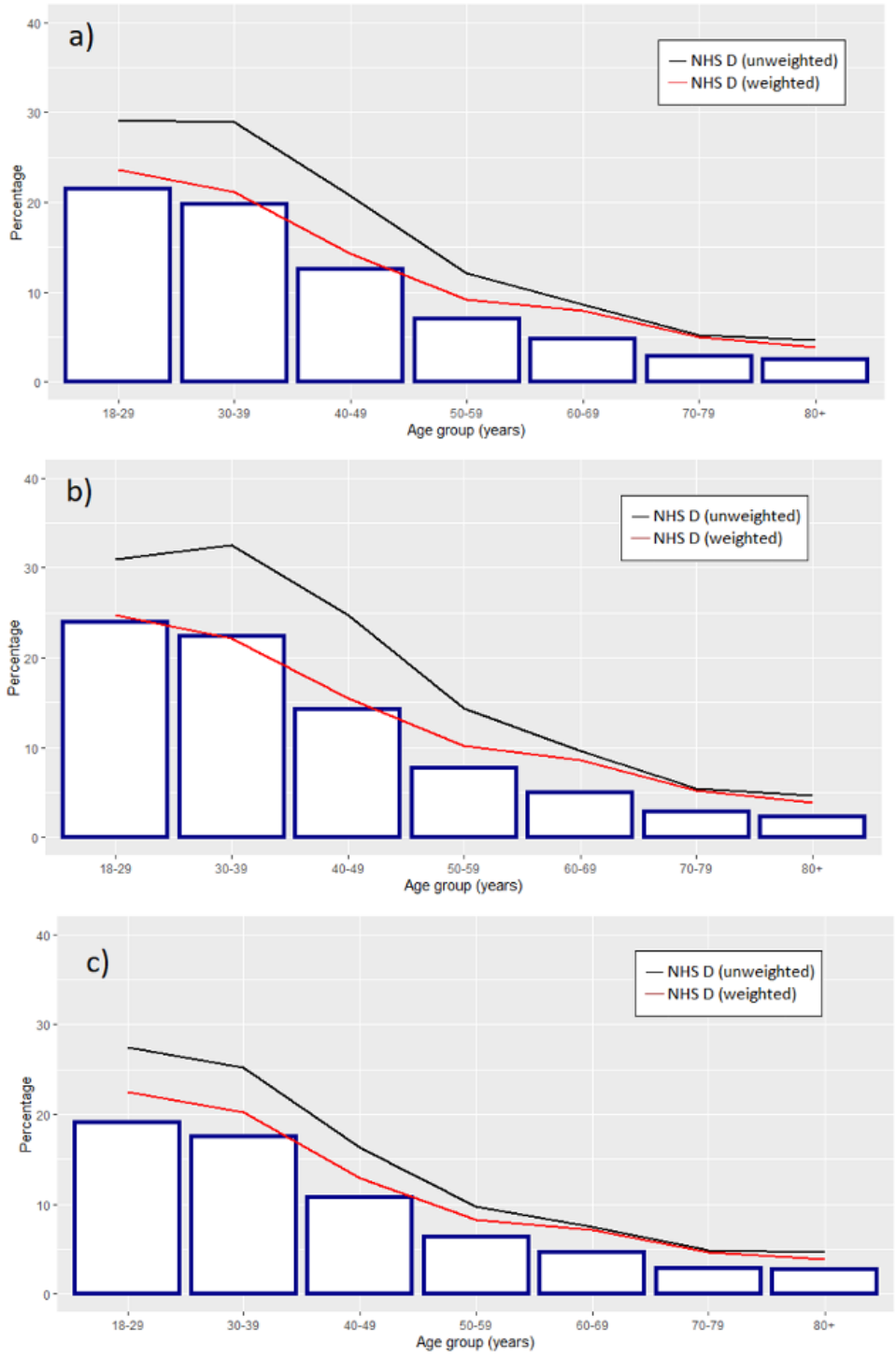
To assess the impact of the denominator correction, we compared the age and sex distribution of the weighted (denominator corrected) population to the unvaccinated patient profile published by the ONS (in individuals aged ≥18 years between 8/12/2021 and 31/12/2021) based on a subset of the national population [20]. Overall, large differences were seen between the unweighted NHS England data and the ONS estimates. These differences were largely minimized when using the denominator-corrected weighted NHS England data. The most overrepresented group in the NHS England data appeared to be those aged 30 - 45 years (Table 1). While most of the overrepresentation in the age groups between 18 and 49 years was seen to have been corrected for when using the weighted data, some degree of overrepresentation was still seen in the age groups between 50 and 80+ years (Table 1, Figure 3, and Multimedia Appendix 1).

Table . Comparison of age distribution of the NHS England data compared to ONS census data before and after weighting.

Age, years	ONS (%)	NHS Digital	
		Unweighted (%)	Weighted (%)
18-25	10.6	9.8	9.7
25-30	8.5	8.4	8.1
30-35	8.6	9.3	8.8
35-40	8.4	9.1	8.6
40-45	7.8	8.6	8.2
45-50	8.2	7.7	7.6
50-55	8.7	8.4	8.5
55-60	8.5	8.4	8.6
60-65	7.2	7.4	7.7
65-70	6.3	6.1	6.4
70-75	6.3	5.5	5.8
75-80	4.5	4.9	5.2
80-85	3.2	3.1	3.3
85-90	2.0	2.0	2.2
90+	1.2	1.3	1.4

XSL•FO
RenderX

Figure 3. Comparison between the age distributions in the NHS England weighted and unweighted data and the ONS published data among unvaccinated individuals in (A) the overall population, (B) males, and (C) females. NHS D: GDPPR data from the NHS Digital Trusted Research Environment.



In our bias analysis, for a hypothetical VE study, we assumed an analytic dataset of 10,000 individual records. In Scenario A (reference case), we correctly excluded all inactive records, analyzing only the 9700 active patients. With 60% vaccination coverage, a 1% event rate among vaccinated and a 5% event rate among unvaccinated individuals, we calculated a true VE of 70%. In Scenario B, we incorrectly included 300 inactive records (3% of total) in the analysis, all classified as unvaccinated with no events. This inflated the unvaccinated denominator to 4180 while keeping the event count unchanged, resulting in an underestimated VE of 67.7%, representing an absolute bias of 2.3% (−3.3% relative bias). In Scenario C, we implemented the weighting approach described in this paper, where inactive records received a mean weight of 0.2 and active records a mean weight of 0.9. This approach effectively reduced the contribution of likely inactive records while maintaining

XSL•FO
RenderX

most of the influence from active records. The weighted analysis yielded a VE estimate of 69.5%, much closer to the true VE seen in Scenario A with only a 0.5% absolute bias (−0.7% relative bias). A higher level of denominator inflation will be associated with a greater degree of absolute bias in VE estimates if inactive records are not accounted for (Multimedia Appendix 2).

Discussion

Main Findings

Our assessment of the NHS England GDPPR data showed that the population size of individuals registered with a GP in England was inflated overall, but specific age groups were underrepresented when compared to the ONS census estimates. The differential nature of the discrepancy could cause certain demographic groups to be over- or under-represented in the NHS England data, which in turn has the potential to influence the results of epidemiological studies that use these data. Possible reasons for "over-registration" and the presence of these inactive records include young, mobile individuals who may have moved out of the country. The NHS number was introduced in 1995; its precise structure and systematic use within NHS information systems were made mandatory in 2009 [24]. Prior to this, multiple registrations were possible. Over-registration has also become much less of a problem since GP2GP transfer of records was introduced in 2011 [9].

Comparison With Other Studies

Inactive electronic medical records are known to be correlated strongly with sociodemographic factors [12,25], which is a finding that our study corroborates. While methods proposed to adjust for this inflated denominator include data-cleaning exercises [12] and identifying an "active patient" denominator through a validation sample [2], the success of these methods relies on selecting a representative validation sample, and given the large sociodemographic and spatial variation in denominator inflation, this may not be feasible. The COVID-19 pandemic period, which saw nationwide engagement with the NHS, could provide a unique opportunity to infer the "active" status of a patient's records. We have described a method to first identify the extent of the denominator inflation by inferring "active" registrations through engagement with the health care system (ie, GP consultation, hospitalization, laboratory test result, COVID-19 testing, or COVID-19 vaccination) during the pandemic period and then to correct for it by reducing the weights of those records that did not have any engagement with the health care system during this period. A similar weighting approach that was derived based on population age and sex distribution has been successfully implemented in a series of COVID-19 VE studies based on Scottish data [17,26,27]. We adapted to the English context by also taking regional variability into account. We note that in cohort profiles for OpenSafely and Clinical Practice Research Datalink, similar differences are seen between the adult age-sex profile and ONS data [28,29]. In this paper, we specifically looked at the NHS data in England; however, this issue is a wider problem that could occur in other secondary health care databases as well.

Implications

Not properly accounting for the inflated denominator in NHS England data could introduce selection bias if it results in follow-up time being differentially accrued in one of the study arms. We illustrated the impact of inactive medical records on a hypothetical VE study in a bias analysis. Inactive records systematically appear as unvaccinated, since they cannot receive documented vaccinations, and without recorded outcomes, since events cannot be documented. This biases the association between vaccination and the outcome of interest by artificially inflating the denominator of unvaccinated individuals without a corresponding increase in events, leading to underestimation of event rates in the unvaccinated group and consequently underestimation of VE. The weighting method that we have proposed seeks to address this bias by decreasing the amount of information contributed to by records with lower weights, which are more likely to be inactive records. Our approach made the observed vaccination and outcome status more closely approximate to the true vaccination and outcome status. The extent of this selection bias will depend on the degree of denominator inflation and also on the sociodemographic factors that the inflation is differential with respect to. The success of the weights in minimizing the impact of selection bias from inactive records will depend on the extent to which they can approximate the true vaccination and outcome status.

We identified an average of 2.4% inflation in the population aged ≥18 years, with this being much higher within specific sociodemographic groups. Indeed, population underrepresentation (as seen in the 65 - 75 y age group) would also result in a similar type of selection bias. Our weighting approach accounts for both underrepresentation and overrepresentation using the same method, ie, reducing the weights of those records without a recent health care system engagement during the COVID-19 pandemic.

Strengths and Limitations

A strength of the weighting approach that we propose is that these individual records are not discarded entirely. The estimated weights can be thought of as confidence scores for each record, and when taken into account in a weighted regression model, they still allow the down-weighted records to contribute information to the model but minimize the impact of any selection bias. The weighting method that we have applied sits within the paradigm of inverse probability weighting methods that are widely used to minimize confounding in epidemiological studies [30]. Current approaches to addressing the impact of inactive records have notable limitations. Complete case analysis that attempts to exclude inactive records often relies on arbitrary rules, such as requiring a minimum number of health care encounters within a specified period. However, such approaches may inadvertently exclude truly healthy individuals with legitimately low health care utilization, introducing another form of selection bias. Alternatively, ignoring the issue by including all records risks substantial underestimation of VE, as our bias analysis demonstrates. Validation samples, where a subset of records undergoes intensive verification through direct patient contact or linkage to external data sources, provide gold-standard identification of inactive records but are

XSL•FO
RenderX

resource-intensive and often infeasible at scale. Probabilistic linkage to external data sources (eg, death registries, migration records) can identify specific causes of inactivity but may miss others and introduce complex data governance challenges. Statistical multiple imputation approaches handle uncertainty in inactive status but rely on correctly specified imputation models and may be computationally intensive for large databases. Our proposed weighting method offers several advantages: it retains all available data, acknowledges uncertainty in inactive status classification, can incorporate diverse predictors of inactivity, and is computationally efficient even with millions of records. However, it requires careful validation of the weighting algorithm and may be sensitive to misspecification of the weighting model. The optimal approach likely depends on specific study contexts, with hybrid methods potentially offering the most robust solution for large-scale VE studies. This is an avenue for future research.

Our method for denominator correction has other limitations as well: It assumes that reference population estimates used to derive the weights are based on perfect information. The limitations of ONS census estimates are well-known [12], and it is indeed possible that sociodemographic determinants of census completion are correlated with GP registration. Further, we do not include ethnicity and social deprivation in our analysis, which is a limitation of our weighting scheme. There is evidence to suggest that ethnicity and social deprivation are other key determinants of over-registration [12], but explicitly including these variables in the weighting would require that marginal frequencies be available for the variables in both the GP records database and the census data, which were not available for this analysis. The IMD score, which is the national mapping of deprivation, uses postcodes as of 2019, which means newer postcodes since then do not have a mapping [31]. Further, ethnicity recording is not usually complete either and may need supplemental mapping [32], which would bring with it additional complexities when analyzed. The robustness of a weighting scheme that does not explicitly account for ethnicity or socioeconomic deprivation will hinge on the extent to which the NHS region can be used as a proxy for the 2 variables. Finally, the weighting system does not distinguish between genuine inactive records versus active records of individuals who may have chosen not to engage with the health care system during the pandemic, therefore, potentially also assigning a low weight to these individuals.

Our method does not correct for inherent limitations of the reference population data used. We have also limited the analyses presented in this paper to the adult population (≥18 y), as there are challenges with using census data to validate primary care records of children. Also, we have only explored variables in primary care and have not allowed for diagnoses made in secondary care (eg, cases of chronic liver disease), which may compensate for the gap in primary care diagnosis.

It is also worth noting that while the widespread health care engagement that was observed in England during the COVID-19 pandemic provided a unique opportunity to assess the active status of health care records, a similar population-wide engagement is unlikely in the near future. The ascertainment of the size of the inactive population and application of our weighting method could potentially also be extended to non-pandemic periods. For instance, the weighting algorithm could incorporate routine health care utilization patterns as predictors of active status, including prescription refill consistency for chronic medications, attendance at preventive screenings, completion of age-appropriate vaccinations, and regular check-ups for specific demographic groups (eg, prenatal visits, well-child visits, and geriatric assessments). Validation could follow a three-pronged approach: (1) temporal validation comparing algorithm performance across different calendar periods, (2) external validation through linkage with a subset of verified active or inactive statuses from administrative sources, and (3) sensitivity analyses examining how varying thresholds of predicted inactivity affect VE estimates. Researchers should calibrate weights to their specific health care system's characteristics, accounting for factors like insurance coverage patterns, referral requirements, and electronic record integration across care settings, which all influence the observable signals of patient activity. This methodological challenge presents a rich opportunity for future research to further develop and validate standardized approaches for addressing bias in epidemiological studies due to the presence of inactive records.

Conclusions

Denominator inflation in large CMR datasets is a known issue. In this paper, we provide tools for identifying an active population in the NHS England data using a system of sampling weights to minimize the selection bias introduced by denominator inflation. We recommend, where appropriate, that epidemiological analyses of large CMR datasets, including the NHS England data, be implemented as weighted regressions taking into account the likelihood of population over- or under-representation. This is especially necessary for closed data access mechanisms on such data where analyses are often performed in specific defined datasets and an assessment of the impact of mis-specifying the denominator may be difficult, or impossible, to undertake.

Acknowledgments

This work was funded by AstraZeneca. We would also like to thank Julian Sherlock for his contributions in preparing the dataset that was used for the analyses described in this paper.

XSL•FO
RenderX

Data Availability

Data underlying the findings described in this manuscript may be obtained in accordance with AstraZeneca's data sharing policy [33].

Authors' Contributions

Conception and design were done by SV, MJ, and SdeL, in discussion with CR; analysis was done by SV; SV, GJ, DK, RW, XF, WM, RT, KST, ST, FDRH, SNA, RB, CR, and SdeL contributed to the interpretation of results; SV, SdeL contributed to the preparation of the first draft of the manuscript; SV, GJ, DK, RW, XF, WM, RT, KST, ST, FDRH, SNA, RB, CR, and SdeL were involved in the review and final manuscript approval.

Conflicts of Interest

FDRH was UK PI for the AZ-sponsored TACKLE trial of a dual COVID monoclonal antibody. SV, WM, and ST are employees of and may hold stock or stock options in AstraZeneca.

Multimedia Appendix 1
Comparison between the age distributions in the NHS England weighted and unweighted data and the ONS published data among unvaccinated individuals in the overall population and stratified by sex.
[DOCX File, 16 KB - publichealth_v11i1e64788_app1.docx]

Multimedia Appendix 2
Methods for bias analysis.
[DOCX File, 18 KB - publichealth_v11i1e64788_app2.docx]

References

1. Dinov ID. Methodological challenges and analytic opportunities for modeling and interpreting big healthcare data. Gigascience 2016;5(1):12. [doi: 10.1186/s13742-016-0117-6] [Medline: 26918190]
2. Robson J, Falshaw M. Audit of preventive activities in 16 inner London practices using a validated measure of patient population, the "active patient" denominator. Healthy Eastenders Project. Br J Gen Pract 1995 Sep;45(398):463-466. [Medline: 7546868]
3. Loper OR, Schultz JM, King TL, Tichy KJ, Kintigh BF, Callaghan DJ. Benefits and challenges of sealing deceased individual immunization records in Immunization Information Systems (IIS): an overview of Iowa's process. J Public Health Manag Pract 2024;30(2):240-243. [doi: 10.1097/PHH.0000000000001858] [Medline: 38153257]
4. Piel FB, Parkes BL, Daby H, Hansell AL, Elliott P. The challenge of opt-outs from NHS data: a small-area perspective. J Public Health (Oxf) 2018 Dec 1;40(4):e594-e600. [doi: 10.1093/pubmed/fdy059] [Medline: 29590471]
5. Weekly national influenza and COVID-19 surveillance report week 43 report (up to week 42 data). : UK Health Security Agency; 2021 URL: https://assets.publishing.service.gov.uk/media/617a825b8fa8f5297eda6806/Weekly_Flu_and_COVID-19_report_w43.pdf [accessed 2025-07-18]
6. de Lusignan S, van Weel C. The use of routinely collected computer data for research in primary care: opportunities and challenges. Fam Pract 2006 Apr;23(2):253-263. [doi: 10.1093/fampra/cmi106] [Medline: 16368704]
7. Martin G, Clarke J, Liew F, et al. Evaluating the impact of organisational digital maturity on clinical outcomes in secondary care in England. NPJ Digit Med 2019;2(1):41. [doi: 10.1038/s41746-019-0118-9] [Medline: 31304387]
8. Schade CP, Sullivan FM, de Lusignan S, Madeley J. e-Prescribing, efficiency, quality: lessons from the computerization of UK family practice. J Am Med Inform Assoc 2006;13(5):470-475. [doi: 10.1197/jamia.M2041] [Medline: 16799129]
9. Ellis B, Howard J, Dedman D, Hawking M. Perceptions on the quality of records received via the GP2GP electronic transfer service: pilot online questionnaire study. J Innov Health Inform 2011;19(4):233-240. [doi: 10.14236/jhi.v19i4.818]
10. Tessier E, Rai Y, Clarke E, et al. Characteristics associated with COVID-19 vaccine uptake among adults aged 50 years and above in England. BMJ Open 2020 May 8;12(3):e055278. [doi: 10.1136/bmjopen-2021-055278] [Medline: 31937653]
11. Wood A, Denholm R, Hollings S, et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. BMJ 2021 Apr 7;373:n826. [doi: 10.1136/bmj.n826] [Medline: 33827854]
12. Burch P, Doran T, Kontopantelis E. Regional variation and predictors of over-registration in English primary care in 2014: a spatial analysis. J Epidemiol Community Health 2018 Jun;72(6):532-538. [doi: 10.1136/jech-2017-210176] [Medline: 29449351]
13. Salisbury C, Macleod J, Egger M, et al. Opportunistic and systematic screening for chlamydia: a study of consultations by young adults in general practice. Br J Gen Pract 2006 Feb;56(523):99-103. [Medline: 16464322]
14. NHS digital. General practice extraction service (GPES) data for pandemic planning and research: a guide for analysts and users of the data. URL: https://digital.nhs.uk/coronavirus/gpes-data-for-pandemic-planning-and-research/guide-for-analysts-and-users-of-the-data [accessed 2025-08-28]

XSL•FO
RenderX

15. Office for national statistics. UK census data - Office for National Statistics. 2021. URL: https://www.ons.gov.uk/census/planningforcensus2021/ukcensusdata [accessed 2025-08-28]
16. Quality and methodology information (QMI) for census 2021. Office for National Statistics. 2021. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/methodologies/qualityandmethodologyinformationqmiforcensus2021#about-this-quality-and-methodology-information-report [accessed 2025-06-30]
17. Vasileiou E, Simpson CR, Robertson C, et al. Effectiveness of first dose of COVID-19 vaccines against hospital admissions in Scotland: national prospective cohort study of 5.4 million people. SSRN Journal 2021. [doi: 10.2139/ssrn.3789264]
18. Kerr S, Bedston S, Cezard G, et al. Undervaccination and severe COVID-19 outcomes: meta-analysis of national cohort studies in England, Northern Ireland, Scotland, and Wales. The Lancet 2024 Feb;403(10426):554-566. [doi: 10.1016/S0140-6736(23)02467-4]
19. Holt D, Elliot D. Methods of weighting for unit non-response. Statistician 1991;40(3):333. [doi: 10.2307/2348286]
20. Office of National Statistics. Coronavirus and vaccination rates in people aged 18 years and over by socio-demographic characteristic and occupation, England. Ons.gov.uk. 2022 Jan 20. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthinequalities/datasets/coronavirusandvaccinationratesinpeopleaged18yearsandoverbysociodemographiccharacteristicandoccupationengland [accessed 2025-06-30]
21. Meeraus W, Joy M, Ouwens M, et al. AZD1222 effectiveness against severe COVID-19 in individuals with comorbidity or frailty: the RAVEN cohort study. J Infect 2024 Apr;88(4):106129. [doi: 10.1016/j.jinf.2024.106129] [Medline: 38431156]
22. NHS Digital. Patients Registered at a GP Practice - October 2021. 2021 Oct. URL: https://digital.nhs.uk/data-and-information/publications/statistical/patients-registered-at-a-gp-practice/october-2021 [accessed 2025-06-30]
23. Office of national statistics. Population Estimates for the UK, England and Wales, Scotland and Northern Ireland - Office for National Statistics. 2021 Jun 25. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2020 [accessed 2025-08-29]
24. Boyd A, Thomas R, Macleod J. NHS numbers and their management systems resource report. 2018 URL: https://closer.ac.uk/wp-content/uploads/CLOSER-resource-NHS-Numbers-and-their-management-systems.pdf [accessed 2025-08-29]
25. Kontopantelis E, Panagioti M, Farragher T, et al. Consultation patterns and frequent attenders in UK primary care from 2000 to 2019: a retrospective cohort analysis of consultation events across 845 general practices. BMJ Open 2021 Dec 20;11(12):e054666. [doi: 10.1136/bmjopen-2021-054666] [Medline: 34930742]
26. Agrawal U, Katikireddi SV, McCowan C, et al. COVID-19 hospital admissions and deaths after BNT162b2 and ChAdOx1 nCoV-19 vaccinations in 2·57 million people in Scotland (EAVE II): a prospective cohort study. Lancet Respir Med 2021 Dec;9(12):1439-1449. [doi: 10.1016/S2213-2600(21)00380-5] [Medline: 34599903]
27. Vasileiou E, Simpson CR, Shi T, et al. Interim findings from first-dose mass COVID-19 vaccination roll-out and COVID-19 hospital admissions in Scotland: a national prospective cohort study. Lancet 2021 May 1;397(10285):1646-1657. [doi: 10.1016/S0140-6736(21)00677-2] [Medline: 33901420]
28. Andrews C, Schultze A, Curtis H, et al. OpenSAFELY: representativeness of electronic health record platform OpenSAFELY-TPP data compared to the population of England. Wellcome Open Res 2022;7:191. [doi: 10.12688/wellcomeopenres.18010.1] [Medline: 35966958]
29. Wolf A, Dedman D, Campbell J, et al. Data resource profile: clinical practice research datalink (CPRD) aurum. Int J Epidemiol 2019 Dec 1;48(6):1740-1740g. [doi: 10.1093/ije/dyz034] [Medline: 30859197]
30. Mansournia MA, Altman DG. Inverse probability weighting. BMJ 2016 Jan 15;352:i189. [doi: 10.1136/bmj.i189] [Medline: 26773001]
31. Ministry of housing, communities & local government. English indices of deprivation 2019: mapping resources. 2019 Sep 26. URL: https://www.gov.uk/guidance/english-indices-of-deprivation-2019-mapping-resources [accessed 2025-08-29]
32. Tippu Z, Correa A, Liyanage H, et al. Ethnicity recording in primary care computerised medical record systems: an ontological approach. J Innov Health Inform 2017 Mar 14;23(4):920. [doi: 10.14236/jhi.v23i4.920] [Medline: 28346128]
33. AstraZeneca. Our Transparency commitments. URL: https://astrazenecagrouptrials.pharmacm.com/ST/Submission/Disclosure [accessed 2025-08-28]

Abbreviations

CMR: computerized medical record
GDPPR: General Practice Extraction Service Data for Pandemic Planning and Research
GP: general practitioner
NHS: National Health Service
ONS: Office of National Statistics
TRE: Trusted Research Environment
VE: vaccine effectiveness

XSL•FO
RenderX

Edited by GJB Sousa; submitted 26.07.24; peer-reviewed by A Lakdawala, A Jamal, RJ Shaw; revised version received 29.07.25; accepted 31.07.25; published 27.10.25.

Please cite as:
Venkatesan S, Joy M, Jamie G, Kar D, Williams R, Fan X, Meeraus W, Tsang RSM, Taylor KS, Taylor S, Hobbs FDR, Anand SN, Robertson C, de Lusignan S
Correcting for the Inflated Adult Population Denominator in an English Nationwide Health Care Cohort: Database Analysis Study
JMIR Public Health Surveill 2025;11:e64788
URL: https://publichealth.jmir.org/2025/1/e64788
doi:10.2196/64788

© Sudhir Venkatesan, Mark Joy, Gavin Jamie, Debasish Kar, Robert Williams, Xuejuan Fan, Wilhelmine Meeraus, Ruby S M Tsang, Kathryn Taylor, Sylvia Taylor, F D Richard Hobbs, Sneha N Anand, Chris Robertson, Simon de Lusignan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 27.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Population Size Estimation of Men Who Have Sex With Men in Low- and Middle-Income Countries: Google Trends Analysis

Carly M Malburg[1,2*], MPH; Steve Gutreuter[1*], PhD; Horacio Ruiseñor-Escudero[1*], MD, PhD; Abu Abdul-Quader[1*], PhD; Wolfgang Hladik[1*], MD, PhD

[1]Division of Global HIV and TB, Centers for Disease Control and Prevention, 1600 Clifton Rd, Atlanta, GA, United States

[2]CDC Global Health Fellowship Program, Public Health Institute, Oakland, CA, United States

[*]all authors contributed equally

Corresponding Author:
Carly M Malburg, MPH
Division of Global HIV and TB, Centers for Disease Control and Prevention, 1600 Clifton Rd, Atlanta, GA, United States

Abstract

Background: Population size estimation (PSE) for key populations is needed to inform HIV programming and policy.

Objective: This study aimed to examine the utility of applying a recently proposed method using Google Trend (GT) internet search data to generate PSE (Google Trends Population Size Estimate [GTPSE]) for men who have sex with men (MSM) in 54 countries in Africa, Asia, the Americas, and Europe.

Methods: We examined GT relative search volumes (representing the relative internet search frequency of specific search terms) for "porn" and, as a comparator term, "gay porn" for the year 2020. We assumed "porn" represents "men" (denominator) while "gay porn" represents a subset of "MSM" (numerator) in each county, resulting in a proportional size estimate for MSM. We multiplied the proportional GTPSE values with the countries' male adult population (15 - 49 years) to obtain absolute size estimates. Separately, we produced subnational MSM PSE limited to countries' (commercial) capitals. Using linear regression analysis, we examined the effect of countries' levels of urbanization, internet penetration, criminalization of homosexuality, and stigma on national GTPSE results. We conducted a sensitivity analysis in a subset of countries (n=14) examining the effect of alternative English search terms, different language search terms (Spanish, French, and Swahili), and alternative search years (2019 and 2021).

Results: One country was excluded from our analysis as no GT data could be obtained. Of the remaining 53 countries, all national GTPSE values exceeded the World Health Organization's recommended minimum PSE threshold of 1% (range 1.2% - 7.5%). For 44 out of 49 (89.8%) of the countries, GTPSE results were higher than Joint United Nations Programme on HIV/AIDS (UNAIDS) Key Population Atlas values but largely consistent with the regional UNAIDS Global AIDS Monitoring results. Substantial heterogeneity across same-region countries was evident in GTPSE although smaller than those based on Key Population Atlas data. Subnational GTPSE values were obtained in 51 out of 53 (96%) countries; all subnational GTPSE values exceeded 1% but often did not match or exceed the corresponding countries' national estimates. None of the covariates examined had a substantial effect on the GTPSE values (R^2 values 0.01 - 0.28). Alternative (English) search terms in 12 out of 14 (85%) countries produced GTPSE>1%. Using non-English language terms often produced markedly lower same-country GTPSE values compared with English with 10 out of 14 (71%) countries showing national GTPSE exceeding 1%. GTPSE used search data from 2019 and 2021, yielding results similar to those of the reference year 2020. Due to a lack of absolute search volume data, credibility intervals could not be computed. The validity of key assumptions, especially who (males and females) searches for porn and gay porn, could not be assessed.

Conclusions: GTPSE for MSM provides a simple, fast, essentially cost-free method. Limitations that impact the certainty of our estimates include a lack of validation of key assumptions and an inability to assign credibility intervals. GTPSE for MSM may provide an additional data source, especially for estimating national-level PSE.

(JMIR Public Health Surveill 2025;11:e58630) doi:10.2196/58630

KEYWORDS

population size estimation; men who have sex with men; MSM; PSE; google trends; HIV; AIDS; programming and policy; internet; porn; gay porn; male adult; geriatric; linear regression; homosexuality; sensitivity analysis; World Health Organization; WHO; epidemiology

XSL•FO
RenderX

Introduction

The Joint United Nations Programme on HIV/AIDS (UNAIDS) estimated that in 2022, about 39 million people were living with HIV worldwide [1]. HIV burden is higher among men who have sex with men (MSM), people who inject drugs, sex workers, and transgender persons, which together are often described as key populations (KP) [1]. KPs and their paying or nonpaying sexual partners may account for 70% of new HIV infections worldwide, with an estimated 80% of new HIV infections outside sub-Saharan Africa (SSA) and 55% of all new infections within SSA [1,2].

Key population size estimation (PSE) is needed to estimate the number of individuals belonging to a KP in a given geographical area [3,4]. PSEs provide the denominator values to inform KP programming and policy [5]. However, PSE is a difficult field and its methods often lack rigor in design or implementation, and the many methods available reflect the lack of an acceptable gold standard [3,6]. Challenges to PSE include lack of sampling frames, mobility, and nondisclosure of KP-defining behaviors [3,4]. Further, most PSE methods produce local estimates whereas national PSE estimates are often obtained through "expert opinion," simple projection, or modeling and less often through national-level empirical data such as direct survey questions or the network scale-up method, both used in general population-based surveys [6,7]. Direct survey questions about KP-defining traits experience reporting bias and require a major effort unless they can be added to an already planned general population survey. The frequent lack of reliable national-level PSE constitutes an even larger challenge compared with the availability of local PSE and complicates national, regional, and global HIV estimation work [3,8-10].

The rise of the internet facilitates web-based activities to improve public health, including in the field of digital epidemiology and infoveillance [11]. Recently, a new PSE method using Google Trends (GT) internet search data was proposed in a proof of concept paper by Card et al [12] GT is a free cloud-based app that displays the relative frequency of user-specified Google search terms as trends across time and user-selected geographical areas [12-14]. Card et al [12] used GT and Canadian census data to estimate the local PSE of MSM in urban and rural locations throughout Canada. Card et al [12] related search terms presumed to be representative of MSM ("gay porn") to that presumed to be representative of the general (male) population ("porn"). By relating these 2 sets of values, Card et al [12] estimated the relative size of MSM in these Canadian towns. To date, no other published PSE exists using this method.

The literature on pornography consumption by sex and sexual orientation is limited and often the MSM population is not represented. However, a major porn website reported that about a third of its visitors globally in 2021 were reportedly women [15,16]. Further, women, regardless of sexual orientation, may also watch gay porn, possibly in substantial numbers [17]. Beyond this, we found no meaningful gray literature or peer-reviewed articles about internet pornography consumption in low- and middle-income countries (LMICs) or pornography consumption by MSM in LMICs. We are also not aware of (gray) literature about the proportion of heterosexual and homosexual men searching Google for (gay) porn in LMICs.

We expanded the literature search to include high-income settings. A study conducted in the United States reported that more men than women consume pornography (92%:68%, respectively) over the span of a year [18]. The study did not report the type of pornography consumed or disaggregate male respondents by sexual orientation or practice [18]. A separate study from Norway with a sample of some 2300 male and female participants suggested that more men than women consume some pornography (94% of men and 68% of women) [19]. However, only 5% (n=106) of participants identified as gay/lesbian/bisexual, no breakdown of sexual orientation by sex was given, and no information on the type of pornography consumed by participants was available [19].

The aim of this study was to examine the potential utility of using GT data to obtain MSM PSE in selected LMICs.

Methods

Preliminary Literature Search

A nonsystematic literature search was conducted to better understand the behavior of pornography consumption of the general population and sexual minorities, by sex, as well as the relative frequencies with which these populations search for (gay) porn in general (via Google) or by directly accessing specific porn sites.

Selection of Countries

We analyzed GT data for a selected set of 54 countries that receive support from the US President's Emergency Plan for AIDS Relief, the US Government's initiative to support global HIV responses, for which information on MSM PSE has been sought [2,20]. These countries are located in SSA (n=29), Asia (n=13), the Americas (n=11), and Ukraine (Tables 1 and 2).

XSL•FO
RenderX

Table . National men who have sex with men (MSM) population size estimation (PSE) for US President's Emergency Plan for AIDS Relief supported countries (n=53) using Google Trends (GT) data for the year 2020[a,b].

Region and country		GT (number of MSM), n	GT, %	UNAIDS[c] GAM[d] regional %, median (IQR)[e]	UNAIDS KP[f] Atlas, %
East Africa[g]				1.67	
	Burundi	48,500	1.77		0.34
	Ethiopia	365,000	1.28		—[h]
	Kenya	276,000	1.99		0.24
	Rwanda	51,300	1.54		0.15
	Tanzania	243,000	1.73		0.35
	Uganda	154,000	1.47		0.23
Southern Africa[i]				1.67	
	Angola	106,000	1.44		—
	Botswana	13,000	2.12		0.43
	Eswatini	4500	1.57		1.38
	Lesotho	10,000	1.71		1.05
	Malawi	52,500	1.16		0.94
	Mozambique	134,000	1.87		0.22
	Namibia	16,500	2.60		—
	South Africa	393,000	2.46		1.94
	Zambia	51,800	1.18		0.15
	Zimbabwe	53,000	1.64		0.71
West Central Africa[j]				1.28 (IQR 0.45 - 1.50)	
	Benin	34,000	1.18		0.20
	Burkina Faso	88,000	1.80		0.07
	Cameroon	148,000	2.29		0.11
	Cote d'Ivoire	166,000	2.68		0.90
	DRC	33,000	1.66		0.98
	Ghana	112,000	1.40		0.69
	Liberia	22,600	1.83		6.04
	Mali	70,500	1.55		0.09
	Nigeria	614,000	1.26		0.49
	Senegal	73,600	1.94		1.38
	Sierra Leone	27,000	1.36		0.16
	Togo	53,100	2.65		0.30
Asia[k]				1.63 (IQR 0.26 - 3.10)	
	Burma	664,000	4.53		1.72
	Cambodia	258,000	5.67		1.93
	India	6,460,000	1.18		0.06
	Indonesia	1,180,000	1.61		1.03
	Kazakhstan	137,000	2.99		1.35
	Kyrgyz Rep.	53,000	3.10		0.99
	Lao PDR	53,000	2.73		2.96

XSL•FO
RenderX

Region and country		GT (number of MSM), n	GT, %	UNAIDS[c] GAM[d] regional %, median (IQR)[e]	UNAIDS KP[f] Atlas, %
	Nepal	83,000	1.19		0.86
	PNG	31,000	1.30		1.58
	Tajikistan	52,000	2.14		0
	Thailand	215,000	1.25		3.08
	Philippines	1,260,000	4.27		2.33
	Viet Nam	1,953,000	7.46		0.98
Europe					
	Ukraine	366,000	3.48	2.11 (IQR 1.75 - 2.49)	1.71
Caribbean[l]				2.71	
	Dominican Rep.	124,000	4.26		4.90
	Guyana	8200	3.60		1.45
	Haiti	108,000	3.60		1.03
	Jamaica	24,000	2.91		5.15
	Trinidad and Tobago	11,000	3.04		—
Central and South America[m]				3.37	
	Brazil	2,960,000	5.18		3.50
	El Salvador	85,000	5.20		3.31
	Guatemala	245,000	5.09		2.42
	Honduras	147,000	5.32		1.48
	Nicaragua	114,000	6.32		1.97
	Panama	81,000	7.23		2.65

[a]These estimates are for descriptive purposes only, to examine issues related to the potential utility of the method proposed by Card et al [12]. They represent the MSM population national population size estimates (percentage of MSM) for the year 2020. The percentage of MSM was calculated by taking the average relative search volume score produced by Google Trends for "gay porn" and dividing it by the average relative search volume score produced by Google Trends for "porn." MSM population size estimate (number of MSM) was calculated by taking the percentage of MSM population size estimate and dividing it by the total male population (ages 15 - 49 years). Key populations (KPs) Atlas percentage of MSM population size estimate was calculated by dividing the absolute MSM population size estimate taken from the United Nations Programme on HIV/AIDS (UNAIDS) KPs Atlas dashboard by the total adult male population (ages 15 - 49 years), and then multiplying by 100. The absolute value difference was calculated by subtracting the GT absolute MSM population size estimate value from the KPs Atlas MSM population size estimate absolute value. All absolute values under 10,000 are rounded to the nearest 100. All other absolute values are rounded to the nearest 1000. UNAIDS Global AIDS Monitoring system (GAM) values are regional values transcribed from the UNAIDS open-source Spectrum 6 guide. The countries used to create these regions and respective values may not be in full alignment with the countries included in the population size estimate analysis, therefore direct 1:1 comparisons should not be made. Max:Min ratio: The ratio based on the largest and smallest PSE % value in each region.

[b] Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[c]UNAIDS: United Nations Programme on HIV/AIDS.

[d]GAM: Global AIDS Monitoring system.

[e]IQR values were included for available regions. Regions without an IQR listed did not have one available.

[f]KP: key population.

[g]Max:Min ratio: 1.6 (GT) and 2.3 (UNAIDS KP).

[h]Not available (data missing for the country).

[i]Max:Min ratio: 2.2 (GT) and 12.9 (UNAIDS KP).

[j]Max:Min ratio: 2.3 (GT) and 86.3 (UNAIDS KP).

[k]Max:Min ratio: 6.3 (GT) and 51.3 (UNAIDS KP).

[l]Max:Min ratio: 1.5 (GT) and 5 (UNAIDS KP).

[m]Max:Min ratio: 1.2 (GT) and 2.4 (UNAIDS KP).

XSL•FO
RenderX

Table . Regional median Google Trends Population Size Estimate, United Nations Programme on HIV/AIDS (UNAIDS) Global AIDS Monitoring system (GAM), and key populations (KP) Atlas for men who have sex with men (MSM) populations for the year 2020.[a]

Region	Median regional percentage MSM population size estimation[b]		
	GT[c], %	UNAIDS GAM, %	UNAIDS KP Atlas, %
Eastern Africa	1.64	1.67	0.24
Southern Africa	1.68	1.67	0.83
West Central Africa	1.73	1.28	0.40
Asia	2.86	1.63	—[d]
Europe	2.86	2.11	1.47
Caribbean	3.60	2.71	3.17
Central & South America	5.26	3.37	2.54

[a]Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[b]Google Trends (GT) and KP Atlas regional estimates only include estimates from included countries with available data (Table 1). UNAIDS GAM data separate regions differently and include countries that vary from our GT or the KP Atlas regional data: UNAIDS GAM includes eastern and southern Africa in 1 estimate and separates Asia and Europe into 2 estimates (1.63% for Asia and the Pacific, 2.11% for Eastern Europe and Central Asia). Region names were not adjusted in the above table to align with GAM data.

[c]GT: Google Trends.

[d]Not available.

Ethical Considerations

No ethics or review board approval or informed consent was obtained or applicable for this work. All data used in this paper are anonymous, aggregate, and publicly available and sourced.

GT-Based Population Size Estimation

GT provides results based on exact search terms, unlike the "topical" search results that Google's main search engine provides. GT does not provide absolute search frequency values; instead, GT offers relative search volume (RSV) values across time (eg, 52 wk) in a specified space (eg, Kenya), ie, it normalizes search frequencies for specific search terms (eg, porn) to a range from 0 to 100, where a search term's maximum frequency (for the specified geographic area and during the specified time frame) is set at 100 and 0 reflects no search for that term [11,13,14]. Importantly, GT allows users to add "comparator" terms (eg, gay porn) next to the main term (eg, porn); the RSV values for such comparator terms are normalized against the main term's RSV values [13,21]. For the purpose of PSE calculation, the main term "porn" may represent all men whereas the comparator term "gay porn" may be viewed as a subset of men who represent the subpopulation of gay men or MSM. To generate an MSM PSE from the RSV values we divide the comparator RSV value (gay porn) by the larger same-time, same-place RSV value (porn).

National Size Estimates

PSE data collection was carried out through GT's application [13]. We applied this analytic approach for the year 2020 using "porn" and "gay porn" as the main and comparator search terms for each of the 54 countries. The time period for data collection was set as the year 2020, the most recent year for which we could obtain all necessary data for this analysis. Weekly RSV values for "porn" and "gay porn" for the year 2020 were exported, summed, and proportional size estimates obtained. For example, for Botswana, the average of the weekly RSV values for "porn" was 78.3, the corresponding average for "gay porn" was 1.66 and the proportional PSE was therefore calculated as 1.66/78.3=2.1%. This was repeated for all countries. We then calculated the absolute Google Trends Population Size Estimate (GTPSE) by multiplying the proportional GTPSE by the total male population aged 15 - 49 years in each country, the most used age range for KPs. The sizes for countries' 15 - 49 year-old male general population in 2020 were obtained through Spectrum (version 6.1, Avenir Health).

Local Size Estimates

GT data can be restricted to subnational areas. Separately from national estimates, for each country, we also attempted to obtain local GTPSE for the political (or, if different, commercial) capital city. Where data were unavailable for the political or commercial capital city, we used data from the district that contained the capital city. The calculation to obtain relative GTPSE was then the same as for the national level. We did not produce absolute subnational GTPSE.

Consistency of GTPSE Results With WHO-Recommended Minimum Estimate

We assessed whether the GTPSE results met the World Health Organization (WHO) and UNAIDS recommendation that national MSM PSE should represent at least 1% of the general adult male population [22,23].

Comparability

We compared the country-level GTPSE against 2 reference data sources used by UNAIDS: the KP Atlas database and the Global AIDS Monitoring system (GAM) [22,24,25]. The KP Atlas database stores countries' self-reported absolute MSM size estimates using a wide range of PSE methods, often projected up to national scale from local estimates, with primary data collected over different periods of time. Proportional KP Atlas

XSL•FO
RenderX

PSE values were computed by dividing the absolute MSM PSE values from the KP Atlas over the male general population (15 - 49 years). UNAIDS' GAM is a global data warehousing system that informs policy and facilitates monitoring, including KP size estimates. Using GAM data, UNAIDS curated a table with regional relative MSM PSE (median and IQR) deemed reasonable.

Covariates Potentially Affecting GTPSE

Overview

We examined the potential effect of select covariates on the relative GTPSE values by performing regression analysis for each covariate. The country-specific covariates we examined included internet penetration, urbanization, stigma, and criminalization of homosexuality. The covariate data are provided in Table S1 in Multimedia Appendix 1; these data were not used to adjust GTPSE values.

Internet Penetration and Urbanization

Internet penetration data were extracted from the World Development Indicators database through the World Bank and the Internet World Statistics database, indicating the percentage of each country's total population with access to the internet. Urbanization data were obtained from the World Development Indicators database through the World Bank, indicating the percent of the total population in each country considered urban [26,27].

Stigma

Country-level stigma values were extracted from the Global Acceptance Index [28]. This index was developed using computer modeling informed by responses to questions that measure attitudes toward lesbian, gay, bisexual, transgender, or intersex people from 11 different global surveys to create a stigma score in 175 countries toward lesbian, gay, bisexual, transgender, or intersex persons. The system scores countries on a scale of 1 to 10; higher scores indicate less stigma [28].

Criminalization

The State-Sponsored Homophobia International Lesbian, Gay, Bisexual, Trans, and Intersex Association report was used to evaluate the effects of criminalization of homosexual orientation or behavior on GTPSE [28]. The report classifies countries based on their level of legal protection or criminalization of sexual orientation and same-sex sexual acts. These classifications, ranging from most severe to most protected, include the death penalty, up to lifelong imprisonment, up to 8 years imprisonment, de facto criminalization, no criminalization or legal protections, limited protections, employment protections, broad protections, and constitutional protections. We converted these classifications into a quantitative ranking ranging from +4 to −4. The most severe classification (death penalty) was assigned the rank value "+4" and descended to the least severe/most protective classification (constitutional protection) with a rank value "−4."

Sensitivity Analysis

Using a subset (n=14) of the 53 countries we performed 3 sensitivity analyses at the national level. The 14 countries were randomly selected among countries with prominent languages being French, Spanish, or Swahili. The first sensitivity analysis probed the effect of select non-English search languages. The 14 countries comprised 4 using Swahili (Kenya, Tanzania, Uganda, Democratic Republic of Congo [DRC]), 5 using French (Cote d'Ivoire, Senegal, Cameroon, Mali, Haiti), and 5 using Spanish (Dominican Republic, Panama, El Salvador, Nicaragua, Honduras) as their national/dominant language. We generated GTPSE using search terms in Swahili ("ngono" and "ngono za mashoga"), French ("porno" and "porno gay"), and Spanish ("porno" and "porno gay") and compared them to the original relative GTPSE values. Using the same 14 countries, the second sensitivity analysis probed the effect of different search terms in English on GTPSE, that is, "sex," "gay sex" as well as "sex," "anal sex" and compared them to the original GTPSE (porn and gay porn). The third sensitivity analysis probed the effect of using different calendar years, ie, (2019 [pre-COVID] and 2021) and compared them to the original 2020 GTPSE values, using the original English language search terms.

Results

GTPSE and Comparability

Of the 54 countries examined, 1 (South Sudan), was omitted for lack of RSV values. All remaining 53 countries had GTPSE exceeding 1% (Table 1), similar to GAM values (all exceeding 1% as well) and compared with KP Atlas values where 24 out of 53 (45%) countries showed values above 1%. GTPSE ranged from 1.16% to 7.46% (median 1.99%, IQR 1.54% - 3.48%), compared with 0.06% to 6.04% (median 0.99%, IQR 0.34 - 1.93%) in the KP Atlas, and 1.38% to 2.82% in GAM regions. In 48 out of 53 (91%) countries, relative GTPSE exceeded estimates in the KP Atlas values; KP Atlas values were larger in 5 countries (DRC, Liberia, Lao People's Democratic Republic [PDR], Thailand, and Jamaica). Absolute differences between GTPSE and KP Atlas ranged from −312,900 (Thailand) to 6,221,800 (India). Table 2 displays regional median GTPSE, ranging from 1.64% (East Africa) to 5.26% (Central/South America), larger in all regions than the corresponding KP Atlas values and largely similar to GAM values in most regions. Table 1 also displays the ratios between the largest and smallest country-level %PSE for each region, separately for GT and KP Atlas values. While substantial variability is seen in all regions and for both data sources (GT and KP Atlas), in all regions the observed heterogeneity was consistently higher for KP Atlas values compared with GT values.

Local GTPSE pertaining to political or commercial capitals or the larger sub-national areas encompassing these are displayed in Table 3. We could obtain local estimates for 51 out of 53 (96%) countries' capital cities; GT did not provide data for Nairobi (Kenya) and Kathmandu (Nepal). Among the 51 cities with estimates, the GTPSE ranged from 0% to 13% (median 2.2%); most cities' estimates (44/51, 86%) exceeded 1%. Five cities yielded noncredible GTPSE values of 0%, including Bujumbura (Burundi), Dodoma (encompassing Dar es Salaam, Tanzania), Ouagadougou (Burkina Faso), Monrovia (Liberia), and Vientiane (Laos PDR). Of the 44 subnational GTPSE with

XSL•FO
RenderX

values >1%, 18 (41%) were below the same-country national GTPSE.

XSL•FO
RenderX

Table . Reported local men who have sex with men (MSM) Google Trends Population Size Estimate (GTPSE) (n=53) in the year 2020.[a]

Region and country		Local area[b]	Relative national GTPSE, %	Relative local GTPSE, %	Absolute percentage difference national and local GTPSE, %
East Africa					
	Burundi	Bujumbura	1.77	0	−1.77
	Ethiopia	Addis Ababa	1.28	1.30	0.02
	Kenya	Nairobi	1.99	—[c]	—
	Rwanda	Kigali	1.54	1.70	0.16
	Tanzania	Dodoma	1.73	0	−1.73
	Uganda	Kampala	1.47	1.60	0.13
Southern Africa					
	Angola	Luanda	1.44	2.04	0.60
	Botswana	Gaborone	2.12	0	−2.12
	Eswatini	Mbabane	1.57	1.48	−0.09
	Lesotho	Maseru	1.71	1.01	−0.70
	Malawi	Lilongwe	1.16	2.24	1.08
	Mozambique	Maputo	1.87	2.04	0.17
	Namibia	Windhoek	2.60	2.53	−0.07
	South Africa	Johannesburg (Gauteng)	2.46	0.99	−1.47
	Zambia	Lusaka	1.18	1.56	0.38
	Zimbabwe	Harare	1.64	1.56	−0.08
West Central Africa					
	Benin	Littoral (Cotonou)	1.18	4.11	2.93
	Burkina Faso	Centre (Ouagadougou)	1.80	0	−1.80
	Cameroon	Littoral (Douala)	2.29	2.47	0.18
	Cote d’Ivoire	Abidjan	2.68	1.01	−1.67
	DRC	Kinshasa	1.66	2.04	0.38
	Ghana	Accra	1.40	1.42	0.02
	Liberia	Monrovia	1.83	0	−1.83
	Mali	Bamako	1.55	2.93	1.38
	Nigeria	Abuja (Federal Capital Terriorty)	1.26	1.44	0.18
	Senegal	Dakar	1.94	2.85	0.91
	Sierra Leone	Freetown	1.36	1.01	−0.35
	Togo	Lome	2.65	2.04	−0.61
Asia					
	Burma	Yangon (Yangon Region)	4.53	4.79	0.26
	Cambodia	Phnom Penh	5.67	5.43	−0.24
	India	New Delhi (Uttar Pradesh)	1.18	1.15	−0.03
	Indonesia	Jakarta	1.61	2.20	0.59
	Kazakhstan	Almaty (Almaty Region)	2.99	5.52	2.53

XSL•FO
RenderX

Region and country		Local area[b]	Relative national GTPSE, %	Relative local GTPSE, %	Absolute percentage difference national and local GTPSE, %
	Kyrgyz Rep.	Bishkek	3.10	3.09	−0.01
	Lao PDR	Vientiane	2.73	0	−2.73
	Nepal	Katmandu/Kantipur	1.19	—	—
	PNG	Port Moresby	1.30	1.01	−0.29
	Tajikistan	Dushanbe	2.14	1.01	−1.13
	Thailand	Bangkok	1.25	3.24	1.99
	Philippines	Manila	4.27	5.51	1.24
	Viet Nam	Hanoi	7.46	4.56	−2.90
Europe					
	Ukraine	Kyiv	3.48	4.14	0.66
Caribbean					
	Dominican Rep.	Santo Domingo	4.26	3.99	−0.27
	Guyana	Georgetown	3.60	3.09	−0.51
	Haiti	Port-au-Prince	3.60	3.09	−0.51
	Jamaica	Kingston (St. Andrew Parish)	2.91	3.38	0.47
	Trinidad and Tobago	Port of Spain	3.04	13	9.96
Central and South America					
	Brazil	São Paulo (State of São Paulo)	5.18	5.92	0.74
	El Salvador	San Salvador	5.20	5.73	0.53
	Guatemala	Guatemala City (Guatemala Department)	5.09	4.91	−0.18
	Honduras	Tegucigalpa (Comayagua)	5.32	8.13	2.81
	Nicaragua	Managua	6.32	5.31	−1.01
	Panama	Panama City	7.23	7.44	0.21

[a]Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[b]Local MSM GTPSE for 53 countries for the year 2020 was calculated by restricting the geographic entity to the desired capital city or commercial hub. Where Google Trends (GT) did not provide data for a given city, we substituted the place name with the largest city by population or by district that had data available in GT. This is noted by listing what was available in GT in parenthesis next to the capital city. Kenya and Nepal were excluded from this analysis due to insufficient regional data available in GT.

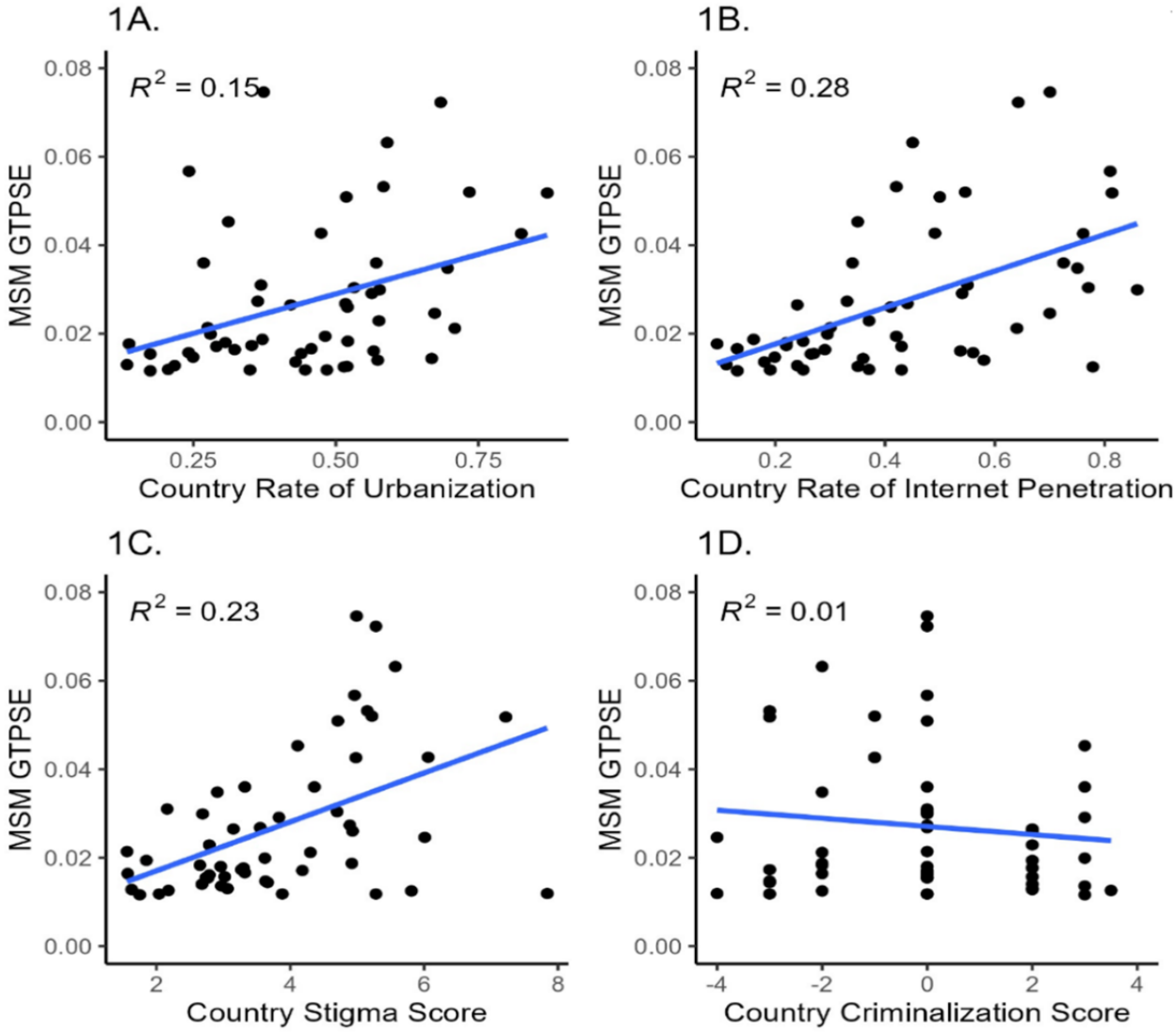
[c]Not available (data missing for that country).

Effect of Covariates

Figure 1A-D displays the correlations between national-level GTPSE and urbanization, internet penetration, stigma, and criminalization. Coefficients ranged from 0.01 (criminalization) to 0.28 (internet penetration).

XSL•FO
RenderX

Figure 1. The linear relationship between the Google Trends national population size estimates and the rate of urbanization in each country (n=53). (A) The linear relationship between the Google Trends national population size estimates and the rate of urbanization in each country (n=53). (B) The linear relationship between the Google Trends national population size estimates and the rate of internet penetration in each country (n=53). (C) The linear relationship between the Google Trends national population size estimates and the level of stigma against LGBTQ+ persons in each country (n=53). (D) The linear relationship between the Google Trends national population size estimates and the degree of criminalization against men who have sex with men population in each country (n=53). LGBTQ+: lesbian, gay, bisexual, transgender, queer, and other identities; MSM: men who have sex with men; GTPSE: Google Trends Population Size Estimate.



Sensitivity Analysis

Table 4 displays how the GTPSE generated from the alternative search terms compares to the original search term GTPSE. In most countries "Porn/Gay Porn" produced higher PSE values compared with "sex/anal sex" (13/14, 93%) as well as compared with "sex/gay sex" (12/14, 86%). For "sex/gay sex," all 14 countries produced estimates exceeding 1%. For "sex/anal sex", 3 out of 14 (21%) countries did not produce estimates reaching the 1% threshold, including Mali for which zero search results were reported for "anal sex."

XSL•FO
RenderX

Table . Sensitivity analysis using alternative search terms in Google Trends to calculate national population size estimations (PSEs) for select US President's Emergency Plan for AIDS Relief countries (n=53) in 2020.[a]

	Original GTPSE[b]	SA alternate search term GTPSE[c]			
	Porn/gay porn PSE	Sex/gay sex PSE	Absolute percentage difference	Sex/anal sex PSE	Absolute percentage difference
Country, %					
Kenya	1.99	1.37	0.62	1.37	0.62
Tanzania	1.73	1.46	0.27	3.54	−1.81
Uganda	1.47	1.38	0.09	1.26	0.21
DRC	1.66	1.55	0.11	1.15	0.51
Cameroon	2.29	1.28	1.01	0.90	1.39
Mali	1.55	1.65	−0.10	0	1.55
Cote d'Ivoire	2.68	1.90	0.78	1.74	0.94
Senegal	1.94	1.50	0.44	0.88	1.06
Haiti	3.60	2.60	1	2.83	0.77
Dominican Rep.	4.26	3.36	0.90	1.83	2.43
Panama	7.23	5.17	2.06	3.71	3.52
El Salvador	5.20	5.34	−0.14	4.19	1.01
Nicaragua	6.32	7.10	−0.78	4.82	1.50
Honduras	5.32	4.85	0.47	2.96	2.36
Median (IQR)	2.49 (1.78-4.97)	1.78 (1.47-4.48)	0.45 (0.10-0.87)	1.79 (1.18-3.40)	1.03 (0.66-1.54)

[a]Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[b]GTPSE: Google Trends Population Size Estimate.

[c]Alternative search terms were chosen based on words that represented the general male population and men who have sex with men subset population in each country (n=53) in the year 2020.

Table 5 shows how GTPSE was generated using alternative language terms compared with the original GT search terms. For Swahili, only 1 country yielded a PSE in that language. All countries using French (n=5), or Spanish (n=5) search terms yielded estimates, all exceeding 1%. All alternative language estimates were lower than the original "porn/gay porn" PSE values.

XSL•FO
RenderX

Table . Sensitivity analysis using alternate national language searches in Google Trends to calculate national population size estimation for select US President's Emergency Plan for AIDS Relief countries (n=14) in 2020.[a]

Language and country		Original GTPSE[b] (English), %	Alternate language term GTPSE, %[c]	Absolute percentage difference, %
Swahili				
	Kenya	1.99	0	1.99
	Tanzania	1.73	0.52	1.21
	Uganda	1.47	0	1.47
	DRC	1.66	0	1.66
French				
	Cameroon	2.29	1.36	0.93
	Mali	1.55	1.07	0.48
	Cote d'Ivoire	2.68	1.35	1.33
	Senegal	1.94	1.28	0.66
	Haiti	3.60	2.23	1.37
Spanish				
	Dominican Rep.	4.26	2.56	1.70
	Panama	7.23	5.14	2.09
	El Salvador	5.20	4.36	0.84
	Nicaragua	6.32	4.13	2.19
	Honduras	5.32	4.07	1.25

[a]Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[b]GTPSE: Google Trends Population Size Estimate.

[c]Alternative language search terms included "ngono/ngono za mashoga" (Swahili), "porno/porno gay" (French), "porno/porno gay" (Spanish).

Table 6 displays how GTPSE generated for alternative years (2019 and 2021) compared with the original 2020 GT searches. All 14 countries in both years produced estimates exceeding 1%. No large discrepancies in PSE between the years were observed; 13 out of 14 in 2019 values were larger than the 2020 values whereas the 2021 values were largely similar to the 2020 values.

XSL•FO
RenderX

Table . Sensitivity analysis for men who have sex with men population size estimates for select US President's Emergency Plan for AIDS Relief supported countries (n=14) using Google Trends data in years 2019 and 2021 compared with the year 2020.[a, b]

	2019 PSE[c], %	2020 PSE, %	2021 PSE, %
Kenya	2.37	1.99	1.99
Tanzania	1.96	1.73	1.85
Uganda	1.73	1.47	1.69
DRC	1.95	1.66	1.60
Cameroon	2.70	2.29	2.25
Mali	2.30	1.55	2.17
Cote d'Ivoire	2.52	2.68	2.23
Senegal	2.54	1.94	1.90
Haiti	4.33	3.60	2.92
Dominican Republic	4.91	4.26	4.34
Panama	9.36	7.23	6.74
El Salvador	6.19	5.20	4.77
Nicaragua	7.31	6.32	4.93
Honduras	6.79	5.32	5.51

[a]2019 and 2021 values were computed in the same way as the reference 2020 estimates.

[b]Absolute values are not provided as Google Trends does not provide absolute search frequency values.

[c]PSE: population size estimation.

Discussion

Principal Findings

Our analysis suggests that national-level MSM GTPSE is feasible in almost all countries. Importantly, all estimates appeared plausible, that is, they exceeded the WHO/UNAIDS suggested minimum threshold of 1%. Heterogeneity of GTPSE across same-region countries was pronounced within all regions yet smaller than the ratios based on the UNAIDS KP Atlas values which contained numerous PSE values well below the 1% threshold.

Our analysis draws on several strengths. We successfully applied the GTPSE method to many low and middle-income countries, suggesting that GTPSE appears to have wide geographic applicability. We compared the values against 2 PSE data sources at UNAIDS, assessed the potential effect of various covariates on GTPSE values, and conducted a sensitivity analysis with varying English search terms, non-English search languages, and different calendar years. Google is the dominant search engine in all countries covered in this analysis, with a market share ranging between 84% and 99% (data shown in Table S1 in Multimedia Appendix 1) [27]. Although no absolute search volume data were available to us, searches for "porn" globally were among the top 20 search terms in 2023 with about 65 million searches globally each month according to one source [29] although this is still well behind the largest porn site-specific searches. GTPSE may emerge as another example of digital public health and epidemiology that includes real-time surveillance of disease outbreaks [30], assessing the impact of global public health days [31], informing health and health policy research [32], or understanding spatiotemporal patterns of dry eye disease [33].

While most local estimates were plausible (>1%), 14% (n=7) did not reach the WHO/UNAIDS minimum threshold, and 2 more locations did not produce a GTPSE value at all due to lack of GT data and how GT organized the subnational data despite some of the affected cities' large population sizes. This is not an uncommon finding, as other PSE methods in active use typically do not meet the WHO/UNAIDs minimum threshold. For a few other country or commercial capital cities with no direct GT data available, such as Johannesburg (South Africa), we could obtain a subnational estimate using the larger district or province within which the city (eg, Johannesburg and Pretoria) are located. This may limit the utility and comparability of such local estimates. About one-third of the local (relative) estimates did not reach or exceed the same country national level estimates, somewhat contrary to our expectation that rural-to-urban migration among MSM may be more pronounced than that of other men and so yielding higher GTPSE values [9]. In Card et al's [12] study on Canadian towns and cities the estimates ranged from 2% to 4% compared with 0% to 13% among the local estimates, whereas the Canadian national estimate was 2.8% compared with 1.2% - 7.5% across all countries we examined. While not a limitation, it is worth noting that weekly RSV data varied widely (data not shown), confirming the recommendation to use GT data for size estimation only over longer time periods, such as a full calendar year.

Limitations

Like most PSE methods, GTPSE has limitations. In particular, the assumptions underlying the GTPSE method deserve close

XSL•FO
RenderX

scrutiny: straight men only search for porn, MSM only search for gay porn, MSM and straight men search for (gay) porn in equal proportions, and women do not search for (gay) porn at all or do not affect the generated GTPSE for MSM. Violations of these assumptions will result in bias if they affect RSV for porn and gay porn to differing extents, hence altering the proportion of porn searches that are directed at gay porn. While the literature from LMIC settings on this topic is very sparse, reports and literature from high-income settings suggest that gay porn is also consumed by heterosexual men and women, suggesting that some bias may be present. Complicating speculations about the magnitude and direction of bias is the fact that specific porn websites' user statistics may not accurately reflect searches for (gay) porn on Google. Women's search behavior on Google regarding gay porn may increase or decrease the GTPSE estimates depending on the frequency relative to searches for just porn.

Regrettably, Google does not provide access to its algorithm generating the RSV data nor can users filter GT searches by age or gender. An inherent limitation in using GT data includes the lack of deduplication in the search data (although repeated searches by the same user within a short time period are not counted multiple times by Google) and the lack of absolute search volume data. Not having access to the absolute search volume data impedes the computation of uncertainty intervals (which in most national settings may be expected to be small due to the large search volumes involved). However, absolute search volume information may eventually be made available by Google and is already offered to some extent by select third-party companies. Absolute search volume data may also inform the choice of search language and even search terms and may facilitate composite GTPSE metrics by incorporating multiple GTPSE metrics stemming from different language search terms. Restricting GTPSE-relevant data to male users may further refine GTPSE values by excluding female users, a limitation our analysis could not overcome. VPN (virtual private network) also has the potential to introduce errors if users select a country other than their place of residence. The adoption of VPN may vary considerably across time and by country, and, among US President's Emergency Plan for AIDS Relief countries. According to one industry website in 2020, VPN was highest in Ukraine (7.9%) and lowest in Kenya (0.5%) [34]. Taken together, these limitations constitute a major source of uncertainty about the bias and precision of GTPSE. For that reason, GPTSE should be regarded as an approximate reference value. Clearly, they do not attain the rigor or transparency of statistically principled estimation from accurately measured data, which the currently best available PSE methods do offer. Additionally, GTPSE may not be feasible for a few countries, perhaps due to poor or little data availability on search terms and frequencies.

GTPSE seems infeasible for size estimation among transgender persons, sex workers, or people who inject drugs. Unlike (gay) porn, where the search is about a web-based product (visual depictions of porn), searches for sex work or clients, transgenderism, or injecting drug use are not directly tied to the internet, and may exhibit a more variable search terminology, and may lack fitting "denominator" search terms (analogous to "porn").

Overall, the GTPSEs often were substantially higher than the KP Atlas estimates but were more closely aligned with the reported GAM regional estimates. The KP Atlas estimates are based on a broad range of PSE methods typically generating local PSE that may or may not be projected to national scale, or summed or averaged across multiple localities, and may refer to various time points (calendar years) and various age ranges. Many KP Atlas based MSM PSE were implausibly low (<1%), suggesting that substantial differences to GTPSE may often be due to KP Atlas underestimates. The regional GAM estimates are based on a more curated database of PSE after excluding estimates with subpar quality and hence of perhaps more trustworthy quality [22]. However, GAM regions do not exactly overlap with the regions we used for GTPSE and the KP Atlas estimates.

The national MSM GTPSE values were robust against varying levels of urbanization, internet penetration, stigma, and criminalization or protection of homosexuality, negating the need for adjustment and increasing comparability across different settings. The largest influence was seen with internet penetration which can be expected to increase over time. In the sensitivity analysis, the largest differences to the original GTPSE values were seen using alternate English language search terms. Among the 14 examined countries, almost half (43%) of the alternate estimates were below 1% and hence considered implausibly low. This indicates that search term selection is important, especially for comparison across time and space. Further exploration may be warranted to evaluate if country or region-specific English or non-English slang terms may produce plausible estimates; however, the limited sensitivity analysis suggests that "Porn/Gay Porn" may be dependable and consistently produces plausible values. The use of similar search terms in French, Spanish, and Swahili yielded universally lower results; Swahili, not a nationally dominant language in most countries, appears particularly unsuitable as it frequently produced 0% PSE values. As most countries display prominent non-English language use, countries may want to consider using the predominant language (used for web searches) when applying this method while considering any language's geographic scope in-country. The results also appeared robust across time (two years affected by the COVID pandemic plus 1-year pre-COVID) as the 2 adjacent years produced plausible and (same country) consistent results. The lack of uncertainty intervals however impeded a more meaningful interpretation of the results from the sensitivity analyses.

Conclusions

Generating national-level PSEs for KPs is challenging for many countries. GTPSE is a simple method with the potential to address this problem efficiently without the need of additional resources. However, the lack of validation of key assumptions and the inability to generate credibility intervals suggest important uncertainty regarding the accuracy and precision of the estimates. Additional research, such as expanding or building on our sensitivity and covariate analysis, to address or better

XSL•FO
RenderX

understand these limitations may further improve the quality and utility of GTPSE for MSM in LMICs.

Acknowledgments

We wish to acknowledge all authors who contributed to this paper. WH and AAQ contributed to the study concept and design. All authors substantially contributed to the study's conduct, or to data analysis and interpretation. All authors contributed to writing, reviewing, or editing the manuscript. We also wish to acknowledge all colleagues within the Center for Disease Control and Prevention who offered insights and alternative perspectives during the analysis and writing of this paper.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the funding agencies. This research publication has been supported by the President's Emergency Plan for AIDS Relief (PEPFAR) through the Centers for Disease Control and Prevention (CDC) under the terms of project number 0900f3eb820619c0. This publication was also supported by Cooperative Agreement number NU2GGH002093 between the Centers for Disease Control and Prevention and the Public Health Institute.

Data Availability

All data generated or analyzed during this study are publicly accessible data (with the exception that only aggregate data can be obtained through Google Trends). Links to where the data can be accessed are included in this published article and its supplementary information files. Further, all relative search volume data points can be reproduced through Google Trends.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary Table 1.
[XLSX File, 18 KB - publichealth_v11i1e58630_app1.xlsx]

References

1. Global HIV & AIDS statistics — fact sheet. UNAIDS. URL: https://www.unaids.org/en/resources/fact-sheet [accessed 2024-12-31]
2. Indicators and questions for monitoring progress on the 2021 Political Declaration on HIV and AIDS — Global AIDS Monitoring 2025. UNAIDS. 2024. URL: https://www.unaids.org/en/resources/documents/2024/global-aids-monitoring-guidelines [accessed 2024-12-31]
3. Neal JJ, Prybylski D, Sanchez T, Hladik W. Population size estimation methods: searching for the holy grail. JMIR Public Health Surveill 2020 Dec 3;6(4):e25076. [doi: 10.2196/25076] [Medline: 33270035]
4. Stone J, Mukandavire C, Boily MC, et al. Estimating the contribution of key populations towards HIV transmission in South Africa. J Int AIDS Soc 2021 Jan;24(1):e25650. [doi: 10.1002/jia2.25650] [Medline: 33533115]
5. Viswasam N, Lyons CE, MacAllister J, et al. The uptake of population size estimation studies for key populations in guiding HIV responses on the African continent. PLoS One 2020;15(2):e0228634. [doi: 10.1371/journal.pone.0228634] [Medline: 32101551]
6. Xu C, Jing F, Lu Y, et al. Summarizing methods for estimating population size for key populations: a global scoping review for human immunodeficiency virus research. AIDS Res Ther 2022 Feb 19;19(1):9. [doi: 10.1186/s12981-022-00434-7] [Medline: 35183203]
7. Teo AKJ, Prem K, Chen MIC, et al. Estimating the size of key populations for HIV in Singapore using the network scale-up method. Sex Transm Infect 2019 Dec;95(8):602-607. [doi: 10.1136/sextrans-2018-053747] [Medline: 31073092]
8. Sabin K, Zhao J, Garcia Calleja JM, et al. Availability and quality of size estimations of female sex workers, men who have sex with men, people who inject drugs and transgender women in low- and middle-income countries. PLOS ONE 2016;11(5):e0155150. [doi: 10.1371/journal.pone.0155150] [Medline: 27163256]
9. Edwards JK, Hileman S, Donastorg Y, et al. Estimating sizes of key populations at the national level: considerations for study design and analysis. Epidemiology 2018 Nov;29(6):795-803. [doi: 10.1097/EDE.0000000000000906] [Medline: 30119057]
10. Bozicevic I, Sharifi H, Haghdoost A, Sabry A, Hermez J. Availability of HIV surveillance data in key populations in the countries of the World Health Organization Eastern Mediterranean Region. Int J Infect Dis 2022 Aug;121:211-216. [doi: 10.1016/j.ijid.2022.05.003] [Medline: 35533830]
11. Mavragani A, Ochoa G. Google Trends in infodemiology and infoveillance: methodology framework. JMIR Public Health Surveill 2019 May 29;5(2):e13439. [doi: 10.2196/13439] [Medline: 31144671]

XSL•FO
RenderX

12. Card KG, Lachowsky NJ, Hogg RS. Using Google Trends to inform the population size estimation and spatial distribution of gay, bisexual, and other men who have sex with men: proof-of-concept study. JMIR Public Health Surveill 2021 Nov 29;7(11):e27385. [doi: 10.2196/27385] [Medline: 34618679]
13. Porn, gay porn - explore. Google Trends. URL: https://trends.google.com/trends/explore?date=2020-01-01%202020-12-31&geo=DZ&q=Porn,Gay%20Porn [accessed 2024-12-31]
14. FAQ about google trends data - trends help. Google. URL: https://support.google.com/trends/answer/4365533?hl=en [accessed 2024-12-31]
15. How many women watch porn? Fight the New Drug. URL: https://fightthenewdrug.org/how-do-men-and-womens-porn-site-searches-differ [accessed 2024-12-31]
16. Pornhub stats reveal women search for more hardcore genres than you might expect. Fight the New Drug. URL: https://fightthenewdrug.org/data-reveals-women-are-searching-hardcore-genres [accessed 2024-12-31]
17. When no one is looking, many women are watching gay porn. NBC News. URL: https://www.nbcnews.com/feature/nbc-out/when-no-one-looking-many-women-are-watching-gay-porn-n894266 [accessed 2024-12-31]
18. Solano I, Eaton NR, O'Leary KD. Pornography consumption, modality and function in a large internet sample. J Sex Res 2020 Jan 2;57(1):92-103. [doi: 10.1080/00224499.2018.1532488]
19. Træen B, Daneback K. The use of pornography and sexual behaviour among Norwegian men and women of differing sexual orientation. Sexologies 2013 Apr;22(2):e41-e48. [doi: 10.1016/j.sexol.2012.03.001]
20. PEPFAR 2022 country and regional operational plan (COP/ROP) guidance for all PEPFAR-supported countries. US Department of State.: PEPFAR; 2022. URL: https://www.state.gov/wp-content/uploads/2022/02/COP22-Guidance-Final_508-Compliant-3.pdf [accessed 2025-01-02]
21. Dangerous inequalities: world AIDS day report 2022. UNAIDS. 2022. URL: https://www.unaids.org/sites/default/files/media_asset/dangerous-inequalities_en.pdf [accessed 2025-01-02]
22. Quick start guide for Spectrum. UNAIDS. 2020. URL: https://www.unaids.org/sites/default/files/media_asset/QuickStartGuide_Spectrum_en.pdf [accessed 2025-01-02]
23. World Health Organization. Recommended population size estimates of men who have sex with men. UNAIDS. 2020. URL: https://www.unaids.org/sites/default/files/media_asset/2020-recommended-population-size-estimates-of-men-who-have-sex-with-men_en.pdf [accessed 2025-01-02]
24. The Key Populations Atlas. UNAIDS. 2022. URL: https://kpatlas.unaids.org/dashboard [accessed 2024-12-31]
25. Key Populations Atlas data sources. UNAIDS - Key Populations Atlas. 2017. URL: https://kpatlas.unaids.org/document/kp_data_sources.pdf [accessed 2025-01-02]
26. Internet world stats - usage and population statistics. ResearchGate. URL: https://www.researchgate.net/publication/258847361_Internet_World_Stats_-_Usage_and_Population_Statistics [accessed 2025-01-02]
27. World development indicators. World Bank Group. URL: https://databank.worldbank.org/source/world-development-indicators [accessed 2024-12-31]
28. Mendos LR, Botha K, Lelis RC, de la Peña EL, Savelev I, Tan D. State-Sponsored Homophobia report. ILGA World. 2020. URL: https://ilga.org/state-sponsored-homophobia-report/ [accessed 2025-01-02]
29. Most searched things on Google in 2024. SimilarWeb. URL: https://www.similarweb.com/blog/marketing/seo/top-keywords/ [accessed 2024-12-31]
30. Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clin Infect Dis 2009 Nov 15;49(10):1557-1564. [doi: 10.1086/630200] [Medline: 19845471]
31. Havelka EM, Mallen CD, Shepherd TA. Using Google Trends to assess the impact of global public health days on online health information seeking behaviour in Central and South America. J Glob Health 2020 Jun;10(1):010403. [doi: 10.7189/jogh.10.010403] [Medline: 32373327]
32. Arora VS, McKee M, Stuckler D. Google Trends: opportunities and limitations in health and health policy research. Health Policy 2019 Mar;123(3):338-341. [doi: 10.1016/j.healthpol.2019.01.001] [Medline: 30660346]
33. Azzam DB, Nag N, Tran J, et al. A novel epidemiological approach to geographically mapping population dry eye disease in the United States through Google Trends. Cornea 2021 Mar 1;40(3):282-291. [doi: 10.1097/ICO.0000000000002579] [Medline: 33177410]
34. NordVPN survey reveals: users still trust free vpns. NordVPN. URL: https://nordvpn.com/blog/nordvpn-usage-survey/ [accessed 2025-01-02]

Abbreviations

GAM: Global AIDS Monitoring system
GT: Google Trend
GTPSE: Google Trend Population Size Estimate
KP: key population
LMIC: low- and middle-income country
MSM: men who have sex with men

XSL•FO
RenderX

PSE: Population size estimation
PSE: population size estimation
RSV: relative search volume
SSA: sub-Saharan Africa
UNAIDS: United Nations Programme on HIV/AIDS
VPN: virtual private network
WHO: World Health Organization

Edited by A Mavragani; submitted 20.03.24; peer-reviewed by E Tuyishime, K Card; revised version received 28.05.24; accepted 28.05.24; published 09.01.25.

Please cite as:
Malburg CM, Gutreuter S, Ruiseñor-Escudero H, Abdul-Quader A, Hladik W
Population Size Estimation of Men Who Have Sex With Men in Low- and Middle-Income Countries: Google Trends Analysis
JMIR Public Health Surveill 2025;11:e58630
URL: https://publichealth.jmir.org/2025/1/e58630
doi:10.2196/58630

© Carly M Malburg, Steve Gutreuter, Horacio Ruiseñor-Escudero, Abu Abdul-Quader, Wolfgang Hladik. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 9.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prevalence and Economic Impact of Acute Respiratory Failure in the Prehospital Emergency Medical Service of the Madrid Community: Retrospective Cohort Study

Ana María Cintora-Sanz[1], MRes, RN; Cristina Horrillo-García[2], MD, PhD; Víctor Quesada-Cubo[3], MD; Ana María Pérez-Alonso[4], MD; Alicia Gutiérrez-Misis[5], MD, PhD

[1]Servicio de Urgencias Médicas de Madrid (SUMMA112), Universidad Autónoma de Madrid, Madrid, Spain

[2]Servicio de Urgencias Médicas de Madrid (SUMMA112), Ventilation Commission, Calle Antracita S/N, Madrid, Spain

[3]Technical Health Research, Madrid, Spain

[4]Servicio de Urgencias Médicas de Madrid (SUMMA112), Research Commission, Madrid, Spain

[5]Department of Medicine, Division of Family Medicine and Primary Care, Clinical Simulation Laboratory, Faculty of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

Corresponding Author:
Cristina Horrillo-García, MD, PhD
Servicio de Urgencias Médicas de Madrid (SUMMA112), Ventilation Commission, Calle Antracita S/N, Madrid, Spain

Abstract

Background: Chronic obstructive pulmonary disease (COPD), congestive heart failure (CHF), and acute pulmonary edema (APE) are serious illnesses that often require acute care from prehospital emergency medical services (EMSs). These respiratory diseases that cause acute respiratory failure (ARF) are one of the main reasons for hospitalization and death, generating high health care costs. The prevalence of the main respiratory diseases treated in a prehospital environment in the prepandemic period and during the COVID-19 pandemic in Spain is unknown. The Madrid Community EMS is a public service that serves all types of populations and represents an epidemiological reference for supporting a population of 6.4 million inhabitants. The high volume of patients treated by Madrid's medical advanced life supports (ALSs) allows us to analyze this little-studied problem.

Objectives: Our goal was to lay the groundwork for comprehensive data collection and surveillance of respiratory failure, with an emphasis on the most prevalent diseases that cause it, an aspect that has been largely overlooked in previous initiatives. By achieving these objectives, we hope to inform efforts to address respiratory failure and establish a standardized methodology and framework that can facilitate expansion to a continuous community-wide registry in Madrid, driving advances in emergency care and care practices in these pathologies. The aim of this retrospective observational study was to determine the pathologies that have mainly caused respiratory failure in patients and required medicalized ALS and to evaluate the cost of care for these pathologies collected through this pilot registry.

Methods: A multicenter descriptive study was carried out in the Madrid Community EMS. The anonymized medical records of patients treated with medical ALS, who received any of the following medical diagnoses, were extracted: ARF not related to chronic respiratory disease, ARF in chronic respiratory failure, exacerbations of COPD, APE, CHF, and bronchospasm (not from asthma or COPD). The prevalence of each pathology, its evolution from 2014 to 2020, and the economic impact of the Medical ALSs were calculated.

Results: The study included 96,221 patients. The most common pathology was exacerbation of COPD, with a prevalence of 0.07% in 2014; it decreased to 0.03% in 2020. CHF followed at 0.06% in 2014 and 0.03% in 2020. APE had a prevalence of 0.01% in 2014, decreasing to 0.005% in 2020 with the pandemic. The greatest economic impact was on exacerbation of COPD in 2015, with an annual cost of €2,726,893 (which equals to US $2,864,628).

Conclusions: COPD exacerbations had the higher prevalence in the Madrid region among the respiratory diseases studied. With the COVID-19 pandemic, the prevalence and costs of almost all these diseases decreased, except for ARF not related to chronic disease. The cost of these pathologies over 5 years was €58,791,031 (US $61,832,879).

(JMIR Public Health Surveill 2025;11:e66179) doi:10.2196/66179

KEYWORDS

acute respiratory failure; COVID-19; chronic obstructive respiratory insufficiency; congestive heart failure; bronchospasm; emergency medical services costs; ambulances; SARS-CoV-2; coronavirus; respiratory; pulmonary; pandemic; economic impact;

XSL•FO
RenderX

observational; Madrid; community; medical records; health records; medical advanced life support; ALS; acute pulmonary edema; chronic obstructive pulmonary disease; COPD; prevalence

Introduction

Chronic obstructive pulmonary disease (COPD) is a public health epidemic. The current state of the art displays a high prevalence and high burden [1]. COPD is the fourth leading cause of death worldwide, accounting for 3.5 million deaths in 2021, approximately 5% of all global deaths [2]. Although systematic reviews on the prevalence of COPD at the hospital level have been published previously [3,4], the literature on current prevalence estimates of COPD exacerbations, treated in prehospital emergency medical services (EMSs), is sparse. In particular, there are few published data on differences in COPD prevalence between rural and urban areas and between countries [5-7]. In this context, relevant and timely information on COPD prevalence in this context is essential to inform, develop, and implement context-appropriate policies and programs for the prevention and optimal control of COPD exacerbations.

Congestive heart failure (CHF) is another highly prevalent disease with significant morbidity and mortality worldwide [8]. There are geographical variations in the epidemiology of CHF [9]. There is a substantial lack of data from the out-of-hospital setting as well, where CHF has different evolution depending on the acute treatment administered [10,11]. Acute pulmonary edema is associated with high mortality. It requires emergency management and usually admission to hospital [7,12,13]. Prospective studies have shown that early treatment applied by prehospital EMS improves short time outcomes and lower hospital readmissions, but the population susceptible to early treatment is unknown [14,15].

COPD, CHF, and acute pulmonary edema (APE) are serious illnesses that often require acute care from EMS [5,6,16]. Respiratory diseases that cause acute respiratory failure (ARF) are one of the main reasons for hospitalization [7,17], generating high health care costs [18,19]. These diseases include bronchospasm as well [20]. Studies on the prevalence and expenditures of prehospital EMSs are scarce [5,21-23]. Consequently, more precise epidemiological data are needed to address these issues and obtain a more accurate picture of the amount of care and the costs generated by these pathologies [24,25] in the prehospital EMS setting.

The main objective of the Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA) project is to determine the prevalence of the most frequent respiratory diseases assisted by the Autonomous Community of Madrid's (CAM) advanced life supports (ALSs). The secondary objective is to measure the EMS economic cost of health care for patients. We present the following article in accordance with the STROBE (Strengthening the Reporting of Observational studies in Epidemiology) reporting checklist (Multimedia Appendix 1).

Methods

Study Design

The SUMMIRA study is an observational, quantitative, multicenter, and cross-sectional study of patients attended by the EMS of the CAM for 6 years.

Sample Size

Between 2014 and 2020, 583,984 persons of the population of the Madrid Community were clinically treated by the CAM ALSs, after a triage in the 112 health emergency telephone number. Based on these medical examinations, the subjects were included in the study. A cohort of 22,085 subjects was identified as having decompensated COPD; 22,085 with a main diagnosis of CHF; 4676 with APE; 17,843 presented ARF not related to chronic respiratory failure (CRF); 6921 were diagnosticated with bronchospasm; and 17,851 were related with ARF in CRF. This study involved a total of 96,221 patients included from the electronic medical records (EMRs).

Study Setting

The CAM has a population of 6.4 million inhabitants from 2014, both urban and rural areas. The CAM has 47 ALSs [26]. The ALSs are responsible for (1) emergency care, namely, patient care at home or on public roads in the Madrid Community and (2) interhospital transfers. The CAM is also in charge of the Helicopter Emergency Medical Service.

Each emergency care is provided by a health professional team formed by an emergency physician, an emergency nurse, and 1 or 2 emergency technicians; these teams assist at home and other places and take care of the patient in the mobile intensive critical unit (ALS).

The EMRs analyzed are those performed by the ALS physicians who attended the pathologies under study. These are the patients attended by the ALS of the Community of Madrid after the patients or their families called 112 (both in the city and in the outskirts).

The EMRs are located in a program named Integrated Management System of the Coordinating Center [27]. Integrated Management System of the Coordinating Center is a highly secure system of clinical data that has received urgent health care from the ALSs of the CAM.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki. This study was reviewed and approved by the Madrid Regional Ethics Committee on Medication Research (dated: June 6, 2022/N° V8.2022). All data used in this study have been deidentified or anonymized to protect the privacy of participants. No information is available that would allow for the direct or indirect identification of individuals. The Madrid Regional Ethics Committee approval covers secondary analysis without additional consent.

XSL•FO
RenderX

Since the study involved the secondary analysis of previously collected data, the ethics committee determined that the project was exempt from further review under the category of anonymized, preexisting data analysis. The study complied fully with international ethical standards and the protection of participants' rights.

All EMRs in this study have been anonymized to protect the privacy of participants. Rigorous measures were taken to ensure the confidentiality of the data and the integrity of the information collected complying, therefore, with the principles regarding data processing referred to in Article 5 of European Union Regulation 2016/679 of the European Parliament and Council of April 27, 2016 (General Data Protection Regulation).

Sampling and Recruitment Procedure

Inclusion Criteria

Patients with EMRs were treated by the CAM ALSs who, according to usual clinical practice, had been diagnosed with ARF not related to CRF, exacerbation of chronic obstructive pulmonary disease (ECOPD), CHF, bronchospasm, or APE according to the codes of the *International Classification of Diseases, Ninth Revision* [9] between 2014 and 2020. Medical diagnostic criteria based on clinical and historical data were followed. The population residing in the CAM per year was taken as the reference population, according to the Spanish National Institute of Statistics [28].

The diagnoses of APE and ECOPD were defined as a "clinical picture" in the prehospital and hospital emergency manuals [29-31], with consensus documents [32,33] and systematic review such as by Bello et al [34] . The diagnostic criteria for COPD and APE are described in Multimedia Appendix 2 and Table 1. The other pathologies were diagnosed following the diagnostic criteria of the 12 Octubre Physician Manual [35] in accordance with another prehospital emergency manual [36] and the Servicio de Urgencias Medicas de Madrid (SUMMA112; Prehospital Emergency Medical Service of Autonomous Community of Madrid) Physician Clinical Guide for Acute Heart Insufficiency (Multimedia Appendix 3).

Table . Prevalence of congestive heart failure, exacerbation of chronic obstructive pulmonary disease (ECOPD), acute pulmonary edema, and bronchospasm acute respiratory insufficiency not related to chronic respiratory failure, of Madrid region population, based to the Madrid Community census [37], and the prehospital emergency medical service´s electronic medical record of Madrid from 2014 to 2020 (n=583,984).

Diseases[a]	2014	2015	2016	2017	2018	2019	2020
Congestive heart failure, n (%)	3593 (0.056)	3696 (0.057)	3651 (0.056)	3484 (0.054)	2895 (0.044)	2828 (0.042)	1938 (0.029)
ECOPD, n (%)	4415 (0.068)	4463 (0.069)	4112 (0.064)	4284 (0.066)	3935 (0.060)	3616 (0.054)	2020 (0.030)
Acute pulmonary edema, n (%)	819 (0.013)	800 (0.012)	725 (0.011)	711 (0.011)	652 (0.010)	616 (0.009)	353 (0.005)
Acute respiratory insufficiency in chronic respiratory failure, n (%)	2872 (0.044)	3034 (0.047)	2834 (0.044)	2797 (0.043)	2399 (0.036)	2362 (0.035)	1553 (0.023)
Bronchospasm, n (%)	1000 (0.015)	1104 (0.017)	1157 (0.018)	1157 (0.018)	1062 (0.016)	979 (0.015)	462 (0.007)
Acute respiratory insufficiency (not related to chronic respiratory failure), n (%)	2648 (0.041)	2736 (0.043)	2717 (0.042)	2573 (0.040)	2349 (0.036)	2289 (0.034)	2531 (0.038)

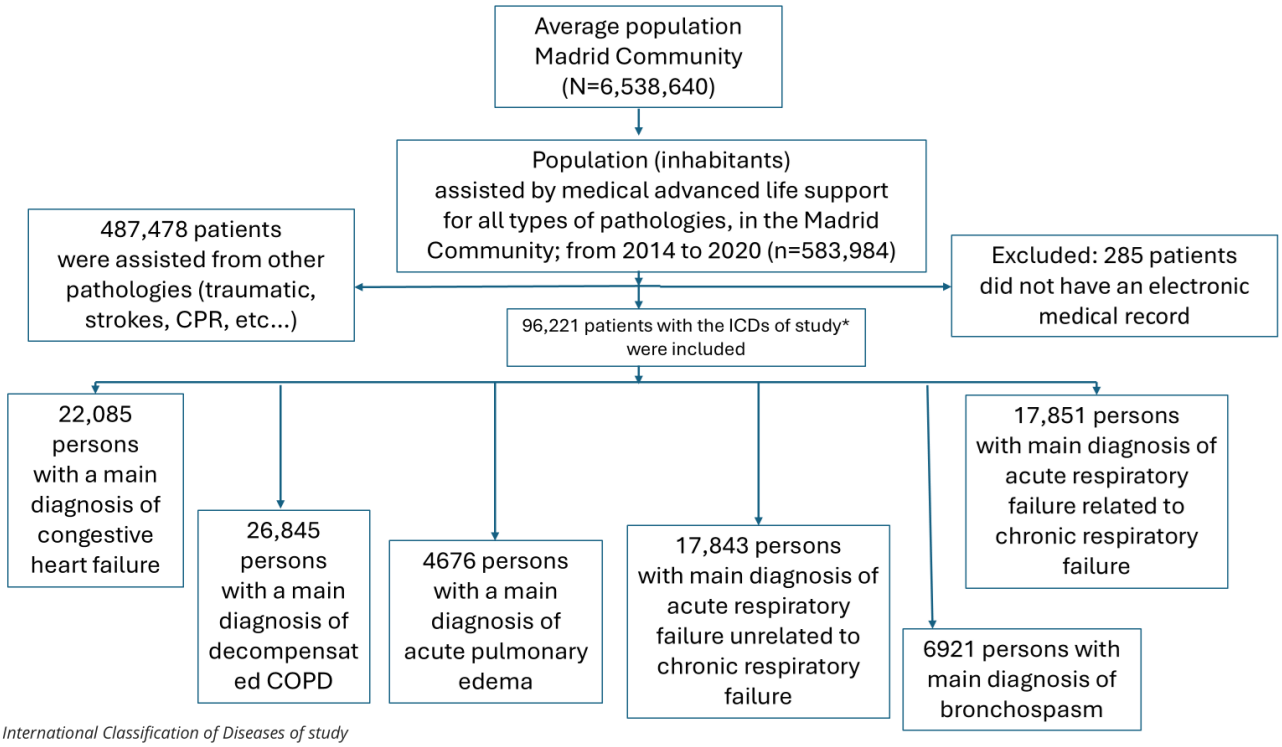
[a]Results from the retrospective observational research Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA).

Exclusion Criteria

Patients without a health identification number or without an adequate record in the basic minimum dataset were excluded. All patients diagnosed with asthma were excluded since they could be included among the patients with ARF related to chronic respiratory disease (Figure 1).

XSL•FO
RenderX

Figure 1. Cohort selection of patients assisted by the Madrid Community emergency medical service from January 1, 2014, to December 31, 2020. Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA) Study. COPD: chronic obstructive pulmonary disease.



Economic Analysis

The prevalence for each pathology was calculated by dividing the total number of individuals who presented the corresponding medical diagnosis in a given year by the population of Madrid at that time, according to the Spanish National Institute of Statistics.

The direct economic cost of CAM EMS care for the pathologies treated was calculated from the expenditure according to public references [38]. The health care cost data were extracted from the public prices for the provision of public health services and activities of the ALSs published in the Official Gazette of the CAM on August 21, 2017 (price list with an application date until December 2023). The reference cost for this study was the average cost of €611 (US $64,261), which is the arithmetic mean of mobilizing and assisting a patient by a terrestrial ALS unit in the public EMS of the Madrid Community [39]. The ALS provides advanced care, as it contains all the electromedical material and personnel that are required to form a mobile intensive care unit. The cost of ALS care (emergency medical services costs) includes physician care, emergency nursing assistance, and all required ventilatory support techniques, depending on the severity of the patient's condition, with pharmacological treatment, oxygen therapy, invasive and noninvasive mechanical ventilation, and transfer to the hospital.

The direct economic impact was obtained by multiplying the cost of care for each pathology treated in the CAM's ALSs by the number of times care was required. All persons within the CAM were able to receive care, as it is included in the universal emergency health care coverage of the Madrid Health Regional Government (public administration). The economic analysis was done in Microsoft Excel 365 and SPSS software (version 26; IBM Corp), with a copyright license from the Madrid Health Regional Government.

Results

Table 1 shows a total of 96,221 patients who were included between 2014 and 2020. The most prevalent pathology among all those studied was COPD, with an initial prevalence of 0.068% (4415/6,454,440) in 2014, which decreased to 0.030% (2020/6,779,888) in 2020. CHF had a prevalence of 0.056% (3593/6,454,440) in 2014, decreasing to 0.029% (1938/6,779,888) in 2020. The prevalence of ARF in CRF was 0.044% (2872/ 6,454,440) in 2014 but decreased to 0.023% (1553/6,779,888) in 2020. The prevalence of bronchospasm (not from asthma or COPD) decreased from 0.015% (1000/6,454,440) in 2014 to 0.007% (462/6,779,888) in 2020. APE's prevalence decreased from 0.013% (819/6,454,440) in 2014 to 0.005% (353/6,779,888) in 2020. The prevalence of ARF in patients not related to CRF decreased from 0.041% (2648/6,454,440) in 2014 to 0.034% (2289/6,663,394) in 2019 and increased to 0.038% (2531/6,779,888) in 2020, reversing this trend of ARF observed, especially in CHF and ECOPD.

The prevalence of APE shows a decreasing pattern from the onset. The other pathologies start with an upward trend, although in the second year, they reverse the trend and gradually decrease. The prevalence of bronchospasm maintains an upward trend until 2017 and from then on it starts to decrease progressively. In 2020, a sharper peak of decrease is assessed in all pathologies, except ARF not related to CRF, the latter being the only one to increase.

A decrease in most respiratory diseases (Table 2) was observed in 2020, both for patients with ECOPD, APE, or bronchospasm

XSL•FO
RenderX

(nonasthmatic, non-COPD) and for patients with CHF (Figures 2A-F and 3), except for ARF not associated with CRF, which corresponded in time and form with the COVID-19 pandemic period [40]. From 2017 on, a decrease in all types of ARF was observed, especially CHF and COPD.

Table 2 shows the economic cost of each pathology and its economic impact in each year. The greatest economic impact was observed in 2015 for ECOPD, with an annual cost of €2,726,893 (US $2,867,982). The estimated cost of all pathologies over 5 years was €58,791,031 (US $61,832,879).

Table . Direct economic impact of the congestive heart failure, exacerbation of chronic obstructive pulmonary disease (ECOPD), acute pulmonary edema (APE), bronchospasm, and acute respiratory failure not related to chronic respiratory failure of the Madrid region population; based on the prehospital Madrid emergency medical service's electronic medical records[a].

ICD-CM[b] type	2014	2015	2016	2017	2018	2019	2020
Congestive heart failure							
Cases	3593	3696	3651	3484	2895	2828	1938
Cost (€)	2,195,323	2,258,256	2,230,761	2,128,724	1,768,845	1,727,908	1,184,118
ECOPD							
Cases	4415	4463	4112	4284	3935	3616	2020
Cost (€)	2,697,565	2,726,893	2,512,432	2,617,524	2,404,285	2,209,376	1,234,220
APE							
Cases	819	800	725	711	652	616	353
Cost (€)	500,409	488,800	442,975	434,421	398,372	376,376	215,683
Acute respiratory failure in chronic respiratory failure							
Cases	2872	3034	2834	2797	2399	2362	1553
Cost (€)	1,754,792	1,853,774	1,731,574	1,708,967	1,465,789	1,443,182	948,883
Bronchospasm							
Cases	1000	1104	1157	1157	1062	979	462
Cost (€)	611,000	674,544	706,927	706,927	648,882	598,169	282,282
Acute respiratory failure (not related to chronic failure)							
Cases	2648	2736	2717	2573	2349	2289	2531
Cost (€)	1,617,928	1,671,696	1,660,087	1,572,103	1,435,239	1,398,579	1,546,441
Total cost (€58,791,031) (€)	9,377,017	9,673,963	9,284,756	9,168,666	8,121,412	7,753,590	5,411,627
Total cost (US $61,832,879) (US $)	9,862,184	10,174,494	9,765,149	9,643,053	8,541,614	8,154,761	5,691,625

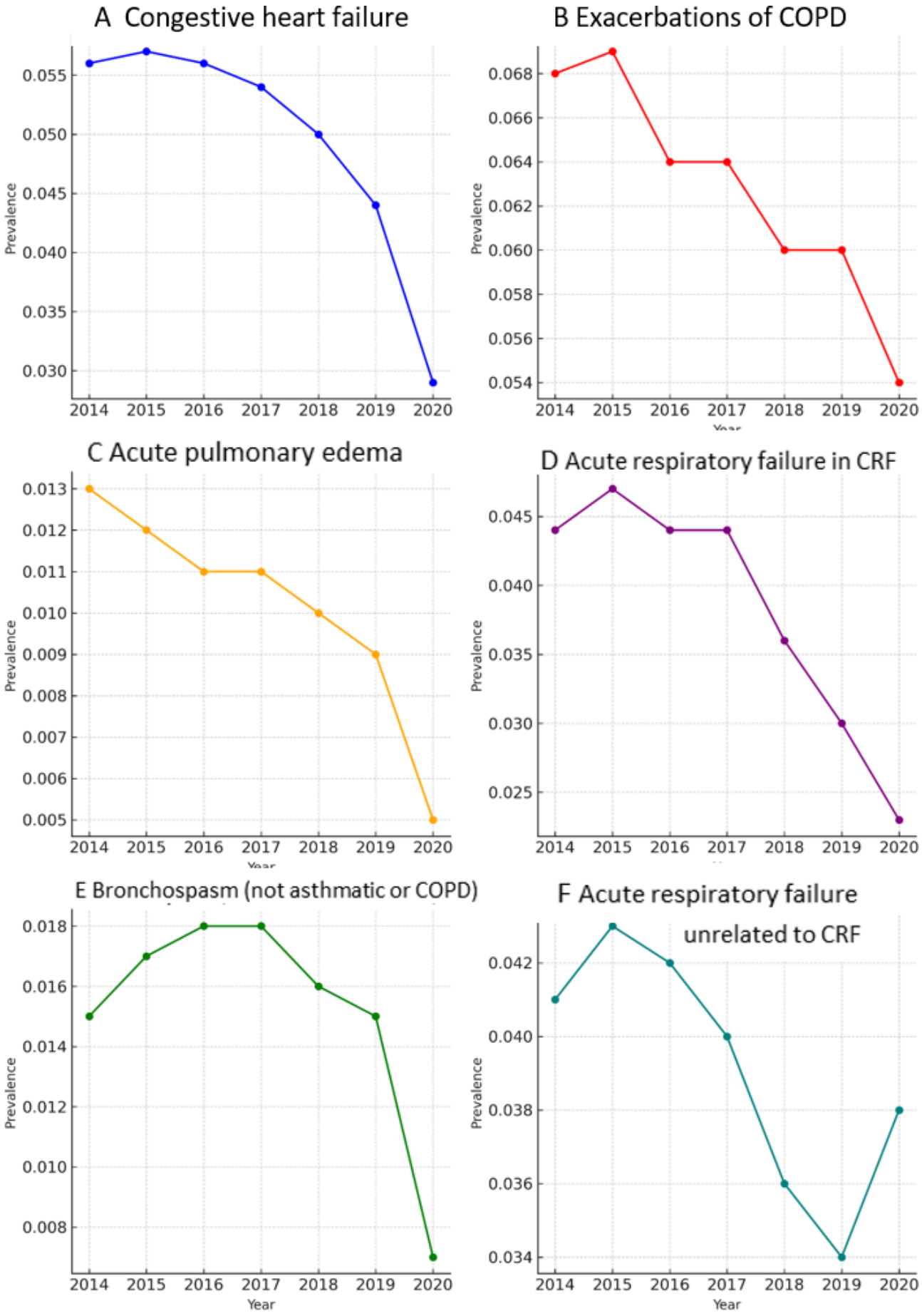
[a]Database from 2014 to 2020 included. Results from the retrospective observational research Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA).

[b]*ICD-CM*: *International Classification of Diseases, Clinical Modification*.

XSL•FO
RenderX

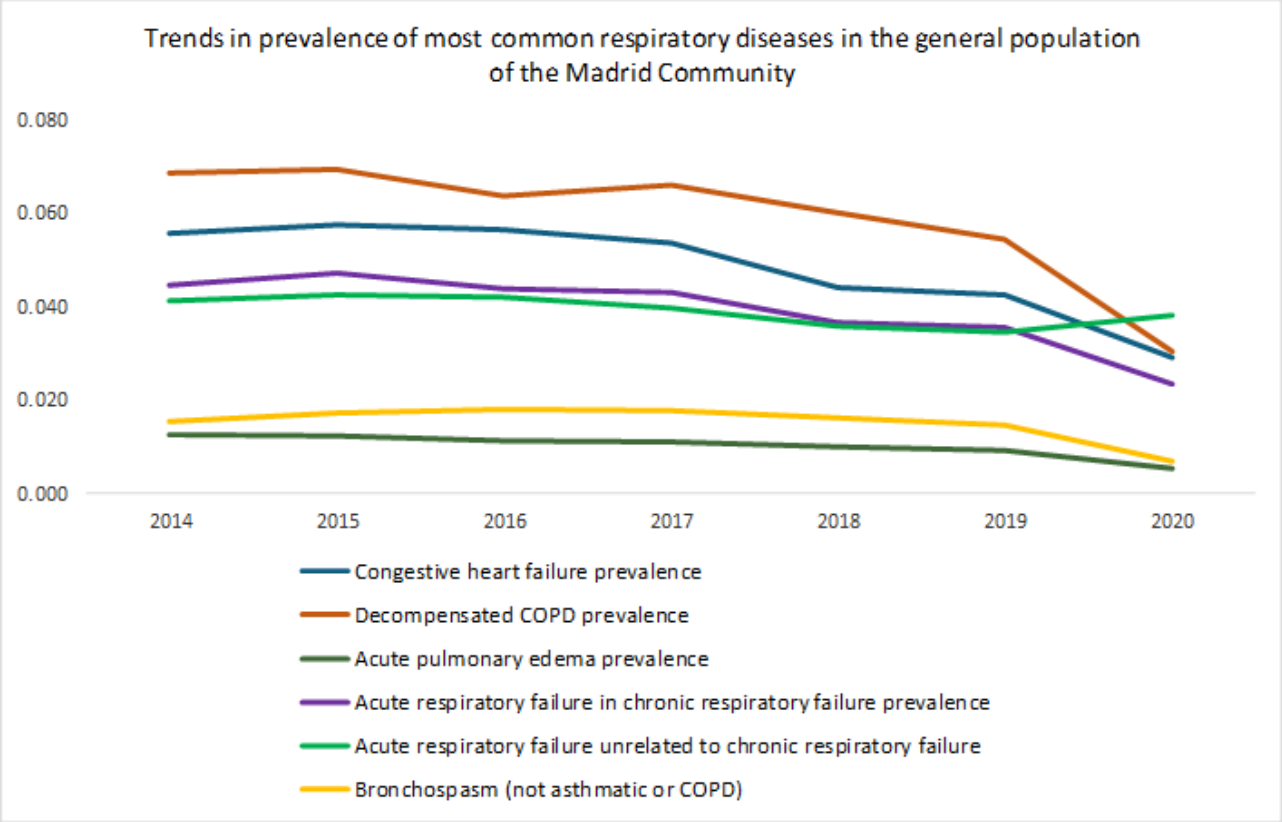
Figure 2. Changes in the prevalence of (A) congestive heart failure, (B) exacerbation of COPD, (C) acute pulmonary edema, (D) acute respiratory failure in chronic respiratory failure, (E) bronchospasm, and (F) acute respiratory insufficiency not related to chronic respiratory failure based on electronic medical record of the prehospital Madrid Community emergency medical service (2014-2020), assisted by prehospital emergency medical services. Results from the retrospective observational research Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA). COPD: chronic obstructive pulmonary disease; CRF: chronic respiratory failure.

XSL•FO
RenderX



XSL•FO
RenderX

Figure 3. Evolution of congestive heart failure, exacerbation of chronic obstructive pulmonary disease, acute pulmonary edema, and acute respiratory failure in chronic respiratory failure prevalences in the Madrid region, based on the general population electronic medical records, assisted by prehospital emergency medical services, from 2014 to 2020. Results from the retrospective observational research Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid (SUMMIRA). COPD: chronic obstructive pulmonary disease.



Discussion

Principal Findings

Among the findings, we can mention that to our knowledge this is the only study in the prehospital EMS that has analyzed the prevalence of these pathologies in Spain and in Europe, so it can serve as a reference for future European studies on this topic. It was measured by the costs categorized by diagnosis. Finally, it is worth mentioning the large sample size, which provides greater reliability to the study. Few studies have quantified the direct costs incurred by ALSs in these pathologies [25,41-45]. A study carried out in Sweden and Belgium quantified a type of direct cost, such as transport for ECOPD, at 3% of the total direct costs [42]. In France, the cost of medical transport per patient was estimated to be €192 - 273 (US $202-287) in 1996 [46]. In this study, the cost of air ambulance transportation, but not that of emergency medical care, was estimated.

In the ALS of the CAM, the demand by patients with dyspnea as the main diagnosis among patients with ECOPD was known, although there were no data to differentiate among ECOPD, bronchospasm, and other processes. Therefore, we needed to accurately calculate the prevalence of ARF in ALS to more precisely estimate the economic impact of these pathologies and their complete clinical picture, which includes emergent prehospital care [47].

The results of the SUMMIRA project have made it possible to estimate the global quantification for prehospital emergency care due to SARS-CoV-2 in the CAM, served by the ALSs of the CAM. The pathologies studied showed a continuously decreasing trend throughout the period analyzed, coinciding with the mortality data related to ECOPD published in the CAM [22]. The only different trend was the peak produced by ARF not associated with chronic failure, which, despite showing a decreasing trend from 2014 to 2019, increased the number of cases in 2020 at the same time as the COVID-19 pandemic in Spain [48].

These other declines previous to 2020 may be the result of many factors, such as regional success of tobacco control measures, improved information and prevention of allergens, better treatment of respiratory and cardiac comorbidities, and a reduction of universally high rates of underdiagnosis of chronic respiratory diseases [25]. The decreasing trend in the number of cases of most of the chronic respiratory diseases studied, such as COPD, CHF, and APE, before 2020, led us to estimate an improvement in disease control in chronic patients to date and a subsequent influence of the cases due to the COVID-19 pandemic [49].

The economic impact of chronic diseases on patients has not been quantified in the most recent COPD studies [7,21,50], as only primary care and hospital expenditure have been taken into account [22,45,51]. A descriptive North American study lasting 5 years with a large sample (166,908 patients) in which patients

XSL•FO
RenderX

with respiratory distress transferred to the hospital was analyzed [6], but the population prevalences were not calculated.

Furthermore, there is a comparative dilemma between the different activities of the EMS's ALSs in different countries. Some of them in Europe, such as the paramedical system, cannot discharge patients since transfer to hospital is part of the work protocol because some of them do not have doctors within health care teams who can discharge patients [52-54]. In this sense, the Spanish system of EMS [55] helps prevent care overload in hospitals since the physicians of ALSs allow patients to remain at home and can be treated there, improving adequately after receiving treatment and thus achieving benefits for both patients and the health system. Considering the number of cases of the diseases found, it would be advisable to evaluate the prevalence of these diseases in other similar European and national emergency health systems to have an updated global vision.

The set of decreasing trends in the prevalence of ECOPD is in line with the trend in industrialized countries [56]. Plausible causes include the greater control of adverse risk factors, as well as the increase in public health measures that protect against respiratory diseases (tobacco, asbestos, protective measures in occupational risks, etc) [18,25,57,58].

Limitations

This work had a shortcoming in the population sample. This cohort did not include the care provided to the patients who go directly to the hospital emergency department; it is a prehospital cohort. The cohort would be bigger if it would include the patients of the Madrid Municipal Emergency Assistance and Rescue Service (SAMUR), though the number of patients treated by them is minimal because their assistance is limited to the public thoroughfare of Madrid capital, and by regular form this type of patient is treated at home and primary care that is covered by SUMMA112.

We have separated the average cost of the assistance provided by the advanced terrestrial life support services of the CAM to estimate the economic cost. Advanced air life support has also treated some patients, which is valued in Table 3.

Table . Case study of congestive heart failure, exacerbation of chronic obstructive pulmonary disease (ECOPD), acute pulmonary edema, bronchospasm, and acute respiratory insufficiency not related to chronic respiratory failure of the Madrid region population per year treated by Servicio de Urgencias Medicas de Madrid Helicopter Emergency Medical Service (SUMMA112 HEMS), from 2014 to 2020[a].

Condition	2015	2016	2017	2018	2019	2020
ECOPD	1 ECOPD	1 ECOPD	2 ECOPD	2 ECOPD	3 ECOPD	1 ECOPD
Acute respiratory failure related to chronic respiratory failure	1 Acute respiratory failure in chronic respiratory failure	2 Acute respiratory failure in chronic respiratory failure	2 Acute respiratory failure related to chronic respiratory failure	1 Acute respiratory failure related to chronic respiratory failure	2 Acute respiratory failure related to chronic respiratory failure	1 Acute respiratory failure related to chronic respiratory failure
Acute pulmonary edema	1 Acute pulmonary edema	1 Acute pulmonary edema	2 Acute pulmonary edema	1 Acute pulmonary edema	3 Acute pulmonary edema	N/A[b]
Congestive heart failure	1 Congestive heart failure	4 Congestive heart failure	2 Congestive heart failure	N/A	N/A	3 Congestive heart failure
Bronchospasm	N/A	1 Bronchospasm	N/A	1 Bronchospasm	2 Bronchospasm	N/A
Acute respiratory failure not related to chronic respiratory failure	N/A	2 Acute respiratory failure not related.to chronic respiratory failure	2 Acute respiratory failure not related to chronic respiratory failure	1 Acute respiratory failure not related to chronic respiratory failure	3 Acute respiratory failure not related to chronic respiratory failure	N/A

[a]As can be valued in the Official Gazette of the Madrid Council [38], the cost per medical helicopter intervention in this period time is €5746 (US $6043).

[b]N/A: not applicable.

Our study shows a reduction in the costs of CHF over the years in EMS. The decrease may be explained by factors such as reduced severity and better treatment of acute coronary syndrome [21]. We have not found any similar bibliography in the prehospital setting. The 2019 Heart Failure Association Atlas reported a heart failure prevalence ranging from ≤12 in Spain [51]. At the European level, there is considerable heterogeneity of data between European countries, ranging from an overall prevalence of 14 per 1000 in the Netherlands [24] to 39 in Slovenia, Lithuania, and Germany but with a common point: EMS use in heart failure is an independent predictor of 30-day mortality [50]. In exacerbations of COPD, the patients have a significant comorbidity burden and experience high rates of hospitalization and mortality [59]. It can lead to several conclusions in the area of public health: (1) the measures taken by primary and specialized care have been effective in control of exacerbations of this type of pathology, and furthermore (2) despite reducing costs, in seasonal outbreaks a greater investment [51] of costs would be needed to reinforce the resources that deal with the ECOPD in the prehospital emergency services.

This study opens the possibility of including these values in the database of the economic impact of chronic respiratory diseases such as COPD, CHF, and APE, which are currently not considered [6-8]. In addition to the registration of their prevalence in the different months of the year, an analysis of the needs required in public health can be made to make investment plans in the times of the year when there is a higher prevalence. Moreover, it could be taken into account when

XSL•FO
RenderX

assessing the costs associated with SARS-CoV-2–related bronchospasm and ARF in prehospital critical care for future respiratory transmission pandemics.

Conclusions

The results found in this study are in line with the general hospitalization downward trend in high-income countries of a decrease in the number of cases from 2014 to 2020 for the most common respiratory diseases treated in hospital emergencies [21,51,56]. The prevalence of ECOPD in the Madrid EMS was higher than the other respiratory diseases studied. In 2020, with the COVID-19 pandemic, a decrease in the prevalence and costs of the pathologies studied was observed, except for the prevalence of ARFs not related to chronic failure. The greatest economic impact was for ECOPD in 2015, with an annual cost of €2,726,893 (US $2,867,982). The estimated cost of these pathologies in 5 years was €58,791,031 (US $61,832,879).

Acknowledgments

The authors thank the support of the medical director and the medical deputy director of the Servicio of Urgencias Médicas of Madrid (SUMMA112); data exploration team of the IT Department of SUMMA112; Mr Juan Pajuelo Ayuso; Mr José Gómez Martínez, and Mrs Nuria Rodríguez Rodil; Candelas Gil, Head of the Medical Library Service of the Autonomous University of Madrid, Spain; and the Primary Care Biomedical Research and Innovation Foundation (FIIBAP). Funding for open-access publishing was provided by the FIIBAP (grant 883345).

This work has been presented as oral communication at the XXXIX Annual Meeting of the Spanish Epidemiology Society (SEE) and at the XIX Congress of the Spanish Society of Public Health and Health Administration (SESPAS). This article sheds light on the changing prevalence of the most frequent respiratory diseases treated by emergency medical services, the influence of the COVID-19 pandemic on them, and the costs involved in their care in the out-of-hospital emergency setting.

Data Availability

The datasets generated analyzed during this study are not publicly available due to privacy requirements and ethical restrictions. Reasonable requests for data access are available by contacting the corresponding author and submitting application documents. Anonymized data from this study are available from the corresponding author upon reasonable request. The descriptive multicenter study that is presented is based on a health care casuistry by filtering the codes of the respiratory pathologies registered in the Electronic Medical Record on the Integrated Management System of the Coordinating Center platform.

Integrated Management System of the Coordinating Center is a computer application that operates under the supervision of the Information Systems Security Office of the Department of Health of the Community of Madrid. It is included in the Integrated Management System of the Prehospital Emergency Medical Service of Autonomous Community of Madrid (SUMMA112) Coordination Center. These data are protected by security systems and guarded by the technical managers of the Information Systems Security Office of the Department of Health of the Community of Madrid (OSSI). There is a high level of security in the treatment and storage of medical records from the Medical Emergency Service of the Madrid Region, which depends on the General Directorate of Information Systems and Health Equipment of the Madrid Health Service.

Authors' Contributions

AMCS, CHG, VQC, AMPA, and AGM were involved in the conception and design of the work. AMCS, CHG, and VQC contributed to the screening process, data extraction, and quality assessment of papers. AMCS, CHG, and VQC interpreted the data and drafted the manuscript. The manuscript was critically reviewed and enriched by CHG, AMPA, and AGM. All authors read and approved the final manuscript and agreed to be accountable for all aspects of this work.

Conflicts of Interest

None declared. All authors have confirmed the maintenance of confidentiality and respect for the rights of the patient in the document of responsibilities of the author, the publication agreement, and the transfer of rights to the journal.

Multimedia Appendix 1
STROBE (Strengthening the Reporting of Observational studies in Epidemiology) reporting checklist.
[DOCX File, 72 KB - publichealth_v11i1e66179_app1.docx]

Multimedia Appendix 2
Exacerbation of chronic obstructive pulmonary disease and acute pulmonary edema diagnosis.
[DOCX File, 67 KB - publichealth_v11i1e66179_app2.docx]

Multimedia Appendix 3

XSL•FO
RenderX

Congestive heart failure diagnosis.
[DOCX File, 60 KB - publichealth_v11i1e66179_app3.docx]

References

1. Ananth S, Hurst JR. ERJ advances: state of the art in definitions and diagnosis of COPD. Eur Respir J 2023 Apr;61(4):2202318. [doi: 10.1183/13993003.02318-2022] [Medline: 36796835]
2. Chronic obstructive pulmonary disease (COPD). World Health Organization. 2024 Nov 6. URL: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd) [accessed 2025-01-03]
3. Adeloye D, Song P, Zhu Y, et al. Global, regional, and national prevalence of, and risk factors for, chronic obstructive pulmonary disease (COPD) in 2019: a systematic review and modelling analysis. Lancet Respir Med 2022 May;10(5):447-458. [doi: 10.1016/S2213-2600(21)00511-7] [Medline: 35279265]
4. Zou W, Ou J, Wu F, et al. Association of mild chronic obstructive pulmonary disease with all-cause mortality: a systematic review and meta-analysis. Pulmonology 2023 Oct 27:S2531-0437(23)00165-4. [doi: 10.1016/j.pulmoe.2023.09.002] [Medline: 38093693]
5. Fitzpatrick D, Duncan EAS, Moore M, et al. Epidemiology of emergency ambulance service calls related to COVID-19 in Scotland: a national record linkage study. Scand J Trauma Resusc Emerg Med 2022 Jan 28;30(1):9. [doi: 10.1186/s13049-022-00995-6] [Medline: 35090527]
6. Prekker ME, Feemster LC, Hough CL, et al. The epidemiology and outcome of prehospital respiratory distress. Acad Emerg Med 2014 May;21(5):543-550. [doi: 10.1111/acem.12380] [Medline: 24842506]
7. Zhou J, Nehme E, Dawson L, et al. Epidemiology, outcomes and predictors of mortality in patients transported by ambulance for dyspnoea: a population-based cohort study. Emerg Med Australas 2023 Feb;35(1):48-55. [doi: 10.1111/1742-6723.14053] [Medline: 35918062]
8. Cardiovascular Research. Global burden of heart failure: a comprehensive and updated review of epidemiology. Oxford Academic. 2022 Feb 12. URL: https://academic.oup.com/cardiovascres/article/118/17/3272/6527627?login=false [accessed 2025-01-03]
9. Savarese G, Becher PM, Lund LH, Seferovic P, Rosano GMC, Coats AJS. Global burden of heart failure: a comprehensive and updated review of epidemiology. Cardiovasc Res 2023 Jan 18;118(17):3272-3287. [doi: 10.1093/cvr/cvac013]
10. Miró Ò, Hazlitt M, Escalada X, et al. Effects of the intensity of prehospital treatment on short-term outcomes in patients with acute heart failure: the SEMICA-2 study. Clin Res Cardiol 2018 Apr;107(4):347-361. [doi: 10.1007/s00392-017-1190-2]
11. Miró Ò, Martínez G, Masip J, et al. Effects on short term outcome of non-invasive ventilation use in the emergency department to treat patients with acute heart failure: a propensity score-based analysis of the EAHFE Registry. Eur J Intern Med 2018 Jul;53:45-51. [doi: 10.1016/j.ejim.2018.03.008] [Medline: 29572091]
12. Purvey M, Allen G. Managing acute pulmonary oedema. Aust Prescr 2017 Apr;40(2):59-63. [doi: 10.18773/austprescr.2017.013]
13. Wang J, Duan J, Zhou L. Incidence of noninvasive ventilation failure and mortality in patients with acute respiratory distress syndrome: a systematic review and proportion meta-analysis. BMC Pulm Med 2024 Jan 22;24(1):48. [doi: 10.1186/s12890-024-02839-8] [Medline: 38254064]
14. Quintana MGM, García CH, Misis AG, et al. Factors associated with in-hospital mortality and readmission in a cohort of patients treated with noninvasive ventilation during emergency department or out-of-hospital emergency care: the VentilaMadrid study. Emerg Rev Soc Esp Med Emerg 2022 Feb;34(1):7-14.
15. Muñoz SS, Canela À, Chao JRC, Doblado I, Fàbrega XJ, Rodríguez JJ. Impact of pre-hospital application of non-invasive mechanical ventilation in acute pulmonary oedema. NIVEMS cohort study. Rev Esp Urgenc Emerg Span J Acute Emerg Care 2024;3(4) [FREE Full text]
16. Sneath E, Bunting D, Hazell W, Tippett V, Yang IA. Pre-hospital and emergency department pathways of care for exacerbations of chronic obstructive pulmonary disease (COPD). J Thorac Dis 2019 Oct;11(Suppl 17):S2221-S2229. [doi: 10.21037/jtd.2019.10.37] [Medline: 31737349]
17. Boers E, Barrett M, Su JG, et al. Global burden of chronic obstructive pulmonary disease through 2050. JAMA Netw Open 2023 Dec 1;6(12):e2346598. [doi: 10.1001/jamanetworkopen.2023.46598] [Medline: 38060225]
18. Duan KI, Birger M, Au DH, Spece LJ, Feemster LC, Dieleman JL. Health care spending on respiratory diseases in the United States, 1996-2016. Am J Respir Crit Care Med 2023 Jan 15;207(2):183-192. [doi: 10.1164/rccm.202202-0294OC] [Medline: 35997678]
19. Boucher PE, Taplin J, Clement F. The cost of ARDS: a systematic review. Chest 2022 Mar;161(3):684-696. [doi: 10.1016/j.chest.2021.08.057] [Medline: 34478719]
20. Peters GA, Cash RE, Goldberg SA, Kolb LM, Ordoobadi AJ, Camargo CA Jr. Emergency medical services management of bronchospasm in the United States: a cross-sectional analysis and nationwide quality assessment. Prehosp Emerg Care 2024;28(2):231-242. [doi: 10.1080/10903127.2023.2220021] [Medline: 37276151]
21. Rosano GMC, Seferovic P, Savarese G, et al. Impact analysis of heart failure across European countries: an ESC-HFA position paper. ESC Heart Fail 2022 Oct;9(5):2767-2778. [doi: 10.1002/ehf2.14076] [Medline: 35869679]

XSL•FO
RenderX

22. White paper on the socio-economic burden of COPD [Article in Spanish]. Weber. 2015. URL: https://weber.org.es/publicacion/libro-blanco-sobre-la-carga-socio-economica-de-la-epoc/ [accessed 2025-01-03]
23. Brown R, Goldstein J, Jensen JL, Travers AH, Carter A. Evaluating the impact of a novel mobile care team on the prevalence of ambulatory care sensitive conditions presenting to emergency medical services in Nova Scotia. Cureus 2023 Apr;15(4):e37280. [doi: 10.7759/cureus.37280] [Medline: 37168216]
24. Sicras-Mainar A, Sicras-Navarro A, Palacios B, Varela L, Delgado JF. Epidemiology and treatment of heart failure in Spain: the HF-PATHWAYS study. Rev Esp Cardiol (Engl Ed) 2022 Jan;75(1):31-38. [doi: 10.1016/j.rec.2020.09.033] [Medline: 33380382]
25. Soriano JB, Kendrick PJ, Paulson KR, et al. Prevalence and attributable health burden of chronic respiratory diseases, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Respir Med 2020 Jun;8(6):585-596. [doi: 10.1016/S2213-2600(20)30105-3]
26. Sacristán FR. El SUMMA 112 uno de los principales baluartes de la sanidad madrileña [Article in Spanish]. Gaceta Médica. 2024. URL: https://gacetamedica.com/politica/el-summa-112-uno-de-los-principales-baluartes-de-la-sanidad-madrilena/ [accessed 2025-01-03]
27. Sistemas y Tecnologías de la Información [Article in Spanish]. URL: https://www.comunidad.madrid/hospital/summa112/profesionales/sistemas-tecnologias-informacion [accessed 2025-01-03]
28. Instituto nacional de estadística [Article in Spanish]. INE. URL: https://www.ine.es/ [accessed 2025-01-02]
29. Manual de ventilación mecánica del SUMMA112 comunidad de madrid [Article in Spanish]. Comunidad de Madrid. URL: https://www.comunidad.madrid/publicacion/ref/50596 [accessed 2025-01-03]
30. Manual de Procedimientos · SAMUR Protección Civil [Article in Spanish]. Madrid. URL: https://www.madrid.es/ficheros/SAMUR/ [accessed 2025-01-03]
31. Bibiano C, Sanz M, Esparza C, Caussade F, Montero M, Puig R. Emergency manual 2014. URL: https://www.urgenciasyemergen.com/nuevo-manual-urgencias-3a-ed-bibiano/ [accessed 2024-12-20]
32. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. American Heart Association Journals. 2022 May 3. URL: https://www.ahajournals.org/doi/epub/10.1161/CIR.0000000000001063 [accessed 2025-01-03]
33. Luján M, Peñuelas Ó, Gómez CC, et al. Sumario de las recomendaciones y puntos clave del Consenso de las Sociedades Científicas Españolas (SEPAR, SEMICYUC, SEMES; SECIP, SENeo, SEDAR, SENP) para la utilización de la ventilación no invasiva y terapia de alto flujo con cánulas nasales en el paciente adulto, pediátrico y neonatal con insuficiencia respiratoria aguda grave [Article in Spanish]. Med Intensiva 2021 Jun;45(5):298-312. [doi: 10.1016/j.medin.2020.08.016]
34. Bello G, De Santis P, Antonelli M. Non-invasive ventilation in cardiogenic pulmonary edema. Ann Transl Med 2018 Sep;6(18):355. [doi: 10.21037/atm.2018.04.39]
35. Manual 12 de Octubre [Article in Spanish]. MSD Profesionales Sanitarios. URL: https://profesionales.msd.es/manual-12-de-octubre/ [accessed 2025-01-03]
36. Manual de actuación clínica en las unidades medicalizadas de emergencia [Article in Spanish]. Portal de Salud de la Junta de Castilla y León. URL: https://www.saludcastillayleon.es/institucion/es/biblioteca/materiales-consejeria-sanidad/buscador/manual-actuacion-clinica-unidades-medicalizadas-emergencia [accessed 2025-01-03]
37. Población de la Comunidad de Madrid (según Censo), por sexo y edad (simple y quinquenal). Series desde el año 1900 | Data bank [Population of the community of Madrid (according to the census), by sex and age (simple and five-year). Series since 1900] [Article in Spanish]. Instituto de Estadística. URL: https://iestadis.edatos.io/statistical-visualizer/visualizer/data.html?resourceType=dataset&agencyId=IECM&resourceId=006_000008&version=~latest#visualization/table [accessed 2025-01-03]
38. Precios públicos por la prestación de los servicios y actividades de naturaleza sanitaria. tasas y precios sanitarios de la comunidad de madrid. BOCM A) disposiciones generales—BOCM https://www.BOCM.es. anexo II asistencia sanitaria en atención primaria, actividad no asistencial, transporte sanitario y asistencia sanitaria móvil en urgencias y emergencias [Article in Spanish]. Comunidad de Madrid. URL: https://sede.comunidad.madrid/medias/precios-publicos-anexos-iiiiii-iv/download [accessed 2025-01-03]
39. UVI Móvil [Article in Spanish]. SUMMA 112. URL: https://www.comunidad.madrid/hospital/summa112/ciudadanos/medios/uvi-movil [accessed 2025-01-03]
40. Impacto socioeconómico del coronavirus (COVID-19) y la cuarentena en España [Article in Spanish]. Statista. URL: https://es.statista.com/estudio/73004/el-coronavirus-covid-19-y-su-impacto-economico-y-social-en-espana/ [accessed 2025-01-03]
41. Briseño DM, Plata MRF, Figueroa M, Padilla RP. La carga económica de la EPOC. Análisis de los costos a nivel internacional [Article in Spanish]. Neumol Cir Torax 2011;70 [FREE Full text]
42. Jansson SA, Backman H, Stenling A, Lindberg A, Rönmark E, Lundbäck B. Health economic costs of COPD in Sweden by disease severity—has it changed during a ten years period? Respir Med 2013 Dec;107(12):1931-1938. [doi: 10.1016/j.rmed.2013.07.012] [Medline: 23910072]
43. Gutiérrez Villegas C, Paz-Zulueta M, Herrero-Montes M, Parás-Bravo P, Madrazo Pérez M. Cost analysis of chronic obstructive pulmonary disease (COPD): a systematic review. Health Econ Rev 2021 Dec;11(1):31. [doi: 10.1186/s13561-021-00329-9]

XSL•FO
RenderX

44. Pincombe A, Schultz TJ, Hofmann D, Karnon J. Economic evaluation of a medical ambulatory care service using a single group interrupted time-series design. J Eval Clin Pract 2023 Mar;29(2):329-340. [doi: 10.1111/jep.13771] [Medline: 36156337]
45. Rehman AU, Hassali MAA, Muhammad SA, Harun SN, Shah S, Abbas S. The economic burden of chronic obstructive pulmonary disease (COPD) in Europe: results from a systematic review of the literature. Eur J Health Econ 2020 Mar;21(2):181-194. [doi: 10.1007/s10198-019-01119-1] [Medline: 31564007]
46. Laurendeau C, Chouaid C, Roche N, Terrioux P, Gourmelen J, Detournay B. Costs of COPD in France: a national database analysis. Value Health 2013 Nov;16(7):A370. [doi: 10.1016/j.jval.2013.08.273]
47. Memorias SUMMA 112 [Article in Spanish]. SUMMA 112. URL: https://www.comunidad.madrid/hospital/summa112/nosotros/memorias-summa-112 [accessed 2025-01-03]
48. Garcia-Carretero R, Vazquez-Gomez O, Gil-Prieto R, Gil-de-Miguel A. Hospitalization burden and epidemiology of the COVID-19 pandemic in Spain (2020-2021). BMC Infect Dis 2023 Jul 18;23(1):476. [doi: 10.1186/s12879-023-08454-y] [Medline: 37464303]
49. Emami A, Javanmardi F, Pirbonyeh N, Akbari A. Prevalence of underlying diseases in hospitalized patients with COVID-19: a systematic review and meta-analysis. Arch Acad Emerg Med 2020;8(1):e35. [Medline: 32232218]
50. Harjola P, Tarvasmäki T, Barletta C, et al. The emergency department arrival mode and its relations to ED management and 30-day mortality in acute heart failure: an ancillary analysis from the EURODEM study. BMC Emerg Med 2022 Feb 14;22(1):27. [doi: 10.1186/s12873-022-00574-z] [Medline: 35164693]
51. Becher PM, Lund LH, Coats AJS, Savarese G. An update on global epidemiology in heart failure. Eur Heart J 2022 Aug 21;43(32):3005-3007. [doi: 10.1093/eurheartj/ehac248] [Medline: 35578978]
52. Chenaitia H, Massa H, Noel C, et al. Paramedics in prehospital emergency medical systems across Europe. Int Paramed Pract 2011 Sep 5;1(1):33-39. [doi: 10.12968/ippr.2011.1.1.33]
53. Sanitäts- und ambulanzdienste [Article in German]. Johanniter Österreich. URL: https://www.johanniter.at/hilfe-angebote/sanitaets-und-ambulanzdienste [accessed 2025-01-03]
54. Dúason S, Ericsson C, Jónsdóttir HL, Andersen JV, Andersen TL. European paramedic curriculum-a call for unity in paramedic education on a European level. Scand J Trauma Resusc Emerg Med 2021 May 31;29(1):72. [doi: 10.1186/s13049-021-00889-z] [Medline: 34059084]
55. Barroeta-Urquiza J, Boada-Bravo N. Los Servicios de Emergencia y Urgencias Médicas Extrahospitalarias En España [Emergency Services and Out-of-Hospital Medical Emergencies in Spain] [Book in Spanish]: Mensor; 2011. URL: https://www.madrid.es/UnidadesDescentralizadas/Emergencias/Samur-PCivil/Samur/ApartadosSecciones/10_DescargasPublicacionesWebsInteres/Ficheros/Los_SEM_en_Espana.pdf [accessed 2024-03-15]
56. 2023 GOLD report. Global Initiative for Chronic Obstructive Lung Disease. URL: https://goldcopd.org/2023-gold-report-2/ [accessed 2025-01-03]
57. Ancochea J, Soriano JB. La EPOC en España al inicio de una nueva década [Article in Spanish]. Arch Bronconeumol 2021 Jan;57(1):1-2. [doi: 10.1016/j.arbres.2020.01.025]
58. Bueno H, Bernal JL, Jiménez-Jiménez V, et al. The clinical outcomes, healthcare resource utilization, and related costs (COHERENT) model. Application in heart failure patients. Rev Esp Cardiol (Engl Ed) 2022 Jul;75(7):585-594. [doi: 10.1016/j.rec.2021.08.009] [Medline: 34688580]
59. Sneath E, Tippett V, Bowman RV, et al. The clinical journey of patients with a severe exacerbation of chronic obstructive pulmonary disease (COPD): from the ambulance to the emergency department to the hospital ward. J Thorac Dis 2022 Dec;14(12):4601-4613. [doi: 10.21037/jtd-22-328] [Medline: 36647501]

Abbreviations

ALS: advanced life support
APE: acute pulmonary edema
ARF: acute respiratory failure
CAM: Autonomous Community of Madrid
CHF: congestive heart failure
COPD: chronic obstructive pulmonary disease
CRF: chronic respiratory failure
ECOPD: exacerbation of chronic obstructive pulmonary disease
EMR: electronic medical record
EMS: emergency medical service
STROBE: Strengthening the Reporting of Observational studies in Epidemiology
SUMMA112: Servicio de Urgencias Medicas de Madrid (Prehospital Emergency Medical Service of Autonomous Community of Madrid)
SUMMIRA: Acute Respiratory Failure Treated by the Prehospital Medical Emergency Service of Autonomous Community of Madrid

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 05.09.24; peer-reviewed by R Brown, S Moitra; revised version received 15.11.24; accepted 17.11.24; published 16.01.25.

Please cite as:
Cintora-Sanz AM, Horrillo-García C, Quesada-Cubo V, Pérez-Alonso AM, Gutiérrez-Misis A
Prevalence and Economic Impact of Acute Respiratory Failure in the Prehospital Emergency Medical Service of the Madrid Community: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e66179
URL: https://publichealth.jmir.org/2025/1/e66179
doi:10.2196/66179

© Ana María Cintora-Sanz, Cristina Horrillo-García, Víctor Quesada-Cubo, Ana María Pérez-Alonso, Alicia Gutiérrez-Misis. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Trends in Early-Onset Colorectal Cancer in Singapore: Epidemiological Study of a Multiethnic Population

Hui Lionel Raphael Chen[1,2], MBBS, MRCS, FRCS; Qingqing Dawn Chong[2,3], MBBS, MRCP, MPH; Brenda Tay[3], BSc; Siqin Zhou[4], BSc, MSc; Evelyn Yi Ting Wong[2,3], MBBS, MRCP, MMed, FAMS; Isaac Seow-En[1,2], MBBS, MRCS, FRCS; Ker Kan Tan[5,6], MBBS, MMED, FRCS, PhD; Yi Wang[6], PhD; Adeline Seow[6], MBBS, MMED, FFPH; Kwong-Wei Emile Tan[1,2], BSc, MBBS, PhD, FRCS; Bee Huat Iain Tan[2,3], MBBS, MRCP, PhD; Sze Huey Tan[4], PhD

[1]Department of Colorectal Surgery, Singapore General Hospital, Singapore, Singapore

[2]Duke-NUS Medical School, Singapore, Singapore

[3]Division of Medical Oncology, National Cancer Centre, Singapore, Singapore

[4]Division of Clinical Trials and Epidemiological Sciences, National Cancer Centre, Singapore, Singapore

[5]Division of Colorectal Surgery, Department of Surgery, National University Hospital, Singapore, Singapore

[6]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

Corresponding Author:
Hui Lionel Raphael Chen, MBBS, MRCS, FRCS
Department of Colorectal Surgery
Singapore General Hospital
20 College Road
Singapore, 169856
Singapore
Phone: 65 63214677
Email: lionel.raphael.chen.h@singhealth.com.sg

Abstract

Background: Colorectal cancer (CRC) incidence and mortality in those aged 50 years and above have decreased over the past 2 decades. However, there is a rising incidence of CRC among individuals under 50 years of age, termed early-onset colorectal cancer (EOCRC). Patients with EOCRC are diagnosed at an advanced stage and may be in more psychosocial, emotional, and financial distress.

Objective: Our study examined the epidemiological shifts in CRC in Singapore, a multiethnic country.

Methods: CRCs diagnosed at age 20 years and above were identified from the Singapore Cancer Registry (SCR) from 1968 to 2019. Patient characteristics included gender, ethnicity, and age of CRC diagnosis. Population information was obtained from the Department of Statistics Singapore (SingStat). Age-specific incidence rates (ASRs) and age-standardized incidence rates (ASIRs) were calculated. The cohort was divided into 3 age groups: 20-49, 50-64, and ≥65 years. Temporal trends in incidence rates were modeled with joinpoint regression. Birth cohort models were fitted using the National Cancer Institute (NCI) age-period-cohort analysis tool. Cancer-specific survival analysis was performed with the Cox proportional hazards model.

Results: In total, 53,044 CRCs were included, and 6183 (11.7%) adults aged 20-49 years were diagnosed with EOCRC. The ASR of EOCRC rose from 5 per 100,000 population in 1968 to 9 per 100,000 population in 1996 at 2.1% annually and rose to 10 per 100,000 population in 2019 at 0.64% annually. The ASR for CRC among adults aged 50-64 years rose at 3% annually from 1968 to 1987 and plateaued from 1987, while the ASR for adults aged 65 years and above rose at 4.1% annually from 1968 to 1989 and 1.3% annually from 1989 to 2003 but decreased from 2003 onwards at 1% annually. The ASR of early-onset rectal cancer increased significantly at 1.5% annually. There was a continued rise in the ASR of EOCRC among males (annual percentage change [APC] 1.5%) compared to females (APC 0.41%). Compared to the 1950-1954 reference birth cohort, the 1970-1984 birth cohort had a significantly higher incidence rate ratio (IRR) of 1.17-1.36 for rectal cancer, while there was no significant change for colon cancer in later cohorts. There were differences in CRC trends across the 3 ethnic groups: Malays had a rapid and persistent rise in the ASR of CRC across all age groups (APC 1.4%-3%), while among young Chinese, only the ASR of rectal cancer was increasing (APC 1.5%). Patients with EOCRC had better survival compared to patients diagnosed at 65 years and above (hazard ratio [HR] 0.73, 95% CI 0.67-0.79, $P<.001$) after adjusting for covariates.

XSL•FO
RenderX

Conclusions: The rise in the incidence of rectal cancer among young adults, especially among Chinese and Malays, in Singapore highlights the need for further research to diagnose CRC earlier and reduce cancer-related morbidity and mortality.

(JMIR Public Health Surveill 2025;11:e62835) doi:10.2196/62835

KEYWORDS

early-onset colorectal cancer; epidemiology; Singapore; joinpoint regression; age-period-cohort; public health; health disparity

Introduction

Colorectal cancer (CRC) presents a significant health care problem as it is the third-most common cancer diagnosed and the second-leading cause of death worldwide. In 2020, CRC accounted for 10% of the worldwide cancer incidence and 9% of cancer deaths. Currently, 2 million new cases are diagnosed annually worldwide, which is projected to rise to 3.2 million per annum by 2040 [1,2].

With the implementation of population-wide, average risk CRC screening for all adults aged 50 years or above, CRC incidence and mortality in this age group have progressively decreased over the past 2 decades. However, there is emerging evidence that the incidence of CRC in people under 50 years of age is rising, especially in high-income countries. Based on the World Health Organization–International Agency of Research on Cancer Incidence in Five Continents Plus database, the incidence of early-onset colorectal cancer (EOCRC) rose in 27 (54%) of 50 countries, with the highest incidence rates of 14-16 per 100,000 person-years seen in the United States, Australia, New Zealand, and South Korea [3]. In United States, since the 1990s, colon and rectal cancer incidence has been increasing at 1.3% and 2.3% per year, respectively, for adults aged 40-49 years [4]. This is in contrast to the declining incidence of CRC among adults aged 55 years or older. Similar trends in colon and rectal cancer incidence were observed in Australia, with an increase of 1.7% and 0.9% per year, respectively, for adults aged 40-49 years [5].

EOCRC may be associated with hereditary syndromes, such as Lynch syndrome and familial adenomatous polyposis, but these only account for 1%-3% of the total number of new CRC cases. The large majority of EOCRC cases are still sporadic and contribute 10%-12% of new CRC cases. EOCRC is more likely to occur in the left colon, with higher rates of mucinous and poorly differentiated cancers, which may be associated with poor prognosis [6]. Patients with EOCRC tend to possess an underappreciation of symptoms; more reluctance to seek medical assistance, leading to delayed diagnosis; and more advanced stage at diagnosis [7]. They present with unique challenges—as younger patients with young children, heavy financial commitments, and fertility preservation [8]. The rise in the incidence of EOCRC may be a consequence of increased endoscopic capacity, which may lower the threshold for performing colonoscopy in younger patients. However, another important factor is the strong birth cohort effect observed worldwide, which may be the result of changes in exposure to risk factors. Although most studies have found EOCRC prevalence to be independent of gender, racial disparities in EOCRC incidence have been described in the United States, with a higher incidence seen in African Americans [9,10].

Singapore has a multiethnic resident population of 4 million, of which 74% are Chinese, 13% Malays, and 9% Indians [11]. In Singapore, CRC has the highest incidence among all cancers for the past several years, which, therefore, imposes a substantial burden in terms of complications, mortality, health care resources, and medical costs [12]. According to the *Singapore Cancer Registry Annual Report 2021*, there were a total of 12,239 new cases of CRC diagnosed from 2017 to 2021. There was a notable rise in the age-standardized incidence rate (ASIR) of CRC from 19.4 per 100,000 population in 1968-1972 to 37.5 per 100,000 population in 2017-2021 among men. A similar trend among females was observed, at 15.4 per 100,000 population in 1968-1972 to 26.9 per 100,000 population in 2017-2021. Among those aged 40-49 years, CRC was the most common cause of cancer among men and the fourth-most common cause of cancer among women. There were differing trends in the ASIRs of cancer among the 3 main ethnic groups. Although the Chinese had the highest ASIR of cancer throughout the years, there is evidence of a closing ethnic gap in cancer incidence between the Chinese and Malays [13]. Although there has been an overall increase in survival rates for CRC over the past decades, the age-specific patterns and effects of gender and ethnicity on the incidence trends in CRC in Singapore are not well studied. Therefore, the aim of this study was to perform a comprehensive analysis of temporal trends in colon and rectal cancer incidence in Singapore.

Methods

Study Design and Data Source

The Singapore Cancer Registry (SCR) is a population-based registry that was established in 1968 and is part of the National Registry of Disease Office. The SCR collects notifications from medical practitioners about all patients with cancer diagnosed in Singapore, as mandated by the National Registry of Diseases Act 2007. The data are internationally comparable and valid, with completeness of about 97.5%. Annual audits of the data are performed to ensure data accuracy of at least 95% as the data are critical to guide cancer prevention policies and cancer research [14].

Patients aged 20 years and older diagnosed with colon cancer (*International Classification of Diseases* [ICD]-9 code 153, ICD-10 code C18) and rectal cancer (ICD-9 codes 154.0, 154.1; ICD-10 codes C19.9, C20.9) from 1968 to 2019 from the SCR were included in this study. Recurrent and metachronous CRCs, premalignant tumors, carcinoma in situ, extranodal colorectal lymphoma, and sarcomas were excluded. Information about the Singapore population was obtained from the Department of Statistics Singapore (SingStat), which provides publicly available data that were used to define the denominator for incidence rates [11]. Patient characteristics included gender,

XSL•FO
RenderX

ethnicity, and age at diagnosis. The tumor, node, and metastasis (TNM) stage at diagnosis was available for patients diagnosed from 2003 onward.

Colon cancers were separated from cancers of the rectosigmoid and rectum, which were grouped together as rectal cancer for analysis. Age-specific incidence rates (ASRs) were calculated as the number of new CRC cases diagnosed each year divided by the population at risk for that age group. The population at risk was defined as the group of individuals susceptible to CRC during the period of interest. ASIRs were calculated as the sum of weighted ASRs using the direct method and the Segi-Doll World Standard as the reference population.

Joinpoint Analysis

The cohort was initially divided into 3 age groups: 20-49, 50-64, and ≥65 years. EOCRC was defined as CRC diagnosed at age 20-49 years. Annual incidence rates from 1968 to 2019 were estimated for the 3 age groups. Joinpoint regression analysis was used to describe the temporal trends in cancer incidence, identify points of change in direction, and estimate the magnitude of the change in rates. Joinpoint analysis estimates the annual percentage change (APC) by assessing observed data and fitting a series of joined straight lines to determine the APC over each period. The changes are tested for statistical significance [15,16]. It assumes that the regression function is piecewise linear, and the segments are continuously connected at unknown change points [17]. Compared to the traditional Monte Carlo permutation method, the data-driven weighted Bayes information criterion (BIC) was used for model selection as it was computationally more efficient yet produced similar results [18].

Age-Period-Cohort Modeling

Birth cohort models were fitted using the National Cancer Institute (NCI) Age Period Cohort web tool [19]. Age-period-cohort modeling provides estimates of parameters that describe relationships between observed incidence rates and age, calendar period, and birth cohort based on age groups and time periods of equal length. There was no information available for those aged ≥70 years for several time points from SingStat population estimates. Therefore, age-period-cohort analysis was limited to those aged 20-69 years. Input data were CRC cases and population counts for 10 five-year time periods (1970-1974, 1975-1979, …, 2015-2019), 10 five-year age groups (20-24, 25-29, …, 65-69 years), and 19 birth cohorts generated starting from the 1905-1909 to the 1995-1999 birth cohort. Cohort effects were presented as incidence rate ratios (IRRs) for a given birth cohort, with the 1950-1954 as the reference birth cohort. Additionally, local drifts estimated the age-specific net APC in incidence rates across the different age groups.

Survival Analysis

Cancer-specific survival was defined as the period from the date of CRC diagnosis to the date of death due to CRC. Patients who died of other causes were censored at their date of death; similarly, patients who were still alive were censored at the date of last follow-up. Clinical characteristics, such as surgical, radiation, and chemotherapy treatments received, were available. Multivariable survival analysis was performed with the Cox proportional hazards model.

Ethical Considerations

The study was approved by the SingHealth Institutional Review Board (IRB 2022/2415) to access and analyze data from the SCR. Informed consent was waived (waiver of consent was approved by the SingHealth IRB) as study data were deidentified.

Results

Characteristics of the Study Population

In total, 53,044 CRCs were included, with 32,880 (62%) colon and 20,164 (38%) rectal cancers. Figure 1 describes the flowchart of identifying and extracting patients from the SCR. Furthermore, 6183 (11.7%) EOCRCs were diagnosed, with 3585 (58%) colon and 2598 (42%) rectal cancers. Clinical characteristics of the study population are described in Table 1. There was a higher proportion of stage 3 (37.4% vs 32.4%) and stage 4 (26.1% vs 24%) CRCs and a lower proportion of stage 2 (21.2% vs 28.1%) CRCs among EOCRCs compared to patients diagnosed at an age of 65 years and above (P<.001).

XSL•FO
RenderX

Figure 1. Flowchart of the patients identified and extracted from the SCR. ICD: *International Classification of Diseases*; SCR: Singapore Cancer Registry.

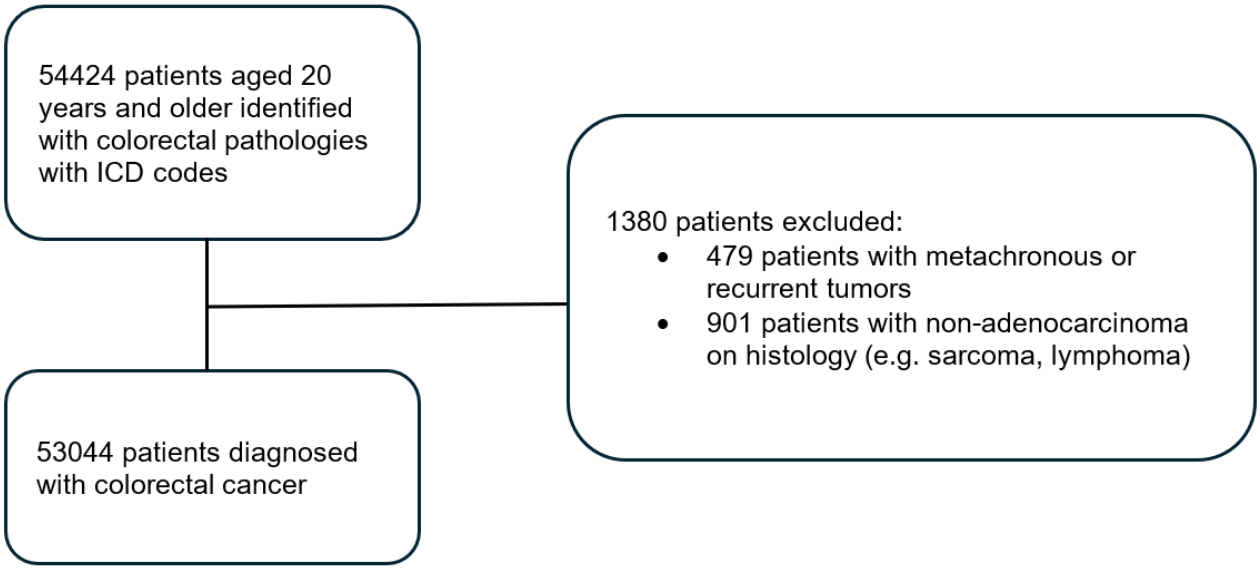


Table 1. Characteristics of CRCs[a] from the SCR[b].

Characteristics	Total cases (N=53,044), n (%)	Group 1 (20-49 years; n=6183), n (%)	Group 2 (50-64 years; n=17,862), n (%)	Group 3 (≥65 years; n=28,999), n (%)
Gender; *P*<.001				
Male	28,429 (53.6)	3235 (52.3)	10,245 (57.4)	14,949 (51.6)
Female	24,615 (46.4)	2948 (47.7)	7617 (42.6)	14,050 (48.4)
Ethnicity; *P*<.001				
Chinese	46,831 (88.3)	5167 (83.6)	15,495 (86.7)	26,169 (90.2)
Indian	1767 (3.3)	305 (4.9)	671 (3.8)	791 (2.7)
Malay	3723 (7.0)	583 (9.4)	1448 (8.1)	1692 (5.8)
Others	723 (1.4)	128 (2.1)	248 (1.4)	347 (1.2)
Tumor site; *P*<.001				
Colon	32,880 (62)	3585 (58.0)	10,332 (57.8)	18,963 (65.4)
Rectum	20,164 (38)	2598 (42.0)	7530 (42.2)	10,036 (34.6)
TNM[c] stage (2003 and beyond); *P*<.001				
1	4436 (15.9)	397 (15.3)	1643 (16.6)	2396 (15.6)
2	7178 (25.7)	551 (21.2)	2299 (23.2)	4328 (28.1)
3	9495 (34)	969 (37.4)	3534 (35.6)	4992 (32.4)
4	6830 (24.4)	677 (26.1)	2448 (24.7)	3705 (24.0)

[a]CRC: colorectal cancer.

[b]SCR: Singapore Cancer Registry.

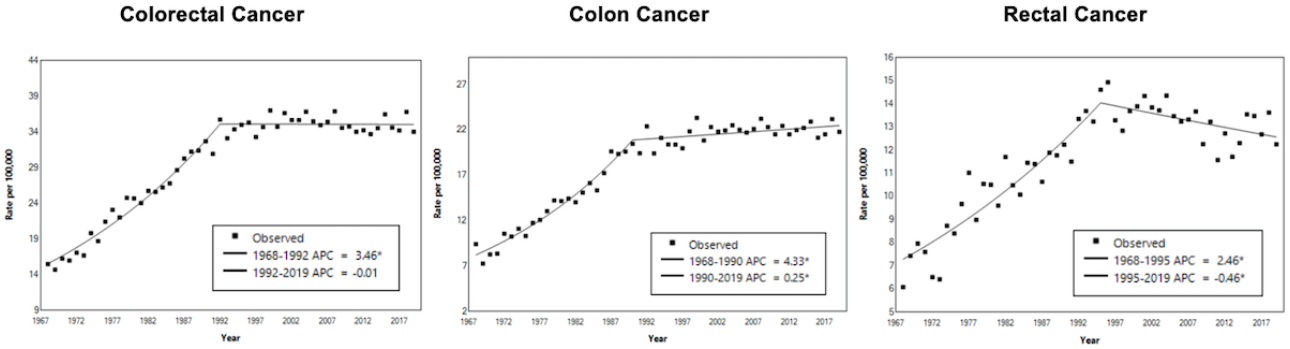
[c]TNM: tumor, node, and metastasis.

Joinpoint Regression Analysis of Incidence Rates of Colorectal, Colon, and Rectal Cancers

The ASIR of CRC rose from 15 in 1968 to 36 per 100,000 population in 1992 at an APC of 3.5% and plateaued (Figure 2). The ASIR of colon cancer rose from 8 in 1968 to 20 per 100,000 population in 1990 at an APC of 4.3% and plateaued. The ASIR of rectal cancer rose from 7 in 1968 to 14 per 100,000 population in 1990 at an APC of 2.5% and decreased to 12 per 100,000 population in 2019 at an APC of –0.5%.

XSL•FO
RenderX

Figure 2. Joinpoint regression of the ASIR of CRC, colon cancer, and rectal cancer. APC: annual percentage change; ASIR: age-standardized incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



The results of joinpoint regression are shown in Table 2 and Figure 3 for CRC, colon cancer, and rectal cancer stratified by age group. For EOCRC, the ASR rose from 5 in 1968 to 9 in 1997 at an APC of 2% and increased to 10 per 100,000 population in 2019 at an APC of 0.6%. For those aged 50-64 years, trends were similar to the ASIR of CRC for the overall Singapore population. For those aged 65 years and above, the ASR of CRC rose from 113 in 1968 to 299 in 2002 at an APC of 4.11% and decreased to 242 per 100,000 population in 2019 at an APC of –1.03%.

The ASR of early-onset colon cancer rose from 3 in 1968 to 5 per 100,000 population in 1995 at an APC of 2.2% and thereafter had a nonsignificant increase at an APC of 0.4%. For those aged 50-64 years, colon cancer trends were similar to the ASIR of CRC for the overall Singapore population. For those aged 65 years and above, the ASR of colon cancer rose from 57 in 1968 to 193 in 2005 at an APC of 4.95% and decreased to 170 per 100,000 population in 2019 at an APC of –0.92%.

The ASR of early-onset rectal cancer rose from 2 in 1968 to 5 per 100,000 population in 2019 at an APC of 1.5% per year. For those aged 50-64 years and 65 years and above diagnosed with rectal cancer, trends were similar to the ASIR of rectal cancer for the overall Singapore population (Table 2 and Figure 3).

Table 2. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] by age group and tumor location. See Multimedia Appendix 1 for a high-resolution version.

Type of cancer and age groups	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC							
20-49 years	6183 (11.6)	1968-1996	2.07[d]	1996-2019	0.64[d]	—[e]	—
50-64 years	17,862 (33.7)	1968-1987	2.97[d]	1987-2019	0.09	—	—
≥65 years	28,999 (54.7)	1968-1989	4.11[d]	1989-2003	1.30[d]	2003-2019	–1.03[f]
Colon cancer							
20-49 years	3585 (10.9)	1968-1995	2.16[d]	1995-2019	0.37	—	—
50-64 years	10,332 (31.4)	1968-1986	3.45[d]	1986-2019	0.29[d]	—	—
≥65 years	18,963 (57.7)	1968-1989	4.95[d]	1989-2005	1.30[d]	2005-2019	–0.92[f]
Rectal cancer							
20-49 years	2598 (12.9)	1968-2019	1.49[d]	—	—	—	—
50-64 years	7530 (37.3)	1968-1992	1.97[d]	1992-2019	–0.34[f]	—	—
≥65 years	10,036 (49.8)	1968-1996	2.67[d]	1996-2019	–1.15[f]	—	—

[a]APC: annual percentage change.

[b]CRC: colorectal cancer.

[c]ASR: age-specific incidence rate.

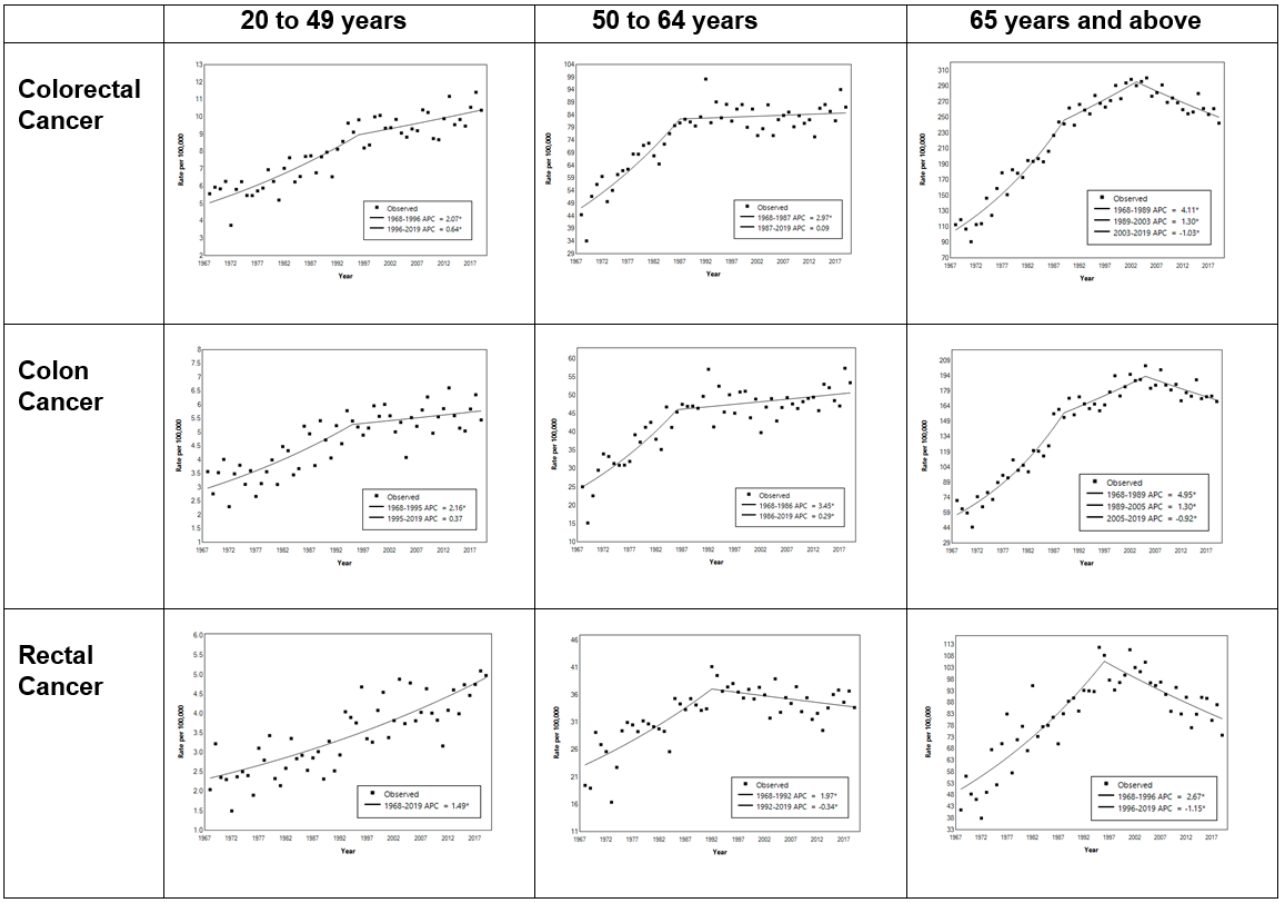
[d]Statistically significant positive APCs.

[e]Not applicable.

[f]Statistically significant negative APCs.

XSL•FO
RenderX

Figure 3. Joinpoint regression of the ASRs of CRC, colon cancer, and rectal cancer by age group and tumor location. APC: annual percentage change; ASR: age-specific incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Effect of Gender

The male population had higher incidence rates for colon and rectal cancers. Similar to the ASIR, for both genders aged 50-64 years and 65 years and above, the incidence of CRC, colon cancer, and rectal cancer rose before it plateaued or decreased in the recent 2 decades (Table 3 and Figures 4 and 5). Among those aged 50-64 years, there was a rapid initial rise in colon cancer among women compared to men (APC 5% vs 3.8%).

For males aged 20-49 years, there was a rising incidence in CRC (APC 1.5%), mainly rectal cancer (APC 1.9%), with a smaller rise in colon cancer (APC 1.1%). For females aged 20-49 years, similar to older females aged 50-64 years, there was a rapid initial rise, particularly in colon cancer (APC 3.8%); however, in recent years, for colon cancer, rectal cancer, and CRC, the APC has been <1% (Table 3 and Figure 5).

XSL•FO
RenderX

Table 3. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] by age group, tumor location, and gender.

Type of cancer and age groups	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC (20-49 years)							
Male	3235 (52.3)	1968-2019	1.48[d]	—[e]	—	—	—
Female	2948 (47.7)	1968-1994	2.10[d]	1994-2019	0.41	—	—
CRC (50-64 years)							
Male	10,245 (57.4)	1968-1989	3.34[d]	1989-2019	0.11	—	—
Female	7617 (42.6)	1968-1980	4.20[d]	1980-2019	0.04	—	—
CRC (≥65 years)							
Male	14,949 (51.6)	1968-1989	3.76[d]	1989-2003	1.79	2003-2019	–0.95
Female	14,050 (48.4)	1968-1981	6.18[d]	1981-1999	2.17[d]	1999-2019	–1.01[f]
Colon cancer (20-49 years)							
Male	1715 (47.8)	1968-2019	1.10[d]	—	—	—	—
Female	1870 (52.2)	1968-1986	3.81[d]	1986-2019	0.55[d]	—	—
Colon cancer (50-64 years)							
Male	5471 (53.0)	1968-1988	3.76[d]	1988-2019	0.34	—	—
Female	4861 (47.0)	1968-1980	4.99[d]	1980-2019	0.19	—	—
Colon cancer (≥65 years)							
Male	9289 (49.0)	1968-1988	4.92[d]	1988-2005	1.64	2005-2019	–0.77
Female	9674 (51.0)	1968-1989	5.04[d]	1989-2004	1.18[d]	2004-2019	–1.04[f]
Rectal cancer (20-49 years)							
Male	1520 (58.5)	1968-2019	1.91[d]	—	—	—	—
Female	1078 (41.5)	1968-2019	0.93[d]	—	—	—	—
Rectal cancer (50-64 years)							
Male	4774 (63.4)	1968-1989	2.61[d]	1992-2019	–0.27	—	—
Female	2756 (36.6)	1968-2019	0	—	—	—	—
Rectal cancer (≥65 years)							
Male	5660 (56.4)	1968-2002	2.18[d]	2002-2019	–1.59[f]	—	—
Female	4376 (43.6)	1968-1995	2.69[d]	1995-2004	–1.55[f]	—	—

[a]APC: annual percentage change.

[b]CRC: colorectal cancer.

[c]ASR: age-specific incidence rate.

[d]Statistically significant positive APCs.

[e]Not applicable.

[f]Statistically significant negative APCs.

XSL•FO
RenderX

Figure 4. Joinpoint regression of ASRs of CRC, colon cancer, and rectal cancer by gender (male: red; female: blue). APC: annual percentage change; ASR: age-specific incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.

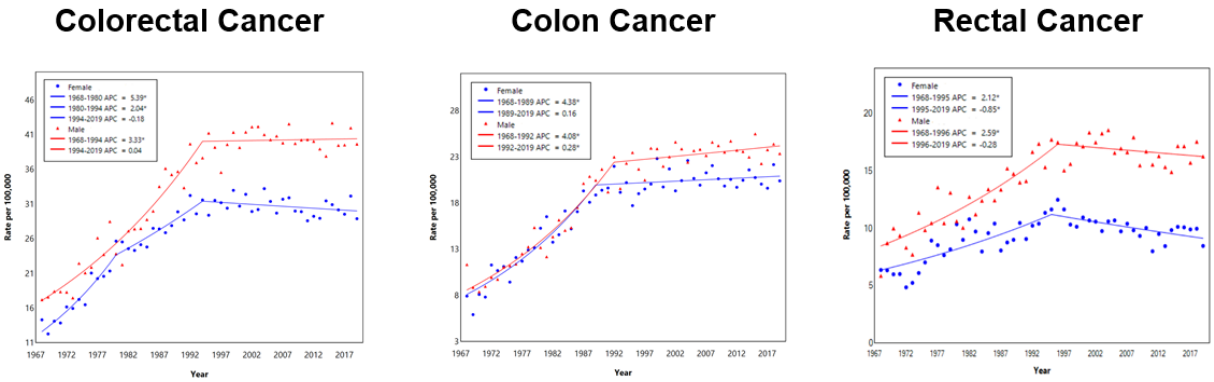
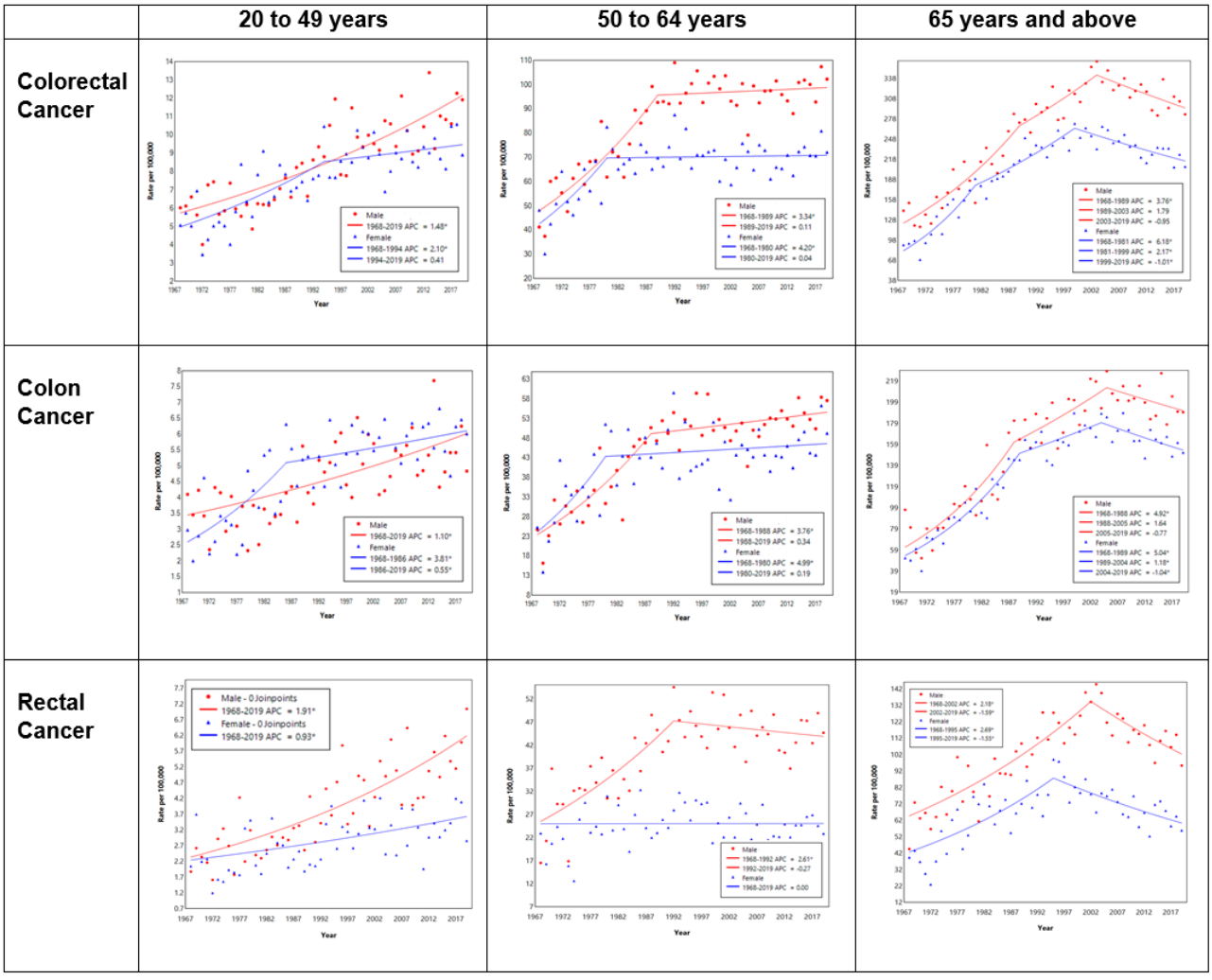


Figure 5. Joinpoint regression of ASRs of CRC, colon cancer, and rectal cancer by age group and gender (male: red; female: blue). APC: annual percentage change; ASR: age-specific incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Effect of Ethnicity

Similar to the ASIR, for Chinese patients aged 50-64 years and 65 years and above, the ASR of CRC, colon cancer, and rectal cancer showed an initial rise but plateaued or decreased in the recent 2 decades. In Malays across all 3 age groups, there was a persistent rise in the ASR of CRC, colon cancer, and rectal cancer (APC 1.4%-3.2%) without evidence of it plateauing. In Indians aged 50-64 years and 65 years and above, there was a smaller persistent rise in the ASR of CRC, colon cancer, and rectal cancer (APC 1%-1.4%) without evidence of it plateauing (Tables 4-6 and Figures 6 and 7).

XSL•FO
RenderX

Table 4. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] for patients diagnosed at age 20-49 years by tumor location and ethnicity.

Tumor location and ethnicity	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC							
Chinese	5167 (85.3)	1968-1999	1.96[d]	1999-2019	0.23	—[e]	—
Indian	305 (5.1)	1968-2019	0.47	—	—	—	—
Malay	583 (9.6)	1968-2019	2.90[d]	—	—	—	—
Colon cancer							
Chinese	2992 (85.4)	1968-1999	1.75[d]	1999-2019	–0.09	—	—
Indian	164 (4.7)	1968-2019	0.03	—	—	—	—
Malay	346 (9.9)	1968-2019	1.96[d]	—	—	—	—
Rectal cancer							
Chinese	2175 (85.2)	1968-2019	1.48[d]	—	—	—	—
Indian	141 (5.5)	1968-2019	0.45	—	—	—	—
Malay	237 (9.3)	1968-2019	1.96[d]	—	—	—	—

[a]APC: annual percentage change.
[b]CRC: colorectal cancer.
[c]ASR: age-specific incidence rate.
[d]Statistically significant positive APCs.
[e]Not applicable.

Table 5. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] for patients diagnosed at age 50-64 years by tumor location and ethnicity.

Tumor location and ethnicity	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC							
Chinese	15,495 (88.0)	1968-1985	3.45[d]	1985-2019	–0.12	—[e]	—
Indian	671 (3.8)	1968-2019	0.95[d]	—	—	—	—
Malay	1448 (8.2)	1968-2019	2.30[d]	—	—	—	—
Colon cancer							
Chinese	9012 (88.4)	1968-1981	5.57[d]	1981-2019	0.13	—	—
Indian	356 (3.5)	1968-2019	1.22[d]	—	—	—	—
Malay	823 (8.1)	1968-2019	2.69[d]	—	—	—	—
Rectal cancer							
Chinese	6483 (87.3)	1968-1988	2.31[d]	1988-2019	–0.42[f]	—	—
Indian	315 (4.3)	1968-2019	0.45	—	—	—	—
Malay	625 (8.4)	1968-2019	1.42[d]	—	—	—	—

[a]APC: annual percentage change.
[b]CRC: colorectal cancer.
[c]ASR: age-specific incidence rate.
[d]Statistically significant positive APCs.
[e]Not applicable.
[f]Statistically significant negative APCs.

XSL•FO
RenderX

Table 6. APC[a] of CRC[b], colon cancer, and rectal cancer ASRs[c] for patients diagnosed at age ≥65 years by tumor location and ethnicity.

Tumor location and ethnicity	Patients, n (%)	Trend 1		Trend 2		Trend 3	
		Years	APC, %	Years	APC, %	Years	APC, %
CRC							
Chinese	26,169 (91.3)	1968-1990	4.32[d]	1990-2002	1.35[d]	2002-2019	–1.47[e]
Indian	791 (2.8)	1968-2019	1.32[d]	—[f]	—	—	—
Malay	1692 (5.9)	1968-2019	2.97[d]	—	—	—	—
Colon cancer							
Chinese	17,254 (92.0)	1968-1990	5.21[d]	1990-2003	1.35[d]	2003-2019	–1.26[e]
Indian	483 (2.6)	1968-2019	1.38[d]	—	—	—	—
Malay	1006 (5.4)	1968-2019	3.24[d]	—	—	—	—
Rectal cancer							
Chinese	8915 (90.0)	1968-1996	2.93[d]	1996-2019	–1.60[e]	—	—
Indian	308 (3.1)	1968-2019	0.99[d]	—	—	—	—
Malay	686 (6.9)	1968-2019	1.88[d]	—	—	—	—

[a]APC: annual percentage change.

[b]CRC: colorectal cancer.

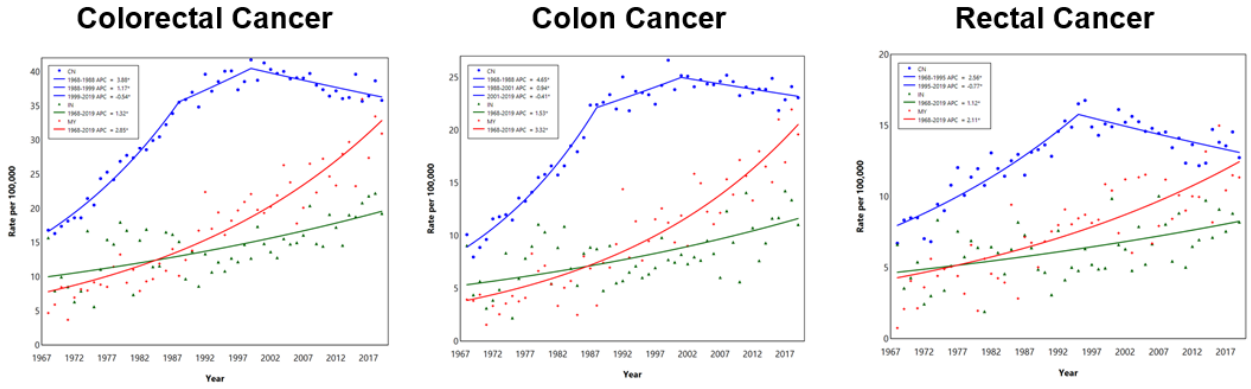
[c]ASR: age-specific incidence rate.

[d]Statistically significant positive APCs.

[e]Statistically significant negative APCs.

[f]Not applicable.

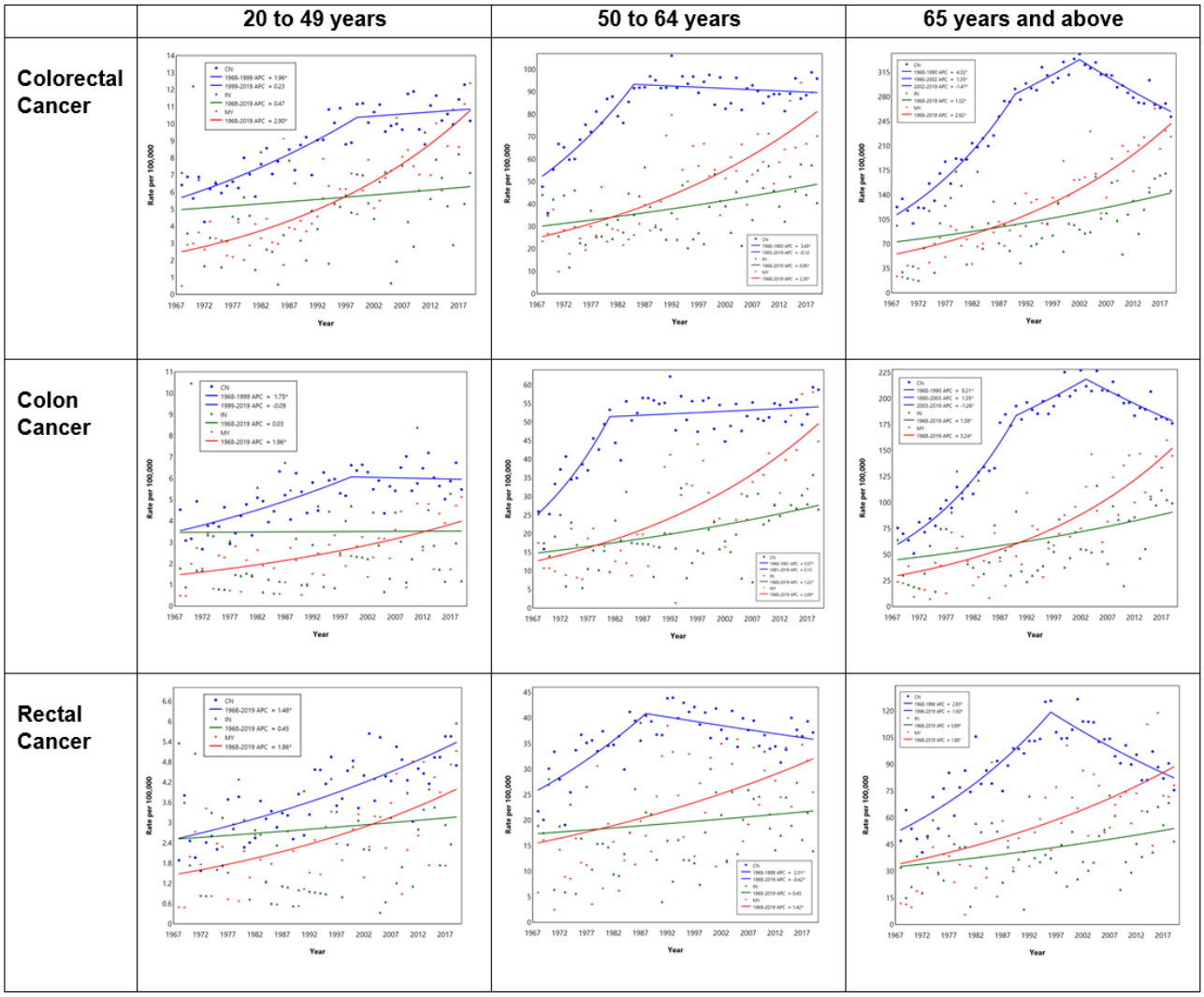
Figure 6. Joinpoint regression of ASIRs of CRC, colon cancer, and rectal cancer by ethnic group (Chinese/CN: blue; Indian/IN: green; Malay/MY: red). APC: annual percentage change; ASIR: age-standardized incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Malay patients aged 20-49 years had the highest rise in the ASR of CRC (APC 2.9%), colon cancer (APC 2%), and rectal cancer (APC 2%). Among Chinese patients aged 20-49 years, the ASR of rectal cancer persistently rose (APC 1.5%), while the ASR of CRC and colon cancer plateaued from 1999. In Indians aged 20-49 years, there was a small and nonsignificant rise in the ASR of CRC (APC 0.5%) and rectal cancer (APC 0.5%), as shown in Tables 4-6 and Figure 7.

XSL•FO
RenderX

Figure 7. Joinpoint regression of ASRs of CEC, colon cancer, and rectal cancer by age group, tumor location, and ethnic group (Chinese/CN: blue; Indian/IN: green; Malay/MY: red). APC: annual percentage change; ASR: age-specific incidence rate; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Age-Period-Cohort Model of IRRs and APCs for CRC, Colon Cancer, and Rectal Cancer

Table 7 and Figure 8 show the IRRs by birth cohort for CRC, colon cancer, and rectal cancer. Compared to the reference birth cohort of 1950-1954, there was a trend toward higher IRRs for colon cancer among the 1970-1990 birth cohorts; however, these changes were not statistically significant. In addition, compared to the reference birth cohort of 1950-1954, the IRRs for rectal cancer among the birth cohorts from 1905 onward were marginally lower. However, there was significant increase in the IRR for rectal cancer among the birth cohorts from 1970 onward, and this peaked in the 1980-1984 birth cohort (IRR 1.35, 95% CI 1.043-1.748). Subsequent birth cohorts from 1985 onward had similar IRRs as the reference birth cohort (Figure 7 and Table 7).

Figure 9 shows the net age-specific APC or local drift for CRC, colon cancer, and rectal cancer incidence rates. There was an increase in the incidence rates for CRC among those aged 25-69 years, with an average APC of 0.61% (95% CI 0.44-0.79) for all age groups over the entire period. There was an increase in incidence rates for colon cancer among those aged 40-69 years, while there was an increase in incidence rates for rectal cancer among those aged 35-69 years.

XSL•FO
RenderX

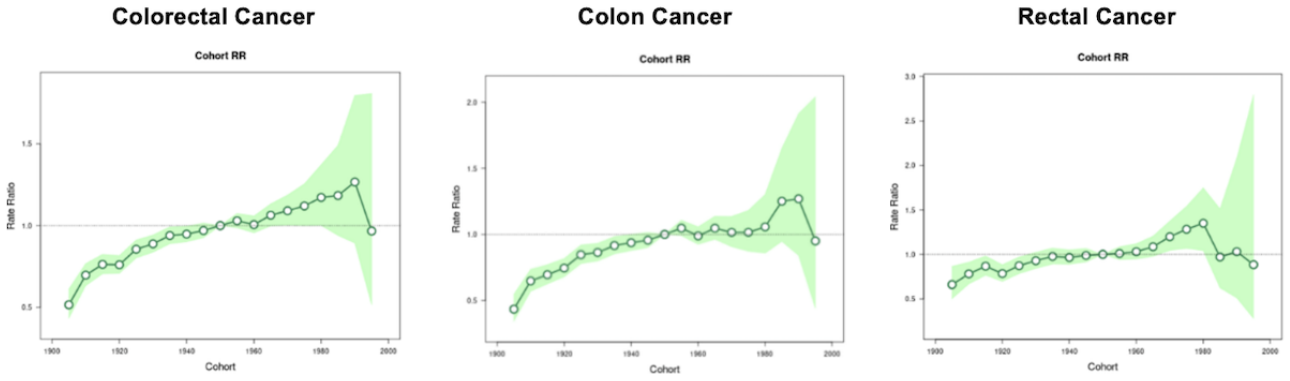
Table 7. IRRs[a] for CRC[b], colon cancer, and rectal cancer by birth cohort.

Birth cohort	CRC IRR (95% CI)	Colon cancer IRR (95% CI)	Rectal cancer IRR (95% CI)
1905-1909	0.516 (0.435-0.613)	0.434 (0.345-0.547)	0.660 (0.503-0.867)
1910-1914	0.697 (0.633-0.768)	0.648 (0.570-0.736)	0.781 (0.668-0.912)
1915-1919	0.763 (0.706-0.825)	0.696 (0.626-0.771)	0.867 (0.765-0.983)
1920-1924	0.761 (0.708-0.817)	0.745 (0.679-0.817)	0.784 (0.696-0.883)
1925-1929	0.856 (0.802-0.912)	0.846 (0.779-0.920)	0.873 (0.784-0.972)
1930-1934	0.889 (0.837-0.943)	0.863 (0.799-0.933)	0.929 (0.842-1.025)
1935-1939	0.939 (0.889-0.993)	0.916 (0.852-0.984)	0.977 (0.892-1.070)
1940-1944	0.948 (0.901-0.997)	0.937 (0.878-1.000)	0.966 (0.889-1.050)
1945-1949	0.970 (0.926-1.015)	0.956 (0.901-1.015)	0.988 (0.915-1.067)
1950-1954	1.000 (1.000)	1.000 (1.000)	1.000 (1.000)
1955-1959	1.030 (0.986-1.075	1.047 (0.991-1.107)	1.010 (0.940-1.086)
1960-1964	1.006 (0.956-1.059))	0.989 (0.925-1.057)	1.030 (0.947-1.120)
1965-1969	1.063 (0.998-1.133)	1.047 (0.964-1.137)	1.086 (0.978-1.205)
1970-1974	1.091 (1.002-1.187)	1.015 (0.907-1.136)	1.199 (1.045-1.375)
1975-1979	1.120 (0.999-1.255)	1.016 (0.873-1.182)	1.28 (1.067-1.541)
1980-1984	1.172 (1.001-1.372)	1.057 (0.858-1.302)	1.35 (1.043-1.748)
1985-1989	1.184 (0.940-1.491)	1.251 (0.947-1.652)	0.971 (0.623-1.512)
1990-1994	1.266 (0.893-1.795)	1.269 (0.841-1.914)	1.031 (0.511-2.080)
1995-1999	0.967 (0.517-1.807)	0.951 (0.443-2.040)	0.884 (0.280-2.792)

[a]IRR: incidence rate ratio.

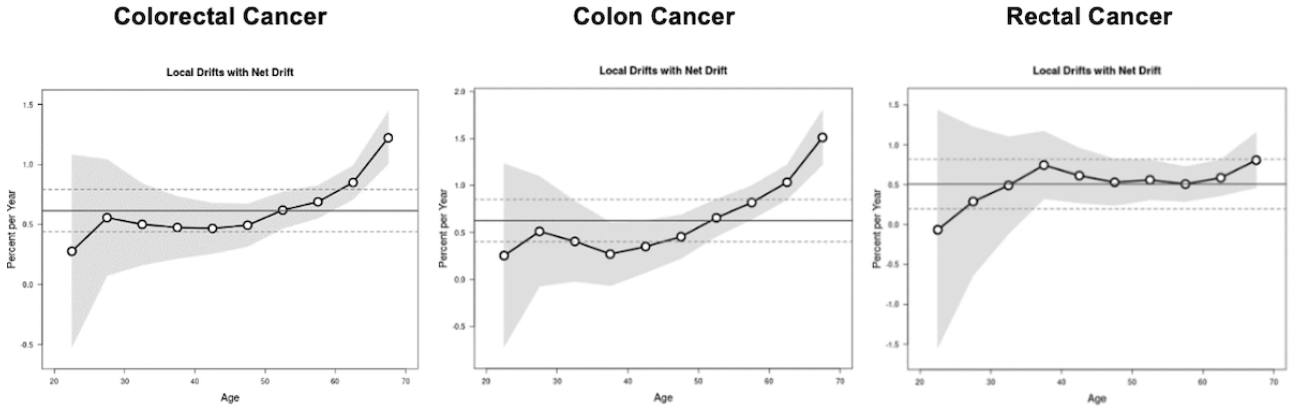
[b]CRC: colorectal cancer.

Figure 8. IRRs by birth cohort for CRC, colon cancer, and rectal cancer. CRC: colorectal cancer; IRR: incidence rate ratio. See Multimedia Appendix 1 for a high-resolution version.



XSL•FO
RenderX

Figure 9. Age-specific APC (local drift) for CRC, colon cancer, and rectal cancer. APC: annual percentage change; CRC: colorectal cancer. See Multimedia Appendix 1 for a high-resolution version.



Survival Outcomes

Patients with EOCRC had better cancer-specific survival compared to patients diagnosed at 65 years and above (hazard ratio [HR] 0.73, 95% CI 0.67-0.79, *P*<.001) after adjusting for the TNM stage, gender, tumor site, surgery, radiotherapy, and chemotherapy (Table 8).

Table 8. Multivariable cancer-specific survival analysis of patients with CRC[a] using the Cox proportional hazards model.

Variables and groups	HR[b] (95% CI)	*P* value
Age (years)		
20-49 vs ≥65	0.73 (0.67-0.79)	<.001
50-64 vs ≥65	0.74 (0.70-0.78)	<.001
TMN[c] stage		
2 vs 1	3.84 (3.23-4.56)	<.001
3 vs 1	11.14 (9.43-13.16)	<.001
4 vs 1	58.70 (49.66-69.38)	<.001
Gender: male vs female	1.06 (1.02-1.11)	.006
Tumor site: rectum vs colon	0.55 (0.52-0.58)	<.001
Treatment		
Surgery vs no surgery	0.43 (0.41-0.46)	<.001
Radiotherapy vs no radiotherapy	1.10 (1.01-1.19)	<.026
Chemotherapy vs no chemotherapy	0.54 (0.51-0.57)	<.001

[a]CRC: colorectal cancer.

[b]HR: hazard ratio.

[c]TNM: tumor, node, and metastasis.

Discussion

Principal Findings

Historically, CRC has been more commonly associated with older age groups, but in recent years, many countries have reported a concerning rise in cases among younger individuals. Although the overall incidence of CRC has been declining in older people, there has been a notable increase in the incidence of EOCRC, particularly in individuals younger than 50 years. In the United States, the overall annual, CRC ASIR has decreased by 46% from 66.2 per 100,000 population in 1985 to 35.7 per 100,000 population in 2019. Notably, CRC is now the most common cause of cancer deaths among individuals less than 50 years old [20]. However, from 2011 through 2019, the CRC incidence rate increased by 1.9% per year for individuals younger than 50 years. Rectal cancer has driven the EOCRC epidemiological trend with an increase of 2% per year. There was a steep decline in incidence among adults aged 50 years and older in the United States at 3%-5% annually in the late 2000s, and this is likely attributed to changing patterns in the risk factors and uptake of CRC screening. Incidence rates have been increasing in adults aged 20-39 years since the mid-1980s and in those aged 40-54 years since the mid-1990s [21]. Similar trends are seen in Canada, the United Kingdom, Australia, New Zealand, Europe, and Asia [5,22-25]. The rising incidence of EOCRC in East Asia may be due to an increase in the prevalence of risk factors, such as smoking, alcoholism, and obesity among young individuals [26]. A population-based study in Taiwan, Korea, Japan, and Hong Kong suggests that

XSL•FO
RenderX

from 1995 to 2014, the pooled incidence of early-onset colon and rectal cancers increased by 2.1% and 3.8%, respectively, among men. Similarly, the pooled incidence of early-onset colon and rectal cancers increased by 2.2% and 4%, respectively, among women [27]. In China, the incidence rate of EOCRC rose from 3.6 to 12.1 per 100,000 population from 1990 to 2019 at a rate of 4.6% per year [25].

Our analysis of incidence trends over time suggests that the rising incidence of EOCRC in Singapore is nuanced. The rising trend for people under 50 years of age is particularly for rectal cancer, especially in males. This trend also affects, in particular, the Malays but also the Chinese, with the Indians less affected. Our study supports the increasing literature that the increase in EOCRC is not confined to the West and may be a worldwide phenomenon. The cause of these observations is still unclear, but it can be explained by the cohort effects. The IRR for rectal cancer has been increasing among the 1970-1984 birth cohorts and suggests that it is the changes in exposure to these birth cohorts that influence the risk of carcinogenesis. High consumption of processed meat, alcohol, a sedentary lifestyle, a lack of dietary fiber consumption, smoking, and obesity are known lifestyle factors associated with CRC. However, these modifiable risk factors are present in all age groups and not specific to young individuals and cannot explain fully the rising incidence among young people [28,29]. It has also been reported that there are similar rates of obesity between EOCRC and older patients with CRC [30]. The increasing prevalence of the use of antibiotics and alterations in the gut microbiome may play a role in the development of CRC through effects of bacterial-derived metabolites and virulence factors. Results of a recent study show an increase in *Fusobacterium* spp. in the microbiome of EOCRC [31]. Understanding the generational shifts in early life exposures may provide further insights into the etiologies of EOCRC [7].

Most tumors in EOCRC are in the sigmoid colon and rectum [32]. The reasons regarding the steep rise in the incidence of rectal compared to colon cancer among young adults are unclear. Understanding the differences in the molecular profile between rectal and colon cancers may be key toward developing preventative strategies. EOCRC in young adults has been associated with synchronous and metachronous tumors, a more poorly differentiated histology, and mucinous and signet ring histology compared to older patients [7,33]. The rates of high microsatellite instability (MSI-H) are higher in EOCRC due to the higher incidence of Lynch syndrome. However, among the microsatellite-stable (MSS) cohort of metastatic CRC, patients with EOCRC have lower rates of *KRAS* and *BRAF* V600 mutations [31]. These observations suggest that the underlying biological process and basis of these clinicopathological conditions in EOCRC are different from those found in people above 50 years old.

The male gender is associated with a higher incidence of CRC compared to females. The incidence of rectal cancer is higher among young males compared to females. In 2019, the incidence rate of early-onset rectal cancer was 6.2 per 100,000 population among men, while it was 3.6 per 100,000 population among women. The higher proportion of men diagnosed with rectal cancer suggests a differential distribution of the tumor location between genders. Although the etiology between the gender differences is uncertain, it may be attributed to the varied exposure in dietary and lifestyle risk factors between both gender groups [34]. Some studies have explored the protective effect of estrogens and progestins against CRC in hormone replacement therapy in postmenopausal women [35,36].

The Chinese population in Singapore had the highest ASIR of CRC when compared to Indians and Malays. Previously, the Asia Pacific Consensus identified Chinese as an ethnic group that is more susceptible to CRC, and this was supported by cohort studies in Singapore and Malaysia [37-40]. In this study, although the incidence initially increased among the Chinese, the incidence rates have plateaued and declined over the past decade, except for early-onset rectal cancer, which continues to rise at 1.5% per year among young Chinese adults. However, the steep increase in the incidence of both colon and rectal cancers among the Malays across all age groups is concerning. This observation is supported by another study of age-period-cohort analyses in the SCR, stratified by ethnicity, that suggests a rising risk of CRC across all periods among the Malay population [41]. The rising incidence of EOCRC predominantly among Malays was also observed in Malaysia [42]. The varied epidemiologic trends across the 3 ethnic groups may be attributed to the differences in health-seeking behavior and lifestyle exposures between the groups in Singapore [41]. Modifiable risk factors, such as the rising rates and high prevalence of obesity among the Malays and the increased consumption of sweetened food, may play a role [43]. Malays were also less likely to participate in cancer screening compared to the Chinese. There is a need for tailored cancer-screening promotion campaigns to narrow the knowledge-behavior gap among the Malays [44]. Socioeconomic indicators, such as educational attainment, homeownership rates, and income levels, have shown improvement across all racial groups in Singapore [45]. Although there is a relatively similar socioeconomic environment in Singapore, the varying rates of the rise of EOCRC among the 3 races suggest that cultural, dietary, or genetic factors may potentially contribute to these differences. Further studies should be conducted that compare the lifestyle factors and social practices among the 3 ethnic groups and evaluate the impact on CRC incidence.

Our study suggests that patients with EOCRC have better cancer-specific survival compared to patients with CRC aged 65 years and older. However, there are conflicting data with regard to the survival outcomes of EOCRC compared to patients with late-onset CRC after adjustment for stage. A survival analysis of 35,713 patients with stage 3 colon cancer from the ACCENT database revealed that patients with EOCRC have improved overall survival, disease-free survival, and survival after recurrence. After adjusting for molecular markers (MMR, KRAS, BRAF), the prognostic impact of age of onset was lost [46]. However, there was a poorer 3-year relapse-free survival among patients with stage 3 EOCRC based on the pooled analyses of 16,349 patients from the International Duration Evaluation of Adjuvant Chemotherapy (IDEA) database. This was despite better treatment adherence and higher administered treatment intensity among the younger patients, hence suggestive of a more aggressive disease biology [47].

XSL•FO
RenderX

The current screening guidelines for CRC in Singapore start at 50 years of age for asymptomatic, average-risk individuals and have no impact on the incidence of early-onset CRC. Recently, the American Cancer Society recommended initiating CRC screening at age 45 years instead of 50 years for average-risk individuals to address the rising CRC incidence among young adults [48]. This recommendation was based on modeling data that suggest that this will prevent 29,400 CRC cases and 11,000 CRC deaths in the United States over the next 5 years [49]. Compared to the rising incidence of colon and rectal cancers in the United States, Europe, and Australia, our study suggests that rectal cancer among young adults is rising. Therefore, it is important for physicians to investigate young adults who present with anorectal symptoms, such as rectal bleeding, tenesmus, anal pain, and changes in bowel habits. Although these findings may suggest a need to re-evaluate the recommended age range for screening and consider extending screening to those 45-49 years of age, the incidence of EOCRC in Singapore remains low, and this strategy may not be cost effective. Despite the launch of the National Colorectal Screening Programme in 2011, the uptake of CRC screening remains low nationally at 41.7% [50,51]. However, flexible sigmoidoscopy among young adults is a less costly alternative as compared to colonoscopy and may capture the majority of EOCRCs since the tumor location of EOCRC is usually in the sigmoid colon or the rectum.

Strengths and Limitations

Although this study provides a comprehensive overview of CRC trends in Singapore, further research into the biological basis of carcinogenesis among young adults and differences in modifiable risk factors among ethnic groups is warranted. Age-period-cohort analysis was limited due to missing population data of those aged 70 years and above. Additionally, the study was limited due to the lack of information about the proportion of hereditary cancer syndromes in the cohort. The prevalence of hereditary cancer syndromes among EOCRCs ranges from 5% to 35% compared to 2%-5% for CRCs overall [52-55]. Although genetic predisposition may be associated with EOCRC, the majority of EOCRCs are still sporadic. Additionally, screening of families with hereditary cancer syndromes may only capture a small proportion of EOCRCs [52]. Future research should focus on identification of risk factors and predictors of EOCRC to determine high-risk groups that should be targeted for screening before 50 years of age. It may be a more viable strategy to develop risk stratification tools that combine family history and blood-based tests to identify high-risk neoplasia among young adults.

Conclusion

The incidence of rectal cancer among people under 50 years of age in Singapore appears to be increasing, especially among the Chinese and Malays. The differences in the epidemiologic trends among the 3 main ethnic groups may be due to varied lifestyle factors and social practices. Although it is possible that the observed increase may be attributable to lifestyle changes in the population over time, further studies are required to determine the underlying etiology. In contrast, CRC rates are decreasing among those above 65 years of age, and this may be due to the impact of the national screening program. These findings suggest the need for further research to diagnose CRC earlier among young adults and reduce the associated cancer-related morbidity and mortality.

Acknowledgments

The authors would like to acknowledge the National Registry of Diseases Office, Health Promotion Board, Singapore, for providing access to colorectal cancer data from the Singapore Cancer Registry. CHLR is supported by the Research Training Fellowship (MOH-001300-00) from the National Medical Research Council Singapore. The work was also supported by the Goh Hak Su Professorship in Colorectal Surgery.

Data Availability

Data collected from the Singapore Cancer Registry are not available publicly.

Authors' Contributions

The authors attest that there was no use of generative artificial intelligence in the generation of text, figures, or other content of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
High-resolution versions of Figures 2-9.
[PDF File (Adobe PDF File), 1739 KB - publichealth_v11i1e62835_app1.pdf]

References

1. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl Oncol 2021 Oct;14(10):101174 [FREE Full text] [doi: 10.1016/j.tranon.2021.101174] [Medline: 34243011]

XSL•FO
RenderX

2. Rabeneck L, Chiu H, Senore C. International perspective on the burden of colorectal cancer and public health effects. Gastroenterology 2020 Jan;158(2):447-452. [doi: 10.1053/j.gastro.2019.10.007] [Medline: 31622620]
3. Sung H, Siegel RL, Laversanne M, Jiang C, Morgan E, Zahwe M, et al. Colorectal cancer incidence trends in younger versus older adults: an analysis of population-based cancer registry data. Lancet Oncol 2024 Dec 11:S1470-2045 [FREE Full text] [doi: 10.1016/S1470-2045(24)00600-4] [Medline: 39674189]
4. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. CA Cancer J Clin 2023;73(3):233-254 [FREE Full text] [doi: 10.3322/caac.21772] [Medline: 36856579]
5. Feletto E, Yu XQ, Lew J, St John DJB, Jenkins MA, Macrae FA, et al. Trends in colon and rectal cancer incidence in Australia from 1982 to 2014: analysis of data on over 375,000 cases. Cancer Epidemiol Biomarkers Prev 2019 Jan;28(1):83-90. [doi: 10.1158/1055-9965.EPI-18-0523] [Medline: 30530848]
6. Patel SG, Karlitz JJ, Yen T, Lieu CH, Boland CR. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. Lancet Gastroenterol Hepatol 2022 Mar;7(3):262-274. [doi: 10.1016/s2468-1253(21)00426-x]
7. Akimoto N, Ugai T, Zhong R, Hamada T, Fujiyoshi K, Giannakis M, et al. Rising incidence of early-onset colorectal cancer - a call to action. Nat Rev Clin Oncol 2021 Apr;18(4):230-243 [FREE Full text] [doi: 10.1038/s41571-020-00445-1] [Medline: 33219329]
8. Lieu CH, Golemis EA, Serebriiskii IG, Newberg J, Hemmerich A, Connelly C, et al. Comprehensive genomic landscapes in early and later onset colorectal cancer. Clin Cancer Res 2019 Oct 01;25(19):5852-5858 [FREE Full text] [doi: 10.1158/1078-0432.CCR-19-0899] [Medline: 31243121]
9. Saraiva MR, Rosa I, Claro I. Early-onset colorectal cancer: a review of current knowledge. World J Gastroenterol 2023 Feb 28;29(8):1289-1303 [FREE Full text] [doi: 10.3748/wjg.v29.i8.1289] [Medline: 36925459]
10. Pan H, Zhao Z, Deng Y, Zheng Z, Huang Y, Huang S, et al. The global, regional, and national early-onset colorectal cancer burden and trends from 1990 to 2019: results from the Global Burden of Disease Study 2019. BMC Public Health 2022 Oct 12;22(1):1896 [FREE Full text] [doi: 10.1186/s12889-022-14274-7] [Medline: 36221047]
11. Resident population by ethnic group, age group and sex dashboard. Department of Statistics Singapore. URL: https://www.singstat.gov.sg/find-data/search-by-theme/population/population-and-population-structure/visualising-data/resident-population-by-ethnic-group-age-group-and-sex-dashboard [accessed 2024-01-02]
12. Lim GH, Chow KY, Lee HP. Singapore cancer trends in the last decade. Singapore Med J 2012 Jan;53(1):3-9 [FREE Full text] [Medline: 22252175]
13. Singapore Cancer Registry Annual Report 2021. National Registry of Diseases Office. 2023 Aug. URL: https://nrdo.gov.sg/publications/cancer/ [accessed 2024-01-02]
14. Fung JWM, Lim SBL, Zheng H, Ho WYT, Lee BG, Chow KY, et al. Data quality at the Singapore Cancer Registry: an overview of comparability, completeness, validity and timeliness. Cancer Epidemiol 2016 Aug;43:76-86. [doi: 10.1016/j.canep.2016.06.006] [Medline: 27399312]
15. Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2001 Feb 15;19(3):335-351. [doi: 10.1002/(sici)1097-0258(20000215)19:3<335::aid-sim336>3.0.co;2-z] [Medline: 10649300]
16. Clegg LX, Hankey BF, Tiwari R, Feuer EJ, Edwards BK. Estimating average annual per cent change in trend analysis. Stat Med 2009 Dec 20;28(29):3670-3682 [FREE Full text] [doi: 10.1002/sim.3733] [Medline: 19856324]
17. Kim H, Luo J, Kim J, Chen H, Feuer EJ. Clustering of trend data using joinpoint regression models. Stat Med 2014 Oct 15;33(23):4087-4103 [FREE Full text] [doi: 10.1002/sim.6221] [Medline: 24895073]
18. Kim H, Chen H, Midthune D, Wheeler B, Buckman DW, Green D, et al. Data-driven choice of a model selection method in joinpoint regression. J Appl Stat 2023;50(9):1992-2013 [FREE Full text] [doi: 10.1080/02664763.2022.2063265] [Medline: 37378270]
19. Rosenberg PS, Check DP, Anderson WF. A web tool for age-period-cohort analysis of cancer incidence and mortality rates. Cancer Epidemiol Biomarkers Prev 2014 Nov;23(11):2296-2302 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-0300] [Medline: 25146089]
20. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74(1):12-49 [FREE Full text] [doi: 10.3322/caac.21820] [Medline: 38230766]
21. Siegel RL, Fedewa SA, Anderson WF, Miller KD, Ma J, Rosenberg PS, et al. Colorectal cancer incidence patterns in the United States, 1974-2013. J Natl Cancer Inst 2017 Aug 01;109(8):djw322 [FREE Full text] [doi: 10.1093/jnci/djw322] [Medline: 28376186]
22. Siegel RL, Torre LA, Soerjomataram I, Hayes RB, Bray F, Weber TK, et al. Global patterns and trends in colorectal cancer incidence in young adults. Gut 2019 Dec;68(12):2179-2185. [doi: 10.1136/gutjnl-2019-319511] [Medline: 31488504]
23. Chambers AC, Dixon SW, White P, Williams AC, Thomas MG, Messenger DE. Demographic trends in the incidence of young-onset colorectal cancer: a population-based study. Br J Surg 2020 Apr;107(5):595-605 [FREE Full text] [doi: 10.1002/bjs.11486] [Medline: 32149386]

XSL•FO
RenderX

24. Brenner DR, Ruan Y, Shaw E, De P, Heitman SJ, Hilsden RJ. Increasing colorectal cancer incidence trends among younger adults in Canada. Prev Med 2017 Dec;105:345-349 [FREE Full text] [doi: 10.1016/j.ypmed.2017.10.007] [Medline: 28987338]
25. Li Q, Yu M, Lv H, Zhang L, Deng Y, Yu H. Burden of early-onset colorectal cancer along with attributable risk factors from 1990 to 2019: a comparative study between China and other G20 countries. BMC Public Health 2023 Jul 31;23(1):1463 [FREE Full text] [doi: 10.1186/s12889-023-16407-y] [Medline: 37525147]
26. Huang J, Lucero-Prisno DE, Zhang L, Xu W, Wong SH, Ng SC, et al. Updated epidemiology of gastrointestinal cancers in East Asia. Nat Rev Gastroenterol Hepatol 2023 May;20(5):271-287. [doi: 10.1038/s41575-022-00726-3] [Medline: 36631716]
27. Sung JJY, Chiu H, Jung K, Jun JK, Sekiguchi M, Matsuda T, et al. Increasing trend in young-onset colorectal cancer in Asia: more cancers in men and more rectal cancers. Am J Gastroenterol 2019 Feb;114(2):322-329. [doi: 10.14309/ajg.0000000000000133] [Medline: 30694865]
28. AlZaabi A, AlHarrasi A, AlMusalami A, AlMahyijari N, Al Hinai K, ALAdawi H, et al. Early onset colorectal cancer: challenges across the cancer care continuum. Ann Med Surg (Lond) 2022 Oct;82:104453 [FREE Full text] [doi: 10.1016/j.amsu.2022.104453] [Medline: 36268309]
29. Lingas EC. Early-onset colon cancer: a narrative review of its pathogenesis, clinical presentation, treatment, and prognosis. Cureus 2023 Sep;15(9):e45404 [FREE Full text] [doi: 10.7759/cureus.45404] [Medline: 37854763]
30. Chen FW, Sundaram V, Chew TA, Ladabaum U. Advanced-stage colorectal cancer in persons younger than 50 years not associated with longer duration of symptoms or time to diagnosis. Clin Gastroenterol Hepatol 2017 May;15(5):728-737.e3 [FREE Full text] [doi: 10.1016/j.cgh.2016.10.038] [Medline: 27856366]
31. Eng C, Jácome AA, Agarwal R, Hayat MH, Byndloss MX, Holowatyj AN, et al. A comprehensive framework for early-onset colorectal cancer research. Lancet Oncol 2022 Mar;23(3):e116-e128. [doi: 10.1016/S1470-2045(21)00588-X] [Medline: 35090673]
32. Bailey CE, Hu C, You YN, Bednarski BK, Rodriguez-Bigas MA, Skibber JM, et al. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975-2010. JAMA Surg 2015 Jan;150(1):17-22 [FREE Full text] [doi: 10.1001/jamasurg.2014.1756] [Medline: 25372703]
33. Muller C, Ihionkhan E, Stoffel EM, Kupfer SS. Disparities in early-onset colorectal cancer. Cells 2021 Apr 26;10(5):1018 [FREE Full text] [doi: 10.3390/cells10051018] [Medline: 33925893]
34. Nguyen SP, Bent S, Chen Y, Terdiman JP. Gender as a risk factor for advanced neoplasia and colorectal cancer: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2009 Jun;7(6):676-681.e1. [doi: 10.1016/j.cgh.2009.01.008] [Medline: 19514116]
35. Newcomb PA, Zheng Y, Chia VM, Morimoto LM, Doria-Rose VP, Templeton A, et al. Estrogen plus progestin use, microsatellite instability, and the risk of colorectal cancer in women. Cancer Res 2007 Aug 01;67(15):7534-7539. [doi: 10.1158/0008-5472.CAN-06-4275] [Medline: 17671225]
36. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA 2002 Jul 17;288(3):321-333. [doi: 10.1001/jama.288.3.321] [Medline: 12117397]
37. Sung JJY, Lau JYW, Young GP, Sano Y, Chiu HM, Byeon JS, Asia Pacific Working Group on Colorectal Cancer. Asia Pacific consensus recommendations for colorectal cancer screening. Gut 2008 Aug;57(8):1166-1176 [FREE Full text] [doi: 10.1136/gut.2007.146316] [Medline: 18628378]
38. Abu Hassan MR, Ismail I, Mohd Suan MA, Ahmad F, Wan Khazim WK, Othman Z, et al. Incidence and mortality rates of colorectal cancer in Malaysia. Epidemiol Health 2016;38:e2016007 [FREE Full text] [doi: 10.4178/epih.e2016007] [Medline: 26971697]
39. Lee HP, Lee J, Shanmugaratnam K. Trends and ethnic variation in incidence and mortality from cancers of the colon and rectum in Singapore, 1968 to 1982. Ann Acad Med Singap 1987 Jul;16(3):397-401. [Medline: 3435003]
40. Wong MC, Ding H, Wang J, Chan PS, Huang J. Prevalence and risk factors of colorectal cancer in Asia. Intest Res 2019 Jul;17(3):317-329 [FREE Full text] [doi: 10.5217/ir.2019.00021] [Medline: 31085968]
41. Burkill S, Tan CS, Lau J, Koh GC, Tan KK. Colorectal cancer incidence in a multi-ethnic population: age-period-cohort analysis of trends in Singapore. Cancer Epidemiol 2021 Dec;75:102046. [doi: 10.1016/j.canep.2021.102046] [Medline: 34653790]
42. Muhammad Nawawi KN, Mokhtar NM, Wong Z, Mohd Azman ZA, Hsin Chew DC, Rehir R, et al. Incidence and clinicopathological features of colorectal cancer among multi-ethnic patients in Kuala Lumpur, Malaysia: a hospital-based retrospective analysis over two decades. PeerJ 2021;9:e12425 [FREE Full text] [doi: 10.7717/peerj.12425] [Medline: 34820182]
43. Zhang Y, Abdin E, Sambasivam R, Shafie S, Roystonn K, Vaingankar JA, et al. Changes in body mass index and its association with socio-demographic characteristics between 2010 and 2016 in Singapore. Front Public Health 2024;12:1374806 [FREE Full text] [doi: 10.3389/fpubh.2024.1374806] [Medline: 38601489]

XSL•FO
RenderX

44. Chan TK, Tan LWL, van Dam RM, Seow WJ. Cancer screening knowledge and behavior in a multi-ethnic Asian population: the Singapore Community Health Study. Front Oncol 2021;11:684917 [FREE Full text] [doi: 10.3389/fonc.2021.684917] [Medline: 34476210]
45. Key socio-economic outcomes across cohorts. Ministry of Finance Singapore. 2019 Oct. URL: https://www.mof.gov.sg/docs/default-source/default-document-library/news-and-publications/press-releases/mof-report.pdf?sfvrsn=5eadc128_2> [accessed 2024-01-02]
46. Jin Z, Dixon JG, Fiskum JM, Parekh HD, Sinicrope FA, Yothers G, et al. Clinicopathological and molecular characteristics of early-onset stage III Colon adenocarcinoma: an analysis of the ACCENT database. J Natl Cancer Inst 2021 Nov 29;113(12):1693-1704 [FREE Full text] [doi: 10.1093/jnci/djab123] [Medline: 34405233]
47. Fontana E, Meyers J, Sobrero A, Iveson T, Shields AF, Taieb J, et al. Early-onset colorectal adenocarcinoma in the IDEA Database: treatment adherence, toxicities, and outcomes with 3 and 6 months of adjuvant fluoropyrimidine and oxaliplatin. J Clin Oncol 2021 Dec 20;39(36):4009-4019 [FREE Full text] [doi: 10.1200/JCO.21.02008] [Medline: 34752136]
48. Wolf AMD, Fontham ETH, Church TR, Flowers CR, Guerra CE, LaMonte SJ, et al. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American Cancer Society. CA Cancer J Clin 2018 Jul;68(4):250-281 [FREE Full text] [doi: 10.3322/caac.21457] [Medline: 29846947]
49. Ladabaum U, Mannalithara A, Meester RGS, Gupta S, Schoen RE. Cost-effectiveness and national effects of initiating colorectal cancer screening for average-risk persons at age 45 years instead of 50 years. Gastroenterology 2019 Jul;157(1):137-148 [FREE Full text] [doi: 10.1053/j.gastro.2019.03.023] [Medline: 30930021]
50. Wong RK, Wong ML, Chan YH, Feng Z, Wai CT, Yeoh KG. Gender differences in predictors of colorectal cancer screening uptake: a national cross sectional study based on the health belief model. BMC Public Health 2013 Jul 23;13:677 [FREE Full text] [doi: 10.1186/1471-2458-13-677] [Medline: 23879593]
51. National Population Health Survey 2023 Report. Ministry of Health Singapore. 2024 Aug. URL: https://www.moh.gov.sg/others/resources-and-statistics/nphs-2023> [accessed 2024-01-02]
52. Zaborowski AM, REACCT Collaborative. Colorectal cancer in the young: Research in Early Age Colorectal Cancer Trends (REACCT) Collaborative. Cancers (Basel) 2023 May 30;15(11):2979 [FREE Full text] [doi: 10.3390/cancers15112979] [Medline: 37296939]
53. Pearlman R, Frankel WL, Swanson B, Zhao W, Yilmaz A, Miller K, Ohio Colorectal Cancer Prevention Initiative Study Group. Prevalence and spectrum of germline cancer susceptibility gene mutations among patients with early-onset colorectal cancer. JAMA Oncol 2017 Apr 01;3(4):464-471 [FREE Full text] [doi: 10.1001/jamaoncol.2016.5194] [Medline: 27978560]
54. Mork ME, You YN, Ying J, Bannon SA, Lynch PM, Rodriguez-Bigas MA, et al. High prevalence of hereditary cancer syndromes in adolescents and young adults with colorectal cancer. J Clin Oncol 2015 Nov 01;33(31):3544-3549 [FREE Full text] [doi: 10.1200/JCO.2015.61.4503] [Medline: 26195711]
55. Goel A, Nagasaka T, Spiegel J, Meyer R, Lichliter WE, Boland CR. Low frequency of Lynch syndrome among young patients with non-familial colorectal cancer. Clin Gastroenterol Hepatol 2010 Nov;8(11):966-971 [FREE Full text] [doi: 10.1016/j.cgh.2010.06.030] [Medline: 20655395]

Abbreviations

APC: annual percentage change
ASR: age-specific incidence rate
ASIR: age-standardized incidence rate
CRC: colorectal cancer
EOCRC: early-onset colorectal cancer
HR: hazard ratio
ICD: International Classification of Diseases
IRR: incidence rate ratio
SCR: Singapore Cancer Registry
SingStat: Department of Statistics Singapore
TNM: tumor, node, and metastasis

XSL•FO
RenderX

Edited by A Mavragani; submitted 09.06.24; peer-reviewed by F Seow, L Zhu, A Schwartz; comments to author 11.09.24; revised version received 02.11.24; accepted 24.12.24; published 14.02.25.

Please cite as:
Chen HLR, Chong QD, Tay B, Zhou S, Wong EYT, Seow-En I, Tan KK, Wang Y, Seow A, Tan KWE, Tan BHI, Tan SH
Trends in Early-Onset Colorectal Cancer in Singapore: Epidemiological Study of a Multiethnic Population
JMIR Public Health Surveill 2025;11:e62835
URL: https://publichealth.jmir.org/2025/1/e62835
doi:10.2196/62835
PMID:

©Hui Lionel Raphael Chen, Qingqing Dawn Chong, Brenda Tay, Siqin Zhou, Evelyn Yi Ting Wong, Isaac Seow-En, Ker Kan Tan, Yi Wang, Adeline Seow, Kwong-Wei Emile Tan, Bee Huat Iain Tan, Sze Huey Tan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Social Media Recruitment in Indigenous and Native American Populations: Challenges in the AI Age

Nadia Diamond-Smith[1,2], MSc, PhD; Alison Comfort[3], PhD; Anna Epperson[4,5], PhD; Alicia R. Riley[6], MPH, MA, PhD; Natalie Beylin[5]; Mary Garcia[4]; Sarah Francis[7]; Lucía Abascal Miguel[1,2], MD, MSc, PhD

[1]Institute for Global Health Sciences, University of California, San Francisco, 550 16th St, San Francisco, CA, United States

[2]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, United States

[3]Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco, San Francisco, United States

[4]Health Sciences Research Institute, University of California Merced, Merced, CA, United States

[5]Department of Psychological Sciences, University of California Merced, Merced, CA, United States

[6]Sociology Department, University of California, Santa Cruz, Santa Cruz, United States

[7]UpSwell, Seattle, WA, United States

Corresponding Author:
Lucía Abascal Miguel, MD, MSc, PhD
Institute for Global Health Sciences, University of California, San Francisco, 550 16th St, San Francisco, CA, United States

Abstract

Using social media recruitment for public health research presents both opportunities and challenges. Despite its increased use, few studies have detailed the practical issues, challenges encountered, and alternative strategies available for social media recruitment. This paper explores strategies for recruiting Indigenous and Native American populations in California for a study on COVID-19 vaccination and social networks. We describe different recruitment approaches, challenges faced, and pros and cons of strategies used to enhance data quality and efficiency, including survey design considerations, Facebook targeting versus use of research panels, quality assurance checks, and decisions around participant incentives. Our local setting involved recruiting Native American and Mesoamerican Indigenous individuals living in California through social media platforms. We highlight key adaptations to survey design, recruitment strategies, and data cleaning processes, noting what approaches that were effective and those that were not. Despite targeted efforts and community collaboration, recruitment was limited, and fraudulent data from bots significantly compromised data quality. Standard Facebook targeting approaches were largely unsuccessful. Our findings suggest that the increasing sophistication of artificial intelligence is becoming a substantial obstacle to authentic participant recruitment through social media. We offer recommendations to improve recruitment of hard-to-reach populations and mitigate AI-related fraud risks in future research.

(JMIR Public Health Surveill 2025;11:e76677) doi:10.2196/76677

KEYWORDS

social media recruitment; indigenous populations; Native American populations; public health research; survey design; data quality

Introduction

Native American and Indigenous populations have historically been underrepresented in health research due to longstanding structural, geographic, and cultural barriers, including limited trust in research institutions, language barriers, digital access challenges, and lack of culturally tailored outreach [1,2]. The rise of social media has sparked interest in its use for recruiting participants in both observational studies and online intervention evaluations. While most research focuses on social media to recruit for interventions, fewer have used it for observational research [3,4]. Social media has been promoted as an effective method for recruiting "hard-to-reach" populations in high-income countries, particularly sexual and gender minorities, with some studies examining effectiveness, cost-efficiency, and representativeness, and some providing guidance on specific approaches or challenges [5-9].

Despite the expansion of digital tools during the COVID-19 pandemic, few studies have described the use of online methods to engage Native American or Mesoamerican Indigenous communities. Some efforts have explored web-based health promotion or survey research with tribal communities, often emphasizing the importance of tribal sovereignty, trust-building, and culturally grounded methods [10,11]. However, few papers have systematically reported the methodological challenges, outcomes and lessons learned of social media recruitment in these population. Moreover, the rise of advanced artificial intelligence (AI) technology has exacerbated these challenges due to the proliferation of bots and fraudulent entries. Thus, there is a gap in practical guidance for researchers aiming to

XSL•FO
RenderX

use social media as a recruitment tool, particularly for hard-to-reach populations in the age of AI.

In this "Notes from the Field," we report our experience recruiting Native American and Mesoamerican Indigenous populations through social media in California for a study examining COVID-19 vaccination and social networks. The parent study aimed to recruit 800 Native American and Mesoamerican Indigenous individuals living in California, as well as 800 participants who were racially and ethnically representative of California's general population. Between January 2024 and February 2025, we recruited a total of 1173 participants, including 209 Native American and Mesoamerican Indigenous individuals. Recruitment approaches included Facebook and Instagram advertisements, outreach through community partner (CBOs), and use of a preformed research panel from Qualtrics (Table 1). Our survey included 50 questions, took approximately 30 minutes to complete, and used validated measures assessing social networks, vaccine attitudes, social media use, and sociodemographic variables. We pilot-tested the tool with students and members of our Indigenous Community Advisory Board to ensure cultural relevance and clarity, and was subsequently revised based on the feedback. In total, we received 5099 survey responses from the general population, of which 3926 were excluded due to suspected poor quality or fraudulent data. Our final sample included 137 (12%) Native American and 72 (6%) Mesoamerican Indigenous participants, far from our initial goal.

Table . Recruitment strategy and output.

Recruitment strategy	Incentive	Responses collected (N=3382)	Responses kept (n=1173)	NA/AI[a] (n=137)	Indigenous (n=72)
Meta ads - general incentive	$30	1212	334 (27%)	39	3
Meta ads - Raffle	Raffle	246	35 (14%)	2	1
Meta ads with Facebook Messenger screener	Raffle	315	95 (30%)	5	5
Qualtrics preformed panel	Set by them	1518	662 (44%)	76	49
CBO[b] distribution	$30	91	47 (52%)	15	14

[a]NA/AI: Native American/ American Indian.

[b]CBO: community-based organization.

Study results were used to examine how social networks influence vaccine decision-making and to inform the design of an intervention to address vaccine hesitancy among these subgroups. This study received institutional review board approval from participating universities and the California Rural Indian Health Board (UCSF IRB #23-38709, CRIHB IRB #2024-003) . It was also guided by a Community Advisory Board composed of Native American and Mesoamerican community members. The lead research team includes individuals from these communities.

This "Notes from the Field" highlights what was effective and what fell short in our methods and approach, and offers recommendations on recruitment strategies, survey design, and post-data collection quality assurance for researchers interested in utilizing social media for recruiting racial and ethnic minority groups in the United States.

Lessons Learned

Survey Design

We designed a survey suitable for a small screen, recognizing that some participants would complete it on mobile devices with limited time [12]. Pilot testing indicated that longer surveys increased dropout rates, prompting us to limit survey completion time to under 20 minutes. Additionally, answer options requiring extensive scrolling caused participant disengagement. To improve retention, we adapted the survey by using concise questions that fit entirely on a single screen.

Based on prior experience, we anticipated the risk of bots submitting fraudulent data to claim incentives and took proactive measures to mitigate this issue. We implemented several "bot-catching" quality checks, which also helped identify disingenuous participants rushing through the survey. Many online survey platforms offer built-in tools to verify human participants, such as CAPTCHAs, a screening mechanism where participants complete easy tasks but are challenging to bots and "invisible" reCAPTCHA's, which humans bypass unnoticed but are designed to trap bots. Recent studies have shown that they are increasingly vulnerable to sophisticated bots and human-assisted workarounds, making them a necessary but insufficient defense when used alone [13]. In our survey, many responses were dropped due to failed CAPTCHA verification; however, approximately 70% (n= 2367/3382) of the responses ultimately deemed ineligible had successfully passed the CAPTCHA screen.

We also experimented with "Honeypot" questions, which are questions that try to trick chatbots or insincere participants rushing who may not be reading the question but simply providing responses. These questions prompted respondents to select a specific answer (eg, the third option). However, we found that bots were able to bypass these as well. Ultimately, open-response fields, such as "other" with a fill-in-the-blank option or short-answer questions, proved more effective. These

XSL•FO
RenderX

fields allowed us to assess whether responses were rational and relevant, to help confirm participant authenticity.

After designing the survey, the next step was programming it using Qualtrics, an electronic survey platform, to integrate screeners, consent forms, and the survey itself. These platforms generate sharable links, making it easy to distribute surveys via social media. However, one drawback is that participants are redirected to an external webpage, often causing drop-offs. A potential workaround is embedding the survey directly within platforms like Facebook Messenger. However, this imposes limitations on the types of questions and branching logic that can be used, requiring simpler formats as these platforms are not specifically designed for complex surveys.

The screening process varied across recruitment strategies. When participants were recruited through Facebook and Instagram and directed to Qualtrics, the screener appeared as the first set of survey questions and included items on residence, age, and race and ethnicity (the race/ethnicity item was omitted when recruitment was open to the general population). Using this approach, only 33 of the 1,212 responses received were deemed ineligible based on the screener. In the Messenger-based approach, participants first completed a brief screener embedded within Facebook Messenger. Those who met eligibility criteria were then provided a link to the Qualtrics survey, which began with the consent form. Approximately 100 individuals were screened using this method, with half (n=50) determined ineligible. For the panel recruitment, we provided eligibility criteria to the panel provider (Qualtrics), who prescreened participants and invited only eligible individuals to participate. These participants were directed straight to the survey's consent form. Across all recruitment strategies, approximately 18 individuals did not consent to participate after reaching the survey.

Recruiting

Social media platforms offer advanced advertising tools that have transformed recruitment strategies. For instance, Meta's Ads Manager allows for efficient creation, monitoring, and optimization of campaigns. One major benefit of using Meta ads is the ability to implement highly targeted recruitment. Campaigns can focus on broad geographic regions or pinpoint specific areas, such as a single zip code. Additionally, targeting criteria can be refined based on factors like language, age, gender, and specific interests, allowing researchers to precisely reach their intended audience.

Using Meta's advertising tools, we targeted recruitment to zip codes with high concentrations of our target population. In collaboration with our community advisory board (CAB), we refined targeting criteria to include interests we hypothesized would resonate with the population, such as the Mexican national soccer team and Ranchera music for individuals with Mexican heritage. Images for ads were selected in partnership with the CAB and with insights from Upswell, they varied depending on the target audience. Campaigns were divided by demographics so we could turn off when we had reached the recruitment target. The objective of the campaign was survey clicks and the ads were active for two-week periods over four months. However, this approach yielded low recruitment rates. A more strategic selection of target interests, informed by deeper community insights or additional testing, could have improved this approach.

To enhance recruitment, we next employed a preformed panel—a curated group of individuals who had previously agreed to participate in research studies. We chose Qualtrics for this purpose, as our survey was already designed within their platform, and our university maintained an existing partnership. While this approach required significant upfront costs (and initial payment to Qualtrics), it proved more cost-effective compared to social media in the end in terms of cost per participant recruited. After filtering out suspected fraudulent responses, we found that recruitment through Meta ads cost approximately 10 times more than using the Qualtrics panel per high quality (real) participant recruited. In addition to cost benefits, the panel led to much faster recruitment compared to social media. However, given that our focal population included hard-to-reach groups (especially the Mesoamerican Indigenous population living in California), the pool of eligible individuals within the panel was quickly exhausted.

Incentives and Suspected Fraudulent Data

Participant compensation was a critical aspect of our recruitment strategy. Initially, based on community partner recommendations, we offered a $30 USD gift card to all participants. However, this led to a high volume of suspected fraudulent data, with 73% (n=884) of responses discarded during the initial data collection phase due to bots and fake submissions. To address this issue, we transitioned to a raffle system, offering a US $100 prize to 1 in 50 respondents. However, recruitment remained limited and fraudulent entries persisted—even after further improving the raffle odds to 1 in 10, recruitment levels did not increase and we continued to encounter suspected fraudulent data.

To balance recruitment needs with data quality, we implemented a system of individual, one-time survey links for each participant. These links could only be used once by one person, and thus, could not be shared or used multiple times over by a chatbot. Participants still received the link through a social media ad, but it could not be circulated further. This allowed us to offer a higher incentive while controlling access to prevent bots and fake entries. However, this method was labor-intensive for both our team and the participants. Interested individuals were required to email or text us to request access, after which we manually provided them with a survey link. While effective in improving data authenticity, this process limited our potential reach. We focused recruitment efforts on controlled channels, including Facebook and WhatsApp groups managed by our community partners, as well as direct contacts. While these methods helped ensure the authenticity of participants, they presented several challenges. Many individuals in our target population did not use email, and the delay between participants expressing interest and receiving their unique survey link further reduced follow-through.

During data collection, we encountered a high volume of suspected fraudulent entries, identified through re-CAPTCHA scores and other quality checks. Many fraudulent entries bypassed these measures in place by Qualtrics, requiring us to

XSL•FO
RenderX

conduct extensive manual data review and cleaning to ensure sample quality. To address these challenges, we reprogrammed the survey screener to run through Facebook Messenger. When users clicked on the ad, it opened an automated chat that included the survey screener, and eligible participants were then provided with the survey link. While this approach reduced bot activity, it came with trade-offs: overall participation decreased, the process was less cost-effective, and the platform's interface presented additional challenges.

Following data collection, we began a systematic data-cleaning process to ensure sample quality. We excluded all responses with a bot likelihood score below 0.5 provided by Qualtrics. Furthermore, we examined the geolocation of respondents' IP addresses to ensure they met our eligibility criteria, discarding those outside our target area. We also removed duplicate IP addresses; however, this process also carried the possibility of excluding legitimate responses from multiple participants within the same household.

To further improve data quality, we applied time filters to discard responses completed too quickly (ie, beyond two standard deviations from the mean). Additionally, collecting email addresses to distribute gift cards provided another opportunity for verification. We manually reviewed these emails for authenticity, flagging suspicious entries characterized by random strings of numbers or letters (eg, la675829@gmail.com).

Conclusions

Our experience with recruitment through social media revealed two key lessons. First, while social media has been effective for recruiting general populations and some specific subgroups, such as sexual and gender minorities, this approach was challenging in engaging our study's hard-to-reach populations—Mesoamerican Indigenous and Native American individuals in California. This was true despite the study team's prior experience working with these populations, expertise in social media recruitment, and input from a community advisory board on survey design and recruitment strategies. Language barriers, as our survey was only available in Spanish and English, may have contributed to these difficulties, highlighting the need for more culturally and linguistically tailored approaches to recruitment.

Second, the prevalence of suspected fraudulent data was more difficult to address than our team's prior experiences collecting data via social media in both the United States and globally. Sophisticated bots are capable of bypassing traditional quality checks, highlighting a need for strategies to differentiate authentic participants from automated responses. Identifying and filtering out fraudulent data required significant time and resources, increasing costs. Thus, it is possible that the age of easy, fast, inexpensive research through social media may be waning with increasingly sophisticated AI developments. To sustain the social media as a beneficial research tool, novel approaches to detect and prevent fraudulent entries are critically needed.

Conflicts of Interest

None declared.

References

1. Bartlett JG, Madariaga-Vignudo L, O'Neil JD, Kuhnlein HV. Identifying indigenous peoples for health research in a global context: a review of perspectives and challenges. Int J Circumpolar Health 2007 Sep;66(4):287-307. [doi: 10.3402/ijch.v66i4.18270] [Medline: 18018843]
2. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. Am J Public Health 2014 Feb;104(2):e16-e31. [doi: 10.2105/AJPH.2013.301706] [Medline: 24328648]
3. Green H, Fernandez R, MacPhail C. Social Media as a platform for recruitment to a national survey during the COVID-19 pandemic: feasibility and cost analysis. JMIR Form Res 2021 Jul 6;5(7):e28656. [doi: 10.2196/28656] [Medline: 34133315]
4. Archer-Kuhn B, Beltrano NR, Hughes J, Saini M, Tam D. Recruitment in response to a pandemic: pivoting a community-based recruitment strategy to facebook for hard-to-reach populations during COVID-19. Int J Soc Res Methodol 2022 Nov 2;25(6):797-808. [doi: 10.1080/13645579.2021.1941647]
5. Martinez O, Wu E, Shultz AZ, et al. Still a hard-to-reach population? Using social media to recruit Latino gay couples for an HIV intervention adaptation study. J Med Internet Res 2014 Apr 24;16(4):e113. [doi: 10.2196/jmir.3311] [Medline: 24763130]
6. King DB, O'Rourke N, DeLongis A. Social media recruitment and online data collection: a beginner's guide and best practices for accessing low-prevalence and hard-to-reach populations. Canadian Psychology / Psychologie canadienne 2014;55(4):240-249. [doi: 10.1037/a0038087]
7. Harfield S, Elliott S, Ramsey L, Housen T, Ward J. Using social networking sites to recruit participants: methods of an online survey of sexual health, knowledge and behaviour of young South Australians. Aust N Z J Public Health 2021 Aug;45(4):348-354. [doi: 10.1111/1753-6405.13117] [Medline: 34097339]

XSL•FO
RenderX

8. Russomanno J, Patterson JG, Jabson Tree JM. Social media recruitment of marginalized, hard-to-reach populations: development of recruitment and monitoring guidelines. JMIR Public Health Surveill 2019 Dec 2;5(4):e14886. [doi: 10.2196/14886] [Medline: 31789598]
9. Wentzell K, Walker HR, Hughes AS, Vessey JA. Engaging social media influencers to recruit hard-to-reach populations. Nurs Res 2021;70(6):455-461. [doi: 10.1097/NNR.0000000000000544] [Medline: 34334700]
10. Stephens D, Peterson R, Singer M, Johnson J, Rushing SC, Kelley A. Recruiting and engaging American Indian and Alaska Native teens and young adults in a SMS help-seeking intervention: lessons learned from the BRAVE study. Int J Environ Res Public Health 2020 Dec 16;17(24):9437. [doi: 10.3390/ijerph17249437] [Medline: 33339260]
11. Reed ND, Bull S, Shrestha U, Sarche M, Kaufman CE. Combating fraudulent participation in urban American Indian and Alaska Native virtual health research: Protocol for Increasing Data Integrity in Online Research (PRIOR). JMIR Res Protoc 2024 Jun 13;13:e52281. [doi: 10.2196/52281] [Medline: 38869930]
12. Philippe Décieux J, Sischka P. Comparing data quality and response behavior between smartphone, tablet, and computer devices in responsive design online surveys. SAGE Journals 2024;14(2). [doi: 10.1177/21582440241252116]
13. Karumathil A, Tripathi R, Indian Institute of Management Kozhikode, Northwestern University. Combating survey bots in online research: an integrative literature review of insights and strategies. THCI 2025;17(1):80-109. [doi: 10.17705/1thci.00218]

Abbreviations

AI: artificial intelligence
CAB: community advisory board
CBO: community-based organization

Edited by M Kopka; submitted 28.04.25; peer-reviewed by AW Kirkpatrick, T Coffin; revised version received 18.08.25; accepted 18.08.25; published 17.09.25.

Please cite as:
Diamond-Smith N, Comfort A, Epperson A, Riley AR, Beylin N, Garcia M, Francis S, Abascal Miguel L
Social Media Recruitment in Indigenous and Native American Populations: Challenges in the AI Age
JMIR Public Health Surveill 2025;11:e76677
URL: https://publichealth.jmir.org/2025/1/e76677
doi:10.2196/76677

© Nadia Diamond-Smith, Alison Comfort, Anna Epperson, Alicia Riley, Natalie Beylin, Mary Garcia, Sarah Francis, Lucía Abascal Miguel. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Amplifying the Voices of Youth for Equity in Wellness and Technology Research: Reflections on the Midwest Youth Wellness Initiative on Technology (MYWIT) Youth Advisory Board

Linnea Laestadius[1], MPP, PhD; Fridarose Hamad[2]; Leena Le[1]; Rosemary Buchtel[3]; Celeste Campos-Castillo[4], PhD

[1]Zilber College of Public Health, University of Wisconsin–Milwaukee, Milwaukee, WI, United States

[2]Department of Political Science, University of Wisconsin–Milwaukee, Milwaukee, WI, United States

[3]College of Natural Science, Michigan State University, East Lansing, MI, United States

[4]Department of Media & Information, Michigan State University, 404 Wilson Road, East Lansing, MI, United States

Corresponding Author:
Celeste Campos-Castillo, PhD
Department of Media & Information, Michigan State University, 404 Wilson Road, East Lansing, MI, United States

Abstract

Incorporating youth perspectives into health research can enhance quality, relevance, and ethics while also providing youth with mentorship, exposure to academic research, and professional development opportunities. This has led to a growing number of youth advisory boards (YABs). However, despite increased attention to YABs, the mentions of YABs remain low in the published research on youth and health. Furthermore, little published work has reflected on the importance of engaging with youth of color in YABs. This is critical both because of the perspectives and insight they bring to the research process and to help close the participation gap in extracurriculars among youth from racial and ethnic minoritized groups. To contribute to the literature on YABs and health equity, we offer an overview and reflection on the development and implementation of the Midwest Youth Wellness Initiative on Technology (MYWIT), a 1-week virtual, financially compensated summer YAB for youth of color aged 13 to 17 years centered on amplifying youth voices on questions related to digital technology and mental health. MYWIT youth advisors successfully codeveloped a novel research question and semistructured interview guide on the topic of navigating social media algorithms. The MYWIT process also highlighted the importance of youth compensation levels, scheduling, recruitment strategies, and overall resource constraints. We hope to encourage researchers to reflect on the value that even short duration YABs can add to the research process and how YABs can be structured to better recruit and support advisors who experience economic, institutional, and structural barriers to participation.

(JMIR Public Health Surveill 2025;11:e69013) doi:10.2196/69013

KEYWORDS

advisory boards; adolescents; social media; qualitative research; community engagement

Introduction

There has been an increasing recognition of the importance of incorporating youth perspectives into the research process to enhance quality, relevance, and ethics while also providing youth with mentorship, exposure to academic research, and professional development opportunities [1]. This has led to a growing number of youth advisory boards (YABs), also known as young people's advisory groups or youth advisory committees, where youth are recruited to help shape one or more stages in the research process. Health-focused researchers in particular have sought to create YABs [1-4], in part due to increased mandates by funders for patient and provider involvement [5].

This also aligns with the promotion of human rights, as Article 12 of the United Nations Convention on the Rights of the Child (UNRC) states that children shall hold the right to freely express their views in all matters affecting them [6]. In 2021, the UNRC clarified that the right to expression also applies to decisions relating to the digital environment [7]. Although the United States has not ratified the Convention, it remains a critical reference point for ethical research and practice. Amplifying youth voices in research on digital technologies and well-being has become particularly pressing as a growing number of policymakers and public figures have called for regulating youth access to social media and internet-enabled smartphones. More research is critically needed in this area, both to unravel the complexity of current evidence on the mental health impacts of digital technologies and to identify strategies for protecting youth without it coming "at the expense of children's other rights" such as freedom of expression, information, and assembly [7]. To ensure evidence-based government and platform policies that promote well-being and health equity, youth perspectives, particularly those of youth from vulnerable and minoritized groups, are essential to crafting research

XSL•FO
RenderX

questions and methods that reflect the lived experiences of youth.

Despite increased attention to YABs, mentions of YABs remain strikingly low in the published research on youth and health [8]. Further, little published work has reflected on the importance of engaging with youth of color in YABs. This is critical both for the perspectives and insight they bring to the research process and to help close the gap between youth from racial and ethnic minoritized and racial and ethnic majority groups in participating in nonsports-related extracurriculars [9]. Specifically, Hispanic youth have relatively low levels of participation in nonsports-related extracurricular activities, and non-Hispanic Black youth have not increased their participation rate in nonsports-related extracurriculars at the same pace as other groups [10].

These disparities in participation matter not just to college and future earnings trajectories [10-12], but more directly to well-being and health outcomes. Youth who participate in nonsports-related extracurriculars are less likely to smoke tobacco, drink alcohol, or use marijuana [10]. Access to these benefits is uneven and difficult to access for those without financial resources. Cuts to public services and a shift to a "pay-to-play" model of extracurriculars skew access toward higher-income families [9,11]. Youth with college-educated parents are also more likely to participate in extracurricular activities [10]. Further, youth participation in the paid workforce has been shown to displace their own participation in extracurricular activities, which may further disadvantage lower youth with fewer economic resources [13]. YABs seeking to incorporate perspectives from youth of color must therefore consider these financial needs.

In short, (1) YABs can offer benefits to youth of color and the advancement of equity-focused research on technology and well-being, and (2) it is critical that YABs financially compensate youth of color for their time to facilitate their participation over competing demands and formally recognize their contributions. Collaborative research with members of the public, particularly with those from racial and ethnic minoritized groups, has a responsibility to center "questions of equity and justice" [14]. To contribute to the literature on YABs and health equity, we offer a brief overview and reflection on the development and implementation of a 1-week virtual, financially compensated summer YAB for youth of color aged 13 to 17 years residing in the US Midwest. We call the YAB the Midwest Youth Wellness Initiative on Technology (MYWIT). The MYWIT is a collaborative effort between faculty at Michigan State University and the University of Wisconsin-Milwaukee. We launched the MYWIT in 2022 via seed funding from the Technology and Adolescent Mental Wellness YAB Initiative at the University of Wisconsin-Madison to help meet the goal of amplifying the voices of youth from racial and ethnic minoritized communities in research on digital technologies and youth mental health. Together, we created a research protocol and a draft semi-structured interview guide, illustrating the value of youth perspectives even in the formative stages of health equity research.

Youth Advisory Boards

YABs are closely related to youth participatory action research (YPAR), but generally stop short of allowing youth to cocreate and participate in all aspects of the research process and may lack the intention of "shifting power structures" inherent to the YPAR orientation [15]. While acknowledging this limitation, YABs offer an advantage in terms of logistical flexibility, allowing for more widespread adoption across a range of methodologies, time frames, and disciplines. They can be used to help develop a single research study or to provide insight and guidance on a broader research area or endeavor [1]. YABs thus have the potential to become a widely adopted tool in public health and medical research. The involvement of target populations in research can improve study feasibility and the usefulness of findings for practice [16,17]. Participating in the research process can also have benefits for youth above and beyond any financial compensation received, including positive outcomes related to leadership (eg, confidence and self-efficacy), academic/career (eg, time management, and oral and written communication), and social support (eg, sense of belonging and community) [2,18]. A YAB can collaborate on tasks such as generating or refining ideas, offering feedback on data collection tools, troubleshooting study challenges, and aiding in interpreting and disseminating findings [1,2,19]. Given the range of scope for YABs, guidelines have been proposed to ensure consistent reporting on how youth contribute to research [8].

However, YABs are not without challenges. Studies of YABs note several barriers to engaging with youth, including navigating ethics approvals, uncertainty about the value that youth could add, concerns about risk and the potential for harm to youth advisors, insufficient time and resources, lack of support from institutions, and recruitment challenges [20,21]. Prior work has also noted difficulties with scheduling, access to technology, attendance, and social dynamics such as a reluctance to share or the formation of cliques [2,19]. Youth may also experience anxiety about engaging with a process that is new to them [22]. Furthermore, YAB conveners must work to ensure that they provide advisors with meaningful experiences and that contributions are felt to be respected and valued rather than tokenistic, even if suggestions are not feasible within the current research environment due to discordance with institutional policies or a lack of funding [17,21,23]. These challenges also highlight the broader issue of "adultism," which privileges the perspectives and norms of adults through bias against youth [24,25]. While YABs are designed to elevate the voices of youth, norms, and expectations may still lead to an environment where youth needs are not centered or their perspectives are not fully respected [17]. Finally, the lack of alignment between the institutional expectations placed on faculty and the extensive work needed to develop and implement a YAB also mean that the work of supporting a standing YAB may need to be taken on by academic staff [23], which necessities substantive funding or other forms of institutional support. These barriers are key to overcoming for YABs to fulfill their potential of working toward addressing health equity.

XSL•FO
RenderX

The Midwest Youth Wellness Initiative on Technology (MYWIT)

Aims

The MYWIT holds the goal of working toward health equity by amplifying the voices of adolescents aged 13 to 17 years who self-identify with communities of color around the Midwest region of the United States. Specifically, the MYWIT program aims to (1) diversify the voices heard within conversations about how technologies impact teen well-being, (2) encourage advisors to share their experiences with technologies and discover new research questions that have never been asked before, (3) train advisors in the scientific method so that they can contribute to shaping research studies and consider a research-related career, and (4) prepare advisors to be leaders by providing them with a platform and skill set to communicate their ideas with others.

Format and Recruitment

In the first year of the MYWIT (ie, 2022), we operated using an academic year model, with online meetings every other month over the course of 12 months. Owing to challenges with scheduling and attendance, as also noted in a prior study [2], we opted to pilot a 1-week virtual summer camp model in 2023, meeting for 3 hours each morning on Zoom between July 17 and 21, 2023. Summer is also a time when extracurricular activities hold particular value due to summer learning loss [26]. A virtual format was selected due to the broad Midwestern catchment of the MYWIT, making a singular in-person meeting impossible. While virtual YABs can lead youth to feel less comfortable sharing their personal reflections and perspectives, prior YABs suggest that this barrier decreases over time as the advisors begin to feel more comfortable with each other [2]. Furthermore, online YABs eliminate the need to worry about how advisors will travel to meetings [1]. The objective was to offer youth an introduction to social science research methods and then support the advisors in collaboratively developing a research question and data collection tool to support future research. The Institutional Review Board at the University of Wisconsin-Milwaukee determined that the work did not qualify as human subject research. As compensation for their time, advisors received a US $100 honorarium if they attended all 5 sessions (with exceptions made for emergency situations). While contact information for a trusted adult was gathered in case of an emergency, the honorarium was distributed directly to each advisor via email in the form of two US $50 Amazon gift cards. All advisors were also offered the opportunity to coauthor this manuscript, which has been noted as a rare but important means of recognizing youth participation [17,27].

In addition to two mid-career faculty leads (one from sociology and one from public health) who served as codirectors, the MYWIT was supported by two undergraduate students (one from political science and one from public health) compensated through a university mechanism to support undergraduate research. All team members were female with a broad range of backgrounds (Latina, Asian, Middle Eastern, and European). In spring 2023, the undergraduate student mentors recruited eligible teens through the dissemination of a recruitment flyer (Figure 1) via personal contacts, outreach to local youth organizations throughout the Midwest, and through university staff that worked with high school recruitment and TRIO Upward Bound programs. Of 12 applications received, 4 youths (3 from Milwaukee, WI and 1 from Ann Arbor, MI) responded to emails and text messages to schedule informal interviews and complete emergency contact information. All advisors were between the ages of 16 and 17 years; 2 identified as male and 2 as female; 3 identified as Black or African-American and 1 as Asian and White. The number of advisors enrolled was lower than our initial target of 7-8. Additionally, no Hispanic youth enrolled as advisors. Notably, the cohort size was larger and more diverse in the first year of the program. It is not clear if the change for this cohort was due to the shift away from the academic year model as the recruitment strategy was similar for both cohorts. One participant indicated a need for technology resources and was provided with a university-owned tablet for the duration of the program. Having devices available and screening for need is critical since limited access to devices is a potential barrier to equitable participation for online YABs [2].

XSL•FO
RenderX

Figure 1. The Midwest Youth Wellness Initiative on Technology (MYWIT) recruitment flyer.

Lesson Topics and Research Output

Faculty and student mentors collaborated on organizing the program through brainstorming sessions informed by prior experiences with teaching research methods and a suggested outline for how to structure each day created by ChatGPT 3.5. The team had prior experience tailoring research design to high-school students; both faculty leads had mentored incoming first-year college students and one of the student mentors had engaged in research with the codirectors in her first summer after high school. Each morning, we started with an icebreaker or game related to technology or wellness for advisors to build rapport and prepare them for the day's activities and topics. We created Canva slides for lessons and used Google Jamboard and Padlet to engage youth advisors in brainstorming, feedback, and collaboration activities. Each day's activities primed the advisors for the next day's topic and equipped them with new

XSL•FO
RenderX

concepts, terminology, and skills related to research. Camera use was encouraged but not required.

The topic for each day can be seen in Table 1. Lessons on research were paired with research output-oriented activities each day. To begin, we asked advisors to brainstorm research topics related to technology in their lives. On day 2, they developed the research question: "How do minority youth navigate social media algorithms to meet their needs (social and consumer needs)?" Specifically, advisors felt that social media algorithms often showed them products not aligned with their needs (eg, haircare products for White hair types) or that targeted them in ways that felt manipulative (eg, ads for colleges that feature diverse students, when most of the student body is White). While it is difficult to study algorithms directly due to their proprietary nature, prior research suggests that social media algorithms may either target or exclude users from racial and ethnic minoritized groups in discriminatory ways and may also privilege social media content from White creators, denying users and creators "equal access to services and opportunities" [28,29]. Thus, the question centered on how youth from racial and ethnic minoritized communities work to navigate social media algorithms to better find the products and communities that meet their needs. Although the undergraduate student mentors and faculty codirectors reflect a broad range of racial and ethnic backgrounds, we note that we lacked concordance with the Black or African-American majority of the advisors. While advisors spoke about racialized experiences on social media, this lack of concordance may have shaped interactions in ways that either encouraged or inhibited discussion about these topics [30].

XSL•FO
RenderX

Table . The Midwest Youth Wellness Initiative on Technology (MYWIT) summer programming, 2023.

MYWIT summer program	Lessons and activities	Research output from advisors
Day 1: Introduction to Mental Health and its Relationship to Technology	1. Introductions and icebreaker 2. Interactive tutorial on current evidence and advisor perspectives on the pros and cons of social media for youth mental health 3. Jamboard brainstorming session: "Now it's your turn to be the researcher. What questions need to be asked? What needs to be explored?"	List of ideas for research studies about technology and youth well-being
Day 2: Research 101	1. Icebreaker 2. Educational session on the scientific method, social science research, and research ethics 3. Activity comparing social science research with natural science research 4. Refinement of brainstorming from Day 1 to reach one shared research question 5. Education session on common social science research and data collection methods 6. Activity for picking a data collection method aligned with the research question	A single open-ended research question to guide the rest of the week: "How do minority youth navigate social media algorithms to meet their needs (social and consumer needs)?" A data collection approach and target population: Semistructured interviews Youth age 13 to 17 years All minoritized groups Social media users (any platform)
Day 3: Developing a Research Plan and Expert Panel	1. Icebreaker 2. Education session on how to craft effective interview questions 3. Padlet brainstorming session for interview questions to respond to research question 4. Advisors led a mock interview with one of the undergraduate student mentors followed by question refinement 5. Expert panel and question and answer sessions with technology and mental health entrepreneurs	List of ideas for semistructured interview questions
Day 4: Finalizing the Research Plan and Program Reflection	1. Icebreaker 2. Final revisions to interview questions 3. Jamboard brainstorming session on what advisors hope will change based on the study and how to recruit participants 4. Padlet reflection on the MYWIT: What did you learn? What did you enjoy? What would you change?	Final list of semistructured interview questions List of study justifications and recruitment strategies Reflections and advice on the youth advisory board process
Day 5: Post Program and College Advice (short session)	1. Guidance on how to highlight the MYWIT participation on college applications and resumes 2. Advice for how to participate in research as an undergraduate student 3. Sharing of information on how to stay in touch and assessing interest in being involved in implementing the proposed study	List of advisors interested in continued work

On days 3 and 4, advisors designed a research protocol and semistructured interview guide that could be implemented in the future (Textbox 1). To provide additional professional development content, the MYWIT also featured an invited panel of entrepreneurs of color who work in the fields of technology and mental health and concluded with a session on how to translate their new research expertise into posthigh–school experiences. With the exception of this final day, which was led fully by the student mentors, each day had one of the codirectors present to support the advisors.

XSL•FO
RenderX

Textbox 1. Interview questions developed by youth advisors on the topic of how youth from racial and ethnic minoritized groups navigate social media algorithms.

1. What is your favorite app?
2. How is your social media feed different now than when you first joined?
3. What kinds of things have you done to personalize your social media feed? How has that worked for you?
4. What kinds of obstacles have you encountered when navigating social media?
5. How does social media affect your mental health?
6. What social media app do you use to get recommendations or inspiration for things you want to buy?
7. What social media app do you use to search for things you want to buy?
8. What kinds of products do you look for on social media?
9. What word or phrase would you use to search for these products?
10. How does it make you feel looking for products on social media that fit your needs?
11. Do you feel like it is hard to find products for you on social media? Why?
12. Have you ever felt excluded by an ad on social media?
13. Can you give an example of a time when you felt excluded by an ad on social media?
14. How do you feel about social media advertisements targeting minorities?
15. What app do you use to search for communities online?
16. How do you use that app to interact with people?
17. How did you find those communities on that app?
18. Do you feel like it is hard to find communities for you online?
19. Have you ever felt left out from communities on social media because of your culture? Can you give an example?
20. How does it make you feel looking for communities on social media to fit your needs?
21. Do you post on social media?
22. What kinds of hashtags or buzzwords do you use to get your posts seen by the people they are meant for?
23. What would you change about the app to make it more inclusive for your community?
24. What would you change about the app to make navigating the app easier?

Experiences and Recommendations

Lessons Learned

At the end of the week, everyone reflected on the process and their recommendations via three questions on Padlet: (1) What did you learn? (2) What did you enjoy and what worked well? (3) What would you change? Additionally, advisors were asked to share further brief reflections via email following the final day. Below, we present the synthesized reflections across the groups (codirectors, student mentors, and advisors). Table 2 provides additional detail on key lessons learned.

XSL•FO
RenderX

Table . Key lessons from the Midwest Youth Wellness Initiative on Technology (MYWIT).

Topic	Key lessons
Recruitment	1. Recognize that recruitment is the most difficult task. 2. Budget extra time for recruitment. 3. Significantly overenroll in anticipation of attrition before the program begins so that the final cohort is larger.[b] 4. Cold calling youth-serving organizations was not effective. 5. Adopt rolling application review until the YAB[a] begins. 6. Reduce the amount of time between the application going live and the start of the YAB[b] 7. Holding advisor interviews can help assess the level of interest and commitment to attend. 8. Inviting advisors' friends to participate did not prove to be a very effective recruitment strategy. 9. Most advisors came from organizations that researchers and youth mentors had direct connections with or from a posting on the Michigan State University's website. Outreach within team member networks is a promising means of recruitment.
Programming and Engagement	1. Interactive icebreakers were important to encourage participation from advisors. 2. From the beginning, have a clear attendance policy (cameras on, how many days they can miss, etc). 3. Change the scope and duration of the program to allow for more in-depth work. Mentors and Codirectors suggested shorter program length with longer days, while Advisors suggested a 2-week program with shorter days.[b] 4. The most reliable communication to ensure daily attendance was texting. Emails seemed to work less frequently. 5. Family or financial commitments may prevent advisors from giving their full attention during sessions. 6. Incentives should be informed by what advisors might make in hourly employment during the duration of the program. 7. Additional resources are needed to support study development and implementation with the same set of advisors. 8. Consider in-person meetings or ways to make online meetings more engaging.[b] 9. Despite obstacles, youth advisors have novel and valuable insights into research goals and processes.

[a]YAB: youth advisory board.

[b]Indicates items based in part on advisor feedback

Advisors

Each of the advisors shared what they learned, what they enjoyed, and what they would change about the experience. Responses highlighted learning how to craft research questions and collect relevant data. Advisors also noted that they valued learning that other youth shared their concerns about digital technologies and enjoyed learning about social media and health. One advisor noted that learning how to develop interview questions was interesting because it taught them how to gather information without making assumptions about research participants. Another stated that they found it a fun, low-stakes entry into scientific research, allowing them to build up confidence for future opportunities. Advisors reported enjoying getting to craft a question that they care about and the panel discussion with Milwaukee, WI technology experts on relevant careers. Some feedback they provided for improvement included having a shorter application time span so that more youth would join and not forget about the program in the middle of the summer, enrolling more youth, and extending the program over 2 weeks and having shorter days instead. Additionally, two advisors indicated that in-person meetings would have been helpful. One explained that they felt the online format made sense for the time, but that having future programs in-person would facilitate conversations, as there were frequent silences during meetings that could be attributed to many of the advisors not turning on their cameras, leading to black screens with no face showing.

Undergraduate Student Mentors

Student mentors reported learning as peer mentors to the youth advisors, specifically enhancing their community engagement and mentorship experience for future careers. Both were nervous about explaining the program in a way that would be interesting and exciting to the advisors but enjoyed the challenge of practicing how to communicate clearly. Seeing how well the youth advisors engaged made the student mentors feel confident in their own communication growth. The students learned that if there are clear explanations and expectations of the youth advisors, they will feel empowered and motivated to participate,

XSL•FO
RenderX

which will enhance the likelihood that the products contribute toward health equity. Student mentors also learned the importance of adjusting communication from emails to phone calls or texts to send reminders to advisors and keep them engaged. Recommendations for change include a later start time for youth to be more engaged and recruiting more advisors to make it a more dynamic experience.

Faculty Codirectors

Codirectors had initially been worried about the low enrollment numbers, which persisted despite proactive efforts to recruit advisors, and the financial incentive offered. Ultimately, the 4 advisors generated valuable insights into research goals and processes, with the momentum of the 1-week schedule potentially offsetting the small group size. It is possible that the smaller-than-anticipated cohort size helped to achieve this positive outcome, as the literature on virtual focus groups indicates that smaller sizes (4 to 6) are preferred to achieve engagement and foster trust in the absence of traditional face-to-face communication [31].

However, not all advisors were able to participate equally during the week. One advisor joined from their place of summer employment and was thus not able to consistently make use of video and audio functions on Zoom. Developing clear guidelines in advance for what counts as full participation is essential. Furthermore, incentives may need to be increased to better reflect what youth could earn in hourly employment during the duration of the program. The loan of a tablet computer to an advisor also posed complexities in terms of institutional policies, with the university requiring that the student sign a document that they would cover the cost of the tablet if it was lost. Accordingly, a tablet that was valued less than the incentive for participation was chosen to avoid creating a potential hardship in the event it was lost or damaged. Considering ways to overcome economic, institutional, and structural barriers to participation is critical for any program seeking to prioritize the voices of youth from minoritized and marginalized communities as a step toward centering equity in technology and wellness research.

Furthermore, while the MYWIT is out of necessity virtual due its broad catchment area, others may wish to explore in-person YABs. Although past works suggest that in-person YABs also struggle with attrition [3,32], initial recruitment may be easier if youth are seeking a more traditional and immersive summer experience. We may also reconsider our policy on camera use being optional during online meetings to help foster a greater sense of community among the advisors. Camera use can be an equity concern, due to camera access, internet speed, and a desire to obscure the residential environment. In addition to the provision of devices when needed, one approach to consider might be to encourage the use of virtual backgrounds to allow cameras to be on, while still maintaining greater privacy [31].

Despite these challenges, the codirectors were extremely pleased with the rapid pace at which advisors picked up key research concepts, the insights advisors offered, and their passion for technology and youth wellness. Undergraduate mentor support was also pivotal for establishing rapport with the advisors. For example, undergraduate mentors helped to troubleshoot communication by suggesting the use of text messages rather than email. While this may appear to be an obvious choice, institutional norms for communication with college students via email (particularly at public institutions, where email is considered a public record) can obscure the actual communication preferences of youth and young adults. This serves as an indication of the critical importance of reflecting on and working to shift “adultism” when engaging with YAB work [24].

The research question that the advisors developed was one that the team had not previously considered, and the proposed interview questions were critical to fully understanding how the advisors operationalized the concept of navigating social media algorithms. Accordingly, the YAB provided significant value to the research process through the development of unique research questions and methods informed by the lived experiences of the advisors and their perceptions of social media, which is key for working toward health equity. Overall, the objective of training advisors in the scientific method so that they can contribute to shaping research studies was achieved and the building blocks for a viable future study were created. Going forward, the MYWIT would benefit from additional financial and time resources to support (1) rapidly developing and launching the resulting study idea, and (2) allowing advisors to participate in actual study implementation.

Next Steps

Sufficient resources are essential to support YAB work. For the MYWIT, we hope to obtain grant funding to support translating the advisor-developed project into a full-scale study, including resources to support youth, either through a YAB to offer guidance at key stages or a YPAR approach to more formally engage youth as partners throughout. A previous study on YABs indicates that advisors, with support and training, can excel at qualitative data collection and analysis [32]. The inability to move directly into the implementation of the ideas should not be seen as an indication that the YAB failed to generate valuable ideas and guidance, but rather an artifact of a small initial budget and the institutional prioritization of grant-funded work. Insufficient financial resources have been noted as one of the most common barriers to engaging youth in the research process [21]. Further reflective of the ways in which the organizational work needed to support a YAB is often misaligned with university expectations for faculty [23], we also hope to obtain the funding needed to hire not just student mentors but also academic staff to support future YAB work. This would also help create stability and continuity in the YAB over time [23]. Finally, we hope to draw on the recruitment lessons outlined in Table 2 to increase youth compensation and adopt new recruitment strategies to create a larger pool of advisors. This will also help to ensure engagement remains robust even if an advisor faces unexpected constraints that prevent their full participation.

Conclusions

Through sharing our experiences with a compensated YAB for youth of color in the US Midwest, our hope is to not just

XSL•FO
RenderX

encourage researchers to reflect on the value that even short-duration YABs can add to the research process, but also how YABs can be structured to better support advisors from racial and ethnic minoritized groups. While prior work has stressed the potential value of YABs and how they might be organized [2,3,8], this second question has received less attention. Sharing lessons, both in terms of what worked and what did not, is critical for identifying how best to leverage the value of YABs in shaping research on technology and health equity.

Acknowledgments

We thank the Midwest Youth Wellness Initiative on Technology (MYWIT) advisors (Nyahsha Stallings, Rosemary Buchtel, and others) for their time and commitment to research. We also want to acknowledge financial support from the Technology and Adolescent Mental Wellness (TAM) grant program at the University of Wisconsin–Madison and the Support for Undergraduate Research Fellows (SURF) program at the University of Wisconsin-Milwaukee.

Conflicts of Interest

None declared.

References

1. Moreno MA, Jolliff A, Kerr B. Youth advisory boards: perspectives and processes. J Adolesc Health 2021 Aug;69(2):192-194. [doi: 10.1016/j.jadohealth.2021.05.001] [Medline: 34078575]
2. Abraham O, Rosenberger CA, Poku VO. Implementing a youth advisory board to inform adolescent health and medication safety research. Res Social Adm Pharm 2023 Apr;19(4):681-685. [doi: 10.1016/j.sapharm.2022.12.003] [Medline: 36528539]
3. Orellana M, Valdez-Soto M, Brockman TA, et al. Creating a pediatric advisory board for engaging youth in pediatric health research: a case study. J Clin Transl Sci 2021 Mar 30;5(1):e113. [doi: 10.1017/cts.2021.399] [Medline: 34221455]
4. Tsang VWL, Chew SY, Junker AK. Facilitators and barriers to the training and maintenance of young persons' advisory groups (YPAGs). Int J Pediatr Adolesc Med 2020 Dec;7(4):166-173. [doi: 10.1016/j.ijpam.2019.10.002] [Medline: 33319014]
5. McCoy MS, Warsh J, Rand L, Parker M, Sheehan M. Patient and public involvement: Two sides of the same coin or different coins altogether? Bioethics 2019 Jul;33(6):708-715. [doi: 10.1111/bioe.12584] [Medline: 30957902]
6. United Nations. Convention on the Rights of the Child. 1989 Nov 20. URL: https://www.ohchr.org/en/instruments-mechanisms/instruments/convention-rights-child [accessed 2025-04-02]
7. Livingstone S, Sylwander KR. There is no right age! The search for age-appropriate ways to support children's digital lives and rights. J Child Media 2025 Jan 2;19(1):6-12. [doi: 10.1080/17482798.2024.2435015]
8. Sellars E, Pavarini G, Michelson D, Creswell C, Fazel M. Young people's advisory groups in health research: scoping review and mapping of practices. Arch Dis Child 2021 Jul;106(7):698-704. [doi: 10.1136/archdischild-2020-320452] [Medline: 33208398]
9. Heath RD, Anderson C, Turner AC, Payne CM. Extracurricular activities and disadvantaged youth: a complicated—but promising—story. Urban Educ (Beverly Hills Calif) 2022 Oct;57(8):1415-1449. [doi: 10.1177/0042085918805797]
10. Meier A, Hartmann BS, Larson R. A quarter century of participation in school-based extracurricular activities: inequalities by race, class, gender and age? J Youth Adolesc 2018 Jun;47(6):1299-1316. [doi: 10.1007/s10964-018-0838-1] [Medline: 29536328]
11. Snellman K, Silva JM, Frederick CB, Putnam RD. The engagement gap. Ann Am Acad Pol Soc Sci 2015 Jan;657(1):194-207. [doi: 10.1177/0002716214548398]
12. Lipscomb S. Secondary school extracurricular involvement and academic achievement: a fixed effects approach. Econ Educ Rev 2007 Aug;26(4):463-472. [doi: 10.1016/j.econedurev.2006.02.006]
13. Crispin LM, Kofoed M. Does time to work limit time to play?: Estimating a time allocation model for high school students by household socioeconomic status. Contemp Econ Policy 2019 Jul;37(3):524-544 [FREE Full text] [doi: 10.1111/coep.12411]
14. Fiske A, Prainsack B, Buyx A. Meeting the needs of underserved populations: setting the agenda for more inclusive citizen science of medicine. J Med Ethics 2019 Sep;45(9):617-622. [doi: 10.1136/medethics-2018-105253] [Medline: 31300487]
15. Ozer EJ, Abraczinskas M, Duarte C, et al. Youth participatory approaches and health equity: conceptualization and integrative review. Am J Community Psychol 2020 Dec;66(3-4):267-278. [doi: 10.1002/ajcp.12451] [Medline: 32969506]
16. Arumugam A, Phillips LR, Moore A, et al. Patient and public involvement in research: a review of practical resources for young investigators. BMC Rheumatol 2023 Mar 9;7(1):2. [doi: 10.1186/s41927-023-00327-w] [Medline: 36895053]
17. Hawke LD, Relihan J, Miller J, et al. Engaging youth in research planning, design and execution: practical recommendations for researchers. Health Expect 2018 Dec;21(6):944-949. [doi: 10.1111/hex.12795] [Medline: 29858526]
18. Anyon Y, Bender K, Kennedy H, Dechants J. A systematic review of youth participatory action research (YPAR) in the United States: methodologies, youth outcomes, and future directions. Health Educ Behav 2018 Dec;45(6):865-878. [doi: 10.1177/1090198118769357] [Medline: 29749267]

XSL•FO
RenderX

19. Flicker S, Guta A, Larkin J, et al. Survey design from the ground up: collaboratively creating the Toronto teen survey. Health Promot Pract 2010 Jan;11(1):112-122. [doi: 10.1177/1524839907309868] [Medline: 18367639]
20. Wadman R, Williams AJ, Brown K, Nielsen E. Supported and valued? A survey of early career researchers' experiences and perceptions of youth and adult involvement in mental health, self-harm and suicide research. Res Involv Engagem 2019;5(1):16. [doi: 10.1186/s40900-019-0149-z] [Medline: 31164992]
21. Faithfull S, Brophy L, Pennell K, Simmons MB. Barriers and enablers to meaningful youth participation in mental health research: qualitative interviews with youth mental health researchers. J Ment Health 2019 Feb;28(1):56-63. [doi: 10.1080/09638237.2018.1521926] [Medline: 30353772]
22. Mawn L, Welsh P, Kirkpatrick L, Webster LAD, Stain HJ. Getting it right! Enhancing youth involvement in mental health research. Health Expect 2016 Aug;19(4):908-919. [doi: 10.1111/hex.12386] [Medline: 26202658]
23. Hohenemser LK, Marshall BD. Utilizing a youth development framework to establish and maintain a youth advisory committee. Health Promot Pract 2002 Apr 1;3(2):155-165. [doi: 10.1177/1524839902003002014]
24. Bell J. Understanding Adultism - A Major Obstacle to Developing Positive Youth-Adult Relationships. 1995. URL: https://actioncivics.scoe.net/pdf/understanding_adultism.pdf [accessed 2025-04-02]
25. Florio E, Caso L, Castelli I. The Adultcentrism Scale in the educational relationship: Instrument development and preliminary validation. New Ideas Psychol 2020 Apr;57:100762. [doi: 10.1016/j.newideapsych.2019.100762]
26. Atteberry A, McEachin A. School's out: the role of summers in understanding achievement disparities. Am Educ Res J 2021 Apr;58(2):239-282. [doi: 10.3102/0002831220937285]
27. Fadiran B, Lee J, Lemminger J, Jolliff A. How our technology use changed in 2020: perspectives from three youths. JMIR Ment Health 2021 Sep 15;8(9):e26154. [doi: 10.2196/26154] [Medline: 34524108]
28. Ali M, Sapiezynski P, Bogen M, Korolova A, Mislove A, Rieke A. Discrimination through optimization. Proc ACM Hum-Comput Interact 2019 Nov 7;3(CSCW):1-30. [doi: 10.1145/3359301]
29. Amarikwa M. Social media platforms' reckoning: the harmful impact of TikTok's algorithm on people of color. SSRN Journal 2023. [doi: 10.2139/ssrn.4349202]
30. Kerstetter K. Insider, outsider, or somewhere between: the impact of insider, outsider, or somewhere between: the impact of researchers' identities on the community-based research process. J Rural Soc Sci 2012;27(2).
31. Lathen L, Laestadius L. Reflections on online focus group research with low socio-economic status African American adults during COVID-19. Int J Qual Methods 2021 Jan;20:16094069211021712. [doi: 10.1177/16094069211021713]
32. Livingood WC, Monticalvo D, Bernhardt JM, et al. Engaging adolescents through participatory and qualitative research methods to develop a digital communication intervention to reduce adolescent obesity. Health Educ Behav 2017 Aug;44(4):570-580. [doi: 10.1177/1090198116677216] [Medline: 27811164]

Abbreviations

MYWIT: Midwest Youth Wellness Initiative on Technology
UNRC: United Nations Convention on the Rights of the Child
YAB: youth advisory board
YPAR: youth participatory action research

Edited by A Mavragani; submitted 19.11.24; peer-reviewed by CJ Sieck, E Gillette, OJ Flores; revised version received 29.01.25; accepted 15.02.25; published 10.04.25.

Please cite as:
Laestadius L, Hamad F, Le L, Buchtel R, Campos-Castillo C
Amplifying the Voices of Youth for Equity in Wellness and Technology Research: Reflections on the Midwest Youth Wellness Initiative on Technology (MYWIT) Youth Advisory Board
JMIR Public Health Surveill 2025;11:e69013
URL: https://publichealth.jmir.org/2025/1/e69013
doi:10.2196/69013

©Linnea Laestadius, Fridarose Hamad, Leena Le, Rosemary Buchtel, Celeste Campos-Castillo. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Notes from the Field

Implementing Social Media Strategies in Community-Partnered HIV Research: Practical Considerations From 3 Ongoing Studies

Elena P Rosenberg-Carlson[1,2], MPH; Tamra B Loeb[3], PhD; Raquel G Hernandez[4,5], MPH, MD; Jane J Lee[6,7], MSW, PhD; Aima A Ahonkhai[8,9,10], MPH, MD; Jessica M Perkins[10,11], PhD; Chunqing Lin[3], PhD; Sung-Jae Lee[3], MPH, PhD; Sharon Hurt[12], MEd; Luke Johnsen[13], DO; Alison B Hamilton[3,14], MPH, PhD

[1]Department of Family Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, United States

[2]Department of Psychiatry and Behavioral Sciences, Stanford Medicine, Stanford, United States

[3]Department of Psychiatry and Biobehavioral Sciences, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, United States

[4]Center for Pediatric Health Equity Research, Johns Hopkins All Children's Hospital, St Petersburg, United States

[5]Department of Pediatrics, Johns Hopkins Medicine, Baltimore, United States

[6]School of Social Work, University of Washington, Seattle, United States

[7]Suzanne Dworak-Peck School of Social Work, University of Southern California, Los Angeles, United States

[8]Division of Infectious Diseases, Massachusetts General Hospital, Boston, United States

[9]Center for AIDS Research, Harvard University, Cambridge, United States

[10]TN Center for AIDS Research, Vanderbilt University Medical Center, Nashville, TN, United States

[11]Department of Human and Organizational Development, Vanderbilt University, Nashville, United States

[12]Street Works, Nashville, TN, United States

[13]Metro Inclusive Health, St Petersburg, FL, United States

[14]VA Center for the Study of Innovation, Implementation, and Policy, VA Greater Los Angeles Healthcare System, Los Angeles, United States

Corresponding Author:
Elena P Rosenberg-Carlson, MPH
Department of Psychiatry and Behavioral Sciences
Stanford Medicine
1070 Arastradero Rd
Stanford, 94304
United States
Phone: 1 415 407 4262
Email: rosenbee@stanford.edu

Abstract

Background: In recent years, social media has emerged as a pivotal tool in implementation science efforts to address the HIV epidemic. Engaging community partners is essential to ensure the successful and equitable implementation of social media strategies. There is a notable lack of scholarship addressing the operational considerations for studies using social media strategies in community-partnered HIV research. This article seeks to bridge this gap by consolidating field notes and practical considerations derived from 3 ongoing NIH-supported studies focused on Ending the HIV Epidemic in the United States.

Objective: This article aims to inform the design, planning, and implementation of operationally effective community-partnered social media strategies in HIV research, ultimately contributing to enhancements in HIV practice and improved outcomes across the HIV prevention and care continua.

Methods: Supported by the University of California, Los Angeles Rapid, Rigorous, Relevant (3R) Implementation Science Hub, the 3 Ending the HIV Epidemic projects convened to form the community-partnered social media campaigns working group. The working group used the Consolidated Framework for Implementation Research to help identify and organize key barriers and facilitators of relevance to implementation of the projects' social media strategies. Given the high degree of interrelatedness across reported factors, the working group thematically synthesized the content into 5 practical considerations to inform use of community-partnered social media strategies in HIV research.

Results: The practical considerations identified by the community-partnered social media campaigns working group include the following: (1) the power and pitfalls of social media platforms (ie, opportunities and challenges inherent to social media

XSL•FO
RenderX

platforms that may affect use of social media strategies in HIV research), (2) messengers and messages matter (ie, ensuring the appropriateness, acceptability, and quality of social media messengers and content), (3) the significance of the sociopolitical environment (ie, characterizing the sociopolitical environment surrounding HIV and its potential impact on implementing social media strategies to reach priority populations), (4) investing in academic-community partnerships (ie, cultivating positive and productive academic-community partnerships to support implementation of social media strategies in HIV research), and (5) the alignment of the institutional environment and research approach (ie, assessing and working to address features of institutional environments that may impact implementation of social media strategies in community-partnered HIV research).

Conclusions: As use of social media in HIV research and practice continues to grow, the practical considerations presented in this paper can help research teams anticipate factors that may impact implementation of community-partnered social media strategies and take early action to mitigate potential challenges. By understanding and addressing the unique challenges and opportunities of social media in community-partnered HIV research, we can leverage these platforms to accelerate progress toward ending the HIV epidemic.

(JMIR Public Health Surveill 2025;11:e73318) doi:10.2196/73318

KEYWORDS

HIV; social media; community partner; academic-community partnerships; Ending the HIV Epidemic

Introduction

Technological advances have transformed the landscape of health interventions, with the use of social media emerging as a pivotal implementation strategy [1]. Social media platforms play an integral role in disseminating information, fostering community engagement and communication, and promoting healthy behaviors [2-4]. In the context of HIV prevention and care, social media is increasingly used to promote awareness and uptake of evidence-based interventions, such as pre-exposure prophylaxis (PrEP) [5,6]. In addition, HIV researchers have leveraged social media in studies aiming to enhance program reach, strengthen coordination across health care sectors, guide clinical decision-making, and improve treatment retention and adherence [7]. Social media is also routinely used as a recruitment tool for HIV-related studies [8,9]. Its capacity to engage diverse populations, particularly those hard to reach due to geographic distance and stigma, underscores the growing relevance of social media in implementation science efforts to address the HIV epidemic [10].

The rapid growth of social media also presents challenges. Platforms such as Facebook (Meta Platforms), Instagram (Meta Platforms), TikTok (ByteDance Platforms), and YouTube (Alphabet Platforms) each have unique characteristics and user demographics, necessitating tailored strategies to connect with target audiences [11,12]. Not all target audiences have equitable access to the internet or familiarity with social media platforms, reflecting a digital divide that can further exacerbate health disparities [13]. Furthermore, protecting the privacy and confidentiality of individuals on social media, particularly in the context of HIV, is paramount. Doing so requires special considerations different from and beyond those typically required in in-person settings [14,15]. The constant evolution of social media technology, algorithms, and policies can also affect the visibility and reach of health intervention content [16]. Involving community partners is essential to effectively mitigate these challenges and ensure successful and equitable implementation of social media strategies in HIV implementation science efforts.

Despite the widespread adoption and implementation of social media, there is a notable lack of scholarship addressing the operational considerations for studies using social media strategies. This article aims to bridge this gap by consolidating field notes from 3 ongoing Ending the HIV Epidemic (EHE) Center for AIDS Research and AIDS Research Center supplement projects. Funded by the National Institutes of Health, these grants focus on enhancing the implementation science knowledge base needed to support the federal EHE initiative using partnered approaches among academic researchers, local health officials, and community groups to address HIV prevention and treatment needs in areas with high HIV transmission [17]. The 3 projects highlighted here each incorporate community-partnered social media strategies in their research to address this public health challenge.

Using the Consolidated Framework for Implementation Research (CFIR) [18] as a conceptual guide, we systematically identified the social media strategy characteristics, outer setting factors, inner setting factors, and processes involved in leveraging social media in each of these ongoing HIV implementation science projects. We aim to assist others engaged in HIV research by identifying common implementation determinants relevant to the use of community-partnered social media strategies across the 3 studies and synthesizing the content into practical considerations that can help researchers anticipate and mitigate potential challenges. Ultimately, these insights are intended to inform the design, planning, and implementation of operationally effective community-partnered social media strategies in HIV research, contributing to enhancements in HIV practice and improved outcomes across the HIV prevention and care continua.

Methods

Study Setting

Overview

The University of California, Los Angeles Rapid, Rigorous, Relevant (3R) Implementation Science Hub (hereafter referred to as the Hub) is funded by a National Institute of Mental Health EHE supplement award to the University of California, Los

XSL•FO
RenderX

Angeles Center for HIV Identification, Prevention, and Treatment Services (grant P30MH058107). The Hub aims to deliver expert technical assistance, coaching, and consultation to funded EHE implementation science projects and develop HIV-related implementation science capacity and collaborations locally, regionally, and nationally. The Hub's unique goal is to provide leadership and support for rapid, rigorous, and relevant (3R) HIV-related implementation research, emphasizing pragmatic study designs and methods that address health equity and produce sustainable solutions. As part of its routine services, the Hub team facilitates all-awardee meetings to provide opportunities for networking, collaboration, and collective problem-solving among its assigned EHE projects. Through these convenings and with the Hub's support, 3 projects that were each using community-partnered social media strategies expressed interest in collaborating on a manuscript to share practical considerations and lessons learned. The leads of these 3 projects (RGH and LJ; JJL; and AAA, JMP, and SH) and members of the Hub team (ABH, CL, EPR-C, SJL, and TBL) formed the community-partnered social media campaigns (CPSMC) working group to pursue this collaborative manuscript. Descriptions of each project and their use of social media are provided in the following subsections (summarized in Table 1).

Table 1. Summary of community-partnered social media campaign working group projects and their use of social media.

Project	Target population	Project goal	Social media strategy goal	Social media strategy description
A	SGDY[a] and AYA[b] in Tampa Bay, Florida	Identify HIV testing and PrEP[c] deserts and inform the development of equity-focused HIV interventions	Recruit participants for data collection activities	Geotargeted advertisements on Instagram as well as on social dating apps (ie, Grindr and Adam4Adam) and Google Ads
B	Latino and Latinx MSM[d] in King County, Washington	Implement and evaluate a culturally relevant social media campaign to increase uptake and use of HIV testing and PrEP	Increase uptake of and engagement in HIV services	Instagram and Facebook campaign in partnership with a local Latinx LGBTQ[e] organization
C	Black men in Shelby County and Davidson County, Tennessee	Promote men's sexual health and wellness and combat HIV stigma in the Black community	Deliver an intervention to address HIV stigma	Black barbers as men's health ambassadors sharing curated posts and videos on their social media platforms (X, Facebook, Instagram, and LinkedIn)

[a]SGDY: sexual and gender diverse youth.

[b]AYA: adolescent and young adult.

[c]PrEP: pre-exposure prophylaxis.

[d]MSM: men who have sex with men.

[e]LGBTQ: lesbian, gay, bisexual, transgender, queer.

Project A: Using Social Media to Recruit Participants for Data Collection Activities

The Mapping Adolescent and Young Adult HIV Testing and PrEP Deserts within the Tampa Bay Region: Promoting Equity Among Sexual and Gender Minority Youth project (principal investigators: RGH and LJ) assesses existing HIV surveillance data, characterizes adolescent and young adult (AYA)–centered HIV care and prevention services (with a focus on PrEP use), and collects and analyzes perspectives from AYA and sexual and gender diverse youth and other key community stakeholders toward promoting equity in HIV treatment and prevention services. Identification of HIV testing and PrEP deserts (broadly defined as areas where the demand or need for HIV testing among AYAs is greater than the supply of AYA-focused HIV and PrEP sites) [19] is a central study outcome that will inform the development of community and clinical interventions. The project uses geotargeted web-based advertisements to recruit AYA participants for ideation workshops focused on understanding the problem of HIV and PrEP deserts in the Tampa Bay, Florida, area, using three types of media channels:

(1) social media applications (ie, Instagram), (2) social dating applications (ie, Grindr and Adam4Adam), and (3) paid search campaigns (ie, Google Ads).

Specific media channels, including the existing social media channels of the project's community partner organization, were selected and advertisement content was developed using feedback from community partners to effectively reach the target population. Content included images of individual AYAs, couples, and relevant medical objects (eg, PrEP and HIV tests) as well as messages focused on empowerment and advocacy for sexual health. Social media content required a distinction between static (preapproved, unchanging materials, eg, stock images) and dynamic (interactive or adaptive materials, eg, trend-driven videos) content for institutional review board (IRB) approval. Dynamic content also included the use of contemporary audio or music and pop-culture trends, as well as strategic hashtags to engage local populations.

Project B: Using Social Media to Increase Uptake of and Engagement in HIV Services

The Implementation of a Culturally Relevant Social Media Campaign to Increase HIV Testing and PrEP uptake in Latinx MSM project (principal investigator: JJL) harnesses the innovative capacity of social media applications in engaging community members for HIV prevention. The study team is implementing and evaluating a culturally relevant and community-informed social media campaign to enhance uptake of HIV testing and PrEP among Latino and Latinx gay, bisexual,

XSL•FO
RenderX

and other men who have sex with men (MSM) in King County, Washington. Project aims include the following: (1) implement the culturally tailored social media (ie, Instagram and Facebook) campaign to facilitate HIV testing and PrEP uptake among Latino and Latinx MSM in partnership with a community organization dedicated to providing HIV and other services to the local Latino and Latinx lesbian, gay, bisexual, transgender, and queer community; (2) evaluate the effectiveness of the campaign in shaping behavioral determinants and HIV testing and PrEP uptake among clients at the organization; and (3) identify the costs and conditions that influence reach, maintenance, and scalability of the campaign to increase HIV testing and PrEP uptake in the Latino and Latinx MSM community.

A community-informed social media campaign was developed specifically for Latino and Latinx MSM and piloted in collaboration with the study's community partner organization, Entre Hermanos, in a previous study [20]. Guided by the unified theory of behavior [21,22], which identifies determinants of health behaviors (ie, normative pressures, beliefs and expectancies, self-concept and image, affect and emotions, and self-efficacy) [23], the study team developed culturally tailored social media content that addresses salient unified theory of behavior constructs and barriers to HIV prevention identified among Latino and Latinx MSM through qualitative interviews and focus groups [24]. The content reflected their specific needs and preferences, and a pilot study demonstrated the feasibility of the social media content in reaching high-risk Latino and Latinx MSM in King County [20]. For this study, the project team implemented the social media campaign on Facebook and Instagram over a 2-month period in Washington State to evaluate its impact on behavioral predictors of HIV testing and PrEP and HIV testing and PrEP use outcomes among Latino and Latinx MSM who seek services at Entre Hermanos.

Project C: Using Social Media to Deliver an Intervention to Address HIV Stigma

The Cutting Out Stigma: A Barbershop-Based Strategy to Reduce Stigma towards Black Men Living with HIV in TN project (principal investigators: AAA, JMP, and SH) involves implementing a barbershop-based health education and multimedia intervention focused on promoting men's sexual health and wellness and combating HIV stigma in the Black community in the 2 counties in Tennessee with the highest HIV prevalence and incidence—Shelby County (home to Memphis) and Davidson County (home to Nashville). This multipart intervention includes the following: (1) training Black barbers as men's health ambassadors to foster important conversations about HIV in the context of sexual health and wellness, and to share relevant community resources with barbershop patrons and (2) launching a multimedia HIV stigma reduction campaign involving print, audiovisual, and social media in and around Black barbershops in Tennessee.

The social media strategy comprises 26 messages that first introduced and branded the participating barbers as men's health ambassadors; provided information about HIV ("Fact Fridays"); provided motivational or educational information about general health in addition to sexual and mental health and wellness; and provided opportunities for barbers to create and share their own live videos ("Cuts and Convos"), pictures, and text ("Choose Your Style, Choose Your Health"). Participating barbers were asked to post curated messages at least once weekly on their social media platforms (X, Facebook, Instagram, and LinkedIn). Barbers were advised to post on prespecified days and times and to use tailored hashtags to optimize campaign reach. Where possible, barbers also were asked to collaborate with or tag the study team's social media account to further promote the reach of the messaging. The last component of the social media strategy involved the creation of short videos that feature some of the participating barbers in a local barbershop discussing men's sexual health and HIV. Barbers play these short videos on the screens in their barbershops at least 3 times weekly and share them via social media. The goal of the social media strategy and broader multimedia campaign is to spark conversations and engagement regarding HIV and HIV-related stigma, and to help reframe the stigmatizing narrative about HIV in the community.

Data Collection

Each project used constructs from the CFIR [18] in its initial project logic model to describe contextual factors germane to its planned project strategies and activities. Given the comprehensive nature of the CFIR and its preexisting application across projects, we used the CFIR innovation, outer setting, and inner setting domains as conceptual anchors and the domain constructs as a framework to help identify and organize key barriers and facilitators of relevance to implementation of the projects' social media strategies.

The academic project leads of the 3 projects (RGH, JJL, and AAA and JMP) began by identifying constructs relevant to their respective projects within the CFIR innovation, outer setting, and inner setting domains and describing how these constructs were operationalized and addressed in the context of implementing their social media strategies. During this process, academic project leads consulted with other members of their research teams as needed to ensure complete identification of key constructs. Academic project leads then engaged in informal conversations with their respective community partners to elicit and document their experiences with strategy implementation in the context of a research process involving academic-community partnerships. Community partners discussed barriers to and facilitators of strategy implementation, as well as benefits and challenges of participating as community partners in the projects. All original insights represented in this paper reflect this within-project examination of the experiences and perspectives of project team members, including academic and community partners engaged in research and implementation activities.

Data Analysis

The initial constructs identified by academic project leads were compiled in a Microsoft Excel matrix and reviewed for common constructs across the projects by the Hub project director (EPR-C). The matrix and common constructs were presented to the CPSMC working group and reviewed, discussed, and refined over 2 CPSMC working group meetings, as well as through opportunities for individual review and written

XSL•FO
RenderX

feedback. Given the high degree of interrelatedness across CFIR constructs and reported factors, after a set of common constructs was compiled, 2 Hub team members (EPR-C and TBL) worked together to thematically synthesize the content into preliminary practical considerations for use of community-partnered social media strategies in HIV research.

Notes from the individual conversations between academic project leads and community partners were reviewed and assessed by EPR-C and TBL for concordance with the preliminary practical considerations. The practical considerations were then modified to integrate community partner perspectives and experiences. The modified set of practical considerations was then reviewed over 2 CPSMC working group meetings, opportunities for individual review and written feedback, and iteration by EPR-C and TBL until consensus was reached.

Ethical Considerations

The 3 projects discussed in this paper are considered research and were, respectively, reviewed and approved by the Johns Hopkins University School of Medicine IRB (approval #IRB00377619; project A), the University of Washington IRB (approval #STUDY00015837; project B), and the Vanderbilt IRB (approval #240024; project C). Ethics approval and consent to participate is not applicable to the initiative described in this paper.

Findings

Overview

The five practical considerations identified by the CPSMC working group to inform use of community-partnered social media strategies in HIV research include the following: (1) the power and pitfalls of social media platforms, (2) messengers and messages matter, (3) the significance of the sociopolitical environment, (4) investing in academic-community partnerships, and (5) the alignment of the institutional environment and research approach. Each consideration and relevant subconsiderations are detailed in the following subsections in the form of synthesized reflections and recommendations from the project teams, drawing on relevant examples from the 3 projects' experiences in the field.

Consideration 1: The Power and Pitfalls of Social Media Platforms

Numerous opportunities and challenges inherent to social media platforms may affect the use of social media strategies in HIV research. While each social media platform has unique characteristics that may impact its appropriateness for a given project, our focus is on highlighting considerations relevant to the use of any platform.

Platforms and User Preferences Constantly Evolve

Each social media platform has unique features that evolve over time, continuously interacting with evolving user preferences and business environments. Shifting cultural trends, social dynamics, and platform features perpetually influence the preferences of diverse and heterogeneous user groups. Due to these constant shifts, with new trends continuously emerging and others dissipating, social media strategies that resonate with a particular community today may be less impactful tomorrow.

Ongoing engagement with the community is essential to stay attuned to these changes and to ensure that a campaign remains responsive and effective. For example, in project B, Latino and Latinx MSM community members suggested using particular hashtags and emojis of current relevance to best connect with their community. Notably, keeping pace with user preferences and behaviors requires a significant investment of time and resources. For these 3 projects, ensuring the continued relevance of particular social media approaches demanded ongoing monitoring and adaptation, which may be constrained by time, personnel, and funding. Moreover, evaluating the impact of social media strategies amidst this dynamic environment poses further obstacles, as the metrics and benchmarks for success may themselves undergo rapid transformation.

Projects May Be Limited by Platform Policies and Guidelines

Social media platforms face their own hurdles in managing the evolving landscape of user-generated content, sometimes resulting in the prohibition of content and social media strategies that particularly resonate with target communities. For example, project B noted that they were blocked from posting the content and images that were most popular with the Latino and Latinx MSM community they were trying to reach, posing a major barrier to implementation of their campaign. Furthermore, platform guidelines are constantly changing and may prohibit specific text and images that were previously acceptable for dissemination, as experienced by project B, potentially necessitating adaptations to planned content. Each social media platform also has unique content specifications, such as word count limits; varying levels of reliance on static images, videos, and text-based content; and restrictions on use of hyperlinks that must be considered in assessing its appropriateness for a given project.

Use of Social Media Introduces Unique Data Use and Security Complexities

Researchers must be aware of the potential risks to the privacy and confidentiality of participants inherent in using these platforms, as the evolving nature of social media introduces uncertainties regarding data use, security, and participant anonymity. Limited guidelines exist for how to best address these risks in the context of research, but individual projects should consider potential ways to mitigate and acknowledge these risks as they develop and implement their use of social media. The CPSMC working group projects highlighted the importance of early communication with IRBs regarding recommended practices for addressing data use and security concerns to avoid extended delays in IRB review and project implementation.

Reaching Your Population of Interest With Social Media

Social media is a powerful tool that can facilitate opportunities to reach populations that have not been effectively reached through traditional clinical and public health outreach, particularly given the widespread popularity and accessibility of social media platforms in the United States. All 3 CPSMC

XSL•FO
RenderX

working group projects emphasized the importance of assessing and addressing the unique needs and preferences of their populations of interest in identifying the appropriate platforms and approaches to most effectively reach them. Furthermore, while social media can facilitate reach to large and diverse communities, a disconnect may exist between the intended audience and actual message recipients. Factors such as algorithmic biases, user engagement patterns, and content sharing dynamics can influence who ultimately receives and interacts with the content. Platforms may offer some targeting options, but they may not always be as granular as desired to reach particular groups (eg, MSM and trans women). Even when these features exist, some individuals might not be open about their identities nor willing to disclose their identities on social media platforms. These limitations may add another layer of complexity to the implementation of social media campaigns for specific communities. While social media platforms may support reaching certain populations, the level of specificity available for targeting may not adequately align with project goals.

The high potential reach of social media platforms may also pose particular challenges for community partners. For example, project C's men's health ambassadors initially raised concerns about how far their messaging may go on social media given their local environment of HIV-related stigma and discrimination. Conversely, while project B's community partners highlighted their excitement about potentially removing the geographic boundaries of their impact via the use of social media, they were not able to do so in practice because their organization only had the capacity to provide local services.

Social Media Is Increasingly Part of Routine Operations in Community Organizations

All 3 projects identified increased use of and investment in social media among community partner organizations in recent years. For example, the clinical organization partnering with researchers on project A noted that they have increased staff dedicated to social media as well as enhanced attention to metrics on the impact of their social media channels. Likewise, the organization has increasingly relied upon social media for outreach to populations of interest, including the AYA populations of focus in project A. The prevalent use of social media by community organizations may help facilitate the use of social media strategies in community-partnered HIV research.

Consideration 2: Messengers and Messages Matter

Identifying Your Messengers

The messengers a research team uses to deliver their social media strategy can have a pronounced impact on strategy success. All 3 CPSMC working group projects gave significant consideration to identifying messengers who would draw the attention and trust of the communities they were seeking to reach. Messengers may include individuals represented as messengers in social media images as well as the message poster or source. For example, project A clearly indicated the source of the content and visibly displayed the logo or name of the trusted partner organization in social media images to help support engagement. The CPSMC working group projects highlighted that if social media messages appear to be coming from a source that is not accepted, relevant, and credible in the eyes of the target audience, the audience is unlikely to positively engage with the content.

Involving Social Media Experts to Ensure High-Quality Content

Social media expertise is a crucial element of successful social media strategy implementation. It may be provided by members of the project team or by external contracted entities. CPSMC working group members noted that public communications and social media education are typically not included in the traditional investigative training pathway, which may result in projects needing to rely more heavily on external expertise. When possible, it is useful to work with experts in social media who have experience communicating about HIV or other potentially sensitive health topics to the population of interest. For example, project C contracted with a Black-owned media and marketing company with expertise in health interventions to help develop their HIV stigma reduction multimedia campaign. As one of their men's health ambassadors noted, "Social media marketing can be effective if it's done really well or people will remove it, block it."

The Importance of Community Input

Community partners are necessary opinion leaders, champions, and external change agents. All 3 projects emphasized that any HIV-focused study intending to implement social media strategies should consider how best to include community partners who reflect and are meaningfully connected to the population they are working to reach throughout their strategy development and implementation process. Community input on social media strategies may also be gathered via focus groups with members of the population a study intends to reach (as was done in project B), pilot-testing, and other idea generation and feedback opportunities with community members. Community partners involved in project B emphasized that "overall, community members were responsive to our campaign because it was designed for Latinos by Latinos." The 3 projects stressed that the quality of social media messaging reflects the care that was taken to design the message to reach, connect with, and inform the population of interest, which requires community input.

While community input is crucial and may offer consistent guidance on certain elements of a social media strategy, it is also important to acknowledge the heterogeneity in communities. All 3 projects observed diverse community member opinions that did not always lead to consensus about the preferred look and feel of their social media content. For example, project A shared that photographs of local youth self-reporting as sexually or gender diverse within an Instagram campaign were considered by some community partners to be denoting too much vulnerability and not demonstrating empowerment. Likewise, project B noted that some participants desired images that "looked like them," which was a key element of their content development process, but given the diversity within the Latino and Latinx MSM community, not everyone always felt represented. Still, all 3 CPSMC working group projects agreed that a community-engaged development process

XSL•FO
RenderX

was invaluable to refining messages and identifying trusted messengers.

Your Content in Context

The context in which an HIV social media strategy is implemented has major implications for how content may be received. The high potential reach unique to the use of social media strategies can pose challenges in contexts where HIV is highly stigmatized. As a men's health ambassador involved in project C shared, some people in their local area of Tennessee "don't want to be attached to [HIV] and don't want to be bothered." While the potential to reach a large audience is generally advantageous for spreading awareness and disseminating information, it may pose additional challenges for research efforts navigating sensitive topics in environments marked by stigma and discrimination. To minimize this risk, the CPSMC working group projects emphasized that researchers using social media strategies should use inclusive and empowering content that presents communities in ways that align with how they desire to be portrayed, being careful to avoid perpetuating stigma.

Relatedly, an important piece of guidance that arose from community input in these projects was to avoid negative framing of disparate health outcomes. Community members emphasized that images and text should have a positive tone that avoids stereotypes or foreboding messages. Paying attention to communities' strengths and resilience may be particularly important for social media strategies focused on marginalized groups, especially in contexts where there is a high level of stigma related to HIV and sexual health. Research teams should also anticipate potential mistrust by community members around targeting specific populations. Studies may need to actively work to establish trust, demonstrate trustworthiness, and consider historical harms in communities of interest to come to a place where their content can be positively received and considered credible by community members.

Consideration 3: The Significance of the Sociopolitical Environment

The sociopolitical environment surrounding HIV and its effect on implementing social media strategies to reach priority populations can have a tremendous impact on project success. To illustrate the importance of considering the sociopolitical environment, these dynamics are explored in the context of each project in the following subsections.

Project A: Social Media as a Recruitment Strategy When Traditional Strategies Are Nonviable Due to Local Attitudes Toward the Target Population

Lesbian, gay, bisexual, transgender, and queer individuals in Florida, as in other states, face worsening disparities in health care access and treatment that heavily influence the implementation of HIV-focused research. Sexually and gender diverse youth within the state are particularly susceptible to severe discrimination, lack of culturally competent care, and marginalization as past HIV-focused treatment and prevention programs were largely closed or minimized in the wake of legislative measures, including the "Don't Say Gay" bill [25] as well as a ban on gender-affirming treatment for youth [26] passed within the state in 2023. As a consequence of these sweeping policy changes, traditional recruitment strategies, such as in-person discussions, use of flyers, and brick-and-mortar-based information, were no longer available for project A, which focused on the inclusion of sexually and gender diverse youth. Furthermore, within the pediatric health care organization involved in the project, clinical programs, such as the adolescent medicine and gender-affirming care clinic, were closed, and clinical and research teams were subject to personal threats for providing care to this population. With local and regional channels to reach and support this highly vulnerable population suspended, social media–based recruitment became the primary and safest method for outreach for project A. This recruitment strategy, while limited to the structure and scope of public-facing platforms, offered the only way by which to communicate information related to sexual health to populations at risk for HIV in a way that protected content from new laws and processes that limit traditional "free speech" opportunities available in other states to engage diverse populations in HIV-focused research.

Project B: Community Acceptability and Supportive Policies Engender a Welcoming Atmosphere for a Social Media Campaign

Project B's decision to develop a social media campaign for Latino and Latinx MSM via their academic-community partnership stemmed from the community's receptiveness to technology-based outreach strategies for HIV prevention. The popularity of social media, coupled with expanded efforts to increase HIV prevention in Washington State and particularly in King County, created a welcoming atmosphere for public initiatives to encourage use of HIV-related services. Specifically, Washington has supportive policies that help eligible people with HIV obtain medications as well as PrEP drug assistance programs that provide financial assistance for people interested in accessing PrEP. Project B's community partner organization was also involved with learning collaboratives as part of King County's EHE efforts, which support health care organizations across the county in better identifying and serving people at risk for HIV. These initiatives demonstrated a readiness and favorable environment for project B and helped facilitate the implementation of its social media campaign.

While perceptions of HIV vary widely across different communities and counties within Washington, progressive policies and supportive programs may contribute to a relatively accepting environment for individuals living with or at risk for HIV. Thus, social media strategies and interventions for HIV prevention may be more feasible in Washington compared to places with more restrictive environments.

Project C: A Social Media Strategy to Address HIV Stigma Plans for Stigma-Based Community Reactions

The broad sociopolitical environment in Tennessee is characterized by various legal and cultural factors that act as barriers to addressing HIV. While pervasive HIV-related stigma has been identified as a key contributor to the disproportionate effects of HIV across the Deep South [27], in Tennessee, this stigma has manifested in some of the most stringent HIV-related criminalization laws in the country. For example, until July

XSL•FO
RenderX

2024, Tennessee was the only US state that imposed a lifetime registration as a "violent sex offender" if convicted of engaging in sex work while living with HIV [28].

Over the last several years, Tennessee has made gradual progress on key HIV prevention and treatment outcomes. However, recent political decisions indicative of continued stigma around HIV and populations most impacted by HIV have the potential to reverse this progress. In 2023, Tennessee rejected millions of federal dollars for HIV prevention, while simultaneously funding prevention efforts from the state's budget and prioritizing new populations (pregnant people, survivors of sex trafficking, and first responders) over populations with the greatest risk and burden of HIV in the state [29]. In the context of project C, the stigma and fear culturally associated with HIV were carefully considered during each step of the social media strategy design and implementation process. In addition, the barbers who became men's health ambassadors as part of the project expressed that social media amplification of the campaign would make some individuals in their communities assume that the barbers have HIV. They said that preparation for how to challenge this assumption would be useful and appreciated. In this way, the broad potential reach of social media was a source of potential concern for project C's strategy implementers who were aware they may become targets of stigma and discriminatory behavior if assumed to have HIV.

Consideration 4: Investing in Academic-Community Partnerships

Identifying the Right Partners

In any academic-community research partnership, cultivating a positive relationship, establishing shared project goals, and identifying the contributions from and benefits for both partners is paramount. For researchers aiming to implement HIV-related social media strategies, this process starts with identifying a community organization that is experienced with and sensitive to HIV-related topics, and trusted by and, where possible, culturally congruent with the population to be reached. It is also helpful for the organization to have social media and communications expertise, ideally specific to reaching the priority population with health messaging. Having a strong previous working relationship with the organization and its leadership is very useful, as was the case with projects A and B, or the study team must invest in building that strong working relationship in preparation for strategy implementation, as in project C [30].

Building Research Literacy

Many community organizations conduct effective programmatic work with populations highly impacted by HIV and have limited knowledge of research. Among the community partners working on these 3 projects, there was variable familiarity and experience with research. For example, while project A's community partner organization, a clinical setting, had previously participated in research, a member of their clinical leadership team shared that "research is just not built into the organization." He also described receiving "blank stares when we talk about data." The CPSMC working group projects noted that researchers seeking to start or strengthen partnerships with community organizations should be prepared for a potential lack of familiarity, comfort, and interest in research among staff. Likewise, key elements of research projects, such as administration of grant funding, project roles, and project timelines, may not be intuitive to community partner staff, which may cause challenges during project initiation and implementation. Sharing the research purpose and components and outlining the practical, administrative elements of project implementation with community partners early on will help build research literacy and lay the foundation for a successful academic-community partnership.

Communicating Effectively

With all partnerships, and particularly when working with community organizations newer to research, it is key for the academic partner to provide clear, transparent, and tailored communication that meets the community partner organization where they are. The academic team should endeavor to communicate the relevance and intended impact of the project as it relates to the community organization's work and avoid using technical jargon. In addition, routinely inviting and responding to questions from community partners is important. All 3 CPSMC working group projects noted that trust is at the core of any positive partnership, and therefore, communication between partners should be timely and reliable.

Pursuing the Win-Win

Often, academic-community partnerships are developed primarily to serve research-related goals set forth by the academic partner. Funds are typically provided to community partners via subcontract to support their participation. However, projects A and C noted that in practice, it can remain challenging to add research activities to a staff person's plate outside of their routine role. A member of the clinical leadership team for project A's community partner organization also noted it can be tricky to assign and supervise research-related roles among existing staff, asking "How do I give feedback on a role they weren't hired for?"

Given these challenges, it is crucial for researchers to work to identify and support benefits of participation for community partners. Researchers may consider eliciting information from staff at their partner organization about what they hope to gain from participating. For example, becoming men's health ambassadors was seen as a growth opportunity for some of the barbers participating in project C. One commented they wanted to be "a part of something greater" by participating in this project to improve HIV-related outcomes. Some also expressed concern about not being integrated into future efforts after "building a team of believers" through this work, presenting an important opportunity for the academic team to consider how to keep them engaged as long-term champions. Ensuring allocation of appropriate financial incentives for community partners is also key, particularly for those who may not have consistent wages paid by the organization, as was the case in project C.

XSL•FO
RenderX

Consideration 5: The Alignment of the Institutional Environment and Research Approach

Overview

The 3 CPSMC working group projects unanimously agreed that while academic-community partnerships have increasingly been used in HIV research to facilitate the development of interventions and strategies tailored to the unique needs of communities most impacted by HIV, this approach remains unfamiliar to many academic institutions and their IRBs. All 3 projects experienced tension between the increased national priority for HIV researchers to conduct community-partnered research and institutional structures that were not created with community-partnered research frameworks in mind nor designed to support community-partnered research. The projects reported struggling with IRB-related policies and infrastructures that hinder the implementation of timely, bidirectional community-partnered research. Additional challenges included institutional restrictions and limited guidance regarding implementing social media strategies as part of these research endeavors. This section highlights the importance of assessing and working to address features of institutional environments that may impact implementation of social media strategies in community-partnered HIV research, exemplified by the experiences of the CPSMC working group projects.

Assessing Institutional Preparedness to Support Your Research

All 3 projects described experiencing significant institutional constraints and generally conservative approaches by institutional IRBs when reviewing their community-partnered HIV research protocols. They commented that despite community engagement and partnership increasingly being included as required elements of proposals for HIV research funding nationally, community-partnered research methodologies are still considered novel and are misaligned with the research infrastructures of many academic institutions. For example, they reported challenges with the inclusion of community partners as "coinvestigators" on IRB protocols, limiting true partnership and coownership of studies. They also highlighted a lack of resources and support from IRBs to help community partners become familiar with institutional processes and requirements, such as training modules on human research protections tailored for community partners.

These challenges were heightened when combining a community-partnered HIV research approach with social media strategies. Although most IRBs are well-attuned to supporting traditional recruitment strategies, projects noted that there appear to be few precedents as to how social media and other digital platforms could be used to connect with potential participants while still protecting participant privacy and confidentiality. In addition, because user interactions with content feed each social media platform's private algorithms that shape future content displays, participants may receive future information on the platform that is uncontrolled by and unpredictable for the research team and their institutional IRB. This uncertainty may create potential challenges for the IRB's efforts to ensure human participants' protection and sometimes results in significant constraints on the research team's social media strategy design and implementation. Furthermore, the CPSMC working group projects found their IRBs to be generally unfamiliar with the use of nonstatic content (eg, videos and reels). They noted that IRBs often do not have guidelines or systems for assessing these dynamic forms of information sharing, which are increasingly popular among social media users. Although social media–based approaches to recruitment, engagement, and intervention delivery are increasingly used in research and have been highly effective in reaching populations most impacted by HIV, the experiences of these 3 projects suggest that institutional IRBs may be underprepared to support implementation of social media strategies.

Navigating the Impact of Your Institutional Environment on Your Research

These challenges place community-partnered HIV research projects implementing social media strategies in situations where projected timelines are almost certain to be delayed pending their institution's readiness to provide relevant support and guidance for their research approach. Two of the 3 CPSMC working group projects experienced major IRB-related delays that halted important study progress for several months. Crucially, these institutional challenges and resulting delays can create delays in payment and confusion and frustration among community partners.

The projects noted that some of this frustration among community partners may have stemmed from a lack of understanding about the function of IRBs and how they work, further reflecting the need for academic partners to invest in building research literacy among all members of their project teams. Likewise, some partners were not familiar with the risks and responsibilities of engaging in human participants research and complying with IRB requirements. Certain requirements were also found off-putting by community partners, such as the requirement to engage in an informed consent process using a document that contained standard IRB language that felt "uncomfortable" or "unnatural." Compounded with restrictions and delays that hindered project progress, project C noted that this discomfort resulted in some community partners questioning the intentions of and need to rely on academic institutions. One partner from project C asked, "Do we always have to wait on something to make a difference?" and "Do we always need some outside source for the community, or do we have what we need?"

The CPSMC working group projects came to consensus in emphasizing that academic partners are ultimately responsible for ensuring community partners are equipped to adhere to institutional policies and procedures and are prepared for administrative processes that may cause delays in project activities and, in some cases, payment to community partners. Given the observed lack of resources and support from IRBs to help community partners become familiar with institutional processes and requirements, researchers must be prepared to provide education about what to expect with respect to the academic institution; subcontracting processes; and the purpose, policies, and procedures of the IRB that are relevant to community partner roles in the project. Researchers should also be prepared to proactively advocate for the needs of their

XSL•FO
RenderX

community partners in an attempt to support strong and sustainable academic-community partnerships.

Discussion

Implications of Findings

The federal EHE initiative has brought renewed attention and investment to addressing the HIV epidemic in the United States. A "whole-of-society" initiative, EHE emphasizes innovative, community-engaged approaches to optimizing HIV research and practice [31]. In community-partnered HIV research, social media is increasingly used as a strategy to reach potential participants and deliver interventions [5-7]. However, it remains a newer approach, and there is little practical guidance available for HIV researchers interested in using social media strategies. This article features the voices and experiences of 3 research teams and their community partners, offering real-world, pragmatic insights from their projects that are more likely to be shared informally among the research community than published in peer-reviewed journals. These insights provide essential guidance to inform the effective implementation of community-partnered social media strategies in HIV research. Five practical considerations were identified, addressing the unique nexus of HIV research, academic-community partnerships, and social media strategies.

While we intend for these notes from the field to offer pragmatic guidance for researchers, this article also serves to spotlight salient implementation barriers commonly identified across 3 distinct HIV research projects. Despite previous calls to address the discordance between institutional structures and community-partnered research approaches [32], the experiences of these 3 projects demonstrate that there remains an urgent need to develop solutions to this ongoing challenge. The barriers experienced by the CPSMC working group projects are largely consistent with those identified in other recent literature. For example, a scoping review of 15 US-based community-engaged research projects by Onakomaiya et al [33] described four key challenges related to working with IRBs: "(1) community partners not being recognized as research partners; (2) cultural competence, language of consent forms, and literacy level of partners; (3) IRBs apply formulaic approaches to community-engaged research; and (4) extensive delays in IRB preparation and approval potentially stifle the relationships with community partners." While there is increased emphasis on community-engaged methods by academic institutions and federal funders, as evidenced by their calls for community-partnered research applications, this ongoing misalignment reflects a culture and infrastructure that are not fully prepared to support community-partnered research. Furthermore, as identified by these 3 projects, these barriers limit researchers' abilities to build strong, sustainable relationships with community partners. Educating IRBs on community-partnered research methodologies, revisiting policies that create barriers to community-partnered research [34], and developing and disseminating human participants research training modules tailored for community partners [35] could help bridge the gap between current institutional infrastructures and community-partnered research approaches.

Likewise, IRBs have limited preparedness to assess and support projects using social media strategies. As noted by the CPSMC working group projects, use of social media in human participants research has raised new ethical and regulatory complexities, and there remains limited guidance for IRBs from federal regulatory bodies. In the absence of that guidance, IRBs and researchers have a responsibility to carefully consider the unique participant privacy and autonomy implications of using social media strategies, balanced with the immense positive potential of using these tools in research endeavors. While protecting research participants is important for engaging in responsible research activities, we are potentially losing creative, innovative, and effective approaches to help end the HIV epidemic by not optimizing the implementation of social media strategies in HIV research. The rubric proposed by Bhatia-Lin et al [36] to assist studies to determine ethical and regulation-consistent use of social media platforms may offer a useful framework for these considerations.

More broadly, the insights shared in this paper highlight the need to continue working toward addressing the historical separation between academia and the community and creating a culture that values and respects building and maintaining strong community partnerships. In the context of implementing social media strategies in community-partnered HIV research, partnership with the community is particularly essential for engaging trusted messengers and developing culturally relevant messages. Likewise, it is critical for research teams to proactively address potential communication and research literacy barriers that may hinder the full realization of their community-partnered HIV research and its potential impacts.

While this article provides practical considerations and guidance regarding implementation of social media strategies in community-partnered HIV research, future research is needed to add to the evidence base on best practices for using social media in HIV research and to continue expanding the growing body of literature on community engagement and partnership in HIV research [37]. Further study is also needed to characterize the unique features of various social media platforms and types of content for use in HIV research, and to assess the effectiveness of community-partnered social media strategies in improving HIV-related outcomes. In addition, developing approaches projects can use to design their social media strategy to match their particular context and research aims would be a useful contribution to the HIV research field and beyond. Though this article focuses on HIV-related research, we also anticipate many of the insights shared here would be transferable to investigators implementing social media strategies and conducting community-partnered research in other fields.

This article offers practical considerations in the form of notes from the field derived from 3 ongoing studies. While instructive, these considerations do not comprehensively represent the many types of social media platforms, content, and strategies used in community-partnered HIV research and are limited to the experiences of the 3 CPSMC working group projects. In addition, we do not capture the effectiveness of these strategies as part of this paper, as all investigations are ongoing. Exemplars of effectiveness studies are available in the literature [38,39]

XSL•FO
RenderX

and can provide useful guidance for determining appropriate outcomes for social media strategies.

Conclusions

As use of social media in HIV research and practice continues to grow, the practical considerations presented in this paper can help studies anticipate factors that may impact implementation of community-partnered social media strategies and take early action to mitigate potential challenges. The potential of social media to reach and engage populations most impacted by HIV is significant, and ongoing research is needed to harness its potential. By understanding and addressing the unique challenges and opportunities of social media in community-partnered HIV research, we can leverage these platforms to accelerate progress toward ending the HIV epidemic.

Acknowledgments

The authors would like to thank all the team members of and contributors to the 3 projects described in this manuscript for making these projects possible, contributing to the insights discussed in this manuscript, and helping advance the national effort to end the HIV epidemic in the United States. We would like to specifically acknowledge key collaborators Martha Zuñiga and Joel Aguirre, who were the lead personnel supporting project B from the project's community partner organization, Entre Hermanos. The development of this manuscript was supported by the University of California, Los Angeles Rapid, Rigorous, Relevant (3R) Implementation Science Hub, funded by the National Institute of Mental Health through an administrative supplement award to the Center for HIV Identification, Prevention, and Treatment Services at the University of California, Los Angeles (grant P30MH058107). Project A was funded by the National Institute of Allergy and Infectious Diseases through an administrative supplement award to the Johns Hopkins Center for AIDS Research (grant P30AI094189). Project B was funded by the National Institute of Allergy and Infectious Diseases through an administrative supplement award to the University of Washington and Fred Hutch Center for AIDS Research (grant P30AI027757). Project C was funded by the National Institute of Allergy and Infectious Diseases through an administrative supplement award to the Tennessee Center for AIDS Research (grant P30AI110527). The funding sources had no role in the design, conduct, or reporting of this manuscript or the studies described herein, nor in the decision to submit the manuscript for publication.

Authors' Contributions

EPR-C facilitated the data collection and analysis approaches. The 3 involved projects contributed primary data (leads RGH and LJ; JJL; and AAA, JMP, and SH). EPR-C, TBL, RGH, JJL, AAA, JMP, CL, and SJL contributed to the analysis. EPR-C and TBL synthesized the data and drafted the manuscript outline. EPR-C, TBL, RGH, JJL, AAA, JMP, and CL contributed to the manuscript draft. All authors read, reviewed, and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Barreto JE, Whitehair CL. Social media and web presence for patients and professionals: evolving trends and implications for practice. PM R 2017 May;9(5S):S98-105. [doi: 10.1016/j.pmrj.2017.02.012] [Medline: 28527508]
2. Madhumathi J, Sinha R, Veeraraghavan B, Walia K. Use of "social media"-an option for spreading awareness in infection prevention. Curr Treat Options Infect Dis 2021;13(1):14-31 [FREE Full text] [doi: 10.1007/s40506-020-00244-3] [Medline: 33519303]
3. Pourkarim M, Nayebzadeh S, Alavian SM, Hataminasab SH. Digital marketing: a unique multidisciplinary approach towards the elimination of viral hepatitis. Pathogens 2022 May 29;11(6):626 [FREE Full text] [doi: 10.3390/pathogens11060626] [Medline: 35745480]
4. Nejad MS, Harooni J, Charkazi A, Shams M, Latifi M. Using social marketing to promote preventive behaviors related to COVID-19. J Lifestyle Med 2021 Jul 31;11(2):52-56 [FREE Full text] [doi: 10.15280/jlm.2021.11.2.52] [Medline: 34549026]
5. Kudrati SZ, Hayashi K, Taggart T. Social media and PrEP: a systematic review of social media campaigns to increase PrEP awareness and uptake among young Black and Latinx MSM and women. AIDS Behav 2021 Dec;25(12):4225-4234 [FREE Full text] [doi: 10.1007/s10461-021-03287-9] [Medline: 33939035]
6. Marcus JL, Sewell WC, Balzer LB, Krakower DS. Artificial intelligence and machine learning for HIV prevention: emerging approaches to ending the epidemic. Curr HIV/AIDS Rep 2020 Jun;17(3):171-179 [FREE Full text] [doi: 10.1007/s11904-020-00490-6] [Medline: 32347446]
7. Avery M, Mills SJ, Stephan E. Real-time monitoring through the use of technology to enhance performances throughout HIV cascades. Curr Opin HIV AIDS 2017 Sep;12(5):488-493. [doi: 10.1097/COH.0000000000000397] [Medline: 28650346]
8. Yuan P, Bare MG, Johnson MO, Saberi P. Using online social media for recruitment of human immunodeficiency virus-positive participants: a cross-sectional survey. J Med Internet Res 2014 May 01;16(5):e117 [FREE Full text] [doi: 10.2196/jmir.3229] [Medline: 24784982]

XSL•FO
RenderX

9. Martinez O, Wu E, Shultz AZ, Capote J, López Rios J, Sandfort T, et al. Still a hard-to-reach population? Using social media to recruit Latino gay couples for an HIV intervention adaptation study. J Med Internet Res 2014 Apr 24;16(4):e113 [FREE Full text] [doi: 10.2196/jmir.3311] [Medline: 24763130]
10. Pagkas-Bather J, Young LE, Chen YT, Schneider JA. Social network interventions for HIV transmission elimination. Curr HIV/AIDS Rep 2020 Oct;17(5):450-457 [FREE Full text] [doi: 10.1007/s11904-020-00524-z] [Medline: 32720253]
11. Darko EM, Kleib M, Olson J. Social media use for research participant recruitment: integrative literature review. J Med Internet Res 2022 Aug 04;24(8):e38015 [FREE Full text] [doi: 10.2196/38015] [Medline: 35925655]
12. Zimmermann BM, Willem T, Bredthauer CJ, Buyx A. Correction: ethical issues in social media recruitment for clinical studies: ethical analysis and framework. J Med Internet Res 2022 Sep 07;24(9):e40848 [FREE Full text] [doi: 10.2196/40848] [Medline: 36074800]
13. Campos-Castillo C, Laestadius LI. Racial and ethnic digital divides in posting COVID-19 content on social media among US adults: secondary survey analysis. J Med Internet Res 2020 Jul 03;22(7):e20472 [FREE Full text] [doi: 10.2196/20472] [Medline: 32568726]
14. Grov C, Westmoreland D, Rendina H, Nash D. Seeing is believing? Unique capabilities of internet-only studies as a tool for implementation research on HIV prevention for men who have sex with men: a review of studies and methodological considerations. J Acquir Immune Defic Syndr 2019 Dec;82 Suppl 3(Suppl 3):S253-S260 [FREE Full text] [doi: 10.1097/QAI.0000000000002217] [Medline: 31764261]
15. Taggart T, Grewe ME, Conserve DF, Gliwa C, Roman Isler M. Social media and HIV: a systematic review of uses of social media in HIV communication. J Med Internet Res 2015 Nov 02;17(11):e248 [FREE Full text] [doi: 10.2196/jmir.4387] [Medline: 26525289]
16. Elliott SA, Dyson MP, Wilkes GV, Zimmermann GL, Chambers CT, Wittmeier KD, et al. Considerations for health researchers using social media for knowledge translation: multiple case study. J Med Internet Res 2020 Jul 23;22(7):e15121 [FREE Full text] [doi: 10.2196/15121] [Medline: 32706653]
17. Ending the HIV Epidemic projects. HIV Implementation Science Coordination Initiative. URL: https://hivimpsci.northwestern.edu/ehe-supplement-projects/ [accessed 2025-08-15]
18. Damschroder LJ, Reardon CM, Widerquist MA, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022 Oct 29;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
19. Vilidnitsky AG, Hernandez RG, Silva C, Fields EL. HIV testing deserts and vulnerability among adolescents and young adults in Tampa Bay, Florida. J Adolesc Health 2024 May;74(5):1012-1018. [doi: 10.1016/j.jadohealth.2024.01.011] [Medline: 38416099]
20. Lee JJ, Aguirre Herrera J, Cardona J, Cruz LY, Munguía L, Leyva Vera CA, et al. Culturally tailored social media content to reach Latinx immigrant sexual minority men for HIV prevention: web-based feasibility study. JMIR Form Res 2022 Mar 16;6(3):e36446 [FREE Full text] [doi: 10.2196/36446] [Medline: 35182416]
21. Jaccard J. A theoretical analysis of selected factors important to health education strategies. Health Educ Monogr 1975;3(2):152-167. [doi: 10.1177/109019817500300201] [Medline: 1150466]
22. Jaccard J, Dodge T, Dittus P. Parent-adolescent communication about sex and birth control: a conceptual framework. New Dir Child Adolesc Dev 2002(97):9-41. [doi: 10.1002/cd.48] [Medline: 14964942]
23. Guilamo-Ramos V, Jaccard J, Dittus P, Collins S. Parent-adolescent communication about sexual intercourse: an analysis of maternal reluctance to communicate. Health Psychol 2008 Nov;27(6):760-769. [doi: 10.1037/a0013833] [Medline: 19025272]
24. Lee JJ, Aguirre J, Munguia L, Robles G, Ramirez Hernandez K, Ramirez JI, et al. Engagement of Latino immigrant men who have sex with men for HIV prevention through eHealth: preferences across social media platforms. Ethn Health 2022 Oct;27(7):1684-1697 [FREE Full text] [doi: 10.1080/13557858.2021.1943322] [Medline: 34152248]
25. Diaz J. Florida's governor signs controversial law opponents dubbed 'Don't Say Gay'. National Public Radio. URL: https://www.npr.org/2022/03/28/1089221657/dont-say-gay-florida-desantis [accessed 2025-02-28]
26. HRC Staff. BREAKING: Florida senate passes extreme gender-affirming care ban. Human Rights Campaign. 2023 Apr 04. URL: https://www.hrc.org/press-releases/breaking-florida-senate-passes-extreme-gender-affirming-care-ban [accessed 2025-02-28]
27. Reif S, Safley D, McAllaster C, Wilson E, Whetten K. State of HIV in the US Deep South. J Community Health 2017 Oct;42(5):844-853. [doi: 10.1007/s10900-017-0325-8] [Medline: 28247067]
28. Tennessee. The Center for HIV Law and Policy. URL: https://www.hivlawandpolicy.org/state-profiles/tennessee [accessed 2025-02-28]
29. Borre ED, Ahonkhai AA, Chi KY, Osman A, Thayer K, Person AK, et al. Projecting the potential clinical and economic impact of HIV prevention resource reallocation in Tennessee. Clin Infect Dis 2024 Dec 17;79(6):1458-1467. [doi: 10.1093/cid/ciae243] [Medline: 38913762]
30. Nwakoby C, Pierce LJ, Crawford R, Conserve D, Perkins J, Hurt S, et al. Establishing an academic-community partnership to explore the potential of barbers and barbershops in the Southern United States to address racial disparities in HIV care

XSL•FO
RenderX

outcomes for black men living with HIV. Am J Mens Health 2023 09;17(1):15579883231152114 [FREE Full text] [doi: 10.1177/15579883231152114] [Medline: 36757054]
31. EHE overview. HIV.gov. URL: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview [accessed 2025-02-28]
32. Hippalgaonkar N, Nguyen RH, Cohn EB, Horowitz J, Waite AW, Mersha T, et al. Are we the problem? A call to action for addressing institutional challenges to engaging community partners in research. Int J Environ Res Public Health 2024 Feb 17;21(2):236 [FREE Full text] [doi: 10.3390/ijerph21020236] [Medline: 38397725]
33. Onakomaiya D, Pan J, Roberts T, Tan H, Nadkarni S, Godina M, et al. Challenges and recommendations to improve institutional review boards' review of community-engaged research proposals: a scoping review. J Clin Transl Sci 2023 Mar 31;7(1):e93. [doi: 10.1017/cts.2023.516] [Medline: 37125052]
34. Windsor L, Benoit E, Kwan P, Tan K, Richmond A. Protection of participants in community-engaged research by institutional review boards: a call for action. Am J Public Health 2024 May;114(S5):S360-S365. [doi: 10.2105/AJPH.2024.307592] [Medline: 38547467]
35. Yonas MA, Jaime MC, Barone J, Valenti S, Documét P, Ryan CM, et al. Community partnered research ethics training in practice: a collaborative approach to certification. J Empir Res Hum Res Ethics 2016 Apr;11(2):97-105 [FREE Full text] [doi: 10.1177/1556264616650802] [Medline: 27241871]
36. Bhatia-Lin A, Boon-Dooley A, Roberts MK, Pronai C, Fisher D, Parker L, et al. Ethical and regulatory considerations for using social media platforms to locate and track research participants. Am J Bioeth 2019 Jun;19(6):47-61 [FREE Full text] [doi: 10.1080/15265161.2019.1602176] [Medline: 31135323]
37. Karris MY, Dubé K, Moore AA. What lessons it might teach us? Community engagement in HIV research. Curr Opin HIV AIDS 2020 Mar;15(2):142-149 [FREE Full text] [doi: 10.1097/COH.0000000000000605] [Medline: 31895141]
38. Cao B, Gupta S, Wang J, Hightow-Weidman LB, Muessig KE, Tang W, et al. Social media interventions to promote HIV testing, linkage, adherence, and retention: systematic review and meta-analysis. J Med Internet Res 2017 Nov 24;19(11):e394 [FREE Full text] [doi: 10.2196/jmir.7997] [Medline: 29175811]
39. Ibrahim K, Kahle EM, Christiani Y, Suryani S. Utilization of social media for the prevention and control of HIV/AIDS: a scoping review. J Multidiscip Healthc 2024;17:2443-2458 [FREE Full text] [doi: 10.2147/JMDH.S465905] [Medline: 38799010]

Abbreviations

3R: University of California, Los Angeles Rapid, Rigorous, Relevant
AYA: adolescent and young adult
CFIR: Consolidated Framework for Implementation Research
CPSMC: community-partnered social media campaigns
EHE: Ending the HIV Epidemic
IRB: institutional review board
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis

Edited by Z Hou; submitted 01.03.25; peer-reviewed by AAS Sawitri, D Dinh; comments to author 25.04.25; revised version received 25.05.25; accepted 05.06.25; published 10.09.25.

Please cite as:
Rosenberg-Carlson EP, Loeb TB, Hernandez RG, Lee JJ, Ahonkhai AA, Perkins JM, Lin C, Lee SJ, Hurt S, Johnsen L, Hamilton AB Implementing Social Media Strategies in Community-Partnered HIV Research: Practical Considerations From 3 Ongoing Studies JMIR Public Health Surveill 2025;11:e73318
URL: https://publichealth.jmir.org/2025/1/e73318
doi:10.2196/73318
PMID:40929715

©Elena P Rosenberg-Carlson, Tamra B Loeb, Raquel G Hernandez, Jane J Lee, Aima A Ahonkhai, Jessica M Perkins, Chunqing Lin, Sung-Jae Lee, Sharon Hurt, Luke Johnsen, Alison B Hamilton. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Assessing the Impact of Frailty on Infection Risk in Older Adults: Prospective Observational Cohort Study

Ya Yang[1*], MPH; Kechun Che[1*], MPH; Jiayan Deng[1], MPH; Xinming Tang[1], MPH; Wenyuan Jing[1], MPH; Xiuping He[1], MPH; Jiacheng Yang[1], MPH; Wenya Zhang[1], MPH; Mingjuan Yin[1], MPH; Congcong Pan[1], MMSc; Xiaoling Huang[2], MPH; Zewu Zhang[3], MPH; Jindong Ni[1], PhD

[1]School of Public Health, Shunde Women and Children's Hospital, Precision Key Laboratory of Public Health, Guangdong Medical University, No.1 Xincheng Road, Dongguan, China

[2]Office of Public Health, Songshan Lake Community Health Service Centre, DongGuan, China

[3]Institute for Infectious Disease Prevention and Control, DongGuan Centre for Disease Control and Prevention, DongGuan, China

[*]these authors contributed equally

Corresponding Author:
Jindong Ni, PhD
School of Public Health, Shunde Women and Children's Hospital, Precision Key Laboratory of Public Health, Guangdong Medical University, No.1 Xincheng Road, Dongguan, China

Related Article:

Correction of: https://publichealth.jmir.org/2024/1/e59762/

(JMIR Public Health Surveill 2025;11:e78866) doi:10.2196/78866

In "Assessing the Impact of Frailty on Infection Risk in Older Adults: Prospective Observational Cohort Study" (JMIR Public Health Surveill. 2024;10:e59762) the authors noted one error.

Affiliations 1 and 2 for authors Ya Yang, Kechun Che, Jiayan Deng, Xinming Tang, Wenyuan Jing, Xiuping He, Jiacheng Yang, Wenya Zhang, Mingjuan Yin, Congcong Pan, and Jindong Ni have been combined, from:

> *School of Public Health, Shunde Women and Children's Hospital, Foshan, China*

And:

> *Precision Key Laboratory of Public Health, Guangdong Medical University, DongGuan, China*

To read as:

> *School of Public Health, Shunde Women and Children's Hospital, Precision Key Laboratory of Public Health, Guangdong Medical University, Dongguan, China*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 11.06.25; this is a non–peer-reviewed article;accepted 11.06.25; published 19.06.25.

Please cite as:
Yang Y, Che K, Deng J, Tang X, Jing W, He X, Yang J, Zhang W, Yin M, Pan C, Huang X, Zhang Z, Ni J
Correction: Assessing the Impact of Frailty on Infection Risk in Older Adults: Prospective Observational Cohort Study
JMIR Public Health Surveill 2025;11:e78866
URL: https://publichealth.jmir.org/2025/1/e78866
doi:10.2196/78866

©Ya Yang, Kechun Che, Jiayan Deng, Xinming Tang, Wenyuan Jing, Xiuping He, Jiacheng Yang, Wenya Zhang, Mingjuan Yin, Congcong Pan, Xiaoling Huang, Zewu Zhang, Jindong Ni. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete

XSL•FO
RenderX

bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Meal Timing and Depression Among Chinese Children and Adolescents: Cross-Sectional Study

Huilun Li[1,2*], MS; Zhaohui Lu[1*], PhD; Erliang Zhang[2*], BS; Jie Zhang[2], BS; Shuheng Cui[2], BS; Masaki Takahashi[3], PhD; Mi Xiang[1,2], PhD

[1]Hainan Branch, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, Sanya, China

[2]School of Public Health, Shanghai Jiao Tong University, 227 S Chongqing Rd, Shanghai, China

[3]Institute for Liberal Arts, Tokyo Institute of Technology, Tokyo, Japan

[*]these authors contributed equally

Corresponding Author:
Mi Xiang, PhD
Hainan Branch, Shanghai Children's Medical Center, School of Medicine, Shanghai Jiao Tong University, Sanya, China

Related Article:

Correction of: https://publichealth.jmir.org/2024/1/e54275

(JMIR Public Health Surveill 2025;11:e73716) doi:10.2196/73716

In "Meal Timing and Depression Among Chinese Children and Adolescents: Cross-Sectional Study" (JMIR Public Health Surveill 2024;10:e54275) the authors noted one error.

The funding information was absent from the final article, which has been added to the Acknowledgements section as follows:

> *This work was supported by The National Natural Science Foundation of China (No. 71804110), Shanghai Science and Technology Development Funds (No. 21QA1405300), Science Foundation for new teachers of Shanghai Jiao Tong University School of Medicine (No. 20X 100040012), Science Foundation of Ministry of Education of China (No.22YJAZH116), and Shanghai Municipal Health Commission (No.: GW-10.1-XK07).*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 10.03.25; this is a non–peer-reviewed article;accepted 10.03.25; published 02.04.25.

Please cite as:
Li H, Lu Z, Zhang E, Zhang J, Cui S, Takahashi M, Xiang M
Correction: Meal Timing and Depression Among Chinese Children and Adolescents: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e73716
URL: https://publichealth.jmir.org/2025/1/e73716
doi:10.2196/73716

© Huilun Li, Zhaohui Lu, Erliang Zhang, Jie Zhang, Shuheng Cui, Masaki Takahashi, Mi Xiang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 2.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Comparison of Virus Watch COVID-19 Positivity, Incidence, and Hospitalization Rates With Other Surveillance Systems: Surveillance Study

Wing Lam Erica Fong[1], MSc; Vincent Grigori Nguyen[1,2,3], PhD; Sarah Beale[1], PhD; Thomas E Byrne[1], PhD; Cyril Geismar[4], PhD; Ellen Fragaszy[1,5], PhD; Jana Kovar[1], PhD; Alexei Yavlinsky[1], PhD; Ibrahim Abubakar[6], PhD; Andrew C Hayward[2,7], MD; Robert W Aldridge[1,8], PhD

[1]Institute of Health Informatics, University College London, 222 Euston Road, London, United Kingdom

[2]Institute of Epidemiology and Health Care, University College London, London, United Kingdom

[3]Department of Population, Policy and Practice, UCL Great Ormond Street Institute of Child Health, London, United Kingdom

[4]NIHR Health Protection Research Unit in Modelling and Health Economics, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom

[5]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, United Kingdom

[6]Faculty of Population Health Sciences, University College London, London, United Kingdom

[7]NIHR Health Protection Research Unit in Health Analytics and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom

[8]Institute for Health Metrics and Evaluation, University of Washington, Seattle, United States

Corresponding Author:
Wing Lam Erica Fong, MSc
Institute of Health Informatics, University College London, 222 Euston Road, London, United Kingdom

Related Article:

Correction of: https://publichealth.jmir.org/2025/1/e69655

Abstract

(JMIR Public Health Surveill 2025;11:e85512) doi:10.2196/85512

In “Comparison of Virus Watch COVID-19 Positivity, Incidence, and Hospitalization Rates With Other Surveillance Systems: Surveillance Study” [1], the authors noted one error.

Coauthor ACH was mistakenly linked to the fourth affiliation, which has now been removed from the author. The following affiliation was added:

> *NIHR Health Protection Research Unit in Health Analytics and Modelling, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, United Kingdom*

This will appear as the seventh affiliation and be attached to ACH, with any subsequent affiliations being renumbered accordingly.

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Fong WLE, Nguyen VG, Beale S, et al. Comparison of virus watch COVID-19 positivity, incidence, and hospitalization rates with other surveillance systems: surveillance study. JMIR Public Health Surveill 2025 Sep 29;11:e69655. [doi: 10.2196/69655] [Medline: 41021969]

XSL•FO
RenderX

Submitted 08.10.25; this is a non–peer-reviewed article;accepted 08.10.25; published 10.10.25.

Please cite as:
Fong WLE, Nguyen VG, Beale S, Byrne TE, Geismar C, Fragaszy E, Kovar J, Yavlinsky A, Abubakar I, Hayward AC, Aldridge RW
Correction: Comparison of Virus Watch COVID-19 Positivity, Incidence, and Hospitalization Rates With Other Surveillance Systems: Surveillance Study
JMIR Public Health Surveill 2025;11:e85512
URL: https://publichealth.jmir.org/2025/1/e85512
doi:10.2196/85512

© Wing Lam Erica Fong, Vincent Grigori Nguyen, Sarah Beale, Thomas E Byrne, Cyril Geismar, Ellen Fragaszy, Jana Kovar, Alexei Yavlinsky, Ibrahim Abubakar, Andrew C Hayward, Robert W Aldridge. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Assessment of Bidirectional and Threshold-Response Associations Between Cognitive Function and Physical Performance: Nationwide Cross-Sectional Study

Huixiu Hu[1]; Lanying Xie[2*]; Yuqing Hao[3*]; Yajie Zhao[4]; Huanhuan Luo[5]; Kang Yu[6]; Chao Sun[1]

[1]Department of Nursing, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, Beijing Hospital, 1 Da Hua Road, DongDan, Dongcheng District, Beijing, China

[2]School of Nursing, Beijing University of Chinese Medicine, Beijing, China

[3]Beijing Hospital, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[4]Department of Cardiology, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, Beijing Hospital, Beijing, China

[5]School of nursing, Peking University, Beijing, China

[6]Department of Clinical Nutrition, Chinese Academy of Medical Science and Peking Union Medical College, Peking Union Medical College Hospital, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Chao Sun
Department of Nursing, National Center of Gerontology, Institute of Geriatric Medicine, Chinese Academy of Medical Science, Beijing Hospital, 1 Da Hua Road, DongDan, Dongcheng District, Beijing, China

Related Article:

Correction of: https://publichealth.jmir.org/2025/1/e80575/

(JMIR Public Health Surveill 2025;11:e88582) doi:10.2196/88582

In "Assessment of Bidirectional and Threshold-Response Associations Between Cognitive Function and Physical Performance: Nationwide Cross-Sectional Study" [1], the authors noted one error in the abstract.

In the original paper, the following information was present:

> *This multicenter study of 20,868 older adults in China analyzed the bidirectional asymmetric and threshold-response associations between cognitive and physical function using logistic or linear regression and restricted cubic spline, with subgroup and interaction analyses.*

This has been revised to the following:

> *This multicenter study of 20,671 older adults in China analyzed the bidirectional asymmetric and threshold-response associations between cognitive and physical function using logistic or linear regression and restricted cubic spline, with subgroup and interaction analyses.*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Hu H, Xie L, Hao Y, et al. Assessment of bidirectional and threshold-response associations between cognitive function and physical performance: nationwide cross-sectional study. JMIR Public Health Surveill 2025 Nov 13;11:e80575. [doi: 10.2196/80575] [Medline: 41232033]

XSL•FO
RenderX

Submitted 27.11.25; this is a non–peer-reviewed article;accepted 27.11.25; published 03.12.25.

Please cite as:
Hu H, Xie L, Hao Y, Zhao Y, Luo H, Yu K, Sun C
Correction: Assessment of Bidirectional and Threshold-Response Associations Between Cognitive Function and Physical Performance: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e88582
URL: https://publichealth.jmir.org/2025/1/e88582
doi:10.2196/88582

© Huixiu Hu, Lanying Xie, Yuqing Hao, Yajie Zhao, Huanhuan Luo, Kang Yu, Chao Sun. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 3.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Building and Developing a Tool (PANDEM-2 Dashboard) to Strengthen Pandemic Management: Participatory Design Study

Carlos Tighe[1], MA; Lem Ngongalah[2], PhD; Alexis Sentís[3], MD, PhD; Francisco Orchard[4], BSc; Gheorghe-Aurel Pacurar[5], Google Agile Project Management Professional; Conor Hayes[6], PhD; Jessica S Hayes[7], PhD; Adrian Toader[8], BSc; Máire A Connolly[7], MB, MPH

[1]Insight, SFI Research Centre for Data Analytics, University of Galway, Galway, Ireland
[2]Trilateral Research Ltd, London, United Kingdom
[3]Epidemiology Department, Epiconcept, Paris, France
[4]IT Department, Epiconcept, Paris, France
[5]PMO, Modus Create, Cluj Napoca, Romania
[6]School of Computer Science, University of Galway, Galway, Ireland
[7]School of Health Sciences, University of Galway, Galway, Ireland
[8]Enterprise Engineering, Modus Create, Cluj Napoca, Romania

Corresponding Author:
Jessica S Hayes, PhD
School of Health Sciences, University of Galway
University of Galway, University Road, Galway, Ireland
Galway, H91 TK33
Ireland
Phone: 353 91524411
Email: jessica.hayes@universityofgalway.ie

Abstract

Background: The COVID-19 pandemic exposed challenges in pandemic management, particularly in real-time data sharing and effective decision-making. Data protection concerns and the lack of data interoperability and standardization hindered the collection, analysis, and interpretation of critical information. Effective data visualization and customization are essential to facilitate decision-making.

Objective: This study describes the development of the PANDEM-2 dashboard, a system providing a standardized and interactive platform for decision-making in pandemic management. It outlines the participatory approaches used to involve expert end users in its development and addresses key considerations of privacy, data protection, and ethical and social issues.

Methods: Development was informed by a review of 25 publicly available COVID-19 dashboards, leading to the creation of a visualization catalog. User requirements were gathered through workshops and consultations with 20 experts from various health care and public health professions in 13 European Union countries. These were further refined by mapping variables and indicators required to fulfill the identified needs. Through a participatory design process, end users interacted with a preprototype platform, explored potential interface designs, and provided feedback to refine the system's components. Potential privacy, data protection, and ethical and social risks associated with the technology, along with mitigation strategies, were identified through an iterative impact assessment.

Results: Key variables incorporated into the PANDEM-2 dashboard included case rates, number of deaths, mortality rates, hospital resources, hospital admissions, testing, contact tracing, and vaccination uptake. Cases, deaths, and vaccination uptake were prioritized as the most relevant and readily available variables. However, data gaps, particularly in contact tracing and mortality rates, highlighted the need for better data collection and reporting mechanisms. User feedback emphasized the importance of diverse data visualization formats combining different data types, as well as analyzing data across various time frames. Users also expressed interest in generating custom visualizations and reports, especially on the impact of government interventions. Participants noted challenges in data reporting, such as inconsistencies in reporting levels, time intervals, the need for standardization between member states, and General Data Protection Regulation concerns for data sharing. Identified risks included ethical concerns (accessibility, user autonomy, responsible use, transparency, and accountability), privacy and data protection (security and access controls and data reidentification), and social issues (unintentional bias, data quality and accuracy, dependency on

XSL•FO
RenderX

technology, and collaborative development). Mitigation measures focused on designing user-friendly interfaces, implementing robust security protocols, and promoting cross-member state collaboration.

Conclusions: The PANDEM-2 dashboard provides an adaptable, user-friendly platform for pandemic preparedness and response. Our findings highlight the critical role of data interoperability, cross-border collaboration, and custom IT tools in strengthening future health crisis management. They also offer valuable insights into the challenges and opportunities in developing IT solutions to support pandemic preparedness.

(JMIR Public Health Surveill 2025;11:e52119) doi:10.2196/52119

KEYWORDS

pandemic preparedness and response; COVID-19; cross-border collaboration; surveillance; data collection; data standardization; data sharing; dashboard; IT system; IT tools

Introduction

Background

The COVID-19 pandemic posed considerable challenges to European countries with widespread variation in response capacity. The COVID-19 pandemic revealed shortcomings in coordination, information sharing, and response strategies among countries, leading to fragmented approaches and delayed actions [1-3]. The lack of standardized data collection and analysis methods hindered the ability to assess the situation comprehensively, to allow data sharing and strengthen or facilitate cross-border collaboration, and to make informed decisions [4,5]. In addition, limited resources and infrastructure strained health care systems, exacerbating the impact of the COVID-19 pandemic. Lessons learned from the COVID-19 experience emphasize the critical need for enhanced preparedness and response efforts [6].

In response to the above challenges, the European Union (EU) has mandated the development and implementation of robust measures for pandemic preparedness and response across member states [7]. This mandate highlights the significance of reinforcing coordination and information exchange, as well as strengthening early detection of threats, monitoring and forecasting of diseases and resources (eg, beds and vaccines), rapid risk assessments, and the capacity of evaluating medical countermeasures (eg, therapeutics) and nonpharmaceutical interventions (eg, social distancing measures such as school closures or lockdowns). The aim is that these activities will build capacity to enhance preparedness and response capabilities to future health threats, including pandemics. Recent studies have demonstrated the potential of digital health dashboards to transform public health policy by providing real-time data analytics for rapid decision-making [8].

The PANDEM-2 IT system provides an innovative solution to address these objectives by offering a standardized and interactive platform for pandemic preparedness training and response to support decision-making on pandemic management. It also offers adaptability for addressing other potential cross-border threats. The system, which is now developed to an advanced prototype level [9,10], enables pandemic managers and stakeholders to acquire the necessary skills, knowledge, and decision-making capabilities required to respond effectively to future pandemics. Furthermore, it is designed to serve as a prototype that can be further developed or adapted to cater to different training needs. The PANDEM-2 system was tested in a 2-day functional exercise to simulate a pandemic caused by a novel strain of avian influenza. The cross-border exercise was conducted between 2 public health emergency operating centers (National Institute of Public Health & Environment, Netherlands, and Robert Koch Institute, Germany) [11]. Several EU public health and first responder agencies, including the Directorate-General for Health and Food Safety's Early Warning Response System simulation platform, were involved in supporting roles. Qualitative and quantitative data were collected to evaluate the performance of the system and its potential as a cross-border preparedness and response tool for health emergencies. Results indicate that the PANDEM-2 IT system holds potential to become a robust training hub for pandemic management in the future, serving as a valuable resource for continuous learning and preparedness.

Objectives

This study aimed to describe the development and potential of the PANDEM-2 dashboard and to outline the methodologies used in obtaining valuable insights and feedback from end users to further refine its development. It also describes how privacy, data protection, and security risks were considered throughout the design of the system. The lessons learned from the development and implementation of the platform can pave the way for future research and development in pandemic management and response technologies and inform the development of similar IT solutions.

Methods

PANDEM-2 System Architecture

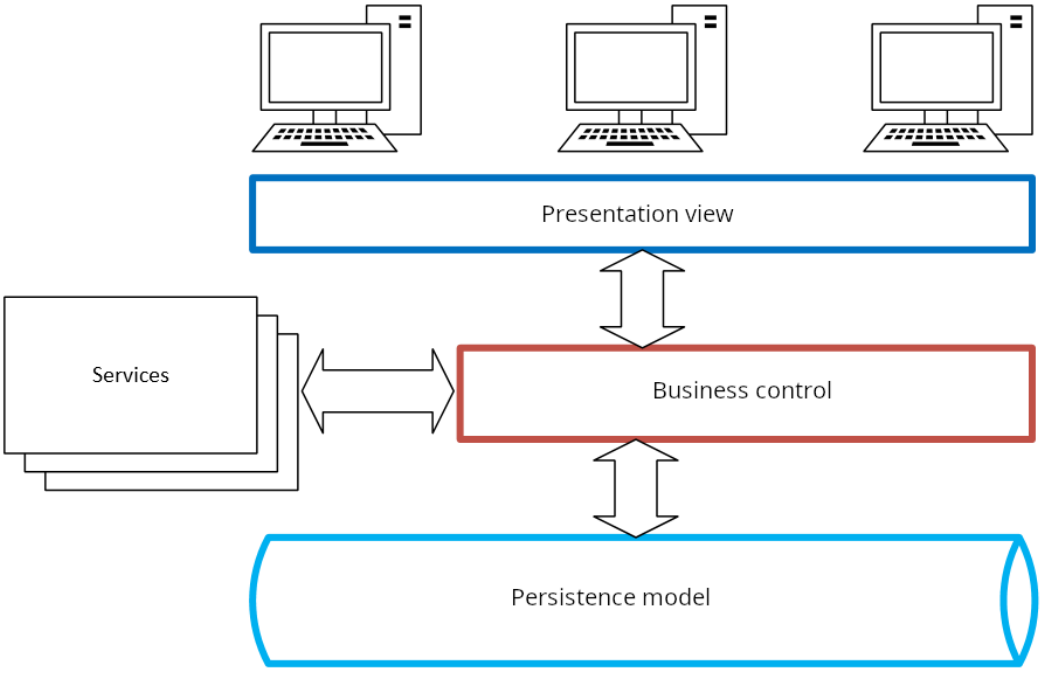
System Design and Components

The PANDEM-2 architecture (Figure 1) is an open-source framework comprising multiple modules and components that are accessed and integrated as web services. Its design and development were guided by the selection of widely supported libraries and adherence to industry-standard development practices and principles. The platform features a multilayered architecture that facilitates communication between its components. The persistence layer, built using MongoDB [12], ensures efficient data retrieval and storage. This layer is accessed through an exposed application programming interface (API) from the NodeJS business layer [13], which facilitates interaction between the MongoDB database and other modules

XSL•FO
RenderX

or external systems. NodeJS is known for its scalability and efficiency and ensures the smooth functioning of the business layer [14]. The presentation layer is developed using the Angular framework [15], which manages the user interface, providing an intuitive and interactive experience for users. Angular was chosen for its robustness, scalability, and extensive library support, which enhances the platform's user-friendliness and maintainability. Angular is a well-maintained framework, with releases every 6 months. Its documentation is comprehensive, containing deprecation, compatibility, and breaking change policies alongside update path guides, ensuring clear upgrades and maintenance. Data visualization within the application is handled by the Highcharts JavaScript library [16]. The use of these stable and well-supported technologies underpins an architectural design that ensures the platform is scalable, adaptable to various presentation contexts, and easy to maintain.

Figure 1. Multilayered architecture of the PANDEM-2 dashboard illustrating the interaction between the different layers, including the persistence layer (MongoDB), business layer (NodeJS), and presentation layer (Angular).



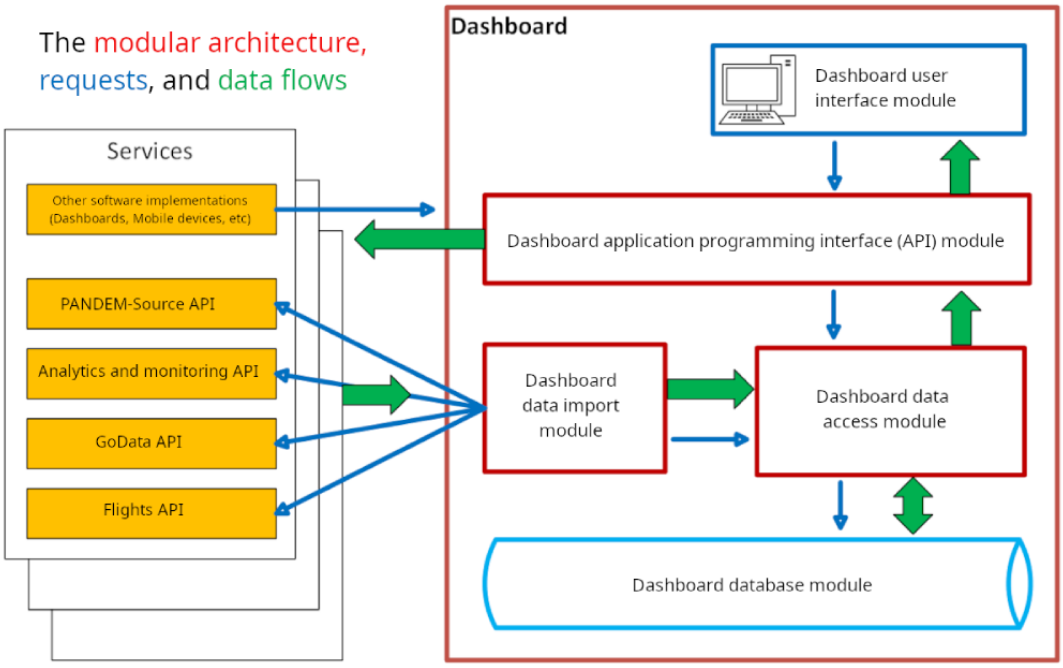
Data Collection and Integration

Data within PANDEM-2 are gathered from various sources via data collectors, which can either be human-operated or automated. The PANDEM-Source module [17], an extract, transform, and load (ETL) pipeline and database was specifically developed for PANDEM-2 to facilitate the continuous collection and integration of heterogeneous surveillance data into a coherent database. PANDEM-Source uses JSON (JavaScript Object Notation)-formatted source definition files to describe each source, including format, monitoring frequency, mapping to PANDEM-2 variables, and coding systems. It performs transcoding by downloading and applying mappings between coding systems, reporting any missing codes or integration errors to a data manager who can update mappings or monitor integration progress. A key challenge was standardizing geographic locations to the Nomenclature of Territorial Units for Statistics (NUTS), a hierarchical system developed by Eurostat that divides the territory of the EU into regions for statistical and reporting purposes [18]. For natural language processing used to detect locations in social media data, we used the GeoNames database, which contains local aliases for geographic locations [19]. To reduce methodological disparities between sources, we prioritized official European Centre for Disease Prevention and Control (ECDC) data where available and relied on existing data collectors like COVID-19 Data Hub [20] when ECDC data were not available. PANDEM-Source allows data managers to tag datasets, enabling the selection of relevant data combinations for export to the PANDEM-2 dashboard. This process, while facilitated by the tool, requires human oversight. To ensure the accuracy of the data acquisition process, manual validation was performed during both the data import and dashboard implementation stages. The declarative syntax used by the tool minimized the risk of errors in data processing. A more formal validation process is currently underway and described in greater detail in a related preprint [21].

The PANDEM-Source module functions as a service with an API providing relevant and preprocessed data to the persistence layer of the dashboard. Data are then processed, exposed through a specific API, and made accessible to the front-end of the dashboard, as well as other interested data publishers (Figure 2).

XSL•FO
RenderX

Figure 2. Dashboard modular architecture demonstrating the flow of data from external sources into the PANDEM-2 dashboard, showing how data is processed through various modules and application programming interfaces (APIs).



The business layer in the dashboard collects data from external modules, such as the World Health Organization GoData implementation [22] and also from the Flights module [23] and the Modelling and Analytics module developed internally as part of the project. It processes these data, performing validations, adaptations, and enrichment based on formulas, to prepare them for display in the presentation layer.

Security Aspects of the PANDEM-2 Dashboard

The PANDEM-2 dashboard infrastructure is secured using Amazon Web Services (AWS) Cloud Security [24]. The AWS environment is monitored using AWS CloudWatch and CloudTrail, ensuring General Data Protection Regulation (GDPR) compliance [25]. Several other AWS services were used to enhance privacy and control network access, including Amazon GuardDuty for continuous threat monitoring [26] and AWS Secret Manager and Key Management Services for data protection [27]. Data are encrypted in transit with Transport Layer Security across the AWS services, and secure sockets layer and transport layer security certificates, are managed with AWS Certificate Manager. A summary of the PANDEM-2 security aspects can be found in Multimedia Appendix 1. The project only hosted publicly available aggregated or simulated data, which addressed data protection concerns for the duration of the project.

Dashboard Design and Development

Exploratory Research and Design Rationale: Why a Dashboard Approach?

A dashboard is a visual display of key analytical indicators arranged on a single screen, enabling real-time monitoring and quick assessment [28]. The goal is to provide stakeholders with relevant analyses to anticipate issues and make decisions in domains with complex multivariate data. For example, a study by Franklin et al [29] modeled the real-time status of a hospital's emergency department as a dashboard to support rapid clinical decision-making and improve throughput. However, the growing use of dashboards has raised concerns about the cognitive effort required to extract actionable information, particularly for users with limited analytic skills [30]. Data storytelling has been proposed as a more intuitive way to communicate information that motivates decision-making [31]. Stakeholders, however, determined that maintaining and updating multiple specific decision narratives would be overly complex. Therefore, we chose a dashboard approach for its flexibility and adaptability across various public health scenarios, including surveillance and training. In addition, a key design requirement was the ability to customize the dashboard view and generate reports, which was seen as a more scalable and sustainable solution compared to data storytelling, given the diverse decision-making needs of different stakeholders. The primary users of the PANDEM-2 dashboard are public health officers and medics typically trained in data analysis.

Review of Existing Dashboards

Responses to the ongoing pandemic provided opportunities to review public health indicators, visualization approaches used, and the design decisions made by different organizations. To gain insights into these aspects, we conducted a review of publicly available dashboards in April and May 2021, operating under the assumption that these reflected the information needs of public health agencies and could be analyzed without an imposition on end users (Multimedia Appendix 2). This review included dashboards recommended by collaborating partners based on their initial user requirements, as well as publicly accessible COVID-19 projects under development (Multimedia Appendix 3). The review assessed the available dashboards, focusing on their aims, target users, and content. We also collated data sources for each dashboard, along with their breakdowns by population, time, location, granularity, and visualization approaches. This analysis provided insights into

XSL•FO
RenderX

how organizations prioritized information and the data views expected by experts. It also highlighted different data interaction models and how data artifacts were aggregated, filtered, modified, and recorded. While the data sources identified during this review were considered for inclusion in the PANDEM-2 database, consortium members' own data eventually proved to be more suitable, especially as larger organizations and governments began openly sharing their information.

Data Visualization Catalog

Capturing and storing visualizations for viewing or referencing during the design process was challenging due to their volume and complexity. Initially, screenshots of each visualization, along with related data, were added to a document (Multimedia Appendix 4). However, as the document grew, this approach became impractical and difficult to reference. To overcome this challenge, a data visualization catalog was created. The catalog was specifically developed for annotating and uploading data visualizations during the design phase of the project. The visualizations became searchable using keywords related to graph types, data types, or indicator types. The catalog served as a directory of existing public health visualization approaches, which the project team could consult during the participatory design process. It is important to note that the Vizualisation Catalogue was used solely as a development tool during the design phase and is not part of the final application.

User Requirements and Participatory Design

Gathering of User Requirements

The design and development of the PANDEM-2 dashboard began with collecting user requirements from members of the project consortium, outlining the necessary functionalities and features to meet users' needs. Participants represented a wide range of domains, including universities, national public health institutes, the Italian Red Cross, the Austrian Red Cross, the National Institute of Medical Emergency (Portugal), a leading institute for trend and technological analysis, and MedTech health organizations. Each consortium partner appointed a contact person to gather internal requirements and communicate them to the project's technical team.

These designated representatives participated in initial user forums that were held during the project's inception, aimed at gathering user requirements from domain experts and capturing their needs and preferences for features in pandemic management software (Multimedia Appendix 5). The group was diverse in age, experience, and gender, with participants ranging from recent postdoctorates to professionals with >30 years of experience, and included software designers, data scientists, epidemiologists, microbiologists, physicians, first responders, and project managers. They also represented a wide geographic spread, including Romania, Italy, Portugal, France, Germany, the Netherlands, Belgium, Sweden, Finland, and Ireland. This diversity ensured a wide range of expertise and perspectives, thus reducing potential design biases.

The requirements gathering process involved 2 workshops held between February and April 2021. During these workshops, users described their ideal dashboard for pandemic management and identified essential visual features for the dashboard. These requirements were then compiled into a comprehensive list and organized into functional units. The participatory design surveys were conducted as part of this process, providing interactive design prototypes and collecting feedback from users on key functionalities (Multimedia Appendix 6).

Generation and Analysis of User Stories

User requirements were assigned to different groups within the technical team, who analyzed and converted them into user stories. A user story describes a feature from the end-user's perspective, detailing how it provides value [32] (eg, "As a policy advisor or health care professional, I want to see cases stratified by age, sex, time, and severity to monitor clusters and track the virus spread over time").

By exploring user requirements and generating user stories, the technical team identified common needs, consolidated similar requests, and clarified key functionalities for the PANDEM-2 system. Each user story generated focused on a primary indicator query and was linked to the initial requirements for traceability. It defined query resources for user interactions, enabling data filtering, aggregation, new data creation, and record management. In addition, each user story considered the time-related and geographic aspects of the indicator within the application. Multimedia Appendix 7 provides an example of a user story.

Prioritization and Refinement of Key Variables and Indicators

To refine and prioritize user requirements for the dashboard, a prioritization process was undertaken. This involved a data availability and prioritization survey, which was conducted with the consortium members using the initial requirements generated in the Gathering of User Requirements section. A matrix was created, mapping all necessary variables and indicators required to fulfill the identified needs. This was followed by an assessment of the priorities and availability of these variables in open and restricted data sources. Users were invited to provide insights into their institution's data availability, including details about the level of information, reporting periodicity, data format, and the geographic level at which the data were accessible (municipal, regional, or national). The questionnaire covered various aspects, such as level of priority, data availability within their institution, the level of detail of publicly available data, adherence to the European Surveillance System (TESSy) format, and the pathogens for which data were collected (Multimedia Appendix 8).

The list of requirements was further divided into 14 variable categories, and each item on the list was classified as either an observation, a characteristic, an indicator, a resource, a document, or a referential. These terminologies are defined in Textbox 1.

XSL•FO
RenderX

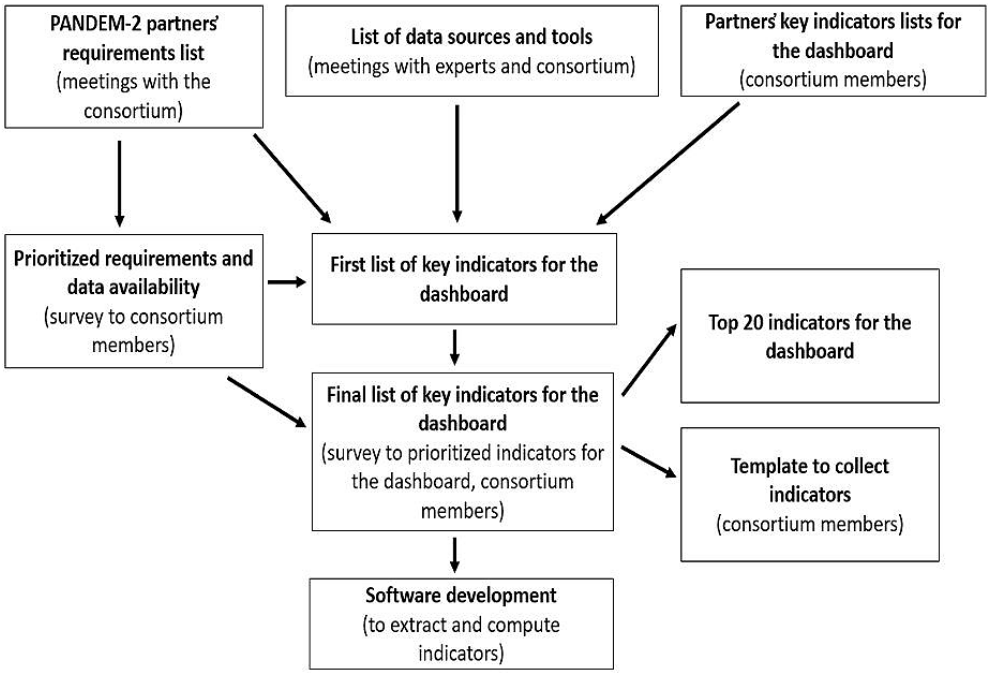
Textbox 1. Definitions of variable categories used in the PANDEM-2 dashboard.

- Observation: Any variable or item that reflects a measurement or observation reported by a data source. These variables are typically numeric and represent concepts like the number of people or resources, or could include nonnumeric items, such as statements (eg, tweets) published by individuals.
- Indicator: Specific observations used to monitor particular questions. They can be categorized as either calculated values with associated methodologies and formulas or directly reported by users.
- Characteristic: Any variable or item which describes or categorizes an observation. It is always linked to variables of the observation type, such as the country, date, or variant of the observed cases.
- Referential: A specific type of characteristic distinguished by having a unique code that can be mapped to other characteristics. For example, municipalities serve as a referential, each possessing a unique code and linking to a country.
- Resources: A specific type of referential that concerns resources needed to handle pandemics. They play an important role in PANDEM-2 for resource modeling features (eg, nurses per bed or material resources). They are associated with observations like the number of available resources.
- Document: Refers to any generic term for a record or written material.

The methods used for defining key indicators and data sources, as outlined earlier, culminated in a final variables list that encompasses both open data sources and data contributed by partners (Multimedia Appendix 9). Alongside this list, data formats and an ETL process were specified to facilitate seamless integration with PANDEM-Source. To support the ETL process, a collaborative effort led to the development of a Microsoft Excel template to accommodate specific data needs for most consortium members. Furthermore, relevant heterogeneous data sources for effective pandemic management, such as Next Generation Sequencing–Laboratory data, social media data, and participatory surveillance data, were identified and corresponding ETL processes were devised for inclusion in the PANDEM-2 database.

After receiving the initial list of requirements and the prioritization feedback, the technical team began further refinement of the requirements. The process was guided by the concept of a minimal viable product (MVP), which focuses on prioritizing essential features to maximize learning with minimal effort [33]. Although the PANDEM-2 system is a prototype, the MVP approach helped in determining where efforts should be concentrated and which requirements should be given priority during the implementation process. Each group in the technical team looked at the requirements list and the prioritization process and created an MVP list. Subsequently, with participation from all Public Health Agencies and first responder partners in the biweekly web-based workshops, the initial requirements, the prioritization list, and the MVPs were combined and deliberated upon. The process of gathering and refining user requirements is illustrated in Figure 3. This process underpinned the development of the PANDEM-Source software used to extract and compute indicators, which were visualized on the PANDEM-2 dashboard. At the end of this prioritization process, a list of key indicators was generated.

Figure 3. Process used to gather partners' requirements list and prioritize partners' key indicator lists for developing the dashboard's core features.



XSL•FO
RenderX

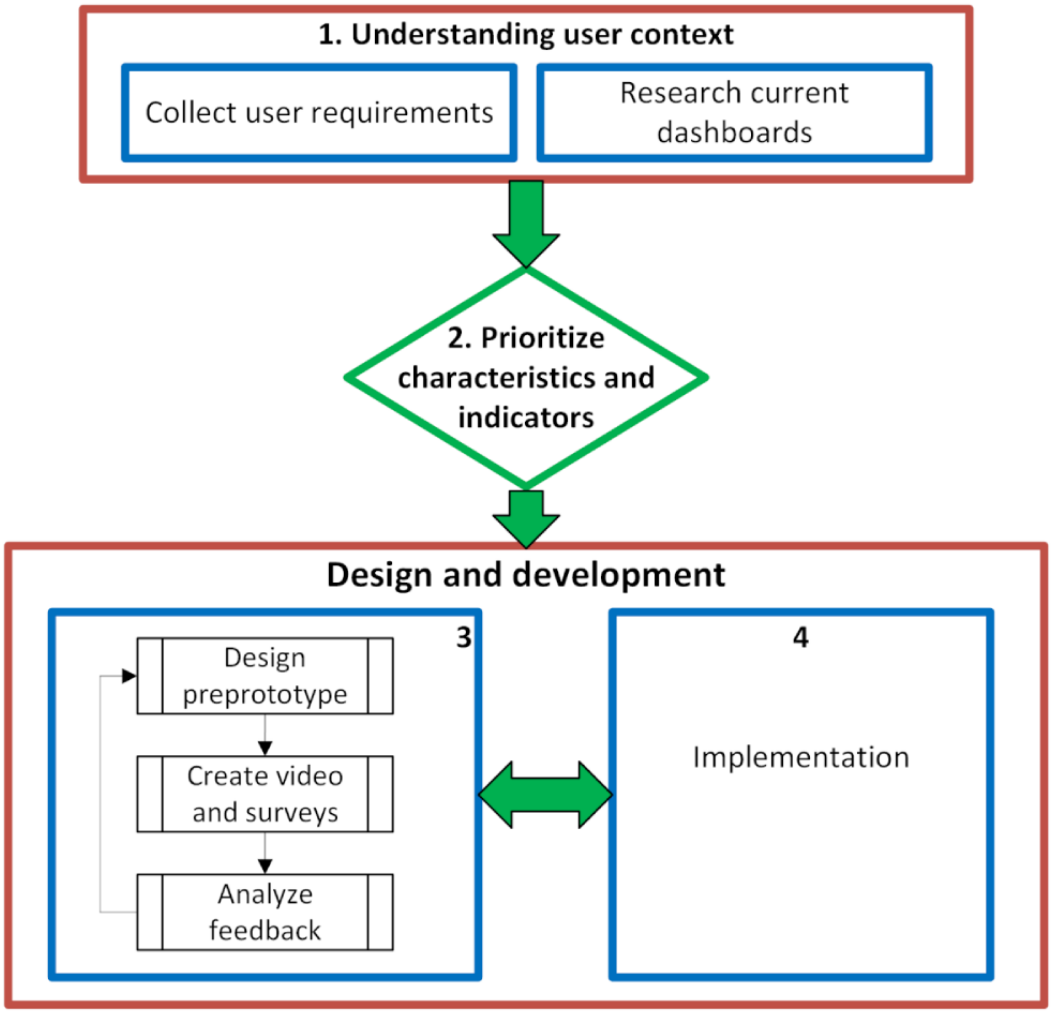
Prototype Design, Participatory Feedback, and Iterative Improvements

The design of the PANDEM-2 dashboard followed an iterative process involving a participatory design approach, with user feedback incorporated throughout the design stages. Participatory design represents a design approach in which the people destined to use the system being designed play a critical role in its design [34]. Generally, participatory design is conducted in face-to-face settings where users are engaged in recording requirements and teasing out functionalities that are necessary to implement them. However, due to the COVID-19 pandemic, these physical meetings were not possible to organize. The participatory design process was therefore adapted to the pandemic context and involved several asynchronous methods, as illustrated in Figure 4. The adapted process allowed relevant stakeholders to participate and enabled collaboration across multiple teams. The analytical components of the dashboard were then designed based on the outcomes of this collaboration and are presented as follows:

1. Understanding the context of use by gathering insights from various stakeholders to understand end-user requirements and identify COVID-19 public dashboards for analysis
2. Prioritization of dashboard characteristics and indicators
3. Asynchronous participatory design (APD) process that involved weekly asynchronous interactions with users. Interactive design prototypes along with video documentation and feedback forms were sent to consortium members, who explored the designs and provided feedback. The primary objective of the APD process was to enable users to interact with a preprototype platform, explore potential interface designs, and provide feedback. This process involved targeted questionnaires, which aimed to explore what graphical representations and interactions were beneficial to users for their respective roles. Participants were presented with a range of potential functionalities linked to priority indicators, which included visual analytics components and interaction models, such as the ability to filter and switch between different data views. Participants shared their feedback on the structure of the pages, the data presented, the data visualizations that were used on each page, and the interactions with those data visualizations. Through this process, users provided feedback on elements to be added, removed, or modified within different components in the system. This feedback was then integrated into subsequent designs.
4. In parallel to the above step, the design prototypes were sent to the software team, who carried out the implementation of designs and improvements toward the final prototype.

XSL•FO
RenderX

Figure 4. Development workflow, including asynchronous participatory design, detailing the iterative feedback process used to incorporate user input into the PANDEM-2 dashboard design. AVG: average.



Impact Assessment

The implementation of ethical, privacy, data protection, and social principles was an integral part of the PANDEM-2 project. The process was led by Trilateral Research, a research and innovation organization specializing in applied ethics, privacy, and data protection in technology development. The project team actively engaged with end users both within and outside the project consortium to identify and mitigate potential risks arising from the tool's development. Following relevant guidance, including the Information Commissioner's Office Privacy Impact Assessment Code of Practice [35] and International Organization for Standardization and International Electrotechnical Commission standard for privacy impact assessments [36], the process mapped information flows, assessed potential risks within the system, and provided recommendations to mitigate the risks. A diverse group of stakeholders were involved in the process, participating in 2 workshops conducted in November 2021 and May 2023. There were 20 participants in each of the workshops, representing a wide geographic spread, including Romania, Italy, Portugal, France, Germany, the Netherlands, Belgium, Sweden, Finland, and Ireland. Those that partook in the workshops came from universities (UCLouvain and University of Galway), national public health institutes (Institute of Public Health in Romania, Public Health Agency of Sweden, the National Institute for Public Health and the Environment in the Netherlands, the Robert Koch Institute, the Portuguese National Emergency Medicine Institute, and the Finnish Institute for Health and Welfare), as well as other organizations such as the Italian and Austrian Red Cross, the Fraunhofer Institute for Technological Trend Analysis, and software companies specializing in health, Epiconcept, and development, Clarisoft. Roles of those taking part included software design and development, data science, epidemiology, microbiology, infectious disease specialists, doctors, researchers, first responders, pandemic managers, and project managers. Notably, participants included a former adviser to the World Health Organization on health security and a former head of the Chemical, Biological, Radiological and Nuclear threat unit for Belgian defense. The initial workshop aimed to raise awareness of ethical, privacy, data protection, and social principles requiring consideration during the development and deployment of the PANDEM-2 system and tools. It also focused on mapping data flows within the PANDEM-2 system and identifying related risks or opportunities. In addition, the workshop explored privacy, social, and ethical issues and opportunities to be considered in different pandemic scenarios, both in the preparedness and response

XSL•FO
RenderX

phases. Possible mitigations for the identified risks were also discussed. The second workshop involved an open and collaborative discussion, allowing stakeholders to contribute insights on mitigation measures, necessary actions, and final recommendations for both project stakeholders and end users.

Ethical Considerations

The PANDEM-2 project adhered to GDPR regulations and Horizon 2020 guidelines. Ethical approvals were obtained from the relevant institutional review boards of all consortium partners involved in data collection and processing. These approvals ensured that the project complied with the highest ethical standards, protecting participants' rights and ensuring the responsible handling of personal data. Informed consent was obtained from all participants before their involvement in the stakeholder engagement activities. The consent procedures ensured that participants clearly understood the nature of the study, the use of their data, and their rights, including the ability to withdraw at any time. Participants were also informed that their data would be used for research purposes only. All personal data were anonymized or pseudonymized to protect participants' identities. Data were securely stored on password-protected servers, accessible only to authorized personnel. Participant identifiers were removed or replaced with pseudonyms to ensure anonymity in the research outputs, preventing any risk of reidentification. No financial compensation was provided to participants. Participation was entirely voluntary, and participants had the right to withdraw from the study at any time without any negative consequences. No identifiable images of participants are included in the study materials or outputs.

Results

System Architecture and Data Integration

Implementation and Configuration

The dashboard was implemented and hosted on the Amazon Cloud as a central data management system at the EU level, accessible to all national health institutes across Europe. To address challenges with semitransparent data access, whereby some but not all data from different institutions can be shared with others, the architecture is not only open access but also adaptable. An adaptable architecture allows for both isolation of data between groups of countries and on-demand installation or deployment in a single health institute.

The dashboard offers flexibility in its configuration, allowing customization in 2 main areas. First, user experience and user-interface attributes, which encompass elements related to user experience and user interface. This involves tailoring aspects such as the language used for interface elements, the design of graphical features, and the overall visual style. Users can personalize the appearance of the dashboard to align with their preferences and needs by adjusting these attributes. Second, internal functional parameters that deal with the technical aspects of the dashboard's operation can be customized. This includes details on how different modules within the system access APIs and specific technical information relevant to the system's functioning. The administration module of PANDEM-2 serves as the control center for managing these parameters, granting users the ability to fine-tune the dashboard's functions and adapt it to their specific requirements.

Data Sources and Key Indicators

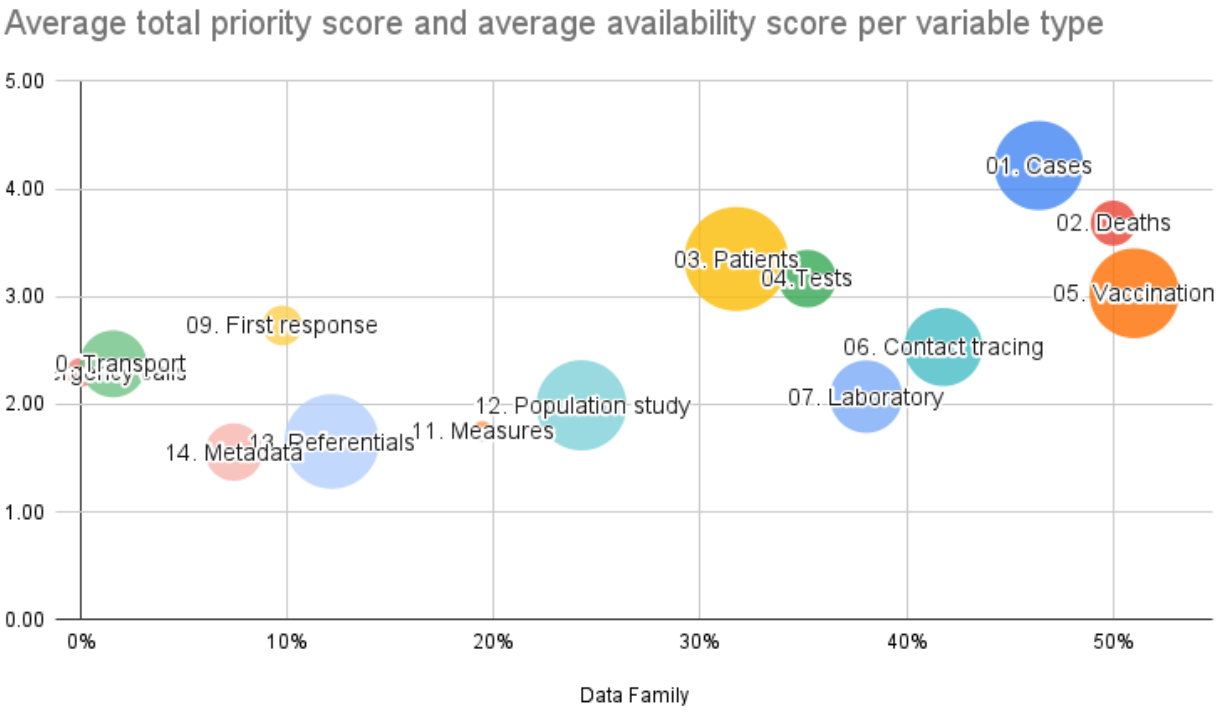
Responses to the data availability and priority survey are shown in Figures 5 and 6, and tables 1 and 2. The 9 participant institutions were questioned about 103 variables classified by groups. For each variable, the respondent answered a series of questions about availability, priority, and format of the associated information. On the basis of these answers, data availability scores were calculated and color-coded, ranging from low (red) to high (green; Figure 5). The variables identified as most relevant were cases, vaccination uptake, deaths, hospital resources, and hospital admissions. Figure 6 shows that the most important variable categories were also the categories with the most available data. Higher priority categories are toward the top of the graph, while high availability categories are toward the right. This observation highlights the importance of prioritizing and ensuring accessibility to critical data categories to support effective pandemic preparedness and response efforts.

XSL•FO
RenderX

Figure 5. Survey results highlighting the availability and prioritization of key data variables for the dashboard (AVG; X). AVG: average; ICU: intensive care unit; N°: number of; PPE: personal protective equipment; Rt:; X: specific pathogen of the pandemic.

Variable type	Variables	# AVG Institution availability score	# AVG Publicly availability score	# AVG of total priority
01. Observation	Confirmed cases	2.3	1.7	6.9
	N° of infected patients per bed type (clinic care, ICU, ventilator)	1.8	1.1	5.0
	People fully vaccinated	1.8	1.1	4.8
	Vaccination status	1.8	1.0	4.0
	People that have received at least 1 dose	1.6	0.9	4.0
	N° of deaths by X	1.7	0.3	4.0
	Suspected cases (possible, probable, and unclassified)	1.6	0.9	3.9
	N° of deaths	2.6	1.3	3.9
02. Characteristic	Cases per (age, sex, comorbidities)	2.2	1.0	4.7
03. Indicator	Incidence rates (last week, two weeks, month)	1.8	0.9	5.4
	Incidence rates (age, sex, comorbidities, variant)	1.1	0.4	5.0
	Rt number	0.6	0.6	4.8
	Positivity rate	1.4	1.4	4.7
	Vaccination progress (proportion of vaccinated, overall, by age and risk group)	1.8	1.2	4.3
04. Resources	Emergency staff	0.3	0.0	4.6
	Vaccination resources (staff, centres, supplies)	1.0	0.8	4.4
	Hospital staff type, eg, intensive care, emergency care, etc	0.8	0.4	4.2
	PPE (stock, type, and need)	0.7	0.6	4.0
	Bed type (ICU, regular clinical wards, emergency)	0.4	0.7	4.0
	Beds/Room occupancy and types, eg, isolation	0.8	0.4	3.8
Grand Total		**1.4**	**0.8**	**4.5**

Figure 6. Comparison of priority versus availability of variable categories, showing the alignment between prioritized variables and their availability across institutions.



Most of the data were reported on a weekly or daily basis, indicating a commitment to timely and up-to-date information (Table 1). A significant proportion of the data were collected and recorded by various institutions and from data sources related to COVID-19. There was limited recording of pandemic-related data before the COVID-19 pandemic. There were also variations in data granularity, reporting frequency, geographic coverage, or other specific details related to the COVID-19 information being collected. On average, only 9.7% (10/103) variables were available in the TESSy [37] format and 26.2% (27/103) only in a different format, a format used by the TESSy (Multimedia Appendix 10).

XSL•FO
RenderX

Table 1. Data reporting frequency and types of pandemic-related data available, highlighting the frequency of reporting (daily, weekly, or monthly) and the prevalence of COVID-19 data from various sources (n=115).

Data category	Reporting variables or items, n (%)
Periodicity	
Weekly	17 (14.7)
Daily	14 (12.1)
Others[a] (specify in comments)	8 (6.9)
Monthly	2 (1.7)
Pathogens	
COVID-19	17 (14.7)
Others (specify in comments)	13 (11.3)
Influenza and COVID-19	8 (6.9)
Influenza	1 (0.8)

[a]The "others" option refers to any reporting frequencies that were neither daily, weekly, nor monthly, which participants were asked to provide qualitatively through comments.

The survey also identified several data gaps, as shown in Table 2. Among these, 2 areas of particular significance were contact tracing and deaths. It was evident that these areas lacked sufficient information despite their high priority in the survey. In addition, there was also a noticeable lack of information concerning resource availability, despite its recognized importance in pandemic management.

Table 2. Data gaps identified during data availability survey, showing critical pandemic-related data, such as contact tracing and deaths that were insufficiently available despite high priority.

Variable category	Average public availability gap
Deaths	1.1
Contact tracing	0.9
Cases	0.8
Laboratory	0.6
Patients	0.5
Vaccination	0.5
Tests	0.3
Referential	0.2
Population study	0.1
Metadata	0.1
First response	0.1
Transport	0.0
Emergency calls	0.0
Measures	0.0

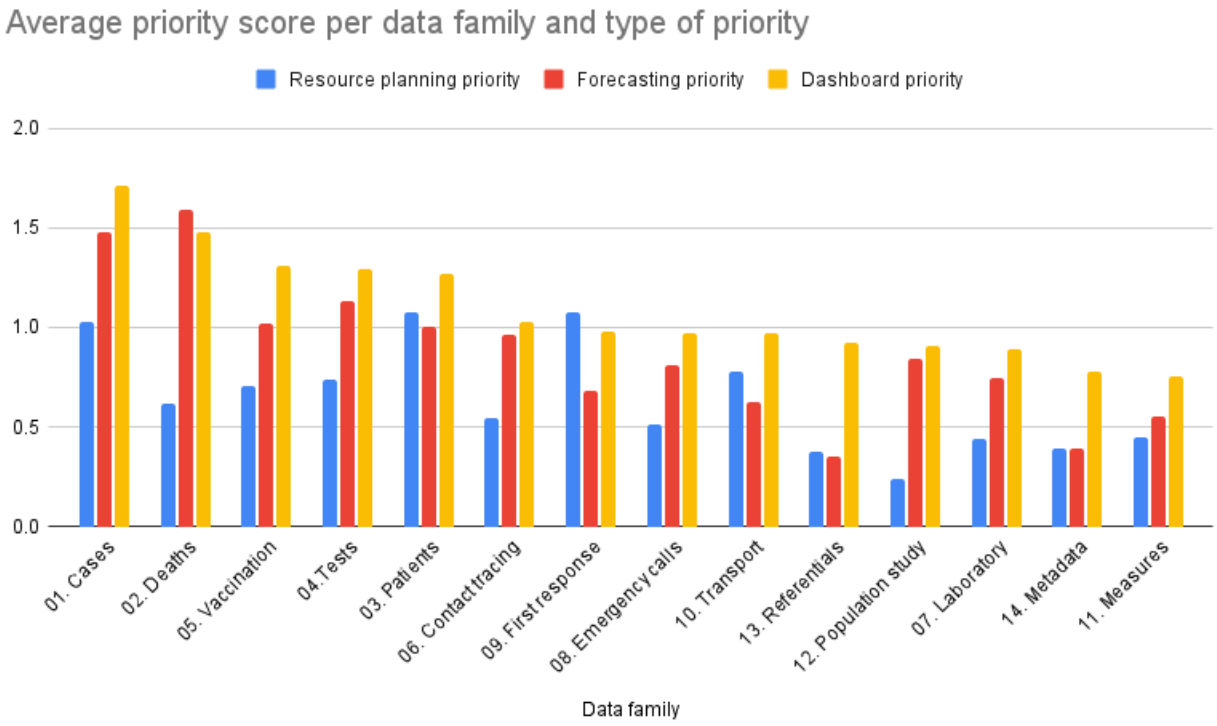
The public availability gap is a metric quantifying the specific types of data accessible at a detailed level but not disclosed publicly. Each participant provided their estimation of variable availability within their institution and at the public level, to the best of their knowledge. The assigned scale attributes 4 as the highest value to individual level, 3 to municipality, 2 to regional, and 1 to country level. In instances where data were unavailable or unknown, a score of 0 was assigned. The data availability gap is determined by calculating the difference between the average availability score at the institution level and at the public level.

Figure 7 illustrates the findings regarding the variables that users prioritized for resource planning, forecasting, and the dashboard. To calculate the priority score for these purposes, participants were asked to rate the importance of each variable from 0 to 3 for the 3 purposes, with 0 being the lowest and 3 being the highest. The results were averaged by variable category. In most cases, dashboard variables held a higher level of importance compared to forecasting or resource planning variables. The spikes observed in the resource planning and forecasting areas highlight the data points of utmost priority for those domains, with particular significance placed on the ability

XSL•FO
RenderX

to forecast deaths and to perform resource planning for patient care at the hospital and first response level.

Figure 7. Prioritization of variable categories for resource planning, forecasting, and the dashboard, showing importance ratings assigned to various data categories by participants.



At the end of the data and indicator prioritization process, a template to collect variables and indicators from nonopen sources (or partner-restricted data) was created.

Dashboard Design and User Engagement

Insights and Outcomes From Dashboard Review

The review of existing dashboards provided several insights that influenced the design of the PANDEM-2 dashboard. The data visualizations within the reviewed dashboards highlighted the prioritization of information by various organizations and identified the data views expected by experts. This review also revealed different data interaction models useful in public health domains and how data artifacts were aggregated, filtered, changed, and recorded. Insights gained from this review helped shape the decision to use consortium members' data as the primary data source, as these proved to be more comprehensive and reliable than many external sources. The review also informed the design features and visualization approaches adopted in the PANDEM-2 dashboard.

Prototype Design and Features

Multimedia Appendix 11 visually illustrates the design of a PANDEM-2 dashboard page. The key components on each page of the dashboard include the following:

1. Overview explanation component. This component includes a page title and provides an explanation of the indicator and the data presented on the page. It serves as a brief summary to provide context and understanding.
2. Indicator cards. These cards display high-priority data related to the indicator. They offer a quick snapshot of the current state of the pandemic, allowing users to grasp important information at a glance. The cards also feature an information button that provides additional details about the data being shown. The highlighted figures show a comparison between the most recent 7 days and the preceding 7 days. The highlighting provides an instant visual understanding of how important indicator variables such as admissions and intensive care unit admissions are changing over time.
3. Map. The map component offers a geographical perspective on the pandemic situation. Users can choose the specific information they wish to view on the map, such as cases, vaccination rates, or other relevant metrics. They can also segment the map by NUTS regions, providing a more localized view.
4. Graphs. The dashboard includes multiple graphs that present detailed information about the indicator. These graphs allow users to explore and filter the data according to their preferences. This feature enables a deeper analysis of the pandemic trends and patterns.
5. User-interface functionality. Users can access the PANDEM-2 system using the interface in the presentation layer. This interface offers multiple possibilities to display and filter information, build reports based on displayed graphical elements and comparisons between different data, and export these reports in many predefined formats, including images, tables, and web service-specific formats.

XSL•FO
RenderX

The final design framework of the PANDEM-2 dashboard was structured to provide an intuitive and interactive experience for users. Each indicator view was conceptualized as a dedicated page, comprising various relevant components. Abstract components were developed that could be adapted to each indicator page, ensuring a unified design and interaction experience across pages. This approach allowed users to intuitively understand the interactions and functionality across different pages once they had become familiar with the system on one page.

Together, these components enhance the accessibility and usability of the PANDEM-2 dashboard, empowering users to gain insights into the pandemic data through both high-level summaries and detailed visualizations. More illustrations of initial and final dashboard designs are shown in Multimedia Appendix 12.

Participatory Design Feedback

The participatory design process led to several modifications and improvements in the PANDEM-2 dashboard prototype. The number of responses to each of the APD feedback forms ranged from 4 to 9.

Feedback from the APD process highlighted the importance of data visualization in diverse formats, combining different data types, and analyzing data across various time frames. Users expressed their interest in generating custom reports and exploring data alongside government interventions (countermeasures):

> *The addition of countermeasures on the emotion/sentiment timelines is good. However, the countermeasures would be easier to see or understand, if they were all on the top part of the graph with the name of each intervention placed next to the line [on the left or right side of the graph]. Now it's difficult to check back to the figure legend below and tell one shade of grey from another. Now there's also a false impression that a countermeasure is somehow linked to the emotions on the same level of the graph* [for example, social distancing with anticipation and disgust]

Because of the feedback received, the final prototype incorporated various features that catered to the specific needs of the stakeholders, such as customizable dashboard views and report generation capabilities. The design also addressed the cognitive effort required to extract actionable information, ensuring that the dashboard remained user-friendly and accessible to all users, regardless of their analytical skills. This feedback proved invaluable for the technical team, offering insights into the preferred indicators and formats for their presentation. In addition, the feedback offered the technical team an understanding of how end users intended to use the data beyond the dashboard, including saving reports, downloading visualizations, and obtaining underlying data. When asked about saving outputs from the dashboard, participants expressed a desire for flexibility, with various preferences emerging. Users wished to save the current state of the page to share with others, download images or PDFs of the page, obtain the data within the page, or capture images of specific sections of the page.

Regarding epidemiological and resource modeling, participants expressed the desire to save the model used, the input parameters for scenario runs, visual outputs, data outputs in alternative formats, and high-level indicators.

Participants' feedback also shed light on data reporting challenges, particularly in terms of disparities in reporting levels and time intervals and the need for standardization across organizations. For instance, a participant highlighted how delays in testing workflows in Finland resulted in challenges when reporting cases and tests:

> *Due to unavoidable delays in testing workflow [sampling, sample shipping and processing, analysis, registration of results] it may not be realistic to report tests and cases 'this week.' In reality [in Finland], numbers of cases and tests from this week cannot be reported before Wednesday [COVID-19] or Friday [other viruses in infectious diseases register] next week. Sentinel surveillance may have even longer delays.*

The APD process further brought to light the complexity of data in public health situations, emphasizing the significance of considering interdependencies. It highlights that the system developed serves as a decision support tool and that the role of public health officials remains crucial in making informed decisions based on the presented information. One participant's feedback from Germany highlighted a specific example of complexity in data representation related to counting beds versus assessing capacity:

> *In Germany we do not really count the beds but the capacity. This is because to be able to add a new patient you need a bed, a room, ventilation, staff etc, or do you mean capacities with the term "bed"?*

This inquiry highlights the importance of clarity and precision in defining data elements to ensure accurate interpretation and decision-making within the tool. Finally, through their engagement with an expanding collection of interactive designs, users were able to identify gaps in the available indicators and provide specific requests for additional information, such as excess mortality data. This iterative process was instrumental in ensuring continuous and active user involvement throughout the project, enabling their valuable contributions to the design and development of the PANDEM-2 system.

Impact Assessment Findings and Mitigation Strategies

The impact assessment process identified 10 potential risks, all of which were deemed to have a low to medium likelihood of occurring. In total, 4 (40%) of the risks were ethical, 2 (20%) regarded privacy and data protection, and 4 (40%) related to social factors. These are summarized in Textbox 2 and reported in more detail elsewhere [38].

XSL•FO
RenderX

Textbox 2. Summary of the identified risks categorized into ethical, privacy and data protection, and social factors.

Ethical risks

- Accessibility and inclusivity: risk of excluding individuals with certain needs.
- User autonomy: risk of limiting end-users' control over their use of the training platform.
- Responsible use: risk of incorrect or unethical use of the training platform.
- Transparency and accountability: risks associated with the ability to account for actions taken or decisions made during the development of the training platform.

Privacy and data protection risks

- Security and access controls: risks relating to the protection of digital information from unwanted actions or unauthorized users, based on commonly accepted security frameworks, such as the confidentiality, integrity, and availability model.
- Data reidentification: risk associated with the correlation or linking of anonymous datasets with other data sources—especially those that have been collected by the state

Social risks

- Unintentional bias: risks associated with biased algorithms, biased data sources, and bias in training content that may lead to systematic errors that create prejudiced or unintended outcomes.
- Data quality and accuracy: risks associated with the ability to make informed decisions based on accurate information.
- Dependency on technology: risk of heavy reliance on the IT platform for critical operations, processes, or decision-making.
- Collaborative development: risk of inadequate collaboration with end users or public health authorities in the development and postproject stages to understand training needs and enhance the platform's capabilities.

The risks were effectively mitigated through several measures throughout the project. Design considerations extended beyond functionality to inclusivity and accessibility. The user interface was designed to accommodate multilingual users, providing options in both English and Romanian with provisions for translation into additional languages. Accessibility for individuals who are visually impaired was ensured through the selection of a color-blind–friendly palette. To facilitate ease of use, user-friendly training materials and guides were developed, offering clear instructions for system navigation. Supplementary features such as tooltips and contextual help were integrated, further enhancing the user experience. In addition, transparency was prioritized through the provision of accessible data sources and modeling assumptions used in the development process. Security was a paramount concern, prompting the adoption of Amazon Cloud, a reputable cloud computing service known for its robust security measures. The commitment to open collaboration and adaptability was evident in the decision to offer all components of the dashboard as open-source software. This approach encourages wider use, fosters adaptability, and invites collaboration across diverse contexts. Furthermore, a comprehensive developer guide in English was provided, serving as a valuable resource for future developers interested in building upon the PANDEM-2 platform or creating similar tools. End-users' feedback was continuously obtained and integrated to ensure the platform's suitability to diverse user needs.

To enhance the accuracy and relevance of training content, public health authorities' expertise was used in content development, and active end users engaged in patient surveillance and treatment assessed the training materials. This collaborative approach guarantees that the training content accurately reflects current information, best practices, and public health guidelines. Training exercises were based on scripts generated by public health authorities, adding objectivity and minimizing subjective biases or arbitrary preferences. These actions collectively enhance the accuracy and scientific validity of the training content.

Discussion

Principal Findings

The PANDEM-2 project aimed to develop an IT prototype system that offers a standardized and interactive platform for pandemic preparedness training and response to support decision-making on pandemic management. This paper outlined the development and potential of the PANDEM-2 dashboard, including methodologies used to gather users' insights to enhance its design and steps taken to identify and address potential risks associated with the technology.

The COVID-19 pandemic brought about an unprecedented surge in pandemic-related data at national, EU, and global levels, presenting a valuable opportunity to identify crucial data needs for effective pandemic response. PANDEM-2 leveraged user surveys to identify essential data requirements, not only for the ongoing pandemic but also for potential future health threats. Despite this increase in data, significant challenges emerged due to data availability, diverse data sources, nonstandardized formats, varying granularity, and data sharing restrictions among EU member states. These obstacles hindered timely data exchange for a coordinated EU-wide response and hampered cross-border response efforts. Addressing these data challenges became a central concern throughout the PANDEM-2 project, and consequently it emerged that it is imperative to proactively anticipate and address data requirements for potential future pandemics.

XSL•FO
RenderX

The EU initiated measures to address data sharing gaps through Regulation 2022/2731 on serious cross-border threats to health [39]. The Commission's responsibility includes defining mechanisms for data exchange, adhering to personal data protection rules and information exchange security. In addition, Article 5 of Regulation 2022/2730 on serious cross-border threats to health specifically refers to the development of the Union health crisis and pandemic plan "to promote effective and coordinated response to cross-border health threats at Union level" [40]. Specifically, article 5, point 5, states, "In order to ensure the operation of the Union preparedness and response plan, the Commission shall conduct stress tests, exercises and in-action and after-action reviews with Member States and update the plan as necessary."

The development process of PANDEM-2 aligns with these measures, both in terms of data sharing and exercises. Lessons learned in the development of the platform offer insights into data analysis, sharing, and reporting methods, effectively supporting public health actions and control measures. The PANDEM-2 platform's design, built on an open architecture and open data sources, ensures security, stability, extensibility, and flexibility, allowing it to adapt to evolving data needs and data sharing practices.

The PANDEM-2 platform is also a training platform, designed to facilitate both national and cross-border simulation exercises. A simulation exercise is a form of training or evaluation of capabilities involving the simulation of an emergency [41]. Operationally, there is the potential for the PANDEM-2 dashboard to be deployed at regional, subnational, national, and EU levels to conduct simulation exercises aimed at stress-testing pandemic response capacity. A recent large-scale functional exercise involving 2 public health emergency operating centers within the PANDEM-2 consortium (the Robert Koch Institute in Germany and the National Institute for Public Health and the Environment in the Netherlands) supports this position. The aim of this exercise was to explore the coordinated response to a large-scale pandemic due to a novel strain of avian influenza in Europe with the objective of testing and evaluating the functionalities of the PANDEM-2 IT system and associated tools. During the 2-day exercise, participants received 44 injects to develop the scenario and provide specific tasks to complete within their country's context. A manuscript outlining the implementation of this exercise is currently under review [11]. Evaluation of the exercise comprised both qualitative (hot-wash group interviews and after-action interviews) and quantitative (questionnaires and interviews), and feedback indicated that the PANDEM-2 dashboard met or exceeded their technical requirements in terms of user experience. Moreover, participants agreed that the dashboard addresses a current gap in national and cross-border training capacity for pandemic preparedness and response.

Alongside addressing data requirements, effective methods for data presentation need to be considered. The PANDEM-2 platform's participatory design approach emphasizes the importance of involving end users in the design of a system. Through this interactive and iterative approach, users can define system requirements, share domain expertise, and identify high-priority indicators for pandemic management. This approach within the PANDEM-2 platform enabled the creation of a design framework that offers adaptability and scalability, ensuring consistent interactions even as indicators evolve.

The participatory design approach also supported the effective knowledge transfer among end users and the technical team. This was crucial to enable the design team to understand the requirements of diverse and specialized domains and to facilitate the design of novel technical solutions within those domains. Ongoing research endeavors should be directed toward enhancing user engagement with intricate software in specialized areas, such as epidemiological and resource models, balancing user-friendly interfaces with necessary complexity, and exploring efficient techniques for presenting voluminous datasets and findings.

Data storytelling has become one such technique for enhancing user engagement and is often seen as an intuitive way to drive decision-making and reduce the risk of information overload. The participatory design process revealed that managing multiple decision narratives would add unnecessary complexity. The dashboard approach chosen provides flexibility across public health scenarios like surveillance and training. The ability to customize views and generate reports makes it a more scalable solution, especially for our primary users, public health officers, and medics trained in data analysis, who require more in-depth exploration than storytelling offers. The dashboard includes filtering, tagging, and prioritization features to address information overload, but managing complex data while meeting the needs of expert users remains a challenge. The participatory design process helped refine key indicators to present concise, relevant information, but balancing exploration and clarity is an ongoing effort in systems like PANDEM-2.

The identification of high-priority indicators and the possibility of consolidating them into a singular variable that captures the status of a pandemic merits in-depth investigation. This strategy could streamline decision-making processes and yield immediate insights into the dynamic nature of the pandemic across its various stages.

Summary

During the project, challenges related to data standardization across consortium members became evident, particularly due to the lack of data in the TESSy format. In addition, many countries had regional divisions that differed from those used by the ECDC (NUTS). End-user experts also highlighted the critical role dashboards play in pandemic preparedness and response. The findings underscore the importance of making data for these indicators accessible for public health purposes.

The project's commitment to responsible innovation and user-centric principles was demonstrated through its robust impact assessment process, which prioritized key elements such as privacy-by-design and adherence to GDPR as well as regulations governing health technology development. This iterative process ensured that the platform remained relevant and user-friendly. Transparency stands as a core value of the PANDEM-2 project, demonstrating its commitment to openness and accountability. All project resources have been made publicly available for interested parties to access and review on

XSL•FO
RenderX

the project website [42] and the European research platform Cordis [43], where stakeholders can find comprehensive information about data sources used, project procedures, and methods.

While the PANDEM-2 dashboard holds promise for effectively managing real-time data during future pandemics, successful integration with existing IT systems within public health agencies will be a necessary step. In the interim, the dashboard stands as an innovative training platform, providing member states the opportunity to assess their pandemic preparedness capacity through collaborative cross-border exercises developed by the PANDEM-2 project. Despite the limitations encountered during the project, our approach and findings provide valuable insights through open-access materials and tools to support public health agencies in EU member states to better prepare for upcoming pandemics.

To ensure comprehensive and transparent reporting of the development process, this study followed the Guidelines for Intervention Development checklist, documenting each stage of the dashboard's design and stakeholder engagement (Multimedia Appendix 13) [44].

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation program (grant 883285). The material presented and views expressed here are the responsibility of the authors only. The EU Commission takes no responsibility for any use made of the information set out. The authors would like to acknowledge the contribution of the PANDEM-2 public health agencies and first responder partners for their valuable input to the design and development process. This publication has emanated from research supported in part by a grant from Science Foundation Ireland (grant 12/RC/2289_P2; Insight SFI Centre for Data Analytics).

Data Availability

The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Security and data protection aspects of the dashboard.
[DOCX File , 18 KB - publichealth_v11i1e52119_app1.docx]

Multimedia Appendix 2
Dashboards reviewed.
[XLSX File (Microsoft Excel File), 39 KB - publichealth_v11i1e52119_app2.xlsx]

Multimedia Appendix 3
List of COVID-19–related projects.
[XLSX File (Microsoft Excel File), 68 KB - publichealth_v11i1e52119_app3.xlsx]

Multimedia Appendix 4
Example of data visualization entered in document.
[PDF File (Adobe PDF File), 225 KB - publichealth_v11i1e52119_app4.pdf]

Multimedia Appendix 5
Checklist for reporting results of the initial user requirements survey.
[DOCX File , 16 KB - publichealth_v11i1e52119_app5.docx]

Multimedia Appendix 6
Checklist for reporting results of internet e-surveys: participatory design.
[DOCX File , 16 KB - publichealth_v11i1e52119_app6.docx]

Multimedia Appendix 7
Example user stories.
[DOCX File , 16 KB - publichealth_v11i1e52119_app7.docx]

XSL•FO
RenderX

Multimedia Appendix 8
Data availability and prioritization survey questions.
[DOCX File , 14 KB - publichealth_v11i1e52119_app8.docx]

Multimedia Appendix 9
Final list of variables by variable category and requirement type.
[DOCX File , 25 KB - publichealth_v11i1e52119_app9.docx]

Multimedia Appendix 10
Existing data in the European Surveillance System format.
[PNG File , 53 KB - publichealth_v11i1e52119_app10.png]

Multimedia Appendix 11
Example design of dashboard page illustrating key components: (1) explanation header; (2) indicator card; (3) map; and (4) graph. NUTS: Nomenclature of Territorial Units for statistics (a hierarchical system for dividing up the economic territory of the European Union). Implemented dashboard page in the PANDEM-2 prototype, reflecting design feedback with features like customizable reports and enhanced data visualization.
[PDF File (Adobe PDF File), 640 KB - publichealth_v11i1e52119_app11.pdf]

Multimedia Appendix 12
Illustrations of initial and final dashboard designs: progression of the map component.
[DOCX File , 541 KB - publichealth_v11i1e52119_app12.docx]

Multimedia Appendix 13
Guidelines for Intervention Development checklist.
[DOCX File , 30 KB - publichealth_v11i1e52119_app13.docx]

References

1. Jit M, Ananthakrishnan A, McKee M, Wouters OJ, Beutels P, Teerawattananon Y. Multi-country collaboration in responding to global infectious disease threats: lessons for Europe from the COVID-19 pandemic. Lancet Reg Health Eur 2021 Oct;9:100221 [FREE Full text] [doi: 10.1016/j.lanepe.2021.100221] [Medline: 34642675]
2. Haldane V, Jung AS, Neill R, Singh S, Wu S, Jamieson M, et al. From response to transformation: how countries can strengthen national pandemic preparedness and response systems. BMJ 2021 Nov 28;375:e067507 [FREE Full text] [doi: 10.1136/bmj-2021-067507] [Medline: 34840139]
3. Lencucha R, Bandara S. Trust, risk, and the challenge of information sharing during a health emergency. Global Health 2021 Feb 18;17(1):21 [FREE Full text] [doi: 10.1186/s12992-021-00673-9] [Medline: 33602281]
4. Badker R, Miller K, Pardee C, Oppenheim B, Stephenson N, Ash B, et al. Challenges in reported COVID-19 data: best practices and recommendations for future epidemics. BMJ Glob Health 2021 May;6(5):e005542 [FREE Full text] [doi: 10.1136/bmjgh-2021-005542] [Medline: 33958393]
5. Tacconelli E, Gorska A, Carrara E, Davis RJ, Bonten M, Friedrich AW, et al. Challenges of data sharing in European covid-19 projects: a learning opportunity for advancing pandemic preparedness and response. Lancet Reg Health Eur 2022 Oct;21:100467 [FREE Full text] [doi: 10.1016/j.lanepe.2022.100467] [Medline: 35942201]
6. Rosenthal A, Waitzberg R. The challenges brought by the COVID-19 pandemic to health systems exposed pre-existing gaps. Health Policy Open 2023 Dec;4:100088 [FREE Full text] [doi: 10.1016/j.hpopen.2022.100088] [Medline: 36536931]
7. Kokki M, Ammon A. Preparing Europe for future health threats and crises - key elements of the European Centre for Disease Prevention and Control's reinforced mandate. Euro Surveill 2023 Jan;28(3):2300033 [FREE Full text] [doi: 10.2807/1560-7917.ES.2023.28.3.2300033] [Medline: 36695487]
8. Katapally TR, Ibrahim ST. Digital health dashboards for decision-making to enable rapid responses during public health crises: replicable and scalable methodology. JMIR Res Protoc 2023 Jun 30;12:e46810 [FREE Full text] [doi: 10.2196/46810] [Medline: 37389905]
9. Tighe C. PANDEM-2 dashboard demo. Zenodo. 2023 Sep 20. URL: https://zenodo.org/records/8363810 [accessed 2024-09-14]
10. PANDEM-2 dashboard. PANDEM-2. URL: https://pandem-2.eu/pandem-2-dashboard/ [accessed 2024-09-14]
11. Houareau C, Spieker C, Grote U, Perseke K, An der Heiden M, Caglar R, et al. The PANDEM-2 Simulation Exercise: Training the Coordinated Response to a Large-Scale Pandemic in 2 European Public Health Emergency Operations Centers. Disaster Medicine and Public Health Preparedness 2025 Jan 14;19(14):e14 [FREE Full text] [doi: 10.1017/dmp.2024.298] [Medline: 39807047]

XSL•FO
RenderX

12. MongoDB homepage. MongoDB. URL: https://www.mongodb.com/ [accessed 2024-08-25]
13. NodeJS homepage. NodeJS. URL: https://nodejs.org/en [accessed 2024-08-28]
14. Chaniotis IK, Kyriakou KI, Tselikas ND. Is Node.js a viable option for building modern web applications? A performance evaluation study. Computing 2014 Mar 28;97:1023-1044. [doi: 10.1007/s00607-014-0394-9]
15. Angular homepage. Angular. URL: https://angular.io [accessed 2024-09-03]
16. Highcharts homepage. Highcharts. URL: https://www.highcharts.com/ [accessed 2024-09-10]
17. Pandem-source. GitHub. URL: https://github.com/Pandem2/pandem-source [accessed 2024-09-10]
18. Overview - Eurostat. Eurostat. URL: https://ec.europa.eu/eurostat/web/nuts/overview [accessed 2024-08-28]
19. GeoNames homepage. GeoNames. URL: https://www.geonames.org [accessed 2024-09-03]
20. COVID-19 data hub homepage. COVID-19 Data Hub. URL: https://covid19datahub.io [accessed 2024-09-10]
21. Chen H, Gala JL, Bonjean M, Ambroise J, Zayed O, Buitelaar P, et al. A COVID-19 European data set to support training in pandemic management. Zenodo. 2023. URL: https://zenodo.org/records/8339303 [accessed 2025-01-06]
22. Go.Data: managing complex data in outbreaks. World Health Organization. URL: https://www.who.int/godata [accessed 2024-09-10]
23. Flightradar24 homepage. Flightradar24. URL: https://www.flightradar24.com [accessed 2024-08-28]
24. Amazon web services cloud security homepage. Amazon Web Services. URL: https://aws.amazon.com/security [accessed 2024-09-03]
25. The AWS shared responsibility model and GDPR. Amazon Web Services. URL: https://aws.amazon.com/blogs/security/the-aws-shared-responsibility-model-and-gdpr/ [accessed 2024-09-10]
26. Amazon GuardDuty. Amazon Web Services. URL: https://aws.amazon.com/guardduty [accessed 2024-09-03]
27. AWS key management service. Amazon Web Services. URL: https://aws.amazon.com/kms/ [accessed 2024-08-28]
28. Few S. Information Dashboard Design: The Effective Visual Communication of Data. Sebastopol, CA: O'Reilly Media, Inc; 2006.
29. Franklin A, Gantela S, Shifarraw S, Johnson TR, Robinson DJ, King BR, et al. Dashboard visualizations: supporting real-time throughput decision-making. J Biomed Inform 2017 Jul;71:211-221 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.024] [Medline: 28579532]
30. Lennerholt C, Van Laere J, Söderström E. Implementation challenges of self-service business intelligence: a literature review. In: Proceedings of the 51st Hawaii International Conference on System Sciences. 2018 Presented at: HICSS 2018; January 3-6, 2018; Hilton Waikoloa Village, HI. [doi: 10.24251/hicss.2018.631]
31. Shao H, Martinez-Maldonado R, Echeverria V, Yan L, Gasevic D. Data storytelling in data visualisation: does it enhance the efficiency and effectiveness of information retrieval and insights comprehension? In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. 2024 Presented at: CHI '24; May 11-16, 2024; Honolulu, HI. [doi: 10.1145/3613904.3643022]
32. Cohn M. User Stories Applied: For Agile Software Development. Boston, MA: Addison-Wesley Publishing Company; 2004.
33. Lenarduzzi V, Taibi D. MVP explained: a systematic mapping study on the definitions of minimal viable product. In: Proceedings of the 42th Euromicro Conference on Software Engineering and Advanced Applications. 2016 Presented at: SEAA 2016; August 31-September 2, 2016; Limassol, Cyprus. [doi: 10.1109/seaa.2016.56]
34. Simonsen J, Robertson T. Routledge International Handbook of Participatory Design. Milton Park, UK: Routledge; 2013.
35. Conducting privacy impact assessments code of practice. Information Commissioner's Office. 2014. URL: https://ico.org.uk/media/about-the-ico/consultations/2052/draft-conducting-privacy-impact-assessments-code-of-practice.pdf [accessed 2024-09-03]
36. ISO/IEC 29134:2017(en) Information technology — security techniques — guidelines for privacy impact assessment. International Organization for Standardization. 2017. URL: https://www.iso.org/obp/ui/#iso:std:iso-iec:29134:ed-1:v1:en [accessed 2024-08-28]
37. The European Surveillance System (TESSy). European Centre for Disease Prevention and Control. 2017 May 17. URL: https://www.ecdc.europa.eu/en/publications-data/european-surveillance-system-tessy [accessed 2024-09-10]
38. Pandemic preparedness and response. European Commission. URL: https://cordis.europa.eu/project/id/883285/results [accessed 2024-09-10]
39. Regulation 2022/2371 - serious cross-border threats to health. EU Monitor. URL: https://www.eumonitor.eu/9353000/1/j4nvk6yhcbpeywk_j9vvik7m1c3gyxp/vlync5alazzz#:~:text=This%20Regulation%20should%20strengthen%20and,European%20Parliament%20and%20of%20the [accessed 2024-09-03]
40. Regulation (EU) 2022/2370 of the European parliament and of the council of 23 November 2022. Official Journal of the European Union. 2022 Dec 6. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R2370&from=EN [accessed 2024-08-28]
41. Simulation exercises. World Health Organization. URL: https://www.who.int/emergencies/operations/simulation-exercises [accessed 2024-09-03]
42. EU pandemic research. PANDEM-2. URL: https://pandem-2.eu/ [accessed 2024-09-10]

XSL•FO
RenderX

43. CORDIS - EU research results. European Commission. URL: https://cordis.europa.eu/ [accessed 2024-09-03]
44. Duncan E, O'Cathain A, Rousseau N, Croot L, Sworn K, Turner KM, et al. Guidance for reporting intervention development studies in health research (GUIDED): an evidence-based consensus study. BMJ Open 2020 Apr 08;10(4):e033516 [FREE Full text] [doi: 10.1136/bmjopen-2019-033516] [Medline: 32273313]

Abbreviations

APD: asynchronous participatory design
API: application programming interface
AWS: Amazon Web Services
ECDC: European Centre for Disease Prevention and Control
ETL: extract, transform, and load
EU: European Union
GDPR: General Data Protection Regulation
MVP: minimal viable product
NUTS: Nomenclature of Territorial Units for Statistics
TESSy: European Surveillance System
WHO: World Health Organization

Edited by A Mavragani; submitted 25.08.23; peer-reviewed by KS Sahu, O Leal Neto; comments to author 18.07.24; revised version received 16.09.24; accepted 23.09.24; published 05.03.25.

Please cite as:
Tighe C, Ngongalah L, Sentís A, Orchard F, Pacurar GA, Hayes C, Hayes JS, Toader A, Connolly MA
Building and Developing a Tool (PANDEM-2 Dashboard) to Strengthen Pandemic Management: Participatory Design Study
JMIR Public Health Surveill 2025;11:e52119
URL: https://publichealth.jmir.org/2025/1/e52119
doi:10.2196/52119
PMID:40053759

©Carlos Tighe, Lem Ngongalah, Alexis Sentís, Francisco Orchard, Gheorghe-Aurel Pacurar, Conor Hayes, Jessica S Hayes, Adrian Toader, Máire A Connolly. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 05.03.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Suspected Mpox Symptoms and Testing in Men Who Have Sex With Men in the United States: Cross-Sectional Study

Kaitlyn Atkins[1,2], MPH, PhD; Thomas Carpino[2], MPH; Amrita Rao[2], ScM, PhD; Travis Sanchez[3], DVM, MPH; O Winslow Edwards[3], MPH; Marissa Hannah[3], MPH; Patrick S Sullivan[3], DVM, PhD; Yasmin P Ogale[4], MSPH, PhD; Winston E Abara[4], MD, PhD; Kevin P Delaney[4], MPH, PhD; Stefan D Baral[2], MPH, MD

[1]Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[2]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[3]Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA, United States

[4]Centers for Disease Control and Prevention, Atlanta, GA, United States

Corresponding Author:
Kaitlyn Atkins, MPH, PhD
Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

Abstract

Background: The 2022 mpox outbreak in the United States disproportionately affected gay, bisexual, and other men who have sex with men (GBMSM). Uptake of mpox testing may be related to symptomology, sociodemographic characteristics, and behavioral characteristics.

Objective: This study aimed to describe suspected mpox symptoms and testing uptake among a sample of GBMSM recruited via the internet in the United States in August 2022.

Methods: We conducted a rapid internet-based mpox survey from August 5 to 15, 2022, among cisgender men 15 years and older who had previously participated in the 2021 American Men's Internet Survey. We estimated the prevalence of suspected mpox symptoms (fever or rash or sores with unknown cause in the last 3 mo) and uptake of mpox testing. We calculated adjusted prevalence ratios (aPRs) and 95% CIs for associations between participant characteristics and suspected mpox symptoms and summarized characteristics of GBMSM reporting mpox testing. Among symptomatic GBMSM who did not receive mpox testing, we described testing self-efficacy, barriers, and facilitators.

Results: Of 824 GBMSM, 126 (15.3%) reported at least 1 mpox symptom in the last 3 months; 58/126 (46%) with rash or sores, 57 (45.2%) with fever, and 11 (8.7%) with both. Increased prevalence of suspected mpox symptoms was associated with condomless anal sex (CAS; aPR 1.53, 95% CI 1.06 - 2.20). Mpox testing was reported by 9/824 GBMSM (1%), including 5 with symptoms. Most GBMSM reporting mpox testing were non-Hispanic White men (7/9 vs 1 Black and 1 Hispanic or Latino man), and all 9 lived in urban areas. Most reported having an sexually transmitted infections test (8/9), 2 or more partners (8/9), CAS (7/9), and group sex (6/9) in the last 3 months. Of those tested, 3 reported living with HIV and all were on treatment, whereas the remaining 6 men without HIV reported current pre-exposure prophylaxis (PrEP) use. Of symptomatic GBMSM who did not report mpox testing, 47/105 (44.8%) had low mpox testing self-efficacy. Among those with low self-efficacy, the most common barriers to testing were not knowing where to get tested (40/47, 85.1%) and difficulty getting appointments (23/47, 48.9%). Among those with high testing self-efficacy (58/105, 55.2%), the most common facilitators to testing were knowing where to test (52/58, 89.7%), convenient site hours (40/58, 69%), and low-cost testing (38/58, 65.5%).

Conclusions: While all GBMSM who reported testing for mpox were linked to HIV treatment or PrEP, those with symptoms but no mpox testing reported fewer such links. This suggests targeted outreach is needed to reduce structural barriers to mpox services among GBMSM in rural areas, Black and Hispanic or Latino GBMSM, and GBMSM living with HIV. Sustaining and scaling community-tailored messaging to promote testing and vaccination represent critical interventions for mpox control among GBMSM in the United States.

(JMIR Public Health Surveill 2025;11:e57399) doi:10.2196/57399

KEYWORDS

mpox; emerging infectious diseases; men who have sex with men; sexual health; testing; monkeypox; epidemic; MSM; United States; rural; cross-sectional study; gay; bisexual; symptomology; sociodemographic; online sample; self-efficacy; rash; fever; HIV; HIV prevention; GBMSM; Black and Hispanic; LGBTQ; Latino; public health; surveillance; sores

XSL•FO
RenderX

Introduction

Human mpox is an orthopox virus first identified in the Democratic Republic of the Congo and endemic to regions of Central and Western Africa, which has caused localized outbreaks for decades [1]. mpox is classified into 2 clades; of the two, clade I mpox is associated with more severe symptoms and is highly lethal, with an estimated 10% mortality rate and several thousands of cases annually in Central Africa, including a 2023‐2024 outbreak in the Democratic Republic of the Congo [2]. While clade I mpox has not been observed outside the African continent, clade II mpox has been occasionally reported internationally, often linked to travel [3]. Although clade II infections are associated with a much lower mortality rate, they have resulted in increasing global morbidity, including a global outbreak in 2022 [3]. From 2022 to 2024, there have been 95,000 confirmed cases of clade II mpox globally, including several deaths, predominantly affecting gay, bisexual, and other men who have sex with men (GBMSM) [1].

The first case in the 2022 mpox outbreak was identified in the United States on May 17, 2022, triggering a nationwide response to identify and monitor new cases and deliver the mpox vaccine [4]. As of March 5, 2024, approximately 32,000 mpox cases have been reported in the United States [1], and over 90% of cases in the United States have occurred among cisgender men, the majority of whom had recent sexual contact with a man [5].

As elsewhere, the 2022US mpox outbreak and its disproportionate impacts on GBMSM have been attributed to connected sexual networks. Moreover, the US mpox outbreak demonstrated disparities in infections associated with structural determinants of health such as racism: communities of color reported higher burdens of infections, but lower uptake of testing and vaccines [1,6]. In studies of testing for sexually transmitted infections (STIs) and COVID-19, lower socioeconomic status has been associated with reduced testing uptake [7,8]. However, few studies to date have examined how structural determinants may have shaped mpox testing in a large, national sample of MSM [9,10].

Beyond structural factors, engagement in mpox screening and testing services may be driven by individual factors, particularly risk perception and behavioral indication for mpox tests [11]. For example, GBMSM who report regular condom use have been found to have lower engagement in testing for HIV [12] and bacterial STIs [13]. Condom use and engagement in other sexual behaviors associated with mpox (such as group sex) [14,15] may therefore be associated with increased engagement in mpox testing.

New infections have slowed significantly since the peak of the outbreak [1,16], but equitable testing remains key to contract tracing, treatment, and linkage to vaccination programs. Understanding potential barriers and facilitators to mpox testing during the peak of the 2022 outbreak can inform strategies for future mpox or other STI outbreaks. The current work aims to fill knowledge gaps around these barriers and facilitators by describing suspected mpox symptoms and testing from an mpox study among a sample of prior participants in the American Men's Internet Survey (AMIS).

Methods

Recruitment

The AMIS mpox study was an internet-based survey conducted from August 5 to 15, 2022 in GBMSM aged 15 years and older who had participated in AMIS 2021. Methods for the annual AMIS have been described elsewhere [17]. Annual AMIS participants were recruited via social media and deemed eligible if they identified as cisgender male, aged ≥15 years, had ever had sex with a man, and lived in the United States. For the mpox study, the sampling frame included GBMSM who had completed the AMIS 2021 survey cycle, had sex with a man in the past year, and had consented to be recontacted for future research (N=2999).

Ethical Considerations

This study received ethical approval from the Emory University Institutional Review Board (IRB00047676). All activities were conducted consistent with US Centers for Disease Control and Prevention policies and regulations. All participants provided informed consent to participate and were not compensated. All study data were deidentified.

Procedures

Potential participants (ie, all AMIS 2021 participants who had consented to be recontacted for future research, N=2999) were contacted via email with an invitation to participate and a secure survey link. Potential participants were contacted up to twice and provided electronic informed consent after eligibility screening. The internet-based survey included questions in English about demographics, sexual behavior, substance use, HIV and STI testing and diagnosis, HIV pre-exposure prophylaxis (PrEP) use, and mpox knowledge, symptoms, testing, diagnosis, vaccination, stigma, and behavior change.

Measures

We measured participant age, race or ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic or Latino, other), US census region (Northeast, Midwest, South, or West), urbanicity based on 2013 National Center for Health Statistics classification [18], health insurance (none, private, public, other), self-reported HIV status, current antiretroviral therapy use (among people living with HIV, [19]) and current PrEP use (among those without HIV).

Participants reported, in the past 3 months, their number of sexual partners (1 or ≥2), participation in group sex, bathhouses or sex clubs, or sex parties, and any illicit or injection drug use. They also reported condomless anal sex (CAS) in the past 12 months, CAS with partners of a different HIV status in the past 3 months, HIV testing in the past 12 months, and STI testing in the past 3 months.

Outcomes

We assessed symptoms suggesting mpox by asking: "In the past three months, have you had a (1: fever or 2: new rash/sores on the skin) and didn't know the cause?" For GBMSM reporting rashes, we asked their location and pain severity (rated 0‐10),

XSL•FO
RenderX

and whether, after GBMSM first noticed the rash or sores, they changed in appearance, number, or location.

We measured mpox testing by self-report. For GBMSM reporting no mpox testing, we assessed agreement (4-point Likert scale) with the statement: "If I thought I had mpox, I could get a test if I wanted it." GBMSM who strongly or somewhat agreed were considered to have high mpox testing self-efficacy, and those who strongly or somewhat disagreed as having low self-efficacy. We also asked participants what made it easy (for those with high self-efficacy) or difficult (for those with low self-efficacy) to get a test; parallel response options included location, eligibility, appointments, using one's own provider, site times, privacy, and cost.

Statistical Analysis

We calculated frequencies for participant characteristics, symptoms suggesting mpox, and mpox testing. We calculated unadjusted prevalence ratios (PRs) and 95% CIs for associations between participant characteristics and symptoms suggesting mpox. Factors found to have a significant association (α<.05) with symptoms were included in a multivariable model to calculate adjusted prevalence ratios (aPRs). The final model included age, race or ethnicity, PrEP use, STI testing, and CAS. We assessed model fit using Hosmer-Lemeshow test, which has been recommended for use with binomially distributed data [20]. We used frequencies to summarize the characteristics of GBMSM reporting mpox testing. Among symptomatic GBMSM who did not report mpox testing, we summarized testing self-efficacy, barriers, and facilitators. All analyses were performed in Stata 15.1 (StataCorp) [21].

Results

Sample Characteristics

Of 2999 GBMSM in the sampling frame, 824 completed the mpox study survey for a response rate of 27.5%. The majority of mpox study participants were aged ≥40 years (478/824, 58%), non-Hispanic White (581/824, 70.9%), residents of urban areas (770/824, 93.4%), and had private health insurance (564/824, 68.6%, Table 1). In the last 3 months, a minority reported CAS with a partner of a different HIV status (220/824, 26.7%), group sex (238/824, 29.1%), or attending a sex club (147/824, 19.7%), or sex party (111/824, 14.9%). Just over two-thirds (575/824, 69.8%) reported past-year HIV testing, and around half (408/824, 49.6%) reported STI testing in the last 3 months. Most (720/824, 87.4%) were not people living with HIV and 44.7% of these (316/720, 44.7%) were currently using PrEP; of 104 people living with HIV in the sample, 99 (96.1%) were currently on HIV treatment.

XSL•FO
RenderX

Table . Characteristics of participants in the American Men's Internet Survey mpox substudy, United States, August 2022 (n=824).

Characteristics		Participants (n=824), n (%)
Age (years)		
	15 - 24	46 (5.6)
	25 - 29	86 (10.4)
	30 - 39	214 (26)
	40 or older	478 (58)
Race or ethnicity		
	Non-Hispanic White	581 (70.9)
	Non-Hispanic Black	87 (10.6)
	Hispanic or Latino	88 (10.7)
	Other	64 (7.8)
US census region		
	Northeast	174 (21.1)
	Midwest	141 (17.1)
	South	305 (37)
	West	204 (24.8)
Urbanicity		
	Rural	54 (6.6)
	Urban	770 (93.4)
Health insurance		
	None	24 (2.9)
	Private	564 (68.6)
	Public	164 (20)
	Other or multiple	70 (8.5)
HIV test, last 12 months		
	No	249 (30.2)
	Yes	575 (69.8)
Bacterial STI[a] test, last 3 months		
	No	410 (49.8)
	Yes	409 (49.6)
HIV status		
	Not living with HIV	720 (87.4)
	Living with HIV	104 (12.6)
Current ART[b] use (n=103)[c]		
	No	4 (3.9)
	Yes	99 (96.1)
Current PrEP[d] use (n=707)		
	No	391 (55.3)
	Yes	316 (44.7)
Group sex, last 3 months		
	No	580 (70.9)
	Yes	238 (29.1)

XSL•FO
RenderX

Characteristics		Participants (n=824), n (%)
Sex club, last 3 months		
	No	599 (80.3)
	Yes	147 (19.7)
Sex party, last 3 months		
	No	634 (85.1)
	Yes	111 (14.9)
CAS[e] with serodifferent partner, last 3 months		
	No	604 (73.3)
	Yes	220 (26.7)

[a]STI: sexually transmitted infection.

[b] ART: antiretroviral therapy.

[c] One participant with HIV did not provide information on ART use.

[d]PrEP: HIV pre-exposure prophylaxis.

[e]CAS: condomless anal sex.

Prevalence of Symptoms Suggesting mpox and Associated Factors

In total, 126/824 (15.3%) participants reported at least 1 mpox symptom in the last 3 months (Table 2): 58/126 (46%) with rash or sores, 57 (45%) with fever, and 11 (9%) with both. In unadjusted analyses, there was decreased prevalence of symptoms suggesting mpox among participants aged ≥40 years versus 15 - 24 years (PR 0.48, 95% CI 0.28 - 0.82) and among Hispanic or Latino versus non-Hispanic White participants (PR 0.44, 95% CI 0.20 - 0.97), and increased prevalence among participants reporting current PrEP use (PR 1.48, 95% CI 1.05 - 2.09), recent STI tests (PR 1.40, 95% CI 1.01 - 1.95), and CAS in the last 12 months (PR 1.64, 95% CI 1.14 - 2.35). In our adjusted model, we failed to reject the Hosmer-Lemeshow null hypothesis (Hosmer-Lemeshow $\chi_8^2$5.01; P=.76), indicating adequate model fit [20]. In adjusted analyses, prevalence of symptoms suggesting mpox remained lower among GBMSM aged ≥40 years versus 15 - 24 years (aPR 0.49, 95% CI 0.28 - 0.83), and higher among those reporting CAS (aPR 1.53, 95% CI 1.06 - 2.20).

XSL•FO
RenderX

Table . Symptoms suggesting mpox and testing among participants in the American Men's Internet Survey mpox substudy, United States, August 2022 (n=824). Bold text indicates measures of association found to be statistically significant at α<.05; that is, confidence intervals that did not cross 1.00.

		At least 1 symptom suggesting mpox[a], last 3 months			Symptoms, no mpox testing	Ever tested for mpox
		Participants (n=126), n (%)	PR[b] (95% CI)	aPR[c] (95% CI)	Participants (n=114), n (%)	Participants (n=9), n (%)
Symptoms suggesting mpox						
	No fever or rash	__[d]	—	—	—	4 (44.4)
	Fever only	57 (45.2)	—	—	53 (46.5)	0 (0)
	Rash or sores only	58 (46)	—	—	52 (45.6)	4 (44.4)
	Fever and rash or sores	11 (8.7)	—	—	9 (7.9)	1 (11.1)
Age (years)						
	15 - 24	12 (9.5)	REF[e]	REF	10 (8.8)	1 (11.1)
	25 - 29	14 (11.1)	0.61 (0.31 - 1.21)	0.64 (0.33 - 1.26)	12 (10.5)	2 (22.2)
	30 - 39	39 (31)	0.68 (0.39 - 1.20)	0.64 (0.37 - 1.12)	36 (31.6)	3 (33.3)
	40 or older	61 (48.4)	0.48 (0.28 - 0.82)	0.49 (0.28 - 0.83)	56 (49.1)	3 (33.3)
Race or ethnicity						
	Non-Hispanic White	90 (71.4)	REF	REF	81 (71.1)	7 (77.8)
	Non-Hispanic Black	14 (11.1)	1.03 (0.62 - 1.73)	1.01 (0.60 - 1.71)	12 (10.5)	1 (11.1)
	Hispanic or Latino	6 (4.8)	0.44 (0.20 - 0.97)	0.41 (0.19 - 0.91)	5 (4.4)	1 (11.1)
	Other	16 (12.7)	1.63 (1.02 - 2.59)	1.44 (0.90 - 2.30)	16 (14)	0 (0)
Urbanicity						
	Nonurban	11 (8.7)	REF	—	10 (8.8)	0 (0)
	Urban	115 (91.3)	0.74 (0.42 - 1.28)	—	104 (91.2)	9 (100)
HIV status						
	Not living with HIV	111 (88.1)	REF	—	102 (89.5)	6 (66.7)
	Living with HIV	15 (11.9)	0.93 (0.56 - 1.53)	—	12 (10.5)	3 (33.3)
Current ART[f] use (n=15)						
	No	0 (0)	—	—	0 (0)	0 (0)
	Yes	15 (100)	—	—	12 (100)	3 (100)
Current PrEP[g] use (n=707)[h], [i]						
	No	50 (45.5)	REF	REF	47 (46.5)	0 (0)
	Yes	60 (54.5)	1.48 (1.05 - 2.09)	1.15 (0.75 - 1.75)	54 (53.5)	6 (100)
STI[j] test, last 3 months						
	No	52 (41.3)	REF	REF	45 (39.5)	1 (11.1)

XSL•FO
RenderX

	At least 1 symptom suggesting mpox[a], last 3 months			Symptoms, no mpox testing	Ever tested for mpox
	Participants (n=126), n (%)	PR[b] (95% CI)	aPR[c] (95% CI)	Participants (n=114), n (%)	Participants (n=9), n (%)
Yes	74 (58.7)	1.40 (1.01 - 1.95)	1.33 (0.96 - 1.85)	69 (60.5)	8 (88.9)
Number of partners, last 3 months					
One	25 (21)	REF	—	23 (21.3)	0 (0)
Two or more	94 (79%)	1.44 (0.95 - 2.17)	—	85 (78.7)	8 (100)
Group sex, last 3 months					
No	81 (64.3%)	REF	—	76 (66.7)	3 (33.3)
Yes	45 (35.7)	1.35 (0.97 - 1.88)	—	38 (33.3)	6 (66.7)
Sex club, last 3 months					
No	88 (74)	REF	—	81 (75)	5 (62.5)
Yes	31 (26)	1.44 (1.00 - 2.08)	—	27 (25)	3 (37.5)
Sex party, last 3 months					
No	96 (80.7)	REF	—	89 (82.4)	5 (62.5)
Yes	23 (19.3)	1.36 (0.91 - 2.05)	—	19 (17.6)	3 (37.5)
CAS[k], last 12 months					
No	36 (28.6)	REF	REF	33 (28.9)	2 (22.2)
Yes	90 (71.4)	1.64 (1.14 - 2.35)	1.53 (1.06 - 2.20)	81 (71.1)	7 (77.8)

[a]mpox: mpox.

[b]PR: prevalence ratio.

[c]aPR: adjusted prevalence ratio.

[d]Not applicable.

[e]REF: reference.

[f]ART: antiretroviral therapy.

[g]PrEP: HIV pre-exposure prophylaxis.

[h]The model with PrEP use was restricted to individuals without HIV, and included all listed covariates except HIV status.

[i]No PR reported given 100% of symptomatic people living with HIV reported current antiretroviral therapy use.

[j]STI: sexually transmitted infection.

[k]CAS: condomless anal sex.

Uptake of mpox Testing

Among 126 symptomatic GBMSM, 114 (90.5%) did not report mpox testing; their characteristics mirrored those of the overall sample (Table 2). mpox testing was reported by 9 participants (1%, Table 2), including 5 with reported symptoms. Most GBMSM reporting mpox testing were non-Hispanic White (7/9 vs 1 Black and 1 Hispanic or Latino man) and all 9 lived in urban areas. Most reported STI testing (8/9), 2 or more partners (8/9), CAS (7/9), and group sex (6/9) in the last 3 months. Three were people living with HIV, all on treatment; the remaining 6 without HIV reported current PrEP use.

Among 69 GBMSM experiencing recent rashes or sores with unknown cause, most located them on hands or arms, feet or legs, and genitals or pelvic area (Table 3). Over half (31/50, 62%) rated their pain as ≤2 on a 10-point scale. Most (38/69, 55%) reported no changes over time; the most common reported change was worsening appearance of rash or sores (22/69, 32%).

XSL•FO
RenderX

Table . Rash or sore location, severity, and changes among men having sex with men in the United States reporting new rashes or sores in the last 3 months (n=69).

		Experienced new rash or sores with unknown cause, last 3 months (n=69), n (%)
Location of rash		
	Hands or arms	26 (37.7)
	Feet or legs	24 (34.8)
	Genitals or pelvic area	19 (27.5)
	Chest	14 (20.3)
	Face	13 (18.8)
	Anus	10 (14.5)
	Mouth	7 (10.1)
Severity of rash or sores		
	0: no pain at all	10 (20)
	1 - 2	21 (42)
	3 - 4	11 (22)
	5 - 6	4 (8)
	7 - 8	4 (8)
	9 - 10: the worst pain possible	0 (0)
Changes in rash or sores		
	No changes	38 (55.1)
	Worsening appearance	22 (31.9)
	Increased number of rashes or sores in the same location	11 (15.9)
	Rashes or sores in new locations	7 (10.1)

Of the 126 symptomatic GBMSM, 105 (83%) did not report mpox testing and were offered questions about testing self-efficacy and barriers and facilitators to testing (Table 4). Among GBMSM with high testing self-efficacy (58/105, 55%), common facilitators to testing were knowing where to test (90%), convenient site hours (69%), and low-cost testing (38/58, 66%). Among those with low testing self-efficacy (47/105, 45%), the most common barriers to testing were not knowing where to get tested (40/47, 85%) and difficulty getting appointments (23/47, 49%).

Table . Characteristics of gay, bisexual, and other men who have sex with men with symptoms suggesting mpox who did not take up mpox testing (n=105).

Barriers or facilitators to testing	Low testing self-efficacy[a] (n=47), n (%)	High testing self-efficacy (n=58), n (%)
(Not) knowing where to get tested	40 (85.1)	52 (89.7)
Getting an appointment	23 (48.9)	29 (50)
Getting tested at own doctor	17 (36.2)	24 (41.4)
Convenience or testing site hours	11 (23.4)	40 (69)
(Not) knowing who is eligible	10 (21.3)	29 (50)
Cost considerations	7 (14.9)	38 (65.5)
Privacy considerations	6 (12.8)	19 (32.8)

[a] Assessed using the statement: “If I thought I had mpox, I could get a test if I wanted it.” Participants who strongly or somewhat agreed were considered to have high mpox testing self-efficacy, and those who strongly or somewhat disagreed as having low self-efficacy.

XSL•FO
RenderX

Discussion

Principal Findings

Our findings show that a large proportion of GBMSM with symptoms suggesting mpox did not access mpox testing and mpox testing was low or zero among GBMSM in rural areas and GBMSM who were younger, Black, Hispanic or Latino, or people living with HIV.

Compared with non-Hispanic White participants, we observed low levels of mpox testing among Black and Hispanic or Latino GBMSM in this study. While this likely reflects our overall sample, which was primarily non-Hispanic White, it is worth noting that this finding aligns with broader disparities not only in mpox cases and vaccine delivery, but also with the majority of US vaccine doses administered to date to non-Hispanic White individuals [1]. Our findings may reflect structural challenges with communication and knowledge of access to and locations of testing, laboratory capacity, and stigma [22]. As others have noted, additional data are needed to ensure equitable access to and communication on mpox testing, including routine collection of demographic data through testing programs [23,24].

There was no reported testing among GBMSM in rural or nonurban areas in our study. While few participants in our study were from rural areas, this parallels findings from an internet-based study also conducted in August 2022, which found that GBMSM in rural areas had lower perceived susceptibility to and severity of mpox, compared with urban GBMSM [25]. The authors of the earlier manuscript found further disparities in rural GBMSM's perceptions of mpox vaccine benefits and in their intention to be vaccinated, amidst potentially inadequate vaccine supply. Our findings likely reflect broader challenges with public health infrastructure in rural areas [26] and increased availability of HIV and STI services in urban areas, which may have facilitated linkages to mpox testing. While there are opportunities to leverage these services for mpox testing, given only half of potentially symptomatic GBMSM were engaged with PrEP and STI services, additional outreach may be needed.

Testing uptake was also low among GBMSM living with HIV. While only about 1 in 5 participants in our study were people living with HIV, we agree with others that tailored interventions can better serve this community with services, including mpox testing, postexposure prophylaxis for mpox exposure, and vaccination [27]. Current US surveillance data suggest that around 40% of mpox diagnoses were among people living with HIV [1]. Given recent evidence that suggests severe mpox in the context of advanced immunosuppression may be an AIDS-defining condition [28], structural interventions to improve access to mpox testing among people living with HIV, such as integrating mpox testing into routine care for people living with HIV, are critical to ensure early identification and treatment of new mpox cases.

We found that the most symptoms suggesting mpox in our sample were rashes or sores, most commonly in the arms, legs, and genitals. Rashes in the arms and legs were reported at a similar rate in our study as has been shown in broader US surveillance data; however, we saw less frequent rashes in the genitals and face [5]. In our study, rashes and sores were reported to be generally mild and commonly did not change after they were first observed. This may explain the relatively low uptake of mpox testing among potentially symptomatic GBMSM, because GBMSM with mild symptoms suggesting mpox may not have felt they needed a test or been eligible for testing at that time [5]. During clinical encounters with GBMSM (eg, STI testing), providers should continue to counsel GBMSM to seek mpox testing and vaccine if they develop unexplained rash or other symptoms consistent with mpox.

Limitations

A key limitation of these analyses is the small sample of GBMSM in our study who reported mpox testing (n=9), precluding advanced statistical analyses comparing the characteristics of GBMSM who tested for mpox with those who did not. Furthermore, we are unable to draw generalizable conclusions based on the characteristics of these 9 individuals. We recruited a convenience sample that is subject to selection bias; our respondents were predominantly non-Hispanic White, elderly, and privately insured and may have increased access to mpox testing services—the sample is thus unlikely to represent all GBMSM in the United States. Our use of self-reported data and recall period of 3 months for most outcomes may have introduced recall bias; furthermore, we are unable to confirm whether those with symptoms suggesting mpox indeed had acquired mpox. Given symptoms were reported in the last 3 months, in some cases, participant responses may have reflected symptoms before the first US mpox case on May 17, 2022. We were also not able to place individuals' locations at the time of symptom onset or testing, which may have affected testing uptake given the nonuniform roll-out of testing nationwide. Regarding symptoms, we only included items assessing the presence of fever and rash or sores; we did not inquire about other symptoms commonly reported among mpox cases such as malaise, headache, nausea or vomiting, pain, itching, or bleeding. Low mpox testing prevalence in this sample may be attributed to perceptions that symptoms, such as fever or rash, were unrelated to mpox and therefore did not require testing. Our findings also reflect the timing of data collection (August 2022); access to mpox testing and other services likely changed as the US outbreak response evolved. For example, testing uptake could have been limited due to sustained increases of telehealth services after the COVID-19 pandemic, particularly among wealthier, insured patients such as the GBMSM who comprised the majority of our sample [29]. Furthermore, access to testing over the course of the outbreak likely varied by geographic location, something we were unable to fully assess in this analysis.

Conclusions

Similar structural barriers that have long been known to increase the risk of other STIs were rapidly replicated during the mpox outbreak in 2022. Sustained disparities fostered by racism, stigma, and limited access to health care in rural settings contribute to ongoing morbidity and mortality for GBMSM and predispose many communities to unexpected new outbreaks and epidemics such as mpox. While the epidemiology of mpox

XSL•FO
RenderX

continues to evolve, transmission risk remains present and characterizing optimal public health interventions for mpox is critical to both prevent future outbreaks and more rapidly respond if they occur.

Acknowledgments

This work was supported by the National Institute of Allergy and Infectious Disease (P30AI050409), National Institute of Mental Health (R01MH132150; R01MH110358; R25MH083620; P30MH136919), and National Institute of Nursing Research (R01NR020437).

Disclaimer

The results presented in this manuscript do not necessarily reflect the views of the National Institutes of Health or the US government.

Data Availability

The datasets analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

TS, KPD, and SDB developed the research questions and designed the study. KA conceived and conducted data analyses for this manuscript, and wrote the manuscript. All authors reviewed the manuscript and provided critical feedback.

Conflicts of Interest

None declared.

References

1. Mpox. Centers for Disease Control and Prevention (CDC). 2023. URL: https://www.cdc.gov/poxvirus/monkeypox/index.html [accessed 2024-12-12]
2. Kibungu EM, Vakaniaki EH, Kinganda-Lusamaki E, et al. Clade I-associated mpox cases associated with sexual contact, the Democratic Republic of the Congo. Emerg Infect Dis 2024 Jan;30(1):172-176. [doi: 10.3201/eid3001.231164] [Medline: 38019211]
3. Van Dijck C, Hoff NA, Mbala-Kingebeni P, et al. Emergence of mpox in the post-smallpox era-a narrative review on mpox epidemiology. Clin Microbiol Infect 2023 Dec;29(12):1487-1492. [doi: 10.1016/j.cmi.2023.08.008] [Medline: 37574113]
4. Minhaj FS, Ogale YP, Whitehill F, et al. Monkeypox outbreak - nine states, May 2022. MMWR 2022 Jun 10;71(23):764-769. [doi: 10.15585/mmwr.mm7123e1] [Medline: 35679181]
5. Kava CM, Rohraff DM, Wallace B, et al. Epidemiologic features of the monkeypox outbreak and the public health response - United States, May 17-October 6, 2022. MMWR Morb Mortal Wkly Rep 2022 Nov 11;71(45):1449-1456. [doi: 10.15585/mmwr.mm7145a4] [Medline: 36355615]
6. Alavian N, Mourad A, Woodhouse EW, et al. Disparities in mpox vaccination among priority populations during the 2022 outbreak. Open Forum Infect Dis 2023 Sep;10(9):fad434. [doi: 10.1093/ofid/ofad434] [Medline: 37662451]
7. Bien-Gund CH, Shaw PA, Agnew-Brune C, et al. HIV self-testing and risk behaviors among men who have sex with men in 23 US cities, 2017. JAMA Netw Open 2022 Dec 1;5(12):e2247540. [doi: 10.1001/jamanetworkopen.2022.47540] [Medline: 36534398]
8. Zimba R, Romo ML, Kulkarni SG, et al. Patterns of SARS-CoV-2 testing preferences in a national cohort in the United States: latent class analysis of a discrete choice experiment. JMIR Public Health Surveill 2021 Dec 30;7(12):e32846. [doi: 10.2196/32846] [Medline: 34793320]
9. Le Forestier JM, Page-Gould E, Chasteen A. Identity concealment may discourage health-seeking behaviors: evidence from sexual-minority men during the 2022 global mpox outbreak. Psychol Sci 2024 Feb;35(2):126-136. [doi: 10.1177/09567976231217416] [Medline: 38215021]
10. Hong C. Mpox on Reddit: a thematic analysis of online posts on mpox on a social media platform among key populations. J Urban Health 2023 Dec;100(6):1264-1273. [doi: 10.1007/s11524-023-00773-4] [Medline: 37580545]
11. Walsh-Buhi ML, Houghton RF, Valdez D, Walsh-Buhi ER. A theory-based assessment of mpox: findings from a nationally representative survey of U.S. adults. PLoS One 2024;19(3):e0299599. [doi: 10.1371/journal.pone.0299599] [Medline: 38489274]
12. Nelson KM, Pantalone DW, Gamarel KE, Carey MP, Simoni JM. Correlates of never testing for HIV among sexually active internet-recruited gay, bisexual, and other men who have sex with men in the United States. AIDS Patient Care STDS 2018 Jan;32(1):9-15. [doi: 10.1089/apc.2017.0244] [Medline: 29232170]

XSL•FO
RenderX

13. de Voux A, Bernstein KT, Kirkcaldy RD, Zlotorzynska M, Sanchez T. Self-reported extragenital chlamydia and gonorrhea testing in the past 12 months among men who have sex with men in the United States-American Men's Internet Survey, 2017. Sex Transm Dis 2019 Sep;46(9):563-570. [doi: 10.1097/OLQ.0000000000001032] [Medline: 31415039]
14. Copen CE, Delaney KP, Agnew-Brune C, et al. Modifications to sexual behaviors associated with mpox (monkeypox) virus transmission among persons presenting for mpox vaccination, Washington, DC, August–October, 2022. Sex Transm Dis 2024 Jan 1;51(1):54-60. [doi: 10.1097/OLQ.0000000000001889] [Medline: 37889944]
15. Sönmez İ, Martínez Riveros H, Folch C, et al. Egocentric sexual network analysis among gay and bisexual men who have sex with men with and without mpox infection. Sex Transm Infect 2023 Dec;99(8):541-547. [doi: 10.1136/sextrans-2023-055901] [Medline: 37704363]
16. Adams C, Kirby AE, Bias M, et al. Detecting mpox cases through wastewater surveillance — United States, August 2022–May 2023. MMWR Morb Mortal Wkly Rep 2024 Jan 18;73(2):37-43. [doi: 10.15585/mmwr.mm7302a3] [Medline: 38236784]
17. Wiatrek S, Zlotorzynska M, Rai R, Sullivan P, Sanchez T. The annual American Men's Internet Survey of behaviors of men who have sex with men in the United States: Key Indicators Report 2018. JMIR Public Health Surveill 2021 Mar 4;7(3):e21812. [doi: 10.2196/21812] [Medline: 33496669]
18. NCHS Urban-Rural Classification Scheme for counties. Centers for Disease Control and Prevention (CDC). 2014. URL: https://www.cdc.gov/nchs/data-analysis-tools/urban-rural.html#:~:text=A%20key%20feature%20of%20the,%E2%80%9D%20metro%20(akin%20to%20suburbs) [accessed 2025-01-06]
19. Carrico AW, Hunt PW, Neilands TB, et al. Stimulant use and viral suppression in the era of universal antiretroviral therapy. J Acquir Immune Defic Syndr 2019 Jan 1;80(1):89-93. [doi: 10.1097/QAI.0000000000001867] [Medline: 30272634]
20. Blizzard L, Hosmer DW. Parameter estimation and goodness-of-fit in log binomial regression. Biom J 2006 Feb;48(1):5-22. [doi: 10.1002/bimj.200410165] [Medline: 16544809]
21. New in Stata 15. StataCorp. 2017. URL: https://www.stata.com/stata15/ [accessed 2025-01-06]
22. Torres TS, Silva MST, Coutinho C, et al. Evaluation of mpox knowledge, stigma, and willingness to vaccinate for mpox: cross-sectional web-based survey among sexual and gender minorities. JMIR Public Health Surveill 2023 Jul 17;9:e46489. [doi: 10.2196/46489] [Medline: 37459174]
23. HHS monkeypox letter. Infectious Diseases Society of America. 2022 Jul 22. URL: https://www.idsociety.org/globalassets/bb-complex-pages/hhs-monkeypox-letter-july-2022-final.pdf [accessed 2025-01-06]
24. Nsoesie EO, Vu C. Monkeypox: where are the racial disparities? Lancet Reg Health Am 2023 Jan;17:100372. [doi: 10.1016/j.lana.2022.100372] [Medline: 36249074]
25. Owens C, Hubach RD. Rural-urban differences in monkeypox behaviors and attitudes among men who have sex with men in the United States. J Rural Health 2023 Mar;39(2):508-515. [doi: 10.1111/jrh.12726] [Medline: 36394371]
26. Lakhani HV, Pillai SS, Zehra M, Sharma I, Sodhi K. Systematic review of clinical insights into novel coronavirus (CoVID-19) pandemic: persisting challenges in U.S. rural population. Int J Environ Res Public Health 2020 Jun 15;17(12):4279. [doi: 10.3390/ijerph17124279] [Medline: 32549334]
27. Maan I, Kohli M, Gilson R. Mpox in people living with HIV. Curr Opin Infect Dis 2024 Feb 1;37(1):1-7. [doi: 10.1097/QCO.0000000000000994] [Medline: 38112084]
28. Mitjà O, Alemany A, Marks M, et al. Mpox in people with advanced HIV infection: a global case series. Lancet 2023 Mar 18;401(10380):939-949. [doi: 10.1016/S0140-6736(23)00273-8] [Medline: 36828001]
29. Wiefels MD, Gmunder KN, Ruiz JW. Patient income level and health insurance correlate with differences in health care utilization during the COVID-19 pandemic. J Public Health Res 2023 Jan;12(1):22799036231160624. [doi: 10.1177/22799036231160624] [Medline: 36911537]

Abbreviations

AMIS: American Men's Internet Survey
aPR: adjusted prevalence ratio
CAS: condomless anal sex
GBMSM: gay, bisexual, and other men who have sex with men
PR: prevalence ratio
PrEP: pre-exposure prophylaxis
STI: sexually transmitted infection

XSL•FO
RenderX

Edited by A Mavragani; submitted 15.02.24; peer-reviewed by M Xiridou, S Chen; revised version received 16.07.24; accepted 06.08.24; published 16.01.25.

Please cite as:
Atkins K, Carpino T, Rao A, Sanchez T, Edwards OW, Hannah M, Sullivan PS, Ogale YP, Abara WE, Delaney KP, Baral SD
Suspected Mpox Symptoms and Testing in Men Who Have Sex With Men in the United States: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e57399
URL: https://publichealth.jmir.org/2025/1/e57399
doi:10.2196/57399

© Kaitlyn Atkins, Thomas Carpino, Amrita Rao, Travis Sanchez, O Winslow Edwards, Marissa Hannah, Patrick S Sullivan, Yasmin P Ogale, Winston E Abara, Kevin P Delaney, Stefan D Baral. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effect of a Short, Animated Storytelling Video on Transphobia Among US Parents: Randomized Controlled Trial

Doron Amsalem[1], MD; Merlin Greuel[2], MD; Shuyan Liu[3,4], PhD; Andrés Martin[5,6], MD, PhD; Maya Adam[2,7,8], MD, PhD

[1]Department of Psychiatry, Columbia University, New York, NY, United States

[2]University Hospital, Heidelberg Institute of Global Health, Heidelberg University, Heidelberg, Germany

[3]Department of Psychiatry and Psychotherapy, Charité—Universitätsmedizin, Campus Charité Mitte, Berlin, Germany

[4]German Center for Mental Health, Deutsche Zentrum für Psychische Gesundheit, Berlin, Germany

[5]Child Study Center, Yale School of Medicine, Yale University, New Haven, CT, United States

[6]Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[7]Stanford Department of Pediatrics, Stanford University School of Medicine, Stanford University, 453 Quarry Rd, Palo Alto, CA, United States

[8]Stanford Center for Digital Health, Stanford University School of Medicine, Stanford University, Stanford, CA, United States

Corresponding Author:
Maya Adam, MD, PhD
University Hospital, Heidelberg Institute of Global Health, Heidelberg University, Heidelberg, Germany

Abstract

Background: Parents play a pivotal role in supporting transgender and gender diverse (TGD) youth. Yet only 35% of TGD youth describe their home as a gender-affirming place. Lack of parental support contributes to recent findings that TGD youth are approximately three times more likely to attempt suicide than their cisgender peers. In contrast, parents' affirmation of their children's gender identity significantly improves their mental health outcomes, by reducing anxiety, depression, and suicidality.

Objective: Addressing the urgent need for effective, scalable interventions, this study evaluates a novel digital approach: short, animated storytelling videos. We hypothesized that our 2.5-minute video intervention would reduce antitransgender stigma, or transphobia, and improve attitudes toward gender diverse children among US parents.

Methods: We recruited 1267 US parents, through the Prolific Academic (Prolific) online research platform, and randomized them into video intervention or control groups. We measured transphobia using the Transgender Stigma Scale, and attitudes toward transgender children using the gender thermometer, before and after watching the video. We compared outcomes between the two groups using 2 × 3 ANOVA. Both groups were invited to return 30 days later for follow-up assessment, before being offered posttrial access to the intervention video, which portrayed an authentic conversation between a mother and her transgender child.

Results: Single exposure to a short, animated story video significantly reduced transphobia and improved attitudes toward transgender children among US parents, immediately post intervention. We observed a significant group-by-time interaction in mean Transgender Stigma Scale scores ($F_{2,1}$=3.7, P=.02) and significant between-group changes when comparing the video and control groups from baseline to post intervention (F_1=27.4, P<.001). Effect sizes (Cohen d) indicated small to moderate immediate changes in response to the 2.5-minute video, though the effect was no longer observed at the 30-day follow-up. Gender thermometer scores revealed significant immediate improvements in the attitudes of participants in the video intervention arm, and this improvement was sustained at the 30-day time point.

Conclusions: Short, animated storytelling is a novel digital approach with the potential to boost support and affirmation of transgender children, by offering authentic insights into the lived experiences of TGD youth. Repeated exposures to such interventions may be necessary to sustain improvements over time. Future studies could test a series of short, animated storytelling videos featuring the lived experiences of several TGD youth. Evaluating the effect of such a series could contribute to the fields of digital health communication and transgender health. Digital approaches, such as short, animated storytelling videos, that support empathy and acceptance of TGD youth could foster a more inclusive society in which every child can thrive.

Trial Registration: AsPredicted.org 159248; https://aspredicted.org/ptmd-3kfs.pdf

(JMIR Public Health Surveill 2025;11:e66496) doi:10.2196/66496

XSL•FO
RenderX

KEYWORDS

public health communication; vulnerable population; stigma reduction; stigma; transphobia; transgender; gender diverse; LGBTQ; parent; mental health; mental illness; transgender children; children; youth; adolescent; storytelling; animation

Introduction

Optimal health outcomes for transgender and gender diverse (TGD) youth depend on the support and affirmation of their parents, as well as other adults with whom they interact [1-5]. Yet, even in their own homes, only 35% of TGD youth feel they are in a gender-affirming place [6]. Lack of parental support contributes to recent findings that TGD youth are 2.7 - 3.5 times more likely to attempt suicide than their cisgender peers [7,8]. Simple, gender-affirming acts, such as consistently respecting a child's pronouns in the home, can dramatically reduce the percent of TGD youth who attempt suicide from 21% to 12% as in a 2023 US national survey [6]. However, reducing stigma toward TGD youth, among adults, remains challenging [9,10]. Stigma acts as a broader social mechanism that contributes to transphobia (prejudice against individuals who defy traditional gender norms) [9,11]. Recent findings indicate that, even within their own homes, only 35% of TGD youth report feeling that they are in a gender-affirming place, with parental transphobia being a key contributing factor [6]. Lack of adult support and affirmation remains one of the leading causes of poor mental health outcomes in TGD youth [2-5,12,13].

This public health problem is aggravated by a growing climate of general societal intolerance in the United States. A record number of anti-LGBTQ (lesbian, gay, bisexual, transgender, queer) bills (more than 400) were introduced across US state legislatures in 2023 [14]. These environmental factors contribute to an alarmingly high rate of suicidality among TGD youth: the 2023 US National Survey on the Mental Health of LGBTQ+ Young People reported that 48% of transgender girls, 56% of transgender boys, and 48% of nonbinary youth had seriously considered attempting suicide in the past year [6]. Overall, TGD youth are 3.5 times more likely to attempt suicide than their cisgender peers [8]. Simple gender-affirming acts from adults, such as consistently respecting a child's pronouns, have been associated with a dramatic reduction in the percent of TGD youth who attempt suicide from 21% to 12% [6]. Even before the COVID-19 pandemic, a scoping review published in the *Journal of Adolescent Health* identified an urgent need for innovative interventions to reduce transphobia toward TGD youth [15]. Pandemic lockdowns, through a loss of community and school support, further threatened the mental health of TGD youth, who found themselves isolated and unsupported [16-18].

Research on transphobia reduction in schools highlights the need for interventions to reduce stigma among all parents, not just those of TGD youth [19]. Not only does affirmation from the general parent community positively impact TGD children directly, it also makes it easier for the parents of TGD children to adopt gender-affirming attitudes toward their children, when these prevail in the broader parent community [20,21]. Importantly, a parent's journey toward acceptance and support of their TGD child begins with them as a member of the general parent community [22]. Only once they have truly accepted the TGD identity of their child will they identify as "a parent of a TGD child." So, efforts to promote acceptance and reduce transphobia in the general parent population supports TGD youth in two ways: (1) a more accepting general parent population will model acceptance for the parents of TGD youth, and (2) a more accepting general parent population results in gender-affirming interactions with TGD youth in school or social settings, thereby creating a more accepting environment in which all children can thrive [19-22].

A significant body of research suggests that social contact-based interventions may be the most effective way to reduce stigma [23-26]. Grounded in intergroup contact theory, this research suggests that exposure to the stories and experiences of transgender individuals can effectively reduce transphobia toward them [27-31]. A known stumbling block of this approach is the selection bias that limits exposure to the stories of TGD youth, since prejudiced individuals generally avoid intergroup contact [30]. Disseminating the stories of TGD youth to the general public, on the platforms where they readily consume information, still presents a major challenge.

In the postpandemic era, social media has become a primary source of health information for the public [32]. A 2021 scoping review in the *Lancet Digital Health* described social media as a "crucial communication tool" for public health [33]. Social media has proven effective for reaching broader and more varied audiences, including hard-to-reach populations [32]. As such, social media presents a powerful, emerging pathway for scaling the stories of TGD youth to the public.

Short, animated storytelling videos, disseminated via social media, emerged as a powerful digital approach to engaging diverse audiences around the world, during the COVID-19 pandemic [34]. Harnessing the appeal of human stories, enhanced with compelling soundtracks, this novel digital approach is easily shareable and thus readily scalable via social media [34,35].

This study investigated the effect of a short, animated storytelling video on transphobia in a sample of US parents. Our hypothesis was that exposure to the authentic narratives of TGD youth, animated to protect their identities, could effectively reduce transphobia and improve attitudes toward gender diverse children, thereby boosting affirmation and support for TGD youth, among parents in the general public.

Methods

Participants and Recruitment Procedure

From January to February 2024, we enrolled parents from the general US population using Prolific Academic, a crowdsourcing platform frequently used to recruit participants in online psychology and medical research [36]. Prolific supports high-quality data collection by ensuring consistency in demographic responses over time, screening for bots, blocking participants who hide their location, and assigning anonymous unique participant IDs [37,38]. For this trial, we recruited

XSL•FO
RenderX

English-speaking parents, aged 18 - 50 years, with at least one child, who were residing in the United States. Participants were recruited through the Prolific Academic platform. Before the initiation of this study, all participants reviewed a detailed information sheet, outlining the rationale for this study as well as all study procedures, and potential risks and benefits of taking part in this study. Those who agreed to participate were directed to complete this study's procedures on Qualtrics, a secure online data-collection platform.

Ethical Considerations

This study, which involved human participants, was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and was approved by the Stanford University internal review board (IRB-#72761) on January 21, 2024. We followed the CONSORT (Consolidated Standards of Reporting Trials) guidelines [39], and this study was registered (#159248) with AsPredicted, a clinical trial registry created in 2015 by the Wharton School at the University of Pennsylvania [40].

Informed consent was obtained from all participants prior to enrollment in this study. Participants were provided with an information sheet that contained a detailed description of this study's purpose, procedures, potential risks, and benefits. The information sheet was reviewed and approved by the Stanford IRB. Participants were informed that participation was voluntary and that they could withdraw at any time without penalty. The consent process was facilitated by Prolific Academic, the academic recruitment platform through which participants were recruited. Participants were paid US $2.40 for participation in this short study, at 2 time points and in accordance with the payment recommendations of Prolific Academic. Participants were assured that their responses would remain confidential and that any identifying data would not be shared with the research team by Prolific. The informed consent process was reviewed and approved by the Stanford IRB (#72761) ensuring compliance with ethical research standards.

Randomization and Study Design

After consenting to take part in this study, participants were randomly assigned to one of two groups: short, animated storytelling video intervention group or a control group that received a simple fact sheet about TGD youth. We used a baseline survey to collect demographic information and we administered validated questionnaires to assess transphobia and attitudes toward TGD youth. We conducted the first round of surveys immediately post exposure to the short, animated storytelling intervention. Follow-up assessments, using the same surveys, were conducted 30 days later. To ensure accuracy and validity of the results, we included attention checks (eg, "For this question, please select the answer 'Neither agree nor disagree.'") and excluded participants who failed these. We also used a timer to confirm that participants in the control group read the TGD fact sheet and that those in the intervention group viewed the short, animated storytelling video before proceeding to the next screen. Each participant received US $2.40 for completing this study.

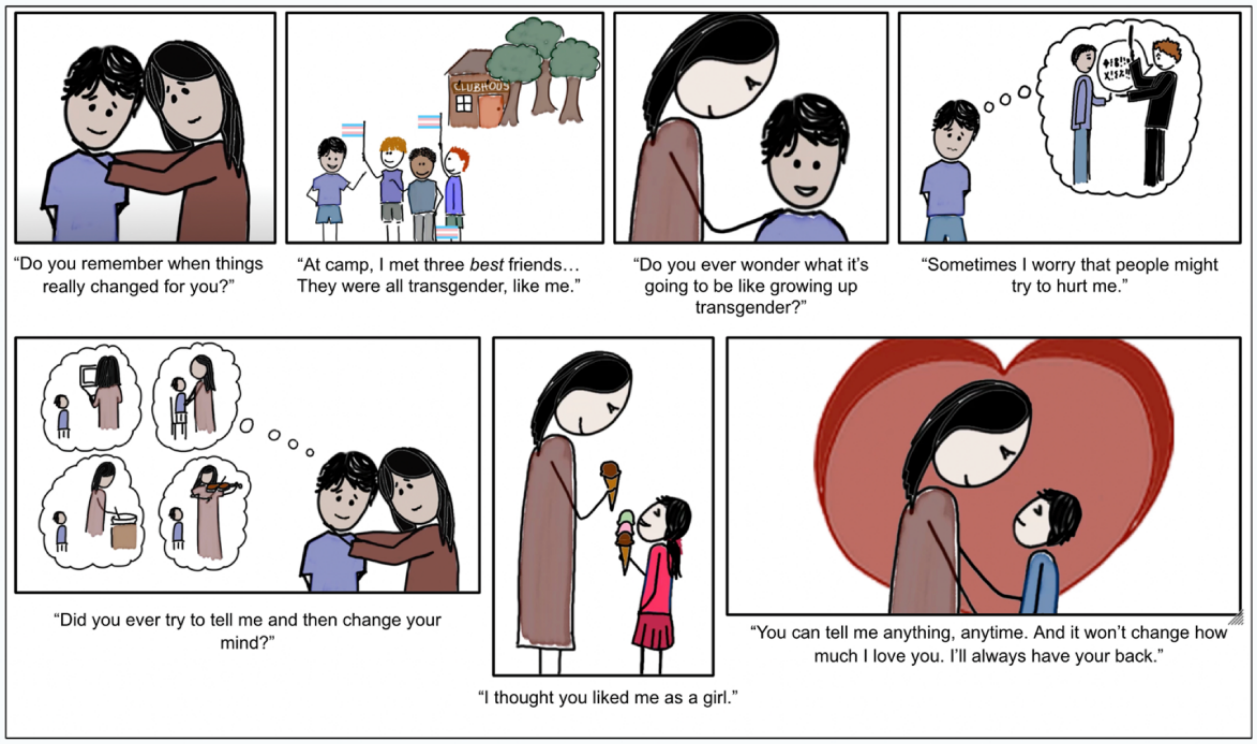
Intervention

Our novel, digital intervention used a short, animated storytelling video to capture an authentic prerecorded conversation between a mother and her transgender child. We chose this video because it follows the principles of social contact-based interventions, which foster identification and emotional engagement. A significant body of research suggests that these principles have proven effective in reducing stigma and transphobia by exposing viewers to the real-life stories and experiences of marginalized individuals [23-30,41]. In the conversation featured in our intervention, a mother asks her son to recall experiences related to his gender identity, including his early hesitations to confide in her that he was transgender. The child then asks his mother about her worries, and she shares these candidly. The identities of the speakers are concealed at all times, and we received written permission to use the audio recording from the speakers. The scenes and characters used to represent the speakers were drawn by hand on a Wacom tablet, by our coauthor (MA), and edited by the Educational Technology Group at Stanford Medicine. Figure 1 shows selected scenes and quotations from the intervention video.

The short, animated video tested in this trial can be previewed on YouTube [42].

XSL•FO
RenderX

Figure 1. Selected scenes and quotations from the short, animated storytelling intervention video shown to reduce transphobia in US parents in this online, randomized controlled trial.



Instruments

The primary outcome was transphobia toward TGD youth, which was measured using the total score of the Transgender Stigma Scale (TSS) developed by Madera et al [43]. We adjusted the wording of the scale items to assess attitudes toward transgender children (eg, "Boys who act like girls should be ashamed of themselves," and "Children should play with toys appropriate to their own sex"). The TSS is scored along a 5-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (5). Higher scores indicate greater transphobia. The TSS is highly reliable and has a Cronbach α of 0.89 [43].

As secondary outcome, we measured attitudes toward TGN children, using the "gender thermometer," a tool we previously developed to assess attitudes regarding sexual orientation and gender diversity [27]. The thermometer provides the following prompt: "Using a scale from zero to 100, please tell us about your personal feelings toward transgender boys or girls. As you do this task, think of an imaginary thermometer. The warmer or more favorable you feel toward the group, the higher the number you should give it. The colder or less favorable you feel, the lower the number. If you feel neither warm nor cold toward the group, rate it 50." We asked respondents about their attitudes toward (1) boys or girls and (2) transgender boys or girls. Higher ratings indicate more positive attitudes toward the group and lower ratings indicate less positive attitudes. Researchers in the fields of political science and psychology have frequently used feeling thermometers to assess attitudes toward diverse groups, including sexual minorities [44].

Analysis

The main outcome measures were changes in total TSS scores and gender thermometer scores, from baseline to post intervention and from baseline to 30-day follow-up. We calculated our sample size based on our previous studies [27,45-48]. We used Pearson chi-square and one-way ANOVA tests to compare sociodemographic characteristics between the short, animated storytelling video intervention and control groups. We used repeated-measures ANOVA tests to compare the mean transphobia score between the groups at 3 time points. Next, we used a one-way ANOVA test to compare changes from baseline scores, at post intervention, and at 30-day follow-up. We used independent *t* tests (2-tailed) to compare gender differences at baseline. Data were analyzed using SPSS (version 29.0; IBM Corp) [49].

Results

Sample Characteristics

We recruited 1267 US parents, through the Prolific Academic online research platform. In total, 1177 parents completed the preintervention assessment after we excluded 40 (3%) who failed attention tests and 50 (4%) who left this study. Of these, 1159 (98%) completed the postintervention assessment and 976 (83%) completed the 30-day follow-up assessment. Figure 2 shows the design and flow of this study.

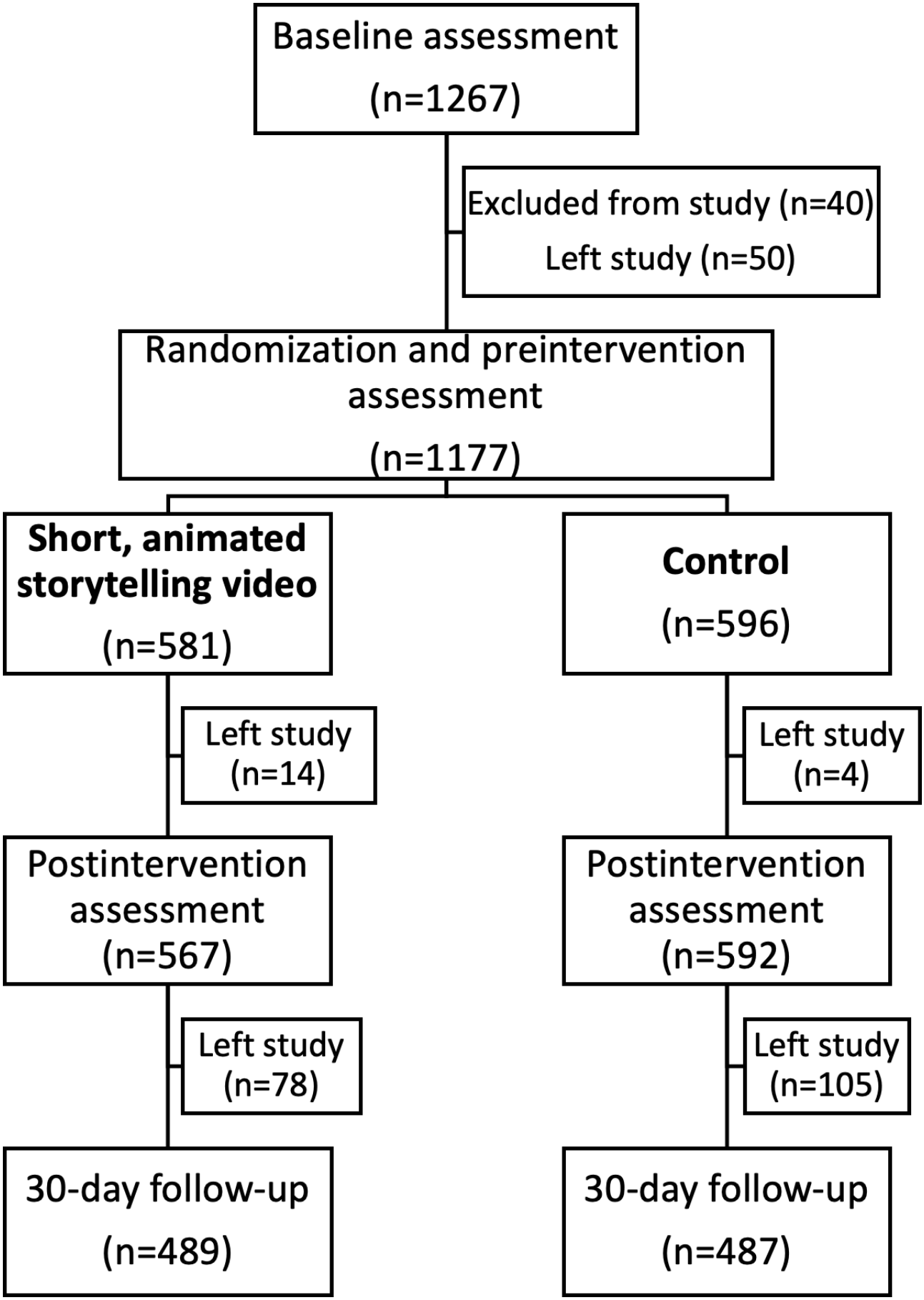
Sociodemographic characteristics and completion rates did not differ significantly between groups (Table 1), nor did baseline characteristics differ between follow-up completers and noncompleters. Mean participant age was 39 (SD 6.9; range 20 - 50) years. More than half of participants self-identified as

XSL•FO
RenderX

female (n=653, 56%), 510 (43%) male, and 14 (1%) transgender or nonbinary. Overall, 110 (9%) participants self-identified as Hispanic, 192 (16%) non-Hispanic Black, 789 (67%) non-Hispanic White, 57 (5%) non-Hispanic Asian, 8 (1%) non-Hispanic Native American, and 18 (2%) other (Table 1).

XSL•FO
RenderX

Figure 2. Study design and flow for our online, randomized controlled trial testing the effect of a short, animated storytelling video intervention on transphobia in US adults.



XSL•FO
RenderX

Table . Demographic characteristics of study participants.

Items		Brief animation (n=581)	Narrative (n=596)	Total (N=1177)	Statistic	
					Pearson χ^2 (*df*)	*P* value
Age (years), mean (SD)		39.1 (7.1)	39 (6.8)	39 (6.9)	0.18 (*1175*)[a]	.86
Gender, n (%)					2.62(*2*)	.27
	Female	336 (58)	317 (53)	653 (56)		
	Male	238 (41)	272 (46)	510 (43)		
	Transgender or nonbinary	7 (1)	7 (1)	14 (1)		
	Prefer not to answer	0 (0)	0 (0)	0 (0)		
Race and ethnicity, n (%)					9.35 (*6*)	.16
	Hispanic	62 (11)	48 (8)	110 (9)		
	Non-Hispanic Asian	32 (6)	25 (4)	57 (5)		
	Non-Hispanic Black	97 (17)	95 (16)	192 (16)		
	Non-Hispanic Native American	3 (1)	5 (1)	8 (1)		
	Non-Hispanic White	378 (65)	411 (69)	789 (66)		
	Non-Hispanic other[b]	6 (1)	12 (2)	18 (2)		
	Prefer not to answer	3 (1)	0 (0)	3 (0)		

[a]Independent *t* tests.

[b]Non-Hispanic other: multiracial (n=15), Native Hawaiian (n=1), Jewish (n=1), and South Asian (n=1).

Intervention Effects

We observed significant differences in outcomes between this study's groups. Figure 3 presents mean TSS scores of this study arms at 3 time points, showing that the control arm changed minimally across time points, in contrast to the short, animated storytelling video intervention arm, which showed a strongly significant decrease in transphobia immediately post intervention. This effect was no longer evident at 30 days post intervention. A 2 × 3 group-by-time ANOVA showed a significant group-by-time interaction in mean TSS scores ($F_{2,1}$=3.7, *P*=.02). One-way ANOVAs revealed significant between-group changes between the short, animated storytelling video and control groups from baseline to post intervention (F_1=27.4, *P*<.001), but not from baseline to the 30-day follow-up.

Figure 4 presents mean gender thermometer scores, by study arm, over time. Figure 4A shows changing attitudes toward transgender boys or girls. Attitudes of participants in the short, animated storytelling video intervention arm improved significantly over time and this improvement was sustained at the 30-day time point. In contrast, we observed a slight improvement in the control arm, which was not significant, and showed a subsequent rebound at the 30-day follow-up.

We did not find a 2 × 3 group-by-time interaction. However, independent *t* tests showed a significant difference between the changes from baseline to post intervention between short, animated storytelling video intervention and control groups (change of 2.2, SD 9.7 vs 0.8, SD 6.5; *t*=3, t_{df}=1151 *P*=.003). We did not observe an effect of the intervention on attitudes toward cisgender boys or girls (Figure 4B).

XSL•FO
RenderX

Figure 3. Mean scores on the TSS between short, animated storytelling video intervention (n=588) and control (n=598) groups at 3 time points in our online RCT measuring the effect of a short, animated storytelling intervention on transphobia in US adults. Follow-up surveys completed after 30 days; higher scores indicate greater transphobia; TSS (range 10 - 50). Repeated measure ANOVA (*P*=.02). RCT: randomized controlled trial; TSS: Transgender Stigma Scale.

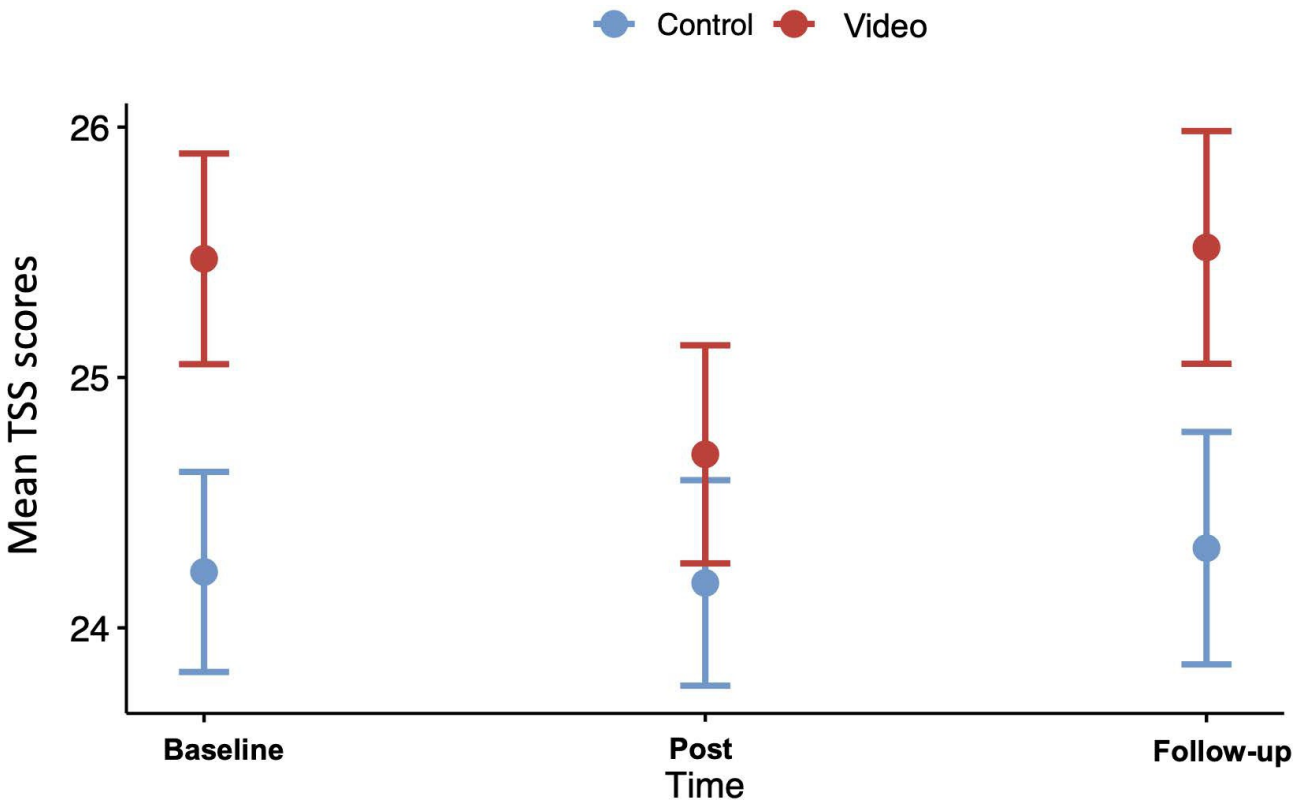
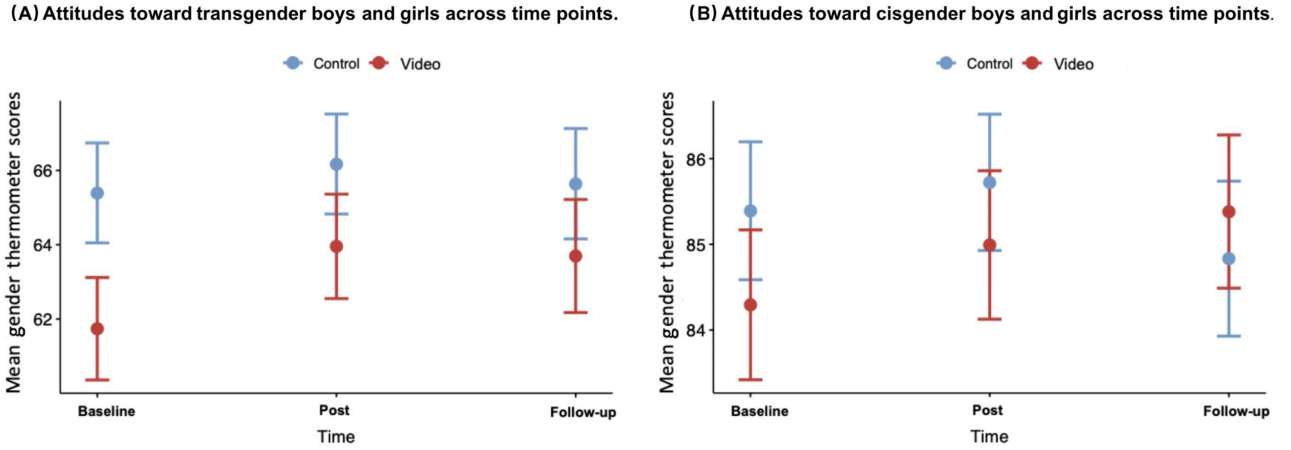


Figure 4. Gender thermometer scores toward (A) transgender children and (B) cisgender children, at 3 time points in our online RCT measuring the effect of a short, animated storytelling intervention on transphobia in US adults. Follow-up surveys completed after 30 days; higher scores indicate more favorable attitudes; gender thermometer (range 0 - 100). RCT: randomized controlled trial.

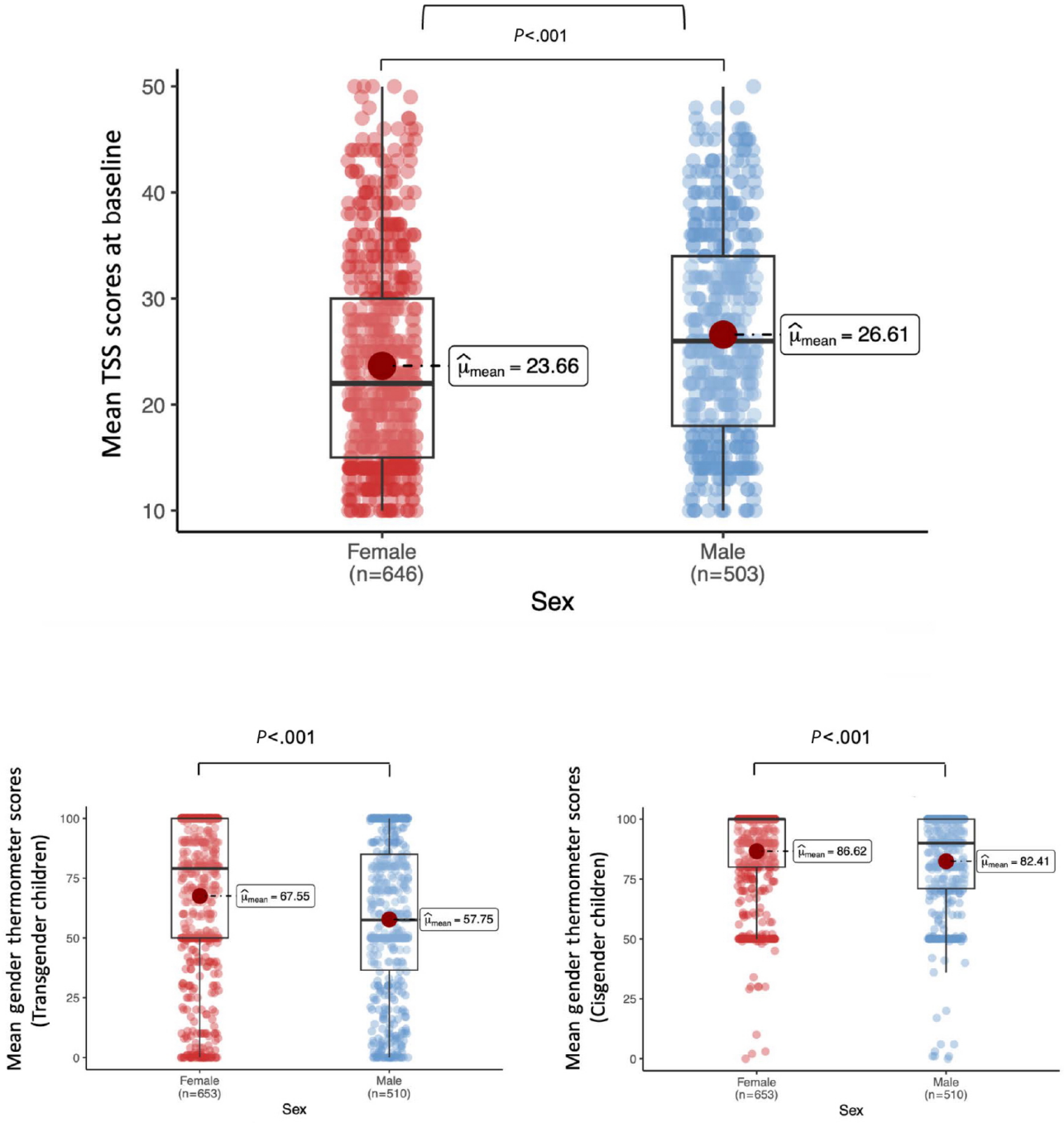


Participant Gender Differences

Figure 5 presents baseline differences between men and women on the TSS (26.6, SD 9.7 vs 23.7, SD 9.8; independent *t* tests: *t*=5.1, *t*df=1161, *P*<.001) and on the gender thermometer both for transgender boys or girls (57.7, SD 32 vs 67.5, SD 33.1, *t*=5.1, *t*df=1161, *P*<.001) and cisgender boys or girls (82.4.7, SD 21 vs 86.6, SD 19.7, *t*=3.5, *t*df=1161 *P*<.001). The observed baseline differences between men and women remained constant across time points in both study groups.

XSL•FO
RenderX

Figure 5. Gender differences (TSS and gender thermometer) at baseline in our online RCT measuring the effect of a short, animated storytelling intervention on transphobia in US adults. Higher scores indicate greater transphobia; TSS (range 10 - 50). Higher scores indicate more positive attitudes; GT (range 0 - 100). GT: gender thermometer; RCT: randomized controlled trial; TSS: Transgender Stigma Scale.



Discussion

In this study, we conducted a large-scale, randomized, controlled experiment to document the potential of a novel, digital approach—short, animated storytelling—to reduce transphobia and improve attitudes toward TGD youth among US parents. Within our study population, we observed significantly reduced transphobia and significantly improved attitudes toward TGD youth, immediately after viewing the 2.5-minute short, animated storytelling video intervention. The effect of this short video was no longer evident one month after the intervention, suggesting that repeated, short exposures to similar short, animated storytelling video content may be necessary to sustain transphobia reductions and improvements in attitudes over time. One potential approach is to develop and disseminate a series of short, animated storytelling videos featuring the authentic lived experiences of several TGD young people. This scalable, digital approach could be implemented through a smartphone-based ecological momentary intervention. Evaluating the effect of such a series could be a meaningful future research direction in the fields of both digital health communication and transgender health.

During the COVID-19 pandemic, our coinvestigator (MA) developed the short, animated storytelling approach, a novel digital modality for scaling health information worldwide. Early short, animated storytelling COVID prevention videos reached

XSL•FO
RenderX

more than 15 million people via social media within 4 months [35]. High voluntary public engagement with short, animated storytelling content was also documented in randomized, controlled trials [50,51]. As short, animated storytelling videos have proven effective for engaging diverse adult audiences, they present a potentially promising new method for engaging the public in the stories of TGD youth. The findings of this study support the potential for using short, animated storytelling to reduce transphobia and elicit empathy. Animation also protects the identities of the TGD storytellers—an important consideration, since many of these young people still live in "stealth" [52].

Intergroup contact theory, and recent research grounded on this theory, support the stigma-reducing potential of social contact with TGD individuals [27,28,30,31,53,54]. To date, interventions designed to reduce transphobia have been difficult to scale broadly to the public because (1) they involve face-to-face or online synchronous interactions, (2) they involve a significant time commitment and some form of active "opting in," thereby self-selecting for people who already have low stigma, or (3) they lack the production value needed to make them highly engaging—one of the keys to scalability [55]. The short, animated storytelling video tested in this study packages a theory-driven approach to stigma reduction (the sharing of human stories) in a readily scalable form of health message (short, animated videos that can be shared via social media) to overcome a critical health challenge (lack of support and affirmation for TGD youth). Anecdotally, and supporting the potential for easy scalability, one participant in our study sent a message to the research team:

> *As the father of a child with 2 trans friends, this video is amazing and I showed it to my daughter, who is 8 and she completely loved it. She asked if we could send it to her friends' parents, who called us and thanked us. So I wanted to let you know this was helpful and well thought. I can't thank you enough!*

In this study, we also observed significant differences in baseline transphobia between men and women. Aligned with prior research, men in our study scored significantly higher on transphobia than did women, and these differences persisted across time points [11,56]. Other researchers have proposed that threatening gender norms can trigger a perceived loss of social status among men more than women [57]. Men may also be more likely to feel that transgender individuals pose a threat to their own gender identity and masculinity [58]. Future interventions could include story-based content that is tailored toward men. Featuring the stories of fathers, who are supporting of and affirming to their transgender children, may be particularly important. We also note that, to date, much of the research on family support of TGD youth focuses on parents, but there is an emerging body of research on the attitudes of other family members (eg, grandparents and siblings) toward transgender youth. If parents are not supportive, these other family members may be able to play a vital role in protecting transgender youth and the approach described here could also work with a wider range of target audiences.

A limitation of this study is the fact that our sample was recruited via an online research platform, suggesting that the sample may not be entirely representative of the general population in the United States. However, considering the intended, digital dissemination route for our intervention (social media), we feel this population is similar enough to our target population to allow us to gain some meaningful insights and draw relevant conclusions.

Our findings highlight the important role of authentic storytelling, packaged in scalable, accessible digital formats, for reducing transphobia toward TGN youth. These results also add to the growing field of research on short, animated storytelling, an innovative digital approach to rapidly scaling health information worldwide. Finally, these findings underscore the potential for light-touch, scalable video content to improve attitudes toward TGN youth within our society. Simply put, adults can be critical determinants of better health outcomes for TGD youth, simply by affirming their right to live and express themselves. Using authentic storytelling to engage and elicit empathy, and social media to broadly reach adults across demographics, we have the potential to catalyze a shift toward affirmation and support within society—a shift that could help children of all gender identities to survive and thrive.

Acknowledgments

This study was funded by the Department of Psychiatry, Columbia University Vagelos College of Physicians & Surgeons and the Stanford Center for Digital Health. The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this paper. We wish to thank the parent and child, who courageously shared their story in the intervention video, with the hope of contributing to a more accepting environment for gender diverse children.

Data Availability

The datasets collected and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

DA and MA led the design, implementation, data collection, and writing of this paper. MG contributed to the design, interpretation of results, and writing of this paper. DA and SL led the analysis, interpretation, and presentation of the results. AM contributed to the design, interpretation, presentation of results, and editing of this paper.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH checklist (V 1.6.1).
[PDF File, 798 KB - publichealth_v11i1e66496_app1.pdf]

References

1. Ghassabian A, Suleri A, Blok E, Franch B, Hillegers MHJ, White T. Adolescent gender diversity: sociodemographic correlates and mental health outcomes in the general population. J Child Psychol Psychiatry 2022 Nov;63(11):1415-1422. [doi: 10.1111/jcpp.13588] [Medline: 35147218]
2. Malpas J, Pellicane MJ, Glaeser E. Family-based interventions with transgender and gender expansive youth: systematic review and best practice recommendations. Transgend Health 2022 Feb;7(1):7-29. [doi: 10.1089/trgh.2020.0165] [Medline: 36644030]
3. Olezeski CL, Kamody RC. Parents matter: considering evidence-based relationship variables within families when working with gender expansive youth. Pract Innov (Wash D C) 2020;5(3):218-229. [doi: 10.1037/pri0000131]
4. Westwater JJ, Riley EA, Peterson GM. What about the family in youth gender diversity? A literature review. Int J Transgend 2019;20(4):351-370. [doi: 10.1080/15532739.2019.1652130] [Medline: 32999622]
5. Kidd KM, Didden E, Harman H, et al. Parents of gender diverse youth: support sought, received, and still needed. J Adolesc Health 2024 Jun;74(6):1078-1087. [doi: 10.1016/j.jadohealth.2023.08.001] [Medline: 37715763]
6. 2023 U.S. National Survey on the Mental Health of LGBTQ+ Young People. The Trevor Project. 2023. URL: https://www.thetrevorproject.org/survey-2023/ [accessed 2024-12-18]
7. Jackman KB, Caceres BA, Kreuze EJ, Bockting WO. Suicidality among gender minority youth: analysis of 2017 youth risk behavior survey data. Arch Suicide Res 2021;25(2):208-223. [doi: 10.1080/13811118.2019.1678539] [Medline: 31642392]
8. Toomey RB, Syvertsen AK, Shramko M. Transgender adolescent suicide behavior. Pediatrics 2018 Oct;142(4):e20174218. [doi: 10.1542/peds.2017-4218] [Medline: 30206149]
9. Hatzenbuehler ML, Pachankis JE. Stigma and minority stress as social determinants of health among lesbian, gay, bisexual, and transgender youth: research evidence and clinical implications. Pediatr Clin North Am 2016 Dec;63(6):985-997. [doi: 10.1016/j.pcl.2016.07.003] [Medline: 27865340]
10. Wittlin NM, Kuper LE, Olson KR. Mental health of transgender and gender diverse youth. Annu Rev Clin Psychol 2023 May 9;19:207-232. [doi: 10.1146/annurev-clinpsy-072220-020326] [Medline: 36608332]
11. Nagoshi JL, Adams KA, Terrell HK, Hill ED, Brzuzy S, Nagoshi CT. Gender differences in correlates of homophobia and transphobia. Sex Roles 2008 Oct;59(7-8):521-531. [doi: 10.1007/s11199-008-9458-7]
12. Rivas-Koehl M, Valido A, Espelage DL, et al. Understanding protective factors for suicidality and depression among U.S. sexual and gender minority adolescents: implications for school psychologists. Sch Psych Rev 2022 May 4;51(3):290-303. [doi: 10.1080/2372966X.2021.1881411]
13. Katz-Wise SL, Sarda V, Line EC, et al. Longitudinal family functioning and mental health in transgender and nonbinary youth and their families. J Child Fam Stud 2024 Apr;33(4):1321-1335. [doi: 10.1007/s10826-023-02738-2] [Medline: 39220671]
14. Choi A. Record number of anti-LGBTQ bills were introduced in 2023. CNN Online.: CNN; 2024. URL: https://edition.cnn.com/politics/anti-lgbtq-plus-state-bill-rights-dg/index.html [accessed 2025-01-14]
15. Parker CM, Hirsch JS, Philbin MM, Parker RG. The urgent need for research and interventions to address family-based stigma and discrimination against lesbian, gay, bisexual, transgender, and queer youth. J Adolesc Health 2018 Oct;63(4):383-393. [doi: 10.1016/j.jadohealth.2018.05.018] [Medline: 30146436]
16. Hawke LD, Hayes E, Darnay K, Henderson J. Mental health among transgender and gender diverse youth: an exploration of effects during the COVID-19 pandemic. Psychol Sex Orientat Gend Divers 2021;8(2):180-187. [doi: 10.1037/sgd0000467]
17. Jones BA, Bowe M, McNamara N, Guerin E, Carter T. Exploring the mental health experiences of young trans and gender diverse people during the COVID-19 pandemic. Int J Transgend Health 2023;24(3):292-304. [doi: 10.1080/26895269.2021.1890301] [Medline: 37519915]
18. Kidd JD, Jackman KB, Barucco R, et al. Understanding the impact of the COVID-19 pandemic on the mental health of transgender and gender nonbinary individuals engaged in a longitudinal cohort study. J Homosex 2021 Mar 21;68(4):592-611. [doi: 10.1080/00918369.2020.1868185] [Medline: 33502286]
19. Bartholomaeus C, Riggs DW. Whole-of-school approaches to supporting transgender students, staff, and parents. Int J Transgenderism 2017;18:361-366. [doi: 10.1080/15532739.2017.1355648]
20. Mangin MM. Transgender Students in Elementary School: Creating an Affirming and Inclusive School Culture: Harvard Education Press Cambridge; 2020.

XSL•FO
RenderX

21. Mariotto M, Albanesi C, Lorusso MM. Encountering transgender and gender-expansive children in school: exploring parents' and teachers' practices through the lens of cisgenderism and adultism. Child Youth Serv Rev 2024 Sep;164:107819. [doi: 10.1016/j.childyouth.2024.107819]
22. Sansfaçon AP, Kirichenko V, Holmes C, et al. Parents' journeys to acceptance and support of gender-diverse and trans children and youth. J Fam Issues 2020 Aug;41(8):1214-1236. [doi: 10.1177/0192513X19888779]
23. Jankowski SE, Yanos P, Dixon LB, Amsalem D. Reducing public stigma towards psychosis: a conceptual framework for understanding the effects of social contact based brief video interventions. Schizophr Bull 2023 Jan 3;49(1):99-107. [doi: 10.1093/schbul/sbac143] [Medline: 36190348]
24. Rani N, Samuel AA. Reducing transphobia: comparing the efficacy of direct and indirect contact. Industrial and Commercial Training 2019 Oct 7;51(7/8):445-460. [doi: 10.1108/ICT-12-2018-0102]
25. Reinke RR, Corrigan PW, Leonhard C, Lundin RK, Kubiak MA. Examining two aspects of contact on the stigma of mental illness. J Soc Clin Psychol 2004 Jun;23(3):377-389. [doi: 10.1521/jscp.23.3.377.35457]
26. Thornicroft G, Mehta N, Clement S, et al. Evidence for effective interventions to reduce mental-health-related stigma and discrimination. The Lancet 2016 Mar;387(10023):1123-1132. [doi: 10.1016/S0140-6736(15)00298-6]
27. Amsalem D, Halloran J, Penque B, Celentano J, Martin A. Effect of a brief social contact video on transphobia and depression-related stigma among adolescents: a randomized clinical trial. JAMA Netw Open 2022 Feb 1;5(2):e220376. [doi: 10.1001/jamanetworkopen.2022.0376] [Medline: 35212749]
28. Boccanfuso E, White FA, Maunder RD. Reducing transgender stigma via an e-contact intervention. Sex Roles 2021 Mar;84(5-6):326-336. [doi: 10.1007/s11199-020-01171-9]
29. Hughto JMW, Reisner SL, Pachankis JE. Transgender stigma and health: a critical review of stigma determinants, mechanisms, and interventions. Soc Sci Med 2015 Dec;147:222-231. [doi: 10.1016/j.socscimed.2015.11.010] [Medline: 26599625]
30. Pettigrew TF. Intergroup contact theory. Annu Rev Psychol 1998;49:65-85. [doi: 10.1146/annurev.psych.49.1.65] [Medline: 15012467]
31. Walch SE, Ngamake ST, Francisco J, Stitt RL, Shingler KA. The attitudes toward transgendered individuals scale: psychometric properties. Arch Sex Behav 2012 Oct;41(5):1283-1291. [doi: 10.1007/s10508-012-9995-6] [Medline: 22810996]
32. Chen J, Wang Y. Social media use for health purposes: systematic review. J Med Internet Res 2021 May 12;23(5):e17917. [doi: 10.2196/17917] [Medline: 33978589]
33. Tsao SF, Chen H, Tisseverasinghe T, Yang Y, Li L, Butt ZA. What social media told us in the time of COVID-19: a scoping review. Lancet Digit Health 2021 Mar;3(3):e175-e194. [doi: 10.1016/S2589-7500(20)30315-0] [Medline: 33518503]
34. Adam M, Bärnighausen T, McMahon SA. Design for extreme scalability: a wordless, globally scalable COVID-19 prevention animation for rapid public health communication. J Glob Health 2020 Jun;10(1):010343. [doi: 10.7189/jogh.10.010343] [Medline: 32411360]
35. Vandormael A, Adam M, Greuel M, et al. The effect of a wordless, animated, social media video intervention on COVID-19 prevention: online randomized controlled trial. JMIR Public Health Surveill 2021 Jul 27;7(7):e29060. [doi: 10.2196/29060] [Medline: 34174778]
36. Peer E, Brandimarte L, Samat S, Acquisti A. Beyond the Turk: alternative platforms for crowdsourcing behavioral research. J Exp Soc Psychol 2017 May;70:153-163. [doi: 10.1016/j.jesp.2017.01.006]
37. Douglas BD, Ewell PJ, Brauer M. Data quality in online human-subjects research: comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA. PLoS One 2023;18(3):e0279720. [doi: 10.1371/journal.pone.0279720] [Medline: 36917576]
38. Palan S, Schitter C. Prolific.ac—a subject pool for online experiments. J Behav Exp Finance 2018 Mar;17:22-27. [doi: 10.1016/j.jbef.2017.12.004]
39. Moher D. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials (Chinese version). J Chin Integr Med 2010 Aug 15;340:701-741. [doi: 10.3736/jcim20100801]
40. Olevska A, Bert B, Ebrahimi L, Schonfelder G, Heinl C. Ensuring reproducible research requires a support infrastructure: the value of public registries to publishers. Sci Editor 2021 Apr 23;44:4-7. [doi: 10.36591/SE-D-4401-4]
41. Walch SE, Sinkkanen KA, Swain EM, Francisco J, Breaux CA, Sjoberg MD. Using intergroup contact theory to reduce stigma against transgender individuals: impact of a transgender speaker panel presentation. J Appl Soc Psychol 2012 Oct;42(10):2583-2605 [FREE Full text] [doi: 10.1111/j.1559-1816.2012.00955.x]
42. You amaze me every day. Maya Adam YouTube Page. 2017 Mar 26. URL: https://www.youtube.com/watch?v=aDihlx776vM [accessed 2024-12-18]
43. Madera SL, Díaz NV, Padilla M, et al. "Just like any other patient": transgender stigma among physicians in Puerto Rico. J Health Care Poor Underserved 2019;30(4):1518-1542. [doi: 10.1353/hpu.2019.0089] [Medline: 31680112]
44. Norton AT, Herek GM. Heterosexuals' attitudes toward transgender people: findings from a national probability sample of U.S. adults. Sex Roles 2013 Jun;68(11-12):738-753. [doi: 10.1007/s11199-011-0110-6]

XSL•FO
RenderX

45. Amsalem D, Jankowski SE, Pagdon S, et al. Selfie videos to reduce stigma and increase treatment seeking among youths: two noninferiority randomized controlled trials. Psychiatr Serv 2023 Mar 1;74(3):229-236. [doi: 10.1176/appi.ps.20220168] [Medline: 36254455]
46. Amsalem D, Markowitz JC, Jankowski SE, et al. Sustained effect of a brief video in reducing public stigma toward individuals with psychosis: a randomized controlled trial of young adults. Am J Psychiatry 2021 Jul;178(7):635-642. [doi: 10.1176/appi.ajp.2020.20091293] [Medline: 33900809]
47. Amsalem D, Martin A. Reducing depression-related stigma and increasing treatment seeking among adolescents: randomized controlled trial of a brief video intervention. J Child Psychol Psychiatry 2022 Feb;63(2):210-217. [doi: 10.1111/jcpp.13427] [Medline: 33821507]
48. Amsalem D, Yang LH, Jankowski S, Lieff SA, Markowitz JC, Dixon LB. Reducing stigma toward individuals with schizophrenia using a brief video: a randomized controlled trial of young adults. Schizophr Bull 2021 Jan 23;47(1):7-14. [doi: 10.1093/schbul/sbaa114] [Medline: 33484269]
49. Pallant J. SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS: Routledge; 2020. [doi: 10.4324/9781003117445]
50. Favaretti C, Adam M, Greuel M, et al. Participant engagement with a short, wordless, animated video on COVID-19 prevention: a multi-site randomized trial. Health Promot Int 2023 Aug 1;38(4):daab179. [doi: 10.1093/heapro/daab179] [Medline: 35137068]
51. Favaretti C, Vandormael A, Hachaturyan V, et al. Participant engagement and reactance to a short, animated video about added sugars: web-based randomized controlled trial. JMIR Public Health Surveill 2022 Jan 24;8(1):e29669. [doi: 10.2196/29669] [Medline: 35072639]
52. Olson KR, Durwood L, DeMeules M, McLaughlin KA. Mental health of transgender children who are supported in their identities. Pediatrics 2016 Mar;137(3):e20153223. [doi: 10.1542/peds.2015-3223] [Medline: 26921285]
53. Cipolletta S, Votadoro R, Faccio E. Online support for transgender people: an analysis of forums and social networks. Health Soc Care Community 2017 Sep;25(5):1542-1551. [doi: 10.1111/hsc.12448] [Medline: 28329916]
54. Massey ZB, Wong NCH, Barbati JL. Meeting the (trans)parent: test of parasocial contact with transgender characters on reducing stigma toward transgender people. Commun Stud 2021 Mar 4;72(2):232-250. [doi: 10.1080/10510974.2021.1876125]
55. Guadagno RE, Rempala DM, Murphy S, Okdie BM. What makes a video go viral? An analysis of emotional contagion and internet memes. Comput Hum Behav 2013 Nov;29(6):2312-2319. [doi: 10.1016/j.chb.2013.04.016]
56. Hatch HA, Warner RH, Broussard KA, Harton HC. Predictors of transgender prejudice: a meta-analysis. Sex Roles 2022 Dec;87(11-12):583-602. [doi: 10.1007/s11199-022-01338-6]
57. Warriner K, Nagoshi CT, Nagoshi JL. Correlates of homophobia, transphobia, and internalized homophobia in gay or lesbian and heterosexual samples. J Homosex 2013;60(9):1297-1314. [doi: 10.1080/00918369.2013.806177] [Medline: 23952924]
58. Harrison BF, Michelson MR. Gender, masculinity threat, and support for transgender rights: an experimental study. Sex Roles 2019 Jan;80(1-2):63-75. [doi: 10.1007/s11199-018-0916-6]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
IRB: internal review board
LGBTQ: lesbian, gay, bisexual, transgender, queer
TGD: transgender and gender diverse
TSS: Transgender Stigma Scale

Edited by A Mavragani; submitted 14.09.24; peer-reviewed by JD Bres, S Haim-Nachum; revised version received 21.11.24; accepted 24.11.24; published 20.01.25.

Please cite as:
Amsalem D, Greuel M, Liu S, Martin A, Adam M
Effect of a Short, Animated Storytelling Video on Transphobia Among US Parents: Randomized Controlled Trial
JMIR Public Health Surveill 2025;11:e66496
URL: https://publichealth.jmir.org/2025/1/e66496
doi:10.2196/66496

© Doron Amsalem, Merlin Greuel, Shuyan Liu, Andrés Martin, Maya Adam. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Social-Ecological Pathways From Sexual Identity to Sleep Among Chinese Women: Structural Equation Modeling Analysis

Chanchan Wu, BSc, MSc, PhD; Pui Hing Chau, BSocSci, PhD; Edmond Pui Hang Choi, BNurs, MPhil, PhD

School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 5/F, Academic Building, 3 Sassoon Road, Pok Fu Lam, Hong Kong, China (Hong Kong)

Corresponding Author:
Edmond Pui Hang Choi, BNurs, MPhil, PhD
School of Nursing, Li Ka Shing Faculty of Medicine, The University of Hong Kong, 5/F, Academic Building, 3 Sassoon Road, Pok Fu Lam, Hong Kong, China (Hong Kong)

Abstract

Background: Women and sexual minority individuals have been found to be at higher risk for experiencing poor sleep health compared to their counterparts. However, research on the sleep health of sexual minority women (SMW) is lacking in China.

Objective: This study aimed to examine sleep quality and social support for Chinese women with varied sexual identities, and then investigate the in-depth relationships between sexual identity and sleep.

Methods: This was a cross-sectional web-based survey. All participants completed a structured questionnaire containing a set of sociodemographic items referring to the social-ecological model of sleep health, the Pittsburgh Sleep Quality Index, the Social Support Rating Scale, and social relationships and environment domains of the World Health Organization Quality of Life-abbreviated short version. Pearson correlation coefficients were used to examine the relationship between sleep quality and social support as well as the two domains of quality of life. Structural equation modeling analysis was used to explore the social-ecological relationships.

Results: A total of 250 cisgender heterosexual women (CHW) and 259 SMW were recruited from July to September 2021. A total of 241 (47.3%) women experienced poor sleep quality and the rate was significantly higher in SMW than in CHW (55.2% vs 39.2%, $P<.001$). Around one-fifth of SMW reported low levels of social support, which was significantly higher than that of CHW (21.6% vs 5.6%, $P<.001$). Pearson correlations showed that overall sleep quality was significantly negatively associated with social support with weak correlations ($r=-0.26$, $P<.001$). The final structural equation modeling analysis with satisfactory fit indices identified 6 social-ecological pathways, showing that alcohol use, objective support, utilization of support, and perceived social relationship and environment quality of life played important roles in the sleep quality of individuals from their sexual identity.

Conclusions: SMW experienced poorer sleep quality compared to CHW. Further research is recommended to address the modifiable factors affecting sleep and then implement tailored sleep improvement programs.

(JMIR Public Health Surveill 2025;11:e53549) doi:10.2196/53549

KEYWORDS

sleep; social support; sexual minority women; social-ecological model; quality of life; structural equation model; Chinese women; China; women; structural equation modeling analysis; sleep quality; sexual identity; survey; heterosexual; cisgender

Introduction

Sleep is a crucial element that significantly impacts health across all populations. As a modifiable behavior, sleep is strongly linked to health and well-being [1], and abnormal sleep has been proven to be significantly associated with an extensive range of adverse health-related outcomes [2,3]. Consequently, sleep health is increasingly recognized as a public health issue that concerns everyone [4,5]. Sleep quality is widely used to evaluate sleep health and includes both objective and subjective aspects of sleep [6]. Many studies have shown that women generally experience poorer sleep quality than age-matched men, possibly due to the influence of sex steroids over sleep [7,8]. Although research on sleep health among vulnerable populations has gradually increased, most studies have been limited to racial or ethnic minorities [7,9-12], and knowledge of sleep health among sexual minority populations is relatively scarce [13].

To date, only 3 reviews have been published on the sleep health of sexual minorities, with all being narrative reviews that contain a very limited number of studies [14-16], reflecting the fact that relevant research is still in its infancy. Overall, these reviews

XSL•FO
RenderX

suggest that sleep health constitutes an unmet health need for sexual minorities, and sleep health disparities related to sexual identity have been widely documented. Consistently, existing studies have recognized that sexual minority individuals experience significantly worse sleep quality than their cisgender and heterosexual peers. Furthermore, women and sexual minority individuals reported more sleep difficulties than did men or heterosexual participants [17], and sexual minority women (SMW) were more likely to report a higher prevalence of poor sleep quality than heterosexual women regardless of race/ethnicity [18-20], especially those in less supportive environments [21]. Additionally, SMW were more vulnerable to experiencing sleep disturbances than both sexual minority and heterosexual men [22], indicating that the sleep health of SMW deserves more attention.

The current evidence on sleep health among sexual minority Chinese is very limited; it remains an understudied area of research. So far, there are only 3 relevant studies, 2 of which are based on the analysis of data from nationwide school-based surveys of adolescents [23,24] and college students [25], indicating that sexual minority status was significantly associated with poor sleep quality. Notably, there is only 1 study that specifically investigated the sleep and discrimination experienced by lesbian, gay, and bisexual individuals in Hong Kong [26], and the results showed that discrimination experienced was associated with greater sleep disturbance, which in turn led to poorer physical and mental health conditions.

Health-related research on sexual minority Chinese also exhibits significant gender disparities, manifested by a substantially higher focus on the male sexual minority population than on female groups [27]. A recently published scoping review that mapped all the scientific literature and gray reports on the health needs of women with same-sex attraction in mainland China determined that Chinese sexual minority women have multiple unmet health needs [28], such as substance abuse, concerning mental health, and sexual and reproductive health. However, little is known about their sleep health, which is a crucial health issue that deserves more attention.

Research on factors affecting sleep has received more attention in recent years [29,30]. Grandner et al [31,32] initially proposed the social-ecological model of sleep health that integrated possible determinants of sleep and sleep-induced health outcomes from a global perspective. This model has been applied to studies in different populations [9,33,34] but has not yet been applied to sexual minority populations. Regarding the determinants of sleep, this model considers from a socio-ecological perspective that sleep may be determined by multiple causes at 3 levels, including the individual level, social level, and societal level.

Individual-level factors contain all aspects directly related to individuals' sleep such as genetics and sleep-related behaviors. A growing body of literature suggests that alcohol use has detrimental effects on sleep [35] by reducing sleep quality [36]. Meanwhile, there is also comprehensive evidence that the pooled association between alcohol use and sleep disorders was not significant [37], suggesting that the relationship between alcohol use and sleep needs further investigation. Similarly, smoking has been identified as another sleep determinant [38], with smokers being more vulnerable to poorer sleep quality [39]. In China, a review targeting sexual minority women concluded that the prevalence of smoking and alcohol use among minority women was much higher than among general women [28]. Therefore, alcohol use and smoking status need to be considered along with sexual identity when exploring individual-level factors that affect sleep quality.

Social-level factors influencing sleep could be socioeconomic status, ethnicity, and social support. Ethnic minorities and individuals from disadvantaged economic backgrounds generally report lower sleep quality [12,13]. In addition, a review of sleep research concluded that financial hardship is associated with poor sleep health in the general population [16]. Notably, many studies have confirmed that social support is a key factor affecting sleep [40,41], with greater social support associated with better sleep outcomes [42-45], while having strained relationships was linked to more troubled sleep [46]. Compared with heterosexual populations, social support has unique functions in sexual minority individuals, but they reported receiving less support [47]. Therefore, when exploring social-level factors affecting sleep quality, conditions such as ethnicity, education, employment, and economic status need to be considered in addition to social support.

Regarding societal-level factors affecting sleep, geography, physical environment, and other aspects of the environment could be grouped into this category [32], and environmental factors may positively or negatively impact sleep [48]. A cohort study also found a link between declining social relationship quality and poor sleep quality [49]. The societal-level factors are closely linked to everyone's livelihood, and therefore it is worth exploring whether individuals are local residents or new migrants to their current residence and whether they are cohabitating with others. The abovementioned factors involved in the social-ecological model might have direct or indirect impacts on sleep, but the associations among these factors have not been fully studied.

Despite the evident vulnerability of women and sexual minorities to poor sleep, there has been little research on their sleep health. Meanwhile, in China, research on the sleep health of sexual minorities, although gradually receiving attention in recent years, is still very limited, and there is currently no research specifically on the sleep of sexual minority women. To help eliminate inequalities in sleep health, a more diverse and inclusive sample, including women of different sexual identities, is greatly needed. Therefore, this study aimed to examine the sleep quality and social support in Chinese adult women, compare them between SMW and cisgender heterosexual women (CHW), then investigate the in-depth relationships between sexual identity and sleep using structural equation modeling (SEM) in the Chinese context.

XSL•FO
RenderX

Methods

Participants and Recruitment

This was a web-based cross-sectional study. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist [50] was used for reporting the study findings. Eligible participants were Chinese women who: (1) were at least 18 years old, (2) self-identified as female or were assigned female gender at birth, and (3) were able to read and understand Mandarin Chinese.

Generally, the minimum sample size for SEM is 200 [51,52], and the sample size needs to be more than 25 times the number of parameters as a rule of thumb [53]. In this study, based on the hypothesized model, it is evident that 10 variables are expected to be entered into the SEM model validation, thus the minimum sample size is 250.

Convenience sampling and respondent-driven sampling methods were used to recruit this relatively hidden population [54,55], by releasing the study poster via 4 popular nongovernmental organizations and encouraging respondents to help recruit potential peers through their network of connections. Details regarding the recruitment procedure have been previously reported [56].

In this study, 524 questionnaires were collected between July and September 2021, of which 15 were excluded for being cisgender men, adolescents, or invalid data. Finally, a total of 509 (97.1%) adult women were enrolled in the study with no missing data.

Ethical Considerations

This study was approved by the Human Research Ethics Committee of the University of Hong Kong (reference number EA210325) on July 8, 2021. Electronic written consent was obtained for each study participant. Data were collected using an online survey platform (Wenjuanxing), and no compensation was provided. All information collected during the study was kept anonymous and strictly confidential.

Measures

Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI) is a commonly used measure of individuals' sleep quality [14,57]. It includes 19 self-rated items, and they generate 7 component scores with an average weight of 0 - 3 points. The Chinese version of the PSQI has been previously validated [58,59] and was used in this study. Specifically, the seven components are: (1) subjective sleep quality, (2) sleep latency, (3) sleep duration, (4) habitual sleep efficiency, (5) sleep disturbances, (6) sleeping medication use, and (7) daytime dysfunction. The sum of the 7 component scores yields 1 global score, ranging from 0 to 21 points, and higher scores represent poorer sleep quality. A global PSQI score >5 indicates poor sleep quality [6,57].

Social Support Rating Scale

The Social Support Rating Scale (SSRS) is a 10-item scale specially designed for the Chinese population with sound reliability and validity [60]. It includes three dimensions of social support: (1) subjective support, (2) utilization of support, and (3) objective support. The full-scale score and dimension scores are the sum of the scores of each item, with higher scores indicating higher levels of social support [61,62]. In addition, a total score of 22 or below could be classified as low level, 23 - 44 as medium level, and 45 - 66 as a high level of social support [63].

World Health Organization Quality of Life – Abbreviated Short Version

The World Health Organization Quality of Life – Abbreviated Short Version (WHOQOL-BREF) is a generic quality of life (Qol) measure comprising 4 domains, and the psychometric properties of its Chinese version have been confirmed [64-66]. According to the social-ecological model of sleep [32], social relationships and environment could be considered as the societal-level factors affecting sleep, but there is little evidence of associations between sleep and these factors. Thus, this study used the social relationship domain and environment domain of the WHOQOL-BREF as societal-level factors of sleep for investigation. Each domain score can be transformed into a score ranging from 4 to 20, with higher scores representing better Qol [64,67].

Sociodemographic Information

All participants completed a set of sociodemographic items that were widely reported in previous studies referring to the social-ecological model of sleep health [31,32], including individual-level factors (sexual identity, age, smoking status, alcohol use status, drug use status), social-level factors (ethnicity, education, employment, income, number of friends, social support, relationship status, bed-sharing situation), and societal-level factors (local resident or not, duration in current residence, cohabitation situation, social relationship and environment Qol).

Statistical Analysis

Descriptive statistics were reported on participants' demographic characteristics, sleep quality, and social support. Independent *t* tests and χ^2 tests were conducted for comparisons between SMW and CHW. Cohen *d* effect size was calculated [68].

Social-Ecological Factors Identification

Pearson correlation coefficients were used to examine the relationship between sleep quality and social support as well as the 2 domains of quality of life, and the correlations were defined as strong (≥0.5), moderate (≥0.3 and <0.5), or weak (<0.3) [68]. Then, independent *t* tests, 1-way ANOVA, and multiple linear regressions were performed to identify factors significantly associated with sleep quality.

Structural Equation Modeling Analysis

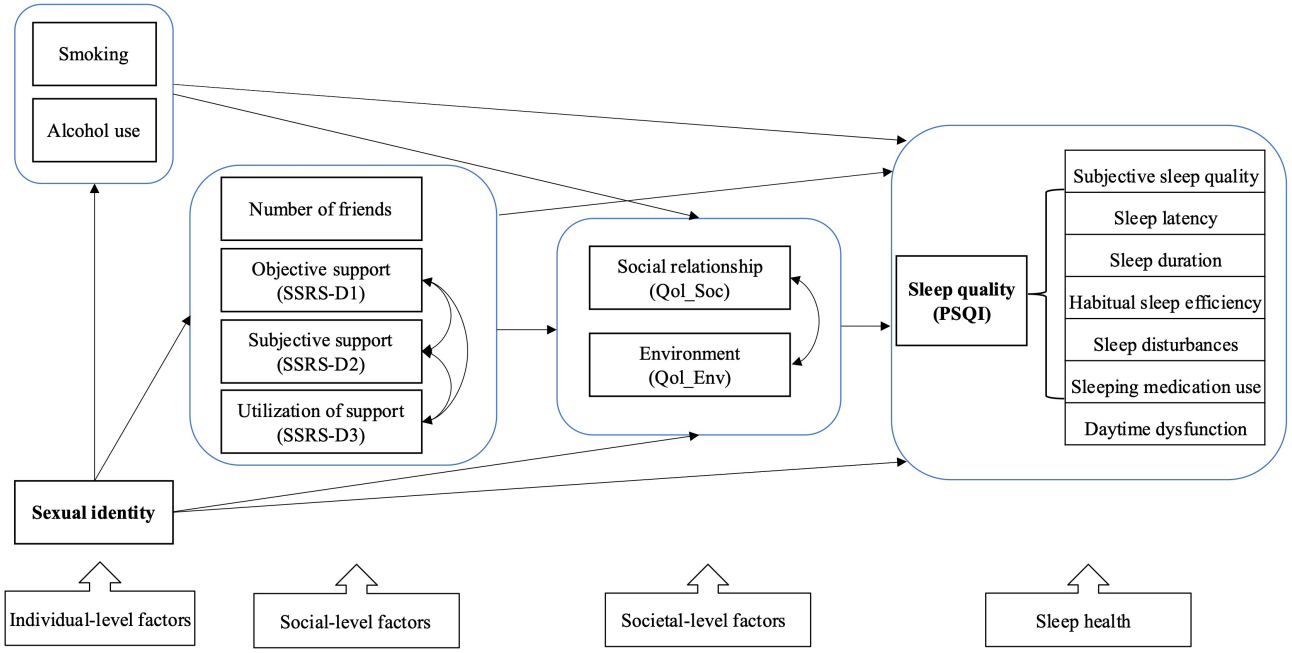
Based on the factors identified previously (which will be further detailed in the Results section), with reference to the socio-ecological model of sleep, the following factors with significant coefficients were simultaneously entered into the SEM analysis, including sexual identity (binary), alcohol use (binary), and smoking (binary) (identified as individual-level factors); the number of friends (categorical) and social support

XSL•FO
RenderX

(continuous) (identified as social-level factors); and social relationship Qol (continuous) and environment Qol (continuous) (identified as societal-level factors). Sleep quality was treated as the latent construct using the 7 domains of the PSQI measure.

The hypothesized model for socioecological influences on sleep based on the social-ecological model of sleep is shown in Figure 1.

Figure 1. Hypothesized model for socioecological influences on sleep (based on the social-ecological model of sleep). PSQI: Pittsburgh Sleep Quality Index; Qol: Quality of life; Qol_Env: Environment domain of Qol; Qol_Soc: Social relationship of Qol; SSRS: Social Support Rating Scale; SSRS-D1: objective support; SSRS-D2: subjective support; SSRS-D3: utilization of support.



All identified variables and the 7 components of sleep quality were included in the SEM model to test the hypothesized relationships among the previously listed variables. Standardized coefficients (β), standard errors, and associated *P* values are reported for all paths of the final model. Model goodness of fit was evaluated using chi-square (χ^2), root mean square error of approximation (RMSEA), comparative fit index (CFI), Tucker-Lewis index (TLI), and standardized root mean square residual (SRMR). The model fit was considered adequate [69,70] when RMSEA (90% CI) ≤0.08, CFI ≥0.90, TLI ≥0.90, and SRMR ≤0.05.

Data were analyzed using SPSS version 28.0 (IBM Corp), and the SEM analysis was performed using Mplus 8.6 software [71]. All significance tests were 2-tailed, and findings with *P* values <.05 were considered statistically significant.

Results

Participant Characteristics

Table 1 presents the sociodemographic characteristics of the 509 women. Approximately half of the participants (259/509, 50.9%) were SMW, and 250 (49.1%) were CHW. The mean age of the overall sample was 25.57 years (SD 5.77), ranging from 18 to 56 years old. Details regarding the sexual orientation and gender identity of participants have been previously reported [56]. Notably, SMW had higher rates of substance use than CHW, and the rates of smoking and alcohol use showed significant differences (27.4% vs 4%, *P*<.001; 73.7% vs 56%, *P*<.001; respectively). Nine participants in total (1.8%) had a history of recreational drug use (4 used cannabis and 5 used methamphetamine), 7 of whom were SMW.

XSL•FO
RenderX

Table . Sociodemographic characteristics of the study sample with comparisons (N=509).

Characteristics	Overall (N=509)	SMW[a] (n=259)	CHW[b] (n=250)
Smoking, n (%); *P* <.001 [c]			
Never smoked or have quit smoking	428 (84.1)	188 (72.6)	240 (96)
Current smoker	81 (15.9)	71 (27.4)	10 (4)
Seldom (<1 time/week)[d]	48 (59.3)	41 (57.7)	7 (70)
Usually (1-7 times/week)[d]	9 (11.1)	9 (12.7)	0
Almost smoke every day[d]	24 (29.6)	21 (29.6)	3 (30)
Alcohol use, n (%); *P* <.001			
Never drank or have quit	178 (35)	68 (26.3)	110 (44)
Current alcohol user	331 (65)	191 (73.7)	140 (56)
Seldom (<1 time/week)[e]	220 (43.2)	111 (42.9)	109 (43.6)
Occasionally (2 - 4 times/month)[e]	88 (17.3)	64 (24.7)	24 (9.6)
Usually (2 - 4 times/week)[e]	17 (3.3)	13 (5)	4 (1.6)
Almost use alcohol every day[e]	6 (1.2)	3 (1.2)	3 (1.2)
Drug use, n (%); *P* =.10			
Never used before	500 (98.2)	252 (97.3)	248 (99.2)
Have used drugs before	9 (1.8)	7 (2.7)	2 (0.8)
Ethnicity, n (%); *P* =.12			
Han people	466 (91.6)	242 (93.4)	224 (89.6)
Others (Muslim, etc)	43 (8.4)	17 (6.6)	26 (10.4)
Education, n (%); *P* <.001			
High school and below	33 (6.5)	21 (8.1)	12 (4.8)
College/bachelor	284 (55.8)	167 (64.5)	117 (46.8)
Graduate degree and above	192 (37.7)	71 (27.4)	121 (48.4)
Monthly income (Chinese yuan)[f], n (%); *P* <.001			
≤1000	163 (32)	88 (54)	75 (46)
1001 - 3000	79 (15.5)	51 (64.6)	28 (35.4)
3001 - 5000	40 (7.9)	23 (57.5)	17 (42.5)
5001 - 7000	33 (6.5)	17 (51.5)	16 (48.5)
7001 - 9000	45 (8.8)	24 (53.3)	21 (46.7)
9001 - 11,000	38 (7.5)	16 (42.1)	22 (57.9)
More than 11,000	111 (21.8)	40 (36)	71 (64)
Relationship, n (%); *P* <.001			
Have a steady partner	251 (49.3)	115 (44.4)	136 (54.4)
Have no steady partner(s)	258 (50.7)	144 (55.6)	114 (45.6)
Bed sharing status, n (%); *P*= .12			
Separate bed in separate room	293 (57.6)	160 (61.8)	133 (53.2)
Separate bed in shared room	76 (14.9)	37 (14.3)	39 (15.6)
Sharing same bed with partner	140 (27.5)	62 (23.9)	78 (31.2)
Local resident or not, n (%); *P*= .52			

XSL•FO
RenderX

Characteristics	Overall (N=509)	SMW[a] (n=259)	CHW[b] (n=250)
Local resident	181 (35.6)	96 (37.1)	85 (34)
Nonlocal resident (migrant)	328 (64.4)	163 (62.9)	165 (66)
Duration in current residence, n (%); *P*= .59			
<3 months	50 (9.8)	27 (10.4)	23 (9.2)
3 - 6 months	17 (3.3)	11 (4.2)	6 (2.4)
7 - 12 months	33 (6.5)	15 (5.8)	18 (7.2)
>1 year	409 (80.4)	206 (79.5)	203 (81.2)
Cohabitation status, n (%); *P* <.001			
Live alone	114 (22.4)	60 (23.2)	54 (21.6)
Live with same-sex partner	41 (8.1)	34 (13.1)	7 (2.8)
Live with opposite-sex partner	75 (14.7)	18 (7)	57 (22.8)
Live with friends	91 (17.9)	42 (16.2)	49 (19.6)
Live with family	162 (31.8)	93 (35.9)	69 (27.6)
Other	26 (5.1)	12 (4.6)	14 (5.6)
Quality of life, mean (SD); *P* <.001			
Social relationship domain	13.62 (3.07)	13.02 (3.35)	14.23 (2.61)
Environment domain	13.74 (2.68)	13.32 (2.90)	14.18 (2.35)

[a]SMW: sexual minority women. The percentages for this column were all calculated with 259 SMW as the denominator.

[b]CHW: cisgender heterosexual women. The percentages for this column were all calculated with 250 CHW as the denominator.

[c]All *P* values were reported by conducting comparisons performing χ^2 tests or independent *t* tests.

[d]The denominator for this row is the number of current smokers (n=81).

[e]The denominator for this row is the number of current alcohol users (n=331).

[f]1 CNY=US $0.136612.

Sleep Quality and Social Support

A total of 241 (47.3%) women experienced poor sleep quality and the rate was significantly higher in SMW than in CHW (55.2% vs 39.2%, *P*<.001). Similarly, independent *t* tests showed that SMW experienced significantly worse overall sleep quality than CHW (*P*<.001). There were statistically significant differences in subjective sleep quality, sleep latency, habitual sleep efficiency, sleep disturbances, and daytime dysfunction between CHW and SMW (Table 2). In addition, 81.5% (415/509) of all participants reported medium levels of social support. Around one-fifth of SMW reported low levels of social support, which was significantly higher than that of CHW (21.6% vs 5.6%, *P*<.001). Table 2 shows that SMW reported significantly lower support in all aspects than CHW (*P*<.01).

XSL•FO
RenderX

Table . Sleep quality and social support of the study sample with comparisons (N=509).

	Overall (N=509)	SMW[a] (n=259)	CHW[b] (n=250)	*P* value	Chi-square (*df*) or Cohen *d*
Sleep quality (PSQI [c]), n (%)					
Good (total score ≤5)	268 (52.7)	116 (44.8)	152 (60.8)	<.001	13.08 (1)
Poor (total score >5)	241 (47.3)	143 (55.2)	98 (39.2)		
Sleep quality (PSQI), mean (SD)					
Subjective sleep quality	1.13 (0.74)	1.22 (0.74)	1.04 (0.72)	.005	0.25
Sleep latency	1.24 (0.98)	1.33 (1.01)	1.14 (0.93)	.02	0.20
Sleep duration	0.42 (0.71)	0.46 (0.75)	0.37 (0.66)	.13	0.14
Habitual sleep efficiency	0.38 (0.78)	0.45 (0.86)	0.30 (0.67)	.03	0.19
Sleep disturbances	1.06 (0.53)	1.12 (0.53)	1.00 (0.53)	.009	0.23
Sleeping medication use	0.17 (0.59)	0.21 (0.66)	0.12 (0.50)	.06	0.17
Daytime dysfunction	1.53 (0.98)	1.68 (0.97)	1.37 (0.96)	<.001	0.32
Total score	5.92 (3.29)	6.48 (3.32)	5.33 (3.16)	<.001	0.36
Social support (SSRS [d]), n (%)					
Low level (total score ≤22)	70 (13.8)	56 (21.6)	14 (5.6)	<.001	36.42 (2)
Medium level (total score 23-44)	415 (81.5)	199 (76.8)	216 (86.4)		
High level (total score >45)	24 (4.7)	4 (1.6)	20 (8)		
Social support (SSRS), mean (SD)					
Objective support	7.55 (2.34)	7.23 (2.35)	7.89 (2.28)	.001	−0.28
Subjective support	16.17 (5.59)	14.54 (5.09)	17.85 (5.60)	<.001	−0.62
Utilization of support	7.49 (1.97)	7.04 (1.88)	7.96 (1.95)	<.001	−0.48
Total score	31.21 (7.96)	28.81 (7.41)	33.70 (7.75)	<.001	−0.65

[a]SMW: sexual minority women.

[b]CHW: cisgender heterosexual women.

[c]PSQI: Pittsburgh Sleep Quality Index.

[d]SSRS: Social Support Rating Scale.

Associations Between Social-Ecological Factors and Sleep

The correlations among sleep quality (PSQI total score and seven component scores) and social support (SSRS total score and three domain scores), social relationship Qol, and environment Qol are reported in Table 3. The total sleep quality was significantly negatively associated with social support and all domains with weak correlations (r ranged from −0.15 to −0.26, P<.001), and it was significantly negatively linked with social relationship Qol (r=−0.38, P<.001) and environment Qol (r=−0.39, P<.001) with moderate correlations.

XSL•FO
RenderX

Table . Correlations between sleep quality, social support, and quality of life.

Sleep quality (PSQI[a])	Social support (SSRS[b])				Quality of life	
	Objective	Subjective	Utilization	Sum score	Social	Environment
Subjective sleep quality						
r	−0.07	−0.17	−0.13	−0.17	−0.23	−0.26
P value	.14	<.001	.005	<.001	<.001	<.001
Sleep latency						
r	−0.13	−0.12	−0.16	−0.16	−0.21	−0.19
P value	.004	.005	<.001	<.001	<.001	<.001
Sleep duration						
r	−0.12	−0.15	−0.13	−0.17	−0.26	−0.28
P value	.007	<.001	.003	<.001	<.001	<.001
Habitual sleep efficiency						
r	−0.12	−0.15	−0.18	−0.18	−0.25	−0.26
P value	.007	<.001	<.001	<.001	<.001	<.001
Sleep disturbances						
r	−0.09	−0.12	−0.13	−0.14	−0.26	−0.29
P value	.04	.009	.003	.001	<.001	<.001
Sleeping medication use						
r	−0.12	−0.12	−0.09	−0.14	−0.10	−0.15
P value	.007	.007	.04	.001	.02	<.001
Daytime dysfunction						
r	−0.02	−0.18	−0.08	−0.15	−0.32	−0.25
P value	.66	<.001	.09	<.001	<.001	<.001
Total score						
r	−0.15	−0.24	−0.21	−0.26	−0.38	−0.39
P value	<.001	<.001	<.001	<.001	<.001	<.001

[a]PSQI: Pittsburgh Sleep Quality Index.

[b]SSRS: Social Support Rating Scale.

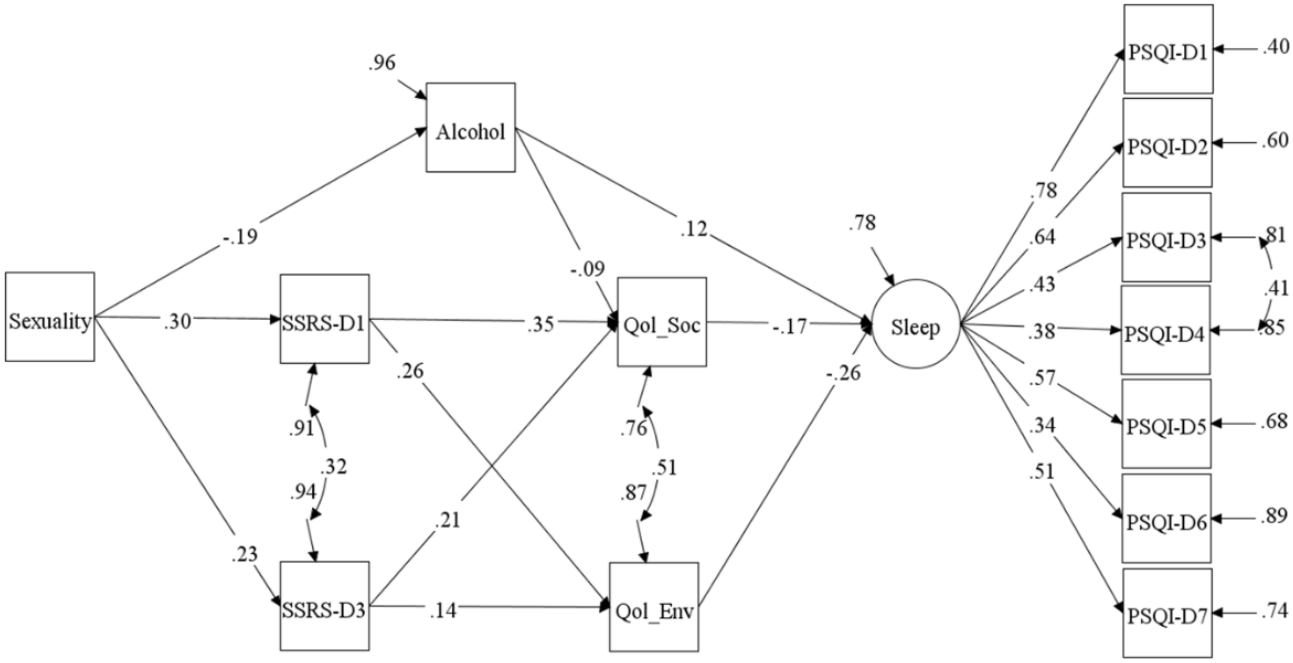
Comparisons of overall sleep quality between different socioecological groups are shown in Table S1 in Multimedia Appendix 1, with significant differences in sleep quality between people with different sexual identities, smoking status, alcohol use status, number of friends, and perceived social support (all *P*<.01). Table S2 in Multimedia Appendix 1 presents the multiple linear regression model of all the significant variables on sleep quality, and SEM was then performed among all these significant variables.

Model Testing

The full initial model including smoking and number of friends failed because these 2 variables were neither directly nor indirectly correlated to sleep in the model and were therefore not included in subsequent model analyses. Contrary to the hypothesized model, none of the aspects of social support were found to have a direct relationship with sleep quality. Furthermore, the model involving the subjective support domain of social support (SSRS-D2) showed no direct or indirect association of this domain with sleep and it was therefore also removed. The final model was then proposed and illustrated in Figure 2, exhibiting robust fit indices (χ^2_{51}=118.80; RMSEA=0.051, 90% CI 0.039-0.063; CFI=0.948, TLI=0.921, SRMR=0.040).

XSL•FO
RenderX

Figure 2. Final structural equation model of this study. Correlations and paths with nonsignificant *P* values are not depicted. PSQI: Pittsburgh Sleep Quality Index; PSQI-D1: subjective sleep quality; PSQI-D2: sleep latency; PSQI-D3: sleep duration; PSQI-D4: habitual sleep efficiency; PSQI-D5: sleep disturbance; PSQI-D6: sleeping medication use; PSQI-D7: daytime dysfunction; Qol: quality of life; Qol_Soc: social relationship of Qol; Qol_Env: environment domain of Qol; SSRS: Social Support Rating Scale; SSRS-D1: objective support; SSRS-D3: utilization of support.



All model standardized coefficients (β), their standard errors, and associated *P* values are reported in Table 4. There were 6 significant paths from sexual identity (sexuality) to sleep quality in the final SEM model. Specifically, alcohol use (β=0.12, *P*=.01), social relationship Qol (β=−0.17, *P*=.004), and environment Qol (β=−0.26, *P*<.001) were direct predictors of sleep quality. Although social support did not directly influence sleep, it may indirectly influence sleep by a ecting the perceived social relationship Qol and environment Qol (β range 0.14 - 0.35, *P*<.01).

XSL•FO
RenderX

Table . Standardized path coefficients and standard errors for all pathways in the final model.

Pathways	β[a]	SE	*P* value
Sexuality → Sleep quality	−0.07	0.05	.17
Sexuality → Alcohol use	−0.19	0.04	<.001
Sexuality → SSRS-D1[b]	0.30	0.04	<.001
Sexuality → SSRS-D3[c]	0.23	0.04	<.001
Sexuality → Qol_Soc[d]	0.03	0.04	.51
Sexuality → Qol_Env[e]	0.05	0.04	.23
Alcohol use → Sleep quality	0.12	0.05	.01
Alcohol use → Qol_Soc	−0.09	0.04	.02
Alcohol use → Qol_Env	0.00	0.04	.98
SSRS-D1 → Sleep quality	−0.03	0.05	.65
SSRS-D1 → Qol_Soc	0.35	0.04	<.001
SSRS-D1 → Qol_Env	0.26	0.05	<.001
SSRS-D3 → Sleep quality	−0.06	0.05	.25
SSRS-D3 → Qol_Soc	0.21	0.04	<.001
SSRS-D3 → Qol_Env	0.14	0.05	.002
Qol_Soc → Sleep quality	−0.17	0.06	.004
Qol_Env → Sleep quality	−0.26	0.06	<.001

[a]β: standardized coefficient.

[b]SSRS-D1: objective support domain of social support.

[c]SSRS-D3: utilization of support domain of social support.

[d]Qol_Soc: social relationship domain of quality of life.

[e]Qol_Env: environment domain of quality of life.

Discussion

Principal Findings

This study was the first to explore the social-ecological pathways from sexual identity to sleep quality. Few studies worldwide have examined the sleep health of SMW, and this investigation was the first to examine the sleep health of Chinese women with diverse sexual identities. This study showed SMW reported significantly worse overall sleep quality relative to CHW, which is consistent with existing reviews [14-16] and findings among Chinese youth [24,25]. Compared with the pooled PSQI mean score of the general Chinese population from a recent meta-analysis [72], our study populations reported higher total scores on the same standardized PSQI scale (5.92 vs 4.32), implying that the female population in this study experienced noticeably poorer sleep quality. Furthermore, compared with the percentage of Korean sexual minority adults (lesbian, gay, and bisexual) reporting poor sleep quality (33.8%) [73], the rate of SMW reporting poor sleep quality in our study (55.2%) was higher, indicating that sleep in SMW requires more attention and improvement.

This study also found that the proportions of current smokers and alcohol users were significantly higher among SMW than those among CHW, which is consistent with findings in China and overseas [28,74]. A scoping review summarized that some Chinese SMW regard smoking and drinking as a means of interacting with other SMW in specific social gatherings. It is also reported that smoking and drinking were the main strategies they used as a sexual minority to cope with marital stress and feeling isolated [28]. These also echo our findings on the mental health of this sample population, where we have found that SMW indeed significantly experienced more psychological symptoms [56]. A national survey conducted in the United States revealed higher rates of substance use disorders among sexual minority adults compared to heterosexual adults, with SMW showing the highest rates [75], suggesting that substance use among women, especially SMW, needs to be given attention as a public health issue.

Alcohol is one of the most commonly used psychoactive substances in different social communities around the world. A growing body of literature has confirmed that alcohol use has deleterious effects on sleep by reducing sleep duration and sleep quality [35,36], which has also been documented in sexual minority men [76] and is consistent with this study. However, a meta-analysis of cohort studies indicates that the relationship between alcohol consumption and sleep disorders is conditional, with pooled analyses finding that general drinking and the incidence of sleep disorder were significantly correlated, while heavy drinking was not [37]. Another review concluded that alcohol use has a bidirectional relationship with sleep continuity disturbance [77]. For instance, insomnia symptoms were

XSL•FO
RenderX

associated with subsequent heavy drinking; conversely, heavy drinking was associated with subsequent insomnia symptoms [78]. Nevertheless, both comparative analyses and the final SEM in this study revealed that drinking alcohol was significantly associated with poor sleep quality compared to not drinking alcohol, suggesting that both alcohol use and sleep in women deserve more attention and tailored intervention. Considering the limitations of the cross-sectional design used in this study, which could not provide evidence of causal relationships between drinking and sleep quality, more longitudinal studies with further subdivision of alcohol use are warranted in the future.

The adverse effects of smoking on sleep are well documented, as evidenced by the fact that smokers are more likely to experience sleeping difficulties, longer sleep latency, and poorer sleep quality [38,39,79], and similar findings were observed in the all-female population included in this study. The possible mechanism regarding the effects of smoking on sleep may be that smoking triggers depression or sleep-related breathing problems, which can lead to poor sleep quality [79]. However, a meta-analysis that included only cohort studies showed that, despite the significant correlation between smoking and the prevalence of insomnia, the difference between smokers and nonsmokers was very small (odds ratio 1.07) [80]. This may partly explain why the smoking variable that differed in between-group comparisons in our study was no longer associated with sleep quality in our final SEM. Future studies should record and subdivide smoking status and use prospective study designs to further explore the relationship between smoking and sleep across sexual identities.

The correlation between social support and sleep has been extensively studied in different populations [42,43,45] but its exploration in sexual minorities remains scarce. Therefore, the strength of this study is to fill this gap by examining the association between social support and sleep quality among Chinese women with diverse sexual identities. Consistent with previous research [49], overall social support was significantly and negatively associated with each component of sleep quality in the current study sample. Moreover, sexual minority individuals generally experience a higher level of poor social relationships, especially in the relatively traditional Chinese culture, where such groups are often invisible and therefore perceive more discrimination and stress [27,28,56]. However, findings from a study of 3 generations showed that feeling stigma was statistically significant in predicting a sleep disorder diagnosis among sexual minority individuals [81].

A mediation analysis of Chinese college students nationwide found that the relationship between sexual orientation and sleep quality was independently and in series mediated by interpersonal relationships and depressive symptoms, and this effect was more robust in men than women [25]. These interpersonal difficulties were related to their perceived stress associated with sleep difficulties. Our findings align with previous studies that found that sexual minority status is significantly associated with poor sleep quality [23], and social support has both direct and indirect correlations with sleep quality [42]. The relationships manifested in this study were that sexual identity was significantly associated with inadequate levels of social support, which in turn was significantly associated with poor levels of quality of life in terms of social relationships and environment, all of which were associated with poor sleep quality. Therefore, the lack of social support for sexual minority women is a significant public health issue, as is their poor sleep health, which deserves more attention.

Our correlation findings suggest that there may be more complex pathways from sexual identity to sleep among Chinese women with different sexual identities. Therefore, based on the holistic perspective proposed by the social-ecological model of sleep [32], this study explored and demonstrated the socioecological pathways from sexual identity to sleep quality for the first time using the multivariate SEM analysis method, providing a comprehensive evidence reference for promoting sleep health. The results of the final fitted model with satisfactory fit indices showed that starting from sexual identity, the following factors played important roles in women's sleep quality: alcohol use; objective social support and utilization of support; and social relationship and environment Qol. Our findings demonstrated the important roles of socioecological factors on sleep quality among women with different sexual identities. Therefore, future research should consider not only modifiable individual-level factors but also the potential influence of social- and societal-level factors when developing and implementing sleep promotion interventions. In particular, future interventions targeting sleep quality in sexual minority women may benefit from incorporating strategies that improve their social support, alcohol use, and the quality of their living environment.

This study has several limitations. First, due to the cross-sectional design, it was not possible to determine the explicit causal relationships among variables, so further research using longitudinal designs is needed. Second, despite the diversity of sexual identities, our sample was an all-female population (self-identified as female or assigned female at birth), so the results might not be generalizable to the broader Chinese sexual minority population. There has also been research indicating that bisexual people [82], those with less education [83], and those who are economically disadvantaged [16] may be more likely to suffer from sleep disorders than other groups. Hence, future studies with more diverse and representative samples including individuals from varied socioeconomic backgrounds are warranted. Third, the social-ecological model of sleep used in this study is only the upstream end of sleep research, and the downstream effects, namely the effects of sleep on specific health outcomes, have not been expanded upon in this study. Therefore, further research is recommended to comprehensively explore sleep and holistic health across diverse populations.

Conclusion

Overall, this study contributes to the existing literature by including SMW and expanding knowledge on sleep health in this population. By exploring socioecological pathways from sexual identity to sleep quality, this study provides a comprehensive understanding of factors that affect sleep health at different levels. These contextual and modifiable factors, based on the socioecological model of sleep, are layered and may interact with one another, providing guidance for future

XSL•FO
RenderX

interventions. In addition, the pathways from sexual identity to sleep identified in this study are complex and multidisciplinary; thus, all stakeholders and sleep professionals across disciplines are encouraged to collaborate and contribute to sleep improvement programs tailored for sexual minority individuals.

Acknowledgments

We are grateful to all the participants for sharing their experiences with us. We also thank TrueSelf Shenzhen, Period Pride China, Wuhan LGBT Center, and r&B bisexual community, as well as all those who shared this survey with their friends, for their contribution to the questionnaire distribution.

Data Availability

The datasets generated during and/or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table S1: Sleep quality of the study sample with comparisons; Table S2: Multiple linear regressions of sleep quality.
[DOCX File, 78 KB - publichealth_v11i1e53549_app1.docx]

References

1. Czeisler CA. Duration, timing and quality of sleep are each vital for health, performance and safety. Sleep Health 2015 Mar;1(1):5-8. [doi: 10.1016/j.sleh.2014.12.008] [Medline: 29073414]
2. Gao C, Guo J, Gong TT, et al. Sleep duration/quality with health outcomes: an umbrella review of meta-analyses of prospective studies. Front Med (Lausanne) 2021;8:813943. [doi: 10.3389/fmed.2021.813943] [Medline: 35127769]
3. Sivakumaran K, Ritonja JA, Palmer N, et al. Effect of sleep disturbance on biomarkers related to the development of adverse health outcomes: a systematic review of the human literature. J Sleep Res 2023 Jun;32(3):e13775. [doi: 10.1111/jsr.13775] [Medline: 36330773]
4. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 1;37(1):9-17. [doi: 10.5665/sleep.3298] [Medline: 24470692]
5. Weaver MD, Barger LK. Sleep health as an issue of public safety. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:489-499.
6. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
7. van de Langenberg SCN, Kocevska D, Luik AI. The multidimensionality of sleep in population-based samples: a narrative review. J Sleep Res 2022 Aug;31(4):e13608. [doi: 10.1111/jsr.13608] [Medline: 35429087]
8. Mong JA, Cusmano DM. Sex differences in sleep: impact of biological sex and sex steroids. Philos Trans R Soc Lond B Biol Sci 2016 Feb 19;371(1688):20150110. [doi: 10.1098/rstb.2015.0110] [Medline: 26833831]
9. Grandner MA, Williams NJ, Knutson KL, Roberts D, Jean-Louis G. Sleep disparity, race/ethnicity, and socioeconomic position. Sleep Med 2016 Feb;18:7-18. [doi: 10.1016/j.sleep.2015.01.020] [Medline: 26431755]
10. Patel NP, Grandner MA, Xie D, Branas CC, Gooneratne N. "Sleep disparity" in the population: poor sleep quality is strongly associated with poverty and ethnicity. BMC Public Health 2010 Aug 11;10:475. [doi: 10.1186/1471-2458-10-475] [Medline: 20701789]
11. Baldwin CM, Ervin AM, Mays MZ, et al. Sleep disturbances, quality of life, and ethnicity: the Sleep Heart Health Study. J Clin Sleep Med 2010 Apr 15;6(2):176-183. [Medline: 20411696]
12. Blanc J, Nunes J, Williams N, Robbins R, Seixas AA, Jean-Louis G. Sleep health equity. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:473-480. [doi: 10.1016/B978-0-12-815373-4.00035-6]
13. Williams N, Jean-Louis G, Blanc J, Wallace DM. Race, socioeconomic position and sleep. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:57-76. [doi: 10.1016/B978-0-12-815373-4.00006-X]
14. Butler ES, McGlinchey E, Juster RP. Sexual and gender minority sleep: a narrative review and suggestions for future research. J Sleep Res 2020 Feb;29(1):e12928. [doi: 10.1111/jsr.12928] [Medline: 31626363]
15. Dolsen EA, Byers AL, Flentje A, et al. Sleep disturbance and suicide risk among sexual and gender minority people. Neurobiol Stress 2022 Nov;21:100488. [doi: 10.1016/j.ynstr.2022.100488] [Medline: 36164391]
16. Patterson CJ, Potter EC. Sexual orientation and sleep difficulties: a review of research. Sleep Health 2019 Jun;5(3):227-235. [doi: 10.1016/j.sleh.2019.02.004] [Medline: 30975620]

XSL•FO
RenderX

17. Patterson CJ, Potter EC. Sleep difficulties of sexual minority adults: evidence from the 2018 National Health Interview Study. Psychol Sex Orientat Gend Divers 2022;9(2):244-253. [doi: 10.1037/sgd0000473]
18. Caceres BA, Hickey KT, Heitkemper EM, Hughes TL. An intersectional approach to examine sleep duration in sexual minority adults in the United States: findings from the Behavioral Risk Factor Surveillance System. Sleep Health 2019 Dec;5(6):621-629. [doi: 10.1016/j.sleh.2019.06.006] [Medline: 31377249]
19. Trinh MH, Agénor M, Austin SB, Jackson CL. Health and healthcare disparities among U.S. women and men at the intersection of sexual orientation and race/ethnicity: a nationally representative cross-sectional study. BMC Public Health 2017 Dec 19;17(1):964. [doi: 10.1186/s12889-017-4937-9] [Medline: 29258470]
20. Galinsky AM, Ward BW, Joestl SS, Dahlhamer JM. Sleep duration, sleep quality, and sexual orientation: findings from the 2013-2015 National Health Interview Survey. Sleep Health 2018 Feb;4(1):56-62. [doi: 10.1016/j.sleh.2017.10.004] [Medline: 29332681]
21. Martin-Storey A, Prickett KC, Crosnoe R. Disparities in sleep duration and restedness among same- and different-sex couples: findings from the American Time Use Survey. Sleep 2018 Aug 1;41(8):zsy090. [doi: 10.1093/sleep/zsy090] [Medline: 29726972]
22. Fricke J, Sironi M. Dimensions of sexual orientation and sleep disturbance among young adults. Prev Med Rep 2017 Dec;8:18-24. [doi: 10.1016/j.pmedr.2017.07.008] [Medline: 28831369]
23. Huang Y, Li P, Lai Z, et al. Role of sleep quality in mediating the relationship between sexual minority status and suicidal behavior among Chinese adolescents. Psychol Res Behav Manag 2018;11:607-615. [doi: 10.2147/PRBM.S186586] [Medline: 30573996]
24. Li P, Huang Y, Guo L, et al. Is sexual minority status associated with poor sleep quality among adolescents? Analysis of a national cross-sectional survey in Chinese adolescents. BMJ Open 2017 Dec 26;7(12):e017067. [doi: 10.1136/bmjopen-2017-017067] [Medline: 29282258]
25. Wu R, Wang W, Li W, et al. Sexual orientation and sleep problem among Chinese college students: mediating roles of interpersonal problems and depressive symptoms. J Affect Disord 2021 Dec 1;295:569-577. [doi: 10.1016/j.jad.2021.08.075] [Medline: 34509072]
26. Chan KKS, Fung WTW. Differential impact of experienced and anticipated discrimination on sleep and health among sexual minorities. Arch Sex Behav 2021 Oct;50(7):3053-3063. [doi: 10.1007/s10508-021-01981-2] [Medline: 34617190]
27. Wu C, Choi EPH, Chau PH. The holistic health status of Chinese homosexual and bisexual adults: a scoping review. Front Public Health 2021;9:710575. [doi: 10.3389/fpubh.2021.710575] [Medline: 34504828]
28. Wei TH, Jiang Y. Mapping the scientific literature on health needs of women with same-sex attraction in Mainland China: a scoping review. LGBT Health 2023 May;10(4):263-277. [doi: 10.1089/lgbt.2022.0119] [Medline: 36595507]
29. Billings ME, Hale L, Johnson DA. Physical and social environment relationship with sleep health and disorders. Chest 2020 May;157(5):1304-1312. [doi: 10.1016/j.chest.2019.12.002] [Medline: 31870910]
30. Etindele Sosso FA, Holmes SD, Weinstein AA. Influence of socioeconomic status on objective sleep measurement: a systematic review and meta-analysis of actigraphy studies. Sleep Health 2021 Aug;7(4):417-428. [doi: 10.1016/j.sleh.2021.05.005] [Medline: 34266774]
31. Grandner MA, Hale L, Moore M, Patel NP. Mortality associated with short sleep duration: the evidence, the possible mechanisms, and the future. Sleep Med Rev 2010 Jun;14(3):191-203. [doi: 10.1016/j.smrv.2009.07.006] [Medline: 19932976]
32. Grandner MA. Social-ecological model of sleep health. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:45-53. [doi: 10.1016/B978-0-12-815373-4.00005-8]
33. Fischer AR, Green SRM, Gunn HE. Social-ecological considerations for the sleep health of rural mothers. J Behav Med 2021 Aug;44(4):507-518. [doi: 10.1007/s10865-020-00189-4] [Medline: 33083923]
34. Li S. Nonrestorative sleep: its evaluation, associated factors and impact. Thesis.: The University of Hong Kong; 2021 URL: https://hub.hku.hk/handle/10722/306966 [accessed 2025-01-13]
35. He S, Taylor BV, Thakur NP, Chakravorty S. Sleep and alcohol use. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:269-281. [doi: 10.1016/B978-0-12-815373-4.00021-6]
36. Ebrahim IO, Shapiro CM, Williams AJ, Fenwick PB. Alcohol and sleep I: effects on normal sleep. Alcohol Clin Exp Res 2013 Apr;37(4):539-549. [doi: 10.1111/acer.12006] [Medline: 23347102]
37. Hu N, Ma Y, He J, Zhu L, Cao S. Alcohol consumption and incidence of sleep disorder: a systematic review and meta-analysis of cohort studies. Drug Alcohol Depend 2020 Dec 1;217:108259. [doi: 10.1016/j.drugalcdep.2020.108259] [Medline: 32927195]
38. Patterson F, Grandner MA, Lozano A, Satti A, Ma G. Transitioning from adequate to inadequate sleep duration associated with higher smoking rate and greater nicotine dependence in a population sample. Addict Behav 2018 Feb;77:47-50. [doi: 10.1016/j.addbeh.2017.09.011] [Medline: 28950118]
39. Patterson F, Ashare R. Improved sleep as an adjunctive treatment for smoking cessation. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:283-301. [doi: 10.1016/B978-0-12-815373-4.00022-8]
40. Gordon AM, Carrillo B, Barnes CM. Sleep and social relationships in healthy populations: a systematic review. Sleep Med Rev 2021 Jun;57:101428. [doi: 10.1016/j.smrv.2021.101428] [Medline: 33596514]

XSL•FO
RenderX

41. Nomura K, Yamaoka K, Nakao M, Yano E. Social determinants of self-reported sleep problems in South Korea and Taiwan. J Psychosom Res 2010 Nov;69(5):435-440. [doi: 10.1016/j.jpsychores.2010.04.014] [Medline: 20955862]
42. Kent de Grey RG, Uchino BN, Trettevik R, Cronan S, Hogan JN. Social support and sleep: a meta-analysis. Health Psychol 2018 Aug;37(8):787-798. [doi: 10.1037/hea0000628] [Medline: 29809022]
43. Krause N, Rainville G. Exploring the relationship between social support and sleep. Health Educ Behav 2020 Feb;47(1):153-161. [doi: 10.1177/1090198119871331] [Medline: 31452389]
44. Jin Y, Ding Z, Fei Y, et al. Social relationships play a role in sleep status in Chinese undergraduate students. Psychiatry Res 2014 Dec 15;220(1-2):631-638. [doi: 10.1016/j.psychres.2014.08.029] [Medline: 25200188]
45. Mesas AE, Peppard PE, Hale L, Friedman EM, Nieto FJ, Hagen EW. Individuals' perceptions of social support from family and friends are associated with lower risk of sleep complaints and short sleep duration. Sleep Health 2020 Feb;6(1):110-116. [doi: 10.1016/j.sleh.2019.08.013] [Medline: 31740375]
46. Ailshire JA, Burgard SA. Family relationships and troubled sleep among U.S. adults: examining the influences of contact frequency and relationship quality. J Health Soc Behav 2012;53(2):248-262. [doi: 10.1177/0022146512446642] [Medline: 22653715]
47. Frost DM, Meyer IH, Schwartz S. Social support networks among diverse sexual minority populations. Am J Orthopsychiatry 2016 Jan;86(1):91-102. [doi: 10.1037/ort0000117] [Medline: 26752447]
48. Jackson CL, Gaston SA. The impact of environmental exposures on sleep. In: Grandner MA, editor. Sleep and Health: Academic Press; 2019:85-103. [doi: 10.1016/B978-0-12-815373-4.00008-3]
49. Stafford M, Bendayan R, Tymoszuk U, Kuh D. Social support from the closest person and sleep quality in later life: evidence from a British birth cohort study. J Psychosom Res 2017 Jul;98:1-9. [doi: 10.1016/j.jpsychores.2017.04.014] [Medline: 28554363]
50. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
51. Crowley SL, Fan X. Structural equation modeling: basic concepts and applications in personality assessment research. J Pers Assess 1997 Jun;68(3):508-531. [doi: 10.1207/s15327752jpa6803_4] [Medline: 9170298]
52. Jackson DL. Revisiting sample size and number of parameter estimates: some support for the N:q hypothesis. Struct Equ Modeling 2003 Jan;10(1):128-141. [doi: 10.1207/S15328007SEM1001_6]
53. Violato C, Hecker KG. How to use structural equation modeling in medical education research: a brief guide. Teach Learn Med 2007;19(4):362-371. [doi: 10.1080/10401330701542685] [Medline: 17935466]
54. Crawford FW, Wu J, Heimer R. Hidden population size estimation from respondent-driven sampling: a network approach. J Am Stat Assoc 2018;113(522):755-766. [doi: 10.1080/01621459.2017.1285775] [Medline: 30828120]
55. Hequembourg AL, Panagakis C. Maximizing respondent-driven sampling field procedures in the recruitment of sexual minorities for health research. SAGE Open Med 2019;7:2050312119829983. [doi: 10.1177/2050312119829983] [Medline: 30800299]
56. Wu C, Chau PH, Choi EPH. Quality of life and mental health of Chinese sexual and gender minority women and cisgender heterosexual women: cross-sectional survey and mediation analysis. JMIR Public Health Surveill 2023 Feb 22;9:e42203. [doi: 10.2196/42203] [Medline: 36811941]
57. Choi EPH, Wan EYF, Kwok JYY, Chin WY, Lam CLK. The mediating role of sleep quality in the association between nocturia and health-related quality of life. Health Qual Life Outcomes 2019 Dec 11;17(1):181. [doi: 10.1186/s12955-019-1251-5] [Medline: 31829192]
58. Tsai PS, Wang SY, Wang MY, et al. Psychometric evaluation of the Chinese version of the Pittsburgh Sleep Quality Index (CPSQI) in primary insomnia and control subjects. Qual Life Res 2005 Oct;14(8):1943-1952. [doi: 10.1007/s11136-005-4346-x] [Medline: 16155782]
59. Liu X, Tang M, Hu L, et al. Reliability and validity of the Pittsburgh sleep quality index. Chin J Neuro Psychiatry 1996;29(2):103-107.
60. Xiao S. The theoretical basis and application of social support rating scale (SSRS). J Clin Psychol Med (Chin) 1994;4(2):98-100.
61. Zhang JY, Shu T, Xiang M, Feng ZC. Learning burnout: evaluating the role of social support in medical students. Front Psychol 2021;12:625506. [doi: 10.3389/fpsyg.2021.625506] [Medline: 33692725]
62. Hou T, Zhang T, Cai W, et al. Social support and mental health among health care workers during coronavirus disease 2019 outbreak: a moderated mediation model. PLoS ONE 2020;15(5):e0233831. [doi: 10.1371/journal.pone.0233831] [Medline: 32470007]
63. Sun J, Sun R, Jiang Y, et al. The relationship between psychological health and social support: evidence from physicians in China. PLoS One 2020;15(1):e0228152. [doi: 10.1371/journal.pone.0228152] [Medline: 31995601]
64. Xia P, Li N, Hau KT, Liu C, Lu Y. Quality of life of Chinese urban community residents: a psychometric study of the mainland Chinese version of the WHOQOL-BREF. BMC Med Res Methodol 2012 Mar 27;12:37. [doi: 10.1186/1471-2288-12-37] [Medline: 22452994]

XSL•FO
RenderX

65. Choi EPH, Wong JYH, Fong DYT. Mental health and health-related quality of life of Chinese college students who were the victims of dating violence. Qual Life Res 2017 Apr;26(4):945-957. [doi: 10.1007/s11136-016-1413-4] [Medline: 27660071]
66. Bonomi AE, Patrick DL, Bushnell DM, Martin M. Validation of the United States' version of the World Health Organization Quality of Life (WHOQOL) instrument. J Clin Epidemiol 2000 Jan;53(1):1-12. [doi: 10.1016/s0895-4356(99)00123-7] [Medline: 10693897]
67. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
68. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Routledge; 2013.
69. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
70. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996;1(2):130-149. [doi: 10.1037//1082-989X.1.2.130]
71. Muthén B, Muthén L. Mplus 86 software. URL: https://www.statmodel.com/index.shtml [accessed 2025-01-13]
72. Chen P, Lam MI, Si TL, et al. The prevalence of poor sleep quality in the general population in China: a meta-analysis of epidemiological studies. Eur Arch Psychiatry Clin Neurosci 2024 Oct;274(7):1-14. [doi: 10.1007/s00406-024-01764-5] [Medline: 38429554]
73. Choo S, Kim R, Lee H, Yi H, Kim R, Kim SS. Association between discrimination and poor sleep health outcomes among 2192 South Korean gay, lesbian, and bisexual adults: a nationwide cross-sectional survey. Sleep Health 2022 Dec;8(6):587-592. [doi: 10.1016/j.sleh.2022.09.006] [Medline: 36220734]
74. Johns MM, Pingel ES, Youatt EJ, Soler JH, McClelland SI, Bauermeister JA. LGBT community, social network characteristics, and smoking behaviors in young sexual minority women. Am J Community Psychol 2013 Sep;52(1-2):141-154. [doi: 10.1007/s10464-013-9584-4] [Medline: 23783884]
75. Kahle EM, Veliz P, McCabe SE, Boyd CJ. Functional and structural social support, substance use and sexual orientation from a nationally representative sample of US adults. Addiction 2020 Mar;115(3):546-558. [doi: 10.1111/add.14819] [Medline: 31599027]
76. Duncan DT, Goedel WC, Mayer KH, et al. Poor sleep health and its association with mental health, substance use, and condomless anal intercourse among gay, bisexual, and other men who have sex with men. Sleep Health 2016 Dec;2(4):316-321. [doi: 10.1016/j.sleh.2016.07.003] [Medline: 29073390]
77. He S, Hasler BP, Chakravorty S. Alcohol and sleep-related problems. Curr Opin Psychol 2019 Dec;30:117-122. [doi: 10.1016/j.copsyc.2019.03.007] [Medline: 31128400]
78. Haario P, Rahkonen O, Laaksonen M, Lahelma E, Lallukka T. Bidirectional associations between insomnia symptoms and unhealthy behaviours. J Sleep Res 2013 Feb;22(1):89-95. [doi: 10.1111/j.1365-2869.2012.01043.x] [Medline: 22978579]
79. Amiri S, Behnezhad S. Smoking and risk of sleep-related issues: a systematic review and meta-analysis of prospective studies. Can J Public Health 2020 Oct;111(5):775-786. [doi: 10.17269/s41997-020-00308-3] [Medline: 32185746]
80. Hu N, Wang C, Liao Y, Dai Q, Cao S. Smoking and incidence of insomnia: a systematic review and meta-analysis of cohort studies. Public Health (Fairfax) 2021 Sep;198:324-331. [doi: 10.1016/j.puhe.2021.07.012] [Medline: 34507139]
81. Segovia M, Sparks PJ. Sexual minorities and sleep health: how does discrimination and stigma influence a sleep disorder diagnosis? Sleep Health 2024 Feb;10(1):41-47. [doi: 10.1016/j.sleh.2023.11.003] [Medline: 38151378]
82. Caceres BA, Hickey KT. Examining sleep duration and sleep health among sexual minority and heterosexual adults: findings from NHANES (2005-2014). Behav Sleep Med 2020;18(3):345-357. [doi: 10.1080/15402002.2019.1591410] [Medline: 30916580]
83. Chu Y, Aune D, Yu C, et al. Temporal trends in sleep pattern among Chinese adults between 2010 and 2018: findings from five consecutive nationally representative surveys. Public Health (Fairfax) 2023 Dec;225:360-368. [doi: 10.1016/j.puhe.2023.10.004] [Medline: 37981452]

Abbreviations

CFI: comparative fit index
CHW: cisgender heterosexual women
PSQI: Pittsburgh Sleep Quality Index
Qol: quality of life
Qol_Env: environment domain of quality of life
Qol_Soc: social relationship domain of quality of life
RMSEA: root mean square error of approximation
SEM: structural equation modeling
SMW: sexual minority women
SRMR: standardized root mean square residual

XSL•FO
RenderX

SSRS: Social Support Rating Scale
TLI: Tucker-Lewis index
WHOQOL-BREF: World Health Organization Quality of Life–abbreviated short version

Edited by A Mavragani; submitted 10.10.23; peer-reviewed by A AL-Asadi, Q Chen; revised version received 22.10.24; accepted 29.10.24; published 21.01.25.

Please cite as:
Wu C, Chau PH, Choi EPH
Exploring Social-Ecological Pathways From Sexual Identity to Sleep Among Chinese Women: Structural Equation Modeling Analysis
JMIR Public Health Surveill 2025;11:e53549
URL: https://publichealth.jmir.org/2025/1/e53549
doi:10.2196/53549

© Chanchan Wu, Pui Hing Chau, Edmond Pui Hang Choi. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparing Multiple Imputation Methods to Address Missing Patient Demographics in Immunization Information Systems: Retrospective Cohort Study

Sara Brown[1], CHES, MPH; Ousswa Kudia[1], MPH; Kaye Kleine[1], MPH, MBA; Bryndan Kidd[2], MS; Robert Wines[2], CPM, CHP, BS; Nathanael Meckes[1], PhD

[1]Scientific Services - Analytics, Scientific Technologies Corporation (United States), 411 S 1st St, Phoenix, AZ, United States

[2]Immunization Services, West Virginia Department of Health and Human Services, Charleston, WV, United States

Corresponding Author:
Sara Brown, CHES, MPH
Scientific Services - Analytics, Scientific Technologies Corporation (United States), 411 S 1st St, Phoenix, AZ, United States

Abstract

Background: Immunization Information Systems (IIS) and surveillance data are essential for public health interventions and programming; however, missing data are often a challenge, potentially introducing bias and impacting the accuracy of vaccine coverage assessments, particularly in addressing disparities.

Objective: This study aimed to evaluate the performance of 3 multiple imputation methods, Stata's (StataCorp LLC) multiple imputation using chained equations (MICE), scikit-learn's Iterative-Imputer, and Python's miceforest package, in managing missing race and ethnicity data in large-scale surveillance datasets. We compared these methodologies in their ability to preserve demographic distribution, computational efficiency, and performed G-tests on contingency tables to obtain likelihood ratio statistics to assess the association between race and ethnicity and flu vaccination status.

Methods: In this retrospective cohort study, we analyzed 2021 - 2022 flu vaccination and demographic data from the West Virginia Immunization Information System (N=2,302,036), where race (15%) and ethnicity (34%) were missing. MICE, Iterative Imputer, and miceforest were used to impute missing variables, generating 15 datasets each. Computational efficiency, demographic distribution preservation, and spatial clustering patterns were assessed using G-statistics.

Results: After imputation, an additional 780,339 observations were obtained compared with complete case analysis. All imputation methods exhibited significant spatial clustering for race imputation (G-statistics: MICE=26,452.7, Iterative-Imputer=128,280.3, Miceforest=26,891.5; *P*<.001), while ethnicity imputation showed variable clustering patterns (G-statistics: MICE=1142.2, Iterative-Imputer=1.7, Miceforest=2185.0; *P*: MICE<.001, Iterative-Imputer=1.7, Miceforest<.001). MICE and miceforest best preserved the proportional distribution of demographics. Computational efficiency varied, with MICE requiring 14 hours, Iterative Imputer 2 minutes, and miceforest 10 minutes for 15 imputations. Postimputation estimates indicated a 0.87% - 18% reduction in stratified flu vaccination coverage rates. Overall estimated flu vaccination rates decreased from 26% to 19% after imputations.

Conclusions: Both MICE and Miceforest offer flexible and reliable approaches for imputing missing demographic data while mitigating bias compared with Iterative-Imputer. Our results also highlight that the imputation method can profoundly affect research findings. Though MICE and Miceforest had better effect sizes and reliability, MICE was much more computationally and time-expensive, limiting its use in large, surveillance datasets. Miceforest can use cloud-based computing, which further enhances efficiency by offloading resource-intensive tasks, enabling parallel execution, and minimizing processing delays. The significant decrease in vaccination coverage estimates validates how incomplete or missing data can eclipse real disparities. Our findings support regular application of imputation methods in immunization surveillance to improve health equity evaluations and shape targeted public health interventions and programming.

(JMIR Public Health Surveill 2025;11:e73916) doi:10.2196/73916

KEYWORDS

multiple imputation; missing data; imputation methods; data science; machine learning; statistical modeling; immunization information system

XSL•FO
RenderX

Introduction

The usage of large datasets obtained from surveillance data and Immunization Information Systems (IIS) has held a vital role in recognizing and comprehending the extent of health disparities and inequities within a population [1,2]. Previous research has shown that there is an association between race and ethnicity and vaccine acceptance and uptake [3-5]. However, as exemplified by the COVID-19 pandemic, race and ethnicity fields in public health surveillance systems and health records are historically underpopulated, thereby limiting a full understanding of the extent of vaccine inequities [2,6,7]. Missing data can affect the capability to effectively describe vaccine coverage and may introduce bias into epidemiologic analyses, particularly when attempting to estimate and address racial and ethnic health inequities, further compromising public health decision-making and resource allocation.

Studies exploring racial and ethnic inequalities traditionally omit individuals with missing demographic information or classify them as "unknown" [8-11]. However, this can miscalculate vaccine coverage rates when stratified by race and ethnicity, leading to biased analyses between different racial and ethnic groups. This selection bias dampens the capacity to accurately measure the true vaccine uptake among underserved populations that are more likely to have racial and ethnic data missing in surveillance datasets [2,12-14]. Multiple imputation for missing data, alternatively, has been shown to be a better method for eliminating missing data and preserving the greatest amount of data [13]. The theoretical foundation and technique for multiple imputation was established by D B Rubin, who developed the structure for managing missing data through the formation of multiple probable datasets [15]. This central process undertakes the innate ambiguity of missing values by generating multiple, complete, and reasonable datasets [16-18].

Following multiple imputation, classic likelihood ratio tests cannot be implemented as is because the final estimates do not come directly from a single model, and as a result, require modification to account for the uncertainty introduced by the imputation process. The multiple imputation likelihood ratio test, developed by Li et al [19] and Meng et al [20], or stacked multiple imputation, introduced by Chan et al [21,22], provides frameworks for hypothesis testing that properly incorporate both within-imputation and between-imputation variability.

Multiple imputations using chained equations (MICE) builds upon Rubin's initial work to handle multifaceted and specific datasets; it is a statistically sound method for managing missing data and has been used in the context of medical research and large, national datasets [23,24]. MICE uses the distribution of the original data to estimate values that signify the ambiguity of the true missing value. This can yield unbiased approximations after an adequate number of imputations, which is contingent on the quality of the dataset.

Machine learning (ML) is increasingly being used to reconcile missing data by offering robust and sophisticated solutions for improving data quality. Different ML techniques use various mathematical processes based on decision trees to predict a specified outcome, all with varying levels of accuracy [25-28]. Similar to traditional MICE, the accuracy and performance of machine learning algorithms to impute missing data rely heavily on the type of missing data being evaluated.

When compared with various ML techniques, MICE had less bias and similar standard errors for parameter estimates, with additional studies demonstrating that combining MICE with ML yields less biased results [26]. Miceforest is a combination of classical epidemiological techniques and ML algorithms. This algorithm uses MICE with light gradient boosting, a tree-based algorithm, to provide a flexible and powerful solution that remains efficient and reliable for managing missing data [29-31].

Recent public health pandemics and outbreaks have emphasized the importance of reliable approaches to missing data in surveillance systems. Studies have demonstrated that imputation methods significantly affect results, with incidence estimates and measures of disparity varying by method, and multiple imputation demonstrating more consistency and efficacy than single or complete case approaches [2,32,33].

Recent attempts have explored Bayesian Surname and Geocoding and Bayesian Improved Surname and Geocoding (BISG) methods to reconcile missing race and ethnicity data in large public health datasets [2,34-36]. While this approach has demonstrated practicality in particular circumstances, its suitability for state-level immunization registries is imperfect. Specifically, name-based imputations can present systematic bias in racially and ethnically diverse populations [37]. Statewide immunization registries contain various naming conventions and regional uniqueness; reliance on a name-based imputation can bolster misclassification bias and lead to misguided epidemiological analyses. As a result, approaches such as multiple imputation using observed covariates may be more suitable for IISs [38-40].

IISs present unique missing data challenges because they are a birth-to-death registry, making them temporally complex. Missing data patterns may vary by birth cohort due to evolving reporting requirements, provider participant rates, state-specific reporting requirements, use of multiple providers, and technology advancements.

However, this research addresses a critical gap identified in the surveillance literature, where, despite the widespread recognition of missing data challenges, few studies have systematically evaluated imputation approaches specifically designed for immunization data. To date, there is little published peer-reviewed literature on methods to address missing data in IIS. A 2024 cross-sectional study improved race and ethnicity data in the NYS laboratory reporting system by integrating IIS data, while Russ et al [41] enhanced IIS completeness by incorporating external records and applying logistic multiple imputation and random forest techniques [42].

This study builds upon established multiple imputation methodologies while addressing the complexities of IIS data structures by adapting recent advancements in handling missing data in EHRs to the specific challenges of population-based immunization surveillance. In this study, we investigate racial and ethnic disparities in flu vaccine uptake in the West Virginia

XSL•FO
RenderX

IIS for the 2021‐2022 flu season. We sought to address potential bias due to missing race and ethnicity data through multiple imputation and test the efficiency and accuracy of 3 established methods to address missing data in IIS and the implications of imputations on flu vaccination coverage data.

Methods

Overview

In this retrospective cohort, patient data for the 2021‐2022 flu vaccination (June 1, 2021, to June 30, 2022) was obtained from the West Virginia Statewide Immunization Information System (WVSIIS). Geo-demographic data, such as urbanicity and Social Vulnerability Index (SVI) status, were calculated using county address data obtained from the IIS. For the purposes of our analysis, urbanicity was defined as a county being either "metropolitan" or "rural." In accordance with the Federal Office of Rural Health Policy (FORHP), counties that had a population of 50,000 or more were considered "metropolitan" and counties with less than 50,000 were considered "rural." County population was obtained from the West Virginia 2020 US Census [43,44]. SVI is calculated by the Centers for Disease Control and Prevention (CDC) and the Agency for Toxic Substances and Disease Registry (ATSDR) at the census tract and county level and represents the potential negative effects of external stresses, such as housing, socioeconomic, and minority statuses have on a community [45].

After a preliminary data cleaning process on a clone of the WVSIIS database, where patients who were marked as deceased, inactive, or retained out of state address were removed, there were a resulting 2,302,036 records with a 2021‐2022 flu vaccine coverage rate of 26%. 15% of records were missing patient race (n=347,633), and 34% (n=780,339) of records were missing patient ethnicity. We addressed missing race and ethnicity data by using 3 different multiple imputation approaches to assess the stability and reliability of our findings.

We first implemented Stata 17's (StataCorp LLC) MICE using the mi impute chained command. Stata allows us to identify individual imputed datasets and pool results, excluding the original dataset with missing values.

We then used the Python scikit-learn's Iterative-Imputer algorithm. Iterative Imputer operates similarly to MICE by using available data in other features to estimate missing values. The imputation is performed in an iterative, round-robin manner, with a regressor to predict the missing values [46].

We applied the Python Miceforest package, which is designed for performing MICE using random forests. Rather than assume a linear relationship between variables, miceforest can capture complex, non-linear patterns and imputes values iteratively, unlike traditional MICE [33].

With each imputation method, we created 15 imputed datasets to stabilize our variance estimates [47,48]. We jointly imputed patient race and ethnicity using age, sex, urbanicity, SVI, county, and flu vaccination status. Selection of these covariates was based on the work done by Zhang et al [40], and the considerable missingness present in other potentially informative variables.

Once missing data had been reconciled, postimputation estimates were applied to WVSIIS flu vaccination data for the 2021‐2022 flu season.

We performed G-tests on contingency tables to obtain likelihood ratio statistics to assess the association between race and ethnicity and flu vaccination status. For MICE, we pooled across individual imputed datasets, and G-test statistics were averaged across imputations. In addition, we calculated between-imputation variance to evaluate the stability of results. Pooling was used in lieu of stacking to accommodate the computational requirements of the statewide immunization data's size. No pooling was required of Iterative-Imputer and Miceforest because they produce a single imputation result.

To assess the proportion estimate, we calculated vaccination proportions with 95% CIs using the Wilson score interval method. This method offers improved coverage properties than the standard normal estimate.

Sensitivity analyses examined vaccination disparity statistical significance, consistency, comparison of G-statistics to determine variances in effect sizes, and whether CIs for vaccination proportions overlap.

The primary engine for Iterative-Imputer, Miceforest, and all postimputation analyses was executed in Python 3.11 using pandas, scipy.stats, and matplotlib. The evaluation framework was intended to manage the multilevel configuration of Stata's MICE while preserving uniformity with Iterative-Imputer and Miceforest. All Python computations were performed on a cloud-based computing cluster with 16-core processors and 128 GB of RAM.

Ethical Considerations

This study consisted of secondary analysis of deidentified data from WVSIIS. The original data collection was conducted as part of state-mandated public health surveillance under W Va Code R § 64-7-6.6, which requires immunization data collection for all individuals residing in West Virginia. As such, it was not subject to institutional review board (IRB) review [49]. The secondary analysis did not involve direct interaction with human candidates and was determined not to constitute human candidate research, qualifying for exemption from IRB review and not requiring informed consent. This study met the jurisdiction requirement for IRB exemption, given the nature of the data used. The data used were compiled for public health surveillance and did not contain personal designations. Consent to access and analyze the data was permitted by the West Virginia Division of Immunization Services. All analyses met the terms of related federal, state, and international data privacy regulations. No compensation was provided or applicable, and no images or materials containing identifiable individual information are included in the manuscript or supplementary files.

Results

MICE

347,633 race categories that were previously missing were imputed after completing MICE. There were 16.5% additional

XSL•FO
RenderX

White records, 16.8% additional Black records, 16.3% additional Asian records, 18.5% additional Indigenous records, 16.8% additional Native Hawaiian or Pacific Islander records, 17.4% additional Multiracial records, and 15.3% additional other records after imputations (Table 1).

780,339 ethnicity categories that were previously missing were imputed after completing MICE. There were 52.4% additional Hispanic or Latino records and 40.5% additional Not Hispanic or Latino records after imputations (Table 1).

After MICE, individual demographics remained proportional to the original dataset distributions (Figures 1 and 2 and Table 2). A chi-square test illustrated that there was a statistically significant difference between the complete case and MICE estimates (P<.001; Table 3). Overall computational time was approximately 14 hours.

For the 2021 - 2022 flu season, the complete case analysis flu vaccine coverage rate was 26%. After MICE, West Virginia had an overall flu vaccine coverage rate of 19%. Flu vaccine coverage rates decreased when stratified by race and ethnicity when compared with a complete case analysis, with the most significant decreases observed in non-Hispanic (NH) White (5%), NH Black (4%), Hispanic or Latino (6%), Asian (7%), Native Hawaiian or Pacific Islander (6%), and Other (19%; Table 4 and Figure 3). After reconciling missing race and ethnicity, an additional 63,984 individuals were included in the analysis that were previously excluded from stratified analyses.

Likelihood ratio tests were used to determine the impact of race and ethnicity on vaccine uptake. Using MICE, race was significantly associated with flu vaccination status (G-statistic range=26,365.427 - 26,612.980; pooled G-statistic=26,452.66; P<.001), with considerable between-imputation variance (4103.38; Table 5). This denotes good stability with a coefficient of variation of 0.24%. For ethnicity, the MICE model also denoted a significant impact (G-statistic range=1055.287 - 1,205.254; pooled G-statistic=1142.23; P<.001), also with substantial between-imputation variance (1277.885) and a moderate coefficient of variation of 3.13% (Table 5). Both of MICE's G-scores demonstrate a large effect.

Table . Assessment of West Virginia Statewide Immunization Information System race and ethnicity demographics using complete case analysis and 3 imputation methods—multiple imputation using chained equations, Iterative-Imputer, and Miceforest.

Variable	Missing, %	Complete case	MICE[a]	Iterative-Imputer	Miceforest
Race	15%				
White, n (%)		1,519,326 (77.7)	1,793,167 (77.8)	1,527,855 (66.4)	1,792,102 (77.8)
Black, n (%)		65,716 (3.36)	77,767 (3.38)	391,756 (17.0)	78,624 (3.42)
Asian, n (%)		14,334 (0.73)	16,886 (0.73)	24,003 (1.04)	16,813 (0.73)
Indigenous, n (%)		7110 (0.36)	8561 (0.37)	7110 (0.31)	9399 (0.41)
Native Hawaiian or Pacific Islander, n (%)		1720 (0.09)	2036 (0.09)	4846 (0.21)	2235 (0.10)
Multiracial, n (%)		2405 (0.12)	2863 (0.12)	2674 (0.12)	2755 (0.11)
Other, n (%)		343,792 (17.6)	400,757 (17.4)	343,792 (14.9)	400,108 (17.4)
Total, N		1,954,403	2,302,036	2,302,036	2,302,036
Ethnicity	34%				
Hispanic or Latino, n (%)		33,652 (2.21)	57,552 (2.50)	33,652 (98.5)	75,031 (3.26)
Not Hispanic or Latino, n (%)		1,488,045 (97.8)	2,244,484 (97.5)	2,268,384 (1.46)	2,227,005 (96.7)
Total, N		1,521,697	2,302,036	2,302,036	2,302,036

[a]MICE: multiple imputation using chained equations.

XSL•FO
RenderX

Figure 1. Flu vaccination coverage proportions with 95% CIs by race across multiple imputation using chained equations, Iterative-Imputer, and Miceforest of the West Virginia Statewide Immunization Information System flu data, 2021 - 2022 (N=2,302,036).

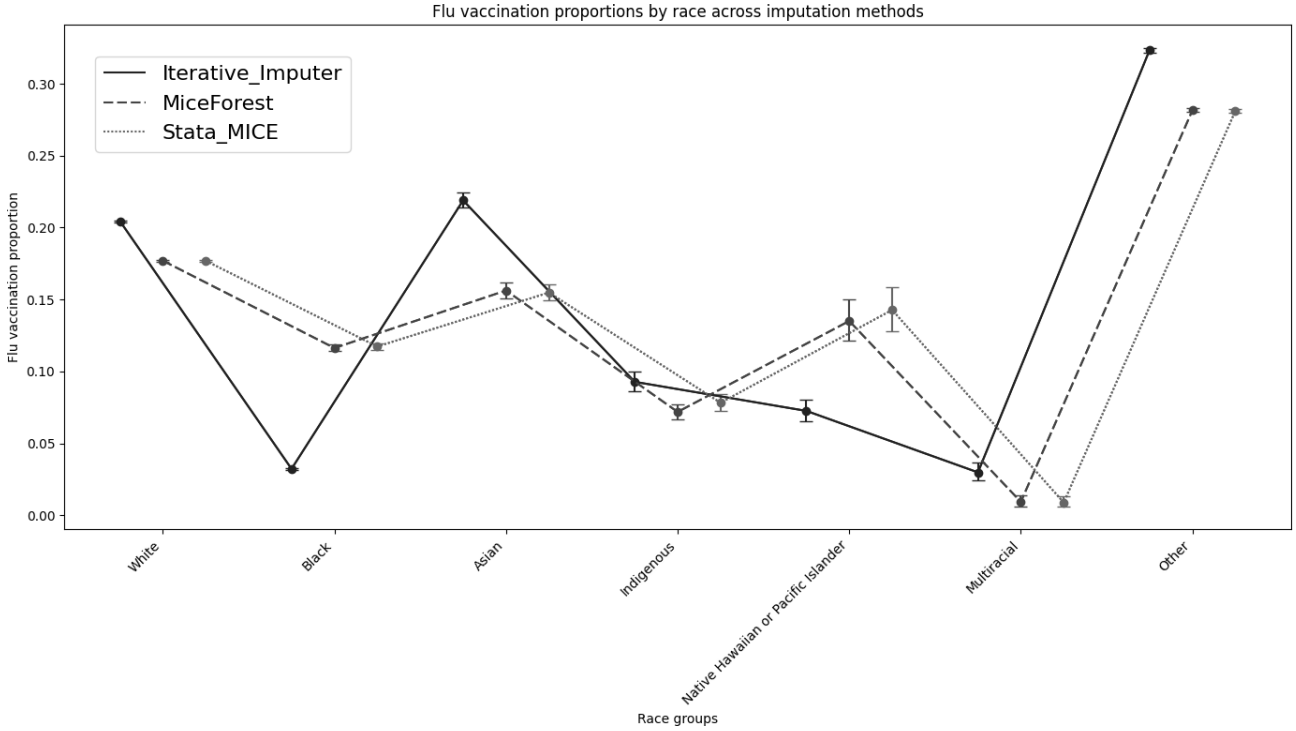
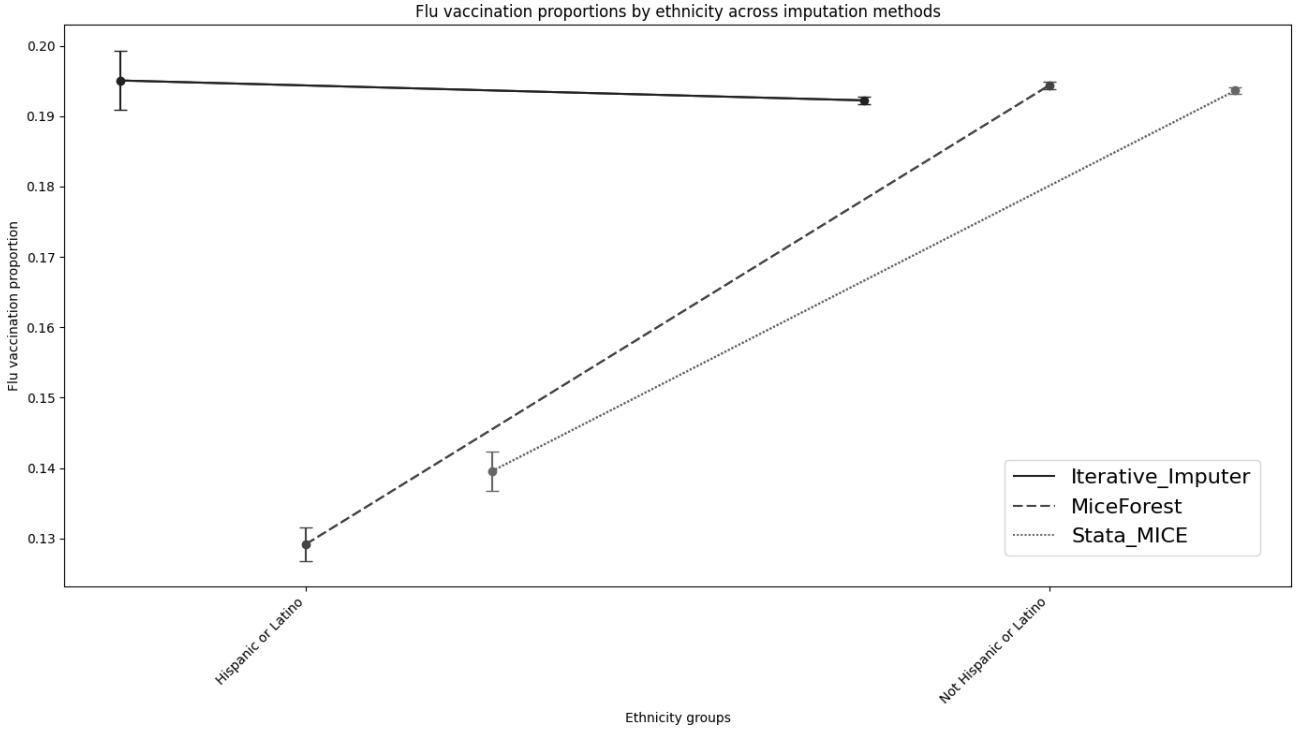


Figure 2. Flu vaccination coverage proportions with 95% CIs by ethnicity across multiple imputation using chained equations, Iterative-Imputer, and Miceforest of the West Virginia Statewide Immunization Information System flu data, 2021 - 2022 (N=2,302,036).



XSL•FO
RenderX

Table . 2021 - 2022 WVSIIS[a] flu vaccination proportions with 95% CIs by race and ethnicity after multiple imputation using chained equations, Iterative-Imputer, and Miceforest.

Variables	MICE[b] (95% CI)	Iterative-Imputer (95% CI)	Miceforest (95% CI)
Race			
White	0.177 (0.176 - 0.178)	0.205 (0.204 - 0.205)	0.177 (0.176 - 0.178)
Black	0.117 (0.115 - 0.120)	0.032 (0.032 - 0.033)	0.116 (0.114 - 0.119)
Asian	0.155 (0.150 - 0.160)	0.219 (0.214 - 0.224)	0.156 (0.151 - 0.162)
Indigenous	0.078 (0.073 - 0.084)	0.093 (0.086 - 0.100)	0.072 (0.067 - 0.077)
Native Hawaiian or Pacific Islander	0.143 (0.128 - 0.159)	0.073 (0.066 - 0.080)	0.135 (0.122 - 0.150)
Multiracial	0.009 (0.006 - 0.013)	0.030 (0.024 - 0.037)	0.009 (0.006 - 0.014)
Other	0.281 (0.280 - 0.282)	0.323 (0.322 - 0.325)	0.282 (0.280 - 0.283)
Ethnicity			
Hispanic or Latino	0.140 (0.137 - 0.142)	0.195 (0.191 - 0.199)	0.129 (0.127 - 0.132)
Not Hispanic or Latino	0.194 (0.193 - 0.194)	0.192 (0.192 - 0.193)	0.194 (0.194 - 0.195)

[a]WVSIIS: West Virginia Statewide Immunization Information System.

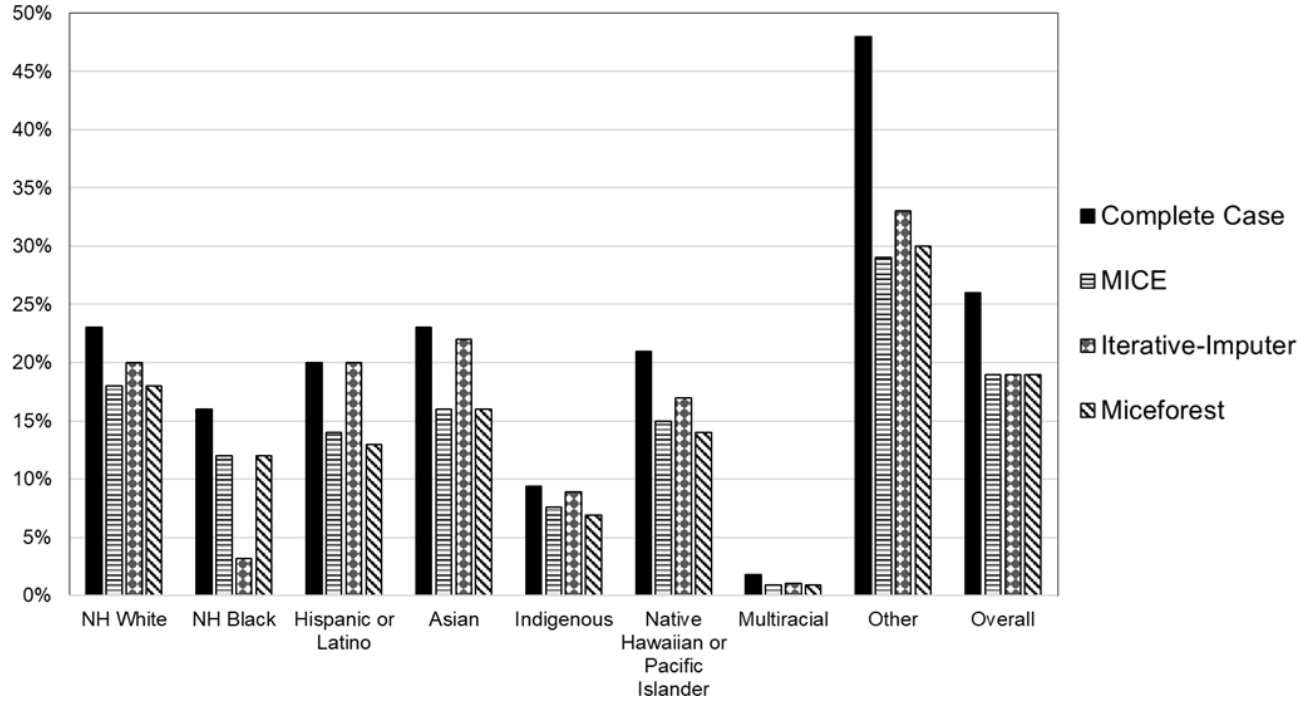
[b]MICE: multiple imputation using chained equations.

Table . Chi-square tests evaluating race and ethnicity distributions in the West Virginia Statewide Immunization Information System among the complete case dataset and the multiple imputation using chained equations-imputed, Iterative-Imputer-imputed, and Miceforest-imputed datasets.

Imputation method	Pearson chi-square	*P* value
MICE[a]	1.1e+04	<.001
Iterative-Imputer	1.8e+05	<.001
Miceforest	1.3e+04	<.001

[a]MICE: multiple imputation using chained equations.

Figure 3. Flu vaccination coverage proportions by race and ethnicity across multiple imputation using chained equations, Iterative-Imputer, and Miceforest of the West Virginia Statewide Immunization Information System flu data compared against the complete case analysis, 2021 - 2022 (N=2,302,036). NH Black: non-Hispanic Black; NH White: non=Hispanic White.



XSL•FO
RenderX

Table . 2021 - 2022 West Virginia Statewide Immunization Information System flu vaccination counts for complete case, multiple imputation using chained equations, Iterative-Imputer, and Miceforest by race and ethnicity.

Race and ethnicity	Complete case (N=378,574), n (%)	MICE[a] (N=442,679), n (%)	Iterative-Imputer (N=442,679), n (%)	Miceforest (N=442,679), n (%)
NH[b] White	274,597 (72.5)	313,530 (71.1)	309,150 (69.8)	312,686 (70.6)
NH[b] Black	7403 (1.96)	8931 (2.02)	12,510 (2.83)	8903 (2.01)
Hispanic or Latino	6565 (1.73)	8031 (1.81)	6565 (1.48)	9692 (2.12)
Asian	1967 (0.52)	2528 (0.57)	5190 (1.17)	2502 (0.57)
Indigenous	495 (0.13)	612 (0.14)	607 (0.14)	605 (0.14)
Native Hawaiian or Pacific Islander	204 (0.05)	246 (0.06)	248 (0.06)	252 (0.06)
Multiracial	16 (0.01)	24 (0.01)	24 (0.01)	22 (0.01)
Other	87,448 (23.1)	108,777 (24.5)	108,385 (24.5)	108,017 (24.4)

[a]MICE: multiple imputation using chained equations.

[b]NH: notHispanicc.

Table . Likelihood ratio test results comparing models with and without race and ethnicity as predictors of flu vaccination. Race and ethnicity were imputed in the West Virginia Statewide Immunization Information System using data collected between 2021 and 2022. Models were fit to datasets imputed using Stata's multiple imputation using chained equations (based on Rubin's rules), Iterative-Imputer (Python scikit-learn), and Miceforest (Python).

Variable and method	G-Statistic	*P* value	Degrees of freedom	Between-Imputation Variance[a]
Race				
Stata MICE[b]	26,452.659	<.001	—[c]	4103.382
Iterative-Imputer	128,280.274	<.001	6	—
Miceforest	26,891.527	<.001	6	—
Ethnicity				
Stata MICE	1142.231	<.001	—	1277.885
Iterative-Imputer	1.683	.19	1	—
Miceforest	2185.012	<.001	1	—

[a]Between-imputation variance is only applicable to Stata's MICE model.

[b]MICE: multiple imputation using chained equations

[c]Not applicable.

Iterative-Imputer

347,633 race categories that were previously missing were successfully imputed after completing the Iterative-Imputer. There were 0.26% additional White, 142.5% additional Black, 50.4% additional Asian, 0% additional Indigenous, 95.2% additional Native Hawaiian or Pacific Islander records, 10.6% additional Multiracial, and 0% additional Other records after imputations (Table 1).

Like MICE, 780,339 ethnicity categories that were previously missing were imputed after completing Iterative-Impute r. There were 0% additional Hispanic or Latino and 4.5% Not Hispanic or Latino records after imputations (Table 1).

After the Iterative Imputer, individual demographics did not remain proportional to the original dataset distributions (Figures 1 and 2 and Table 2). A chi-square test illustrated that there was a statistically significant difference between the complete case and the Iterative-Imputer estimation (*P*<.001; Table 3). Overall computational time was 2 minutes.

For the 2021 - 2022 flu season, the complete case analysis flu vaccine coverage rate was 26%. Using Iterative-Imputer, after imputations, West Virginia had an overall flu vaccine coverage rate of 19%. Flu vaccine coverage rates decreased when stratified by race and ethnicity when compared with a complete case analysis, with the most significant decreases observed in NH White (3%), NH Black (13%), Native Hawaiian or Pacific Islander (4%), and other (15%; Figure 3 and Table 4). After reconciling missing race and ethnicity, an additional 63,984 individuals were included in the analysis that were previously excluded from stratified analyses.

Likelihood ratio tests were used to determine the impact of race and ethnicity on vaccine uptake. Applying Iterative Imputer, race was significantly associated with flu vaccination status (G-statistic=128,280.27, *P*<.001), demonstrating a massive

XSL•FO
RenderX

effect (Table 5). For ethnicity, the Iterative-Imputer model produced a nonsignificant result (G-statistic=1.68, *P*=.195; Table 5).

Miceforest

347,633 race categories that were previously missing were imputed after completing Miceforest. There were 16.5% additional White records, 17.9% additional Black records, 15.9% additional Asian records, 27.7% additional Indigenous records, 26.0% additional Native Hawaiian or Pacific Islander records, 13.6% additional Multiracial records, and 15.1% additional Other records after imputations (Table 1).

780,339 ethnicity categories that were previously missing were imputed after completing Miceforest. There were 76.1% additional Hispanic or Latino records and 39.8% additional Not Hispanic or Latino records after imputations (Table 1).

After Miceforest, individual demographics remained proportional to the original dataset distributions (Figures 1 and 2 and Table 2). A chi-square test illustrated that there was a statistically significant difference between the complete case and Miceforest estimates (*P*<.001; Table 3). Overall computational time was 10 minutes.

For the 2021 - 2022 flu season, the complete case analysis flu vaccine coverage rate was 26%. Using miceforest, after imputations, West Virginia had an overall flu vaccine coverage rate of 19%. Flu vaccine coverage rates decreased when stratified by race and ethnicity when compared with a complete case analysis, with the most significant decreases observed in NH White (5%), NH Black (4%), Hispanic or Latino (7%), Asian (7%), Native Hawaiian or Pacific Islander (7%), and Other (18%; Figure 2 and Table 4). After reconciling missing race and ethnicity, an additional 63,984 individuals were included in the analysis that were previously excluded from stratified analyses.

Likelihood ratio tests were used to determine the impact of race and ethnicity on vaccine uptake. Using Miceforest, race was significantly associated with flu vaccination status (G-statistic=26,891.53; *P*<.001), demonstrating a large effect (Table 5). For ethnicity, the Miceforest model remained robust and was significantly associated with flu vaccination status (G-statistic=2185.01; *P*<.001; Table 5).

Discussion

Principal Findings

Our results highlight that the imputation method can profoundly change research findings. For race, all 3 methods show highly significant results and conclude that there are significant race disparities in flu vaccination after addressing missing data. G-tests performed on race and ethnicity and flu vaccination contingency tables showed consistent evidence of significant racial disparities across all imputation methods (*P*<.001). However, for ethnicity, Iterative-Imputer failed to distinguish any significant disproportions, while MICE and Miceforest found strong associations.

Of the 3 methods used to address missing data, we found that MICE and Miceforest had the best model fit and effect sizes. MICE and Miceforest were able to reconcile missing data with little bias, produced more stable and reliable imputations, and retained better preservation of original data relationships, producing similar conclusions; subsequently allowing for more sophisticated post hoc analyses. Iterative Imputer exhibited the most substantial deviations from the complete case analysis. While the majority of race and ethnicity categories retained proportions similar to those observed before imputation, the Black race category had a notable 142% increase post-Iterative-Imputer. Though all methods identify disparities for race, Iterative-Imputer suggests effects that are 5 times stronger than MICE and Miceforest (Iterative-Imputer G-statistic=128,280; MICE G-statistic=26,453; Miceforest G-statistic=26,892). Alternatively, Iterative-Imputer suggests no significant (*P*=.19) ethnicity disparities, while MICE and Miceforest distinguish strong disparities. Iterative-Imputer's results suggest that the model may be overfitting that data, intensifying present correlations, or generating false patterns in the imputed dataset. The significant change within a singular racial and ethnic group carries several potential implications: it may lead to neglecting crucial health disparities or greatly overvaluing disparities.

While MICE and Miceforest demonstrated more reliable and moderate effect sizes, the most striking difference between MICE and Miceforest was overall computational time. Large population surveillance-based datasets can be computationally expensive to impute [50]. While traditional MICE and Miceforest were proportionally similar to the complete case analysis and had similar G-statistics, the computational load was a significant limitation for further epidemiological analyses. Since Stata 17 operates locally, executing MICE required approximately 14 hours, using between 92% and 98% of the total central processing unit (CPU) capacity of the local system. In contrast, as Iterative-Imputer and Miceforest were processed on a cloud-based drive, the total computational time was reduced to under 10 minutes. Cloud-based computing extends several advantages, including enhanced scalability, reduced dependence on local servers, and the capability to allocate computational jobs efficiently [51,52]. As datasets become larger and more complex, they require larger memory and CPUs become necessary, and consequently, there will be computational limitations. By divesting resource-intensive tasks to the cloud, researchers and public health officials can alleviate processing constraints, enabling analogous performance of multiple tasks and minimizing downtime. This improved efficiency not only accelerates analyses but also improves overall productivity by removing lengthy processing delays.

Compared with the complete case analysis, the imputed dataset was 45% larger, and the estimated 2021 - 2022 flu vaccination coverage rate decreased from 26% to 19%. This resulted in distinct divergences in vaccination rates from those seen in the complete case analysis. Stratified flu vaccination coverage rates declined across all racial and ethnic categories following imputation. This is both expected, due to denominator inflation, and consistent with existing literature, which suggests that enhancements in data completeness often uncover underlying racial inequities [2,6-8,10,14,53].

XSL•FO
RenderX

In addition, the presence of the "Other" category may be artificially influencing the race missingness rate. Reports indicate that the increasing size of the "Other" category may be influenced by individuals choosing not to disclose their race, varying cultural perceptions of race, or vaccine providers selecting "Other" rather than leaving the race field blank [54,55]. Those who select the "Other" category get combined into a single, dissimilar category that does not reflect accurate identities and makes results for this group difficult to accurately decipher.

These findings align with the broader imputation literature, which demonstrates that not only does reconciling missing race and ethnicity data reveal under-reported disparities, but that imputation methodology can significantly influence the magnitude of these inequities. Similar patterns have been observed in findings from Labgold et al [2], Dorabawila et al [42], and Russ et al [41], where imputing missing race and ethnicity revealed greater disparities in health outcomes than previously understood, but the severity of these inequities fluctuated contingent upon parameters placed on the imputation method. This is consistent with theoretical bases proposing that the creation and assumptions of statistical models and algorithms can either dilute or intensify disparities, conditional on how missingness is managed and whose data are omitted [2,30,39,56]. The absence of demographic data in IISs carries significant public health implications. Missing data limits our capability to correctly assess and monitor vaccination coverage, potentially leading to an incomplete or inaccurate understanding of disparities in vaccination rates among different vulnerable groups. Consequently, public health interventions and policies may be based on flawed assumptions, undermining their effectiveness. Moreover, the inability to identify and address disparities hinders efforts to promote equity within communities. Poor data quality can also result in inefficient allocation of resources, both human and financial, further exacerbating inequities. Finally, inadequate data can erode public trust in health systems, as decisions based on incomplete information may be perceived as unreliable or inequitable. However, Miceforest successfully imputed missing values in a large public health dataset in the most time-efficient manner, and therefore, we were able to mitigate potential bias and increase our statistical power in our analyses. By allowing more complete and representative datasets, Miceforest encourages more equitable vaccination surveillance and permits fuller public health decision-making in a timely manner. State public health departments can use this technique to augment incomplete data to identify vaccine coverage gaps more accurately, concentrate on programming and resource allocation in underserved communities, and assess whether interventions are aiding populations who need it most. Our findings emphasize how multiple imputation is not only a statistical solution but a means for progressing equity in immunization programs and interventions.

There were several limitations to our approach. Due to high levels of missingness of variables in the dataset, there were limited informative variables for the imputations. The WVSIIS population denominator is higher than the census population, which can skew flu vaccination rates lower than they are. Race and ethnicity are self-reported, and they may not reflect the reality of demographic distribution in the state, especially with growing nuances of racial and ethnic identity.

Data are the foundation of all effective public health interventions, and missing data can reduce our comprehension of vaccination coverage, recognizing disparities, and creating successful public health programs. Without properly addressing missing data, vulnerable groups continue to be underserved. As demonstrated, different methods for reconciling missing data result in different assumptions regarding the data and the process. However, using Miceforest to reconcile missing demographic data poses a potential solution, offering a flexible, fast, and iterative approach that can improve data completeness while preserving underlying distributions and mitigating potential bias.

Future efforts should validate imputation values against other population-based surveillance systems, such as census data and claims data. In addition, because our imputations focused on one state's IIS, supplementary efforts should examine the performance of various multiple imputation methods across more diverse populations and geographies and examine the optimal frequency for updating imputation models. Further, based on our results, we recommend several suggestions for enhancing both the epidemiological approach and practice in immunization surveillance. Though more resource-intensive, preventing missing data initially is the most effective measure to handle missing data, especially for surveillance data that necessitates system-wide resolutions. This can be achieved through standardized race and ethnicity data collection protocols and data collection training in state IISs. Imputation should be a temporary answer for a larger, systematic problem. Public health agencies should be transparent in the imputation methods used in reports and present both complete case and imputed estimates to demonstrate the impact of missing data on vaccine equity reports.

Conclusions

Both MICE and Miceforest offer flexible and reliable approaches for imputing missing demographic data in IISs, outperforming Iterative-Imputer with regard to bias mitigation and distributional accuracy. These findings highlight that imputation method selection can greatly affect research outcomes, with repercussions for both statistical validity and public health decision-making and trust. While MICE and Miceforest yielded comparable effect sizes and preserved demographic proportions, MICE's substantial computational demands limit its scalability for large datasets, while Miceforest's ability to leverage cloud-based computing enhances efficiency by offloading resource-intensive tasks, enabling parallel execution, and minimizing processing delays.

Addressing missing data is both a methodological necessity and a public health imperative when handling large surveillance data. By improving data completeness and analytical precision, multiple imputation methods such as Miceforest can help illuminate disparities, reveal disparities, inform resource allocation, and promote equity in immunization and public health programs.

XSL•FO
RenderX

Acknowledgments

The results presented in this manuscript represent the collaborative efforts across the West Virginia Department of Health and Human Services, the West Virginia Immunization Department, and STChealth, West Virginia Statewide Immunization Information System (WVSIIS). The authors would like to extend their gratitude to Sawyer Koops, MS, Ilyssa Simmons, MPH, Kyle Freese, PhD, and countless other state, contract, federal, and STChealth staff whose expertise, leadership, and dedication were instrumental in this effort.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to data in WVSIIS being confidential and exempt from public disclosure but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Noppert GA, Zalla LC. Who counts and who gets counted? Health equity in infectious disease surveillance. Am J Public Health 2021 Jun;111(6):1004-1006. [doi: 10.2105/AJPH.2021.306249] [Medline: 33950717]
2. Labgold K, Hamid S, Shah S, et al. Estimating the unknown: greater racial and ethnic disparities in COVID-19 burden after accounting for missing race and ethnicity data. Epidemiology (Sunnyvale) 2021 Mar 1;32(2):157-161. [doi: 10.1097/EDE.0000000000001314] [Medline: 33323745]
3. Kazeminia M, Afshar ZM, Rajati M, Saeedi A, Rajati F. Evaluation of the acceptance rate of COVID-19 vaccine and its associated factors: a systematic review and meta-analysis. J Prev (2022) 2022 Aug;43(4):421-467. [doi: 10.1007/s10935-022-00684-1] [Medline: 35687259]
4. Andersen JA, Gloster E, Hall S, et al. Associations between COVID-19 vaccine uptake, race/ethnicity, and political party affiliation. J Behav Med 2023 Jun;46(3):525-531. [doi: 10.1007/s10865-022-00379-2] [Medline: 36417011]
5. Mahmud SM, Xu L, Hall LL, et al. Effect of race and ethnicity on influenza vaccine uptake among older US medicare beneficiaries: a record-linkage cohort study. Lancet Healthy Longev 2021 Mar;2(3):e143-e153. [doi: 10.1016/S2666-7568(20)30074-X] [Medline: 36098112]
6. Spangler KR, Levy JI, Fabian MP, et al. Missing race and ethnicity data among COVID-19 cases in Massachusetts. J Racial Ethn Health Disparities 2023 Aug;10(4):2071-2080. [doi: 10.1007/s40615-022-01387-3] [Medline: 36056195]
7. Krieger N, Testa C, Hanage WP, Chen JT. US racial and ethnic data for COVID-19 cases: still missing in action. The Lancet 2020 Nov;396(10261):e81. [doi: 10.1016/S0140-6736(20)32220-0]
8. Yoon P, Hall J, Fuld J, et al. Alternative methods for grouping race and ethnicity to monitor COVID-19 outcomes and vaccination coverage. MMWR Morb Mortal Wkly Rep 2021 Aug 13;70(32):1075-1080. [doi: 10.15585/mmwr.mm7032a2] [Medline: 34383729]
9. Stokes EK, Zambrano LD, Anderson KN, et al. Coronavirus disease 2019 case surveillance - United States, January 22-May 30, 2020. MMWR Morb Mortal Wkly Rep 2020 Jun 19;69(24):759-765. [doi: 10.15585/mmwr.mm6924e2] [Medline: 32555134]
10. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2 infection in the United States. Nat Commun 2020 Sep 9;11(1):4507. [doi: 10.1038/s41467-020-18272-4] [Medline: 32908126]
11. Lu PJ, Hung MC, Srivastav A, et al. Surveillance of vaccination coverage among adult populations -United States, 2018. MMWR Surveill Summ 2021 May 14;70(3):1-26. [doi: 10.15585/mmwr.ss7003a1] [Medline: 33983910]
12. Valente TW. Data collection and management. In: Evaluating Health Promotion Programs: Oxford University Press; 2002:123-146.
13. Weston BW. Blind spots: biases in prehospital race and ethnicity recording. Prehosp Emerg Care 2023;27(8):1072-1075. [doi: 10.1080/10903127.2023.2175089] [Medline: 36735657]
14. Sholle ET, Pinheiro LC, Adekkanattu P, et al. Underserved populations with missing race ethnicity data differ significantly from those with structured race/ethnicity documentation. J Am Med Inform Assoc 2019 Aug 1;26(8-9):722-729. [doi: 10.1093/jamia/ocz040] [Medline: 31329882]
15. Rubin DB. Multiple imputations in sample surveys—a phenomenological bayesian approach to nonresponse. Presented at: Proceedings of the Survey Research Methods Section of the American Statistical Association; 1978 p. 20-28.
16. Rubin DB. Bayesian inference for causal effects: the role of randomization. Ann Statist 1978;6(1). [doi: 10.1214/aos/1176344064]
17. Rubin DB. Inference and missing data. Biometrika 1976 Dec;63(3):581. [doi: 10.2307/2335739]
18. Rubin DB. Multiple Imputation After 18+ Years. J Am Stat Assoc 1996 Jun;91(434):473. [doi: 10.2307/2291635]

XSL•FO
RenderX

19. Li KH, Raghunathan TE, Rubin DB. Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution. J Am Stat Assoc 1991 Dec;86(416):1065. [doi: 10.2307/2290525]
20. Meng XL, Rubin DB. Performing likelihood ratio tests with multiply-imputed data sets. Biometrika 1992;79(1):103-111. [doi: 10.1093/biomet/79.1.103]
21. Chan KW, Meng XL. Multiple improvements of multiple imputation likelihood ratio tests. STAT SINICA 2022. [doi: 10.5705/ss.202019.0314]
22. Chan KW. General and feasible tests with multiply-imputed datasets. Ann Statist 2022;50(2). [doi: 10.1214/21-AOS2132]
23. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res 2011 Mar;20(1):40-49. [doi: 10.1002/mpr.329] [Medline: 21499542]
24. Bouhlila DS, Sellaouti F. Multiple imputation using chained equations for missing data in TIMSS: a case study. Large-scale Assess Educ 2013 Dec;1(1). [doi: 10.1186/2196-0739-1-4]
25. Wang H, Tang J, Wu M, Wang X, Zhang T. Application of machine learning missing data imputation techniques in clinical decision making: taking the discharge assessment of patients with spontaneous supratentorial intracerebral hemorrhage as an example. BMC Med Inform Decis Mak 2022 Jan 13;22(1):13. [doi: 10.1186/s12911-022-01752-6] [Medline: 35027065]
26. Getz K, Hubbard RA, Linn KA. Performance of multiple imputation using modern machine learning methods in electronic health records data. Epidemiology (Sunnyvale) 2023 Mar 1;34(2):206-215. [doi: 10.1097/EDE.0000000000001578] [Medline: 36722803]
27. Emmanuel T, Maupong T, Mpoeleng D, Semong T, Mphago B, Tabona O. A survey on missing data in machine learning. J Big Data 2021;8(1):140. [doi: 10.1186/s40537-021-00516-9] [Medline: 34722113]
28. Zhang X, Yan C, Gao C, Malin BA, Chen Y. Predicting missing values in medical data via XGBoost regression. J Healthc Inform Res 2020 Dec;4(4):383-394. [doi: 10.1007/s41666-020-00077-1] [Medline: 33283143]
29. Zhao X, Shen W, Wang G, Schwenker F, Schwenker F. Early prediction of sepsis based on machine learning algorithm. Comput Intell Neurosci 2021;2021(1):6522633. [doi: 10.1155/2021/6522633] [Medline: 34675971]
30. Shao L, Chen W. Coal and gas outburst prediction model based on miceforest filling and PHHO–KELM. Processes (Basel) 2023;11(9):2722. [doi: 10.3390/pr11092722]
31. AnotherSamWilson. miceforest: Fast, Memory Efficient Imputation with LightGBM.: GitHub; 2020 Aug 30. URL: https://github.com/AnotherSamWilson/miceforest [accessed 2025-08-19]
32. Pham HT, Do T, Baek J, et al. Handling missing data in COVID-19 incidence estimation: secondary data analysis. JMIR Public Health Surveill 2024 Aug 20;10:e53719. [doi: 10.2196/53719] [Medline: 39166439]
33. Feng S, Hategeka C, Grépin KA. Addressing missing values in routine health information system data: an evaluation of imputation methods using data from the Democratic Republic of the Congo during the COVID-19 pandemic. Popul Health Metr 2021 Nov 4;19(1):44. [doi: 10.1186/s12963-021-00274-z] [Medline: 34736462]
34. Branham DK, Finegold K, Chen L, et al. Trends in missing race and ethnicity information after imputation in HealthCare.gov marketplace enrollment data, 2015-2021. JAMA Netw Open 2022 Jun 1;5(6):e2216715. [doi: 10.1001/jamanetworkopen.2022.16715] [Medline: 35687340]
35. Scroggins JK, Hulchafo II, Topaz M, Cato K, Barcelona V. Addressing bias in preterm birth research: the role of advanced imputation techniques for missing race and ethnicity in perinatal health data. Ann Epidemiol 2024 Jun;94:120-126. [doi: 10.1016/j.annepidem.2024.05.003] [Medline: 38734192]
36. Sorbero ME, Euller R, Kofner A, Elliott MN. Imputation of race and ethnicity in health insurance marketplace enrollment data, 2015-2022 open enrollment periods. Rand Health Q 2022 Nov;10(1):4. [doi: 10.7249/RRA1853-1]
37. Elliott MN, Morrison PA, Fremont A, McCaffrey DF, Pantoja P, Lurie N. Using the census bureau's surname list to improve estimates of race/ethnicity and associated disparities. Health Serv Outcomes Res Method 2009 Jun;9(2):69-83. [doi: 10.1007/s10742-009-0047-1]
38. Kozlowski D, Murray DS, Bell A, et al. Avoiding bias when inferring race using name-based approaches. PLoS ONE 2022;17(3):e0264270. [doi: 10.1371/journal.pone.0264270] [Medline: 35231059]
39. Zhang Y, Kissin DM, Liao KJ, DeSantis CE, Yartel AK, Gutman R. Multiple imputation of missing race/ethnicity information in the national assisted reproductive technology surveillance system. J Womens Health (Larchmt) 2024 Mar;33(3):328-338. [doi: 10.1089/jwh.2023.0267] [Medline: 38112534]
40. Zhang G, Rose CE, Zhang Y, et al. Multiple imputation of missing race and ethnicity in CDC COVID-19 case-level surveillance data. Int J Stat Med Res 2022 Jan 28;11:1-11. [doi: 10.6000/1929-6029.2022.11.01] [Medline: 35368775]
41. Russ S, Bramley J, Liu Y, Boyce I. Bolstering the measurement of racial inequity of COVID-19 vaccine uptake. Vaccines (Basel) 2023 Apr 21;11(4):876. [doi: 10.3390/vaccines11040876] [Medline: 37112788]
42. Dorabawila V, Hoen R, Hoefer D. Leveraging multiple administrative data sources to reduce missing race and ethnicity data: a descriptive epidemiology cross-sectional study of COVID-19 case relative rates. J Racial Ethn Health Disparities 2024 Oct 22;22. [doi: 10.1007/s40615-024-02211-w] [Medline: 39436568]
43. How we define rural. Health Resource & Services Administration.: HRSA; 2025 Feb. URL: https://www.hrsa.gov/rural-health/about-us/what-is-rural [accessed 2025-06-05]
44. West Virginia: 2020 Census. United States Census Bureau. 2021 Aug 25. URL: https://www.census.gov/library/stories/state-by-state/west-virginia.html [accessed 2025-06-05]

XSL•FO
RenderX

45. CDC/ATSDR Social Vulnerability Index 2020 Database West Virginia. Agency for Toxic Substances and Disease Registry. Geospatial Research, Analysis, and Services Program. Centers for Disease Control and Prevention. URL: https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html [accessed 2025-06-05]
46. Iterativeimputer. scikit-learn. URL: https://scikit-learn.org/stable/modules/generated/sklearn.impute.IterativeImputer.html#sklearn.impute.IterativeImputer [accessed 2024-11-14]
47. Raghunathan TE, Solenberger PW, Hoewyk J. IVEware imputation and Variance Estimation Software User Guide.: Survey Research Center, Institute for Social Research University of Michigan URL: https://www.src.isr.umich.edu/wp-content/uploads/iveware/v0.1/Documentation/ive_user.pdf [accessed 2023-01-04]
48. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007 Sep;8(3):206-213. [doi: 10.1007/s11121-007-0070-9] [Medline: 17549635]
49. IIS policies: West Virginia. Centers for Disease Control and Prevention. 2024 Aug 9. URL: https://www.cdc.gov/iis/policy-legislation/west-virginia.html [accessed 2025-07-01]
50. Al-Jumaili AHA, Muniyandi RC, Hasan MK, Paw JKS, Singh MJ. Big data analytics using cloud computing based frameworks for power management systems: status, constraints, and future recommendations. Sensors (Basel) 2023 Mar 8;23(6):2952. [doi: 10.3390/s23062952] [Medline: 36991663]
51. Lebeda FJ, Zalatoris JJ, Scheerer JB. Government cloud computing policies: potential opportunities for advancing military biomedical research. Mil Med 2018 Nov 1;183(11-12):e438-e447. [doi: 10.1093/milmed/usx114] [Medline: 29425378]
52. Ahmadi M, Aslani N. Capabilities and advantages of cloud computing in the implementation of electronic health record. Acta Inform Med 2018;26(1):24-28. [doi: 10.5455/aim.2018.26.24-28] [Medline: 29719309]
53. Coelho R, Rocha R, Hone T. Improvements in data completeness in health information systems reveal racial inequalities: longitudinal national data from hospital admissions in Brazil 2010-2022. Int J Equity Health 2024 Jul 18;23(1):143. [doi: 10.1186/s12939-024-02214-3] [Medline: 39026324]
54. Holloway KR, Radack J, Barreto A, et al. The “other” race category on birth certificates and its impact on analyses of preterm birth inequity. J Perinatol 2025 Mar;45(3):372-377. [doi: 10.1038/s41372-024-02123-x] [Medline: 39304729]
55. Woolverton GA, Marks AK. “I just check ‘other’”: evidence to support expanding the measurement inclusivity and equity of ethnicity/race and cultural identifications of US adolescents. Cultur Divers Ethnic Minor Psychol 2023 Jan;29(1):64-73. [doi: 10.1037/cdp0000360] [Medline: 34351178]
56. Ansari B, Hart-Malloy R, Rosenberg ES, Trigg M, Martin EG. Modeling the potential impact of missing race and ethnicity data in infectious disease surveillance systems on disparity measures: scenario analysis of different imputation strategies. JMIR Public Health Surveill 2022 Nov 9;8(11):e38037. [doi: 10.2196/38037] [Medline: 36350701]

Abbreviations

ATSDR: Agency for Toxic Substances and Disease Registry
BISG: Bayesian Improved Surname and Geocoding
CDC: Centers for Disease Control and Prevention
CPU: central processing unit
FORHP: Federal Office of Rural Healthy Policy
IIS: immunization information systems
IRB: institutional review board
MICE: multiple imputation using chained equations
ML: machine learning
NH: non-Hispanic
SVI: Social Vulnerability Index
WVSIIS: West Virginia Statewide Immunization Information System

Edited by A Mavragani, T Sanchez; submitted 13.03.25; peer-reviewed by KW Chan, L Ou; revised version received 01.07.25; accepted 04.07.25; published 26.08.25.

Please cite as:
Brown S, Kudia O, Kleine K, Kidd B, Wines R, Meckes N
Comparing Multiple Imputation Methods to Address Missing Patient Demographics in Immunization Information Systems: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e73916
URL: https://publichealth.jmir.org/2025/1/e73916
doi:10.2196/73916

XSL•FO
RenderX

© Sara Brown, Ousswa Kudia, Kaye Kleine, Bryndan Kidd, Robert Wines, Nathanael Meckes. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A "Pandemic-Proof" Methodology for Outbreak Detection Adapted From COVID-19's Impact on Notifications of Infectious Diseases in the Netherlands: Surveillance Study

Babette van Deursen[1,2], MSc; Stijn Raven[1,2], PhD; Wolfer van den Bosch[2], MSc; Cornelia H M van Jaarsveld[1], PhD; Aura Timen[1], PhD

[1]Department of Primary and Community Care, Research Institute for Medical Innovation, Radboud University Medical Center, Geert Grooteplein 21, Nijmegen, The Netherlands

[2]Department of Infectious Disease Control, Public Health Service region Utrecht, Zeist, The Netherlands

Corresponding Author:
Babette van Deursen, MSc
Department of Primary and Community Care, Research Institute for Medical Innovation, Radboud University Medical Center, Geert Grooteplein 21, Nijmegen, The Netherlands

Abstract

Background: Reporting of notifiable infectious diseases was overall impacted by the COVID-19 pandemic. This could affect disease surveillance and thus, outbreak detection, as it is based on historical data. For effective outbreak detection, it is crucial that communicable disease control professionals can rely on accurate and valuable alarm thresholds to ensure timely and adequate response to potential outbreaks.

Objective: In this study, we take the first steps in the development of a methodology that adjusts for the impact of the COVID-19 pandemic on the number of notifications of notifiable infectious diseases and provides corrected alarm thresholds for outbreak detection. We identify which infectious diseases were affected by the COVID-19 pandemic, assess the duration of these effects, and explore potential correction methods to improve outbreak detection.

Methods: We analyzed cases of 25 notifiable infectious diseases reported from 2015 - 2023 in the Netherlands. Negative binomial regression was used to calculate the incidence rate ratios for each period: pre-COVID-19 pandemic, COVID 2020, COVID 2021, COVID 2022, and post-COVID-19 pandemic. To address the decrease in notifications during the COVID-19 pandemic, we tested 3 correction methods: (1) recoding COVID-19 years as missing; (2) imputing with the last pre-COVID-19 pandemic observation; and (3) imputing the historical moving average.

Results: Between 2015 - 2023, a total of 74,990 notifications were reported in the Netherlands, of which 9,836 notifications (13%) occurred during the COVID-19 pandemic. Malaria, typhoid fever, hepatitis A, meningococcal infection, paratyphoid fever, Q-fever, shigellosis, measles, mumps, and pertussis had significantly lower notifications during the COVID-19 pandemic, but the duration and magnitude of the effect differed among the infections. The effect of the COVID-19 pandemic on the notifications of malaria (incidence rate ratio [IRR] 0.17, 95% CI 0.08 - 0.36) and typhoid fever (IRR 0.12, 95% CI 0.02 - 0.55) was only seen in 2020, while the notifications for measles (IRR 0.34, 95% CI 0.15 - 0.75) and pertussis (IRR 0.34, 95% CI 0.20 - 0.60) were still significantly lower in the post-COVID-19 pandemic period. In addition, the newly calculated alarm thresholds showed a noticeable difference compared with the original unadjusted alarm thresholds. However, the variation among the 3 different corrected alarm thresholds was not substantial.

Conclusions: During the COVID-19 pandemic, notifications of 10 infectious diseases declined significantly. The duration of this decline varied among infections, highlighting the need for pattern-specific adjustments. Our study demonstrates that accounting for the reduced notifications impacts alarm threshold calculations for outbreak detection. We therefore recommend amending the alarm thresholds to account for this impact to ensure reliable outbreak detection, so that communicable disease control professionals can act in a timely and data-driven manner. The next step in the development of the "pandemic-proof" methodology is to determine which correction method is most suitable. Further validation by communicable disease control professionals is essential to assess the applicability of the adjusted alarm thresholds for outbreak detection.

(JMIR Public Health Surveill 2025;11:e73953) doi:10.2196/73953

KEYWORDS

public health surveillance; COVID-19 pandemic; infectious diseases; epidemiology; outbreak detection; routine surveillance data

XSL•FO
RenderX

Introduction

In the Netherlands, it is mandatory under the Dutch Public Health Act for all medical doctors and laboratories to notify certain infectious diseases to the public health services (PHS), who then report it to the National Institute for Public Health and the Environment (RIVM) [1]. Among others, the RIVM is responsible for the surveillance of those notifiable infectious diseases at the national level and the PHS at a regional level. A critical aspect of surveillance is outbreak detection, for which various methods establish an alarm threshold for the number of notifications that signal for a potential outbreak detection [2-6]. Regarding these methods, robust historical data are essential for optimizing alarm thresholds accuracy. This allows communicable disease control professionals to implement preventive actions or control measures in a timely manner, and at the same time minimizing false positive signals, to effectively interrupt transmission and thus contain an outbreak.

The COVID-19 pandemic had an impact on the notifications of other infectious diseases. Several countries worldwide reported less notifications than expected during the pandemic on all types of infectious diseases, not only respiratory infections [7-16]. Nonpharmaceutical interventions (NPIs) that were implemented to decrease the transmission of SARS-CoV-2 most likely influenced the transmission of other infectious diseases, especially respiratory infections. Some studies suggest that SARS-CoV-2 was in competition with other respiratory infections, which suppressed the transmission of the lesser dominant respiratory infection [17-19]. In addition, the pandemic disrupted health care systems and altered health-seeking behavior, resulting in underdiagnosing and underreporting and thus, a decrease in notifications of other infectious diseases [20].

Since outbreak detection thresholds are based on historical data and considering reports of the impact of the COVID-19 pandemic on notifications of other infectious diseases globally, it is essential to investigate what the effect of the COVID-19 pandemic was on the number of notifications. This includes investigating which infectious diseases were affected and the duration of these effects, and whether these trends align with the aforementioned hypothesized causes. This is crucial for understanding the impact of the COVID-19 pandemic on outbreak detection, as it still remains unknown. To our knowledge, no previous studies have been conducted on how to adjust for the impact of the COVID-19 pandemic on the historical data and thus, establish accurate alarm thresholds for outbreak detection. The development of a methodology that adjusts for the impact of the COVID-19 pandemic on the number of notifications of notifiable infectious diseases requires a stepwise approach: first identifying which infectious diseases were affected by the COVID-19 pandemic, then developing a novel methodology, and finally validating the novel methodology for its applicability in effective outbreak detection. This future "pandemic-proof" methodology will provide corrected historical averages and alarm threshold as indicators of changes in infectious disease trends, to monitor and signal for upcoming infectious disease outbreaks.

In this study, we focus on the first steps of this process toward a "pandemic-proof" methodology by examining which infectious diseases were affected by the COVID-19 pandemic, the duration of these effects, and exploring approaches to correct for them in outbreak detection. The validation of any correction method is beyond the scope of this study, as it first requires a comprehensive understanding of the impact of the COVID-19 pandemic on infectious disease notifications, as well as of the potential correction methods for addressing this.

Methods

Study Design and Data Sources

To determine which notifiable infectious diseases were affected by the COVID-19 pandemic on the number of notifications in the Netherlands, we used anonymized routinely collected data on all 45 notifiable infectious diseases from the database "Osiris" collected by National Institute of Public Health and Environment. Notifiable clusters of infections, such as Methicillin-resistant Staphylococcus aureus clusters, were left out of the dataset. All notifications (except COVID-19 pandemic) between January 2015 and December 2023 based on the date of disease onset were included. In addition, information on age, sex, region (the regional distribution is the PHS catchment area), and type of infection for each notification was extracted from the database. Infectious diseases with lacking historical data (eg, notifiable since 2019 or later on) were excluded from the analysis (n=4). Furthermore, we also excluded infectious diseases that were only notified 20 times or less during the entire study period (n=16). We investigated the notifications of 25 infectious diseases.

Definition of Study Periods

To investigate the duration of the COVID-19 pandemic effect, we split our study period into different periods: pre-COVID-19 pandemic, COVID 2020, COVID 2021, COVID 2022, and post-COVID-19 pandemic. Pre-COVID-19 pandemic was defined as January 2015-February 2020, as the first case in the Netherlands was reported at the end of February 2020 and the first national NPIs were implemented mid-March. COVID 2020 was defined as March 2020-December 2020, COVID 2021 included the whole year of 2021, COVID 2022 was defined as January 2022-August 2022 and post-COVID-19 pandemic was defined as September 2022-December 2023. We chose September 2022 as the ending date of the COVID-19 pandemic, as the COVID-19 stringency index was at its lowest for the Netherlands and all NPIs were not in place anymore [21]. Since then, there were no NPIs implemented in the Netherlands.

Hypothesis Development

First, we hypothesized what the potential impact of the COVID-19 pandemic was on the number of notifications and trends for every notifiable infectious disease. We based our hypotheses on 4 factors: effect of NPIs (eg, physical distancing and travel bans), variants of SARS-CoV-2 interference (competing between infections), demographic determinants (ie, region, age, and gender), health care use and access to diagnostics. All factors were reviewed by the team with extensive knowledge on infectious disease control (BvD,

XSL•FO
RenderX

infectious disease epidemiologist; SR, medical doctor in communicable disease control; CHMvJ, epidemiologist; and AT, professor and medical doctor in communicable disease control) and discrepancies were discussed until consensus was reached. The latter potential factor (access to diagnostics) was additionally reviewed by a microbiologist. The effect of the NPIs was leading in this decision-making, whether we expected an effect on the notifications or not (Multimedia Appendix 1).

Statistical Analysis

For the descriptive analysis, we calculated the age, sex, and region distribution for each infection in each study period (pre-COVID-19 pandemic, COVID 2020, COVID 2021, COVID 2022, and post-COVID-19 pandemic). To examine which infectious diseases were affected and the duration of these effects, we performed negative binomial regression to determine the incidence rate ratios (IRR) and their respective 95% CI, using pre-COVID-19 pandemic as reference period. The included months of the reference period corresponded to the respective months of the investigated study period to account for seasonality. For example, the COVID 2020 period was compared with the historical months March until December from the years 2015 - 2019. We defined the impact of the COVID-19 pandemic on infectious disease notifications as a significantly lower IRR than the pre-COVID period, with the effect beginning in the COVID 2020 period, as we considered that substantial effects of the pandemic should be visible in that period. Effects that emerged later are more difficult to directly attribute to the COVID-19 pandemic.

Threshold Recalibration

For the exploration of recalibrating the alarm threshold, we reviewed literature and expert opinions on the correction for the decrease in notifications. The correction was integrated into an alarm threshold for outbreak detection, which is mostly based on the method of Stroup that focuses on the moving average [3]. The moving average for a month is calculated by the means of the preceding, current, and upcoming month of the last 5 years, with a corresponding moving SD and moving SE. For example, the moving average of July 2021 is calculated by the means of the months June, July, and August of the years 2016 - 2020. If there was an outbreak in one of those months (defined as: moving average+[2*SD]), the number of cases was set back to the moving average of that month. Thus, we calculated the alarm threshold with the moving average that is corrected for outbreaks, from now on called baseline. The original alarm threshold itself is calculated as: baseline+(1.96 * SQRT[SD^2+SE^2]).

After assessing which notifiable infectious diseases were affected, we tested the following 3 methods for correction: (1) treating disease counts as missing during the correction period; (2) replacing disease counts with the last value (outbreak corrected) for the same month in the year preceding the correction period; and (3) replacing disease counts with the baseline value (as mentioned above) for the same month in the year preceding the correction period. We calculated new thresholds for infectious diseases impacted by the COVID-19 pandemic, focusing on those with a decrease in notifications, and only for the corresponding affected period. We used the original alarm threshold method (as described above), as a reference. We refer to this further on as the unadjusted threshold method, as it is not corrected for a decreased number of notifications.

Ethical Considerations

We used anonymized notification data collected by the RIVM in accordance with the Dutch Public Health Act. The Medical Ethics Committee of Radboudumc declared that an ethical approval was not required for this study, as the study did not fall under the scope of the Medical Research Involving Human Subjects Act.

Results

Overview

In the entire study period (2015 - 2023), 74,990 notifications were reported of the 25 included infectious diseases. Of those 74,990 notifications, 73% (n=54,430) were notified in the pre-COVID-19 pandemic period, 13% (n=9836) during the COVID-19 pandemic, and 14% (n=10,724) in the post-COVID-19 pandemic period. Notification numbers and characteristics for each notifiable infectious disease are provided in (Multimedia Appendix 2). We found no difference regarding age, sex, or region distribution between the periods.

Affected by the COVID-19 Pandemic

We hypothesized that the notifications of Creutzfeldt-Jakob disease, Q-fever, hantavirus, listeriosis, and psittacosis were not affected by the COVID-19 pandemic (Multimedia Appendix 1). For all other 20 notifiable infectious diseases, we expected a decrease in notifications due to implementation of the NPIs (egsimilar transmission routes), interference of other infections, changing demographics, change in health care seeking behavior or use of diagnostics.

Based on the IRRs, we found that the number of notifications of 11 infectious diseases was significantly affected by the COVID-19 pandemic (Table 1 and Figure 1). However, the duration of affected months differed. In Figure 1, we present the IRRs per COVID-19 period for all infectious diseases, indicating whether a significant effect was found and how this effect evolved over time.

XSL•FO
RenderX

Table . Incidence rate ratios (IRRs) for notifications in the COVID-19 pandemic and post-COVID-19 pandemic periods compared with pre-COVID-19 pandemic for 25 notifiable infectious diseases in the Netherlands, January 2015-December 2023.

Infectious disease	Pre-COVID-19 pandemic (January 2015–February 2020)		COVID 2020 (March-December)		COVID 2021 (January-December)		COVID 2022 (January-August)		Post-COVID-19 pandemic (September 2022–December 2023)	
	Notifications, n	IRR (95% CI)	Notifications, n	IRR (95% CI)	Notifications, n	IRR (95% CI)	Notifications, n	IRR (95% CI)	Notifications, n	IRR (95% CI)
Effect during COVID 2020										
Typhoid fever	106	Ref.[a]	2	0.12 (0.02 - 0.55)	15	0.73 (0.32 - 1.66)	93	6.55 (2.89 - 14.86)	32	1.17 (0.59 - 2.30)
Malaria	1281	Ref.	37	0.17 (0.08 - 0.36)	154	0.62 (0.32 - 1.17)	138	0.87 (0.40 - 1.91)	330	1.00 (0.57 - 1.75)
Effect during COVID 2020 and 2021										
Hepatitis A	900	Ref.	29	0.19 (0.09 - 0.41)	78	0.44 (0.23 - 0.86)	60	0.54 (0.24 - 1.20)	185	0.80 (0.45 - 1.41)
Meningococcal infection	837	Ref.	44	0.34 (0.16 - 0.72)	33	0.20 (0.10 - 0.41)	51	0.47 (0.21 - 1.06)	175	0.81 (0.46 - 1.44)
Paratyphoid fever	210	Ref.	4	0.11 (0.03 - 0.36)	17	0.41 (0.19 - 0.91)	19	0.66 (0.27 - 1.62)	67	1.24 (0.67 - 2.29)
Q-fever	92	Ref.	5	0.30 (0.10 - 0.93)	6	0.33 (0.12 - 0.94)	8	0.65 (0.22 - 1.86)	8	0.34 (0.14 - 0.84)
Shigellosis	2483	Ref.	129	0.30 (0.15 - 0.61)	221	0.46 (0.24 - 0.86)	243	0.80 (0.37 - 1.74)	807	1.26 (0.72 - 2.19)
Hib[b]	174	Ref.	58	2.25 (1.07 - 4.74)	67	2.06 (1.04 - 4.07)	36	2.00 (0.85 - 4.69)	78	1.74 (0.94 - 3.20)
Effect during COVID 2020 – 2022										
Mumps	452	Ref.	20	0.27 (0.12 - 0.61)	1	0.01 (0.00 - 0.10)	7	0.13 (0.04 - 0.37)	93	0.80 (0.44 - 1.44)
Pertussis	29,037	Ref.	485	0.10 (0.05 - 0.20)	79	0.01 (0.00 - 0.03)	56	0.01 (0.00 - 0.03)	2570	0.34 (0.20 - 0.60)
Measles	139	Ref.	0	<0.001 (0.00 - 0.00)	0	<0.001 (0.00 - 0.00)	1	0.04 (0.01 - 0.34)	12	0.34 (0.15 - 0.75)
No effect										
Legionellosis	2722	Ref.	413	0.85 (0.43 - 1.70)	670	1.26 (0.67 - 2.35)	405	1.17 (0.54 - 2.52)	1162	1.66 (0.95 - 2.88)
Brucellosis	29	Ref.	2	0.42 (0.09 - 2.05)	2	0.37 (0.08 - 1.78)	3	0.88 (0.21 - 3.74)	7	0.94 (0.35 - 2.52)
Diphtheria	14	Ref.	2	0.83 (0.16 - 4.31)	0	<0.001 (0.00 - 0.00)	1	1.00 (0.10 - 9.75)	21	5.81 (2.43 - 13.89)
Hepatitis B acute	561	Ref.	78	0.85 (0.41 - 1.74)	72	0.66 (0.34 - 1.29)	50	0.71 (0.31 - 1.60)	129	0.89 (0.50 - 1.60)

XSL•FO
RenderX

Infectious disease	Pre-COVID-19 pandemic (January 2015–February 2020)		COVID 2020 (March-December)		COVID 2021 (January-December)		COVID 2022 (January-August)		Post-COVID-19 pandemic (September 2022–December 2023)	
	Notifications, n	IRR (95% CI)	Notifications, n	IRR (95% CI)	Notifications, n	IRR (95% CI)	Notifications, n	IRR (95% CI)	Notifications, n	IRR (95% CI)
Hepatitis B chronic	5356	Ref.	553	0.64 (0.32 - 1.28)	741	0.71 (0.38 - 1.33)	519	0.76 (0.36 - 1.64)	1129	0.82 (0.47 - 1.42)
Leptospirosis	423	Ref.	57	0.71 (0.34 - 1.48)	58	0.69 (0.35 - 1.37)	51	0.99 (0.43 - 2.25)	158	1.45 (0.81 - 2.58)
Listeriosis	482	Ref.	83	1.03 (0.50 - 2.10)	102	1.09 (0.56 - 2.09)	69	1.11 (0.50 - 2.47)	129	1.04 (0.58 - 1.86)
Psittacosis	342	Ref.	67	1.17 (0.57 - 2.43)	56	0.89 (0.45 - 1.76)	33	0.78 (0.34 - 1.82)	103	1.17 (0.65 - 2.11)
STEC[c]	2707	Ref.	286	0.60 (0.30 - 1.20)	481	0.90 (0.48 - 1.69)	387	1.03 (0.48 - 2.22)	772	1.11 (0.63 - 1.93)
Tuberculosis	4189	Ref.	506	0.73 (0.37 - 1.45)	673	0.83 (0.44 - 1.54)	423	0.74 (0.34 - 1.59)	925	0.86 (0.49 - 1.49)
Creutzfeldt-Jakob classic	154	Ref.	22	0.85 (0.38 - 1.93)	24	0.80 (0.38 - 1.70)	10	0.54 (0.20 - 1.48)	50	1.26 (0.67 - 2.38)
Hantavirus	179	Ref.	15	0.48 (0.20 - 1.14)	39	1.12 (0.55 - 2.28)	7	0.27 (0.09 - 0.79)	34	0.74 (0.38 - 1.43)
Hepatitis C acute	293	Ref.	33	0.70 (0.32 - 1.51)	23	0.40 (0.19 - 0.84)	17	0.44 (0.18 - 1.10)	33	0.44 (0.23 - 0.84)
iGAS[d]	1268	Ref.	110	0.59 (0.29 - 1.20)	127	0.53 (0.28 - 1.00)	358	1.91 (0.88 - 4.11)	1715	5.24 (3.01 - 9.12)

[a]Ref.: Reference.

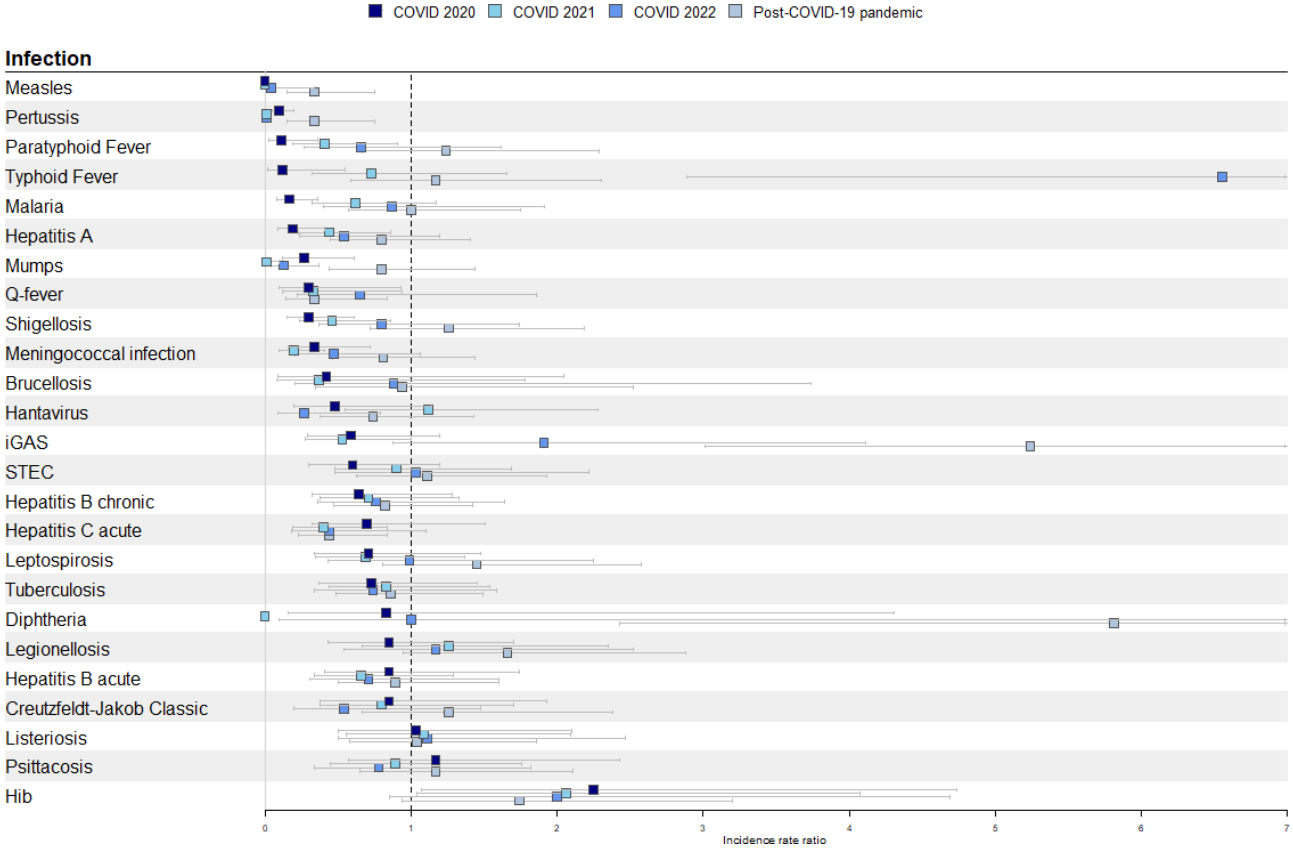
[b]Hib: *Haemophilus influenzae* type b.

[c]STEC: Shiga toxin-producing *E. coli*.

[d]iGAS: invasive group A streptococcal.

XSL•FO
RenderX

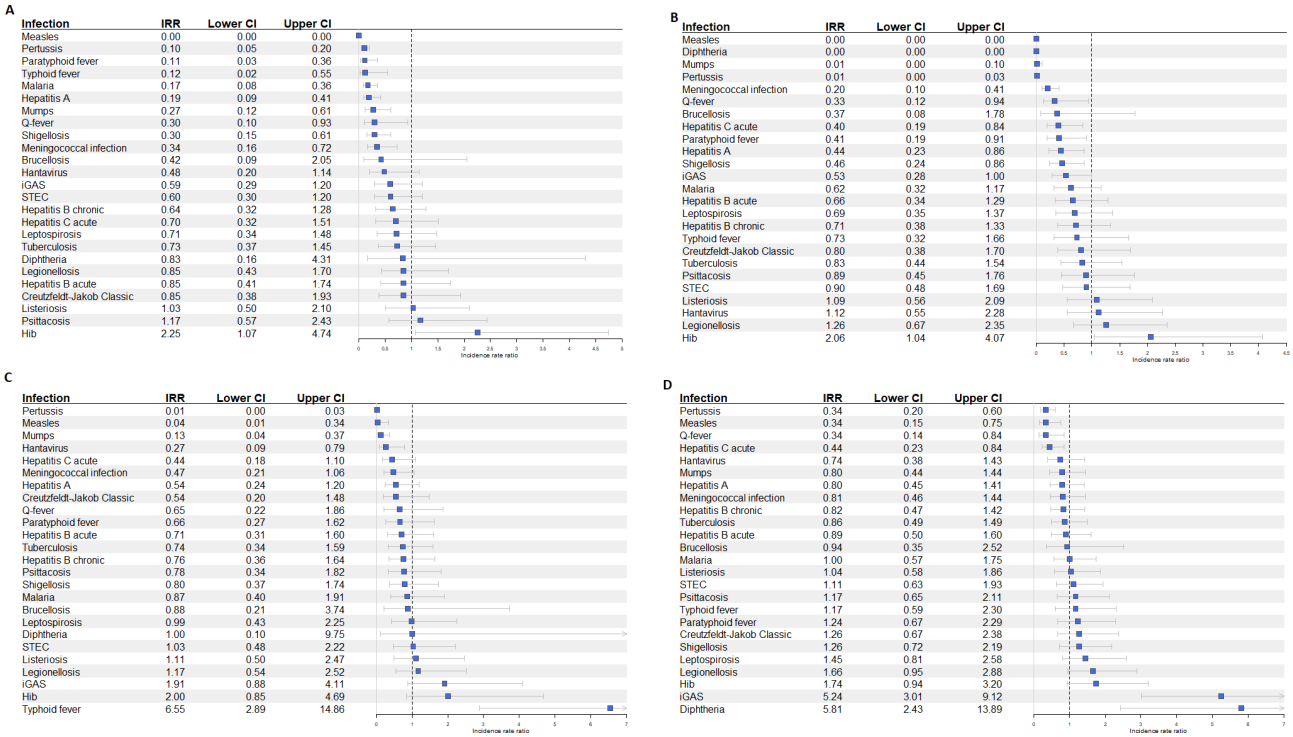
Figure 1. Incidence rate ratios for 25 notifiable infectious diseases plotted by the COVID-19 pandemic periods (2020, 2021, 2022 and post-COVID-19 pandemic), with the pre-COVID-19 pandemic period (January 2015-February 2020) as reference, the Netherlands. Hib: *Haemophilus influenzae* type b; iGAS: invasive group A streptococcal; STEC: Shiga toxin-producing *E. coli*.



Malaria (IRR 0.17, 95% CI 0.08 - 0.36) and typhoid fever (IRR 0.12, 95% CI 0.02 - 0.55) were significantly less notified during COVID 2020 and were back on pre-COVID-19 pandemic levels in 2021 (Figure 2). Typhoid fever was in the COVID 2022 period significantly more reported compared with the pre-COVID-19 pandemic period, which was related to a large outbreak (IRR 6.55, 95% CI 2.89 - 14.86) [22]. Infectious disease notifications affected in 2020 and 2021 were hepatitis A, meningococcal infection, paratyphoid fever, Q-fever, and shigellosis. *Haemophilus influenzae* type b (Hib) was also affected in 2020 and 2021; however, contradictory to other infections, the number of notifications increased significantly in those periods (IRR 2.25, 95% CI 1.07 - 4.74 and IRR 2.06, 95% CI 1.04 - 4.07, respectively). This significant increase was not reported in the COVID 2022 period (Figure 2).

XSL•FO
RenderX

Figure 2. IRRs for 25 notifiable infectious diseases for each COVID-19 pandemic period compared with pre-COVID-19 pandemic period (January 2015-February 2020): (A) COVID 2020 period; (B) COVID 2021 period; (C) COVID 2022 period; and (D) Post-COVID-19 pandemic period, the Netherlands. Hib: *Haemophilus influenzae* type b; iGAS: invasive group A streptococcal; IRR: incidence rate ratio; STEC: Shiga toxin-producing *E. coli*.



Mumps, measles, and pertussis were affected during the entire duration of the COVID-19 pandemic, with significantly less notifications reported (Table 1). Furthermore, the notifications for measles and pertussis continued to be low in the post-COVID-19 pandemic period as well (IRR 0.34, 95% CI 0.15 - 0.75 and IRR 0.34, 95% CI 0.20 - 0.60, respectively). Our hypotheses on whether we expected an effect of the COVID-19 pandemic on the notifications did not match with the found effects for Q-fever and Hib.

Our hypotheses on whether we expected an effect of the COVID-19 pandemic on the notifications did not match with the found effects for Q-fever and Hib.

Not Affected by the COVID-19 Pandemic

Based on our definition of the impact of the COVID-19 pandemic, we found no effect on the notifications of the other 14 infectious diseases: legionellosis, brucellosis, diphtheria, hepatitis B acute, hepatitis B chronic, leptospirosis, listeriosis, psittacosis, Shiga toxin-producing *E. coli*, tuberculosis, Creutzfeldt-Jakob disease, hantavirus, hepatitis C acute, and invasive Group A streptococcal disease. However, we expected an effect in ten of them (Multimedia Appendix 1). Our hypotheses for hantavirus, listeriosis, psittacosis, and Creutzfeldt-Jakob disease were correct.

Correcting for the Impact of COVID-19 Pandemic

As the notifications of malaria, typhoid fever, hepatitis A, meningococcal infection, paratyphoid fever, Q-fever, shigellosis, mumps, measles, and pertussis were significantly less during the COVID-19 pandemic, we calculated their adjusted alarm thresholds for each correction method. The alarm thresholds for outbreak detection do not require adjustment for Hib, as Hib notifications increased during the COVID-19 pandemic, and correcting for outbreaks is standard practice in the original method.

The results of the different correction methods are presented in (Figures 3 and 4, Multimedia Appendices 3-10). We found that there is a difference in the number of notifications between the unadjusted and corrected alarm thresholds for exceeding the alarm threshold; however, no clear difference was observed between the different corrected alarm thresholds. The difference in the number of exceedances of the alarm threshold per method compared with the original unadjusted method was only seen for malaria and mumps (Multimedia Appendix 11). We observed that the corrected alarm thresholds are influenced by the duration and magnitude of the effect of the COVID-19 pandemic, as well as the number of notifications. Over the observed study period, we highlight 2 noteworthy results.

XSL•FO
RenderX

Figure 3. Number of monthly notifications of pertussis with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, the Netherlands, January 2015-December 2023.

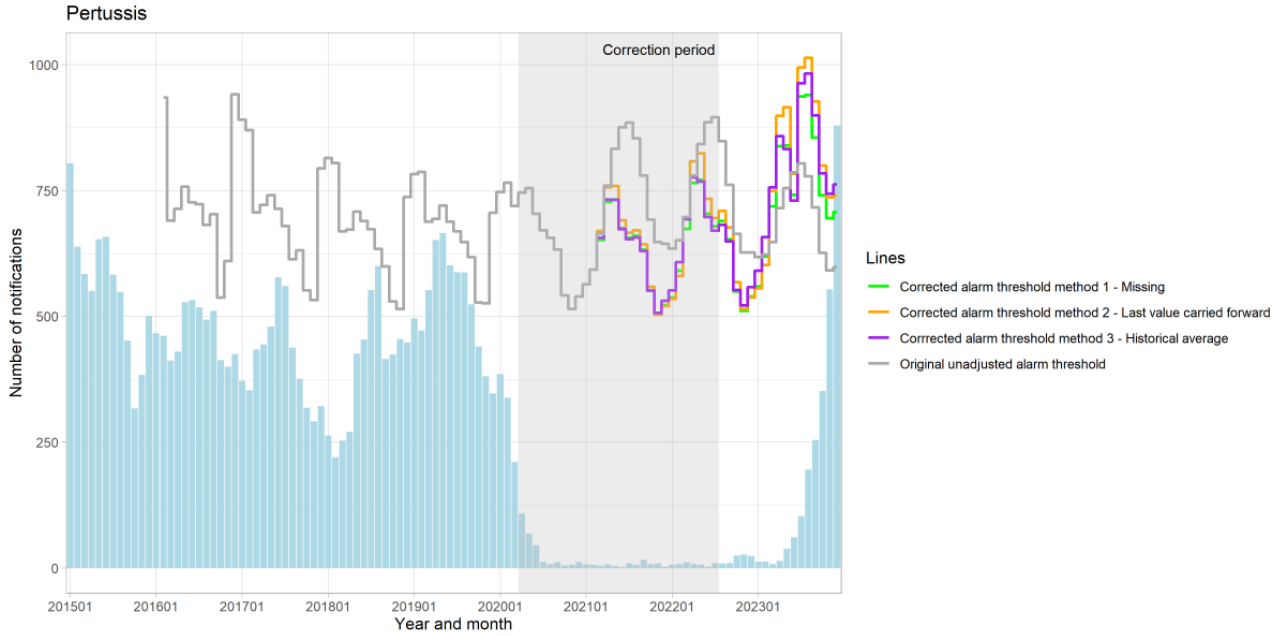
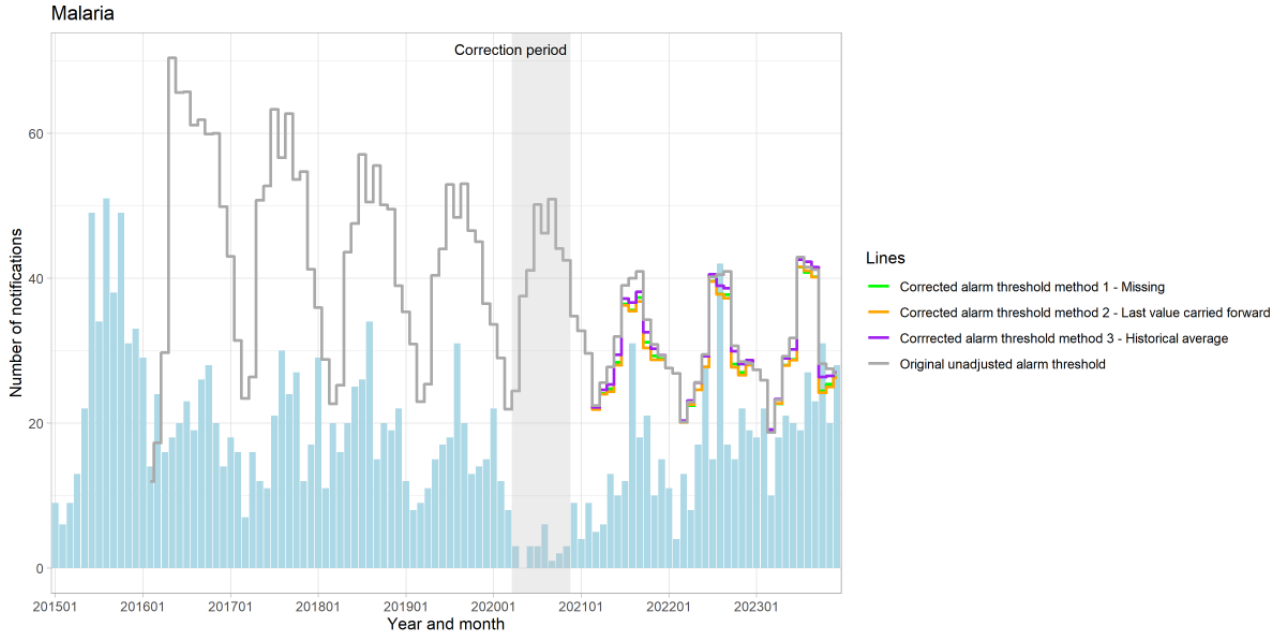


Figure 4. Number of monthly notifications of malaria with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, the Netherlands, January 2015-December 2023.



First, we found that the corrected alarm thresholds for hepatitis A, measles, mumps, paratyphoid fever, and Q-fever were generally higher than their original unadjusted alarm thresholds. The alarm thresholds for typhoid fever showed minimal variation across methods. In contrast, the corrected alarm thresholds were at some points in time lower than the unadjusted alarm thresholds for pertussis, meningococcal infection, and shigellosis. For instance, during the COVID-19 pandemic, the corrected alarm thresholds for pertussis were lower, with a difference of 200 notifications compared with the unadjusted alarm threshold (Figure 3 and Multimedia Appendix 12). This indicates that the correction already impacts the alarm thresholds during the COVID-19 pandemic.

Second, the length of the correction period also varies depending on the duration of the effects of the COVID-19 pandemic on specific disease notifications, which in turn influences the corrected alarm thresholds. Malaria, for example, was only corrected for the period March-December 2020 during which notifications significantly decreased (Figure 4 and Multimedia Appendix 13). As the correction period is longer, the immediate effect of the correction seems more visible compared with a shorter correction period, such as seen with pertussis (Figure 3 and Multimedia Appendix 12). Furthermore, the disparity between the unadjusted and corrected thresholds seems to gradually diminish earlier over time, compared with the infectious diseases with a longer correction period.

XSL•FO
RenderX

Discussion

Principal Findings

This study shows that the COVID-19 pandemic had a significant impact on the notifications of other infectious diseases, and therefore on the surveillance and especially a noteworthy effect on the alarm thresholds for outbreak detection. To take timely preventive actions and implement control measures to contain outbreaks, an accurate alarm threshold for outbreak detection is essential. Malaria, typhoid fever, hepatitis A, meningococcal infection, paratyphoid fever, Q-fever, shigellosis, measles, mumps, and pertussis had significantly lower notifications during the COVID-19 pandemic, but the duration of the effect differed among the infections. Only measles, mumps, and pertussis were affected during the whole duration of the COVID-19 pandemic, and for measles and pertussis, this effect continued in the post-COVID-19 pandemic period. Interestingly, the notifications of Hib increased significantly during the COVID-19 pandemic. In addition, the newly calculated alarm thresholds for infectious diseases with decreased notifications showed a noticeable difference compared with the unadjusted alarm thresholds. However, the variation among the 3 different corrected alarm thresholds was not substantial.

Comparison With Previous Work

Decrease in notifications of malaria, typhoid fever, hepatitis A, meningococcal infection, paratyphoid fever, shigellosis, measles, mumps, and pertussis was expected based on the hypotheses. This was already observed in other countries for malaria, hepatitis A, meningococcal infection, shigellosis, measles, mumps, and pertussis [7-13,15]. In our study, the notifications of Q-fever unexpectedly decreased, while this concerns a zoonotic disease, and the presence of *Coxiella burnetii* in the reservoir in the Netherlands was unlikely to be affected by the COVID-19 pandemic and the NPIs. Furthermore, the number of notifications of the other 14 investigated infectious diseases showed no impact from the COVID-19 pandemic. However, other studies reported a decrease in notification numbers of brucellosis, hepatitis B (acute and chronic), hepatitis C, legionellosis, leptospirosis, listeriosis, and tuberculosis, possibly due to differences in surveillance systems and case definitions [7,8,10,11,13].

Surprisingly, we found a significant increase in the Hib notifications during COVID 2020 and 2021. We expected that Hib would decrease, as it has a similar transmission route as SARS-CoV-2, and therefore, the NPIs could have impacted Hib transmission as well. Besides, other studies described a decrease in Hib notifications during the COVID-19 pandemic [11,23,24]. According to the study of Steens et al [25], the observed increase in incidence in the Netherlands could be explained by methodological, behavioral, biological, or vaccine-related factors; however, the conclusive reason still remains unknown.

In exploring the post-COVID-19 period, we found that the number of notifications for measles and pertussis remained significantly lower compared with the pre-COVID-19 period. This prolonged effect on the notifications was also observed in Poland [26]. This is most likely due to the prolonged unintended effects of NPIs, leading to reduced circulation of these infectious diseases [27]. It highlights the importance of a robust surveillance system, as a resurgence of cases is likely to occur in the future. In the Netherlands, we have experienced a pertussis outbreak in early 2024, exceeding pre-COVID-19 levels (data not shown).

Our study shows that correcting for the impact of the COVID-19 pandemic on the decreased notifications of other notifiable infectious diseases has an effect on the alarm threshold for outbreak detection. The levels of the original unadjusted alarm thresholds differed from the other alarm threshold methods, which were corrected for the decrease in notifications. Other studies that observed a significant decrease in their notifications briefly mention the impact for surveillance; however, none of them propose a concrete solution [11,28]. To our knowledge, this is the first study that investigates whether correction for the decrease in notifications has an effect on outbreak detection and investigates multiple correction methods. Based on our findings, the initial difference among the investigated alarm threshold methods was not evident. This could be explained by the delayed effect of the correction, as it calculates with historical values and could appear visibly later on in the period of 2025 - 2027. Nevertheless, the effect was already strongly visible for the infectious diseases with high notification numbers and a longer correction period, such as pertussis. This underlines the need for pattern-specific corrections. However, implementing a different outbreak detection method for every infectious disease would be too complex in a PHS setting and would require more data points than are available, especially for infectious diseases with low notification numbers. We therefore propose a general methodology that incorporates pattern-specific corrections based on the timing of impact. For example, the alarm threshold for malaria would only be corrected for the COVID 2020 period, while the alarm threshold for pertussis will be corrected for the whole COVID-19 pandemic period. The next step will be the validation by communicable disease control professionals to assess the applicability of the corrected alarm thresholds for outbreak detection.

The significant and prolonged effect of the COVID-19 pandemic on notifiable infectious diseases underlines the necessity of correction for the decrease in notifications. As it is important for communicable disease professionals to act validly in a timely manner. Not correcting for the decrease of notifications during the COVID-19 pandemic could lead to false positive exceedances of the alarm threshold, thereby consuming time and other resources of communicable disease control professionals. The findings also suggest that certain outbreaks could potentially go undetected, and consequently, no action would be taken, with all its associated consequences. Furthermore, as for pandemic preparedness, data-driven decision and policy-making is crucial. Consequently, a robust surveillance system is a necessity. This study shows that correction for a decrease in notifications affects outbreak detection for other infectious diseases, even during an ongoing pandemic. In conclusion, it is recommended to account for periods of underreporting or reduced notifications to ensure reliable outbreak detection. Future work should explore which correction method is the most appropriate for outbreak detection in the PHS setting.

XSL•FO
RenderX

Limitations

One of the limitations of this study is the inability to determine the factors driving the decrease in notifications based on this analysis. We hypothesized that the NPIs had an effect on the notifications, as the interference of the SARS-CoV-2 variants, possible shift in demographic determinants, and difference in health care use and access to diagnostics during the COVID-19 pandemic. We did not have any data on vaccination status, health care usage, and number of diagnostic requests; therefore, we could not identify if the notifications were affected by one of these factors specifically. However, we did not see any difference in the distribution of demographic determinants among the notifications. Other studies suggest that the decrease is likely due to a combination of factors [7,10,11]. In addition, exploring the causes of this decrease was beyond the scope of this study.

Another limitation is that we had a simplified definition for the post-COVID-19 pandemic period. We defined September 2022 as the end of the COVID-19 pandemic in the Netherlands, as the COVID-19 Stringency Index was at its lowest. One could argue that the post-COVID-19 pandemic period started later on, which could lead to underestimating the effect in the post-COVID-19 pandemic period. An alternative approach is to evaluate whether the monthly value is significantly below the historical average, similar to the methodology used for outbreak correction. As it will correct for short-term effects, this could create an alarm threshold that is more sensitive, which can be preferred in infectious disease control. Furthermore, assessing a monthly or even weekly correction approach would be valuable during a future pandemic, as it allows continuous adjustments for the ongoing impact rather than defining it retrospectively. This option was not assessed in our study, as we believe it would require more data and a prolonged study period. However, we recommend including this in future validation studies to further develop the “pandemic-proof” methodology.

Finally, we did not perform an interrupted time-series analysis for determining the trend of the notifications, as seen in other studies. This analysis required more data points than were available for some infectious diseases. Nevertheless, the negative binomial regression proved to be a robust alternative, as it accounts for zero notifications. To assess sensitivity, we compared the IRR between the pre-COVID-19 pandemic period and the entire COVID-19 pandemic period (2020 - 2022) and found no notable differences to the IRRs of the individual COVID-19 pandemic periods. Our main focus was to identify which infectious diseases were affected and the duration of these effects, which was adequately addressed using the negative binomial regression.

Conclusions

The COVID-19 pandemic and the related pandemic measures are likely to have had a considerable effect on the notifications of other infectious diseases, affecting both surveillance and outbreak detection. The length of the effect also differed among the infectious diseases, necessitating pattern-specific corrections. Our study shows that correcting for the decrease in notifications has an effect on the alarm thresholds for outbreak detection. Further research is required to determine which correction method is most suitable and applicable for effective outbreak detection, to ensure communicable disease control professionals act in a timely and data-driven manner.

Acknowledgments

The authors acknowledge Dr M Nabuurs-Franssen for providing expert insights on the impact of the COVID-19 pandemic on healthcare access and diagnostic use. The authors also extend their gratitude to Dr R Akkermans and Dr R Pijnacker for their valuable statistical advice. In addition, they would like to thank the National Institute of Public Health and Environment for providing the data. The authors acknowledge the research infrastructure provided by the Dutch Collaborative Academic Practice for Public Health Infectious Diseases. This project was funded by the Ministry of Health, Welfare and Sport of the Dutch Government within the program "Strengthening infectious disease control and pandemic preparedness of the regional Public Health Services." The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The artificial intelligence language model ChatGPT was used exclusively to enhance the manuscript’s readability and to correct linguistic and spelling errors.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overview of hypothesized effects of factors on the number of notifications of listed infections during the COVID-19 pandemic, the Netherlands.
[XLSX File, 13 KB - publichealth_v11i1e73953_app1.xlsx]

Multimedia Appendix 2

XSL•FO
RenderX

Descriptives of 25 notifiable infectious diseases (74,990 notifications) by the COVID-19 periods (pre-COVID, 2020, 2021, 2022, and post-COVID), 2015-2023, the Netherlands.
[XLSX File, 27 KB - publichealth_v11i1e73953_app2.xlsx]

Multimedia Appendix 3
Number of monthly notifications of Hepatitis A with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, The Netherlands, January 2015-December 2023.
[PNG File, 49 KB - publichealth_v11i1e73953_app3.png]

Multimedia Appendix 4
Number of monthly notifications of measles with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, The Netherlands, January 2015-December 2023.
[PNG File, 45 KB - publichealth_v11i1e73953_app4.png]

Multimedia Appendix 5
Number of monthly notifications of mumps with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, The Netherlands, January 2015-December 2023.
[PNG File, 46 KB - publichealth_v11i1e73953_app5.png]

Multimedia Appendix 6
Number of monthly notifications of paratyphoid fever with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, The Netherlands, January 2015-December 2023.
[PNG File, 46 KB - publichealth_v11i1e73953_app6.png]

Multimedia Appendix 7
Number of monthly notifications of Q-fever with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, The Netherlands, January 2015-December 2023.
[PNG File, 45 KB - publichealth_v11i1e73953_app7.png]

Multimedia Appendix 8
Number of monthly notifications of typhoid fever with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, The Netherlands, January 2015-December 2023.
[PNG File, 43 KB - publichealth_v11i1e73953_app8.png]

Multimedia Appendix 9
Number of monthly notifications of meningococcal infection with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, The Netherlands, January 2015-December 2023.
[PNG File, 49 KB - publichealth_v11i1e73953_app9.png]

Multimedia Appendix 10
Number of monthly notifications of shigellosis with the unadjusted threshold and 3 corrected threshold methods for outbreak detection, The Netherlands, January 2015-December 2023.
[PNG File, 46 KB - publichealth_v11i1e73953_app10.png]

Multimedia Appendix 11
Number of exceedances of the outbreak detection threshold for the unadjusted alarm threshold and the 3 correction methods, March 2020-December 2023.
[PDF File, 53 KB - publichealth_v11i1e73953_app11.pdf]

Multimedia Appendix 12
Unadjusted alarm threshold and the alarm thresholds of the 3 correction methods for pertussis, March 2021-December 2023, the Netherlands
[XLSX File, 13 KB - publichealth_v11i1e73953_app12.xlsx]

Multimedia Appendix 13

XSL•FO
RenderX

Unadjusted alarm threshold and the alarm thresholds of the 3 correction methods for malaria, March 2021-December 2023, the Netherlands.
[XLSX File, 12 KB - publichealth_v11i1e73953_app13.xlsx]

References

1. Bijkerk P, Fanoy EB, Kardamanidis K, et al. To notify or not to notify: decision aid for policy makers on whether to make an infectious disease mandatorily notifiable. Euro Surveill 2015;20(34):30003. [doi: 10.2807/1560-7917.ES.2015.20.34.30003] [Medline: 26530302]
2. Farrington CP, Andrews NJ, Beale AD, Catchpole MA. A statistical algorithm for the early detection of outbreaks of infectious disease. J R Stat Soc Ser A Stat Soc 1996;159(3):547. [doi: 10.2307/2983331]
3. Stroup DF, Wharton M, Kafadar K, Dean AG. Evaluation of a method for detecting aberrations in public health surveillance data. Am J Epidemiol 1993 Feb 1;137(3):373-380. [doi: 10.1093/oxfordjournals.aje.a116684] [Medline: 8452145]
4. Unkel S, Farrington CP, Garthwaite PH, Robertson C, Andrews N. Statistical methods for the prospective detection of infectious disease outbreaks: a review. J R Stat Soc Ser A Stat Soc 2012 Jan 1;175(1):49-82. [doi: 10.1111/j.1467-985X.2011.00714.x]
5. Morbey RA, Elliot AJ, Charlett A, Verlander NQ, Andrews N, Smith GE. The application of a novel “rising activity, multi-level mixed effects, indicator emphasis” (RAMMIE) method for syndromic surveillance in England. Bioinformatics 2015 Nov 15;31(22):3660-3665. [doi: 10.1093/bioinformatics/btv418] [Medline: 26198105]
6. Noufaily A, Enki DG, Farrington P, Garthwaite P, Andrews N, Charlett A. An improved algorithm for outbreak detection in multiple surveillance systems. Stat Med 2013 Mar 30;32(7):1206-1222. [doi: 10.1002/sim.5595] [Medline: 22941770]
7. Facciolà A, Laganà A, Genovese G, et al. Impact of the COVID-19 pandemic on the infectious disease epidemiology. J Prev Med Hyg 2023 Sep;64(3):E274-E282. [doi: 10.15167/2421-4248/jpmh2023.64.3.2904] [Medline: 38125993]
8. Geng MJ, Zhang HY, Yu LJ, et al. Changes in notifiable infectious disease incidence in China during the COVID-19 pandemic. Nat Commun 2021 Nov 26;12(1):6923. [doi: 10.1038/s41467-021-27292-7] [Medline: 34836947]
9. Nash K, Lai J, Sandhu K, et al. Impact of national COVID-19 restrictions on incidence of notifiable communicable diseases in England: an interrupted time series analysis. BMC Public Health 2022 Dec 12;22(1):2318. [doi: 10.1186/s12889-022-14796-0] [Medline: 36510205]
10. Ullrich A, Schranz M, Rexroth U, et al. Impact of the COVID-19 pandemic and associated non-pharmaceutical interventions on other notifiable infectious diseases in Germany: An analysis of national surveillance data during week 1-2016 - week 32-2020. Lancet Reg Health Eur 2021 Jul;6(100103):100103. [doi: 10.1016/j.lanepe.2021.100103] [Medline: 34557831]
11. van de Berg S, Charles T, Dörre A, Katz K, Böhm S. Epidemiology of common infectious diseases before and during the COVID-19 pandemic in Bavaria, Germany, 2016 to 2021: an analysis of routine surveillance data. Euro Surveill 2023 Oct;28(41):2300030. [doi: 10.2807/1560-7917.ES.2023.28.41.2300030] [Medline: 37824248]
12. Zhou Q, Hu J, Hu W, Li H, Lin GZ. Interrupted time series analysis using the ARIMA model of the impact of COVID-19 on the incidence rate of notifiable communicable diseases in China. BMC Infect Dis 2023 Jun 5;23(1):375. [doi: 10.1186/s12879-023-08229-5] [Medline: 37316780]
13. Sohail A, Cheng AC, McGuinness SL, Leder K. The epidemiology of notifiable diseases in Australia and the impact of the COVID-19 pandemic, 2012-2022. BMC Glob Public Health 2024 Jan 2;2(1):1. [doi: 10.1186/s44263-023-00029-y] [Medline: 39681889]
14. Hausmann J, Dörre A, Katz K, van de Berg S. Lifting COVID-19-associated non-pharmaceutical interventions: potential impact on notifications of infectious diseases transmitted from person to person in 2022 in Bavaria, Germany. Front Public Health 2024;12(1437485):1437485. [doi: 10.3389/fpubh.2024.1437485] [Medline: 39148653]
15. Pijnacker R, Mughini-Gras L, Verhoef L, van den Beld M, Franz E, Friesema I. Impact of non-pharmaceutical interventions during the COVID-19 pandemic on pathogens transmitted via food in the Netherlands. Epidemiol Infect 2024 Oct 22;152:e130. [doi: 10.1017/S0950268824000815] [Medline: 39434376]
16. Zhang W, Wu Y, Wen B, et al. Non-pharmaceutical interventions for COVID-19 reduced the incidence of infectious diseases: a controlled interrupted time-series study. Infect Dis Poverty 2023 Mar 9;12(1):15. [doi: 10.1186/s40249-023-01066-3] [Medline: 36895021]
17. Piret J, Boivin G. Viral interference between respiratory viruses. Emerg Infect Dis 2022 Feb;28(2):273-281. [doi: 10.3201/eid2802.211727] [Medline: 35075991]
18. Yuan H, Yeung A, Yang W. Interactions among common non-SARS-CoV-2 respiratory viruses and influence of the COVID-19 pandemic on their circulation in New York City. Influenza Other Respir Viruses 2022 Jul;16(4):653-661. [doi: 10.1111/irv.12976] [Medline: 35278037]
19. Matera L, Manti S, Petrarca L, et al. An overview on viral interference during SARS-CoV-2 pandemic. Front Pediatr 2023;11(1308105):1308105. [doi: 10.3389/fped.2023.1308105] [Medline: 38178911]
20. Frey A, Tilstra AM, Verhagen MD. Inequalities in healthcare use during the COVID-19 pandemic. Nat Commun 2024 Feb 29;15(1):1894. [doi: 10.1038/s41467-024-45720-2] [Medline: 38424038]
21. Mathieu E, Ritchie H, Rodés-Guirao L, et al. COVID-19 pandemic. Our World in Data. 2020. URL: https://ourworldindata.org/coronavirus [accessed 2025-08-7]

XSL•FO
RenderX

22. Ooms D, de Vries A, Koedijk FD, et al. Large outbreak of typhoid fever on a river cruise ship used as accommodation for asylum seekers, the Netherlands, 2022. Euro Surveill 2024 Feb;29(5):2300211. [doi: 10.2807/1560-7917.ES.2024.29.5.2300211] [Medline: 38304948]
23. Kadambari S, Goldacre R, Morris E, Goldacre MJ, Pollard AJ. Indirect effects of the covid-19 pandemic on childhood infection in England: population based observational study. BMJ 2022 Jan 12;376:e067519. [doi: 10.1136/bmj-2021-067519] [Medline: 35022215]
24. Shaw D, Abad R, Amin-Chowdhury Z, et al. Trends in invasive bacterial diseases during the first 2 years of the COVID-19 pandemic: analyses of prospective surveillance data from 30 countries and territories in the IRIS Consortium. Lancet Digit Health 2023 Sep;5(9):e582-e593. [doi: 10.1016/S2589-7500(23)00108-5] [Medline: 37516557]
25. Steens A, Stanoeva KR, Knol MJ, Mariman R, de Melker HE, van Sorge NM. Increase in invasive disease caused by Haemophilus influenzae b, the Netherlands, 2020 to 2021. Euro Surveill 2021 Oct;26(42):34676819. [doi: 10.2807/1560-7917.ES.2021.26.42.2100956] [Medline: 34676819]
26. Rzymski P, Zarębska-Michaluk D, Parczewski M, et al. The burden of infectious diseases throughout and after the COVID-19 pandemic (2020-2023) and Russo-Ukrainian war migration. J Med Virol 2024 May;96(5):38712743. [doi: 10.1002/jmv.29651] [Medline: 38712743]
27. Oh KB, Doherty TM, Vetter V, Bonanni P. Lifting non-pharmaceutical interventions following the COVID-19 pandemic - the quiet before the storm? Expert Rev Vaccines 2022 Nov;21(11):1541-1553. [doi: 10.1080/14760584.2022.2117693] [Medline: 36039786]
28. Dougherty B, Smith CR, Forrest RO, et al. Considerations for enteric disease surveillance in a post-COVID-19 pandemic world: a Canadian perspective. Foodborne Pathog Dis 2023 Apr;20(4):121-122. [doi: 10.1089/fpd.2023.0033.editorial] [Medline: 37062810]

Abbreviations

Hib: *Haemophilus influenzae* type b
IRR: incidence rate ratio
NPI: nonpharmaceutical intervention
PHS: public health service
RIVM: National Institute for Public Health and the Environment

Edited by A Mavragani, T Sanchez; submitted 14.03.25; peer-reviewed by A Adeoye, C El-Hayek, C br; revised version received 30.06.25; accepted 01.07.25; published 26.08.25.

Please cite as:
van Deursen B, Raven S, van den Bosch W, van Jaarsveld CHM, Timen A
A "Pandemic-Proof" Methodology for Outbreak Detection Adapted From COVID-19's Impact on Notifications of Infectious Diseases in the Netherlands: Surveillance Study
JMIR Public Health Surveill 2025;11:e73953
URL: https://publichealth.jmir.org/2025/1/e73953
doi:10.2196/73953

© Babette van Deursen, Stijn Raven, Wolfer van den Bosch, Cornelia H M van Jaarsveld, Aura Timen. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Validation of Prediction Models for Perceived and Unmet Mental Health Needs in the Canadian General Population: Model-Based Synthetic Estimation Study

Jianli Wang[1,2], PhD; Heather Orpana[3,4], PhD; André Carrington[5], PhD; George Kephart[1,6], PhD; Helen-Maria Vasiliadis[7], PhD; Benjamin Leikin[8], BSc

[1]Department of Community Health and Epidemiology, Faculty of Medicine, Dalhousie University, 5790 University Avenue, Halifax, NS, Canada

[2]Department of Psychiatry, Faculty of Medicine, Dalhousie University, Halifax, NS, Canada

[3]Centre for Surveillance and Applied Research, Public Health Agency of Canada, Ottawa, ON, Canada

[4]School of Psychology, University of Ottawa, Ottawa, ON, Canada

[5]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

[6]School of Health Administration, Faculty of Health, Dalhousie University, Halifax, NS, Canada

[7]Department of Community Health Sciences, Faculty of Medicine and Health, Université de Sherbrooke, Sherbrooke, QC, Canada

[8]Community Health and Wellness Branch, Ottawa Public Health, Ottawa, ON, Canada

Corresponding Author:
Jianli Wang, PhD
Department of Community Health and Epidemiology, Faculty of Medicine, Dalhousie University, 5790 University Avenue, Halifax, NS, Canada

Abstract

Background: Research has shown that perceptions of a mental health need are closely associated with service demands and are an important dimension in needs assessment. Perceived and unmet mental health needs are important factors in the decision-making process regarding mental health services planning and resources allocation. However, few prediction tools are available to be used by policy and decision makers to forecast perceived and unmet mental health needs at the population level.

Objective: We aim to develop prediction models to forecast perceived and unmet mental health needs at the provincial and health regional levels in Canada.

Methods: Data from 2018, 2019, and 2020 Canadian Community Health Survey and Canadian Urban Environment were used (n=65,000 each year). Perceived and unmet mental health needs were measured by the Perceived Needs for Care Questionnaire. Using the 2018 dataset, we developed the prediction models through the application of regression synthetic estimation for the Atlantic, Central, and Western regions. The models were validated in the 2019 and 2020 datasets at the provincial level and in 10 randomly selected health regions by comparing the observed and predicted proportions of the outcomes.

Results: In 2018, a total of 17.82% of the participants reported perceived mental health need and 3.81% reported unmet mental health need. The proportions were similar in 2019 (18.04% and 3.91%) and in 2020 (18.1% and 3.92%). Sex, age, self-reported mental health, physician diagnosed mood and anxiety disorders, self-reported life stress and life satisfaction were the predictors in the 3 regional models. The individual based models had good discriminative power with C statistics over 0.83 and good calibration. Applying the synthetic models in 2019 and 2020 data, the models had the best performance in Ontario, Quebec, and British Columbia; the absolute differences between observed and predicted proportions were less than 1%. The absolute differences between the predicted and observed proportion of perceived mental health needs in Newfoundland and Labrador (−4.16% in 2020) and Prince Edward Island (4.58% in 2019) were larger than those in other provinces. When applying the models in the 10 selected health regions, the models calibrated well in the health regions in Ontario and in Quebec; the absolute differences in perceived mental health needs ranged from 0.23% to 2.34%.

Conclusions: Predicting perceived and unmet mental health at the population level is feasible. There are common factors that contribute to perceived and unmet mental health needs across regions, at different magnitudes, due to different population characteristics. Therefore, predicting perceived and unmet mental health needs should be region specific. The performance of the models at the provincial and health regional levels may be affected by population size.

(JMIR Public Health Surveill 2025;11:e66056) doi:10.2196/66056

KEYWORDS

population risk prediction; development; validation; perceived mental health need; unmet mental health need

XSL•FO
RenderX

Introduction

Mental disorders are prevalent and have a considerable impact on the people who experience them, as well as families, communities and society as a whole. To enhance the provision of timely and appropriate mental health services for people who are in need, effective health services planning should be adaptable and responsive to the changing needs and emerging opportunities. Research has shown that many people living with a mental disorder do not receive the services due to the barriers of availability, accessibility, and acceptability [1]. On the other hand, many people with symptoms of mental disorders that do not meet clinical criteria may also benefit from intervention and these people may actively seek treatment to prevent the symptoms from escalating to the level of a disorder [2]. These phenomena reflect different types of mental health care needs and the extent to which the mental health needs are being met.

Mental health need is difficult to define. Bradshaw [3] proposed six need types: (1) normative (presence of a mental disorder diagnosis), (2) felt (subjective perception of a mental health problem), (3) expressed (demand for mental health service), (4) comparative (population inequities in mental health), (5) medical (treatable disease), and (6) social (restoring quality of life). There is evidence to suggest that perceptions of a mental health need (felt need) are closely associated with service demands [4] and are an important dimension in needs assessment [5]. Data-driven evidence about perceived and unmet mental health needs at provincial, state and regional, or county levels can be very helpful in building the business case of demand [6]. However, population health data about perceived and unmet mental health needs have not been well adopted for this purpose [7].

Unmet health need is “the absence of sufficient or appropriate care and services.” [8] These needs are commonly assessed as being for information, medication, counseling or therapy, or another type of help. A need is considered to be fully met when a person receives help meeting all their expectations. In other cases, the help may only partly fulfill the need, or no support is provided at all and the need is unmet [9]. Based on 2012 Canadian Community Health Survey (CCHS)–mental health, 17.3% of Canadians aged 15 years and older reported having mental health needs in the past year, and about one-third of those with a mental health care need reported that it was unmet [9,10]. Addressing unmet needs is important, as people with untreated anxiety or depression are at higher risk of experiencing poor outcomes including persistence of symptoms, delayed recovery, poor personal and occupational functioning, and recurrence of these problems [11]. Unmet health care need has been identified to be an important indicator of health care access within various health care systems [12,13], and is a form of inefficiency that has economic implications, according to Mental Health Commission of Canada [14]. Both perceived and unmet mental health needs constitute the central part of demand which is a critical element in the process of health resources allocation.

Decision makers and health service planners have detailed information about existing mental health resources including available infrastructure, services and workforce at the provincial and regional levels. To allocate and use these resources efficiently, the magnitude of mental health needs and unmet needs in the community is critical information in the decision-making process. The ability to closely monitor and forecast the trends of perceived and unmet needs in populations can greatly facilitate this decision-making process. The objective of this study was to develop prediction models for estimating and forecasting perceived and unmet mental health needs at the provincial and health regional levels in Canada.

Methods

Study Design

For the objective of this study, data from the 2018, 2019, and 2020 CCHS and Canadian Urban Environment (CANUE) were used. CCHS is an annual national population health survey conducted by Statistics Canada to gather data on health status, health care use, and health determinants at the health region levels of geography. The CCHS covers the household population from 12 years of age and older living in the 10 provinces and 3 territories. The CCHS sample is selected using different frames according to the age group. For the adult population (18 y and older), the sample of households is selected from an area frame. For the youth population (12 to 17 y old) a list frame is used to select persons. The area frame used by the Labour Force Survey is used as a sampling frame for the adult population. The Labour Force Survey uses a 2-stage sample design. In the first stage, a sample of primary sampling units, corresponding to geographical regions called clusters, is selected. In each selected primary sampling unit, a sample of dwellings is drawn at the second stage. To sample persons for the youth population between the ages of 12 and 17 years, the CCHS uses a list frame created from the Canadian Child Benefit files. The sample for the youth population is selected from a list of individuals. For adult participants, 1 person is selected per household using varying probabilities taking into account the age and the household composition. Each year the CCHS collects information on approximately 65,000 Canadians (60,000 aged 18 y and older and 5000 aged 12 to 17 y) [15]. The CCHS uses 2 separate computer-assisted interviewing applications to collect data, 1 for telephone interviews (computer-assisted telephone interviewing) and 1 for personal interviews (computer-assisted personal interviewing). Approximately 25% of these completed cases were conducted in person using computer-assisted personal interviewing, and the other 75% were conducted over the phone using computer-assisted telephone interviewing. Starting in 2015, the CCHS undertook significant redesign in sampling and the content [15]. Data about the outcome variables of the proposed study (perceived and unmet mental heath needs) were not collected in 2015, 2016, and 2017 CCHS, therefore, we will analyze the 2018, 2019, and 2020 CCHS data.

CANUE is a Canadian Institutes of Health Research–funded collaboration that focuses on developing robust methods for producing measures of urban form that capture a wide range of characteristics for every postal code in Canada [16]. It has produced a unique repository of standardized metrics of urban, suburban, and rural characteristics. CANUE shares a range of data that can be directly linked with population health surveys

XSL•FO
RenderX

through postal codes, including neighborhood-level material and social deprivation [17], marginalization [18], and walkability. CANUE is updated every 5 years. The most recent CANUE update was in 2016. The CANUE data can be obtained from CANUE consortium at no costs for research, and is linkable to the CCHS by postal codes.

Perceived and Unmet Mental Health Needs

In the CCHS, perceived and unmet mental health needs were measured by the Meadows et al [19] Perceived Needs for Care Questionnaire (PNCQ). The PNCQ was designed and field tested for the Australian National Survey for Mental Health and Wellbeing which was commissioned by the Federal Government of Australia [19]. The PNCQ assessed four types of help for problems with emotions, mental health, or the use of alcohol or drugs: (1) information about problems, treatments, or services; (2) medication; (3) counseling or therapy; and (4) other mental health services. Respondents were asked which types of help they had received in the past 12 months. For each type received, they were asked if they felt they had received enough. For each type of help not received, they were asked if they felt it was needed. Based on the PNCQ, the following binary outcome variables can be derived: [9,20]

- *Perceived mental health needs:* having reported "yes" to a perceived need of any type of help.
- *Unmet mental health needs:* For any type of help for which a respondent had a perceived need, having reported not receiving any or enough of that type of help. Interrater reliability of the PNCQ (κ value) was 0.62 [19]. Validity testing using homo-method agreements and hetero-method agreement demonstrated that the PNCQ has good discriminative validity [19]. The PNCQ has been used to assess perceived and unmet mental health needs in a number of population health surveys conducted in Australia [19], Canada [21], the United States [22], and the Netherlands [20].

Candidate Predictors

Overview

Guided by the Andersen Health Behavior model [23], we selected potential predictors from the core component of the CCHS which are consistently administered over time in all provinces and territories. This ensured that health regions can readily ascertain the predictor profile of their regions to make an estimation and prediction when PNCQ is not administered in the CCHS.

From the CCHS core component, we selected the predisposing factors (sex, gender, age, marital status, race or ethnicity, rural or urban residence, first official language spoken, immigration, household food security, self-reported life stress, work stress, life satisfaction, and sense of belonging), enabling factors (educational level, household income, employment status, and insurance coverage), and need factors (general health, comorbid physical conditions, self-reported mental health, mood, and anxiety, and primary care service use). Mental health problems such as depression, anxiety, and psychological distress are not part of the core component; data about these variables are not available in every CCHS and for every province and territory. Therefore, these variables were not selected. A number of studies have shown that health behaviors (smoking [24], problematic substance use [25,26], and physical inactivity [27,28]) are associated with depression, anxiety, and psychological distress. We selected these health behaviors from the core component as candidate predictors. From the CANUE, neighborhood-level social and material deprivation [17], marginalization [18], and walkability will be selected as enabling factors.

Model Development

The 2018 CCHS data were used to develop the prediction models. Given the vast geographic area of Canada, we developed the models by regions: Atlantic region (Newfoundland and Labrador, Prince Edward Island, Nova Scotia, and New Brunswick), Central region (Ontario and Quebec), Western region (Manitoba, Saskatchewan, Alberta, and British Columbia). For each region, we developed prediction models for perceived and unmet mental health needs through the application of regression synthetic estimation developed by the UK's Office of National Statistics [29], which is a type of small area estimation methodology. The same methodology has been used to estimate state and area-level prevalence of severe mental illness in the United States [30]. The regression synthetic estimation method involves several steps: (1) The development of a prediction model of the outcome at the individual level. (2) The coefficients derived from the foregoing model are used with a parallel set of predictors on the area level, using data obtained from CCHS, and coded using the same categories as used to estimate the individual-level model and to compute national and area-level estimates. (3) Internal and external validation.

To develop the predictive models for the outcomes at the health region level, multilevel, random intercept binary logistic regression models were used to analyze the perceived and unmet need for mental health services (level 1) nested within health regions (level 2). A backward selection method was used to identify the model with the best calibration and discrimination. The decisions of model selection were initially based on the changes in the values of Akaike information criterion and Bayesian information criterion [31].

Because the goal of models was to assist decision makers and service planners to estimate perceived and unmet needs at the population level (not to be used by clinicians to identify high risk individuals), we focused more on the calibration of the model, instead of discrimination, when it came to model performance. Discrimination is the ability of a prediction model to separate those who experienced the outcome events from those who did not, by predicting higher versus lower probabilities, respectively. We quantified this by calculating the area under the receiver operating characteristics curve, which is equal to the C statistic. Calibration measures how closely predicted outcomes agree with actual outcomes. For this, we used the calibration slope by which a value of 1 indicates perfect model calibration with data; less than 1 is overconfident and greater than 1 is underconfident [32]. Additionally, we used calibration plots to visually compare the mean predicted risk of the outcomes versus the observed risk (the cumulative fraction

XSL•FO
RenderX

of events) by decile risk groups so that the overall calibration and the areas with over or under prediction can be identified.

The second step was synthetic estimation of the proportions at the population level with perceived and unmet mental health needs, which consisted of 2 stages. First, population proportions for predictor variables that were used in the initial modeling were estimated. For instance, if 5 household income categories are used in the initial modeling (<CAD $30,000, CAD $30,000 - $49,000, CAD $49,000 - $60,000, CAD $60,000 - $80,000, and CAD $80,000+; <US $21,175.80, US $21,175.80-$34,587.14, US $34,587.14-$42,351.60, US $42,351.60-56,468.80, and US $56,468.80+, respectively), the population proportion of participants in each of the same 5 income ranges were estimated from the CCHS. The regression coefficients were then applied to the corresponding proportions in the dataset, and to calculate the logit estimates for each province, which were then converted into probabilities, giving the predicted proportions of perceived and unmet mental health needs in the province. These procedures were repeated with data aggregated at the provincial and health regional levels. All analyses were conducted using bootstrap weights provided by Statistics Canada to account for the unequal sampling probabilities and design effects.

Validation

As internal validation, we used CCHS-2018 data to compare the predicted proportions (step 2 above) and the observed proportions. To ensure the validity of the predictive models, it is also important that the developed models are externally validated in related but independent populations. As external validation, we applied the developed synthetic models in CCHS-2019 and CCHS-2020, to examine the performance of the models, to investigate how the pandemic may affect the performance of the models by comparing the model predicted (synthetic) proportions and the observed proportions. Additionally, we randomly selected 10 subprovincial health regions from over 100 across the country. We applied the regional models directly to the health regions and compared the observed and predicted proportions of perceived and unmet mental health needs. All analyses were conducted using the bootstrap weights provided by Statistics Canada to account for the sampling and design effects of the CCHS.

Ethical Considerations

This is a secondary data analysis using the deidentified data collected by Statistics Canada. The data analysis was carried out at one of the Research Data Centers across the country, located on the campus of Dalhousie University. Researchers who conducted data analysis had to go through security clearances and become “deemed employees” of Statistics Canada. The results were vetted by data analysts of the Research Data Center to ensure privacy and confidentiality. CCHS participants gave informed consent to Statistics Canada for use of their data in accordance with the Statistics Act and no compensation was provided to the CCHS participants. For these reasons, ethics review was waived by the health research ethics committee of Dalhousie University.

Results

Overview

The 2018 CCHS data were used to develop the prediction models. The demographic and socioeconomic characteristics of the participants in the 2018 CCHS are presented in Table 1. In the 2018 CCHS, 17.82% reported perceived mental health need and 3.81% reported unmet mental health need. In 2019, the proportion of perceived and unmet mental health needs was 18.04% and 3.91%, respectively; in 2020, the proportion of perceived and unmet mental health needs was 18.1% and 3.92%, respectively.

XSL•FO
RenderX

Table . The demographic and socioeconomic characteristics of the participants in the 2018 cross-sectional Canadian Community Health Survey.

Variables	Values
Men, weighted %	49.4
Women, weighted %	50.6
Age (years), mean (SD)	45.7 (0.14)
Married, common law, or partner, weighted %	58.2
Single, weighted %	30.1
Divorced, separated, or widowed, weighted %	11.7
<High school, weighted %	17.4
High school, weighted %	22.6
≥College or university, weighted %	60
Employed, weighted %	67.2
Unemployed, weighted %	22.4
<15 or >75 years old[a], weighted %	10.4
Immigrants, weighted %	27.1
Nonimmigrants, weighted %	72.9
White, weighted %	75.5
Non-White, weighted %	24.5
Rural residence, weighted %	17.2
Urban residence, weighted %	82.8
Have insurance for medication, weighted %	19.7
No, weighted %	80.3

[a]Participants aged <15 or 75+ years were not eligible for the employment question.

Predicting Perceived Mental Health Need

The individual based models with the best performance for predicting perceived mental health need are in supplemental tables (Tables S1 [Atlantic region], S2 [Central region], and S3 [Western region] in Multimedia Appendix 1). Sex, age, self-reported mental health, physician diagnosed mood and anxiety disorders, and self-reported life stress and life satisfaction were the predictors in the 3 regional models. Perceived work stress and household food insecurity were in the models for the Atlantic and Central regions, but not for the Western region. However, immigrants, smoking, or problematic drinking and material deprivation were predictors in the models for the Atlantic and the Western region, not for the Central region. The individual based models had good discriminative power with C statistics over 0.83 (Tables S4-S6 in Multimedia Appendix 1) and good calibration by comparing observed and predicted proportions.

Converting the individual based models into synthetic estimation models, the predicted proportion of perceived mental health need in each province was estimated. The absolute difference between the observed and predicted proportions in 2018 was less than 1% for all provinces (Tables 2-4), except Saskatchewan (3.37%). This is not surprising as the models were developed using the 2018 CCHS data. To validate the models, we applied the synthetic models directly in 2019 and 2020 data. The extent to which the predicted or synthetic estimates agree with observed proportions varied by regions (Tables 2-4). The models had the best performance in Ontario and British Columbia (absolute difference <1% in both 2019 and 2020), followed by Quebec (absolute difference=1.54% in 2019 and 0.54% in 2020). The models for Nova Scotia, New Brunswick, Saskatchewan, and Alberta had 1 absolute difference being less than 1% and the other greater than 2%. The absolute differences between observed and predicted proportions in Newfoundland and Labrador (−4.16% for perceived needs in 2020) and Prince Edward Island (4.58% for perceived needs in 2019) were larger than those in other provinces. When applying the models in the 10 selected health regions, the data showed that the model performed well in health regions in Ontario (first 2 digits with 35) and in Quebec (first 2 digits with 24), supported by the relatively small absolute differences between observed and predicted proportions (Table 5). Health region 1304 had the largest absolute differences in 2019 and 2020.

XSL•FO
RenderX

Table . The observed and predicted proportions of perceived and unmet mental health needs in the Atlantic provinces in the 2018, 2019, and 2020 Canadian Community Health Survey.

		Observed proportion (%)	Predicted proportion (%)	Absolute difference[a] (%)
Newfoundland and Labrador				
Perceived need				
	2018	14.9	14.57	−0.33
	2019	16.24	14.04	−2.2
	2020	19.35	15.19	−4.16
Unmet need				
	2018	2.43	2.07	−0.36
	2019	2.93	1.94	−0.99
	2020	2.43	1.96	−0.47
Prince Edward Island				
Perceived need				
	2018	14.37	14.23	−0.14
	2019	20.92	16.34	−4.58
	2020	16.89	14.93	−1.96
Unmet need				
	2018	3.63	3.96	0.33
	2019	2.74	4.15	1.41
	2020	2.6	3.85	1.25
Nova Scotia				
Perceived need				
	2018	19.81	19.15	−0.66
	2019	23.58	20.44	−3.14
	2020	20.56	21.33	0.77
Unmet need				
	2018	3.43	3.52	0.09
	2019	5.78	3.55	−2.23
	2020	4.23	3.3	−0.93
New Brunswick				
Perceived need				
	2018	17.66	18.44	0.78
	2019	17.22	19.97	2.75
	2020	20	20.83	0.83
Unmet need				
	2018	3.05	2.85	−0.2
	2019	3.85	3.52	−0.33
	2020	3.71	3.42	−0.29

[a]Absolute difference=predicted proportion–observed proportion.

XSL•FO
RenderX

Table . The observed and predicted proportions of perceived and unmet mental health needs in Ontario and Quebec in the 2018, 2019, and 2020 Canadian Community Health Survey.

			Observed proportion (%)	Predicted proportion (%)	Absolute difference[a] (%)
Ontario					
	Perceived need				
		2018	16.93	16.63	−0.30
		2019	17.35	16.94	−0.41
		2020	17.97	17.44	−0.53
	Unmet need				
		2018	3.96	3.98	0.02
		2019	4.08	4.09	0.01
		2020	4	4.29	0.29
Quebec					
	Perceived need				
		2018	17.08	17.65	−0.43
		2019	16.02	17.56	1.54
		2020	16.72	17.26	0.54
	Unmet need				
		2018	3.25	3.26	0.01
		2019	3.01	3.21	0.2
		2020	3.36	3.09	−0.27

[a]Absolute difference=predicted proportion−observed proportion.

XSL•FO
RenderX

Table . The observed and predicted proportions of perceived and unmet mental health needs in the Western provinces in the 2018, 2019, and 2020 Canadian Community Health Survey.

			Observed proportion (%)	Predicted proportion (%)	Absolute difference[a] (%)
Manitoba					
	Perceived need				
		2018	18.8	19.52	0.72
		2019	18.86	19.56	0.7
		2020	16.87	20.13	3.26
	Unmet need				
		2018	3.59	4.18	0.59
		2019	5.11	4.12	−0.99
		2020	4.25	4.43	0.18
Saskatchewan					
	Perceived need				
		2018	16.57	19.94	3.37
		2019	18.61	21.14	2.53
		2020	15.98	21.31	5.33
	Unmet need				
		2018	3.1	4.03	0.93
		2019	2.84	4.23	1.39
		2020	4.07	4.16	0.09
Alberta					
	Perceived need				
		2018	21.73	20.77	−0.96
		2019	20.02	20.91	0.89
		2020	18.92	21.49	2.57
	Unmet need				
		2018	4.52	4.41	−0.11
		2019	3.66	4.43	0.77
		2020	3.72	4.54	0.82
British Columbia					
	Perceived need				
		2018	18.39	18.38	−0.01
		2019	19.14	18.86	−0.28
		2020	20.05	20.01	−0.04
	Unmet need				
		2018	4.28	4.16	−0.12
		2019	4.9	N/A[b]	N/A
		2020	4.87	4.51	−0.36

[a]Absolute difference=predicted proportion–observed proportion.

[b]N/A: household food insecurity data were not collected in British Columbia in 2019.

XSL•FO
RenderX

Table . The observed and predicted proportions of perceived mental health needs by selected health regions in 2018, 2019, and 2020 Canadian Community Health Survey.

Health regions and years		Observed proportion (%)	Predicted proportion (%)	Absolute difference[a] (%)
1301				
	2018	17.42	19.08	1.66
	2019	19.81	22.47	2.66
	2020	22.89	22.48	−0.41
1304				
	2018	11.86	16.12	4.26
	2019	6.59	15.23	8.64
	2020	9.45	19.61	10.16
2404				
	2018	14.22	16.04	1.82
	2019	14.01	15.83	1.82
	2020	10.78	14.72	3.94
3544				
	2018	20.12	19.59	−0.53
	2019	20.4	18.7	−1.7
	2020	20.13	20.36	0.23
3570				
	2018	12.57	14.91	2.34
	2019	16.29	15.36	−0.93
	2020	17.97	15.87	−2.1
3536				
	2018	14.99	15.9	0.91
	2019	17.4	16.16	−1.24
	2020	17.21	16.38	−0.83
5913				
	2018	15.17	18.28	3.11
	2019	N/A[b]	N/A	N/A
	2020	24.29	18.8	−5.49
5921				
	2018	13.75	18.66	4.91
	2019	N/A	N/A	N/A
	2020	21.24	20.54	−0.7
5922				
	2018	15.8	18.44	2.64
	2019	N/A	N/A	N/A
	2020	19.77	19.46	−0.31
5943				
	2018	17.88	17.71	−0.17
	2019	N/A	N/A	N/A
	2020	23.37	17.84	−5.53

[a]Absolute difference=predicted proportion–observed proportion.

XSL•FO
RenderX

[b]N/A: estimates were not available due to lack of data on household food insecurity.

Predicting Unmet Mental Health Need

The individual based models with the best performance for predicting unmet mental health need are in Tables S7 (Atlantic region), S8 (Central region), and S9 (Western region) in Multimedia Appendix 1. As seen from the tables, sex, age, marital status, self-reported mental health, self-reported life stress, and low sense of belonging to the community were the common predictors in the models for the 3 regions. Household income was a predictor only in the model for Central region; professional diagnosed mood disorders, problematic drinking, and smoking and social deprivation were the factors specific for the model for the Atlantic region. The individual based models had good discriminative power with C statistics over 0.77 (Tables S10-S12 in Multimedia Appendix 1) and good calibration by comparing observed and predicted proportions.

The individual based models were converted into synthetic estimation models, and the predicted proportion of perceived mental health need in each province was estimated (Tables 2-4). The data demonstrated that the models for unmet mental health need performed well in most of the provinces with the absolute difference between observed and predicted proportions being less than 1%, except for Prince Edward Island, Nova Scotia, and Saskatchewan. When applying the models in the 10 selected health regions, the data showed that the models calibrated well with data in all selected health regions with absolute difference between observed and predicted proportions being less than 1% (Table S13 in Multimedia Appendix 1).

Discussion

Principal Results

The 3 waves of CCHS showed that the proportion of the Canadian population that perceived the need for mental health care was between 17.82% and 18.1%, and about 4% reported unmet mental health need. This study demonstrated the feasibility of integrating individual and community level data to build informative synthetic models for perceived and unmet mental health care needs at the population level. The models performed well in predicting the outcomes at both the provincial and health regional levels, particularly in populous provinces such as Ontario, Quebec, and British Columbia. The absolute differences between the observed and predicted proportions of perceived and unmet mental health needs in Ontario, Quebec, and British Columbia were less than 1%. For the rest of the provinces, the results showed that the models at provincial and health regional levels predicting unmet mental health needs had better calibration than the models predicting perceived mental health need.

Limitations

This study had several limitations. First, although the predictors in the models were associated with the outcomes, causal inferences cannot be made. The goal of the prediction models is to identify a key set of factors that in combination are best predictive of the outcomes. The models are not to test hypothesis or make inferences about etiology. Second, the relationships between the selected factors and the outcomes are complex. The logistic regression model is a linear function. Although we found no evidence of interactions among the selected predictors, nonlinear relationships between some predictors and the outcomes are still possible. Future studies may test if models using machine learning techniques have better performance. Third, the selection of candidate predictors was limited by what CCHS collected. It is possible that other factors are associated with the outcomes of this study, but were not measured in CCHS, therefore could not be examined in this study. Finally, the models appeared to perform well at regional and provincial levels. Relatively large absolute differences between predicted and observed proportions of perceived and unmet needs were observed in some health regions (eg, health region 1304). Caution should be exercised for using them at the health region level.

Comparison With Prior Work

The CCHS used a nationwide representative sample of the Canadian household population. However, we found that a model developed for the whole population performed well only in the most populous provinces such as Ontario and Quebec, but not in the rest of the provinces. Developing models for specific regions appeared to be a better approach. The regional models have some common predictors, including sex, age, self-reported mental health, physician diagnosed mood and anxiety disorders, self-reported life stress and life satisfaction for perceived mental health need, and sex, age, marital status, self-reported mental health, self-reported life stress, and low sense of belonging to the community for unmet mental health need. These factors have been found to be significantly associated with mental health problems [33] in the literature. On the other hand, the coefficients associated with these factors varied by regional models and these models also contained predictors that are regionally specific. This finding indicates that there are common factors in the mechanisms underlying perceived and unmet mental health needs across geographic regions and populations. The extent to which these factors influence perceived and unmet mental health needs may differ due to the population characteristics of these provinces and regions. The differences in population characteristics and the distribution of the outcomes across the regions may also contribute to the finding that the models contained region specific predictors.

We have not found studies on predicting perceived and unmet mental health needs at the population level. Therefore, a direct comparison with previous studies is not possible. There are several studies predicting county or state level suicide and severe mental illness. For example, Kandula et al [34] modelled county-level suicide risk in the United States using county-level predictors derived from 8 different databases of different sources (government programs, health surveys, or private organizations). For some predictors such as the prevalence of major depressive episode, only state-level estimates were available and these estimates were extrapolated to the counties [34]. Hudson [30] explored the utility of regression synthetic estimation model that incorporated individual data from the National Comorbidity

XSL•FO
RenderX

Survey, census and hospital administrative data to predict state-level prevalence of severe mental illness. The advantages of these population risk prediction models are the use of community level predictors from existing sources or published research and the ability of adapting the models in local context. Notably, our study used the regression synthetic estimation modeling approach. We used CCHS data. The use of a single data source may improve the efficiency of data analysis, data access, and eventually of the decision-making process.

One critical element of building risk prediction models is assessing model performance and model validation. This is to ensure that the developed model is accurate and has good performance in different populations or at different time periods. In this study, we developed the models using the 2018 CCHS data, and validated the models using the data from 2019 and 2020 CCHS. Furthermore, we validated the models in a random sample of health regions. These models were designed to estimate population proportions and to identify regions with high levels of mental health need and unmet mental health need; these models were not to be used by clinicians to identify high-risk individuals. Therefore, the focus of model performance assessment can be different. Kandula et al [34] used symmetric proportional error (observed deaths–predicted deaths)/(observed deaths+predicted deaths) to quantify model calibration. Hudson [30] calculated the absolute difference between the predicted and observed prevalence of severe mental illness. In this study, we calculated the absolute difference between the observed and model predicted proportions of perceived and unmet mental health needs. The absolute difference indicates the extent to which the model prediction deviates from the observed value. However, there is no consensus about what an optimal threshold of absolute difference should be. There may be other indicators that are useful for assessing the performance of population level models. Consultations with policy and decision makers (ie, end users of the models) would be helpful to understand what indicators are informative about model performance, and the level of the model error that is acceptable.

The results of this study are expected to have implications for population mental health planning. Few would deny that resources allocation should be partly driven by needs, and needs assessments typically require the knowledge of potential changes in prevalence estimates and in local population profiles, for example, their demographics, diagnoses, and mental health services use. The prediction models developed by this study will allow decision makers and mental health services planners to forecast the proportions of perceived and unmet mental health needs in the years to come at the provincial (state) and health regional (county) levels based on the potential changes in local population profiles. Such profile changes may be estimated using health administrative data and national population census data. Additionally, region-specific estimates can help categorize health regions—for example, regions with relatively stable mental health needs especially those that remained in the highest or lowest groups, or regions in which the largest year-to-year changes are observed—and hence help identify areas with a greater need of preventive resources, or conversely identify areas where interventions seem to be effective.

Conclusion

In conclusion, accurate forecast of perceived and unmet mental health needs in the population can allow policy and decision makers and mental health services planners to categorize regions or communities that are at high need and to monitor changes so that they may mobilize resources and interventions to the right populations and the right places at the right time. Regularly collected population health data such as those from the CCHS are readily accessible to policy and decision makers and mental health services planners. The models are particularly useful for service planners at the health regional level because population health surveys usually do not contain sufficient numbers of participants at the health regional level. Future studies are needed to identify methods to improve prediction for regions with small numbers of residents.

Acknowledgments

This study was supported by a grant from the Canadian Institutes of Health Research.

Data Availability

The Canadian Community Health Survey (CCHS) data were collected and are maintained by Statistics Canada. The deidentified raw data can only be accessed at a Research Data Center after becoming a “deemed employee” of Statistics Canada. However, Statistics Canada has prepared carefully modified Public Use Microdata Files (PUMFs) of the CCHS which ensure that individuals cannot be identified, and which are in the public domain.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File, 35 KB - publichealth_v11i1e66056_app1.docx]

References

XSL•FO
RenderX

1. Wang J. Perceived barriers to mental health service use among individuals with mental disorders in the Canadian general population. Med Care 2006 Feb;44(2):192-195. [doi: 10.1097/01.mlr.0000196954.67658.95] [Medline: 16434920]
2. Kessler RC, Zhao S, Blazer DG, Swartz M. Prevalence, correlates, and course of minor depression and major depression in the National Comorbidity Survey. J Affect Disord 1997 Aug;45(1-2):19-30. [doi: 10.1016/s0165-0327(97)00056-6] [Medline: 9268772]
3. Bradshaw J. The conceptualization and measurement of need: a social policy perspective. In: Popay J, Williams G, editors. Researching the People's Health: Routledge; 2005:57-70 URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9780203996782-12/conceptualization-measurement-need-social-policy-perspective [accessed 2025-02-05]
4. Wichstrøm L, Belsky J, Jozefiak T, Sourander A, Berg-Nielsen TS. Predicting service use for mental health problems among young children. Pediatrics 2014 Jun;133(6):1054-1060. [doi: 10.1542/peds.2013-3184] [Medline: 24819574]
5. Stevens A, Gabbay J. Needs assessment needs assessment. Health Trends 1991;23(1):20-23. [Medline: 10113881]
6. Roos NP, Fransoo R, Carrière KC, et al. Needs-based planning: the case of Manitoba. CMAJ 1997 Nov 1;157(9):1215-1221. [Medline: 9361640]
7. Vigo D, Jones W, Dove N, et al. Estimating the prevalence of mental and substance use disorders: a systematic approach to triangulating available data to inform health systems planning. Can J Psychiatry 2022 Feb;67(2):107-116. [doi: 10.1177/07067437211006872] [Medline: 33827278]
8. Carr W, Wolfe S. Unmet needs as sociomedical indicators. Int J Health Serv 1976;6(3):417-430. [doi: 10.2190/MCG0-UH8D-0AG8-VFNU] [Medline: 955751]
9. Sunderland A, Findlay LC. Perceived need for mental health care in Canada: results from the 2012 Canadian Community Health Survey-Mental Health. Health Rep 2013 Sep;24(9):3-9. [Medline: 24258361]
10. Urbanoski K, Inglis D, Veldhuizen S. Service use and unmet needs for substance use and mental disorders in Canada. Can J Psychiatry 2017 Aug;62(8):551-559. [doi: 10.1177/0706743717714467] [Medline: 28616935]
11. Wang J. A longitudinal population-based study of treated and untreated major depression. Med Care 2004 Jun;42(6):543-550. [doi: 10.1097/01.mlr.0000128001.73998.5c] [Medline: 15167322]
12. Lasser KE, Himmelstein DU, Woolhandler S. Access to care, health status, and health disparities in the United States and Canada: results of a cross-national population-based survey. Am J Public Health 2006 Jul;96(7):1300-1307. [doi: 10.2105/AJPH.2004.059402] [Medline: 16735628]
13. Chen J, Hou F. Unmet needs for health care. Health Rep 2002;13(2):23-34. [Medline: 12743954]
14. Strengthening the case for investing in Canada's mental health system: economic considerations. Mental Health Commission of Canada. 2017. URL: https://www.mentalhealthcommission.ca/wp-content/uploads/drupal/2017-03/case_for_investment_eng.pdf [accessed 2025-02-07]
15. Canadian community health survey (CCHS) annual component user guide: 2018 and 2017-2018 microdata file. Statistics Canada. 2019. URL: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3226
16. CANUE Data Portal. URL: https://www.canuedata.ca/ [accessed 2025-02-05]
17. Pampalon R, Hamel D, Gamache P, Raymond G. A deprivation index for health planning in Canada. Chronic Dis Can 2009;29(4):178-191. [Medline: 19804682]
18. Matheson FI, Dunn JR, Smith KLW, Moineddin R, Glazier RH. Development of the Canadian Marginalization Index: a new tool for the study of inequality. Can J Public Health 2012 Apr 30;103(8 Suppl 2):S12-S16. [doi: 10.1007/BF03403823] [Medline: 23618065]
19. Meadows G, Harvey C, Fossey E, Burgess P. Assessing perceived need for mental health care in a community survey: development of the Perceived Need for Care Questionnaire (PNCQ). Soc Psychiatry Psychiatr Epidemiol 2000 Sep;35(9):427-435. [doi: 10.1007/s001270050260] [Medline: 11089671]
20. Prins M, Meadows G, Bobevski I, et al. Perceived need for mental health care and barriers to care in the Netherlands and Australia. Soc Psychiatry Psychiatr Epidemiol 2011 Oct;46(10):1033-1044. [doi: 10.1007/s00127-010-0266-3] [Medline: 20686887]
21. Thomas S, Wannell B. Combining cycles of the Canadian Community Health Survey. Health Rep 2009 Mar;20(1):53-58. [Medline: 19388369]
22. Nam B, Wilcox HC, Hilimire M, DeVylder JE. Perceived need for care and mental health service utilization among college students with suicidal ideation. J Am Coll Health 2018;66(8):713-719. [doi: 10.1080/07448481.2018.1434779] [Medline: 29384469]
23. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [Medline: 7738325]
24. Fluharty M, Taylor AE, Grabski M, Munafò MR. The association of cigarette smoking with depression and anxiety: a systematic review. Nicotine Tob Res 2017 Jan;19(1):3-13. [doi: 10.1093/ntr/ntw140] [Medline: 27199385]
25. Escelsior A, Belvederi Murri M, Corsini GP, et al. Cannabinoid use and self-injurious behaviours: a systematic review and meta-analysis. J Affect Disord 2021 Jan 1;278:85-98. [doi: 10.1016/j.jad.2020.09.020] [Medline: 32956965]
26. Wang J, Patten SB. Prospective study of frequent heavy alcohol use and the risk of major depression in the Canadian general population. Depress Anxiety 2002;15(1):42-45. [doi: 10.1002/da.1084] [Medline: 11816052]

XSL•FO
RenderX

27. Bailey AP, Hetrick SE, Rosenbaum S, Purcell R, Parker AG. Treating depression with physical activity in adolescents and young adults: a systematic review and meta-analysis of randomised controlled trials. Psychol Med 2018 May;48(7):1068-1083. [doi: 10.1017/S0033291717002653] [Medline: 28994355]
28. Mammen G, Faulkner G. Physical activity and the prevention of depression: a systematic review of prospective studies. Am J Prev Med 2013 Nov;45(5):649-657. [doi: 10.1016/j.amepre.2013.08.001] [Medline: 24139780]
29. Heady P, Clarke P, Brown G, et al. Model-based small area estimation series no. 2. small area estimation project report. : Office for National Statistics; 2003.
30. Hudson CG. Validation of a model for estimating state and local prevalence of serious mental illness. Int J Methods Psychiatr Res 2009 Dec;18(4):251-264. [doi: 10.1002/mpr.294] [Medline: 20014156]
31. Vrieze SI. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods 2012 Jun;17(2):228-243. [doi: 10.1037/a0027127] [Medline: 22309957]
32. Hilden J, Habbema JD, Bjerregaard B. The measurement of performance in probabilistic diagnosis. II. Trustworthiness of the exact values of the diagnostic probabilities. Methods Inf Med 1978 Oct;17(4):227-237. [Medline: 366335]
33. Arango C, Dragioti E, Solmi M, et al. Risk and protective factors for mental disorders beyond genetics: an evidence-based atlas. World Psychiatry 2021 Oct;20(3):417-436. [doi: 10.1002/wps.20894] [Medline: 34505386]
34. Kandula S, Martinez-Alés G, Rutherford C, et al. County-level estimates of suicide mortality in the USA: a modelling study. Lancet Public Health 2023 Mar;8(3):e184-e193. [doi: 10.1016/S2468-2667(22)00290-0] [Medline: 36702142]

Abbreviations

CANUE: Canadian Urban Environment
CCHS: Canadian Community Health Survey
PNCQ: Perceived Needs for Care Questionnaire

Edited by A Mavragani; submitted 02.09.24; peer-reviewed by LK Chua, W Wang; revised version received 14.01.25; accepted 22.01.25; published 19.02.25.

Please cite as:
Wang J, Orpana H, Carrington A, Kephart G, Vasiliadis HM, Leikin B
Development and Validation of Prediction Models for Perceived and Unmet Mental Health Needs in the Canadian General Population: Model-Based Synthetic Estimation Study
JMIR Public Health Surveill 2025;11:e66056
URL: https://publichealth.jmir.org/2025/1/e66056
doi:10.2196/66056

© JianLi Wang, Heather Orpana, André Carrington, George Kephart, Helen-Maria Vasiliadis, Benjamin Leikin. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 19.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Strategies to Increase Response Rate and Reduce Nonresponse Bias in Population Health Research: Analysis of a Series of Randomized Controlled Experiments during a Large COVID-19 Study

Christina J Atchison[1], PhD; Nicholas Gilby[2], BA; Galini Pantelidou[2], MSc; Sam Clemens[2], BA; Kevin Pickering[2], MSc; Marc Chadeau-Hyam[1], PhD; Deborah Ashby[1], PhD; Wendy S Barclay[3], PhD; Graham S Cooke[3], PhD; Ara Darzi[4], MD; Steven Riley[1], PhD; Christl A Donnelly[1,5], ScD; Helen Ward[1], PhD; Paul Elliott[1], PhD

[1]School of Public Health, Imperial College London, London, United Kingdom

[2]Ipsos, London, United Kingdom

[3]Department of Infectious Disease, Imperial College London, Norfolk Place, London, United Kingdom

[4]Institute of Global Health Innovation, Imperial College London, South Kensington Campus, London, United Kingdom

[5]Department of Statistics, University of Oxford, Oxford, United Kingdom

Corresponding Author:
Paul Elliott, PhD
School of Public Health, Imperial College London, London, United Kingdom

Abstract

Background: High response rates are needed in population-based studies, as nonresponse reduces effective sample size and bias affects accuracy and decreases the generalizability of the study findings.

Objective: We tested different strategies to improve response rate and reduce nonresponse bias in a national population–based COVID-19 surveillance program in England, United Kingdom.

Methods: Over 19 rounds, a random sample of individuals aged 5 years and older from the general population in England were invited by mail to complete a web-based questionnaire and return a swab for SARS-CoV-2 testing. We carried out several nested randomized controlled experiments to measure the impact on response rates of different interventions, including (1) variations in invitation and reminder letters and SMS text messages and (2) the offer of a conditional monetary incentive to return a swab, reporting absolute changes in response and relative response rate (95% CIs).

Results: Monetary incentives increased the response rate (completed swabs returned as a proportion of the number of individuals invited) across all age groups, sex at birth, and area deprivation with the biggest increase among the lowest responders, namely teenagers and young adults and those living in more deprived areas. With no monetary incentive, the response rate was 3.4% in participants aged 18 - 22 years, increasing to 8.1% with a £10 (US $12.5) incentive, 11.9% with £20 (US $25.0), and 18.2% with £30 (US $37.5) (relative response rate 2.4 [95% CI 2.0-2.9], 3.5 [95% CI 3.0-4.2], and 5.4 [95% CI 4.4-6.7], respectively). Nonmonetary strategies had a modest, if any, impact on response rate. The largest effect was observed for sending an additional swab reminder (SMS text message or email). For example, those receiving an additional SMS text message were more likely to return a completed swab compared to those receiving the standard email-SMS approach, 73.3% versus 70.2%: percentage difference 3.1% (95% CI 2.2%-4.0%).

Conclusions: Conditional monetary incentives improved response rates to a web-based survey, which required the return of a swab test, particularly for younger age groups. Used in a selective way, incentives may be an effective strategy for improving sample response and representativeness in population-based studies.

(JMIR Public Health Surveill 2025;11:e60022) doi:10.2196/60022

KEYWORDS

study recruitment; response rate; population-based research; COVID-19; SARS-CoV-2; web-based questionnaires

Introduction

In population-based studies, a high response rate from a representative sample may reduce nonparticipation bias, increase the generalizability, and improve the accuracy of study estimates [1]. However, achieving this goal is challenging, both due to the difficulty in contacting and then engaging eligible participants [2]. For example, UK Biobank, a population-based

XSL•FO
RenderX

cohort study with stored biological samples from half a million participants aged 40 - 69 years in the United Kingdom, achieved an overall response rate of 5.5% [3], which was lower in men, younger people, and those living in more deprived areas [4]. The impact of nonresponse and nonrepresentativeness on the generalizability of disease prevalence and incidence rates in the UK Biobank has been widely debated [5,6].

It is important to address low or falling response rates to reduce the likelihood of systematic biases that may affect study estimates [7]. While weighting is commonly applied to correct for differential participation, it may fail to correct bias if the responders in a particular subgroup of the population are not representative of that subgroup as a whole. Furthermore, weighting to correct for observed biases worsens precision (reducing the effective sample size) [8].

Systematic reviews that have evaluated interventions to increase response rates in surveys have concluded that monetary incentives are more effective than nonmonetary incentives [9-11], although findings were inconsistent concerning web-based surveys in educational research [12]. Some studies have found incentives can increase response among under-represented sociodemographic groups, such as those with low incomes, those with low education, single parents, and minority ethnic groups, potentially reducing nonresponse bias [13], while others show mixed results [9].

Other strategies that have been shown to improve response rates in surveys have included the use of SMS text message reminders to enhance the contact method of letters and emails [14,15], using alternative motivational statements in invitation letters [16], and changing the font color of text [17]. In a United Kingdom–based study investigating the effects of augmenting the contact strategy of letters and emails with SMS text messages for a web questionnaire, the findings indicated that SMS text messages did not help to significantly increase response rates overall, although some subgroups benefited from them, such as younger panel members and those with an irregular response pattern [15].

The Real-time Assessment of Community Transmission-1 (REACT-1) study was one of the largest population surveillance studies in the world. Across 19 rounds between May 1, 2020 and March 31, 2022, it provided timely prevalence estimates of SARS-CoV-2, the virus that causes COVID-19, from random cross-sectional samples of the population in England [18,19].

Response rate varied between 11.7% and 30.5% and, like in many population surveys, varied across demographic groups [19]. For financial reasons, we could not issue more than 845,000 invitation letters by mail, so we could only achieve the minimum desired sample size adopted from round 12 (May 20 to June 7, 2021) of 100,000 by improving response [18]. The observed nonresponse biases meant REACT-1 was under-representing groups with lower vaccination rates and where COVID-19 prevalence was highest; thus, we were likely underestimating the true population prevalence despite our attempts to correct for such biases by use of weighting on known demographic variables [20,21]. Here we present results of experiments nested within the REACT-1 study to test the effectiveness of different strategies to increase response rates and participation of groups with a lower propensity to take part.

Methods

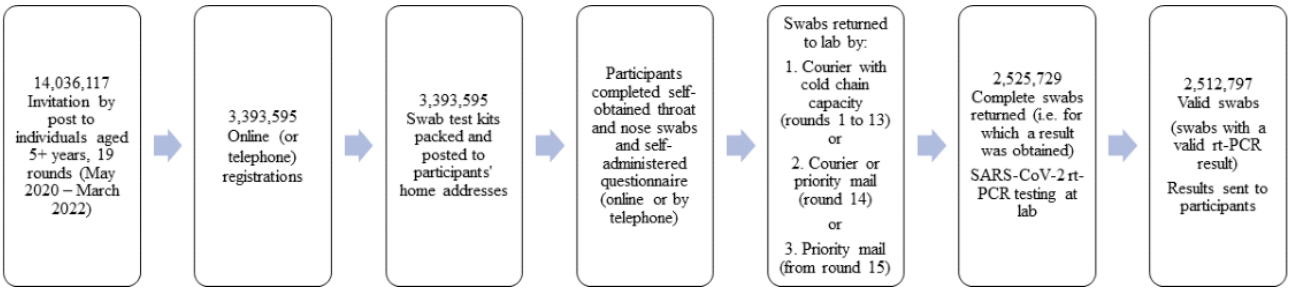
The REACT-1 Study

Methods for the study, including sample size calculations, are described in detail elsewhere [18,19]. In summary, at approximately monthly intervals, between 395,020 and 841,227 people were sent personalized invitations by mail to take part. For children (5-17 years old), the invitation was sent to or via the parents or guardians. Individuals aged 5 years and older were randomly sampled from the National Health Service (NHS) list of patients in England (with near-universal population coverage) across all 316 Lower Tier Local Authorities [18,19]. This list includes the name, address, date of birth, and sex of everyone registered with a general practitioner in England. Invitees who registered (most digitally, some by telephone) for the study received a kit by mail with instructions on how to take a throat and nose swab and send it for SARS-CoV-2 testing using reverse transcriptase polymerase chain reaction (rt-PCR). Swabs were transferred to laboratories for processing, initially being picked up by courier with cold chain capacity (rounds 1-13 and part of 14) or sent by priority mail (part of round 14 and subsequent rounds). Participants were also asked to complete a self-administered web-based or telephone questionnaire [18,19].

Over the 19 study rounds, we sent out 14,036,117 invitations, 3,393,595 registrations were made, and 2,525,729 completed swabs were returned (ie, for which a laboratory result was obtained) (Figure 1). Of these swabs, 2,512,797 (99.5% of completed swabs returned) were considered valid for analysis in REACT-1 (swabs with a valid rt-PCR result) [19]. A swab with a valid rt-PCR result was a swab for which a “cycle threshold” (Ct) value could be obtained. Therefore, not all swabs tested by the lab were considered valid. Overall, 12,932 (0.5% of completed swabs returned) were considered invalid and rejected. Reasons included inadequate sample volume, contamination during sample collection, inappropriate sample storage, or inappropriate sample transportation. All analyses in this paper are based on completed swabs returned (ie, for which a laboratory result was obtained, n=2,525,729), thus including swabs for which a Ct value could not be obtained but excluding swabs returned unused.

XSL•FO
RenderX

Figure 1. REACT-1 study process over 19 rounds of data collection: England, May 1, 2020 to March 31, 2022. Overall, across 19 rounds, we report the number of invitations sent, the number of participants registered, the number of swab test kits sent out, the number of completed swabs returned (ie, for which a laboratory result was obtained) and the number of valid swabs (swabs with a valid rt-PCR result). REACT-1: REal-time Assessment of Community Transmission-1; rt-PCR: reverse transcriptase Polymerase Chain Reaction; SARS-CoV-2: Severe Acute Respiratory Syndrome CoronaVirus 2.



All the experiments carried out to improve response rate were randomized trials, enabling an unbiased assessment of the impact of the changed survey procedure compared to a control group. Due to funding constraints, initial experiments focused on approaches which would not materially affect the survey budget, before turning to an experiment with monetary incentives.

Swab Reminder and Tailored Letter or SMS Experiments

In each round of REACT-1, those registering for a swab test were, where necessary, sent at least one reminder to complete the swab test and return it, to maximize the number of swabs returned. In round 3 (July 24 to August 11, 2020) we conducted an experiment to establish the optimal use of email and SMS text message swab return reminders, with participants randomly allocated to the experimental conditions (Table 1).

Table . Round 3 swab reminder experimental conditions, England, July 24 to August 11, 2020.

Condition	Reminder on day 4 after swab test kit received	Reminder on day 6 after swab test kit received	Reminder on day 8 after swab test kit received	Sample, n
Control group	Email (SMS if no email address)	SMS	None	11,194
Experimental group A	SMS	Email (SMS if no email address)	None	11,154
Experimental group B	Email (SMS if no email address)	SMS	Email (SMS if no email address)	96,337
Experimental group C	SMS	Email (SMS if no email address)	SMS	96,305

The tailored letters or SMS experiments are summarized in Table 2. Further details are available in Multimedia Appendix 1. The experiments tested whether it was possible to increase participation by different types of conditions (Textbox 1).

XSL•FO
RenderX

Table . Rounds 9, 11, and 12 registration invitation letter experimental conditions and rounds 10 and 12 SMS registration reminder experimental conditions, England, February 4 to June 7, 2021.

Age and letter or SMS type	Additional content for experiment (actual additional content used in bold text)	Sample, n
Round 9[a] (≥70 years)		
Standard invitation letter Adult	None	37,037
Experiment invitation letter A	"**It is still important to take part in this study if you have received a vaccination from COVID-19 or expect to be vaccinated in the near future**. Your participation will help DHSC assess the impact of the vaccines on COVID-19 infection rates." As well as a new sub-heading "**COVID-19 Testing Study: Take part to help measure COVID-19 infection rates among those aged 70 and over**."	37,037
Experiment invitation letter B	As per Experiment Letter A with additional line "**Older people are a vulnerable group, so we need your help to monitor prevalence**. It is still important to take part in this study if you have received a vaccination from COVID-19 or expect to be vaccinated in the near future."	37,036
Round 9 (5 - 12 years)		
Standard invitation letter Child (addressed to parent)	None	24,009
Experiment invitation letter C	**"We need to know how many children and young people have COVID-19, and how easily the new variant spreads amongst them."**	24,009
Experiment invitation letter D	As per Experiment Letter C with new sub-heading **"COVID-19 Testing Study: Take part to help measure how easily COVID-19 spreads among children and young people."**	24,008
Round 9 (all)		
Standard registration reminder letter	Blue text used	306,012
Experiment registration reminder letter E	Red text used	305,041
Round 11[b] (≥18 years)		
Standard invitation reminder letter	None	178,828
Experiment invitation reminder letter A	New content asking participants to take a test to help prevent the spread of COVID-19 and explaining that taking part would help the Government work out the best way to manage the pandemic. Also mentioned testing for new variants, that the study compared people who had been vaccinated with those who had not, and that taking part would help inform the vaccine strategy and help to avoid lockdowns.	178,809
Round 12[c] (all)		
Standard invitation final reminder letter	Double-sided	169,845
Shorter invitation final reminder letter	Single-sided	342,191
Round 10[d] (all)		
Standard SMS first reminder	Unchanged **"The study is closing soon, please register by 18 March if you want to take part."**	50,000

XSL•FO
RenderX

Age and letter or SMS type	Additional content for experiment (actual additional content used in bold text)	Sample, n
Experiment first SMS reminder	New SMS content "**Taking part will help inform decisions about the best time to lift restrictions**."	430,283
Round 11 (all)		
Standard SMS second reminder	Unchanged "**The study is closing soon, please register by 3pm on 22 April if you want to take part**."	127,028
Experiment second SMS reminder	New SMS content "**Taking part will help monitor infection rates and new variants of the virus**."	127,028
Round 12 (all)		
Standard SMS first reminder	Unchanged "**Taking part will help inform decisions about the best time to lift restrictions**."	321,042
Experiment first SMS reminder	New SMS Content **"Taking part will help monitor infection rates and new variants of the virus."**	155,683
Standard SMS second reminder	Unchanged **"Taking part will help monitor infection rates and new variants of the virus."**	272,836
Experiment second SMS reminder	New SMS Content **"Last chance to help monitor variants in your area."**	136,026

[a]Round 9 (Feb 4-23, 2021)

[b]Round 11 (Apr 15 to May 3, 2021)

[c]Round 12 (May 20 to Jun 7, 2021)

[d]Round 10 (Mar 11-30, 2021)

Textbox 1. Types of conditions tested in the tailored letters or SMS experiments.

- Using additional content in the invitation letter, tailored for the oldest and youngest age groups.
- Using color, additional content, and varying the length of the reminder letter.
- Using additional content in the SMS reminder.

Incentives Experiment

In round 15 (October 19 to Nov 5, 2021), conditional incentives (£10 [US $12.5], £20 [US $25.0], or £30 [US $37.5] gift vouchers for returning a completed swab test) were tested in a randomized controlled trial for all age groups except 5 - to 12-year-olds.

The process for obtaining consent in REACT-1 for children was undertaken differently based on participant age at the time of the invitation [18]. For 5 - to 12-year-olds, the parent or guardian was contacted via letter and asked to consent on behalf of the child. Therefore, we did not include the 5 - to 12-year-olds in the trial, as the sampled child would not be making the decision to take the swab test, and their parent would be incentivized, raising ethical and reputational concerns. For 13 - to 17-year-olds, the parent or guardian received a letter addressed to them, asking them to pass on an enclosed invitation letter addressed to their sampled child if they agreed for their child to take part in the study. As such, children aged 13 - 17 years were able to decide whether to consent to the study and take the swab test. In addition, those aged 13 to 15 years were asked at registration to confirm the name of the parent or guardian who had given them permission to take part. This was not required for those aged 16 - 17 years, as in UK health research, the Health Research Authority states that young people over 16 are presumed capable of giving consent on their own behalf [22].

Participants were randomly allocated to experimental and control groups: (1) £10 (US $12.5) conditional incentive (n=10,900), (2) £20 (US $25.0) conditional incentive (n=10,900), (3) £30 (US $37.5) conditional incentive (only for 18 - to 32-year-olds) (n=1750), and (4) control group (n=23,500). Further details of the sample size calculations are available in Multimedia Appendix 1.

XSL•FO
RenderX

The £30 (US $37.5) incentive was limited to the 18‐ to 32-year-olds because the response rate in REACT-1 was lowest among this age group. Also, there is evidence that incentives can be more effective among younger age groups [23]. Those in this age group were of particular interest as they were less likely to be vaccinated, had more social contact (and therefore were more likely to be at risk of infection), and had been particularly impacted by the pandemic (in terms of well-being, education, and employment) [21,24]. It was decided to test offering a larger (£30 [US $37.5]) incentive to this age group (and not the other age groups) to overcome their higher reluctance to take part and to better represent this group in the achieved sample. Based on the same rationale, we oversampled younger age groups to maximize the statistical power we had to detect an increase in the response rate due to the use of incentives among these groups.

The primary outcome was overall swab response rate, ie, the number of completed swabs returned (referred to as swabs returned forthwith) as a proportion of the number of invitations sent. For those invited, we knew age, sex at birth, and score from an area-level index of multiple deprivation, the Index of Multiple Deprivation (IMD) 2019 [25]. Participants were classified by quintiles of the deprivation score based on their residential postcode.

We used COVID-19 vaccination status (the proportion who had received at least one vaccine dose) as a proxy for attitudes to health behaviors and health care access, hypothesizing that REACT-1 responders would be more likely to be vaccinated than those who did not, indicating a responder bias. Thus, the difference in vaccination status at registration between the experimental and control groups was used as a crude indicator of how incentives might improve response rates in individuals less likely to participate in research, beyond sociodemographic characteristics. We also compared the COVID-19 vaccination status of those who returned a swab with the achieved population vaccination rate for that age group as a whole. To obtain information on dates of received COVID-19 vaccine doses, participant study data were linked to their NHS records from NHS Digital (now NHS England) on COVID-19 vaccination events [26] using their unique NHS number and other personal identifiers. This was only possible for study participants who had consented to data linkage. The source of vaccination data for the population vaccination rates was the NHS National Immunization Management System [27].

Ethical Considerations

The study was ethically approved by the South Central-Berkshire B Research Ethics Committee (IRAS ID: 283787). Participants provided informed consent when they registered for the study, and all data were handled securely in accordance with a detailed privacy notice. Collected data were deidentified; the data used in this study were anonymous and did not contain any personally identifiable information. Participants had the ability to opt out anytime during the research period. The study did not provide any specific compensation other than the monetary gift vouchers for returning a completed swab test as set out in the study's incentives experiment described above.

Statistical Analysis

Data analysis was conducted using IBM SPSS Statistics (version 28). As the incentives experiment was skewed toward younger age groups, swab response rates for sex at birth and area deprivation (IMD) were calculated with age-standardized weighting using 2021-based population estimates for England [28]. The percentage point difference (95% CI) and independent 2-tailed *t*-tests were used to show the absolute difference in swab response rates between the experimental and control groups and were also used to show the absolute difference in vaccination rates at registration between the experimental and control groups. Using multivariable logistic regression, we tested the impact of each of the incentive conditions on swab response rate by age, sex at birth, and area deprivation (relative response rate [RRR] with 95% CI). The reference group was the no-incentive condition—eg, the response rates for females in the £10 (US $12.5), £20 (US $25.0), and £30 (US $37.5) incentives groups were compared to females in the no-incentive group (£0 [US $0.0]). We tested interaction terms for age, sex at birth, and area deprivation by incentive (incentive*age, incentive*sex at birth, and incentive*IMD), which can be interpreted as testing whether the estimated effects of incentives on swab response rates differ by each of these 3 covariates.

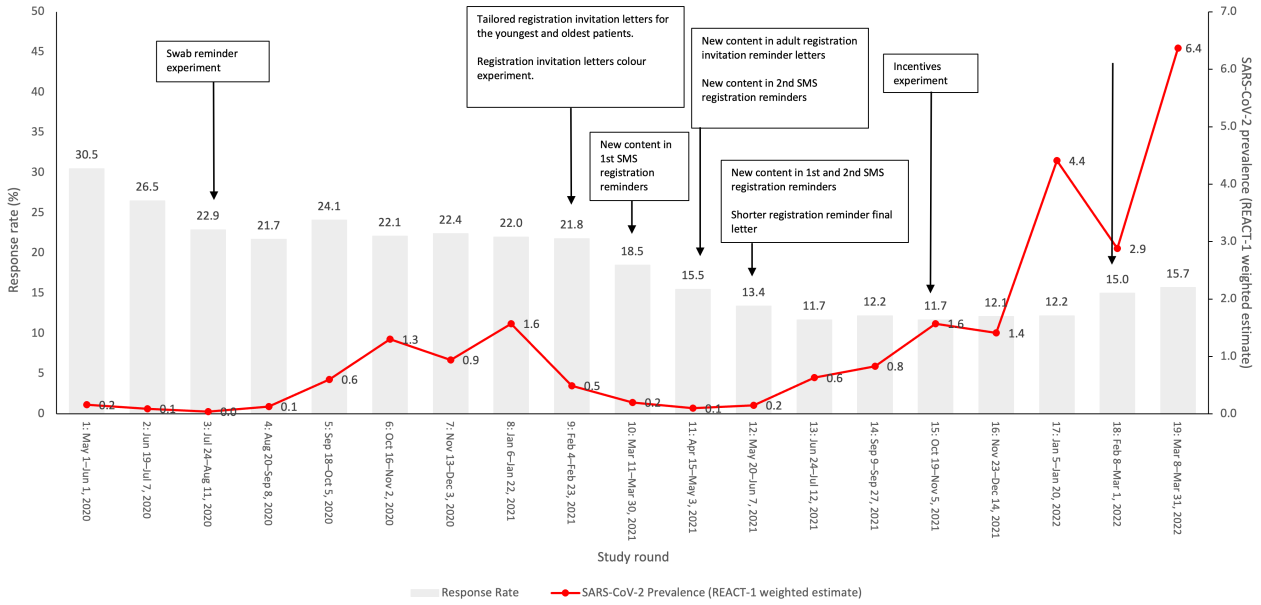
Results

Overview

Overall, 24.2% (3,393,595/14,036,117) of invitees registered for the study, and 74% (2,512,797/3,393,595) of those registered returned valid swabs, giving an overall response rate for the REACT-1 study (number of valid swabs/number of invitations) of 17.9% (2,512,797/14,036,117) [19]. Whilst the rate at which registered participants returned valid swabs remained relatively stable across rounds (range 67.2%-78.9%), response rates varied more widely, ranging from 11.7% in rounds 13 (98,233/841,227) (June 24 to July 12, 2021) and 15 (100,112/859,184) (October 19 to November 5, 2021) to 30.5% in round 1 (120,620/395,020) (May 1 to June 1, 2020, during the first lockdown in England) (Figure 2). The following groups were relatively underrepresented: younger people, men, ethnic minorities, and those living in the most deprived areas (comparing achieved sample profiles with population profiles) (Table S1 in Multimedia Appendix 2).

XSL•FO
RenderX

Figure 2. REACT-1 study timeline over 19 rounds of data collection showing response rates, SARS-CoV-2 prevalence (weighted), and timing of experiments to improve response. England, May 1, 2020 to March 31, 2022. REACT-1 Response Rate: number of valid swabs returned/number of invitations. We report weighted SARS-CoV-2 swab-positivity prevalence for individuals aged 5 years and older from all rounds of the REACT-1 study. REACT-1: REal-time Assessment of Community Transmission-1; SARS-CoV-2: Severe Acute Respiratory Syndrome CoronaVirus 2.



Swab Reminder and Tailored Letter or SMS Experiments

Table S2 in Multimedia Appendix 2 summarizes the results of the swab reminder and tailored letter or SMS experiments. Sending an additional reminder (email or SMS) to those who registered resulted in a small increase in response rate: those receiving a third swab reminder (experimental groups B and C) were more likely to return a completed swab compared to those receiving the standard Email-SMS approach (group B vs control: 73% vs 70.2%, percentage difference 2.8% [95% CI 1.9%-3.7%]; group C vs control 73.3% vs 70.2%, percentage difference 3.1% [95% CI 2.2%-4%]).

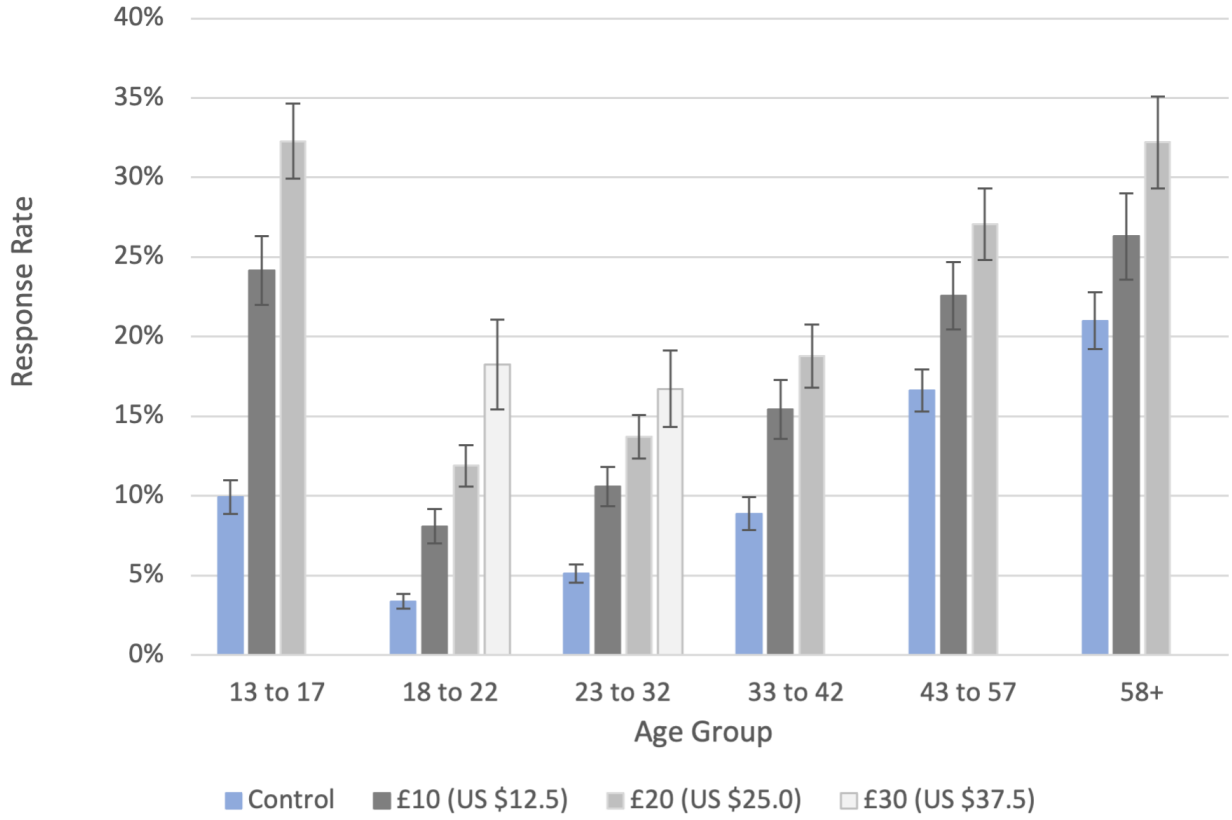
In round 9 (February 4-23, 2021), both experimental invitation letters A and B had a small but positive impact on response rate in participants aged ≥70 years of 0.9% (95% CI 0.2%-1.5%) and 1.2% (95% CI 0.6%-1.8%) percentage difference, respectively, compared to the standard invitation letter. For participants aged 5 - 12 years, experiment letter C generated a slightly higher response rate compared to the standard letter (16.6% vs 15.9%; percentage difference 0.7% (95% CI 0.1%-1.4%)). In round 11 (April 15 to May 3, 2021) and round 12 (May 20 to June 7, 2021), the experimental invitation reminder letters had a small positive impact on response rate compared to the standard letters: round 11 (new content), 5.6% vs 5.4%, percentage difference 0.2% (95% CI 0%-0.3%); round 12 (shorter), 2.3% vs 1.6%, percentage difference 0.8% (95% CI 0.7%-0.8%). We saw no effect on response rate for any of the other nonmonetary strategies (Table S2 in Multimedia Appendix 2).

Incentives Experiment

The conditional monetary incentives increased the response rate across all age groups but were particularly effective among the lowest responding groups, those aged 13 - 17 years and 18 - 22 years (Figure 3 and Tables S3 and S4 in Multimedia Appendix 2). Table 3 shows the RRR for each incentive level by age, sex at birth, and area deprivation. The higher the monetary value of the incentive, the higher the response rate. For example, in participants aged 18 - 22 years, the response rate in the control group was 3.4% (95% CI 2.9%-3.8%), increasing to 8.1% (95% CI 7.0%-9.2%), 11.9% (95% CI 10.6%-13.2%), and 18.2% (95% CI 15.4%-21.1%) with £10 (US $12.5), £20 (US $25.0), and £30 (US $37.5) incentives, respectively. The largest relative increase was with the £30 (US $37.5) incentive in 18 - to 22-year-olds (RRR 5.4 [95% CI 4.4-6.7]) (Table 3). All incentive conditions led to a greater increase in response rate in younger age groups. The £20 (US $25.0) incentive led to a greater increase in the more deprived areas, RRR 2.7 (95% CI 2.2-3.3) for the most deprived quintile and RRR 1.8 (95% CI 1.6-2.1) for the least deprived.

XSL•FO
RenderX

Figure 3. Swab response rates and 95% CIs for the intervention and control groups in the incentives experiment in round 15, England, October 19 to November 5, 2021. Note: Participants randomly allocated to experimental and control groups. (1) £10 (US $12.5) conditional incentive (n=10,900), (2) £20 (US $25.0) conditional incentive (n=10,900), (3) £30 (US $37.5) conditional incentive (only for 18- to 32-year-olds) (n=1750), and (4) control group (n=23,500).



XSL•FO
RenderX

Table . Variation in relative response rates (RRR) and 95% CI for the interventions compared to the control group by age, sex at birth, and area deprivation (IMD[a]) in the incentives experiment in round 15, England, October 19 to November 5, 2021.

	£10 (US $12.5) RRR (95% CI)	£20 (US $25.0) RRR (95% CI)	£30 (US $37.5) RRR (95% CI)
Age[b] (years)			
13-17	2.4 (2.1-2.8)	3.3 (2.6-3.7)	N/A
18-22	2.4 (2.0-2.9)	3.5 (3.0-4.2)	5.4 (4.4-6.7)
23-32	2.1 (1.7-2.4)	2.7 (2.3-3.1)	3.3 (2.7-3.9)
33-42	1.7 (1.5-2.0)	2.1 (1.8-2.5)	—[c]
43-57	1.4 (1.2-1.5)	1.6 (1.5-1.8)	—
58+	1.3 (1.1-1.4)	1.5 (1.4-1.7)	—
P value for interaction between incentive and age	<.001	<.001	<.001
Sex at birth[bd]			
Male	1.4 (1.3-1.6)	1.8 (1.7-2.1)	3.7 (3.0-4.7)
Female	1.5 (1.4-1.6)	1.8 (1.6-2.0)	3.7 (3.1-4.4)
P value for interaction between incentive and sex at birth	.37	.96	.68
IMD[bd]			
1—most deprived	1.8 (1.5-2.3)	2.7 (2.2-3.3)	4.8 (3.3-7.0)
2	1.6 (1.3-1.9)	1.9 (1.6-2.3)	4.0 (2.9-5.5)
3	1.4 (1.2-1.6)	1.5 (1.3-1.8)	3.4 (2.5-4.8)
4	1.3 (1.1-1.5)	1.7 (1.5-1.9)	3.2 (2.4-4.3)
5—least deprived	1.5 (1.3-1.7)	1.8 (1.6-2.1)	3.5 (2.6-4.8)
P value for interaction between incentive and IMD	.38	.01	0.72

[a]IMD: Index of Multiple Deprivation

[b]Reference group, ie, the reference group for each row is the no incentive condition. For example, the RRR for female £10 (US $12.5), female £20 (US $25.0), and female £30 (US $37.5) is versus female £0 (US $0.0). *P* value for main effect of incentive on response rate for all row comparisons <.001.

[c]Not applicable

[d]Age-standardized weighting applied to calculate swab response rate with the control group totals used as the sample profiles.

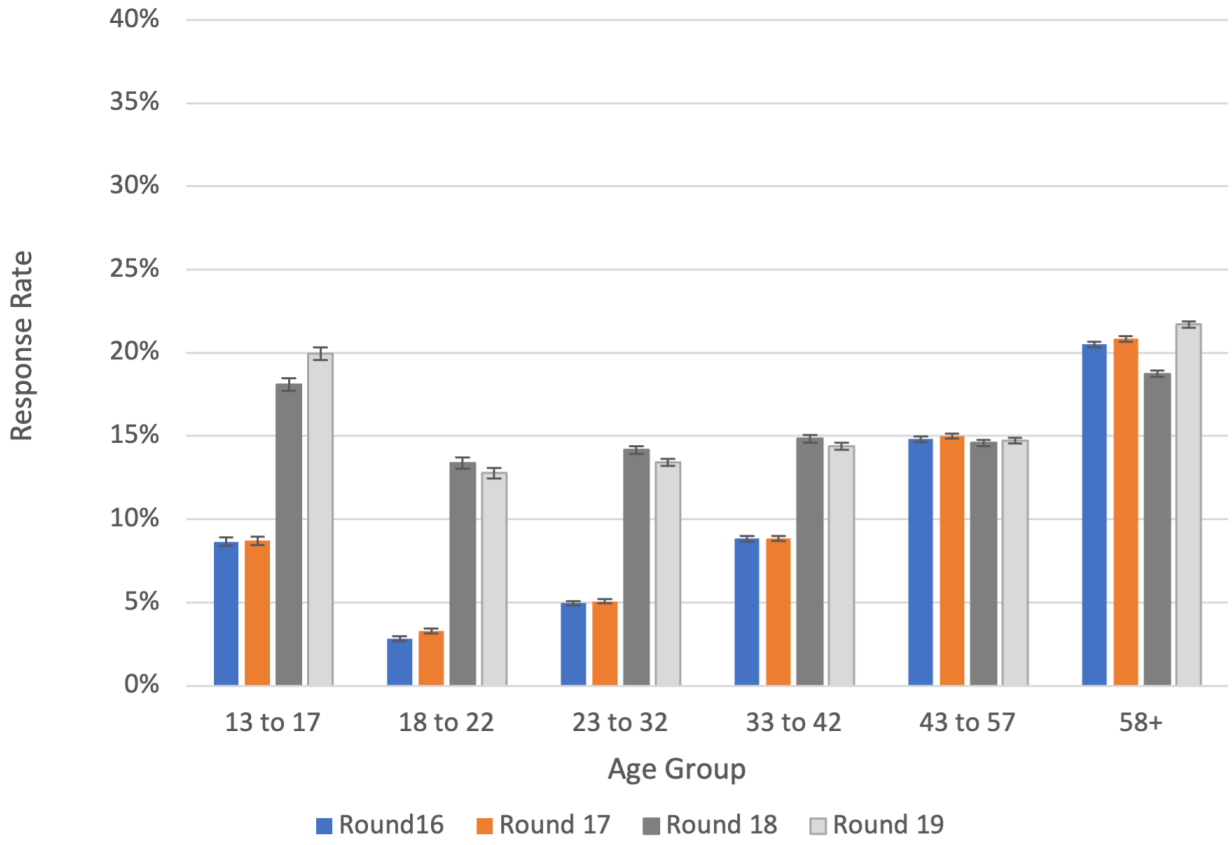
Following the results of the selective use of incentives in round 15, they were introduced in rounds 18 (Feb 8-Mar 1, 2022) and 19 (Mar 8-Mar 31, 2022). For returning their completed test, those aged 13 - 17 and 35 - 44 years were offered a gift voucher worth £10 (US $12.5), while those aged 18 - 34 years were offered a voucher worth £20 (US $25.0). In these final 2 rounds, this had the effect of increasing the swab response rate in these groups and was associated with less variation in response rate by age (Figure 4), suggesting that the selective use of incentives reduced participation bias by age.

Table S5 in Multimedia Appendix 2 shows the effective sample sizes and sample efficiency for each round of REACT-1. The effective sample size measures the size of a (unweighted) simple random sample that would achieve the same precision (standard error) as the design used. The efficiency of a sample is given by the ratio of the effective sample size to the actual sample size. Rounds 18 and 19, where selective use of incentives was used, saw the fourth and second highest (respectively) effective sample sizes of any REACT-1 round, and the highest sample efficiency for any REACT-1 round.

Overall, vaccination rates were higher in REACT-1 participants than in the general population (Tables S6 and S7 in Multimedia Appendix 2). For example, by October 24, 2021, just over 3 quarters of 18 - to 22-year-olds had received at least one vaccine dose nationally (Table S6 in Multimedia Appendix 2) [29], lower than the 84.0% (95% CI 78.1%-88.6%) in the round 15 (October 19-November 5, 2021) control group for that age (Table S7 in Multimedia Appendix 2). With the incentives that proportion declined to 82.1% (95% CI 76.0%-86.8%), 73.9% (95% CI 68.6%- 78.7%), and 75.9% (95% CI 67.9%-82.5%) for £10 (US $12.5), £20 (US $25.0), and £30 (US $37.5), respectively, suggesting that the selective use of incentives reduced participation bias in relation to vaccination status as a proxy for health behaviors.

XSL•FO
RenderX

Figure 4. Swab response rates and 95% CIs for round 16 (November 23 to December 14, 2021), round 17 (January 5-20, 2022), and rounds 18 (February 8 to March 1, 2022) and 19 (March 8-31, 2022) in which incentives were used selectively, England. Note: Incentive amounts used in rounds 18 and 19: £10 (US $12.5) for 13- to 17-year-olds and 35- to 44-year-olds, £20 (US $25.0) for 18- to 34-year-olds, and no incentives for other age groups.



Discussion

Principal Findings

In this large population-based study of the prevalence of SARS-CoV-2 infection in England, we tested several measures to increase response rates and reduce nonresponder bias. We found that changes to the wording of letters, timing, and numbers of reminders made only limited differences to response rates, with a maximum increase in response rate of 3.1 percentage points for additional swab reminders sent to people who had already registered for the study. Sending an additional reminder, regardless of its form (SMS or email), increased response. This is consistent with other studies in the literature [30]. These reminder strategies may have helped slow, but did not halt, the decline in response rates over time observed during REACT-1. Nonetheless, these findings informed the swab reminder strategy and invitation letter wording in later rounds. In contrast, the offer of a financial incentive conditional on the return of a completed swab made a more substantial difference of up to 22.3 percentage points and was particularly effective in those with a lower propensity to respond: younger age groups and those living in more deprived areas. Similarly, incentives increased the return of completed swabs by unvaccinated individuals so that COVID-19 vaccination rates were more in keeping with those in the general population at the time. Thus, the selective use of incentives may reduce nonresponder bias in relation to factors of interest in population health research beyond sociodemographic characteristics.

The selective use of incentives was subsequently adopted from round 18, making the achieved sample more representative by age, with a reduction in age-based variation in response rates. Previous research suggests that ethnic minorities [31], individuals living in more deprived areas [32], those in urban areas [33], and the youngest and oldest age groups [34,35] are the least likely to respond in general population surveys. Using incentives selectively allowed us, at modest cost, to increase recruitment among such groups and hence increase the effective sample size; thus, in round 18, the effective sample size was over 10,000 greater compared to round 17, even though we received circa 7000 fewer swabs. We were able to reduce the number of invitations sent out while achieving a similar number of completed swabs returned as in earlier rounds when response rates were higher.

Using incentives selectively has been tried in UK social surveys previously and is common practice in the United States, where studies show they are cost-effective, improve response, and reduce bias [11,13,23,36]. From an ethical perspective, in the selective use of (versus universal) incentives, it was important to consider not only issues of equity but also cost and the public interest in continuing to obtain high-quality data, covering all sectors of society, to monitor the spread of a serious disease. This needs to be balanced against the possible disappointment of some participants who learn others are being offered a (larger)

XSL•FO
RenderX

incentive. These considerations might apply to many population-based surveys. We accept that the argument for using incentives selectively may have been more persuasive in the context of REACT-1, a study to measure the spread of SARS-CoV-2 during the pandemic, the policy responses to which had far-reaching consequences for the way of life of every person in England.

Both unconditional and conditional financial incentives have been shown to significantly increase response rates to both postal and web-based surveys [37,38]. Although unconditional incentives appear to have the largest effect, the conditional approach is more cost-effective [37,38]. Unconditional incentives have been used in social surveys in the United Kingdom, and in experiments in how to increase response rates [39,40]. Unconditional incentives were not an option for REACT-1 due to the constraints of the survey budget.

Limitations

In terms of limitations, it was not possible to ascertain the extent to which noncontact (ie, the intended recipient did not receive the invitation letter) accounted for nonresponse. Such situational factors, for example, not informing their general practitioner of a change in address or having moved with no forwarding address (shown to be greater for young adults and lower socioeconomic groups) [32] will not be affected by the experimental conditions; therefore, our estimates of effect are likely conservative, as invitations sent out do not necessarily mean that invitations were received. In addition, the unique circumstances of carrying out such assessments of response rates during a global pandemic may not "read across" to other less pressing issues.

Conclusions

We achieved small improvements in response rates by varying the number, order, and content of invitations and reminders but much larger effects were seen through the use of monetary incentives. Lessons learnt from the REACT-1 study may help inform the design and implementation of future population-based surveys where the intent is to obtain as representative a sample as possible and to reduce nonresponse bias at reasonable cost. The results suggest selectively using incentives with younger and more deprived individuals may be justifiable to achieve these ends.

Acknowledgments

This work was supported by the Department of Health and Social Care in England. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of this article. Infrastructure support for this research was provided by the National Institute for Health (NIHR) Imperial Biomedical Research Centre (BRC). PE is director of the Medical Research Council (MRC) Centre for Environment and Health (MR/L01341X/1 and MR/S019669/1). PE acknowledges support from Health Data Research UK, the NIHR Imperial Biomedical Research Centre; NIHR Health Protection Research Units in Chemical and Radiation Threats and Hazards, and Environmental Exposures and Health; the British Heart Foundation Centre for Research Excellence at Imperial College London (RE/18/4/34215); and the UK Dementia Research Institute at Imperial College London (MC_PC_17114). HW acknowledges support from an NIHR Senior Investigator Award, the Wellcome Trust (205456/Z/16/Z), the NIHR Imperial Biomedical Research Centre, the NIHR School of Public Health Research, and the NIHR Applied Research Collaboration North West London. CAD and SR acknowledge support from MRC Centre for Global Infectious Disease Analysis (reference MR/R015600/1), jointly funded by the UK MRC and the UK Foreign, Commonwealth & Development Office (FCDO), under the MRC/FCDO Concordat agreement. CAD acknowledges support from the NIHR Health Protection Research Unit (HPRU) in Emerging and Zoonotic Infections (NIHR200907), and the NIHR funded Vaccine Efficacy Evaluation for Priority Emerging Diseases (PR-OD-1017-20007). SR acknowledges support from the NIHR HPRU in Respiratory Infections, Wellcome Trust (200861/Z/16/Z, 200187/Z/15/Z), and Centres for Disease Control and Prevention (US, U01CK0005-01-02). MCH acknowledges support from Cancer Research UK, Population Research Committee Project grant "Mechanomics" (grant 22184). MCH acknowledges support from the H2020-EXPANSE (Horizon 2020 grant 874627) and H2020-LongITools (Horizon 2020 grant 874739).We thank key collaborators on this work: K Beaver, S Finlay, J Kennedy, D Peskett, G Welch, and K Ward at Ipsos. We also thank the REACT Public Advisory Panel who provided input which changed the wording of the invitation letters and reminders and provided advice on the design and equity concerns in relation to the incentives experiment.

Data Availability

The datasets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

NG, GP, SC, SR, HW, and PE designed the study. NG conducted the analyses. KP provided statistical support. CJA, NG, HW, and PE designed the analytical strategy and helped to interpret the findings. CJA, NG, and HW conducted the literature review. CJA, NG, HW and PE drafted the manuscript. HW, GSC, CJA, AD, CAD, MCH, SR, DA, WSB, and PE provided study oversight. AD and PE obtained funding. All authors have reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Supplementary methods.
[DOCX File, 19 KB - publichealth_v11i1e60022_app1.docx]

Multimedia Appendix 2
Supplementary tables.
[DOCX File, 55 KB - publichealth_v11i1e60022_app2.docx]

References

1. Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. Public Opin Q 2008;72(2):167-189. [doi: 10.1093/poq/nfn011]
2. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
3. Allen N, Sudlow C, Downey P, et al. UK Biobank: current status and what it means for epidemiology. Health Policy Technol 2012 Sep;1(3):123-126. [doi: 10.1016/j.hlpt.2012.07.003]
4. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Allen NE. OP41: the representativeness of the UK biobank cohort on a range of sociodemographic, physical, lifestyle and health-related characteristics. J Epidemiol Community Health 2016 Sep;70(Suppl 1):A26. [doi: 10.1136/jech-2016-208064.41]
5. Swanson JM. The UK Biobank and selection bias.. Lancet 2012 Jul;380(9837):110. [doi: 10.1016/S0140-6736(12)61179-9]
6. Keyes KM, Westreich D. UK Biobank, big data, and the consequences of non-representativeness. Lancet 2019 Mar 30;393(10178):1297. [doi: 10.1016/S0140-6736(18)33067-8] [Medline: 30938313]
7. Andrade C. Sample size and its importance in research. Indian J Psychol Med 2020;42(1):102-103. [doi: 10.4103/IJPSYM.IJPSYM_504_19] [Medline: 31997873]
8. Gelman A, Carlin J. Poststratification and weighting adjustments. In: Groves R, Dillman D, Eltinge J, Little R, editors. Survey Nonresponse: Wiley; 2000:289-302.
9. Pforr K, Blohm M, Blom AG, et al. Are Incentive effects on response rates and nonresponse bias in large-scale, face-to-face surveys generalizable to Germany? Evidence from ten experiments. Public Opin Q 2015;79(3):740-768. [doi: 10.1093/poq/nfv014]
10. Lange C, Finger JD, Allen J, et al. Implementation of the European health interview survey (EHIS) into the German health update (GEDA). Arch Public Health 2017;75:40. [doi: 10.1186/s13690-017-0208-6] [Medline: 28936356]
11. Singer E, Hoewyk JV, Gebler N, Raghunathan TE, Mcgonagle KA. The effect of incentives on response rates in interviewer-mediated surveys. J Off Stat 1999;15(2):217-230.
12. Wu MJ, Zhao K, Fils-Aime F. Response rates of online surveys in published research: a meta-analysis. Comput Hum Behav Rep 2022 Aug;7:100206. [doi: 10.1016/j.chbr.2022.100206]
13. Nicolaas G CE, Davies B. The use of incentives to recruit and retain hard-to-get populations in longitudinal studies. NatCen Soc Res 2019 May 31.
14. Mavletova A, Couper MP. Mobile web survey design: scrolling versus paging, SMS versus e-mail invitations. J Surv Stat Methodol 2014 Dec;2(4):498-518. [doi: 10.1093/jssam/smu015]
15. Cabrera-Álvarez P, Lynn P. Text messages to facilitate the transition to web-first sequential mixed-mode designs in longitudinal surveys. J Surv Stat Methodol 2024 Jun 1;12(3):651-673. [doi: 10.1093/jssam/smae003]
16. Lynn P. Targeted appeals for participation in letters to panel survey members. PUBOPQ 2016;80(3):771-782. [doi: 10.1093/poq/nfw024]
17. Taylor KSM, Counsell CE, Harris CE, Gordon JC, Fonseca SCC, Lee AJ. In a randomized study of envelope and ink color, colored ink was found to increase the response rate to a postal questionnaire. J Clin Epidemiol 2006 Dec;59(12):1326-1330. [doi: 10.1016/j.jclinepi.2006.03.011] [Medline: 17098576]
18. Riley S, Atchison C, Ashby D, et al. REal-time Assessment of Community Transmission (REACT) of SARS-CoV-2 virus: study protocol. Wellcome Open Res 2020;5(200):200. [doi: 10.12688/wellcomeopenres.16228.2] [Medline: 33997297]
19. Elliott P, Whitaker M, Tang D, et al. Design and implementation of a National SARS-CoV-2 Monitoring Program in England: REACT-1 study. Am J Public Health 2023 May;113(5):545-554. [doi: 10.2105/AJPH.2023.307230] [Medline: 36893367]
20. Ward H, Whitaker M, Flower B, et al. Population antibody responses following COVID-19 vaccination in 212,102 individuals. Nat Commun 2022;13(1):907. [doi: 10.1038/s41467-022-28527-x]
21. Eales O, Wang H, Haw D, et al. Trends in SARS-CoV-2 infection prevalence during England's roadmap out of lockdown, January to July 2021. PLoS Comput Biol 2022 Nov;18(11):e1010724. [doi: 10.1371/journal.pcbi.1010724] [Medline: 36417468]
22. Adolescent consent. NHS Health Research Authority. 2022. URL: https://www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/adolescent-consent/

XSL•FO
RenderX

#:~:text=Young%20people%20over%20the%20age%20of%2016%20are,consent%20from%20their%20parents%20provided%20they%20are%20competent [accessed 2024-12-31]
23. Wong E. Incentives in Longitudinal Studies, CLS Working Paper 2020/1: UCL Centre for Longitudinal Studies; 2020.
24. Ward H, Atchison C, Whitaker M, et al. SARS-CoV-2 antibody prevalence in England following the first peak of the pandemic. Nat Commun 2021 Feb 10;12(1):905. [doi: 10.1038/s41467-021-21237-w] [Medline: 33568663]
25. Comparisons of all-cause mortality between european countries and regions: january to june 2020. Office for National Statistics. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/comparisonsofallcausemortalitybetweeneuropeancountriesandregions/januarytojune2020 [accessed 2024-12-31]
26. COVID-19 vaccine data. NHS Digital. 2022. URL: https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/directions-and-data-provision-notices/data-provision-notices-dpns/covid-19-vaccine-data [accessed 2022-12-31]
27. Tessier E, Edelstein M, Tsang C, et al. Monitoring the COVID-19 immunisation programme through a national immunisation management system - England's experience. Int J Med Inform 2023 Feb;170:104974. [doi: 10.1016/j.ijmedinf.2022.104974] [Medline: 36577202]
28. Population and household estimates, england and wales: census 2021. Office for National Statistics. 2022. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/populationandhouseholdestimatesenglandandwales/census2021 [accessed 2024-12-31]
29. COVID-19 vaccinations archive. 2021. URL: https://www.england.nhs.uk/statistics/statistical-work-areas/covid-19-vaccinations/covid-19-vaccinations-archive [accessed 2024-12-31]
30. Sammut R, Griscti O, Norman IJ. Strategies to improve response rates to web surveys: a literature review. Int J Nurs Stud 2021 Nov;123:104058. [doi: 10.1016/j.ijnurstu.2021.104058] [Medline: 34454334]
31. D. Royal K. Surveying Black or African American populations: challenges and solutions in medicine and education. Educ Med J 2019 Jun;11(2):59-61. [doi: 10.21315/eimj2019.11.2.8]
32. Goodman A, Gatward R. Who are we missing? Area deprivation and survey participation. Eur J Epidemiol 2008;23(6):379-387. [doi: 10.1007/s10654-008-9248-0] [Medline: 18409005]
33. Inpatient experience survey 2018: technical report. Scottish Government. 2018 Aug 28. URL: https://www.gov.scot/publications/inpatient-experience-survey-2018-technical-report/ [accessed 2024-12-31]
34. Harrison S, Alderdice F, Henderson J, Redshaw M, Quigley MA. Trends in response rates and respondent characteristics in five National Maternity Surveys in England during 1995-2018. Arch Public Health 2020;78(46):46. [doi: 10.1186/s13690-020-00427-w] [Medline: 32509303]
35. Klein DJ, Elliott MN, Haviland AM, et al. Understanding nonresponse to the 2007 Medicare CAHPS survey. Gerontologist 2011 Dec;51(6):843-855. [doi: 10.1093/geront/gnr046] [Medline: 21700769]
36. Yu S, Alper HE, Nguyen AM, et al. The effectiveness of a monetary incentive offer on survey response rates and response completeness in a longitudinal study. BMC Med Res Methodol 2017 Apr 26;17(1):77. [doi: 10.1186/s12874-017-0353-1] [Medline: 28446131]
37. Young JM, O'Halloran A, McAulay C, et al. Unconditional and conditional incentives differentially improved general practitioners' participation in an online survey: randomized controlled trial. J Clin Epidemiol 2015 Jun;68(6):693-697. [doi: 10.1016/j.jclinepi.2014.09.013] [Medline: 25450450]
38. Abdelazeem B, Hamdallah A, Rizk MA, et al. Does usage of monetary incentive impact the involvement in surveys? A systematic review and meta-analysis of 46 randomized controlled trials. PLoS One 2023;18(1):e0279128. [doi: 10.1371/journal.pone.0279128] [Medline: 36649255]
39. Office for National Statistics. Surveys and provision of monetary voucher for participation. 2018. URL: https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/surveysandprovisionofmonetaryvoucherforparticipation [accessed 2024-12-31]
40. Office for National Statistics. Crime survey for england and wales: technical note (Y/E march 2022). 2023. URL: https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/crimeandjustice/methodologies/crimeandjusticemethodology/csewtechnicalnoteyemarch2022.pdf [accessed 2024-12-31]

Abbreviations

Ct: cycle threshold
IMD: Index of Multiple Deprivation
NHS: National Health Service
REACT-1: Real-time Assessment of Community Transmission-1
RRR: relative response rate
rt-PCR: reverse transcriptase polymerase chain reaction

XSL•FO
RenderX

Edited by A Mavragani; submitted 29.04.24; peer-reviewed by D Falk, P Lynn; revised version received 12.11.24; accepted 13.11.24; published 09.01.25.

Please cite as:
Atchison CJ, Gilby N, Pantelidou G, Clemens S, Pickering K, Chadeau-Hyam M, Ashby D, Barclay WS, Cooke GS, Darzi A, Riley S, Donnelly CA, Ward H, Elliott P
Strategies to Increase Response Rate and Reduce Nonresponse Bias in Population Health Research: Analysis of a Series of Randomized Controlled Experiments during a Large COVID-19 Study
JMIR Public Health Surveill 2025;11:e60022
URL: https://publichealth.jmir.org/2025/1/e60022
doi:10.2196/60022

©Christina J Atchison, Nicholas Gilby, Galini Pantelidou, Sam Clemens, Kevin Pickering, Marc Chadeau-Hyam, Deborah Ashby, Wendy S Barclay, Graham S Cooke, Ara Darzi, Steven Riley, Christl A Donnelly, Helen Ward, Paul Elliott. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 9.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Investigating Social Network Peer Effects on HIV Care Engagement Using a Fuzzy-Like Matching Approach: Cross-Sectional Secondary Analysis of the N2 Cohort Study

Cho-Hee Shrader[1,2], MPH, PhD; Dustin T Duncan[1], MS, ScD; Redd Driver[3], PhD; Juan G Arroyo-Flores[4], PhD; Makella S Coudray[5], MPH, PhD; Raymond Moody[6], PhD; Yen-Tyng Chen[7], PhD; Britt Skaathun[8], PhD; Lindsay Young[9], PhD; Natascha del Vecchio[10], MS; Kayo Fujimoto[11], PhD; Justin R Knox[12,13,14], MA, PhD; Mariano Kanamori[15], PhD; John A Schneider[16,17], MPH, MD

[1]Department of Epidemiology, Columbia University, 722 W 168th St, New York, NY, United States

[10]A Place for Rover, Seattle, WA, United States

[11]Center for Health Promotion and Prevention Research, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, United States

[12]Department of Psychiatry, Columbia University Irving Medical Center, Columbia University, New York, NY, United States

[13]New York State Psychiatric Institute, HIV Center for Clinical and Behavioral Studies, New York, NY, United States

[14]Department of Sociomedical Sciences, Columbia University, New York, NY, United States

[15]Department of Public Health Sciences, School of Medicine, University of Miami, Miami, FL, United States

[16]Department of Medicine, University of Chicago, Chicago, IL, United States

[17]Department of Public Health Sciences, University of Chicago, Chicago, IL, United States

[2]College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, United States

[3]New York Department of Mental Health and Hygiene, New York, NY, United States

[4]Fors Marsh, Arlington, VA, United States

[5]Population Health Sciences, University of Central Florida, Orlando, FL, United States

[6]University of Connecticut, Hartford, CT, United States

[7]Edward J. Bloustein School of Planning and Public Policy, Rutgers University, New Brunswick, NJ, United States

[8]School of Medicine, University of California San Diego, San Diego, CA, United States

[9]School for Communication and Journalism, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Cho-Hee Shrader, MPH, PhD
Department of Epidemiology, Columbia University, 722 W 168th St, New York, NY, United States

Abstract

Background: Social network data are essential and informative for public health research and implementation as they provide details on individuals and their social context. For example, health information and behaviors, such as HIV-related prevention and care, may disseminate within a network or across society. By harmonizing egocentric and digital networks, researchers may construct a sociocentric-like “fuzzy” network based on a subgroup of the population.

Objective: We aimed to generate a more complete sociocentric-like “fuzzy” network by harmonizing alternative sources of egocentric and digital network data to examine relationships between participants in the Neighborhoods and Networks (N2) cohort study. Further, we examined network peer effects of the status-neutral HIV care continuum cascade.

Methods: Data were collected from January 2018 to December 2019 in Chicago, Illinois, United States, from a community health center and via peer referral sampling as part of the N2 cohort study, comprised of Black sexually minoritized men and gender expansive populations. Participants provided sociodemographics, social networks, sexual networks, mobile phone contacts, and Facebook friends list data. Lab-based information about the HIV care continuum cascade was also collected. We used an experimental approach to develop and test a fuzzy matching algorithm to construct a more complete network across social, sexual, phone, and Facebook networks using R and Excel. We calculated social network centrality measures for each of these networks and then described the HIV care continuum within the context of each network. We then used Spearman correlation and a network autocorrelation model to examine social network peer effects with HIV status and care engagement.

Results: A total of 412 participants resulted in 2054 network connections (ties) across the confidant and sexual partner social networks (participants=387; ties=445), peer referral network (participants=412; ties=362), phone contacts (participants=273;

XSL•FO
RenderX

ties=362), and Facebook network (participants=144; ties=1383), reaching the entire study sample in one fully connected "fuzzy" network. Results from the individual networks' autocorrelation model suggest there are no peer effects on status-neutral HIV care engagement. Results from the final fuzzy-like sociocentric network autocorrelation model, adjusted for HIV serostatus, suggest that participants who were proximate to network members engaged in HIV care were significantly more likely to be engaged in care (ρ=0.128, SE 0.064; *P*=.045).

Conclusions: Using alternative sources of network data allowed us to fuzzy match a more complete network: fuzzy matching may identify hidden ties among participants that were missed by examining alternative sources of network data separately. Although sociocentric studies require significant resources to implement, more complete sociocentric-like networks may be generated using a fuzzy match approach that leverages egocentric, peer referral, and digital networks. Enriching offline networks with digital network data may provide insights into characteristics and norms that egocentric approaches may not be able to capture.

International Registered Report Identifier (IRRID): RR2-10.2196/48548

(JMIR Public Health Surveill 2025;11:e64497) doi:10.2196/64497

KEYWORDS

social network analysis; HIV; social media; social environment; sexual and gender minorities; African Americans; social network; social networking; public health; HIV care; health center; peer referral sampling; neighborhoods and networks; N2; Facebook; sexual health; mobile phone; fuzzy; innovation; digital health; mHealth; eHealth; surveillance; mobile health

Introduction

Social epidemiology is the "branch of epidemiology concerned with the way that social structures, institutions, and relationships influence health." [1] Within social epidemiology, social network models focus on how social networks can influence attitudes and knowledge surrounding health as they are a powerful force for influencing health behaviors [2,3]. As health information and behaviors can disseminate within a network, or across society [4], network science, a set of theories and methods for examining social relationships, has garnered interest among researchers, including social scientists and public health researchers, who are seeking to describe social phenomena [2,5-8]. Within network science, there is a fundamental dichotomy between sociocentric and egocentric network characterization [8]. Sociocentric study designs collect information about a "complete" network among members of a population or cohort of interest, whereas egocentric studies focus on the network of contacts that immediately surround a focal individual (ego) and often collect information from subgroups sampled from a larger population [1,9].

Sociocentric approaches are one of many frameworks for examining social network data; however, they are often perceived by researchers as the "gold standard" of network science in that they presumably include all nodes and ties within a network. The complexity and resource demands of collecting such data have led social scientists to use egocentric study designs, which are considered more pragmatic in that their elicitation can be more easily worked into traditional survey data collection. Despite their pragmatism, however, egocentric networks are limited as they intentionally ignore potential connections between ego networks and they tend to represent only a subsample of an existing network [5]. Given these tensions between the resource demands of sociocentric network data collection and the limits of egocentric network data collection, there is a need to find alternative sources of network data that could be collected from individuals without knowledge of who else is in the network and could be used to locate "hidden" connections among members of a focal population.

Individuals are often thought to have upwards of 1500 connections; crafting an egocentric survey design to capture all personal networks or a sociocentric approach to include all individuals bound within a network would be burdensome and analytically complex, although not impossible [10]. One way of bolstering traditional egocentric research and circumventing the steep demands of sociocentric designs is by incorporating digital epidemiology, which includes digital methods and data. Digital data includes mobile phone network data, data derived from the internet, and social media network data [11]. In the United States, cell phone and internet use are ubiquitous, with 97% and 95% of US adults owning a cell phone [12] and using the internet [13], respectively. Further, the majority of US adults frequent social networking websites such as YouTube (83%) and Facebook (68%) [14].

By harmonizing egocentric and digital networks, researchers may construct a "fuzzy" network based on a subgroup of the population. Although this more complete network would not be a "pure" sociocentric network as described by Laumann's boundary specification [10,15], it could offer hidden ties that a purely egocentric network could not unveil. Deterministic linkage, a technique used to link records that may be less than a 100% match, could be used to triangulate egocentric data from multiple sources (eg, digital and phone networks) into a fuzzy network [16]. This matching compares all records of one dataset to another and creates matching pairs using prespecified linkage specifications such as phoenetic name and age, which are then manually validated to ensure a true match. Social network researchers could use deterministic linkage algorithms to explore if a study participant's nominated tie (ie, alter) could be another participant in the study (ie, ego) [16]. Including additional types of networks such as mobile phonebook networks and online social networks could further aid the construction of a more complete network of a subgroup of a population with variable tie typologies that together may provide a more expansive understanding of social influence.

XSL•FO
RenderX

Using fuzzy matching and digital networks to generate connections across egocentric networks could provide insights into "hidden" relationships and social influence dynamics that are commonly excluded due to the limitations of egocentric approaches. Thus, our aim was to construct a fuzzy sociocentric-like network using egocentric confidant and sexual networks, peer referral ties, and digital networks (mobile phone and Facebook contacts), then to identify correlations between each of these networks and the fuzzy sociocentric-like network and status-neutral HIV care engagement among a cohort of Black sexually minoritized and gender expansive people (SGM) in Chicago [17]. This group is a relatively underreached group that continues to experience HIV-related health disparities at a disproportionate rate; approximately 1 in 3 Black sexually minoritized men are estimated to be diagnosed with HIV by age 50 if current rates persist [18,19]. As pre-exposure prophylaxis (PrEP) can reduce HIV acquisition by up to 99% [4-10], and antiretroviral treatment (ART) effectively eliminates forward transmission of HIV [11,12], together they can effectively end the HIV epidemic among Black SGM. However, Black SGM are not being reached by or using PrEP and ART at optimal rates to curb this HIV epidemic. Black SGM experience poor HIV-related outcomes in the United States due to historical and structural barriers and intersectional discrimination [20-23], which have manifested as reduced access to PrEP and ART. Although supportive social relationships are shown to promote PrEP and ART uptake [24], there is scant research that examines personal networks to identify the role of social networks on status-neutral HIV outcomes. A status-neutral approach proposes the same engagement approach for HIV prevention and care, regardless of HIV serostatus, and focuses instead on the common final state of engagement in HIV prevention (ie, PrEP) or care (ie, ART) activities [25].

Previous studies have suggested that social networks may be leveraged to surmount health disparities [26-31]; thus, this experimental approach is novel and important in reaching health equity for this marginalized group. Given that the data in the present study were collected 6 years ago, harmonization provides an opportunity to enhance their relevance and applicability by integrating them into a fuzzy-like sociocentric network. At the time that these data were collected, they were cutting edge, and despite 6 years passing since the collection of this data, HIV-related disparities in Chicago have exhibited stability over time, suggesting that the information derived from this dataset may still be applicable. To our knowledge, this robust dataset is the most updated and comprehensive social network dataset of Black SGM available since the uConnect cohort in Chicago [28], and it presents a unique opportunity for secondary analyses on a persistent health disparity (HIV) that disproportionately impacts Black SGM. For example, although HIV diagnoses in Chicago decreased by 7.7% from 2018 to 2022, Black non-Hispanic people (46.7% vs 47.1%) and men who have sex with men (60% vs 60.7%) continued to disproportionately comprise the same percentage of new HIV diagnoses while transgender women experienced an increase in new HIV diagnoses by 12.2%. This necessitates a better understanding of how networks may play a role in HIV prevention and care. Thus, the aim of this paper was to generate a more complete sociocentric-like "fuzzy" network by harmonizing alternative sources of egocentric and digital network data to examine relationships between participants in the Neighborhoods and Networks (N2) cohort study. Further, we aimed to examine the network peer effects of the status-neutral HIV care continuum cascade using our sociocentric-like "fuzzy" network.

Methods

Study Sample and Recruitment

Data were collected at baseline from January 2018 to December 2019 from the Chicago site of the multisite N2 cohort study, a prospective longitudinal study. Participants were recruited from a community health clinic and via peer referral in the South Side of Chicago, Illinois, United States. The community health clinic provides next-generation testing and notification, integrated biobehavioral HIV prevention, and community mobilization with the aim of eliminating new HIV transmission in the United States by 2041. The recruitment strategy was designed to leverage clients of the clinic and participants from existing research projects. We used a probabilistic variant of peer referral sampling that allowed participants to refer up to 6 members of their social networks. Participant inclusion criteria were the following: (1) Black or African American racial identity, (2) 16 years old or older, (3) assignment of male sex at birth, (4) residency in the Chicago Metropolitan Statistical Area (MSA), (5) having at least one sexual encounter with a cisgender man or transgender woman in the past year, (6) willingness to wear a GPS device, and (7) intention to remain in the Chicago MSA during the course of the study (~2 years). Additional information about the parent N2 study can be found elsewhere [32]. Given that the data in this study were collected 6 years ago, this analysis provides an opportunity to enhance their relevance by integrating them into a fuzzy-like sociocentric network. This paper was prepared using the Strengthening the Report of Observational Studies in Epidemiology (STROBE) guidelines [33].

Data Collection

A study team member trained in quantitative and network data collection obtained written informed consent from participants and then collected data using interviewer-administered computer-assisted assessments in a private room at the study site. We collected the following data from each participant: sociodemographics, confidant and sexual partner information, mobile phone contact lists, and Facebook friends lists.

Sociodemographic data about the participant and each confidant and sexual partner included first name, last name, gender identity (cisgender man, cisgender woman, transgender or another gender expansive identity), race (Black or not Black), Latine ethnicity (Latine, not Latine), and age (in years). Sociodemographic data were collected using a network inventory using Qualtrics. We downloaded participants' phone contact information and Facebook network information. We attempted the download of phone contacts and Facebook friends one time, and if the download was not successful then we considered that participants' data to be missing at random.

XSL•FO
RenderX

Analysis

Overview

We developed and tested a fuzzy matching algorithm (ie, deterministic linking algorithm) with a set of identifiers that provided a match result across confidant, sexual, phone contact, and Facebook networks. We used R (version 12.1.402; R Foundation for Statistical Computing [34]) and Microsoft Excel (Microsoft Corp) for data analysis and matching. The boundaries of the final complete sociocentric-like fuzzy network were only those participants enrolled in the N2 cohort study and did not include those individuals that participants may have named as being in their social networks. We describe the following characteristics for each of the networks: connected components (number of unique networks within the larger network), network density (proportion of nodes that are connected relative to all potential connections), degree (average number of nodes that a node reports knowing added to average number of nodes who also know the node), diameter (maximum distance between the furthest pair of nodes), radius (maximum shortest distance between pairs of nodes), betweenness centrality (how often a node acts as a bridge along the shortest path between two nodes), and characteristic path length (average shortest distance for all possible pairs of network nodes). We conducted an ANOVA to assess the differences in mean betweenness centrality based on HIV care cascade outcome (ie, not using PrEP, using PrEP, not virally suppressed, virally suppressed).

Matching Confidants With Sexual Partners

To identify confidants who were also sexual partners, we collected these data separately. Thus, participants named up to 5 people they considered to be their closest social network members using the prompt, "Who is the person you most talk to about things that are important to you?" Participants also named up to 5 people with whom they had sex with in the past 6 months. Participants indicated whether they had sex with a confidant and if a sexual partner was also a confidant. Egocentric network members included those who were confidants or sexual partners.

Matching Social Network Members With Participants Enrolled in N2

We adapted a previously used deterministic linkage algorithm [16]. We first phoneticized first and last names to standardize names in the instance of incorrect or multiple spellings of names. Thus, each name was given a unique phonetic code. We then used a *for loop* in R to match each social network member with each participant (dyad) based on the following selection criteria: phoneticized first name (both individuals have the same phoneticized first name), phoneticized last name (both individuals have the same phoneticized last name), gender identity (both individuals are homophilous on gender), race (both individuals within the pair being matched are homophilous on race), Latine ethnicity (both individuals within the pair being matched are homophilous on Latine ethnicity), and age (both individuals within the pair being matched are within 5 years of each other, $age_i \pm 5$ years to age_j).

Following a similar algorithm as in previous work, we created an autoscoring algorithm ranking the certainty of each dyad being compared [35]. We then manually reviewed each potential match of dyads to identify whether social network members named in the inventory could be matched to participants enrolled in the study. We decided that if dyads fit the following criteria, they were the same person: both members of the potential dyadic match (1) had the same phoneticized first name, phoneticized last name, or phoneticized first and last name; (2) were either both cisgender women, were both cisgender men, were both transgender-identified, or one member of the potential match dyad was transgender-identified and another member of the potential match dyad was a cisgender man; (3) were both Black or were both not Black; (4) and were within 5 years of age. We decided to match individuals who were indicated as being transgender with those who were indicated as being a cisgender man because we recognized that not all individuals may be "out" with their social networks and because gender can be fluid, and great strides in gender identity can be made in a short time. In addition, we considered a match on Latine identity to contribute toward the manual match algorithm; however, if both members of a dyad were not Latine, this did not negate the potential match.

Identifying Ties Between Participants Using Peer Referral Ties

We identified ties between participants using the peer referral ties during enrollment. Thus, we could see which participants were referred by an existing participant in the study.

Matching Participants Based on Mobile Phone Contact Data

We identified matches between participants' phone contact data and other participants in the study based on the phone number provided at baseline data collection.

Matching Participants Based on Facebook Friends Lists

Facebook data were cleaned using Python (including the packages *beautifulsoup4.9.3*, *soupsieve2.2.1*, and *pandas1.2.4*) to parse through the files each participant's information was stored in. The only available information for a Facebook friend was their display name—there was no additional identifying information for friends lists such as an identifier or URL. For this reason, we assumed any common Facebook friend was one held in common between participants, rather than assuming that 2 different people with the same name had friendships among the study population.

Network Peer Effects on HIV Engagement in Care Analysis

We investigated if there was an association between individuals' HIV-related care and their network members' engagement in status-neutral HIV-related care using a multistep process. We first identified the participant's HIV serostatus (ie, 0=not living with HIV; 1=living with HIV) and their network members' mean serostatus (ie, proportion of members living with HIV). We then identified the participant's engagement in status-neutral HIV-related care (ie, 0=not engaged in care or not using ART/PrEP; 1=engaged in care or virally suppressed or using PrEP) and their network members' mean engagement. For both of these outcomes (ie, HIV serostatus and engagement in

XSL•FO
RenderX

status-neutral HIV-related care), we conducted the following steps:

- Step 1: Spearman correlation to examine binary correlations between an individual's outcome and the mean of their network's outcome.
- Step 2: Unadjusted linear regression to examine associations between an individual's outcome and the mean of their network's outcome.
- Step 3: Quadratic Assignment Procedure (QAP) correlation to examine the relationship between an individual's outcome and the mean of their network's outcome. This allowed for the identification of the product-moment correlation between the adjacency matrices of graphs for an individual's outcome and the mean of their network's outcome [36]. We used the R environment packages *igraph* [37] and *statnet* [38] to use the *gcor*() function for analysis.
- Step 4a: Linear Network Autocorrelation Model (LNAM) to examine the relationship between an individual's outcome and the mean of their network's outcome [39-41]. This model "assumes that social influence occurs when individuals are exposed to a behavior by their network contacts." [42] We used the R environment package *sna* and the *lnam*() function [43].
- Step 4b: LNAM to identify the social cross-sectional peer effects of a participant's engagement in status-neutral HIV-related care within each of the networks, adjusting for their social network members' HIV serostatus. We used the R environment package *sna* and the *lnam*() function [43].

The network autocorrelation model assumes continuous dependent variables; however, our outcome variables are binary. Because there is no established statistical model that can account for the network autocorrelation for the Bernoulli distribution of the error term, we used the LNAM. This model assumes a normal distribution of error terms with a null value and variance in the identity matrix. Due to this limitation, we added the supplemental analysis of conducting logistic regression as we could not adjust for the limitation of not accounting for potential network autocorrelation. Within the network autocorrelation model [39,44-46], we specified the network effects model, that is defined in the following formula:

$$y=\rho Wy+X\beta+\varepsilon,\varepsilon\sim N(0,\sigma 2I)$$

in which y is a vector for values of our outcome variable (ie, engagement in status-neutral HIV-related care), and X is a matrix of values for the N actors. The β term is a vector of regression coefficients, ρ is a scalar estimate of the autocorrelation parameter, and W is a matrix of weights, which is the allocation of peer effects among network members (eg, engagement in HIV-related care). We also fit the multiple network autocorrelation model [40],

We also fit the multiple network autocorrelation model [40]

which was adjusted to include HIV serostatus (Xβ), as our unadjusted model does not include HIV serostatus Xβ.

Network Visualization

We used Cytoscape to visualize participants' peer referral ties; sexual, social, phone contact, and Facebook networks; and the final fuzzy match networks [47].

Ethical Considerations

The N2 cohort study protocol for the Chicago site was approved by the Biological Sciences Division/University of Chicago Medical Center Institutional Review Board (IRB16-1419) and Columbia University (AAAS7654). Written informed consent was obtained from the study participants. Participants were compensated US $150 for their time and US $20 for each person they referred to the study. A Certificate of Confidentiality from the National Institute of Health was obtained to protect study participants' privacy.

Results

Overview

A total of 412 Black SGM enrolled in the N2 cohort study. Of participants, 387 provided confidant and sexual network data, 273 provided mobile phone contact data, and 144 provided Facebook data. Of these networks, the Facebook network had the highest number of undirected ties (n=1383; undirected means that if one person in the network reported knowing the other, then they were considered to know each other, regardless of the direction of the tie), the fewest connected components, the highest density (13.4% of all potential ties), highest degree (n=19.21), and lowest network diameter (n=5), network radius (n=3), and average characteristic path length (length=2.28). The fuzzy match network consisted of 400 unique nodes and 12 isolates. There was a total of 2054 undirected edges in one centralized component after removing isolates. Its density was 0.026, with an average degree of 10.27. The network diameter was 9, the radius was 5, and the characteristic path length was 3.49. An overview of each of the networks, including nodes, undirected ties, and other network characteristics, can be found in Table 1.

XSL•FO
RenderX

Table . Overview of social network characteristics among 412 Chicago-based Black sexually minoritized and gender expansive people in the Neighborhoods and Networks (N2) cohort study (2018 - 2019), with isolates removed, by network type.

	Confidant and sexual network	Peer referral network	Phone contact only	Facebook only	Entire sociocentric network (fuzzy)
Participants who provided data, n	387	412	273	144	412
Egos in network, n	379	344	245	144	400
Undirected ties within network, n	445	362	362	1383	2054
Connected components, n	10	19	12	1	1
Network density, mean	0.007	0.012	0.015	0.134	0.026
Network degree, mean	2.42	1.99	3.18	19.21	10.27
Network diameter, n	19	18	12	5	9
Network radius, n	11	9	7	3	5
Betweenness centrality, mean	0.08	0.03	0.03	0.009	0.006
Characteristic path length, mean	7.77	7.64	4.93	2.28	3.49
Number of HIV-negative people (%)	203 (52.5)	189 (45.8)	163 (59.7)	79 (54.9)	224 (54.4)
Number of people engaged in care (%)	200 (51.7)	194 (47.1)	144 52.7)	62 43.1)	179 (43.4)

Matching Social Confidants With Sexual Partners

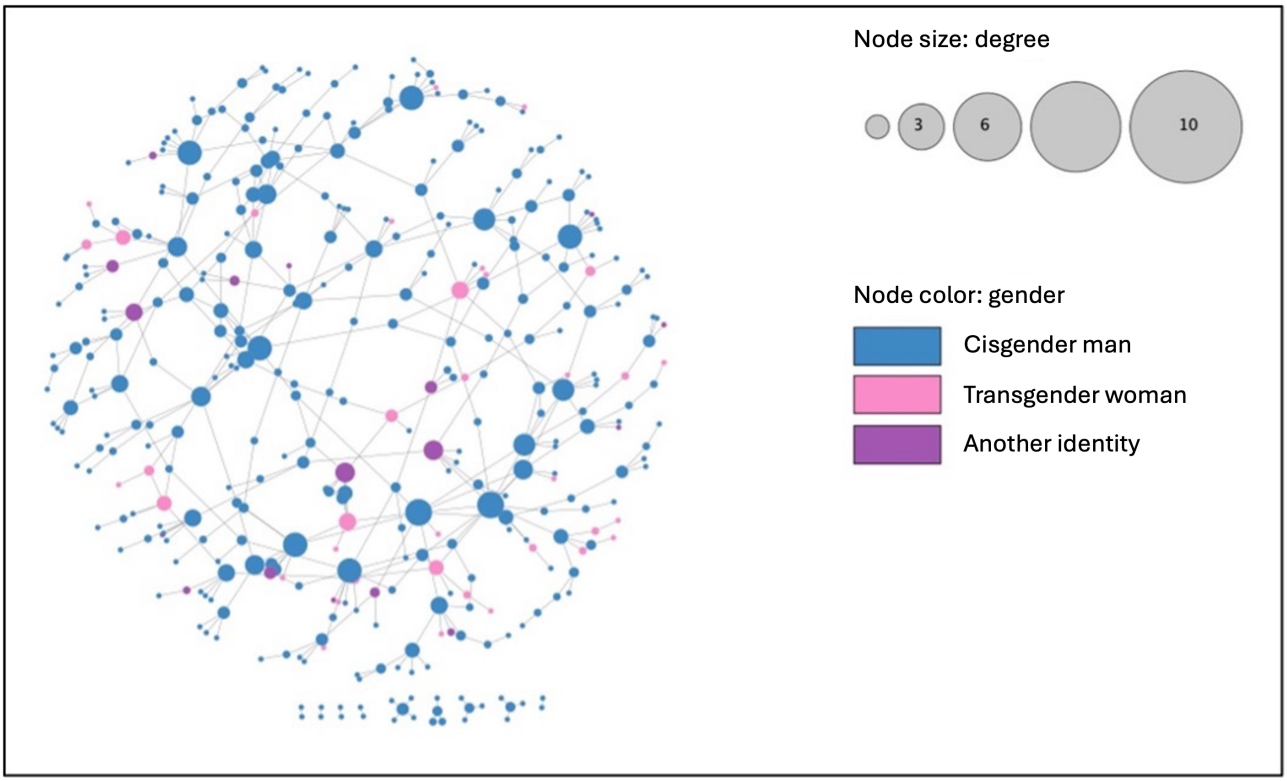
Of 412 participants, 387 completed the social network inventory. Participants named up to 5 social network alters (n=1036; mean 2.73, SD 1.32) and 5 sexual partner alters (n=1017; mean 2.57, SD 1.58). We identified an overlap of 136 confidants who were also sexual partners, for a total of 1917 unique alters.

Matching Social Network Members With Participants Enrolled in N2

There were a total of 1917 unique alters. When considering that 412 participants were enrolled in the N2 study, this means that there was a potential of 169,332 directional matches our analyses could identify (ie, A knowing B is treated as separate from B knowing A; between 412 participants, there are 84,666 potentially unique ties as 412 × 411 = 169,332). We found that there were 8020 potential "fuzzy" matches. Upon manual validation, we found that there were 445 undirected ties between the 379 nodes in this network. There were 10 connected components, and the network density was 0.007, indicating it was a sparse network. Average degree was 2.42 and betweenness centrality was 0.028. The network diameter was 19 steps, the radius was 11 steps, and the characteristic path length was 7.77 steps. The egocentric network can be observed in Figure 1. There was no statistical association between betweenness centrality and the HIV care cascade (F_3=0.242; P=.87).

XSL•FO
RenderX

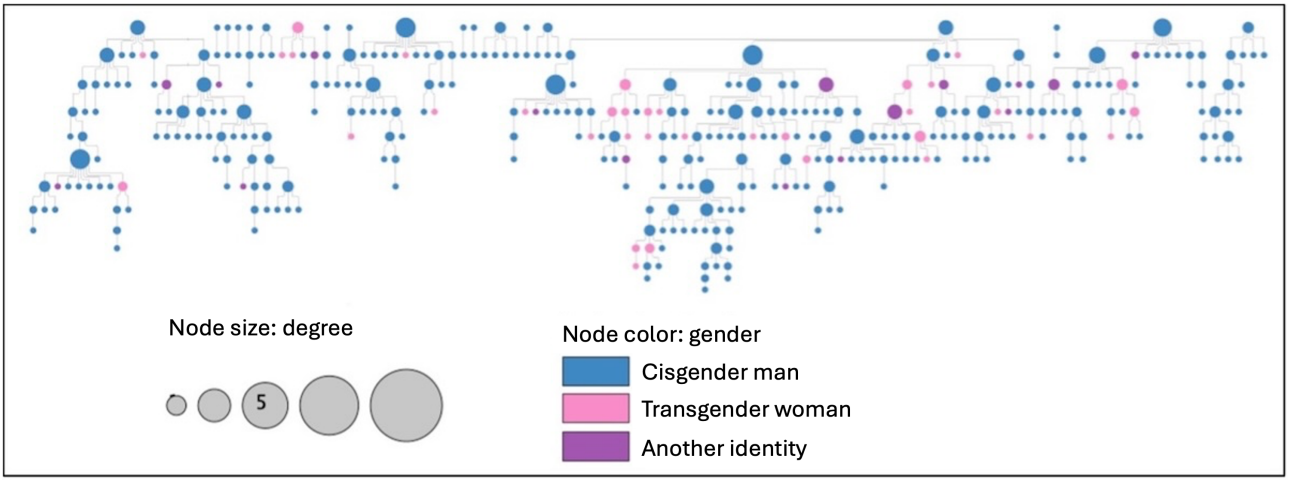
Figure 1. Egocentric confidant and sexual partner network of Chicago Black sexually minoritized and gender expansive people from the Neighborhoods and Networks (N2) cohort study (n=379), edge-weighted spring embedded layout by edge-betweenness.



Identifying Ties Between Participants Using the Peer Referral Sampling Chain

We identified 363 unique ties between 344 participants from the peer referral ties. There were 19 connected components (ie, 19 seeds), and the network density was 0.012. The average degree was 1.99 and betweenness centrality was 0.080. The network diameter was 18 steps, the radius was 9 steps, and the characteristic path length was 7.64 steps. The peer referral tie network is visualized in Figure 2 using a *yFiles* hierarchic layout to display the chain [48]. There was no statistical association between betweenness centrality and the HIV care cascade (F_3=2.00; P=.11).

Figure 2. Peer referral tie network of Chicago Black sexually minoritized and gender expansive people from the Neighborhoods and Networks (N2) cohort study (n=344), yFiles hierarchic layout.

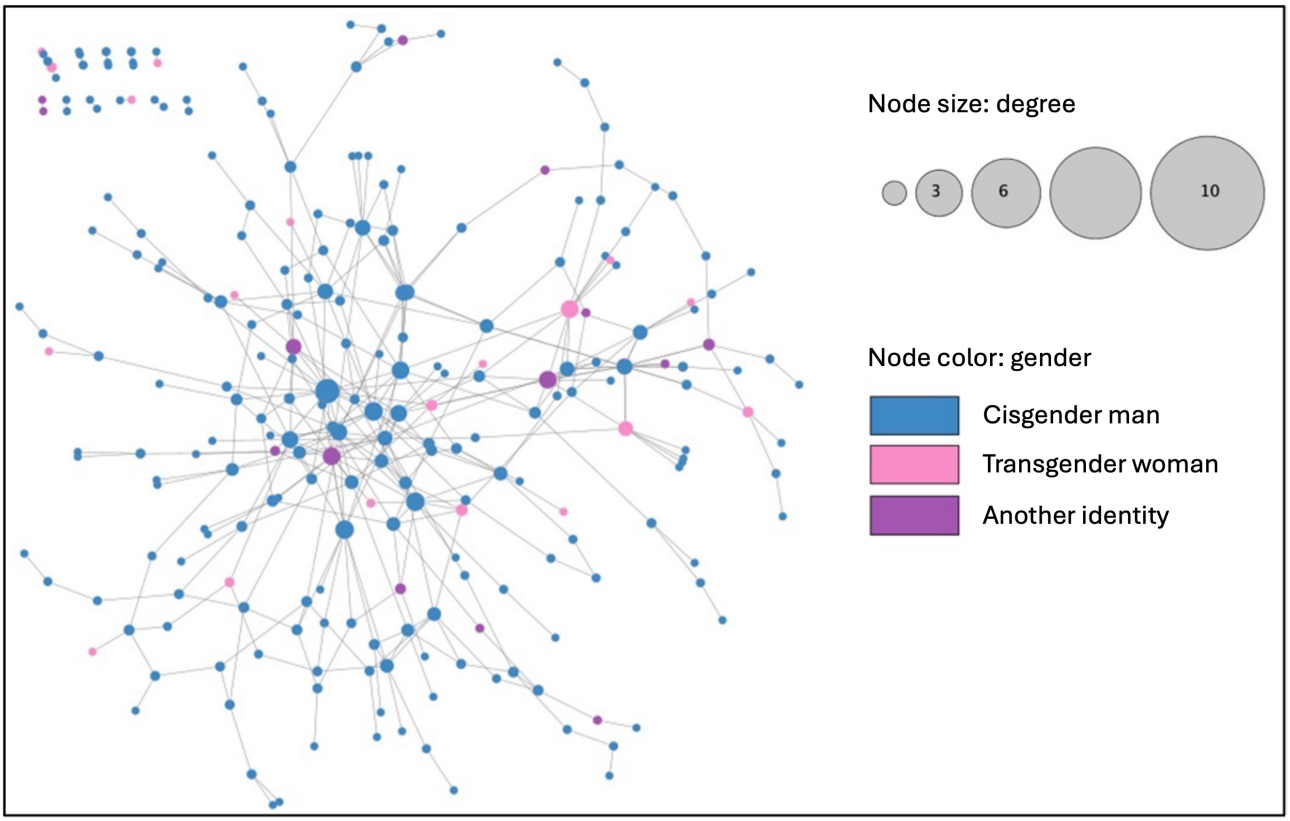


Matching Participants Based on Mobile Phone Contact Data

Of the participants, 294 provided consent to share their mobile phone contact data. There were 95,831 phone number entries, of which 51,671 were valid phone numbers (ie, in a 10-digit format). We found that there were 416 unique ties among 270 participants from the mobile phone contact data. Figure 3 displays a visualization of the mobile phone contact network. The average degree was 3.18 and betweenness centrality was 0.03. There was no statistical association between betweenness centrality and the HIV care cascade (F_3=1.03; P=.39).

XSL•FO
RenderX

Figure 3. Phone contact network of Chicago Black sexually minoritized and gender expansive people from the Neighborhoods and Networks (N2) cohort study (n=245), edge-weighted spring embedded layout by edge-betweenness.

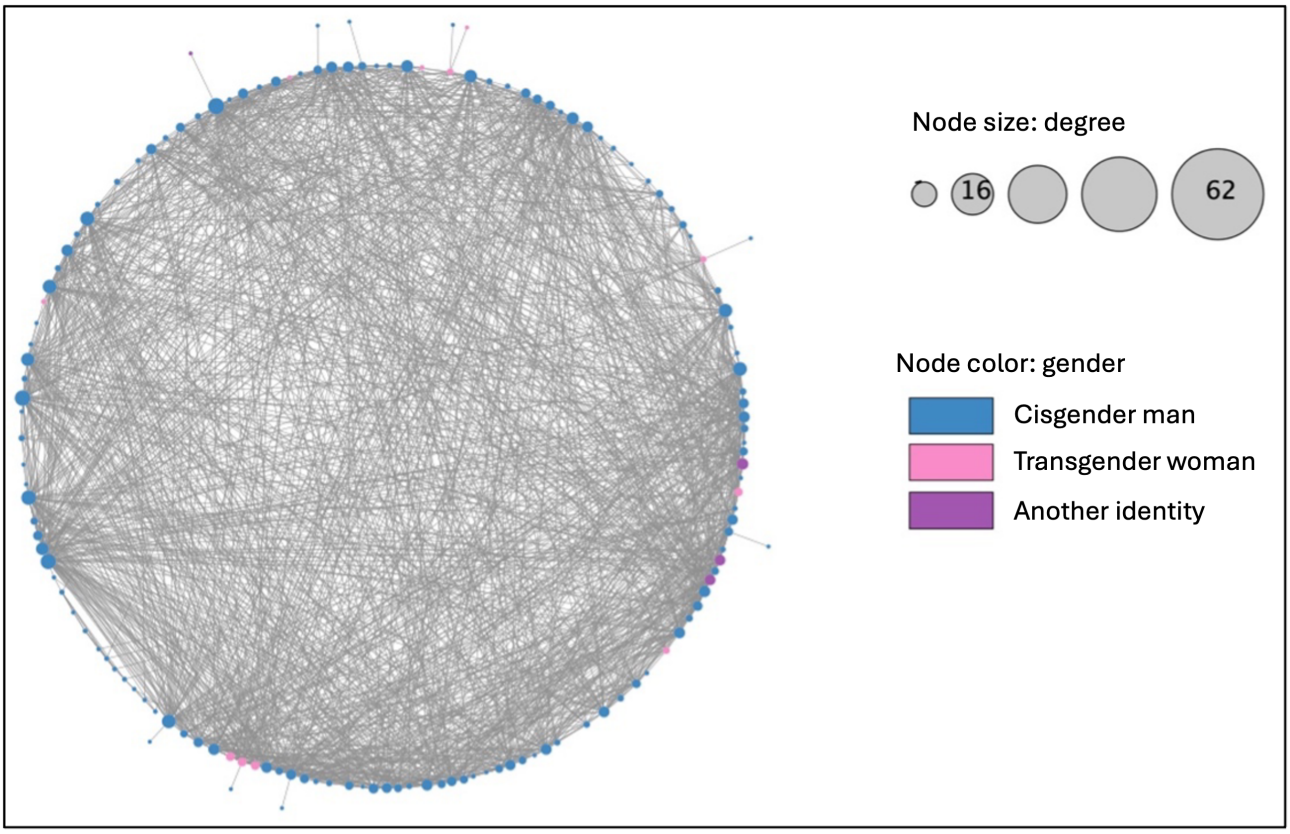


Matching Participants Based on Facebook Data

We downloaded the Facebook friends list of 144 participants. Due to refusal and interruptions in the download process, we were not able to download data for 268 participants. There were a total of 18,255 Facebook friendship ties, posts, and comments. Among these, we identified 1383 undirected ties between N2 participants. There was one connected component with a density of 0.134. The average degree was 19.21 and betweenness centrality was 0.009. The network diameter was 5 steps, the radius was 3 steps, and the characteristics path length was 2.28 steps. The Facebook-only network is visualized in Figure 4. There was no statistical association between betweenness centrality and the HIV care cascade (F_3=1.194; P=.31).

XSL•FO
RenderX

Figure 4. Facebook network of Chicago Black sexually minoritized and gender expansive people from the Neighborhoods and Networks (N2) cohort study (n=144), circular layout.

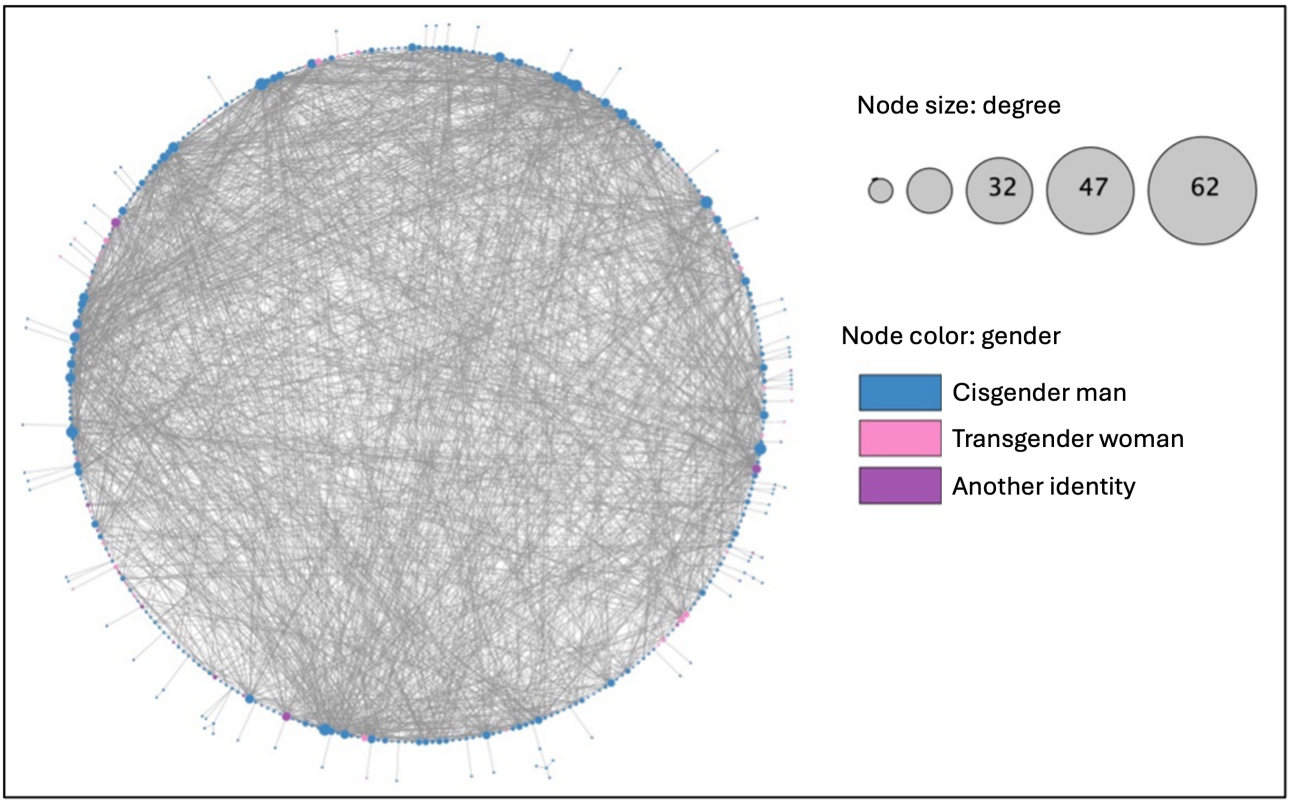


Fuzzy Match Network Description

Once all networks were combined, we identified 400 participants who were included in the fuzzy-like sociocentric network. There were 2054 unique ties among participants. The density of our fuzzy match network was 0.025, and there was one connected component. The average degree was 10.21 and betweenness centrality was 0.006. The network diameter was 9 steps, the radius was 5 steps, and the characteristic path length was 3.49 steps. The fuzzy match sociocentric network is visualized in Figure 5. There was a statistically significant difference between betweenness centrality and the HIV care cascade (F_3=5.837; P<.001). We conducted a post hoc comparison using the Tukey Honestly Significant Difference test and found that the not virally suppressed group had higher betweenness centrality measures relative to the group not using PrEP (Diff=0.006, 95% CI 0.002 to 0.009; P<.05); the PrEP-using group had lower betweenness centrality measures relative to the not virally suppressed group (Diff=−0.005, 95% CI −0.009 to −0.007; P<.05); and the virally suppressed group had lower betweenness centrality relative to the not virally suppressed group (Diff=−0.005, 95% CI −0.009 to –0.001; P<.01). The other 3 conditions were not statistically significant (P>.05).

XSL•FO
RenderX

Figure 5. Sociocentric fuzzy match network of Chicago Black sexually minoritized and gender expansive people from the Neighborhoods and Networks (N2) cohort study (n=400), circular layout.



Findings of the Correlation, Quadratic Assignment Procedure Correlation, Regression Models, and Network Autocorrelation Models

We found a low correlation and QAP correlation between HIV status and care engagement between participants and their network members (Table 2). Only the regression model for engagement in care was found to be statistically significant; participants had a 2.07 higher odds ratio (95% CI 1.09 - 3.97; *P*=.03) of engagement in care with higher network-level engagement in care. None of the unadjusted linear network autocorrelation models were significant when examining HIV status and care engagement for each of the different networks (Table 2, models a-e). For the linear network autocorrelation models adjusted for HIV status (Table 2, models f-j), HIV status had a significant negative association with care engagement for all network models: confidant and sexual (β=−0.268, SE 0.048; *P*<.001), peer referral (β=−0.281, SE 0.049; *P*<.001), phone contact–only (β=−0.231, SE 0.056; *P*<.001), Facebook-only (β=−0.252, SE 0.079; *P*=.001), and the combined sociocentric network (β=−0.256, SE 0.047; *P*<.001). Results from the linear network autocorrelation model for the combined sociocentric network adjusted for HIV status showed that care engagement had network clustering. Thus, participants who were proximate to network members who were engaged in care were significantly more likely to be engaged in care (ρ=0.128, SE 0.064; *P*=.045).

XSL•FO
RenderX

Table . Results of the unadjusted and adjusted network autocorrelation assessing the peer effects of status-neutral HIV outcomes among 412 Chicago-based Black sexually minoritized and gender expansive people in the Neighborhoods and Networks (N2) cohort study (2018 - 2019), by network type.

	Confidant and sexual network	Peer referral network	Phone contact only	Facebook only	Entire sociocentric network (fuzzy)
Spearman correlation between participant and average score of their network members					
HIV serostatus	0.054	0.018	0.021	−0.119	0.004
Engagement in care	0.014	0.043	0.068	0.149	0.114
Logistic regression analyses between participant and average score of their network members					
HIV serostatus	OR 1.30, 95% CI 0.68 - 2.46; *P*=.43	OR 1.08, 95% CI 0.70 - 1.64; *P*=.74	OR 1.02, 95% CI 0.52 - 2.38; *P*=.77	OR 0.38, 95% CI 0.09 - 1.46; *P*=.17	OR 1.02, 95% CI 0.54 - 1.93; *P*=.94
Engagement in care	OR 1.07, 95% CI 0.55 - 2.10; *P*=.84	OR 1.20, 95% CI 0.77 - 1.88; *P*=.44	OR 1.41, 95% CI 0.68 - 2.92; *P*=.36	OR 3.51, 95% CI 0.86 - 15.53; *P*=.09	OR 2.07, 95% CI 1.09 - 3.97; *P*=.03[a]
Quadratic Assignment Procedure correlation between participant and difference in the average score of their network members					
HIV serostatus	−0.003	−0.003	−0.004	−0.007	−0.004
Engagement in care	−0.006	−0.000	−0.000	0.022	0.011
Unadjusted linear network autocorrelation models					
Individual HIV serostatus	Model a: ρ=0.077, SE 0.061; *P*=.21	Model b: ρ=0.048, SE 0.054; *P*=.38	Model c: ρ=0.065, SE 0.070; *P*=.36	Model d: ρ=−0.174, SE 0.139; *P*=.21	Model e: ρ=−0.003, SE 0.067; *P*=.97
Engagement in care	Model a: ρ=0.016, SE 0.056; *P*=.77	Model b: ρ=0.029, SE 0.052; *P*=.58	Model c: ρ=0.051, SE 0.067; *P*=.45	Model d: ρ=0.210, SE 0.136; *P*=.12	Model e: ρ=0.114, SE 0.066; *P*=.08
Linear network autocorrelation model, adjusted for HIV serostatus					
HIV serostatus (covariate)	Model f: β=−.268, SE 0.048; *P*<.001[a]	Model g: β=−.281, SE 0.049; *P*<.001[a]	Model h: β=−.231, SE 0.056; *P*<.001[a]	Model i: β=−.252, SE 0.079; *P*=.001[a]	Model j: β=−.256, SE .047; *P*<.001[a]
Engagement in care (ρ)	Model f: ρ=0.018, SE 0.054; *P*=.74	Model g: ρ=0.048, SE 0.052; *P*=.35	Model h: ρ=0.043 SE 0.065; *P*=.51	Model i: ρ=0.253, SE 0.13; *P*=.055	Model j: ρ=0.128, SE 0.064; *P*=.045[a]

[a]*P*<.05.

Discussion

Principal Findings

Sociocentric network study designs are resource heavy to conduct and primary data collection is difficult. To address this gap, we aimed to develop and test a matching algorithm consisting of multiple types of nonsociocentric networks. We generated a fuzzy matched sociocentric-like network based on the egocentric social networks, egocentric sexual networks, peer referral ties, egocentric mobile phone contact data, and Facebook network present within the N2 cohort study of Black SGM. In addition to analyzing a fuzzy match algorithm, we found that sociocentric networks and digital networks have high utility in unveiling hidden connections. More specifically, we found that in the "fuzzy" network, bridging (ie, betweenness centrality) differed based on the HIV care cascade and engagement in HIV-related care had an autocorrelative effect; these findings were not identified in the partial networks. To our knowledge, this is one of the first studies to harmonize multiple networks to construct a larger fuzzy network, to examine the peer effects of network-level, status-neutral HIV care at the individual level, and to prioritize Black SGM social networks. Our methods and findings are novel and contribute toward the fields of social and digital epidemiology by introducing a fuzzy matching algorithm that can identify hidden ties among participants within different types of networks to better understand the larger sociocentric network.

The initial peer referral chain identified 362 existing ties within our network; however, by including alternative types of networks such as social networks and sexual partner networks, phone contact data, and Facebook friends lists, we were able to identify a total of 2054 undirected ties among 400 participants. The egocentric network quintupled the network degree and doubled density while introducing an additional >50 participants into the network. Whereas the peer referral network had 19 connected components, the final and more complete network brought together these networks into one connected component. In addition, the network diameter decreased from 18 to 9. This means that, while the most distant pairs of nodes required 18 steps at maximum to reach each other, this number halved to 9 steps in the fuzzy match network. In addition, the network radius decreased from 9 to 5, meaning it did not take more than 5 steps for anyone in the network to reach someone else. The characteristic path length, or what Milgram (1967) [49] referred to as "degrees of separation," also halved from 7.64 to 3.49 steps from the peer referral ties to the fuzzy network. Although the Chicago MSA is indeed a small world, and our population may be too homophilous for a pair of individuals to be selected at random, our findings support the idea of the "small world problem" or the probability that 2 random people selected from a larger population will know each other. Our findings support

XSL•FO
RenderX

the need for study designs using peer referral to include additional estimators to account for the interdependence of participants, as these networks may not be as apparent. Further, our study supports the utility of using a fuzzy match algorithm to construct a larger sociocentric network based on other networks, which scientists may collect at a lower burden. Our findings must be interpreted with caution as the network autocorrelation parameter was small. In real-world settings, this may suggest that there is weak social influence within the fuzzy-like network, networks are diverse or mixed and have low homophily, or individual variability may be more important than network influence.

The digital data networks identified the most ties and had superior network density, degree, diameter, radius, and characteristic path length relative to the egocentric confidant and sexual networks and the peer referral network. Interestingly, the characteristic path length of the Facebook network, 2.28 steps, is lower than reported in another study, which found pairs of random people on Facebook to have a characteristic path length of 4 steps [50]. Although our study included mobile phone contact networks and Facebook networks, future studies may have better success in identifying networks using other types of digital networks based on their priority population. For example, studies prioritizing people aged 18 - 29 years may have success using Instagram as 78% of people in the 18 - 29 year age group use Instagram, which is higher than other age groups (59% of people aged 30 - 49 years, 35% of people aged 50 - 64 years, and 15% of people aged ≥65 years) [14]. If studies hope to reach Black populations, TikTok or Instagram may be the best social networking sites as 49% of US Black adults report using TikTok (relative to 28% of White adults, 49% of Hispanic adults, and 29% of Asian adults) and 46% report using Instagram (relative to 43% of White adults, 58% of Hispanic adults, and 57% of Asian adults). In addition, after the COVID-19 pandemic stay-at-home orders in the United States, social media platforms became an online "home" for many SGM [51,52]; however, further study on social media use by young (ie, 16 - 34 years old) Black SGM is warranted.

Our findings suggest that, while we did not identify differences in betweenness centrality based on a participant's HIV care cascade in the partial networks, we did so in the fuzzy-like network. More specifically, we found that participants who were virally suppressed had lower betweenness centrality measures relative to individuals in the other HIV care cascades (ie, people who did not use PrEP, people who used PrEP, and people who were not virally suppressed). This means that, relative to individuals who are HIV-negative or not virally suppressed, participants who are virally suppressed appear to be a bridge in the network less often. This could be due to their lower likelihood to engage in behaviors with enhanced vulnerability to HIV such as condomless sex [53], which could also reduce their network connections. People who are virally suppressed may also be embedded in clinical networks instead of other social, sexual, or other community networks, as they are engaged in their HIV-related care. People who are virally suppressed may not only have smaller networks but also have networks that are more closely knit, which could reduce their bridging. Alternatively, people living with HIV may also experience stigma from HIV-negative people, which could reduce their network bridging potential. Additionally, research can explore the relationship between bridging or betweenness centrality and HIV care cascade outcomes.

Our findings must be considered within their context. There is a dearth of studies comparing influence in digital and in-person networks, and it is difficult to ascertain whether endogenous influence and social norms can be transmitted through digital or in-person interactions. Additionally, the weight social media influencers and content creators may have on social norms surrounding health and how social media algorithms "push" specific content through are largely unexplored. Further, we were unable to ascertain which relationships may be more important to intervene on. Finally, within social epidemiology and social network studies, the contextual factors of the sample as well as the geographic setting are important to consider as findings may (or may not) apply to other marginalized populations or geographic settings. For example, Black SGM in Chicago, Illinois, experience intersectional discrimination [20-23] and health disinvestment within their communities [54], which negatively impact health outcomes including HIV-related outcomes. Within this population, social norms may influence access to resources, support, and other opportunities [55-58]. These findings may translate to other Black SGM born in the United States, but may not translate to Black SGM born elsewhere who migrated to the United States or Black SGM living in cities with a more robust health infrastructure. Future studies could explore these topics.

Sociocentric data, the "gold standard" of social network data, are difficult to collect due to resource demands and complex study designs. However, constructing sociocentric networks from digital data may surmount challenges such as limitations in data collection as asking participants to list everyone that they know may be subject to measurement issues like recall bias. Although our inclusion of digital data expanded the breadth of our sociocentric network, the depth, or the influence of individuals in digital networks, is unknown. In addition, although we found that the inclusion of digital data resulted in a model that found statistical significance in HIV-related behaviors, it is unclear if this model is more precise or accurately describes the data. Thus, to see if the quality or strength of digital networks may be similar to that of in-person networks, future studies identifying the quality of these ties are needed. Another way to do so could be to analyze a behavior that is more susceptible to network structure, such as direct interactions among participants (eg, likes, comments, reactions). Although our findings support the collection of additional network data, from an ethical perspective, this may increase the likelihood that a participant may be identified as part of a network that one may not want to be identified with. These concerns must be considered in study designs.

Limitations

Network analyses are often not perceived to be generalizable because they are representative of one network at a given time point; however, we argue that the purpose of this analysis is not to generalize our findings to the larger population but rather to explore novel fuzzy matching methods to enrich social network

XSL•FO
RenderX

research by identifying opportunities to harmonize multiple types of data to triangulate a more robust and complete network. Differentiating selection from peer effects in this analysis may manifest differently based on the mechanisms at play and the networks being considered—for this, Simulation Investigation for Empirical Network Analysis (SIENA) models may allow for the distinguishing between these 2 mechanisms. In addition, participants were asked to provide their perceptions of the sociodemographic information of their network members. Race and ethnicity are complicated social identities that can be incorrectly perceived by others. This could introduce misclassification as participants may incorrectly estimate their network member's sociodemographic information and introduce bias into our findings by creating incorrect linkages between participants and creating errors in the network structure [9]. This could then trickle down and impact our tests of differences and associations and, if nonrandom, create confounding effects. Because the majority of participants felt close or very close (>80%) to their social members, we believe that misclassification bias may be minimized. A future study could also assess the accuracy of this fuzzy matching analysis. This type of data validation could include a brief survey or interviews with a random sample of the network to authenticate relationships. Another limitation is the number of participants for whom we were able to download complete Facebook data (n=144, 35%). If we were able to include Facebook data from additional participants, we could have potentially constructed a larger fuzzy network. However, because we were able to identify participants who did not provide their Facebook data from the Facebook friends lists of current participants, we believe this bias may be minimized. Future studies can opt to collect less Facebook data, such as only collecting friendship networks, instead of collecting data on group membership, comments, and other Facebook interactions. Finally, we used a cross-sectional linear network autocorrelation, a model that assumes continuous dependent variables, to analyze binary variables. The use of this model assumes that associations are linear, which may not hold for binary variables that constrain probabilities within the (0,1) range. The use of the linear network autocorrelation model may introduce misestimation of effect sizes due to nonlinearity, predictions outside the valid probability range, and heteroskedasticity. To address this bias, we have added supplemental analyses of unadjusted logistic regressions to test the effects of the network exposure to HIV serostatus and engagement in care. To establish peer effects, longitudinal data and a robust set of covariates could be used instead. Our findings may not be capturing peer effects, but rather homophily in the select effect (ie, participants deliberately seek others similar in their HIV care to be friends with). As our study was a pilot approach, our findings on autocorrelative effects may be interpreted with caution as they are marginally significant correlations.

Although not a limitation, a notable feature of this analysis was the time intensity of the match between participants and the social alters and sexual partners listed by other participants. This match of respondents to social network members and sexual network members generated over 8000 potential matches. We manually validated each match to determine if it was a true match or a false-positive match. This was a time-consuming analysis that unveiled <1% of ties of the final network. A machine learning algorithm, if trained appropriately, may alternatively be used to identify matches.

Conclusions

Digital networks, which we operationalized as phone contact and Facebook networks, were the most productive in identifying ties and may be best to characterize social networks. Our experimental approach identified opportunities to characterize networks more thoroughly and accurately. Future studies can explore how to operationalize and examine the influence of ties across various networks. In addition, there are added benefits to including digital networks to egocentric networks as they can identify hidden ties. There are clearly differential impacts of digital data networks on health and HIV outcomes; future studies could more thoroughly assess these networks. Thus, fuzzy matching may be an effective means to identify hidden ties among participants within different types of egocentric and digital networks to better understand the larger sociocentric network, including among vulnerable populations.

Acknowledgments

DTD and JAS report funding from the National Institute on Mental Health (R01MH112406, principal investigators [PIs]: DTD and JAS) and the Centers for Disease Control and Prevention under the Minority HIV/AIDS Research Initiative (U01PS005122, PI: DTD). CHS and RM's efforts were supported by the National Institute on Drug Abuse (T32DA031099; PIs: Deborah Hasin and Silvia Martins). CHS was supported by the National Institute of Allergy and Infectious Diseases (T32AI114398; PI: Andrea Howard). YTC and JAS were supported by the National Institute on Drug Abuse (R03DA053161, PIs: YTC and JAS). RD's efforts were supported by the National Institute of Mental Health (T32MH019139, PI: Theo Sandfort). JRK's effort on this project was funded by NIH grants K01AA028199, R01DA054553, and R21DA053156. BS was supported by DA049665. RM was supported by a grant from the National Center for Injury Prevention and Control, Centers for Disease Control and Prevention to the Center for Injury Epidemiology and Prevention at Columbia University (R49CE003096, PI: Charles Branas). The University of Chicago authors were supported in part by the National Institute on Drug Abuse (U2CDA050098, PI: JAS). LY's efforts were supported by the National Institute of Child Health and Human Development (R00HD094648, PI: LY). Generative AI was not used to prepare this manuscript.

XSL•FO
RenderX

Data Availability

Data may be requested from the second and senior authors with reasonable request and explanation and may be provided under the condition of the institution's institutional review board approval. The code can be requested from the first author, and will be considered upon reasonable request and explanation and under the condition of the institution's institutional review board approval.

Conflicts of Interest

None declared.

References

1. Berkman LF, Kawachi I, Glymour MM. Social Epidemiology: Oxford University Press; 2014. [doi: 10.1093/med/9780195377903.001.0001]
2. Valente T. Social Networks and Health: Models, Methods, and Application: Oxford University Press; 2010. [doi: 10.1093/acprof:oso/9780195301014.001.0001]
3. Freeman L. The development of social network analysis. In: A Study in the Sociology of Science 2004, Vol. 1:159-167.
4. Valente TW. Network models and methods for studying the diffusion of innovations. In: Models and Methods in Social Network Analysis 2005, Vol. 28:98-116. [doi: 10.1017/CBO9780511811395.006]
5. Borgatti SP, Mehra A, Brass DJ, Labianca G. Network analysis in the social sciences. Science 2009 Feb 13;323(5916):892-895. [doi: 10.1126/science.1165821] [Medline: 19213908]
6. Luke DA, Harris JK. Network analysis in public health: history, methods, and applications. Annu Rev Public Health 2007;28:69-93. [doi: 10.1146/annurev.publhealth.28.021406.144132] [Medline: 17222078]
7. Ramos-Vidal I. Social network analysis in public health research: a systematic review. Rev Salud Publica (Bogota) 2019 Jan 1;21(1):109-114. [doi: 10.15446/rsap.V21n1.68530] [Medline: 33206917]
8. Perry BL, Pescosolido BA, Borgatti SP. Egocentric Network Analysis: Foundations, Methods, and Models: Cambridge University Press; 2018, Vol. 44.
9. Harris AS, Findley MG. Is ethnicity identifiable? Lessons from an experiment in South Africa. J Conflict Resolut 2014;58(1):4-33. [doi: 10.1177/0022002712459710]
10. Perry BL, Roth AR. On the boundary specification problem in network analysis: an update and extension to personal social networks. In: Personal Networks:431-443. [doi: 10.1017/9781108878296.031]
11. Park HA, Jung H, On J, Park SK, Kang H. Digital epidemiology: use of digital data collected for non-epidemiological purposes in epidemiological studies. Healthc Inform Res 2018 Oct;24(4):253-262. [doi: 10.4258/hir.2018.24.4.253] [Medline: 30443413]
12. Mobile phone ownership over time. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile [accessed 2024-02-09]
13. Poushter J, Gubbala S, Austin S. 8 charts on technology use around the world. Pew Research Center. URL: https://www.pewresearch.org/short-reads/2024/02/05/8-charts-on-technology-use-around-the-world [accessed 2024-02-09]
14. Americans' social media use. Pew Research Center. URL: https://www.pewresearch.org/internet/2024/01/31/americans-social-media-use [accessed 2024-09-24]
15. Laumann EO, Marsden PV, Prensky D. The boundary specification problem in network analysis. In: Research Methods in Social Network Analysis 1989, Vol. 61.
16. Zhu Y, Matsuyama Y, Ohashi Y, Setoguchi S. When to conduct probabilistic linkage vs. deterministic linkage? A simulation study. J Biomed Inform 2015 Aug;56:80-86. [doi: 10.1016/j.jbi.2015.05.012] [Medline: 26004791]
17. Pagkas-Bather J, Motley DN, Schneider JA. Health equity for Black sexually minoritised men through status-neutral HIV care. Lancet HIV 2024 Apr;11(4):e268-e272. [doi: 10.1016/S2352-3018(24)00032-8] [Medline: 38430923]
18. Hess KL, Hu X, Lansky A, Mermin J, Hall HI. Lifetime risk of a diagnosis of HIV infection in the United States. Ann Epidemiol 2017 Apr;27(4):238-243. [doi: 10.1016/j.annepidem.2017.02.003] [Medline: 28325538]
19. Singh S, Hu X, Hess KL, Iqbal K. Estimating lifetime risk of a diagnosis of HIV infection among MSM: United States, 2017-2021. Presented at: CROI; Mar 2-6, 2024; Denver, Colorado URL: https://www.natap.org/2024/CROI/croi_243.htm [accessed 2025-04-28]
20. English D, Carter JA, Boone CA, et al. Intersecting structural oppression and Black sexual minority men's health. Am J Prev Med 2021 Jun;60(6):781-791. [doi: 10.1016/j.amepre.2020.12.022] [Medline: 33840546]
21. Duncan DT, Callander D, Bowleg L, et al. Intersectional analysis of life stress, incarceration and sexual health risk practices among cisgender Black gay, bisexual and other men who have sex with men in the Deep South of the US: the MARI Study. Sex Health 2020 Feb;17(1):38-44. [doi: 10.1071/SH19062] [Medline: 31972125]
22. Crenshaw KW. On Intersectionality: Essential Writings: The New Press; 2017.
23. Bowleg L. "Once you've blended the cake, you can't take the parts back to the main ingredients": Black gay and bisexual men's descriptions and experiences of intersectionality. Sex Roles 2013 Jun;68(11-12):754-767. [doi: 10.1007/s11199-012-0152-4]

XSL•FO
RenderX

24. Lauby JL, Marks G, Bingham T, et al. Having supportive social relationships is associated with reduced risk of unrecognized HIV infection among black and Latino men who have sex with men. AIDS Behav 2012 Apr;16(3):508-515. [doi: 10.1007/s10461-011-0002-3] [Medline: 21805191]
25. Bouris A, Jaffe K, Eavou R, et al. Project nGage: results of a randomized controlled trial of a dyadic network support intervention to retain young Black men who have sex with men in HIV care. AIDS Behav 2017 Dec;21(12):3618-3629. [doi: 10.1007/s10461-017-1954-8] [Medline: 29079949]
26. Shrader CH, Stoler J, Arroyo-Flores J, et al. Geographic disparities in availability of Spanish-language PrEP services among Latino sexual minority men in south Florida. J Immigr Minor Health 2023 Apr;25(2):374-381. [doi: 10.1007/s10903-022-01412-x] [Medline: 36264402]
27. Shrader C, Duncan D, Chen YT, et al. Family members and emotional closeness are associated with more frequent HIV and STI conversations: the N2 cohort study in Chicago. AIDS Behav 2024;28(7):2463-2475. [doi: 10.1007/s10461-024-04348-5]
28. Schneider J, Cornwell B, Jonas A, et al. Network dynamics of HIV risk and prevention in a population-based cohort of young Black men who have sex with men. Net Sci 2017 Sep;5(3):381-409. [doi: 10.1017/nws.2016.27]
29. Schneider J, Michaels S, Bouris A. Family network proportion and HIV risk among black men who have sex with men. J Acquir Immune Defic Syndr 2012 Dec 15;61(5):627-635. [doi: 10.1097/QAI.0b013e318270d3cb] [Medline: 23011395]
30. Chen YT, Duncan DT, Issema R, et al. Social-environmental resilience, PrEP uptake, and viral suppression among young Black men who have sex with men and young Black transgender women: the neighborhoods and networks (N2) study in Chicago. J Urban Health 2020 Oct;97(5):728-738. [doi: 10.1007/s11524-020-00425-x] [Medline: 32468507]
31. Pagkas-Bather J, Young LE, Chen YT, Schneider JA. Social network interventions for HIV transmission elimination. Curr HIV/AIDS Rep 2020 Oct;17(5):450-457. [doi: 10.1007/s11904-020-00524-z] [Medline: 32720253]
32. Duncan DT, Hickson DA, Goedel WC, et al. The social context of HIV prevention and care among Black men who have sex with men in three U.S. cities: the neighborhoods and networks (N2) cohort study. Int J Environ Res Public Health 2019 May 30;16(11):1922. [doi: 10.3390/ijerph16111922] [Medline: 31151275]
33. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. Ann Intern Med 2007 Oct 16;147(8):W163-W194. [doi: 10.7326/0003-4819-147-8-200710160-00010-w1] [Medline: 17938389]
34. The R Stats Package. URL: https://stat.ethz.ch/R-manual/R-devel/library/stats/html/00Index.html [accessed 2023-06-14]
35. Young LE, Ramachandran A, Schumm LP, Khanna AS, Schneider JA. The potential of online social networking data for augmenting the study of high-risk personal networks among young men who have sex with men at-risk for HIV. Soc Networks 2020 Oct;63:201-212. [doi: 10.1016/j.socnet.2020.06.003] [Medline: 33100488]
36. Hubert L, Schultz J. Quadratic assignment as a general data analysis strategy. Brit J Math & Statis 1976 Nov;29(2):190-241. [doi: 10.1111/j.2044-8317.1976.tb00714.x]
37. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal, Complex Systems 2006;1695:1-9 [FREE Full text]
38. Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. statnet: software tools for the representation, visualization, analysis and simulation of network data. J Stat Softw 2008;24(1):1548-7660. [doi: 10.18637/jss.v024.i01] [Medline: 18618019]
39. Leenders R. Modeling social influence through network autocorrelation: constructing the weight matrix. Soc Networks 2002 Jan;24(1):21-47. [doi: 10.1016/S0378-8733(01)00049-1]
40. Doreian P. Network autocorrelation models: problems and prospects. In: Spatial Statistics: Past, Present, Future: Institute of Mathematical Geography; 1989:369-389.
41. Doreian P. Estimating linear models with spatially distributed data. Sociol Methodol 1981;12:359. [doi: 10.2307/270747]
42. Fujimoto K, Chou CP, Valente TW. The network autocorrelation model using two-mode data: affiliation exposure and potential bias in the autocorrelation parameter. Soc Networks 2011 Jul 1;33(3):231-243. [doi: 10.1016/j.socnet.2011.06.001] [Medline: 21909184]
43. Butts CT. Social network analysis with SNA. J Stat Softw 2008;24:1-51. [doi: 10.18637/jss.v024.i06]
44. Dow MM, Burton ML, White DR, Reitz KP. Galton's problem as network autocorrelation. Am Ethnol 1984 Nov;11(4):754-770. [doi: 10.1525/ae.1984.11.4.02a00080]
45. Ord K. Estimation methods for models of spatial interaction. J Am Stat Assoc 1975 Mar;70(349):120-126. [doi: 10.1080/01621459.1975.10480272]
46. Mizruchi MS, Neuman EJ. The effect of density on the level of bias in the network autocorrelation model. Soc Networks 2008 Jul;30(3):190-200. [doi: 10.1016/j.socnet.2008.02.002]
47. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003 Nov;13(11):2498-2504. [doi: 10.1101/gr.1239303] [Medline: 14597658]
48. yFiles layouts. Cytoscape User Manual. 2024. URL: https://manual.cytoscape.org/en/stable/Navigation_and_Layout.html [accessed 2025-04-28]
49. Milgram S. The small world problem. Psychol Today 1967;2(1):60-67 [FREE Full text]

XSL•FO
RenderX

50. Backstrom L, Boldi P, Rosa M, Ugander J, Vigna S. Four degrees of separation. 2012 Jun 22 Presented at: WebSci '12; Jun 22-24, 2012; Evanston, IL. [doi: 10.1145/2380718.2380723]
51. Simpson E, Semaan B. For you, or for "you"? Everyday LGBTQ+ encounters with TikTok. Proc ACM Hum Comput Interact 2021;4(CSCW3):1-34. [doi: 10.1145/3432951]
52. Hiebert A, Kortes-Miller K. Finding home in online community: exploring TikTok as a support for gender and sexual minority youth throughout COVID-19. J LGBT Youth 2023 Oct 2;20(4):800-817. [doi: 10.1080/19361653.2021.2009953]
53. Mattson CL, Freedman M, Fagan JL, et al. Sexual risk behaviour and viral suppression among HIV-infected adults receiving medical care in the United States. AIDS 2014 May 15;28(8):1203-1211. [doi: 10.1097/QAD.0000000000000273] [Medline: 25000558]
54. Nuriddin A, Mooney G, White AIR. Reckoning with histories of medical racism and violence in the USA. Lancet 2020 Oct 3;396(10256):949-951. [doi: 10.1016/S0140-6736(20)32032-8] [Medline: 33010829]
55. Chen YT, Shrader CH, Duncan DT, et al. Using GPS-defined venue-based affiliation networks among Black sexually minoritized men and transgender women to identify locations for HIV prevention interventions. Ann Epidemiol 2024 Aug;96:80-87. [doi: 10.1016/j.annepidem.2024.07.001] [Medline: 38971348]
56. Shrader CH, Duncan DT, Chen YT, et al. Latent profile patterns of network-level norms and associations with individual-level sexual behaviors: the N2 cohort study in Chicago. Arch Sex Behav 2023 Aug;52(6):2355-2372. [doi: 10.1007/s10508-023-02555-0] [Medline: 36877319]
57. Shrader CH, Dt D, Driver R, et al. Social network characteristics associated with more frequent HIV and STI prevention conversations: the N2 cohort study in Chicago. AIDS Behav 2024 Jul;28(7):2463-2475. [doi: 10.1007/s10461-024-04348-5] [Medline: 38703338]
58. Moline T, Duncan DT, Knox J, et al. Neighborhood factors as correlates of alcohol use in the N2 cohort study of Black sexually minoritized men and transgender women. J Urban Health 2024 Dec 20;2024:1-13. [doi: 10.1007/s11524-024-00942-z] [Medline: 39704912]

Abbreviations

ART: antiretroviral treatment
LNAM: Linear Network Autocorrelation Mode
MSA: Metropolitan Statistical Area
N2: Neighborhoods and Networks
PrEP: pre-exposure prophylaxis
QAP: Quadratic Assignment Procedure
SGM: sexually minoritized and gender expansive people
STROBE: Strengthening the Report of Observational Studies in Epidemiology

Edited by A Mavragani; submitted 18.07.24; peer-reviewed by CH Chan, OL Neto, R Xu; revised version received 17.03.25; accepted 18.03.25; published 08.05.25.

Please cite as:
Shrader CH, Duncan DT, Driver R, Arroyo-Flores JG, Coudray MS, Moody R, Chen YT, Skaathun B, Young L, del Vecchio N, Fujimoto K, Knox JR, Kanamori M, Schneider JA
Investigating Social Network Peer Effects on HIV Care Engagement Using a Fuzzy-Like Matching Approach: Cross-Sectional Secondary Analysis of the N2 Cohort Study
JMIR Public Health Surveill 2025;11:e64497
URL: https://publichealth.jmir.org/2025/1/e64497
doi:10.2196/64497

© Cho-Hee Shrader, Dustin T Duncan, Redd Driver, Juan G Arroyo-Flores, Makella S Coudray, Raymond Moody, Yen-Tyng Chen, Britt Skaathun, Lindsay Young, Natascha del Vecchio, Kayo Fujimoto, Justin R Knox, Mariano Kanamori, John A Schneider. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Leveraging Electronic Health Records in International Humanitarian Clinics for Population Health Research: Cross-Sectional Study

Sarah Draugelis[1], BSc; Jessica Hunnewell[2], MPH; Sam Bishop[3,4], BFA; Reena Goswami[5], MD; Sean G Smith[1,6], BBA, BSN, MSc; Philip Sutherland[1], MD; Justin Hickman[1], MD; Donald A Donahue[1,7,8,9], DHEd, MBA, MSJ; George A Yendewa[10,11], MD, MPH; Amir M Mohareb[2,5,12], MD

[1]Team fEMR, St. Clair Shores, MI, United States

[10]Department of Medicine, Case Western Reserve University, Cleveland, OH, United States

[11]Division of Infectious Diseases and HIV Medicine, University Hospitals Cleveland Medical Center, Cleveland, OH, United States

[12]Department of Medicine, Harvard Medical School, Boston, MA, United States

[2]Center for Global Health, Massachusetts General Hospital, 125 Nashua St, #722, Boston, MA, United States

[3]University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

[4]Global Response Medicine, Marco Island, FL, United States

[5]Department of Medicine, Massachusetts General Hospital, Boston, MA, United States

[6]Critical-Care Professionals International, Graham, FL, United States

[7]World Association for Disaster and Emergency Medicine, Madison, WI, United States

[8]Beth Israel Deaconess Medical Center, Boston, MA, United States

[9]Faculty of Medicine, Sigmund Freud Private University, Vienna, Austria

Corresponding Author:
Amir M Mohareb, MD
Center for Global Health, Massachusetts General Hospital, 125 Nashua St, #722, Boston, MA, United States

Abstract

Background: As more humanitarian relief organizations are beginning to use electronic medical records in their operations, data from clinical encounters can be leveraged for public health planning. Currently, medical data from humanitarian medical workers are infrequently available in a format that can be analyzed, interpreted, and used for public health.

Objectives: This study aims to develop and test a methodology by which diagnosis and procedure codes can be derived from free-text medical encounters by medical relief practitioners for the purposes of data analysis.

Methods: We conducted a cross-sectional study of clinical encounters from humanitarian clinics for displaced persons in Mexico between August 3, 2021, and December 5, 2022. We developed and tested a method by which free-text encounters were reviewed by medical billing coders and assigned codes from the *International Classification of Diseases, Tenth Revision* (*ICD-10*) and the *Current Procedural Terminology* (*CPT*). Each encounter was independently reviewed in duplicate and assigned *ICD-10* and *CPT* codes in a blinded manner. Encounters with discordant codes were reviewed and arbitrated by a more experienced medical coder, whose decision was used to determine the final *ICD-10* and *CPT* codes. We used chi-square tests of independence to compare the *ICD-10* codes for concordance across single-diagnosis and multidiagnosis encounters and across patient characteristics, such as age, sex, and country of origin.

Results: We analyzed 8460 encounters representing 5623 unique patients and 2774 unique diagnosis codes. These free-text encounters had a mean of 20.5 words per encounter in the clinical documentation. There were 58.78% (4973/8460) encounters where both coders assigned 1 diagnosis code, 18.56% (1570/8460) encounters where both coders assigned multiple diagnosis codes, and 22.66% (1917/8460) encounters with a mixed number of codes assigned. Of the 4973 encounters with a single code, only 11.82% (n=588) had a unique diagnosis assigned by the arbitrator that was not assigned by either of the initial 2 coders. Of the 1570 encounters with multiple diagnosis codes, only 3.38% (n=53) had unique diagnosis codes assigned by the arbitrator that were not initially assigned by either coder. The frequency of complete concordance across diagnosis codes was similar across sex categories and ranged from 30.43% to 46.05% across age groups and countries of origin.

Conclusions: Free-text electronic medical records from humanitarian relief clinics can be used to develop a database of diagnosis and procedure codes. The method developed in this study, which used multiple independent reviews of clinical encounters, appears to reliably assign diagnosis codes across a diverse patient population in a resource-limited setting.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e66223) doi:10.2196/66223

KEYWORDS

refugee; population health; disaster medicine; humanitarian clinic; electronic health record; Fast Electronic Medical Record; fEMR

Introduction

The global population of forcibly displaced persons, including refugees, asylum seekers, and internally displaced persons, has surpassed 100 million people [1]. Both human-made and natural disasters can break down traditional centers of health care delivery, which may leave displaced persons with gaps in medical care [2]. Displaced persons who leave their countries of origin are often highly mobile between the time of displacement and resettlement [1], which can create challenges for continuity of care, including management of chronic medical conditions, provision of psychiatric and mental health services, and prevention and treatment of communicable diseases.

Humanitarian organizations that provide medical care for displaced persons do so with few resources, limited technological infrastructure, and high patient and staff turnover [3]. Staff turnover may pose serious challenges to maintaining continuity of care for displaced persons. Health information, if properly leveraged, can guide public health interventions for these populations [4,5]. More humanitarian relief organizations are transitioning their operations from paper to electronic medical records, but clinical documentation in ad hoc relief clinics is typically not standardized, nor conducive to analysis [2]. Specifically, clinicians working in resource-limited settings, particularly in disaster response scenarios, may not have time to consistently document all aspects of history-taking and the physical examination, and they may not themselves formally assign *International Classification of Diseases, Tenth Revision* (*ICD-10*) codes to patient encounters [4,5]. Quality of care in these settings can be greatly enhanced if data from clinical documentation can be effectively analyzed and interpreted [6,7]. Electronic health records designed for austere environments may also be an untapped resource to enhance population health and guide public health interventions, for example, by contributing to systems of syndromic surveillance for early warning of infectious disease outbreaks in humanitarian settings [7-9].

Past research has evaluated the ability for free-text clinical documentation to be analyzed with high efficiency, both by manual review and using natural language processing [10,11]. Studies using machine learning methods have used free-text clinical documentation, mostly from high-income settings, to categorize symptom presentations and classify diagnosed chronic diseases [11,12]. However, few studies have evaluated free-text clinical documentation in humanitarian relief settings where clinicians are under particularly high time pressures and operate with limited resources. There is a growing recognition that clinical documentation from humanitarian relief settings could provide much-needed data to measure the outcomes and improve the performance of humanitarian action [13]. To address this gap, we developed and tested a methodology of assigning medical diagnosis codes to free-text medical records collected from a humanitarian clinic at the Mexico-US border.

Methods

Study Design

We conducted a cross-sectional study of clinical encounters from humanitarian clinics for displaced persons at the Mexico-US border between August 3, 2021, and December 5, 2022. There has been a rise over the past decade in the number of displaced persons who transit through the Rio Grande Valley, which is situated near the Eastern coast of the Mexico-US border. Volunteer clinicians with Global Response Medicine, an international humanitarian relief organization, documented medical encounters using Fast Electronic Medical Record (fEMR), a digital medical record system specifically designed for disaster response [14]. Details of a similar clinical operation are separately described [15]. We created a program in which trained volunteer medical coders reviewed free-text medical records from this clinical site and assigned *ICD-10* and *Current Procedural Terminology* (*CPT*) codes. Our analysis included all the encounters with completed medical triage in the camp (ie, the patient registered in the electronic system) and which were assigned a medical diagnosis by the coders.

Ethical Considerations

All fEMR records used in this study were deidentified and made Health Insurance Portability and Accountability Act–compliant to protect patient privacy. Informed consent and participant compensation for research were not possible as all data were deidentified. No research participant was recruited for this study as it was a review of medical records. We did not use machine learning, artificial intelligence, language models, or any similar technology for the purposes of information gathering, analyses, content creation, manuscript writing, or manuscript editing. This study was approved by the fEMR board of directors and the Mass General Brigham Institutional Review Board (2022P003273).

Training of Medical Coders

Coders underwent a rigorous onboarding process before participating in the project. As a prerequisite, they were required to achieve a score of 70% accuracy or higher on a structured entry examination, as determined by experienced medical coders who designed the training. This examination used a subset of medical records from the study dataset to assess the coder's ability to accurately interpret and assign codes in alignment with the project's standards. In addition, all coders were required to complete training in Health Insurance Portability and Accountability Act to ensure compliance with patient privacy and data security regulations.

To support coders in their tasks and ensure uniformity in the coding process, they were provided with access to a

XSL•FO
RenderX

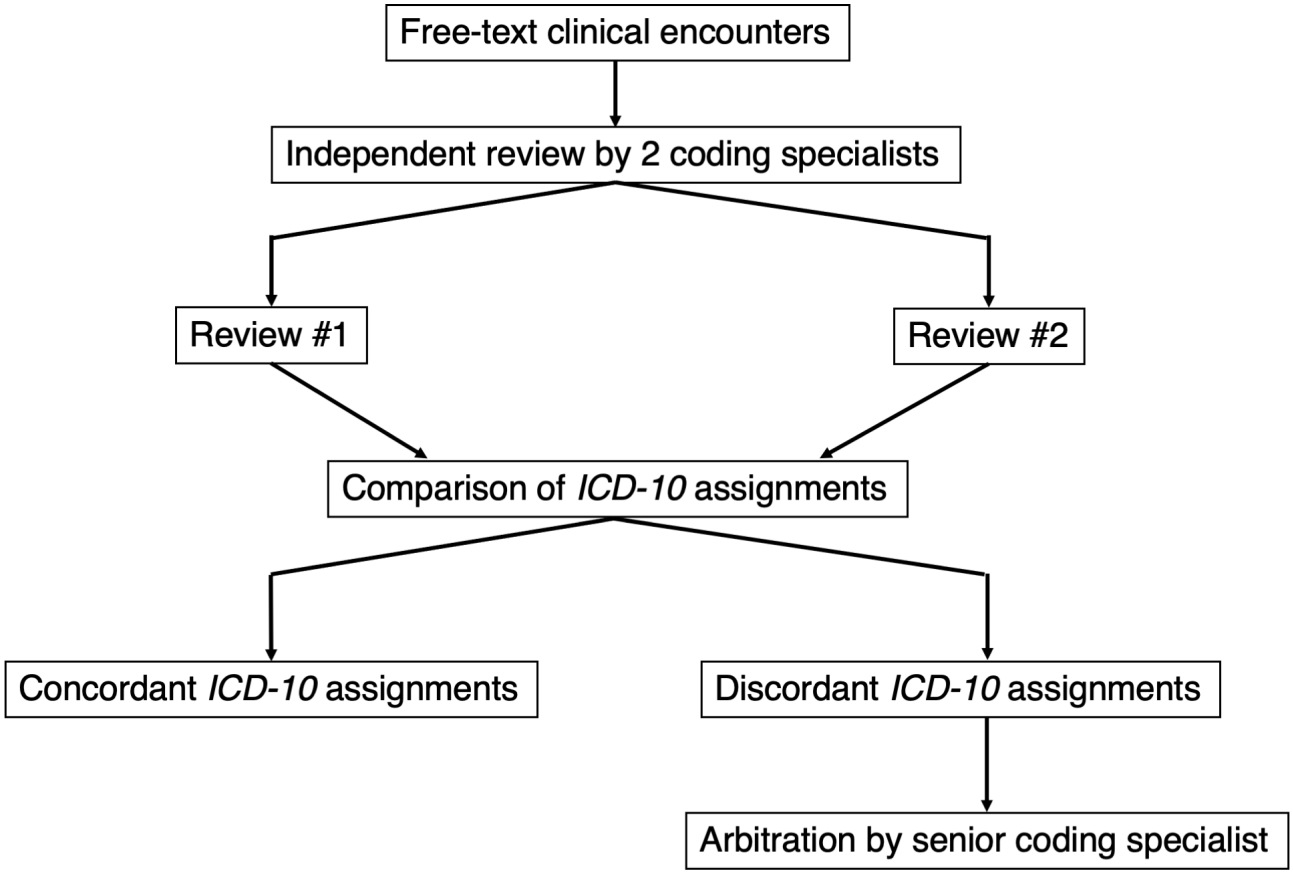
comprehensive set of resources. These resources included detailed instructions on the coding process, examples of coded records, and a manual specifically tailored to the unique aspects of the dataset used in this study. This preparation was designed to standardize the approach across coders, reducing variability and enhancing the reliability of the coding process.

Assignment of Diagnosis Codes

To ensure consistency and reliability in the assignment of *ICD-10* and *CPT* codes, a structured process was developed for the assignment of coders. All coders were randomly assigned blocks of 100 clinical records to review and code. Each block was to be completed within 30 days of assignment. Once a coder submitted their completed block, they were assigned the next available block of 100 records in the dataset. This approach minimized the potential for systematic biases related to the order of record review, such as rater fatigue. Importantly, there were no penalties for exceeding the 30-day completion period, nor incentives for finishing ahead of schedule, ensuring that coders could focus on accuracy rather than speed.

Each record was independently reviewed by 2 coders who assigned *ICD-10* and *CPT* codes in a manner blinded to each other. We compared the codes for concordance, which we defined as an exact match in both the *ICD-10* codes. When discrepancies were identified, the arbitration process was initiated. Specifically, a third, more experienced coder, who had undergone additional training in advanced medical coding, reviewed the original free-text record along with the codes assigned by both initial coders. This coder then independently determined the final coding assignment based on established guidelines and standards for *ICD-10* and *CPT* coding. The final arbitration was documented along with necessary notes detailing the rationale for code selection to facilitate quality assurance and reproducibility in future applications of the process (Figure 1). Quarterly team calls were conducted to review coding processes and provide clarification on complex cases. In addition, team members engaged in ongoing, near-daily discussions through a dedicated group chat platform to address questions and ensure consistency in coding practices.

Figure 1. Workflow for assignment of *ICD-10* codes to free-text clinical encounters from humanitarian relief clinics in Mexico, 2021-2022. *ICD-10*: *International Classification of Diseases, Tenth Revision*.



Statistical Analysis

We evaluated the content of the encounters by measuring the mean, median, and IQR of words in each free-text encounter. We measured concordance between coders by calculating the proportion of encounters assigned the same *ICD-10* diagnosis code. We separately analyzed encounters in which both coders assigned a single diagnosis, both coders assigned multiple diagnoses, and at least 1 coder assigned multiple diagnoses. In cases of multiple diagnoses, we measured complete concordance (defined as all identical diagnoses), partial concordance (defined as at least 1 identical diagnosis), and complete discordance (defined as no identical diagnoses). We used chi-square tests of independence to compare concordance across single-diagnosis and multidiagnosis encounters. We determined the most common diagnoses and analyzed concordance stratified by patient characteristics using chi-square tests of independence for comparison across age, sex, and country of origin. We also

XSL•FO
RenderX

used chi-square testing to compare concordance across encounters with different lengths of free-text documentation. We used R 2024.04.2 for all calculations.

Results

From an original sample of 8546 encounters, we found that 99.0% (n=8460) were assigned at least 1 *ICD-10* code and therefore met inclusion criteria for analysis. These represented 5623 unique patients and 2774 unique diagnosis codes (Table 1). The diagnoses included communicable and noncommunicable illnesses (Table 2). The most common diagnoses included management of hypertension, diabetes mellitus, headaches, and encounters to receive prescription refills for chronic conditions. Common diagnoses that were likely related to infectious diseases during this study period included cough, fever, and urinary tract infection. The mean word count for free-text encounters was 20.5 words, with a median of 14 (IQR 2 - 28) words per encounter. A sample of the free-text clinical documentation and diagnosis code assignments are shown in Table 3. There were 58.78% (4973/8460) encounters where both coders assigned 1 diagnosis code, 18.56% (1570/8460) encounters where both coders assigned multiple diagnosis codes, and 22.66% (1917/8460) encounters with a mixed number of codes assigned.

Table . Concordance of assigned *International Classification of Diseases, Tenth Revision* (*ICD-10*) diagnosis codes across demographic variables in free-text clinical encounters from humanitarian relief clinics in Mexico, 2021 - 2022 (N=8460 encounters).

Characteristic	Proportion	Complete concordance	*P* value
Assigned diagnosis codes, n (%)			<.001
Single, both coders	4973 (58.78)	3096 (62.26)	
Multiple, both coders	1570 (18.56)	215 (13.69)	
Mixed, single and multiple	1917 (22.66)	N/A[a]	
Sex, n (%)			.15
Female	5203 (61.50)	1985 (38.15)	
Male	3257 (38.50)	1294 (39.73)	
Age group (years), n (%)			<.001
0 - 5	827 (9.78)	350 (42.32)	
6 - 10	776 (9.17)	333 (42.91)	
11 - 20	538 (6.36)	210 (39.03)	
21 - 30	1770 (20.92)	740 (41.81)	
31 - 40	2827 (33.42)	1118 (39.55)	
>40	1722 (20.35)	528 (30.66)	
Country of origin, n (%)			.003
Brazil	235 (2.78)	101 (42.98)	
Chile	215 (2.54)	99 (46.05)	
Cuba	46 (0.54)	14 (30.43)	
El Salvador	846 (10.00)	343 (40.54)	
Guatemala	601 (7.10)	220 (36.61)	
Haiti	3157 (37.32)	1328 (42.07)	
Honduras	2126 (25.13)	760 (35.75)	
Mexico	247 (2.92)	98 (39.68)	
Unknown or other	987 (11.67)	316 (32.02)	
Free-text length quartiles, n (%)			<.001
0 - 1 words (first quartile)	2095 (24.76)	1572 (75.04)	
2 - 28 words (second and third quartile)	4179 (49.40)	1407 (33.67)	
29 - 231 words (fourth quartile)	2186 (25.84)	300 (13.72)	

[a]N/A: not applicable.

XSL•FO
RenderX

Table . Most frequently assigned diagnoses in 8460 free-text clinical encounters from humanitarian relief clinics in Mexico, 2021 - 2022[a].

Diagnosis	*ICD-10*[b] code	Frequency, n (%)
Primary hypertension	I10	1234 (14.59)
Type 2 diabetes mellitus	E11.9	639 (7.55)
Headache	R51.9	561 (6.63)
Repeat prescription	Z76.0	542 (6.41)
Cough	R05.9	500 (5.91)
Abdominal pain	R10.9	424 (5.01)
History of diseases of the circulatory system	Z86.79	346 (4.09)
Urinary tract infection	N39.0	345 (4.08)
Pain	R52	329 (3.89)
Fever	R50.9	305 (3.61)

[a]Incomplete records and those with nonspecific diagnosis codes (ie, Z00 codes) were excluded from this table.

[b]*ICD-10*: *International Classification of Diseases, Tenth Revision.*

XSL•FO
RenderX

Table . Sample free-text clinical encounters with assigned *International Classification of Diseases, Tenth Revision* (*ICD-10*) codes from humanitarian relief clinics in Mexico, 2021-2022.

Sex	Age group (years)	Country of origin	Free-text clinical documentation	Coder 1 diagnoses	Coder 2 diagnoses	Arbitrator diagnoses
Male	30-39	Mexico	"Sore throat × 7 days, green mucous, has been taking paracetamol × 7 days. Fever past 2 days. Last tylenol taken 11 pm. No allergies. Taking 2 other medications but can't identify. Note HR. Double checked."	J35.1, J02.9, R50.9	Z00.00	J35.1, J02.9, R50.9
Male	40-49	Honduras	"Pain and swelling in R eye and R neck and R side of head. Sun makes it worse. More than 1 year. Took ibuprofen."	H10.31, R05.9	I10, H10.31	H10.31, I10
Female	60-69	Honduras	"Pt reports being seen here Tuesday for her foot pain and given pain medications at that time that have been working. Pt instructed to continue taking those meds PRN."	R53.1, M79.673, R03.0	R53.1, M79.671	R26.89
Male	40-49	Nicaragua	"Cough productive of yellow phlegm and cold x4 days, afebrile 35.1 here, throat pain, no body aches, no vomiting or diarrhea, no sick contacts. large tonsils without exudates or erythema, lungs clear, nontoxic."	J11.1, R05.9	J11.1, J35.1	R05.8, J00, R07.0
Female	40-49	Haiti	"Patient with diabetes and no insulin for 8 days. previously on metformin but didn't tolerate secondary to nausea. now feels bad with blurred vision and headache. also has h/o ovarian cysts and some abdominal pain. PMH: DM, ovarian cysts."	E11.9, N83.209, Z79.84	E08.620, E11.9	E11.9, N83.209, T38.3× 6A

Of the 4973 encounters with a single code, 62.26% (n=3096) were concordant and 37.74% (n=1877) were discordant. Of the 1877 discordant encounters, there were 68.67% (n=1289) of instances in which 1 coder was concordant with the arbitrator, and 31.33% (n=588) of instances that had a unique diagnosis assigned by the arbitrator which was not assigned by either coder. Therefore, only 11.82% (588/4973) single diagnosis encounters had a diagnosis that was not assigned by either of the initial 2 coders.

XSL•FO
RenderX

Of the 1570 encounters with multiple diagnosis codes, 13.69% (n=215) had complete concordance between the 2 coders, 64.65% (n=1015) were partially concordant, and 21.66% (n=340) were completely discordant. Of the 340 completely discordant encounters, there were 22.94% (n=78) of instances in which 1 coder was concordant with the arbitrator, 61.47% (n=209) of instances in which 1 coder was partially concordant with the arbitrator, and 15.59% (n=53) of instances that had unique diagnosis codes assigned by the arbitrator which were not assigned by either coder. Therefore, only 3.38% (53/1570) of multidiagnosis encounters had a diagnosis that was not assigned by either of the initial 2 coders.

The most frequent diagnoses in concordant clinical encounters were hypertension, acute upper respiratory tract infection, acute nasopharyngitis, viral infection, urinary tract infection, cough, and diabetes mellitus. The most frequent diagnoses in discordant clinical encounters were hypertension, headache, repeat prescription, diabetes mellitus, cough, abdominal pain, and history of diseases of the circulatory system. The frequency of concordance was not significantly different across sex categories: concordance was observed between 38.15% (1985/5203) of encounters with female patients and 39.73% (1294/3257) of encounters with male patients (P=.34). Concordance in diagnosis codes in encounters across age strata ranged from 30.66% to 42.91%. Clinical encounters with older patients (those older than 40 years) had lower concordance between coders than those with other age groups (P<.001). Concordance ranged from 30.43% to 46.05% in encounters across different countries of origin (Table 1).

Encounters with fewer words were more frequently coded in a concordant manner than encounters with more words (Table 1). Specifically, encounters with fewer than 2 words were coded concordantly 75.04% (1572/2095) of the time whereas encounters with greater than 28 words were coded concordantly 13.72% (300/2186) of the time. We found that *ICD-10* codes with less specificity (eg, Z00: Encounter for a General Examination Without Abnormal Findings) were more frequently assigned to encounters with 0 or 1 words.

Discussion

In this analysis of clinical records from ad hoc humanitarian clinics serving displaced persons migrating in Mexico, we developed a system to assign medical diagnosis codes to clinical encounters. This system relied on entry-level medical coders to independently review encounters and then more senior medical coders to arbitrate differences between their assigned *ICD-10* codes. Our dataset included 8460 clinical encounters over a 16-month study period, representing 5623 unique patients. We found that the first 2 medical coders demonstrated concordance in more than 60% of their assignments of single- and multidiagnosis encounters. In cases where they differed, additional review by a third, more experienced coder allowed for resolution of most of the discordance. This novel method can be applied to other circumstances where it is necessary to review medical information collected through the course of medical relief operations.

Our analysis suggests that this approach is reliable across most patient demographic categories. As could be expected, encounters with multiple diagnoses were less frequently coded in a completely concordant manner than encounters with single diagnoses. Encounters with older patients in our sample had a lower proportion of their diagnosis codes completely concordant between assignments. This could be indicative of more comorbidities or more complex medical diagnoses in patients who are older. It is also possible that older adults had more frequent multidiagnosis encounters than younger patients, which could produce discrepancies in the diagnosis assigned based on the free-text medical encounter. While many of the same diagnoses were frequently assigned in both concordantly and discordantly coded encounters, some nonspecific diagnoses (eg, abdominal pain and history of disease of circulatory system) were more commonly assigned in discordantly coded encounters. There was also variation observed between the proportion of encounters with concordant diagnosis assignments across patients of different countries of origin; however, these differences were relatively small. It is possible that there are differences in complexity of the clinical documentation between people with different medical conditions and migration experiences. It is also possible that linguistic or communication barriers were present for some patients based on their country of origin, which may have impacted the clarity with which clinicians documented their evaluation.

We also noted a difference in how medical encounters were coded based on the length of free-text documentation in the medical record. Interestingly, clinical encounters with more text were less frequently coded in a concordant manner than encounters with fewer words. This may be because in encounters with few words, the clinicians frequently just stated the diagnosis rather than any of the clinical findings (eg, “pneumonia” or “hypertension”) whereas lengthy documentation in other encounters may describe several findings that could be coded in different ways. It is also likely that clinical encounters with lengthy documentation may systematically differ from less thoroughly documented encounters as the former may include greater complexity in their presentation, disease process, or management considerations.

To address these challenges, future adaptations of this approach should prioritize enhanced training for coders focused on complex cases, particularly those involving multidiagnosis encounters or nonspecific presentations common in low-resource settings. Future adaptations could also incorporate training on culturally specific presentations of diseases and variations in medical terminology across demographic groups [16]. This could help coders better understand and interpret free-text medical records originating from diverse populations. Collaborating with clinicians to improve the clarity and consistency of free-text documentation would further reduce ambiguities and support more accurate assignments, although this may not be feasible in many austere environments. Finally, periodic audits of coding concordance stratified by demographic characteristics could provide valuable feedback to ensure equity in coding across diverse patient populations.

There is growing use of electronic health records among relief organizations serving displaced persons worldwide [11,12,14].

XSL•FO
RenderX

Past reports from humanitarian settings have identified challenges in the implementation of *ICD-10* coding into existing health records, including the time required for data entry, staff turnover, and the operational capacity for data analysis [13,17,18]. The approach used in this study can be used to characterize medical diagnoses from electronic health record systems in disaster response and other austere settings with a high level of consistency [15]. The strengths of this approach include the reliability in coding and the ability to have records reviewed and coded remotely while relief operations are ongoing. It also has applicability to diverse geographic and sociopolitical contexts where clinicians work in low-resource environments. It should be noted that we have no way to ascertain diagnosis accuracy (ie, the "true" illnesses and conditions present in the clinical encounters studied). Our method of completing multiple reviews of each encounter presents challenges for efficiency with high-volume clinical operations. Future work may leverage large language models to more efficiently code clinical encounters, but this would require validation before implementation [12]. As with other operations involving electronic health records, patient privacy and data security are also paramount concerns [16].

There are limitations in this analysis and approach. First, given that this analysis was conducted retrospectively, we cannot compare the accuracy of diagnostic codes with the reference standard of clinician designation, nor is this information readily available from clinics in austere environments. In addition, because of the limited diagnostic capabilities in the clinical sites studied, most of the diagnoses were based on the bedside examination (ie, history and examination), which resulted in nonspecific diagnoses being assigned.

In summary, we developed a system to assign *ICD-10* codes to thousands of clinical encounters for displaced people in transit at the Mexico-US border, with a variety of clinical presentations for a diverse patient population. The simple methodological approach described in this report holds potential to be adapted to other disaster response and resource-limited settings to advance the provision of clinical care and population health research. Scale-up of the electronic health record technology described in this study can improve patient safety by allowing program managers to monitor quality of care of diagnoses across patient populations. It can also provide critical information to public health officials regarding risk and clusters of transmissible infections, trauma and injuries, and outbreak preparedness.

Acknowledgments

The authors wish to thank the clinical and project staff who worked in the relief clinics and the coding staff at fEMR (Fast Electronic Medical Record).

Data Availability

Aggregate forms of the datasets generated during and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Global Trends report 2023. United Nations High Commissioner for Refugees. 2024. URL: https://www.unhcr.org/global-trends-report-2023 [accessed 2024-07-31]
2. Lind K, Gerdin M, Wladis A, Westman L, von Schreeb J. Time for order in chaos! A health system framework for foreign medical teams in earthquakes. Prehosp Disaster Med 2012 Feb;27(1):90-93. [doi: 10.1017/S1049023X11006832] [Medline: 22591933]
3. Jafar AJN, Norton I, Lecky F, Redmond AD. A literature review of medical record keeping by foreign medical teams in sudden onset disasters. Prehosp Disaster Med 2015 Apr;30(2):216-222. [doi: 10.1017/S1049023X15000102] [Medline: 25659602]
4. Abir M, Mostashari F, Atwal P, Lurie N. Electronic health records critical in the aftermath of disasters. Prehosp Disaster Med 2012 Dec;27(6):620-622. [doi: 10.1017/S1049023X12001409] [Medline: 23009722]
5. Were MC. Challenges in digital medicine applications in under-resourced settings. Nat Commun 2022;13(1):3020. [doi: 10.1038/s41467-022-30728-3]
6. Lal A, Ashworth HC, Dada S, Hoemeke L, Tambo E. Optimizing pandemic preparedness and response through health information systems: lessons learned from Ebola to COVID-19. Disaster Med Public Health Prep 2022 Feb;16(1):333-340. [doi: 10.1017/dmp.2020.361] [Medline: 33004102]
7. Kruse CS, Stein A, Thomas H, Kaur H. The use of electronic health records to support population health: a systematic review of the literature. J Med Syst 2018 Sep 29;42(11):214. [doi: 10.1007/s10916-018-1075-6] [Medline: 30269237]
8. Syzdykova A, Malta A, Zolfo M, Diro E, Oliveira JL. Open-source electronic health record systems for low-resource settings: systematic review. JMIR Med Inform 2017 Nov 13;5(4):e44. [doi: 10.2196/medinform.8131] [Medline: 29133283]

XSL•FO
RenderX

9. Jarrett SY, Redfern A, Li J, et al. A novel tool for assessing pediatric emergency care in low- and middle-income countries: a pilot study. Int J Emerg Med 2025 Jan 16;18(1):15. [doi: 10.1186/s12245-024-00802-2] [Medline: 39819298]
10. Sanderson AL, Burns JP. Clinical documentation for intensivists: the impact of diagnosis documentation. Crit Care Med 2020 Apr;48(4):579-587. [doi: 10.1097/CCM.0000000000004200] [Medline: 32205605]
11. Sheikhalishahi S, Miotto R, Dudley JT, Lavelli A, Rinaldi F, Osmani V. Natural language processing of clinical notes on chronic diseases: systematic review. JMIR Med Inform 2019 Apr 27;7(2):e12239. [doi: 10.2196/12239] [Medline: 31066697]
12. Koleck TA, Dreisbach C, Bourne PE, Bakken S. Natural language processing of symptoms documented in free-text narratives of electronic health records: a systematic review. J Am Med Inform Assoc 2019 Apr 1;26(4):364-379. [doi: 10.1093/jamia/ocy173] [Medline: 30726935]
13. Altare C, Weiss W, Ramadan M, Tappis H, Spiegel PB. Measuring results of humanitarian action: adapting public health indicators to different contexts. Confl Health 2022 Oct 14;16(1):54. [doi: 10.1186/s13031-022-00487-5] [Medline: 36242013]
14. Draugelis SD, Brown EC, Donahue DA, et al. Development and implementation of an electronic health record system for use in humanitarian emergencies, disaster response, and conflict zones. PLOS Glob Public Health 2025;5(1):e0003124. [doi: 10.1371/journal.pgph.0003124] [Medline: 39879195]
15. Reynolds CW, Cheung AW, Draugelis S, et al. Epidemiology of asylum seekers and refugees at the Mexico-US border: a cross-sectional analysis from the migrant settlement camp in Matamoros, Mexico. BMC Public Health 2024 Feb 16;24(1):489. [doi: 10.1186/s12889-024-17947-7] [Medline: 38365627]
16. Chiesa V, Chiarenza A, Mosca D, Rechel B. Health records for migrants and refugees: a systematic review. Health Policy 2019 Sep;123(9):888-900. [doi: 10.1016/j.healthpol.2019.07.018] [Medline: 31439455]
17. Mesmar S, Talhouk R, Akik C, et al. The impact of digital technology on health of populations affected by humanitarian crises: recent innovations and current gaps. J Public Health Pol 2016 Nov;37(S2):167-200. [doi: 10.1057/s41271-016-0040-1]
18. Broadwin C, Zeidan W, Siam M, et al. Advancing refugee health data management: the implementation of ICD-11 in UNRWA's primary care system. Int J Environ Res Public Health 2024 Aug 26;21(9):1121. [doi: 10.3390/ijerph21091121] [Medline: 39338004]

Abbreviations

***CPT*:** *Current Procedural Terminology*
fEMR: Fast Electronic Medical Record
***ICD-10*:** *International Classification of Diseases, Tenth Revision*

Edited by A Mavragani; submitted 08.10.24; peer-reviewed by AD Kuday, R Mozer; revised version received 20.02.25; accepted 21.02.25; published 17.04.25.

Please cite as:
Draugelis S, Hunnewell J, Bishop S, Goswami R, Smith SG, Sutherland P, Hickman J, Donahue DA, Yendewa GA, Mohareb AM
Leveraging Electronic Health Records in International Humanitarian Clinics for Population Health Research: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e66223
URL: https://publichealth.jmir.org/2025/1/e66223
doi:10.2196/66223

© Sarah Draugelis, Jessica Hunnewell, Sam Bishop, Reena Goswami, Sean G Smith, Philip Sutherland, Justin Hickman, Donald A Donahue, George A Yendewa, Amir M Mohareb. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 17.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Opportunities and Challenges for an Organizational Digital Public Health Strategy in a Provincial Public Health Program in Canada: Qualitative Description of Practitioner Perspectives

Ihoghosa Iyamu[1,2], MD, MDICHA, PhD; Devon Haag[1], MSc; Anna Carson[3], MSc; Ivy Wang[4], MM; Colin King[4], MSc; Ian Roe[1], MMIE; Kristy Kerr[5], MPH; Sofia Bartlett[1,2], PhD; Geoffrey McKee[1,2], MD, MPH; Mark Gilbert[1,2], MD, MHSc

[1]BC Centre for Disease Control, Vancouver, BC, Canada

[2]School of Population and Public Health, University of British Columbia, 2206 East Mall, Vancouver, BC, Canada

[3]Independent Consultancy, Vancouver, BC, Canada

[4]The ClearView Group - Solutions for Health, Delta, BC, Canada

[5]Pacific Public Health Foundation, Vancouver, BC, Canada

Corresponding Author:
Ihoghosa Iyamu, MD, MDICHA, PhD
BC Centre for Disease Control, Vancouver, BC, Canada

Abstract

Background: The digital transformation of health services accelerated during the pandemic. While "digital health" strategies were created, they paid minimal attention to public health services like health promotion, disease surveillance, emergency preparedness, and health protection.

Objective: This study aimed to inform a digital public health (DPH) strategy at the British Columbia Centre for Disease Control (BCCDC) and explored public health practitioners' perspectives on challenges and opportunities of integrating digital technologies into public health functions within the organization.

Methods: In this qualitative description, we conducted 18 focus groups (FGs) between January and June 2023, drawing practitioners from 9 organizational subunits of the BCCDC including population and public health, environmental health, clinical services, vaccine-preventable diseases, communications, knowledge translation, data analytics, and Indigenous health (2 FGs per subunit). Discussions explored practitioners' application of digital technologies in their public health work, focusing on challenges encountered during implementation (current state FGs) and perceived opportunities (future state FGs). Sessions were audio-recorded, and detailed field notes were taken. Thematic analysis was conducted, comparing perspectives across groups using constant comparative techniques.

Results: We identified 3 themes. First, "bridging existing inequities—an opportunity and a challenge contingent on public trust" described participants' excitement about opportunities for DPH to disrupt historical inequities if centered on trust and reconciliation, while recognizing current digital transformation efforts risk exacerbating existing inequities with the digital divide. Second, "a sense of disconnect between "digital" and "public health" functions" described perceptions of DPH as being out of scope of core public health duties, requiring new competencies and navigation of complex organizational policies for which support is suboptimal. Third, "balancing the need for responsive DPH with necessary reactivity" highlighted practitioners' yearnings for a proactive DPH strategy rather than current issue-based reactive approaches. Participants suggest that a centralized systematic program can help achieve this goal.

Conclusions: A cohesive, systematic, and proactive organizational strategy for DPH is critical to enable equity-focused digital transformation. Such a strategy can bridge perceived disconnects between digital and public health functions through organizational supports like competency development and streamlined policies that can better support public health practitioners to integrate digital technologies into their work.

(JMIR Public Health Surveill 2025;11:e72588) doi:10.2196/72588

KEYWORDS

digital health; digital public health; implementation science; digital health strategy

XSL•FO
RenderX

Introduction

Uptake of digital technologies for public health functions accelerated in the past decade, especially during the COVID-19 pandemic [1-3]. These technologies have been used to streamline real-time data processing, analysis, and reporting for health surveillance and emergency response [1,2,4]. They have also been used to expand the reach of contextually relevant and personalized health promotion and protection services through social media, mobile apps, and other similar technologies [1,5-7]. In 2022, Canada's Chief Public Health Officer identified medical and digital health technology as a key part of strengthening Canada's public health systems, recognizing its potential to improve the speed, efficiency, and impact of public health programs [8]. The strategy includes using digital technologies to strengthen early warning systems, improve supply chains and laboratory services through real-time tracking, support evidence-based decision making, and promote adherence to public health interventions via social media [8,9]. At the same time, it highlights the importance of carefully evaluating these technologies in light of legal and ethical concerns [8].

Current digital health strategies primarily focus on clinical applications, often overlooking the broader public health implications of digital transformation [10]. For instance, British Columbia's digital health strategy emphasizes a trusted, digitally enabled health system aimed at patient empowerment, provider experience, system connectivity, and operational efficiency—largely within clinical settings [10,11]. However, the rapid adoption of digital technologies across sectors like commerce, education, and social services has significant public health implications [10,12]. Social media and the digital information environment also contribute to health inequities, particularly through misinformation and disinformation, which disproportionately affect the youngest, oldest, and least digitally literate populations [10,12-14]. Digital access and literacy influence key health determinants, shaping employment, education, and income, which in turn affect health behaviors, access, and outcomes [10,12,15,16]. A recent expert consensus highlights how digital transformation intersects with social, commercial, and political determinants of health [17]. Recognizing this, BC's 2024 Population and Public Health Framework formally identifies the digital environment as a key determinant of health requiring public health action [18].

Yet, few public health agencies have developed digital public health (DPH) strategies that not only consider the role digital technologies play in strengthening public health functions, but also better position them to respond to the public health implications of society's digital transformation as described [10]. This project was conducted as part of efforts to inform a DPH strategy within a provincial public health organization in Canada. This provincial organization is responsible for providing health surveillance, disease and injury prevention, diagnosis and treatment of diseases of public health importance, health promotion (including healthy public policy), and emergency response functions. This project was conducted in part to align organizational public health action with the broader British Columbia digital health strategy and the 2024 British Columbia population and public health framework, seeking to better position the organization and its workforce to address this evolving public health need. Therefore, we explored public health practitioners' perspectives on the challenges and opportunities of applying digital technologies to public health functions within our organizational context.

Methods

Study Design

This was a qualitative description of public health practitioners' perspectives on the opportunities and challenges of leveraging digital technologies to operationalize public health functions [19]. Our aim was to generate meaningful insights to inform the deliberations regarding the development of a DPH strategy within our organization. Similar qualitative approaches have been used to guide health services planning and policy formulation [20]. Our inquiry was grounded in a pragmatic and constructivist orientation, enabling us to draw from relevant theory while remaining participant-driven [21]. We built on the British Columbia Centre for Disease Control's (BCCDC) strategic plan, which identifies "promoting digital health and clinical services" as a transformational priority, situating our study within the context of organizational change [10,22]. To inform our exploration of contextual factors shaping digital transformation, we drew on Weiner's theory of organizational readiness for change, with attention to organizational policies, strategic priorities, perceived value of the change (change valence), and available resources [23]. Our working definitions of DPH were derived from a scoping review by our research team, alongside an internal review and national consultation process -Textbox 1 [22].

The project was led by public health practitioners (DH, GM, IR, II, and MG) with experience designing, implementing, and evaluating DPH interventions within the organization in collaboration with relevant health systems partners [24]. The research team also included private sector consultants with expertise in digital transformation across various health contexts (IW and CK) and public health researchers with qualitative research experience in applied health systems settings (II and AC). All authors were part of a DPH working group established with representation across the organization to guide the implementation of the project.

We acknowledge that our diverse professional roles and previous engagement with digital transformation influenced our philosophical stance and study design. Our constructivist approach was informed by an understanding that participants' perspectives are shaped by their specific roles, work environments, and organizational objectives [19]. To mitigate potential bias and enhance reflexivity, data collection was conducted by consultants external to the organization, and analysis was led by an independent qualitative researcher who maintained a hands-off approach to support analytic independence. The full team engaged in regular reflective discussions on the emerging findings, using these sessions to critically examine how our positionality as insiders and researchers may have influenced interpretation. These ongoing reflective practices were integral to maintaining transparency and rigor in our approach.

XSL•FO
RenderX

Textbox 1. Operational definition of digital public health.

Definition of digital public health

- Digital public health is the use of digital technology to transform the delivery of public health functions in people-centered ways that optimize health outcomes for all. It brings together expertise in health, technology, and other disciplines to improve the health of populations, while addressing modern public health challenges that are amplified through the rapid and widespread uptake of digital technologies.
- Digital public health uses digital technologies to design, implement, evaluate, and scale up interventions to address public health challenges and promote health equity. It includes a wide range of digital technologies and applications, including data analytics, mobile health, social media, virtual health, and others.

Study Setting, Sampling, and Recruitment

This project was conducted at the BCCDC, a program of the Provincial Health Services Authority responsible for provincial and national leadership in disease surveillance, detection, treatment, prevention, and health promotion. The program is organized into subunits called service areas including clinical prevention services and its 2 clinical programs (sexually transmitted infections and tuberculosis clinics); data and analytics services (DAS); environmental health services (EHS); immunization program and vaccine-preventable diseases; public health response (PHR); Population and Public Health and cross-cutting programs including the Chee Mamuk Indigenous Health Program and the communications and knowledge translation (KT) Program. This project was informed by broader organizational strategic planning goals of promoting digital health and clinical services.

Participants were recruited from across key organizational subunits (service areas and programs) to ensure a range of operational perspectives on DPH. Recruitment followed a purposive sampling approach, with invitations shared through team leads and internal email lists. All staff within the selected units were eligible to participate, and participation was voluntary. Due to privacy and organizational sensitivities, particularly given the small size and specialized nature of many teams, we do not report detailed participant characteristics such as roles or years of experience.

Data Collection

Between February and April 2023, we conducted 2 rounds of focus groups (FGs-18 in all) with each of the service areas as described. FGs lasted an average of 52 minutes (range: 39 - 64 min). In the first round, we explored challenges with applying digital technologies to public health functions in the current organizational context. In the second round, we explored potential opportunities and priorities for a desired future state of DPH within the organization. FG guides were adapted to each service area, with initial sessions informing slight adaptations to ensure that discussions elicited context-specific insights while maintaining comparability across groups Multimedia Appendix 1. FGs were facilitated by (DH, IW, and CK) who have experience leading group discussions. FGs were conducted on Microsoft Teams and all sessions were auto-transcribed, with transcripts reviewed for accuracy to mitigate potential technology-related issues. Detailed field notes were also recorded during and after each FG.

Ethical Considerations

This project did not require formal ethics approval as it was classified as quality improvement given that it was under the jurisdiction of a public health authority and did not have research as a primary goal [25]. However, we followed basic ethical principles [26]. We obtained client consent, including verbal assent for recordings. To protect privacy, no identifiable information was collected. Participants could skip questions or withdraw at any time. No honorarium was provided, as FGs occurred during paid work hours.

Data Analyses

We conducted a reflexive thematic analysis on the combined dataset from both rounds of FGs, following Braun and Clarke's recommendations [27,28]. First, verbatim transcripts and field notes from both rounds of FGs were reviewed for accuracy and edited as necessary. All transcripts were imported into QSR NVivo (version 12) for data management and analysis [29]. The initial coding was performed by a primary analyst (AC), who closely read and reread transcripts to identify salient ideas and patterns in the data. A codebook was developed inductively and iteratively refined through regular consultations with the DPH working group Multimedia Appendix 2. A second analyst (II) reviewed the coding tree and assigned codes to ensure accuracy.

Throughout the coding and analysis process, we held weekly meetings with the DPH working group to discuss emerging codes, reflect on interpretive questions, and consider how organizational context influenced participants' perspectives. We used triangulation across field notes, participant quotes, and group discussions to ensure credibility and conducted crosstabulations to explore patterns across different roles and areas of practice. Themes were developed, refined, named, and described through a process that balanced theoretical sensitivity with fidelity to participants' voices. The final themes reflect the contextual realities, expectations, and concerns of practitioners engaged in DPH-related initiatives. In line with the qualitative descriptive approach, our analysis focused on capturing rich, context-specific insights without aiming for data saturation [19]. Our reporting adheres to the consolidated criteria for reporting qualitative research (COREQ) Checklist 1 [30].

Results

Description of FG Participants

Table 1 describes participants recruited from various service areas and programs in 2 rounds of FGs. Overall, there were 60 unique participants across 18 FGs, representing 6 service areas and 5 programs within BCCDC. Participants worked at varying

XSL•FO
RenderX

levels including frontline clinical services, middle level, and executive-level management.

We identified 3 themes which are described in detail below, including representative quotes as appropriate.

Table . Description of focus group participation by Sub-Unit and round–British Columbia Centre for Disease Control (2023).

Focus group	Number of participants	
	Round 1	Round 2
Clinical prevention services line–sexually transmitted infections program	8	7
Clinical prevention services line-tuberculosis program	3	5
Data and analytics services line	6	5
Environmental health services	5	6
Immunization program and vaccine-preventable diseases	3	4
Public health response	6	4
Population and public health	4	10
Chee mamuk indigenous health program	9	4
Communications and knowledge translation programs	3	4

Bridging Existing Inequities: An Opportunity and a Challenge Contingent on Public Trust

Practitioners described digital transformation as an opportunity to improve equity of public health services among historically marginalized populations. This was especially recognized for Indigenous peoples in Canada for whom additional opportunities for reconciliation exist. Practitioners emphasized that realizing the potential of digital transformation is contingent on establishing public trust which is influenced by the reach and effectiveness of current DPH efforts, the accessibility of digital services and resources, the timeliness of current efforts to bridge existing digital divides, and perceived benefits of digital transformation to various populations. Priorities for equity and reconciliation were especially emphasized for Indigenous communities who may be marginalized by current digital strategies. One participant said

> *The truth of the matter is, there's a lot of communities who are still existing on those older systems and like to me, it's I guess for where I would like to currently sit on this conversation, is sitting in that tension of those who have and those who do not have, because at the end of the day, the people who are most going to pay for not being digitally connected are Indigenous people and that's already knowing like we literally already know that and none are. But in the health systems, we're actually going more to virtual and digital and we're ensuring that we leave Indigenous people out.* [Indigenous health program, Session 2]

While recognizing the potential for increased accessibility, reach, equity, effectiveness, and efficiency of public health services through digital transformation, service areas emphasized concerns about further marginalizing populations without dependable digital access and literacy. Practitioners described scenarios where Indigenous, rural, and remote communities, people with lower socioeconomic status, and people experiencing language barriers or visual impairments are already being marginalized by our current digital transformation efforts. For example, one participant described a situation where Indigenous Elders seeking vaccines were caught in a technological dilemma

> *We had elders that were reaching out to the 1 - 800 number looking for vaccination. Only to call the number, sit on the line for 45 minutes, and then for them to get on to the line of like saying, "Hey, you have to go to this website" and you're like, "Well, we don't have a laptop, nor do we have Internet. We were told to call this number." And they're like, "Yeah, you have to go to this website." And so they were basically just pointing at each other.* [Indigenous health program, Session 1]

Carefully considering the benefits of digital transformation of public health alongside limitations of the digital divide was described as an important way to avoid creating an inherently inequitable 2-tiered system. Practitioners also suggested that fostering trust in DPH is dependent on broader sociocultural contexts. They described various negative community perceptions of public activities including data collection through intervention, to planning implementation and KT. The need to address these perceptions in meaningful ways was described as fundamental to the organization's ability to make progress within the fast-paced technological landscape

> *Privacy is a constant struggle and it feels like there's often the implicit assumption that we are trying to use the data to do terrible things and you know changing the conversation and it's starting to change for sure, but changing the conversation to reframe it as the things we can do to benefit the population when*

XSL•FO
RenderX

> *the data exist and are accessible to us as opposed to the, you know, the various purposes that that we might put the data to you and you know if I had a blue sky world, you know that would be the first thing that everybody thought of when we asked for data or made a request for data was "wow, think of what good these people could do if they had access to those data.* [EHS, Session 2]

Practitioners highlighted the importance of recognizing how socioeconomic status, geographical location, and access to health care services can both facilitate and hinder public trust in DPH, encouraging careful consideration of DPH's role to address the potential harms of the broader digital transformation of society and new inequities resulting from the process (ie, the digital determinants of health)

> *It would be a miss if we didn't consider our mandate to address some of the harms and facilitators that there are from digital technology adoption and the broader community in the work that we do. I mean just looking at how certain technologies-through inequitable uptake- can exacerbate underlying inequities within society and how that addresses or how that can exacerbate the social determinants of health.* [Population and public health, Session 2]

Practitioners also emphasized the need to provide concurrent non-digital options for public health interventions to promote equity. The following quote provides important context on addressing the digital divide as a 2-fold process

> *The digital divide, I would characterize as two separate issues. The first one is people who just simply do not have digital access at all and that could be anything from having their own personal contact number, but it could also be like not having financial resources to have Internet connectivity. Because even people who have devices actually often do not have access because of the cost of data in Canada, or they don't actually have consistent access to the Internet… And so, I think that this is actually something that's a very kind of important gap and challenge for us to be thinking about, but also acknowledging. And then the second part of the digital divide is people who actually have access to these things, but do not have the digital literacy to be able to actually fully utilize them, and I would put a bunch of people into that kind of category. You know, myself included.* [Sexually transmitted infections, Session 2]

Practitioners suggested that implementing strategies to address these issues can foster trust building among diverse populations, potentially enhancing the acceptability, efficacy, and efficiency of DPH.

A Sense of Disconnect Between "Digital" and Public Health Functions

The second theme highlighted a perceived disconnect between the staff's daily work and the envisioned or idealized practice of "doing DPH work." Practitioners described how engaging with digital health and technologies was often viewed as "being in excess of assigned roles." While often described with some frustration, practitioners also expressed excitement about the potential of DPH skill sets and practices to improve the reach, efficiency, and effectiveness of internal and external services of the organization. For example, this participant said

> *I'm kind of thinking, wow, if we had a tool to do this, that would be great. I have all these potential future uses in mind now but I have no idea if there is such a thing. But you know, if there's something that could help us do the work better or, you know, in new ways and reach more people, then that's really exciting.* [PHR, Session 1]

Practitioners described being unable to effectively formulate digital solutions to health systems problems they encounter. They emphasized needing in-service training and centralized supports to address competency gaps. These competency gaps were described as contributory to feelings of frustration with systems inefficiencies which exacerbate current workloads:

> *My experience trying to submit my [cloud-based data platform] application was excruciating. Like I'm a pretty smart person, but not having expertise made that a horrible, very, very long process. Whereas if I had someone that was used to doing that, skilled in doing that, they could have sat down with me and spent an hour. So, trying to make each individual group become an expert, or at least skilled in doing this as opposed to having central resources that already have those skills and do it all the time, I think is really inefficient and it leaves opportunities on the table.* [EHS, Session 1]

As described, practitioners acknowledged competency gaps but also emphasized the role of centralized support in filling capacity gaps for DPH services given current workloads and a sense that DPH services are not a core responsibility. For example, this participant said:

> *We can give information, but it's not on us because we literally don't have the capacity, like we just don't. And it would make sense […] like, where's that digital team?* [Indigenous health program, Session 2]

Practitioners also suggested creating novel positions with technology-specific job descriptions, with comparable remuneration rates to similar roles in the private sector. This was described as a prerequisite for effective DPH transformation, involving a re-evaluation of current human resource classification systems and securing additional funding

> *I think that is one thing that basically requires creating a new stream of jobs within BCCDC […]. Just say like we don't have these jobs and have been working on even bringing one person on board and we have been just kind of like trying to pigeonhole these people into our existing positions and that also kind of like goes back to our comp and class system, which is a little bit archaic. It takes quite a bit of time to get a new position classified and so on, so it's at multiple levels. First is just kind of like the realization to think through, these are skills that are needed for*

XSL•FO
RenderX

> *the future of work that we do at [BCCDC]… and then the second is creating these job descriptions, having funding allocated to hire these people and then create an ecosystem–whether it's through partnerships with academics and the scientists here to provide that environment where people are successful in implementing their skills and providing solutions to our problems.* [DAS, Session 2]

Practitioners suggested internal capacity-building to effectively use existing digital technologies (eg, limited knowledge of Microsoft Office suite tools, graphic design, and tools to support building and maintaining internet-based courses and manuals) while prioritizing the need to hire in-house resources to support more advanced digital and analytics projects.

Balancing the Need for Responsive DPH With Necessary Reactivity

This theme emphasizes the need to balance systems' reactivity with responsivity through an approach that fosters a proactive DPH strategy across service areas and programs. Here, reactivity was used to describe issue-based and unplanned implementation of digital technology in response to emergent public health issues as compared with responsivity, which describes proactive and planned strategies for considering DPH within broader public health responses. Practitioners described creative, adaptive, and individual approaches to implementing digital tools and programs, both internally and externally, and described needing a centralized approach at the BCCDC to guide DPH transformation. Practitioners critiqued overly reactive systems built on assumptions of public health's inherent function of reacting to emerging threats. This was said to result in multiple and inconsistent communication channels with public and providers, potentially limiting the organization's public health impact. For example, one practitioner said

> *So, in terms of messaging information and from a communicable disease, rapid response perspective, being able to push out consistent messaging on a new emerging [threat], whether it's a measles outbreak or an upcoming heat [wave] because we don't have that one-stop shop […] It's just–there's too many ways to get slightly different information and then we end up potentially not having the most impactful inputs or support to that group of providers.* [PHR, Session 2]

This reactive approach often leads to rapidly developed DPH interventions described as frequently untested, duplicative of previous efforts, and contributing to fragmented DPH systems. These fragmented systems were described as potentially wasteful and challenging to integrate into existing systems, ultimately undermining confidence in DPH initiatives. To reduce this reactivity, one participant suggested promoting broader access to and understanding of emerging digital tools across the organization

> *It's not just someone to help create that vision with using all these tools, but also the access to the right resources in a timely manner to learn what else is coming. I could dream up a lot of things, but I can only go so far because I don't know what else is in the [IT subunit] world. The funny thing is, like the texting solution that worked on the COVID side, I was the technology expert, way back in the day, we had created the concept before, and it just sat on the shelf.* [EHS, Session 2]

While the organization may not anticipate all potential technological shifts, setting the groundwork for information management (eg, privacy guidelines and standards for novel datasets), communication, and centralized support will provide needed foundations for service areas to work from. Perspectives of unclear DPH foundations are compounded by the current ad-hoc approach to DPH relying on staff motivation, as illustrated in the following quote about the lack of institutional knowledge in creating online courses for health care providers

> *I mean, going back to the conversation we're having about the courses that are built on [software name]. Like we don't necessarily have any training. People just have taken it upon themselves to figure it out.* [STI, Session 1]

Practitioners' narratives emphasized the need for a systematic and centralized program to lead the organization's effort in a more responsive and proactive approach to DPH. This also involves enhancing data, clinical, and operational interoperability among health authorities, governmental levels, and other health systems partners. Participants highlighted the challenges they experienced with provincial interoperability

> *These...solutions within [health authority (HA1)] and then later on within [HA2] and then for the regular communicable disease and other diseases. We have [software 1] which is being used by one or two health authorities. Then we have [software 2] within [HA3] and then we have [software 3] within [HA1] and then a different system in [HA4]. Like we don't have a good mechanism to capture information and then a mechanism to develop some of these solutions on the fly, because in public health you need to have some nimbleness, like when COVID emerged.* [DAS, Session 2]

Interoperability challenges were also said to result from complex and inconsistent data access protocols, agreements, and platforms that are cumbersome to navigate at an interinstitutional level, while acknowledging ongoing work to address these barriers.

> *There just isn't provincially consistent and reliable structures for sharing data across groups, and we get hung up in, you know, information–like establishing information sharing agreements can take years in some cases, and it really impedes our ability to do the work that we need to do. So that's kind of a constant frustration.* [EHS, Session 1]

Practitioners highlighted opportunities to leverage and repurpose existing technologies and platforms across service areas, automate tasks (eg, data entry and cleaning, communication with patients and providers), and improve communication, policies, and training opportunities as potential solutions aligned with public health needs. As noted by this participant, lengthy

XSL•FO
RenderX

approval processes do not align with priorities for nimble responsiveness or necessary reactivity.

> *The thought is that there needs to be a generally rapid response if something is out there, and currently, because of the communication approval process, there's very little that's rapid. So, I think there's a lot of like bits feeding into, you know, even if we do decide that we need to change how we manage mis- or disinformation, we're not set up for success in that at the moment.* [KT, Session 2]

Clinical staff also noted technological gaps that could alleviate frustrations in patient-facing services if addressed. Comprehensively integrating more flexible and automated services was highlighted as a change that could improve system responsivity. For instance, while acknowledging external constraints (ie, interinstitutional interoperability), the clinical tuberculosis program offered suggestions to improve internal responsivity through specific channels

> *There's the appointment reminders piece, like the automated appointment reminders piece. There's the fact that we still mail things out. So, there's no direct fax capability like there is with basically any EMR that you buy off the shelf. We don't have a texting platform. We don't have a scheduler that actually allows us to provide and schedule the patients in the way in which we would like to schedule them–like we have an incredibly rigid scheduler.* [Tuberculosis, Session 1]

Practitioners emphasized needing robust mechanisms for information, data exchange, and collaboration among partners within the DPH landscape. This includes advocacy for standardized protocols and frameworks that facilitate seamless data sharing and communication among health authorities and government entities at various levels. However, prioritizing efforts to enhance the accessibility, timeliness, and utility of DPH tools (including digital clinical tools) is easier to achieve internally and will build in the desired system responsivity.

Participants also highlighted concerns about the sustainability of DPH transformation processes, including the infrequency of comprehensive and full-cycle planning that considers evaluation and knowledge translation for DPH interventions. Insufficient evaluation was said to limit the organization's ability to demonstrate impact and secure ongoing funding for DPH interventions. For example, participants emphasized the importance of rigorous evaluation of DPH interventions

> *That evaluation component I think is really, really important and something our team really wants to try and get sorted. But again, capacity is a huge component of that for us, so.* [KT, Session 2]

> *Whatever we decide to move forward with, we want to have a fairly robust evaluation strategy built into that to make sure it's gonna be effective and be of use for the target population.* [Immunization program and vaccine-preventable diseases, Session 2]

Addressing these issues requires a shift toward a more structured and comprehensive DPH transformation approach within the organization, integrating systematic DPH planning, program design with clear goals, effective implementation, evaluation, and knowledge translation. Moreover, establishing rigorous evaluation frameworks is essential to generating evidence on the efficacy and effectiveness (including cost-effectiveness) of DPH interventions, potentially enhancing their long-term success and sustainability. Participants considered the importance of interdisciplinary academic and industry partnerships in addressing these 'full cycle' approaches to DPH interventions

> *There were folks at the University of Waterloo, and their computer sciences and health management system. And they wanted to work with us. But we got busy in the pandemic. Similarly, here at UBC and SFU. So, there are so many things–like they could develop apps, they could come up with completely new ideas because we don't usually work in that environment. And our sphere of thought processes revolves mainly around epidemiology and statistics. And we are not engaging in computer science. And some of the digital people who work on some of these solutions from their point of view and if you bring it that kind of like cross-disciplinary approach, we may be able to come up with the solutions which are a little bit better suited than what we have right now.* [DAS, Session 2]

Discussion

Principal Findings

In this qualitative descriptive study of practitioners' perspectives on the challenges and opportunities of DPH, we identified 3 interconnected themes. These align with broader digital health strategies and reflect constructs from organizational change theory, including change valence, change efficacy, organizational norms and procedures, and perceptions of contextual resources [23,31].

First, practitioners' expressed narratives about bridging existing inequities - an opportunity and a challenge contingent on public trust. This reflected practitioners' optimism about the potential for DPH to advance equity, effectiveness, and reconciliation while being concerned about its capacity to widen the digital divide. This emphasized building and maintaining public trust as a foundational element of DPH. Second, practitioners reported a sense of disconnect between "digital" and public health functions. DPH was often viewed as outside the traditional scope of public health work, reflecting concerns about limited digital competencies and fragmented systems. Practitioners described needing centralized supports, targeted in-service training, and restructured human resource systems aligned with emerging DPH needs. Third, practitioners' discourse regarding balancing the need for responsive DPH with necessary reactivity described challenges with overly reactive and somewhat ad-hoc approaches to digital transformation across the organization. Digital initiatives were often issue-specific and fragmented. These accounts reflect a reactive organizational culture, underscoring needed shifts toward more proactive, integrated approaches. Practitioners emphasized needing improved internal and external coordination within the DPH ecosystems. Effective

XSL•FO
RenderX

communication, interoperability, and rigorous evaluation were identified as essential components of a sustainable DPH system.

Our findings contribute to the limited but growing body of literature on organizational transformation in public health settings [22,32]. While digital health implementation is often studied in clinical contexts, few studies have examined how public health organizations navigate structural, cultural, and operational shifts in the context of digital transformation [10]. Echoing findings from clinical settings, practitioners in this study expressed optimism about the potential of digital tools to improve access, coordination, and outcomes [32]. However, they also raised concerns about the absence of coherent organizational strategies, limited digital infrastructure, and persistent gaps in workforce capacity [33-37]. Our findings reinforce the need for digital transformation efforts to be accompanied by robust organizational capacity-building strategies. Gaps in digital competencies, which range from basic digital literacy to specialized technical skills, undermine practitioners' ability to engage with and sustain digital initiatives [33,37,38]. Our findings also align with studies demonstrating public trust and engagement as a core part of digital transformation strategies and can be facilitated through proactive communication and transparent partnerships [33,35,37,38].

This study also illustrates tensions public health practitioners must navigate while facilitating the digital transformation of public health within broader societal digital transformation. Practitioner perspectives reflect a pre-existing institutional focus on equity and reconciliation [18]. However, perceived tensions emphasize calls for a techno-realistic equity-focused approach to digital transformation [9,39]. Further, practitioners' sense of disconnect between DPH and core public health functions is explained by expressed concerns about change efficacy for transformation [23]. These concerns exist at both individual and system levels, echoing similar concerns identified in our prior research [23,38,40]. Change efficacy concerns also demonstrate potential misalignments between the prevailing organizational culture and a view of DPH as integral to core public health functions [23,31,37].

This misalignment may be reinforced by current issue-focused and ad-hoc approaches to deploying digital tools to address specific short-term public health concerns [37,38]. Other studies suggest this approach, combined with organizational bureaucracies, reinforces the apparent fragmentation of digital systems and limits the potential for digital transformation of public health [31,37,41]. Within this context, incentives for privacy and infrastructure workarounds that are dependent on an enthusiastic few continue to complicate perceived fragmentation [22,37,41]. These findings demonstrate the role of proactive leadership capable of shaping organizational norms, culture, and processes to support long-term digital transformation, while also navigating short-term operational demands [23,31]. Such efforts must support practitioners to see DPH as integral to long-term visions, rather than peripheral to their mandate.

Implications for Public Health Practice and Research

Our study has implications for organizational DPH strategies. Given existing recommendations for people-centered digital health strategies guided by integrated leadership, financial, organizational, human, and technological resources, it is reassuring that our findings are well aligned with British Columbia's broader digital health strategy which centers trust and equity [11,42,43]. Building on this will require better governance and collaboration among partners and teams across the organization and served communities to not only build trusted integrated systems but also prevent duplication of efforts. These efforts must occur alongside ongoing advocacy for equitable digital access and literacy [44,45]. Moreover, our findings of a disconnect emphasize the need for an integrated review of public health training programs and their curricula to more appropriately reflect the role of digital transformation in contemporary public health, starting with basic general digital competencies with opportunities for specialization [46]. Our findings emphasize the need for a cohesive strategy to train public health practitioners across the organization on an ongoing basis to ensure they are up to date with modern digital technologies. Our findings also highlight the need for a revamp of current hiring practices with updated job descriptions (especially in a transdisciplinary context) and new public health roles for digital specialists (eg, highly skilled data scientists, computer scientists, and artificial intelligence and machine learning experts) with clearer career progression plans and competitive compensation packages [37,42].

Finding the need to balance responsivity with necessary reactivity highlights longstanding tensions in public health that significantly affected the public health response during the COVID-19 pandemic [8,41]. While the pan-Canadian and the British Columbia digital health strategies seek to implement people-centered interoperable systems, our study suggests adapting and extending this same vision across the organization to include all digital systems and supports [41,42]. Such an approach will require a centralized DPH program with digital specialists to support effective communication across the organization, maintain a repository of digital interventions and identify opportunities for shared costs, support ongoing practitioner training, support surveillance and attention to the growing risk of digital determinants of health, support rigorous evaluation of DPH interventions and ensure governance of these systems [37]. However, we must be careful not to allow these resources to constitute barriers to innovation while helping to transition the organization from reactive to more proactive and people-centered DPH systems [42]. Effective implementation will require not only technical resources but also strong organizational readiness, including alignment between leadership, workforce capacity, and operational structures that can sustain digital transformation over time [23].

Findings from this study should be considered based on its strengths and limitations. First, this study uniquely focuses on practitioner perspectives in real-world DPH contexts. Much of the literature has focused on mainly clinical and health systems. Our study explores the realities of public agencies that are reliant on internal and external process and data flows necessary for public health action across varied jurisdictions. It also accounts for the growing role of digital determinants of health. Our study comprehensively explores experiences of practitioners within a provincial public health organization, reflecting perspectives

XSL•FO
RenderX

from similarly sized organizations which support comparable mandates. However, care must be taken to transfer findings from this study to other jurisdictions, given the uniqueness of the organization (ie, a provincial role) and the health systems within which the organization is located. We have also not accounted for community and partner perspectives in this study and our recommendations. Further research is needed to appropriately capture these perspectives to inform a comprehensive DPH strategy. Research is also required to understand the models through which recommendations for centralized supports, ongoing training, and systems revisions will become operable.

Conclusions

While practitioners share excitement about the opportunities for DPH to facilitate equitable, efficient, and reconciliatory public health action, there are shared concerns about DPH being out of scope of core public health work, digital competency gaps, and a reactive approach to digital transformation within the organization that results in fragmented systems that limit DPH's potential impacts. Organizational DPH strategies are needed to support practitioners to achieve equity through DPH by providing a cohesive, systematic, and proactive approach to digital transformation through clear and careful communication with partners and communities, updated human resource and in-service training policies that support the organization with engaging and retaining the right digital talent, and creating centralized supports to ensure a more responsive approach to DPH.

Acknowledgments

The authors would like to thank public health practitioners of the British Columbia Centre for Disease Control who contributed their time and experiences towards this project. ChatGPT was used in copyediting the manuscript. All original content was generated by the authors and outputs were reviewed by the authors. ChatGPT had no role in study design, data collection, and analysis. The authors received no specific funding for this project. II is supported by the Canadian Institutes for Health Research (CIHR) Health Systems Impact Fellowship (Award number 521442), Michael Smith Health Research BC Trainee Award (Award number – HSIF-2024- - 04465) and the CIHR Canadian HIV Trials Network (CTN+) post-doctoral fellowship. MG is supported by the CIHR Applied Public Health Chair Award [47].

Data Availability

The datasets generated or analyzed during this study are not publicly available due because they were generated as part of organizational improvement efforts, but are available from the corresponding author on reasonable request.

Ihoghosa Iyamu: Conceptualization; Methodology; Validation; Formal Analysis – Review; Writing – Original Draft;
Devon Haag: Conceptualization; Project Administration; Funding Acquisition; Resources; Writing – Review & Editing;
Anna Carson: Formal Analysis; Writing – Review & Editing;
Ivy Wang: Data Curation; Project Administration; Formal Analysis – Participation; Writing – Review & Editing;
Colin King: Data Curation; Formal Analysis – Participation; Writing – Review & Editing;
Ian Roe: Conceptualization – Support; Data Interpretation; Writing – Review & Editing;
Kristy Kerr: Conceptualization – Support; Data Interpretation; Writing – Review & Editing;
Sofia Bartlett: Data Interpretation; Writing – Review & Editing;
Geoffrey McKee: Conceptualization – Support; Data Interpretation; Writing – Review & Editing;
Mark Gilbert: Investigation; Supervision; Data Interpretation; Writing – Review & Editing.

Conflicts of Interest

II serves on the board of the Public Health Association of British Columbia. II, DH, IR, GM and MG are affiliated with the BCCDC – the subject of this project. II, DH, MG support online sexual health services at the BCCDC. No other conflicts declared.

Multimedia Appendix 1
Focus group discussion guide.
[DOCX File, 24 KB - publichealth_v11i1e72588_app1.docx]

Multimedia Appendix 2
Codebook.
[DOCX File, 20 KB - publichealth_v11i1e72588_app2.docx]

Checklist 1
COREQ (COnsolidated criteria for REporting Qualitative research) Checklist.
[DOCX File, 19 KB - publichealth_v11i1e72588_app3.docx]

XSL•FO
RenderX

References

1. Budd J, Miller BS, Manning EM, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
2. Gunasekeran DV, Tseng R, Tham YC, Wong TY. Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. NPJ Digit Med 2021 Feb 26;4(1):40. [doi: 10.1038/s41746-021-00412-9] [Medline: 33637833]
3. Gunasekeran DV, Tham YC, Ting DSW, Tan GSW, Wong TY. Digital health during COVID-19: lessons from operationalising new models of care in ophthalmology. Lancet Digit Health 2021 Feb;3(2):e124-e134. [doi: 10.1016/S2589-7500(20)30287-9] [Medline: 33509383]
4. Dion M, AbdelMalik P, Mawudeku A. Big data and the global public health intelligence network (GPHIN). Can Commun Dis Rep 2015 Sep 3;41(9):209-214. [doi: 10.14745/ccdr.v41i09a02] [Medline: 29769954]
5. Al-Dmour H, Masa'deh R, Salman A, Abuhashesh M, Al-Dmour R. Influence of social media platforms on public health protection against the COVID-19 pandemic via the mediating effects of public health awareness and behavioral changes: integrated model. J Med Internet Res 2020 Aug 19;22(8):e19996. [doi: 10.2196/19996] [Medline: 32750004]
6. Amiri P, Karahanna E. Chatbot use cases in the Covid-19 public health response. J Am Med Inform Assoc 2022 Apr 13;29(5):1000-1010. [doi: 10.1093/jamia/ocac014] [Medline: 35137107]
7. Kickbusch I, Holly L. Addressing the digital determinants of health: health promotion must lead the charge. Health Promot Int 2023 Jun 1;38(3):daad059. [doi: 10.1093/heapro/daad059] [Medline: 37264549]
8. Public Health Agency of Canada. The chief public health officer of canada's report on the state of public health in canada 2021: a vision to transform canada's public health system. : Ottawa, Canada: Public Health Agency of Canada; 2021 URL: https://www.canada.ca/en/public-health/corporate/publications/chief-public-health-officer-reports-state-public-health-canada/state-public-health-canada-2021.html [accessed 2025-07-21]
9. Gómez-Ramírez O, Iyamu I, Ablona A, et al. On the imperative of thinking through the ethical, health equity, and social justice possibilities and limits of digital technologies in public health. Can J Public Health 2021 Jun;112(3):412-416. [doi: 10.17269/s41997-021-00487-7] [Medline: 33725332]
10. Iyamu I, McKee G, Haag D, Gilbert M. Defining the role of digital public health in the evolving digital health landscape: policy and practice implications in Canada. Health Promot Chronic Dis Prev Can 2024 Feb;44(2):66-69. [doi: 10.24095/hpcdp.44.2.04] [Medline: 38353941]
11. BC Ministry of Health. BC digital health strategy: a digitally enabled system trusted by all who use it. : Victoria, BC: BC Ministry of Health; 2024. URL: https://www2.gov.bc.ca/assets/gov/health/about-bc-s-health-care-system/health-priorites/bc-digital-health-strategy/bcdhs.pdf [accessed 2024-06-20]
12. Kickbusch I, Piselli D, Agrawal A, et al. The Lancet and Financial Times Commission on governing health futures 2030: growing up in a digital world. The Lancet 2021 Nov;398(10312):1727-1776. [doi: 10.1016/S0140-6736(21)01824-9]
13. Fung KK, Hung SSL, Lai DWL, Shum MHY, Fung HW, He L. Access to information and communication technology, digital skills, and perceived well-being among older adults in Hong Kong. Int J Environ Res Public Health 2023 Jun 23;20(13):6208. [doi: 10.3390/ijerph20136208] [Medline: 37444058]
14. Sina E, Boakye D, Christianson L, Ahrens W, Hebestreit A. Social media and children's and adolescents' diets: a systematic review of the underlying social and physiological mechanisms. Adv Nutr 2022 Jun 1;13(3):913-937. [doi: 10.1093/advances/nmac018] [Medline: 35218190]
15. Jahnel T, Dassow HH, Gerhardus A, Schüz B. The digital rainbow: digital determinants of health inequities. Digit Health 2022;8:20552076221129093. [doi: 10.1177/20552076221129093] [Medline: 36204706]
16. van Kessel R, Wong BLH, Clemens T, Brand H. Digital health literacy as a super determinant of health: More than simply the sum of its parts. Internet Interv 2022 Mar;27:100500. [doi: 10.1016/j.invent.2022.100500] [Medline: 35242586]
17. van Kessel R, Seghers LE, Anderson M, et al. A scoping review and expert consensus on digital determinants of health. Bull World Health Organ 2025 Feb 1;103(2):110-125H. [doi: 10.2471/BLT.24.292057] [Medline: 39882497]
18. BC Ministry of Health. British columbia's population and public health framework: strengthening public health. : Victoria, BC: BC Ministry of Health; 2024 URL: https://www2.gov.bc.ca/assets/gov/health/about-bc-s-health-care-system/public-health/pph-framework/bc_population_and_public_health_framework.pdf [accessed 2025-07-21]
19. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
20. Morgenstern JD, Rosella LC, Daley MJ, Goel V, Schünemann HJ, Piggott T. "AI's gonna have an impact on everything in society, so it has to have an impact on public health": a fundamental qualitative descriptive study of the implications of artificial intelligence for public health. BMC Public Health 2021 Jan 6;21(1):40. [doi: 10.1186/s12889-020-10030-x] [Medline: 33407254]
21. Kaushik V, Walsh CA. Pragmatism as a research paradigm and its implications for social work research. Soc Sci (Basel) 2019;8(9):255. [doi: 10.3390/socsci8090255]
22. Iyamu I, Xu AXT, Gómez-Ramírez O, et al. Defining digital public health and the role of digitization, digitalization, and digital transformation: scoping review. JMIR Public Health Surveill 2021 Nov 26;7(11):e30399. [doi: 10.2196/30399] [Medline: 34842555]

XSL•FO
RenderX

23. Weiner BJ. A theory of organizational readiness for change. Implement Sci 2009 Oct 19;4:67. [doi: 10.1186/1748-5908-4-67] [Medline: 19840381]
24. Gilbert M, Haag D, Hottes TS, et al. Get checked… where? the development of a comprehensive, integrated internet-based testing program for sexually transmitted and blood-borne infections in British Columbia, Canada. JMIR Res Protoc 2016 Sep 20;5(3):e186. [doi: 10.2196/resprot.6293] [Medline: 27649716]
25. Provincial Health Services Authority. PHSA Project Sorting Tool. 2024. URL: https://rc.bcchr.ca/redcap/surveys/?s=HNWAAKFF97 [accessed 2025-01-14]
26. Government of Canada IAP on RE. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2. 2023. URL: https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2022.html [accessed 2024-05-05]
27. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021 Jul 3;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
28. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
29. QSR International Pty Ltd. NVivo qualitative data analysis software (version 14). 2020 URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home
30. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
31. Batras D, Duff C, Smith BJ. Organizational change theory: implications for health promotion practice. Health Promot Int 2016 Mar;31(1):231-241. [doi: 10.1093/heapro/dau098] [Medline: 25398838]
32. Odone A, Buttigieg S, Ricciardi W, Azzopardi-Muscat N, Staines A. Public health digitalization in Europe. Eur J Public Health 2019 Oct 1;29(Supplement_3):28-35. [doi: 10.1093/eurpub/ckz161] [Medline: 31738441]
33. Wong BLH, Maaß L, Vodden A, et al. The dawn of digital public health in Europe: Implications for public health policy and practice. Lancet Reg Health Eur 2022 Mar;14:100316. [doi: 10.1016/j.lanepe.2022.100316] [Medline: 35132399]
34. Hickey G, Dunne C, Maguire L, McCarthy N. An exploration of practitioners' experiences of delivering digital social care interventions to children and families during the COVID-19 pandemic: mixed methods study. JMIR Form Res 2023 Apr 28;7:e43498. [doi: 10.2196/43498] [Medline: 36888555]
35. Turnbull SL, Dack C, Lei J, et al. Barriers and facilitators to use of digital health tools by healthcare practitioners and their patients, before and during the COVID-19 pandemic: a multimethods study. BMJ Open 2024 Mar 5;14(3):e080055. [doi: 10.1136/bmjopen-2023-080055] [Medline: 38448080]
36. Roy I, Salles J, Neveu E, et al. Exploring the perspectives of health care professionals on digital health technologies in pediatric care and rehabilitation. J Neuroeng Rehabil 2024 Sep 12;21(1):156. [doi: 10.1186/s12984-024-01431-9] [Medline: 39261920]
37. Iyamu I, Gómez-Ramírez O, Xu AX, et al. Challenges in the development of digital public health interventions and mapped solutions: Findings from a scoping review. Digit Health 2022;8:20552076221102255. [doi: 10.1177/20552076221102255] [Medline: 35656283]
38. Iyamu I, Ramachandran S, Chang HJ, et al. Considerations for adapting digital competencies and training approaches to the public health workforce: an interpretive description of practitioners' perspectives in Canada. BMC Public Health 2025 Jan 10;25(1):122. [doi: 10.1186/s12889-024-21089-1] [Medline: 39794767]
39. Rodriguez JA, Lyles CR. Strengthening digital health equity by balancing techno-optimism and techno-skepticism through implementation science. NPJ Digit Med 2023 Nov 2;6(1):203. [doi: 10.1038/s41746-023-00954-0] [Medline: 37919418]
40. Ramachandran S, Chang HJ, Worthington C, et al. Digital competencies and training approaches to enhance the capacity of practitioners to support the digital transformation of public health: rapid review of current recommendations. JMIR Public Health Surveill 2024 Sep 9;10:e52798. [doi: 10.2196/52798] [Medline: 39248660]
41. Expert Advisory Group. Pan-canadian health data strategy: towards a world-class health data system. : Ottawa, Canada: Public Health Agency of Canada; 2022 URL: https://www.canada.ca/en/public-health/corporate/mandate/about-agency/external-advisory-bodies/list/pan-canadian-health-data-strategy-reports-summaries/expert-advisory-group-report-03-toward-world-class-health-data-system.html [accessed 2025-07-21]
42. World Health Organization. Global strategy on digital health 2020-2025. : Geneva: World Health Organization; 2021 URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2025-07-21]
43. US Centers for Disease Control and Prevention. CDC global digital health strategy. 2022 URL: https://www.cdc.gov/global-health/topics-programs/global-digital-health-strategy.html [accessed 2025-07-21]
44. Azzopardi-Muscat N, Sørensen K. Towards an equitable digital public health era: promoting equity through a health literacy perspective. Eur J Public Health 2019 Oct 1;29(Supplement_3):13-17. [doi: 10.1093/eurpub/ckz166] [Medline: 31738443]
45. Lyles CR, Wachter RM, Sarkar U. Focusing on Digital Health Equity. JAMA 2021 Nov 9;326(18):1795-1796. [doi: 10.1001/jama.2021.18459] [Medline: 34677577]
46. Iyamu I, Ramachandran S, Chang HJ, et al. Considerations for adapting digital competencies and training approaches to the public health workforce: an interpretive description of practitioners' perspectives in Canada. . 2024. [doi: 10.21203/rs.3.rs-5320781/v1]

XSL•FO
RenderX

47. Canadian institutesof health research. 2022 Applied Public Health Chairs. URL: https://cihr-irsc.gc.ca/e/53651.html [accessed 2025-08-11]

Abbreviations

BCCDC: BC Centre for Disease Control
COREQ: Consolidated criteria for reporting qualitative research
DAS: data and analytics services
DPH: digital public health
EHS: environmental health services
FGs: focus groups
KT: knowledge translation
PHR: public health response

Edited by A Mavragani, T Sanchez; submitted 12.02.25; peer-reviewed by I Wilson, K Elechi; revised version received 09.05.25; accepted 19.05.25; published 12.08.25.

Please cite as:
Iyamu I, Haag D, Carson A, Wang I, King C, Roe I, Kerr K, Bartlett S, McKee G, Gilbert M
Opportunities and Challenges for an Organizational Digital Public Health Strategy in a Provincial Public Health Program in Canada: Qualitative Description of Practitioner Perspectives
JMIR Public Health Surveill 2025;11:e72588
URL: https://publichealth.jmir.org/2025/1/e72588
doi:10.2196/72588

© Ihoghosa Iyamu, Devon Haag, Anna Carson, Ivy Wang, Colin King, Ian Roe, Kristy Kerr, Sofia Bartlett, Geoffrey McKee, Mark Gilbert. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Building the Workforce's Capacity to Support the Digital Transformation of Public Health: Environmental Scan of Training Programs for Digital Technologies in Public Health

Ihoghosa Iyamu[1,2], MD, MDICHA, PhD; Swathi Ramachandran[2], BDS, MSc; Hsiu-Ju Chang[2], MSc; Andre Kushniruk[3], BSc, BA, MSc, PhD; Francisco Ibáñez-Carrasco[4], PhD; Catherine Worthington[5], PhD; Hugh Davies[1], PhD; Geoffrey McKee[1], MD, MPH; Adalsteinn Brown[4], DPhil; Mark Gilbert[1,2], MHSc, MD

[1]School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

[2]BC Centre for Disease Control, 655 West 12th Avenue, Vancouver, BC, Canada

[3]School of Health Information Science, University of Victoria, Victoria, BC, Canada

[4]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[5]School of Public Health and Social Policy, University of Victoria, Victoria, BC, Canada

Corresponding Author:
Mark Gilbert, MHSc, MD
School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

Abstract

Background: The digital transformation of society and public health has created an urgent need for new competencies to address evolving and contemporary public health challenges. While some public health institutions and schools worldwide have begun responding through various training programs and approaches, many have yet to do so. A clearer understanding of the current training landscape can inform more coordinated efforts to update curricula and strengthen digital competency within the public health workforce.

Objective: This study aimed to map and describe existing digital public health (DPH) training programs, identifying common curricula content, disciplinary involvement, and training approaches. It also aimed to identify gaps and opportunities for curricular adaptation.

Methods: This environmental scan was conducted in 2 stages, drawing on guidance from studies by Rowel et al and Wilburn et al. First, we performed a systematic search of DPH training programs, followed by interviews with selected program directors to explore their program design and implementation. The scan emphasized a transdisciplinary lens, consistent with the evolving nature of DPH. Between March and May 2023, we searched Google and public health association directories to identify degree programs and courses (as part of degree awarding programs) focused on building capacity for using digital technologies in public health. We then conducted semi-structured interviews with 4 directors of identified programs exploring program characteristics and the inter- or transdisciplinary partnerships essential to their design. Search data were summarized using narrative synthesis, while content analysis was applied to the interview data.

Results: Overall, 58 DPH training programs were identified, categorized into 3 groups: public health data science (29/58, 50%), public health informatics (16/58, 28%), and a mix of programs exploring digital competencies (13/58, 22%) related to project management and addressing the digital determinants of health. Interviews revealed that motivation for developing interdisciplinary DPH programs stemmed from the need to meet evolving job market demands and respond to calls for curricular renewal among professional bodies. Effective design and delivery were supported by academic–industry partnerships, which aimed to cultivate professionals with depth in public health and breadth in digital competencies. These programs drew on diverse disciplinary perspectives from academia, the public sector, and private industry. However, sustaining such partnerships was challenged by the need to negotiate shared priorities, reconcile differing viewpoints, and secure ongoing funding.

Conclusions: This global scan of DPH training programs found a strong focus on data-centric competencies, with less emphasis on digital skills for health promotion, leadership, and addressing digital determinants of health. Bridging these gaps requires a stepwise approach: integrating digital competencies into existing curricula, offering stand-alone programs for specialized skills, and strengthening partnerships to navigate funding and administrative barriers while promoting equity-driven, interdisciplinary collaboration.

(JMIR Public Health Surveill 2025;11:e73088) doi:10.2196/73088

XSL•FO
RenderX

KEYWORDS

digital health technology; public health schools; professional competence; public health education for professionals; digital health; competency-based education; health workforce

Introduction

The digital transformation of society, including public health, was significantly accelerated during the COVID-19 pandemic [1-3]. Digital technologies have been adopted to enhance the efficiency, effectiveness, and reach of essential public health functions, such as real-time precision disease surveillance using big data and artificial intelligence, targeted health promotion via social media, apps, and wearables, and environmental health protection informed by sensor data [2-5]. Digital access and literacy have recently emerged as important determinants of health, directly and indirectly shaping public health outcomes [6-8]. For example, limited digital access and literacy may restrict social and economic opportunities, which in turn limit access to digitally mediated health and social services [8,9]. Conversely, for those with digital access, exposure to misinformation, disinformation, and exploitative practices by big tech platforms can adversely impact health and health behaviors [8,10,11].

To harness these opportunities and address the evolving challenges, the contemporary public health workforce must develop new digital competencies [12,13]. In our previous research, we conducted a rapid literature review and identified 45 unique digital competency statements that cut across all 7 categories of the 2008 Core Competencies for Public Health in Canada framework [12]. We also identified a potentially new competency category related to digital data, data systems management, and governance [12]. While these competencies are not exhaustive, especially with rapidly developing digital technologies, they represent aspirations of public health practitioners and educators to address existing competency and curricular gaps [14,15]. For instance, the Association of Schools for Public Health in the European Region (ASPHER) task force for digital transformation has recommended updates to the public health curriculum across themes such as digital literacy, health data collection and analysis, health data management and governance, ethics and regulation of digital transformations in society, and understanding the infosphere and spread of information over digital networks [16]. They emphasize the need for interdisciplinary and transdisciplinary academic communities that understand public health, as well as computer science, data science, and other sciences relevant for digital public health (DPH) [16,17]. Ongoing efforts by the Public Health Agency of Canada, including commissioning the National Collaborating Centers for Public Health to update the Core Competencies for Public Health in Canada, are also examples of work to address these competency needs [18].

It remains unclear how these competency aspirations and curricular recommendations have been implemented in public health training programs and courses. Gaining insight into how these competencies have shaped curricular updates can guide targeted recommendations for digital competency development and curriculum design in Canada. This would support the broader vision of transforming public health and enable schools of public health to make necessary adaptations [13]. In addition, with well-documented challenges in inter- and transdisciplinary practice, there is limited understanding of the disciplinary perspectives incorporated into these programs and how educators navigate diverse viewpoints in practice [17,19].

Therefore, our main objective was to map, explore, and describe training programs designed to integrate and enhance the use of digital technologies specifically for public health and its subspecialties around the world. Specifically, we examined common curricular content, the academic disciplines involved in program design and delivery, and the training approaches used to implement these curricula. We also identified opportunities for curriculum adaptation and strategic pivots in response to emerging needs and gaps in current DPH curricula and training programs. This work is part of a broader study supporting the digital transformation of public health by informing competency development, encouraging interdisciplinary collaboration, and guiding future curriculum planning [12,20]. Ultimately, this work seeks to better prepare public health practitioners to engage with and lead within an increasingly digital health landscape.

Methods

Study Design

The study was an environmental scan conducted following a 2-stage approach: a systematic search for DPH programs aligned with the study objectives and semistructured interviews, with program directors or faculty members from programs emphasizing inter- or transdisciplinary perspectives in their curriculum planning and delivery [21,22].

Systematic Search for DPH Program Websites

Two members of the research team (II and SR) conducted independent searches using Google's advanced search (google.com/advanced_search) without country restrictions. Search terms included ("digital public health" OR "digital health" OR "digitalization" OR "digital transformation" OR "data science" OR "social media" OR "artificial intelligence" OR "machine learning") AND ("school of public health" OR "college of public health" OR "department of public health" OR "division of public health" OR "institute of public health") AND (degree OR program OR course). Thereafter, the search was repeated, restricting returns to countries identified as prominent contributors to DPH literature, including Australia, Canada, China, France, Germany, Italy, Sweden, the United Kingdom, and the United States of America (Multimedia Appendix 1) [1,17]. We reviewed the first 100 returns for each search. Thereafter, we searched popular global online directories of schools of public health, including the Council on Education for Public Health, ASPHER, and the Public Health Agency of Canada, to identify potential programs or courses for review. All relevant returns were reviewed against the inclusion and

XSL•FO
RenderX

exclusion criteria (Table 1). All searches were conducted between March 24, 2023, and May 16, 2023.

Table . Program inclusion and exclusion —environmental scan of digital public health training programs (2023).

Domain	Inclusion criteria	Exclusion criteria
Population	Focus on public health students, public health informaticians and analysts, practitioners (including public health physicians), public health researchers, and decision- and policy-makers.	Exclusive focus on clinicians, programmers, or other related technologists, informaticians, or other allied health system practitioners.
Concept	Programs focused on building capacity for designing, implementing, and evaluating digital technologies in any of the domains of public health or across public health.	Programs focused on the broader field of digital health or a mostly clinical perspective (eg, telemedicine or virtual health in a clinical context).
Context of training program (ie, context of application)	Population and public health	Clinical or solely clinical contexts
Program type	Established degree awarding training programs at undergraduate and graduate levels or courses undertaken as a requirement for a degree awarding public health program.	Short courses, workshops, webinars, and courses not related to a public health program.
Timeline	Must be currently available for enrollment	Previously available programs that have been discontinued or truncated.

Program Selection

We adopted a definition of an eligible DPH training program as any undergraduate or graduate-level degree-awarding public health training programs that were currently available for enrollment and entirely focused on applying digital technologies to all or to specific public health domains as defined in the Canadian Public Health Association's conceptual framework for public health (Table 1) [23]. We also included course modules that are part of degree-granting public health programs. In addition, programs with a broader medical focus were included if they explicitly addressed public health, such as through modules relevant to public health as described within their overall objectives. We excluded programs and courses that are exclusively focused on clinical and clinical informatics perspectives and those offered as short courses (ie, stand-alone course modules not offered as part of a degree-awarding program). One member of the research team (SR) reviewed search returns based on our inclusion and exclusion criteria (Table 1), with another researcher (II) reviewing the list of included programs and courses for accuracy based on the inclusion and exclusion criteria.

Data Collection

We reviewed the websites for each included training program and course, identifying program documents and training curricula available on their websites. Data from these documents were extracted using pretested data extraction forms (Multimedia Appendix 2) adapted from the guidelines for reporting evidence-based practice educational interventions and teaching [24]. Extracted data included basic program information (eg, program name, country, city, and year of commencement) and program descriptors (eg, target audience, level of training, educational materials used, characteristics of the instructor(s), and underpinning theories ([ie, educational or field-specific theory(ies) guiding the program]). For program websites not originally in English, Google Translate was used to access the documents before extraction.

Semi-Structured Interviews

We reviewed online documents for each included program to identify evidence of a clear mandate for inter- and transdisciplinary training. This included mission statements, faculty from diverse disciplines, and explicitly stated inter- or transdisciplinary partnerships. For each qualifying program, we invited program directors or staff involved in program design, delivery, and evaluation to participate in semi-structured interviews. Additional eligibility criteria included being an established graduate- or undergraduate-level degree-awarding program with English-speaking staff. Recruitment emails were sent using contact details from program websites, and a convenience sample of staff who responded to invitations was recruited. Each eligible program received up to 3 reminders. Using a snowball sampling approach, we identified an additional program that was not on the original list, as it did not have a listed degree-awarding program on its website. However, their strong emphasis on transdisciplinary DPH training and traditional public health training focused on DPH research justified their inclusion in the interviews. Interviews were conducted with consenting participants over Zoom Online (Zoom Communications Inc.). We used a topic guide that explored characteristics of the DPH training programs, academic and practice disciplines involved in designing and delivering the program curricula, characteristics of these inter-/transdisciplinary collaborations, and lessons learned (Multimedia Appendix 3). The topic guide was developed based on initial findings from our rapid review of DPH competency recommendations, insights from previous DPH-related literature reviews, and consultations with education leads from schools of public health who served as co-investigators on the project [12,17]. All interviews were conducted by II, who is trained in qualitative health research methods, supported by SR who received training on the data collection protocol and made detailed field notes. All interviews were audio recorded and automatically transcribed using Zoom's auto-transcribe function. Transcripts were reviewed for accuracy and corrected as

XSL•FO
RenderX

appropriate. Data analyses were concurrently conducted alongside interviews. Each interview lasted on average 47 minutes (range 39 - 53 min).

Ethical Considerations

This study was reviewed by the University of British Columbia's Behavioral Research Ethics Board (ethics #H22-03153). We obtained written informed consent from each participant at least 24 hours before the interviews via email and assigned each participant an identification number, which was included in all interview transcripts and field notes (Multimedia Appendix 4). All study materials were stripped of personal identifiers before analyses. Participants did not receive an honorarium.

Data Analyses

We summarized data from the systematic search using descriptive statistics (frequencies and proportions) in Microsoft Excel and conducted a narrative synthesis. For the semistructured interviews, transcripts and field notes were imported into QSR NVivo version 1.3 for data management and analysis [25]. We conducted a content analysis exploring key descriptions of programs, inter/transdisciplinary collaborations, and lessons learned from implementing such collaborations within the program context. Our analyses followed an inductive process beginning with initial coding, categorization, and interpretation of findings [26,27]. Emergent codes and categories were reviewed with the research team consisting of public health educators and DPH experts whose views informed ongoing refinement and interpretation of codes and categories generated during analysis [27]. Throughout the analyses, II and SR made reflexive memos about our perspectives of the data [28]. These memos also informed the analysis and the final report draft.

Results

DPH Programs and Courses Identified From Systematic Search

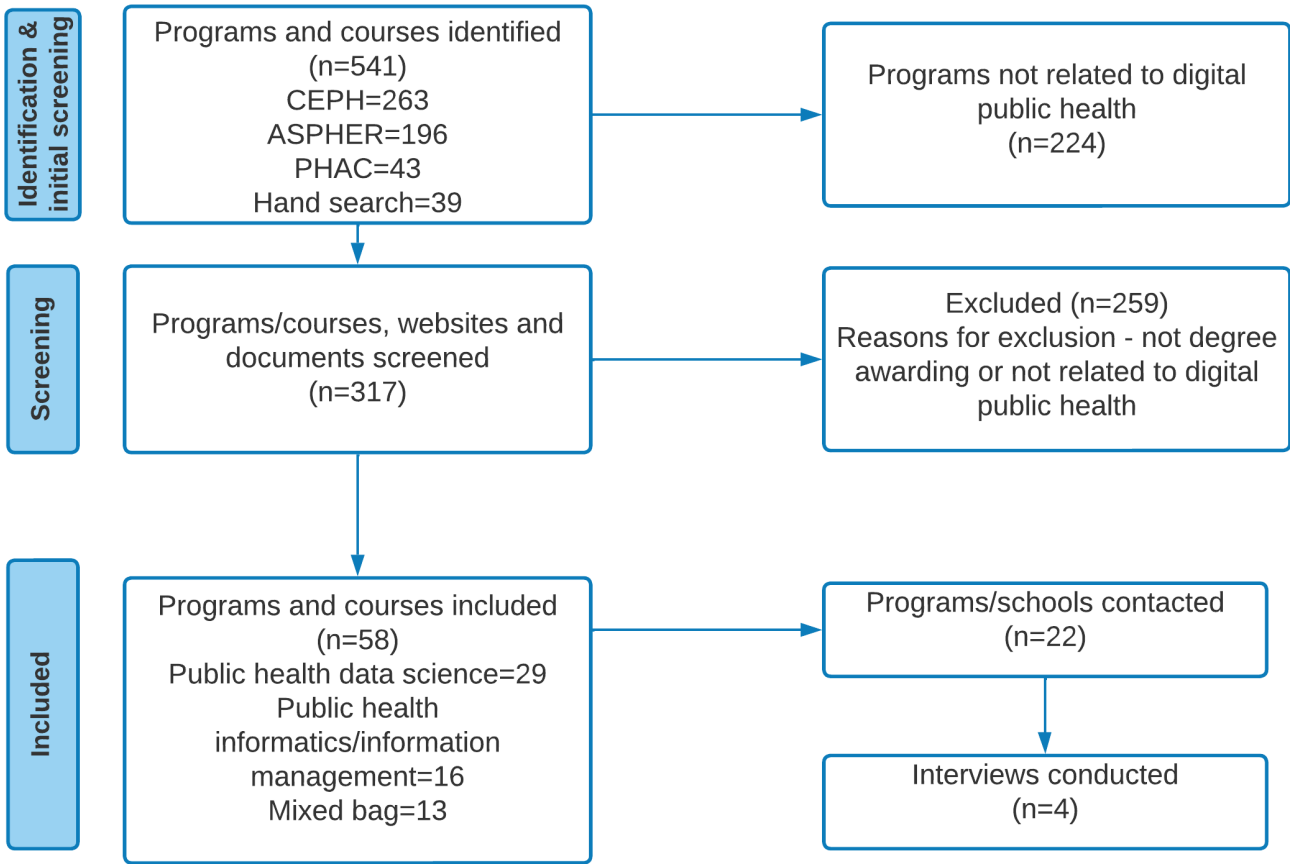
Overview of Included Programs

Overall, we identified 58 DPH training programs and courses offered as requirements for a degree awarding program applying digital technologies for public health functions (Figure 1. Of 58 programs, 35 (60.3%) were from universities and colleges in the United States (Table 2), 46 (79.3%) were at the master's degree level, 25 (43.1%) were delivered in-person (including options for applied components such as practicums and internships), and 47 (81%) did not explicitly describe theoretical foundations of their programs and courses (Multimedia Appendix 5). Regarding foundational theories underpinning programs, only 11 (18.9%) of 58 explicitly described being built on modern statistical theories, including probability and Bayesian frameworks.

Identified programs focused on 3 main DPH categories: public health data science (29/58, 50%) programs focused on using modern statistical and analytic approaches to harness digital data streams to provide evidence for public health decision making; public health informatics (16/58, 28%) programs focused on using information, computer science, and technology to create and manage the infrastructure and systems needed for digital data streams and analytic techniques; and a mix of programs and courses (13/58, 22%) focused on various digital competencies, including digital leadership, project management, digital health communication, digital transformation leadership, digital design and implementation, digital monitoring and evaluation, and tackling the digital determinants of health. There was significant overlap between programs across all 3 categories. However, the first 2 focused on a previously identified competency category related to digital data, data systems management, and governance.

XSL•FO
RenderX

Figure 1. Flow diagram of the program search and selection process.



XSL•FO
RenderX

Table . Characteristics of included DPH training programs and mandatory courses—environmental scan (2023).

Program characteristics	Frequency (N=58), n (%)
Country of program	
USA	35 (60.3)
Canada	6 (0.3)
UK	4 (6.9)
Australia	2 (3.4)
Germany	2 (3.4)
India	2 (3.4)
Others	7 (2.1)
Level of training	
Undergraduate	4 (6.9)
Master's	46 (79.3)
PhD (Doctoral level)	1 (1.7)
Combined Master's and Doctoral training	4 (6.9)
Others	3 (5.2)
Foundational theories	
Not stated	47 (81.0)
Biostatistics/statistics, probability, or related theories	10 (17.2)
People centeredness[a]	1 (1.7)
Partnerships reported (ie, public and private sector partners)	
Not stated	51 (87.9)
Industry partners (eg, health departments and health analytics consulting groups)	4 (6.9)
Health industry and university partnerships	3 (5.2)
Delivery	
Not stated	12 (20.7)
Online[b]	9 (15.5)
Blended	12 (20.7)
In person	25 (43.1)

[a]Emphasizes person-centredness as an organizing framework, in contrast to traditional informatics approaches.

[b]Includes online-delivered programs that incorporate in-person practicums or internships.

Curricula Content and Training Approaches

Identified programs had diverse learning objectives, foundational components, and emphasized various public health domains and functions. The majority of the public health data science programs and courses explored subjects related to public health sciences, biostatistics, computer science, and variants of these core content areas as presented as data science training. Most of these programs focused on public health and health services research, epidemiology, disease surveillance, and prevention. Public health informatics programs leveraged information science, computer science, data management, and health management information systems as applied to public health generally. The mixed bag of programs and courses leveraged varied digital content applied in health promotion, program development and evaluation, epidemiology, disease surveillance, and public health more generally. The majority of the programs (33/58) were delivered online, while others used blended learning (12/58) or in-person training (12/58). Most of the in-person training involved a blend of coursework and internships or practicums in real-world settings. The majority of these were focused on public health data science and public health informatics. Importantly, courses within the mixed-bag category leveraged industry–academic partnerships to facilitate content delivery as appropriate.

Disciplines Involved in Training Programs

While most programs highlighted their multi-, inter-, or transdisciplinary perspectives using broad statements, 17/58 programs (29.3%) did not clearly describe the disciplinary perspectives involved in the design and delivery of the training programs and courses. Where specified, programs predominantly

XSL•FO
RenderX

indicated specific disciplines, including public health sciences (including specific fields such as epidemiology and biostatistics). Beyond public health sciences, programs highlighted the involvement of statistics, computer science, information technology and informatics, health care administration, mathematics, life and health sciences, media and communication, health economics, law bioethics, psychology, anthropology, management, and so on (Multimedia Appendix 5).

Qualitative Interview Results—Considerations for Interdisciplinarity

Overview

A total of 4 program directors or faculty representing 4 of the included programs participated in interviews. Interviews were with program directors of programs in Canada, Australia, Germany, and the Netherlands. From the interviews, we identified 4 main qualitative categories that characterize interdisciplinary partnerships and offer insights into implementing transdisciplinary DPH programs in real-world settings. Each category is supported by exemplar quotes.

Motivation For Interdisciplinary DPH Programs

This reflected program directors' impetus for developing programs focused on interdisciplinarity. First, programs recognize the need to align their curricula and training with evolving digital trends in public health, responding to job market competency requirements and current course offering gaps based on previous research and professional experiences that emphasize the need for applied skills in inter- and transdisciplinary settings. Moreover, program directors' interdisciplinary training was also a key motivation for developing and implementing interdisciplinary programs that reflect their understanding of the current job market. For example, one participant said:

> *I am pure technologist by training, and being a digital health researcher, my research findings suggested that there is a need to have consumer education in public health... So, I wouldn't say that I'm an expert in public health that's not my area of expertise, but I've learned a lot in the past few years since I've been here...* [Participant 2]

Program directors emphasized the need to align their programs with curricular updates from professional bodies (such as ASPHER and the Australian Digital Health Agency) that have issued recommendations for digital competencies. They also noted that interdisciplinary programs were designed to address students' expressed learning needs, emphasizing the importance of experiences that closely reflect real-world working conditions.

Design and Delivery of Interdisciplinary Programs

Program directors primarily designed their programs with a focus on desired student outcomes. A common aspiration was to develop *T-shaped professionals*—individuals with broad knowledge across digital technology fields and deep expertise in at least one public health discipline. For most programs, their design was informed by active partnerships with industry practitioners, including private sector partners. For example, some program directors described program design supervised by an advisory board consisting of public and private industry partners that highlighted competencies needed in the workplace. Programs were structured around competencies aimed at fostering interdisciplinary collaboration, as well as practical understanding of digital regulatory frameworks, leadership in digital technology, design thinking, and advanced analytics.

While directors acknowledged the limited availability of competency frameworks in this area, they often described building curricula based on their vision for student outcomes. This involved inviting key partners to contribute and leveraging their various disciplinary expertise areas to align with the program's goals:

> *I think what really helped us work together was having this clear vision of you know this is who the ideal graduate is. This is what we want them to be able to do in the workplace, and then taking that vision and really breaking it apart, right? And then the data scientists could really say, okay, if I put myself in a position of having to work with this person, what would I want them to know? And I could put myself in the position of okay. If this is a person that's going to advise me on governance issues, you know, hospital policy, these kinds of things. What would I want them to tell me about technology, right? ... So, that allowed us to again retreat in our, you know, disciplinary silos. But at least work on the curriculum as a team, right? But again, it was from our perspectives...* [Participant 1]

Regarding training approaches, all interviewed programs described prioritizing experiential learning using a combination of foundational class-based courses and practicums or internships within public health organizations demonstrating interdisciplinary work. For instance, one of the public health data science programs is implemented with triads within the program, combining students of varied backgrounds (eg, public health sciences, computer science, and biostatistics) to solve real-world problems within health systems.

Characteristics of Interdisciplinary and Transdisciplinary Partnerships Supporting Identified Programs

Program directors suggested that interdisciplinary partnerships are built off a core group of leaders who can engage other disciplines based on the goals for the "T-shaped professional." Partnerships span multiple disciplines, including public health and medicine (eg, biostatistics, epidemiology, public health sciences, clinical medicine, psychology); applied sciences and technology (eg, computer science, data science, human–computer interaction, informatics, and cybersecurity); and the arts and humanities (eg, law, bioethics, media and communications, and anthropology). These partnerships essentially were prioritized based on the competencies for the envisioned T-shaped professional, starting within the internal academic environment first, extending to public and private sector partners as needed, including the formation of advisory boards as described. In addition, these partnerships are leveraged

XSL•FO
RenderX

for curricular content delivery to align student experiences with real-world experiences and expectations:

> *We have a what we called a stakeholder advisory board as well where we have some people from industry and we had different workshops where we invited people to present you know what they had been doing what they're forward or line of thinking is ...* [Participant 4]

Challenges With Implementing Interdisciplinary and Transdisciplinary Partnerships

Program directors highlighted a range of challenges associated with establishing transdisciplinary partnerships, as well as those stemming from administrative processes or the implementation of innovative programs within established schools. In terms of partnerships, challenges included negotiating shared commitment and perspectives among diverse experts with differing methodologies, views on knowledge, and levels of commitment to public health principles. These issues were particularly pronounced in academic–industry collaborations, especially with private entities. Concerns centered on the potential for noncritical, techno-optimistic implementation of technologies and the risk of exploitation within such partnerships. Participants emphasized the importance of having a clear understanding of the principles driving the operations of private entities. For example, one participant stated:

> *I mean, they're [private entities] obviously a big player in this because they are substantially driving the field if we think about all these health apps. So far, they are actually shaping some of the stuff that perhaps others should shape but it's just because they're doing it and they're not asking too much too many questions... They just want to produce products... Nevertheless, we would certainly continue also to, you know, aim to work with them. You find some yeah enlightened groups there to work with [and] not get this feeling of you know exploitation in some ways or going on their own in the wrong direction this is really important ...* [Participant 4]

Program directors also described difficulties in establishing a shared vocabulary and a common definition of "digital public health" among partners. Funding and resource challenges were frequently mentioned, with some programs relying heavily on grant funding to support their programs. These resource constraints influenced decisions regarding the depth versus breadth trade-offs in curriculum design and the need to compensate instructors for their time. In addition, resourcing challenges impacted decisions to limit the number of partnerships pursued for curriculum implementation.

Additional challenges included difficulties in gaining accreditation, particularly within traditional public health education systems that may undervalue the goal of developing T-shaped professionals. These challenges were addressed by securing leadership buy-in and institutional support. Program directors also emphasized the difficulty of creating curricula responsive to the rapid evolution of digital technology. To address this, they focused on fundamental principles such as systems thinking, privacy, and data governance, while incorporating applied training. This approach ensures students interact with current tools to better understand and apply these foundational principles.

Discussion

Principal Findings

In this environmental scan of DPH programs, we identified 58 programs and courses on digital competencies worldwide, the majority of which were in high-income countries. These initiatives addressed digital competencies across 3 main areas. The first 2 were related to public health data science and public health informatics, which build competencies for digital data, data systems management, and governance, consistent with our previous research showing this as a focus for digital competencies in public health [12]. Furthermore, we identified a third category of DPH programs and courses leveraging varied digital content applied in health promotion, program development and evaluation, epidemiology, disease surveillance, and public health. This category also included programs and courses aimed at understanding and addressing the digital determinants of health. For the interdisciplinary programs, most have aligned with external pressures from job market trends, student demands, and curricular recommendations for interdisciplinary practice [15,16]. While most programs emphasize applied learning within interdisciplinary frameworks to develop T-shaped professionals, challenges such as funding constraints and difficulties in achieving shared public health commitments and vocabulary remain significant barriers.

To the best of our knowledge, this is one of the first global scans of DPH programs and courses. Previous jurisdiction-specific scans, such as one in Germany, indicated that only about 20% of public health programs integrate digital competencies across their curricula [29]. While the German scan highlighted a strong focus on digital data analytics and informatics, it noted limited attention to social and equity perspectives in DPH—findings consistent with our scan. However, our global scan identified a few programs beginning to address these social perspectives, including understanding and tackling the digital determinants of health [29,30]. In Canada, while no specific environmental scans have been conducted, literature reviews emphasize the need for competency updates to support the deployment of artificial intelligence and other DPH interventions in public health [12,31]. Our study extends the literature by characterizing the nature of inter- and transdisciplinary partnerships, which are critical for designing and delivering such training programs [17,29,32]. The geographic distribution of programs and courses we identified shows a concentration in high-income countries, highlighting significant digital competency gaps in low- and middle-income countries. In many low- and middle-income country contexts, limited digital competency within the public health workforce, often coupled with low self-efficacy, is a key barrier to the uptake of digital technologies [33,34]. This is a gap requiring adaptive models that can leverage open-access curricula and a blend of learning approaches that are adapted to resource-constrained settings.

XSL•FO
RenderX

Given global recognition of the public health implications of societal digital transformation, our study offers valuable guidance to update public health workforce training programs [8]. In Canada, the Chief Public Health Officer's vision for transforming public health prioritizes the use of digital technology [13]. Our findings show that schools of public health are adapting to this global vision in nonstandardized ways. They have focused primarily on data-centric fields such as data science and public health informatics for surveillance, epidemiology, and research through biostatistics and computer sciences [12,29,31]. Furthermore, few programs have comprehensively addressed broader DPH issues, such as digital health promotion, program evaluation, leadership, new media communications, and the digital determinants of health [12,29,35]. Contemporary challenges, such as mis- and disinformation exacerbated by digital transformation, underscore the need to expand efforts in these areas [32,35-37].

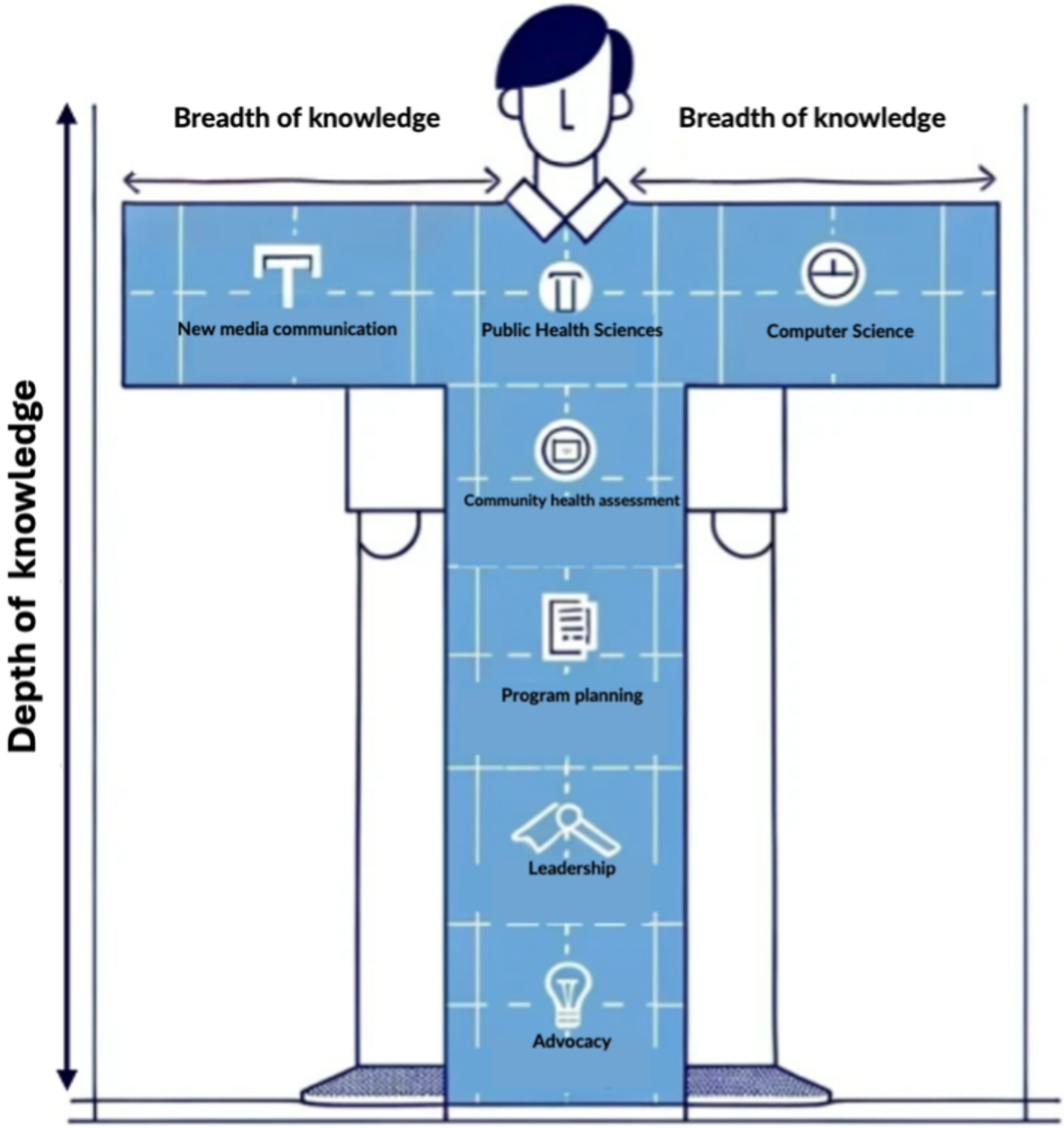
While the ASPHER task force on digital transformation has recommended integrating digital determinants of health into curricula, their suggestion to avoid stand-alone courses does not fully support the development of inter- and transdisciplinary skills required for specialized DPH roles [16,38]. Public health professionals require both foundational knowledge through curricular integration and deeper expertise through stand-alone programs or courses [38]. A stepwise approach is needed to review and integrate digital competencies into existing public health curricula as core competencies for all professionals, while also offering specialized training opportunities [38]. This approach must prioritize inter- and transdisciplinary frameworks to develop "T-shaped professionals" equipped to navigate complex challenges while ensuring opportunities for applied training in real-world contexts (Figure 2) [38-40].

However, adopting this inter- and transdisciplinary paradigm requires addressing challenges such as partnerships, administrative barriers, and funding constraints within schools of public health [40,41]. Funding is essential for baseline integration and the design of specialized DPH programs. Advocacy and collaboration with organizations aligned with transdisciplinary public health goals can catalyze this process, as has been demonstrated by programs such as the Canadian Health Systems Impact Fellowship [42]. Establishing transdisciplinary programs requires engaging motivated partners first, gradually expanding to include other collaborators with clearly defined roles and expectations based on desired graduate competencies [39-41]. Engagement with public and private health sectors and digital actors must follow a principles-driven approach that prioritizes equity and health over commercialization and techno-optimism [9,43].

Reflexivity, ongoing dialog, and program co-creation with partners are essential for the effective development and implementation of these programs. Public sector collaboration is crucial to ensuring programs align with health system needs. Educators must clearly define the value and roles of program graduates to enhance their real-world impact. In addition, curricula should prioritize evaluation using clear indicators to measure effects on partner organizations, workforce competencies, and public health outcomes. As many of the identified programs are relatively new, formal evaluations were considered beyond the scope of this study. We recognize that assessing outcomes will be essential for understanding their effectiveness and for securing long-term funding and institutional support [41].

XSL•FO
RenderX

Figure 2. Depiction of a T-shaped professional.



Strengths and Limitations

The findings from this study should be interpreted considering its strengths and limitations. To our knowledge, this is the first systematic documentation of DPH training programs globally. Our approach combined a review of available programs with interviews with training program staff, ensuring both rigor and a deeper understanding of program contexts. We also aimed for an expansive review by including all eligible programs worldwide. However, our reliance on Google Search and the use of English-only terms in the search strategy may have introduced bias, as search algorithms are location specific and prioritize popular programs. To mitigate this, we used Google Translate where appropriate to ensure returns in languages other than English were not excluded. In addition, this study did not comprehensively review all nondigital courses in public health training programs to assess whether current coursework incorporates DPH perspectives. Other similar environmental scans adopted this approach and showed approximately 20% of programs have considered DPH mainly using similar approaches as identified in our review [29]. Our assumptions of digital programs and courses being hosted within the schools of public health, as demonstrated in our search, may have inadvertently missed important DPH programs hosted in other allied schools or programs leveraging cross-listed courses. We also acknowledge limitations in the sample size for the interview with only 4 participants responding out of 22 invitations sent. This limited us from making any interjurisdictional comparisons. While our coding was inductive and did not follow a predetermined framework, the limited sample size led to

XSL•FO
RenderX

categories that closely mirrored the main areas outlined in the topic guide. Nevertheless, the breadth and depth of conversations in the interviews ensure that we can gain useful insights from other programs while encouraging careful appraisal of readers' own context when considering transferability of findings. Further research could include more broad qualitative inquiry of international perspectives, but our work represents an important first step.

Conclusions

This environmental scan of DPH training programs highlights the predominance of data-centric competencies, such as public health data science and informatics, while identifying critical gaps in addressing competencies for digital health promotion, program evaluation, leadership, and digital determinants of health. Despite efforts by schools of public health to align with global and national priorities, progress remains fragmented and geographically limited. To address these gaps, we recommend a stepwise approach that integrates digital competencies as core elements in public health curricula while developing stand-alone programs to cultivate specialized skills. Emphasizing inter- and transdisciplinary frameworks is essential for building "T-shaped professionals" capable of navigating complex public health challenges, but overcoming barriers related to funding, administrative constraints, and partnerships will require sustained advocacy, principles-driven collaboration, and robust evaluation frameworks. These efforts must prioritize equity and health outcomes over commercial imperatives, ensuring that DPH training equips the workforce to respond effectively to the challenges and opportunities of societal digital transformation.

Acknowledgments

The authors would like to express their gratitude to the program directors who freely contributed their time for interviews. This research was supported by the Canadian Institutes for Health Research (CIHR) through the Chief Public Health Officer Catalyst Grant for Transforming Public Health (grant AWD-022946). II is supported by the CIHR Health Systems Impact Fellowship (Award number 521442), Michael Smith Health Research BC Trainee Award (Award number HSIF-2024-04465), and the CIHR Canadian HIV Trials Network (CTN+) postdoctoral fellowship.

Data Availability

The datasets used and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

II and MG conceptualized the study and drafted the study protocol, which was reviewed by HJC, AK, AB, FIC, HD, and GM; II and SR conducted data collection; II and SR analyzed the data and reviewed outputs with all authors; SR and II drafted the manuscript; MG secured funding for the study, and all authors critically reviewed and approved the final version of this manuscript.

Conflicts of Interest

II serves on the board of the Public Health Association of British Columbia. All other authors declare no conflicts of interest.

Multimedia Appendix 1
List of countries included in Google advanced search.
[DOCX File, 17 KB - publichealth_v11i1e73088_app1.docx]

Multimedia Appendix 2
Data extraction form for environmental scan.
[DOCX File, 15 KB - publichealth_v11i1e73088_app2.docx]

Multimedia Appendix 3
Interview guide for semistructured interviews.
[DOCX File, 15 KB - publichealth_v11i1e73088_app3.docx]

Multimedia Appendix 4
Consent form for interviewers.
[PDF File, 145 KB - publichealth_v11i1e73088_app4.pdf]

Multimedia Appendix 5
Summary of all included programs.
[DOCX File, 51 KB - publichealth_v11i1e73088_app5.docx]

XSL•FO
RenderX

References

1. Iyamu I, Xu AXT, Gómez-Ramírez O, et al. Defining digital public health and the role of digitization, digitalization, and digital transformation: scoping review. JMIR Public Health Surveill 2021 Nov 26;7(11):e30399. [doi: 10.2196/30399] [Medline: 34842555]
2. Budd J, Miller BS, Manning EM, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
3. Gunasekeran DV, Tseng R, Tham YC, Wong TY. Applications of digital health for public health responses to COVID-19: a systematic scoping review of artificial intelligence, telehealth and related technologies. NPJ Digit Med 2021 Feb 26;4(1):40. [doi: 10.1038/s41746-021-00412-9] [Medline: 33637833]
4. Odone A, Buttigieg S, Ricciardi W, Azzopardi-Muscat N, Staines A. Public health digitalization in Europe. Eur J Public Health 2019 Oct 1;29(Supplement_3):28-35. [doi: 10.1093/eurpub/ckz161] [Medline: 31738441]
5. Dion M, AbdelMalik P, Mawudeku A. Big data and the global public health intelligence network (GPHIN). Can Commun Dis Rep 2015 Sep 3;41(9):209-214. [doi: 10.14745/ccdr.v41i09a02] [Medline: 29769954]
6. van Kessel R, Wong BLH, Clemens T, Brand H. Digital health literacy as a super determinant of health: more than simply the sum of its parts. Internet Interv 2022 Mar;27:100500. [doi: 10.1016/j.invent.2022.100500] [Medline: 35242586]
7. van Kessel R, Seghers LE, Anderson M, et al. A scoping review and expert consensus on digital determinants of health. Bull World Health Organ 2025 Feb 1;103(2):110-125H. [doi: 10.2471/BLT.24.292057] [Medline: 39882497]
8. Kickbusch I, Piselli D, Agrawal A, et al. The Lancet and Financial Times Commission on governing health futures 2030: growing up in a digital world. The Lancet 2021 Nov;398(10312):1727-1776. [doi: 10.1016/S0140-6736(21)01824-9]
9. Gómez-Ramírez O, Iyamu I, Ablona A, et al. On the imperative of thinking through the ethical, health equity, and social justice possibilities and limits of digital technologies in public health. Can J Public Health 2021 Jun;112(3):412-416. [doi: 10.17269/s41997-021-00487-7] [Medline: 33725332]
10. Sina E, Boakye D, Christianson L, Ahrens W, Hebestreit A. Social media and children's and adolescents' diets: a systematic review of the underlying social and physiological mechanisms. Adv Nutr 2022 Jun 1;13(3):913-937. [doi: 10.1093/advances/nmac018] [Medline: 35218190]
11. Sina E, Buck C, Ahrens W, et al. Digital media exposure and cognitive functioning in European children and adolescents of the I.Family study. Sci Rep 2023 Nov 1;13(1):18855. [doi: 10.1038/s41598-023-45944-0] [Medline: 37914849]
12. Ramachandran S, Chang HJ, Worthington C, et al. Digital competencies and training approaches to enhance the capacity of practitioners to support the digital transformation of public health: rapid review of current recommendations. JMIR Public Health Surveill 2024 Sep 9;10:e52798. [doi: 10.2196/52798] [Medline: 39248660]
13. Public Health Agency of Canada. The chief public health officer of canada's report on the state of public health in canada 2021: a vision to transform canada's public health system. : Public Health Agency of Canada; 2021 Dec 1924–7087 URL: https://www.canada.ca/en/public-health/corporate/publications/chief-public-health-officer-reports-state-public-health-canada/state-public-health-canada-2021.html [accessed 2022-03-01]
14. Ubaldi B, Welby B, Chauvet L. The OECD framework for digital talent and skills in the public sector. 2021 45 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/04/the-oecd-framework-for-digital-talent-and-skills-in-the-public-sector_f6fb7838/4e7c3f58-en.pdf [accessed 2025-08-13] [doi: 10.1787/4e7c3f58-en]
15. WHO-ASPHER competency framework for the public health workforce in the European region world health organization regional office for Europe. World Health Organization. 2020. URL: https://iris.who.int/handle/10665/347866 [accessed 2024-06-29]
16. Hrzic R, Buttigieg S, Wong BLH, et al. Reflecting digital transformations in public health curricula. Public Health Rev 2025;46:1608237. [doi: 10.3389/phrs.2025.1608237] [Medline: 40144365]
17. Iyamu I, Gómez-Ramírez O, Xu AX, et al. Challenges in the development of digital public health interventions and mapped solutions: findings from a scoping review. Digit Health 2022;8:20552076221102255. [doi: 10.1177/20552076221102255] [Medline: 35656283]
18. National collaborating centres for public health. 2024 Core Competencies for Public Health in Canada. 2023. URL: https://nccph.ca/projects/2024-core-competencies-for-public-health-in-canada/ [accessed 2024-11-22]
19. Wickson F, Carew AL, Russell AW. Transdisciplinary research: characteristics, quandaries and quality. Futures 2006 Nov;38(9):1046-1059. [doi: 10.1016/j.futures.2006.02.011]
20. Iyamu I, Ramachandran S, Chang HJ, et al. Empowering Canada's Workforce for the Digital Tranformation of Public Health: Recommendations for Digital Competencies, Curricular and Traning Updates. : Digital & Sexual Health Initiative, University of British Columbia Centre for Disease Control; 2025 URL: https://med-fom-digitalpublichealth.sites.olt.ubc.ca/files/2025/04/Digital-competencies-report-final-2.pdf [accessed 2025-05-25]
21. Wilburn A, Vanderpool RC, Knight JR. Environmental scanning as a public health tool: Kentucky's human papillomavirus vaccination project. Prev Chronic Dis 2016 Aug 18;13:E109. [doi: 10.5888/pcd13.160165] [Medline: 27536901]
22. Rowel R, Moore ND, Nowrojee S, Memiah P, Bronner Y. The utility of the environmental scan for public health practice: lessons from an urban program to increase cancer screening. J Natl Med Assoc 2005 Apr;97(4):527-534. [Medline: 15868772]

XSL•FO
RenderX

23. Public health: a conceptual framework. : Canadian Public Health Association; 2017 URL: https://www.cpha.ca/sites/default/files/uploads/resources/cannabis/cpha_public_health_conceptual_framework_e.pdf [accessed 2025-10-02]
24. Phillips AC, Lewis LK, McEvoy MP, et al. Development and validation of the guideline for reporting evidence-based practice educational interventions and teaching (GREET). BMC Med Educ 2016 Sep 6;16(1):237. [doi: 10.1186/s12909-016-0759-1] [Medline: 27599967]
25. QSR International Pty Ltd. NVivo Qualitative Data Analysis Software. 2020. URL: https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home/ [accessed 2025-08-13]
26. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
27. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
28. Finlay L. Reflexivity: an essential component for all research? British Journal of Occupational Therapy 1998 Oct;61(10):453-456. [doi: 10.1177/030802269806101005]
29. Albrecht J, Maaß L, Tokgöz P, et al. Wie viel Digital Public Health steckt in Public-Health-Studiengängen? Eine systematische Modulhandbuchanalyse von Vollzeitstudiengängen an öffentlichen Hochschulen und Universitäten in Deutschland [Article in German]. Bundesgesundheitsbl 2024 Mar;67(3):339-350. [doi: 10.1007/s00103-024-03844-2] [Medline: 38436689]
30. Maaß L, Dockweiler C, Hocke-Bolte Z, et al. Digital Public Health in Deutschland: Status quo, Herausforderungen und Zukunftsperspektiven: Ein Positionspapier des Fachbereichs Digital Public Health der Deutschen Gesellschaft für Public Health e [Article in German]. V Bundesgesundheitsbl 2024 Dec 5. [doi: 10.1007/s00103-024-03989-0]
31. Fisher S, Rosella LC. Priorities for successful use of artificial intelligence by public health organizations: a literature review. BMC Public Health 2022 Nov 22;22(1):2146. [doi: 10.1186/s12889-022-14422-z] [Medline: 36419010]
32. Iyamu I, McKee G, Haag D, Gilbert M. Defining the role of digital public health in the evolving digital health landscape: policy and practice implications in Canada. Health Promot Chronic Dis Prev Can 2024 Feb;44(2):66-69. [doi: 10.24095/hpcdp.44.2.04] [Medline: 38353941]
33. Shiferaw KB, Tilahun BC, Endehabtu BF. Healthcare providers' digital competency: a cross-sectional survey in a low-income country setting. BMC Health Serv Res 2020 Nov 9;20(1):1021. [doi: 10.1186/s12913-020-05848-5] [Medline: 33168002]
34. Wang M, Huang K, Li X, et al. Health workers' adoption of digital health technology in low- and middle-income countries: a systematic review and meta-analysis. Bull World Health Organ 2025 Feb 1;103(2):126-135F. [doi: 10.2471/BLT.24.292157] [Medline: 39882495]
35. Kickbusch I, Holly L. Addressing the digital determinants of health: health promotion must lead the charge. Health Promot Int 2023 Jun 1;38(3):daad059. [doi: 10.1093/heapro/daad059] [Medline: 37264549]
36. Ferrara E, Cresci S, Luceri L. Misinformation, manipulation, and abuse on social media in the era of COVID-19. J Comput Soc Sci 2020;3(2):271-277. [doi: 10.1007/s42001-020-00094-5] [Medline: 33251373]
37. Wardle C, Singerman E. Too little, too late: social media companies' failure to tackle vaccine misinformation poses a real threat. BMJ 2021 Jan 21;372:n26. [doi: 10.1136/bmj.n26] [Medline: 33478950]
38. Iyamu I, Ramachandran S, Chang HJ, et al. Considerations for adapting digital competencies and training approaches to the public health workforce: an interpretive description of practitioners' perspectives in Canada. BMC Public Health 2025 Jan 10;25(1):122. [doi: 10.1186/s12889-024-21089-1] [Medline: 39794767]
39. Demirkan H, Spohrer JC. Commentary—cultivating T-shaped professionals in the era of digital transformation. Serv Sci 2018 Mar;10(1):98-109. [doi: 10.1287/serv.2017.0204]
40. Caputo F, Cillo V, Fiano F, Pironti M, Romano M. Building T-shaped professionals for mastering digital transformation. J Bus Res 2023 Jan;154:113309. [doi: 10.1016/j.jbusres.2022.113309]
41. Lawrence MG, Williams S, Nanz P, Renn O. Characteristics, potentials, and challenges of transdisciplinary research. One Earth 2022 Jan;5(1):44-61. [doi: 10.1016/j.oneear.2021.12.010]
42. McMahon M, Bornstein S, Brown A, Simpson LA, Savitz L, Tamblyn R. Training for health system improvement: emerging lessons from Canadian and US approaches to embedded fellowships. Healthc Policy 2019 Oct;15(SP):34-48. [doi: 10.12927/hcpol.2019.25981] [Medline: 31755858]
43. Rodriguez JA, Lyles CR. Strengthening digital health equity by balancing techno-optimism and techno-skepticism through implementation science. NPJ Digit Med 2023 Nov 2;6(1):203. [doi: 10.1038/s41746-023-00954-0] [Medline: 37919418]

Abbreviations

ASPHER : Association of Schools for Public Health in the European Region
DPH: digital public health

XSL•FO
RenderX

Edited by C Argyropoulos; submitted 24.02.25; peer-reviewed by D Naidoo, K Ikuyinminu, TL Rose; revised version received 13.08.25; accepted 08.09.25; published 15.10.25.

Please cite as:
Iyamu I, Ramachandran S, Chang HJ, Kushniruk A, Ibáñez-Carrasco F, Worthington C, Davies H, McKee G, Brown A, Gilbert M Building the Workforce's Capacity to Support the Digital Transformation of Public Health: Environmental Scan of Training Programs for Digital Technologies in Public Health
JMIR Public Health Surveill 2025;11:e73088
URL: https://publichealth.jmir.org/2025/1/e73088
doi:10.2196/73088

© Ihoghosa Iyamu, Swathi Ramachandran, Hsiu-Ju Chang, Andre Kushniruk, Francisco Ibáñez-Carrasco, Catherine Worthington, Hugh Davies, Geoffrey McKee, Adalsteinn Brown, Mark Gilbert. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 15.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Estimating Chronic Hepatitis B Prevalence and Undiagnosed Proportion in Canada, 2007-2021: Mathematical Framework Development

Julien Smith-Roberge[1], PhD; Farinaz Forouzannia[1], PhD; Abdullah Hamadeh[1], PhD; Zeny Feng[2], PhD; Nashira Popovic[3], PhD; William W L Wong[1], PhD

[1]School of Pharmacy, University of Waterloo, 10A Victoria St. S., Kitchener, ON, Canada

[2]Department of Mathematics and Statistics, University of Guelph, Guelph, ON, Canada

[3]Centre for Communicable Diseases and Infection Control, Public Health Agency of Canada, Ottawa, ON, Canada

Corresponding Author:
William W L Wong, PhD
School of Pharmacy, University of Waterloo, 10A Victoria St. S., Kitchener, ON, Canada

Abstract

Background: Chronic hepatitis B (CHB) remains a significant health burden for at least a hundred thousand Canadians. The government of Canada has endorsed the global strategy to eliminate hepatitis as a public health threat by 2030, but effectively targeting public health interventions is complicated by the silent nature of the disease, which can remain asymptomatic for decades.

Objective: This study develops a framework to estimate the prevalence of CHB and the proportion of the infected population that remains undiagnosed. We apply the proposed method to national data from Canada, from 2007 to 2021.

Methods: We infer the prevalence and undiagnosed proportion of CHB by fitting a mathematical state-transition model, based on the natural history of CHB, to observed CHB-related events. Data for the calibration were obtained from the Public Health Agency of Canada and Statistics Canada.

Results: We estimate the national prevalence of CHB in Canada in 2021 to be 0.483% (95% CI 0.443%-0.535%). The corresponding percentage of undiagnosed cases was estimated to be 54.3% (95% CI 50.9%-58.9%).

Conclusions: The estimates of CHB prevalence obtained via our method are in line with previous estimates obtained from national seroprevalence studies. More specialized estimates, stratified by province or age cohort, may be achievable with detailed health administrative data.

(JMIR Public Health Surveill 2025;11:e66309) doi:10.2196/66309

KEYWORDS

hepatitis B; prevalence; undiagnosed proportion; mathematical modeling; estimate; Canada; mathematical framework; chronic hepatitis B; CHB; hepatitis; undiagnosed; diagnosis; estimates; prediction; state transition model; liver; liver transplant; hepatocellular carcinoma; decompensated cirrhosis

Introduction

Chronic hepatitis B (CHB) is a silent disease. In its early stages, which can last decades, it is largely asymptomatic. Because of this, those with the disease can remain undiagnosed for years until they are alerted to their condition by late-stage complications: decompensated cirrhosis (DC), hepatocellular carcinoma (HCC), and liver failure. It was estimated that hepatitis B caused 1.1 million deaths worldwide in 2022 [1].

In 2016, Canada endorsed the Global Strategy to eliminate hepatitis as a public health threat by 2030 [2]. It has already made progress toward this goal; a seroprevalence study published in 2013 estimated the national prevalence of CHB to be 0.4%, well below the global average [3]. This is in large part thanks to Canada's robust vaccination program. In 2021, vaccine coverage was 89% for those aged 14 years or older [4]. However, many challenges remain: hepatitis B was one of the most frequently refused vaccines in 2021, third behind only varicella and rotavirus, and the lack of a nationally standardized vaccine schedule means that children who move across provincial lines may miss vaccination due to conflicting schedules [4,5]. Moreover, this vaccination program leaves several key populations underprotected, notably immigrants from countries with high hepatitis B prevalence and children exposed prior to vaccination, a risk that is elevated in provinces that do not have an infant vaccination program. The same 2013 seroprevalence study found that the CHB prevalence among the foreign-born population was 1.6%, 4 times the national

XSL•FO
RenderX

average [3], and 90% of unvaccinated infants born to hepatitis B virus (HBV)–infected mothers will develop CHB within 6 months [6]. The risk of contracting CHB also increases with exposure to infected blood and other body fluids, such as through needle sharing or unprotected sex [1].

These at-risk groups comprise a significant portion of Canada's population. According to the most recent Canadian census published in 2021 [7], immigrants represent 23% of the country's total population. Many of these originate from regions where HBV is either highly or moderately endemic, notably China and the Philippines, which, along with India, have been the top 3 sources of immigrants to Canada since 1991. As a result, a significant proportion of the CHB population in Canada is comprised of immigrants. Furthermore, Canada's vaccination program leaves a significant portion of the population underprotected. Many of the provinces still perform adolescent vaccination instead of birth dose vaccination [8]. Other gaps exist in the country's prevention strategy. For example, in Ontario, prenatal screening for HBV was not universal, with a coverage of 92.7% between 2012 and 2016. These gaps resulted in at least 139 children born between 2003 and 2013 being infected with HBV before being eligible for adolescent vaccination [9].

Meeting Canada's elimination goals will require further evidence-based interventions, such as screening and community outreach, rooted in robust estimates of CHB prevalence to track the impact and the efficacy of these investments in public health. In this paper, we present a model-based framework for estimating the prevalence and undiagnosed proportion of CHB. Our approach follows in the line of similar back-calculation methods used for other diseases with long incubation periods, such as hepatitis C [10,11] and HIV [12]. Similar methods have also been used to estimate the prevalence and undiagnosed fraction of COVID-19 [13]. To our knowledge, the present study represents the first attempt to apply such back-calculation techniques to CHB.

Methods

Overview

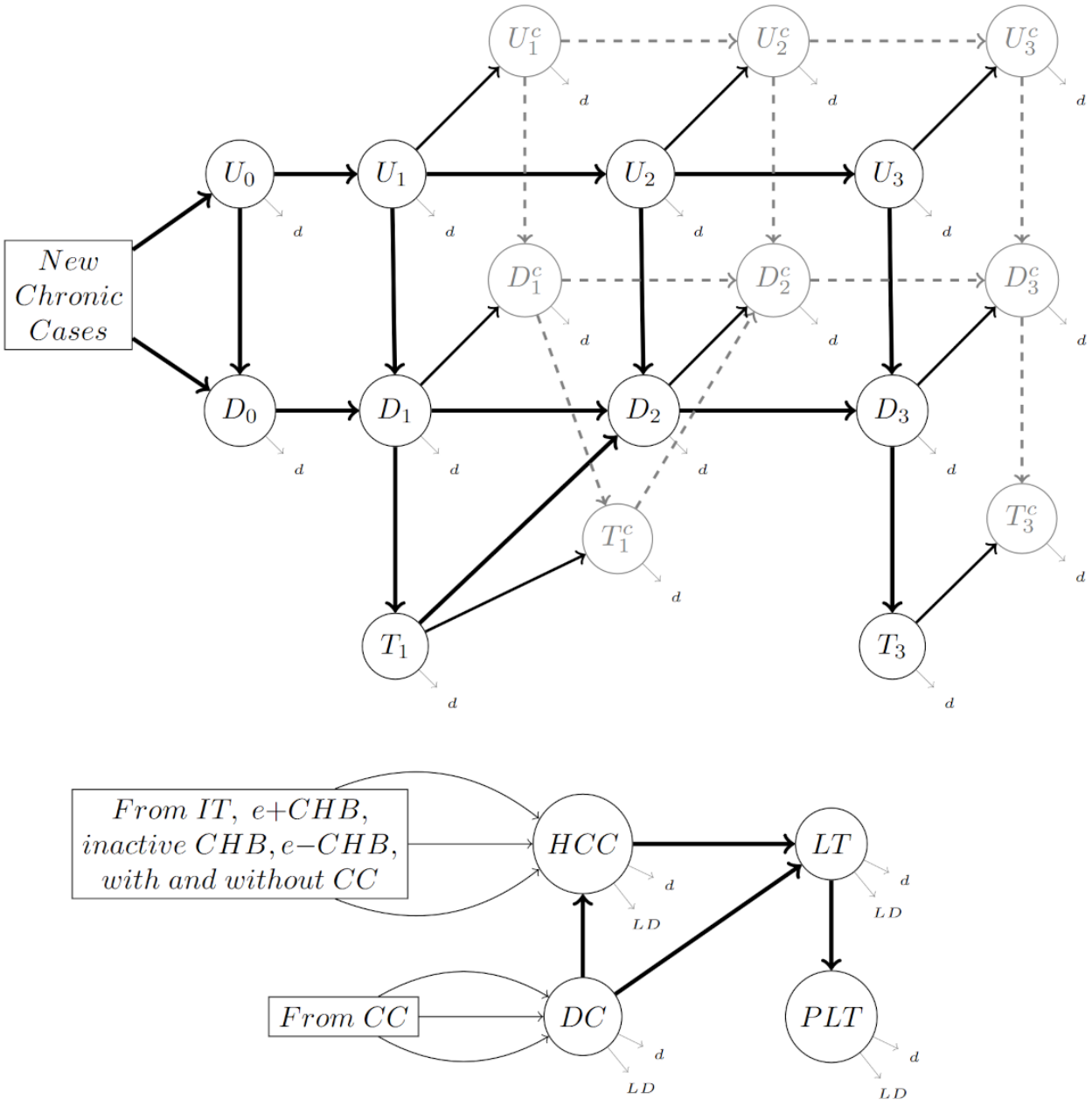
In this section, we describe the mathematical framework we use to estimate the prevalence of CHB in populations of interest. We developed a state-transition model of CHB progression based on the natural history of the disease, building on earlier work in [14]. The model includes several late-stage clinical health states, such as DC and HCC, which are more readily diagnosed than the asymptomatic stages of CHB. Indeed, we assume that the symptoms of these late-stage clinical health states are severe enough that they will be diagnosed within 1 year. Using a Bayesian calibration method based on the Metropolis-Hastings algorithm, we used publicly available diagnosis data for these late-stage complications, in combination with diagnosis data for CHB itself, to back-calculate the prevalence and undiagnosed fraction of CHB.

Natural History Model

We developed a state transition model to describe the infection, progression, and treatment process, following the natural history of CHB. A schematic of this model is presented in Figure 1. We assume that people infected with CHB will remain in exactly one state for the duration of a given calendar year. Once per year, individuals may transition to a different state based on probabilities derived from the parameters in Table 1 and Multimedia Appendix 1. Some of these state transitions, such as the transition from an undiagnosed state to a diagnosed state, correspond to observable events recorded by public health officials. Our goal is to infer the number of individuals in each state by calibrating the model to reproduce the time evolution of these observables.

XSL•FO
RenderX

Figure 1. A schematic representation of the state-transition model. CC: compensated cirrhosis; CHB: chronic hepatitis B; DC: decompensated cirrhosis; HCC: hepatocellular carcinoma; LT: liver transplant; PLT: post–liver transplant.



XSL•FO
RenderX

Table . Unknown model parameters. The minimum and maximum constrain the range explored during the calibration process.

Parameter	Description	Min	Max
dl/r	Annual probability of CHB[a] diagnosis	0.005	0.5
dcc	Annual probability of CC[b]-specific diagnosis	0	0.1
daa	Probability of newly infected individuals being diagnosed with CHB within the first year	0	0.3
μl/r	Annual CHB incidence	0	0.1
rhcc0	Adjustment to CHB-related HCC[c] risk	0	1

[a]CHB: chronic hepatitis B.

[b]CC: compensated cirrhosis

[c]HCC: hepatocellular carcinoma.

The structure of our state-transition model is depicted in Figure 1. For clarity, the early stages in the top portion of the diagram are organized along x-, y-, and z-axes based on their nature. The states are organized in rows and denoted by U, D, and T, corresponding to the cascade of care, specifically, the undiagnosed and untreated states, for those who have CHB but are unaware of their infection; the diagnosed and untreated states, for those with a positive CHB diagnosis who have not initiated treatment; and diagnosed and treated states, for those receiving treatment for their CHB infection. Each of these basic states is further subdivided into sub-stages, organized into columns, that represent the natural history of CHB. These sub-stages are denoted using a subscript. The subscript 1 denotes chronic hepatitis B e-antigen (HBeAg) + hepatitis B (hepatitis B surface antigen [HBsAg]: positive, HBeAg: positive, viral load: high, ALT level: high). The subscript 2 denotes Inactive Hepatitis B (HBsAg: positive, HBeAg: negative, viral load: low, ALT level: normal). The subscript 3 denotes Chronic HBeAg-Hepatitis B (HBsAg: positive, HBeAg: negative, viral load: high, ALT level: high).

Finally, each of these states can exist with or without compensated cirrhosis (CC), which, when present, is represented with a superscript C. In order to cover the whole natural history of CHB, we also include 2 states, U0 and D0, corresponding to the immune tolerant state (HBsAg: positive, HBeAg: positive, viral load: high, ALT level: normal).

People infected with CHB may progress into advanced stages of liver disease, which are depicted in the bottom half of Figure 1. These advanced stages are categorized into the following states: HCC, DC, liver transplant (LT), and post–liver transplant (PLT).

We define the LT state to last at most 1 year. Transplant recipients who have survived more than 1 year after their liver transplantation enter the PLT state. Additionally, we assume that individuals in these later states have a probability of mortality greater than the background all-cause mortality. This is represented by arrows to a liver death state, LD. Background mortality is represented by an arrow to a death state, d. A full list of model states is included in Multimedia Appendix 2.

Finally, we assume that in the year *t*, there are u(t) new CHB infections. These new infections include observed infections that are diagnosed within the first year, in which case they will be directed to the state D0, as well as unobserved cases that remain undiagnosed and go to U0.

Model Transition Probabilities

The transition probabilities of the model depend on a mix of known and unknown parameters. The known parameters are obtained from the literature and are presented in Multimedia Appendix 1. The unknown parameters are presented in (Table 1). We will estimate these unknown parameters using our calibration procedure, which is described in later sections of this paper.

Because CHB is initially asymptomatic, we assume that the probability of receiving a diagnosis is constant across all the undiagnosed states U_i. In principle, these diagnosis probabilities can vary significantly from year to year. As a simplifying approximation, we assume that this trend is linear over the period being modeled, with a diagnosis probability of dl in the first year being modeled, tinit, and a final diagnosis probability of dr in the last year modeled, tfin. For notational convenience, we denote this time-dependent quantity dl/r(t). For undiagnosed states with compensated cirrhosis, we add an annual probability of CC-specific diagnosis, dcc, to allow for the fact that compensated cirrhosis may likely produce symptoms, which increases the probability of diagnosis. Thus, the diagnosis probability for an individual in one of the UiC states would be dl/r(t)+dcc. Finally, we apply a different diagnosis probability for the first year of infection, daa. In any given year, daau(t) new cases will be added to the D0 state, with the remaining (1−daa)u(t) cases being added to the U0 state. This first-year diagnosis probability allows us to account for CHB-infected individuals who seek screening following high-risk behavior or immigrants who disclose their health status on entering the country.

Much like the diagnosis probability, we also assume that CHB incidence, μ, is a linear function of time. Following the same notational convention as earlier, the number of new cases in a given year is given by u(t)=μl/r(t)·Pop(t), where Pop(t) is the

XSL•FO
RenderX

population of the cohort being studied. We note that several of the parameter extrema, such as a CHB incidence of 0.1, are set implausibly high to grant the model the flexibility to explore all plausible parameter values.

The final calibration parameter, rhcc0, is an adjustment parameter that aims to account for uncertainty in the probability of developing HCC. The annual probabilities of developing HCC, as reported in the literature, are compiled in Multimedia Appendix 1.

Finally, we note that these model parameters, both known and unknown, represent the underlying probabilities associated with any basic state transition. In principle, simulating the full state transition model represented by Figure 1 would also require knowledge of joint probabilities. In order to obtain a tractable model, we make the simplifying assumption that each of these basic events is independent.

Model Assumptions

We highlight some key simplifying assumptions made during the development of the model (Textbox 1).

Textbox 1. Model assumptions.

- Incidence rate is a linear function of time and does not explicitly depend on other factors, such as vaccine uptake.
- The age distribution of the population with chronic hepatitis B (CHB) matches that of the general population. In particular, age-related quantities, such as the background mortality rate, are the same in both populations.
- There is no net migration of infected individuals after arrival. Moreover, because CHB has no readily available cure (spontaneous loss of hepatitis B surface antigen [HBsAg] is possible, but rare), we assume that the number of cases can only decrease by death.
- Treatment uptake and efficacy remained constant over the observed period, as did survival rates of late-stage complications.

Model Dynamics

For each of the model states (U0, D0, U1, ..., HCC, DC, LT, and PLT), the state-transition model generates a time series for the expected number of individuals in each state. We denote these by U0(t), D0(t), ..., and PLT(t), for t=tinit,...,tfin. Additionally, we record the number of individuals transitioning between states. These simulated time series—the number of individuals in each state and the number of individuals transitioning between states—are used to calculate 4 *observables*, which are used to calibrate the model. The first of these is the number of people receiving treatment, T1(t)+T3(t)+T1C(t)+T3C(t). Based on clinical guidelines [15], treatment is mainly offered to infected individuals in HBeAg+ and HBeAg- states, and thus our model does not include an inactive treatment state, T2. The remaining observables are the number of CHB diagnoses, the number of HCC diagnoses, and the number of late HCC diagnoses, which we define as any transition from an undiagnosed state, Ui-/C, to the HCC state. These diagnosis events correspond to transitions between undiagnosed states to diagnosed states. The details of these computations can be found in (Multimedia Appendix 2).

Data Sources

The observables described earlier are compared to 3 data sources: the annual number of HCC diagnoses, obtained from Statistics Canada [16], which we will denote YHCCTotal(t); the annual number of CHB diagnoses, obtained from the Public Health Agency of Canada's (PHAC) Notifiable Diseases Online [17], denoted YCHB(t); and CHB treatment data, which is based on prescription data obtained from IQVIA via collaboration with PHAC, denoted YTreat(t). The CHB and treatment data can be directly compared to their corresponding observables, but the other 2 observables require additional assumptions on the data before they can be used.

The HCC figures from Statistics Canada include all instances of HCC, regardless of cause. Following [14], the estimated number of CHB-related HCC cases is 9% of total HCC cases, $YHCC(t) = 0.09 \cdot YHCCTotal(t)$.

In [18], an estimated 1.07% of CHB diagnoses (198 of 18,434) occur within ±6 months of an HCC diagnosis. Thus, we assume the number of late HCC diagnoses per year is given by

$$.YLateHCC(t)=0.0107 \cdot YCHB(t)$$

Reporting HBV cases is mandatory. Therefore, we can be confident that the Notifiable Diseases Online database accurately reflects the number of HBV diagnoses. However, although HBV diagnoses must be reported, the chronic versus acute distinction is not reportable in all provinces and territories. Unspecified cases were apportioned based on the chronic-to-acute ratio observed in the specified cases. We note that although the rate of HBV diagnosis per se may be biased (not all people who are infected will seek medical attention, and therefore the Notifiable Diseases database might systematically underrepresent populations less likely to interact with the health care system), the structure of our mathematical model accounts for this possibility. Bias or underreporting in the diagnoses of late-stage complications represents a greater source of uncertainty in the model. As such, the HCC data from Statistics Canada represents the most significant source of data-related uncertainty in our analysis. The data for model calibration are available in Multimedia Appendix 3.

Model Calibration

The model contains several unknown quantities that need to be inferred via model calibration. These come in 2 broad categories: the unknown parameters from Table 1, and the initial population sizes for each of the model states.

If the initial population distributions in the year tinit deviate too much from the true values, the calibration converges to a spurious local optimum. In order to avoid this, we use a 2-stage calibration process. We begin by assuming a prevalence of 0.4%, in line with earlier seroprevalence results [3], and distribute these among the states in the model based on

XSL•FO
RenderX

percentages used in previous modeling studies [19]. We then calibrated the model against data from a subset of provinces—Alberta, British Columbia, Ontario, and Saskatchewan—whose diagnosis, treatment, and HCC data were the most complete and selected the population distribution that yielded the closest match to their data. These population distributions then served as our initial distribution in all subsequent calibrations.

In order to obtain an initial estimate of incidence, μl/r, we assume the system is in approximate equilibrium and apply the same 0.4% seroprevalence estimate. This yields a preliminary estimate of μl/r≈0.00004.

In principle, the early diagnosis parameter, daa, can take any value between 0 and 1, though in practice we know that values close to 1 are unrealistic. We restrict our search range to 0≤daa≤0.3 [18].

Calibration Algorithm

Following [10], we use a 2-stage calibration algorithm, which is designed to account for the uncertainty in the literature-derived parameters. Let vlit=[q01,...,dld4] be a set of the literature-derived model parameters, ie, a set of parameters sampled from the ranges in Multimedia Appendix 1. Similarly, let vcal= [dl,...,rhcc0,U0(tinit),...,PLT(tinit)] denote a set of the calibration parameters, ie, a set of the unknown model parameters from Table 1 and a set of initial population estimates. Taken together, a single pair (vlit,vcal) determines the initial conditions and transition probabilities, and thus fully specifies the behavior of the state transition model. The goal of our calibration is to generate pairs that yield a good fit with the available data.

Our 2-stage calibration algorithm is described as follows:

Stage 1: We begin by generating k sets of the literature-derived parameters,

[v1lit,…,vklit], by sampling uniformly from the ranges in Multimedia Appendix 1. For each sample vilit, we calibrate the model using the Metropolis-Hastings MCMC algorithm [20], using the likelihood function described in Multimedia Appendix 2, until the Markov chain converges to a stationary distribution, P(vcal|Y,vilit). Discarding the burn-in, we obtain a set of n samples for the unknown parameters, [v1cal,…,vncal], sampled from this distribution.

Stage 2: The samples generated in the first stage are conditioned on a choice of vilit. Thus, in order to obtain a full set of model parameters, we perform a second sampling step. We first sample uniformly from the literature-derived parameters obtained in step 1, [v1lit,…,vklit], and then sample uniformly from the associated distribution obtained via the MCMC algorithm. The result is a single combined set of parameters, (vlit,vcal), sampled from the distribution P(vlit,vcal|Y).

For each pair (vlit,vcal) derived in this way, the model produces a time series for each of the model states, U0(t), D0(t), ..., PLT(t). In order to describe the typical behavior of our model, we sample multiple pairs (vlit,vcal) and compute the means of the associated time series, $\bar{U0}(t)$, $\bar{D0}(t)$, ..., $\bar{PLT}(t)$, as well as the corresponding standard deviations.

We obtain a prevalence estimate by summing over all model states and dividing by the total population size.

$$\text{CHB prevalence}(t)=[\bar{U0}(t)+\bar{D0}(t)+...+\bar{PLT}(t)]/\text{Pop}(t)$$

The undiagnosed fraction is computed as the ratio of all undiagnosed states and the total CHB.

U

$$\text{Undiagnosed fraction}(t)=\bar{U0}(t)+\bar{U1}(t)+...+\bar{U3C}(t)\bar{U0}(t)+\bar{D0}(t)+...+\bar{PLT}(t)$$

Using this 2-stage calibration algorithm, we first estimate the national prevalence. We then conduct a stratified analysis for male and female populations. Finally, we conduct a one-way sensitivity analysis to explore the uncertainty associated with these estimates.

Ethical Considerations

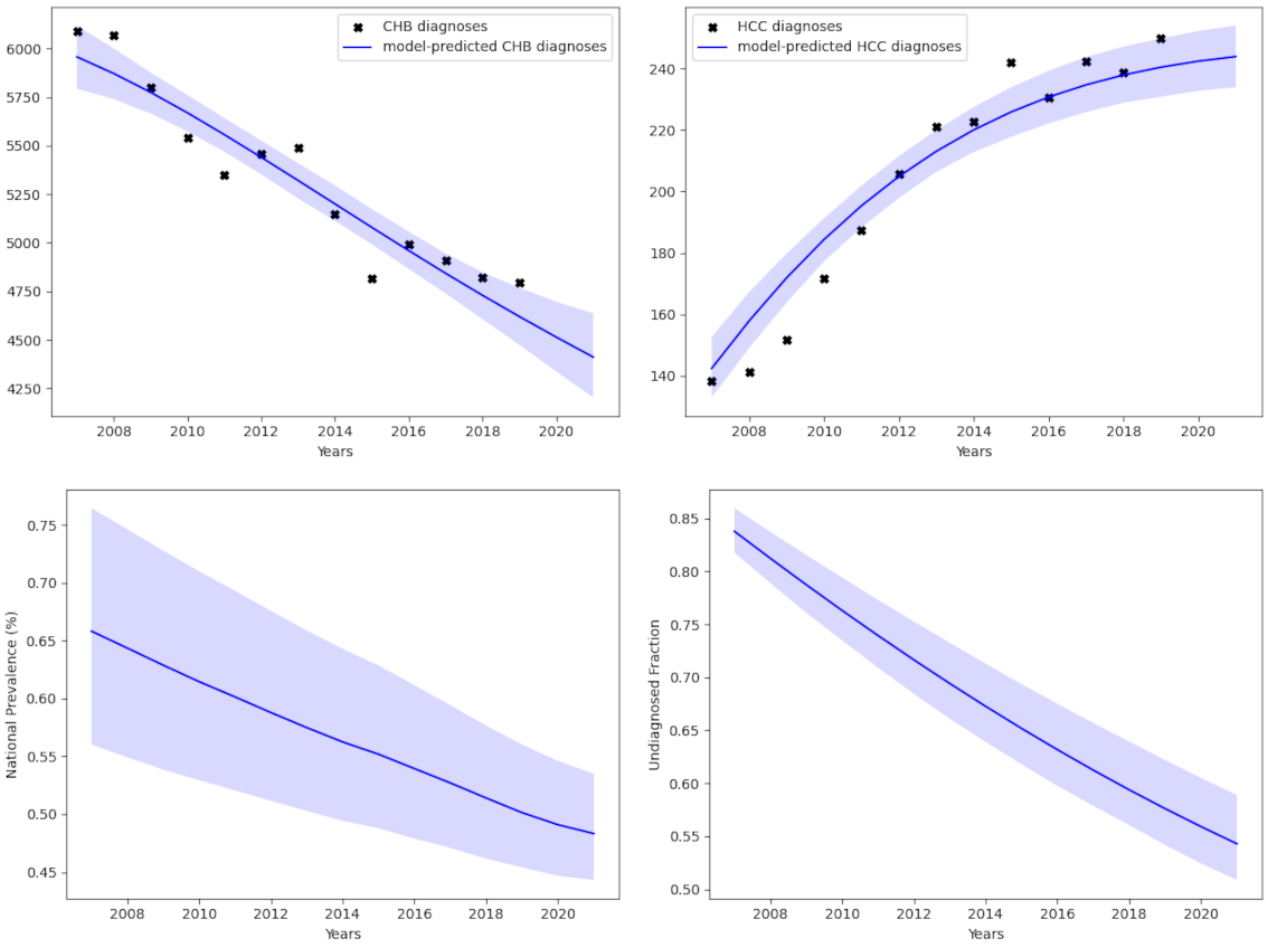
This study uses publicly deidentified data collected by Public Health Agency of Canada and Statistics Canada. The analysis procedure has been approved by the Research Board at the University of Waterloo (#43995).

Results

After completing the calibration process outlined earlier, the algorithm identified a set of model parameters and initial conditions that can closely match with the observed statistics on CHB diagnosis and HCC diagnosis (Figure 2). In order to produce such a fit, the national prevalence of CHB in 2021 was estimated to be 0.483% (95% CI 0.443% - 0.535%), while the proportion undiagnosed in 2021 was estimated to be 54.3% (95% CI 50.9% - 58.9%; Figure 2). The model had a deviance information criterion of 53.9.

XSL•FO
RenderX

Figure 2. Estimates of the national chronic hepatitis B (CHB) prevalence across all cohorts. The top 2 plots show the model fit compared to data for CHB diagnoses (top left) and CHB-attributed hepatocellular carcinoma (HCC) diagnoses (top right). The bottom 2 plots accordingly display the estimates of prevalence (bottom left) and of the undiagnosed proportion (bottom right).

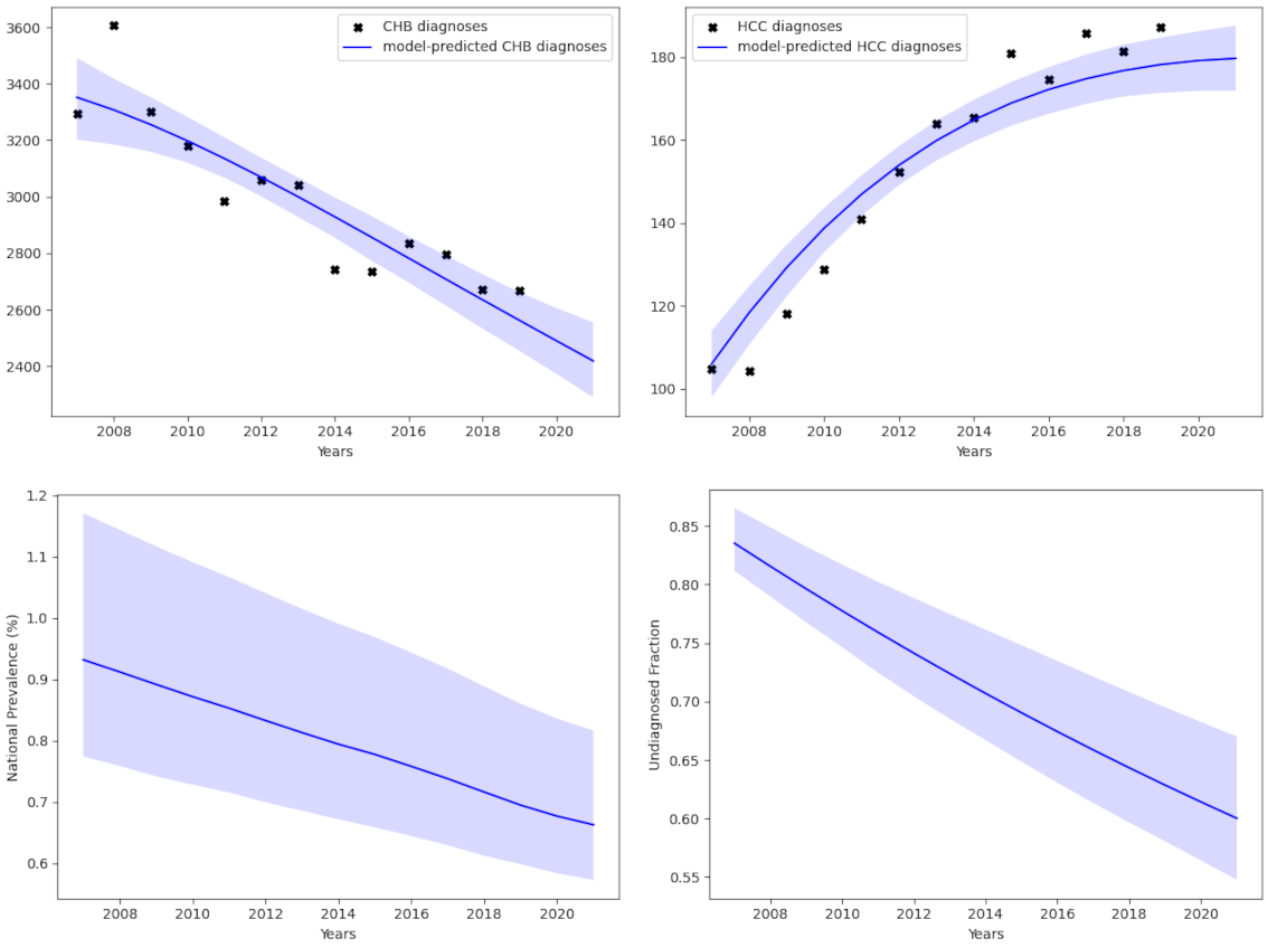


Stratified Analyses

We conducted a second calibration stratified by sex. Again, the algorithm identified a set of model parameters and initial conditions that closely match the observed statistics on CHB diagnosis and HCC diagnosis for the male population (Figure 3) and the female population (Figure 4). In order to produce these fits, the stratified analyses suggest a slightly higher prevalence of 0.663% (95% CI 0.572% - 0.816%) in the male population, while the proportion undiagnosed in 2021 was estimated to be 60.0% (95% CI 54.8% - 67.0%; Figure 3). In the female population, we obtain a slightly lower prevalence of 0.382% (95% CI 0.347% - 0.424%), while the proportion undiagnosed in 2021 was estimated to be 48.8% (95% CI 44.2% - 54.0%; Figure 4). The male and female fits had deviance information criteria of 9.0 and 73.2, respectively. The differences in prevalence obtained from this calibration process are consistent with the known sexual dimorphism in CHB [21].

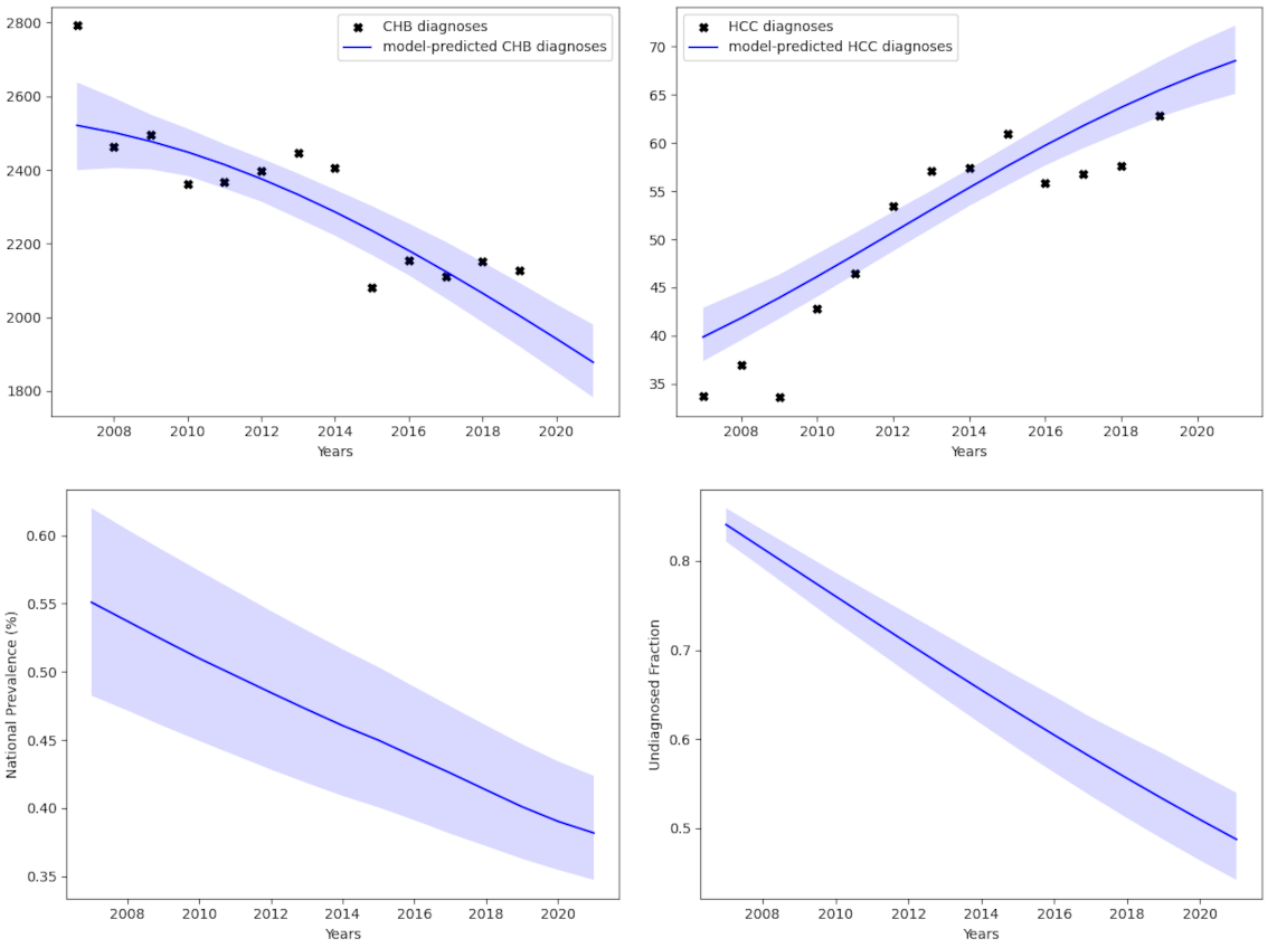
XSL•FO
RenderX

Figure 3. Estimates of the chronic hepatitis B (CHB) prevalence in males. The top two plots show the model fit compared to data for CHB diagnoses (top left) and CHB-attributed hepatocellular carcinoma (HCC) diagnoses (top right). The bottom 2 plots accordingly display the estimates of prevalence (bottom left) and of the undiagnosed proportion (bottom right).



XSL•FO
RenderX

Figure 4. Estimates of the chronic hepatitis B (CHB) prevalence in females. The top two plots show the model fit compared to data for CHB diagnoses (top left) and CHB-attributed hepatocellular carcinoma (HCC) diagnoses (top right). The bottom 2 plots accordingly display the estimates of prevalence (bottom left) and of the undiagnosed proportion (bottom right).



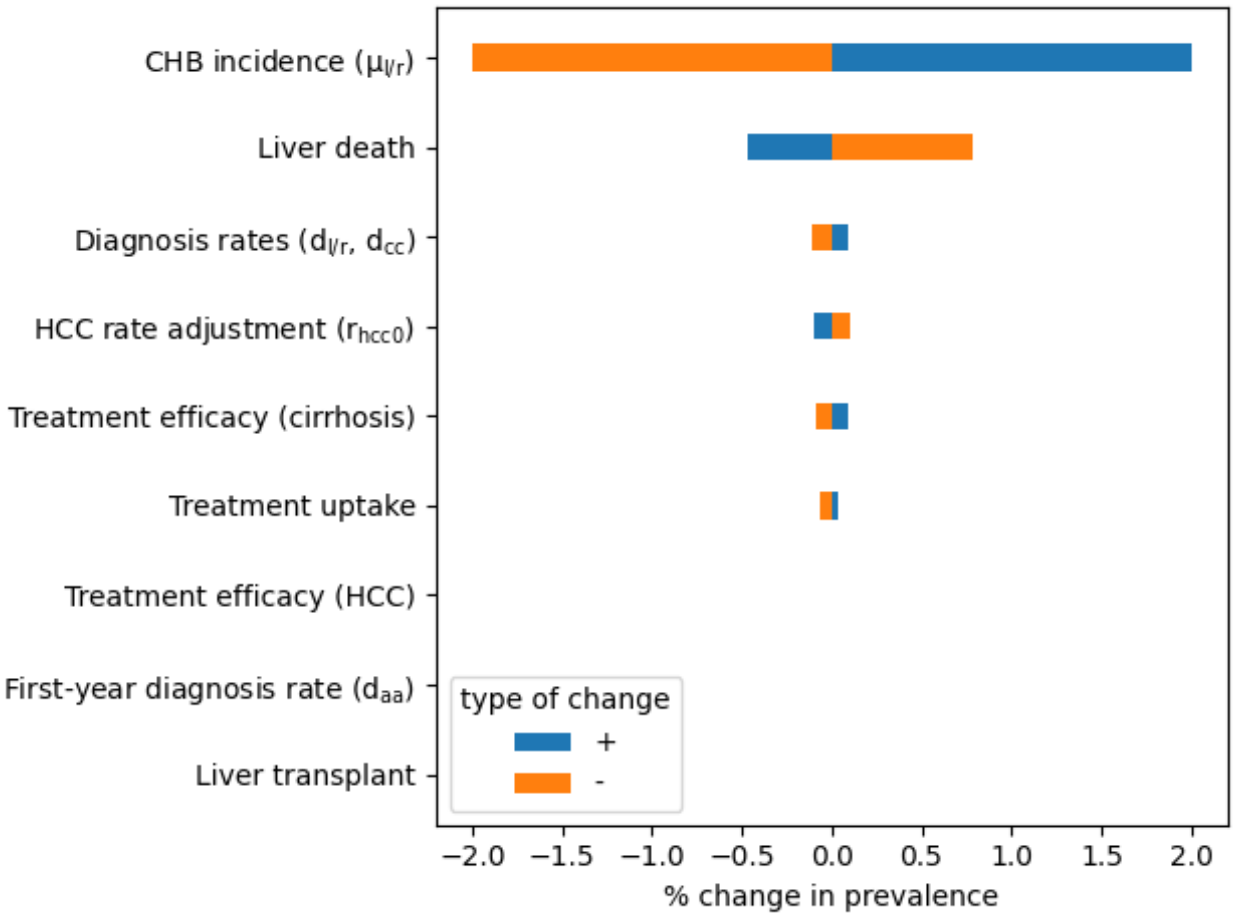
In addition to stratifying by sex, we also attempted an analysis based on age cohort. However, because late-stage complications are highly skewed toward older individuals, the dataset for the younger cohort was too small and noisy to make reliable inferences. We intend to address this in future work with better health administrative data.

Sensitivity Analysis

We performed a 1-way sensitivity analysis of the calibrated model, varying the parameters in Table 1 to determine which components of the model caused the greatest variation in the estimates of the prevalence and undiagnosed fraction. Parameters were varied by ±50%, and we measured the resulting change in the prevalence and diagnosed fraction, averaged over the time period studied (2007 - 2021). The results are presented in Figures 5 and 6.

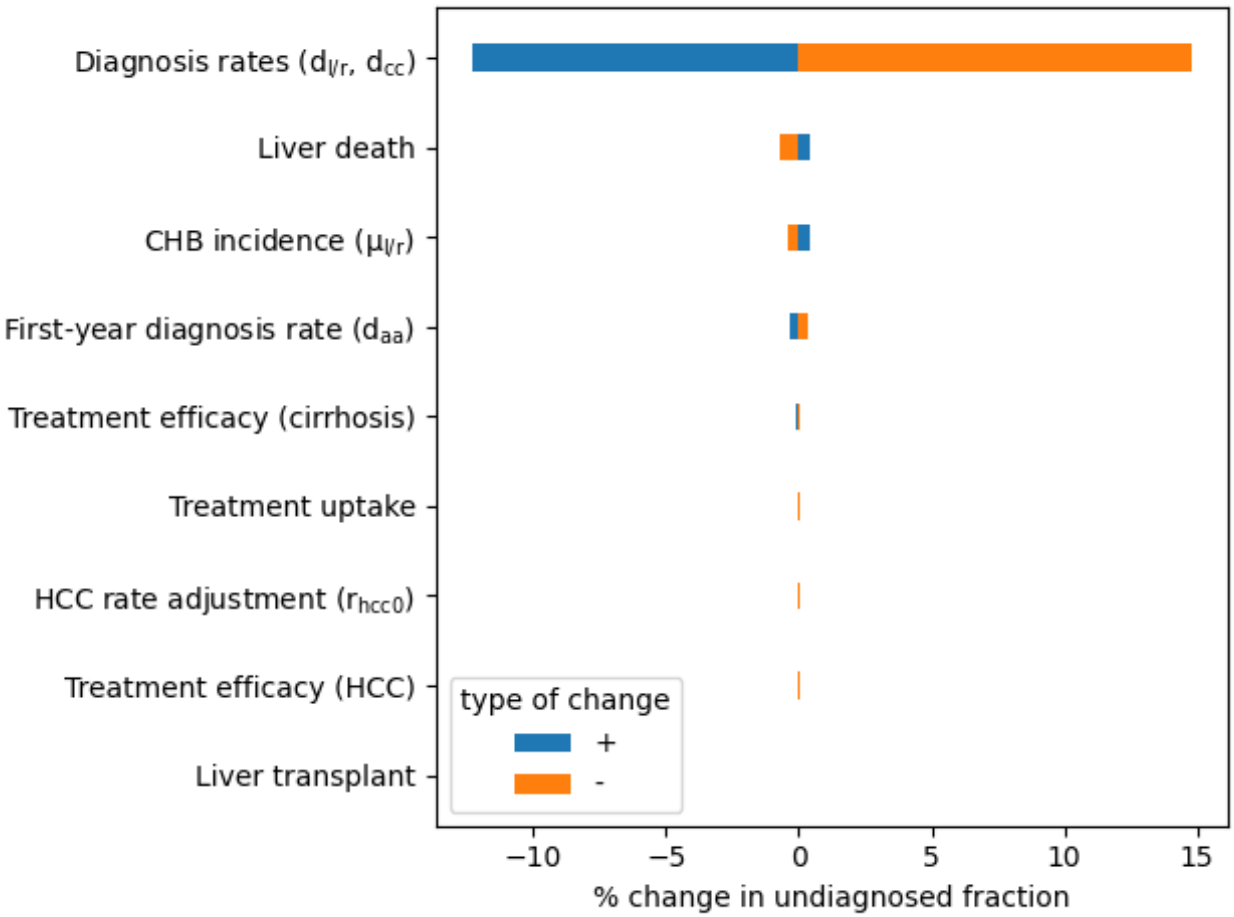
XSL•FO
RenderX

Figure 5. One-way sensitivity analysis for the estimated prevalence of chronic hepatitis B (CHB). HCC: hepatocellular carcinoma.



XSL•FO
RenderX

Figure 6. One-way sensitivity analysis for the estimated undiagnosed proportion of CHB infections. CHB: chronic hepatitis B; HCC: hepatocellular carcinoma.



As expected, the yearly CHB incidence has the largest impact on prevalence. Other parameter changes impact prevalence indirectly by changing the mortality rate. This effect was seen most strongly in the probability of liver death associated with late-stage complications. Changing the diagnosis probabilities or HCC incidence also increased prevalence, presumably because of their downstream effects on CHB-related mortality. Similarly, increasing the treatment uptake and treatment effectiveness increased prevalence due to the associated drop in mortality. The effects of daa on prevalence were negligible.

Changes in the diagnosis probabilities had the largest effect on the undiagnosed fraction, followed by the probability of liver death, CHB incidence, and daa, which had comparable effects. Changes in events downstream of diagnosis, such as HCC incidence or treatment effectiveness, did not change the overall prevalence enough to have a significant impact on the undiagnosed proportion.

Discussion

Principal Results

CHB has a long incubation period, often progressing undetected until the onset of late-stage complications. This creates significant uncertainties in the prevalence and undiagnosed proportion and poses challenges to policymakers who wish to measure the efficacy of public health interventions or track progress toward elimination targets. We have proposed a model-based method to estimate these quantities. Using limited CHB and HCC diagnosis data, we estimate prevalence in 2021 to be 0.483% (95% CI 0.443% - 0.535%). These results are in line with previous 0.4% seroprevalence estimates [3]. Additionally, the proposed method provides estimates of the fraction of CHB cases that remain undiagnosed, in this case 54.3% (95% CI 50.9% - 58.9%). Again, this is in line with previous estimates. The same Canadian seroprevalence study indicated that 46% of respondents who tested positive for a present HBV infection knew of their infected status.

These numbers are comparable to estimates obtained in other countries. In 2022, approximately 205,549 individuals in Australia were living with CHB, representing 0.78% of the population. It is estimated that 27.9% of these remain undiagnosed [22]. The prevalence of hepatitis B varies across European Union/European Economic Area countries, with estimates ranging from 0.1% in Ireland to 4.4% in Romania [23]. The proportion of undiagnosed infections also varies widely, ranging between 45% and 85%, highlighting gaps in national testing programs [24]. Although our results are focused on Canada, other high-income, immigrant-receiving countries like the United Kingdom and Australia may also benefit from applying the methods developed in this paper.

XSL•FO
RenderX

A 1-way sensitivity analysis identifies CHB incidence as the greatest factor influencing overall prevalence, followed by the probability of liver death in late-stage clinical health states, followed by treatment efficacy against cirrhosis. The undiagnosed fraction was sensitive to changes in diagnosis probabilities and CHB incidence, but treatment efficacy did not impact prevalence enough to have a significant effect on the undiagnosed proportion.

Limitations

The estimation method we have presented has several advantages. For one, it is systematic: the analysis we have performed can, in principle, be extended to any cohort for which data are available. Additionally, the 2-step calibration method accounts for uncertainty in the model parameters that are derived from the literature, yielding statistically robust estimates. However, a few important limitations constrain its use. We made the simplifying assumption that the CHB incidence is a linear function of time, but the 1-way sensitivity analysis identified CHB incidence as the parameter with the largest impact on prevalence. Therefore, the assumption of linearity greatly limits the time resolution of the model. As such, the time series presented in this paper should be taken as indicators of overall prevalence trends, rather than a detailed year-by-year accounting. In particular, the model incidence as currently formulated cannot accommodate sudden changes in vaccine uptake or immigration rates. This latter limitation may be addressed by immigrant sub-cohort modeling, though we currently lack the data required to perform this analysis. There is also considerable uncertainty surrounding HCC incidence. For one, the HCC data we use for calibration include HCC from all causes. Following [14], we estimate that hepatitis B accounts for 9% of HCC cases, but this estimate does not account for annual variation. Moreover, as with all cancers, HCC incidence is known to be highly age-dependent, but the available HCC data were not detailed enough to calibrate a state transition model that accounts for age. This latter limitation makes it impossible to reliably stratify our analysis by age cohort. Fortunately, both of these HCC-related limitations can be addressed with more accurate health administration data.

Conclusions

With these limitations in mind, our results point to some notable public health concerns. Despite Canada's 2016 endorsement of ambitious elimination targets, our 2021 CHB prevalence estimate does not indicate a significant decrease over previous estimates from 2013. Hepatitis B continues to be a substantial health burden for at least a hundred thousand Canadians. Of these, a significant proportion remains undiagnosed and untreated, putting them at increased risk of serious complications. Our analysis suggests that measures to decrease CHB incidence, such as vaccination and screening, have the greatest potential to decrease the overall disease burden. Based on our current data limitations, further stratified analysis is not possible. Our next step will be accessing health administrative data to overcome this barrier, with the aim of producing more reliable and detailed estimates.

We have presented a method to estimate the prevalence of CHB, using a state transition model based on the natural history of the disease. The estimates are obtained via a 2-stage calibration process that incorporates multiple sources of uncertainty, yielding results that are robust to perturbations of the model. This calibration process relies only on diagnosis data, which are already collected by the health care system, making it an attractive alternative to seroprevalence studies and other methods that require the collection of novel data. Conversely, a lack of high-quality health administrative data significantly limits the accuracy of the model. This is especially true among small populations and younger age cohorts, as younger individuals will typically not have developed a large and statistically robust pattern of late-stage complications. With more data, these estimates can further be stratified by region or by age cohort, allowing them to inform the targeting of public health interventions.

Acknowledgments

This study was supported by the Public Health Agency of Canada and Canadian Institutes of Health Research grants (178734 and 190792). The authors would like to express appreciation to Dr Curtis Cooper and Dr Carla S Coffin for their valuable suggestions and insights during the conduct of this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Literature-derived model parameters.
[DOCX File, 26 KB - publichealth_v11i1e66309_app1.docx]

Multimedia Appendix 2
Model dynamics and likelihood.
[DOCX File, 23 KB - publichealth_v11i1e66309_app2.docx]

Multimedia Appendix 3
Calibration data.

XSL•FO
RenderX

[DOCX File, 18 KB - publichealth_v11i1e66309_app3.docx]

References

1. Hepatitis B: Fact Sheet. World Health Organization. 2024. URL: https://www.who.int/en/news-room/fact-sheets/detail/hepatitis-b [accessed 2024-04-30]
2. Message from the Minister of Health - World Hepatitis Day. Government of Canada. 2016. URL: https://www.canada.ca/en/health-canada/news/2016/07/message-from-the-minister-of-health-world-hepatitis-day.html [accessed 2024-04-30]
3. Rotermann M, Langlois K, Andonov A, Trubnikov M. Seroprevalence of hepatitis B and C virus infections: results from the 2007 to 2009 and 2009 to 2011 Canadian Health Measures Survey. Health Rep 2013 Nov;24(11):3-13. [Medline: 24259199]
4. Highlights from the 2021 childhood National Immunization Coverage Survey. Government of Canada. 2023. URL: https://www.canada.ca/en/public-health/services/immunization-vaccines/vaccination-coverage/2021-highlights-childhood-national-immunization-coverage-survey.html [accessed 2024-04-30]
5. Provincial and territorial routine and catch-up vaccination schedule for infants and children in Canada. Government of Canada. 2023. URL: https://www.canada.ca/en/public-health/services/provincial-territorial-immunization-information/provincial-territorial-routine-vaccination-programs-infants-children.html [accessed 2024-04-30]
6. Tran TT. Management of hepatitis B in pregnancy: weighing the options. Cleve Clin J Med 2009 May;76 Suppl 3(Suppl 3):S25-S29. [doi: 10.3949/ccjm.76.s3.06] [Medline: 19465706]
7. Table 98-10-0349-01 Place of birth and period of immigration by gender and age. Statistics Canada. 2023. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=9810034901 [accessed 2025-03-25]
8. Ontario Agency for Health Protection and Promotion (Public Health Ontario). Immunization Coverage Report for School-Based Programs in Ontario: 2019-20, 2020-21 and 2021-22 School Years with Impact of Catch-up Programs. : King's Printer for Ontario; 2023.
9. Biondi MJ, Marchand-Austin A, Cronin K, et al. Prenatal hepatitis B screening, and hepatitis B burden among children, in Ontario: a descriptive study. Can Med Assoc J 2020 Oct 26;192(43):E1299-E1305. [doi: 10.1503/cmaj.200290] [Medline: 33106301]
10. Hamadeh A, Feng Z, Krahn M, Wong WWL. A model-based framework for chronic hepatitis C prevalence estimation. PLoS One 2019;14(11):e0225366. [doi: 10.1371/journal.pone.0225366] [Medline: 31751393]
11. Trubnikov M, Yan P, Archibald C. Estimated prevalence of Hepatitis C Virus infection in Canada, 2011. Can Commun Dis Rep 2014 Dec 18;40(19):429-436. [doi: 10.14745/ccdr.v40i19a02] [Medline: 29769874]
12. Longini IM Jr, Byers RH, Hessol NA, Tan WY. Estimating the stage‐specific numbers of HIV infection using a markov model and back‐calculation. Stat Med 1992 Jan;11(6):831-843. [doi: 10.1002/sim.4780110612]
13. Hamadeh A, Feng Z, Niergarth J, Wong WW. Estimation of COVID-19 period prevalence and the undiagnosed population in Canadian Provinces: model-based analysis. JMIR Public Health Surveill 2021 Sep 9;7(9):e26409. [doi: 10.2196/26409] [Medline: 34228626]
14. Kim J. Model-based estimation of the prevalence of chronic hepatitis B in Canada and cost-effectiveness analysis of implementing a universal birth hepatitis B vaccination program in Ontario, Canada. : University of Waterloo; 2022.
15. Terrault NA, Lok ASF, McMahon BJ, et al. Update on prevention, diagnosis, and treatment of chronic hepatitis B: AASLD 2018 hepatitis B guidance. Hepatology 2018 Apr;67(4):1560-1599. [doi: 10.1002/hep.29800] [Medline: 29405329]
16. Table 13-10-0111-01—Number and rates of new cases of primary cancer, by cancer type, age group and sex. Statistics Canada. 2024. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310011101 [accessed 2024-04-30]
17. Notifiable Diseases Online. Public Health Agency of Canada. 2024. URL: https://diseases.canada.ca/notifiable/ [accessed 2024-04-30]
18. Lapointe-Shaw L, Chung H, Holder L, et al. Diagnosis of chronic hepatitis B pericomplication: risk factors and trends over time. Hepatology 2021 Jun;73(6):2141-2154. [doi: 10.1002/hep.31557] [Medline: 32931613]
19. Wong WWL, Woo G, Jenny Heathcote E, Krahn M. Cost effectiveness of screening immigrants for hepatitis B. Liver Int 2011 Sep;31(8):1179-1190. [doi: 10.1111/j.1478-3231.2011.02559.x] [Medline: 21745300]
20. Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 1970 Apr 1;57(1):97-109. [doi: 10.1093/biomet/57.1.97]
21. Brown R, Goulder P, Matthews PC. Sexual dimorphism in chronic hepatitis B virus (HBV) infection: evidence to inform elimination efforts. Wellcome Open Res 2022;7:32. [doi: 10.12688/wellcomeopenres.17601.2] [Medline: 36212217]
22. Prevalence and epidemiology of hepatitis B. Australasian Society for HIV, Viral Hepatitis and Sexual Health Medicine (ASHM). 2022. URL: https://hepatitisb.org.au/prevalence-and-epidemiology-of-hepatitis-b/ [accessed 2025-03-25]
23. Systematic review on hepatitis B and C prevalence in the EU/EEA. European Centre for Disease Prevention and Control. 2016. URL: https://www.ecdc.europa.eu/en/publications-data/systematic-review-hepatitis-b-and-c-prevalence-eueea [accessed 2025-03-25]
24. ECDC: around 9 million Europeans are affected by chronic hepatitis B or C. European Centre for Disease Prevention and Control. 2017. URL: https://www.ecdc.europa.eu/en/news-events/ecdc-around-9-million-europeans-are-affected-chronic-hepatitis-b-or-c [accessed 2025-03-25]

XSL•FO
RenderX

Abbreviations

CC: compensated cirrhosis
CHB: chronic hepatitis B
DC: decompensated cirrhosis
HBeAg: hepatitis B e-antigen
HBsAg: hepatitis B surface antigen
HBV: hepatitis B virus
HCC: hepatocellular carcinoma
LT: liver transplant
PHAC: Public Health Agency of Canada
PLT: post–liver transplant

Edited by VA Adepoju; submitted 16.09.24; peer-reviewed by C Cohen, TH Su; revised version received 09.04.25; accepted 02.05.25; published 20.08.25.

Please cite as:
Smith-Roberge J, Forouzannia F, Hamadeh A, Feng Z, Popovic N, Wong WWL
Estimating Chronic Hepatitis B Prevalence and Undiagnosed Proportion in Canada, 2007-2021: Mathematical Framework Development
JMIR Public Health Surveill 2025;11:e66309
URL: https://publichealth.jmir.org/2025/1/e66309
doi:10.2196/66309

© Julien Smith-Roberge, Farinaz Forouzannia, Abdullah Hamadeh, Zeny Feng, Nashira Popovic, William W L Wong. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 20.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Antivirals Nirmatrelvir-Ritonavir and Molnupiravir in Viral Sepsis: Retrospective Cohort Study

Teddy Tai Loy Lee[1], BPharm, MPH; Alex Chang-Hao Lyu[1], MBBS; Ting Ting Jiang[1], MBBS; Sunny Ching Long Chan[1], BSc, MStat; Crystal Ying Chan[2], PhD, DFPH; Edmond Tsz Fung Yip[3], BEng; Luke Yik Fung Luk[3], BSc; Joshua Wing Kei Ho[3], PhD; Kevin Wang Leong So[1], BPharm; Omar Wai Kiu Tsui[1], MPH; Man Lok Lam[1], MBBS; Shi Yeow Lee[1], MBBS; Tafu Yamamoto[4], MBChB; Chak Kwan Tong[5], MBBS; Man Sing Wong[6], PhD; Eliza Lai Yi Wong[2], PhD; Abraham Ka Chung Wai[1], MD; Timothy Hudson Rainer[1], MD

[1]Department of Emergency Medicine, The University of Hong Kong, G06, 5 Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

[2]Centre for Health Systems & Policy Research, JC School of Public Care & Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[3]School of Biomedical Sciences, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)

[4]Yan Chai Hospital, Hong Kong Hospital Authority, Hong Kong, China (Hong Kong)

[5]The University of Hong Kong, Hong Kong, China (Hong Kong)

[6]Department of Land Surveying and Geo-informatics, Faculty of Construction and Environment, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

Corresponding Author:
Abraham Ka Chung Wai, MD
Department of Emergency Medicine, The University of Hong Kong, G06, 5 Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

Abstract

Background: Viral infections, including those leading to sepsis, are common but often overlooked in clinical practice, yet the treatment strategies for viral sepsis remain inadequately defined.

Objective: This study aims to investigate the effectiveness of antivirals nirmatrelvir-ritonavir and molnupiravir in the treatment of culture-negative sepsis.

Methods: This retrospective cohort study was conducted across public hospitals in Hong Kong. We included patients diagnosed with COVID-19 between February 22, 2022, and June 30, 2023, who had no secondary bacterial or fungal infections. Propensity score matching was used to assess the efficacy of the antivirals nirmatrelvir-ritonavir and molnupiravir in patient subgroups with or without organ dysfunction at hospital admission, including circulatory shock, respiratory failure, acute kidney injury, coagulopathy, acute liver impairment, a composite of all organ dysfunctions, or no organ dysfunction. Key outcomes were in-hospital mortality and length of stay, reported as hazard ratios (HR) and mean differences, respectively.

Results: The study included 15,599 COVID-19 patients with a mean age of 75.1 (SD 15.9) years. Molnupiravir treatment was associated with a significantly lower risk of mortality in patients in both the presence of any organ dysfunction (HR 0.75, 95% CI 0.58 to 0.96) and without organ dysfunction (HR 0.29, 95% CI 0.15-0.56). Nirmatrelvir-ritonavir was associated with decreased mortality with respiratory failure (absolute risk difference: 9.5%, 95% CI 6.26-12.72) and without organ dysfunction (HR 0.17, 95% CI 0.05-0.56). Antivirals also reduced the length of hospital stay; nirmatrelvir-ritonavir reduced length of stay in respiratory failure by an average of 3.37 (95% CI 2.32-4.42) days, acute kidney injury by 7.25 (95% CI 2.97-11.52) days, and coagulopathy by 7.04 (95% CI 2.99-4.05) days. Molnupiravir reduced the length of stay in acute kidney injury by an average of 6.7 (95% CI 2.39-11.08) days and coagulopathy by 5.68 (95% CI 1.20-10.16) days.

Conclusions: Antivirals reduced mortality among hospitalized COVID patients, with the greatest reduction observed in patients without organ dysfunction. Antivirals were also effective in reducing the length of hospital stay.

(JMIR Public Health Surveill 2025;11:e72124) doi:10.2196/72124

KEYWORDS

organ dysfunction; viral infection; sepsis; antivirals; COVID-19

XSL•FO
RenderX

Introduction

Sepsis, a syndrome characterized by organ dysfunction resulting from a dysregulated response to infection, poses a significant mortality risk [1]. While bacterial and fungal infections are well-recognized causes of sepsis, viral infections—especially in older people and immunocompromised individuals—are often overlooked despite their potential to cause severe illness [2].

Extensive literature has primarily focused on bacterial sepsis, with early antibiotic administration being the cornerstone of its treatment [3,4]. However, the emergence of COVID-19, caused by the SARS-CoV-2, has sparked renewed interest in the role of viral infections in sepsis. Despite ongoing vaccination efforts that have reduced the severity and mortality associated with COVID-19 [5], the disease continues to significantly impact vulnerable populations, particularly older people.

A meta-analysis revealed that the majority of COVID-19 patients admitted to the ICU met the Sepsis-3 criteria [6]. During the early phases of the pandemic, COVID-19 accounted for 1 in 6 sepsis cases [7]. COVID-19 is associated with a range of organ dysfunctions, including circulatory shock, respiratory failure, acute kidney injury, liver impairment, and coagulopathy [8]. Yet, the use of antivirals in patients across varying organ dysfunction presentations remains inadequately understood. On a related note, recent studies support nirmatrelvir-ritonavir use in other critical conditions, such as in patients with advanced kidney disease [9], in hematologic malignancies [10], and in immunosuppressed patients [11], where antivirals reduced length of stay, shortened time to symptom resolution, and reduced viral load.

Viral sepsis remains underdiagnosed due to its lack of specific clinical manifestations [12]. Early research indicates that viral sepsis often occurs without secondary bacterial and fungal infections; for instance, 76% of blood and sputum cultures from COVID-19 patients were negative for bacteria and fungi [13].

Currently, guidelines offer limited recommendations for the use of antivirals in patients with viral sepsis [14,15], highlighting the need for further research to establish effective treatment protocols. This study aims to evaluate the effectiveness of antiviral treatment in a cohort of COVID-19 patients who had negative blood culture results during their hospital stay, thereby contributing to the evidence base for managing viral sepsis.

Methods

Cohort Identification

All patients included in this study were identified using the Clinical Data Analysis and Reporting System (CDARS), a comprehensive territory-wide electronic health care database encompassing all public hospitals managed by the Hospital Authority, which serves a population of 7.1 million. CDARS has been reliably used for numerous large-scale population-based studies [16], including research on antiviral use for COVID-19 [17,18], where diagnoses are stored using *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes. Each patient's data, including prescription, admission, and diagnosis, were linked using a unique anonymous identifier. Patients were potentially included in this study if they had related service records, such as admission and medication use within CDARS.

Study Design

This retrospective cohort study used CDARS to investigate culture-negative viral infection among COVID-19 inpatients. The study compared the effect of antiviral treatments (nirmatrelvir-ritonavir or molnupiravir) versus no antiviral treatment (control) on 30-day in-hospital mortality and length of stay, across various conditions of organ dysfunction.

Eligibility Criteria

The cohort selection is detailed in Figure S1 in Multimedia Appendix 1. Eligible patients were those aged 18 years and older with a diagnosis of COVID-19, identified using *ICD-9-CM* code 519.8(8). Patients admitted between February 22, 2022, and June 30, 2023, were included in the study. The baseline time was defined as the date of hospital admission. To meet the culture-negative criteria, patients had to have at least one blood culture record during admission. Within the first 21 days of admission, no organism should have been identified from the culture, or only "negative organisms" should have been considered irrelevant to the viral infection course, as defined in Table S1 in Multimedia Appendix 1. Only patients who were culture-negative were included. As the study focus is on sepsis, other cultures, such as sputum culture, were not used to fulfill the culture-negative criteria; moreover, viral co-infections were not excluded, as they are within the scope of the study.

Organ Dysfunction

Patients were categorized into one of 6 organ dysfunction groups (circulatory shock, respiratory failure, acute kidney injury, acute liver impairment, coagulopathy, composite organ dysfunction) based on the presence of organ dysfunction within the first 3 days of admission, as defined in Table S2 in Multimedia Appendix 1. These groups use the same method used to identify organ dysfunction according to the Hospital Toolkit for Adult Sepsis Surveillance released by the US Center for Disease Control and Prevention [19]. Patients without any organ dysfunction during the entire follow-up period and without death in the first 3 days of hospitalization were categorized into the no organ dysfunction group, which represents COVID-19 patients with mild to moderate illness severity.

Exposure

The study examined the treatment strategy of using antivirals versus not using antivirals. Specifically, there were two primary anti-exposure interests: (1) nirmatrelvir-ritonavir and (2) molnupiravir. Each antiviral group was compared with a control group consisting of patients who did not receive any antiviral treatment during their hospital stay. We used a new-user design, including only patients who had received antiviral treatment within the 3 days of admission [20], and excluded those who began antiviral treatment before admission. Controls were defined as patients who did not use antivirals at any point during the study period. To ensure the selection of new users, we included only patients who had not received antivirals within

XSL•FO
RenderX

7 days prior to admission and excluded those concurrently using both antivirals.

Follow-Up Period and Outcomes

Patients were followed until one of the following events occurred: death, discharge, or 28 days after admission. The primary outcome of interest was in-hospital mortality and length of stay, defined by the time difference between the date of admission and the end of follow-up.

Covariates

Baseline covariates included: age, sex, residency in care homes for older people, health status over the past 2 years as indicated by the Charlson Comorbidity Index, prior antibiotic use, the calendar year and month of COVID-19 diagnosis, and socioeconomic status as indicated by the Social Deprivation Index (SDI) of their residential address [21,22]. An SDI quartile of Q4 indicated the most deprived, while Q1 indicated the least deprived. To account for evolving COVID-19 characteristics due to emerging variants, the month of admission was included in the propensity score model. Patients with missing data were excluded from analyses.

Statistical Analysis

We hypothesized that antiviral use (nirmatrelvir-ritonavir or molnupiravir) would be associated with a decreased risk of mortality and a shorter length of stay compared with no antiviral use. Propensity score matching was used to balance baseline characteristics among the groups, using a 1:2 ratio with the nearest neighbor algorithm. A caliper width of 0.2 was chosen to optimize matching [23]. Covariate balance postmatching was evaluated using standardized mean differences (SMD), with a SMD value <0.2 considered indicative of balance. Descriptive statistics were reported as mean (SD) and count (percentage [%]) wherever appropriate. Estimates for in-hospital mortality were presented as hazard ratios (HR) with 95% CIs, derived from a univariate Cox proportional hazards model. The proportional hazards assumption was tested using the Schoenfeld Residual test, which confirmed that the assumption was met. Length of stay estimates were expressed as mean difference with 95% CI. Absolute risk difference, reported as percentages (%) with 95% CI, is calculated as the difference between absolute risks of antiviral and control groups, which are individually calculated as (number of events / number of patients × 100%).

Additionally, 2 sensitivity analyses were conducted: excluding patients who received adjuvant therapies such as remdesivir, interferon-beta 1b, tocilizumab, baricitinib, and plasma infusions during hospitalization; and a landmark analysis of in-hospital mortality on the 3rd day of admission.

All analyses were performed using R (version 4.4.1; R Foundation for Statistical Computing). All tests were 2-tailed, and a *P* value <.05 was considered statistically significant. Generative artificial intelligence was not used in any portion of the composition of the manuscript.

Ethical Considerations

This study received approval from the Hospital Authority Hong Kong West Cluster/The University of Hong Kong Institutional Review Board (UW 20 - 112) and was carried out in accordance with the Helsinki Declaration. The requirement for obtaining informed consent from patients was waived by the institutional review board, as all patient data collection and extraction was anonymized. No compensation was received for human participants' research, and no identification of individual participants is possible in the study materials.

Results

Study Population

Following the exclusion of patients who did not meet the culture-negative criteria or fit within defined organ dysfunction groups, our analysis encompassed 15,599 patients (Table 1). The mean age of the cohort was 75.10 (SD 15.86) years. Among those identified with organ dysfunction, 511 patients experienced circulatory shock, 5236 had respiratory failure, 1053 had acute kidney injury, 880 had coagulopathy, and 162 exhibited acute liver impairment. Notably, respiratory failure frequently co-occurred with other organ dysfunctions, manifesting in 386 patients (76.1%) with circulatory shock, 632 (60.5%) patients with acute kidney injury, 521 (59.5%) patients with coagulopathy, and 76 (46.9%) patients with acute liver impairment. A total of 5695 patients had at least one type of organ dysfunction upon admission, classifying them under the composite organ dysfunction group.

XSL•FO
RenderX

Table . Baseline characteristics of COVID-19 patients before propensity score matching.[a]

	Circulatory shock (n=511)	Respiratory failure (n=5236)	Acute kidney injury (n=1053)	Coagulopathy (n=880)	Acute liver impairment (n=162)	No organ dysfunction (n=7757)	Composite organ dysfunction (n=5695)
Sex (male), n (%)	351 (68.7)	3117 (59.5)	638 (60.6)	528 (60.0)	112 (69.1)	3930 (50.7)	3384 (59.4)
Age (years), mean (SD)	73.22 (13.83)	79.14 (13.08)	76.73 (13.96)	76.58 (13.66)	72.96 (16.09)	72.16 (17.44)	78.28 (13.50)
Admission of aged care resident, n (%)	60 (11.7)	1189 (22.7)	164 (15.6)	118 (13.4)	14 (8.6)	1280 (16.5)	1224 (21.5)
Charlson Comorbidity Index, n (%)							
0	53 (10.4)	132 (2.5)	34 (3.2)	31 (3.5)	12 (7.4)	805 (10.4)	166 (2.9)
1	30 (5.9)	187 (3.6)	50 (4.7)	30 (3.4)	8 (4.9)	553 (7.1)	232 (4.1)
2	93 (18.2)	522 (10.0)	140 (13.3)	113 (12.8)	23 (14.2)	1058 (13.6)	621 (10.9)
3	144 (28.2)	950 (18.1)	208 (19.8)	188 (21.4)	41 (25.3)	1551 (20.0)	1073 (18.8)
4+	222 (43.4)	3445 (65.8)	621 (59.0)	518 (58.9)	78 (48.1)	3790 (48.9)	3603 (63.3)
Social Deprivation Index, n (%)							
Q1[b]	22 (4.3)	754 (14.4)	121 (11.5)	96 (10.9)	26 (16.0)	1020 (13.1)	779 (13.7)
Q2	30 (5.9)	1795 (34.3)	395 (37.5)	325 (36.9)	58 (35.8)	2794 (36.0)	2011 (35.3)
Q3	93 (18.2)	1899 (36.3)	408 (38.7)	342 (38.9)	56 (34.6)	2853 (36.8)	2068 (36.3)
Q4	222 (43.4)	788 (15.0)	129 (12.3)	117 (13.3)	22 (13.6)	1090 (14.1)	837 (14.7)
Previous antibiotic use, n (%)	341 (66.7)	3413 (65.2)	603 (57.3)	497 (56.5)	87 (53.7)	3480 (44.9)	3540 (62.2)
Antiviral use upon admission, n (%)							
Nirmatrelvir-ritonavir	37 (7.2)	490 (9.4)	92 (8.7)	88 (10.0)	19 (11.7)	2487 (32.1)	596 (10.5)
Molnupiravir	35 (6.8)	329 (6.3)	123 (11.7)	109 (12.4)	23 (14.2)	1596 (20.6)	482 (8.5)
Coexisting organ dysfunction upon admission, n (%)							
Circulatory shock	—[c]	386 (7.4)	189 (18.1)	116 (13.3)	18 (11.1)	—	—
Respiratory failure	386 (76.1)	—	632 (60.5)	521 (59.5)	76 (46.9)	—	—
Acute kidney injury	189 (37.0)	632 (12.1)	—	344 (39.1)	32 (19.8)	—	—
Coagulopathy	116 (22.7)	521 (10.0)	344 (32.8)	—	37 (22.8)	—	—
Acute liver impairment	18 (3.5)	76 (1.5)	32 (3.1)	37 (4.2)	—	—	—

[a]COVID-19 patients were retrospectively identified from electronic health care records in Hong Kong. The most common organ dysfunction presenting in patients was respiratory failure (n=5236), followed by acute kidney injury (n=1053) and coagulopathy (n=880). Most patients had used antibiotics before receiving antiviral treatment.

[b]Q: Quartile.

[c]Not applicable.

Propensity Score Matching

To ensure comparability between treatment groups, propensity score matching was conducted, resulting in 14 well-balanced cohorts, with all matched variables achieving an SMD of less than 0.2. The distribution of propensity score postmatching is depicted in Figure S2 in Multimedia Appendix 1.

In-Hospital Mortality

In-hospital mortality data, along with cohort sizes and death counts across comparator groups, are detailed in Table 2. Antiviral use was associated with significantly decreased mortality in patients without organ dysfunction, for both nirmatrelvir-ritonavir (HR 0.17, 95% CI 0.05-0.56) and molnupiravir (HR 0.29, 95% CI 0.15-0.56). Molnupiravir was

XSL•FO
RenderX

associated with a significantly reduced mortality hazard (HR 0.75, 95% CI 0.58-0.96) among patients with composite organ dysfunction, and a marginally significant reduction (HR 0.69, 95% CI 0.46-1.02) among patients with respiratory failures. Nirmatrelvir-ritonavir was also associated with reduced absolute mortality risk among patients with respiratory failure (absolute risk difference: 9.5%, 95% CI 6.26%-12.72%) and composite organ dysfunction (absolute risk difference: 6.0%, 95% CI 3.18%-8.91%). In other subgroups of patients with individual organ dysfunction, the reductions in mortality were not statistically significant and had wide CIs.

Table . In-hospital mortality grouped by organ dysfunction and use of antiviral medication after propensity score matching.

Treatment	Antiviral				Control					
	Patients, n	Events, n	Patient-days, n	Absolute risk, %	Patients, n	Events, n	Patient-days, n	Absolute risk, %	Absolute risk difference, % (95% CI)	Hazard ratio (95% CI)
Nirmatrelvir-ritonavir										
Circulatory shock	34	5	501	14.7	66	18	1120	27.3	12.6 (−3.47 to 28.60)	0.67 (0.25 to 1.83)
Respiratory failure	490	34	3789	6.9	980	161	10,879	16.4	9.5 (6.26 to 12.72)[a]	0.73 (0.51 to 1.06)
Acute kidney injury	87	17	1468	19.5	172	52	4149	30.2	10.7 (−0.10 to 21.48)	0.77 (0.44 to 1.34)
Coagulopathy	88	13	1570	14.8	173	33	4304	19.1	4.3 (−5.14 to 13.75)	0.98 (0.52 to 1.88)
Acute liver impairment	14	0	148	0.0	26	3	409	11.5	11.5 (−0.74 to 23.82)	—[b]
No organ dysfunction	1581	3	6957	0.2	1661	47	13,152	2.8	2.6 (1.81 to 3.47)[a]	0.17 (0.05 to 0.56)[a]
Composite organ dysfunction	725	71	7637	9.8	1446	229	20,966	15.8	6.0 (3.18 to 8.91)[a]	0.88 (0.67 to 1.14)
Molnupiravir										
Circulatory shock	34	7	619	20.6	65	11	958	16.9	−3.69 (−20.03 to 12.70)	1.21 (0.47 to 3.13)
Respiratory failure	327	33	3657	10.1	653	97	7226	14.9	4.8 (0.51 to 9.02)	0.69 (0.46 to 1.02)
Acute kidney injury	123	20	2360	16.3	242	73	6273	30.2	13.9 (5.19 to 22.62)[a]	0.77 (0.44 to 1.34)
Coagulopathy	109	18	2110	16.5	213	50	5333	23.5	7.0 (−2.04 to 15.96)	0.90 (0.53 to 1.55)
Acute liver impairment	23	4	254	17.4	43	7	698	16.3	−1.09 (−20.13 to 17.91)	0.70 (0.43 to 1.15)
No organ dysfunction	1533	10	8927	0.7	1801	64	14,568	3.59	2.9 (1.96 to 3.85)[a]	0.29 (0.15 to 0.56)[a]
Composite organ dysfunction	619	83	9036	13.4	1237	225	18,641	18.2	4.8 (1.34 to 8.22)[a]	0.75 (0.58 to 0.96)[a]

[a]Denotes statistically significant. Absolute risk difference is calculated as the difference between absolute risks of antiviral and control groups, which are individually calculated as (number of events / number of patients×100%). Reduced in-hospital mortality was seen for nirmatrelvir-ritonavir users in respiratory failure, composite organ dysfunction, and no organ dysfunction; and for molnupiravir users in acute kidney injury, composite organ dysfunction and no organ dysfunction. Due to small sample size in the acute liver impairment group using nirmatrelvir-ritonavir, hazard ratio was not calculated for that group.

[b]Not available.

XSL•FO
RenderX

Length of Stay

Our analysis demonstrated that the use of both nirmatrelvir-ritonavir and molnupiravir was associated with a reduced length of hospital stay for patients with various types of organ dysfunction at admission, detailed in Table 3. Specifically, nirmatrelvir-ritonavir significantly reduced the length of stay in patients with respiratory failure by an average (mean) of 3.37 (95% CI 2.32-4.42) days; acute kidney injury by 7.25 (95% CI 2.97-11.52) days; coagulopathy by 7.04 (95% CI 2.44-11.64) days, and in those without organ dysfunction by 3.52 (95% CI 2.99-4.05) days. Additionally, it reduced the length of stay by 3.97 days in patients with composite organ dysfunction (95% CI 2.77-5.17). However, no significant reduction was observed in patients with circulatory shock (2.23, 95% CI −4.54 to 9.00 days) and acute liver impairment (5.16, 95% CI −0.96 to 11.28 days).

Table . Length of stay (days) grouped by organ dysfunction and use of antiviral medication.

Treatment	Antiviral, mean (SD)	Control, mean (SD)	Mean difference (95% CI)
Nirmatrelvir-ritonavir			
Circulatory shock	14.74 (16.26)	16.97 (15.66)	2.23 (−4.54 to 9.00)
Respiratory failure	7.73 (8.62)	11.1 (11.46)	3.37 (2.32 to 4.42)[a]
Acute kidney injury	16.87 (12.58)	24.12 (22.3)	7.25 (2.97 to 11.52)[a]
Coagulopathy	17.84 (13.7)	24.88 (23.96)	7.04 (2.44 to 11.64)[a]
Acute liver impairment	10.57 (8.36)	15.73 (10.22)	5.16 (−0.96 to 11.28)
No organ dysfunction	4.40 (4.17)	7.92 (10.18)	3.52 (2.99 to 4.05)[a]
Composite organ dysfunction	10.5 (11.2)	14.5 (17.0)	3.97 (2.77 to 5.17)[a]
Molnupiravir			
Circulatory shock	18.21 (22.53)	14.74 (14.86)	−3.47 (−12.07 to 5.14)
Respiratory failure	11.18 (11.26)	11.07 (10.8)	−0.12 (−1.60 to 1.36)
Acute kidney injury	19.19 (17.53)	25.92 (24.03)	6.73 (2.39 to 11.08)[a]
Coagulopathy	19.36 (16.94)	25.04 (23.34)	5.68 (1.20 to 10.16)[a]
Acute liver impairment	11.04 (7.85)	16.23 (13.82)	5.19 (−0.14 to 10.52)
No organ dysfunction	5.82 (5.93)	8.09 (9.99)	2.27 (1.72 to 2.81)[a]
Composite organ dysfunction	14.6 (14.9)	15.1 (17.9)	0.47 (−1.06 to 2.01)

[a]Denotes statistically significant. Nirmatrelvir-ritonavir reduced length of stay in respiratory failure, acute kidney injury; while molnupiravir reduced length of stay in acute kidney injury and coagulopathy.

Molnupiravir also showed a significant reduction in the length of stay for patients with acute kidney injury by an average (mean) of 6.7 (95% CI 2.39-11.08) days, coagulopathy by 5.68 (95% CI 1.20-10.16) days, and in those without organ dysfunction by 2.27 (95% CI 1.72-2.81) days. However, no significant reduction was observed for circulatory shock (−3.47, 95% CI −12.07 to 5.14 days), respiratory failure (−0.12, 95% CI −1.60 to 1.36 days), acute liver impairment (5.19, 95% CI −0.14 to 10.52 days), or composite organ dysfunction (0.47, 95% CI −1.06 to 2.01 days).

Sensitivity Analysis

Results excluding patients who received adjuvant therapy are described in Tables S3 and S4 in Multimedia Appendix 1.

The sensitivity analysis revealed reduced mortality associated with nirmatrelvir-ritonavir in respiratory failure (absolute risk difference: 8.3%, 95% CI 2.4%-14.2%) and molnupiravir use in patients with coagulopathy (absolute risk difference: 13.0%, 95% CI 2.2%-23.8%). Consistent with the main analysis, both antivirals were associated with decreased mortality in patients without organ dysfunction and composite organ dysfunction.

Nirmatrelvir-ritonavir use was associated with a decreased length of stay in patients with respiratory failure by an average of 8.4 (95% CI 5.1-11.7) days, acute kidney injury by 11.6 (95% CI 4.1-19.1) days, coagulopathy by 12.8 (95% CI 4.7-20.9) days, and those without organ dysfunction by 3.7 (95% CI 3.1-4.3) days. Molnupiravir was associated with a decreased length of stay in patients with acute kidney injury by an average of 8.1 (95% CI 1.5-14.7) days, coagulopathy by 7.8 (95% CI 0.2-15.4) days, and those without organ dysfunction by 2.2 (95% CI 0.5-6.2) days. Nirmatrelvir-ritonavir was associated with decreased length of stay in patients with composite organ dysfunction by an average of 9.8 (95% CI 7.3-12.2) days, while molnupiravir was associated with an average reduction of 3.4 (95% CI, 0.5-6.2) days in the length of stay among patients with composite organ dysfunction.

Results of the landmark analysis 3 days after hospitalization are available in Table S5 in Multimedia Appendix 1. Consistent with the main analysis, nirmatrelvir-ritonavir was associated with decreased mortality in respiratory failure (HR 0.43, 95% CI 0.22-0.82) and composite organ dysfunction (HR 0.42, 95% CI 0.22-0.80), as well as lower risk of circulatory shock

XSL•FO
RenderX

(absolute risk difference: 13.7%, 95% CI 3.1-24.4). Molnupiravir was similarly associated with decreased mortality in respiratory failure (HR 0.19, 95% CI 0.08-0.47) and composite organ dysfunction (HR 0.19, 95% CI 0.08-0.48).

Discussion

Principal Findings

In our study of patients diagnosed with culture-negative viral sepsis, antiviral use was associated with reduced mortality among those without any organ dysfunction, with respiratory failure, or with composite organ dysfunction. The strongest reduction in mortality was observed among patients without any organ dysfunction. The use of these antivirals was associated with reduced length of stay across most types of organ dysfunction, except for circulatory shock and acute liver impairment. Importantly, the decreased length of stay was consistent regardless of whether patients received adjuvant therapies including remdesivir, interferon-beta 1b, tocilizumab, baricitinib, or plasma infusions. These findings build on our recent research using an inverse probability weighting approach to evaluate the impact of antivirals in preventing the onset of new organ dysfunction [24].

Our results are in line with previous studies that have investigated the utility of treating COVID-19-associated organ dysfunction. For patients hospitalized with COVID-19, both nirmatrelvir-ritonavir and molnupiravir have been associated with a decreased risk of developing new organ dysfunctions, such as circulatory shock and respiratory failure [24]. Additionally, earlier research has shown that antivirals can reduce the length of hospital stay and accelerate viral clearance times [17,25,26].

The 2021 Surviving Sepsis Campaign COVID-19 guideline did not provide specific recommendations for the newly introduced nirmatrelvir-ritonavir and molnupiravir [14]. They did recommend intravenous remdesivir due to its ability to reduce hospital stay and shorten time to clinical recovery, although it did not reduce mortality benefit across different disease severity strata [27]. In our study, the antivirals nirmatrelvir-ritonavir and molnupiravir showed a mortality benefit for those with mild to moderate disease, with stronger effects in patients without presenting organ dysfunction.

The Infectious Diseases Society of America Guidelines on the Treatment and Management of Patients with COVID-19 advocate for the timely administration of antivirals, including nirmatrelvir-ritonavir and molnupiravir, within 5 to 7 days of symptom onset [15]. This period is deemed most effective due to high viral load and immature immune response characteristic of early infection. These antivirals are specifically recommended for patients with mild to moderate COVID-19 [28,29]. However, many patients are admitted with pre-existing organ dysfunction, indicative of an excessive inflammatory response typical of moderate to severe disease, where the efficacy of antivirals diminishes. Our findings highlight that organ dysfunction constitutes a significant proportion of COVID-19 inpatient deaths. While anti-inflammatory treatments, such as corticosteroids or tocilizumab, are recommended for critically ill COVID-19 patients [14], our data suggest that antivirals still have clinical utility by reducing the length of hospital stay and hastening recovery.

For patients with suspected sepsis but unconfirmed infection, it is recommended to consider alternative diagnoses and the discontinuation of antibiotics [14]. Given the high prevalence of COVID-19, particularly among older adults, it is crucial to consider COVID-19 as a potential source of sepsis when bacterial, parasitic, or fungal causes of infection are not established.

This study has several limitations. First, the study did not account for the varying severity of organ dysfunction due to the heterogeneous clinical presentation of sepsis in undifferentiated COVID-19 patients. It is plausible that antivirals could demonstrate a mortality benefit in patients with mild organ dysfunction. Second, the study assumed that all patients initiated on antivirals completed the full course of treatment. Thirdly, this study excluded patients who did not have blood culture records and those with positive blood culture records; thus, the results may not be generalizable to all COVID-19 inpatients. Last, COVID-19 diagnoses were identified from *ICD-9* codes. Nevertheless, CDARS has previously demonstrated high coding accuracy and data completeness in prior studies [30,31].

Conclusion

Using electronic health data representative of Hong Kong residents, we investigated the effect of nirmatrelvir-ritonavir and molnupiravir on reducing mortality and hospital stay among patients with or without early-onset organ failures. Their association with reduced mortality and hospital stay underscores their potential role in expediting clinical recovery.

Acknowledgments

This study was supported by the Center for Health Systems and Policy Research, which was founded with support from the Tung's foundation, and in part by the AIR@InnoHK of the Innovation and Technology Commission. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author (AKCW) on reasonable request.

XSL•FO
RenderX

Authors' Contributions

AKCW and TTLL conceived the study. TTLL, ACHL, TTJ, and SC performed the study. TTLL, ACHL, and TTJ carried out data analysis. TTLL, ACHL, TTJ, SCLC, CYC, ETFY, LYFL, JWKH, KWLS, OWKT, MLL, SYL, TY, CKT, MSW, ELYW, AKCW, and THR contributed to the revision, and read and approved the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary publication material.
[DOCX File, 187 KB - publichealth_v11i1e72124_app1.docx]

References

1. Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). JAMA 2016 Feb 23;315(8):801-810. [doi: 10.1001/jama.2016.0287] [Medline: 26903338]
2. Wünsch K, Anastasiou OE, Alt M, et al. COVID-19 in elderly, immunocompromised or diabetic patients-from immune monitoring to clinical management in the hospital. Viruses 2022 Apr 1;14(4):746. [doi: 10.3390/v14040746] [Medline: 35458476]
3. Font MD, Thyagarajan B, Khanna AK. Sepsis and septic shock - basics of diagnosis, pathophysiology and clinical decision making. Med Clin North Am 2020 Jul;104(4):573-585. [doi: 10.1016/j.mcna.2020.02.011] [Medline: 32505253]
4. Im Y, Kang D, Ko RE, et al. Time-to-antibiotics and clinical outcomes in patients with sepsis and septic shock: a prospective nationwide multicenter cohort study. Crit Care 2022 Jan 13;26(1):19. [doi: 10.1186/s13054-021-03883-0] [Medline: 35027073]
5. Fiolet T, Kherabi Y, MacDonald CJ, Ghosn J, Peiffer-Smadja N. Comparing COVID-19 vaccines for their characteristics, efficacy and effectiveness against SARS-CoV-2 and variants of concern: a narrative review. Clin Microbiol Infect 2022 Feb;28(2):202-221. [doi: 10.1016/j.cmi.2021.10.005] [Medline: 34715347]
6. Karakike E, Giamarellos-Bourboulis EJ, Kyprianou M, et al. Coronavirus disease 2019 as cause of viral sepsis: a systematic review and meta-analysis. Crit Care Med 2021 Dec 1;49(12):2042-2057. [doi: 10.1097/CCM.0000000000005195] [Medline: 34259663]
7. Shappell CN, Klompas M, Chan C, et al. Use of electronic clinical data to track incidence and mortality for SARS-CoV-2-associated sepsis. JAMA Netw Open 2023 Sep 5;6(9):e2335728. [doi: 10.1001/jamanetworkopen.2023.35728] [Medline: 37773495]
8. Nakayama R, Bunya N, Tagami T, et al. Associated organs and system with COVID-19 death with information of organ support: a multicenter observational study. BMC Infect Dis 2023 Nov 20;23(1):814. [doi: 10.1186/s12879-023-08817-5] [Medline: 37986049]
9. Contreras Nieves M, Anand S, Thomas IC, et al. Safety and effectiveness of nirmatrelvir-ritonavir in patients with advanced kidney dysfunction and COVID-19. Am J Kidney Dis 2025 Aug;86(2):202-211. [doi: 10.1053/j.ajkd.2025.02.603] [Medline: 40204015]
10. Aiello TF, Peyrony O, Chumbita M, et al. Real-life comparison of antivirals for SARS-CoV-2 omicron infection in patients with hematologic malignancies. Influenza Other Respir Viruses 2024;18(3):e13264. [doi: 10.1111/irv.13264] [Medline: 38468434]
11. Zhong X, Wang C, Huang L, et al. Evaluation of the efficacy and safety of nirmatrelvir/ritonavir co-administration inpatients with rheumatic disease infected with SARS-CoV-2: a real-world study. Front Pharmacol 2023;14:1288402. [doi: 10.3389/fphar.2023.1288402] [Medline: 38125894]
12. Gu X, Zhou F, Wang Y, Fan G, Cao B. Respiratory viral sepsis: epidemiology, pathophysiology, diagnosis and treatment. Eur Respir Rev 2020 Sep 30;29(157):200038. [doi: 10.1183/16000617.0038-2020] [Medline: 32699026]
13. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020 Mar 28;395(10229):1054-1062. [doi: 10.1016/S0140-6736(20)30566-3] [Medline: 32171076]
14. Alhazzani W, Evans L, Alshamsi F, et al. Surviving sepsis campaign guidelines on the management of adults with coronavirus disease 2019 (COVID-19) in the ICU: first update. Crit Care Med 2021 Mar 1;49(3):e219-e234. [doi: 10.1097/CCM.0000000000004899] [Medline: 33555780]
15. Bhimraj A, Morgan RL, Shumaker AH, et al. Infectious diseases society of America guidelines on the treatment and management of patients with COVID-19. Clin Infect Dis 2020 Apr 27;78(7):e83-e102. [doi: 10.1093/cid/ciaa478] [Medline: 32338708]
16. Tsoi MF, Chung MH, Cheung BMY, Lau CS, Cheung TT. Epidemiology of gout in Hong Kong: a population-based study from 2006 to 2016. Arthritis Res Ther 2020 Sep 4;22(1):204. [doi: 10.1186/s13075-020-02299-5] [Medline: 32887668]

XSL•FO
RenderX

17. Wai AKC, Chan CY, Cheung AWL, et al. Association of molnupiravir and nirmatrelvir-ritonavir with preventable mortality, hospital admissions and related avoidable healthcare system cost among high-risk patients with mild to moderate COVID-19. Lancet Reg Health West Pac 2023 Jan;30:100602. [doi: 10.1016/j.lanwpc.2022.100602] [Medline: 36212676]
18. Wang H, Wei Y, Hung CT, et al. Association of nirmatrelvir–ritonavir with post-acute sequelae and mortality in patients admitted to hospital with COVID-19: a retrospective cohort study. Lancet Infect Dis 2024 Oct;24(10):1130-1140. [doi: 10.1016/S1473-3099(24)00217-2] [Medline: 38710190]
19. Hospital toolkit for adult sepsis surveillance. Centers for Disease Control and Prevention. 2018. URL: https://stacks.cdc.gov/view/cdc/132387 [accessed 2025-07-25]
20. Yoshida K, Solomon DH, Kim SC. Active-comparator design and new-user design in observational studies. Nat Rev Rheumatol 2015 Jul;11(7):437-441. [doi: 10.1038/nrrheum.2015.30] [Medline: 25800216]
21. Wong CM, Ou CQ, Chan KP, et al. The effects of air pollution on mortality in socially deprived urban areas in Hong Kong, China. Environ Health Perspect 2008 Sep;116(9):1189-1194. [doi: 10.1289/ehp.10850] [Medline: 18795162]
22. Estimation of variations in health care needs between small areas in HK using routine statistics on morbidity, mortality and socioeconomic and demographic characteristics of the population. : Health Services Research Committee, Hong Kong Hospital Authority; 1999.
23. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharm Stat 2011;10(2):150-161. [doi: 10.1002/pst.433] [Medline: 20925139]
24. Wai AKC, Lee TTL, Chan SCL, et al. Association of molnupiravir and nirmatrelvir-ritonavir with reduced mortality and sepsis in hospitalized omicron patients: a territory-wide study. Sci Rep 2023 May 15;13(1):7832. [doi: 10.1038/s41598-023-35068-w] [Medline: 37188726]
25. Wang Y, Zhao D, Chen X, Liu X, Xiao W, Feng L. The effect of nirmatrelvir-ritonavir on viral clearance and length of hospital stay in patients infected with SARS-CoV-2 omicron variants. Influenza Other Respir Viruses 2023 Feb;17(2):e13095. [doi: 10.1111/irv.13095] [Medline: 36843224]
26. Goswami H, Alsumali A, Jiang Y, et al. Cost-effectiveness analysis of molnupiravir versus best supportive care for the treatment of outpatient COVID-19 in adults in the US. Pharmacoeconomics 2022 Jul;40(7):699-714. [doi: 10.1007/s40273-022-01168-0]
27. Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the treatment of covid-19 - final report. N Engl J Med 2020 Nov 5;383(19):1813-1826. [doi: 10.1056/NEJMoa2007764] [Medline: 32445440]
28. Fact sheet for patients, parents, and caregivers: emergency use authorization (EUA) of paxlovid for coronavirus disease 2019 (COVID-19). : Pfzier Labs; 2024 URL: https://www.fda.gov/media/155051/download [accessed 2025-09-11]
29. Fact sheet for patients and caregivers: emergency use authorization (EUA) of LAGEVRIO. : Merck & Co, Inc; 2024 URL: https://www.fda.gov/media/155055/download [accessed 2025-09-11]
30. Gao L, Leung MTY, Li X, et al. Linking cohort-based data with electronic health records: a proof-of-concept methodological study in Hong Kong. BMJ Open 2021 Jun 22;11(6):e045868. [doi: 10.1136/bmjopen-2020-045868] [Medline: 34158297]
31. Chan JSK, Lakhani I, Lee TTL, et al. Cardiovascular outcomes and hospitalizations in asian patients receiving immune checkpoint inhibitors: a population-based study. Curr Probl Cardiol 2023 Jan;48(1):101380. [doi: 10.1016/j.cpcardiol.2022.101380] [Medline: 36031015]

Abbreviations

CDARS : Clinical Data Analysis and Reporting System
HR: hazard ratio
***ICD-9-CM*:** *International Classification of Diseases, Ninth Revision, Clinical Modification*
SDI : Social Deprivation Index
SMD : standardized mean difference

Edited by A Mavragani, T Sanchez; submitted 04.02.25; peer-reviewed by FWT Cheng, KM Jia; revised version received 25.07.25; accepted 04.08.25; published 18.09.25.

Please cite as:
Lee TTL, Lyu ACH, Jiang TT, Chan SCL, Chan CY, Yip ETF, Luk LYF, Ho JWK, So KWL, Tsui OWK, Lam ML, Lee SY, Yamamoto T, Tong CK, Wong MS, Wong ELY, Wai AKC, Rainer TH
Effectiveness of Antivirals Nirmatrelvir-Ritonavir and Molnupiravir in Viral Sepsis: Retrospective Cohort Study
JMIR Public Health Surveill 2025;11:e72124
URL: https://publichealth.jmir.org/2025/1/e72124
doi:10.2196/72124

XSL•FO
RenderX

© Teddy Tai Loy Lee, Alex Chang-Hao Lyu, Ting Ting Jiang, Sunny Ching Long Chan, Crystal Ying Chan, Edmond Tsz Fung Yip, Luke Yik Fung Yip, Joshua Wing Kei Ho, Kevin Wang Leong So, Omar Wai Kiu Tsui, Man Lok Lam, Shi Yeow Lee, Tafu Yamamoto, Chak Kwan Tong, Man Sing Wong, Eliza Lai Yi Wong, Abraham Ka Chung Wai, Timothy Hudson Rainer. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 18.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Shifts in Influenza and Respiratory Syncytial Virus Infection Patterns in Korea After the COVID-19 Pandemic Resulting From Immunity Debt: Retrospective Observational Study

Minah Park[1], PhD; Won Suk Choi[2], MD; Benjamin J Cowling[3,4], PhD

[1]Department of Health Convergence, Ewha Womans University, Seoul, Republic of Korea

[2]Division of Infectious Diseases, Department of Internal Medicine, Korea University College of Medicine, Seoul, Republic of Korea

[3]WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, G/F Patrick Manson Building (North Wing), 7 Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

[4]Laboratory of Data Discovery for Health Limited, Hong Kong Science and Technology Park, New Territories, China (Hong Kong)

Corresponding Author:
Benjamin J Cowling, PhD
WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, G/F Patrick Manson Building (North Wing), 7 Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

Abstract

Background: Nonpharmaceutical interventions (NPIs) such as mask-wearing and social distancing during the COVID-19 pandemic significantly reduced the transmission of common respiratory viruses, including influenza virus and respiratory syncytial virus (RSV). As NPIs were stopped, concerns emerged about “immunity debt,” which suggests that limited natural exposure to pathogens may have increased susceptibility and severity, particularly among young children. However, despite growing attention, the postpandemic impact of NPIs on epidemiologic patterns and shifts in age-specific disease burden remains underexplored.

Objective: This study aims to investigate, using national surveillance data, the repercussions of the COVID-19 pandemic on the epidemiology and clinical burden of influenza virus and RSV infections in Korea, with an emphasis on the influence of NPIs on the incidence and clinical severity of these infections, particularly among young children.

Methods: We analyzed weekly virologic, outpatient, and inpatient surveillance data on influenza virus and RSV infections from the Korea Disease Control and Prevention Agency from 2017 to 2024, covering the prepandemic, pandemic, and postpandemic periods. Time-series analyses were conducted to examine changes in seasonality and to estimate age-specific incidence and clinical severity of influenza virus and RSV infections before and after the COVID-19 pandemic.

Results: In the postpandemic seasons, both RSV and influenza virus infections showed disrupted seasonality with delayed and prolonged epidemics. While the overall burden of both viruses was comparable to that for prepandemic periods, there was a notable shift in the age distribution of severe cases. Among influenza–associated hospital admissions, the proportion of school-aged children (7 - 18 years) doubled, rising from 14% (1,814/12,660) in 2019/20 to 28% (2,176/7,755) in 2022/23. Hospitalization rates in this age group also increased significantly, from 46.8 to 64.4 per 100,000 among children aged 7 - 12 years, and from 16.4 to 30.0 per 100,000 among those aged 13 - 18 years. For RSV infections, the burden shifted most prominently to young children aged 1 - 6 years, whose share of hospital admissions rose from 48% (5,789/11,969) to 61% (7,316/12,011) over the same period. This age group also experienced the largest rise in RSV-associated hospitalization rates, increasing from 230.8 to 357.5 per 100,000 between the 2019/20 and 2022/23 seasons.

Conclusions: The patterns of influenza virus and RSV infections in Korea following the COVID-19 pandemic reveal distinct shifts in timing, severity, and the age groups that were most affected. Postpandemic influenza and RSV activity in Korea showed delayed and prolonged epidemics, with shifts in age-specific disease burden rather than an overall increase. Substantial increases in susceptibility and severity among young children for RSV infections and older children for influenza virus infections suggest lingering immunity gaps from reduced exposures during the pandemic. These effects may be further compounded by declining influenza vaccine uptake among children following the pandemic. Our findings underscore the importance of ongoing surveillance and targeted public health measures to manage respiratory viruses in the postpandemic era.

(JMIR Public Health Surveill 2025;11:e68058) doi:10.2196/68058

KEYWORDS

influenza; RSV; immunity debt; epidemiology; public health

XSL•FO
RenderX

Introduction

Nonpharmaceutical interventions (NPIs), including mask-wearing and social distancing, were widely implemented during the COVID-19 pandemic to mitigate the spread of SARS-CoV-2. While effective in controlling COVID-19, these measures also led to an unprecedented decline in the circulation of other respiratory viruses, including influenza and respiratory syncytial virus (RSV) [1-3].

The concept of "immunity debt," which suggests that reduced exposure to common pathogens during the pandemic may have increased susceptibility and the risk of more severe outcomes upon re-exposure, has gained growing attention as an important factor affecting the postpandemic resurgence and altered epidemiology of respiratory viruses. This theory is particularly relevant to influenza virus and RSV infections [1,4-7], as early-life exposures are critical for developing immune defenses. For example, a recent modeling study demonstrated that prolonged NPIs delayed natural RSV exposure, resulting in a higher burden of infection among young children, most of whom are typically exposed to RSV by the age of 2 [8].

Numerous studies from Europe and North America have reported increased hospitalizations and shifts in age-specific severity for influenza virus and RSV infections following the COVID-19 pandemic [9-12]. A cohort study of 2809 children hospitalized with RSV in the United States found a significant increase in the proportion of children aged 2 to younger than 5 years (34.2% in 2022‐23 vs 23.4% in prepandemic seasons) [9], suggesting increased susceptibility and severity in this age group that may be linked to immunity debt. Another study of patients with RSV infections 5 years or younger in the United States reported an increase in the median age of children with RSV (11.3 mo in 2022‐23 vs 6.8 mo in prepandemic seasons) and children with severe RSV infections requiring advanced respiratory support (6.9 mo in 2022‐23 vs 4.6 mo in prepandemic seasons for high-flow nasal cannula) [10]. A cohort study in Italy also showed that RSV-associated hospitalizations significantly increased in postpandemic seasons, particularly among older age groups, whose risk of severe RSV increased by 4.7 times in 2021‐22 [11]. However, limited data are available from Asian countries that implemented more stringent and prolonged NPIs compared to many Western countries. Recent research [6] has shown that the intensity of NPIs is positively associated with the magnitude of subsequent immunity debt, with countries enforcing stricter measures experiencing larger surges in respiratory infections following the relaxation of NPIs.

As such, this study aims to fill this gap by analyzing shifts in influenza and RSV epidemiology and clinical burden before and after the COVID-19 pandemic using nationwide surveillance data in Korea, where strict NPIs, including prolonged mask mandates and rigorous social distancing policies, remained in place until as late as March 2023, and investigating how immunity debt may have manifested in this unique context.

By directly comparing influenza virus and RSV infections (often jointly discussed in the context of the "tripledemic" alongside COVID-19) across pre- and postpandemic periods, this study aims to provide a more comprehensive understanding of pandemic-driven shifts in respiratory virus patterns. Furthermore, it aims to offer novel insights into the concept of immunity debt, with a particular focus on differences in age-specific impact of influenza virus and RSV infections in the postpandemic period.

Methods

Study Design

We conducted a retrospective observational study to examine shifts in seasonality and age-specific incidence and severity of influenza virus and RSV infections using national surveillance data. Time-series approach was used to characterize patterns of influenza virus and RSV activities and associated hospitalization rates over the study period. Age-specific proportions of hospitalizations were compared between pre- and postpandemic, and influenza virus and RSV infections.

Data Sources

We analyzed weekly virologic, outpatient, and inpatient surveillance data collected from September 3, 2017 (wk 36) to August 31, 2024 (wk 35) through the Influenza and Respiratory Viruses Surveillance System (KINRESS) by the Korea Disease Control and Prevention Agency (KDCA). KINRESS is a nationwide sentinel surveillance system designed to monitor influenza and other respiratory viruses, including RSV. As of February 2025, KINRESS integrates data from 300 sentinel clinics (outpatient surveillance), 18 public health laboratories in each city or province testing respiratory specimens from 106 clinics (virologic surveillance), and 221 hospitals (hospitalization surveillance) across Korea. The KDCA publishes Weekly Sentinel Surveillance Reports on its website, which are available to the public. These reports provide weekly updates on 23 Category IV infectious diseases that are considered less severe but highly prevalent, such as influenza, RSV, and other acute respiratory infections. For this study, we extracted data from these reports, including the proportion of influenza-like illness (ILI) consultations from outpatient surveillance and the proportion of respiratory specimens testing positive for influenza virus (by subtype) and RSV from virologic surveillance coordinated through the Korea Respiratory Virus Integrated Surveillance System (K-RISS), to estimate weekly influenza and RSV activity. K-RISS conducts real-time RT-PCR on respiratory specimens collected from participating clinics to test for 8 common respiratory pathogens, including influenza virus and RSV, and publishes weekly updates on the KDCA's Weekly Sentinel Surveillance Reports. We also used the age-specific number of hospital admissions with confirmed influenza virus or RSV infections from inpatient surveillance, which is publicly available from the KDCA online portal, to estimate age-specific proportions and incidence rates of hospital admissions associated with influenza or RSV.

Statistical Analysis

Each season was defined from week 36 of one year to week 35 of the following year, as per KINRESS. We then categorized these 7 seasons into three distinct periods as follows: prepandemic (2017/18 to 2019/20; 3 seasons), pandemic

XSL•FO
RenderX

(2020/21; 1 season), and postpandemic (2021/22 to 2023/24; 3 seasons). For this study, we consider the 2021/22 season as postpandemic since it followed the government's "Living with COVID-19" plan, a phased approach aimed at progressively lifting NPIs and returning to normal life, which began on November 1, 2021, although some control measures did continue through to mid-2022, and there was a period in early 2022 where some measures were temporarily re-introduced to manage a large surge in Omicron cases [13].

As with our previous study on influenza [14], we multiplied the proportion of patients with ILI by the proportion of respiratory specimens testing positive for each influenza type and subtype or RSV to derive a weekly indicator of influenza or RSV incidence. The proportion of patients with ILI was calculated as the percentage of total outpatient visits that meet the KDCA's ILI case definition (ie, fever ≥38°C with cough or sore throat). The proportion of respiratory specimens testing positive for influenza or RSV was calculated based on the total number of real-time RT-PCR tests done in each week. We conducted a time-series analysis using the proxies to describe trends and seasonal patterns of influenza and RSV during the past 7 years, from 2017/18 to 2023/24. The weekly count of hospitalizations associated with influenza virus and RSV was collated by age group for each season to compare the clinical severity across pre- and postpandemic periods. We then determined influenza- and RSV-associated hospitalization rates per 100,000 persons using age-specific mid-year population estimates for 2017‐2024 retrieved from Statistics Korea [15] to adjust for demographic shifts and enable comparisons across age groups. All analyses were conducted using R version 4.1.1 (R Foundation for Statistical Computing).

Ethical Considerations

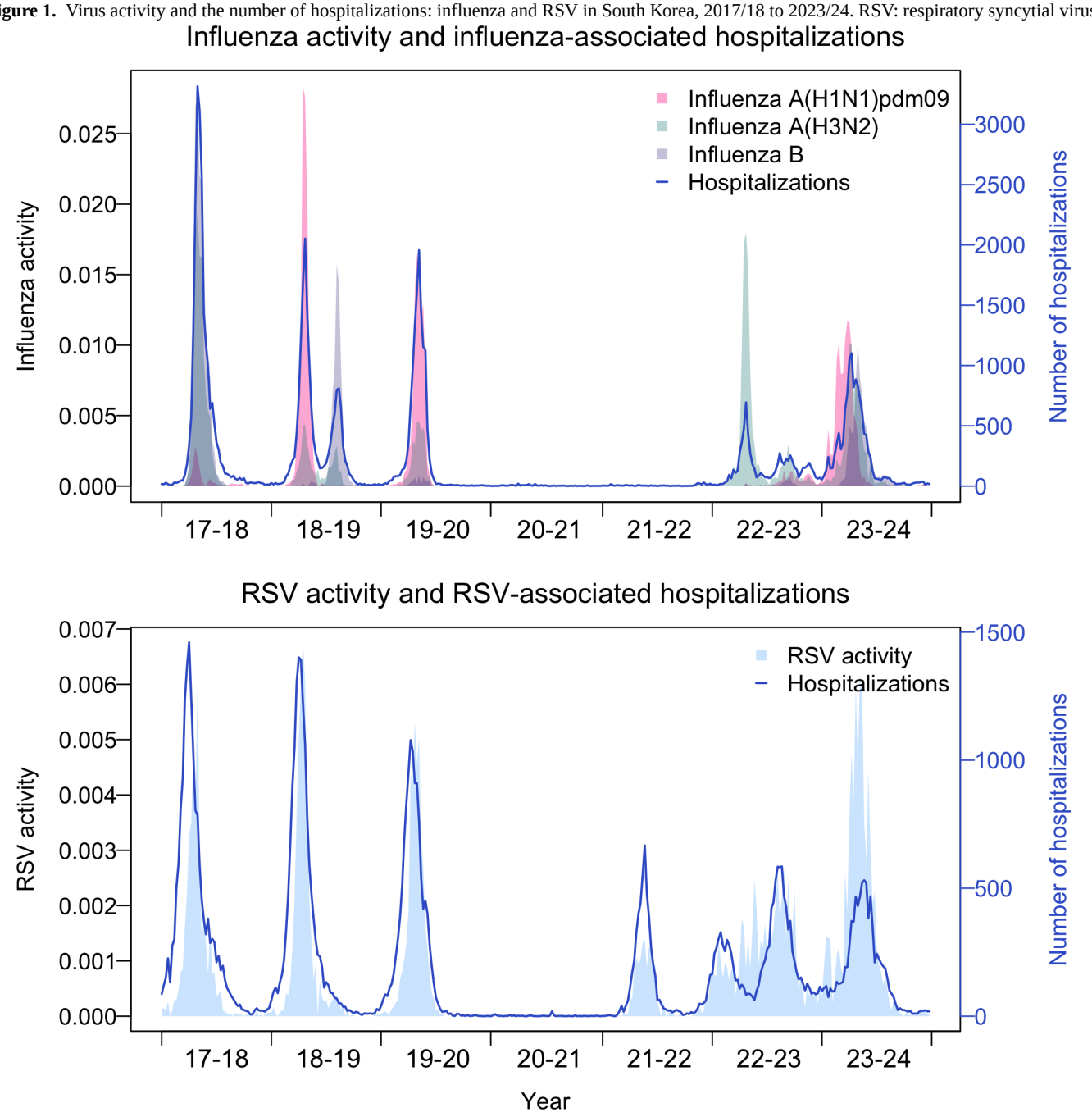
This study used aggregated, routinely collected surveillance data that contained no personally identifiable information and was therefore granted an exemption from ethics review by the Institutional Review Board of Ewha Womans University (ref no. 2025‐0096). Individual informed consent from participants was not required.

Results

In prepandemic seasons, influenza virus and RSV infections followed a similar seasonal pattern, both peaking in winter, with RSV epidemics typically occurring 3‐4 weeks earlier than influenza (wk 49‐50 for RSV and 52‐2 for influenza, respectively), as illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Virus activity and the number of hospitalizations: influenza and RSV in South Korea, 2017/18 to 2023/24. RSV: respiratory syncytial virus.



Following the pandemic, RSV infections reappeared in the 2021/22 season with a delayed epidemic peak in week 5, which occurred 7 - 8 weeks later than usual. In the 2022/23 season, RSV had 2 distinct peaks: an early peak in week 41, 8 - 9 weeks ahead of the typical timing, and a second peak in week 17 amid a prolonged epidemic. Influenza took one more season than RSV to reappear in the 2022/23 season with a typical winter peak in week 1.

The number of influenza- and RSV-associated hospitalizations from 2017/18 to 2023/24 is shown in Multimedia Appendix 1. The overall burden was substantially reduced in postpandemic seasons. Between the 2019/20 and 2020/21 seasons, there was a significant reduction in the number of hospital admissions associated with influenza virus (n=12,660 to n=211; 98.3%) and RSV (n=11,969 to n=80; 99.3%) across all age groups (Multimedia Appendix 1). In prepandemic seasons (2017/18 to 2019/20), influenza virus and RSV infections caused a mean of 17,000 (SD 4483) and 14,583 (SD 2441) hospital admissions, respectively. In postpandemic seasons (2021/22 to 2023/24), RSV infections caused a mean of 8728 (SD 3319) hospital admissions, which is only about 60% of the admissions from prepandemic seasons.

A similar pattern was observed in influenza- and RSV-associated hospitalization rates (Table 1). The overall influenza-associated hospitalization rate significantly declined from 24.7 per 100,000 in the 2019/20 season to nearly zero in the pandemic and early postpandemic years, but bounced back to 23.9 per 100,000 by the 2023/24 season, indicating a near full recovery to prepandemic levels. As for RSV, hospitalization rates dropped significantly across all age groups during the pandemic before rebounding to near pre-pandemic levels in the 2022/23 season. The overall RSV hospitalization rate substantially declined from

XSL•FO
RenderX

23.3 per 100,000 persons to just 0.2, then to 10.5 in the 2020/21 and 2021/22 seasons, respectively. By the 2022/23 season, the rate had recovered to 23.4 per 100,000, closely mirroring prepandemic levels.

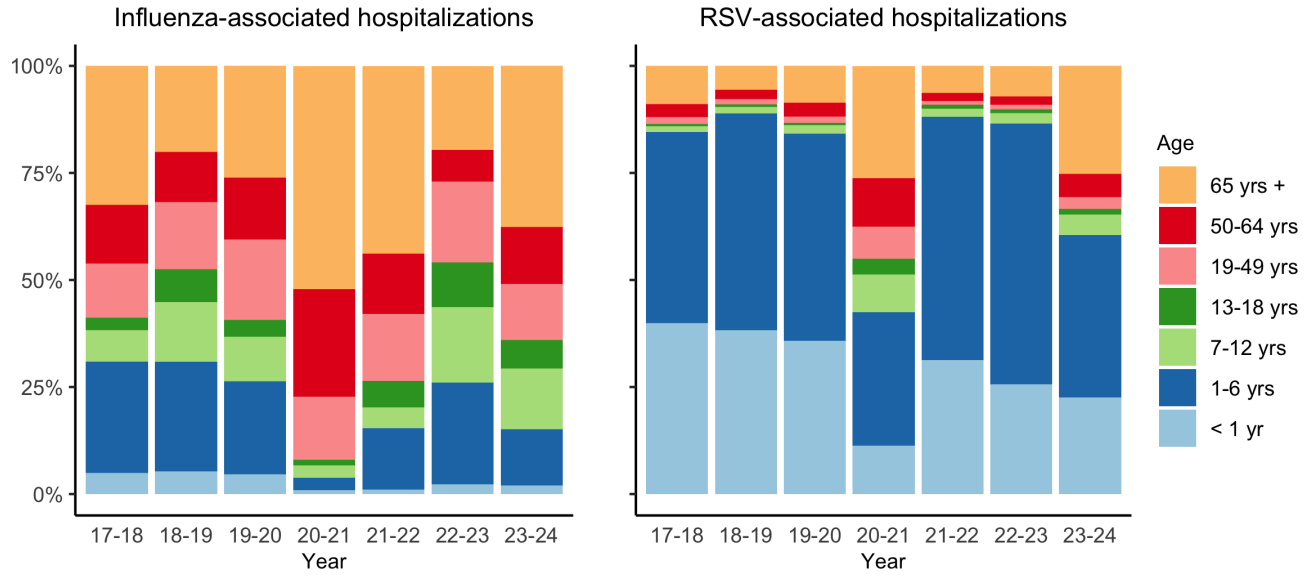
Table . Influenza- and respiratory syncytial virus–associated hospitalization rates (per 100,000 persons).

Season	Peak (wk)	Seasonal hospitalization rates (per 100,000 persons)						
		All age groups	<1 y	1 - 6 y	7 - 12 y	13 - 18 y	19 - 64 y	≥65 y
Influenza								
2017/18	1	42.2	287.5	205.6	57.5	18.6	16.3	99.4
2018/19	52	32.7	267.1	162.4	83.9	40.6	13.2	45.7
2019/20	2	24.7	189	109.5	46.8	16.4	12.0	42.9
2020/21	49	0.4	0.7	0.3	0.2	0.1	0.2	1.4
2021/22	33	0.6	1.2	1.9	0.5	0.6	0.3	1.5
2022/23	1	15.1	70.3	89.9	49.6	28.8	5.9	16.9
2023/24	50	24.3	103.8	86.6	64.4	30	9.6	49.4
RSV[a]								
2017/18	49	32.8	1811.4	275.2	8	2.5	2.2	21.3
2018/19	49	29.2	1725.5	287.6	7.8	2.6	1.5	11.4
2019/20	50	23.3	1396.2	230.8	8.4	2.2	1.6	13.4
2020/21	13	0.2	3.2	1.1	0.2	0.1	0	0.3
2021/22	5	10.5	648.1	138.0	3.5	1.8	0.4	4
2022/23	18	23.4	1236.1	357.5	10.3	3.6	1.1	9.6
2023/24	4	17.2	843.1	177.5	15.2	4.2	2.1	23.5

[a]RSV: respiratory syncytial virus.

For influenza virus infections, more older children aged 7 - 12 years and 13 - 18 years experienced severe outcomes requiring hospitalization postpandemic (Figure 2 and Multimedia Appendix 1). In the 2022/23 season, the proportion of children aged 7 - 12 years was 18% (1,372/7,755), a significant increase from 10% (1,323/12,66) in the 2019/20 season. Similarly, the proportion for children aged 13 - 18 years increased from 4% (491/12,660) in the 2019/20 to 10% (804/7,755) in the 2022/23 season. Similarly, children aged 7 - 12 years and 13 - 18 years experienced the greatest increase in influenza-associated hospitalization rates postpandemic. For children aged 7 - 12 years, the rate rose from 46.8 per 100,000 in the 2019/20 season to 64.4 per 100,000 in the 2023/24 season, while for those aged 13 - 18 years, it increased from 16.4 to 30.0 per 100,000. On the other hand, the hospitalization rates for infants (<1 y) and children aged 1 - 6 years declined (Table 1).

Figure 2. Influenza- and respiratory syncytial virus–associated hospitalizations by age group (%). RSV: respiratory syncytial virus.



XSL•FO
RenderX

For RSV infections, the proportion of young children aged 1 - 6 years among RSV-associated hospitalizations increased post pandemic (Figure 2 and Multimedia Appendix 1). In the 2021/22 season, this age group accounted for 57% (3,057/5,375) of RSV hospitalizations, up from 48% (5,789/11,969) in the 2019/20 season. By the 2022/23 season, this figure rose further to 61% (7,316/12,011). This contrasts with all other age groups, whose proportions either remained the same or decreased in postpandemic seasons. Similarly, age-specific hospitalization rates show that the most significant increase in RSV-associated hospitalizations postpandemic occurred in children aged 1 - 6 years (Table 1). This group saw their hospitalization rate rise from 231 per 100,000 in the 2019/20 season to 358 per 100,000 in the 2022/23 season, marking the greatest increase among all age groups. Meanwhile, infants younger than 1 year of age, who had the highest RSV burden prepandemic, experienced a considerable decline, with the rate dropping from 1400 per 100,000 in the 2019/20 season to 1,240 per 100,000 in 2022/23.

Discussion

Principal Findings

The widespread implementation of NPIs during the COVID-19 pandemic led to substantial reductions in the transmission of influenza virus and RSV in Korea, as it did elsewhere [1-3]. As the NPIs were progressively relaxed in late 2021, RSV returned almost immediately, followed by influenza a year later. We identified substantial changes in RSV and influenza activity, with both viruses having delayed and prolonged epidemics in postpandemic seasons. However, this prolonged activity was not linked to an increased burden. In fact, our analysis revealed a lower overall epidemiological burden compared to previous years. This was contrary to the feared “triple-demic” [16]. The lessened impact could be due to a step change in awareness of infectious disease prevention, such as more frequent hand-washing and mask-wearing after the pandemic.

An important observation in our study is the significantly increased proportion of young children aged 1 - 6 years among RSV-associated hospital admissions in the postpandemic years. This may indicate that the susceptibility to severe infections, especially in young children, may be higher than in prepandemic seasons. This could be partly due to the accumulated “immunity debt,” where more young children are susceptible to severe outcomes because they have not developed immunity through prior exposures during the pandemic [4,5]. In particular, our findings suggest that immunity debt was more significant for RSV infections, particularly in children aged 1 - 6 years. They were the only age group that experienced an increased severity, with an increased proportion among all RSV-related hospital admissions in the first 2 postpandemic seasons. This is consistent with a recent US study [17], which found a significantly higher RSV burden in children aged 2 - 5 years, having almost 4.86 times as likely to be hospitalized in 2022 compared to prepandemic years, whereas younger children, such as infants aged 0 - 5 months, were only 1.77 times as likely. Primary RSV infections tend to be more severe than re-infections [18], and by delaying primary infections for one or two years, clinical severity may have become elevated in slightly older children who had avoided RSV infection during the COVID-19 pandemic.

Our analysis revealed that the proportion of older children aged 7 - 18 years increased among influenza-associated hospital admissions in the postpandemic years. It may be partly due to the delayed exposure to influenza virus infections during the pandemic, which may have resulted in reduced immunity in these school-age children. School closures, remote learning, and reduced social interactions during the pandemic likely disrupted the usual circulation of influenza, leading to fewer infections in younger populations. As restrictions eased and normal social interactions resumed, these children who had lower pre-existing immunity became more susceptible to the virus, contributing to higher hospitalization rates in this age group.

In addition, changes in vaccination coverage or the effectiveness of influenza vaccination in these years may also have played a role in the observed increase in hospitalizations. In Korea, all children aged 6 months to 13 years are currently eligible to receive free annual influenza vaccination under the National Immunization Program. While vaccination uptake among children remains relatively high compared to other countries, it has been declining in recent years, dropping from 78.5% in the 2019/20 season (prepandemic) to 72.7% in 2020/21 (pandemic), and further to 69.5% in 2023/24 (postpandemic). This slight decline in influenza vaccination coverage, combined with increased susceptibility due to immunity debt, may have contributed to the rise in influenza-associated hospitalization rates, particularly among school-aged children, who consistently have the lowest vaccination rates (82.5% for children aged 6 to 59 months of age vs 49.2% for children aged 13 years). The new RSV monoclonal antibody nirsevimab is not yet available in Korea for the prevention of RSV infections in infants.

Furthermore, while both influenza virus and RSV have returned as NPIs eased, the timing and clinical burden of these rebounds varied. RSV returning a year earlier than influenza might suggest that NPIs had less effect on RSV. RSV is thought to spread more through direct contact with contaminated surfaces and is more stable on these surfaces than influenza virus, allowing it to survive longer outside the human body [19]. This stability, combined with the virus’s prolonged shedding in infected individuals, especially children, could make RSV transmission more resilient to NPIs focused on airborne spread, such as masking and social distancing. In contrast, influenza’s shorter infectious period and reliance on airborne transmission make it more susceptible to these interventions, leading to a more effective reduction in its spread than RSV.

Our study contributes to the growing body of evidence on the postpandemic epidemiology of respiratory viruses and highlights the need for age-specific surveillance and vaccination approaches in the evolving landscape of respiratory infectious diseases. While similar trends of early and intensified RSV and influenza activity have been reported in other regions, the Korean experience offers important comparative insights due to the country’s prolonged and comprehensive implementation of NPIs. The distinct timing and magnitude of respiratory virus resurgence observed in Korea highlight how variations in public

XSL•FO
RenderX

health interventions and population behaviors can differentially influence viral epidemiology. By providing data from an East Asian setting, our study expands the geographic scope of evidence on immunity debt and informs future policy planning for vaccination strategies, hospital resource allocation, and public health preparedness at the global level.

Limitations

Our study has several limitations. First, our analyses are based on weekly virologic, outpatient, and weekly surveillance data. We did not have additional information, such as serologic data, which would allow a more comprehensive picture of the incidence and severity of infections in the community, as well as estimate population immunity. Second, changes in the burden of severe influenza virus or RSV infections in postpandemic seasons could be partially associated with increases in testing after the pandemic linked to increased laboratory capacity and attention to viral infections [20]. However, it is unlikely that increased testing fully explains our observations, particularly for RSV. In Korea, RSV diagnostic testing is generally less actively conducted than for influenza or COVID-19 due to the absence of a standard antiviral treatment, although testing is more common in infants who are at higher risk of severe illness. While increased testing postpandemic could contribute to higher reported cases, our focus on hospitalization rates rather than raw incidence mitigates this effect, as severe outcomes are less influenced by testing frequency. Furthermore, the diagnostic methods for RSV, such as rapid antigen tests, PCR, and viral culture, have not significantly changed before or after the COVID-19 pandemic. Thus, despite the pandemic's impact on health care, the frequency and methods of RSV testing in Korea likely remained consistent, suggesting that this factor is unlikely to fully explain the observed severity shifts. This is a challenge shared by similar studies adapting to heightened postpandemic surveillance. Third, changes in health-seeking behaviors after the pandemic must also be accounted for when interpreting results from the study. The COVID-19 pandemic significantly altered how individuals accessed health care services, often delaying or avoiding visits to hospitals or clinics due to fear of exposure, public health restrictions, or overwhelmed health care systems. This shift may have resulted in the underreporting of cases in the surveillance data, affecting the comparability of pre- and postpandemic disease burden. Although altered health-seeking behaviors during the pandemic may have led to underreporting, our reliance on inpatient surveillance data, which reflects severe cases necessitating care, likely minimizes this bias. As restrictions eased post-2021, health care access likely stabilized, suggesting this limitation is more relevant to the pandemic peak than the postpandemic trends we emphasize. Future studies could refine these findings by integrating health care use data. Finally, increased awareness of respiratory infections may have heightened case detection postpandemic, but our emphasis on hospitalization trends, which should be less influenced by awareness than outpatient visits, anchors our conclusions regarding clinical severity. This awareness, potentially reducing transmission through sustained preventive behaviors, may even offset inflated detection, aligning with the lower overall burden we observed. This reflects a broader postpandemic shift in public health consciousness. Our study's limitations, including potential biases from increased testing, shifting health-seeking behaviors, and heightened awareness, are tempered by our focus on hospitalization data and the consistency of diagnostic practices in Korea. These factors, common to postpandemic research, introduce uncertainty but do not negate the distinct shifts in influenza and RSV patterns we identified. Future analyses adjusting for these variables could further enhance our understanding of the role of immunity debt.

Conclusions

The patterns of influenza virus and RSV infections in Korea following the COVID-19 pandemic reveal distinct shifts in timing, severity, and affected age groups. While the resurgence of RSV and influenza was anticipated with the relaxation of NPIs, the delayed and prolonged epidemic patterns observed in the postpandemic seasons did not correspond with an increased overall disease burden. Instead, we found specific increases in susceptibility among younger children for RSV infections and older children for influenza virus infections, potentially due to immunity gaps arising from limited exposure during the pandemic. These findings highlight the lasting impacts of pandemic-related disruptions on population immunity and respiratory disease transmission patterns. Ongoing surveillance and targeted public health measures remain critical to understanding and managing seasonal respiratory viruses in a postpandemic context.

Acknowledgments

We are grateful to the Korea Disease Control and Prevention Agency (KDCA) for providing the weekly respiratory virus surveillance data.

Authors' Contributions

MP, WSC, and BJC designed the study. MP collected and conducted all statistical analyses. MP and BJC wrote the first draft. WSC reviewed and commented on the manuscript.

Conflicts of Interest

WSC has consulted for Moderna, Sanofi Pasteur, and CSL, and has been an investigator in clinical trials sponsored by SK Bioscience, GlaxoSmithKline, Pfizer, Moderna, Sanofi Pasteur, and Janssen. BJC has consulted for AstraZeneca, Fosun Pharma, GlaxoSmithKline, Haleon, Moderna, Novavax, Pfizer, Roche, and Sanofi Pasteur.

XSL•FO
RenderX

Multimedia Appendix 1
Number of influenza- and respiratory syncytial virus–associated hospitalizations (n, %).
[PDF File, 112 KB - publichealth_v11i1e68058_app1.pdf]

References

1. Bardsley M, Morbey RA, Hughes HE, et al. Epidemiology of respiratory syncytial virus in children younger than 5 years in England during the COVID-19 pandemic, measured by laboratory, clinical, and syndromic surveillance: a retrospective observational study. Lancet Infect Dis 2023 Jan;23(1):56-66. [doi: 10.1016/S1473-3099(22)00525-4] [Medline: 36063828]
2. Sullivan SG, Carlson S, Cheng AC, et al. Where has all the influenza gone? The impact of COVID-19 on the circulation of influenza and other respiratory viruses, Australia, March to September 2020. Euro Surveill 2020 Nov;25(47):33243355. [doi: 10.2807/1560-7917.ES.2020.25.47.2001847] [Medline: 33243355]
3. Hamid S, Winn A, Parikh R, et al. Seasonality of respiratory syncytial virus - United States, 2017-2023. MMWR Morb Mortal Wkly Rep 2023 Apr 7;72(14):355-361. [doi: 10.15585/mmwr.mm7214a1] [Medline: 37022977]
4. Hatter L, Eathorne A, Hills T, Bruce P, Beasley R. Respiratory syncytial virus: paying the immunity debt with interest. Lancet Child Adolesc Health 2021 Dec;5(12):e44-e45. [doi: 10.1016/S2352-4642(21)00333-3] [Medline: 34695374]
5. Billard MN, Bont LJ. Quantifying the RSV immunity debt following COVID-19: a public health matter. Lancet Infect Dis 2023 Jan;23(1):3-5. [doi: 10.1016/S1473-3099(22)00544-8] [Medline: 36063827]
6. Chen L, Guo Y, López-Güell K, et al. Immunity debt for seasonal influenza after the COVID-19 pandemic and as a result of nonpharmaceutical interventions: an ecological analysis and cohort study. Adv Sci (Weinh) 2025 May;12(20):e2410513. [doi: 10.1002/advs.202410513] [Medline: 40186616]
7. Munro AP, House T. Cycles of susceptibility: Immunity debt explains altered infectious disease dynamics post -pandemic. Clin Infect Dis 2024 Oct 11:ciae493. [doi: 10.1093/cid/ciae493] [Medline: 39390969]
8. Kutsaya A, Teros-Jaakkola T, Kakkola L, et al. Prospective clinical and serological follow-up in early childhood reveals a high rate of subclinical RSV infection and a relatively high reinfection rate within the first 3 years of life. Epidemiol Infect 2016 Jun;144(8):1622-1633. [doi: 10.1017/S0950268815003143] [Medline: 26732801]
9. Rao S, Armistead I, Messacar K, et al. Shifting epidemiology and severity of respiratory syncytial virus in children during the COVID-19 pandemic. JAMA Pediatr 2023 Jul 1;177(7):730-732. [doi: 10.1001/jamapediatrics.2023.1088] [Medline: 37184852]
10. Winthrop ZA, Perez JM, Staffa SJ, McManus ML, Duvall MG. Pediatric respiratory syncytial virus hospitalizations and respiratory support after the COVID-19 pandemic. JAMA Netw Open 2024 Jun 3;7(6):e2416852. [doi: 10.1001/jamanetworkopen.2024.16852] [Medline: 38869896]
11. Lastrucci V, Pacifici M, Puglia M, et al. Seasonality and severity of respiratory syncytial virus during the COVID-19 pandemic: a dynamic cohort study. Int J Infect Dis 2024 Nov;148(107231):107231. [doi: 10.1016/j.ijid.2024.107231] [Medline: 39222708]
12. Hoy G, Maier HE, Kuan G, et al. Increased influenza severity in children in the wake of SARS-CoV-2. Influenza Other Respir Viruses 2023 Jul;17(7):e13178. [doi: 10.1111/irv.13178] [Medline: 37492240]
13. Announcement of three-phase plan for returning to normal: press conference [Article in Korean]. Korea Ministry of Health and Welfare. 2021. URL: https://www.mohw.go.kr/board.es?mid=a10503010100&bid=0027&act=view&list_no=368300&tag=&nPage=186 [accessed 2024-09-01]
14. Park M, Wu P, Goldstein E, Joo Kim W, Cowling BJ. Influenza-associated excess mortality in South Korea. Am J Prev Med 2016 Apr;50(4):e111-e119. [doi: 10.1016/j.amepre.2015.09.028] [Medline: 26610897]
15. Resident registration population by city/county/gender/age (1 year) population trend survey, Statistics Korea. KOrean Statistical Information Service (KOSIS). 2024. URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B040M1&conn_path=I3 [accessed 2025-02-25]
16. Burki TK. Circulation of influenza, RSV, and SARS-CoV-2: an uncertain season ahead. Lancet Respir Med 2021 Oct;9(10):e103. [doi: 10.1016/S2213-2600(21)00364-7] [Medline: 34370976]
17. Suss RJ, Simões EAF. Respiratory syncytial virus hospital-based burden of disease in children younger than 5 years, 2015-2022. JAMA Netw Open 2024 Apr 1;7(4):e247125. [doi: 10.1001/jamanetworkopen.2024.7125] [Medline: 38635270]
18. Borchers AT, Chang C, Gershwin ME, Gershwin LJ. Respiratory syncytial virus--a comprehensive review. Clin Rev Allergy Immunol 2013 Dec;45(3):331-379. [doi: 10.1007/s12016-013-8368-9] [Medline: 23575961]
19. Tang JW, Loh TP. Correlations between climate factors and incidence--a contributor to RSV seasonality. Rev Med Virol 2014 Jan;24(1):15-34. [doi: 10.1002/rmv.1771] [Medline: 24421259]
20. Petros BA, Milliren CE, Sabeti PC, Ozonoff A. Increased pediatric respiratory syncytial virus case counts following the emergence of SARS-CoV-2 can be attributed to changes in testing. Clin Infect Dis 2024 Jun 14;78(6):1707-1717. [doi: 10.1093/cid/ciae140] [Medline: 38602423]

Abbreviations

ILI: influenza-like illness

XSL•FO
RenderX

K-RISS: Korea Respiratory Virus Integrated Surveillance System
KDCA: Korea Disease Control and Prevention Agency
NPI: nonpharmaceutical intervention
RSV: respiratory syncytial virus

Edited by A Mavragani, T Sanchez; submitted 30.10.24; peer-reviewed by AA Sayiner, TM Magalhaes Moreira; revised version received 04.06.25; accepted 05.06.25; published 23.07.25.

Please cite as:
Park M, Choi WS, Cowling BJ
Shifts in Influenza and Respiratory Syncytial Virus Infection Patterns in Korea After the COVID-19 Pandemic Resulting From Immunity Debt: Retrospective Observational Study
JMIR Public Health Surveill 2025;11:e68058
URL: https://publichealth.jmir.org/2025/1/e68058
doi:10.2196/68058

© Minah Park, Won Suk Choi, Benjamin J Cowling. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 23.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Greater Improvements in Vaccination Outcomes Among Black Young Adults With Vaccine-Resistant Attitudes in the United States South Following a Digital Health Intervention: Latent Profile Analysis of a Randomized Control Trial

Noah Mancuso[1,2], MSPH; Jenna Michaels[1], MS; Erica N Browne[1], MS; Allysha C Maragh-Bass[3], MPH, PhD; Jacob B Stocks[4], MSc; Zachary R Soberano[4], BA; C Lily Bond[4], MPH; Ibrahim Yigit[4], MA, PhD; Maria Leonora G Comello[5], MS, PhD; Margo Adams Larsen[6], PhD; Kathryn E Muessig[4], PhD; Audrey Pettifor[5], MPH, PhD; Lisa B Hightow-Weidman[4], MPH, MD; Henna Budhwani[4], MPH, PhD; Marie C D Stoner[1], MPH, PhD

[1]Women's Global Health Imperative, Research Triangle Institute (RTI) International, 3040 East Cornwallis Road, Research Triangle Park, NC, United States

[2]Department of Epidemiology, Emory University, Atlanta, GA, United States

[3]FHI 360, Durham, NC, United States

[4]Institute on Digital Health and Innovation, College of Nursing, Florida State University (FSU), Tallahassee, FL, United States

[5]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[6]Virtually Better, Inc, Decatur, GA, United States

Corresponding Author:
Noah Mancuso, MSPH
Women's Global Health Imperative, Research Triangle Institute (RTI) International, 3040 East Cornwallis Road, Research Triangle Park, NC, United States

Abstract

Background: Negative attitudes toward vaccines and suboptimal vaccination rates among African American and Black (Black) Americans have been well documented, due to a history of medical racism and human rights violations in the United States. However, digital health interventions (DHI) have been shown to address racial disparities in several health outcomes, such as cardiovascular disease, HIV, and maternal health. The Tough Talks COVID (TT-C) study was a randomized controlled trial of a DHI designed to empower Black young adults in the United States South to make informed, autonomous decisions about COVID-19 vaccine uptake by addressing structural barriers and misinformation about vaccines.

Objective: Our objective was to identify subgroups of Black young adults with various vaccine attitudes at baseline and determine the subgroups for which the TT-C DHI was most impactful.

Methods: Black young adults aged 18‐29 years in Alabama, Georgia, and North Carolina who were unvaccinated or insufficiently vaccinated against COVID-19 completed three online surveys over three months (N=360). Latent profile analysis was used to identify subgroups based on general vaccine attitudes at baseline, including hesitancy, confidence, knowledge, conspiracy beliefs, and mistrust. Logistic regression was used to examine the associations between latent profiles and vaccine uptake, and linear regression was used to examine changes in vaccine attitudes at three months post-randomization. Modification of the TT-C DHI's effects was assessed by latent profiles.

Results: Three latent profiles emerged: vaccine-receptive (n=124), vaccine-neutral (n=155), and vaccine-resistant (n=81). Political affiliation, income, social support, and recent flu vaccination differed significantly between the three subgroups (P<.05). Vaccine uptake was not significantly different by subgroup, and the TTC-DHI did not have differing effects on uptake across subgroups. However, the DHI had the strongest effect—with statistically significant measures of association (P<.05) and interaction P values (P<.10)—among the baseline vaccine-resistant and vaccine-neutral subgroups compared to the vaccine-receptive subgroups at three months in improving vaccine hesitancy, confidence, and conspiracy beliefs at three months: vaccine-resistant difference: −0.40 (−0.76 to −0.37), 0.39 (0.02 to 0.75), and −0.47 (−0.86 to −0.09); vaccine neutral difference: −0.36 (−0.52 to −0.19), 0.35 (0.18 to 0.51), and −0.24 (−0.44 to −0.03). The DHI had no effects on these outcomes among the vaccine-receptive subgroup.

Conclusions: Our findings revealed subgroups of Black young adults in the United States South with different vaccination attitudes, for which the TT-C intervention had differing effects. Black young adults who are vaccine-resistant or vaccine-neutral

XSL•FO
RenderX

may experience larger gains from a digital vaccine intervention. Future work aimed at improving vaccination outcomes could target these populations to maximize resource efficiency and drive the greatest improvements in vaccine outcomes.

Trial Registration: ClinicalTrials.gov NCT05490329; https://clinicaltrials.gov/study/NCT05490329

(JMIR Public Health Surveill 2025;11:e67370) doi:10.2196/67370

KEYWORDS

COVID-19; mHealth; African American and Black; young adults; vaccination; mobile health

Introduction

The World Health Organization declared the outbreak of novel COVID-19, a global pandemic in early 2020, caused by SARS-CoV-2 [1]. By the end of the year, two COVID-19 vaccines had been issued emergency use authorizations in the United States [2,3]. At the time, nearly three in five Americans reported intentions to be vaccinated for COVID-19; however, African American or Black (Black) Americans reported intentions to vaccinate approximately 30% lower than that of White Americans and 50% lower than that of Asian Americans [4]. Hesitancy toward COVID-19 vaccination was particularly high among Black young adults and residents of the United States South [5,6], which translated into lower uptake of the COVID-19 vaccine and a disproportionately high mortality rate for Black Americans early in the pandemic [7,8].

Negative attitudes toward vaccines and suboptimal vaccination rates among Black Americans have been influenced by a history of medical racism and human rights violations in the United States [9-14]. Continued harmful experiences with the health care system and government motivate hesitancy toward public health programs, such as vaccination, as a self-protective coping mechanism [15]. Studies have consistently shown this hesitancy, with reported lower vaccine uptake among Black communities for the annual flu shot due to distrust of medical institutions and concerns for safety [16-18]. Additionally, disparities in human papillomavirus vaccination have been observed among young Black girls, which is reflected in higher rates of cervical cancer incidence and mortality later in life [19,20]. During the COVID-19 pandemic, these experiences were exacerbated for Black young adults by the spread of misinformation and conspiracy theories online and on social media [21,22].

Digital health interventions (DHIs) have been successful in addressing racial disparities in health by creating culturally tailored messaging, inclusive programming, and more accessible services for disease prevention, ranging from cardiovascular disease to HIV to maternal health [23-27]. DHIs are particularly helpful for young adults, given the high proportion of Americans aged 18‐28 years with smartphones who report using web-based news and social media sites to access information about COVID-19 [28,29]. However, few DHIs to date have been developed specifically for COVID-19 vaccination among Black young adults in the United States South.

The Tough Talks COVID (TT-C) study was a randomized controlled trial of a DHI designed to empower Black young adults in the United States South to make informed, autonomous decisions about COVID-19 vaccine uptake (ClinicalTrials.gov NCT05490329) [30]. The TT-C intervention was cocreated using community-based participatory research methods to address structural barriers and misinformation about vaccines through interactive activities and digital storytelling [31,32]. Assuming heterogeneity among the study population, we aimed to use a latent profile analysis (LPA) to identify subgroups of Black young adults with different baseline vaccine attitudes and identify subgroups for whom TT-C was most impactful, to allow for better-informed community-based policies, practices, and messaging related to vaccine uptake and attitudes.

Methods

Participants

We used data from the TT-C randomized control trial, which included Black young adults aged 18‐29 years who were unvaccinated or insufficiently vaccinated against COVID-19 in Alabama, Georgia, and North Carolina. Participants were recruited online or via social media and randomized 1:1 to either the TT-C DHI, a self-directed asynchronous app developed to enable Black young adults in Southern United States to make decisions about COVID-19 vaccine receipt via nonstigmatizing and tailored messaging, or an information standard-of-care control consisting of COVID-19 vaccine materials adapted from the Centers for Disease Control and Prevention for three months. Surveys were completed at baseline, one month, and three months after randomization. Additional details about the study, which was conducted between March and December 2023, have been published previously [30] .

Latent Indicators

Five baseline variables related to general vaccine attitudes at baseline were used as latent indicators in the LPA: vaccine hesitancy, vaccine confidence, vaccine-related conspiracy beliefs, vaccine knowledge, and group-based medical mistrust. Each variable was assessed using a previously validated scale with a standardized continuous score as input using z-scores, such that the mean score transformed to zero to remove the impact of any baseline differences in means or variances across the variables before they were used in the LPA. This approach helps mitigate the influence of baseline distribution differences in the clustering process by making the variables comparable. The measures used in this LPA were prespecified in the study protocol and tested in formative surveys with Black young adults [30,33]. Vaccine hesitancy was a mean score of a 9-item scale with Likert responses ranging from 1‐5, where higher scores indicated greater hesitancy [34]. Vaccine confidence was a mean score of a 7-item scale with Likert responses ranging from 1‐5, where higher scores indicated greater confidence [35]. Vaccine-related conspiracy beliefs were a mean score of an adapted 5-item scale with Likert responses ranging from 1‐4,

XSL•FO
RenderX

where higher scores indicated greater belief in conspiracy [36]. Vaccine knowledge was a sum score of an adapted 4-item scale with "yes" or "no" responses for each item, where higher scores indicated greater knowledge [37]. Group-based medical mistrust was a sum score of a 12-item scale with Likert responses ranging from 1 - 5, where higher scores indicated greater mistrust [38]. For mean scores, if a response was missing for more than one scale item, the mean score was set as missing and not calculated. For summed scores, if a response was missing for any scale item, the summed score was set as missing and not calculated.

Additional Measures

Once latent profiles were established, differences in several baseline sociodemographic variables and other measures were assessed between the profiles. These included self-reported: sex assigned at birth, gender, age, education (ie, 9th-12th grade; no diploma; high school graduate or GED completed; some college level or technical or vocational degree; bachelor's degree; other advanced degree), political affiliation (ie, Democrat, Republican, Independent, other), employment status (eg, full-time; part-time; student; not working; disabled, permanently or temporarily), annual income, insurance coverage (eg, public, private, none), social support (continuous scale) [39], and previous flu and human papillomavirus vaccination.

Outcomes

The primary outcome was uptake of a COVID-19 vaccine anytime during the 3-month follow-up period, based on uploaded vaccine cards or, if not uploaded, self-report at one month and three months. Secondary outcomes included vaccine hesitancy, vaccine confidence, vaccine-related conspiracy beliefs, and vaccine knowledge, which were reported at 3 months. These secondary measures were assessed in the same manner as the baseline latent indicators explained above. Latent profiles were used to assess effect measure modification (EMM) of the TT-C DHI on both primary and secondary outcomes.

Statistical Analysis

Latent profile analysis allows us to gain insights into "hidden" subgroups in a study sample using person-centered mixture modeling. This approach models how indicator variables relate to each other by estimating different profiles (or classes) and determining the best-fitting model [40,41]. In LPA, model parameters include the probability of an individual belonging to a given profile, and each profile can have its own set of means and variances.

We evaluated two different model specifications: (1) equal variance and covariance across profiles and (2) varied variances and covariances across profiles. For each specification, we tested models with two, three, and four profiles. Model fit was assessed using an analytical hierarchy process based on the following information criteria: Akaike information criterion, Bayesian information criterion, Approximate Weight of Evidence, consistent likelihood criterion, and Kullback information criterion [42]. Lower values indicate a better fit for each criterion. Additional criteria including the sample size-adjusted Bayesian information criterion, entropy, and the bootstrap likelihood ratio (BLRT) were used to assess overall model quality. For these criteria, a lower sample size-adjusted Bayesian information criterion, higher entropy, and statistically significant BLRT indicate a good fit. The best model was selected using an analytical hierarchical process that balanced low values of Akaike information criterion, Bayesian information Criterion, Approximate Weight of Evidence, consistent likelihood criterion, and Kullback information criterion, which indicate a more parsimonious model, with fewer parameters while still accounting for data complexity, with higher entropy to ensure distinct profile separation and a significant BLRT to confirm robustness [42]. After model selection, we assigned study participants to the profile with the highest posterior probability based on their pattern of responses across the variables. Descriptive statistics were then calculated for the sociodemographic and additional variables within each profile. Differences in continuous variables across profiles were tested using ANOVA, while differences in categorical variables were assessed with the χ^2 test of independence.

The primary outcome of vaccine uptake at three months was assessed using a logistic regression model. The odds ratio approximates the risk ratio, as the probability of vaccine uptake was not common (<15%). For secondary outcomes—vaccine hesitancy, vaccine confidence, and vaccine-related conspiracy beliefs—we used linear regression to determine differences in outcomes at three months. Poisson regression was used to test for differences in the secondary outcome of vaccine knowledge score at three months. The main effects of the TT-C DHI intervention on vaccine uptake and vaccine attitudes will be reported in the primary manuscript of the study (Budhwani et al, PhD, unpublished data, 2024). The EMM of the effect of the TT-C DHI intervention on primary and secondary outcomes by profile group was assessed in this study by including the study arm and latent profile at baseline as both main effects and interaction terms. The EMM assesses whether an intervention effect (ie, the effect of the intervention on the outcome in the intervention group compared to the control group) differs across a variable of interest (in this case, the LPA groups). Interaction terms with a P value <.10 were considered indicators of significant effect measure modification. For all models, the state of residence was controlled to account for its inclusion in the stratified randomization study design. We did not control for other variables as they were balanced by the randomized design. All analyses were conducted using R software (version 4.4.1; R Foundation for Statistical Computing) [43].

Sensitivity Analysis

A sensitivity analysis was conducted using latent class analysis (LCA), typically used with categorical variables, where all indicator variables were dichotomized at appropriate cut-points for each validated scale [44]. We tested LCA models with two, three, and four classes and used similar criteria to assess fit. We then compared the number of classes identified by LCA to the number of profiles identified by LPA, as well as the total variance explained and general sociodemographic differences across the classes/profiles. We chose to report LPA as the main analysis to maintain the continuous nature of the originally validated measures used as indicator variables and because of the added model flexibility of LPA with varying variances and covariances across profiles.

XSL•FO
RenderX

Ethical Considerations

All study materials and procedures were reviewed and approved by the Institutional Review Boards of the University of North Carolina, Chapel Hill (21 - 1746) and Florida State University (STUDY00003617). Informed consent was collected digitally from all study participants before data collection and randomization. All data were deidentified to safeguard participant information.

Results

A total of 360 Black young adults were included in this study, of whom 76% (n=272) identified as cisgender or transgender women. The median age was 24 (IQR 21 - 27) years, and approximately one-third of participants (n=122) had a bachelor's degree or higher. Overall vaccine uptake at month 3 was low (n=21) and did not significantly differ between the intervention and control arms (odds ratio [OR] 1.88, 95% CI 0.76 to 4.69).

Latent Profile Analysis

Using the prespecified fit criteria with the analytic hierarchy process, we found that the model with three profiles and varied variances and covariances across profiles provided the best fit (Table 1). The profiles in Figure 1 were named based on the mean values of the five indicator variables, all of which were significantly different (P<.001) between the profiles (Table 2). The vaccine-receptive group (Profile 2) included 124 participants (34%) and had the lowest mean scores for medical mistrust, conspiracy beliefs, and vaccine hesitancy and the highest mean score for vaccine confidence and knowledge of vaccines compared to other profiles. The vaccine-resistant group (Profile 3) included 81 participants (23%) and had the highest mean score for medical mistrust, conspiracy beliefs, and vaccine hesitancy and the lowest mean score for confidence and knowledge of vaccines compared to other profiles. The vaccine-neutral group (Profile 1) included 155 participants (43%) and had mean scores between those of the vaccine-receptive and vaccine-resistant groups for all indicator variables.

Table . Fit indices for the latent profile analysis of Model 1 (equal variance and covariance across profiles) and Model 2 (varied variances and covariances across profiles) for 2 - 4 profiles of baseline vaccine attitudes in the TT-C[a] study

Model	Number of profiles	AIC[b]	BIC[c]	AWE[d]	CLC[e]	KIC[f]	SABIC[g]	Entropy	BLRT[h] *P* value
1	2	4439.42	4501.59	4642.03	4409.15	4458.42	4450.83	0.87	<.001
1	3	4245.05	4330.55	4524.38	4202.71	4270.05	4260.75	0.83	<.001
1	4	4218.36	4327.17	4574.38	4163.96	4249.36	4238.34	0.80	<.001
2	2	3889.83	4049.16	4412.27	3809.06	3933.83	3919.09	0.61	<.001
2	3	3844.34	4035.28	4634.79	3721.77	3909.34	3888.59	0.65	<.001
2	4	3843.89	4180.44	4916.48	3693.39	3943.89	3917.12	0.71	<.001

[a]TT-C: Tough Talks COVID.

[b]AIC: Akaike information criterion.

[c]BIC: Bayesian information criterion.

[d]AWE: Approximate Weight of Evidence.

[e]CLC: consistent likelihood criterion.

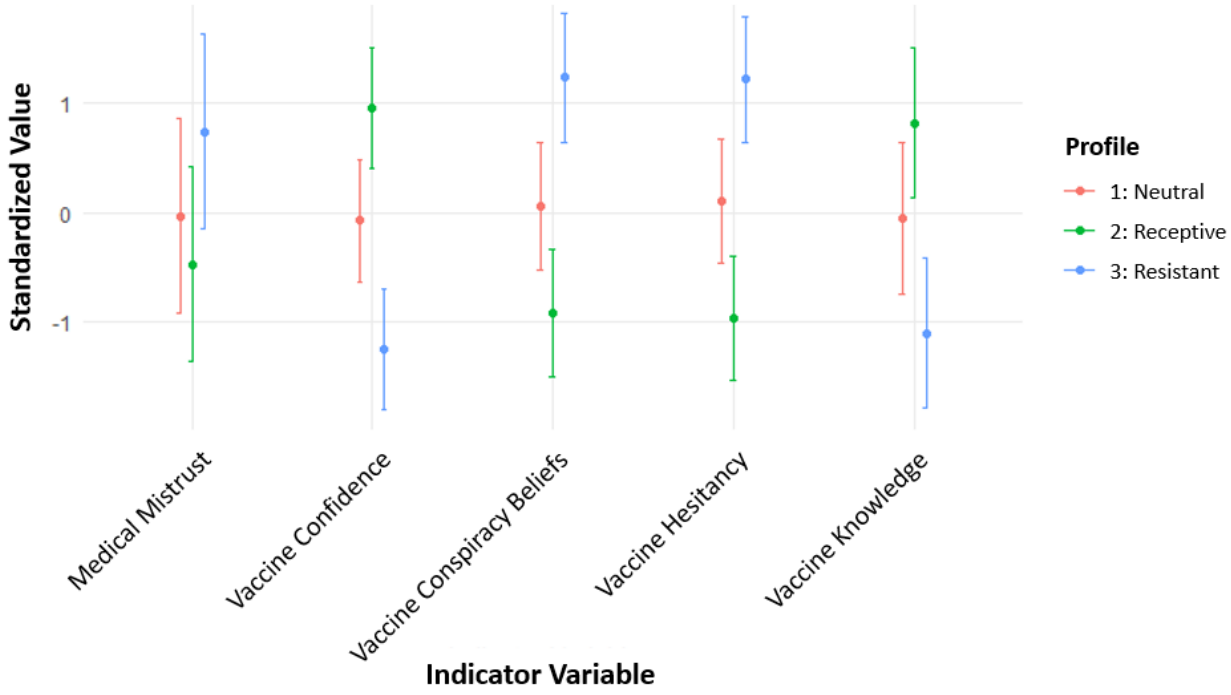
[f]KIC: Kullback information criterion.

[g]SABIC: Sample size-adjusted Bayesian information criterion.

[h]BLRT: Bootstrap likelihood ratio.

XSL•FO
RenderX

Figure 1. Standardized values of each indicator variable included in the three-profile model that best fit the latent profile analysis of baseline vaccine attitudes in the TT-C study. Points represent the mean standardized value for each profile, with bars showing ± 1 SD. The vaccine-resistant group (Profile 3 in blue) showed the highest medical mistrust, vaccine conspiracy beliefs, and vaccine hesitancy, coupled with the lowest vaccine confidence and knowledge. TT-C: Tough Talks COVID.



XSL•FO
RenderX

Table . Sociodemographic and baseline variables of TT-C[a] study participants by vaccine attitude profile.

Baseline variable	Profile 2 Vaccine-receptive (n=124)	Profile 1 Vaccine-neutral (n=155)	Profile 3 Vaccine-resistant (n=81)	*P* value
Vaccine hesitancy (ranges from 0 - 5)				<.001
Median (IQR)	1.8 (1.4-2.0)	2.5 (2.2-2.8)	3.3 (3.0-3.8)	
Mean (SD)	1.7 (0.3)	2.5 (0.5)	3.4 (0.6)	
Vaccination knowledge score (ranges from 0 - 4)				<.001
Median (IQR)	4 (3-4)	2 (2-3)	1 (0-1)	
Mean (SD)	3.4 (0.6)	2.2 (0.9)	0.8 (0.8)	
Vaccine confidence (ranges from 0 - 5)				<.001
Median (IQR)	4.1 (4.0-4.4)	3.4 (3.1-3.7)	2.6 (2.1-2.9)	
Mean (SD)	4.2 (0.4)	3.4 (0.4)	2.5 (0.5)	
Vaccine conspiracy beliefs (ranges from 0 - 5)				<.001
Median (IQR)	1.6 (1.2-1.8)	2.2 (2.0-2.6)	3.1 (3.0-3.6)	
Mean (SD)	1.6 (0.4)	2.3 (0.4)	3.2 (0.5)	
Medical mistrust (ranges from 0 - 60)				<.001
Median (IQR)	32 (29-35)	35 (32-38)	39 (36-42)	
Mean (SD)	31.9 (4.6)	34.6 (5.0)	39.4 (6.3)	
Randomization arm, n (%)				.20
Intervention	55 (44)	78 (50)	47 (58)	
Control	69 (56)	77 (50)	34 (42)	
Sex assigned at birth, n (%)				.08
Male	19 (15)	28 (18)	14 (17)	
Female	105 (85)	127 (82)	67 (83)	
Gender identity, n (%)				.70
Man (cis or trans)	17 (14)	25 (16)	15 (19)	
Woman (cis or trans)	93 (75)	119 (77)	60 (74)	
Gender fluid, gender queer, gender nonconforming, nonbinary, agender or other gender	14 (11)	11 (7)	6 (7)	
Age (years)				.05
Median (IQR)	23.0 (20.0-27.0)	24.0 (21.0-26.0)	24.0 (23.0-27.0)	
Mean (SD)	23.3 (3.7)	23.6 (3.3)	24.5 (3.1)	
Hispanic/Latino ethnicity, n (%)	5 (4)	9 (6)	2 (3)	.50
Flu vaccination (ever), n (%)				.05
No	23 (19)	33 (21)	27 (33)	
Yes	97 (78)	114 (74)	48 (59)	
Flu vaccination (past 12 months), n (%)				.01
No	55 (57)	72 (63)	40 (83)	
Yes	40 (41)	33 (29)	7 (15)	

XSL•FO
RenderX

Baseline variable	Profile 2 Vaccine-receptive (n=124)	Profile 1 Vaccine-neutral (n=155)	Profile 3 Vaccine-resistant (n=81)	*P* value
Human papillomavirus (any), n (%)				.12
No	23 (19)	40 (26)	25 (31)	
Yes	83 (67)	86 (55)	39 (48)	
Human papillomavirus vaccination (full series), n (%)				.20
No	8 (9.6)	16 (19)	4 (10)	
Yes	73 (88)	62 (72)	33 (85)	
Political affiliation, n (%)				.001
Democrat	74 (60)	70 (45)	27 (33)	
Republican	2 (1.6)	6 (3.9)	3 (3.7)	
Independent	27 (22)	46 (30)	28 (35)	
Other	8 (6.5)	3 (1.9)	1 (1.2)	
Social support (ranges from 0 - 5), n (%)				.02
Median (IQR)	4.2 (3.5-4.8)	3.9 (3.0-4.7)	3.8 (2.6-4.8)	
Mean (SD)	4.1 (0.8)	3.8 (0.9)	3.6 (1.1)	
Annual income (USD), n (%)				.01
≤$24,999	31 (25)	63 (42)	32 (40)	
$25,000 - $49,999	41 (34)	46 (29)	24 (30)	
$50,000 - $74,999	20 (16)	17 (11)	9 (11)	
$75,000 - $99,999	11 (8.9)	11 (7.1)	3 (3.7)	
≥$100,000	12 (9.7)	5 (3.2)	1 (1.2)	
Insurance coverage, n (%)				<.001
None	12 (9.7)	18 (12)	12 (15)	
Private	81 (65)	66 (43)	23 (28)	
Public	27 (22)	59 (38)	40 (49)	
Education level, n (%)				.20
9th to 12th grade, no diploma	8 (6.5)	6 (3.9)	2 (2.5)	
High school graduate or GED completed	22 (18)	42 (27)	24 (30)	
Some college level/technical/vocational degree	45 (36)	50 (32)	36 (44)	
Bachelor’s degree	36 (29)	39 (25)	13 (16)	
Other advanced degree (Master’s, Doctoral degree)	12 (9.7)	17 (11)	5 (6.2)	
Employment status, n (%)				.30
Working, full-time	54 (44)	64 (41)	35 (43)	
Working, part-time	20 (16)	19 (12)	12 (15)	
Not working	14 (11)	26 (17)	11 (14)	
Student	34 (27)	38 (25)	14 (17)	
Disabled, permanently or temporarily	0 (0)	4 (2.6)	3 (3.7)	

XSL•FO
RenderX

Baseline variable	Profile 2 Vaccine-receptive (n=124)	Profile 1 Vaccine-neutral (n=155)	Profile 3 Vaccine-resistant (n=81)	*P* value
Other	2 (1.6)	3 (1.9)	5 (6.2)	

[a]TT-C: Tough Talks COVID

Individuals in the vaccine-receptive group was significantly more likely to have been vaccinated for the flu in the past 12 months (n=40, 41%) compared with those in the vaccine-neutral (n=33, 29%) and the vaccine-resistant (n=7, 15%) groups (*P*=.01). Additionally, the vaccine-receptive group was more likely to report Democratic political affiliation (n=74, 60%) compared to the vaccine-neutral (n=70, 45%) and vaccinate-resistant (n=27, 33%) groups (*P*=.001). Reported annual income and insurance coverage were highest among the vaccine-receptive group; however, there were no significant differences in educational attainment or employment status across profiles. The vaccine-receptive group also tended to be younger (mean age, 23.3, SD 3.7) years and reported more social support (mean score, 4.1, SD 0.8) compared with both the vaccine-neutral (mean age, 23.6, SD 3.3 years; mean social support score, 3.8, SD, 0.9) and vaccine-resistant (mean age, 24.5, SD 3.1 years; mean social support score, 3.6, SD 1.1) groups (*P*=.05).

EMM by Baseline Vaccine Attitude Profile

Effect measure modification analysis showed that the TT-C DHI did not have a different effect on vaccine uptake across latent profile subgroups (Table 3). For secondary outcomes, EMM showed that the TT-C DHI had differing effects among those identified as vaccine-resistant compared with those classified as vaccine-receptive at baseline for the outcomes of vaccine hesitancy (interaction *P*=.07), vaccine confidence (interaction *P*=.06), and vaccine-related conspiracy beliefs (interaction *P*=.04) at 3 months (Table 3). There were also differing effects between vaccine-neutral and vaccine-receptive groups from baseline for the outcomes of vaccine hesitancy (interaction *P*=.05) and vaccine confidence (interaction *P*=.05) at 3 months. The vaccine-resistant group showed the greatest effect from the TT-C DHI, with a reduced mean vaccine hesitancy score of −0.40 (95% CI −0.76 to −0.37), an increased mean vaccine confidence score of 0.39 (95% CI 0.02 to 0.75), and a reduced conspiracy belief score of −0.47 (95% CI −0.86 to −0.09). The vaccine-neutral subgroup experienced moderate effects, including a reduced mean vaccine hesitancy score of −0.36 (95% CI −0.52 to −0.19) and an increased mean vaccine confidence score of 0.35 (95% CI 0.18 to 0.51). The TT-C DHI had no significant effect on any of the secondary vaccine-related outcomes among the vaccine-receptive subgroup.

XSL•FO
RenderX

Table . Results of the effect measure modification assessment for the impact of the TT-C[a] DHI[b] on primary and secondary outcomes at month 3 for the three vaccine attitude profiles at baseline.

Outcomes	Odds ratio (95% CI) or Diff (95% CI)[c]	Interaction *P* value[d]
Primary outcomes		
Vaccine uptake (3 months)		.79
Vaccine-resistant	1.81 (0.41-7.90)	
Vaccine-neutral	not enough data	
Vaccine-receptive	1.32 (0.35-5.04)	
Secondary outcomes		
Vaccine hesitancy (3 months)		
Vaccine-resistant	−0.40 (−0.76 to −0.37)	.07
Vaccine-neutral	−0.36 (−0.52 to −0.19)	.05
Vaccine-receptive	−0.08 (−0.26 to 0.09)	_[e]
Vaccine confidence (3 months)		
Vaccine-resistant	0.39 (0.02-0.75)	.06
Vaccine-neutral	0.35 (0.18-0.51)	.05
Vaccine-receptive	0.06 (−0.11 to 0.24)	–
Vaccine knowledge (3 months)		
Vaccine-resistant	0.90 (0.27-1.53)	.12
Vaccine-neutral	0.81 (0.41-1.20)	.17
Vaccine-receptive	0.16 (0.07-0.40)	–
Vaccine-related conspiracy beliefs (3 months)		
Vaccine-resistant	−0.47 (−0.86 to −0.09)	.04
Vaccine-neutral	−0.24 (−0.44 to −0.03)	.36
Vaccine-receptive	−0.09 (−0.26 to 0.08)	–

[a]TTC: Tough Talks COVID

[b]DHI: digital health interventions

[c]Odds ratios (95% CI) are presented for primary outcomes and Diff (95% CI) for secondary outcomes.

[d]Reference for the interaction *P*-value is the vaccine-receptive subgroup with significance indicated by $P<.10$

[e]not applicable

Sensitivity Analysis

When using the prespecified fit criteria with the analytic hierarchy process, we found the LCA model with the best fit to include three classes, similar to the LPA results. Similar to our LPA results, the three classes represented subgroups of study participants who were either: (1) vaccine-receptive, with low mean scores for mistrust, conspiracy beliefs, or hesitancy and high mean scores for knowledge and confidence; (2) vaccine-resistant, with high mean scores for mistrust, conspiracy beliefs, and hesitancy and low mean scores for knowledge or confidence; or (3) somewhere in between the two, as vaccine-neutral. The classes from the LCA had similar differences as the profiles in the LPA analysis with respect to political affiliation, income, insurance coverage, and flu vaccination. Thus, LCA results reinforce the findings from the LPA.

Discussion

Our LPA of Black young adults participating in an RCT for the TT-C DHI found three distinct subgroups of individuals with different baseline vaccine attitudes. The three subgroups ranged from vaccine-receptive to vaccine-resistant and differed on several sociodemographic variables, such as political affiliation, income, and social support. While there were no significant differences in vaccine uptake by intervention arm between the three subgroups, we identified modification whereby the TT-C DHI had the largest improvement in hesitancy, confidence, and conspiracy beliefs at 3 months among the subgroups who were vaccine-resistant and vaccine-neutral at baseline compared to those who were vaccine-receptive.

To our knowledge, this is the first study to use LPA to explore the efficacy of DHIs for improving vaccine outcomes among Black young adults in the United States South, a population understudied and underserved in vaccine research. Our results

XSL•FO
RenderX

from TT-C align with those of several social media and video-based educational DHIs across different populations, which showed the biggest improvements in vaccine outcomes among those with poor vaccine attitudes at baseline [45,46]. Because the TT-C DHI was cocreated with youth from the community and incorporated digital storytelling—a method well suited for marginalized communities [32]—we believe we prevented experiences of "backfiring" or reverse effects seen in more generalized vaccine-related DHIs where vaccine attitudes worsened among those with poor attitudes at baseline [47,48]. Additionally, the vaccine-receptive subgroup in this study had high confidence and knowledge in vaccines, with little hesitancy, mistrust, or conspiracy beliefs despite being insufficiently vaccinated. These baseline scores left limited room for improvement at month 3, which likely explains why TT-C differs from some other vaccine DHIs that saw the largest improvements in vaccine outcomes among those with more positive attitudes at baseline [49,50].

The success of this community-based participatory research project in developing a DHI for Black young adults in the United States South mirrors that of other DHIs aimed at acknowledging and addressing the historical injustices and ongoing inequities that drive distrust within communities of color, while using culturally appropriate messages and existing community infrastructure to address concerns about COVID-19 vaccine safety and access [51-55]. By using digital storytelling to convey accurate, relatable, and culturally resonant messages, DHIs can foster trust and engagement in positive vaccine beliefs, which can serve as a model for improving vaccination among other underserved or vaccine-resistant populations, both for COVID-19 and other vaccines. This work provides a blueprint for public health initiatives to adapt and implement culturally sensitive vaccine interventions that help build the capacity to identify inaccurate or misleading information online and to encourage individuals to make informed decisions. However, it also highlights the challenges of translating changed attitudes into actions.

Despite the significant improvements seen in vaccine attitudes, the lack of impact from the TTC-DHI on actualized vaccination behavior among any subgroup of Black young adults in this study may be attributed to a number of external factors. One key factor was the availability and access of COVID-19 vaccines during the study period of late 2023, when primary series promotion was shifting to boosters. Changing public health guidelines and the short 3-month timeframe in the study may have resulted in a lack of opportunity for behavioral action to get vaccinated, as the recommended frequency of boosters was greater than 3 months and it may have been less likely for clinics to offer the primary vaccination series. Indeed, only 21 (6.4%) participants received a new COVID-19 vaccine over the 3-month follow-up period. Additionally, the Health Belief Model has shown that a change in attitudes is not always enough to effect behavior change; certain conditions, such as perceived risk of infection and resource mobilization, are also required [56,57]. Lastly, other external influences, such as ongoing misinformation and continued structural inequalities in access, may have dampened the direct impact of the intervention on behavior, even as attitudes shifted positively.

This study draws on a large sample of Black young adults from the United States South, and we used a comprehensive range of validated vaccine scales as constructs to conduct the LPA. However, there are several limitations to this work. First, the landscape of COVID-19 science was changing quickly while the app was being developed and implemented (2021 - 2024); although app content was responsive to these changes, not all components were adjustable once developed, which may limit the generalizability and overall effectiveness of app messaging. This evolving context may have influenced participants' attitudes and behaviors in ways that are difficult to fully capture, as new information or policy changes could have overshadowed the effects of the intervention itself. Second, information about vaccine attitudes was collected via self-report, which may be misreported due to social desirability bias. Lastly, the use of LPA does not guarantee proper creation or assignment of subgroups and overlap between identified profiles may impact findings. However, our sensitivity analysis resulted in the same number of subgroups with similar distributions and effect measures, strengthening our confidence in these results and our study conclusions.

This analysis found that Black young adults who are vaccine-resistant or vaccine-neutral may experience larger improvements in vaccine outcomes when using DHIs like TT-C. Future researchers and policymakers can use this knowledge to create better-informed community-based policies, practices, and messaging related to vaccine uptake and attitudes while understanding subpopulations to target for resource efficiency and to drive the biggest improvement in vaccine outcomes. However, vaccine-receptive individuals may experience other benefits from these types of DHIs not assessed, such as improved linkage to vaccine resources or social support and thus should not be excluded. Vaccine-receptive individuals could also be leveraged as peer leaders or vaccine advocates in their respective communities to further increase uptake and acceptability of DHIs among Black young adults in the United States South.

Acknowledgments

The research reported in this publication was supported by the National Institute of Minority Health Disparities of the National Institutes of Health under award number R01MD016834. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

XSL•FO
RenderX

Disclaimer

Research reported in this publication was supported by the National Institute on Minority Health And Health Disparities of the National Institutes of Health under Award Number R01MD016834. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

Conceptualization: NM, LHW, HB, MS
Data curation: JM, JS
Formal analysis: NM, JM, EB, MS
Funding acquisition: LHW, HB
Investigation: AMB, JS, ZS, LB, MLC, MAL, KM, LHW, HB, MS
Methodology: NM, EB, MS
Project administration: JS, ZS, LB, LHW, HB
Validation: IY, EB
Writing - original draft: NM, JM, EB, MS
Writing - review & editing: NM, JM, EB, AMB, JS, ZS, LB, IY, MLC, MAL, KM, AP, LHW, HB, MS

Conflicts of Interest

None declared.

References

1. Timeline of WHO's response to COVID-19. World Health Organization. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline [accessed 2024-10-02]
2. Oliver SE, Gargano JW, Marin M, et al. The Advisory Committee on Immunization Practices' Interim Recommendation for use of Moderna COVID-19 vaccine — United States, December 2020. MMWR Morb Mortal Wkly Rep 2021 Jan 1;69(5152):1653-1656. [doi: 10.15585/mmwr.mm695152e1] [Medline: 33382675]
3. Parums DV. Editorial: First full regulatory approval of a COVID-19 vaccine, the BNT162b2 Pfizer-BioNTech vaccine, and the real-world implications for public health policy. Med Sci Monit 2021 Sep 6;27:e934625. [doi: 10.12659/MSM.934625] [Medline: 34483336]
4. Funk C, Tyson A. Intent to get a COVID-19 vaccine rises to 60% as confidence in research and development process increases. : Pew Research Center; 2020. URL: https://www.pewresearch.org/science/2020/12/03/intent-to-get-a-covid-19-vaccine-rises-to-60-as-confidence-in-research-and-development-process-increases/ [accessed 2024-10-02]
5. Gurley S, Bennett B, Sullivan PS, et al. COVID-19 vaccine perceptions, intentions, and uptake among young adults in the United States: prospective college-based cohort study. JMIR Public Health Surveill 2021 Dec 15;7(12):e33739. [doi: 10.2196/33739] [Medline: 34847054]
6. Willis DE, Presley J, Williams M, Zaller N, McElfish PA. COVID-19 vaccine hesitancy among youth. Hum Vaccin Immunother 2021 Dec 2;17(12):5013-5015. [doi: 10.1080/21645515.2021.1989923] [Medline: 34715003]
7. Gold JAW, Rossen LM, Ahmad FB, et al. Race, ethnicity, and age trends in persons who died from COVID-19 - United States, May-August 2020. MMWR Morb Mortal Wkly Rep 2020 Oct 23;69(42):1517-1521. [doi: 10.15585/mmwr.mm6942e1] [Medline: 33090984]
8. Cheng Z, Li Y. Racial and ethnic and income disparities in COVID-19 vaccination among Medicare beneficiaries. J Am Geriatr Soc 2022 Sep;70(9):2638-2645. [doi: 10.1111/jgs.17920] [Medline: 35639044]
9. Nuriddin A, Mooney G, White AIR. Reckoning with histories of medical racism and violence in the USA. Lancet 2020 Oct 3;396(10256):949-951. [doi: 10.1016/S0140-6736(20)32032-8] [Medline: 33010829]
10. Bogart LM, Dong L, Gandhi P, et al. COVID-19 vaccine intentions and mistrust in a national sample of Black Americans. J Natl Med Assoc 2022 Jan;113(6):599-611. [doi: 10.1016/j.jnma.2021.05.011] [Medline: 34158171]
11. Tsui J, Martinez B, Shin MB, et al. Understanding medical mistrust and HPV vaccine hesitancy among multiethnic parents in Los Angeles. J Behav Med 2023 Apr;46(1-2):100-115. [doi: 10.1007/s10865-022-00283-9] [Medline: 35107656]
12. Bogart LM, Ojikutu BO, Tyagi K, et al. COVID-19 related medical mistrust, health impacts, and potential vaccine hesitancy among Black Americans living with HIV. J Acquir Immune Defic Syndr 2021 Feb 1;86(2):200-207. [doi: 10.1097/QAI.0000000000002570] [Medline: 33196555]
13. Wall LL. J. Marion Sims and the vesicovaginal fistula: historical understanding, medical ethics, and modern political sensibilities. Female Pelvic Med Reconstr Surg 2018;24(2):66-75. [doi: 10.1097/SPV.0000000000000546] [Medline: 29474277]
14. Alang S, McAlpine DD, Hardeman R. Police brutality and mistrust in medical institutions. J Racial Ethn Health Disparities 2020 Aug;7(4):760-768. [doi: 10.1007/s40615-020-00706-w] [Medline: 31989532]

XSL•FO
RenderX

15. Bogart LM, Takada S, Cunningham WE. Medical mistrust, discrimination, and the domestic HIV epidemic. In: Ojikutu BO, Stone VE, editors. HIV in US Communities of Color: Springer International Publishing; 2021:207-231. [doi: 10.1007/978-3-030-48744-7_12]
16. Shon EJ, Lee L. Effects of vaccine literacy, health beliefs, and flu vaccination on perceived physical health status among under/graduate students. Vaccines (Basel) 2023 Mar 30;11(4):765. [doi: 10.3390/vaccines11040765] [Medline: 37112677]
17. Jamison AM, Quinn SC, Freimuth VS. "You don't trust a government vaccine": narratives of institutional trust and influenza vaccination among African American and White adults. Soc Sci Med 2019 Jan;221:87-94. [doi: 10.1016/j.socscimed.2018.12.020] [Medline: 30576982]
18. Quinn SC, Hilyard KM, Jamison AM, et al. The influence of social norms on flu vaccination among African American and White adults. Health Educ Res 2017 Dec 1;32(6):473-486. [doi: 10.1093/her/cyx070] [Medline: 29220514]
19. Cohen CM, Wentzensen N, Castle PE, et al. Racial and ethnic disparities in cervical cancer incidence, survival, and mortality by histologic subtype. J Clin Oncol 2023 Feb 10;41(5):1059-1068. [doi: 10.1200/JCO.22.01424] [Medline: 36455190]
20. Cook RL, Zhang J, Mullins J, et al. Factors associated with initiation and completion of human papillomavirus vaccine series among young women enrolled in Medicaid. J Adolesc Health 2010 Dec;47(6):596-599. [doi: 10.1016/j.jadohealth.2010.09.015] [Medline: 21094437]
21. Skafle I, Nordahl-Hansen A, Quintana DS, Wynn R, Gabarron E. Misinformation about COVID-19 vaccines on social media: rapid review. J Med Internet Res 2022 Aug 4;24(8):e37367. [doi: 10.2196/37367] [Medline: 35816685]
22. Gabarron E, Oyeyemi SO, Wynn R. COVID-19-related misinformation on social media: a systematic review. Bull World Health Organ 2021 Jun 1;99(6):455-463A. [doi: 10.2471/BLT.20.276782] [Medline: 34108756]
23. Jack BW, Bickmore T, Yinusa-Nyahkoon L, et al. Improving the health of young African American women in the preconception period using health information technology: a randomised controlled trial. Lancet Digit Health 2020 Sep;2(9):e475-e485. [doi: 10.1016/S2589-7500(20)30189-8] [Medline: 33328115]
24. Cyriac J, Jenkins S, Patten CA, et al. Improvements in diet and physical activity-related psychosocial factors among African Americans using a mobile health lifestyle intervention to promote cardiovascular health: the FAITH! (Fostering African American Improvement in Total Health) app pilot study. JMIR Mhealth Uhealth 2021 Nov 12;9(11):e28024. [doi: 10.2196/28024] [Medline: 34766917]
25. Zhou ES, Ritterband LM, Bethea TN, Robles YP, Heeren TC, Rosenberg L. Effect of culturally tailored, internet-delivered cognitive behavioral therapy for insomnia in Black women: a randomized clinical trial. JAMA Psychiatry 2022 Jun 1;79(6):538-549. [doi: 10.1001/jamapsychiatry.2022.0653] [Medline: 35442432]
26. Mancuso N, Mansergh G, Stephenson R, et al. Factors associated with mobile app-based ordering of HIV self-test kits among men who have sex with men in Atlanta, Detroit and New York City: an exploratory secondary analysis of a randomized control trial. J Int AIDS Soc 2023 May;26(5):e26100. [doi: 10.1002/jia2.26100] [Medline: 37221941]
27. Stoner MCD, Smith L, Ming K, et al. Results from a pilot study of an automated directly observed therapy intervention using artificial intelligence with conditional economic incentives among young adults with HIV. J Acquir Immune Defic Syndr 2024 Jun 1;96(2):136-146. [doi: 10.1097/QAI.0000000000003397] [Medline: 38363868]
28. Faverio M. Share of those 65 and older who are tech users has grown in the past decade.: Pew Research Center; 2022. URL: https://www.pewresearch.org/short-reads/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ [accessed 2024-10-02]
29. Hill PW, Diamond J, Spiegel AN, VanWormer E, Leadabrand M, McQuillan J. Accuracy of COVID-19 relevant knowledge among youth: number of information sources matters. PLOS ONE 2022;17(12):e0267871. [doi: 10.1371/journal.pone.0267871] [Medline: 36574374]
30. Budhwani H, Maragh-Bass AC, Tolley EE, et al. Tough Talks COVID-19 digital health intervention for vaccine hesitancy among Black young adults: protocol for a hybrid type 1 effectiveness implementation randomized controlled trial. JMIR Res Protoc 2023 Feb 13;12(1):e41240. [doi: 10.2196/41240] [Medline: 36689557]
31. Wallerstein N, Duran B, Oetzel JG, Minkler M. Community-Based Participatory Research for Health: Advancing Social and Health Equity, 3rd edition: Wiley; 2017. URL: https://www.wiley.com/en-dk/Community-Based+Participatory+Research+for+Health%3A+Advancing+Social+and+Health+Equity%2C+3rd+Edition-p-9781119258858 [accessed 2024-10-09]
32. de Jager A, Fogarty A, Tewson A, Lenette C, Boydell K. Digital storytelling in research: a systematic review. TQR 2017;22(10):2548-2582. [doi: 10.46743/2160-3715/2017.2970]
33. Stoner MCD, Tweedy D, Comello MGL, et al. Using narratives to inform the development of a digital health intervention related to COVID-19 vaccination in Black young adults in Georgia, North Carolina and Alabama. Vaccine (Auckl) 2022 Nov 15;40(48):6908-6916. [doi: 10.1016/j.vaccine.2022.10.027] [Medline: 36280559]
34. Shapiro GK, Tatar O, Dube E, et al. The vaccine hesitancy scale: psychometric properties and validation. Vaccine (Auckl) 2018 Jan 29;36(5):660-667. [doi: 10.1016/j.vaccine.2017.12.043] [Medline: 29289384]
35. Gilkey MB, Reiter PL, Magnus BE, McRee AL, Dempsey AF, Brewer NT. Validation of the vaccination confidence scale: a brief measure to identify parents at risk for refusing adolescent vaccines. Acad Pediatr 2016;16(1):42-49. [doi: 10.1016/j.acap.2015.06.007] [Medline: 26300368]

XSL•FO
RenderX

36. Shapiro GK, Holding A, Perez S, Amsel R, Rosberger Z. Validation of the vaccine conspiracy beliefs scale. Papillomavirus Res 2016 Dec;2:167-172. [doi: 10.1016/j.pvr.2016.09.001] [Medline: 29074176]
37. Zingg A, Siegrist M. Measuring people's knowledge about vaccination: developing a one-dimensional scale. Vaccine (Auckl) 2012 May 28;30(25):3771-3777. [doi: 10.1016/j.vaccine.2012.03.014] [Medline: 22445808]
38. Shelton RC, Winkel G, Davis SN, et al. Validation of the group-based medical mistrust scale among urban Black men. J Gen Intern Med 2010 Jun;25(6):549-555. [doi: 10.1007/s11606-010-1288-y] [Medline: 20195782]
39. Sherbourne CD, Stewart AL. The MOS social support survey. Social Science & Medicine 1991 Jan;32(6):705-714. [doi: 10.1016/0277-9536(91)90150-B]
40. Muthén B, Muthén LK. Integrating person-centered and variable-centered analyses: growth mixture modeling with latent trajectory classes. Alcohol Clin Exp Res 2000 Jun;24(6):882-891. [doi: 10.1111/j.1530-0277.2000.tb02070.x] [Medline: 10888079]
41. Jason L, Glenwick D. Handbook of Methodological Approaches to Community-Based Research: Qualitative, Quantitative, and Mixed Methods: Oxford University Press; 2016. [doi: 10.1093/med:psych/9780190243654.001.0001]
42. Akogul S, Erisoglu M. An approach for determining the number of clusters in a model-based cluster analysis. Entropy (Basel) 2017;19(9):452. [doi: 10.3390/e19090452]
43. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: https://www.r-project.org/ [accessed 2024-08-01]
44. Scotto Rosato N, Baer JC. Latent class analysis: a method for capturing heterogeneity. Soc Work Res 2012 Mar 1;36(1):61-69. [doi: 10.1093/swr/svs006]
45. Vandormael A, Adam M, Greuel M, et al. The effect of a wordless, animated, social media video intervention on COVID-19 prevention: online randomized controlled trial. JMIR Public Health Surveill 2021 Jul 27;7(7):e29060. [doi: 10.2196/29060] [Medline: 34174778]
46. Bender FL, Rief W, Brück J, Wilhelm M. Effects of a video-based positive side-effect information framing: an online experiment. Health Psychol 2023 Aug;42(8):593-602. [doi: 10.1037/hea0001263] [Medline: 36780276]
47. van Stekelenburg A, Schaap G, Veling H, Buijzen M. Investigating and improving the accuracy of US citizens' beliefs about the COVID-19 pandemic: longitudinal survey study. J Med Internet Res 2021 Jan 12;23(1):e24069. [doi: 10.2196/24069] [Medline: 33351776]
48. Freeman D, Loe BS, Yu LM, et al. Effects of different types of written vaccination information on COVID-19 vaccine hesitancy in the UK (OCEANS-III): a single-blind, parallel-group, randomised controlled trial. Lancet Public Health 2021 Jun;6(6):e416-e427. [doi: 10.1016/S2468-2667(21)00096-7] [Medline: 33991482]
49. Johnson V, Butterfuss R, Kim J, Orcutt E, Harsch R, Kendeou P. The "Fauci Effect": Reducing COVID-19 misconceptions and vaccine hesitancy using an authentic multimodal intervention. Contemp Educ Psychol 2022 Jul;70:102084. [doi: 10.1016/j.cedpsych.2022.102084] [Medline: 35765462]
50. Amazeen MA, Krishna A, Eschmann R. Cutting the bunk: comparing the solo and aggregate effects of prebunking and debunking Covid-19 vaccine misinformation. Sci Commun 2022 Aug;44(4):387-417. [doi: 10.1177/10755470221111558]
51. Strully KW, Harrison TM, Pardo TA, Carleo-Evangelist J. Strategies to address COVID-19 vaccine hesitancy and mitigate health disparities in minority populations. Front Public Health 2021;9:645268. [doi: 10.3389/fpubh.2021.645268] [Medline: 33968884]
52. Ford KL, West AB, Bucher A, Osborn CY. Personalized digital health communications to increase COVID-19 vaccination in underserved populations: a double diamond approach to behavioral design. Front Digit Health 2022;4:831093. [doi: 10.3389/fdgth.2022.831093] [Medline: 35493533]
53. Koskan A, Larkey L, Todd M, Kim SW. Protocol for a community-based digital storytelling pilot intervention to reduce Hispanic parents' vaccine hesitancy to immunize their children against COVID-19. PLOS ONE 2024;19(3):e0299787. [doi: 10.1371/journal.pone.0299787] [Medline: 38502659]
54. Castellon-Lopez YM, Klomhaus AM, Garcia C, et al. MivacunaLA (MyshotLA): a community-partnered mobile phone intervention to improve COVID-19 vaccination behaviors among low-income, Spanish-speaking, and immigrant Latino parents or caregivers. Vaccines (Basel) 2024 May 8;12(5):511. [doi: 10.3390/vaccines12050511] [Medline: 38793762]
55. AuYoung M, Rodriguez Espinosa P, Chen WT, et al. Addressing racial/ethnic inequities in vaccine hesitancy and uptake: lessons learned from the California alliance against COVID-19. J Behav Med 2023 Apr;46(1-2):153-166. [doi: 10.1007/s10865-022-00284-8] [Medline: 35066696]
56. Limbu YB, Gautam RK, Pham L. The Health Belief Model applied to COVID-19 vaccine hesitancy: a systematic review. Vaccines (Basel) 2022 Jun 18;10(6):973. [doi: 10.3390/vaccines10060973] [Medline: 35746581]
57. Zampetakis LA, Melas C. The health belief model predicts vaccination intentions against COVID-19: a survey experiment approach. Appl Psychol Health Well Being 2021 May;13(2):469-484. [doi: 10.1111/aphw.12262] [Medline: 33634930]

Abbreviations

BLRT: bootstrap likelihood ratio
DHI: digital health intervention

XSL•FO
RenderX

EMM: effect measure modification
LCA: latent class analysis
LPA: latent profile analysis
TT-C: Tough Talks COVID

Edited by A Mavragani; submitted 15.10.24; peer-reviewed by E Gavor, K Adegoke, MGC Iwunwa, Q Fang, Y Xie; revised version received 27.02.25; accepted 04.03.25; published 16.04.25.

Please cite as:
Mancuso N, Michaels J, Browne EN, Maragh-Bass AC, Stocks JB, Soberano ZR, Bond CL, Yigit I, Comello MLG, Larsen MA, Muessig KE, Pettifor A, Hightow-Weidman LB, Budhwani H, Stoner MCD
Greater Improvements in Vaccination Outcomes Among Black Young Adults With Vaccine-Resistant Attitudes in the United States South Following a Digital Health Intervention: Latent Profile Analysis of a Randomized Control Trial
JMIR Public Health Surveill 2025;11:e67370
URL: https://publichealth.jmir.org/2025/1/e67370
doi:10.2196/67370

© Noah Mancuso, Jenna Michaels, Erica N Browne, Allysha C Maragh-Bass, Jacob B Stocks, Zachary R Soberano, C Lily Bond, Ibrahim Yigit, Maria Leonora G Comello, Margo Adams Larsen, Kathryn E Muessig, Audrey Pettifor, Lisa B Hightow-Weidman, Henna Budhwani, Marie C D Stoner. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 16.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Associations of Violence Against Women With Comorbid Symptoms of Depression and Anxiety Among Left-Behind Women in Rural China: Cross-Sectional Study

Leying Hou[1], BM; Zeyu Luo[1], BM; Weidi Sun[1], BM; Jiayao Ying[1], BM; Jing Wu[1], BM; Shiyi Shan[1], BM; Wen Liu[2], MPH; Peige Song[1], PhD

[1]Center for Clinical Big Data and Statistics of the Second Affiliated Hospital Zhejiang University School of Medicine, School of Public Health Zhejiang University School of Medicine, No. 866 Yuhangtang Road, Xihu District, Hangzhou City, Zhejiang Province, Hangzhou, China

[2]Department of General Practice and Primary Care, The University of Melbourne, Melbourne, Victoria, Australia

Corresponding Author:
Peige Song, PhD
Center for Clinical Big Data and Statistics of the Second Affiliated Hospital Zhejiang University School of Medicine, School of Public Health Zhejiang University School of Medicine, No. 866 Yuhangtang Road, Xihu District, Hangzhou City, Zhejiang Province, Hangzhou, China

Abstract

Background: Violence against women (VAW) is a major public health and human rights concern with profound mental health consequences. However, the association between specific VAW forms and mental health, particularly among left-behind women in rural China, remains unclear.

Objective: This study aimed to identify the associations of VAW with depression, anxiety, and comorbid symptoms and to explore the potential roles of resilience and social support.

Methods: The cross-sectional survey was conducted in Y City, Henan Province, China, in July 2023. A multistage stratified random sampling method was used to recruit left-behind women, resulting in a final sample of 1503 participants. Data on participants and their VAW were collected through a face-to-face questionnaire survey. The forms of VAW assessed were nonpartner violence (NPV) and intimate partner violence (IPV; including remote IPV). Depressive symptoms were evaluated using the 10-item Center for Epidemiological Studies Depression Scale, while anxiety symptoms were assessed with the Generalized Anxiety Disorder-7. The comorbid symptoms of depression and anxiety (CDA) were ascertained as the simultaneous presence of depressive and anxiety symptoms. A multivariable logistic regression model was used to estimate the odds ratio and 95% CIs. A 4-way decomposition analysis was conducted to test the mediation roles and interactions of resilience and social support between VAW and mental health outcomes. Population attributable fractions and pathway-specific population attributable fractions were calculated to estimate the burden of mental health outcomes attributable to VAW.

Results: Lifetime VAW (adjusted odds ratio [aOR] 1.84, 95% CI 1.32 - 2.54) was associated with an increased risk of CDA. Women who were exposed to lifetime IPV (aOR 1.84, 95% CI 1.32 - 2.56), remote IPV (aOR 2.79, 95% CI 1.60 - 4.74), and NPV (aOR 2.63, 95% CI 1.58 - 4.26) had an increased likelihood of reporting CDA. Similar associations could also be found for depressive symptoms and anxiety symptoms. In the 4-way decomposition analysis for VAW and CDA, mediation effects of low resilience and social support were statistically significant (P<.05), whereas none of the interactions reached significance (P>.05). The pure mediation proportion was 28.2% for the low resilience and 18.6% for the social support between VAW and CDA. A total of 20.8% of CDA cases, 15.1% of depressive symptoms cases, and 22.7% of anxiety symptoms cases were attributable to VAW. Among these, low resilience accounted for 7.2% and low social support accounted for 4.7% of CDA cases as mediators.

Conclusions: Lifetime VAW, including IPV (and remote IPV) and NPV, shows significant associations with CDA and depressive and anxiety symptoms among rural left-behind women in China. The associations are partly mediated by low resilience and social support. Targeted strategies, including efforts to reduce violence against rural left-behind women, enhance their resilience and strengthen their social support networks, are urgently needed.

(JMIR Public Health Surveill 2025;11:e72064) doi:10.2196/72064

KEYWORDS

left-behind women; violence; depressive and anxiety symptoms; resilience; social support

XSL•FO
RenderX

Introduction

With rapid urbanization and economic reforms, rural left-behind women—those whose husbands migrate to urban areas for employment—are becoming increasingly common in many low- and middle-income countries [1-3]. In China, over 50 million women are left behind in rural areas due to internal migration [4]. This demographic shift is shaped by structural inequalities, such as women's disadvantaged position, underdeveloped social services, and cultural norms like familism, which often prioritize family needs over individual desires [5]. As a result, left-behind women bear a "triple burden" of caring for children, supporting older family members, and managing agricultural responsibilities in the absence of their spouses [6]. The prolonged separation from their spouses, combined with socioeconomic pressures, raises concerns about their mental health [3].

Mental disorders, particularly depression and anxiety, have become increasingly prevalent and are now major global mental health concerns [7]. Previous studies showed that depression and anxiety frequently co-occur, sharing common symptoms and risk factors like adverse life events [8]. Comorbid symptoms of depression and anxiety (CDA) are more debilitating than either condition alone, leading to greater psychosocial impairment and poorer health outcomes [9]. Among left-behind women, these disorders are of particular concern due to the compounded stressors they face. The absence of a spouse, social isolation, and the potential for increased economic pressure can create an environment that fosters psychological distress [10,11]. Thus, attention to the mental health of left-behind women is essential, particularly given the limited availability of mental health services in rural settings [12].

A crucial, yet often overlooked, determinant of mental health is violence against women (VAW). VAW is a pervasive global issue encompassing physical, emotional, and sexual abuse, with intimate partner violence (IPV) being the most commonly studied form [13,14]. Although the physical absence of a husband might suggest reduced exposure to IPV, left-behind women are susceptible to experiencing remote IPV, a form of abuse facilitated by technology, where geographically distant partners exert control through harassment, threats, or emotional abuse via phone calls, messaging, or social media [15,16]. Moreover, nonpartner violence (NPV) should be given attention among left-behind women since they experience prolonged periods of living alone or with other family members [17,18]. A growing body of literature demonstrates that exposure to VAW, particularly IPV, is strongly associated with adverse mental health outcomes, including depression and anxiety [19,20]. For example, for women in Afghanistan under the Taliban government, exposure to traumatic events such as violence can contribute to depressive and anxiety symptoms, ultimately diminishing their quality of life [21-23]. However, significant gaps remain in our understanding of how different forms of VAW, specifically IPV, remote IPV, and NPV, impact mental health outcomes among left-behind women.

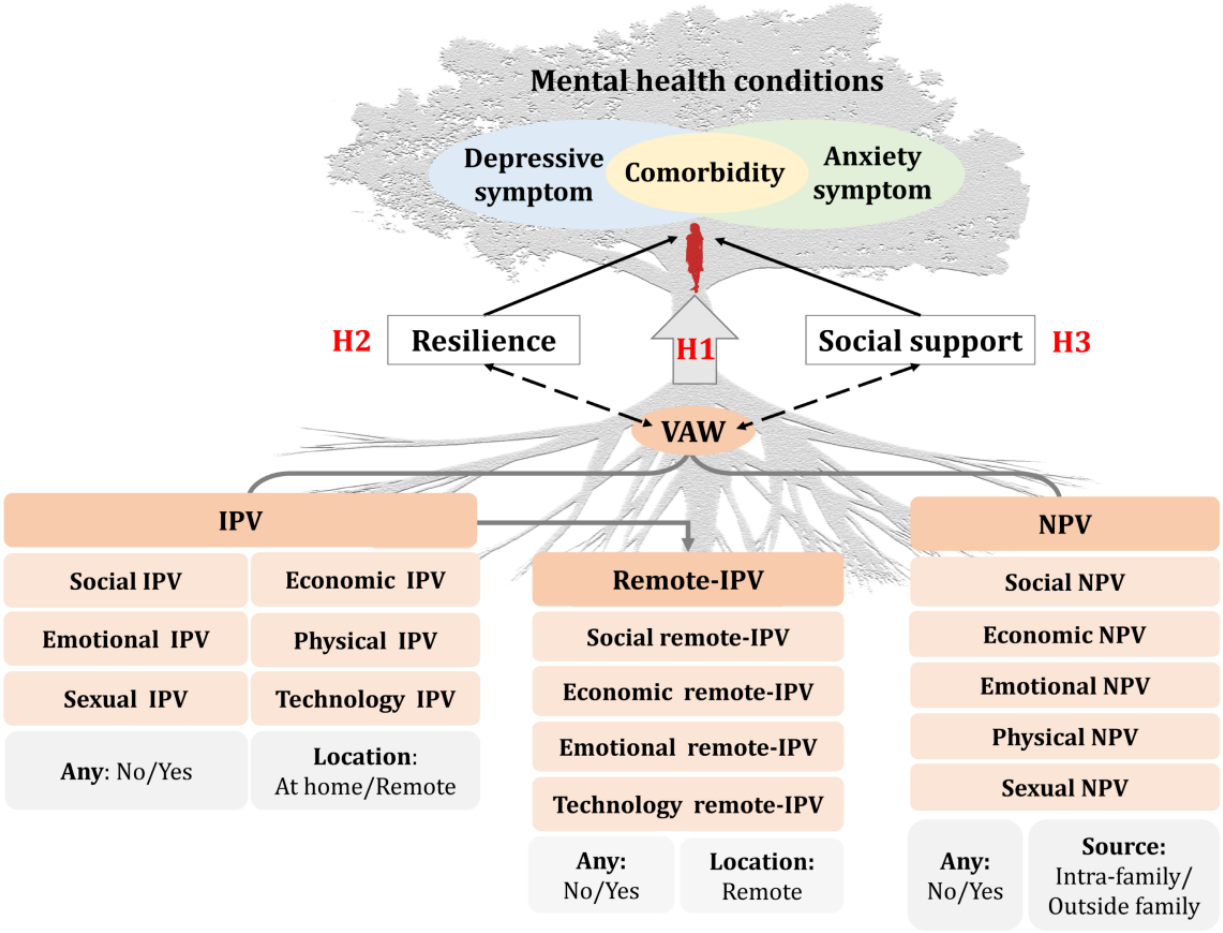
Resilience is the ability to adapt positively in the face of adversity [24]. Social support, similarly, is considered a positive role in improving mental health outcomes [25]. Despite their known protective effects, the roles of resilience and social support as mediators or moderators in the relationship between VAW and mental health, particularly among rural left-behind women, remain unclear.

This study aimed to address these gaps by investigating the associations of VAW with depressive, anxiety, and comorbid symptoms among left-behind women in rural China, and to explore the potential roles of resilience and social support. Specifically, this study sought to test the following three primary hypotheses (Figure 1):

1. Hypothesis 1 (H1): VAW is positively associated with depressive and anxiety symptoms, with different forms, types, and periods of violence exhibiting various associations.
2. Hypothesis 2 (H2): Resilience serves as a mediator and moderates the association between VAW and mental health conditions.
3. Hypothesis 3 (H3): Social support acts as a mediator and moderates the association between VAW and mental health conditions.

XSL•FO
RenderX

Figure 1. Conceptual framework illustrating the hypothesized associations between various forms and types of violence and mental health outcomes among rural left-behind women in China, with resilience and social support as potential mediators. In this figure, the tree trunk represents mental health status composed of depressive and anxiety symptoms, while the roots symbolize various forms of violence as underlying factors. The tree as a whole illustrates how these root-level exposures may influence the development of mental health outcomes. H1: hypothesis 1; H2: hypothesis 2; H3: hypothesis 3; IPV: intimate partner violence; NPV: nonpartner violence; VAW: violence against women.



Methods

Study Design and Participants

The cross-sectional survey was conducted in Y City, Henan Province, China, in July 2023, where there is frequent labor migration to urban areas for work (Figure S1 in Multimedia Appendix 1). For participant recruitment, a multistage stratified random sampling method was used. First, 7 towns were chosen: 3 low-, 2 medium-, and 2 high-economic towns. Second, 102 survey sites were selected within these towns. With support from local Women's Federation staff or social workers of these selected survey sites, participants were invited to complete a face-to-face questionnaire administered by trained field investigators in meeting rooms of village committee service centers. Data on women's sociodemographic and lifestyle information, violence exposure, and mental health outcomes were self-reported and collected following the World Health Organization recommendations to maximize participant safety and disclosure [26]. To ensure data quality, daily debriefing sessions were held with all field investigators to address potential questions arising from interviews based on their submitted questionnaires. The statistics team performed daily checks on antifraud questions, minimum completion time thresholds, and logical consistency checks (eg, implausible age values).

Eligible participants were rural married women aged 18 years or older with normal communication and understanding ability, whose husbands had been absent for at least 6 months within the past year. A total of 1525 questionnaires were finally retained for the survey, of which 1519 were valid. Participants with missing or ambiguous data on violence (n=3) or mental health outcomes (n=13) were excluded, resulting in a final sample of 1503 women (Figure S1 in Multimedia Appendix 1).

Ethical Considerations

The study protocol was approved by the Ethics Review Committee of Zhejiang University School of Public Health (approval number: ZGL202306-5). All participants provided written informed consent for participation and publication. Study data were anonymized prior to analysis. Each participant received a health gift package valued at approximately ¥20 (approximately US $2.80). No identifiable images of participants are included in the manuscript or Multimedia Appendix files.

XSL•FO
RenderX

Assessment of VAW

The questionnaire used to assess various forms of VAW in this study was developed based on the World Health Organization's VAW Instrument [27]. The adapted questionnaire was validated through a 2-round Delphi survey with 18 experts, ensuring its credibility. The lifetime VAW was divided into 2 primary forms according to the sources: IPV (violence from partners) and NPV (violence from other family members or outside the family) (Figure 1). Given that technology-driven, geographically distant abuse is unique for rural left-behind women, remote IPV was specifically measured. IPV was measured using 25 items across 6 types (social IPV, economic IPV, emotional IPV, physical IPV, sexual IPV, and technology-facilitated IPV). Remote IPV was assessed using 16 items across 4 types (social remote IPV, economic remote IPV, emotional remote IPV, technology-facilitated remote IPV). NPV was measured through 19 items across 5 types (social NPV, economic NPV, emotional NPV, physical NPV, and sexual NPV) (Table S1 in Multimedia Appendix 1).

Assessment of Mental Health Outcomes

Depressive symptoms were measured using the 10-item Center for Epidemiological Studies Depression scale [28], which demonstrated good internal consistency (Cronbach α=0.80) [29]. Participants were asked how frequently they experienced a series of depressive symptoms over the past week, with response options ranging from 0 (rarely or none of the time) to 3 (most or all of the time). The total 10-item Center for Epidemiological Studies Depression Scale score ranged from 0 to 30, with a cutoff score of 10 indicating the presence of depressive symptoms [30].

Anxiety symptoms were assessed using the generalized anxiety disorder-7 scale [31,32], which also showed excellent internal consistency (Cronbach α=0.92) [33]. The Generalized Anxiety Disorder-7 includes 7 items rated on a 4-point scale: 0 (not at all) to 3 (nearly every day), with a cutoff score of 5 indicating anxiety symptoms [34].

CDA was defined as the simultaneous presence of both depressive and anxiety symptoms. In this study, CDA, depressive symptoms, and anxiety symptoms were the primary mental health outcomes of interest.

Assessment of Resilience and Social Support

Resilience was measured using the 10-item Connor-Davidson Resilience Scale [24,35,36], with a Cronbach α of 0.98. Items were rated on a 0 - 4 scale, with higher scores indicating greater resilience (range 0 - 40). Social support was measured using the Social Support Rating Scale, including objective support (3 items), subjective support (4 items), and utilization of support (3 items), with a total score of 66 points. The Social Support Rating Scale has demonstrated acceptable internal consistency and reliability, with a Cronbach α of 0.70 in this study [37].

Ascertainment of Covariates

Covariates include women's sociodemographic, lifestyle, and left-behind related variables. The age of women was assessed as a continuous variable. Education was classified into primary school and below, middle school, and high school and above. Work was categorized into no job, farmer, formal job, casual job, or others. Annual household income in CNY was divided into below ¥30,000, ¥30,001 - 50,000, and ¥50,001 and above (all currency values in Chinese yuan [¥] are based on the conversion rate of ¥7.1 = US $1, the approximate exchange rate at the time of the study). An unhealthy lifestyle consists of smoking, drinking, and poor sleep quality. BMI was calculated as weight divided by the square of height and categorized into 3 levels (normal weight or under: <23.9 kg/m^2, overweight: 23.9 - 27.8 kg/m^2, and obesity: ≥27.9 kg/m^2) [38]. Chronic disease history was self-reported physician diagnosis of the common noncommunicable diseases, including hypertension, dyslipidemia, diabetes or hyperglycemia, gastrointestinal disorders, liver diseases, heart diseases, stroke, urological disorders, respiratory system diseases, arthritis, or rheumatism. Length of being left behind was divided into below 3 years, 3 - 6 years, 6 years and above.

Statistical Analysis

Descriptive statistics summarized rural left-behind women's characteristics by CDA, depressive, and anxiety symptoms. Skewed continuous variables were reported as medians (IQRs) and compared via the Wilcoxon rank-sum test. Categorical variables were presented as frequencies (%) and analyzed using Pearson chi-square or Fisher exact tests.

Multivariable logistic regressions were used to analyze associations of any VAW and its forms (IPV [including remote IPV] and NPV) as well as specific types with CDA, depressive, and anxiety symptoms. Models were adjusted for age, education, occupation, household income, healthy lifestyle, BMI, chronic disease history, and length of being left behind. Restricted cubic splines were used to investigate the dose-response associations between the number of VAW items and mental health outcomes.

A 4-way decomposition analysis was performed to quantify how much of the association between VAW and mental health outcomes could be attributed to mediation (the sum of the mediated interaction and the pure indirect effect) or interaction (the sum of the reference interaction and the mediated interaction) of low resilience and social support [39]. In addition, separate mediation and additive and multiplicative interaction analyses were also performed. Detailed methods are described in Multimedia Appendix 1.

Population attributable fractions (PAFs) were calculated to estimate the proportion of CDA, depressive symptoms, and anxiety symptoms cases that could potentially be attributable to VAW based on the following equation[40]: PAF=Pe(RR-1)1+PeRR-1. Our study replaced risk ratios (RRs) in the formula with odds ratios (ORs) obtained in the above logistic regression. The fraction of participants being exposed to the risk factor (Pe) and the OR of the outcome associated with the corresponding factors were derived from the study. The 95% CI of PAF was calculated using bootstrap [41]. Furthermore, we explored pathway-specific PAFs of potential mediators between VAW and outcomes, which indicated the relative reduction in outcome prevalence if a certain mediating pathway was disabled. The "graphPAF" R package was used to obtain pathway-specific PAFs [42,43].

XSL•FO
RenderX

All statistical analyses were performed using R software (version 4.2.1; R Foundation for Statistical Computing). Statistical significance was considered to be <.05 with a 2-sided, or 95% CIs of OR that did not cross 1.00.

Results

Characteristics

The characteristics of rural left-behind women stratified by CDA are shown in Table 1. Of the 1503 rural left-behind women, 12.8% (n=192) had CDA. Women with CDA were more likely to experience VAW (43.8% vs 29.4%), IPV (40.1% vs 26.6%), remote IPV (12.0% vs 4.4%), and NPV (14.6% vs 5.9%), compared with those without CDA (all *P*<.05). For VAW types, similar comparisons are also evident.

XSL•FO
RenderX

Table . Characteristics of Chinese rural left-behind women by comorbid symptoms of depression and anxiety, based on a cross-sectional study conducted in China in 2023.

Characteristic	Overall (n=1503)	No (n=1311)	CDA[a] (n=192)	*P* value[b]
Sociodemographic				
Age (year), median (IQR)	52.0 (43.0-57.0)	52.0 (43.0-57.0)	52.0 (42.0-56.3)	.97
Education, n (%)				.24
Primary school and below	526 (35.0)	449 (34.2)	77 (40.1)	
Middle school	788 (52.4)	693 (52.9)	95 (49.5)	
High school and above	189 (12.6)	169 (12.9)	20 (10.4)	
Occupation, n (%)				.20
No job	686 (45.6)	591 (45.1)	95 (49.5)	
Farmer or formal job	510 (33.9)	443 (33.8)	67 (34.9)	
Casual job or others	307 (20.4)	277 (21.1)	30 (15.6)	
Household income (¥)[c], n (%)				.001
Below 30,000	296 (19.7)	244 (18.6)	52 (27.1)	
30,001 - 50,000	531 (35.3)	456 (34.8)	75 (39.1)	
50,001 and above	676 (45.0)	611 (46.6)	65 (33.9)	
Smoking status, n (%)				.10
No smoking	1494 (99.4)	1305 (99.5)	189 (98.4)	
Smokers	9 (0.6)	6 (0.5)	3 (1.6)	
Drinking status, n (%)				.13
No drinking	1412 (93.9)	1227 (93.6)	185 (96.4)	
Drinkers	91 (6.1)	84 (6.4)	7 (3.6)	
Sleep quality, n (%)				<.001
Good	681 (45.3)	637 (48.6)	44 (22.9)	
Average	425 (28.3)	376 (28.7)	49 (25.5)	
Poor	397 (26.4)	298 (22.7)	99 (51.6)	
BMI (kg/m^2), n (%)				.42
Normal weight and under (<23.9)	715 (47.6)	618 (47.1)	97 (50.5)	
Overweight (23.9 - 27.8)	571 (38.0)	498 (38.0)	73 (38.0)	
Obesity (≥27.9)	217 (14.4)	195 (14.9)	22 (11.5)	
Chronic disease history, n (%)				<.001
No	902 (60.0)	812 (61.9)	90 (46.9)	
Yes	601 (40.0)	499 (38.1)	102 (53.1)	
Length of being left behind, n (%)				<.001
Below 3 years	244 (16.2)	195 (14.9)	49 (25.5)	
3 - 6 years	241 (16.0)	206 (15.7)	35 (18.2)	
6 years and above	1,018 (67.7)	910 (69.4)	108 (56.3)	
Resilience scores, median (IQR)	41.0 (37.0-45.0)	41.0 (37.0-45.0)	38.0 (33.0-42.0)	<.001

XSL•FO
RenderX

Characteristic	Overall (n=1503)	No (n=1311)	CDA[a] (n=192)	*P* value[b]
Social support scores, median (IQR)	30.0 (23.0-36.0)	30.0 (25.0-36.0)	20.0 (16.0-26.0)	<.001
VAW[d]				
Any VAW, n (%)				<.001
No	1,033 (68.7)	925 (70.6)	108 (56.3)	
VAW	470 (31.3)	386 (29.4)	84 (43.8)	
VAW number, median (IQR)	0.0 (0.0-1.0)	0.0 (0.0-1.0)	0.0 (0.0-3.0)	<.001
Class of VAW number, n (%)				<.001
0	1,033 (68.7)	925 (70.6)	108 (56.3)	
1	196 (13.0)	177 (13.5)	19 (9.9)	
2	100 (6.7)	84 (6.4)	16 (8.3)	
≥3	174 (11.6)	125 (9.5)	49 (25.5)	
Any IPV[e], n (%)				<.001
No	1,077 (71.7)	962 (73.4)	115 (59.9)	
IPV	426 (28.3)	349 (26.6)	77 (40.1)	
IPV type, n (%)				
Social IPV	88 (5.9)	63 (4.8)	25 (13.0)	<.001
Economic IPV	40 (2.7)	32 (2.4)	8 (4.2)	.17
Emotional IPV	189 (12.6)	148 (11.3)	41 (21.4)	<.001
Physical IPV	301 (20.0)	238 (18.2)	63 (32.8)	<.001
Sexual IPV	32 (2.1)	22 (1.7)	10 (5.2)	.01
Technology IPV	7 (0.5)	4 (0.3)	3 (1.6)	.05
Any remote IPV, n (%)				<.001
No	1422 (94.6)	1253 (95.6)	169 (88.0)	
Remote IPV	81 (5.4)	58 (4.4)	23 (12.0)	
Remote IPV type, n (%)				
Social remote IPV	50 (3.3)	34 (2.6)	16 (8.3)	<.001
Economic remote IPV	23 (1.5)	19 (1.4)	4 (2.1)	.52
Emotional remote IPV	26 (1.7)	20 (1.5)	6 (3.1)	.13
Technology remote IPV	7 (0.5)	4 (0.3)	3 (1.6)	.05
Any NPV[f], n (%)				<.001
No	1398 (93.0)	1234 (94.1)	164 (85.4)	
NPV	105 (7.0)	77 (5.9)	28 (14.6)	
NPV type, n (%)				
Economic NPV	8 (0.5)	5 (0.4)	3 (1.6)	.07
Emotional NPV	89 (5.9)	62 (4.7)	27 (14.1)	<.001
Physical NPV	24 (1.6)	18 (1.4)	6 (3.1)	.11
Sexual NPV	1 (0.1)	1 (0.1)	0 (0.0)	>.99

XSL•FO
RenderX

Characteristic	Overall (n=1503)	No (n=1311)	CDA[a] (n=192)	*P* value[b]
Social NPV	2 (0.1)	2 (0.2)	0 (0.0)	>.99

[a]CDA: comorbid symptoms of depression and anxiety.

[b]*P* values represent statistical measurements of comparing "No" and "CDA" using Wilcoxon Rank Sum test, Pearson chi-square test, or Fisher exact test.

[c]Chinese ¥1=US $0.14.

[d]VAW: violence against women.

[e]IPV: intimate partner violence.

[f]NPV: nonpartner violence.

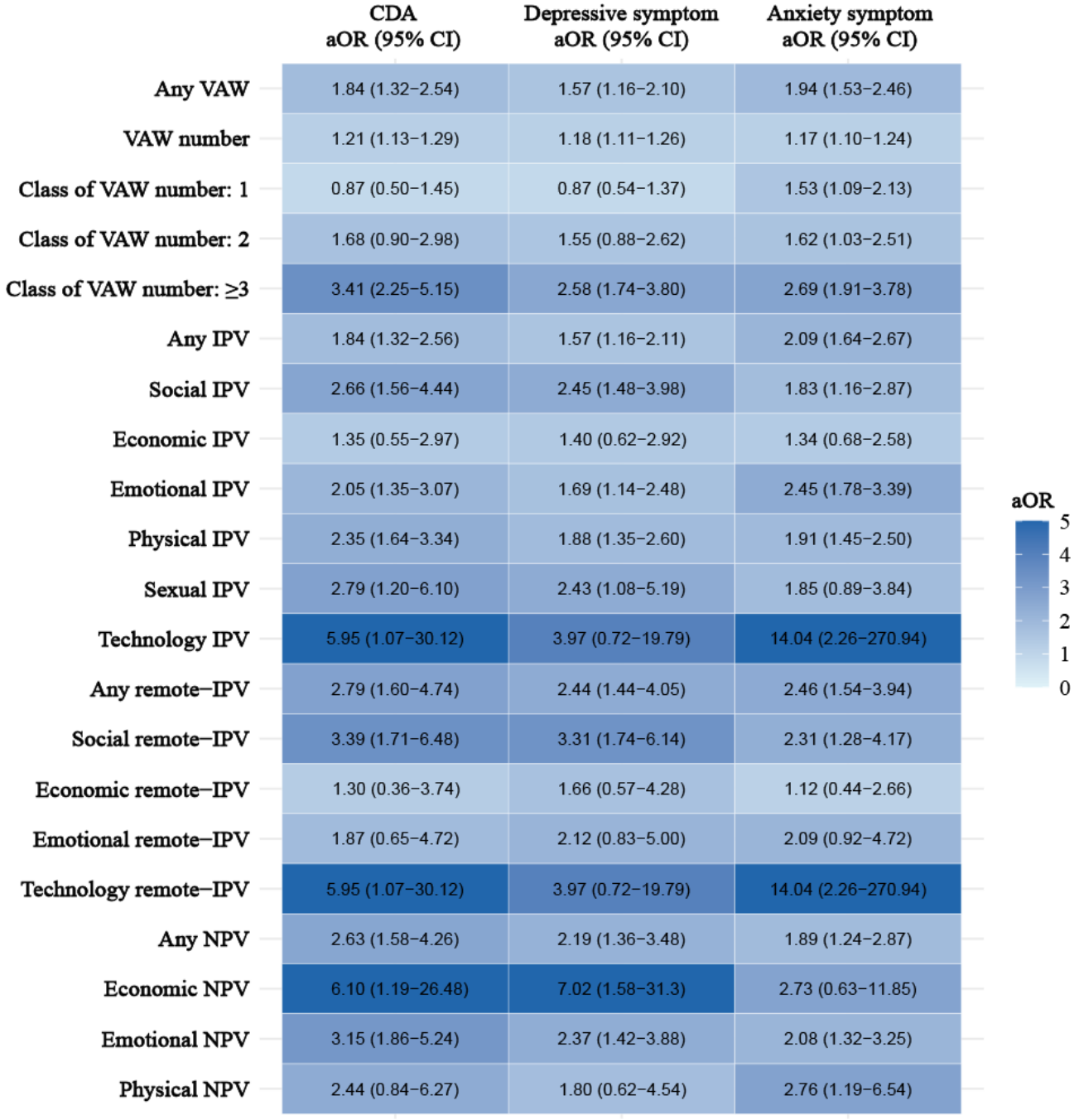
Associations Between VAW and Mental Health Outcomes

In the multivariable logistic models presented in Figure 2 and Table S2 in Multimedia Appendix 1, compared with no VAW, VAW exposure was associated with an increased CDA risk (adjusted odds ratio [aOR] 1.84, 95% CI 1.32 - 2.54), and this risk rose with each additional VAW item (aOR 1.21, 95% CI 1.13 - 1.29). When VAW was categorized by item count, women with 3 or more VAW items showed a markedly higher risk of CDA (aOR 3.41, 95% CI 2.25 - 5.15). Similar associations were also observed for depressive and anxiety symptoms in relation to VAW (Figure 2 and Table S2 in Multimedia Appendix 1). Restricted cubic spline curves exhibited dose-response associations between increased VAW numbers and elevated risks of mental health outcomes (Figure S2 in Multimedia Appendix 1).

For the 3 forms of VAW, lifetime IPV (aOR 1.84, 95% CI 1.32 - 2.56), remote IPV (aOR 2.79, 95% CI 1.60 - 4.74), and NPV (aOR 2.63, 95% CI 1.58 - 4.26) were all associated with higher odds of CDA. Among the types of IPV, only economic IPV did not show a significant association with mental health outcomes. Social remote IPV (aOR 3.39, 95% CI 1.71 - 6.48), technology-facilitated remote IPV (aOR 5.95, 95% CI 1.07 - 30.12), economic NPV (aOR 6.10, 95% CI 1.19 - 26.48), and emotional NPV (aOR 3.15, 95% CI 1.86 - 5.24) were all associated with higher odds of CDA. Similar findings were observed for depressive symptoms and anxiety symptoms, respectively (Figure 2 and Table S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 2. Heat maps of aOR from multivariable logistic regression analyses examining associations between various forms and types of violence and mental health outcomes among Chinese rural left-behind women, based on a cross-sectional study conducted in China in 2023. A darker color in a gradient from white to blue shows a higher aOR of corresponding outcomes. All models were adjusted for age, education, occupation, household income, unhealthy lifestyle (smoking, drinking, poor sleep quality), BMI, chronic disease history, and length of being left behind. aOR: adjusted odds ratio; CDA: comorbid symptoms of depression and anxiety; IPV: intimate partner violence; NPV: nonpartner violence; VAW: violence against women.



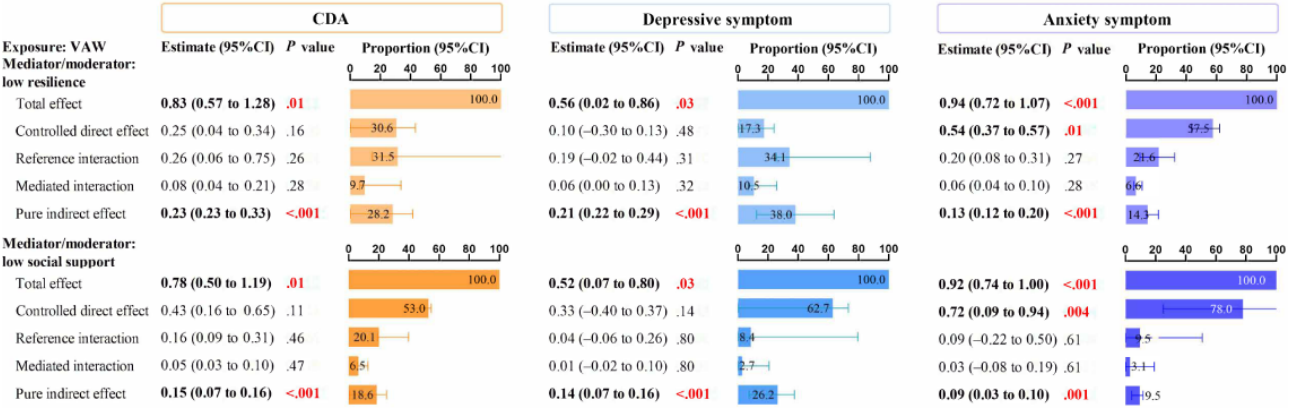
Mediation and Interaction Effects

In a 4-way decomposition analysis, all mediation effects were statistically significant (P<.05), whereas none of the interactions reached significance (P>.05). Overall, 69.4% of the association between VAW and CDA could be explained by low resilience, with 37.9% of this effect attributable to total mediation. Similarly, 45.2% of the association between VAW and CDA could be explained by low social support, with 25.1% attributable to total mediation (Figure 3 and Table S3 in Multimedia Appendix 1).

Separate mediation and interaction analyses validated similar conclusions. Mediation analysis also showed that low resilience and low social support partially mediated the associations between VAW and mental health outcomes (Table S4 in Multimedia Appendix 1). While the additive and multiplicative interactions of "VAW*resilience" and "VAW*social support" on mental health outcomes were not significant, as the 95% CI of relative excess risk due to interaction crossed 0 and the 95% CI of multiplicative scales crossed 1 (Table S5 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 3. Four-way decomposition analysis of the associations between violence and mental health outcomes with low resilience or social support as mediators among Chinese rural left-behind women, based on a cross-sectional study conducted in China in 2023. Controlled direct effect indicates association not caused by mediation or interaction, reference interaction indicates association due to interaction only, mediated interaction indicates association due to mediation and interaction, and pure indirect effect indicates association due to mediation only. The total mediation proportion comprises the combined effects of the mediated interaction and the pure indirect effect. Similarly, the total interaction proportion is determined by the sum of the reference interaction and the mediated interaction. All models were adjusted for age, education, occupation, household income, unhealthy lifestyle (smoking, drinking, and poor sleep quality), BMI, chronic disease history, and length of being left behind. CDA: comorbid symptoms of depression and anxiety; VAW: violence against women.

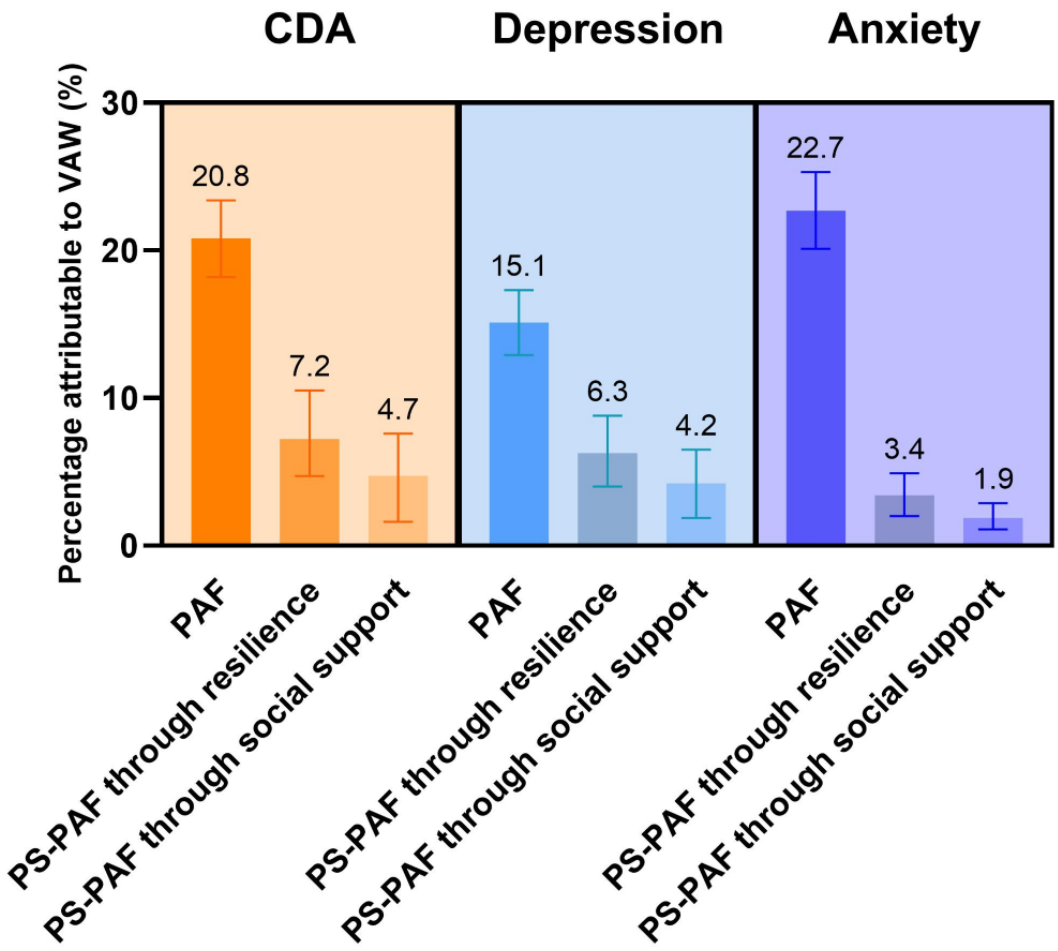


Population Attributable Fractions

A total of 20.8% of CDA cases, 15.1% of depressive symptoms cases, and 22.7% of anxiety symptoms cases were attributable to VAW (Figure 4). Among the 3 forms of VAW, the PAF of IPV on depressive symptoms (13.9%) is lower than on anxiety symptoms (23.6%), while the PAF of NPV on depressive symptoms (7.7%) is higher than on anxiety symptoms (5.9%) (Table S6 in Multimedia Appendix 1). The pathway-specific PAFs of CDA, depressive symptoms, and anxiety symptoms attributable to VAW through resilience or social support are presented (Figure 4 and Table S6 in Multimedia Appendix 1). As a mediator, resilience accounted for 7.2%, 6.3%, and 3.4% of CDA, depressive symptoms, and anxiety symptoms with VAW, respectively. The pathway-specific PAFs of social support between VAW and CDA, depressive symptoms, and anxiety symptoms were 4.7%, 4.2%, and 1.9%, respectively.

XSL•FO
RenderX

Figure 4. PAF and PS-PAF for violence on mental health outcomes for Chinese rural left-behind women, based on a cross-sectional study conducted in China in 2023. CDA: comorbid symptoms of depression and anxiety; PAF: population attributable fraction; PS-PAF: pathway-specific population attributable fraction; VAW: violence against women.



Discussion

Principal Findings

Chinese rural left-behind women's exposure to lifetime VAW, including IPV (and remote IPV) and NPV, is strongly associated with adverse mental health outcomes, including CDA, depressive, and anxiety symptoms. These associations were partially mediated by low resilience and social support, though interactions were nonsignificant. VAW accounted for a substantial proportion of CDA, depressive, and anxiety symptoms, with differing attributable fractions for IPV and NPV, and resilience and social support partially mediated these attributable proportions.

This study aligns with previous research, reinforcing the association between lifetime VAW and adverse mental health outcomes, including its cumulative effects [44,45]. The left-behind situation among women in rural areas has lasted several decades in China due to disproportionate male labor migration to advanced economic regions. In such a context, women are particularly vulnerable without support from their husbands. Stereotypical gender norms and prevailing beliefs that family matters are private in rural areas constrain women's autonomy and perpetuate violence [46,47]. Moreover, this vulnerability can be further exacerbated by women's economic dependence and limited mental health services resources, further deepening their mental health issues [48]. These associations between VAW and mental health issues can also be supported by biological evidence. Repeated violence exposure may lead to dysregulation of the hypothalamic-pituitary-adrenal axis, a critical pathway in the body's stress response system, contributing to long-term adverse health effects [49]. Moreover, violence exposure can alter neural circuits involved in emotional learning, resulting in impaired memory, heightened threat vigilance, and exaggerated emotional responses to environmental stimuli [50].

The increased risk of adverse mental health outcomes for women exposed to lifetime IPV and NPV is well established and extended by our findings [20,51,52]. Remote IPV has also been found to be an important factor affecting left-behind women's mental health, calling attention to technology-facilitated remote IPV [16,53]. All IPV types, except economic IPV, were significantly associated with mental health outcomes. The lack of a significant association for economic IPV might be due to the socioeconomic context of rural left-behind women, many of whom are in lower economic strata and depend primarily on

XSL•FO
RenderX

their husbands for financial support. Additionally, social and technology-facilitated remote IPV, along with economic and emotional NPV, was associated with higher CDA odds. Our findings expand previous research by comprehensively assessing a broader range of VAW forms, types, and their associations with mental health outcomes.

The associations between VAW and mental health outcomes were partially mediated by low resilience and social support, aligning with previous research. Resilience is thought to buffer the psychological effects of violence, given its ability to access various resources to adapt positively in the face of adversity [54,55]. A study found that empowering women to develop resilience-based coping strategies significantly mediated the impact of IPV on mental health outcomes [56]. Social support is also a vital source of empowerment for women provided by a network of relationships, which helps them cope with stress and maintain well-being. Consistent with our findings, research shows that higher social support buffers the negative impacts of adverse events, helping individuals maintain physical and mental well-being [57,58]. This highlights the importance of strengthening social support networks and fostering resilience as potential interventions to reduce the psychological burden of VAW among rural left-behind women.

Our study estimated that 20.8% of CDA cases, 15.1% of depressive symptoms cases, and 22.7% of anxiety symptoms cases were attributable to VAW. Depressive symptoms were predominantly attributed to IPV, while anxiety symptoms were mainly attributed to NPV. These findings underscore the significant mental health burden associated with VAW, emphasizing the need for clinicians, psychotherapists, and social workers to recognize it as a critical risk factor in rural left-behind women with depression or anxiety. Additionally, pathway-specific PAF revealed that psychological resilience accounted for a greater fraction of mental health outcomes attributable to VAW than social support. This suggests that enhancing resilience may be more effective in reducing the mental health impact of VAW than solely focusing on social support. These findings underscore the importance of incorporating both resilience and social support into public health interventions to address the mental health consequences of VAW.

Implications for Policy

This study highlights the urgent need for targeted strategies to reduce VAW, foster resilience, and strengthen social support networks through collaborative efforts from communities, health services, and local governments. Such strategies can help to improve mental health outcomes for rural left-behind women in China. For example, implementing educational campaigns to raise prevention awareness of VAW, providing psychological skills training to build resilience, and establishing women's support groups to strengthen social networks might be practical and effective community-based strategies. Besides, we call for early detection of various forms of VAW and routine mental health screening in both community and health care settings, which can facilitate timely identification and support, such as psychological counseling. Moreover, consistently implementing poverty alleviation and economic empowerment for women living in rural areas in China can further enhance women's resilience.

Future Research

Future research should explore prevention and intervention strategies for addressing VAW among left-behind women in rural China. Additionally, intervention studies evaluating the effectiveness of such strategies and policies would provide actionable guidance for improving the mental health of women experiencing violence.

Limitations

This study also has several limitations. First, the focus on rural left-behind women may limit the generalizability. Second, the use of self-reported data for VAW exposure and covariates introduces potential recall bias, which may underestimate the prevalence of VAW. Third, as a cross-sectional study, this research cannot establish causation, though consistent and cumulative associations suggest potential causal pathways. Finally, the available covariates could not fully capture all relevant factors, such as variables reflecting body fat composition and physical activity, which were not included.

Conclusions

Our study demonstrates that rural left-behind women who experienced lifetime VAW, including IPV (and remote IPV) and NPV, had significantly higher risks of CDA, depressive and anxiety symptoms, with low resilience and inadequate social support partially mediating these associations. Targeted strategies, including reducing VAW, fostering resilience, and strengthening social support networks, are urgently needed to improve mental health outcomes among rural left-behind women. Future research on the development of VAW prevention and intervention strategies, along with mental health support services, could aid in efforts to mitigate the prevalence of VAW and mental health issues for women in rural areas.

Acknowledgments

We appreciate the substantial time and effort from the participants who contributed to this study. We thank all project staff for their sincere contribution to data collection. We thank the Sexual Violence Research Initiative for funding and providing approval of the Knowledge for Action to End Violence Against Women and Violence Against Children Program, through a key project titled "Uncovering adverse childhood experiences and violence against left-behind women in China: life course pathways and priority setting." The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are not publicly available due to restrictions on the confidential information but are available from the corresponding author on reasonable request.

Authors' Contributions

PS and WL were involved in the study design. WL coordinated the data collection. LH managed and analyzed the data and prepared the first draft. LH reviewed and edited the manuscript, with comments from ZL, WS, JY, JW, SS, WL, and PS. LH had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. All authors were involved in revising the paper and gave final approval of the submitted versions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables and figures.
[DOCX File, 363 KB - publichealth_v11i1e72064_app1.docx]

References

1. Hugo G. Migration and Women's Empowerment Women's Empowerment and Demographic Processes: Oxford University Press; 2000:287-317.
2. Yi J, Zhong B, Yao S. Health-related quality of life and influencing factors among rural left-behind wives in Liuyang, China. BMC Womens Health 2014 May 14;14:67. [doi: 10.1186/1472-6874-14-67] [Medline: 24886024]
3. Jin Y, Qin Q, Zhan S, Yu X, Liang L, Huang F. Depressive symptoms were prevalent among left-behind women in Ma'anshan, China. J Nerv Ment Dis 2016 Mar;204(3):226-232. [doi: 10.1097/NMD.0000000000000444] [Medline: 26694074]
4. Huang H, Liu S, Cui X, Zhang J, Wu H. Factors associated with quality of life among married women in rural China: a cross-sectional study. Qual Life Res 2018 Dec;27(12):3255-3263. [doi: 10.1007/s11136-018-1944-y] [Medline: 30032339]
5. Ye J, Wu H, Rao J, Ding B, Zhang K. Left-behind women: gender exclusion and inequality in rural-urban migration in China †. J Peasant Stud 2016 Jul 3;43(4):910-941. [doi: 10.1080/03066150.2016.1157584]
6. Hu J, Khan MS. Modeling the factors influencing the quality of life of rural left-behind women in Guangxi, China. RGSA 2024;18(2):e05714. [doi: 10.24857/rgsa.v18n2-149]
7. Chaudhari KS, Dhapkas MP, Kumar A, Ingle RG. Mental disorders – a serious global concern that needs to address. IJPQA 2024;15(2):973-978. [doi: 10.25258/ijpqa.15.2.66]
8. Cummings CM, Caporino NE, Kendall PC. Comorbidity of anxiety and depression in children and adolescents: 20 years after. Psychol Bull 2014 May;140(3):816-845. [doi: 10.1037/a0034733] [Medline: 24219155]
9. Felez-Nobrega M, Haro JM, Koyanagi A. Multimorbidity, depression with anxiety symptoms, and decrements in health in 47 low- and middle-income countries. J Affect Disord 2022 Nov;317:176-184. [doi: 10.1016/j.jad.2022.08.110]
10. Morgan A, Boxall H. Social isolation, time spent at home, financial stress and domestic violence during the COVID-19 pandemic. Trends and Issues in Crime and Criminal Justice 2020:609. [doi: 10.52922/ti04855]
11. Wang J, Lloyd-Evans B, Giacco D, et al. Social isolation in mental health: a conceptual and methodological review. Soc Psychiatry Psychiatr Epidemiol 2017 Dec;52(12):1451-1461. [doi: 10.1007/s00127-017-1446-1] [Medline: 29080941]
12. Liang D, Mays VM, Hwang WC. Integrated mental health services in China: challenges and planning for the future. Health Policy Plan 2018 Jan 1;33(1):107-122. [doi: 10.1093/heapol/czx137] [Medline: 29040516]
13. Declaration on the elimination of violence against women. : United Nations; 1993.
14. Oram S, Fisher HL, Minnis H, et al. The Lancet Psychiatry Commission on Intimate Partner Violence and mental health: advancing mental health services, research, and policy. Lancet Psychiatry 2022 Jun;9(6):487-524. [doi: 10.1016/S2215-0366(22)00008-6] [Medline: 35569504]
15. Bansal V, Rezwan M, Iyer M, et al. A scoping review of technology-facilitated gender-based violence in low- and middle-income countries across Asia. Trauma Violence Abuse 2024 Jan;25(1):463-475. [doi: 10.1177/15248380231154614] [Medline: 36869803]
16. Rogers MM, Fisher C, Ali P, Allmark P, Fontes L. Technology-facilitated abuse in intimate relationships: a scoping review. Trauma Violence Abuse 2023 Oct;24(4):2210-2226. [doi: 10.1177/15248380221090218] [Medline: 35537445]
17. Abrahams N, Devries K, Watts C, et al. Worldwide prevalence of non-partner sexual violence: a systematic review. Lancet 2014 May 10;383(9929):1648-1654. [doi: 10.1016/S0140-6736(13)62243-6] [Medline: 24529867]
18. John NA. Exploring the linkages between women's paid and unpaid work and their experiences of intimate partner and non-partner violence in Nepal. Fem Econ 2020 Oct 1;26(4):89-113. [doi: 10.1080/13545701.2020.1828601]

XSL•FO
RenderX

19. Ellsberg M, Jansen H, Heise L, Watts CH, Garcia-Moreno C. Intimate partner violence and women's physical and mental health in the WHO multi-country study on women's health and domestic violence: an observational study. Lancet 2008 Apr 5;371(9619):1165-1172. [doi: 10.1016/S0140-6736(08)60522-X] [Medline: 18395577]
20. White SJ, Sin J, Sweeney A, et al. Global prevalence and mental health outcomes of intimate partner violence among women: a systematic review and meta-analysis. Trauma Violence Abuse 2024 Jan;25(1):494-511. [doi: 10.1177/15248380231155529] [Medline: 36825800]
21. Mohammadi AQ, Neyazi A, Habibi D, Mehmood Q, Neyazi M, Griffiths MD. Female education ban by the Taliban: a descriptive survey study on suicidal ideation, mental health, and health-related quality of life among girls in Afghanistan. J Public Health (Oxf) 2024 Aug 25;46(3):e439-e447. [doi: 10.1093/pubmed/fdae111] [Medline: 38936836]
22. Neyazi A, Padhi BK, Mohammadi AQ, et al. Depression, anxiety and quality of life of Afghan women living in urban areas under the Taliban government: a cross-sectional study. BMJ Open 2023 Aug 3;13(8):e071939. [doi: 10.1136/bmjopen-2023-071939] [Medline: 37536972]
23. Neyazi A, Mohammadi AQ, Razaqi N, Satapathy P, Mehmood Q, Neyazi M. Correlation of violence with anxiety and sleep disturbance among Talibs (mosque students) in Afghanistan: a cross-sectional study. Sleep Epidemiology 2024 Dec;4:100075. [doi: 10.1016/j.sleepe.2024.100075]
24. Connor KM, Davidson JRT. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). Depress Anxiety 2003;18(2):76-82. [doi: 10.1002/da.10113] [Medline: 12964174]
25. Acoba EF. Social support and mental health: the mediating role of perceived stress. Front Psychol 2024;15:1330720. [doi: 10.3389/fpsyg.2024.1330720] [Medline: 38449744]
26. Ellsberg M, Heise L, Peña R, Agurto S, Winkvist A. Researching domestic violence against women: methodological and ethical considerations. Stud Fam Plann 2001 Mar;32(1):1-16. [doi: 10.1111/j.1728-4465.2001.00001.x] [Medline: 11326453]
27. Schraiber LB, Latorre MDR, França I Jr, Segri NJ, D'Oliveira A. Validity of the WHO VAW study instrument for estimating gender-based violence against women. Rev Saude Publica 2010 Aug;44(4):658-666. [doi: 10.1590/s0034-89102010000400009] [Medline: 20676557]
28. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (center for epidemiologic studies depression scale). Am J Prev Med 1994;10(2):77-84. [Medline: 8037935]
29. Boey KW. Cross-validation of a short form of the CES-D in Chinese elderly. Int J Geriatr Psychiatry 1999 Aug;14(8):608-617. [doi: 10.1002/(sici)1099-1166(199908)14:8<608::aid-gps991>3.0.co;2-z] [Medline: 10489651]
30. Zhou L, Ma X, Wang W. Relationship between cognitive performance and depressive symptoms in Chinese older adults: the China health and retirement longitudinal study (CHARLS). J Affect Disord 2021 Feb 15;281:454-458. [doi: 10.1016/j.jad.2020.12.059] [Medline: 33360747]
31. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
32. Sun J, Liang K, Chi X, Chen S. Psychometric properties of the generalized anxiety disorder scale-7 Item (GAD-7) in a large sample of Chinese adolescents. Healthcare (Basel) 2021 Dec 9;9(12):1709. [doi: 10.3390/healthcare9121709] [Medline: 34946435]
33. He X, Li C, Qian J, Cui H, Wu W. Reliability and validity of a generalized anxiety scale in general hospital outpatients. Shanghai Arch Psychiatry 2010;22(4):200-203. [doi: 10.3969/j.issn.1002-0829.2010.04.002]
34. Wang S, Zheng X, Huang J, Liu J, Li C, Shang H. Sleep characteristics and risk of Alzheimer's disease: a systematic review and meta-analysis of longitudinal studies. J Neurol 2024 Jul;271(7):3782-3793. [doi: 10.1007/s00415-024-12380-7] [Medline: 38656621]
35. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the Connor-davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. J Trauma Stress 2007 Dec;20(6):1019-1028. [doi: 10.1002/jts.20271] [Medline: 18157881]
36. Wang L, Shi Z, Zhang Y, Zhang Z. Psychometric properties of the 10-item Connor-Davidson Resilience Scale in Chinese earthquake victims. Psychiatry Clin Neurosci 2010 Oct;64(5):499-504. [doi: 10.1111/j.1440-1819.2010.02130.x] [Medline: 20923429]
37. Peng L, Zhang J, Li M, et al. Negative life events and mental health of Chinese medical students: the effect of resilience, personality and social support. Psychiatry Res 2012 Mar 30;196(1):138-141. [doi: 10.1016/j.psychres.2011.12.006] [Medline: 22405636]
38. Flegal KM, Kit BK, Graubard BI. Body mass index categories in observational studies of weight and risk of death. Am J Epidemiol 2014 Aug 1;180(3):288-296. [doi: 10.1093/aje/kwu111] [Medline: 24893710]
39. VanderWeele TJ. A unification of mediation and interaction: a 4-way decomposition. Epidemiology (Sunnyvale) 2014 Sep;25(5):749-761. [doi: 10.1097/EDE.0000000000000121] [Medline: 25000145]
40. Khosravi A, Nazemipour M, Shinozaki T, Mansournia MA. Population attributable fraction in textbooks: time to revise. Glob Epidemiol 2021 Nov;3:100062. [doi: 10.1016/j.gloepi.2021.100062] [Medline: 37635714]
41. Efron B, Tibshirani RJ. An Introduction to the Bootstrap: Chapman and Hall/CRC; 1994:1-456. [doi: 10.1201/9780429246593]

XSL•FO
RenderX

42. O'Connell MM, Ferguson JP. Pathway-specific population attributable fractions. Int J Epidemiol 2022 Dec 13;51(6):1957-1969. [doi: 10.1093/ije/dyac079] [Medline: 35536313]
43. Ferguson J, O'Connell M. Estimating and displaying population attributable fractions using the R package: graphPAF. Eur J Epidemiol 2024 Jul;39(7):715-742. [doi: 10.1007/s10654-024-01129-1] [Medline: 38971917]
44. Mootz JJ, Fennig M, Giusto A, Mumey A, Greene CM, Wainberg ML. Interventions addressing family violence and mental illness or substance use in low- and middle-income countries: a systematic review. Glob Ment Health (Camb) 2023;10:e71. [doi: 10.1017/gmh.2023.62] [Medline: 38024805]
45. Rees S, Silove D, Chey T, et al. Lifetime prevalence of gender-based violence in women and the relationship with mental disorders and psychosocial function. JAMA 2011 Aug 3;306(5):513-521. [doi: 10.1001/jama.2011.1098] [Medline: 21813429]
46. Lei L. The patriarchal trap: the village wives left behind amid Chinese rural–urban migration. Sage Open 2023 Jan;13(1). [doi: 10.1177/21582440221144245]
47. Wang C, Tang J. Ritualized law and livelihood fragility of left-behind women in rural China. Int J Environ Res Public Health 2020 Jun 17;17(12):4323. [doi: 10.3390/ijerph17124323] [Medline: 32560353]
48. Wu H, Ye J. Hollow lives: women left behind in rural China. J Agrar Change 2016;16(1):50-69. [doi: 10.1111/joac.12089]
49. Neigh GN, Gillespie CF, Nemeroff CB. The neurobiological toll of child abuse and neglect. Trauma Violence Abuse 2009 Oct;10(4):389-410. [doi: 10.1177/1524838009339758] [Medline: 19661133]
50. McLaughlin KA, Sheridan MA, Lambert HK. Childhood adversity and neural development: deprivation and threat as distinct dimensions of early experience. Neurosci Biobehav Rev 2014 Nov;47:578-591. [doi: 10.1016/j.neubiorev.2014.10.012] [Medline: 25454359]
51. Dokkedahl SB, Kirubakaran R, Bech-Hansen D, Kristensen TR, Elklit A. The psychological subtype of intimate partner violence and its effect on mental health: a systematic review with meta-analyses. Syst Rev 2022 Aug 10;11(1):163. [doi: 10.1186/s13643-022-02025-z] [Medline: 35948921]
52. Devries K, Watts C, Yoshihama M, et al. Violence against women is strongly associated with suicide attempts: evidence from the WHO multi-country study on women's health and domestic violence against women. Soc Sci Med 2011 Jul;73(1):79-86. [doi: 10.1016/j.socscimed.2011.05.006] [Medline: 21676510]
53. Patel U, Roesch R. The prevalence of technology-facilitated sexual violence: a meta-analysis and systematic review. Trauma Violence Abuse 2022 Apr;23(2):428-443. [doi: 10.1177/1524838020958057] [Medline: 32930064]
54. Ungar M. Resilience, trauma, context, and culture. Trauma Violence Abuse 2013 Jul;14(3):255-266. [doi: 10.1177/1524838013487805] [Medline: 23645297]
55. Grych J, Hamby S, Banyard V. The resilience portfolio model: understanding healthy adaptation in victims of violence. Psychol Violence 2015;5(4):343-354. [doi: 10.1037/a0039671]
56. Mengo C, Small E, Black B. Intimate partner violence and women's mental health: the mediating role of coping strategies among women seeking help from the police. J Interpers Violence 2021 Jan;36(1-2):527-551. [doi: 10.1177/0886260517729402] [Medline: 29294903]
57. Hou F, Cerulli C, Wittink MN, Caine ED, Qiu P. Using confirmatory factor analysis to explore associated factors of intimate partner violence in a sample of Chinese rural women: a cross-sectional study. BMJ Open 2018 Feb 2;8(2):e019465. [doi: 10.1136/bmjopen-2017-019465] [Medline: 29420233]
58. Coker AL, Smith PH, Thompson MP, McKeown RE, Bethea L, Davis KE. Social support protects against the negative effects of partner violence on mental health. J Womens Health Gend Based Med 2002 Jun;11(5):465-476. [doi: 10.1089/15246090260137644] [Medline: 12165164]

Abbreviations

aOR: adjusted odds ratio
CDA: comorbid symptoms of depression and anxiety
IPV: intimate partner violence
NPV: nonpartner violence
OR: odds ratio
PAF: population attributable fraction
RR: risk ratio
VAW: violence against women

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 03.02.25; peer-reviewed by A Neyazi, E Fedele; revised version received 16.06.25; accepted 18.06.25; published 07.08.25.

Please cite as:
Hou L, Luo Z, Sun W, Ying J, Wu J, Shan S, Liu W, Song P
Associations of Violence Against Women With Comorbid Symptoms of Depression and Anxiety Among Left-Behind Women in Rural China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e72064
URL: https://publichealth.jmir.org/2025/1/e72064
doi:10.2196/72064

© Leying Hou, Zeyu Luo, Weidi Sun, Jiayao Ying, Jing Wu, Shiyi Shan, Wen Liu, Peige Song. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 7.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Contribution of Travelers to Plasmodium Vivax Malaria in South West Delhi, India: Cross-Sectional Survey

Deepali Savargaonkar[1], MBBS, MPH; Bina Srivastava[1], MSc; Chander Prakash Yadav[2], PhD; Mrigendra Pal Singh[3], PhD; Anup Anvikar[1,4], MBBS, MD; Amit Sharma[5], PhD; Himmat Singh[1,4*], PhD; Abhinav Sinha[1,4*], MBBS, MRES, MSc, MD, PhD

[1]ICMR-National Institute of Malaria Research, Sector 8, Dwarka, New Delhi, India
[2]ICMR-National Institute of Cancer Prevention and Research, Noida, India
[3]ICMR-National Institute for Research in Tribal Health, Jabalpur, India
[4]Academy of Scientific and Innovative Research, Ghaziabad, India
[5]International Centre for Genetic Engineering and Biotechnology, New Delhi, India
[*]these authors contributed equally

Corresponding Author:
Abhinav Sinha, MBBS, MRES, MSc, MD, PhD
ICMR-National Institute of Malaria Research, Sector 8, Dwarka, New Delhi, India

Abstract

Background: India is committed to malaria elimination by the year 2030. According to the classification of malaria endemicity, the National Capital Territory of Delhi falls under category 1, with an annual parasite incidence of <1, and was targeted for elimination by 2022. Among others, population movement across states is one of the key challenges for malaria control, as it can result in imported malaria, thus introducing local transmission in an area nearing elimination.

Objective: This descriptive study attempts to assess the contribution of such imported *Plasmodium vivax* cases to the malaria burden in South West Delhi (SWD).

Methods: A cross-sectional study was carried out at the fever clinic of the Indian Council of Medical Research-National Institute of Malaria Research in SWD from January 2017 to December 2019. Demographic and travel history data were recorded for all *P vivax* confirmed malaria cases diagnosed at the fever clinic. Vector and fever surveys along with reactive case detection were conducted in SWD and Bulandshahr district of Uttar Pradesh, 1 of the 6 geographical sources for a high number of imported malaria cases.

Results: A total of 355 *P vivax* malaria cases were reported during the study period. The proportion of imported cases was 63% (n=222). Of these, 96% (n=213) of cases were from Uttar Pradesh. The distribution of malaria cases revealed that imported cases were significantly associated with travel during the transmission season compared with that in the nontransmission season. Entomological and fever surveys and reactive case detection carried out in areas visited by imported *P vivax* malaria cases showed the presence of adults and larvae of *Anopheles* species and *P vivax* parasitemia.

Conclusions: Population movement is a key challenge for malaria elimination. Although additional *P vivax* infections and vector mosquitoes were detected at places visited by the imported malaria cases, the inability to detect the parasite in mosquitoes and the possibility of relapses associated with *P vivax* limit the significance of malaria associated with the travel. However, there remains a need to address migration malaria to prevent the introduction and re-establishment of malaria in areas with very low or 0 indigenous cases.

(JMIR Public Health Surveill 2025;11:e50058) doi:10.2196/50058

KEYWORDS

malaria; Plasmodium vivax; imported malaria; population movement; transmission; elimination; India

Introduction

Malaria is a parasitic disease transmitted by the bites of infected female *Anopheles* mosquitoes. The estimated number of malaria cases worldwide in 2021 was 247 million spread over 84 countries with 619,000 deaths. Although the World Health Organization (WHO) South East Asia region comprised ≈2% of the estimated global burden, India contributed to ≈79% of these cases, with a preponderance of *Plasmodium vivax* over *Plasmodium falciparum* [1].

Migration, both international and within-country, is a recognized social health determinant of multiple diseases across the globe, and malaria is no exception. Various factors put the migrating

XSL•FO
RenderX

population at risk of contracting malaria, and these include their socioeconomic, living, working, and transit conditions [2]. The risk of malaria is also to the host communities that provide shelter to the migrants, particularly when the migration is along an epidemiological gradient from a high-burden to a low-burden or nonendemic area, putting malaria elimination efforts at risk [3-5]. Several pieces of evidence of this have been documented in the context of international migration [6-8] and within-country migration [9,10], including that in India [11].

India is committed to malaria elimination by 2030 and has formulated the National Framework for Malaria Elimination that classifies Indian states and union territories into 4 categories from 0 to 3, with category 3 being the highest-burden areas with an annual parasite incidence (API) of ≥1 per thousand persons at risk [12]. To achieve the elimination goal in the desired time frame, special focus needs to be given to the identified challenges by the National Center for Vector Borne Diseases Control (NCVBDC). Population size and migration are recognized as important challenges for malaria elimination, apart from asymptomatic parasite reservoirs, low-density infections, and parasite- and vector-resistance [13-15]. The movement of populations across and within Indian states is one of the key challenges in malaria control [12], and particularly, the migration of workers in large numbers from rural areas to cities has been reported in India [16]. Similarly, malaria cases among the mobile population contribute to a high percentage of total malaria cases in many countries [17] and have been a recognized challenge for malaria elimination [18]. Hence, it is important to carry out the screening and treatment for malaria in mobile populations for control and elimination of malaria in endemic areas, especially National Framework for Malaria Elimination category 1 areas (with API<1), and for prevention of the re-establishment of local transmission of malaria [12] in areas that have eliminated malaria (category 0 with 0 indigenous cases).

Despite being the capital of India that attracted >100,000 migrants each year since 2012 [19], no study on malaria in the migrant population in Delhi has yet been reported. Delhi falls under category 1 (API<1 per 1000 persons at risk), and its neighboring state, Uttar Pradesh (UP), is in category 2 with an API of less than 1 but with some districts having an API of 1 or more. With such a magnitude of migration, there remains a sustained threat of the introduction of *Plasmodium* infection by infected migrant populations from high-burden areas to the areas in categories 0 and 1 [16].

As Delhi and its neighboring states have *P vivax* as the predominant *Plasmodium* species causing malaria, this study was therefore carried out to assess the contribution of imported *P vivax* cases to the *P vivax* malaria burden in South West Delhi (SWD) by tracking the travel history of infected patients diagnosed at the fever clinic of Indian Council of Medical Research-National Institute of Malaria Research (ICMR-NIMR). Additionally, the study also aimed to identify mosquito breeding habitats and the presence and types of mosquito species in areas where these imported malaria cases resided/visited, and to detect additional *Plasmodium* infections/malaria cases through reactive case detection (RACD) and fever surveys in selected areas visited by these patients. These entomological and fever surveys and RACD were done to further identify any ongoing transmission in the areas visited by patients with imported malaria to support the hypothesis that these are indeed the cases likely to import malaria from areas with ongoing transmission to SWD.

Methods

Ethical Considerations

The study was approved by the Institutional Ethics Committee of ICMR-NIMR (ECR/NIMR/EC/2015/507 and ECR/NIMR/EC/2019/175). Informed consent was obtained from all human participants who were involved in the study. The participant identifier data were anonymized. No compensation was provided to the participants.

Study Sites, Samples, and Definitions

The study was initiated at the fever clinic at ICMR-NIMR, SWD, and later expanded to include the prominent catchment areas of the fever clinic of SWD and 6 villages of 1 selected district (Bulandshahr) of UP. Incoming febrile patients were screened for *Plasmodium* infection at the fever clinic of ICMR-NIMR from January 2017 to December 2019. Basic demographic data including age, gender, history of fever, and travel details (if any) were recorded using a paper-based structured questionnaire. The parasitological diagnosis at the clinic was performed by microscopy. Thick and thin blood smears were prepared, stained using the Jaswant Singh–Bhattacharji stain [20] and examined under 100× magnification. Microscopy was performed independently by 2 trained microscopists. In case of a discrepancy, a third trained microscopist examined the smears, and consensus observation by 2 trained microscopists was considered final. Those diagnosed with malaria were treated as per the national drug policy [21].

History of travel during the preceding 30 days of fever, including the places visited and duration of stay, was verbally elicited through a calendar-based recall method among all patients with malaria. Obtained travel history was used to classify the patients into 3 nonoverlapping categories: no travel history (patients neither traveled in or out of their residence in SWD), returning travelers (patients who were residing in SWD and traveled outside Delhi but came back), and incoming travelers (patients who were not residing in SWD but transiently traveled to SWD). Returning travelers were further classified based on the duration of stay outside Delhi into those returning to SWD within 7 days and those returning between 7 and 30 days. Similarly, incoming travelers were also classified into those who came to SWD within and beyond 30 days of fever.

Since the incubation period for malaria is 7 to 30 days [22-25], returning travelers with malaria who returned to SWD between 7 and 30 days of fever onset and incoming travelers who entered SWD within 30 days of fever onset were classified as imported malaria cases (acquired infection outside Delhi) for this study. The rest of the patients with malaria were considered to have indigenous infections.

Imported malaria cases were further investigated for the exact village and district of travel based on their recorded travel

XSL•FO
RenderX

history. Thus, to further investigate whether the malaria cases were imported or indigenous, malaria vector (anopheline) surveys were carried out both in SWD and in the villages the imported malaria cases traveled to. These surveys were carried out by field workers adequately trained in entomology in catchment areas (Raj Nagar and Bagdola; every month from September 2018 to December 2019) of the fever clinic of ICMR-NIMR predominantly reporting malaria and also in villages (once in October 2019) of Bulandshahr district of UP state. Fever surveys and RACD were also carried out in these villages of Bulandshahr, UP. Bulandshahr district was preferred out of the 6 districts that showed significant sources of imported malaria in UP based on the burden of imported cases, logistic convenience, and operational feasibility. All 6 villages of Bulandhahr districts that had epidemiologically relevant travel connections with the imported malaria cases were surveyed.

The vector survey included the collection of mosquitoes (adults and larvae) from the houses of reported cases and their surrounding houses, species identification, and enumeration of mosquito breeding habitats. Resting adult mosquito collection was conducted in households of 6 villages of Bulandshahr district during early morning (6 AM to 8 AM) using hand aspirators. The larval collection was also done in each village from all water-bearing sites, that is, ponds, ditches, large cement tanks, drains, and seepages, in peridomestic and domestic areas of each village. In the Raj Nagar catchment locality of SWD, the houses were searched from 7 AM to 9 AM, and larval collection was also conducted simultaneously if searched houses and containers were found positive for larval presence from domestic and peridomestic water bodies and other sites including overhead tanks, large open water bodies, tires, coolers, bird pots, flowerpots, iron containers, and solid wastes in urban catchment areas of the clinic and from domestic and peridomestic containers in the houses of reported cases. The collected larvae were reared in an insectary separately up to their emergence to identify the mosquito species. Identification of species was done following the standard taxonomic key as described by Christophers [26]. Adult mosquito collection was also done using the hand catch method [27], and the collected *Anopheles* mosquitoes were screened for the presence of malaria parasites (*P falciparum*, *P vivax*, *Plasmodium malariae*, and *Plasmodium ovale*), through polymerase chain reaction (PCR) [28], in pools of mosquitoes made village-wise and species-wise. To estimate the critical density of malaria vectors, per man-hour density (PMHD) was calculated as the number of anophelines collected per hour by an insect collector using the formula:

$$PMHD = \frac{\text{No. of mosquitoes collected}}{\text{Time spent (in hours)} \times \text{no. of insect collectors}}$$

Fever surveys and RACD were carried out in October 2019 in 6 villages of the Bulandshahr (UP) with support from the local health personnel (Accredited Social Health Activists and Health Inspectors). Fever camps were organized at a central location in each village, and the local health personnel informed the villagers about the camp and motivated them to visit. Incoming febrile cases were screened for malaria by using a rapid diagnostic test (SD Bio Line Malaria Ag P.f / P.v, Standard Diagnostics, Inc, Republic of Korea) as per the manufacturer's instructions. All febrile cases were treated symptomatically, and patients with malaria were treated as per the national drug policy. RACD was done as described by the WHO [29], and blood smears were prepared from apparently healthy individuals in and around the household of the index cases. The smears were examined for the presence of malaria parasites at ICMR-NIMR, Delhi (as described previously), and the results were communicated to the concerned health personnel for further management.

Data Entry and Statistical Analysis

All the collected data were entered in a Microsoft Excel 2016 spreadsheet and presented as proportions (percentages), medians, and ranges, where appropriate. The strength of association was estimated using a chi-square test, and a *P* value of less than .05 was considered statistically significant.

Results

Overview

A total of 14,748 fever cases were screened for malaria by microscopy from January 2017 to December 2019. The 3-year period prevalence of malaria was 2.4% (364/14,748). Out of these 364 cases, 355 (97.5%) were *P vivax* mono-infections, 8 *P falciparum* mono-infections (2.1%), and 1 mixed infection of *P vivax* and *P falciparum*. There was male predominance among patients with fever (59%) as well as patients with *P vivax* malaria (71%). More than half of the patients with *P vivax* malaria were in the age group of 15 - 29 years (183/355, 52%) with a median age of 22 years. The parasite burden ranged from 63 to 206,187 parasites per microliter of blood.

Imported Malaria Burden

Out of the 250 *P vivax* cases with a travel history (250/355, 70%), 186 (74%) cases were returning travelers and the remaining 64 (26%) cases were incoming travelers (Figure 1). However, relevant travel history to be able to label them as imported cases was available from 63% (222/355) of patients. Out of these imported *P vivax* cases, 173 (78%) cases were returning and 49 (22%) cases were incoming travelers. Ninety-five percent (212/222) of the imported cases were from UP and 142 of them (142/212, 67%) had traveled to 1 of the 6 districts of UP viz Bareilly, Badaun, Aligarh, Hathras, Bulandshahr, and Mainpuri. The remaining 10 imported cases had traveled to Uttarakhand (3; Dehradun and Hardwar), Rajasthan (3; Bikaner, Bundi, and Sawai Madhopur), Madhya Pradesh (2; Gwalior), Haryana (1; Gurugram), and Punjab (1; Sri Muktsar Sahib), as shown in Figure 2. Out of 9 patients with *P falciparum* malaria (including 1 mixed infection), 8 had a travel history.

XSL•FO
RenderX

Figure 1. Travel history among 355 patients with *Plasmodium vivax* malaria. Based on the epidemiologically relevant travel history, cases associated with travel were categorized into "imported" (222; shown in red font) and indigenous (133; gold colored font) cases. The geographical distribution (states) of imported cases is also mentioned.

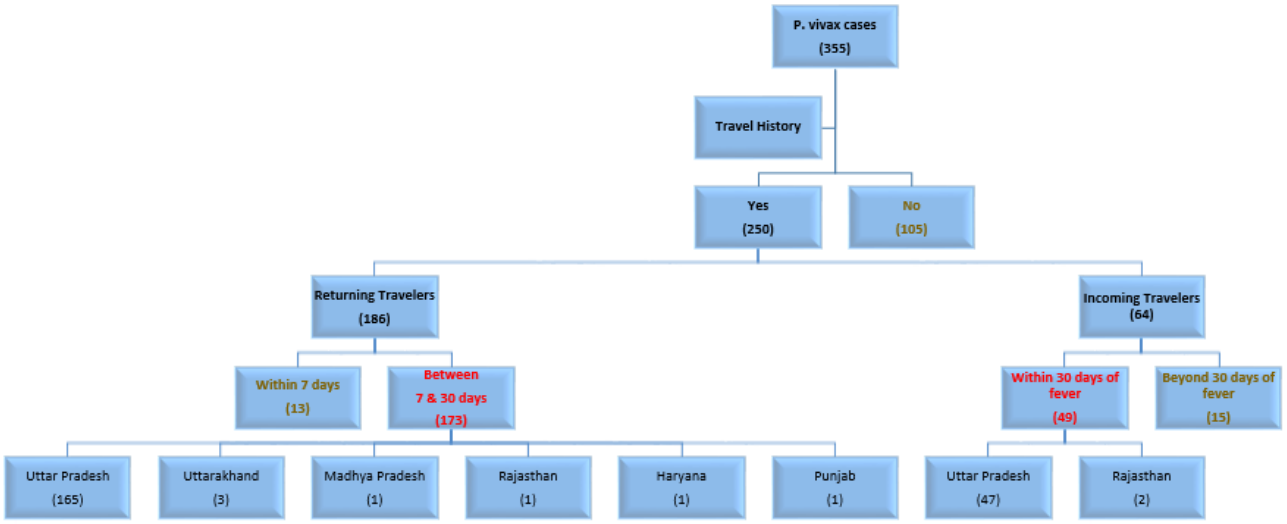
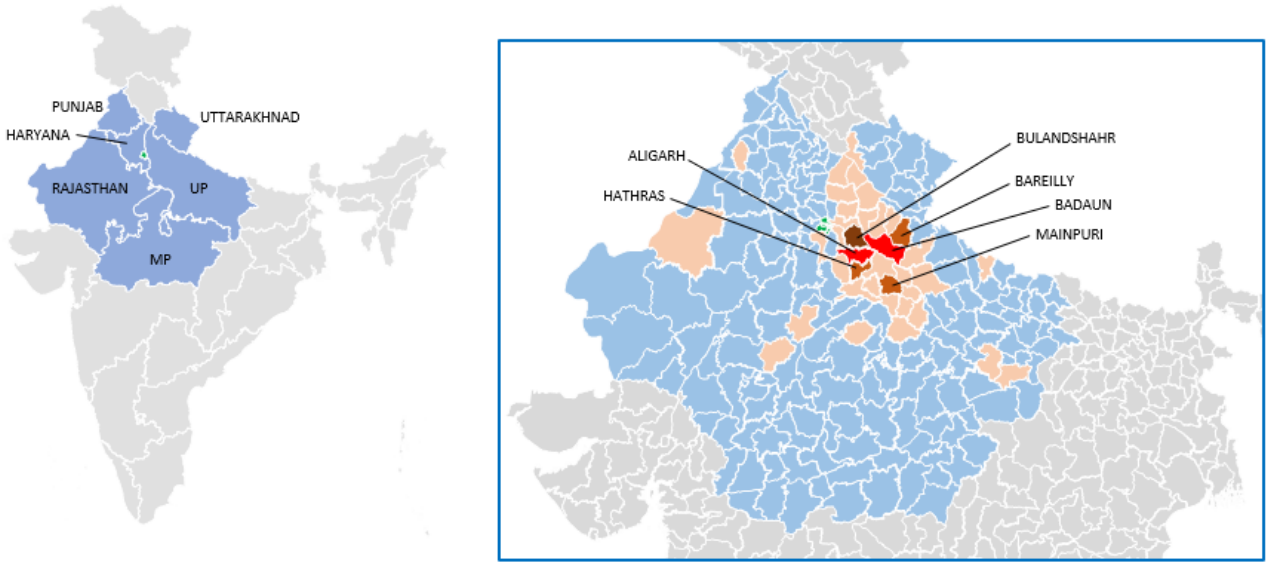


Figure 2. Areas traveled to by *Plasmodium vivax* malaria imported cases. The areas (states and union territories of India) are shown in blue (Punjab, Chandigarh, Haryana, Uttarakhand, UP, MP, and Rajasthan) whereas Delhi (study site) is shown in green. The zoomed-in image of the map in the inset shows further administrative breakdown of these 7 states and union territories (as districts) in blue. The districts within these 7 states and union territories, which are associated with the travel history of imported cases, are colored based on the number of imported cases contributed by each district: light orange (1 - 5 cases); dark orange (5 - 15 cases); darker orange (15 - 25 cases), and red (>25 cases). It is evident that UP has 3 dark orange districts: Hathras (10 cases), Mainpuri (12 cases), and Bareilly (15 cases); 1 darker orange district: Bulandshahr (25 cases); and 2 red districts: Aligarh (37 cases) and Badaun (47 cases). MP: Madhya Pradesh; UP: Uttar Pradesh.



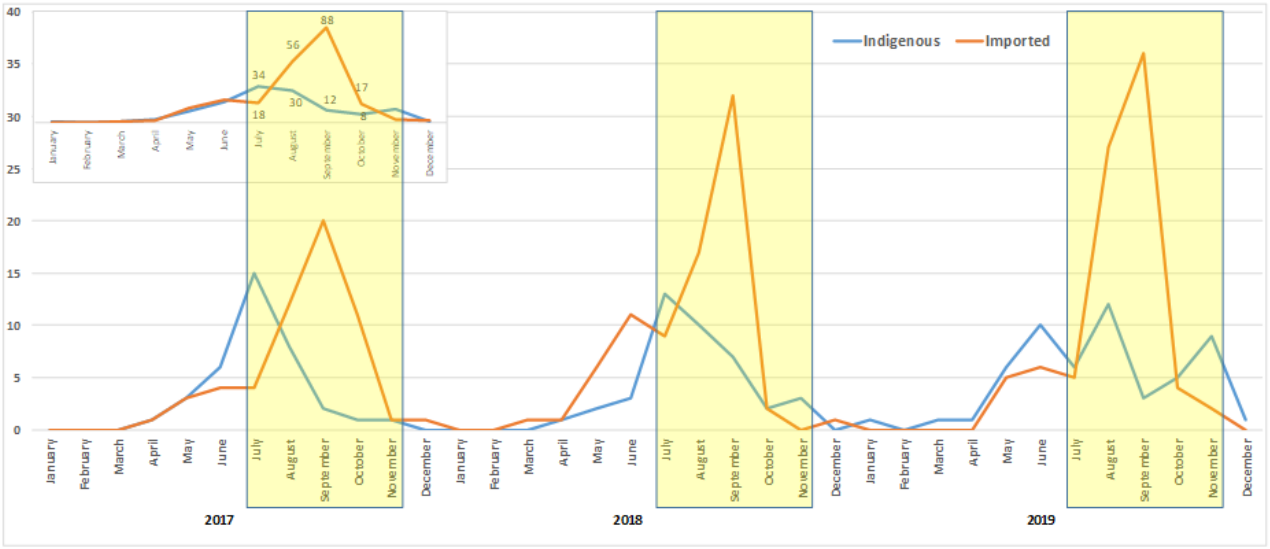
The reasons for travel in *P vivax* malaria cases included visiting their native residence in various states (mainly for returning travelers); education or employment (for incoming travelers); and visiting relatives, family, and friends during festivals (for returning and incoming travelers) as many Indian and regional festivals temporally coincide with the malaria transmission season. The minimum length of stay outside Delhi among travelers was 1 day while the maximum stay was of 111 days. Travel history was reported by the family members of the patients with malaria as well. There were 24 families with at least 2 members probably acquiring malaria after traveling.

A majority of the *P vivax* cases (279/355, 79%) were detected during the transmission season, that is, July to November. The nontransmission season (December to June) contributed to the remaining 21% (76/355) of cases.

During the transmission season (July to November; 2017 - 2019), the proportion of imported *P vivax* malaria cases diagnosed at the fever clinic was higher (65%) than that in the nontransmission season (53%; December to June), as shown in Figure 3, and the difference was statistically significant (χ^2_1=4.04; *P*=.031) at a 95% confidence level.

XSL•FO
RenderX

Figure 3. *Plasmodium vivax* malaria cases reported in Delhi (2017 - 2019). The figure shows the number of *P vivax* cases, indigenous (blue) and imported (imported), as reported each month and each year during the study period (January 2017 to December 2019) at the Indian Council of Medical Research-National Institute of Malaria Research fever clinic. The cumulative month-wise data from 2017 to 2019 are shown in the inset. The yellow semitransparent rectangles show the transmission season.



Entomological Surveys

Adult mosquito collections from the draining and catchment areas (Bagdola and Raj Nagar areas) of the fever clinic in SWD showed 4 mosquito species, with a low prevalence of *Anopheles* mosquitoes during the survey (September 2018 to December 2019). Out of 573 adult mosquitoes collected from 3395 houses, only 7 (1%) were *Anopheles stephensi* (Table 1), and all of them were found during the malaria transmission season (September). No other species of *Anopheles* were found. The majority of mosquitoes (443/573, 78%) were *Aedes aegypti* in the households.

XSL•FO
RenderX

Table . Month-wise adult mosquito collection in Bagdola and Raj Nagar, South West Delhi, between September 2018 and December 2019. Transmission season is highlighted in gold. The proportion of mosquito species identified out of the total is shown as a percentage (in parentheses).

		Houses checked, n	Mosquito species identified				
			Aedes aegypti, n (%)	*Aedes albopictus*, n (%)	*Anopheles stephensi*, n (%)	*Culex quinquefasciatus*, n (%)	Total, n
Month and year							
	September 2018[a]	316	28 (37)	0 (0)	1 (1)	46 (62)	75
	October 2018[a]	97	14 (70)	0 (0)	0 (0)	6 (30)	20
	November 2018[a]	402	63 (73)	0 (0)	0 (0)	23 (27)	86
	December 2018	220	16 (84)	1 (5)	0 (0)	2 (11)	19
	January 2019	112	1 (100)	0 (0)	0 (0)	0 (0)	1
	February 2019	70	0 (0)	0 (0)	0 (0)	0 (0)	0
	March 2019	179	2 (40)	0 (0)	0 (0)	3 (60)	5
	April 2019	60	13 (100)	0 (0)	0 (0)	0 (0)	13
	May 2019	344	10 (63)	0 (0)	0 (0)	6 (37)	16
	June 2019	367	89 (87)	0 (0)	0 (0)	13 (13)	102
	July 2019[a]	352	53 (75)	0 (0)	0 (0)	18 (25)	71
	August 2019[a]	270	61 (91)	0 (0)	6 (9)	0 (0)	67
	September 2019[a]	301	44 (90)	5 (10)	0 (0)	0 (0)	49
	October 2019[a]	211	40 (100)	0 (0)	0 (0)	0 (0)	40
	November 2019[a]	33	9 (100)	0 (0)	0 (0)	0 (0)	9
	December 2019	61	0 (0)	0 (0)	0 (0)	0 (0)	0
Total		3395	443 (78)	6 (1)	7 (1)	117 (20)	573

[a]Transmission season.

On the other hand, a total of 9 species of mosquitoes (5 anopheline, 2 *Aedes*, 1 *Culex*, and 1 *Armigeres* species) were collected from 6 villages in the Bulandshahr district (Table 2) of UP. Out of the 416 adult mosquitoes collected, 126 (30%) were *Anopheles* with *Anopheles subpictus* (21%) and *Anopheles culicifacies* (5%) dominating the anopheline burden. *Culex quinquefasciatus* was the major mosquito species found in rural Bulandshahr (235/416, 57%). Among the anophelines, *A subpictus* (PMHD 16.61) was the most abundant species followed by the main rural vector *A culicifacies* (PMHD 4.25) and *A stephensi* (PMHD 1.35).

XSL•FO
RenderX

Table . Adult mosquito collection and species identification in the surveyed 6 villages of Bulandshahr district of Uttar Pradesh during October 2019. The proportion of different mosquito species identified out of the total is shown as a percentage (in parentheses).

		Mosquito species identified									
		Aedes aegypti, n (%)	*Aedes albopictus*, n (%)	*Anopheles stephensi*, n (%)	*Anopheles annularis*, n (%)	*Anopheles culicifacies*, n (%)	*Anopheles nigerrimus*, n (%)	*Anopheles subpictus*, n (%)	*Culex quinquefasciatus*, n (%)	*Armigeris subalbatus*, n (%)	Total, n
Villages											
	Adauli	1 (0.7)	0 (0)	0 (0)	1 (0.7)	9 (7)	4 (3)	30 (22)	86 (63)	6 (4)	137
	Lakhaoti	0 (0)	1 (2)	0 (0)	4 (6)	5 (8)	0 (0)	12 (18)	26 (40)	18 (27)	66
	Shikarpur (Kot Kalan 1)	4 (6)	2 (3)	5 (8)	0 (0)	4 (6)	0 (0)	6 (9)	45 (68)	0 (0)	66
	Kutubpur	0 (0)	0 (0)	0 (0)	1 (4)	0 (0)	0 (0)	5 (19)	20 (77)	0 (0)	26
	Mustafabad Daduwa	0 (0)	0 (0)	2 (5)	0 (0)	2 (5)	1 (2)	0 (0)	18 (41)	21 (47)	44
	Dinoul	0 (0)	0 (0)	0 (0)	0 (0)	2 (3)	0 (0)	33 (43)	40 (52)	2 (3)	77
Total		5 (1)	3 (0.7)	7 (2)	6 (1)	22 (5)	5 (1)	86 (21)	235 (57)	47 (11)	416

A total of 24 collected anopheline mosquito pools (22 mosquitoes in 6 pools of *A culicifacies*, 7 mosquitoes in 2 pools of *A stephensi*, 3 mosquitoes in 3 pools of *Anopheles annularis*, 86 mosquitoes in 11 pools of *A subpictus*, and 5 mosquitoes in 2 pools of *Anopheles nigerrimus*) were tested by PCR for the presence of malaria parasites; however, none of the pools was found positive for the presence of malaria parasites.

The vector survey to identify mosquito breeding habitats in Raj Nagar and Bagdola catchment localities revealed that out of 14,333 containers (including large containers, cemented tanks, and underground tanks) checked, *Anopheles* breeding was found only in 8 containers that included coolers, overhead tanks, cement tanks, and iron containers. There were no large water bodies in the surrounding area of the survey.

A mosquito breeding habitat survey in the 6 villages of Bulandshahr found that out of the 203 water-holding containers and water bodies, 51 (25%) had *Anopheles* breeding. Major breeding sites included drains (4/7, 57%), canals (1/2, 50%), ponds (3/6, 50%), and pits (9/20, 45%). Other sites where breeding was found included domestic and peridomestic water bodies (26/123, 21%), cemented ground tanks (cattle tanks; 5/26, 19%), and rice fields (3/19, 16%). *A culicifacies* were found mostly in canal banks and village ponds whereas *C quinquefasciatus* in sewages.

Fever Surveys

Camp-based fever surveys in the 6 villages identified 86 persons with fever, with only 1 person testing positive for *P vivax* by the rapid diagnostic tests. RACD from 22 asymptomatic persons around the households of the index cases revealed 5 additional cases (5/22, 23%) of *P vivax* by microscopy.

Discussion

Principal Findings

Between January 2017 and December 2019, 355 monoinfected *P vivax* cases were reported, and out of them, 63% (n=223) could be categorized as possible imported malaria cases based on relevant travel history, thus forming a major burden of reported malaria cases in SWD. The study also detected 5 additional *P vivax* cases through RACD done in villages visited by the imported cases and identified malaria vectors of anopheline species and their breeding habitats in such areas.

The distribution of malaria cases reported in the fever clinic at ICMR-NIMR revealed that the malaria cases were more likely to be imported than indigenous and occur in transmission season. The period July to November is considered to be the malaria transmission season in Delhi, while December to June is considered a nontransmission season [30].

Although 67% of the *P vivax* cases were imported, being associated with relevant travel history, the remaining 37% of indigenous cases could be associated with possible local transmission of *P vivax* in SWD, as suggested by the presence of anophelines in Delhi (this study) and the reported presence of malaria vectors in Delhi [31,32]. *P vivax* malaria cases during the nontransmission season or in nontravelers might also be recurrences or relapses due to the activation of hypnozoites from the liver [30].

Recent travel within the country is associated with malaria in various studies [22,23]. This study showed that the proportion of males was more than females among imported as well as indigenous malaria cases and a similar trend was seen in patients with fever as well. In similar studies, men traveling away from home in the last 30 days were reported to be strongly associated with malaria in Ethiopia [24,25].

Although Delhi shares its borders with the state of UP in the east and the state of Haryana in the remaining directions, we

XSL•FO
RenderX

observed that ≈96% of the imported cases were from UP. Data highlights of the census of India in 2001 and 2011 show that Delhi receives a higher number of migrants (≈50% of the total in-migrants) from UP versus that from Haryana (≈10% of in-migrants) [33]. With >20-fold higher malaria burden in UP (than in Haryana), the findings of >95% of cases being imported from UP are explainable [12]. Reasons for migration to Delhi are cited to be due to employment, business, education, marriage, etc [33,34]. The reasons for travel reported during this study were festivals, farming, and visits to relatives. Those visiting friends and relatives in malaria-endemic areas have been reported to be at high risk of contracting malaria [35,36].

Many districts in UP contributed to the imported *P vivax* cases in SWD (Figure 2); however, 6 UP districts contributed 10 or more cases: Hathras (10 cases), Mainpuri (12 cases), Bareilly (15 cases), Bulandshahr (25 cases), Aligarh (37 cases), and Badaun (47 cases). Further investigations (vector and fever surveys) were carried out in 6 villages of Bulandshahr district only due to reasons explained earlier. Bulandshahr district of UP, located southeast of Delhi, is situated between the Ganga and Jamuna rivers, which are the major rivers in North India. The soil is very fertile where mainly sugarcane, and wheat are grown. Irrigation facilities are also well-developed and this area is canal-irrigated as well [37] which makes the area highly mosquitogenic.

During vector surveillance in 6 villages of Bulandshahr, 25% of the water bodies were positive for anopheline larval breeding, and 5 species of adult *Anopheles* mosquito were found during adult mosquito collections. Unlike Bulandshahr, where almost every village had ponds, canals, and ample water in surrounding areas providing sufficient opportunities for the breeding of anophelines, Delhi is highly urbanized and lands are not available for ponds and crop fields. In comparison to Bulandshahr, the catchment areas of the fever clinic (Raj Nagar and Bagdola localities) of SWD had a very low prevalence of *Anopheles*. Only 1 species, that is, *A stephensi* was present in these localities in contrast to Bulandshahr where 5 species of *Anopheles* were collected out of which 2 were major malaria vectors, that is, *A stephensi* and *A culicifacies*. Larval surveys suggested that urban and rural areas have different breeding habitats. In villages, natural water bodies like ponds, canals, pits, and crop fields were more prominent and harbored more breeding than the peridomestic and domesticated containers in contrast to the urban areas where natural breeding sites are limited and were confined to peridomestic and domestic containers only. Mosquito species like *A stephensi* and *A aegypti* are adapted to breed in such urban areas whereas *A culicifacies* mostly breed in outdoor natural water habitats like canal banks, village ponds, etc and *C quinquefasciatus* is found in sewage water.

Among malaria vectors, *A culicifacies* was found to be the dominating mosquito species along with an efficient malaria-transmitting vector, *A stephensi*. However, the month of the survey (October) had low vector density, which may be due to the low ambient temperature (20 - 25 °C) during the survey period. Further, the mosquitoes that were collected from the villages of Bulandshahr district did not show parasite positivity by PCR. This may be due to multiple factors, including the very short period of vector survey (30 d), a limited number of vectors collected toward the end of transmission (October), and the difference in time of mosquito collection and case reporting in the clinic, as the vector survey was carried out as a response for tracking ongoing transmission in areas previously visited by imported malaria cases.

The 6 districts of UP that contributed most to imported cases in SWD had API (2018) of 0.06 (Bulandshahr and Mainpuri), 0.1 (Hathras and Aligarh), 5.5 (Badaun), and 7.3 (Bareilly), whereas the API of Delhi during this period was 0.02 [38]. A survey was therefore, carried out in Bulandshahr wherein 1 out of 86 febrile cases (fever survey) and 5 out of 22 afebrile persons (RACD) were identified with *P vivax* infections which signifies that further studies are needed to assess the extent of asymptomatic *Plasmodium* infection and its role in transmission in such areas.

The prevalence of malaria was found lower in the camp-based fever surveys compared with the prevalence reported from the fever clinic in SWD. This may be because the camp-based fever surveys were carried out during October, which marks the end of the transmission season and therefore may have had a lower number of cases. Further, the catchment area of the camp included a village whereas the fever clinic at SWD has a much larger and densely populated catchment area.

The regions nearing malaria elimination tend to have a heterogeneous endemicity, with foci of high burden and areas with no endogenous malaria transmission. For eliminable diseases such as malaria, within-country migration is a recognized but understudied challenge in such geographically heterogeneous transmission to sustain zero-burden and prevent reintroduction and re-establishment of transmission [23,39]. Such regions often lack a robust surveillance system to deal with imported cases besides treating them, and there appears to be a lack of documented cross-reporting and targeted intervention in the foci where the infections probably originated.

This study is therefore important, as it attempted to comprehensively investigate imported malaria cases, and may be adopted and locally adapted as an implementation model in similar areas with no or few locally acquired malaria cases.

Limitations

There is an obvious limitation of this study that a limited geographical area for fever and vector surveillance was selected. Nevertheless, the study shows the presence of malaria transmission in areas where patients with malaria reporting to the fever clinic had traveled. The study was also limited by the possibility of recall bias of study participants correctly recalling the exact dates of travel for both imported and indigenous patients with malaria. The investigators, however, tried to extract the near-exact dates by relating travel to the locally relevant cultural events, festivals, and other contextual events. Misclassification bias (incorrect classification of imported malaria) could have stemmed from the possibilities of recurrences and relapses of *P vivax* infections acquired before the study period. The study did not use available molecular methods to differentiate recurrent versus new infection and therefore could not account for *P vivax* relapses. However, the

XSL•FO
RenderX

possibility that only up to 40% of *P vivax* infections in the study area (that too in the nontransmission season) could be due to possible relapses [30], nonavailability of molecular methods to confidently differentiate recurring infections from new infections, and random possibilities of recurrence in both the imported and indigenous patients with malaria may have compensated for this limitation.

Last, only 2 methods for mosquito collection were used. The hand catch method using an aspirator was the only method adopted for the estimation of vector density. For larval collection, dips were taken from water bodies for assessment of breeding. No other method was adopted for mosquito collection, and this might have underestimated the frequency of vectors and their possible infection with *Plasmodium*, because the PCR results did not show any vector positivity.

Conclusions

A significant burden (63%) of *P vivax* malaria reported in SWD was found to be imported from UP. Malaria transmission possibilities (multiple breeding sites suggesting stable breeding ground of anophelines) were higher in Bulandshahr than in SWD. Indigenous cases in SWD are also a concern, as adult vectors were also found in the area. Despite the detection of additional *P vivax* cases following RACD in Bulandshahr and vector-breeding sites being identified, the conclusion that the imported cases really acquired infections from the surveyed areas in Bulandshahr is limited by the correct recall of travel or fever dates and the possibility of relapses due to *P vivax*.

Way Forward

The study reiterates that population movement is a key challenge for malaria elimination, particularly in areas with very low or 0 indigenous malaria cases, and investigations of the potential role of travelers in introducing malaria and its further spread are definitely needed. Since the epidemiology of migration malaria is contextual, appropriate tailor-made measures are needed, both at the sites where imported cases are detected and in areas where these infections might have been acquired. In addition, effective “information, education, and communication” activities to educate travelers regarding the potential risks of travel-associated malaria and its prevention should be undertaken.

Acknowledgments

The authors acknowledge the support provided by the state programme officer of Uttar Pradesh and district malaria officer of Bulandshahr. The authors are thankful to the technical staff of fever clinic and vector biology division of ICMR-NIMR for their field support. The study did not receive any specific funding from a dedicated grant and was funded by Indian Council of Medical Research-National Institute of Malaria Research (ICMR-NIMR), New Delhi, out of its intramural funds.

Disclaimer

The views expressed in the submitted article are our own and not an official position of the institution or funder. No generative artificial intelligence was used in any portion of the manuscript writing.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

DS, HS, and A Sinha conceived the idea, carried out clinical work and fieldwork, and wrote the first draft. BS carried out laboratory work. CPY and MPS carried out analysis. AA and A Sinha analyzed and interpreted the data and reviewed the manuscript. A Sharma reviewed the manuscript. HS and A Sinha contributed equally as co-corresponding authors.

Conflicts of Interest

None declared.

References

1. World malaria report 2022. : World Health Organization; 2022 URL: https://www.who.int/publications/i/item/9789240064898 [accessed 2023-06-01]
2. Chapter 2 – migration and migrants: a global overview. World Migration Report 2018. : International Organization for Migration (IOM); 2017 URL: https://onlinelibrary.wiley.com/doi/10.1002/wom3.2 [accessed 2024-04-12] [doi: 10.1002/wom3.2]
3. Lu G, Zhang D, Chen J, et al. Predicting the risk of malaria re-introduction in countries certified malaria-free: a systematic review. Malar J 2023 Jun 6;22(1):175. [doi: 10.1186/s12936-023-04604-4] [Medline: 37280626]
4. Requena-Méndez A, Cattaneo P, Bogale RT, et al. Malaria parasite prevalence among migrants: a systematic review and meta-analysis. Clin Microbiol Infect 2023 Dec;29(12):1528-1537. [doi: 10.1016/j.cmi.2023.09.010] [Medline: 37739263]
5. Cohen JM, Smith DL, Cotter C, et al. Malaria resurgence: a systematic review and assessment of its causes. Malar J 2012 Apr 24;11:122. [doi: 10.1186/1475-2875-11-122] [Medline: 22531245]

XSL•FO
RenderX

6. Pousibet-Puerto J, Lozano-Serrano AB, Soriano-Pérez MJ, et al. Migration-associated malaria from Africa in southern Spain. Parasit Vectors 2021 May 7;14(1):240. [doi: 10.1186/s13071-021-04727-0] [Medline: 33962647]
7. Rodrigues PT, Valdivia HO, de Oliveira TC, et al. Human migration and the spread of malaria parasites to the new world. Sci Rep 2018 Jan 31;8(1):1993. [doi: 10.1038/s41598-018-19554-0] [Medline: 29386521]
8. Wångdahl A, Wyss K, Saduddin D, et al. Severity of Plasmodium falciparum and non-falciparum malaria in travelers and migrants: a nationwide observational study over 2 decades in Sweden. J Infect Dis 2019 Sep 13;220(8):1335-1345. [doi: 10.1093/infdis/jiz292] [Medline: 31175365]
9. Sorichetta A, Bird TJ, Ruktanonchai NW, et al. Mapping internal connectivity through human migration in malaria endemic countries. Sci Data 2016 Aug 16;3:160066. [doi: 10.1038/sdata.2016.66] [Medline: 27529469]
10. Prothero RM. Migration and malaria risk. Health Risk Soc 2001 Mar 1;3(1):19-38. [doi: 10.1080/13698570123413]
11. Tyagi BK. Trans-border migration and malaria in desert populations. In: Desert Malaria: Springer; 2023. [doi: 10.1007/978-981-19-7693-3_21]
12. Country Office for India, World Health Organization. WHO Country Office for India. 2016. URL: https://iris.who.int/handle/10665/246096 [accessed 2023-05-31]
13. Sharma S, Verma R, Yadav B, Kumar A, Rahi M, Sharma A. What India can learn from globally successful malaria elimination programmes. BMJ Glob Health 2022 Jun;7(6):e008431. [doi: 10.1136/bmjgh-2022-008431] [Medline: 35760440]
14. Rahi M, Sharma A. Malaria control initiatives that have the potential to be gamechangers in India's quest for malaria elimination. Lancet Reg Health Southeast Asia 2022 Jul;2:100009. [doi: 10.1016/j.lansea.2022.04.005] [Medline: 37383297]
15. Khan N, Awasthi G, Das A. How can the complex epidemiology of malaria in India impact its elimination? Trends Parasitol 2023 Jun;39(6):432-444. [doi: 10.1016/j.pt.2023.03.006] [Medline: 37031071]
16. Kumar A. Some considerable issues concerning malaria elimination in India. J Vector Borne Dis 2019;56(1):25-31. [doi: 10.4103/0972-9062.257770] [Medline: 31070162]
17. Malaria in migrants and mobile populations. World Health Organization. URL: https://www.who.int/malaria/areas/high_risk_groups/migrants_mobile_populations/en [accessed 2023-05-31]
18. Karim MA, Kabir MM, Siddiqui MA, Laskar MSI, Saha A, Naher S. Epidemiology of imported malaria in Netrokona district of Bangladesh 2013-2018: analysis of surveillance data. Malar Res Treat 2019;2019:6780258. [doi: 10.1155/2019/6780258] [Medline: 31312425]
19. Migrated population in delhi from 2003 to 2022. Statista. URL: https://www.statista.com/statistics/1400863/india-migration-in-delhi/ [accessed 2023-05-31]
20. Singh J, Bhattacharji LM. Rapid staining of malarial parasites by a water soluble stain. Ind Med Gaz 1944 Mar;79(3):102-104. [Medline: 29012018]
21. Anvikar AR, Arora U, Sonal GS, et al. Antimalarial drug policy in India: past, present & future. Indian J Med Res 2014 Feb;139(2):205-215. [Medline: 24718394]
22. Poumerol G, Wilder-Smith A, editors. Vaccines for routine and selective use: rabies. In: International Travel and Health: World Health Organization; 2012. URL: https://www.who.int/publications/i/item/9789241580472 [accessed 2023-05-31]
23. Lynch CA, Bruce J, Bhasin A, Roper C, Cox J, Abeku TA. Association between recent internal travel and malaria in Ugandan highland and highland fringe areas. Trop Med Int Health 2015 Jun;20(6):773-780. [doi: 10.1111/tmi.12480] [Medline: 25689689]
24. Alemu K, Worku A, Berhane Y, Kumie A. Men traveling away from home are more likely to bring malaria into high altitude villages, northwest Ethiopia. PLoS One 2014;9(4):e95341. [doi: 10.1371/journal.pone.0095341] [Medline: 24748159]
25. Yukich JO, Taylor C, Eisele TP, et al. Travel history and malaria infection risk in a low-transmission setting in Ethiopia: a case control study. Malar J 2013 Jan 24;12:33. [doi: 10.1186/1475-2875-12-33] [Medline: 23347703]
26. Christophers SR. The Fauna of British India, Including Ceylon and Burma. Diptera, Vol. Iv, Family Culicidae, tribe Anophelini: Taylor & Francis; 1933.
27. Manual on Practical Entomology in Malaria / Prepared by the WHO Division of Malaria and Other Parasitic Diseases: World Health Organization Division of Malaria and Other Parasitic Diseases; 1995. URL: https://iris.who.int/handle/10665/42481 [accessed 2024-12-04]
28. Tassanakajon A, Boonsaeng V, Wilairat P, Panyim S. Polymerase chain reaction detection of Plasmodium falciparum in mosquitoes. Trans R Soc Trop Med Hyg 1993;87(3):273-275. [doi: 10.1016/0035-9203(93)90124-9] [Medline: 8236389]
29. Malaria surveillance, monitoring & evaluation: a reference manual. : World Health Organization; 2018 URL: https://www.who.int/publications/i/item/9789241565578 [accessed 2024-12-04]
30. Adak T, Sharma VP, Orlov VS. Studies on the Plasmodium vivax relapse pattern in Delhi, India. Am J Trop Med Hyg 1998 Jul;59(1):175-179. [doi: 10.4269/ajtmh.1998.59.175] [Medline: 9684649]
31. Savargaonkar D, Nagpal BN, Srivastava B, Anvikar AR, Valecha N. The footprints of relapsing malaria in southwest Delhi, India. J Vector Borne Dis 2015 Dec;52(4):287-292. [Medline: 26714507]
32. Singh RK, Mittal PK, Kumar G, Dhiman RC. Insecticide susceptibility status of Aedes aegypti and Anopheles stephensi larvae against temephos in Delhi, India. Int J Mosq Res 2014;1(3):69-73.

XSL•FO
RenderX

33. Census of India 2001. Data highlights: migration tables (D1, D1 (Appendix), D2 and D3 tables). The Hindu Centre For Politics and Public Policy. 2001. URL: https://www.thehinducentre.com/publications/policy-watch/article33219541.ece/binary/data_highlights_D1D2D3.pdf [accessed 2023-05-31]
34. Sarkar P. An overview of out-migration from Uttar Pradesh using census 2011. J Migr Aff 2020;2(2):58. [doi: 10.36931/jma.2020.2.2.58-66]
35. Casuccio A, D'Angelo C, Casuccio N, Di Carlo P, Immordino P. Visiting friends and relatives (VFRs) travelers and imported malaria in the Palermo district (Sicily). Ann Ist Super Sanita 2014;50(4):369-374. [doi: 10.4415/ANN_14_04_13] [Medline: 25522079]
36. Heywood AE, Zwar N, Forssman BL, et al. The contribution of travellers visiting friends and relatives to notified infectious diseases in Australia: state-based enhanced surveillance. Epidemiol Infect 2016 Dec;144(16):3554-3563. [doi: 10.1017/S0950268816001734] [Medline: 27574034]
37. Preliminary draft report. District survey report for (planning & execution of) minor mineral excavation. : Directorate of Geology & Mining; 2016 URL: https://cdn.s3waas.gov.in/s3091d584fced301b442654dd8c23b3fc9/uploads/2019/02/2019021130.pdf [accessed 2021-01-11]
38. Directorate of National Vector Borne Diseases Control Programme, 22 Sham Nath Marg, DELHI-110054. Annual report 2018. National Center for Vector Borne Diseases Control (NCVBDC). 2018. URL: https://nvbdcp.gov.in/Doc/Annual-Report-2018.pdf [accessed 2024-12-04]
39. World malaria report 2023. : World Health Organization; 2023 URL: https://www.who.int/publications/i/item/9789240086173 [accessed 2024-12-04]

Abbreviations

API: annual parasite incidence
ICMR-NIMR: Indian Council of Medical Research-National Institute of Malaria Research
NCVBDC: National Center for Vector Borne Diseases Control
PCR: polymerase chain reaction
PMHD: per man-hour density
RACD: reactive case detection
SWD: South West Delhi
SWD: South West Delhi
UP: Uttar Pradesh
WHO: World Health Organization

Edited by A Mavragani, T Sanchez; submitted 18.06.23; peer-reviewed by J Campbell, P Sahu; revised version received 16.05.24; accepted 23.05.24; published 08.01.25.

Please cite as:
Savargaonkar D, Srivastava B, Yadav CP, Singh MP, Anvikar A, Sharma A, Singh H, Sinha A
Contribution of Travelers to Plasmodium Vivax Malaria in South West Delhi, India: Cross-Sectional Survey
JMIR Public Health Surveill 2025;11:e50058
URL: https://publichealth.jmir.org/2025/1/e50058
doi:10.2196/50058

© Deepali Savargaonkar, Bina Srivastava, Chander Prakash Yadav, Mrigendra Pal Singh, Anup Anvikar, Amit Sharma, Himmat Singh, Abhinav Sinha. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Potentially Inappropriate Use of Transdermal Fentanyl in Working-Age and Older Adult Populations with Non-Cancer Pain: Nationwide Cross-Sectional Study

Seunghyun Cheon, PharmD, PhD; Seo-Yong Choi, MS; Kyu Ri Kim, BS; Han-Gon Choi, PhD; Jee-Eun Chung, PhD

College of Pharmacy and Institute of Pharmaceutical Science and Technology, Hanyang University, 55 Hanyangdahak-ro, Ansan, Republic of Korea

Corresponding Author:
Jee-Eun Chung, PhD
College of Pharmacy and Institute of Pharmaceutical Science and Technology, Hanyang University, 55 Hanyangdahak-ro, Ansan, Republic of Korea

Abstract

Background: As a potent opioid analgesic, fentanyl transdermal patches (FTDs) have been widely used in patients with moderate to severe pain. However, increasing concerns about the opioid epidemic have made it important to strengthen the rational use and management of FTDs.

Objective: Our objective was to investigate the trends in the use of FTD and to evaluate potentially inappropriate FTD use in patients with noncancer pain, based on established evaluation criteria, referring various recommendations and guidelines.

Methods: A nationwide cross-sectional study was conducted using a national insurance claims database from 2014 to 2020. The study included patients who were prescribed FTDs at least once a year in an outpatient setting, while excluding cancer or pediatric patients. To identify potentially inappropriate use, we developed evaluation criteria based on the established recommendations for the safe use of prescription opioids in patients with noncancer pain and assessed each patient's compliance. The working-age and older adult groups were compared to evaluate the differences in FTD use, and modified Lorenz curves and Gini coefficients were used to assess the equality of FTD use.

Results: A total of 5386 patients and their 19,800 reimbursements were included in the study. The number of patients with FTDs decreased from 58.6 to 53.7 per 100,000 registrants between 2014 and 2020. Meanwhile, the number of reimbursements increased by 7.4%, from 203.2 to 218.1 per 100,000 registrants during the same period. The working-age patients had an average of 3.9 reimbursements per year, with an average morphine milligram equivalent per day (MME/day) of 64.4 for each reimbursement. The older adult patients had an average of 3.5 reimbursements per year, and their average MME/day was 47.9. As a result of applying the evaluation criteria, 567 (24.5%) working-age patients and 531 (17.3%) older adult patients were identified as potentially inappropriate FTD users. Among patients with multiple FTD prescriptions, the working-age group with potentially inappropriate FTD use had significantly higher estimated MME/day than the older adult group (P<.001). The modified Lorenz curve showed that more than 70% of the total FTDs used in the working-age group were consumed by those with potentially inappropriate FTD use, while older adult group with potentially inappropriate FTD use accounted for less than 50% of the total older adult patients. The working-age patients also had a higher Gini coefficient than the older adult group, indicating unequal use of FTDs (0.461 vs 0.406).

Conclusions: Although the number of patients receiving FTD prescriptions has decreased during the study period, the total amount of FTDs consumed increased, suggesting that caution is warranted. This study also highlights the potential for inappropriate FTD use in working-age patients. Further research is needed to quantify and qualify the risk factors in patients with potentially inappropriate use, given the clinical rationale associated with prescribing FTDs.

(JMIR Public Health Surveill 2025;11:e63960) doi:10.2196/63960

KEYWORDS

transdermal fentanyl; prescription opioid; potentially inappropriate use; claim data; working age; older adult; Lorenz curves; Gini coefficients

XSL•FO
RenderX

Introduction

Background

In recent years, the opioid crisis has emerged as a significant global health challenge [1,2], with an increasing number of patients with opioid use disorders (OUDs) leading to substantial socioeconomic burdens. While the exact cause of the opioid crisis remains controversial, the US government attributes it primarily to lenient management and chronic use of prescription opioids [3-5]. Wei et al [6] reported that approximately 65% of adult patients under 65 years diagnosed with OUD or overdose in the United States had used prescription opioids. Several studies have also examined the patterns of prescription opioid use in different countries, highlighting concerns about excessive consumption and the potential for opioid crises in several regions [7-10]. An analysis of trends in global prescription opioid consumption using pharmaceutical sales data found excessively high use in some countries [11], while another study using VigiBase raised concerns about the possibility of an opioid crisis in several countries [12], and Ju et al's [13] study documented a 4.0% annual increase in global prescription opioid usage.

Similarly, the use of prescription opioids has steadily increased in South Korea [7,13,14]. The implementation of the Narcotics Information Management System (NIMS) since 2018 has aimed to monitor the entire process from production to distribution, prescription, and dispensing of scheduled drugs, including prescription opioids [15]. However, approximately 15% of noncancer patients receiving prescription opioids in South Korea were evaluated to have been prescribed inappropriately [16]. This percentage is comparable to the 18% of Medicaid beneficiaries in the United States who were found to have received potentially inappropriate opioid prescriptions before the declaration of the opioid crisis [17]. In addition, the incidence of opioid abuse in noncancer patients with chronic prescription opioid use was reported to be 19.9 per 100 person-years [18]. The illegal use of FTDs among adolescents has emerged as a social concern [19]. In response to these problems, the Korea Ministry of Food and Drug Safety established the recommendations for safe use of prescription opioids for noncancer pain, based on recommendations from various countries such as the Centers for Disease Control and Prevention (CDC) guidelines [20].

Fentanyl and its analogues have been particularly highlighted in the recent opioid epidemic due to their potency and potential for abuse. Despite reductions in the dispensing of other prescription opioids in the United States, fentanyl prescriptions have shown minimal decline or even increases in certain health care settings [21]. In Australia, a significant percentage (4%) of deaths attributed to prescription opioids were linked to fentanyl alone, with a significant portion of fentanyl-related deaths associated with prescribed fentanyl [22,23]. Despite the high risk of abuse, the fentanyl prescriptions have increased rapidly, especially in high-income countries [11,16], and research on the appropriateness of FTD use in the real world remains limited. The most commonly prescribed form of fentanyl is the fentanyl transdermal patch (FTD). FTDs are most widely used to control chronic pain because of their extended-release property that steadily releases the drug over 72 hours to retain a high plasma concentration [24]. Therefore, it is necessary to investigate whether the use and management of FTDs for medical purposes has been appropriate. Given the differences in health profiles, such as the prevalence of chronic pain and prescription opioid use, between older adults and working-age adults [25,26], it is essential to analyze FTD use in these 2 groups separately. Furthermore, opioid abuse and its socioeconomic impacts among working-age adults have become increasingly problematic in recent years [27,28]. Comparing these 2 groups might provide a more precise understanding of FTD use patterns.

Purpose

We aimed to investigate the trend of FTD usage and evaluate the potentially inappropriate FTD use in patients with noncancer pain based on established evaluation criteria by referring to various recommendations and guidelines.

Methods

Study Design

We conducted a nationwide cross-sectional study using the national health insurance claim database. This study was conducted in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cross-sectional studies (Multimedia Appendix 1).

Data Sources

We used the Health Insurance Review and Assessment Service-National Patient Sample (HIRA-NPS) claim database from 2014 to 2020. The HIRA-NPS claim database is derived from the national health insurance system, which covers approximately 98% of the South Korean population [29]. It is constructed through stratified random sampling based on sex and age group from all individuals who enrolled in the Korean national health insurance by the HIRA and provided to researchers. The HIRA-NPS datasets contain information on approximately 1,400,000 registrants per year for 2014 - 2018 (sampling ratio, 3%) and 1,000,000 registrants per year for 2019 - 2020 (sampling ratio, 2%). The information of registrants in the HIRA-NPS includes an encrypted personal identification code, reimbursement codes, sex, age, diagnoses, prescription information, and health care institution information [30].

Study Patients

Patients who had been prescribed FTDs at least once a year in an outpatient setting were included. To investigate FTD use for noncancer pain only, the participants with a history of cancer diagnosis were excluded from each year of data. Pediatric patients younger than 20 years were also excluded. The patients were classified into working-age (<65 years of age) and older adult (≥65 years of age) groups. Subsequently, each group was divided into a single-prescription group and a multiple-prescription group: working-age patients with a single FTD prescription group (WS), working-age patients with multiple FTD prescriptions group (WM), older adult patients with a single FTD prescription group (OS), and older adult

XSL•FO
RenderX

patients with multiple FTD prescription group (OM). It was evaluated whether the recommendations for the evaluation criteria established in this study were followed, resulting in each group being divided into a group with appropriate FTD use and a group with potentially inappropriate FTD use: WS with appropriate FTD use (WS-A), WS with potentially inappropriate FTD use (WS-I), WM with appropriate FTD use (WM-A), WM with potentially inappropriate FTD use (WM-I), OS with appropriate FTD use (OS-A), OS with potentially inappropriate FTD use (OS-I), OM with appropriate FTD use (OM-A), and OM with potentially inappropriate FTD use (OM-I).

The Evaluation Criteria to Identify Potentially Inappropriate FTD Use

We developed evaluation criteria to identify potentially inappropriate FTD use for noncancer pain, referring to the recommendations for safe use of prescription opioids by Korea Ministry of Food and Drug Safety and a literature review, which included the US CDC and Canadian guidelines [20,31-36]. Patients who received FTD once were evaluated against 2 criteria and those with multiple prescriptions were assessed using 5 criteria (Table 1). Compliance with all evaluation criteria indicated appropriate FTD use, whereas deviations from any criteria were indicative of potentially inappropriate FTD use.

The duration of FTD use per patch was considered to be 3 days, and cases where the period of each prescription was 90 days or less (ie, 30 patches or less) were evaluated as compliance with the recommendations. For patients with 2 or more prescriptions, even if the total consecutive use exceeded 90 days, it was deemed that the recommendations were followed as long as each prescription was within 90 days. To assess the dose for all types of opioids co-prescribed with FTD, we unified the doses of prescription opioids into MME. For FTDs, the MME per patch was calculated by multiplying the dose (mcg/h), the MME conversion factor (2.4), and the approved use period per patch (3 d). The MME/day was then calculated by dividing the MME by the longest prescription days among the co-prescribed opioids. For patients with multiple prescriptions, the estimated MME/day (eMME/day) was calculated considering the interval between visits. To evaluate the overlap in the prescription period covered by 2 consecutive prescriptions, patients were allowed to visit the health care institution up to 3 days earlier than scheduled. Flexibility in annual prescription periods was allowed based on national health insurance policies in South Korea. Overlapping prescriptions of up to 30 days within a 6-month period were permitted. Flexibility in prescription periods was granted with a 16% allowance for annual prescription periods, from the first visit to the last visit each year, of less than 182 days (6 mo), and a 30-day allowance for periods exceeding 182 days. Details of the evaluation method are provided in Multimedia Appendix 2: Methodology of evaluating FTD use using the established criteria.

Table . The established evaluation criteria for identifying potentially inappropriate use of fentanyl transdermal patch in patients with noncancer pain.

Criteria	Recommendations
For patients with a single FTD[a] prescription per year	
Period	• Prescribed for less than 90 days per prescription
Dose[b]	• Prescribed at a dose of 90 MME[c]/day or less
For patients with multiple FTD prescriptions per year	
Period	• Prescribed for less than 90 days per prescription
Dose[b]	• Prescribed at a dose of 90 eMME[d]/day or less
Overlapping	• Allowed up to 3 days of overlapping prescriptions for each prescription period • Allowed up to 16% (maximum 30 days) of overlapping prescriptions for the annual prescription period
Multi-institutions	• Prescribed by no more than 3 health care institutions in 1 year

[a]FTD: fentanyl transdermal patch.

[b]Dose was calculated by incorporating all types of opioids co-prescribed with FTD

[c]MME: morphine milligram equivalent.

[d]eMME:estimated morphine milligram equivalent.

Statistical Analysis

Descriptive statistics were used to assess the clinical characteristics of the participants included in this study by group. The continuous variables, such as the number of reimbursements per patient, MME/day, annual FTD usage, and annual prescription opioid usage, were represented as means with SDs. We performed the Shapiro-Wilk test for normality and the Levene test for homogeneity of variance. The statistical methods used were ANOVA with the Tukey HSD for data satisfying both normality and homogeneity and the Kruskal-Wallis H test with the Dunn test for data violating normality. The categorical variables, including sex, age group, and diagnoses, were calculated as percentages or frequencies, and statistical

XSL•FO
RenderX

significance was evaluated using χ^2 test [37]. Additionally, a modified Lorenz curve was used to compare the working-age and older adult groups to evaluate the difference in FTD use [38]. In each group, modified Lorenz curves were drawn for FTD use from the top 1% to 100% of the 4 groups divided according to number of prescriptions and evaluation results. To determine the equality of FTD use across groups, we also plotted a classic Lorenz curve and calculated the Gini coefficient [39]. All data processing and statistical analysis were performed using R software (version 3.5.1; R Project for Statistical Computing).

Ethical Considerations

This study followed the principles of the Declaration of Helsinki and was approved by the Ethics Committee of Hanyang University (HYUIRB-202307-008), and the informed consent was waived off by the review boards due to the nature of this research. All data in the manuscript and supplementary materials were anonymized in accordance with ethical standards, ensuring no personally identifiable information could be discerned.

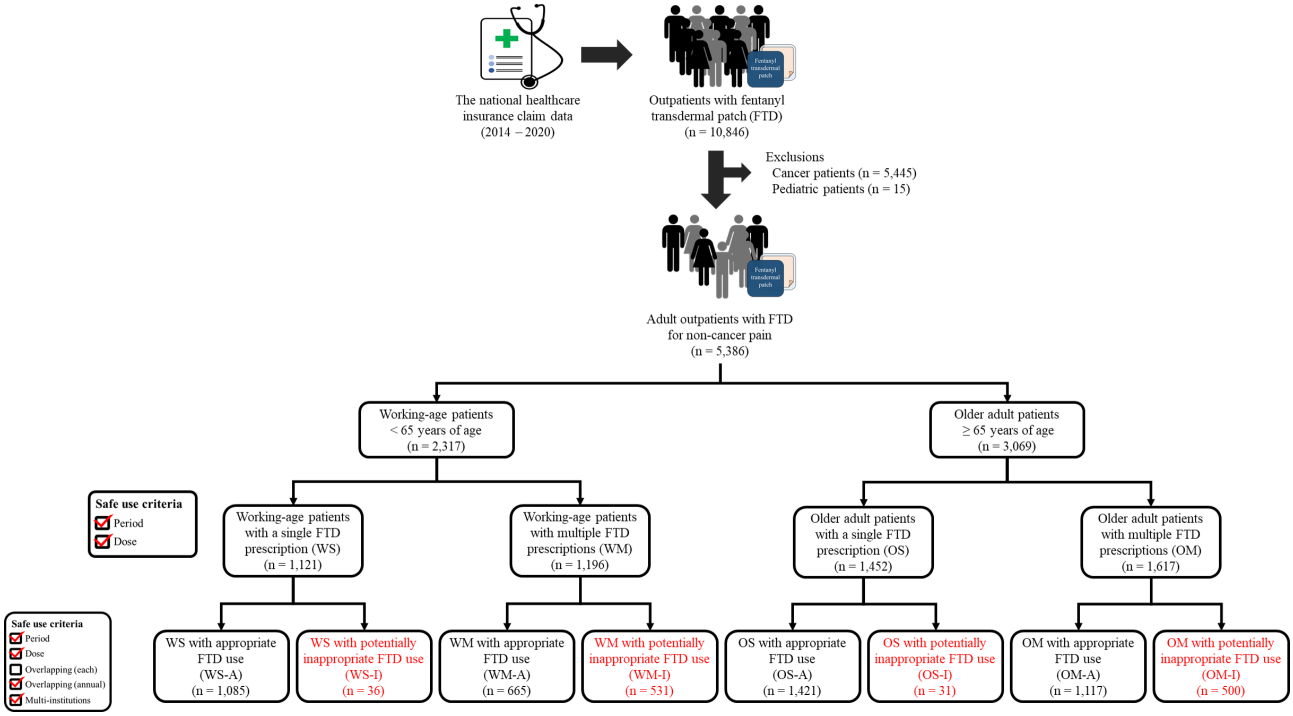
Results

Trend in FTD Use

The insurance claim data of 9,284,003 registrants were evaluated, and 10,846 of these patients used FTDs in an outpatient setting. Following the exclusion of patients with diagnoses for cancer and pediatric patients, 5386 patients and their 19,800 reimbursements were included in the study. The study schematic diagram is shown in Figure 1.

The trends in the use of FTD are shown in Table 2. The number of patients with FTDs for noncancer pain increased from 58.6 per 100,000 registrants in 2014 to 61.3 per 100,000 registrants in 2018 (+4.6% over 4 years), and then decreased to 53.7 per 100,000 registrants in 2020 (–12.6% over 2 years). Meanwhile, the number of reimbursements increased by 7.4%, from 203.2 per 100,000 registrants in 2014 to 218.1 per 100,000 registrants in 2020, and by 1.4% in annual percent change (APC) terms. It indicated an increase in the number of FTD prescriptions per patient. In fact, the number of annual reimbursements per patient increased from 3.5 to 4.1 during the study period, and the average annual FTD usage per patient increased from 3216.7 MME to 4625.7 MME. Approximately 40% (2173/5401) of the study patients were male, which was less than females, but the male proportion steadily increased from 36.0% (21.1/58.6 per 100,000 registrants) in 2014 to 45.1% (24.2/53.7 per 100,000 registrants) in 2020. The proportion of older adult patients aged 65 years or older increased from 55.2% (32.3/58.6 per 100,000 registrants) in 2014 to 59.7% (32.1/53.7 per 100,000 registrants) in 2020.

Figure 1. A schematic diagram of the patient classification process in the present nationwide cross-sectional study for evaluating fentanyl transdermal patch (FTD) use for noncancer pain in South Korea from 2014 to 2020. Patients were categorized into 8 groups by annual prescription frequency and compliance with evaluation criteria. OM: older adult patients with multiple FTD prescriptions; OM-A: older adult patients with multiple FTD prescriptions with appropriate FTD use; OM-I: older adult patients with multiple FTD prescriptions with potentially inappropriate FTD use; OS: older adult patients with a single FTD prescription; OS-A: older adult patients with a single FTD prescription with appropriate FTD use; OS-I: older adult patients with a single FTD prescription with potentially inappropriate FTD use; WM: working-age patients with multiple FTD prescriptions; WM-A: working-age patients with multiple FTD prescriptions with appropriate FTD use; WM-I: working-age patients with multiple FTD prescriptions with potentially inappropriate FTD use; WS: working-age patients with a single FTD prescription; WS-A: working-age patients with a single FTD prescription with appropriate FTD use; WS-I: working-age patients with a single FTD prescription with potentially inappropriate FTD use.



XSL•FO
RenderX

Table . Annual trends in the use of fentanyl transdermal patches among patients with noncancer pain in South Korea, based on nationwide claim data from 2014 to 2020.

	Calendar year						
	2014	2015	2016	2017	2018	2019	2020
Number of patients, n per 100,000 registrants	58.6	57.4	56.9	61.0	61.3	54.6	53.7
Gender, n per 100,000 registrants (%)							
Male	21.1 (36.0)	21.8 (37.9)	22.8 (40.1)	25.2 (41.3)	25.4 (41.4)	22.6 (41.4)	24.2 (45.1)
Female	37.5 (64.0)	35.6 (62.1)	34.1 (59.9)	35.8 (58.7)	35.9 (58.6)	32.0 (58.6)	29.5 (54.9)
Age, n per 100,000 registrants (%)							
Working age	26.3 (44.8)	25.0 (43.5)	24.7 (43.4)	26.4 (43.2)	26.5 (43.3)	22.2 (40.7)	21.6 (40.3)
Older adult	32.3 (55.2)	32.4 (56.5)	32.2 (56.6)	34.6 (56.8)	34.8 (56.7)	32.4 (59.3)	32.1 (59.7)
Number of reimbursements, n per 100,000 registrants	203.2	201.8	213.3	222.5	219.6	216.9	218.1
Number of reimbursements, n per patient[a]	3.5 (5.1)	3.5 (4.1)	3.8 (4.8)	3.6 (4.8)	3.6 (5.2)	4.0 (5.3)	4.1 (6.1)
Average annual FTD usage, MME[b]	3216.7 (6063.2)	3664.4 (6564.4)	3840.9 (7240.9)	4480.4 (19,917.1)	4189.0 (10,968.7)	4361.9 (8281.8)	4625.7 (7478.1)

[a]The number of reimbursements per patient and annual FTD usage are presented as mean (SD). The percentages are calculated based on the number of patients and reimbursements calculated in each year.

[b]MME: morphine milligram equivalents.

Patterns of FTD Use by Age Group

Patients were classified into working-age and older adult groups, comprising 2317 and 3069 patients (25.0 and 33.1 per 100,000 registrants), respectively. The male proportion was 50.6% (1172/2317) in working-age group, while 32.3% (990/3069) of older adult patients were male. There was a statistically significant difference between the 2 groups (P<.001). The working-age group had an average of 3.9 reimbursements per year, with an average MME/day of 64.4 for each reimbursement. The older adult group had an average of 3.5 reimbursements per year, and their average MME/day was 47.9, which was significantly lower than that of the working-age group (P=.005). The most common morbidity in both groups was diseases of the musculoskeletal system and connective tissue, accounting for 50.0% (4539/9075) in the working-age group and 61.0% (6537/10,725) in the older adult group. In the working-age group, the FTD prescriptions for the diseases of the circulatory system were less common (2.4% vs 5.5%), but for the diseases of the nervous system and diseases of the digestive system, the FTD prescriptions were more frequent compared to the older adult group (8.6% vs 6.8%; 4.2% vs 1.6%, respectively).

Dividing the groups by the number of FTD prescriptions, the WM had an average of 6.7 reimbursements, and the average annual total amount of FTDs used was approximately 18.7 times greater than the total for the WS (8938.9 vs. 477.5 MME). The OM showed a smaller difference: an average of 5.7 reimbursements and 10.2 times more FTD total usage than OS. Calculating MME/day for each prescription, WS, WM, OS, and OM were prescribed an average dose of 39.7, 67.9, 38.3, and 49.4 MME/day, respectively (Table 3).

XSL•FO
RenderX

Table . The clinical characteristics of 4 groups stratified by age and prescription frequency among patients with noncancer pain prescribed fentanyl transdermal patch in South Korea from 2014 to 2020.

	Working-age group			Older adult group			*P* value[a]
	WS	WM	Total	OS	OM	Total	
Number of patients, n	1121	1196	2317	1452	1617	3069	-
Gender, n (%)[b]							<.001
Male	539 (48.1)	633 (52.9)	1172 (50.6)	418 (28.8)	572 (35.4)	990 (32.3)	
Female	582 (51.9)	563 (47.1)	1145 (49.4)	1034 (71.2)	1045 (64.6)	2079 (67.7)	
Number of reimbursements, n	1121	7954	9075	1452	9273	10,725	-
Number of reimbursements, n per patient[c]	1.0 (0.0)	6.7 (6.8)	3.9 (5.7)	1.0 (0.0)	5.7 (5.2)	3.5 (4.4)	.005
Annual usage, MME[d]							
FTD	477 (564)	8939 (20,247)	4845 (15,151)	580 (562)	5890 (7320)	3378 (5950)	<.001
Prescription opioids	555 (824)	12,043 (24,828)	6485 (18,744)	651 (846)	7380 (14,401)	4196 (10,994)	<.001
Prescription opioid usage per reimbursement, MME/day	39.7 (24.9)	67.9 (53.5)	64.4 (51.7)	38.3 (23.7)	49.4 (37.2)	47.9 (35.9)	<.001
Primary diagnosis, n (%)[b]							<.001
Diseases of the musculoskeletal system and connective tissue	672 (59.9)	3867 (48.6)	4539 (50.0)	859 (59.2)	5678 (61.2)	6537 (61.0)	
Injury, poisoning and certain other consequences of external causes	201 (17.9)	719 (9.0)	920 (10.1)	280 (19.3)	833 (9.0)	1113 (10.4)	
Diseases of the nervous system	35 (3.1)	748 (9.4)	783 (8.6)	45 (3.1)	687 (7.4)	732 (6.8)	
Diseases of the digestive system	30 (2.7)	352 (4.4)	382 (4.2)	24 (1.7)	148 (1.6)	172 (1.6)	
Diseases of the circulatory system	23 (2.1)	194 (2.4)	217 (2.4)	59 (4.1)	530 (5.7)	589 (5.5)	
Others	160 (14.3)	2074 (26.2)	2234 (24.7)	185 (12.6)	1397 (15.1)	1582 (14.7)	

[a]*P* value means the comparison between the total working-age and older adult groups.

[b]The percentages are calculated based on the number of patients and reimbursements calculated in each group.

[c]The number of reimbursements per patient, annual usage and prescription opioid usage per reimbursement are presented as mean (SD)

[d]MME: morphine milligram equivalent

Identification of Potentially Inappropriate FTD Use

Potentially inappropriate FTD use was assessed in 24.5% (567/2317) of working-age patients and 17.3% (531/3069) older adult patients. Among them, 67 were prescribed FTD once (WS-I, 36; OS-I, 31) and 1031 used FTDs more than twice (WM-I, 531; OM-I, 500). The proportion of male patients in WM-I was 56.7% (301/531), meaning that while 47.6% (301/633) of male patients in the WM were evaluated as having potentially inappropriate FTD use, only 40.9% (230/563) of

XSL•FO
RenderX

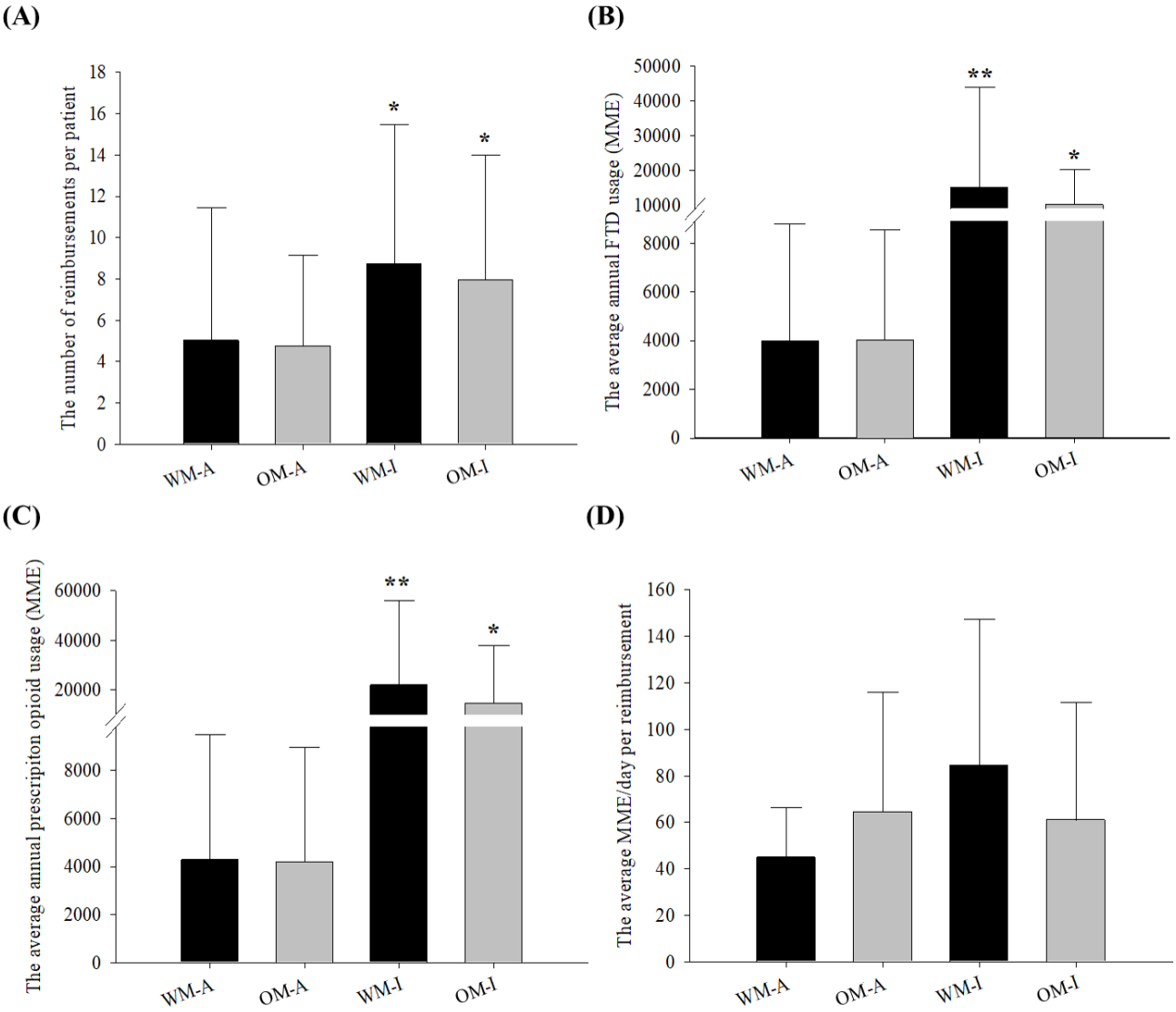
females were potentially inappropriate FTD users. In older adult patients with multiple FTD prescriptions, 36.5% (209/572) and 27.8% (291/1045) of males and females, respectively, were assessed as potentially inappropriate FTD users, which was lower than that of the WM.

The proportions of patients identified as potentially inappropriate FTD users for each criterion. For 2 single-prescription groups, high doses of prescriptions were the main reason for being evaluated as potentially inappropriate: 3.1% (35/1121) of WS and 2.1% (30/1421) of OS. The primary reason for potentially inappropriate evaluation in the 2 multiple-prescription groups was overlapping of each prescription (33.6% of WM; 402/1196 and 25.8% of OM; 417/1617), followed by high-dose prescriptions and overlapping of annual prescriptions (17.8% of WM; 213/1196 and 9.8% of OM; 159/1617). Regarding evaluation criterion for dose, the proportion showed the greatest difference at 24.2% (290/1196) for working-age patients and 11.7% (189/1617) for older adult patients. Patients with multiple health care institution visits and those who received long-term treatment were both less than 1% in each group.

Figure 2 presents a comparative analysis of characteristics related to FTD and prescription opioid usage across 4 multiple-prescription groups. As shown in **Figure 2A**, both the WM-I and OM-I groups had significantly more numerous reimbursements per patient than the corresponding group with appropriate FTD use (WM-I, 8.7 vs WM-A, 5.0; OM-I, 7.9 vs OM-A 4.8) but there were no significant differences between the WM-I and OM-I. Similar trends were observed for the average annual FTD usage and prescription opioid usage (**Figure 2B and 2C**). The WM-I had a significantly higher average annual FTD usage (3.8 times; $P<0.001$) and prescription opioid usage (5.1 times; $P<.001$) than those of the WM-A. The OM-I also had significantly higher average annual FTD use and prescription opioid use compared to the OM-A ($P<.001$). In addition to the comparison with potentially inappropriate users, the WM-I had significantly higher average annual FTD and prescription opioid usage than the OM-I ($P<.001$). For the average MME/day per reimbursement, the WM-I showed significantly greater MME/day than the WM-A, while the MME/day of OM-I was lower than that of OM-A, although there was no significant difference (**Figure 2D**). Comparing the eMME/day for multiple-prescription groups to MME/day, the appropriate groups showed a lower eMME/day than their average MME/day (WM-A, 32.7 eMME/day vs 44.9 MME/day; OM-A, 29.3 eMME/day vs 64.4 MME/day), while potentially inappropriate groups showed a higher value of eMME/day than their MME/day (WM-I, 149.4 eMME/day vs 84.4 MME/day; OM-I, 85.6 eMME/day vs 60.9 MME/day). Among inappropriate groups, WM-I had significantly higher eMME/day than OM-I ($P<.001$).

XSL•FO
RenderX

Figure 2. Comparison of the characteristics of 4 groups with multiple fentanyl transdermal patch prescriptions, stratified by age and evaluation outcome, in a nationwide cross-sectional study of patients with noncancer pain in South Korea using claim data from 2014 to 2020. Panel A shows the number of reimbursements per patient; Panel B shows the average annual FTD usage; Panel C shows the average annual prescription opioid usage; and Panel D shows the average MME/day per reimbursement. Black bars represent working-age patients and the older adult patients are represented by gray bars. *Means $P<.001$ compared to the corresponding group with appropriate FTD use. **Means $P<.001$ compared to the other groups. MME: morphine milligram equivalent; OM-A: older adult patients with multiple FTD prescriptions with appropriate FTD use; OM-I: older adult patients with multiple FTD prescriptions with potentially inappropriate FTD use; WM-A: working-age patients with multiple FTD prescriptions with appropriate FTD use; WM-I: working-age patients with multiple FTD prescriptions with potentially inappropriate FTD use.

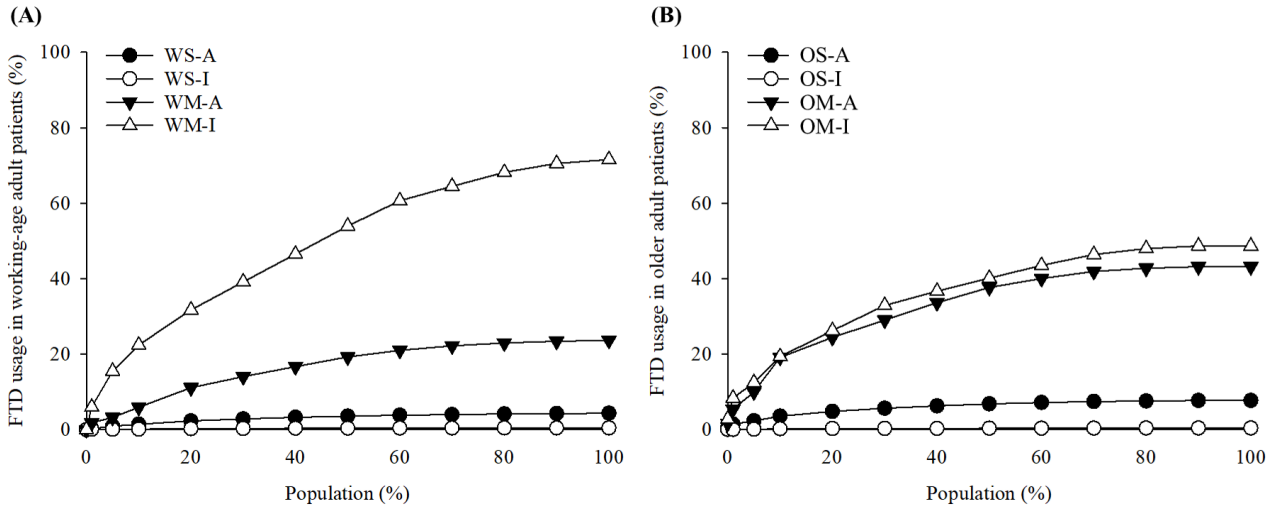


Inequality Assessment of FTD Use

The total amount of FTDs used in working-age patients and older adult patients was 11,226,138 MME and 10,365,939 MME, respectively. To confirm the distribution of FTD usage for each group, a modified Lorenz curve was plotted against the total amount of FTDs (Figure 3). The WM-I accounted for more than 70% (8,034,172/11,226,138) of the total FTDs used in working-age patients. In particular, the amount of FTDs used by the top 5% of the WM-I was 3.6 times the amount of FTDs used by patients in the top 5% of the other 3 groups. With only 36 patients in the WS-I, the use in this group was only 0.44% of that in the working-age patients. In the older adult patients, the OM-I also used the highest proportion of FTDs, but unlike the working-age group, there was only a small difference between the OM-I and OM-A (48.7%; 5,045,751/10,365,939 vs 43.2%; 4,477,804/10,365,939). The OS-I also had a very low portion of 0.41%, as there were only 31 patients in the OS-I. Moreover, Gini coefficients were calculated for the classic Lorenz curves of all groups. The working-age patients had a greater Gini coefficient than the older adult group, indicating unequal use of FTDs (0.461 vs 0.406). Each of the working-age groups showed greater inequality than the corresponding older adult groups.

XSL•FO
RenderX

Figure 3. A modified Lorenz curve illustrating the distribution of fentanyl transdermal patch (FTD) usage among patients with noncancer pain in South Korea from 2014 to 2020, stratified by evaluation outcome and prescription frequency. Panel A presents data for the working-age group, and Panel B presents data for older adult group. OM-A: older adult patients with multiple FTD prescriptions with appropriate FTD use; OM-I: older adult patients with multiple FTD prescriptions with potentially inappropriate FTD use; OS-A: older adult patients with a single FTD prescription with appropriate FTD use; OS-I: older adult patients with a single FTD prescription with potentially inappropriate FTD use; WM-A: working-age patients with multiple FTD prescriptions with appropriate FTD use; WM-I: working-age patients with multiple FTD prescriptions with potentially inappropriate FTD use; WS-A: working-age patients with a single FTD prescription with appropriate FTD use; WS-I: working-age patients with a single FTD prescription with potentially inappropriate FTD use.



Discussion

Principal Findings and Comparison With Previous Works

This study is the first study to assess the rational use of FTDs with recommendations for safe use of prescription opioids using a nationwide database. This study identified 567 (24.5%) working-age patients and 531 (17.3%) older adult patients as potentially inappropriate FTD users over a 7-year period.

The number of FTD users for noncancer pain in South Korea exhibited an increase from 2014 to 2018, followed by a subsequent decline from 2019, resulting in an overall reduction of 8.4% during the study period. The decline in the number of patients prescribed FTDs in 2018 coincides with the implementation of the NIMS in South Korea. The NIMS is a system designed to monitor and report on the manufacturing, distribution, prescribing, and dispensing of all controlled pharmaceutical drugs in the country [15]. It is intended to serve some of the functions of the prescription drug monitoring program in the United States, potentially contributing to the decrease in unnecessary use of FTDs. However, the number of FTD prescriptions per 1000 registrants increased over the study period. The increase rate found in this study is consistent with the APC of 1.8% for all types of opioid prescriptions in South Korea from 2013 to 2019 [7]. The trend in the number of FTD prescriptions per patient, rather than per population, is more rapidly increasing, with an APC of 2.6%. The average annual amount of FTD usage per patient increased at an APC of 6.9%, suggesting that the increase in FTD usage amount is driven not only by an increase in the number of FTD prescriptions per patient but also by an increase in the FTD usage per prescription. These changes in the number of patients and FTD prescribing trends were also supported by the results that the decrease in the number of patients with a single FTD prescription and the increase in the number of patients with multiple FTD prescriptions. Additionally, the increase in FTD usage (total, 13.6 times; working age, 18.7 times; older adult, 10.2 times) surpassed the increase in the average number of reimbursements per patient (total, 6.1 times; working age, 6.7 times; older adult, 5.7 times) between the 2 groups. According to a previous study, prescription opioid usage in South Korea increased at an APC of 30.3% from 2015 to 2019, while in the United States and Canada, it declined during the same period by 14.1% and 11.6% per year, respectively [12]. Another study reported that the amount of fentanyl and its derivatives used per patient in the United States decreased by 65.5% from 2010 to 2019 [40]. Despite the lower population-based prescription opioid usage in South Korea compared to global levels in 2019 [13], approximately 20% of patients using FTDs for noncancer pain were deemed at risk for potentially inappropriate use of prescription opioids, including FTDs.

We developed evaluation criteria based on practice guidelines from different countries and endeavored to reflect what clinicians might encounter in the real world. For instance, the prescription period was assessed only in terms of the number of days covered by each prescription, regardless of the duration of consecutive use. It is based on the assumption that clinicians evaluate the necessity of FTD use for each patient during the re-prescription process. Overlapping prescriptions were evaluated by setting a margin of flexibility, allowing 3 days per prescription and 16% per year, considering the practical problems in the real world. For patients with multiple FTD prescriptions, eMME/day was calculated and used. The MME/day is the patient's prescribed opioid dose divided by the estimated treatment duration based on the authorized dosage, while the eMME/day is the prescription opioid dose divided by the time until the next visit, assuming the patient has taken all of the prescribed opioids prior to the visit. If the patient is compliant with a prescribed regimen, the MME/day and

XSL•FO
RenderX

eMME/day should be equal. The eMME/day may be higher than MME/day for patients who visit earlier than scheduled, and eMME/day may be lower than MME/day for those who visit later than scheduled.

This study showed that a higher proportion of working-age patients was identified as potentially inappropriate FTD users than older adult patients (24.5% vs 17.3%). Comparing the 2 groups with the multiple FTD prescriptions, 44.4% of the working-age patients were evaluated as having potentially inappropriate FTD use, while only 30.9% of the older adult patients were classified into the OM-I. Although the numbers of patients in the WM-I and OM-I were similar, the total annual amount of FTD use was 1.6 times higher in the WM-I than in the OM-I. In addition, there were more patients evaluated in WM-I than in OM-I as potentially inappropriate FTD users due to high doses. Compared to the corresponding groups with appropriate FTD use, the WM-I used 3.0 times that of the WM-A patients, while the OM-I used only 1.1 times that of the OM-A. These findings suggest that the WM-I patients frequently used FTDs outside practice guidelines more than the OM-I patients. The comparison of Gini coefficients calculated by plotting the classic Lorenz curve also revealed that the FTD use was relatively more equally distributed in the older adult patients than in the working-age patients. Additionally, in both WM-A and OM-A, MME/day was higher than the newly calculated eMME/day, but in both WM-I and OM-I, eMME/day was higher than MME/day. This suggests that patients who used FTD appropriately may use them longer than prescribed periods, while the patients having a risk of potentially inappropriate FTD use may take a higher than prescribed dose of FTDs or hoard spare FTDs. The difference from MME/day to eMME/day was greater in the WM-I than in the OM-I. There was a 1.4 times larger increase in OM-I patients (mean difference of 24.7), but a 1.8 times larger increase in WM-I patients (mean difference 64.9). These differences also support a more potentially inappropriate FTD use in working-age patients. These findings are consistent with a result of meta-analysis finding that the risk of opioid overdose was greater in adults aged 25 - 54 than in the older adult [28]. Furthermore, death rates from prescription opioid overdoses were also reported to be higher in younger patients than in older patients [41]. Prescription opioid-related death in young adults might have a significant impact in terms of years of life lost, increasing the public health burden [42].

The strength of this study is that we focused on the potentially inappropriate FTD use, which has recently been assessed as having the highest risk of misuse among prescription opioids. Moreover, this study used national claim data to investigate changes in FTD use for noncancer pain and assessments of potentially inappropriate FTD use simultaneously, which resulted in more consistent results.

Limitations

This study has several limitations. First, it was a cross-sectional analysis based on health care insurance claims data, which means it only included data for which FTD claims were submitted by health care institutions. Consequently, data regarding nonclaimed use of FTDs were not captured in this study. Second, despite several adjustments having been made to the evaluation criteria, there remains a potential risk that patients who are deemed by clinicians to require high doses or long-term use of FTDs for pain management may be classified as inappropriate users. Third, the FTD usage was assessed solely based on prescriptions, without verification of actual patient consumption. This limitation suggests that patient adherence may have influenced our results.

Conclusions

In South Korea, although the number of patients receiving FTD prescriptions has decreased, the total amount of FTDs consumed has increased in recent years, suggesting that caution is warranted. The present study demonstrated the potential for inappropriate FTD use in working-age patients compared with older adult patients. While our findings don't imply that FTDs should be restricted to those patients who desperately need them, health care providers should be alert to the risk of prescription opioids and the opioid epidemic and promote the appropriate use of prescription opioids. Further research is needed to quantify and qualify the risk factors in patients having a risk of potentially inappropriate use, considering the clinical rationale associated with prescribing FTDs.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2023 - 00217123 and 2022R1F1A1075439).

Data Availability

The datasets used and/or analyzed during the present study are available from the Health Insurance Review and Assessment on reasonable request. The authors are not authorized to share the raw data voluntarily.

Authors' Contributions

JEC involved in the conceptualization of the study; SC and SYC involved in data curation; formal analysis was done by SC; funding acquisition was done by JEC; SC, SYC, KRK, HGC, and JEC contributed to investigation; HGC and JEC are responsible for methodology and project administration; resources were arranged by HGC and JEC; SC, SYC, and KRK are responsible for software; supervision was done by JEC; validation was done by SC and JEC; visualization was done by SC and KRK; SC contributed to writing – original draft; JEC contributed to writing – review and editing.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Checklist.
[PDF File, 139 KB - publichealth_v11i1e63960_app1.pdf]

Multimedia Appendix 2
Methodology of evaluating FTD use using the established criteria: (A), calculation of MME/day for FTDs; (B), MME/day calculation for co-prescribed opioids; (C), eMME/day calculation in a potentially inappropriate FTD use case; (D), eMME/day calculation in an appropriate FTD use case; (E), overlapping evaluation (annual).
[PNG File, 230 KB - publichealth_v11i1e63960_app2.png]

References

1. Dennen A C, Blum K, Braverman R E, et al. How to combat the global opioid crisis. CPQ Neurol Psychol 2023;5(4):93. [Medline: 36812107]
2. Gottschalk M. The opioid crisis: the war on drugs is over. Long live the war on drugs. Annu Rev Criminol 2023 Jan 27;6(1):363-398 [FREE Full text] [doi: 10.1146/annurev-criminol-030421-040140]
3. Alpert A, Evans WN, Lieber EMJ, Powell D. Origins of the opioid crisis and its enduring impacts. Q J Econ 2022 May;137(2):1139-1179. [doi: 10.1093/qje/qjab043] [Medline: 35832727]
4. Bonnie RJ, Kesselheim AS, Clark DJ. Both urgency and balance needed in addressing opioid epidemic: a report from the National Academies of Sciences, Engineering, and Medicine. JAMA 2017 Aug 1;318(5):423-424. [doi: 10.1001/jama.2017.10046] [Medline: 28715551]
5. Stoicea N, Costa A, Periel L, Uribe A, Weaver T, Bergese SD. Current perspectives on the opioid crisis in the US healthcare system: a comprehensive literature review. Medicine (Baltimore) 2019 May;98(20):e15425. [doi: 10.1097/MD.0000000000015425] [Medline: 31096439]
6. Wei YJJ, Chen C, Fillingim R, Schmidt SO, Winterstein AG. Trends in prescription opioid use and dose trajectories before opioid use disorder or overdose in US adults from 2006 to 2016: a cross-sectional study. PLoS Med 2019 Nov;16(11):e1002941. [doi: 10.1371/journal.pmed.1002941] [Medline: 31689302]
7. Cho NR, Chang YJ, Lee D, Kim JR, Ko DS, Choi JJ. Trends in opioid prescribing practices in South Korea, 2009–2019: are we safe from an opioid epidemic? PLoS ONE 2021;16(5):e0250972. [doi: 10.1371/journal.pone.0250972] [Medline: 33979378]
8. Wang JJ, Chu YR, Teng SF, Chu CC, Ho CH, Chu LL. Prevalence of opioid prescriptions in Taiwan (2008-2018). J Chin Med Assoc 2022 May 1;85(5):603-609. [doi: 10.1097/JCMA.0000000000000720] [Medline: 35353736]
9. Hurtado I, García-Sempere A, Peiró S, Sanfélix-Gimeno G. Increasing trends in opioid use from 2010 to 2018 in the Region of Valencia, Spain: a real-world, population-based study. Front Pharmacol 2020;11:612556. [doi: 10.3389/fphar.2020.612556] [Medline: 33362564]
10. Guillou-Landreat M, Quinio B, Le Reste JY, et al. Analgesic opioid misuse and opioid use disorder among patients with chronic non-cancer pain and prescribed opioids in a pain centre in France. Int J Environ Res Public Health 2021 Feb 21;18(4):2097. [doi: 10.3390/ijerph18042097] [Medline: 33670004]
11. Jayawardana S, Forman R, Johnston-Webber C, et al. Global consumption of prescription opioid analgesics between 2009-2019: a country-level observational study. EClinicalMedicine 2021 Dec;42:101198. [doi: 10.1016/j.eclinm.2021.101198] [Medline: 34820610]
12. Robert M, Jouanjus E, Khouri C, Fouilhé Sam-Laï N, Revol B. The opioid epidemic: a worldwide exploratory study using the WHO pharmacovigilance database. Addiction 2023 Apr;118(4):771-775. [doi: 10.1111/add.16081] [Medline: 36331523]
13. Ju C, Wei L, Man KKC, et al. Global, regional, and national trends in opioid analgesic consumption from 2015 to 2019: a longitudinal study. Lancet Public Health 2022 Apr;7(4):e335-e346. [doi: 10.1016/S2468-2667(22)00013-5] [Medline: 35366408]
14. Kim J, Shin SJ, Yoon J, et al. Recent trends in opioid prescriptions in Korea from 2002 to 2015 based on the Korean NHIS-NSC cohort. Epidemiol Health 2022;44:e2022029. [doi: 10.4178/epih.e2022029] [Medline: 35209705]
15. Lee IH, Kim SY, Park S, Ryu JG, Je NK. Impact of the narcotics information management system on opioid use among outpatients with musculoskeletal and connective tissue disorders: quasi-experimental study using interrupted time series. JMIR Public Health Surveill 2024 Feb 21;10:e47130. [doi: 10.2196/47130] [Medline: 38381481]
16. Noh Y, Heo KN, Yu YM, Lee JY, Ah YM. Trends in potentially inappropriate opioid prescribing and associated risk factors among Korean noncancer patients prescribed non-injectable opioid analgesics. Ther Adv Drug Saf 2022;13:20420986221091001. [doi: 10.1177/20420986221091001] [Medline: 35509350]

XSL•FO
RenderX

17. Jayawardhana J, Abraham AJ, Perri M. Opioid analgesics in Georgia Medicaid: trends in potential inappropriate prescribing practices by demographic characteristics, 2009-2014. J Manag Care Spec Pharm 2018 Sep;24(9):886-894. [doi: 10.18553/jmcp.2018.24.9.886] [Medline: 30156455]
18. Kim S, Suh HS. A population-based study on the risk of prescription opioid abuse in patients with chronic opioid use and cost-effectiveness of prescription drug monitoring program using a patient simulation model in South Korea. Int J Drug Policy 2023 Feb;112:103953. [doi: 10.1016/j.drugpo.2023.103953] [Medline: 36645947]
19. Lee JH. Still a threatening opioid wave: it is time for the protagonist to step up!. Korean J Pain 2023 Jan 1;36(1):1-3. [doi: 10.3344/kjp.22403] [Medline: 36581596]
20. The Korean Ministry of Food and Drug Safety (KMFDS). The standards for safe use of prescription opioids. 2021. URL: https://www.mfds.go.kr [accessed 2024-06-02]
21. Khouja T, Tadrous M, Matusiak L, Suda K. Opioid prescribing in United States Health Systems, 2015 to 2019. Value Health 2021 Sep;24(9):1279-1284. [doi: 10.1016/j.jval.2021.04.1274] [Medline: 34452707]
22. Roxburgh A, Hall WD, Gisev N, Degenhardt L. Characteristics and circumstances of heroin and pharmaceutical opioid overdose deaths: comparison across opioids. Drug Alcohol Depend 2019 Dec 1;205:107533. [doi: 10.1016/j.drugalcdep.2019.06.035] [Medline: 31704378]
23. Roxburgh A, Nielsen S. Twenty-year trends in pharmaceutical fentanyl and illicit fentanyl deaths, Australia 2001-2021. Int J Drug Policy 2022 Nov;109:103854. [doi: 10.1016/j.drugpo.2022.103854] [Medline: 36150355]
24. Costantino RC, Gressler LE, Onukwugha E, et al. Initiation of transdermal fentanyl among US commercially insured patients between 2007 and 2015. Pain Med 2020 Oct 1;21(10):2229-2236. [doi: 10.1093/pm/pnaa091] [Medline: 32377671]
25. Daveluy A, Micallef J, Sanchez-Pena P, et al. Ten-year trend of opioid and nonopioid analgesic use in the French adult population. Br J Clin Pharmacol 2021 Feb;87(2):555-564. [doi: 10.1111/bcp.14415] [Medline: 32496599]
26. Rikard SM, Strahan AE, Schmit KM, Guy GP Jr. Chronic pain among adults - United States, 2019-2021. MMWR Morb Mortal Wkly Rep 2023 Apr 14;72(15):379-385. [doi: 10.15585/mmwr.mm7215a1] [Medline: 37053114]
27. Gomes T, Tadrous M, Mamdani MM, Paterson JM, Juurlink DN. The burden of opioid-related mortality in the United States. JAMA Netw Open 2018 Jun 1;1(2):e180217. [doi: 10.1001/jamanetworkopen.2018.0217] [Medline: 30646062]
28. Brady JE, Giglio R, Keyes KM, DiMaggio C, Li G. Risk markers for fatal and non-fatal prescription drug overdose: a meta-analysis. Inj Epidemiol 2017 Dec;4(1):24. [doi: 10.1186/s40621-017-0118-7] [Medline: 28762157]
29. Kim L, Kim JA, Kim S. A guide for the utilization of Health Insurance Review and Assessment Service National Patient Samples. Epidemiol Health 2014;36:e2014008. [doi: 10.4178/epih/e2014008] [Medline: 25078381]
30. Cheon S, Oh SH, Kim JT, Choi HG, Park H, Chung JE. Nutrition therapy by nutrition support team: a comparison of multi-chamber bag and customized parenteral nutrition in hospitalized patients. Nutrients 2023 May 29;15(11):2531. [doi: 10.3390/nu15112531] [Medline: 37299494]
31. Dowell D, Ragan KR, Jones CM, Baldwin GT, Chou R. CDC clinical practice guideline for prescribing opioids for pain - United States, 2022. MMWR Recomm Rep 2022 Nov 4;71(3):1-95. [doi: 10.15585/mmwr.rr7103a1] [Medline: 36327391]
32. Chou R, Hartung D, Turner J, Blazina I, Chan B, Levander X, et al. Opioid Treatments for Chronic Pain [Internet]: Rockville (MD): Agency for Healthcare Research and Quality; 2020. [Medline: 32338848]
33. Busse JW, Craigie S, Juurlink DN, et al. Guideline for opioid therapy and chronic noncancer pain. CMAJ 2017 May 8;189(18):E659-E666. [doi: 10.1503/cmaj.170363] [Medline: 28483845]
34. Healthcare Improvement Scotland. The Scottish intercollegiate guidelines network 50 (SIGN 50): a guideline developer's handbook. 2019. URL: https://www.sign.ac.uk [accessed 2024-06-02]
35. The National Academy of Medicine of Korea. Guidance for Prevention of Opioid Analgesics Abuse. 2021. URL: https://www.namok.or.kr [accessed 2024-06-02]
36. White AG, Birnbaum HG, Schiller M, Tang J, Katz NP. Analytic models to identify patients at risk for prescription opioid abuse. Am J Manag Care 2009 Dec;15(12):897-906. [Medline: 20001171]
37. Hailemeskel Abebe T. The derivation and choice of appropriate test statistic (Z, t, F and Chi-square test) in research methodology. ML 2019;5(3):33. [doi: 10.11648/j.ml.20190503.11]
38. Birt J, Johnston J, Nelson D. Exploration of claims-based utilization measures for detecting potential nonmedical use of prescription drugs. J Manag Care Spec Pharm 2014 Jun;20(6):639-646. [doi: 10.18553/jmcp.2014.20.6.639] [Medline: 24856602]
39. Sitthiyot T, Holasut K. A simple method for measuring inequality. Palgrave Commun 2020;6(1). [doi: 10.1057/s41599-020-0484-6]
40. Stemrich RA, Weber JV, McCall KL, Piper BJ. Pronounced declines in dispensed licit fentanyl, but not fentanyl derivatives. Res Social Adm Pharm 2022 Jun;18(6):3046-3051. [doi: 10.1016/j.sapharm.2021.08.001] [Medline: 34419378]
41. Scholl L, Seth P, Kariisa M, Wilson N, Baldwin G. Drug and opioid-involved overdose deaths — United States, 2013–2017. MMWR Morb Mortal Wkly Rep 2019;67(5152):1419-1427. [doi: 10.15585/mmwr.mm675152e1]
42. Castaldelli-Maia JM, Wang YP, Brunoni AR, et al. Burden of disease due to amphetamines, cannabis, cocaine, and opioid use disorders in South America, 1990-2019: a systematic analysis of the Global Burden of Disease Study 2019. Lancet Psychiatry 2023 Feb;10(2):85-97. [doi: 10.1016/S2215-0366(22)00339-X] [Medline: 36697127]

XSL•FO
RenderX

Abbreviations

APC: annual percent change
CDC: Centers for Disease Control and Prevention
eMME: estimated morphine milligram equivalents
FTD: fentanyl transdermal patch
HIRA: Health Insurance Review and Assessment Service
MME: morphine milligram equivalent per day
NIMS: Narcotics Information Management System
NPS: National Patient Sample
OM: older adult patients with multiple FTD prescriptions
OM-A: older adult patients with multiple FTD prescriptions with appropriate FTD use
OM-I: older adult patients with multiple FTD prescriptions with potentially inappropriate FTD use
OS: older adult patients with a single FTD prescription
OS-A: older adult patients with a single FTD prescription with appropriate FTD use
OS-I: older adult patients with a single FTD prescription with potentially inappropriate FTD use
OUD: opioid use disorder
PDMP: prescription drug monitoring program
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WM: working-age patients with multiple FTD prescriptions
WM-A: working-age patients with multiple FTD prescriptions with appropriate FTD use
WM-I: working-age patients with multiple FTD prescriptions with potentially inappropriate FTD use
WS: working-age patients with a single FTD prescription
WS-A: working-age patients with a single FTD prescription with appropriate FTD use
WS-I: working-age patients with a single FTD prescription with potentially inappropriate FTD use

Edited by A Mavragani, X Li; submitted 04.07.24; peer-reviewed by I Mircheva, K Mckelvey; revised version received 09.04.25; accepted 11.04.25; published 28.05.25.

Please cite as:
Cheon S, Choi SY, Kim KR, Choi HG, Chung JE
Potentially Inappropriate Use of Transdermal Fentanyl in Working-Age and Older Adult Populations with Non-Cancer Pain: Nationwide Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e63960
URL: https://publichealth.jmir.org/2025/1/e63960
doi:10.2196/63960

© Seunghyun Cheon, Seo-Yong Choi, Kyu Ri Kim, Han-Gon Choi, Jee-Eun Chung. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.5.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Impact of Point-of-Care Testing for Influenza on Antimicrobial Stewardship in UK Primary Care: Nested Cohort Study

Uy Hoang[1], MD; Jessica Smylie[1], BA; Elizabeth Button[1], MSc; Jack Macartney[1], BSc; Cecilia Okusi[1], MSc; Rachel Byford[1], BA; Filipa Ferreira[1], PhD; Charis Xie[2], MSc; Mark Joy[1], PhD; Tristan Clark[3], MD; Simon de Lusignan[1,4], MBBS

[1]Nuffield Department of Primary Care Health Sciences, Medical Sciences Division, University of Oxford, Gibson Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, United Kingdom

[2]Barts and The London School of Medicine and Dentistry, Wolfson Institute of Population Health, Queen Mary University of London, London, United Kingdom

[3]Faculty of Medicine, School of Clinical and Experimental Sciences, University of Southampton, Southampton, United Kingdom

[4]Royal College of General Practitioners, London, United Kingdom

Corresponding Author:
Simon de Lusignan, MBBS
Nuffield Department of Primary Care Health Sciences, Medical Sciences Division, University of Oxford, Gibson Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, United Kingdom

Abstract

Background: Influenza and respiratory syncytial virus (RSV) predominantly circulate during the winter season and cause acute respiratory illness (ARI). Deploying molecular point-of-care testing (POCT) in primary care can inform whether a patient presenting with an ARI has influenza or RSV. An early virological diagnosis could facilitate appropriate use of antivirals and enable better antimicrobial stewardship.

Objective: This study aimed to report the impact of POCT for influenza and RSV on antimicrobial prescribing, including antiviral therapy in primary care.

Methods: The impact of POCT for influenza on antimicrobial stewardship (PIAMS) in UK primary care was a nested cohort study undertaken from January 20 to May 31, 2023, after the period of peak virus circulation, within practices that contribute data to the English sentinel network. People presenting with ARI had a nasopharyngeal swab performed and were tested for influenza and RSV with a molecular POCT analyzer located within the practice. Data on antimicrobial prescribing and other study outcomes were collected by linking information from the analyzer to coded data from the patient's computerized medical record.

Results: In total, 323 swabs were collected from 10 PIAMS study practices. In total, 59.7% (197/323) of swabbed patients were female, and the mean age was 37.28 (SD 25.05) years. Furthermore, 2.9% (9/323) of all swabs were positive, with 0.3% (1/323) positive for influenza A, 1.6% (5/323) positive for influenza B, and 0.9% (3/323) positive for RSV. In total, 80 patients were prescribed antibiotics 7 days following POCT testing. There were no instances of antiviral prescribing in the 7 days post testing. A statistically significant difference in antibiotic prescribing given a positive POCT result compared with a negative test was not found with an unadjusted odds ratio (OR) of 7 days post testing. A statistically significant difference in antibiotic prescribing given a positive POCT result compared with a negative test was not found with an unadjusted OR of 1.54 (95% CI 0.38 - 6.30; *P*=.55) and adjusted OR of 1.21 (95% CI 0.00 - 1.78).

Conclusions: This study illustrates the risk of having a narrow study window; our observation period was not aligned with when influenza was circulating. The peak of weekly incidence of influenza in the sentinel network was in the last week of 2022, and RSV was circulating before this. Further evidence is needed to assess the impact of POCT on antimicrobial prescribing. The viruses tested for using POCT could be aligned with the circulating viruses identified by the sentinel network.

International Registered Report Identifier (IRRID): RR2-10.2196/46938

(JMIR Public Health Surveill 2025;11:e72322) doi:10.2196/72322

KEYWORDS

medical records systems; computerized; influenza; point of care systems; primary care; implementation; outcome assessment; health care

XSL•FO
RenderX

Introduction

Accurate, rapid molecular point-of-care testing (POCT) has the potential to (1) improve clinical decision-making regarding the use of antibiotics and antivirals, (2) improve patient outcomes due to the early appropriate use of antivirals, and (3) provide better information to inform sentinel surveillance and clinical research including studies of vaccine effectiveness and real-world trials [1,2].

For patients with influenza infection, early diagnosis and administration of antivirals may improve clinical outcomes [3,4]. They may also limit symptom duration and spread to household contacts, and newer antivirals for influenza and respiratory syncytial virus (RSV), such as Baloxavir, have been shown to improve the time to resolution of symptoms and reduce complications in high-risk patients [5].

Currently, only a small proportion of patients with acute respiratory illness (ARI) undergo diagnostic microbiological testing before receiving treatments in primary care [6], and there is evidence of widespread variations in antimicrobial prescribing practices [7]. This is important as prescribing in primary care accounts for about 8% of National Health Service expenditure, which is equivalent to over £9 billion (US $12.25 billion) per year, with just over £220 million (US $299.53 million) being spent on antimicrobials [8]. Inappropriate prescribing of antimicrobials and unwarranted variation in prescribing can contribute to an increase in antimicrobial-resistant strains and patient adverse events in the short and long term [9].

We have previously shown that in a prepandemic context, it is feasible to undertake POCT for influenza in primary care in England, with promising impacts on antimicrobial use and comparable estimates of influenza vaccine effectiveness to published data [10-12]. Although its impact on more severe outcomes, such as hospitalization and mortality, following infection was not reported.

With the ending of widespread national testing for SARS-CoV-2 and other respiratory tract infections (RTIs) in the United Kingdom in March 2022 [13], and with high levels of circulating influenza in a post–COVID-19 health service during autumn in 2022 compared with 2021, there was a need to revisit questions about the feasibility of implementing rapid diagnosis of influenza during the expected peak of viral circulation from October 2022 to May 2023 and its impact on clinical management in terms of improved antimicrobial stewardship.

We aimed to deploy POCT during October 2022 to May 2023. The impact of POCT for influenza on antimicrobial stewardship (PIAMS) took place between January 20, 2023, and May 31, 2023, after the peak of virus circulation.

Methods

Study Design, Setting, and Population

This cohort study was nested within the English National Sentinel Surveillance Network managed by the Oxford Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC).

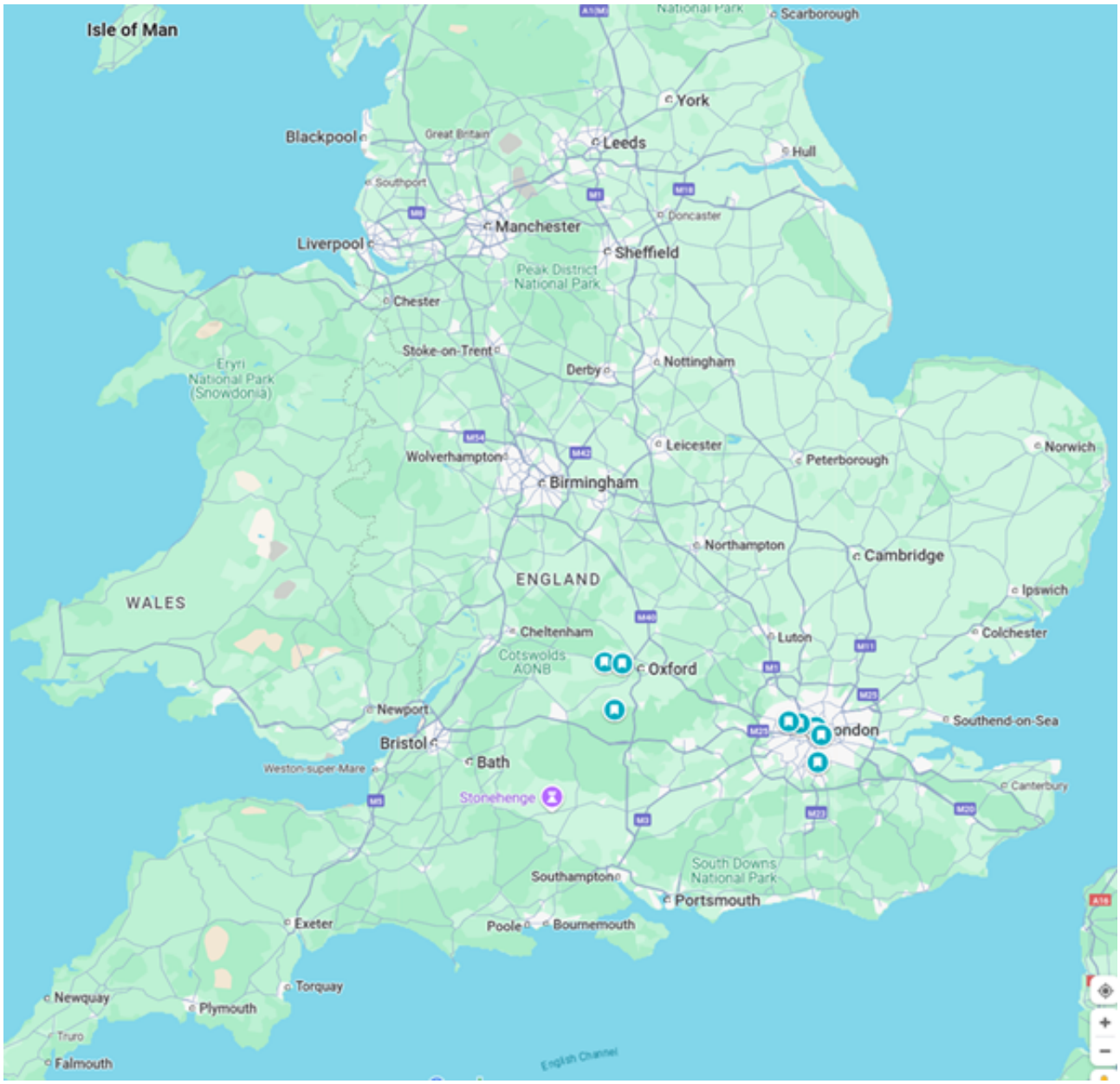
The RSC network of over 2000 primary care practices in England is generally representative of the English population [14] and serves as the English national infectious disease surveillance network. It has been providing weekly data extracts for over 50 years, which are used to monitor trends in infectious disease and investigate real-world vaccine and treatment effectiveness [15]. A subset of practices within the network undertake virology swabbing for testing at the UK Health Security Agency's reference laboratory [14].

All practices that contribute data to the English National Sentinel Network were invited to participate in the PIAMS study. In total, 10 practices were selected (Figure 1). We prioritized practices within the network with the capacity to undertake point-of-care influenza testing and who had previously been involved in SARS-CoV-2 POCT through the Rapid Community Testing for COVID-19 study [16]. Those practices that had a history of less than 80% complete data returns during the previous winter season were excluded. The sample size was influenced by an earlier nested cohort study of respiratory POCT undertaken before the pandemic in the United Kingdom, which used 12 primary care practices [10,17].

Each participant received training about the study, including hands-on training on how to administer a swab test and how to use the POCT analyzers.

XSL•FO
RenderX

Figure 1. Location of general practices in the point-of-care testing influenza on antimicrobial stewardship (PIAMS) study, a nested cohort study undertaken within the English National Sentinel Surveillance Network between January 20, 2023 (International Organization for Standardization week 3, 2023), and May 31, 2023 (International Organization for Standardization week 22, 2023).



Case Definition of Eligible Patients

All patients registered within the PIAMS study practice and showing symptoms of influenza-like illness (ILI), ARI, or fever of higher than 37.5 °C were eligible for the study if they consented to participate. We used the European Centre for Disease Prevention and Control case definitions of ILI and ARI for this study. We used the following exclusion criteria: (1) the patient has an opt-out code on their medical record, and (2) the patient declined informed consent.

Face-to-Face Recruitment of Eligible Patients and POCT for Influenza

We undertook opportunistic swab sampling for this study, with potential participants being identified from those registered patients who presented to the PIAMS study practices with respiratory symptoms described in the case definition. No screening or eligibility assessment was undertaken. Blinding of participants and researchers was not undertaken.

Eligible patients or their parents or legal guardians were approached by a practice general practitioner (GP) or research nurse to explain the study and ask for consent to take part when they presented for a face-to-face consultation at the practice.

After obtaining consent, a nasopharyngeal swab was taken by a suitably qualified and experienced GP or research nurse. For those who do not attend the practice in person, a self-test kit was sent to their home.

The swab was inoculated in a test kit and tested with the POCT analyzer as soon as possible after being taken. The results were available to the clinician in less than 20 minutes.

The POCT test we used was the cobas liat analyzer, manufactured by Roche Diagnostics International [18]. This

XSL•FO
RenderX

POCT analyzer is an automated multiplex polymerase chain reaction system, with previous studies demonstrating excellent performance comparable with gold standard laboratory assays, with sensitivity/specificity in the region of 100%/97.1% for influenza A, 97.8%/99.7% for influenza B, and 94.2%/99% for RSV when fresh prospectively collected samples are tested [19,20]. It has Conformité européenne (CE) marking in the European Union [21]. In the United States, it is approved by the Federal Drug Administration [22].

All eligible patients for this study were seen face-to-face.

Study Outcomes and Data Sources

Antimicrobial prescribing and other study outcomes for those who had been swabbed were obtained by linking information from the POCT with data from the patient's computerized medical record in primary care. A pseudonymized National Health Service number was used to allow the linkage of these datasets and to ensure patients' records were kept confidential.

Data for the study are held on dedicated secure servers within the Oxford-RCGP Clinical Informatics Digital Hub trusted research environment. The research group's secure network is situated behind a firewall within the university's network. To protect privacy and confidentiality, only study staff or associated members of the research group who have been appropriately trained and approved by the Head of Department can access the data from secure workstations or secure laptops with encrypted drives. All staff members of the research group working within the team must work from secure workstations or secure laptops with encrypted drives within the research group's secure network. A risk assessment of the physical security of the research group's offices and server room has been conducted by the building and facilities manager, the faculty information technology service manager, and the research group's information governance lead. The university is compliant with the Data Protection Act and UK General Data Protection Regulation and has systems for technical and organizational controls for information security, including a university-level information security and governance group, chaired by the university senior information risk owner. The research group's private network has its own system-level security policy and is tested for vulnerabilities annually.

Statistical Analysis

To quantify the impact of POCT for influenza on antimicrobial prescribing in primary care, we present the odds ratio (OR) of antimicrobial prescribing given a positive POCT result compared with a negative POCT test. Unadjusted OR was calculated by dividing the odds of antimicrobial prescribing in those with a positive POCT result group by the odds of antimicrobial prescribing in those with a negative POCT result. We used established methods to calculate the standard error and 95% CI for the OR [23], as well as the *P* value for significance [24]. The unadjusted OR for antibiotic prescribing and antiviral prescribing was calculated separately. We also used logistic regression to calculate adjusted ORs taking into account patient demographics (age, sex, ethnicity, and socioeconomic status as measured using the Index of Multiple Deprivation), urban-rural classification, and smoking status, factors known to be associated with antimicrobial prescribing in patients presenting with respiratory illness [25]. Missing values were excluded.

Ethical Considerations

The study was reviewed and approved by the English National Research Ethics Committee (reference 21/YH/0077) and Integrated Research Application System (reference 292961), dated October 5, 2022.

Study practices were given a stipend to cover the costs of training staff members and hosting the study. A small remuneration was also provided to practices for each POCT swab to cover the additional time taken during each consultation to undertake swabbing for this study. Patients were not remunerated for taking part in this study. Informed consent was undertaken by a trained practice GP or research nurse for all patients who took part in this study. A pseudonymized extract of information from consented patients was analyzed for this study.

Results

PIAMS Practice Recruitment

In total, 10 practices were recruited for the PIAMS study with a total registered population size of 144,426. The demographic profile of the PIAMS practices is illustrated in Table 1.

XSL•FO
RenderX

Table . Characteristics of study practices taking part in the point-of-care testing influenza on antimicrobial stewardship (PIAMS) study, a nested cohort study undertaken within the English National Sentinel Surveillance Network between January 20, 2023 (International Organization for Standardization week 3, 2023), and May 31, 2023 (International Organization for Standardization week 22, 2023).

Practice or characteristics	1	2	3	4	5	6	7	8	9	10	England or Wales
Practice size[a], n	18,128	18,324	8909	7305	9434	16,435	15,672	10,541	19,125	20,553	59,620,100
Female[a], n (%)	9344 (51.5)	8056 (44.0)	4354 (48.9)	3560 (48.7)	4606 (48.8)	8142 (49.5)	7961 (50.8)	5090 (48.3)	9099 (47.6)	10,483 (51.0)	30,420,202 (51.0)
Ethnicity, n/N (%)[b]											
White	8669/8953 (96.8)	1628/13,588 (12.0)	1917/3196 (60.0)	3748/7612 (49.2)	3748/7612 (49.2)	5844/10,678 (54.7)	5196/5398 (96.2)	4566/8483 (53.8)	5745/18,754 (30.6)	4622/4870 (94.9)	[illegible] (86.0)
Mixed or multiple ethnic groups	112/8953 (1.3)	347/13,588 (2.6)	156/3196 (4.9)	408/7612 (5.4)	408/7612 (5.4)	624/10,678 (5.8)	83/5398 (1.5)	568/8483 (6.7)	6.7 (1250/18,754 (6.7)	77/4870 (1.6)	[illegible] (2.2)
Asian or Asian British	132/8953 (1.5)	10,041/13,588 (73.9)	778/3196 (24.3)	2089/7612 (27.4)	2089/7612 (27.4)	1487/10,678 (13.9)	88/5398 (1.6)	1421/8483 (16.8)	6534/18,754 (34.8)	110/4870 (2.3)	[illegible] (7.5)
Black, African, Caribbean, or Black British	31/8953 (0.3)	1285/13,588 (9.5)	249/3196 (7.8)	1073/7612 (14.1)	1073/7612 (14.1)	2365/10,678 (22.1)	14/5398 (0.3)	1454/8483 (17.1)	4760/18,754 (25.4)	43/4870 (0.9)	[illegible] (3.3)
Other ethnic group	9/8953 (0.1)	287/13,588 (2.1)	96/3196 (3.0)	294/7612 (3.9)	294/7612 (3.9)	358/10,678 (3.4)	20/5398 (0.4)	474/8483 (5.6)	465/18,754 (2.5)	18/4870 (0.4)	[illegible] (1.0)
IMD[c] Decile[d], n	10	4	3	2	2	2	10	1	5	8	—[e]

[a]Registered practice population as of September 2023.

[b]Ethnic group statistics from the 2011 Census for England and Wales at postcode sector level.

[c]IMD: Index of Multiple Deprivations.

[d]IMD decile: from 1=the most deprived 10% geographic area to 10=the least deprived 10% of geographic area.

[e]Not applicable.

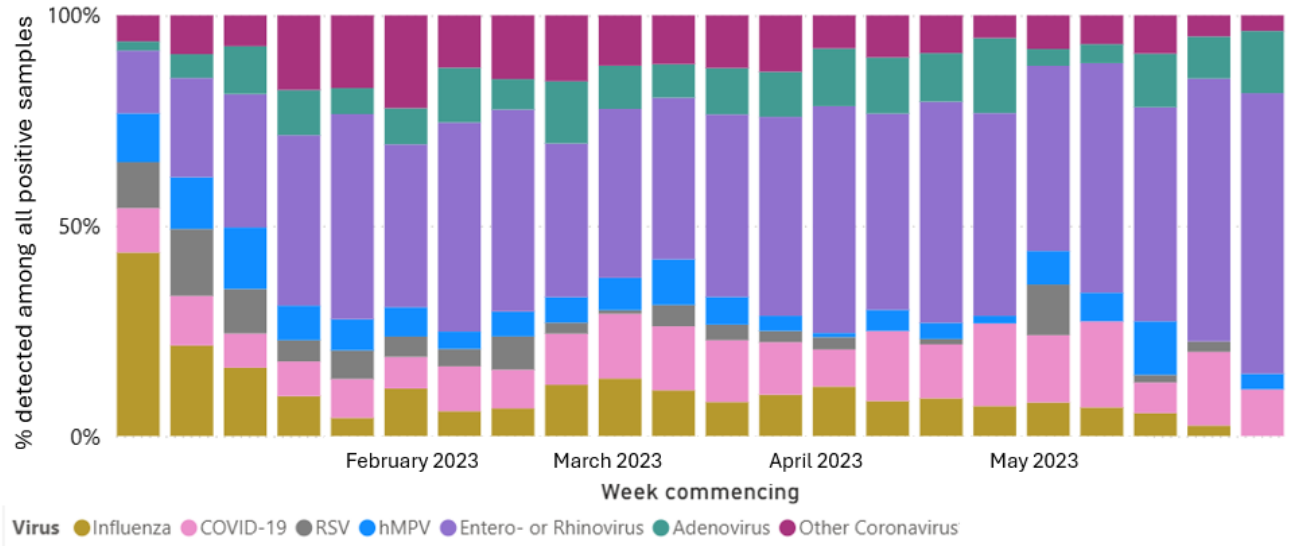
Swabbing Rates in PIAMS Practices

The study started on January 20, 2023 (International Organization for Standardization [ISO] week 3, 2023). In total, 323 swabs were collected from PIAMS study practices until May 31, 2023 (ISO week 22, 2023). Multimedia Appendix 1 illustrates the number of POCT swabs collected in the PIAMS study by week.

Swabbing rates varied considerably between PIAMS study practices from 4.9 to 75.2 swabs per 1000 patients with eligible symptoms (Multimedia Appendix 2).

XSL•FO
RenderX

Figure 2. English national sentinel virology surveillance swab for respiratory syncytial virus (RSV), influenza, SARS CoV-2, and other respiratory viruses between 2022 and 2023 [26]. hMPV: human metapneumovirus.



Summary Demographics of All Swabbed Patients

Table 2 illustrates further detailed analysis of the demographics of those swabbed, showing that 59.4% (192/323) were female, and the mean age of those swabbed was 37 years. A total of 46.4% (150/323) samples were taken from patients of White ethnicity and 39% (126/323) samples taken from patients of Asian, Black, or mixed ethnicity.

XSL•FO
RenderX

Table . Demographic profile of the patients swabbed in the point-of-care testing influenza on antimicrobial stewardship (PIAMS) study, a nested cohort study undertaken within the English National Sentinel Surveillance Network between January 20, 2023 (International Organization for Standardization week 3, 2023), and May 31, 2023 (International Organization for Standardization week 22, 2023).

Demographic characteristics	Swabs (N=323), n (%)
Age band (years)	
<1	11 (3.4)
1 - 4	43 (13.3)
5 - 14	27 (8.4)
15 - 24	24 (7.4)
25 - 44	89 (27.6)
45 - 64	77 (23.8)
65 - 74	27 (8.4)
75 - 84	18 (5.6)
85+	7 (2.2)
Sex	
Female	192 (59.4)
Male	131 (40.6)
Ethnicity	
White	150 (46.4)
Asian	83 (25.7)
Black	35 (10.8)
Mixed	8 (2.5)
Other	10 (3.1)
Unknown	37 (11.5)
Index of Multiple Deprivation quintile	
1 (most deprived)	42 (13)
2	105 (32.5)
3	52 (16.1)
4	45 (13.6)
5 (least deprived)	79 (24.6)
Urban-rural classification	
City and Town	68 (21.1)
Conurbation	198 (61.3)
Rural	57 (17.7)
Smoking status	
Active smoker	23 (7.4)
Ex-smoker	50 (15.5)
Nonsmoker	147 (46.4)

Swab Positivity

In total, swab positivity in the PIAMS study was 2.8% (9/323), with influenza A positivity at 0.3% (1/323), influenza B positivity at 1.6% (5/323), and RSV positivity at 0.9% (3/323). Multimedia Appendix 3 illustrates the proportion of swabs that were RSV and influenza positive in PIAMS practices compared with practices in the English National Sentinel Surveillance Network by week.

Effects of POCT Results on Antimicrobial Prescribing 7 Days After POCT Testing

In total, there were 80 instances of antibiotic prescribing 7 days following POCT testing and no instances of antiviral prescribing 7 days following POCT testing in the PIAMS study. Table 3

XSL•FO
RenderX

illustrates the number of cases that prescribed antibiotics given the POCT result.

Table . Number of cases that prescribed antibiotics 7 days following point-of-care testing swab results in the point-of-care testing influenza on antimicrobial stewardship (PIAMS) study, a nested cohort study undertaken within the English National Sentinel Surveillance Network between January 20, 2023 (International Organization for Standardization week 3, 2023), and May 31, 2023 (International Organization for Standardization week 22, 2023).

POCT[a] virology swab result	Antibiotic prescribed within 7 days of POCT virology swab result		Total
	Yes	No	
Positive, n	3	6	9
Negative, n	77	237	314
Total, n	80	243	323

[a]POCT: point-of-care testing.

The unadjusted OR for antibiotic prescribing given a positive POCT result was 1.54 (95% CI 0.38 - 6.30; *P*=.55) compared with a negative POCT result. The adjusted OR for antibiotic prescribing, taking into account differences in age, sex, ethnicity, socioeconomic status as measured using the Index of Multiple Deprivation, urban-rural classification, and smoking status was 1.21 (95% CI 0.00 - 1.78).

Discussion

Main Study Findings

A total of 10 general practices with a combined registered list size of 144,426 patients participated in this study. They integrated POCT into their clinical workflow, collecting 323 samples. Furthermore, 59.4% (192/323) of the samples were from female patients, and 39% (126/323) samples taken were from patients of Asian, Black, or Mixed ethnicity. The mean age of those swabbed was 37.28 (SD 25.05) years. In addition, 2.8% (9/323) of the swabs collected were positive, with influenza A positivity at 0.3% (1/323), influenza B positivity at 1.6% (5/323), and RSV positivity at 0.9% (3/323). Of the 9 POCT virology swab-positive cases, 33% (3/9) received antibiotics. Of the 314 POCT virology swab negative cases, 24.5% (77/314) received antibiotics. A statistically significant difference in antibiotic prescribing given a positive POCT result compared with a negative test was not found with an OR of 1.54 (95% CI 0.38 - 6.30; *P*=.55).

Implications of Our Findings

We have shown that in a postpandemic health service, POCT for respiratory viruses can be integrated into primary care workflows, although there was a wide variation in the rate of virological swabbing between practices. Our qualitative substudy identified 2 distinct POCT swabbing workflows—one led by clinicians and another managed by research nurses or health care assistants [27].

Key factors that influenced the adoption of each POCT swabbing workflow included the usability of the technology, the skill mix of primary care staff within the practice, the perceived ease of integration of POCT into routine clinical workflows, the availability of comprehensive staff training, the organizational readiness for change, and collective buy-in from all stakeholders [27].

The degree to which these different POCT swabbing workflow models were adopted could have accounted for the widespread differences in swabbing rates seen.

In addition, we have illustrated in Multimedia Appendix 3 that the number of swabs in the PIAMS practices and swab positivity found from POCT virology swabbing broadly reflected what was happening in the English national sentinel system over the same weeks [26]. Figure 2, from the English National Sentinel Surveillance Network, illustrated that RSV rates peaked in the early weeks of 2023, which was also seen in PIAMS practices (Multimedia Appendix 3), where RSV positive swabs were seen between ISO weeks 3 and 5, 2023.

However, our findings make it difficult to offer any recommendations on the impact of POCT on antimicrobial stewardship, as the number of patients prescribed antibiotics and antiviral medications following POCT virology swabbing results was low. Thus, our ORs for prescribing antibiotics given a positive POCT result compared with a negative POCT result were nonsignificant. OR for prescribing antivirals was not calculable as no antivirals were prescribed during the study. The absence of antiviral prescribing is of note despite the receipt of POCT-positive swab samples. This may be due to the only timing of the study at the end of the influenza season when there was a predominance of influenza B cases, which are much less likely to receive antiviral treatments despite evidence that these cases have comparable clinical outcomes to influenza A cases [28] and clinical guidelines recommending their use in POCT confirmed cases of influenza B [29].

Comparison With Existing Literature

Our ORs for antibiotic prescribing given a positive POCT swabbing result of 1.54 (95% CI 0.38 - 6.30; *P*=.55) contrasts with an earlier study conducted within the sentinel network before the pandemic in 2019, which found an OR of 0.4 (95% CI 0.2-0.8; *P*=.01) for antibiotic prescribing given a positive POCT result compared with a negative test, suggesting that antibiotic prescribing was less likely given a positive influenza POCT result compared with a negative result [11].

However, our results are consistent with a systematic review and meta-analysis of POCT in ambulatory care before the pandemic in 2019 which suggested that POCTs had no effect

XSL•FO
RenderX

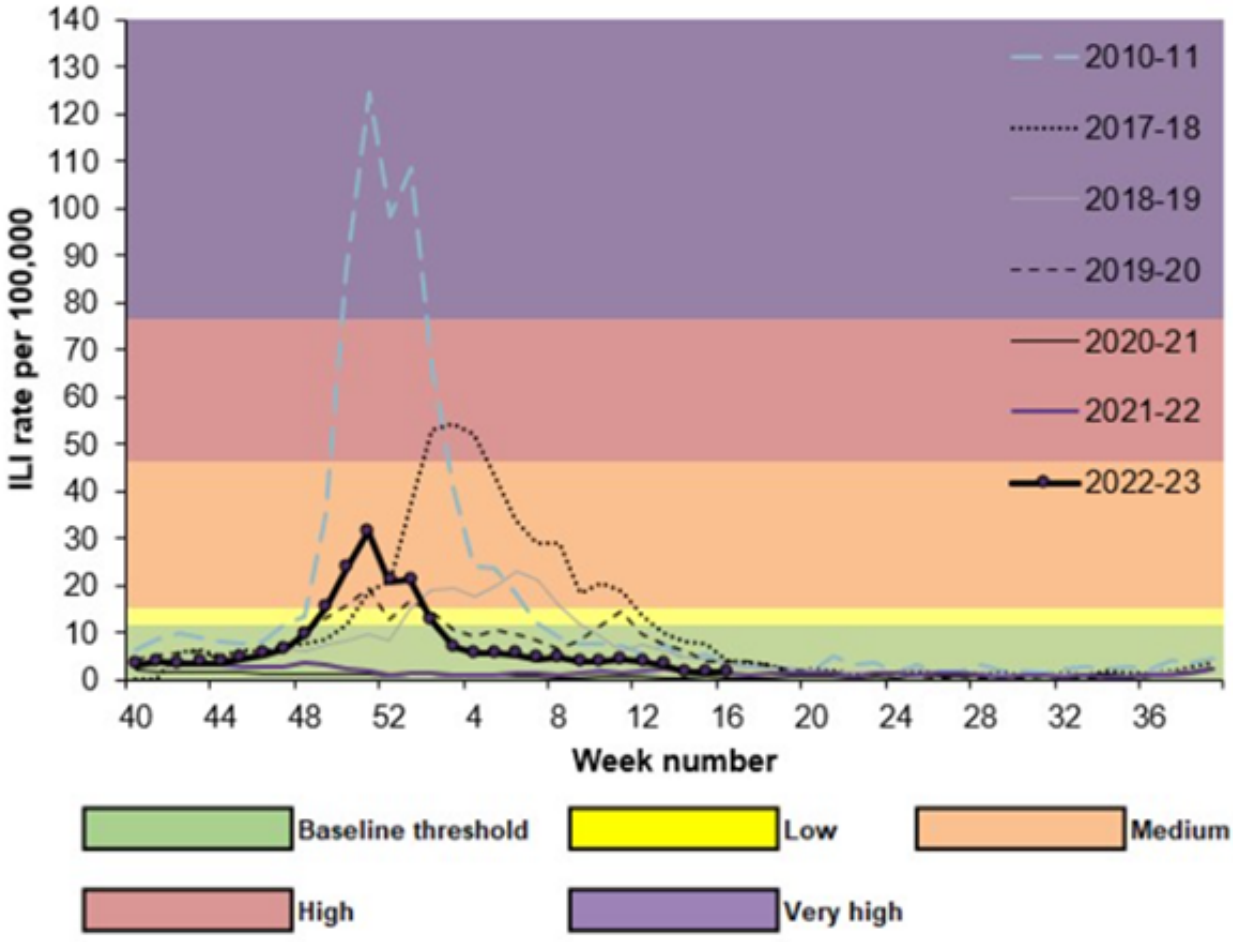
on antibiotic prescribing rates (relative risk 0.97, 95% CI 0.82-1.15; I^2=70%) [30].

No further systematic reviews have been undertaken of POCT on influenza in the postpandemic primary care context, although research has suggested that overall antimicrobial prescribing for RTIs in the community reduced significantly by 12.4% during the pandemic winter season (December 202o to February 2021) compared with the prepandemic winter season (December 2019 to February 2020) [31-33], although there was a slight uptick in antimicrobial prescribing in 2022. Antimicrobial prescribing in primary care typically accounts for 80% (29/36.4) of total antibiotic prescribing in England [8], of which 46% (39.6/80.6) are prescribed for RTIs [34].

Limitations of the Study

The main weakness of our study was the small sample size as a result of the small number of practices included in the study and an earlier start to the seasonal influenza epidemic in October 2022 to May 2023 [35]. This was earlier than seen in previous years and is comparable with the peak of ILI seen in 2010 - 2011 (Figure 3). A further weakness of our study was the lack of testing for respiratory viruses other than influenza and RSV. As illustrated in Figure 2, other respiratory viruses, such as COVID-19, were circulating in the community during our study. The lack of POCT analysis for microbes other than influenza and RSV, which could cause eligible symptoms and which may have had an impact on antimicrobial prescribing, could reduce the strength of our study to detect an effect of POCT on patient management in primary care. Some patients may also have had additional virological swabs sent to the reference laboratory for testing; however, this information was not available in this study and is unlikely to affect the prescribing of antimicrobials, given the significant delay between virological testing and the receipt of a result from the reference laboratory versus POCT.

Figure 3. Weekly all-age GP influenza-like illnessv (ILI) rates in the English National Sentinel Surveillance Network for 2022 - 2023 and past seasons [35].



Conclusion

The practice of performing rapid testing for suspected viral illnesses had become an accepted norm for patients and clinicians alike during the COVID-19 pandemic [36]. As seasonal patterns of community spread of respiratory viruses are re-established following the pandemic, it is important to re-evaluate the impact of novel methods for rapid diagnosis and clinical or public health management of common respiratory viruses such as POCT. POCT tests might usefully be aligned with what the RSC sentinel network says is circulating.

Our study was performed immediately post pandemic and was disrupted by an earlier start to the influenza circulation in

XSL•FO
RenderX

October 2022 to May 2023. Further research is needed to study the impact of POCT on clinical management in primary care, including its effects on antibiotic and antiviral stewardship and the cost-benefits of POCT in postpandemic UK general practice. This study illustrates the risk of having a narrow study window. Research teams planning studies of POCT testing associated with viruses that circulate seasonally should avoid narrow observation windows or risk low rates of identification of their target viruses.

Acknowledgments

We are thankful toward participating practices for sharing data and patients and their parents or carers for volunteering to participate in this study; Alice Williams and Carole Aspden for their support on this study; Charis Xie for her work on the point-of-care testing for influenza on antimicrobial stewardship (PIAMS) qualitative substudy; Magentus and EMIS for their collaboration with pseudonymized data extraction; and Elissa Robbins, Karen Gilliam, Susie Ochoa, and Babar Javed from Roche Diagnostics, for their support in undertaking this study. We acknowledge Professor Maria Zambon from the UK Health Security Agency for her advice about the use of POCT in the National Health Service. The University of Oxford sponsored this study and reviewed and proposed edits to the study protocol and study documents before submission for ethics approval. This collaborative study is funded by Roche Diagnostics through an investigator-initiated grant (funder's reference SUB-000815). The manuscript was also reviewed by Roche Diagnostics before submission.

Data Availability

Data cannot be shared publicly because of it is owned by the Oxford–Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) and its participating practices. Data are available from the University of Oxford Institutional Data Access Committee for researchers who meet the criteria for access to confidential data. The data underlying the results presented in the study are available from the corresponding author on reasonable request. Further enquires about the RCGP RSC network and data requests can be done by contacting the corresponding author.

Conflicts of Interest

SdL is director of the Oxford Royal College of General Practitioners (RCGP) Research and Surveillance Centre. Through his university, he has received funding for vaccine related research from AstraZeneca, GSK, Sanofi, Seqirus, and Takeda; he has been a member of advisory boards for AstraZeneca, Sanofi, and Seqirus. UH has undertaken continuing professional development podcasts funded by Seqirus and has been a member of advisory boards for Jansen. TC has undertaken consultancy work and sat on advisory boards for Roche. TC is also a member of an Independent Data Monitoring Committee (IDMC) for a Roche-sponsored study.

Multimedia Appendix 1

Number of point-of-care testing swabs collected by week in the point-of-care testing influenza on antimicrobial stewardship (PIAMS) study, a nested cohort study undertaken within the English National Sentinel Surveillance Network between January 20, 2023 (International Organization for Standardization week 3, 2023), and May 31, 2023 (International Organization for Standardization week 22, 2023).

[JPG File, 28 KB - publichealth_v11i1e72322_app1.jpg]

Multimedia Appendix 2

Swabbing rate per 1000 patients with eligible symptoms in point-of-care testing influenza on antimicrobial stewardship (PIAMS) study practices, a nested cohort study undertaken within the English National Sentinel Surveillance Network between January 20, 2023 (International Organization for Standardization week 3, 2023), and May 31, 2023 (International Organization for Standardization week 22, 2023). Eligible symptoms include patients registered within PIAMS study practice who were showing symptoms of influenza like illness (ILI), acute respiratory illness (ARI) or fever of higher than 37.5 °C.

[PNG File, 20 KB - publichealth_v11i1e72322_app2.png]

Multimedia Appendix 3

Proportion of swabs that were respiratory syncytial virus and influenza positive in the point-of-care testing influenza on antimicrobial stewardship (PIAMS) study practices, a nested cohort study undertaken within the English National Sentinel Surveillance Network between January 20, 2023 (International Organization for Standardization week 3, 2023), and May 31, 2023 (International Organization for Standardization week 22, 2023), compared all practices with the English national sentinel network by week.

[JPG File, 62 KB - publichealth_v11i1e72322_app3.jpg]

References

XSL•FO
RenderX

1. Leland DS, Ginocchio CC. Role of cell culture for virus detection in the age of technology. Clin Microbiol Rev 2007 Jan;20(1):49-78. [doi: 10.1128/CMR.00002-06] [Medline: 17223623]
2. Zavattoni M, Percivalle E, Cattaneo E, Revello MG, Torsellini M, Gerna G. Optimized detection of respiratory viruses in nasopharyngeal secretions. New Microbiol 2003 Apr;26(2):133-140. [Medline: 12737194]
3. Osterholm MT, Kelley NS, Sommer A, Belongia EA. Efficacy and effectiveness of influenza vaccines: a systematic review and meta-analysis. Lancet Infect Dis 2012 Jan;12(1):36-44. [doi: 10.1016/S1473-3099(11)70295-X] [Medline: 22032844]
4. Jefferson T, Jones MA, Doshi P, et al. Neuraminidase inhibitors for preventing and treating influenza in adults and children. Cochrane Database Syst Rev 2014 Apr 10;2014(4):CD008965. [doi: 10.1002/14651858.CD008965.pub4] [Medline: 24718923]
5. Ison MG, Portsmouth S, Yoshida Y, Shishido T, Hayden F, Uehara T. LB16. Phase 3 trial of baloxavir marboxil in high-risk influenza patients (CAPSTONE-2 study). Open Forum Infect Dis 2018 Nov 26;5(suppl_1):S764-S765. [doi: 10.1093/ofid/ofy229.2190]
6. Linder JA, Nieva HR, Blumentals WA. Antiviral and antibiotic prescribing for influenza in primary care. J Gen Intern Med 2009 Apr;24(4):504-510. [doi: 10.1007/s11606-009-0933-9] [Medline: 19225847]
7. Ashworth M, Charlton J, Ballard K, Latinovic R, Gulliford M. Variations in antibiotic prescribing and consultation rates for acute respiratory infection in UK general practices 1995-2000. Br J Gen Pract 2005 Aug;55(517):603-608. [Medline: 16105368]
8. English surveillance programme for antimicrobial utilisation and resistance (ESPAUR) report 2018. : Public Health England; 2018 URL: https://www.gov.uk/government/publications/english-surveillance-programme-antimicrobial-utilisation-and-resistance-espaur-report [accessed 2025-07-09]
9. Armstrong D, Ashworth M, Dregan A, White P. The relationship between prior antimicrobial prescription and meningitis: a case-control study. Br J Gen Pract 2016 Apr;66(645):e228-e233. [doi: 10.3399/bjgp16X684313] [Medline: 26965030]
10. de Lusignan S, Hoang U, Liyanage H, et al. Using point of care testing to estimate influenza vaccine effectiveness in the English primary care sentinel surveillance network. PLoS ONE 2021;16(3):e0248123. [doi: 10.1371/journal.pone.0248123] [Medline: 33705452]
11. de Lusignan S, Hoang U, Liyanage H, et al. Integrating molecular point-of-care testing for influenza into primary care: a mixed-methods feasibility study. Br J Gen Pract 2020 Aug;70(697):e555-e562. [doi: 10.3399/bjgp20X710897] [Medline: 32661013]
12. de Lusignan S, Hoang U, Liyanage H. Feasibility of point-of-care testing for influenza within a national primary care sentinel surveillance network in england: protocol for a mixed methods study. JMIR Preprints. Preprint posted online on Mar 28, 2019. [doi: 10.2196/preprints.14186]
13. Iacobucci G. Covid-19: who will be eligible for free testing from 1 April? BMJ 2022 Mar 25;376:o807. [doi: 10.1136/bmj.o807]
14. Leston M, Elson WH, Watson C, et al. Representativeness, vaccination uptake and COVID clinical outcomes 2020-21 in the UK's oxford-RCGP research and surveillance network: cohort profile. (preprint). JMIR Public Health Surveill. p. 39141. [doi: 10.2196/preprints.39141]
15. de Lusignan S, Correa A, Smith GE, et al. RCGP Research and Surveillance Centre: 50 years' surveillance of influenza, infections, and respiratory conditions. Br J Gen Pract 2017 Oct;67(663):440-441. [doi: 10.3399/bjgp17X692645] [Medline: 28963401]
16. Nicholson BD, Hayward G, Turner PJ, et al. Rapid community point-of-care testing for COVID-19 (RAPTOR-C19): protocol for a platform diagnostic study. Diagn Progn Res 2021 Feb 8;5(1):4. [doi: 10.1186/s41512-021-00093-8] [Medline: 33557927]
17. de Lusignan S, Hobbs FR, Sheikh A. Lessons from the English primary care sentinel network's response to the COVID-19 pandemic. Lancet Infect Dis 2024 Jan;24(1):14-16. [doi: 10.1016/S1473-3099(23)00736-3] [Medline: 38040007]
18. Ling L, Kaplan SE, Lopez JC, Stiles J, Lu X, Tang YW. Parallel validation of three molecular devices for simultaneous detection and identification of Influenza A and B and respiratory syncytial viruses. J Clin Microbiol 2018 Mar;56(3):e01691-17. [doi: 10.1128/JCM.01691-17] [Medline: 29263204]
19. Gibson J, Schechter-Perkins EM, Mitchell P, et al. Multi-center evaluation of the cobas ® Liat ® Influenza A/B & RSV assay for rapid point of care diagnosis. J Clin Virol 2017 Oct;95(5-9):5-9. [doi: 10.1016/j.jcv.2017.08.004]
20. Gosert R, Naegele K, Hirsch HH. Comparing the cobas Liat Influenza A/B and respiratory syncytial virus assay with multiplex nucleic acid testing. J Med Virol 2019 Apr;91(4):582-587. [doi: 10.1002/jmv.25344] [Medline: 30345524]
21. Cobas® Influenza A/B & RSV. Roche Diagnostics. URL: https://diagnostics.roche.com/global/en/products/lab/cobas-influenza-a-b-and-rsv-rmd-liat-infl-002.html#productSpecs [accessed 2025-05-19]
22. Roche receives CLIA waiver for cobas® Influenza A/B & RSV test for the cobas® Liat®. PR Newswire. 2016. URL: https://www.prnewswire.com/news-releases/roche-receives-clia-waiver-for-cobas-influenza-ab--rsv-test-for-the-cobas-liat-300306666.html [accessed 2025-05-19]
23. Altman DG. Practical Statistics for Medical Research, 1st edition: New York: Chapman and Hall/CRC; 1990.
24. Sheskin D. Handbook of Parametric and Non-Parametric Statistical Procedures, 5th edition 2011.

XSL•FO
RenderX

25. McKay R, Mah A, Law MR, McGrail K, Patrick DM. Systematic review of factors associated with antibiotic prescribing for respiratory tract infections. Antimicrob Agents Chemother 2016 Jul;60(7):4106-4118. [doi: 10.1128/AAC.00209-16] [Medline: 27139474]
26. RSC: public health data. Royal College of General Practitioners. URL: https://www.rcgp.org.uk/representing-you/research-at-rcgp/research-surveillance-centre/public-health-data [accessed 2025-05-19]
27. Xie CX, Hoang U, Smylie J, et al. Qualitative evaluation of a molecular point-of-care testing study for influenza in UK primary care. BJGP Open 2024 Dec;8(4):38981656. [doi: 10.3399/BJGPO.2024.0112] [Medline: 38981656]
28. Davido B, Lemarie B, Gault E, et al. Comparison between clinical outcomes in influenza A and B Infections: a multicenter retrospective cohort study. CMI Communications 2025 Mar;2(1):105072. [doi: 10.1016/j.cmicom.2025.105072]
29. Guidance on use of antiviral agents for the treatment and prophylaxis of seasonal influenza, version 11. : UK Health Security Agency; 2021 URL: https://assets.publishing.service.gov.uk/media/62209cd38fa8f549097b87ec/ukhsa-guidance-antivirals-influenza-11v4.pdf
30. Lee JJ, Verbakel JY, Goyder CR, et al. The clinical utility of point-of-care tests for influenza in ambulatory care: a systematic review and meta-analysis. Clin Infect Dis 2019 Jun 18;69(1):24-33. [doi: 10.1093/cid/ciy837] [Medline: 30285232]
31. Andrews A, Bou-Antoun S, Guy R, Brown CS, Hopkins S, Gerver S. Respiratory antibacterial prescribing in primary care and the COVID-19 pandemic in England, winter season 2020-21. J Antimicrob Chemother 2022 Feb 23;77(3):799-802. [doi: 10.1093/jac/dkab443] [Medline: 34897486]
32. Yang YT, Zhong X, Fahmi A, et al. The impact of the COVID-19 pandemic on the treatment of common infections in primary care and the change to antibiotic prescribing in England. Antimicrob Resist Infect Control 2023 Sep 16;12(1):102. [doi: 10.1186/s13756-023-01280-6] [Medline: 37717030]
33. McCloskey AP, Malabar L, McCabe PG, Gitsham A, Jarman I. Antibiotic prescribing trends in primary care 2014-2022. Res Social Adm Pharm 2023 Aug;19(8):1193-1201. [doi: 10.1016/j.sapharm.2023.05.001] [Medline: 37183105]
34. Dolk FCK, Pouwels KB, Smith DRM, Robotham JV, Smieszek T. Antibiotics in primary care in England: which antibiotics are prescribed and for which conditions? J Antimicrob Chemother 2018 Feb 1;73(suppl_2):ii2-ii10. [doi: 10.1093/jac/dkx504] [Medline: 29490062]
35. Surveillance of influenza and other seasonal respiratory viruses in the UK, winter 2022 to 2023. GOV.UK. 2024. URL: https://www.gov.uk/government/statistics/annual-flu-reports/surveillance-of-influenza-and-other-seasonal-respiratory-viruses-in-the-uk-winter-2022-to-2023 [accessed 2024-11-22]
36. Bevan I, Stage Baxter M, Stagg HR, Street A. Knowledge, attitudes, and behavior related to COVID-19 testing: a rapid scoping review. Diagnostics (Basel) 2021 Sep 15;11(9):1685. [doi: 10.3390/diagnostics11091685] [Medline: 34574026]

Abbreviations

ARI: acute respiratory illness
GP: general practitioner
ILI: influenza-like illness
ISO: International Organization for Standardization
OR: odds ratio
PIAMS: point-of-care testing for influenza on antimicrobial stewardship
POCT: point-of-care testing
RCGP: Royal College of General Practitioners
RSC: Research and Surveillance Centre
RSV: respiratory syncytial virus
RTI: respiratory tract infection

Edited by A Mavragani, T Sanchez; submitted 07.02.25; peer-reviewed by D Umoru, I Cheung; revised version received 19.05.25; accepted 21.05.25; published 15.07.25.

Please cite as:
Hoang U, Smylie J, Button E, Macartney J, Okusi C, Byford R, Ferreira F, Xie C, Joy M, Clark T, de Lusignan S
The Impact of Point-of-Care Testing for Influenza on Antimicrobial Stewardship in UK Primary Care: Nested Cohort Study
JMIR Public Health Surveill 2025;11:e72322
URL: https://publichealth.jmir.org/2025/1/e72322
doi:10.2196/72322

© Uy Hoang, Jessica Smylie, Elizabeth Button, Jack Macartney, Cecilia Okusi, Rachel Byford, Filipa Ferreira, Charis Xie, Mark Joy, Tristan Clark, Simon de Lusignan. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org),

XSL•FO
RenderX

15.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessment of Influenza Testing Distribution in the United States for the 2021-2022, 2022-2023, and 2023-2024 Influenza Seasons: Online Cross-Sectional Study

Autumn Gertz[1], MS; Juliana B Sopko[1], MPH; Christopher Remmel[1], MM; Benjamin Rader[1,2], MPH, PhD; John S Brownstein[1,2], PhD

[1]Boston Children's Hospital, 300 Longwood Ave, Boston, MA, United States

[2]Harvard Medical School, Boston, MA, United States

Corresponding Author:
Autumn Gertz, MS
Boston Children's Hospital, 300 Longwood Ave, Boston, MA, United States

Abstract

Background: Effective surveillance of seasonal influenza is crucial to understanding disease burden and impact. Traditional surveillance accounts for those who interact with the health care system, including those who are testing for diseases like influenza. However, care seeking and testing are not as common with influenza and can lead to bias. Better understanding who is being captured by current surveillance methods can help further knowledge around influenza and identify areas of improvement in surveillance, disease mitigation, and intervention efforts.

Objective: This study aimed to examine who is testing for influenza amongst a United States representative survey population, across three seasons influenza seasons spanning 2021 to 2024.

Methods: Outbreaks near me (ONM) is a participatory surveillance system that, in partnership with SurveyMonkey, conducted a web-based, weekly cross-sectional survey. ONM Survey data from three influenza seasons was used in this study: 2021 - 2022, 2022 - 2023, and 2023 - 2024. Tested for influenza was defined as a "yes" response to "In the past 30 days, have you been tested for influenza (flu)?" Descriptive proportions applying survey weights reflecting US census targets were produced to understand which demographic groups were testing for influenza. A weighted multivariate logistic regression was conducted for influenza testing by income, adjusting for other demographics and COVID-19 testing. Descriptive proportions and multivariate regressions were conducted by influenza season.

Results: In total, 940,172 responses were collected, with similar amounts in 2021 - 2022 (n=335,964) and 2022 - 2023 (n=334,584), and slightly less in 2023 - 2024 (n=269,624). Generally, low levels of influenza testing were reported in each season at 4.2%, 9.1%, and 8.9%, respectively. Weighted proportions of those who tested for influenza only and no other diseases (like COVID-19) were even lower (0.4%, 971/335,964; 1.5%, 4,382/334,584; and 2.0%, 4579/269,624; respectively). Broadly, those who had lower income tested for influenza at progressively higher proportions. A similar trend was observed season to season with education level as well. Across the 3 observed influenza seasons, lower household annual income (under US $15,000) was associated with higher odds of testing for influenza (2021 - 2022: adjusted odds ratio [AOR] 1.41, 95% CI 1.34 - 1.48; 2022 - 2023: AOR 1.42, 95% CI 1.35 - 1.49; 2023 - 2024: AOR 1.25, 95% CI 1.18 - 1.34), while those with higher incomes (over US $150,000) were less likely to have been tested for influenza (2021 - 2022: AOR 0.64, 95% CI 0.55 - 0.86; 2022 - 2023: AOR 0.82, 95% CI 0.73 - 0.91; 2023 - 2024: AOR 0.66, 95% CI 0.56 - 0.76).

Conclusions: Within this study population, individuals who fall within lower-income brackets tested for influenza more than their higher-income counterparts. In all 3 seasons spanning 2021 - 2024, lower income was associated with higher proportions of influenza testing and an increased likelihood of having tested for influenza in the past 30 days. These trends suggest that populations that may experience more barriers to care are not only accessing influenza testing but doing so differently than groups that historically access care.

(JMIR Public Health Surveill 2025;11:e76459) doi:10.2196/76459

KEYWORDS

influenza; flu testing; syndromic surveillance; health survey; influenza-like illness

XSL•FO
RenderX

Introduction

Effective surveillance of respiratory illnesses is crucial to early outbreak detection and robust response efforts [1]. In the United States, surveillance of respiratory disease, such as influenza, has largely been conducted via encounters with the health care system [2]. Laboratory-confirmed influenza cases, outpatient influenza-like illness (ILI), and mortality data help public health researchers gauge the burden of influenza [2]. However, variable access to medical care and health care usage suggests that traditional surveillance is only capturing a part of the population [3]. Influenza severity and burden have been observed to vary across demographics, particularly that racial and ethnic minorities continue to experience higher hospitalization rates for influenza compared to White people [4]. Influenza vaccination rates have also varied across racial and ethnic groups in the United States [4].

During the COVID-19 pandemic, a robust and rapid expansion of testing occurred. Throughout the pandemic, governments turned to large-scale testing efforts to help mitigate the spread of COVID-19 [5]. Early surveillance efforts relied heavily upon traditional data streams like testing centers and hospital data [6]. However, as the pandemic progressed, changes such as the introduction of at-home rapid antigen tests impacted testing, care-seeking behaviors, and surveillance strategy [7,8]. An uptick in at-home testing for COVID-19 was observed, particularly among individuals who identified as White, had higher education attainment, and higher annual income, despite efforts to increase availability and access [9,10]. While such efforts did not explicitly include other respiratory illnesses like influenza, they may have led to secondary changes. For example, test administration within hospitals evolved over the pandemic to assess incidence of pathogens causing acute respiratory illnesses [11]. More recently, multiplex assays have been approved and used to assess single samples for both influenza and COVID-19 [12,13].

In addition to advancements like at-home and multiplex testing, digital disease surveillance methods like participatory surveillance can augment traditional disease surveillance [7,8,14,15]. Outbreaks near me (ONM) is an online participatory syndromic surveillance system that allows for the general public to report disease and health-related information in the United States [16]. ONM enables near-real-time collection and aggregation of data related to disease burden, vaccination coverage, and symptomatology [16]. In addition, participants are asked to answer questions related to age, gender, race, income, and occupation.

In this study, we leverage the previously validated ONM cross-sectional survey to assess the population testing for influenza in the 2021 - 2022, 2022 - 2023, and 2023 - 2024 influenza seasons [10,17]. Specifically, we assess which demographic groups were likely to get tested for influenza and examine if there were any changes in the seasons following the COVID-19 pandemic. ONM is well situated to capture and evaluate any trends and changes to influenza testing, to further the understanding of the impacts and burden of influenza. Here we use the ONM survey to assess who is testing for influenza and identify which populations are represented, or not, in existing influenza surveillance.

Methods

Survey Population

The ONM survey was deployed on the SurveyMonkey digital platform. Participants were recruited via end page river sampling, a sampling method that prompts an optional survey after completion of an unrelated survey on SurveyMonkey [18]. With over two million daily visitors to SurveyMonkey, the survey is able to capture a random sample of US residents, collecting between 5000 - 15,000 responses per week. Recruitment excludes those younger than 13 years due to sampling and privacy concerns. An IP address is used to mitigate the same individual from being prompted to complete the survey more than once [18]. The survey collected robust demographic data via multiple choice and multiselect questions. The data collected via this survey have been previously validated in other studies [10,17]. Data from the 2021 - 2022, 2022 - 2023, and 2023 - 2024 influenza seasons are included in this analysis. A testable version of the survey from the 2023 - 2024 season is available [19]. Influenza seasons are defined as Morbidity and Mortality Weekly Report (MMWR) week 40 (approximately October 1) through March 31, with exact dates of collection in Table S1 in Multimedia Appendix 1. Influenza seasons were truncated to March 31, rather than MMWR week 20, for consistency as data collection ended March 31, 2024.

Analysis

Respondents were asked “In the past 30 days, have you been tested for influenza (flu)?” with participants being able to select either “yes” or “no.” A series of demographic questions were asked, and we collapsed the self-reported industry responses into 2 categories, health care workers and non–health care workers. “In the past 30 days, have you been tested for COVID-19?” was captured as well. Self-reports of at-home tests are included in these “yes” responses. For descriptive analyses, we assessed 2 influenza testing groups: tested for influenza and only tested for influenza. Tested for influenza included anyone who answered “yes” to “In the past 30 days, have you been tested for influenza (flu)?” Those only tested for influenza comprised a subset of those tested for influenza, which excluded anyone who also reported having been tested for influenza and excluded anyone who also reported “yes” to testing for COVID-19 as well. This group does not exclude anyone who was tested for respiratory illnesses beyond COVID-19. An inverse probability of survey selection weighting scheme targeting the US census estimated population was applied to all results in the study [10,17,18,20]. Influenza positive results were defined as a response of “positive” to “What were the results of your influenza (flu) test?” A weighted multivariate logistic regression model was conducted for each influenza season to assess the likelihood of having been tested for influenza. The models assessed the likelihood of having been tested for influenza in the past thirty days by income, adjusting for age, gender, race, education level, health insurance type, health care worker status, and having been tested for COVID-19 in the past 30 days. Descriptive proportions and regression

XSL•FO
RenderX

models were assessed for each influenza season, and not across the study period in its entirety. RStudio (version 2024.12.1+563; Posit) was used to conduct the analyses.

Ethical Considerations

This study was approved by Boston Children's Hospital institutional review board for secondary use of data and received a waiver of informed consent (IRB-P00023700). Participation in the survey was voluntary and participants were not compensated. Limited identifiable data were collected in the study, and all data used were deidentified.

Results

Survey Population

During the study period, a total of 940,172 responses were collected. The weighted demographic breakdowns of the 2021 - 2022, 2022 - 2023, and 2023 - 2024 influenza seasons were similar and were similar to the US population and were thus deemed representative [20]. In all 3 seasons, more respondents identified as female than as male (Table 1). Across all seasons, the largest proportion of respondents was the group aged 65 years and older (19.3%, 69,893/335,964; 19.5%, 60,415/334,584; and 19.5%, 40,477/269,624; respectively); the smallest proportion was of those younger than 18 years (7.1%, 4457/335,964; 6.5%, 3745/334,584; and 7.1%, 2997/269,624; respectively). Across all seasons, the largest proportion of respondents was of those who completed high school or earned their General Educational Development diplomas (33.3%, 57,899/335,964; 33.2%, 51,033/334,584; and 32.8%, 36,531/269,624; respectively), and the smallest proportion was of those who did not complete high school (8.9%, 10,668/335,964; 8.7%, 9871/334,584; and 9.4%, 7893/269,624; respectively). The majority of respondents received insurance via a plan through their spouse, parent, or employer (39.8%, 151,845/335,964; 36.9%, 99,236/334,584; and 35.0%, 54,920/269,624; respectively), and the smallest proportion of participants received insurance through TRICARE (1.9%, 6333/335,964; 1.9%, 4737/334,584; and 1.8%, 2672/269,624; respectively). The complete results of weighted proportions for industry are available in Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Weighted demographic characteristics of the study population across 3 influenza seasons. An online cross-sectional study of US residents 13 and older was carried out to assess self-reported demographics and influenza testing in the last 30 days for the 2021 - 2022 (n=335,964), 2022 - 2023 (n=334,584), and 2023 - 2024 (n=269,624) seasons. This table displays the demographics of the study population. Proportion uses an inverse probability of survey selection weighting scheme targeting the US census. Numbers are unweighted and reflect the number of individuals in each demographic group during each influenza season.

Demographics	Season[a]					
	2021 - 2022 (N=335,964)		2022 - 2023 (N=334,584)		2023 - 2024 (N=269,624)	
	N	Weighted %	N	Weighted %	N	Weighted %
Gender						
Male	129,233	47.1	100,980	47.1	66,401	47.1
Female	200,684	50.6	160,850	50.8	105,135	50.6
Not listed or nonconforming	6047	2.3	5276	2.1	4094	2.2
Race						
White	217,482	62.1	160,169	61.8	102,610	61.7
Black	46,759	16.7	39,214	16.9	26,417	17.1
Hispanic	43,561	12.5	35,636	12.6	23,210	12.5
Asian	16,029	5.7	12,326	5.4	9119	5.4
No Answer	12,133	3.0	87,239	3.3	108,268	3.3
Education						
Did not complete high school	10,668	8.9	9871	8.7	7893	9.4
High school or General Educational Development diploma	57,899	33.3	51,033	33.2	36,531	32.8
Associate's degree	32,592	9.1	25,602	9.5	16,269	9.4
Some college	66,617	19.5	51,577	19.4	33,148	19.2
College graduate	91,664	18.2	70,289	18.3	44,833	18.3
Postgraduate degree	76,524	10.9	58,734	10.9	36,956	10.9
Age (years)						
13 - 17	4457	7.1	3745	6.5	2997	7.1
18 - 24	21,420	11.2	16,655	11.1	11,628	10.9
25 - 34	43,489	16.7	33,772	16.8	22,153	16.7
35 - 44	56,378	15.5	44,276	15.6	28,582	15.5
45 - 54	67,123	14.9	51,021	15.0	33,070	14.9
55 - 64	73,204	15.4	57,222	15.5	36,723	15.5
65 and older	69,893	19.3	60,415	19.5	40,477	19.5
Income (US $)						
Under $15,000	30,200	12.8	25,095	13.4	16,376	14.2
Between $15,000 and $29,999	35,390	13.2	26,565	12.9	16,002	12.4
Between $30,000 and $49,999	46,445	15.4	34,833	15.4	20,741	15.3
Between $50,000 and $74,999	50,944	15.1	36,881	15.0	22,096	15.1
Between $75,000 and $99,999	42,361	11.5	29,940	11.3	17,337	11.0
Between $100,000 and $150,000	52,730	12.9	37,042	12.9	21,364	12.9
Over $150,000	55,466	12.2	28,954	12.2	22,557	12.2
No Answer	22,423	6.9	105,274	6.9	133,151	7.0
Insurance						
Plan through your spouse, parent, or employer	151,845	39.8	99,236	36.9	54,920	35.0

XSL•FO
RenderX

Demographics	Season[a]					
	2021 - 2022 (N=335,964)		2022 - 2023 (N=334,584)		2023 - 2024 (N=269,624)	
Other (please specify)	20,200	6.7	8921	3.8	—[b]	—
Plan you purchased yourself	37,674	10.3	27,863	10.5	16,670	10.5
Medicare	67,835	21.5	56,337	23.1	35,013	23.3
Medicaid or Medical	26,810	9.9	24,050	11.7	15,039	11.9
TRICARE	6333	1.9	4737	1.9	2672	1.8
Not covered by health insurance	19,231	7.7	13,817	7.1	9405	8.0
No Answer	6034	2.2	99,623	5.1	135,905	9.5
Industry						
Non–health care[c]	76,586	78.7	133,084	79.0	177,869	79.9
Health care	25,770	21.3	44,387	21.0	57,672	20.1

[a]Please refer to Table S1 in Multimedia Appendix 1 for the specific date ranges of the 2021-2022, 2022-2023, and 2023-2024 seasons.

[b]Not applicable.

[c]Non–health care is a condensed group that consists of all occupations except health care. To see the complete breakdown of non–health care occupations, refer to Table S2 in Multimedia Appendix 1.

Influenza Testing

During the 2021 - 2022 season, 4.2% (11,733/335,964) of the population had tested for influenza and 0.4% (971/335,964) had only tested for influenza. Slightly more tested during 2022 - 2023 at 9.1% (26,875/334,584) and 1.5% (4382/334,584) only tested for influenza. In the 2023 - 2024 season, 8.9% (21,809/269,624) of the population tested for influenza and 2.0% (4579/269,624) had only tested for influenza. The proportion of individuals who had not completed high school and tested for influenza and only tested for influenza was higher than the other educational groups. Those who made under $15,000 reported more testing in both tested for influenza (7.1%, 2002/335,964; 15.0%, 3640/209,646; and 14.2%, 2319/269,624; respectively) and only tested for influenza (0.8%, 217/335,964; 2.3%, 536/334,584; and 2.7%, 426/269,624; respectively) than higher-income groups (Table 2). Overall, the proportion of individuals with Medicaid or Medical tested for influenza (6.7%, 1659/335,964; 14.4%, 3357/334,584; and 13.2%, 1985/269,624; respectively) and only tested for influenza (0.6%, 130/335,964; 2.1%, 487/334,584; and 2.6%, 376/269,624; respectively) at higher levels (Table 2). There was an observed increase in the proportion of those who did not complete high school and who only tested for influenza (truncated due to low numbers: 2.3%, 251/334,584 and 3.3%, 252/269,624) from 2021 to 2024 (Table 2). Table S3 in Multimedia Appendix 1 contains influenza testing among all the industry options.

XSL•FO
RenderX

Table . Weighted demographic characteristics of individuals who tested for influenza and only tested for influenza across 3 seasons. An online cross-sectional study of US residents aged 13 years and older was carried out to assess self-reported influenza testing in the last 30 days for the 2021 - 2022 (n=335,964), 2022 - 2023 (n=334,584), and 2023 - 2024 (n=269,624) seasons. This table included the numbers and proportions of those who reported testing for influenza in each season. Survey weights to reflect US census targets were applied to proportions only.

Demographics		Season											
		2021 - 2022				2022 - 2023				2023 - 2024			
		Tested for influenza[a]		Only tested for influenza[b]		Tested for influenza[a]		Only tested for influenza[b]		Tested for influenza[a]		Only tested for influenza[b]	
		n	Weighted % (MOE[c])	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)
Gender													
	Male	4463	4.4 (4.3 - 4.6)	392	0.4 (0.4 - 0.4)	8041	9.3 (9.1 - 9.6)	1238	1.4 (1.3 - 1.5)	5146	8.8 (8.5 - 9.1)	1062	1.8 (1.7 - 2.0)
	Female	6976	3.9 (3.7 - 4.0)	544	0.3 (0.3 - 0.4)	12,862	8.8 (8.6 - 9.0)	2128	1.5 (1.4 - 1.6)	8280	8.9 (8.7 - 9.2)	1612	1.9 (1.7 - 2.0)
	Not listed or non-conforming	294	5.0 (4.3 - 5.8)	__[d]	—	591	11.4 (10.2 - 12.6)	—	—	454	11.2 (9.7 - 12.7)	—	—
Race													
	White	5639	3.1 (3.0 - 3.2)	378	0.2 (0.2 - 0.3)	9733	7.2 (7.0 - 7.4)	1432	1.1 (1.1 - 1.2)	6006	6.9 (6.6 - 7.1)	1023	1.3 (1.2 - 1.4)
	Black	2798	6.8 (6.5 - 7.1)	265	0.7 (0.6 - 0.8)	4982	13.9 (13.4 - 14.4)	842	2.2 (2.0 - 2.4)	3441	14.2 (13.6 - 14.8)	719	3.1 (2.8 - 3.4)
	Hispanic	2298	6.1 (5.8 - 6.4)	221	0.6 (0.5 - 0.7)	4176	12.5 (12.0 - 13.0)	711	2.1 (1.9 - 2.3)	2611	11.8 (11.2 - 12.4)	578	2.5 (2.3 - 2.8)
	Asian	462	3.4 (3.0 - 3.8)	—	—	748	7.4 (6.8 - 8.1)	181	1.7 (1.4 - 2.0)	523	7.8 (6.9 - 8.7)	182	2.5 (1.9 - 3.0)
	No answer	536	4.7 (4.2 - 5.2)	—	—	7236	10.1 (9.5 - 10.7)	1216	1.6 (1.4 - 1.9)	9228	10.7 (9.9 - 11.5)	2077	2.5 (2.1 - 2.9)
Education													
	Did not complete high school	684	5.7 (5.1 - 6.2)	—	—	1359	12.4 (11.5 - 13.3)	251	2.3 (1.9 - 2.8)	1034	12.1 (11.0 - 13.2)	252	3.3 (2.7 - 3.9)
	High school or General Educational Development diploma	2971	5.2 (5.0 - 5.5)	270	0.5 (0.4 - 0.6)	5630	10.9 (10.5 - 11.3)	696	1.9 (1.7 - 2.1)	3917	10.6 (10.1 - 11.1)	804	2.2 (2.0 - 2.4)
	Associate's degree	1335	4.2 (3.9 - 4.5)	104	0.3 (0.3 - 0.4)	2253	9.0 (8.5 - 9.5)	330	1.3 (1.1 - 1.5)	1409	8.5 (7.9 - 9.1)	291	1.8 (1.5 - 2.1)
	Some college	2569	4.0 (3.8 - 4.2)	222	0.4 (0.3 - 0.4)	4625	9.0 (8.6 - 9.3)	727	1.3 (1.2 - 1.5)	2939	9.0 (8.5 - 9.4)	541	1.7 (1.5 - 1.9)

XSL•FO
RenderX

Demographics	Season											
	2021 - 2022				2022 - 2023				2023 - 2024			
	Tested for influenza[a]		Only tested for influenza[b]		Tested for influenza[a]		Only tested for influenza[b]		Tested for influenza[a]		Only tested for influenza[b]	
	n	Weighted % (MOE[c])	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)
College graduate	2409	2.7 (2.6 - 2.8)	175	0.2 (0.2 - 0.2)	4464	6.5 (6.2 - 6.7)	699	1.0 (0.9 - 1.1)	2645	6.3 (5.9 - 6.6)	523	1.2 (1.1 - 1.4)
Postgraduate degree	1765	2.4 (2.3 - 2.6)	118	0.2 (0.1 - 0.2)	3163	5.7 (5.4 - 6.0)	475	0.9 (0.8 - 1.0)	1936	5.7 (5.4 - 6.1)	346	1.0 (0.9 - 1.2)
Age (years)												
13 - 17	201	4.2 (3.4 - 5.0)	—	—	396	10.1 (8.7 - 11.5)	101	2.1 (1.4 - 2.8)	352	10.8 (9.0 - 12.6)	131	3.9 (2.7 - 5.0)
18 - 24	1086	5.4 (5.1 - 5.8)	110	0.5 (0.4 - 0.7)	1735	10.8 (10.1 - 11.5)	334	2.0 (1.7 - 2.4)	1184	10.7 (9.8 - 11.6)	271	2.4 (2.0 - 2.8)
25 - 34	2319	6.2 (5.9 - 6.5)	170	0.5 (0.4 - 0.6)	3992	12.8 (12.3 - 13.3)	571	1.8 (1.6 - 2.0)	2297	11.1 (10.5 - 11.7)	419	1.9 (1.6 - 2.1)
35 - 44	2471	4.9 (4.7 - 5.2)	164	0.4 (0.3 - 0.4)	4533	10.8 (10.4 - 11.2)	713	1.8 (1.6 - 2.0)	2820	10.5 (10.0 - 11.0)	518	1.9 (1.7 - 2.1)
45 - 54	2392	3.9 (3.7 - 4.1)	185	0.3 (0.3 - 0.4)	4040	8.3 (8.0 - 8.7)	588	1.3 (1.1 - 1.4)	2689	8.7 (8.2 - 9.1)	473	1.6 (1.4 - 1.8)
55 - 64	1857	2.8 (2.7 - 3.0)	142	0.2 (0.2 - 0.3)	3724	6.9 (6.6 - 7.2)	607	1.1 (1.0 - 1.2)	2432	7.2 (6.8 - 7.5)	505	1.5 (1.3 - 1.7)
65 and older	1407	2.3 (2.1 - 2.4)	175	0.3 (0.3 - 0.4)	3074	5.6 (5.4 - 5.9)	537	1.0 (0.9 - 1.1)	2106	5.7 (5.4 - 6.1)	440	1.2 (1.1 - 1.4)
Income (US $)												
Under 15,000	2002	7.1 (6.7 - 7.5)	217	0.8 (0.7 - 0.9)	3640	15.0 (14.4 - 15.6)	536	2.3 (2.0 - 2.5)	2319	14.2 (13.5 - 14.9)	426	2.7 (2.3 - 3.0)
Between 15,000 and 29,999	1782	5.4 (5.1 - 5.8)	162	0.5 (0.4 - 0.6)	2999	12.0 (11.5 - 12.5)	527	2.2 (2.0 - 2.5)	1863	11.7 (11.0 - 12.4)	350	2.2 (1.9 - 2.5)
Between 30,000 and 49,999	1879	4.4 (4.2 - 4.7)	158	0.4 (0.3 - 0.5)	3247	9.5 (9.1 - 9.9)	496	1.4 (1.2 - 1.5)	1993	9.7 (9.2 - 10.2)	411	2.1 (1.8 - 2.3)
Between 50,000 and 74,999	1687	3.7 (3.5 - 3.9)	145	0.3 (0.2 - 0.4)	2897	8.0 (7.6 - 8.4)	471	1.3 (1.1 - 1.4)	1706	7.8 (7.3 - 8.3)	367	1.7 (1.5 - 1.9)
Between 75,000 and 99,999	1206	3.2 (3.0 - 3.4)	—	—	2041	7.4 (7.0 - 7.8)	325	1.3 (1.1 - 1.4)	1160	7.1 (6.6 - 7.6)	236	1.6 (1.4 - 1.9)

XSL•FO
RenderX

Demographics		Season											
		2021 - 2022				2022 - 2023				2023 - 2024			
		Tested for influenza[a]		Only tested for influenza[b]		Tested for influenza[a]		Only tested for influenza[b]		Tested for influenza[a]		Only tested for influenza[b]	
		n	Weighted % (MOE[c])	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)
	Between 100,000 and 150,000	1288	2.8 (2.6 - 2.9)	—	—	2175	6.3 (6.0 - 6.6)	344	1.1 (1.0 - 1.2)	1209	6.1 (5.7 - 6.6)	224	1.3 (1.1 - 1.5)
	Over 150,000	1287	2.8 (2.6 - 3.0)	—	—	2183	6.3 (5.9 - 6.6)	295	0.9 (0.8 - 1.0)	1211	6.0 (5.6 - 6.4)	190	1.2 (1.0 - 1.4)
	No answer	601	3.3 (3.0 - 3.6)	—	—	7693	7.1 (6.6 - 7.6)	1388	1.4 (1.2 - 1.7)	10348	7.2 (6.6 - 7.9)	2375	1.9 (1.9 - 1.6)
Insurance													
	Plan through your spouse, parent, or employer	4180	3.2 (3.0 - 3.3)	265	0.2 (0.2 - 0.3)	6536	7.1 (6.9 - 7.4)	947	1.1 (1.0 - 1.2)	3604	7.0 (6.7 - 7.3)	626	1.4 (1.3 - 1.6)
	Other (please specify)	819	4.5 (4.2 - 4.9)	—	—	789	9.2 (8.4 - 10.0)	166	1.8 (1.5 - 2.2)	—	—	—	—
	Plan you purchased yourself	1515	4.7 (4.4 - 5.0)	141	0.4 (0.3 - 0.5)	2446	9.1 (8.7 - 9.6)	404	1.5 (1.4 - 1.7)	1478	9.8 (9.2 - 10.4)	298	1.9 (1.6 - 2.2)
	Medicare	2263	4.2 (4.0 - 4.4)	242	0.4 (0.4 - 0.5)	4470	9.5 (9.2 - 9.8)	706	1.6 (1.4 - 1.7)	2843	9.4 (9.0 - 9.8)	545	1.9 (1.7 - 2.1)
	Medicaid or Medical	1659	6.7 (6.3 - 7.1)	130	0.6 (0.4 - 0.7)	3357	14.4 (13.8 - 15.0)	487	2.1 (1.9 - 2.3)	1985	13.2 (12.5 - 13.9)	376	2.6 (2.3 - 3.0)
	TRICARE	262	4.8 (4.1 - 5.5)	—	—	501	11.6 (10.4 - 12.8)	—	—	232	8.9 (7.5 - 10.4)	—	—
	Not covered by health insurance	777	4.7 (4.3 - 5.1)	—	—	1112	8.6 (8.0 - 9.2)	181	1.3 (1.1 - 1.6)	753	7.9 (7.2 - 8.7)	162	1.6 (1.3 - 2.0)
	No answer	257	4.7 (4.0 - 5.5)	—	—	7664	9.3 (8.6 - 10.1)	1406	1.9 (1.6 - 2.3)	10,914	9.3 (8.6 - 9.9)	2517	2.3 (2.0 - 2.7)
Industry													
	Health care	2166	4.3 (4.1 - 4.5)	147	0.3 (0.3 - 0.4)	3851	9.7 (9.4 - 10.1)	555	1.4 (1.3 - 1.6)	2145	9.1 (8.6 - 9.6)	424	1.9 (1.6 - 2.0)

XSL•FO
RenderX

Demographics	Season											
	2021 - 2022				2022 - 2023				2023 - 2024			
	Tested for influenza[a]		Only tested for influenza[b]		Tested for influenza[a]		Only tested for influenza[b]		Tested for influenza[a]		Only tested for influenza[b]	
	n	Weighted % (MOE[c])	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)	n	Weighted % (MOE)
Non-health care	6269	4.2 (4.1 - 4.3)	478	0.3 (0.3 - 0.4)	10,772	9.0 (8.8 - 9.3)	1745	1.5 (1.4 - 1.6)	6192	8.9 (8.6 - 9.2)	1203	1.7 (1.6 - 1.9)

[a]"Tested for influenza" includes all individuals who answered "Yes" to "In the past 30 days, have you been tested for influenza (flu)?"

[b]"Only Tested for influenza" includes only individuals who answered "Yes" to "In the past 30 days have you been tested for influenza (flu)?" and excluded an answer of "Yes" to "In the past 30 days, have you been tested for COVID-19?"

[c]MOE: Margin of Error.

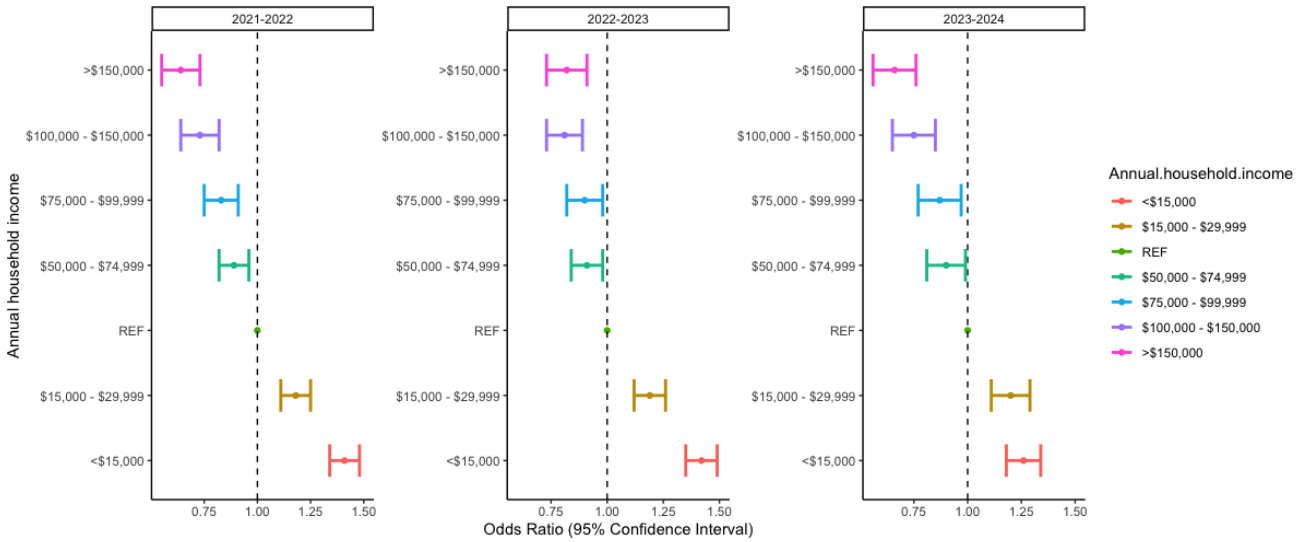
[d]Suppressed values are indicated by "- -" for groups with an N lower than 100.

Overall, positive results to influenza tests were low, ranging from 8.8% (1019/11,733) in 2021 - 2022 to 12.0% (3228/26,875) in 2022 - 2023 and 12.8% (2811/21,892) in 2023 - 2024. Across seasons, the 25 - to 34-year age group was the largest proportion of influenza positive results (27.0%, 241/1019; 28.6%, 631/2811; and 22.0%, 334/2811; respectively), followed by the subsequent 35 - to 44-year age group (17.9%, 220/1019; 18.9%, 639/2811; and 19.0%, 407/3228; respectively). A full output of weighted proportions of influenza positive results among influenza testers by age is located in Table S4 in Multimedia Appendix 1.

Multivariate Logistic Regressions

Likelihood to have been tested for influenza in the past 30 days decreased consistently as income increased, when adjusting for other variables. Those below the reference annual household income (US $30,000 and US $49,999) had higher odds of having been tested for influenza than those with higher incomes (Figure 1). As household income increased, the odds of having been tested for influenza decreased. This trend was observed consistently across the 2021 - 2022, 2022 - 2023, and 2023 - 2024 influenza seasons, with slightly variant magnitudes (Figure 1). Those with an income under US $15,000 were most associated (adjusted odds ratio [AOR] 1.41, 95% CI 1.34 - 1.48; AOR 1.42, 95% CI 1.35 - 1.49; and AOR 1.25, 95% CI 1.18 - 1.34; respectively) with being tested for influenza, and those with an income over US $150,000 were least associated (AOR 0.64, 95% CI 0.55 - 0.86; AOR 0.82, 95% CI 0.73 - 0.91; and AOR 0.66, 95% CI 0.56 - 0.76; respectively) with being tested for influenza in the influenza seasons of 2021 - 2022, 2022 - 2023, and 2023 - 2024. Complete results of the regressions can be found in Table S5 in Multimedia Appendix 1.

Figure 1. AOR for likelihood to have been tested for influenza by income, controlling for demographics and likelihood to test for COVID-19. An online cross-sectional study of US residents aged 13 years and older was carried out to assess self-reported influenza testing in the last 30 days for the 2021 - 2022 (n=335,964), 2022 - 2023 (n=209,646), and 2023 - 2024 (n=269,624) influenza seasons. A weighted multivariate logistic regression was used to assess likelihood of influenza testing by income when controlling for demographics and COVID-19 testing. In all three seasons, those with lower incomes were observed to have a higher likelihood of testing for influenza, and those with higher incomes were less likely. AOR: adjusted odds ratio.



XSL•FO
RenderX

Discussion

Principal Findings

Our objective for this study was to assess who among an online, cross-sectional survey was tested for influenza between 2021 - 2024. In each of the included influenza seasons, only a small amount of the population reported testing for influenza. We found higher proportions of lower income, lower educational attainment, and Black and Hispanic individuals among those who self-reported being tested for influenza in the past 30 days. Furthermore, those who made under US $15,000 a year were the most likely to report being tested for influenza, and those who made over US $150,000 were the least likely, across all 3 seasons. The observation of lower income and higher likelihood of influenza testing differs from other studies, which have found higher income individuals generally have perceived better access and uptake of respiratory disease testing, particularly for COVID-19 [10,21].

Increases in influenza testing over the study period were observed, though levels remained small compared to the overall sample. Advances in diagnostics, such as at-home and multiplex testing, may contribute to increases in testing for influenza in the 2023 - 2024 season [7,8]. Here we found individuals who did not complete high school, made under US $15,000, and had income-assisted insurance had higher proportions of influenza testing. One potential explanation is that the population testing for influenza may need evidence of a health care visit, positive test result, or diagnosis to receive sick time or paid time off benefits. It is documented that people without paid sick leave benefits are less likely to receive influenza vaccination, use emergency care, or visit a health care professional when sick with respiratory illness [22,23]. This population is also likelier to include younger individuals, females, Hispanic individuals, those with a lower educational level, and those in certain occupations like blue-collar or farm jobs [23]. Our finding that lower income individuals are testing more may also be indicating that higher income counterparts have flexibility with their sick leave, such as remote work options [24-27]. Another contributing factor could be that populations known to experience severe influenza may be tested more due to higher hospitalization rates, existing comorbidities, or other factors leading to more interaction with health care during an influenza infection than other population groups [4].

Adjusting our regression model to account for health care worker status and COVID-19 testing further supports these findings as it accounts for those who may also require influenza testing for work. Similarly, adjusting for COVID-19 testing in the past 30 days indicates these trends were not driven by COVID-19 testing during the pandemic. Overall levels of influenza testing were consistently lower than those of COVID-19 testing throughout the study. Furthermore, the population reporting testing for influenza varied from the observed population of COVID testers during a similar time, which tended to be higher income among other demographics [5,6,10,11]. This difference further highlights that there are potential economic differences and impacts in testing for either influenza or COVID-19 during 2021 - 2024. The overlap of our study and the pandemic does introduce nonnormal behaviors, interest, and access to testing. Further study of these trends would be insightful to see if they persist.

This study encompasses strengths in its methodology that lend to the validity of the observed results. Data used for this study was collected from a previously validated, nationally representative survey, allowing us to generalize the observed results [10,17,18,20]. This validation and the weighting scheme that enables findings to encompass census targets minimizes bias about those who participated differing from the US population broadly. The inclusion of data spanning 3 influenza seasons further supports the empirically observed trends as they were consistent across time. Furthermore, the study period captures any potential differences season-to-season and any changes due to changes in the COVID-19 pandemic [28].

This study is also subject to limitations. First, being cross-sectional, the trends observed are insights into different individuals at certain points in time, rather than longitudinally assessing a consistent cohort each season. This means that our insights are broader and not assessing specific individuals' influenza testing and income over time. In addition, access to reliable internet is required for conducting the survey so populations that do not have reliable access are not captured. Similar to many health surveys, older women with higher income and higher educational attainment were more likely to complete the survey. Furthermore, health surveys generally capture an audience who may be more health-minded than the general population. The survey may also be subject to misclassification as participants are asked not only to self-report but to recall testing that occurred in the last 30 days and not an immediate time frame. Moreover, participants may not know what illnesses they were tested for, particularly in a clinical setting.

Conclusion

As public health continually seeks to improve the prevention and mitigation of respiratory diseases, it is vital that we understand trends in who is testing for disease, and the implications for surveillance and response. Different access and uptake of testing by disease may be influenced by economic factors like employment, income, and benefits. These differences are important to understand for future public health efforts, such as modifying messaging around influenza testing, increasing affordable access to influenza testing, and creating policies to ease the burden of testing across illnesses when it is required to receive sick leave. Further studies are needed to understand the full scope of social and economic factors influencing testing behavior for respiratory diseases in the United States. In addition, studying differences across respiratory, rather than pan-respiratory diseases, may provide insights that support targeted interventions. As public health continues to strive for more robust disease mitigation and response, it becomes increasingly important not only to understand differences in access to testing, but disparities in who has the need for testing and official diagnoses and the implications on the health care system and disease surveillance overall.

XSL•FO
RenderX

Acknowledgments

The authors wish to acknowledge the following members of the Boston Children's Hospital team for their support on this project: Kara Sewalk, John Lawless, Amelia Thyen, and Mahlet Kelecha. We also want to acknowledge the SurveyMonkey team for their support in data collection. This work was supported by The Centers for Disease Control and Prevention (grant numbers 75D30123C16026 and 75D30121C11606), National Philanthropic Trust, and Tides Foundation (grant number TF2008-094054).

Data Availability

The datasets analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional tables.
[DOCX File, 38 KB - publichealth_v11i1e76459_app1.docx]

References

1. Al-Tawfiq JA, Zumla A, Gautret P, et al. Surveillance for emerging respiratory viruses. Lancet Infect Dis 2014 Oct;14(10):992-1000. [doi: 10.1016/S1473-3099(14)70840-0] [Medline: 25189347]
2. Brammer L, Budd A, Cox N. Seasonal and pandemic influenza surveillance considerations for constructing multicomponent systems. Influenza Other Respir Viruses 2009 Mar;3(2):51-58. [doi: 10.1111/j.1750-2659.2009.00077.x] [Medline: 19496841]
3. 2021 National Healthcare Quality and Disparities Report. Agency for Healthcare Research and Quality (US). 2021. URL: http://www.ncbi.nlm.nih.gov/books/NBK578529 [accessed 2025-03-06]
4. Black CL, O'Halloran A, Hung MC, et al. Vital Signs: Influenza Hospitalizations and Vaccination Coverage by Race and Ethnicity-United States, 2009-10 Through 2021-22 Influenza Seasons. MMWR Morb Mortal Wkly Rep 2022 Oct 28;71(43):1366-1373. [doi: 10.15585/mmwr.mm7143e1] [Medline: 36302226]
5. Savard Lamothe A, Gabet M, Richard Z, et al. A Descriptive Comparison of Mass Testing During the COVID-19 Pandemic in Montreal, Paris, Bamako, and Recife. Int J Public Health 2022;67:1604992. [doi: 10.3389/ijph.2022.1604992] [Medline: 36213140]
6. Miller MF, Shi M, Motsinger-Reif A, Weinberg CR, Miller JD, Nichols E. Community-Based Testing Sites for SARS-COV-2 - United States, March 2020-November 2021. MMWR Morb Mortal Wkly Rep 2021 Dec 10;70(49):1706-1711. [doi: 10.15585/mmwr.mm7049a3] [Medline: 34882655]
7. FDA Authorizes First Over-the-Counter At-Home Test to Detect Both Influenza and COVID-19 Viruses. Food and Drug Administration. 2024 Aug 9. URL: https://www.fda.gov/news-events/press-announcements/fda-authorizes-first-over-counter-home-test-detect-both-influenza-and-covid-19-viruses [accessed 2025-03-06]
8. Home test to treat program extends nationwide. National Institutes of Health. 2023. URL: https://www.nih.gov/news-events/news-releases/home-test-treat-program-extends-nationwide [accessed 2025-03-06]
9. Hannay E, Pai M. Breaking the cycle of neglect: building on momentum from COVID-19 to drive access to diagnostic testing. EClinicalMedicine 2023 Mar;57:101867. [doi: 10.1016/j.eclinm.2023.101867] [Medline: 36820097]
10. Rader B, Gertz A, Iuliano AD, et al. Use of At-Home COVID-19 Tests - United States, August 23, 2021-March 12, 2022. MMWR Morb Mortal Wkly Rep 2022 Apr 1;71(13):489-494. [doi: 10.15585/mmwr.mm7113e1] [Medline: 35358168]
11. Talbot HK, Martin ET, Gaglani M, et al. Coronavirus disease 2019 (COVID-19) versus influenza in hospitalized adult patients in the United States: Differences in demographic and severity indicators. Clin Infect Dis 2021 Dec 16;73(12):2240-2247. [doi: 10.1093/cid/ciab123] [Medline: 34050659]
12. Multiplex assays authorized for simultaneous detection of influenza viruses and SARS-cov-2 by FDA. CDC. 2025. URL: https://www.cdc.gov/flu/hcp/testing-methods/flu-covid19-detection.html [accessed 2025-06-26]
13. Health C for D and R Influenza Diagnostic Tests. Food and Drug Administration. 2024. URL: https://www.fda.gov/medical-devices/in-vitro-diagnostics/influenza-diagnostic-tests [accessed 2025-06-26]
14. McNeil C, Verlander S, Divi N, Smolinski M. The Landscape of Participatory Surveillance Systems Across the One Health Spectrum: Systematic Review. JMIR Public Health Surveill 2022 Aug 5;8(8):e38551. [doi: 10.2196/38551] [Medline: 35930345]
15. Wójcik OP, Brownstein JS, Chunara R, Johansson MA. Public health for the people: participatory infectious disease surveillance in the digital age. Emerg Themes Epidemiol 2014;11(1):7. [doi: 10.1186/1742-7622-11-7] [Medline: 24991229]
16. Home. Outbreaks Near Me. URL: https://outbreaksnearme.org/us/en-US [accessed 2024-03-31]
17. Rader B, White LF, Burns MR, et al. Mask-wearing and control of SARS-CoV-2 transmission in the USA: a cross-sectional study. Lancet Digit Health 2021 Mar;3(3):e148-e157. [doi: 10.1016/S2589-7500(20)30293-4] [Medline: 33483277]

XSL•FO
RenderX

18. Survey methodology: how surveymonkey gets its data. SurveyMonkey. URL: https://www.surveymonkey.com/mp/survey-methodology/ [accessed 2025-03-06]
19. Outbreaks near me: help track the COVID-19 pandemic. SurveyMonkey. URL: https://www.surveymonkey.com/r/TLXVW86 [accessed 2025-03-06]
20. ACS summary file table-based format. United States Census Bureau. URL: https://www.census.gov/programs-surveys/acs/data/summary-file.html [accessed 2025-07-01]
21. Ali SH, Tozan Y, Jones AM, Foreman J, Capasso A, DiClemente RJ. Regional and socioeconomic predictors of perceived ability to access coronavirus testing in the United States: results from a nationwide online COVID-19 survey. Ann Epidemiol 2021 Jun;58:7-14. [doi: 10.1016/j.annepidem.2021.03.001] [Medline: 33691088]
22. Lamsal R, Napit K, Rosen AB, Wilson FA. Paid Sick Leave and Healthcare Utilization in Adults: A Systematic Review and Meta-analysis. Am J Prev Med 2021 Jun;60(6):856-865. [doi: 10.1016/j.amepre.2021.01.009] [Medline: 33785275]
23. Zhai Y, Santibanez TA, Kahn KE, Black CL, de Perio MA. Paid sick leave benefits, influenza vaccination, and taking sick days due to influenza-like illness among U.S. workers. Vaccine (Auckl) 2018 Nov 19;36(48):7316-7323. [doi: 10.1016/j.vaccine.2018.10.039] [Medline: 30361122]
24. Federal employees - sick leave for personal medical needs. US Office of Personnel Management. URL: https://www.opm.gov/policy-data-oversight/pay-leave/leave-administration/fact-sheets/personal-sick-leave/ [accessed 2025-03-06]
25. Who will bear the brunt of lockdown policies? evidence from tele-workability measures across countries. International Monetary Fund. URL: https://tinyurl.com/28x7jpze [accessed 2025-03-06]
26. Hill HD. Paid Sick Leave and Job Stability. Work Occup 2013 May 1;40(2). [doi: 10.1177/0730888413480893] [Medline: 24235780]
27. How to talk to your employer about taking time off for family and medical reasons. US Department of Labor. URL: https://www.dol.gov/agencies/whd/fmla/how-to-talk-to-your-employer-about-leave [accessed 2025-03-06]
28. The Changing Threat of COVID-19 National Center for Immunization and Respiratory Diseases. Centers for Disease Control and Prevention. 2025 Jan 31. URL: https://www.cdc.gov/ncird/whats-new/changing-threat-covid-19.html [accessed 2025-03-07]

Abbreviations

AOR: adjusted odds ratio
ILI: influenza-like illness
MMWR: Morbidity and Mortality Weekly Report
ONM: Outbreaks Near Me

Edited by A Mavragani, T Sanchez; submitted 23.04.25; peer-reviewed by J Jiang, O Vlaicu; revised version received 14.07.25; accepted 15.07.25; published 26.08.25.

Please cite as:
Gertz A, Sopko JB, Remmel C, Rader B, Brownstein JS
Assessment of Influenza Testing Distribution in the United States for the 2021-2022, 2022-2023, and 2023-2024 Influenza Seasons: Online Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e76459
URL: https://publichealth.jmir.org/2025/1/e76459
doi:10.2196/76459

© Autumn Gertz, Juliana B Sopko, Christopher Remmel, Benjamin Rader, John S Brownstein. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

COVID-19 Testing Equity in New York City During the First 2 Years of the Pandemic: Demographic Analysis of Free Testing Data

Daniel Rosenfeld[1,2], MSci; Sean Brennan[1,3], MPH; Andrew Wallach[1,4], MD; Theodore Long[1,5], MHS, MD; Chris Keeley[1], MUP; Sarah Joseph Kurien[1], MPA

[1]NYC Test & Treat Corps, The New York City Health + Hospitals Corporation, 50 Water St., New York, NY, United States

[2]Department of Housing Stabilization, The Massachusetts Executive Office of Housing and Livable Communities, Boston, MA, United States

[3]Memorial Sloan Kettering Cancer Center, New York, NY, United States

[4]Division of General Internal Medicine and Clinical Innovation, New York University Grossman School of Medicine, New York, NY, United States

[5]Department of Population Health, New York University, New York, NY, United States

Corresponding Author:
Daniel Rosenfeld, MSci
NYC Test & Treat Corps, The New York City Health + Hospitals Corporation, 50 Water St., New York, NY, United States

Abstract

Background: COVID-19 has caused over 46,000 deaths in New York City, with a disproportional impact on certain communities. As part of the COVID-19 response, the city has directly administered over 6 million COVID-19 tests (in addition to millions of indirectly administered tests not covered in this analysis) at no cost to individuals, resulting in nearly half a million positive results. Given that the prevalence of testing, throughout the pandemic, has tended to be higher in more affluent areas, these tests were targeted to areas with fewer resources.

Objective: This study aimed to evaluate the impact of New York City's COVID-19 testing program; specifically, we aimed to review its ability to provide equitable testing in economically, geographically, and demographically diverse populations. Of note, in addition to the brick-and-mortar testing sites evaluated herein, this program conducted 2.1 million tests through mobile units to further address testing inequity.

Methods: Testing data were collected from the in-house Microsoft SQL Server Management Studio 18 Clarity database, representing 6,347,533 total tests and 449,721 positive test results. These tests were conducted at 48 hospital system locations. Per capita testing rates by zip code tabulation area (ZCTA) and COVID-19 positivity rates by ZCTA were used as dependent variables in separate regressions. Median income, median age, the percentage of English-speaking individuals, and the percentage of people of color were used as independent demographic variables to analyze testing patterns across several intersecting identities. Negative binomial regressions were run in a Jupyter Notebook using Python.

Results: Per capita testing inversely correlated with median income geographically. The overall pseudo r^2 value was 0.1101 when comparing hospital system tests by ZCTA against the selected variables. The number of tests significantly increased as median income fell (SE 1.00000155; $P<.001$). No other variables correlated at a significant level with the number of tests (all P values were >.05). When considering positive test results by ZCTA, the number of positive test results also significantly increased as median income fell (SE $1.57e^{-6}$; $P<.001$) and as the percentage of female residents fell (SE 0.957; $P=.001$). The number of positive test results by ZCTA rose at a significant level alongside the percentage of English-only speakers (SE 0.271; $P=.03$).

Conclusions: New York City's COVID-19 testing program was able to improve equity through the provision of no-cost testing, which focused on areas of the city that were disproportionately impacted by COVID-19 and had fewer resources. By detecting higher numbers of positive test results in resource-poor neighborhoods, New York City was able to deploy additional resources, such as those for contact tracing and isolation and quarantine support (eg, free food delivery and free hotel stays), early during the COVID-19 pandemic. Equitable deployment of testing is feasible and should be considered early in future epidemics or pandemics.

(JMIR Public Health Surveill 2025;11:e52972) doi:10.2196/52972

KEYWORDS

COVID-19 testing; health disparities; equity in testing; New York City; socioeconomic factors; testing accessibility; health care inequalities; demographic analysis; COVID-19 mortality; coronavirus; SARS-CoV-2; pandemic; equitable testing; cost; poor neighborhood; resources

XSL•FO
RenderX

Introduction

In March 2020, New York City (NYC) emerged as one of the global epicenters of SARS-CoV-2 infections and COVID-19 deaths. This was roughly 2 months after the first detected case in the United States and 3 months following the first case in Wuhan, China [1]. Early in the pandemic, within the United States and especially within NYC, testing availability was extremely limited. In fact, there were no positive test results recorded in NYC until the beginning of March 2020 [2]. Over the ensuing 3 years, over 45,000 deaths and over 3.2 million infections were documented in NYC—a wave of mortality that contributed to a decline of 4.6 years in life expectancy from 2019 to 2020 alone [3].

It should be noted that COVID-19 mortality and SARS-CoV-2 infection rates have varied widely based on several spatial and socioeconomic factors [4]. US counties with higher poverty rates had higher COVID-19 case numbers and related death rates, and racial and ethnic minority individuals are at increased vulnerability for COVID-19 when considering both infection rates and mortality rates [5,6]. While life expectancy declined by 4.6 years citywide, it declined by 6 years for Hispanic and Latinx New Yorkers (to 77.3 years), by 5.5 years for Black New Yorkers (to 73 years), and by 3 years for White New Yorkers (to 80.1 years) [3]. These mortality disparities were not limited to race. Individuals with chronic conditions, such as obesity and diabetes, faced higher mortality rates than those faced by the general population as a whole, while socioeconomic factors, such as poverty, housing overcrowding, the effects of historical residential racial segregation, and an inability to perform some jobs remotely, have intersected with the aforementioned chronic health issues to create a perfect storm of increased COVID-19 mortality in particular neighborhoods [7].

Following the initial lack of testing at the onset of the pandemic, the federal and state governments sought to develop and implement a mass testing strategy. Once tests were made available, significant barriers to accessing testing remained in terms of costs and the geographic locations of testing sites. The NYC Health + Hospitals Corporation (H+H)—the largest public health system in the United States, with 11 acute care facilities, 6 diagnostic and treatment centers, and more than 50 neighborhood health centers (mostly, but not entirely, in economically depressed neighborhoods)—became the largest free testing provider in NYC. The NYC Department of Health and Mental Hygiene has always tracked health and socioeconomic disparities. This background provided the basis for establishing the TRIE (Taskforce on Racial Inclusion & Equity) neighborhoods [8]. These were neighborhoods that had economic disadvantages and were disproportionately impacted by COVID-19. We leveraged the neighborhoods, which essentially involved combining former disparity data related to the impact of COVID-19, to help set our strategy for allocating testing sites. By providing access to no-cost COVID-19 testing, we sought to understand whether our equitably distributed testing program had a relationship with health outcomes and whether the no-cost testing option can be used to increase testing in underserved areas.

Methods

Data Sources

Data on COVID-19 testing that was directly administered by the NYC H+H were collected from the in-house NYC H+H Microsoft SQL Server Management Studio 18 Clarity database, representing 6,347,533 total tests and 449,721 positive test results. These tests were sourced from the in-house NYC H+H instance of Epic Systems; these did not include tests conducted by external testing vendors contracted by the NYC H+H or tests conducted by the school testing program that mandated surveillance testing in NYC public schools from 2020 to 2022. Demographic data were taken from 2020 US Census Bureau data [9]. Citywide testing data came from the NYC coronavirus data repository [10].

Experimental Design

We initially sought to run a Poisson regression analysis comparing counts of tests and positive test results to demographic variables while using population as an offset; however, we found the data to be overdispersed and chose to use negative binomial regressions. In terms of the variables we chose to analyze, we were influenced by two studies from King County, Washington, and NYC that, among other methods, analyzed testing rates against demographic variables. The King County, Washington, study used a variable called “people of color (POC),” into which POC racial data were collapsed due to collinearity among POC groups [11]. Given similar patterns of collinearity in our data, we chose to do this in our study. The NYC study used a number of different demographic variables, such as income, education, and housing overcrowding, to calculate a single socioeconomic index score [2]. Although this is one way to handle collinearity between socioeconomic variables, such as income and poverty, we chose to use median income as our main predictor of socioeconomic status. As variables related to COVID-19 mortality and testing rates, we included age and sex breakdowns, as being older and being male positively correlate with higher mortality. As a measure of presumed difficulty in accessing health care, the percentage of English speakers was included.

Ethical Considerations

We used deidentified patient data. The Biomedical Research Alliance of New York (BRANY) Institutional Review Board (IRB) determined that our research in this paper did not constitute research involving human subjects that is regulated by US Department of Health and Human Services or US Food and Drug Administration regulations, and it was therefore not subject to further BRANY IRB review. The BRANY study ID is 23-15-564, the sponsor ID is RH2023-814, and our category was determined as “15-Not Human Subject Research.” We were advised by BRANY that we did not have to create a National Clinical Trial number with the US National Library of Medicine ClinicalTrials.gov website because our data-only study used preexisting data.

Analysis

Regressions were run in a Jupyter Notebook (Project Jupyter) using Python. Data visualizations were created by using Tableau

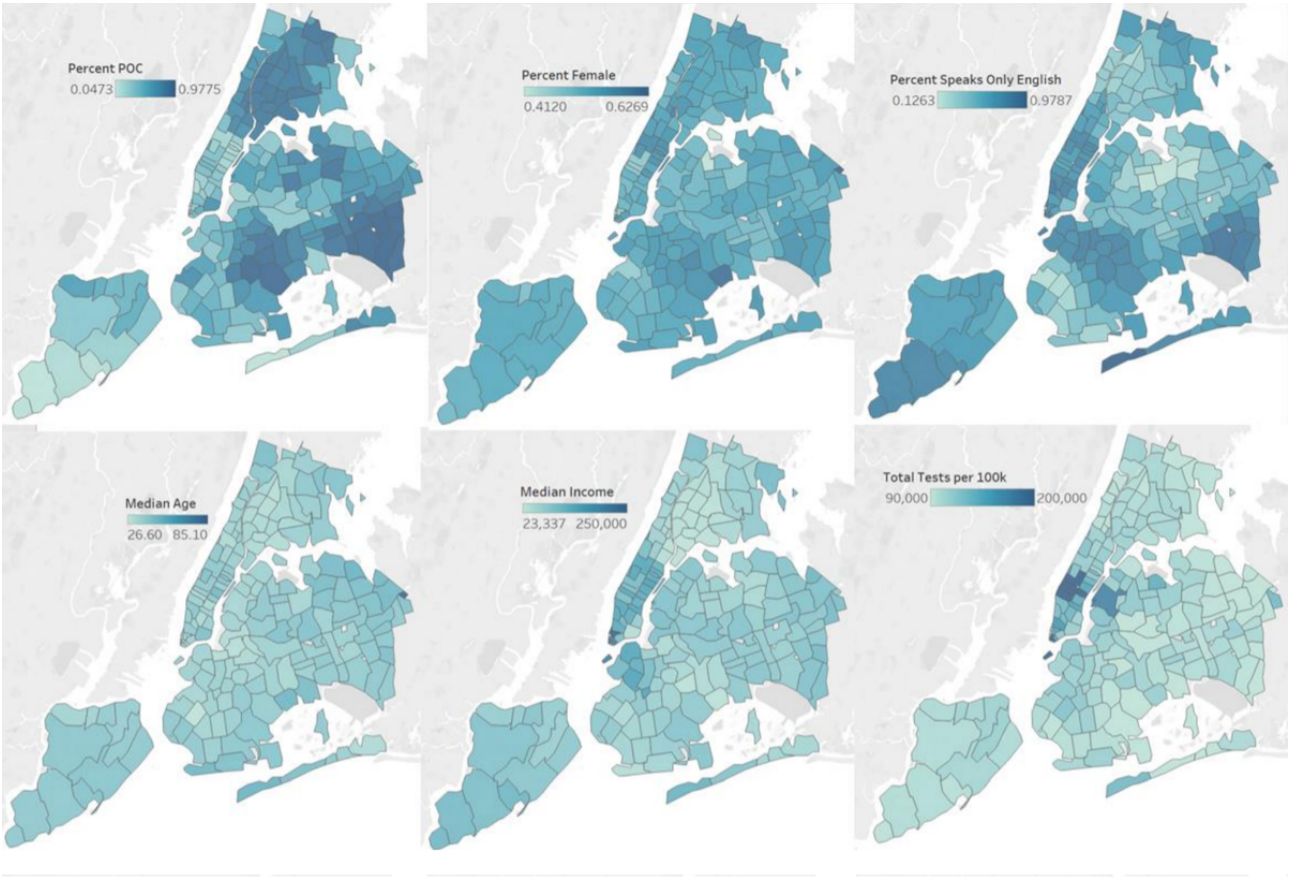
XSL•FO
RenderX

(Tableau Software LLC). A selection of data visualizations can be found in the *Results* section. As this was a spatial analysis, we took care to map each of our demographic variables, testing data points, and background health outcome data. We mapped out "Percent POC," "Percent Female," "Percent Speaks Only English," "Median Age," "Median Income," and "COVID-19 death rate" as demographic and health outcome variables by zip code tabulation area (ZCTA). To spatially display our testing data, we mapped the variables "Total Citywide Tests per 100,000," "NYC H+H COVID-19 Tests," "NYC H+H COVID-19 tests per 100,000," "NYC H+H positive tests per 100,000," and "NYC H+H COVID-19 test positivity rate" also by ZCTA. We also created scatter plots to illustrate the relationships between these variables.

Results

The overall pseudo r^2 value was 0.1101 when comparing NYC H+H tests by ZCTA against the selected variables. The number of NYC H+H tests significantly increased as median income fell (exponentiated β=.99998, SE 1.00000155; *P*<.001). No other variables correlated at a significant level with the number of tests (all *P* values were >.05). When considering positive test results by ZCTA, the number of positive test results also significantly increased as median income fell (exponentiated β=.99998, SE $1.57e^{-6}$; *P*<.001) and as the percentage of female residents fell (exponentiated β=.04014, SE 0.957; *P*=.001). Positive test results by ZCTA rose at a less significant level alongside the percentage of English-only speakers (β=1.81630, SE 0.271; *P*=.03). No other predictor variables correlated with positive test results by ZCTA (all *P* values were >.05). *P*<.05 was chosen as a standard biomedical level of significance for geographic area–level research. A selection of data visualizations depicting our results are shown in Figures 1-6.

Figure 1. Selected demographic, testing, and mortality variables by zip code tabulation areas (clockwise from upper left): "Percent POC," "Percent Female," "Percent Speaks Only English," "Total Tests per 100k," "Median Income" (US $), and "Median Age" (years). POC: people of color.



XSL•FO
RenderX

Figure 2. Selected demographic, testing, and mortality variables by zip code tabulation areas continued (clockwise from upper left): "NYC H+H COVID-19 Tests," "NYC H+H Positive Tests," "NYC H+H COVID-19 test positivity rate," "COVID-19 death rate," "NYC H+H positive tests per 100,000," and "NYC H+H COVID-19 tests per 100,000." NYC H+H: New York City Health + Hospitals Corporation.

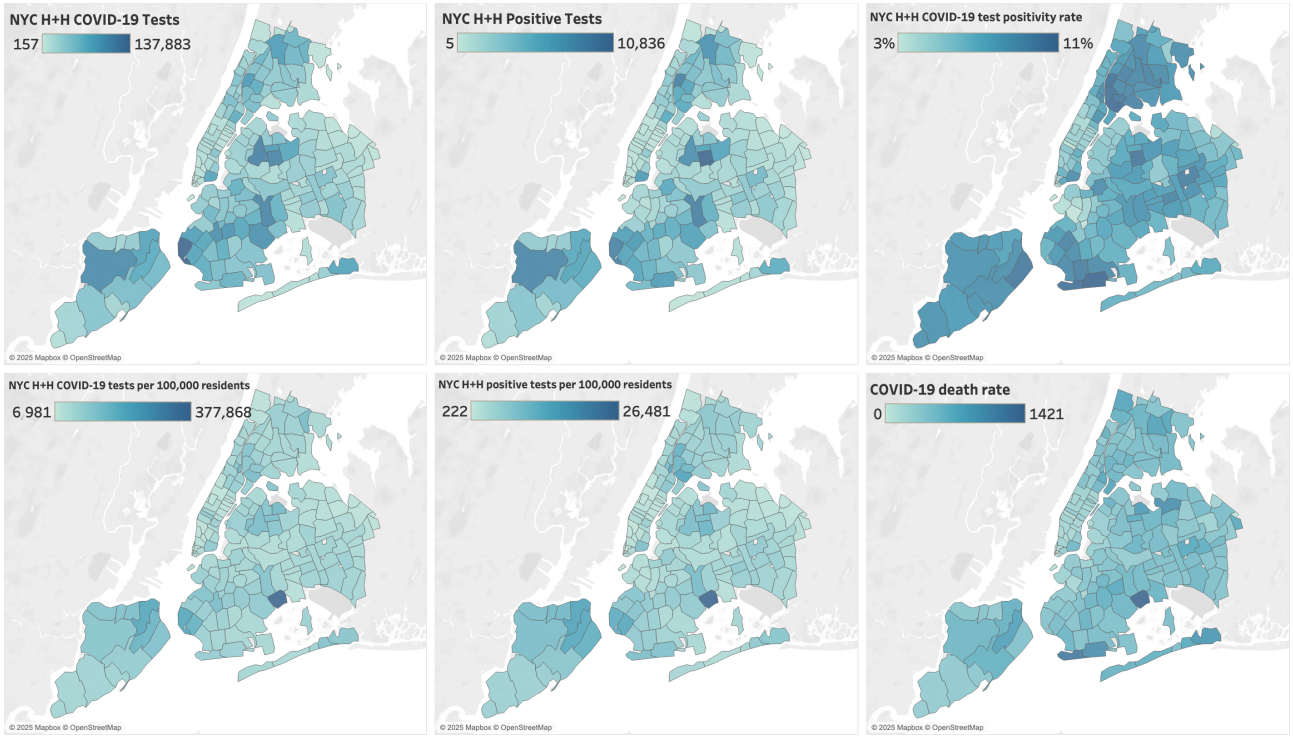
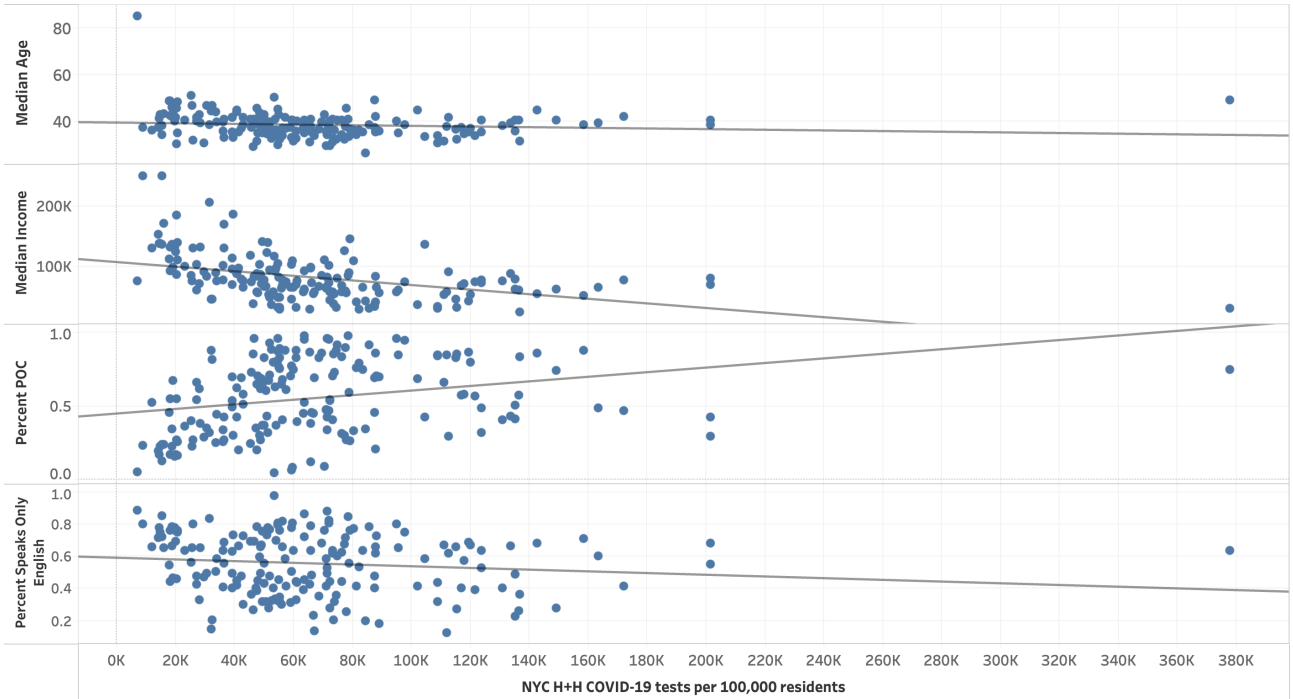


Figure 3. Selected demographics compared to NYC H+H COVID-19 tests per 100,000 residents by zip code tabulation area (from top to bottom): "Median Age" (years), "Median Income" (US $), "Percent POC," and "Percent Speaks Only English" by NYC H+H COVID-19 tests per 100,000 residents. NYC H+H: New York City Health + Hospitals Corporation; POC: people of color.



XSL•FO
RenderX

Figure 4. Selected demographics compared to NYC H+H COVID-19 test positivity rate by zip code tabulation area (from top to bottom): "Median Age" (years), "Median Income" (US $), "Percent POC," and "Percent Speaks Only English" by NYC H+H test positivity rate. NYC H+H: New York City Health + Hospitals Corporation; POC: people of color.

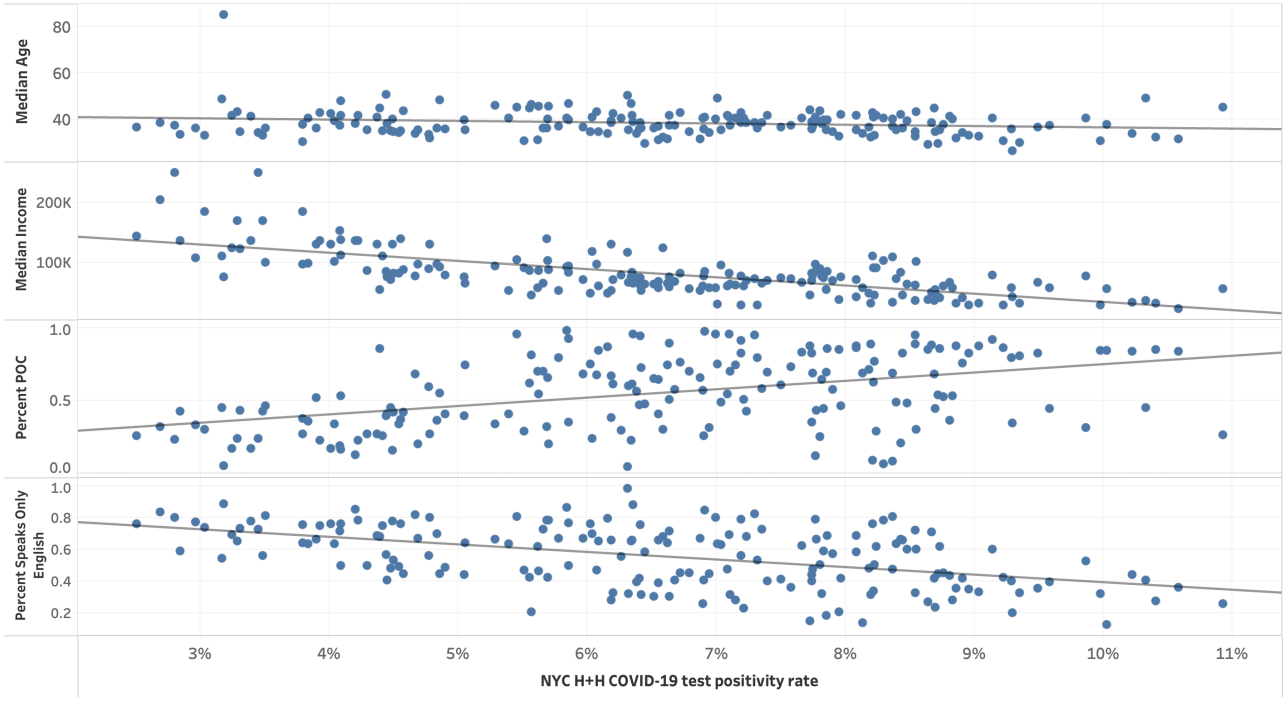
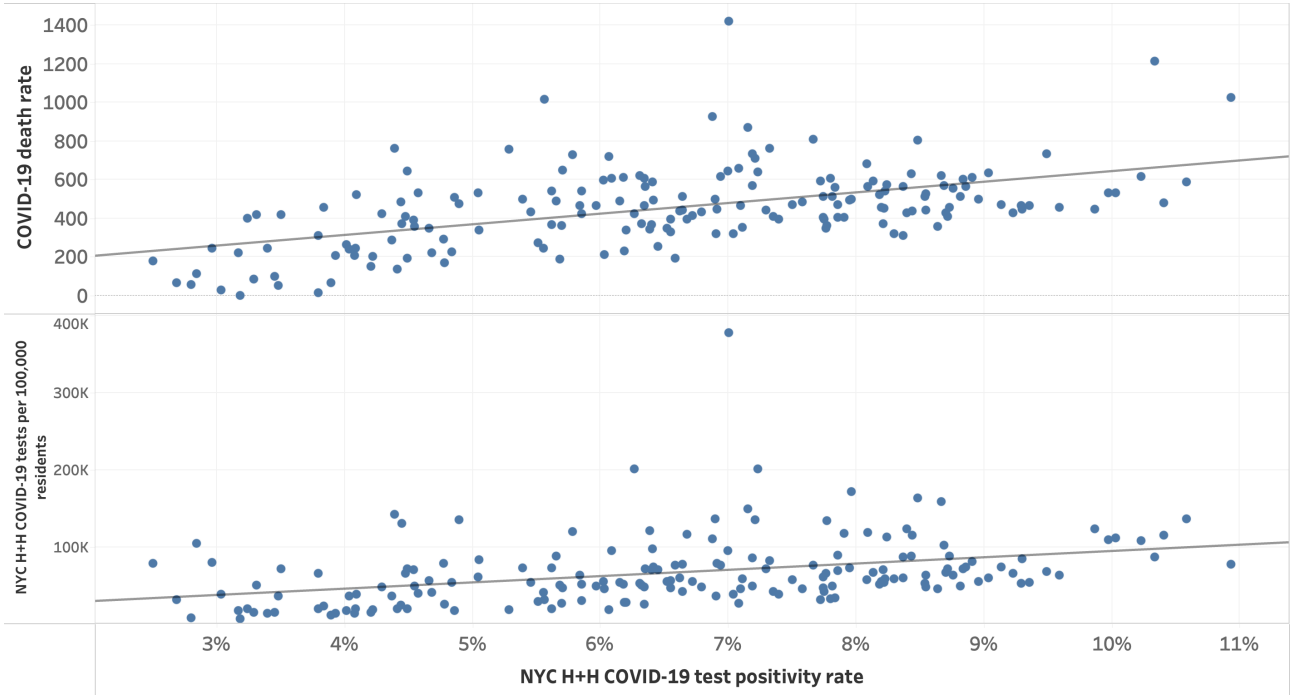


Figure 5. Selected test and death rates compared to NYC H+H COVID-19 tests per 100,000 residents by zip code tabulation area (from top to bottom): NYC H+H COVID-19 test positivity rate and COVID-19 death rate by NYC H+H COVID-19 tests per 100,000 residents. NYC H+H: New York City Health + Hospitals Corporation.

XSL•FO
RenderX

Figure 6. Selected test and death rates compared to NYC H+H COVID-19 positivity rate by zip code tabulation area (from top to bottom): COVID-19 death rate and NYC H+H COVID-19 tests per 100,000 residents by NYC H+H COVID-19 positivity rate. NYC H+H: New York City Health + Hospitals Corporation.



Discussion

The devastatingly high mortality rates seen in NYC during the height of the first COVID-19 wave in April 2020 were likely linked to undetected cases as a result of limited testing capabilities at the time. Roughly half of all deceased NYC patients with COVID-19 died during the first months of the first wave, making comparative analysis somewhat difficult in NYC when compared to other regions. Approximately 6 million COVID-19 tests were completed by the NYC H+H between early 2020 and late 2022, though most testing was done following the first wave in 2020 due to the time and resources required to set up testing sites. In our analysis, the number of NYC H+H no-cost COVID-19 tests increased as the median income of a discrete neighborhood fell—a trend that stood in opposition to overall citywide testing trends, wherein the highest testing rates were observed in wealthier areas of the city. The number of positive test results at the NYC H+H also rose as median income fell in a discrete neighborhood, even when controlling for other variables. This suggests that NYC H+H no-cost COVID-19 tests were equitably distributed in targeted increased-risk neighborhoods of NYC. This was important, considering the disproportionate negative health impacts of the pandemic on these neighborhoods, which already had a preexisting vulnerability to a range of adverse health outcomes.

Consistent and easily accessible large-scale testing may be a crucial element of successful strategies for avoiding additional mortality. Our research echoes that of some other studies that found an inverse relationship between COVID-19 positivity and test availability [11]. With regard to the racial and income equity issues that have been seen in COVID-19 testing within NYC and the United States, our findings on tests administered by the NYC H+H were more mixed for demographic variables, such as "Percent POC," that otherwise might be related to COVID-19 mortality risk. Our study had limitations with regard to interpreting disparities in COVID-19 testing rates across space and disparities in time to COVID-19 mortality. COVID-19 testing was extremely limited early in the pandemic due to the lack of readily available testing, whereas from 2022 to 2023, at-home rapid testing became ubiquitous but came with the caveat that results were rarely reported. In addition, as we compared COVID-19 testing rates and results between geographic and demographic variables, we recognized the pitfalls that may arise when equating jurisdiction-level data with individuals who live within those jurisdictions. NYC is an extremely diverse city with substantial income inequality and variability even within the smallest geographic neighborhoods. The relationship between COVID-19 test positivity and the percentage of female residents could suggest that ZCTAs with a higher percentage of male residents had higher positivity rates, and this information could be used to inform future testing campaign strategies (eg, targeting areas with a higher percentage of male residents).

We take the inverse relationship between NYC H+H testing rates and median income as a positive sign that the no-cost COVID-19 testing program was able to penetrate areas of the city that were likely neglected by other testing providers. Despite the devastating consequences from the COVID-19 pandemic in NYC, this targeted, equitably distributed testing plan could have broader public health implications if similar outreach strategies are used in future pandemics.

XSL•FO
RenderX

Acknowledgments

We thank Kym Neck, Anna Gilbert, Cornelia Mills, Taehan Lee, Sharhzad Divsalar, Chi Gao, Cara Feldkamp, George Zhuo, Alexander Dobranic, Christina Pili, and George Cowan from the New York City Health + Hospitals Corporation (NYC H+H) for their help in completing this paper. We also thank Raffaella Hart, Laura Donohue, Linda Reuter, and Svetlana Abramova from Biomedical Research Alliance of New York (BRANY) for their assistance in the institutional review board review process, as well as Dr Bob Nolan, who provided guidance during the submission process.

Data Availability

All underlying data by zip code tabulation area can be found on GitHub (Multimedia Appendix 1) [12,13].

Conflicts of Interest

None declared.

Multimedia Appendix 1
Python code.
[DOCX File, 14 KB - publichealth_v11i1e52972_app1.docx]

References

1. Thompson CN, Baumgartner J, Pichardo C, et al. COVID-19 outbreak - New York City, February 29-June 1, 2020. MMWR Morb Mortal Wkly Rep 2020 Nov 20;69(46):1725-1729. [doi: 10.15585/mmwr.mm6946a2] [Medline: 33211680]
2. Lieberman-Cribbin W, Tuminello S, Flores RM, Taioli E. Disparities in COVID-19 testing and positivity in New York City. Am J Prev Med 2020 Sep;59(3):326-332. [doi: 10.1016/j.amepre.2020.06.005] [Medline: 32703702]
3. 2020 vital statistics report reveals toll of COVID-19 on New Yorkers and life expectancy. City of New York. 2023 Apr 7. URL: https://www.nyc.gov/site/doh/about/press/pr2023/2020-vital-statistics-report-reveals-toll-of-covid-19.page [accessed 2025-06-13]
4. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis 2020 May;20(5):533-534. [doi: 10.1016/S1473-3099(20)30120-1] [Medline: 32087114]
5. Liao TF, De Maio F. Association of social and economic inequality with coronavirus disease 2019 incidence and mortality across US counties. JAMA Netw Open 2021 Jan 4;4(1):e2034578. [doi: 10.1001/jamanetworkopen.2020.34578] [Medline: 33471120]
6. Parolin Z, Lee EK. The role of poverty and racial discrimination in exacerbating the health consequences of COVID-19. Lancet Reg Health Am 2022 Mar;7:100178. [doi: 10.1016/j.lana.2021.100178] [Medline: 35018358]
7. Bilal U, Tabb LP, Barber S, Diez Roux AV. Spatial inequities in COVID-19 testing, positivity, confirmed cases, and mortality in 3 U.S. cities : an ecological study. Ann Intern Med 2021 Jul;174(7):936-944. [doi: 10.7326/M20-3936] [Medline: 33780289]
8. About Taskforce on Racial Inclusion & Equity. City of New York. URL: https://www.nyc.gov/site/trie/about/about.page [accessed 2024-08-07]
9. S2502: demographic characteristics for occupied housing units. American Community Survey. 2022: ACS 5-year estimates subject tables. United States Census Bureau. 2022. URL: https://data.census.gov/table/ACSST5Y2022.S2502?q=S2502 [accessed 2025-03-02]
10. NYC Department of Health and Mental Hygiene. coronavirus-data/data-by-modzcta.csv at master · nychealth/coronavirus-data. GitHub. URL: https://github.com/nychealth/coronavirus-data/blob/master/totals/data-by-modzcta.csv [accessed 2022-09-29]
11. Seto E, Min E, Ingram C, Cummings BJ, Farquhar SA. Community-level factors associated with COVID-19 cases and testing equity in King County, Washington. Int J Environ Res Public Health 2020 Dec 18;17(24):9516. [doi: 10.3390/ijerph17249516] [Medline: 33353095]
12. testingequitydata / 2022-12-02 updated testing equity python code.ipynb. GitHub. URL: https://github.com/rosenfed/testingequitydata/blob/main/2022-12-02%20updated%20testing%20equity%20python%20code.ipynb [accessed 2025-02-27]
13. testingequitydata / 2022-10-12_acsimport.xlsx. GitHub. URL: https://github.com/rosenfed/testingequitydata/blob/main/2022-10-12_acsimport.xlsx [accessed 2025-02-27]

Abbreviations

BRANY: Biomedical Research Alliance of New York
H+H: Health + Hospitals Corporation
IRB: Institutional Review Board
NYC: New York City

XSL•FO
RenderX

POC: people of color
TRIE: Taskforce on Racial Inclusion & Equity
ZCTA: zip code tabulation area

Edited by A Mavragani; submitted 20.09.23; peer-reviewed by F Kaiksow, S Mathew; revised version received 08.09.24; accepted 10.09.24; published 13.03.25.

Please cite as:
Rosenfeld D, Brennan S, Wallach A, Long T, Keeley C, Kurien SJ
COVID-19 Testing Equity in New York City During the First 2 Years of the Pandemic: Demographic Analysis of Free Testing Data
JMIR Public Health Surveill 2025;11:e52972
URL: https://publichealth.jmir.org/2025/1/e52972
doi:10.2196/52972

© Daniel Rosenfeld, Sean Brennan, Andrew Wallach, Theodore Long, Chris Keeley, Sarah Joseph Kurien. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 13.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Predictors of Loneliness and Psychological Distress in Older Adults During the COVID-19 Pandemic: National Cross-Sectional Study

Rahela Orlandini[1,2], MSN; Antonela Matana[1], PhD; Deana Švaljug[3], PhD; Ivana Gusar[4], PhD; Vesna Antičević[1], PhD

[1]Faculty of Health Sciences, University of Split, Ruđera Boškovića 35, Split, Croatia

[2]School of Medicine, University of Split, Split, Croatia

[3]Faculty of Health Studies, University of Rijeka, Rijeka, Croatia

[4]Department of Health Studies, University of Zadar, Zadar, Croatia

Corresponding Author:
Rahela Orlandini, MSN
Faculty of Health Sciences, University of Split, Ruđera Boškovića 35, Split, Croatia

Abstract

Background: The COVID-19 pandemic has significantly affected the mental health of older adults, particularly through increased loneliness and psychological distress. While various contributing factors have been studied, the role of preference for solitude as a potential predictor and mediator remains poorly understood.

Objective: This national cross-sectional study aimed to examine predictors of loneliness and psychological distress among older adults during the pandemic, with a specific focus on preference for solitude and its mediating role between pandemic-specific stressors and self-efficacy.

Methods: A total of 2053 Croatian residents aged 65 years and older were recruited using snowball sampling. Validated instruments were used, including the UCLA Loneliness Scale, Preference for Solitude Scale, Clinical Outcomes in Routine Evaluation-Outcome Measure, General Self-Efficacy Scale, and Pandemic-Specific Stressors Questionnaire for Older Adults. Hierarchical regression and path analysis were used, with statistical significance set at $P<.05$.

Results: For loneliness, the final model explained 30.7% of the variance, with a coefficient of determination (R^2) of 0.307 and a root mean square error of 0.708 ($P<.001$). Significant predictors included marital status (eg, never married: B=0.390, $P<.001$; psychological problems: B=0.020, $P<.001$; functionality: B=−0.037, $P<.001$; and social distancing: B=0.014, $P<.001$). Preference for solitude was also a significant predictor of loneliness (B=0.011; $P<.001$). For psychological distress, the final model explained 30.8% of the variance (R^2=0.308; root mean square error=14.872; $P<.001$). Self-efficacy emerged as the strongest negative predictor of distress (B=−1.066; $P<.001$), whereas preference for solitude was a positive predictor (B=2.403; $P<.001$). The variable "spending several hours alone per day" was associated with lower levels of distress (B=−3.509; $P=.003$), while "frequent or superficial interactions with acquaintances (eg, at least once a week)" were related to higher distress (B=4.321, $P=.002$). Path analysis revealed that both social distancing and exposure to infection had a significant direct effect on self-efficacy: negative in the case of social distancing (β=−.391; $P<.001$) and positive for exposure to infection (β=.386; $P<.001$). However, preference for solitude did not significantly mediate either relationship, as indicated by nonsignificant indirect effects (β=−.006; $P=.09$ and β=.002; $P=.48$, respectively).

Conclusions: Psychological problems and reduced functionality emerged as the strongest predictors of loneliness among older adults during the pandemic. Self-efficacy was the most important protective factor against psychological distress. Although preference for solitude may have adaptive benefits, in the context of this study, it was associated with increased loneliness and distress during enforced isolation. These findings suggest that public health interventions should balance respect for individual preferences with the provision of active support for vulnerable populations during crises.

(JMIR Public Health Surveill 2025;11:e78728) doi:10.2196/78728

KEYWORDS

older adults; COVID-19; loneliness; psychological distress; preference for solitude; pandemic-specific stressors; self-efficacy

Introduction

In today's uncertain times, it is essential to continually improve the well-being of older adults by learning from past experiences, such as the COVID-19 pandemic.

During the COVID-19 pandemic, periods of social isolation highlighted how personal preference for solitude (PS) versus feelings of loneliness can impact mental health. Differentiating solitude from social isolation and loneliness is essential for studying psychological well-being in older adults. Solitude is

XSL•FO
RenderX

the objective state of being alone, which can feel positive, negative, or mixed, while social isolation and loneliness are commonly associated with negative emotions. Social isolation is the objective lack of social connections, and loneliness is the subjective perception of insufficient social contact; both are associated with depression, anxiety, and stress (as cited by Peng et al [1]). This paper explores the effects of involuntary social isolation and solitude resulting from pandemic-related measures on loneliness and the mental health of older adults.

Loneliness and psychological distress are significant mental health concerns, particularly among older adults during global crises [2]. Due to social isolation during the COVID-19 lockdown, older adults experienced significantly higher levels of loneliness [3,4] and psychological distress [5] compared with the pre-COVID period. This is supported by studies showing substantial increases in measurement scores [3,4]. Parlapani et al [6] demonstrate that older adults are not uniformly vulnerable to psychological distress during COVID-19; however, their capacity to adapt to adversity varies considerably. This variation depends on social, economic, and individual factors, aligning with previous findings on predictors of loneliness and psychological distress during the pandemic. There is substantial evidence indicating similar factors contribute to both loneliness and psychological distress.

Regarding sociodemographic factors, some research found that being female was associated with higher rates of distress and loneliness [7-9], while others did not observe this trend [10,11]. Similarly, some studies have found that advanced age in older adults is associated with greater loneliness [12-14] and psychological distress [9], but other studies reported the opposite [8,15]. Being unpartnered and living alone were consistent risk factors for loneliness [14,16] and psychological distress [17], while environmental factors such as housing type and rurality remain less studied. For older adults in Croatia, factors contributing to loneliness and stress included social isolation, movement restrictions, and intensive media coverage of the pandemic [18]. A pilot study on Croatian citizens found that older adults reported less loneliness during the pandemic if they lived with others, were more emotionally stable, had fewer psychological problems and pandemic-specific stressors, and engaged in fewer risky behaviors [19]. This study is now being expanded to a representative national sample of all older Croatian adults. In addition to loneliness, where the same predictors from the pilot study were retained, researchers are also examining predictors of psychological distress, focusing on time spent alone during lockdown, PS, and self-efficacy. Research shows that lifestyle changes, such as increased isolation and reduced social activities, disrupted older adults' routines and harmed their mental health. One study found that those confined for more than 50 days were more likely to experience depression, with fear of infection linked to greater confusion, fatigue, tension, and reduced vigor [20]. Social isolation has also been connected to a wide range of adverse outcomes in older adults, including lower quality of life and cognitive decline. Other negative effects include loss of muscle mass, multimorbidity, and increased disability [21].

Studies also highlight the role of self-efficacy and personal perceptions in shaping psychological distress among older adults during lockdowns, emphasizing the need for adaptive coping strategies to support mental health. One study found that self-efficacy in the general population shaped coping mechanisms and distress levels during the initial lockdown [22]. A longitudinal study in Spain showed that older adults with more positive self-perceptions of aging reported less distress [9]. Research in Catalonia on frail older adults identified prelockdown depressive symptoms and engagement in leisure activities as important predictors of distress. Although self-efficacy was not directly measured, personal resources and coping strategies were found to be essential for reducing distress [23].

In addition to examining the well-established predictors of loneliness and psychological distress, this study introduces PS as a potential contributing factor to mental health outcomes in older adults. Although PS has been studied in other age groups [24,25], its impact on loneliness and psychological distress among older adults during periods of social isolation has remained under-researched. This gap in the literature presents an opportunity to explore the role of PS in older adults' mental well-being. PS, defined as the voluntary pursuit of time alone that is both enjoyable and productive [26,27], may affect well-being during periods of extreme social distancing. PS typically involves voluntary social isolation, where a person may enjoy solitude without experiencing loneliness, just as one can feel lonely even when surrounded by others [28]. This concept highlights the distinction between being alone and feeling lonely, which is essential for understanding how different social dynamics affect mental health. Research on the effects of PS reveals a notably inconsistent and complex picture. Some studies suggest that enjoying solitude can reduce negative emotions and improve emotional well-being [29]. However, other research underscores that the impact of solitude on well-being is highly context-dependent and not straightforward. For older adults, certain dimensions of solitude—for example, feeling productive when alone—have been linked to greater life satisfaction and increased positive affect [30,31]. Conversely, PS is also frequently associated with feelings of loneliness, which can undermine subjective well-being [30]. This contradictory evidence illustrates a puzzling inconsistency—while PS might buffer some negative effects of social isolation, it does not consistently translate into better mental health outcomes. This phenomenon, termed the "preference for solitude paradox" [32], highlights that PS does not always provide the expected psychological benefits, especially during times of widespread isolation, such as global crises [27]. Overall, these mixed findings emphasize the urgent need for a deeper investigation into how and when PS influences mental health in older adults.

The empirical findings presented in the "Introduction" section highlight gaps in existing knowledge that will be addressed by answering the research aims.

First, this study explored personal and situational factors contributing to loneliness in older adults during COVID-19, extending regional findings to a national level. Second, it examined how pandemic-induced changes in social routines (eg, spending time alone or socializing with others), PS, and self-efficacy predicted psychological distress. Finally, the study

XSL•FO
RenderX

investigated the mediating role of PS in the relationship between pandemic stressors and self-efficacy, focusing on its contribution to maintaining functionality during the pandemic.

Methods

Overview

This study used a cross-sectional, survey-based research design with a nonrandomized sampling method. The study design, data collection, and reporting followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for cross-sectional studies [33].

Participants

Participants were Croatian citizens of both genders, aged 65 years and older, living in Croatia. To define older adults, we used a cutoff age of 65 years [34]. The sample comprised 2053 participants.

As shown in Table 1, the sample was mostly composed of older adults, with a majority being female, residing in urban areas, and either married or widowed. More specifically, the median age of the participants was 74 (IQR 69-81) years. Regarding gender distribution, women were more represented in the sample (n=1324, 64.49%) than men (n=729, 35.51%). Most participants lived in urban areas (n=1585, 76.57%), while a smaller proportion lived in rural areas (n=468, 22.61%). In terms of marital status, about half of the participants were married (n=1000, 48.71%). A significant proportion were widowed for over 1 year (n=745, 36.29%), while smaller percentages were recently widowed (less than 3 months: n=65, 3.17%; 3 months to 1 year: n=56, 2.73%). In addition, 97 (4.73%) of participants had never been married, and 90 (4.38%) were divorced.

Table . Sociodemographic characteristics of participants in a cross-sectional study of older adults in Croatia during the COVID-19 pandemic (2021).

Variables	Descriptive statistic
Age (years), median (IQR)	74.0 (69-81)
Gender, n (%)	
Man	729 (35.51)
Women	1324 (64.49)
Place of residence, n (%)	
Urban	1585 (76.57)
Rural	468 (22.61)
Marital status, n (%)	
Married	1000 (48.71)
Divorced	90 (4.38)
Never married	97 (4.73)
Widowed (<3 months)	65 (3.17)
Widowed (3 months to 1 year)	56 (2.73)
Widowed (>1 year)	745 (36.29)

Procedure

Participants were recruited from community settings using a snowball sampling technique [35,36]. Initially, health students from 6 Croatian cities (Zagreb, Split, Rijeka, Zadar, Dubrovnik, and Osijek), each guided by a university professor, were instructed to identify 6 individuals aged 65 years and older among their relatives, ensuring inclusion of both sexes. Eligible participants were community-dwelling older adults who were able to understand the study purpose and complete the questionnaire independently. Each recruited participant was then asked to invite up to 6 peers from their own social circles (friends or relatives) who met the same criteria. Data collection continued until the target sample size (n=2053) was achieved, calculated using the Cochrane formula for continuous data [37].

Before participation, potential respondents were informed about the study aims, the voluntary nature of participation, and the guarantee of anonymity and confidentiality. Anonymity of the respondents was ensured by placing each completed questionnaire in a sealed envelope. After filling out the questionnaires, respondents returned the sealed envelopes to the student. The students then delivered the envelopes to their supervising professor, who subsequently forwarded them to the principal investigator. The research was conducted between April and October 2021 in nonclinical, community-based environments (primarily private households).

Measures

The following measures were used to collect data and address the research objectives:

Revised UCLA Loneliness Scale (Croatian Short-Form Version)

This 7-item scale measures loneliness as a general unidimensional construct, defined as an unpleasant emotional experience that arises when an individual's needs for intimacy, love, and belonging are unmet. Participants rate each item on a 5-point Likert scale, ranging from 1 (not at all true of me) to 5 (completely true of me). The total score was calculated by

XSL•FO
RenderX

averaging all item scores and dividing by the number of items [38]. The unidimensional structure and high reliability of the scale have been confirmed in both international and Croatian samples. In the Croatian sample, Cronbach α was 0.84 [39].

PS Scale (Croatian Adapted Version)

This 7-item questionnaire measures the degree of PS, with each item rated on a 5-point Likert scale. According to Burger [26], the scale demonstrated internal consistency (Cronbach α=0.70-0.73) and adequate validity for measuring PS. Further validation by Cramer and Lake [40] yielded similar internal consistency (Cronbach α=0.74-0.75) and test-retest reliability (r=0.76) over a 6- to 8-week interval. In the Croatian version, reliability coefficients ranged from Cronbach α=0.74-0.79 [41].

Clinical Outcomes in Routine Evaluation–Outcome Measure (CORE-OM; Croatian Version)

This instrument assesses psychological distress through 34 self-report items covering 4 dimensions: subjective well-being (4 items), problems or symptoms (12 items), daily functioning (12 items), and risky behaviors (6 items). These domains represent different aspects of psychological distress and dysfunction. Participants rate how frequently they experienced the described feelings during the past 2 weeks, on a scale from 0 (never) to 4 (almost always). Higher scores indicate greater psychological distress, with 8 positively worded items reverse-scored [42]. The Croatian version demonstrated high internal consistency (Cronbach α=0.92 [43]).

General Self-Efficacy Scale (GSE; Croatian Version)

This 10-item scale assesses an individual's belief in their ability to handle new and challenging tasks across different contexts. Items are rated on a 5-point scale, from 1 (not at all true of me) to 5 (completely true of me). For example, the statement "I am confident that I can successfully handle unexpected situations" reflects general self-efficacy, with higher scores indicating stronger self-belief. The scale has demonstrated positive psychometric properties [44]. A review by Scholz et al [45] reported internal consistency coefficients across samples and countries ranging from 0.75 to 0.91. Evidence supports a unidimensional factor structure [46]. In the Croatian version, Cronbach α ranged from 0.74 to 0.85 [47].

Pandemic-Specific Stressors Questionnaire for Older Adults (PSQ-OA)

This questionnaire measures pandemic-specific stressors among older adults. It includes 8 statements rated on a 5-point scale from 0 (not at all) to 4 (extremely), where respondents assess their level of concern related to the pandemic. Exploratory factor analysis revealed 2 factors: social distancing behaviors (5 items) and exposure to infection (3 items). The scale demonstrated good reliability (Cronbach α=0.88 [19]).

General Data Questionnaire

This questionnaire collected demographic information, such as gender, age, place of residence, and marital status, as well as questions about time spent alone and social interactions. It also included items on neurological conditions (used as exclusion criteria) and questions about social behavior during quarantine.

Statistical Analysis

Descriptive measures were first used to summarize demographic data, with frequencies and percentages calculated for categorical variables such as gender, marital status, time spent alone, socializing with others, and place of residence. Means and SDs were used for continuous variables such as age, loneliness, psychological distress, and self-efficacy. Hierarchical regression analysis was conducted in a stepwise manner to assess the contribution of various predictors to the dependent variables of loneliness and psychological distress. The predictors included demographic factors, psychological factors, and social behaviors during the lockdown. Model fit was assessed using R^2, F statistics, and P values. In addition, path analysis was used to examine the mediating contribution of PS (mediator) in the relationship between pandemic-specific stressors (social distancing behaviors and exposure to infection as predictors) and self-efficacy as the dependent variable. Path coefficients, SEs, and CIs were reported for direct, indirect, and total effects. The significance level was set at P<.05 for all tests.

Use of Artificial Intelligence Tools

ChatGPT (OpenAI) was used only as a supportive technical tool in this study. Specifically, it was used to provide suggestions for improving the accuracy of English grammar and phrasing. The authors carefully verified, edited, and approved all artificial intelligence–generated content and take full responsibility for the final paper.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and was approved by 7 ethics committees (Table 2). All participants provided written informed consent before participation. One copy of the consent form was kept by the research participant, and a second was stored separately from the questionnaire to maintain anonymity. Data collection and privacy protection were conducted in accordance with the General Data Protection Regulation (2016/679). All data were anonymized before analysis, and no identifying information was collected, stored, or reported. Participants did not receive any financial or other compensation for their participation in the study. No images or additional materials contain identifiable information about study participants.

XSL•FO
RenderX

Table . List of ethical approvals for a cross-sectional study of older adults in Croatia during the COVID-19 pandemic (2021).

University Ethics Committee	Approval number	Date of approval
School of Medicine, University of Split	Class: 003-08/20-03/0005; registration number (reg no): 2181-198-03-04-20-0083	June 30, 2020
Faculty of Health Sciences, University of Split	Class: 001-01/20-01/0009; reg no: 2181-228-07-20-0005	November 13, 2020
Faculty of Health Studies, University of Rijeka	Class: 602-04/21-01/85; reg no: 2170-57-5-07-1-21-1	March 13, 2021
University of Dubrovnik	Reg no: 524/21	April 30, 2021
University of Applied Health Sciences, Zagreb	Class: 602-04/21-18/222; reg no: 251-379-10-21-02	May 18, 2021
University of Zadar	Class: 114-06/21-01/15; reg no: 2198-1-79-62-21-02	May 25, 2021
Faculty of Dental Medicine and Health, Josip Juraj Strossmayer University of Osijek	Class: 602-04/20-08/02; reg no: 2158/97-97-07-21-08	June 2, 2021

Results

As outlined in the first research aim, Table 3 shows the predictive value of sociodemographic and psychological factors, as well as PS and pandemic-related stressors, on loneliness in older adults during the COVID-19 pandemic. The first step in the model included an initial set of sociodemographic predictors, explaining 5.4% of the variance in loneliness, suggesting that while the model provides a small but statistically significant contribution, much of the variance remains unexplained. In the second step, psychological predictors were introduced, significantly improving the model's explanatory power. The increase in R^2 (from 0.054 to 0.294) indicates that these new variables accounted for an additional 24% of the variance. This represents the largest improvement across all steps. The third step introduced pandemic-specific stressors, leading to a modest yet statistically significant improvement in model accuracy (R^2=0.302; P<.001), contributing an additional 0.8%. The final step added PS as an additional predictor, resulting in a small but significant increase in explained variance (R^2=0.307; P<.001), accounting for only 0.5%.

XSL•FO
RenderX

Table . Hierarchical regression analysis of demographic and psychological variables, pandemic-specific stressors, and preference for solitude as predictors of loneliness among older adults during the COVID-19 pandemic (2021).

Predictor	Step 1 (B[a])	Step 2 (B)[a]	Step 3 (B)[a]	Step 4 (B)[a]
Demographic variable				
Age	–0.002	–0.006[b]	–0.006[b]	–0.005[b]
Gender (women)	–0.023	–0.072[b]	–0.085[b]	–0.081[b]
Marital status				
Divorced	0.219[b]	0.048	0.067	0.026
Never married	0.676[c]	0.407[c]	0.417[c]	0.390[c]
Widowed (<3 months)	0.581[c]	0.324[c]	0.321[c]	0.292[c]
Widowed (3 months-1 year)	0.416[c]	0.128	0.142	0.137
Widowed (>1 year)	0.274[c]	0.189[c]	0.191[c]	0.180[c]
Place of residence (rural)	–0.154[c]	–0.108[c]	–0.103[c]	–0.100[b]
Psychological variable (CORE[d])				
Functionality	—	–0.041[c]	–0.039[c]	–0.037[c]
Problems	—	0.023[c]	0.020[c]	0.020[c]
Subjective well-being	—	–0.014	–0.012	–0.011
Risky behaviors	—	–5.49×10-4	–0.002	–0.001
PSQ[e]				
Social distancing	—	—	0.012[c]	0.014[c]
Exposure to infection	—	—	0.009	0.008
Preference for solitude				0.011[c]
R[f]	0.232	0.542	0.550	0.554
R^2[g]	0.054	0.294	0.302	0.307
RMSE[h]	0.826	0.714	0.710	0.708
F-statistic	14.472	70.78	62.99	60.256
P value	<.001	<.001	<.001	<.001

[a]B: unstandardized regression coefficient.

[b]*P*<.05.

[c]P<.01.

[d]CORE: Clinical Outcomes in Routine Evaluation.

[e]PSQ: pandemic-specific stressor.

[f]R: correlation coefficient.

[g]R^2: coefficient of determination.

[h]RMSE: root mean square error.

Regarding the individual contribution of each predictor in the final step of the regression analysis, age was found to have a small but significant negative relationship with loneliness (unstandardized regression coefficient [B]=−0.005; *P*=.05), indicating that individuals in older groups experienced lower levels of loneliness. Gender also emerged as a significant factor, with women reporting lower loneliness compared with men (B=−0.081; *P*<.02). Marital status played an important role in loneliness levels, where living alone was significantly associated with increased loneliness (B=0.390; *P*<.001), as was being recently widowed (B=0.292; *P*=.002 and B=0.180; *P*<.001, respectively). A significant association was also found between place of residence and loneliness, with individuals living in rural areas reporting lower levels of loneliness compared with those in urban settings (B=−0.100; *P*=.008).

Psychological factors were strongly related to loneliness, where higher levels of psychological problems were significantly

XSL•FO
RenderX

associated with greater loneliness (B=0.020; *P*<.001), while higher functionality was related to lower loneliness (B=−0.037; *P*<.001).

Further, increased social distancing was significantly associated with higher loneliness (B=0.014; *P*<.001), suggesting that restrictions on social interactions contributed to feelings of isolation.

Finally, PS was significantly related to higher loneliness (B=0.011; *P*<.001), suggesting that while some individuals may choose solitude, it can still contribute to feelings of loneliness (Table 3).

Furthermore, as stated in the second research aim, the study examined the contribution of lockdown activities (spending time alone or with others), PS, and self-efficacy to psychological distress levels among older adults. The hierarchical regression analysis assessed how various social and psychological factors during lockdown predicted psychological distress, as measured by Clinical Outcomes in Routine Evaluation total scores (Table 4).

XSL•FO
RenderX

Table . Hierarchical regression analysis of frequency of socialization, preference for solitude, and self-efficacy as predictors of psychological distress among older adults during the COVID-19 pandemic (2021).

Predictor	Step 1 (B)[a]	Step 2 (B)[a]	Step 3 (B)[a]
Time spent alone			
Being alone (only at night)	–1.656	–1.460	–1.827
Being alone (only during the day)	–1.820	–2.510	–3.705
Being alone (> half a day)	0.961	1.039	1.006
Being alone (several hours a day)	–5.173[b]	–4.908[b]	–3.509[b]
Being alone (never longer than 1 hour)	–4.613[b]	–3.946[b]	–2.968[c]
Socializing with others during lockdown			
Close friends (>1 week)	–0.915	–1.076	–1.054
Close friends (at least once a week)	–0.659	–0.701	–1.536
Close friends (2 - 3 times a month)	–0.174	–0.325	–1.631
Close friends (once per month)	0.822	0.673	–1.141
Close friends (once in 6 months)	–0.457	–0.813	–5.156[c]
Close friends (very rare)	5.172[b]	5.058[b]	0.845
Acquaintances (multiple times per week)	1.042	1.050	2.496
Acquaintances (at least once a week)	2.432	2.471	4.321[b]
Acquaintances (2 - 3 times a month)	2.444	2.154	4.543[b]
Acquaintances (once per month)	3.845[c]	3.506[c]	4.357[b]
Acquaintances (once in 6 months)	1.135	0.984	4.163[c]
Socializing with acquaintances (very rare)	7.069[b]	6.750[b]	6.690[b]
Last visit			
Family member (1 month ago)	0.202	0.412	3.895
Family member (1 week ago)	–1.604	–1.371	1.359
Family member (yesterday or today)	–3.396	–3.093	0.481
Another relative	4.083[b]	3.950[b]	3.962[b]
Friends	0.335	0.061	1.127
Nonfamily member	1.374	1.213	3.353[c]
Preference for solitude	—	1.593[b]	2.403[b]
Self-efficacy			–1.066[b]
R[d]	0.299	0.308	0.555
R^2[e]	0.089	0.095	0.308
RMSE[f]	17.059	17.011	14.872
F statistic	8.631	8.835	36.156
P value	<.001	<.001	<.001

[a]B: unstandardized regression coefficient.

XSL•FO
RenderX

[b]$P<.01$.

[c]$P<.05$.

[d]R: correlation coefficient.

[e]R^2: coefficient of determination.

[f]RMSE: root mean square error.

The results consistently showed an improvement in the model's ability to explain psychological distress among older adults during the lockdown across the different regression steps. In Step 1, which included variables related to time spent alone and the frequency of social interactions, the model accounted for 8.9% of the variance. Step 2 introduced PS, slightly improving the model's explanatory power to 9.5%. Although this was a modest increase, it indicated that preference toward solitude contributed uniquely to psychological outcomes. The final model (Step 3), which introduced self-efficacy, led to a substantial increase in explanatory power, explaining 30.8% of the variance and underscoring the pivotal role of personal psychological resources in predicting distress levels during the lockdown period. Self-efficacy showed a strong negative relationship with psychological distress (B=−1.066; $P<.001$), indicating that older adults who believe in their capacity to handle challenges tended to report significantly lower levels of distress. Despite the prevalence of self-efficacy, several other predictors remained significant. Specifically, being alone for several hours per day and never being alone longer than an hour were both significantly associated with lower distress (B=−3.509; $P=.003$ and B=−2.968; $P=.01$, respectively), suggesting that a short amount of time spent alone may have protective effects on the mental health of older adults. On the contrary, rare contacts with close friends, once every 6 months, were associated with lower distress (B=−5.156; $P=.02$), highlighting the emotional importance of regular, meaningful friendships. Further, both frequent and infrequent interactions with acquaintances (weekly to rarely) were consistently associated with higher distress (B value varied from 4.163 to 6.690; all $P<.05$), implying that interactions lacking emotional closeness or confidence may not be beneficial and could even contribute to distress. Visits from distant family members (B=3.962; $P<.001$) and nonfamily members (B=3.353; $P=.04$) were also significant predictors of distress, suggesting these social contacts may have been stress-inducing under lockdown conditions. Finally, PS remained a significant positive predictor of distress (B=2.403; $P<.001$), suggesting that individuals who generally enjoy being alone may have still experienced emotional loneliness during enforced isolation.

The third research aim investigated the mediating role of PS in the relationship between pandemic stressors (social distancing behaviors and exposure to infection) and self-efficacy among older adults. Social distancing behaviors and exposure to infection were included as predictor variables in the path analysis, while self-efficacy was considered the outcome variable. PS, or an individual's tendency to seek time alone, was examined as a mediator that may either buffer or amplify the effects of these stressors on self-efficacy (Table 5; Figure 1).

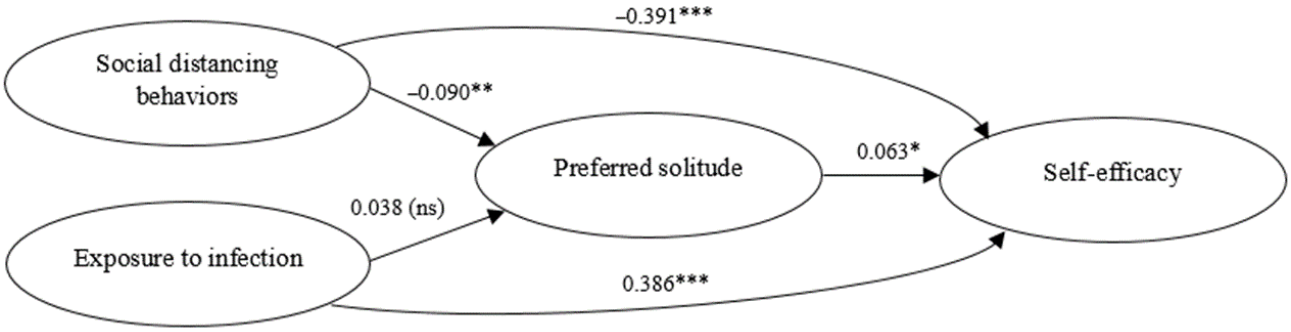
Table . Path analysis of the mediating role of preference for solitude in the relationship between pandemic-specific stressors and self-efficacy among older adults during the COVID-19 pandemic (2021). Criterion: self-efficacy; Mediator: preference for solitude.

Path and effect	Social distancing behaviors as a predictor			Exposure to infection as a predictor			Path coefficients		
	Direct effects	Indirect effects	Total effects	Direct effects	Indirect effects	Total effects	Social distancing behaviors and preference for solitude	Exposure to infection and preference for solitude	Preference for solitude and self-efficacy
β[a] (95% CI)	–0.391 (–0.477 to –0.303)	–0.006 (–0.016 to –0.000)	–0.396 (–0.487 to –0.312)	0.386 (–0.477 to –0.303)	0.002 (–0.003 to 0.016)	0.388 (0.252 to 0.534)	–0.090 (–0.161 to –0.023)	0.038 (–0.067 to 0.158)	0.063 (–0.002 to 0.123)
SE	0.043	0.003	0.043	0.069	0.003	0.069	0.033	0.052	0.029
P value	<.001	.09	<.001	<.001	.48	<.001	.006	.46	.03

[a]β: standardized regression coefficient

XSL•FO
RenderX

Figure 1. Path model showing direct and indirect effects of pandemic-specific stressors on self-efficacy, with preference for solitude as a mediator. *P<.05; **P<.01; ***P<.001; ns: nonsignificant.



Social distancing behaviors were negatively associated with self-efficacy, showing a strong and significant direct effect (β=−.391; P<.001). The indirect pathway through PS was small and not significant (β=−.006; P=.09), resulting in a total negative effect of social distancing on self-efficacy (β=−.396; P<.001). By contrast, exposure to infection demonstrated a strong positive direct association with self-efficacy (β=.386; P<.001). Its indirect effect through PS was negligible and nonsignificant (β=.002; P=.48), yielding a total positive effect (β=.388; P<.001). Furthermore, PS was positively related to self-efficacy (β=.063; P=.03), suggesting that individuals with a higher PS reported slightly greater levels of self-efficacy. The path from social distancing to PS was significant and negative (β=−.090; P=.006), whereas the path from exposure to infection to PS was nonsignificant (β=.038; P=.46). Overall, the findings indicate that while pandemic stressors exerted significant direct effects on self-efficacy, PS contributed only a small positive direct effect and did not act as a strong mediator in these relationships.

Discussion

Principal Findings

This study aimed to evaluate how sociodemographic and psychological factors, along with exposure to pandemic-related stressors and PS, were associated with loneliness among older adults during the COVID-19 pandemic. Second, it explored how pandemic-induced changes in social routines, self-efficacy, and PS predicted levels of psychological distress in this population. Finally, it examined whether PS mediated the relationship between pandemic stressors and self-efficacy.

The findings of this study closely mirrored those of the earlier pilot study [19], highlighting consistent risk factors for loneliness among older adults during the COVID-19 pandemic. Similar to previous international research, predictors such as living alone [11], widowhood [48], psychological distress [2], and social isolation [49] were reaffirmed as significant contributors to loneliness. Older adults and women reported lower levels of loneliness, possibly due to caregiving roles and strong intergenerational ties that are common in Croatian society. This contrasts with previous research showing that older women are more vulnerable to loneliness [17,50]. In our study, men appeared more at risk, likely due to weaker social or family engagement. Similarly, rural residents reported lower loneliness than urban residents, opposing earlier findings [14], possibly due to rural self-reliance and reduced dependence on frequent social interaction. This study also showed lower loneliness among older adults, contrary to some previous studies [51,52]. This may reflect cultural or contextual factors that help explain differences observed in the Croatian sample, as identified in earlier research. For example, older adults in Croatia often benefit from strong family ties and intergenerational support [53]. They may also be more resilient due to past experiences with hardship, such as war or economic instability [54], or may feel more comfortable with smaller, stable social networks that are less affected by crises [55].

This study also examined the role of PS in older adults' mental health and self-efficacy during the COVID-19 pandemic, revealing a complex relationship. Previous studies have shown that PS can support autonomy, emotional regulation, and self-reflection [56,57]. However, recent research [27,58], including this study, suggests that these benefits are context dependent. Our findings show that a higher PS is associated with increased loneliness and psychological distress during enforced isolation. This aligns with previous research suggesting that even voluntary solitude can become emotionally burdensome under restrictive conditions. This tension highlights the need to distinguish between voluntary solitude sought for restoration and involuntary solitude experienced during social constraints (eg, pandemic lockdowns [27]). Theoretically, PS may be protective only when individuals have the agency to choose when and how they are alone. When external circumstances remove that choice, the same preference may leave individuals more vulnerable, especially if it co-occurs with limited social support [32].

PS showed a small but statistically significant association with older adults' well-being, accounting for only limited variance in loneliness and self-efficacy. In contrast, self-efficacy emerged as a stronger and more consistent predictor, reinforcing its role as an important psychological resource in later life. This finding is consistent with previous research highlighting the protective effects of self-efficacy against loneliness [59]. It may help buffer distress through various mechanisms, including enhanced coping, reduced helplessness, and a greater sense of control—factors emphasized in Bandura's social cognitive theory [60]. While PS may offer certain benefits, its effects appear modest compared with the pronounced impact of self-efficacy. Therefore, psychosocial interventions may include elements that support PS but with realistic expectations of its

XSL•FO
RenderX

overall influence. Intervention programs for older adults should prioritize the development of self-efficacy through mastery experiences, meaningful engagement, coping skills training, and supportive social environments. Such approaches are more likely to reduce loneliness and improve emotional well-being in later life.

The path analysis showed that pandemic-related experiences, including social distancing and infection exposure, had the strongest impact on self-efficacy. In contrast, PS had only a minor effect. No studies directly link PS with self-efficacy among older adults in the context of the pandemic. However, existing research [61-64] highlights the dominant role of situational factors such as lockdowns and isolation in shaping psychological outcomes, including self-efficacy. This supports our finding that dispositional factors such as PS are less central in explaining self-efficacy during the pandemic.

Another hierarchical regression analysis in this study revealed several significant predictors of psychological distress (measured by Clinical Outcomes in Routine Evaluation scores). The findings indicate that lower self-efficacy is an important risk factor for psychological distress among older adults, reinforcing existing evidence of its protective role in mental health [61-64]. PS also emerged as a significant contributor to distress, even after accounting for self-efficacy and social behaviors. This finding indicates that solitude may function as a vulnerability factor in this context [27,32]. Moderate amounts of time spent alone were associated with lower distress, suggesting a potential restorative value of solitude. In contrast, limited or superficial social interactions—particularly with acquaintances—were linked to higher levels of psychological strain. This pattern underscores the importance of relationship quality over quantity during periods of social disruption. Our findings resonate with previous research showing that emotionally close ties offer more effective emotional support than weaker social connections, particularly in times of crisis [65,66]. Weak ties, by contrast, may fail to provide meaningful support and, in some cases, may even intensify feelings of loneliness or stress [67]. Interventions aimed at reducing psychological distress in older populations should not simply encourage more social interaction but should instead prioritize the development of deeper, emotionally meaningful connections.

This study, based on a large national sample of older adults, makes a valuable scientific contribution. It highlights important risk and protective factors for mental health and self-efficacy in later life during a global crisis. One of the main contributions lies in examining the underexplored construct of PS, which sheds light on the complex role PS plays in shaping mental health outcomes during enforced social isolation. Although often beneficial, a PS may increase loneliness and distress during crises, particularly when social autonomy is limited. The study also reinforces the importance of psychological well-being and social connectedness as central to maintaining mental health and self-efficacy in older populations during disruptive events.

In practical terms, the findings can inform public health policies and guide targeted interventions aimed at preventing mental health decline among older adults during future pandemics or similar emergencies. Recognizing that loneliness and distress can arise even in individuals who prefer solitude is crucial, as it highlights the need for nuanced strategies that respect individual preferences while maintaining strong social support systems. In addition, the role of self-efficacy as a buffer against distress is significant, highlighting the importance of empowering older adults with clear information, accessible support services, and opportunities to maintain independence during crises. Interventions tailored to rural and urban differences and designed to be sensitive to gender- and age-specific vulnerabilities could further improve outcomes.

Limitations

Despite its strengths, the study has several limitations. The use of snowball sampling was efficient in reaching a large number of older adults across multiple Croatian cities. However, it may have introduced selection bias and limited the representativeness of the findings. Participants were likely drawn from similar networks, resulting in a more homogeneous sample in terms of socioeconomic status, education, health awareness, and other factors. This may restrict the generalizability of the results, particularly to more vulnerable or socially isolated older adults who may have been underrepresented. This sampling method was chosen due to the lockdown conditions at the time of data collection, as epidemiological measures severely limited older adults' ability to leave their homes. Snowball sampling was the only feasible way to reach them under those circumstances. Furthermore, the cross-sectional design prevents causal inferences regarding the observed relationships. Future research should consider using probability-based sampling to enhance representativeness, especially among marginalized or hard-to-reach populations. In addition, future research should use longitudinal designs, as they can help test causality by establishing temporal order, clarifying whether changes in solitude precede changes in psychological distress or vice versa. Such designs allow researchers to track how variables evolve and examine whether shifts in one predict shifts in another. By controlling for baseline levels and assessing potential mediators or moderators, such as social support or health status, longitudinal studies can provide deeper insight into the mechanisms underlying these relationships and offer stronger causal inferences. Finally, qualitative approaches may offer deeper insights into how older adults experience and navigate solitude, stress, and resilience during crises, supporting the development of more targeted mental health interventions.

Conclusion

This study examined how sociodemographic, psychological, and pandemic-related factors affected loneliness and distress in older adults during COVID-19. It also explored whether PS mediated the relationship between pandemic stressors and self-efficacy. Most of these factors affected loneliness, accounting for about one-third of its variance. This finding supports and extends previous regional findings using a nationally representative sample. Changes in social routines, PS, and self-efficacy also predicted psychological distress, explaining roughly one-third of its variance. Spending a moderate amount of time alone was linked to lower distress, while limited or superficial social interactions increased it. Self-efficacy emerged as a strong protective factor against

XSL•FO
RenderX

psychological distress. Although PS had only a small mediating role between pandemic stressors and self-efficacy, it still affected mental health. Overall, situational factors had the greatest impact on mental well-being. Future interventions should focus on strengthening self-efficacy and promoting meaningful social connections among older adults during times of crisis.

Data Availability

The datasets generated and analyzed during this study are not publicly available because the participants' informed consent did not include permission for public data sharing, but the data may be obtained from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: RO
Data curation: AM, RO
Formal analysis: AM, RO, VA
Investigation: DS, IG, RO
Methodology: RO, VA
Project administration: RO
Resources: DS, IG, RO
Supervision: RO, VA
Validation: AM, VA
Visualization: AM, DS, IG, RO, VA
Writing—original draft: RO, VA
Writing—review and editing: AM, DS, IG, RO, VA

Conflicts of Interest

None declared.

References

1. Peng X, Jiang D, Tse DCK, Ho YW, Lay J. Momentary solitude and affective experiences: the moderating role of trait mindfulness. Appl Psychol Health Well Being 2025 Jun;17(3):e70052. [doi: 10.1111/aphw.70052] [Medline: 40503637]
2. Sams N, Fisher DM, Mata-Greve F, et al. Understanding psychological distress and protective factors amongst older adults during the COVID-19 pandemic. Am J Geriatr Psychiatry 2021 Sep;29(9):881-894. [doi: 10.1016/j.jagp.2021.03.005] [Medline: 33867224]
3. Krendl AC, Perry BL. The impact of sheltering in place during the COVID-19 pandemic on older adults' social and mental well-being. J Gerontol B Psychol Sci Soc Sci 2021 Jan 18;76(2):e53-e58. [doi: 10.1093/geronb/gbaa110] [Medline: 32778899]
4. Novotný JS, Gonzalez-Rivas JP, Kunzová Š, et al. Risk factors underlying COVID-19 lockdown-induced mental distress. Front Psychiatry 2020;11:603014. [doi: 10.3389/fpsyt.2020.603014] [Medline: 33424666]
5. Hülür G, Macdonald B. Rethinking social relationships in old age: digitalization and the social lives of older adults. Am Psychol 2020;75(4):554-566. [doi: 10.1037/amp0000604] [Medline: 32378949]
6. Parlapani E, Holeva V, Nikopoulou VA, Kaprinis S, Nouskas I, Diakogiannis I. A review on the COVID-19-related psychological impact on older adults: vulnerable or not? Aging Clin Exp Res 2021 Jun;33(6):1729-1743. [doi: 10.1007/s40520-021-01873-4] [Medline: 33999378]
7. Wong SYS, Zhang D, Sit RWS, et al. Impact of COVID-19 on loneliness, mental health, and health service utilisation: a prospective cohort study of older adults with multimorbidity in primary care. Br J Gen Pract 2020 Nov;70(700):e817-e824. [doi: 10.3399/bjgp20X713021] [Medline: 32988955]
8. Choi EY, Farina MP, Wu Q, Ailshire J. COVID-19 social distancing measures and loneliness among older adults. J Gerontol 2022 Jul 5;77(7):e167-e178. [doi: 10.1093/geronb/gbab009]
9. Losada-Baltar A, Martínez-Huertas J, Jiménez-Gonzalo L, et al. Longitudinal correlates of loneliness and psychological distress during the lockdown situation due to COVID-19. Effects of age and self-perceptions of aging. The Journals of Gerontology 2022 Apr 1;77(4):652-660. [doi: 10.1093/geronb/gbab012]
10. Cihan FG, Gökgöz Durmaz F. Evaluation of COVID-19 phobia and the feeling of loneliness in the geriatric age group. Int J Clin Pract 2021 Jun;75(6):e14089. [doi: 10.1111/ijcp.14089] [Medline: 33570809]
11. Polenick CA, Perbix EA, Salwi SM, Maust DT, Birditt KS, Brooks JM. Loneliness during the COVID-19 pandemic among older adults with chronic conditions. J Appl Gerontol 2021 Aug;40(8):804-813. [doi: 10.1177/0733464821996527] [Medline: 33641513]

XSL•FO
RenderX

12. Lehtisalo J, Palmer K, Mangialasche F, Solomon A, Kivipelto M, Ngandu T. Changes in lifestyle, behaviors, and risk factors for cognitive impairment in older persons during the first wave of the coronavirus disease 2019 pandemic in Finland: results from the FINGER study. Front Psychiatry 2021;12:624125. [doi: 10.3389/fpsyt.2021.624125] [Medline: 33643095]
13. Macdonald B, Hülür G. Well-being and loneliness in Swiss older adults during the COVID-19 pandemic: the role of social relationships. Gerontologist 2021 Feb 23;61(2):240-250. [doi: 10.1093/geront/gnaa194] [Medline: 33258898]
14. Kirkland SA, Griffith LE, Oz UE, et al. Increased prevalence of loneliness and associated risk factors during the COVID-19 pandemic: findings from the Canadian Longitudinal Study on Aging (CLSA). BMC Public Health 2023 May 12;23(1):872. [doi: 10.1186/s12889-023-15807-4] [Medline: 37170234]
15. Kobayashi LC, O'Shea BQ, Kler JS, et al. Cohort profile: the COVID-19 coping study, a longitudinal mixed-methods study of middle-aged and older adults' mental health and well-being during the COVID-19 pandemic in the USA. BMJ Open 2021 Feb 10;11(2):e044965. [doi: 10.1136/bmjopen-2020-044965] [Medline: 33568377]
16. Lara E, Matovic S, Vasiliadis HM, et al. Correlates and trajectories of loneliness among community-dwelling older adults during the COVID-19 pandemic: a Canadian longitudinal study. Arch Gerontol Geriatr 2023 Dec;115:105133. [doi: 10.1016/j.archger.2023.105133] [Medline: 37499331]
17. Robb CE, de Jager CA, Ahmadi-Abhari S, et al. Associations of social isolation with anxiety and depression during the early COVID-19 pandemic: a survey of older adults in London, UK. Front Psychiatry 2020;11:591120. [doi: 10.3389/fpsyt.2020.591120] [Medline: 33132942]
18. Tucak Junaković I. The impact of the COVID-19 pandemic on older adults: results of a descriptive study. Contemp Psychol 2021;24(1):77-92. [doi: 10.21465/2021-SP-241-04]
19. Orlandini R, Lušić Kalcina L, Antičević V. Understanding loneliness in older adults during the pandemic: predictors and questionnaire validation. Diseases 2025 Feb 4;13(2):45. [doi: 10.3390/diseases13020045] [Medline: 39997052]
20. Andrade A, D'Oliveira A, Dos Santos KM, et al. Impact of social isolation caused by the COVID-19 pandemic on the mood profile of active and sedentary older adults: physical activity as a protective factor. Front Public Health 2023;11:1221142. [doi: 10.3389/fpubh.2023.1221142] [Medline: 37849723]
21. Tosato M, Ciciarello F, Zazzara MB, et al. Lifestyle changes and psychological well-being in older adults during COVID-19 pandemic. Clin Geriatr Med 2022 Aug;38(3):449-459. [doi: 10.1016/j.cger.2022.05.002] [Medline: 35868665]
22. Ritchie L, Cervone D, Sharpe BT. Goals and self-efficacy beliefs during the initial COVID-19 lockdown: a mixed methods analysis. Front Psychol 2021;11:559114. [doi: 10.3389/fpsyg.2020.559114]
23. Castellano-Tejedor C, Pérez LM, Soto-Bagaria L, et al. Correlates to psychological distress in frail older community-dwellers undergoing lockdown during the COVID-19 pandemic. BMC Geriatr 2022 Dec;22(1):516. [doi: 10.1186/s12877-022-03072-w]
24. Chen WY, Yan L, Yuan YR, Zhu XW, Zhang YH, Lian SL. Preference for solitude and mobile phone addiction among Chinese college students: the mediating role of psychological distress and moderating role of mindfulness. Front Psychol 2021;12:750511. [doi: 10.3389/fpsyg.2021.750511]
25. Wang JM, Rubin KH, Laursen B, Booth-LaForce C, Rose-Krasnor L. Preference-for-solitude and adjustment difficulties in early and late adolescence. J Clin Child Adolesc Psychol 2013 Nov 1;42(6):834-842. [doi: 10.1080/15374416.2013.794700]
26. Burger JM. Individual differences in preference for solitude. J Res Pers 1995 Mar;29(1):85-108. [doi: 10.1006/jrpe.1995.1005]
27. Ren D, Stavrova O, Evans AM. Does dispositional preference for solitude predict better psychological outcomes during times of social distancing? Beliefs and reality. J Pers 2023 Dec;91(6):1442-1460. [doi: 10.1111/jopy.12821] [Medline: 36748170]
28. Tucak Junaković I, Nekić M, Burić I. Loneliness and preferred solitude as predictors of self-rated health of adults [Article in Croatian]. Medica Jadertina 2013;43(1-2):11-21 [FREE Full text]
29. Toyoshima A, Sato S. Examination of the relationship between preference for solitude and emotional well-being after controlling for the effect of loneliness. Shinrigaku Kenkyu 2015 Jun;86(2):142-149. [doi: 10.4992/jjpsy.86.13234] [Medline: 26182490]
30. Toyoshima A, Kusumi T. Examining the relationship between preference for solitude and subjective well-being among Japanese older adults. Innov Aging 2022;6(1):igab054. [doi: 10.1093/geroni/igab054] [Medline: 35146129]
31. Smith JL, Thomas V, Azmitia M. Happy alone? Motivational profiles of solitude and well-being among senior living residents. Int J Aging Hum Dev 2023 Apr;96(3):312-334. [doi: 10.1177/00914150221112283] [Medline: 35861793]
32. Sakurai R, Sakurai M, Suzuki H, Fujiwara Y. Preference for solitude paradox: the psychological influence of social isolation despite preference. J Affect Disord 2024 Nov 15;365:466-473. [doi: 10.1016/j.jad.2024.08.020] [Medline: 39142587]
33. Elm EV, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD]
34. Alterovitz SSR, Mendelsohn GA. Relationship goals of middle-aged, young-old, and old-old Internet daters: an analysis of online personal ads. J Aging Stud 2013 Apr;27(2):159-165. [doi: 10.1016/j.jaging.2012.12.006] [Medline: 23561281]
35. Coleman JS. Relational analysis: the study of social organizations with survey methods. Hum Organ 1958 Dec;17(4):28-36. [doi: 10.17730/humo.17.4.q5604m676260q8n7]
36. Goodman LA. Snowball sampling. Ann Math Statist 1961 Mar;32(1):148-170. [doi: 10.1214/aoms/1177705148]

XSL•FO
RenderX

37. Cochran WG. Sampling Techniques: Johan Wiley & Sons Inc; 1977.
38. Russell D, Peplau LA, Cutrona CE. The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence. J Pers Soc Psychol 1980 Sep;39(3):472-480. [doi: 10.1037//0022-3514.39.3.472] [Medline: 7431205]
39. Lacković-Grgin K, Penezić Z, Nekić M. Kratka verzija UCLA skale usamljenosti [Short version of the UCLA Loneliness Scale]. In: Lacković-Grgin K, Proroković A, Ćubela V, Penezić Z, editors. Zbirka Psihologijskih Skala i Upitnika, 1st edition: Filozofski fakultet u Zadru; 2002:77-78 URL: https://morepress.unizd.hr/books/press/catalog/book/28 [accessed 2025-11-09]
40. Cramer KM, Lake RP. The preference for Solitude Scale: psychometric properties and factor structure. Pers Individ Dif 1998 Feb;24(2):193-199. [doi: 10.1016/S0191-8869(97)00167-0]
41. Lacković-Grgin K, Nekić M. Adaptirana skala preferirane samoće [Adapted Solitude Preference Scale]. In: Lacković-Grgin K, Proroković A, Ćubela V, Penezić Z, editors. Zbirka Psihologijskih Skala i Upitnika, 1st edition: Filozofski fakultet u Zadru; 2002:23-24 URL: https://morepress.unizd.hr/books/press/catalog/book/28 [accessed 2025-11-09]
42. Evans C, Mellor-Clark J, Margison F, et al. CORE: Clinical Outcomes in Routine Evaluation. J of Mental Hlth 2000 Jun 1;9(3):247-255. [doi: 10.1080/713680250]
43. Jokić-Begić N, Lauri Korajlija A, Jurin T, Evans C. Factor structure, psychometric properties, and cut-off value of the Croatian translation of the CORE-OM questionnaire. Psychological Topics 2014;23(2):265-288 [FREE Full text]
44. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright SC, Johnston M, editors. Causal and Control Beliefs (Measures in Health Psychology: A User's Portfolio): NFER - Nelson; 1995:35-37.
45. Scholz U, Doña BG, Sud S, Schwarzer R. Is general self-efficacy a universal construct? Psychometric findings from 25 countries. Eur J Psychol Assess 2002;18(3):242. [doi: 10.1027//1015-5759.18.3.242]
46. Schwarzer R, Bassler J, Kwiatek P, Schroder K, Zhang JX. The assessment of optimistic self‐beliefs: comparison of the German, Spanish, and Chinese versions of the General Self‐Efficacy Scale. Applied Psychology 1997 Jan 1;46(1):69-88. [doi: 10.1080/026999497378557]
47. Ivanov L, Penezić Z. Skala opće samoefikasnosti [Generalized self-efficacy scale]. In: Lacković-Grgin K, Proroković A, Ćubela V, Penezić Z, editors. Zbirka Psihologijskih Skala i Upitnika, 1st edition: Filozofski fakultet u Zadru; 2002:6-7 URL: https://morepress.unizd.hr/books/press/catalog/book/28 [accessed 2025-11-09]
48. Dahlberg L, Andersson L, McKee KJ, Lennartsson C. Predictors of loneliness among older women and men in Sweden: a national longitudinal study. Aging Ment Health 2015 May 4;19(5):409-417. [doi: 10.1080/13607863.2014.944091]
49. Kilic D, Aslan G, Ata G, Bakan ABS. Relationship between the fear of COVID-19 and social isolation and depression in elderly individuals. Psychogeriatrics 2023 Mar;23(2):222-229. [doi: 10.1111/psyg.12924] [Medline: 36516958]
50. Savage RD, Wu W, Li J, et al. Loneliness among older adults in the community during COVID-19: a cross-sectional survey in Canada. BMJ Open 2021 Apr 2;11(4):e044517. [doi: 10.1136/bmjopen-2020-044517] [Medline: 33811054]
51. Khan MSR, Kadoya Y. Loneliness during the COVID-19 pandemic: a comparison between older and younger people. Int J Environ Res Public Health 2021 Jul 25;18(15):7871. [doi: 10.3390/ijerph18157871] [Medline: 34360164]
52. O'Shea BQ, Finlay JM, Kler J, Joseph CA, Kobayashi LC. Loneliness among US adults aged ≥55 early in the COVID-19 pandemic: findings from the COVID-19 coping study. Public Health Rep 2021;136(6):754-764. [doi: 10.1177/00333549211029965] [Medline: 34283657]
53. Plužarić J, Barać I, Ilakovac V, et al. Connectedness and successful aging of older adults in Croatia. Sustainability 2023;15(14):10843. [doi: 10.3390/su151410843]
54. Perinić Lewis A, Škarić-Jurić T, Despot Lučanin J, Smolić Š. The lasting impact of war experiences on quality of life in long-lived retirement homes residents: the birth cohort 1906–1928. Ageing Soc 2024 Mar;44(3):568-596. [doi: 10.1017/S0144686X22000319]
55. Charles ST, Carstensen LL. Social and emotional aging. Annu Rev Psychol 2010;61(1):383-409. [doi: 10.1146/annurev.psych.093008.100448] [Medline: 19575618]
56. Nguyen TVT, Ryan RM, Deci EL. Solitude as an approach to affective self-regulation. Pers Soc Psychol Bull 2018 Jan;44(1):92-106. [doi: 10.1177/0146167217733073]
57. Fan Z, Cai H, Shi X, Yu N, Chen L. The positive solitude scale (PS): psychometric properties among Chinese older. BMC Public Health 2024 Aug 9;24(1):2162. [doi: 10.1186/s12889-024-19237-8] [Medline: 39123169]
58. Michinov E, Michinov N. Stay at home! When personality profiles influence mental health and creativity during the COVID-19 lockdown. Curr Psychol 2023;42(7):5650-5661. [doi: 10.1007/s12144-021-01885-3] [Medline: 34092985]
59. Lee JW, Nersesian PV, Suen JJ, et al. Loneliness is associated with lower coping self-efficacy among older adults. J Appl Gerontol 2023 Feb;42(2):270-279. [doi: 10.1177/07334648221129858] [Medline: 36178675]
60. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
61. Ruotolo F, Ruggiero G, Cattaneo Z, et al. Psychological reactions during and after a lockdown: self-efficacy as a protective factor of mental health. Int J Environ Res Public Health 2023 Aug 29;20(17):6679. [doi: 10.3390/ijerph20176679] [Medline: 37681819]

XSL•FO
RenderX

62. Ab Malik N, Mohd RH, Abdul Kadir NB. Knowledge, self-efficacy, and preventive behavior of COVID-19 outbreak among community in Malaysia. Asia Pac J Public Health 2021 Jul;33(5):674-675. [doi: 10.1177/10105395211012542] [Medline: 33938242]
63. Guidry JPD, Perrin PB, Bol N, et al. Social distancing during COVID-19: threat and efficacy among university students in seven nations. Glob Health Promot 2022 Mar;29(1):5-13. [doi: 10.1177/17579759211051368] [Medline: 34702108]
64. Fida R, Paciello M, Watson D, Nayani R. The protective role of work self-efficacy on wellbeing during COVID-19 pandemic: results from a longitudinal year-long study. Pers Individ Dif 2022 Oct;197:111760. [doi: 10.1016/j.paid.2022.111760] [Medline: 35677891]
65. Lee SS, Shim Y, Choi J, Choi I. Paradoxical impacts of social relationship on well-being during the COVID-19 pandemic. J Happiness Stud 2023;24(2):745-767. [doi: 10.1007/s10902-022-00614-2] [Medline: 36686601]
66. Kulcar V, Bork-Hüffer T, Schneider AM. Getting through the crisis together: do friendships contribute to university students' resilience during the COVID-19 pandemic? Front Psychol 2022;13:880646. [doi: 10.3389/fpsyg.2022.880646] [Medline: 35651553]
67. Tibbetts M, Epstein-Shuman A, Leitao M, Kushlev K. A week during COVID-19: online social interactions are associated with greater connection and more stress. Comput Hum Behav Rep 2021;4:100133. [doi: 10.1016/j.chbr.2021.100133] [Medline: 34693074]

Abbreviations

PS: preference for solitude
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani; submitted 08.06.25; peer-reviewed by K Nagar, Ž Kralj; revised version received 22.09.25; accepted 24.09.25; published 14.11.25.

Please cite as:
Orlandini R, Matana A, Švaljug D, Gusar I, Antičević V
Predictors of Loneliness and Psychological Distress in Older Adults During the COVID-19 Pandemic: National Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e78728
URL: https://publichealth.jmir.org/2025/1/e78728
doi:10.2196/78728

© Rahela Orlandini, Antonela Matana, Deana Švaljug, Ivana Gusar, Vesna Antičević. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 14.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Influenza Vaccination Coverage and Determinants of New Vaccinations During the COVID-19 Pandemic in Spain (ENE-COVID): Nationwide Population-Based Study

Miguel Ángel de la Cámara[1,2*], PhD; Nerea Fernández de Larrea-Baz[1,3*], PhD; Roberto Pastor-Barriuso[1,3], PhD; Amparo Larrauri[1,3], PhD; Pablo Fernández-Navarro[1,3], PhD; Marina Pollán[1,3*], PhD; Beatriz Pérez-Gómez[1,3*], PhD; ENE-COVID Study Group

[1]National Centre for Epidemiology, Instituto de Salud Carlos III, Monforte de Lemos, 5, Madrid, Spain

[2]Department of Educational Sciences, University of Alcalá, Alcalá de Henares, Spain

[3]Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), Madrid, Spain

[*]these authors contributed equally

Corresponding Author:
Nerea Fernández de Larrea-Baz, PhD
National Centre for Epidemiology, Instituto de Salud Carlos III, Monforte de Lemos, 5, Madrid, Spain

Abstract

Background: Influenza vaccination coverage is commonly suboptimal. However, the COVID-19 pandemic and consequent high exposure to health information may have changed population attitudes toward this vaccination.

Objective: The aim of this study is to describe influenza vaccine uptake in Spain during the first influenza season following the start of the COVID-19 pandemic compared to the previous one and identify characteristics associated with vaccination among those previously unvaccinated.

Methods: This was a population-based study of 28,987 adults included in influenza vaccination target groups (≥65 years old, with risk conditions, living with someone with risk conditions, health care workers, security or emergency workers) who were participants in the nationwide Seroepidemiological Survey of SARS-CoV-2 Infection in Spain (ENE-COVID) study. Information on vaccination and sociodemographic, health, and COVID-19–related factors was collected by interview. Coverage change from 2019 to 2020 and standardized prevalences of vaccination in 2020 among the population unvaccinated in 2019 were estimated using logistic model–based methods.

Results: Coverage rose from 31.4% (95% CI 30.5% - 32.2%) to 46.8% (95% CI 45.8% - 47.8%). People ≥65 years old showed the highest uptake in both periods (58.3%, 95% CI 56.8% - 59.8% and 74.8%, 95% CI 73.5% - 76.1%), while health care workers had the greatest increase (22%, 95% CI 17.8% - 26.2%). Among people unvaccinated in 2019, factors associated with vaccination in 2020 were age, female sex, higher education, Spanish nationality, multimorbidity, being a former smoker, obesity, contact with COVID-19 cases, living with older adults, living in provinces with low COVID-19 incidence, wearing a face mask during family meetings, and using surgical/FFP2 masks.

Conclusions: This study provides nationwide representative estimates of influenza vaccination coverage, which clearly increased between 2019 and 2020 in the 5 target groups. However, coverage goals were attained only in the ≥65 year old group, highlighting the importance of reinforcing influenza vaccination. Our detailed results on determinants of vaccination provide some clues to tailor vaccination strategies.

(JMIR Public Health Surveill 2025;11:e60658) doi:10.2196/60658

KEYWORDS

influenza prevention; influenza vaccination; vaccination coverage; vaccination determinants; COVID-19 pandemic; influenza; vaccination; influenza vaccine; Spain; population-based study; health information; adult; ENE-COVID; epidemiology; sociodemographic; infection prevention; public health

Introduction

Seasonal influenza is a global health problem, with almost one billion cases every year and 3 - 5 million severe cases, causing 290,000 - 650,000 deaths worldwide [1]. Influenza vaccination is the leading safe and effective pharmacological intervention available to prevent and reduce the impact of this annual pandemic, particularly in more susceptible groups [2], such as older people, pregnant women, and people with underlying comorbidities, as well as those in selected occupations, including

XSL•FO
RenderX

health care workers (HCWs). However, in most countries, the established goal of 70% - 75% vaccination coverage for these target groups is not achieved [3-5].

Several factors may drive vaccine uptake. In addition to health policy and promotion strategies or socioeconomic factors (eg, household income or town size) [6], some individual characteristics seem to be associated to influenza vaccination [7,8]. Higher coverage has been reported in men, older people, people with higher education, those who make frequent visits to the doctor, or those with risk factors for severe influenza. Risk perception of influenza disease and an individual's opinion about the efficacy and safety of the vaccine usually play a role in the decision to be vaccinated [7]. However, in 2020, a new actor came into play: the deadly COVID-19 pandemic, caused by another respiratory virus. That year, the World Health Organization emphasized the need for influenza vaccination [9], concerned by the additional burden of cocirculation of both viruses for vulnerable populations and health systems. Its recommendations were adopted by different countries, with reinforced influenza vaccination campaigns during the 2020/2021 season. These strengthened campaigns, together with the fear of a higher severity of potential influenza and SARS-CoV-2 coinfections and the commitment of the population to help to avoid a new overload of the health care system, probably contributed to increased influenza vaccination acceptance, as has been observed in some studies [10-15]. However, most of these studies relied on convenience samples, mostly gathered through web-based questionnaires, or explored only a limited number of population characteristics.

In 2020, the Spanish Ministry of Health and the Institute of Health Carlos III, in collaboration with the health services of all Spanish regions, launched the Seroepidemiological Survey of SARS-CoV-2 Infection in Spain (ENE-COVID), a nationwide, representative, population-based seroepidemiological survey, with more than 60,000 participants, to quantify SARS-CoV-2 infection prevalence [16]. In the fourth round of the ENE-COVID, in November 2020, self-reported information on influenza vaccination was gathered, together with a wide range of sociodemographic, health, and COVID-19–related epidemiological data.

Taking advantage of this sound design, the aims of this study are to provide influenza vaccination coverage estimates among the adult population corresponding to the main target groups in Spain in 2019 and 2020 (ie, the influenza season prior to the start of the COVID-19 pandemic and the first influenza season following the start of the pandemic) and to identify a wide range of factors related to the change from not being vaccinated to influenza vaccine uptake.

Methods

Study Design and Population

ENE-COVID is a nationwide population-based cohort study designed to provide representative seroprevalence figures for SARS-CoV-2 infection in Spain for the noninstitutionalized population [16]. Briefly, 35,885 households were selected through a stratified 2-stage sampling, with strata formed by cross-classifying the 52 Spanish provinces and autonomous cities with 4 categories of municipality size. A total of 1500 census tracts were selected with probability proportional to their size, and then 24 households were randomly sampled within each census tract. Households were contacted by telephone and all residents were invited to participate. Participation required answering an epidemiological questionnaire and being assessed for SARS-CoV-2 antibodies by a point-of-care test. The survey was repeated on 4 occasions for the same cohort between April and November 2020 [17].

For this study, we analyzed data from the fourth follow-up round, conducted in November 16-29, 2020, which included 51,409 participants. Among them, adult participants in any of the following influenza vaccination target groups were included in the present analysis (28,987 participants): (1) people ≥65 years old; (2) people aged 18 - 64 years reporting health conditions that increase their risk of severe influenza, including diabetes; chronic lung, liver, or kidney disease; asthma, cardiovascular disease, hematological tumour, any other cancer in the last five years, immunosuppression treatment, or severe obesity; (3) people living with someone reporting any of these health conditions; (4) HCWs, including health care or sociosanitary workers, and at-home professional caregivers; and (5) security and emergency workers. For participants that pertained to more than one group, the following order of preference was applied: age ≥65 years, risk conditions, HCW, living with someone with a risk condition, and security and emergency workers. Other population groups included in the international influenza vaccination recommendations, such as pregnant women or children with certain conditions, were not addressed because the questionnaire information did not allow us to comprehensively identify participants in those groups.

Epidemiological Information

The epidemiological questionnaire included information on sociodemographic characteristics (registered sex, age, nationality, education, employment status, and job sector), health-related characteristics (self-reported height, weight, smoking habit, health status, chronic conditions, and officially recognized disability, that is, having received a certificate of any degree of disability—physical, sensory, intellectual, or mental disability—after evaluation by the technical teams of the corresponding regional authority), and COVID-19–related factors (self-reported COVID-19 pneumonia or hospitalization, contacts with symptomatic or confirmed cases, and behaviors related to possible exposure to SARS-CoV-2 such as mask wearing or attending family events). We classified personal experience of COVID-19 as no infection, not severe infection (positive antibody test or self-reported positive polymerase chain reaction or antigen test, without pneumonia or hospitalization), or severe infection (pneumonia or hospitalization). Additionally, we considered household and contextual information: household size and age composition (presence of children under 3 years of age or adults over 65 years of age), autonomous community (administrative level at which regional influenza vaccination campaigns are organized), municipality size, average personal income in the census tract relative to each province, and provincial 7-day cumulative

XSL•FO
RenderX

incidence of COVID-19 at the start of the field work (November 16-22, 2020).

Self-reported influenza vaccination was collected by asking the participants if they had been vaccinated in the fall of 2020 or intended to do so and if they had received the vaccine in the previous year. Since the 2020 - 2021 vaccination campaign was still ongoing at the time of interview, participants who reported that they intended to be vaccinated were considered as vaccinated in our main analyses (a sensitivity analysis excluding these participants was also performed).

Statistical Analysis

Influenza vaccination coverage in 2019 and 2020 was calculated as the proportion of participants who reported having been vaccinated or intended to do so in that year. Vaccination coverage by sociodemographic, health, and COVID-19–related characteristics was standardized to the overall distribution of influenza vaccination target groups in the entire study population and compared by estimating standardized differences between 2019 and 2020. We also computed target group–standardized prevalences of influenza vaccination by cross-classifying 2019 and 2020 data (not vaccinated in any season, vaccinated only in 2019, vaccinated both years, or only vaccinated in 2020) overall and by participant characteristics. Logistic model-based standardization methods were used for this purpose [18].

To assess factors associated with a change from 2019 to 2020 favouring vaccination, we calculated the prevalence of new influenza vaccination in 2020 among participants unvaccinated in 2019. Using logistic model–based standardization methods [18], the prevalence of new vaccinees by participant characteristics was standardized to the overall distribution of influenza vaccination groups, sex, age, education, and autonomous community in the unvaccinated target population in 2019. Standardized prevalence differences and ratios for new influenza vaccination were estimated across categories of participant characteristics. Lastly, to assess potential effect modifications of prespecified characteristics on new vaccination by target group, we included target group-by-covariate interactions in logistic models and estimated standardized prevalence ratios (SPR) within each target group. We tested for homogeneity using Wald tests.

All statistical analyses used sampling weights to correct for the different selection probabilities by province and the distinct response rates by sex, age, and census tract income, and they accounted for stratification and clustering by household and census tract when computing CIs and tests [16]. Analyses were performed in Stata (version 16; StataCorp) and graphics were produced in R (version 4; R Foundation for Statistical Computing).

Ethical Considerations

The institutional review board of the Institute of Health Carlos III approved the study (register number PI 39_2020) and written informed consent was obtained from all study participants. Data were deidentified before statistical analysis. No compensation was provided to participants.

Results

Figure 1 summarizes the participation flowchart for the fourth round of ENE-COVID. The 28,987 adults included in the study population were distributed by vaccination target groups: aged ≥65 years (10,279/28,987, 35.5%), living with someone with risk conditions (8190/28,987, 28.2%), aged <65 years with a risk condition (8000/28,987, 27.6%), HCW (2172/28,987, 7.5%), and security and emergency workers (346/28,987, 1.2%). Among them, 15,802 (54.5%) were women (mean age 56.7, SD 17.6 years) and 13,185 (45.5%) were men (mean age 56.3, SD 17.4 years).

XSL•FO
RenderX

Figure 1. Flowchart of participants in the fourth round of the ENE-COVID study, November 16 - 29, 2020, Spain. ENE-COVID: Seroepidemiological Survey of SARS-CoV-2 Infection in Spain.

XSL•FO
RenderX

Table 1 summarizes influenza vaccination coverage in 2019 and 2020 by target group and participants' characteristics. Overall coverage in 2019 was 31.4% (95% CI 30.5% - 32.2%) and it was 46.8% (45.8% - 47.8%) in 2020, corresponding to an absolute increase of 15.4% (14.1% - 16.8%) and a relative increase of 49% (44%-54%). By target group, the highest uptake was among adults ≥65 years old, with 58.3% (56.8% - 59.8%) in 2019, achieving 74.8% (73.5% - 76.1%) in 2020, while the lowest figures were observed among people living with someone with risk conditions (2019: 9.3%, 8.4% - 10.1%; 2020: 21.1%, 19.7% - 22.4%) and security/emergency workers (2019: 8.6%, 4.7% - 12.4%; 2020: 21%, 15.2% - 26.8%). All groups had a marked increase in 2020, especially HCWs (22.0%, 17.8% - 26.2%).

XSL•FO
RenderX

Table . Influenza vaccination coverage in 2019 and 2020[a] and coverage differences among adults of selected vaccination target groups[b] (Spain, November 2020).

	Number	2019 coverage, % (95% CI)[c]	2020 coverage, % (95% CI)[c]	Standardized difference, % (95% CI)[c]
Total	28,987	31.4 (30.5 - 32.2)	46.8 (45.8 - 47.8)	15.4 (14.1 - 16.8)
Influenza-related characteristics				
Influenza vaccination target group				
≥65 years	10,279	58.3 (56.8 - 59.8)	74.8 (73.5 - 76.1)	16.5 (14.5 - 18.5)
<65 years with a risk condition[d]	8000	25.0 (23.7 - 26.3)	41.4 (39.8 - 43.0)	16.4 (14.3 - 18.4)
Health care workers	2172	23.3 (20.6 - 26.0)	45.3 (42.2 - 48.5)	22.0 (17.8 - 26.2)
Living with someone with a risk condition[d]	8190	9.3 (8.4 - 10.1)	21.1 (19.7 - 22.4)	11.8 (10.2 - 13.4)
Security and emergency workers	346	8.6 (4.7 - 12.4)	21.0 (15.2 - 26.8)	12.4 (5.5 - 19.4)
Number of influenza vaccination criteria				
1	16,583	28.0 (26.8 - 29.2)	44.3 (43.0 - 45.5)	16.3 (14.5 - 18.0)
2	9321	32.7 (31.3 - 34.0)	48.5 (46.9 - 50.2)	15.9 (13.8 - 18.0)
3	3083	40.9 (38.4 - 43.3)	56.6 (53.4 - 59.8)	15.7 (11.7 - 19.8)
Sociodemographic characteristics				
Sex				
Men	13,185	30.9 (29.9 - 31.9)	45.2 (43.9 - 46.4)	14.3 (12.7 - 15.9)
Women	15,802	31.8 (30.7 - 32.8)	48.2 (47.0 - 49.4)	16.4 (14.9 - 18.0)
Age group (years)				
18 - 34	4059	10.7 (9.4 - 12.0)	20.2 (18.5 - 22.0)	9.5 (7.3 - 11.8)
35 - 59	11,920	17.1 (16.1 - 18.0)	33.0 (31.7 - 34.3)	15.9 (14.3 - 17.5)
60 - 64	2729	34.7 (32.1 - 37.2)	57.1 (54.5 - 59.8)	22.5 (18.8 - 26.2)
65 - 69	3339	46.8 (44.3 - 49.3)	70.8 (68.8 - 72.9)	24.1 (20.8 - 27.3)
70 - 74	2799	56.1 (53.4 - 58.8)	74.3 (71.8 - 76.8)	18.2 (14.6 - 21.9)
75 - 79	1988	64.9 (62.1 - 67.7)	75.8 (73.3 - 78.4)	11.0 (7.1 - 14.8)
≥80	2153	74.0 (71.3 - 76.8)	81.0 (78.6 - 83.3)	6.9 (3.3 - 10.5)
Nationality				
Spanish	28,172	31.6 (30.7 - 32.4)	47.2 (46.1 - 48.2)	15.6 (14.2 - 16.9)
Other	815	24.0 (19.9 - 28.1)	35.3 (31.1 - 39.5)	11.3 (5.4 - 17.2)
Education				
Primary or less	7830	35.2 (33.7 - 36.6)	49.0 (47.3 - 50.7)	13.8 (11.6 - 16.1)
High school/vocational education	15,191	29.0 (27.9 - 30.1)	44.5 (43.3 - 45.8)	15.5 (13.8 - 17.2)
University	5745	31.9 (30.0 - 33.7)	49.9 (48.1 - 51.7)	18.0 (15.4 - 20.6)
Employment status				
Actively working	11,730	23.8 (22.2 - 25.4)	40.7 (38.9 - 42.4)	16.9 (14.5 - 19.3)
Student	974	19.0 (14.3 - 23.6)	29.8 (25.0 - 34.6)	10.8 (4.1 - 17.5)
Retired	10,636	39.1 (37.4 - 40.8)	57.3 (55.3 - 59.3)	18.2 (15.6 - 20.8)
Unemployed/other	5647	29.5 (27.8 - 31.1)	44.6 (42.7 - 46.5)	15.1 (12.6 - 17.7)
On-site activity				

XSL•FO
RenderX

	Number	2019 coverage, % (95% CI)[c]	2020 coverage, % (95% CI)[c]	Standardized difference, % (95% CI)[c]
No	16,267	34.7 (33.6 - 35.9)	50.6 (49.3 - 51.9)	15.8 (14.1 - 17.6)
On-site work	10,847	25.8 (24.2 - 27.3)	43.4 (41.8 - 45.0)	17.6 (15.4 - 19.8)
Other on-site activity	1873	25.0 (21.9 - 28.1)	36.3 (33.0 - 39.5)	11.3 (6.8 - 15.8)
Household and residence characteristics				
Living with older adults				
No	18,778	29.6 (28.6 - 30.6)	44.6 (43.4 - 45.7)	15.0 (13.5 - 16.4)
Yes	10,195	34.3 (33.0 - 35.7)	51.5 (49.8 - 53.2)	17.2 (15.0 - 19.4)
Living with children <3 years				
No	27,792	31.3 (30.4 - 32.1)	46.8 (45.7 - 47.8)	15.5 (14.1 - 16.9)
Yes	1195	33.6 (29.4 - 37.7)	46.8 (42.9 - 50.7)	13.2 (7.6 - 18.9)
Household size				
1 - 2 people	11,816	32.2 (31.0 - 33.4)	48.5 (47.0 - 49.9)	16.3 (14.4 - 18.1)
≥3 people	17,171	30.6 (29.5 - 31.8)	45.7 (44.4 - 46.9)	15.0 (13.3 - 16.8)
Municipality size (population)				
>100,000	8727	32.0 (30.5 - 33.4)	47.6 (45.9 - 49.3)	15.7 (13.5 - 17.8)
20,000 - 100,000	8495	30.5 (29.1 - 31.9)	46.1 (44.3 - 47.9)	15.5 (13.3 - 17.8)
5000 - 20,000	6075	30.2 (28.4 - 31.9)	46.0 (44.2 - 47.8)	15.9 (13.4 - 18.3)
<5000	5690	33.0 (31.4 - 34.5)	46.8 (44.3 - 49.4)	13.9 (10.9 - 16.9)
Relative income of the residential area				
Under 25th percentile	7757	30.6 (29.1 - 32.1)	46.2 (44.1 - 48.2)	15.6 (13.1 - 18.1)
25th to 50th percentile	7272	33.1 (31.4 - 34.7)	47.2 (45.3 - 49.1)	14.2 (11.6 - 16.7)
50th to 75th percentile	6517	30.0 (28.6 - 31.4)	44.7 (42.8 - 46.5)	14.7 (12.3 - 17.0)
Over 75th percentile	7441	31.7 (30.0 - 33.5)	49.1 (47.1 - 51.1)	17.3 (14.7 - 20.0)
Health-related characteristics				
Self-rated health				
Very good or good	24,451	29.9 (29.0 - 30.8)	45.7 (44.6 - 46.8)	15.8 (14.4 - 17.2)
Bad or very bad	4536	38.4 (36.7 - 40.1)	52.9 (51.0 - 54.9)	14.6 (12.0 - 17.1)
Chronic diseases				
None	10,720	22.5 (21.1 - 23.9)	38.1 (36.5 - 39.8)	15.6 (13.5 - 17.8)
1	7646	29.9 (28.6 - 31.3)	47.1 (45.5 - 48.6)	17.1 (15.1 - 19.2)
2 - 3	8451	37.2 (35.7 - 38.7)	53.9 (52.1 - 55.7)	16.6 (14.3 - 19.0)
≥4	2170	49.9 (47.1 - 52.6)	63.9 (61.0 - 66.8)	14.1 (10.1 - 18.1)
Disability				
No	26,688	30.7 (29.8 - 31.6)	46.1 (45.1 - 47.2)	15.4 (14.0 - 16.8)
Yes	2118	39.8 (37.2 - 42.4)	55.7 (52.8 - 58.7)	16.0 (12.1 - 19.9)
BMI				
Underweight and normal weight	10,986	29.6 (28.5 - 30.8)	45.0 (43.6 - 46.4)	15.3 (13.5 - 17.2)
Overweight	11,425	31.7 (30.6 - 32.9)	47.2 (45.8 - 48.6)	15.5 (13.6 - 17.3)
Obesity	6571	33.5 (32.0 - 35.0)	49.4 (47.7 - 51.1)	15.9 (13.6 - 18.1)
Tobacco use				
Never	14,848	32.0 (30.9 - 33.0)	46.7 (45.4 - 48.0)	14.8 (13.1 - 16.4)

XSL•FO
RenderX

	Number	2019 coverage, % (95% CI)[c]	2020 coverage, % (95% CI)[c]	Standardized difference, % (95% CI)[c]
Former	8073	34.3 (33.0 - 35.7)	52.5 (50.9 - 54.0)	18.1 (16.1 - 20.2)
Current	6064	25.0 (23.5 - 26.5)	39.9 (38.1 - 41.7)	14.9 (12.5 - 17.2)
COVID-19–related characteristics				
Personal history				
No infection	24,800	31.5 (30.6 - 32.4)	47.0 (46.0 - 48.1)	15.5 (14.1 - 16.9)
Not-severe infection	3445	30.3 (28.2 - 32.4)	44.7 (42.0 - 47.3)	14.4 (11.0 - 17.7)
Pneumonia or hospitalization	742	32.0 (28.2 - 35.7)	49.1 (44.2 - 53.9)	17.1 (10.9 - 23.2)
Contact with case				
No contact	18,584	31.9 (30.9 - 32.9)	46.6 (45.4 - 47.8)	14.7 (13.1 - 16.2)
Noncohabitant symptomatic contact	1254	26.4 (23.2 - 29.6)	44.7 (41.1 - 48.2)	18.3 (13.5 - 23.1)
Noncohabitant case contact	4325	30.6 (28.6 - 32.5)	47.7 (45.8 - 49.6)	17.1 (14.4 - 19.9)
Household symptomatic contact	1380	29.9 (26.3 - 33.5)	46.4 (42.8 - 50.0)	16.5 (11.4 - 21.5)
Household case contact	3444	31.9 (29.6 - 34.2)	47.8 (45.0 - 50.6)	15.9 (12.3 - 19.6)
Cumulative provincial 7-day incidence (per 100,000)				
<100	2331	27.5 (25.3 - 29.7)	41.9 (39.1 - 44.7)	14.4 (10.8 - 17.9)
100 - 200	18,378	32.0 (31.0 - 33.0)	47.3 (46.1 - 48.5)	15.3 (13.7 - 16.9)
200 - 300	7447	30.2 (28.5 - 31.9)	46.5 (44.3 - 48.7)	16.3 (13.5 - 19.0)
≥300	831	35.1 (31.2 - 39.0)	49.9 (45.4 - 54.3)	14.8 (8.9 - 20.7)
Number of people in close contact				
None	3020	33.7 (31.3 - 36.1)	47.1 (44.3 - 49.9)	13.4 (9.7 - 17.1)
1 - 2	8788	32.3 (31.0 - 33.6)	48.7 (47.1 - 50.3)	16.4 (14.3 - 18.5)
3 - 5	11,372	31.4 (30.2 - 32.7)	47.0 (45.6 - 48.3)	15.5 (13.7 - 17.4)
6 - 9	4328	29.0 (27.1 - 30.8)	45.5 (43.5 - 47.5)	16.5 (13.7 - 19.2)
≥10	1479	26.9 (23.5 - 30.2)	39.4 (36.1 - 42.7)	12.5 (7.9 - 17.2)
Facemask type used				
Cloth	4312	28.4 (26.6 - 30.2)	40.5 (38.6 - 42.4)	12.1 (9.5 - 14.7)
Surgical	18,449	30.8 (29.8 - 31.8)	46.4 (45.2 - 47.6)	15.6 (14.1 - 17.2)
FFP2	6170	34.7 (33.1 - 36.2)	52.2 (50.4 - 54.0)	17.5 (15.1 - 19.9)
Facemask use in family meetings				
No	9850	29.1 (27.9 - 30.4)	43.0 (41.6 - 44.4)	13.8 (11.9 - 15.7)
Sometimes	2618	30.3 (28.1 - 32.5)	46.1 (43.6 - 48.7)	15.9 (12.5 - 19.2)
Yes	11,852	33.2 (31.9 - 34.4)	49.3 (47.8 - 50.7)	16.1 (14.2 - 18.0)
Did not meet	4667	32.4 (30.7 - 34.1)	49.9 (47.8 - 52.1)	17.5 (14.8 - 20.2)
Indoor bar attendance				
<Once/month	21,629	32.5 (31.5 - 33.4)	47.9 (46.7 - 49.0)	15.4 (13.9 - 16.9)
1 - 3 times/month	3938	27.6 (25.7 - 29.5)	44.0 (41.9 - 46.0)	16.4 (13.6 - 19.2)
≥Once/week	3420	28.3 (26.1 - 30.4)	43.5 (41.2 - 45.9)	15.3 (12.1 - 18.5)
Indoor family meetings with >10 people				
Never	25,078	31.7 (30.8 - 32.6)	47.3 (46.2 - 48.4)	15.6 (14.2 - 17.0)

XSL•FO
RenderX

	Number	2019 coverage, % (95% CI)[c]	2020 coverage, % (95% CI)[c]	Standardized difference, % (95% CI)[c]
1 - 5 times	3098	28.7 (26.4 - 31.1)	43.3 (40.9 - 45.8)	14.6 (11.2 - 18.0)
>5 times	811	30.7 (26.3 - 35.1)	44.4 (40.1 - 48.7)	13.7 (7.6 - 19.8)
Indoor events, most people without mask				
No	26,482	31.6 (30.7 - 32.4)	47.2 (46.1 - 48.3)	15.6 (14.2 - 17.0)
Yes	2505	29.3 (26.8 - 31.8)	42.9 (40.3 - 45.5)	13.6 (10.1 - 17.2)
Use of public transport				
No	24,045	31.8 (30.9 - 32.7)	47.4 (46.3 - 48.5)	15.7 (14.2 - 17.1)
Yes	4942	29.8 (28.1 - 31.5)	44.3 (42.4 - 46.3)	14.6 (12.0 - 17.2)

[a]The 2020 campaign was not finished when the survey took place; in this table, participants answering that they had still not received the vaccine but intended to be vaccinated were considered vaccinated.

[b]Selected target groups: population older than 65 years, health care or sociosanitary workers or at-home caregivers, population with some risk condition, living with someone with some risk condition, and security and emergency workers.

[c]Standardized to the overall distribution of the influenza vaccination target group in the whole study population, except coverage for age groups that were not standardized.

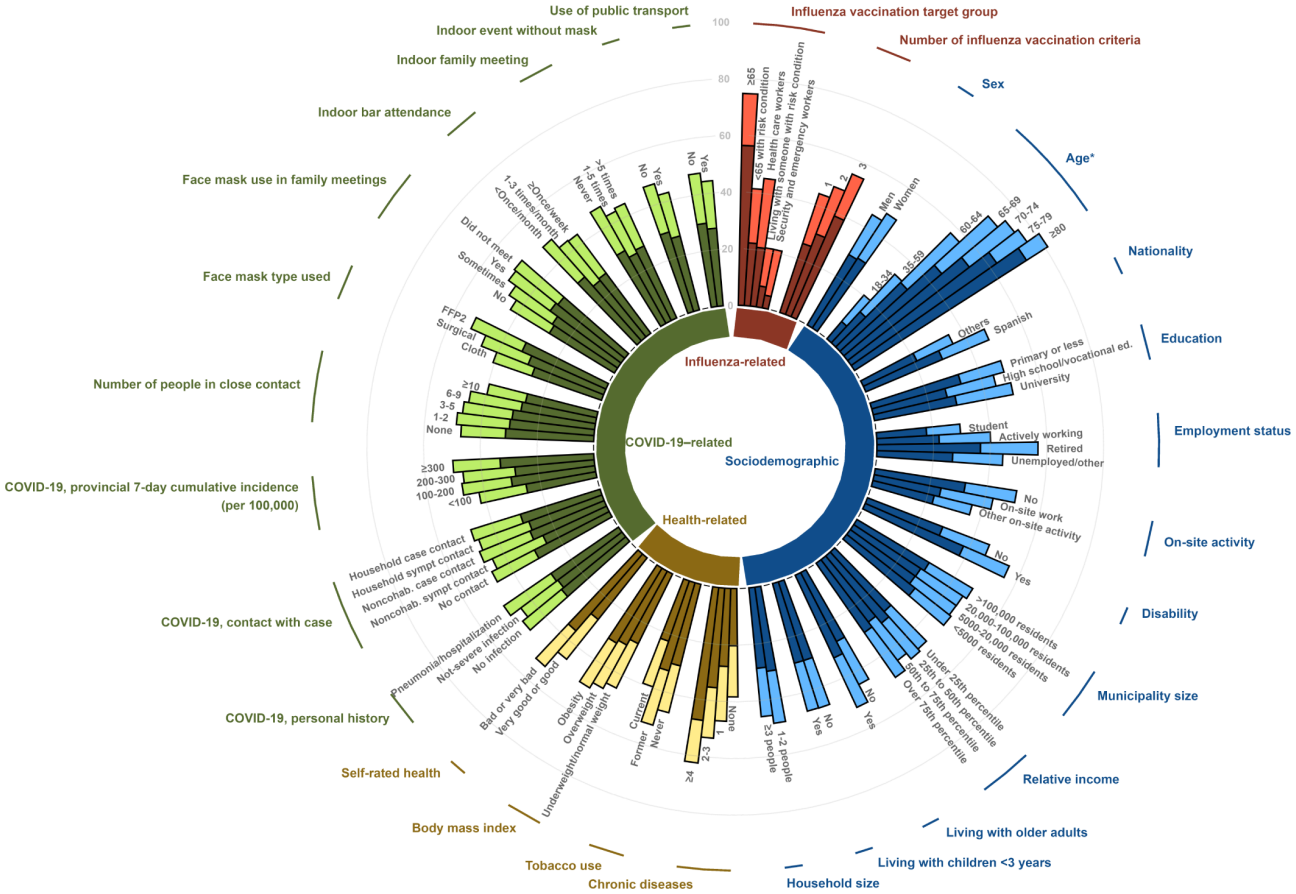
[d]Risk conditions: diabetes; chronic lung, liver, or kidney disease; asthma, cardiovascular disease, hematological neoplasm, any other cancer in the last 5 years, immunosuppressor treatment, or severe obesity.

According to sociodemographic characteristics, in both influenza seasons, the standardized prevalence of vaccine uptake was higher among those with lower or higher educational level (compared to intermediate level), Spanish nationality, retired, or those living with older people, without clear differences by municipality size. On the other hand, in 2020, coverage was higher among women, and also somewhat higher in those residing in wealthier areas. Focusing on health-related factors, in both seasons, coverage was higher in those with bad or very bad self-rated health, disability, or obesity, and increased with the number of chronic diseases. Regarding tobacco use, the highest prevalence of vaccinees and the highest absolute increase in coverage (18.1%, 95% CI 16.1% - 20.2%) corresponded to former smokers. For COVID-19–related factors, people reporting pneumonia, contact with noncohabitant suspected or confirmed cases, using an FFP2 face mask, or not attending family meetings showed a higher absolute increase in coverage, while no differences were observed by provincial COVID-19 incidence.

Figure 2 shows a general picture of vaccination coverage in 2020, highlighting the contribution of the new vaccinees (those not receiving the vaccine in 2019), by participant characteristics, standardized to the overall vaccination target group composition. In all, 17.6% (95% CI 16.9% - 18.3%) of the participants were new vaccinees, ranging from 8.9% (7.1% - 10.7%) in the oldest age group to around 25% for those aged 60 - 69 years and HCWs (Table S1 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 2. Prevalences of influenza vaccination, standardized by target group distribution in the overall population, by influenza-related, sociodemographic, health-related, and COVID-19–related factors. Dark colors represent the proportion of participants vaccinated in both the 2019 and 2020 campaigns, and light colors represent those vaccinated only in 2020 (new vaccinees). Noncohab: noncohabitant. *Prevalence by age not standardized.



One of 4 participants unvaccinated in 2019 opted for influenza vaccination in 2020 (Table 2). This change was most prevalent, once standardized by sex, education, and autonomous community, among participants older than 65 years (44.8%, 95% CI 42.2% - 47.3%), followed by HCWs (30.0%, 95% CI 26.7% - 33.2%). On the other hand, among those living with someone with risk conditions, only 15% of unvaccinated individuals changed their mind (95% CI 13.8% - 16.2%). Regarding other factors, our results, additionally standardized by age and influenza vaccination target group, showed that this behavior change was more frequent in participants meeting more than one influenza vaccination criteria ($SPR_{\text{two vs one}}$: 1.10, 95% CI 1.01 - 1.19), women (SPR: 1.15, 95% CI 1.08 - 1.22), older participants ($SPR_{\text{65-69 vs 18-34}}$: 3.66, 95% CI 3.20 - 4.19), those with higher education ($SPR_{\text{university vs primary}}$: 1.28, 95% CI 1.15 - 1.42), those with chronic diseases ($SPR_{\geq 4 \text{ vs none}}$: 1.35, 95% CI 1.15 - 1.58), former smokers ($SPR_{\text{former vs never}}$: 1.13, 95% CI 1.05 - 1.22), or those living with older people (SPR: 1.19, 95% CI 1.11 - 1.28). With respect to COVID-19–related factors, SARS-CoV-2 infection was not clearly associated with influenza vaccination, while being aware of previous contact with COVID-19 cases increased new vaccination ($SPR_{\text{nonhousehold case contact vs no contact}}$: 1.14, 95% CI 1.04 - 1.24). Curiously, local incidence of COVID-19 at interview was inversely related to the proportion of new vaccinees ($SPR_{\text{highest vs lowest}}$: 0.79, 95% CI 0.59 - 1.05). Regarding compliance with COVID-19 prevention measures, the use of masks in family meetings ($SPR_{\text{always vs no}}$: 1.19, 95% CI 1.09 - 1.30) and the type of mask worn ($SPR_{\text{FFP2 vs cloth}}$: 1.43, 95% CI 1.28 - 1.60) showed a remarkable association with vaccination.

XSL•FO
RenderX

Table . Prevalence of new influenza vaccination in 2020[a] among participants unvaccinated in 2019, by sociodemographic, health-related, and COVID-19–related factors.

	Number	Crude prevalence, % (95% CI)	Standardized prevalence, % (95% CI)[b]	Standardized prevalence difference, % (95% CI)[b]	Standardized prevalence ratio, (95% CI)[b]
Total	19,334	25.7 (24.7 to 26.7)			
Influenza-related characteristics					
Influenza vaccination target group					
≥65 years	4214	44.1 (41.8 to 46.4)	44.8 (42.2 to 47.3)	Reference	Reference
<65 years with a risk condition	5880	25.7 (24.0 to 27.3)	25.7 (24.1 to 27.4)	–19.0 (–21.9 to –16.1)	0.58 (0.53 to 0.62)
Health care workers	1621	31.8 (28.6 to 35.1)	30.0 (26.7 to 33.2)	–14.8 (–18.9 to –10.7)	0.67 (0.59 to 0.76)
Living with someone with a risk condition	7309	15.0 (13.8 to 16.2)	14.9 (13.7 to 16.2)	–29.8 (–32.6 to –27.0)	0.33 (0.30 to 0.37)
Security and emergency workers	310	17.6 (11.8 to 23.3)	19.3 (13.3 to 25.4)	–25.4 (–32.0 to –18.9)	0.43 (0.31 to 0.59)
Number of influenza vaccination criteria					
1	12,992	21.3 (20.3 to 22.3)	24.7 (23.5 to 25.8)	Reference	Reference
2	5244	33.7 (31.6 to 35.7)	27.0 (25.2 to 28.9)	2.4 (0.2 to 4.5)	1.10 (1.01 to 1.19)
3	1098	42.8 (38.2 to 47.5)	29.2 (25.1 to 33.3)	4.6 (0.3 to 8.8)	1.19 (1.02 to 1.37)
Sociodemographic characteristics					
Sex					
Men	8852	22.7 (21.4 to 24.0)	23.8 (22.5 to 25.1)	Reference	Reference
Women	10,482	28.5 (27.3 to 29.7)	27.3 (26.2 to 28.5)	3.5 (2.1 to 5.0)	1.15 (1.08 to 1.22)
Age (years)					
18 - 34	3574	13.8 (12.1 to 15.4)	13.4 (11.8 to 15.0)	Reference	Reference
35 - 59	9748	21.8 (20.5 to 23.0)	21.5 (20.3 to 22.7)	8.1 (6.3 to 9.9)	1.60 (1.42 to 1.81)
60 - 64	1798	38.0 (34.7 to 41.2)	38.6 (35.3 to 41.9)	25.2 (21.5 to 28.9)	2.87 (2.48 to 3.33)
65 - 69	1831	48.4 (45.2 to 51.6)	49.2 (46.0 to 52.5)	35.8 (32.2 to 39.4)	3.66 (3.20 to 4.19)
70 - 74	1183	45.9 (41.8 to 49.9)	47.1 (43.0 to 51.3)	33.7 (29.3 to 38.1)	3.51 (3.03 to 4.07)
75 - 79	675	37.1 (32.3 to 41.8)	38.8 (33.9 to 43.8)	25.4 (20.2 to 30.6)	2.89 (2.43 to 3.44)
≥80	525	34.3 (28.7 to 40.0)	37.0 (30.9 to 43.1)	23.5 (17.2 to 29.9)	2.75 (2.24 to 3.38)
Nationality					
Spanish	18,692	26.1 (25.0 to 27.1)	25.9 (24.9 to 26.9)	Reference	Reference
Other	642	16.3 (12.6 to 20.1)	19.7 (15.5 to 23.9)	–6.2 (–10.4 to –2.0)	0.76 (0.61 to 0.94)
Education					
Primary or less	3858	32.6 (30.4 to 34.8)	23.4 (21.6 to 25.3)	Reference	Reference
High school/vocational education	11,250	22.9 (21.7 to 24.1)	24.8 (23.5 to 26.0)	1.3 (–0.8 to 3.4)	1.06 (0.97 to 1.15)
University	4226	27.3 (25.3 to 29.2)	30.0 (28.0 to 31.9)	6.5 (3.8 to 9.2)	1.28 (1.15 to 1.42)
Household and residence characteristics					
Living with older adults					
No	13,660	22.9 (21.8 to 24.0)	24.3 (23.2 to 25.4)	Reference	Reference
Yes	5671	32.9 (31.0 to 34.8)	29.0 (27.2 to 30.7)	4.7 (2.7 to 6.6)	1.19 (1.11 to 1.28)
Household size					

XSL•FO
RenderX

	Number	Crude prevalence, % (95% CI)	Standardized prevalence, % (95% CI)[b]	Standardized prevalence difference, % (95% CI)[b]	Standardized prevalence ratio, (95% CI)[b]
1 - 2 people	6615	32.9 (31.2 to 34.7)	26.2 (24.7 to 27.8)	Reference	Reference
≥3 people	12,719	22.1 (21.0 to 23.2)	25.4 (24.1 to 26.6)	–0.9 (–2.8 to 1.1)	0.97 (0.90 to 1.04)
Municipality size (population)					
>100,000	5776	26.8 (25.2 to 28.4)	26.0 (24.5 to 27.5)	Reference	Reference
20,000 - 100,000	5829	24.9 (23.0 to 26.9)	25.9 (24.0 to 27.8)	–0.1 (–2.6 to 2.3)	0.99 (0.91 to 1.09)
5000 - 20,000	4167	25.1 (22.9 to 27.2)	25.4 (23.4 to 27.5)	–0.6 (–3.1 to 1.9)	0.98 (0.89 to 1.08)
<5000	3562	24.6 (21.8 to 27.3)	24.5 (21.3 to 27.6)	–1.6 (–5.1 to 2.0)	0.94 (0.82 to 1.08)
Relative income of the residential area					
Under 25th percentile	5226	25.4 (23.1 to 27.7)	27.1 (24.8 to 29.3)	Reference	Reference
25th to 50th percentile	4776	24.4 (22.7 to 26.2)	25.0 (23.2 to 26.8)	–2.1 (–5.0 to 0.8)	0.92 (0.83 to 1.03)
50th to 75th percentile	4424	24.2 (22.2 to 26.2)	23.9 (22.0 to 25.7)	–3.2 (–6.1 to –0.2)	0.88 (0.79 to 0.99)
Over 75th percentile	4908	28.8 (26.8 to 30.9)	26.9 (25.0 to 28.7)	–0.2 (–3.1 to 2.7)	0.99 (0.89 to 1.11)
Health-related characteristics					
Self-rated health					
Very good or good	17,029	24.9 (23.8 to 26.0)	25.4 (24.4 to 26.5)	Reference	Reference
Bad or very bad	2305	31.8 (29.3 to 34.2)	27.4 (25.2 to 29.6)	2.0 (–0.4 to 4.4)	1.08 (0.99 to 1.18)
Chronic diseases					
None	8740	19.8 (18.6 to 21.0)	22.9 (21.2 to 24.5)	Reference	Reference
1	5130	28.9 (27.2 to 30.7)	26.9 (25.2 to 28.6)	4.0 (1.6 to 6.5)	1.18 (1.07 to 1.30)
2 - 3	4644	31.9 (29.9 to 33.8)	28.2 (26.1 to 30.2)	5.3 (2.5 to 8.2)	1.23 (1.10 to 1.38)
≥4	820	38.3 (33.9 to 42.7)	30.8 (26.8 to 34.9)	8.0 (3.4 to 12.5)	1.35 (1.15 to 1.58)
Disability					
No	18,184	25.3 (24.3 to 26.3)	25.5 (24.5 to 26.5)	Reference	Reference
Yes	1150	32.1 (28.5 to 35.7)	28.3 (24.9 to 31.8)	2.8 (–0.7 to 6.4)	1.11 (0.98 to 1.26)
BMI					
Underweight and normal weight	7879	23.4 (22.0 to 24.7)	25.1 (23.6 to 26.6)	Reference	Reference
Overweight	7351	26.4 (24.9 to 27.9)	25.2 (23.8 to 26.7)	0.1 (–1.8 to 2.1)	1.01 (0.93 to 1.09)
Obesity	4101	29.4 (27.4 to 31.3)	27.7 (25.8 to 29.6)	2.6 (0.3 to 4.9)	1.10 (1.01 to 1.20)
Tobacco use					
Never	9584	25.6 (24.3 to 26.9)	25.5 (24.1 to 26.8)	Reference	Reference
Former	4903	32.8 (30.9 to 34.7)	28.8 (27.0 to 30.6)	3.4 (1.3 to 5.4)	1.13 (1.05 to 1.22)
Current	4847	19.3 (17.7 to 20.9)	22.7 (21.0 to 24.5)	–2.7 (–4.8 to –0.7)	0.89 (0.82 to 0.97)
COVID-19–related characteristics					
Personal history					
No infection	16,510	25.8 (24.8 to 26.9)	26.0 (24.9 to 27.0)	Reference	Reference
Not-severe infection	2361	23.5 (20.8 to 26.3)	23.2 (20.6 to 25.9)	–2.7 (–5.4 to 0.0)	0.89 (0.80 to 1.00)

XSL•FO
RenderX

	Number	Crude prevalence, % (95% CI)	Standardized prevalence, % (95% CI)[b]	Standardized prevalence difference, % (95% CI)[b]	Standardized prevalence ratio, (95% CI)[b]
Pneumonia or hospitalization	463	31.0 (24.8 to 37.1)	28.6 (22.9 to 34.3)	2.6 (–3.1 to 8.4)	1.10 (0.90 to 1.35)
Contact with case					
No contact	11,685	26.2 (24.9 to 27.4)	24.5 (23.3 to 25.7)	Reference	Reference
Noncohabitant symptomatic contact	974	24.1 (20.3 to 28.0)	26.4 (22.4 to 30.5)	1.9 (–2.2 to 6.1)	1.08 (0.92 to 1.26)
Noncohabitant case contact	3226	25.5 (23.3 to 27.7)	27.8 (25.6 to 30.0)	3.3 (0.9 to 5.8)	1.14 (1.04 to 1.24)
Household symptomatic contact	1035	23.7 (20.1 to 27.3)	27.4 (23.6 to 31.3)	2.9 (–1.0 to 6.9)	1.12 (0.97 to 1.29)
Household case contact	2414	25.3 (22.3 to 28.4)	27.5 (24.4 to 30.6)	3.0 (–0.2 to 6.2)	1.12 (1.00 to 1.26)
Provincial 7-day cumulative incidence (per 100,000)					
<100	1654	22.6 (19.8 to 25.4)	33.7 (27.3 to 40.1)	Reference	Reference
100 - 200	12,249	25.8 (24.5 to 27.0)	25.4 (24.1 to 26.8)	–8.3 (–15.0 to –1.5)	0.75 (0.62 to 0.93)
200 - 300	4925	26.5 (24.4 to 28.7)	24.2 (21.7 to 26.7)	–9.5 (–16.6 to –2.3)	0.72 (0.57 to 0.90)
≥300	506	27.0 (21.2 to 32.8)	26.6 (20.3 to 32.8)	–7.1 (–15.5 to 1.3)	0.79 (0.59 to 1.05)
Face mask type used					
Cloth	3207	18.0 (16.3 to 19.8)	20.9 (19.0 to 22.8)	Reference	Reference
Surgical	12,183	26.1 (24.9 to 27.3)	25.5 (24.3 to 26.7)	4.7 (2.5 to 6.8)	1.22 (1.11 to 1.35)
FFP2	3924	31.1 (29.0 to 33.3)	29.8 (27.8 to 31.9)	9.0 (6.2 to 11.7)	1.43 (1.28 to 1.60)
Face mask use in family meetings					
No	6898	21.9 (20.5 to 23.3)	22.9 (21.5 to 24.3)	Reference	Reference
Sometimes	1873	24.8 (22.0 to 27.6)	26.7 (23.9 to 29.6)	3.8 (0.7 to 7.0)	1.17 (1.03 to 1.32)
Yes	7634	27.8 (26.1 to 29.5)	27.3 (25.6 to 28.9)	4.4 (2.2 to 6.5)	1.19 (1.09 to 1.30)
Did not meet	2929	31.1 (28.6 to 33.6)	28.0 (25.7 to 30.2)	5.1 (2.5 to 7.6)	1.22 (1.11 to 1.35)
Indoor bar attendance					
<Once/month	13,812	27.0 (25.8 to 28.2)	26.0 (24.9 to 27.2)	Reference	Reference
1 - 3 times/month	2980	23.0 (20.8 to 25.2)	25.2 (22.9 to 27.5)	–0.8 (–3.3 to 1.7)	0.97 (0.88 to 1.07)
≥Once/week	2542	21.8 (19.4 to 24.2)	24.2 (21.8 to 26.6)	–1.8 (–4.4 to 0.8)	0.93 (0.84 to 1.04)
Indoor family meetings with >10 people					
Never	16,461	26.6 (25.5 to 27.7)	25.9 (24.9 to 27.0)	Reference	Reference
1 - 5 times	2299	20.6 (18.4 to 22.9)	23.9 (21.5 to 26.3)	–2.0 (–4.6 to 0.5)	0.92 (0.83 to 1.02)
>5 times	574	21.7 (17.0 to 26.5)	25.1 (20.4 to 29.7)	–0.9 (–5.6 to 3.9)	0.97 (0.80 to 1.17)
Indoor events, most people without mask					
No	17,450	26.4 (25.3 to 27.4)	25.9 (24.8 to 26.9)	Reference	Reference
Yes	1884	19.9 (17.4 to 22.4)	23.9 (21.1 to 26.7)	–1.9 (–4.9 to 1.0)	0.93 (0.82 to 1.05)

[a]The 2020 campaign was not finished when the survey took place; in this table, participants answering that they had still not received the vaccine but intended to be vaccinated were considered vaccinated.

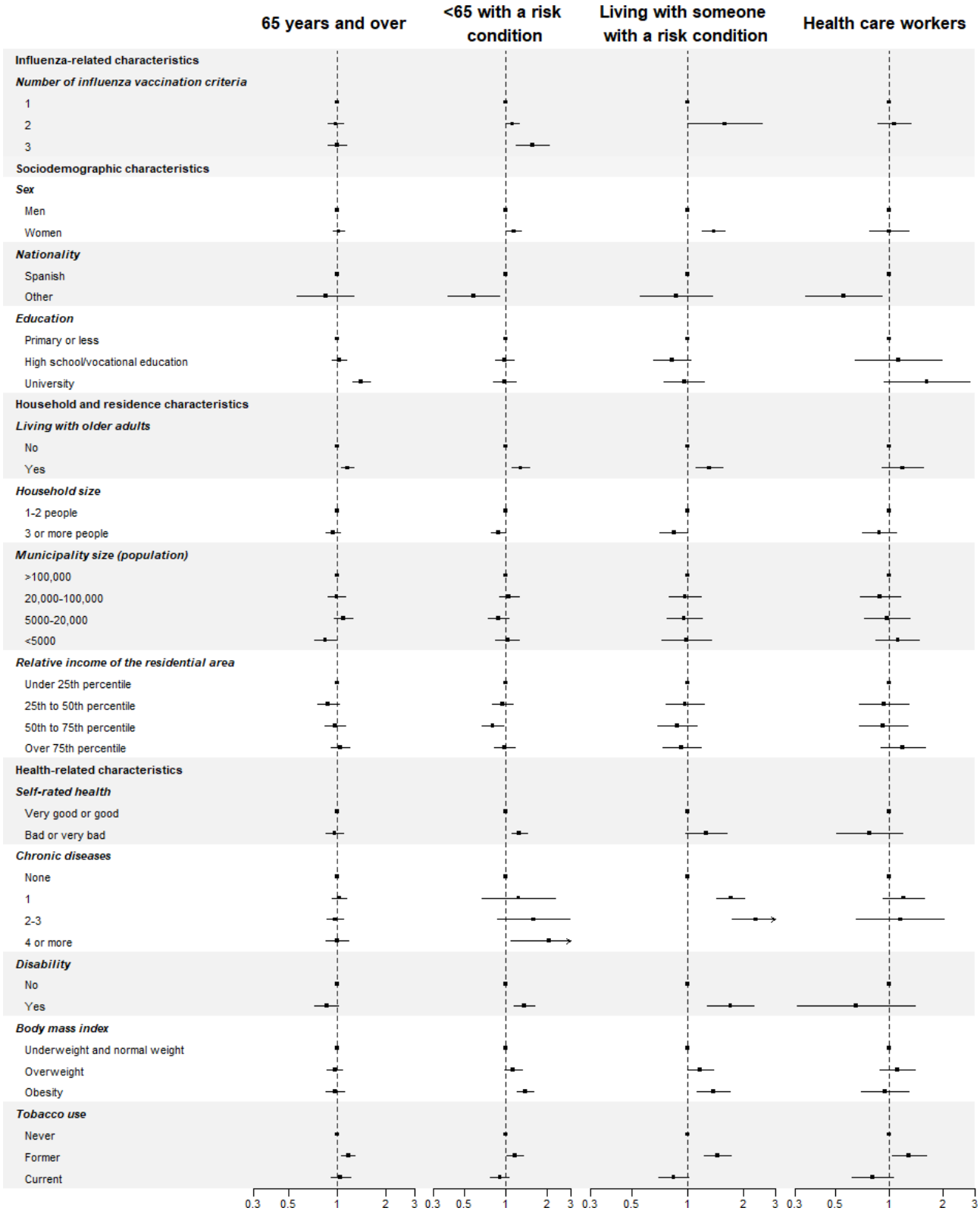
[b]Standardized to the overall distribution of influenza vaccination group, sex, age, education, and autonomous community in the unvaccinated target population in 2019, except for influenza vaccination groups and age groups that were not mutually standardized.

Some of these associations differed by target group (Figure 3, Figure 4, Table S2 in Multimedia Appendix 1). Number of influenza vaccination criteria, female sex, and poorer health status were related to a higher prevalence of new vaccinees mainly in people younger than 65 years with risk conditions or those living with someone at risk (*P* heterogeneity ranged from

XSL•FO
RenderX

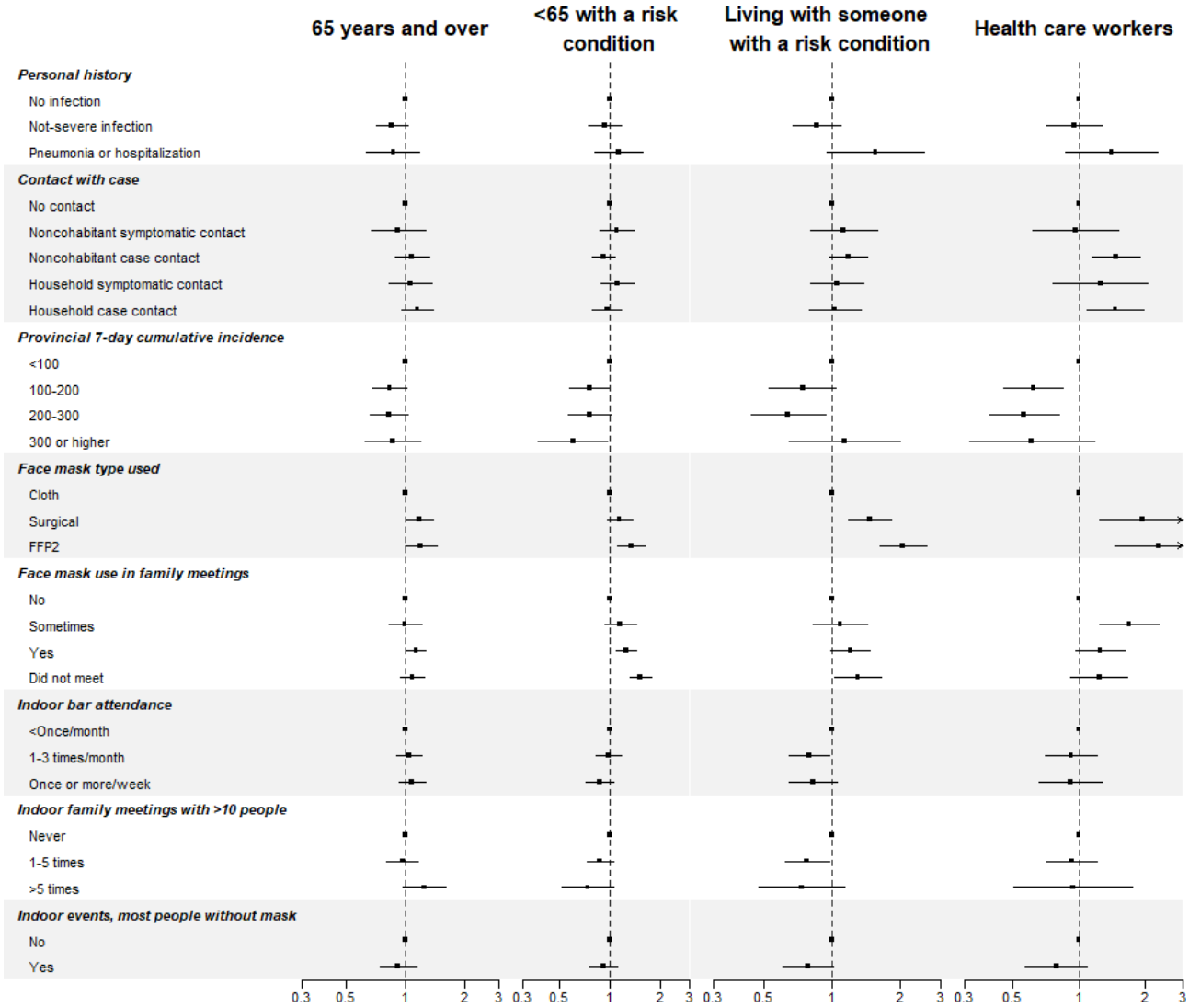
<.001 to .01). In contrast, university education showed a higher proportion of new vaccinees in people older than 65 years (SPR: 1.41, 95% CI 1.24 - 1.60) and in HCWs (SPR: 1.62, 0.94 - 2.79; *P* heterogeneity=.003). With respect to COVID-19–related factors (Figure 4), contact with COVID-19 cases was associated with new vaccinations mainly in HCWs (*P* heterogeneity=.10). The higher prevalence of new influenza vaccine uptake among those wearing surgical and FFP2 face masks compared to cloth masks was stronger in HCWs and people living with someone with risk conditions (*P* heterogeneity=.003).

Figure 3. Standardized prevalence ratios of new influenza vaccination in 2020 among participants unvaccinated in 2019 across influenza-related, sociodemographic, and health-related factors, by target group.



XSL•FO
RenderX

Figure 4. Standardized prevalence ratios of new influenza vaccination in 2020 among participants unvaccinated in 2019 across COVID-19–related factors, by target group.



Sensitivity analyses excluding participants reporting their intention to be vaccinated in the 2020‐2021 campaign but who had not yet received the vaccine led to a 5% lower vaccination coverage estimate in 2020 and a 6% reduction in the absolute proportion of newly vaccinated. No remarkable differences were observed between both approaches in the reported associations with explanatory factors except for a slightly smaller increase in coverage in the group of individuals under 65 years old with risk conditions compared to the other target groups and a higher difference in the proportion of new vaccinees depending on the cumulative incidence in the province of residence (Tables S3-S4 in Multimedia Appendix 1).

Discussion

Principal Findings

One of the few positive side effects of the COVID-19 pandemic may have been the increase in the uptake of influenza vaccination, probably influenced by a higher awareness of the risk associated with respiratory diseases and the desire to avoid an infection that might be mistaken for COVID-19 and avoid any new saturation of the health care system. In this study, taking advantage of the ENE-COVID study design, we provide nationally representative estimates and a description of factors associated with influenza vaccination coverage in the main target groups in Spain in the influenza season before the appearance of SARS-CoV-2 (2019/2020) and compare it to the first season after its emergence (2020/2021). Our data show a generalized increase in vaccination, especially among HCWs. In addition, we showed that among people unvaccinated in 2019, older people, women, more educated groups, those with comorbidities, or those living with older adults were more prone to accept the vaccine in 2020. Some COVID-19–related factors were also found to be associated with the decision to vaccinate against seasonal influenza, such as the perceived risk of SARS-CoV-2 infection, reflected here by the awareness of having had contact with COVID-19 cases, the type of masks used, their use in family meetings, and not attending family/friend meetings. It should be noted that this study was conducted in November 2020, when the influenza vaccination campaign was ongoing, the number of COVID-19 cases was quickly rising (the third COVID-19 epidemic wave), COVID-19 vaccines were still not authorized, and more than 40,000 COVID-19–related deaths had occurred in the country.

XSL•FO
RenderX

During the 2020/2021 influenza season, a historical reduction in influenza transmission was observed worldwide, including in Spain [19,20]. The circulation of the highly transmissible SARS-CoV-2 in a fully susceptible population probably displaced other respiratory viruses, including influenza, which, together with COVID-19 mitigation measures, probably had a relevant role, further aided by the increase in influenza vaccination. Fear of unknown but potentially devastating effects of a "twindemic" [21,22] may have favored the adoption of preventive behaviors, including influenza vaccination, as suggested in other countries [14,23,24]. The reinforced public health influenza vaccination campaign implemented that year, with specific messages targeted at HCWs, likely contributed to increased coverage. Our data allowed us to quantify this effect, showing that, in Spain, vaccination coverage among the main target groups augmented from 31.4% in 2019 to 46.8% in 2020, breaking the downward trend observed in Spain [15,25] and other European countries in previous years [3].

Comparison With Prior Work

In 2020/2021, coverage was around 75% among older people, attaining the recommended goal established by the World Health Organization and the European Council [5]. This figure is higher than the 66% published by the Spanish Ministry of Health [15], although they are not fully comparable, as they come from different data sources (self-reported versus regional health information systems) and there are 2 regions not included in the estimation provided by the Ministry. For the other target groups, even though still remaining far from the goal [5,15], coverage figures also showed a marked improvement, which was particularly noteworthy among HCWs, a group heavily affected by the pandemic that year. The burden experienced during the first COVID-19 epidemic wave, when the health care system in Spain became overwhelmed, could have boosted this practice, though it was not maintained in 2021 and 2022 [15]. This finding, also observed in other European countries [26], contrasts with the stable evolution in the 65 years and older population, which may reflect the success of the vaccination strategy for this age group, which included the simultaneous administration of the influenza vaccine and the COVID-19 booster in a context of high COVID-19 vaccine effectiveness. On the other hand, among HCWs, risk perception may have decreased once effective COVID-19 vaccines became available, reducing their willingness to receive the influenza vaccine. Still, the drop in coverage did not reach the low levels seen before the pandemic [27]. HCWs reported that important reasons driving influenza vaccine uptake in 2020/2021 were self-protection, protecting family, and protecting patients [26,28]. Highlighting these benefits of vaccination in campaigns aimed at promoting influenza vaccine uptake among HCWs may contribute to their effectiveness, which is especially relevant, as vaccination is a basic measure to protect at-risk people that HCWs work closely with and HCWs' advice may influence the decisions of others.

Focusing on new vaccinations, 1 of 4 people in target groups that were unvaccinated in 2019 reported vaccination in 2020. This figure is similar to estimations based on vaccination intention in the early months of the pandemic [29,30] and is consistent with the results of an Italian survey [14], where 20.4% of adults reported that their intention to receive the influenza vaccination in 2020/2021 was due to the COVID-19 pandemic. Our data showed that the increase in vaccine uptake was not evenly distributed. New vaccination was more frequent in older adults, in women, among Spaniards, and among those with higher education, which is consistent with findings in the prepandemic literature [31]. In addition, current smokers were less likely to change to receiving the vaccine, which could reflect generally lower engagement with healthy lifestyles. Regarding household composition, living with older people was associated with higher vaccination uptake in 2020. Households are an important setting for the transmission of respiratory diseases, with risk of infection from a household contact being up to 38% [32,33].

Pandemic-related factors, such as worries about COVID-19, have been described in the literature as being associated with the intention to receive the influenza vaccine [34]. Our results showed that people who might be considered more cautious, as reflected by their reported attitude toward the prevention of SARS-CoV-2 transmission, showed higher probability of being newly vaccinated. These prudent behaviors may have been prompted by what has been called cues to action, which include both internal and environmental triggers that may activate people's motivation to engage in healthy activities [30]. Overall, these findings also suggest that both protection of others and self-protection represent motivations for vaccination. HCWs are a clear example of this [26,28]. In our study, contact with confirmed COVID-19 cases showed a stronger association with opting for influenza vaccination in 2020 among HCWs than in other target groups.

One of the main strengths of this study is the use of data from the nationwide population-based ENE-COVID study, based on a big sample representative of the noninstitutionalized population in Spain. The high participation rate, as well as the poststratification weights used in the analyses, allowed us to diminish the risk of bias in our estimations. We have analyzed many factors, several of them not usually available at a national level, potentially related to influenza vaccination in people from 5 target groups, allowing for a deeper understanding of the determinants of influenza vaccination. In addition, we evaluated factors associated with the change in influenza vaccine uptake just after the pandemic and provided standardized estimates that took into account possible confounders.

Limitations

This study also had some limitations. First, people living in nursing homes (around 4% of the population aged 65 years and older in Spain) [35] were not included in the ENE-COVID study; consequently, our prevalence figures might be slightly underestimated, as vaccination among institutionalized older people is generally higher. Second, potential inaccuracies in the population register could have precluded contact with current residents in some of the selected households, potentially introducing some degree of selection bias. Additionally, most of the information was self-reported and may have been subject to recall and social desirability biases. In addition, data were collected before the 2020/2021 influenza vaccination campaign had finished (it was extended from the first week of October

XSL•FO
RenderX

2020 to the end of January 2021) and, therefore, for the 2020/2021 influenza season, some participants (7% of the sample) reported intention to be vaccinated rather than vaccination, which could overestimate vaccination rates. However, intention has been considered as a good proxy for vaccine uptake [34] and our sensitivity analysis excluding participants that intended to receive the vaccine showed similar associations with the studied factors. The standardized method used, based on logistic models, assumed no relevant effect modification in the exposure odds ratio by strata of the confounding factors. Therefore, we cannot rule out some degree of model misspecification, but we expect this method to provide stable results and leave little residual confounding. Regarding classification of participants into the target groups of people younger than 65 years with a risk condition or those living with someone with a risk condition, the questionnaire did not collect information on all pathologies that are included as risk conditions for influenza vaccination (ie, neurological diseases, nonmalignant blood diseases, asplenia, causes of immunodepression other than pharmacological treatments, and celiac and inflammatory bowel diseases). Likewise, we were not able to analyze some target groups, such as pregnant women. Another limitation is that we did not consider possible regional differences in changes in coverage and the factors associated with this change, thus limiting the direct extrapolation of our findings to the regional level. Finally, our results should be interpreted in the context of the Spanish National Health System, where influenza vaccination is free of charge for people in the recommended groups, which may reduce their generalizability to contexts with a different health care system. Future studies based on vaccination registries combined with clinical or self-reported information or carried out in specific settings (eg, nursing homes) could overcome some of these limitations.

Conclusions

A large increase in influenza vaccination coverage was observed between 2019 and 2020 in the main target groups in Spain, particularly in HCWs, although vaccination goals were only attained in the group of people aged ≥65 years. In this group, this historic rise in coverage broke the declining trend observed in the previous years. However, now that the COVID-19 pandemic is under control, sustaining or improving these figures may prove challenging. Identifying the components of the 2020 reinforced influenza vaccination campaign that contributed to its effectiveness can inform future campaigns. Additionally, the presented crude and standardized estimates of the characteristics of the population not vaccinated in the past that were associated with becoming vaccinated in a posterior campaign may represent valuable insights to guide targeted interventions and vaccination strategies.

Acknowledgments

This work was supported by the Spanish Ministry of Health; the Institute of Health Carlos III (Ministry of Science and Innovation); and the National Health System, including the Health Services of all autonomous communities and autonomous cities. Funding institutions were involved in the study logistics but had no role in the analysis, interpretation of the data, writing of the report, or decision to submit this manuscript. We thank all participants and the nurses, general practitioners, administrative staff, and other health care workers who collaborated in the undertaking of this study. This study is the result of the efforts of many professionals and the trust and generosity of more than 61,000 participants who understood the importance of providing time, information, and biological samples to learn about the COVID-19 pandemic in Spain.

We thank the members of the ENE-COVID Study Group: Pilar Aparicio Azcárraga, Faustino Blanco, Rodrigo Gutiérrez Fernández, Mariano Martín, Saturnino Mezcua Navarro, Marta Molina, Juan F Muñoz-Montalvo, Matías Salinero Hernández, Jose L Sanmartín, Manuel Cuenca-Estrella, José León Paniagua, Raquel Yotti, Nerea Fernández de Larrea Baz, Pablo Fernández-Navarro, Roberto Pastor-Barriuso, Beatriz Pérez-Gómez, Marina Pollán, Ana Avellón, Giovanni Fedele, Aurora Fernández-García, Jesús Oteo Iglesias, María Teresa Pérez Olmeda, Israel Cruz, Maria Elena Fernández Martínez, Francisco D Rodríguez-Cabrera, Miguel A Hernán, Susana Padrones Fernández, José Manuel Rumbao Aguirre, José M Navarro Marí, Begoña Palop Borrás, Ana Belén Pérez Jiménez, Manuel Rodríguez-Iglesias, Ana María Calvo Gascón, María Luz Lou Alcaine, Ignacio Donate Suárez, Oscar Suárez Álvarez, Mercedes Rodríguez Pérez, Margarita Cases Sanchís, Carlos Javier Villafáfila Gomila, Lluis Carbo Saladrigas, Adoración Hurtado Fernández, Antonio Oliver, Elías Castro Feliciano, María Noemí González Quintana, José María Barrasa Fernández, María Araceli Hernández Betancor, Melisa Hernández Febles, Leopoldo Martín Martín, Luis-Mariano López López, Teresa Ugarte Miota, Inés De Benito Población, María Sagrario Celada Pérez, María Natalia Vallés Fernández, Tomás Maté Enríquez, Miguel Villa Arranz, Marta Domínguez-Gil González, Isabel Fernández-Natal, Gregoria Megías Lobón, Juan Luis Muñoz Bellido, Pilar Ciruela, Ariadna Mas i Casals, Maria Doladé Botías, M Angeles Marcos Maeso, Dúnia Pérez del Campo, Antonio Félix de Castro, Ramón Limón Ramírez, Maria Francisca Elías Retamosa, Manuela Rubio González, María Sinda Blanco Lobeiras, Alberto Fuentes Losada, Antonio Aguilera, German Bou, Yolanda Caro, Noemí Marauri, Luis Miguel Soria Blanco, Isabel del Cura González, Montserrat Hernández Pascual, Roberto Alonso Fernández, Paloma Merino-Amador, Natalia Cabrera Castro, Aurora Tomás Lizcano, Cristóbal Ramírez Almagro, Manuel Segovia Hernández, Nieves Ascunce Elizaga, María Ederra Sanz, Carmen Ezpeleta Baquedano, Ana Bustinduy Bascaran, Susana Iglesias Tamayo, Luis Elorduy Otazua, Rebeca Benarroch Benarroch, Jesús Lopera Flores, and Antonia Vázquez de la Villa.

XSL•FO
RenderX

Disclaimer

This article presents independent research. The views expressed are those of the authors and not necessarily those of the Instituto de Salud Carlos III.

Data Availability

The datasets used in this article are not readily available to preserve the confidentiality of the participants but the ENE-COVID study has established a procedure for data request, with a Scientific Board that evaluates these petitions and guarantees the safeguard of participants' rights under the limits imposed by the ethical committee. Requests should be addressed to direccion-enecovid@isciii.es.

Authors' Contributions

BP-G, AL, and MP developed the original idea. MAC, NFL, RP-B, BP-G, and MP designed the analysis plan. MAC, BP-G, NFL, RP-B, and PF-N verified the underlying data and were in charge of statistical analyses and table and figure design. Other authors included in the ENE-COVID group contributed to data acquisition, laboratory analyses, and quality control at their respective regions and/or at the national level. MAC wrote the initial draft. All authors contributed to interpreting the results and critically revising the draft, agreed on the final version, and agreed to be accountable for the work.
MAC and NFL, MP and BP-G contributed equally in their corresponding roles

Conflicts of Interest

None declared.

Multimedia Appendix 1
Tables S1-S4
[PDF File, 918 KB - publichealth_v11i1e60658_app1.pdf]

References

1. Global Influenza Strategy 2019-2030.: World Health Organization; 2019 URL: https://iris.who.int/handle/10665/311184 [accessed 2025-06-17]
2. Vaccines against influenza WHO position paper – November 2012. Wkly Epidemiol Rec 2012 Nov 23;87(47):461-476. [Medline: 23210147]
3. Jorgensen P, Mereckiene J, Cotter S, Johansen K, Tsolova S, Brown C. How close are countries of the WHO European Region to achieving the goal of vaccinating 75% of key risk groups against influenza? Results from national surveys on seasonal influenza vaccination programmes, 2008/2009 to 2014/2015. Vaccine (Auckl) 2018 Jan 25;36(4):442-452. [doi: 10.1016/j.vaccine.2017.12.019] [Medline: 29287683]
4. Healthy people 2030. Increase the proportion of people who get the flu vaccine every year — IID-09. US Department of Health and Human Services. Office of Disease Prevention and Health Promotion. 2021. URL: https://health.gov/healthypeople/objectives-and-data/browse-objectives/vaccination/increase-proportion-people-who-get-flu-vaccine-every-year-iid-09 [accessed 2021-04-06]
5. Seasonal Influenza Vaccination and Antiviral Use in EU/EEA Member States – Overview of Vaccine Recommendations for 2017–2018 and Vaccination Coverage Rates for 2015–2016 and 2016–2017 Influenza Seasons.: European Centre for Disease Prevention and Control; 2018. [doi: 10.2900/721517]
6. Endrich MM, Blank PR, Szucs TD. Influenza vaccination uptake and socioeconomic determinants in 11 European countries. Vaccine (Auckl) 2009 Jun;27(30):4018-4024. [doi: 10.1016/j.vaccine.2009.04.029]
7. Yeung MPS, Lam FLY, Coker R. Factors associated with the uptake of seasonal influenza vaccination in adults: a systematic review. J Public Health 2016 Dec 2;38(4):746-753. [doi: 10.1093/pubmed/fdv194]
8. Kan T, Zhang J. Factors influencing seasonal influenza vaccination behaviour among elderly people: a systematic review. Public Health 2018 Mar;156:67-78. [doi: 10.1016/j.puhe.2017.12.007] [Medline: 29408191]
9. WHO SAGE seasonal influenza vaccination recommendations during the COVID-19 pandemic. Interim guidance. World Health Organization. 2020. URL: https://www.who.int/publications/m/item/who-sage-seasonal-influenza-vaccination-recommendations-during-the-covid-19-pandemic [accessed 2020-03-01]
10. Prada-García C, Fernández-Espinilla V, Hernán-García C, et al. Attitudes, perceptions and practices of influenza vaccination in the adult population: results of a cross-sectional survey in Spain. Int J Environ Res Public Health 2022 Sep 5;19(17):11139. [doi: 10.3390/ijerph191711139] [Medline: 36078854]
11. Prada-García C, Toquero-Asensio M, Fernández-Espinilla V, et al. The impact of the COVID-19 pandemic on influenza vaccination attitudes and actions in Spain's adult population. Vaccines (Basel) 2023 Sep 23;11(10):1514. [doi: 10.3390/vaccines11101514] [Medline: 37896918]

XSL•FO
RenderX

12. Domnich A, Grassi R, Fallani E, et al. Changes in attitudes and beliefs concerning vaccination and influenza vaccines between the first and second COVID-19 pandemic waves: a longitudinal study. Vaccines (Basel) 2021 Sep 13;9(9):1016. [doi: 10.3390/vaccines9091016] [Medline: 34579253]
13. Sulis G, Basta NE, Wolfson C, et al. Influenza vaccination uptake among Canadian adults before and during the COVID-19 pandemic: an analysis of the Canadian Longitudinal study on Aging (CLSA). Vaccine (Auckl) 2022 Jan 24;40(3):503-511. [doi: 10.1016/j.vaccine.2021.11.088] [Medline: 34916102]
14. Domnich A, Cambiaggi M, Vasco A, et al. Attitudes and beliefs on influenza vaccination during the COVID-19 pandemic: results from a representative Italian survey. Vaccines (Basel) 2020 Nov 30;8(4):33266212. [doi: 10.3390/vaccines8040711] [Medline: 33266212]
15. Portal estadístico. Área de inteligencia de gestión. Ministerio de Sanidad. 2021. URL: https://pestadistico.inteligenciadegestion.sanidad.gob.es/publicoSNS/I/sivamin/informe-de-evolucion-de-coberturas-de-vacunacion-por-vacuna [accessed 2025-06-12]
16. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet 2020 Aug 22;396(10250):535-544. [doi: 10.1016/S0140-6736(20)31483-5] [Medline: 32645347]
17. Pastor-Barriuso R, Pérez-Gómez B, Oteo-Iglesias J, et al. Design and implementation of a nationwide population-based longitudinal survey of SARS-CoV-2 infection in Spain: the ENE-COVID study. Am J Public Health 2023 May;113(5):525-532. [doi: 10.2105/AJPH.2022.307167] [Medline: 36893375]
18. Greenland S. Model-based estimation of relative risks and other epidemiologic measures in studies of common outcomes and in case-control studies. Am J Epidemiol 2004 Aug 15;160(4):301-305. [doi: 10.1093/aje/kwh221] [Medline: 15286014]
19. Olsen SJ, Azziz-Baumgartner E, Budd AP, et al. Decreased influenza activity during the COVID-19 pandemic-United States, Australia, Chile, and South Africa, 2020. Am J Transplant 2020 Dec;20(12):3681-3685. [doi: 10.1111/ajt.16381] [Medline: 33264506]
20. Informe anual de vigilancia de infección respiratoria aguda en España. Temporada 2020-2021. SIVIRA. : Instituto de Salud Carlos III; 2021 URL: https://cne.isciii.es/documents/d/cne/informe_vigilancia_ira_4020-3921_16nov2021-pdf [accessed 2025-06-12]
21. Chotpitayasunondh T, Fischer TK, Heraud JM, et al. Influenza and COVID-19: what does co-existence mean? Influenza Other Respir Viruses 2021 May;15(3):407-412. [doi: 10.1111/irv.12824] [Medline: 33128444]
22. Iacobucci G. Covid-19: risk of death more than doubled in people who also had flu, English data show. BMJ 2020 Sep 23;370:m3720. [doi: 10.1136/bmj.m3720] [Medline: 32967850]
23. Sherman SM, Sim J, Amlôt R, et al. Intention to have the seasonal influenza vaccination during the COVID-19 pandemic among eligible adults in the UK: a cross-sectional survey. BMJ Open 2021 Jul 13;11(7):e049369. [doi: 10.1136/bmjopen-2021-049369] [Medline: 34257095]
24. Bachtiger P, Adamson A, Chow JJ, Sisodia R, Quint JK, Peters NS. The impact of the COVID-19 pandemic on the uptake of influenza vaccine: UK-wide observational study. JMIR Public Health Surveill 2021 Apr 14;7(4):e26734. [doi: 10.2196/26734] [Medline: 33651708]
25. Sarriá-Santamera A, Rofriguez-Cabrera F, González-Soriano P, Pinilla-Navas L, Corral-Canto T. Determinantes de la vacunación de la gripe en personas de más de 65 años: 20 años no es nada [Article in Spanish]. Rev Esp Salud Pública 2019:e201905022 [FREE Full text]
26. Bianchi FP, Stefanizzi P, Di Lorenzo A, et al. Attitudes toward influenza vaccination in healthcare workers in Italy: a systematic review and meta-analysis. Hum Vaccin Immunother 2023 Dec 15;19(3):2265587. [doi: 10.1080/21645515.2023.2265587] [Medline: 37849235]
27. Portal estadístico (SIVAMIN). Ministerio de Sanidad. 2024. URL: https://pestadistico.inteligenciadegestion.sanidad.gob.es/publicoSNS/I/sivamin/sivamin [accessed 2024-02-01]
28. Fuster-Casanovas A, Vidal-Alaball J, Bonet-Esteve A, Muñoz-Miralles R, Catalina QM. Acceptance or rejection of vaccination against influenza and SARS-CoV2 viruses among primary care professionals in Central Catalonia. A cross-sectional study. Vaccine X 2023 Aug;14:100290. [doi: 10.1016/j.jvacx.2023.100290] [Medline: 37008959]
29. Bonet-Esteve A, Muñoz-Miralles R, Gonzalez-Claramunt C, Rufas AM, Cruz XP, Vidal-Alaball J. Influenza vaccination during the coronavirus pandemic: intention to vaccinate among the at-risk population in the Central Catalonia Health Region (VAGCOVID). BMC Fam Pract 2021 May 1;22(1):84. [doi: 10.1186/s12875-021-01434-8] [Medline: 33931039]
30. Sturm L, Kasting ML, Head KJ, Hartsock JA, Zimet GD. Influenza vaccination in the time of COVID-19: a national U.S. survey of adults. Vaccine (Auckl) 2021 Apr 1;39(14):1921-1928. [doi: 10.1016/j.vaccine.2021.03.003] [Medline: 33715898]
31. Nagata JM, Hernández-Ramos I, Kurup AS, Albrecht D, Vivas-Torrealba C, Franco-Paredes C. Social determinants of health and seasonal influenza vaccination in adults ≥65 years: a systematic review of qualitative and quantitative data. BMC Public Health 2013 Apr 25;13(1):388. [doi: 10.1186/1471-2458-13-388] [Medline: 23617788]
32. Tsang TK, Lau LLH, Cauchemez S, Cowling BJ. Household transmission of influenza virus. Trends Microbiol 2016 Feb;24(2):123-133. [doi: 10.1016/j.tim.2015.10.012] [Medline: 26612500]

XSL•FO
RenderX

33. Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE. Household transmission of SARS-CoV-2: a systematic review and meta-analysis. JAMA Netw Open 2020 Dec 1;3(12):e2031756. [doi: 10.1001/jamanetworkopen.2020.31756] [Medline: 33315116]
34. Kong G, Lim NA, Chin YH, Ng YPM, Amin Z. Effect of COVID-19 pandemic on influenza vaccination intention: a meta-analysis and systematic review. Vaccines (Basel) 2022 Apr 13;10(4):606. [doi: 10.3390/vaccines10040606] [Medline: 35455354]
35. Abellán-García A, Aceituno-Nieto P, Fernández-Morales I, Ramiro-Fariñas D, Pujol-Rodríguez R. Una estimación de la población que vive en residencias de mayores. : CSIC - Instituto de Economía, Geografía y Demografía (IEGD); 2020 URL: https://digital.csic.es/handle/10261/240875 [accessed 2025-06-17]

Abbreviations

ENE-COVID: Seroepidemiological Survey of SARS-CoV-2 Infection in Spain
HCW: health care worker
SPR: standardized prevalence ratio

Edited by A Mavragani; submitted 17.05.24; peer-reviewed by A Nakhostin-Ansari, A Sljivo; revised version received 05.12.24; accepted 05.12.24; published 01.07.25.

Please cite as:
de la Cámara MÁ, Fernández de Larrea-Baz N, Pastor-Barriuso R, Larrauri A, Fernández-Navarro P, Pollán M, Pérez-Gómez B, ENE-COVID Study Group
Influenza Vaccination Coverage and Determinants of New Vaccinations During the COVID-19 Pandemic in Spain (ENE-COVID): Nationwide Population-Based Study
JMIR Public Health Surveill 2025;11:e60658
URL: https://publichealth.jmir.org/2025/1/e60658
doi:10.2196/60658

© Miguel Ángel de la Cámara, Nerea Fernández de Larrea-Baz, Roberto Pastor-Barriuso, Amparo Larrauri, Pablo Fernández-Navarro, Marina Pollán, Beatriz Pérez-Gómez, ENE-COVID Study Group. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 1.7.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Meeting Global Health Needs via Infectious Disease Forecasting: Development of a Reliable Data-Driven Framework

Ravikiran Keshavamurthy[1,2], BVSc, MVSc, PhD; Karl T Pazdernik[1,3], PhD; Colby Ham[1], BS; Samuel Dixon[1], MS; Samantha Erwin[1], MS, PhD; Lauren E Charles[1,2], DVM, MS, PhD

[1]Artificial Intelligence and Data Analytics Division, Pacific Northwest National Laboratory, 902 Battelle Boulevard, Richland, WA, United States

[2]Paul G. Allen School for Global Health, Washington State University, Pullman, WA, United States

[3]Department of Statistics, North Carolina State University, Raleigh, NC, United States

Corresponding Author:
Lauren E Charles, DVM, MS, PhD
Artificial Intelligence and Data Analytics Division, Pacific Northwest National Laboratory, 902 Battelle Boulevard, Richland, WA, United States

Abstract

Background: Infectious diseases (IDs) have a significant detrimental impact on global health. Timely and accurate ID forecasting can result in more informed implementation of control measures and prevention policies.

Objective: To meet the operational decision-making needs of real-world circumstances, we aimed to build a standardized, reliable, and trustworthy ID forecasting pipeline and visualization dashboard that is generalizable across a wide range of modeling techniques, IDs, and global locations.

Methods: We forecasted 6 diverse, zoonotic diseases (brucellosis, campylobacteriosis, Middle East respiratory syndrome, Q fever, tick-borne encephalitis, and tularemia) across 4 continents and 8 countries. We included a wide range of statistical, machine learning, and deep learning models (n=9) and trained them on a multitude of features (average n=2326) within the One Health landscape, including demography, landscape, climate, and socioeconomic factors. The pipeline and dashboard were created in consideration of crucial operational metrics—prediction accuracy, computational efficiency, spatiotemporal generalizability, uncertainty quantification, and interpretability—which are essential to strategic data-driven decisions.

Results: While no single best model was suitable for all disease, region, and country combinations, our ensemble technique selects the best-performing model for each given scenario to achieve the closest prediction. For new or emerging diseases in a region, the ensemble model can predict how the disease may behave in the new region using a pretrained model from a similar region with a history of that disease. The data visualization dashboard provides a clean interface of important analytical metrics, such as ID temporal patterns, forecasts, prediction uncertainties, and model feature importance across all geographic locations and disease combinations.

Conclusions: As the need for real-time, operational ID forecasting capabilities increases, this standardized and automated platform for data collection, analysis, and reporting is a major step forward in enabling evidence-based public health decisions and policies for the prevention and mitigation of future ID outbreaks.

(JMIR Public Health Surveill 2025;11:e59971) doi:10.2196/59971

KEYWORDS

disease forecasting; machine learning; deep learning; epidemiology; One Health; decision-making; data visualization

Introduction

The frequency and magnitude of infectious disease (ID) events have seen a drastic incline in the past few decades, mainly attributed to climate change, urbanization, and globalization [1]. These events often involve the emergence of novel infectious agents or the re-emergence of a previously known pathogen. Timely and accurate prediction of such ID events is crucial for decision makers to decrease associated mortality, morbidity, and economic losses [2]. However, the complex and unpredictable nature of pathogen ecology makes forecasting the spatiotemporal dynamics of IDs a challenging task [3].

Digital information related to ID events is being generated and shared faster than ever before. Additionally, information associated with disease occurrence, such as meteorology, socioeconomics, demographics, land use, agriculture, social media, and internet trends, is often readily available [4]. To keep up with this unprecedented amount of data being generated, machine learning (ML)– and deep learning (DL)–based algorithms are being adopted by the research community. These methods have shown to be better at detecting cryptic patterns

XSL•FO
RenderX

arising from interactions between multiple features, which are difficult, often impossible at times, to uncover with conventional prediction methods [5].

The time-series models have been previously used for forecasting ID events including brucellosis [6], Middle East respiratory syndrome [7], campylobacteriosis [8], and Q fever [5]. However, over the last decade, especially after the COVID-19 pandemic outbreak, considerable progress has been made in the field of ID surveillance, as the disease diagnosis, reporting, and intelligence-sharing infrastructure continue to grow on a global scale [9]. Dixion et al [5] compared the performance of various forecasting approaches across several diseases and countries and found that tree-based techniques had better predictive performance compared to statistical and DL techniques. Integrating diverse large-scale epidemiological data from multiple sources has further enhanced the accuracy and utility of disease prediction models [10]. Ensemble forecasting techniques, which merge predictions from multiple models, offer notable advantages over single-model approaches [11]. For example, Reich et al [12] used real-time multimodel ensembles for seasonal influenza in the United States, while Ma et al [13] demonstrated the effectiveness of integrating internet search data along with ensemble forecasting techniques to jointly forecast COVID-19 and influenza-like illnesses, underscoring the value of diverse data sources in public health forecasting.

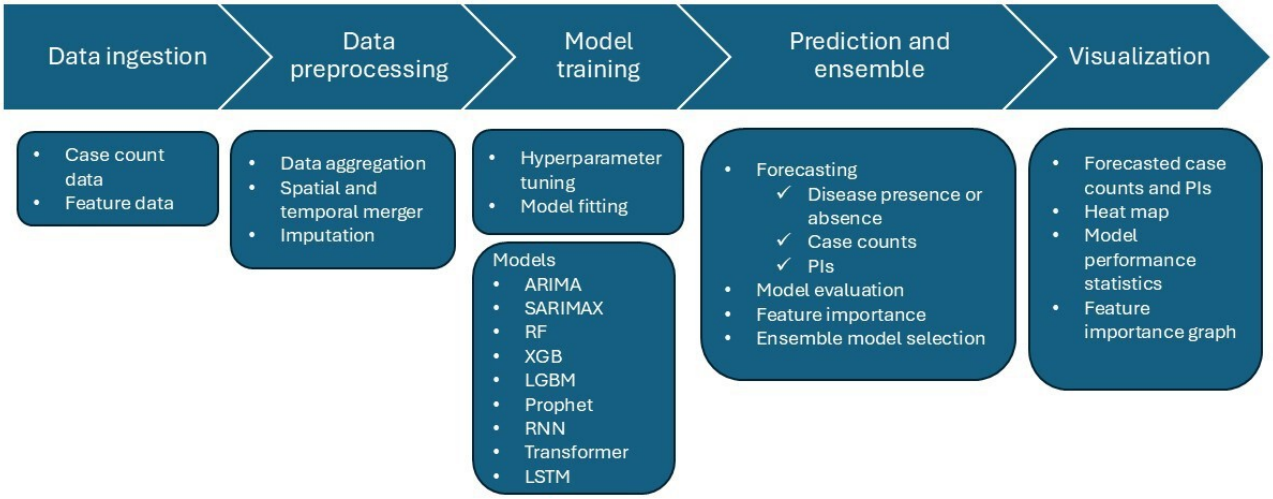
Despite several advantages, ML and DL techniques do come with a number of limitations, making them less desirable in the real-world operational setting. We conducted a scoping systematic review to determine the status and advances made in the field of ID prediction using ML and DL, which were published elsewhere [9]. One of our focuses in this systematic review was to identify if and how researchers were incorporating the necessary methods required to eventually enable the operational deployment of their ID forecasting models. Through this work, we identified multiple shortcomings in published modeling techniques that could hinder their ability to support operational biopreparedness and decision-making. Even though most of the studies focused on increasing the prediction accuracy of their models, which is critical for operational deployment, they did not address other important decision-making metrics such as computational efficiency, uncertainty quantification, interpretability, and generalizability. Modern ML and DL techniques are performance-driven, that is, they aim to generate better predictive or classification accuracy by minimizing errors [14]. As a result, other metrics that are crucial for operational decision-making are often overlooked by the scientific community. First, as the structural and functional complexities of the forecasting models evolve, the computational resources required to train such models also grow considerably. In addition, these complexities result in an expanded number of model hyperparameters that can be individually tuned using numerous possible techniques. This ever-increasing complexity gives rise to an even larger problem space with highly variable results. Second, since most of these techniques are nonparametric in nature, model uncertainties are not inherently estimated. Hence, model uncertainties are often overlooked even though these estimates are crucial for systematic and transparent decision-making [15]. Third, ML and DL models are considered black boxes, meaning that their internal logic and inner workings are mostly hidden from the end user [9]. Consequently, verifying and understanding the rationale behind the model forecasts are difficult and often neglected. With the above considerations in mind, we developed an ID forecasting pipeline that holistically focuses on these often neglected yet essential performance metrics required for operational decision support. By incorporating this pipeline into an interactive dashboard, we created a capability that allows users to easily visualize the ID forecasting results for use in operational planning and decision-making.

Methods

Overview

The graphical flowchart illustrating our forecasting pipeline including data ingest, preprocessing, model training, prediction, and visualization is presented in Figure 1.

Figure 1. Graphical flowchart illustrating the infectious disease forecasting pipeline. ARIMA: autoregressive integrated moving average; LGBM: light gradient boosting machine; LSTM: long short-term memory; PI: prediction interval; RF: random forest; RNN: recurrent neural network; SARIMAX: seasonal autoregressive integrated moving average with exogenous factors; XGB: extreme gradient boosting.



XSL•FO
RenderX

Case Count Data

The model outcome variable was case counts collected from EpiArchive [16] and Food and Agriculture Organization's Emergency Prevention System-I [17]. We included the most consistent, available, and interpretable disease case counts from these sources. These data also included locations that reported only 0 disease cases. In instances where only the disease presence was reported without any information about the exact number of cases, the region, or the date of occurrence, such disease-location combinations were excluded from the analysis. We included all the diseases that spanned across multiple countries or regions and had a consistent reporting of disease events for our study period. Data were collected at the regional-level resolution for each country from January 2014 to December 2018, which contained case counts at either the weekly or monthly temporal resolution.

Feature Data

We collected feature data from a variety of open sources across the One Health landscape (Table 1). Broadly, the features included historic case counts (case count lags up to 6 months), country census data, socioeconomic factors, health data, agricultural trade data, landscape, and climate records. The feature information varied in spatial and temporal resolutions; data were typically available at monthly or yearly resolutions and covered either the country or regional level. A complete list of countries, regions, and the total number of features collected for each region is presented in Multimedia Appendix 1.

Table . The input feature data types, source, feature description by name, geographic location, geographic resolution, time period, and periodicity. The table was originally published by Dixon et al [5].

Data type	Individual features	Geographic location	Geographic resolution	Time period	Periodicity
Case counts [16]	Incidences of select human diseases	Countries of interest	Region level	2009-2018	Daily
Political borders [18]	Geopolitical borders (country and within the country)	Countries of interest	Region level	2018	Single instance
Climate [19,20]	Air temperature, humidity, precipitation, soil moisture, and wind speed	Global	Gridded 0.25 °×0.25 °, 1 °×1 °	2012 - 2018	Monthly
Gross domestic product [21]	Gross domestic product	Global	Country level	Varies	Yearly
Elevation [22]	Digital elevation map	Global	43,200×17,200 (30 arc seconds)	N/A[a]	N/A
Mortality [23]	Deaths by country, year, sex, age group, and cause of death	Global	Country level	2009 - 2018	Yearly
Municipal waste [24]	Municipal waste generation and treatment	Countries of interest	Country level	2009 - 2017	Yearly
Sociopolitical and physical data [25]	Socioeconomic and political attributes	Global	Varies by country; 1: 10-110 m	2019	Single instance
Population [26]	Population by age intervals by location	Global	Country level	2009 - 2015	Every 5 years
Population density [27]	Population density	Global	30 arc seconds	2009 - 2015	Every 5 years
Water potability and treatment [24]	Freshwater resources, available water, wastewater treatment plant capacity, surface water	Countries of interest	Country level	2009 - 2017	Yearly

[a]N/A: not applicable.

Data Preparation and Imputation

The harvested raw data contained information at varied spatial and temporal resolutions in different file formats. These raw data were harmonized to create the final datasets containing monthly case counts and input features for each region within the country. In situations where the feature data at the regional level were missing, the features aggregated by an average at the national level were used. Any features with more than 20% missing values across all regions and dates were not included as a predictor variable. The remaining missing values were imputed temporally using spline and forward-fill techniques for monthly and yearly feature data, respectively. Some features could not easily be imputed temporally through these techniques but were still sufficiently prevalent across regions to include in the models. In these instances, the majority of the missing data (less than 20% overall) came from only a few select regions. For these special cases, we performed geoimputation using the

XSL•FO
RenderX

k-nearest neighbor method and restricted input to only regional feature data within the same country. We used distance-based weights for the features with 5 neighboring samples for imputation. The imputation techniques were implemented using the "pandas" and "skforecast" libraries in Python (Python Software Foundation) [26].

Model Choice

Overview

In our analysis, we included the following models discussed in detail below: statistical time-series models (ie, autoregressive integrated moving average [ARIMA], seasonal autoregressive integrated moving average with exogenous factors [SARIMAX], and Prophet), tree-based models (ie, random forest [RF], extreme gradient boosting [XGB], and light gradient boosting machine [LBGM]), and DL models (recurrent neural network [RNN], long short-term memory [LSTM], and transformer).

Statistical Time-Series Models

The ARIMA model is a traditional statistical technique used in time-series forecasting. These models use a linear, regression-type equation, in which the predictors are lags of the dependent variable or lags of the forecast errors. The SARIMAX is an extension of ARIMA that takes seasonal trends into account within time-series data. Because SARIMAX can also accommodate exogenous features, it is often preferable to traditional ARIMA. However, SARIMAX models may fail to converge if those exogenous features are highly correlated. To overcome this issue, the least absolute shrinkage and selection operator (LASSO) regression was used for feature selection and dimensionality reduction. These LASSO-selected features were later used as input variables for SARIMAX. The modeling techniques were implemented using "statsmodels" and "sklearn" libraries in Python [28,29].

Prophet is a decomposable time-series model that uses 3 main components to make its forecasts, namely, trend, seasonality, and holidays [30]. The trend component models the nonperiodic changes in the time series. Seasonality models the periodic, or seasonal, changes in the time series; this can be an important part of a disease forecasting model, as many diseases have seasonal trends. The Prophet model was implemented using the "darts" library in Python [31].

Tree-Based ML Models

Tree-based models make predictions by creating decision trees that divide the feature space using a series of binary decision thresholds. The RF is an extension of the decision tree method that uses an ensemble of decision trees to increase performance and reduce the risk of overfitting the training data. The XGB is an implementation of a decision tree that uses stochastic gradient boosting to sequentially improve prediction during training. Because of this methodology, XGB achieves superior accuracy in prediction tasks. LGBM is similar to XGB; however, it also applies a novel sampling method and feature bundling process that allows the method to get comparable performance to XGB but uses far less memory and trains much faster. These modeling techniques were implemented using "skforecast" library in Python [32].

DL Models

The RNN is built on DL architecture with a hidden state that memorizes sequential data. This added process provides the model with an understanding of temporal information in a time-series sequence. The LSTM is an extension of RNN with additional parameters that incorporate different scales of temporal information, which allow the model to effectively use both long- and short-term temporal data in its prediction. Hence, these models are routinely used in disease forecasting tasks.

Transformers [33] are another form of DL architecture that has been recently outperforming other models in many applications of ML. These models leverage a structure known as "attention," which allows them to model dependence in both short and long temporal lags. Unfortunately, these models are especially slow to train and often require very large amounts of data and computational power. The DL techniques were implemented as an encoder-record architecture using "darts" library in Python.

Model Training and Prediction

The model training involved hyperparameter tuning and model fitting, whereas model testing involved generating 1-step ahead forecasts (ie, monthly) and their prediction intervals (PIs). When creating predictions, all feature data and disease case counts up until the time of prediction were included in the model. The model training included hyperparameter tuning using a 5-fold time-series cross-validation split and model fitting. First, the models were trained using 2014 - 2016 data and tested on 2017 data. Subsequently, the model was retrained using 2014 - 2017 data and tested on 2018 data in a sliding window manner. A detailed description of model-specific hyperparameters and their search space used for model training is presented in Multimedia Appendix 1.

Model Evaluation

The forecasting models were evaluated in two ways: (1) F_1-score to evaluate the ability of a model to predict the presence or absence of a disease in each region (ie, did the region report 1 or more cases during the testing period) and (2) mean absolute error (MAE) for the subset of predictions where the disease was detected in a region (case counts greater than 0). The multimetric approach was adopted because of the high frequency of time-location pairs that had no disease present. Without separating these results, the MAE metric was largely driven by the results of the high-frequency disease and regions. Finally, we created ensemble models by selecting the best-performing technique based on the MAE of the testing dataset. Other metrics such as precision, recall, negative predictive value, true negatives, false positives, false negatives, and true positives were also estimated but not considered when building our ensemble model.

Uncertainty Quantification and Interpretability

The 95% PIs for our forecasted values were estimated to account for forecasting uncertainties. For statistical models (ARIMA, SARIMAX, and Prophet), PIs were readily obtained along with the model forecasts as a natural consequence of the model construction. However, PIs were not directly available for ML and DL techniques. Hence, we used Python packages that

XSL•FO
RenderX

retroactively produced PIs using alternate techniques. The PIs for tree-based models (RF, XGB, and LGBM) were estimated by a bootstrapping technique using "skforecast" library, whereas for DL models (RNN, LSTM, and transformer), PIs were calculated by a nonparametric method known as quantile regression using the "Darts" library [31,32]. For each model forecast, we estimated the coverage probability of the 95% PIs (ie, instances where the PIs surrounded the true value) as a measure of calibration assessment for the prediction uncertainties.

Feature importance, based on Shapley Additive Explanation (SHAP) values for the best-performing models, was used as a generalized approach to interpretability for tree-based and DL models [34]. SARIMAX was the only statistical model that allowed for feature input and interpretability, which was provided by coefficient estimates from the model. However, the SARIMAX model was unable to incorporate the large number of features available for each region. To address this limitation, a feature selection approach was applied using LASSO regression to reduce the model's input variables.

The historic case count data are important and most frequently used predictive features in the ID prediction domain [9]. In our analysis, we used case count lags up to 6 months. To estimate the relative importance of these case count lags compared to other features, we calculated 2 metrics, namely, the feature ratio and mean reciprocal rank (MRR). A threshold of the top 10 features, as deemed by the models, was set to estimate the feature ratio and MRR. The feature ratio was the total number of lag case count features divided by the total number of features (up to the top 10 features). The MRR, on the other hand, also considers the position of the first relevant item in the ranked list, that is, case count lags. The MRR was defined as the mean of reciprocal ranks of case count lags across all features. The values range from 0 to 1, with a higher score signifying the greater relative feature importance rank of case count lags compared to other features for obtaining model forecasts.

Spatial Generalizability Using Transfer Learning

We used the transfer learning framework to test the models for their ability to produce accurate forecasts in regions that were fairly new to the disease. These regions were strategically picked by maximizing the global coverage and following the stratified random sampling framework. First, only the country-disease combinations where case count data were available were selected. This process also included the regions that reported 0 disease cases during our study period. Next, one region within the selected countries was randomly picked and denoted as the target region. Then, we selected a similar region to the target region to train the transfer learning models based on (1) the target and similar region being part of the same country and (2) the target and similar region containing comparable case counts over the last 6 months. The similar region pretrained model was then used to forecast the disease case counts of the target region, and performance was assessed.

Ethical Considerations

Data used in this study are open source and do not identify individual information, either directly or indirectly. Therefore, this research was exempted from ethical review.

Results

Overview

The summary statistics of the disease-location combinations aggregated at the country level are presented in Table 2, whereas a detailed breakdown of this summary for both test and train split is shown in Multimedia Appendix 1.

Table . Median (IQR) regional case counts per country and disease for the entire dataset from January 2014 to December 2018.

Disease	Country							
	Australia, median (IQR)	Germany, median (IQR)	Israel, median (IQR)	Japan, median (IQR)	Norway, median (IQR)	Saudi Arabia, median (IQR)	Sweden, median (IQR)	United States, median (IQR)
Brucellosis	3 (0-29)	11 (0-42)	30 (0-1072)	0 (0-3)	0 (0-2)	—[a]	1 (0-2)	4 (0-25)
Campylobacteriosis	7904 (0-43,205)	148,616 (50,644-197,616)	998 (0-7666)	—	306 (0-1456)	—	1103 (0-10,815)	1159 (0-25,327)
MERS[b]	0 (0-0)	—	—	0 (0-0)	—	3 (0-71)	—	—
Q fever	55 (0-974)	28 (0-769)	46 (0-103)	0 (0-1)	—	—	—	5 (0-26)
Tick-borne encephalitis	—	15 (0-897)	—	0 (0-2)	—	—	—	—
Tularemia	—	—	0 (0-0)	0 (0-1)	15 (0-69)	—	—	4 (0-89)

[a]Not available.

[b]MERS: Middle East respiratory syndrome.

There were 757 disease-region combinations ran through 9 different models in this study. As the first step, we calculated F_1-scores to determine the ability of the forecasting models to detect the presence or absence of a disease in a given region (Table 3). Overall, tree-based models (ie, XGB, RF, and LGBM) and Prophet consistently had better and comparable F_1-scores across the diseases. The DL models (ie, RNN, transformer, and LSTM) had significantly lower F_1-scores. Patterns were consistent across all diseases in each region. The additional evaluation metrics for model and disease combination including

XSL•FO
RenderX

precision, recall, negative predictive value, true negatives, false positives, false negatives, and true positives are presented in Multimedia Appendix 1.

Table . F_1-score for all models forecasting the presence or absence of disease for each region in 2017 and 2018.

Disease	Model								
	ARIMA[a]	LGBM[b]	LSTM[c]	Prophet	RF[d]	RNN[e]	SARIMAX[f]	Transformer	XGB[g]
Brucellosis	0.65	0.67	0.55	0.69	0.67	0.56	0.66	0.59	0.67
Campylobacteriosis	0.93	0.96	0.96	0.97	0.96	0.95	0.91	0.94	0.96
MERS[h]	0.98	0.90	0.60	0.90	0.90	0.61	0.95	0.56	0.90
Q fever	0.78	0.78	0.64	0.79	0.78	0.67	0.74	0.65	0.78
Tick-borne encephalitis	0.68	0.98	0.59	0.97	0.98	0.55	0.95	0.54	0.98
Tularemia	0.54	0.73	0.49	0.71	0.73	0.50	0.72	0.55	0.73

[a]ARIMA: autoregressive integrated moving average.

[b]LGBM: light gradient boosting machine.

[c]LSTM: long short-term memory.

[d]RF: random forest.

[e]RNN: recurrent neural network.

[f]SARIMAX: seasonal autoregressive integrated moving average with exogenous factors.

[g]XGB: extreme gradient boosting.

[h]MERS: Middle East respiratory syndrome.

Next, we assessed the performance of each model for the subset of regions, where the total case counts were greater than 0 (ie, the disease was present) based on MAE. The model with the lowest MAE was inconsistent across diseases (Table 4).

Table . MAE[a] (95% CI) for each disease aggregated across all locations where the disease was present in 2017 and 2018.

Disease	Model								
	ARIMA[b], MAE (95% CI)	LGBM[c], MAE (95% CI)	LSTM[d], MAE (95% CI)	Prophet, MAE (95% CI)	RF[e], MAE (95% CI)	RNN[f], MAE (95% CI)	SARIMAX[g], MAE (95% CI)	Transformer, MAE (95% CI)	XGB[h], MAE (95% CI)
Brucellosis	0.8 (0.5 - 1.1)	0.8 (0.5 - 1.1)	0.84 (0.5 - 1.14)	0.9 (0.5 - 1.2)	0.8 (0.5 - 1.1)	0.9 (0.5 - 1.2)	1.1 (0.7 - 1.6)	0.8 (0.5 - 1.6)	0.8 (0.4 - 1.1)
Campylobacteriosis	237.9 (1.1 - 474.9)	43.8 (35.3 - 52.3)	45.2 (36.3 - 54.12)	43.4 (34.9 - 51.8)	39.8 (32.3 - 47.2)	46.6 (37.7 - 55.5)	243.7 (74.9 - 412.6)	50.9 (40.3 - 61.6)	43.1 (34.2 - 52.1)
Q fever	28.9 (25.4 - 83.1)	1.5 (0.9 - 2.1)	1.18 (0.8 - 1.6)	1.4 (0.9 - 1.9)	1.5 (0.8 - 2.1)	1.3 (0.9 - 1.8)	1.9 (1.2 - 2.7)	1.4 (0.9 - 2.0)	1.2 (0.9 - 1.5)
Tick-borne encephalitis	2.0 (0.6 - 3.3)	2.4 (0.7 - 4.1)	2.61 (0.7 - 4.53)	1.3 (0.5 - 2.2)	2.4 (0.6 - 4.1)	2.6 (0.7 - 4.6)	1.6 (0.6 - 2.6)	2.9 (0.7 - 5.2)	2.4 (0.6 - 4.1)
Tularemia	0.6 (0.4 - 0.7)	0.6 (0.5 - 0.7)	0.58 (0.5 - 0.71)	0.6 (0.5 - 0.7)	0.6 (0.5 - 0.7)	0.6 (0.4 - 0.7)	0.7 (0.5 - 0.9)	0.6 (0.5 - 0.7)	0.6 (0.5 - 0.7)

[a]MAE: mean absolute error.

[b]ARIMA: autoregressive integrated moving average.

[c]LGBM: light gradient boosting machine.

[d]LSTM: long short-term memory.

[e]RF: random forest.

[f]RNN: recurrent neural network.

[g]SARIMAX: seasonal autoregressive integrated moving average with exogenous factors.

[h]XGB: extreme gradient boosting.

The MAE of the best-performing model for each disease across all locations (with case counts greater than 0) is presented in Table 5, which we define as our “ensemble” model. Using this technique and when averaging across all regions, prediction

XSL•FO
RenderX

years, and diseases, we observed that the ensemble model had the lowest MAE compared to other models (Table 6). While the difference does not appear to be a large amount, the difference is considerable, and the magnitude is visually minimized by the large error in the ARIMA and SARIMAX models. All the models included in our study had similar MAE except for ARIMA and SARIMAX since they had a few outlier predictions that skewed their MAE.

Table . Ensemble (best performing) models for each disease averaged across all regions where the disease was present in 2017 and 2018.

Disease	Best model	MAE[a] (95% CI)
Brucellosis	LGBM[b]	0.8 (0.5 - 1.1)
Campylobacteriosis	RF[c]	39.8 (32.3 - 47.2)
Q fever	LSTM[d]	1.2 (0.8 - 1.6)
Tick-borne encephalitis	Prophet	1.4 (0.5 - 2.2)
Tularemia	ARIMA[e]	0.6 (0.4 - 0.7)

[a]MAE: mean absolute error.

[b]LGBM: light gradient boosting machine.

[c]RF: random forest.

[d]LSTM: long short-term memory.

[e]ARIMA: autoregressive integrated moving average.

Table . MAE[a] (95% CIs) for each model aggregated across all diseases and regions where the disease was present in 2017 and 2018.

Model	MAE (95% CI)
ARIMA[b]	115.2 (5.2 - 225.2)
Ensemble	8.7 (6.5 - 23.9)
LGBM[c]	20.8 (16.5 - 25.1)
LSTM[d]	21.4 (16.9 - 25.9)
Prophet	20.6 (16.3 - 24.9)
RF[e]	19.0 (15.2 - 22.8)
RNN[f]	22.1 (17.6 - 26.7)
SARIMAX[g]	113.1 (34.6 - 191.6)
Transformer	24.2 (18.8 - 29.5)
XGB[h]	20.5 (16.0 - 25.0)

[a]MAE: mean absolute error.

[b]ARIMA: autoregressive integrated moving average.

[c]LGBM: light gradient boosting machine.

[d]LSTM: long short-term memory.

[e]RF: random forest.

[f]RNN: recurrent neural network.

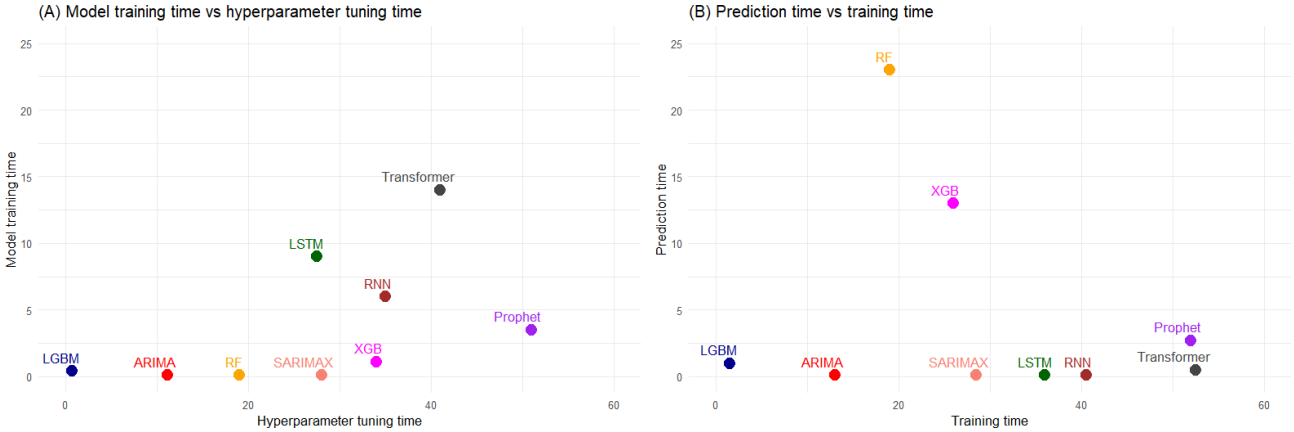
[g]SARIMAX: seasonal autoregressive integrated moving average with exogenous factors.

[h]XGB: extreme gradient boosting.

The average training time, which included hyperparameter tuning and model fitting, along with prediction time (ie, the time to generate 1-step ahead forecasts and their PIs) for each model averaged across all diseases, regions, and periods are presented in Figure 2. The LGBM followed by ARIMA were the most time-efficient models. Although RF and XGB had a training time comparable with other models, their average prediction time was considerably higher mainly because of the increased amount of time required by them to generate 95% PIs using the bootstrapping technique.

XSL•FO
RenderX

Figure 2. (A) Average hyperparameter tuning and model fitting time and (B) average training and prediction time for each model averaged across all diseases, locations, and periods. ARIMA: autoregressive integrated moving average; LGBM: light gradient boosting machine; LSTM: long short-term memory; RF: random forest; RNN: recurrent neural network; SARIMAX: seasonal autoregressive integrated moving average with exogenous factors; XGB: extreme gradient boosting.

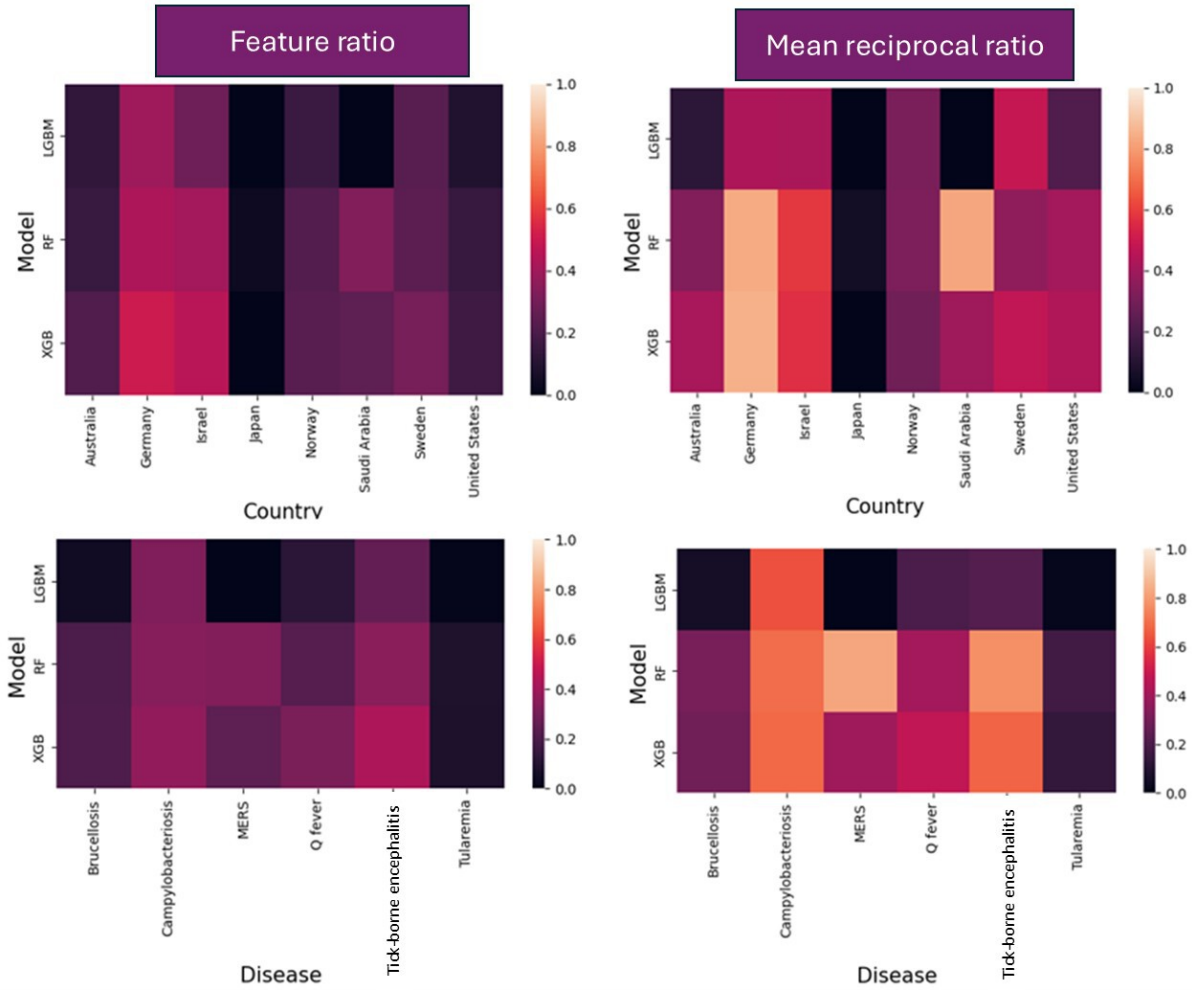


Feature Importance

For the tree-based models (ie, LGBM, RF, and XGB), we selected up to 10 top features based on SHAP values that the model deemed important and calculated feature ratio and MRR to show how lag case count may be impacting our predictive performance (Figure 3). Overall, the feature ratios were low (<0.4) when aggregated across countries and diseases, suggesting that non–lag case count features were more important for forecasting compared to case lag data. Campylobacteriosis and tick-borne encephalitis had higher MRR compared to other diseases, indicating that lag case count features were ranked relatively higher among the top 10 features for these diseases. However, the low feature ratio and MRR for many diseases highlight how the additional feature variables can be valuable in disease forecasting, especially for less prevalent diseases. A complete breakdown of feature ratios and MRRs for each country-disease combination is presented in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 3. Relative importance of case count lags to produce forecasts for tree-based models quantified as feature ratios and mean reciprocal ranks of the most important features (up to the top 10) aggregated across countries (top row) and diseases (bottom row). Higher scores signify the greater relative importance of case count lags compared to other features for obtaining model forecasts. LGBM: light gradient boosting machine; MERS: Middle East respiratory syndrome; RF: random forest; XGB: extreme gradient boosting.



The coverage probability of the forecasting models represented by the percent coverage of the true observations by 95% PIs is presented in Table 7. The LGBM had a 94.4% (n=14,707) coverage, indicating that the PIs were well calibrated, even better than traditional statistical models like ARIMA and SARIMAX. On the other hand, the DL models had less than 10% coverage, highlighting the fact that these models were not thoroughly calibrated. This observation was true when the coverage percentage was broken down across disease and country combinations in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Coverage percentage of the 95% prediction intervals of the forecasting models[a].

Model	Coverage, n (%)
ARIMA[b]	14,582 (93.6)
LGBM[c]	14,707 (94.4)
LSTM[d]	1277 (8.2)
Prophet	13,086 (84)
RF[e]	13,756 (88.3)
RNN[f]	1246 (8)
SARIMAX[g]	13,959 (89.6)
Transformer	1262 (8.1)
XGB[h]	13,491 (86.6)

[a]The ideal coverage percentage is 95%.

[b]ARIMA: autoregressive integrated moving average.

[c]LGBM: light gradient boosting machine.

[d]LSTM: long short-term memory.

[e]RF: random forest.

[f]RNN: recurrent neural network.

[g]SARIMAX: seasonal autoregressive integrated moving average with exogenous factors.

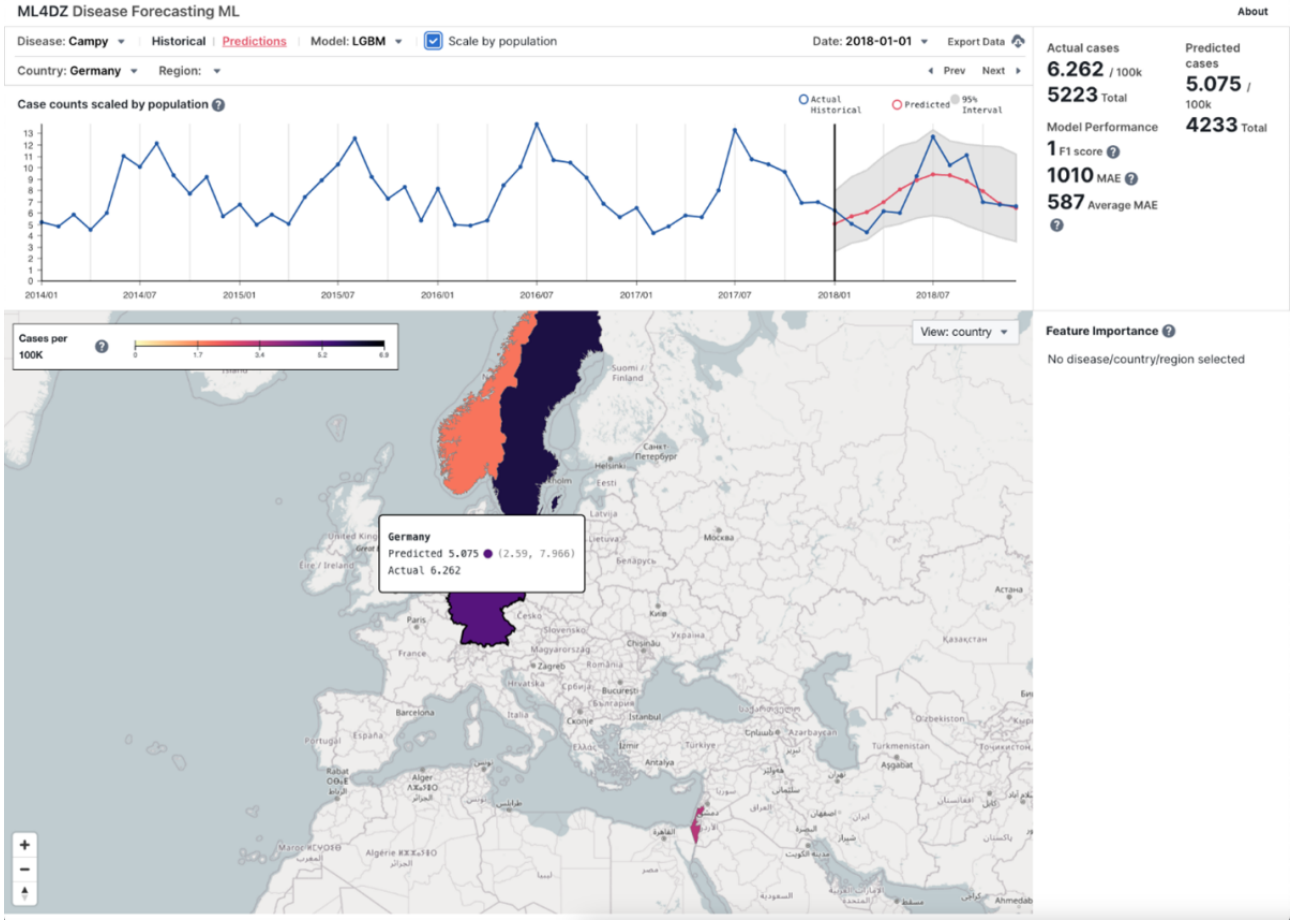
[h]XGB: extreme gradient boosting.

Results Visualization

We created an interactable dashboard to visualize the predictions made by forecasting models (Figure 4). The information presented in the user interface includes a time series of actual and predicted values with 95% PI, model performance metrics, and feature importance values for each disease and country or region. In Figure 5, campylobacteriosis in Schleswig-Holstein, Germany, for the year 2018 using the LGBM model is presented as an example, showing uncertainty quantification by PIs and feature importance as mean SHAP values. The lagging case count feature for the past 6 periods was the most important feature for the model, but variables related to global health spending and monetary export value of cattle also ranked high in feature importance.

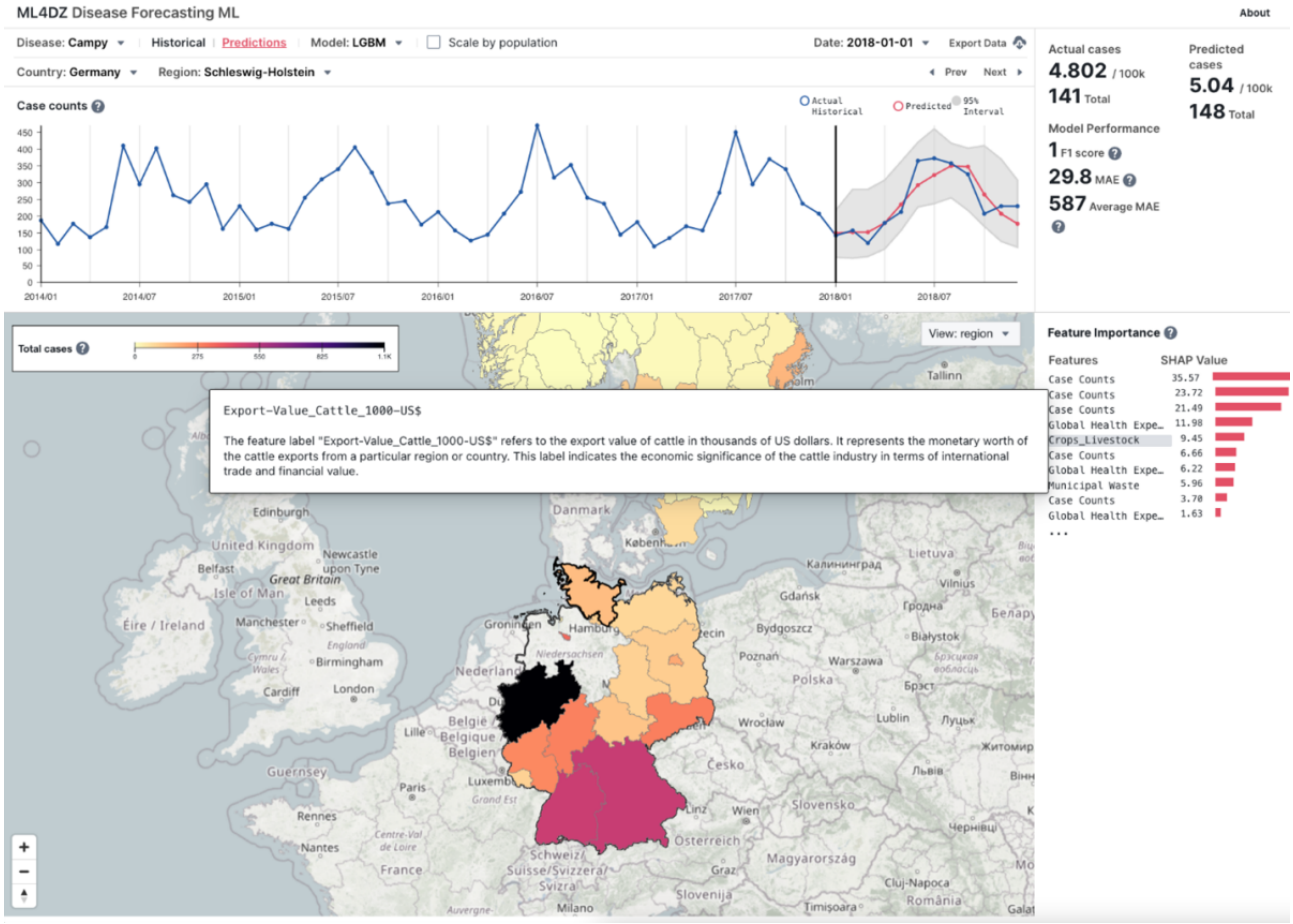
XSL•FO
RenderX

Figure 4. Dashboard in prediction view with values based on case counts scaled by population. LGBM: light gradient boosting machine; MAE: mean absolute error; ML: machine learning.



XSL•FO
RenderX

Figure 5. Dashboard with disease ("campy"=campylobacteriosis) and region (Schleswig-Holstein, Germany) selected in prediction view, which includes model performance statistics in summary and model feature importance list. Pop-out box provides a detailed explanation of the clicked and highlighted feature. LGBM: light gradient boosting machine; MAE: mean absolute error; ML: machine learning.



Spatial Generalizability Using Transfer Learning

The spatial generalizability for the forecasting models was tested based on the performance of the pretrained ensemble model from the location with a comparable disease pattern (ie, similar region) to forecast the disease case counts of the selected region (ie, target region). The model tested (ensemble or best model) varied by location and included statistical (ARIMA: n=3) and tree-based models (LGBM: n=2 and XGB: n=1). In most cases, the MAE of a similar region was close to and often lower than the MAE of the target region (Table 8). However, campylobacteriosis forecasting in Kalmar, Sweden, was an exception, where the model MAE increased from 5.01 to 16.98 when Västmanland feature data were used for training. It is important to note that these are pairwise comparisons of a single metric calculated across a single region and do not include confidence bounds to determine if any differences are statistically significant. However, given that over half of the disease-region models produced an equal or lower MAE when testing in a new region, if a power analysis was conducted assuming a binary outcome and a continued trend, no amount of data would be sufficient to show that the transfer learning results would be significantly worse.

XSL•FO
RenderX

Table . The mean absolute error (MAE) values representing spatial generalizability of transfer learning models[a].

Disease	Country	Target region (cases in the past 6 months)	Ensemble model	Similar region (cases in the past 6 months)	MAE of baseline forecast	MAE of transfer learning forecast
Brucellosis	Israel	Jerusalem (34)	LGBM[b]	Akko (32)	5.41	2.3
Campylobacteriosis	Sweden	Kalmar (122)	XGB[c]	Västmanland (124)	5.01	16.98
MERS[d]	Saudi Arabia	Bisha (0)	ARIMA[e]	Eastern Province (0)	0.03	0.02
Q fever	Australia	New South Wales (37)	ARIMA	Queensland (48)	2.85	2.00
Tick-borne encephalitis	Germany	Saxony-Anhalt (3)	ARIMA	Brandenburg (3)	0.20	0.25
Tularemia	Japan	Ibaraki (0)	LGBM	Saga (0)	0.00	0.00

[a]For baseline forecasts, the models were trained and tested on target regions. For transfer learning forecasts, the models were trained on similar regions and tested on target regions.

[b]LGBM: light gradient boosting machine.

[c]XGB: extreme gradient boosting.

[d]MERS: Middle East respiratory syndrome.

[e]ARIMA: autoregressive integrated moving average.

Discussion

Principal Findings

As ML and DL techniques gain popularity in the ID forecasting domain, they are being extensively used across a wide range of pathogens with diverse ecology, geographic, and temporal scales. However, there is limited consensus among the forecasting community regarding the application and reporting of these results in a manner that lends themselves useful for operational decision-making in real-world circumstances [9]. In this study, we present a universal pipeline for analyzing data, generalizing the results, and automating reporting. We also address crucial operational metrics such as forecasting accuracy, computational efficiency, spatiotemporal generalizability, uncertainty quantification, and interpretability, which are essential in an operational scenario. This generalized and automated analytic pipeline is a major step toward addressing operational demands in the ID forecasting domain and to better inform public health and veterinary policies. We included 6 IDs with diverse transmission dynamics spanning 8 countries and 213 regions. Additionally, we trained our models using a broad range of data from demographic, geographic, climatic, and socioeconomic factors within the One Health landscape. This comprehensive approach enables our analysis to capture the intricate interplay of variables that drive infectious disease presence and transmission, more closely reflecting the complex realities observed in the real world.

Model Performance

We assessed forecasting model accuracies by their ability to detect the presence of a disease based on F_1-scores and by their ability to accurately forecast case counts when present using MAE. This 2-step approach prevented the inflation of model accuracy by the regions that did not encounter disease presence during our study period. Though we did not find a single best model suitable for all disease, region, and country combinations, the tree-based models were consistently better than all other models at detecting both disease presence and actual case counts with better F_1-scores and MAE values, respectively. Our ensemble technique, which selects the best-performing model for each disease-location pair, demonstrates superior forecasting performance by reducing prediction noise compared to single-model approaches, aligning with previous research [35-37]. While ideal ensemble methods would dynamically update to identify the best model at each time stamp, this demands extensive data, computational resources, and continuous validation. To address these challenges, we opted for a simpler approach by selecting the best model based on MAE for the testing dataset, minimizing resource and data requirements. However, ensemble techniques still require substantial computational power and analytic pipelines, posing challenges for low- and middle-income countries with limited resources. Investments in infrastructure, skill development, and data-sharing initiatives are crucial to fully leverage ML and DL techniques for ID forecasting in such settings.

Computational Efficiency

The computation time required for ID forecasting can broadly be split into model training and model prediction time. In our analysis, much of the training time was spent on hyperparameter tuning. Autoregressive and tree-based models took less time to tune than DL models. On the other hand, tree-based models (ie, RF and XGB) took considerably longer to produce predictions compared to the other models. This increase in time was mainly because of the time required to compute bootstrap PIs. The LGBM was an exception, as it used the least computational time for both training and testing compared to all the other models. Since our forecasting pipeline performed stepwise forecasting for each month with models retrained only once a year, more time was spent on predicting rather than model training. Hence, when building a forecasting pipeline where the

XSL•FO
RenderX

models are retrained once a year and predictions are made every month, it is computationally optimal to choose models with lower prediction time rather than training time.

The accuracy of modeling techniques can drastically differ with the amount of time and computational resources spent on fine-tuning the model. Here, we chose the initial hyperparameter values for optimization based on their relative importance and time required for overall analysis with the goal of achieving comparable results. It is important to find a balance between model complexity and computational efficiency according to the individual operational needs and available resources.

Uncertainty Quantification

Since the effectiveness of health policies and operational decision-making is driven by the accuracy of forecasts, policy makers need to know the credibility of the models in the form of prediction uncertainties [38]. In traditional statistical techniques, such as ARIMA and SARIMAX, forecasting uncertainties are computed based on probability theory and a set of statistical assumptions and, therefore, are readily available. However, generating such PIs is rather complicated in ML and DL compared to traditional methods due to additional uncertainties associated with noise distributions, hyperparameters, overparameterization, and optimization that should be accounted for [39]. Therefore, most of the ID forecasting studies fail to report their model uncertainty despite their recent popularity [9]. In our study, we used bootstrapping and probabilistic methods to compute PIs around our point estimates of ML and DL methods. Such bounds are crucial to assessing future uncertainties, making operational decisions, planning policies for a range of possible outcomes, and comparing different forecast models thoroughly [40]. The LGBM was the best-calibrated model with almost 95% of true observations covered by 95% of PIs, followed by other tree-based and statistical models. The DL models had the least desirable coverage along with their inferior predictive performance and narrow prediction band. This may not be surprising, though as there is evidence in the literature that uncertainty estimates around DL predictions often fail to capture the true data distribution [41].

Spatiotemporal Generalizability

Overall, the predictive performance of transfer learning models was comparable to their base models as presented in Table 8. For example, using the Akko model to predict case counts of brucellosis in Jerusalem created an increase in predictive accuracy. This suggests the potential for spatial generalizability of our ensemble, that is, an ability to predict existing diseases in new regions with low error. However, campylobacteriosis forecasting in Kalmar, Sweden, was an exception, where the model MAE increased when Västmanland feature data were used for training. Since campylobacteriosis tends to have higher case counts, the difference in MAE cannot be directly comparable to the other diseases with much lower case counts and naturally lower MAE values. There are many possible reasons for the difference seen in the results. For example, the similar region model may not have contained the necessary similarities to the target region required to enable a lower prediction error, such as differences in regional management practices. The ensemble model (best-performing model for disease-location pair) was different between the test cases and included both statistical (only case counts) and tree-based (many features) models. More advanced transfer learning techniques that include multiple input features to identify similarity and emphasize model retraining by reallocating model weights between target and similar regions could be considered to improve spatial generalizability [2,42]. Regardless, these results are encouraging, and a more comprehensive study is warranted.

Model Interpretability

We included interpretability for each model in the form of feature importance, a critical metric that is often neglected in the ID forecasting domain [9]. Interpretability provides insight into the decision-making process of the prediction systems, which is crucial for implementing appropriate disease mitigation strategies. Our scoping systematic review of ID prediction models showed that historic case counts are the most commonly used input feature in the ID forecasting domain, while the other important predictors, such as climate, demographics, socioeconomics, and geography, are often neglected [9]. In this study, we estimated lower feature ratio and MRR for lag case count features across countries and diseases, indicating that the non–case count features were equally important, if not more, compared to case count lags. This study suggests that including data related to the full disease ecology is critical for obtaining accurate, reliable, and interpretable ID forecasts. For example, cattle export was one of the important informative non–case count features for forecasting campylobacteriosis in Schleswig-Holstein, Germany. A direct causal association between the disease and cattle exports cannot be made just by a feature importance plot. However, campylobacteriosis is the most commonly reported bacterial food-borne gastrointestinal infection in the European Union that is closely associated with the dairy industry [43]. Our results suggest that high dairy and other agriculture activities, including cattle trading in Schleswig-Holstein, could play an important role in the disease prevalence in the region and, if adjusted, could result in disease mitigation.

Data Visualization

To make the best-performing models and corresponding predictions accessible to decision makers, we developed an interactive data visualization dashboard. This dashboard provides the raw data, the model predictions with accompanying 95% PIs and feature importance, overall model performance, and an interactive global map. This dashboard allows a user to visualize forecasts for up to a year into the future in any country or region where the data exist and can be used to inform control measures to reduce the spread of an ID within a country or region.

Conclusions

Our study provides a generalized platform to analyze and report ID forecasts with an emphasis on analytical accuracy, computational efficiency, uncertainty quantification, interpretability, and generalizability. These 5 aspects are crucial in determining the forecasting approach optimal for each situation's operational needs. While all the forecasting

XSL•FO
RenderX

techniques come with their own strengths and weaknesses, choosing an optimal approach is usually a tradeoff between computational efficiency, model complexity, and forecasting accuracy. Recognizing and addressing such nuances will facilitate the use of ID forecasts in an operational environment for better preparedness and response during an ID emergency.

Acknowledgments

This work was funded by the Defense Threat Reduction Agency (project CB11029).

Data Availability

The datasets generated or analyzed during this study are publicly available with website links in the Methods section, including Table 1.

Authors' Contributions

LEC and KTP acquired the funding. RK, KTP, SD, SE, and LEC conceptualized and designed the study. RK, CH, SD, and SE analyzed the data. RK, CH, and KTP created visualizations. RK drafted the manuscript. RK, KTP, CH, and LEC reviewed and edited the manuscript. All authors agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary figures and tables.
[DOCX File, 395 KB - publichealth_v11i1e59971_app1.docx]

References

1. Bloom RM, Buckeridge DL, Cheng KE. Finding leading indicators for disease outbreaks: filtering, cross-correlation, and caveats. J Am Med Inform Assoc 2007;14(1):76-85. [doi: 10.1197/jamia.M2178] [Medline: 17068353]
2. Keshavamurthy R, Charles LE. Predicting Kyasanur forest disease in resource-limited settings using event-based surveillance and transfer learning. Sci Rep 2023 Jul 8;13(1):11067. [doi: 10.1038/s41598-023-38074-0] [Medline: 37422454]
3. Woolhouse M. How to make predictions about future infectious disease risks. Philos Trans R Soc Lond B Biol Sci 2011 Jul 12;366(1573):2045-2054. [doi: 10.1098/rstb.2010.0387] [Medline: 21624924]
4. Bansal S, Chowell G, Simonsen L, Vespignani A, Viboud C. Big data for infectious disease surveillance and modeling. J Infect Dis 2016 Dec 1;214(Suppl 4):S375-S379. [doi: 10.1093/infdis/jiw400] [Medline: 28830113]
5. Dixon S, Keshavamurthy R, Farber DH, Stevens A, Pazdernik KT, Charles LE. A comparison of infectious disease forecasting methods across locations, diseases, and time. Pathogens 2022 Jan 29;11(2):185. [doi: 10.3390/pathogens11020185] [Medline: 35215129]
6. Shen L, Jiang C, Sun M, et al. Predicting the spatial-temporal distribution of human brucellosis in Europe based on convolutional long short-term memory network. Can J Infect Dis Med Microbiol 2022;2022:7658880. [doi: 10.1155/2022/7658880] [Medline: 35967090]
7. Da'ar OB, Ahmed AE. Underlying trend, seasonality, prediction, forecasting and the contribution of risk factors: an analysis of globally reported cases of Middle East Respiratory Syndrome Coronavirus. Epidemiol Infect 2018 Aug;146(11):1343-1349. [doi: 10.1017/S0950268818001541] [Medline: 29886854]
8. Adnan M, Gao X, Bai X, et al. Potential early identification of a large Campylobacter outbreak using alternative surveillance data sources: autoregressive modelling and spatiotemporal clustering. JMIR Public Health Surveill 2020 Sep 17;6(3):e18281. [doi: 10.2196/18281] [Medline: 32940617]
9. Keshavamurthy R, Dixon S, Pazdernik KT, Charles LE. Predicting infectious disease for biopreparedness and response: a systematic review of machine learning and deep learning approaches. One Health 2022 Dec;15:100439. [doi: 10.1016/j.onehlt.2022.100439] [Medline: 36277100]
10. Chae S, Kwon S, Lee D. Predicting infectious disease using deep learning and big data. Int J Environ Res Public Health 2018 Jul 27;15(8):1596. [doi: 10.3390/ijerph15081596] [Medline: 30060525]
11. Oidtman RJ, Omodei E, Kraemer MUG, et al. Trade-offs between individual and ensemble forecasts of an emerging infectious disease. Nat Commun 2021 Sep 10;12(1):5379. [doi: 10.1038/s41467-021-25695-0] [Medline: 34508077]
12. Reich NG, McGowan CJ, Yamana TK, et al. Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the U.S. PLOS Comput Biol 2019 Nov;15(11):e1007486. [doi: 10.1371/journal.pcbi.1007486] [Medline: 31756193]
13. Ma S, Ning S, Yang S. Joint COVID-19 and influenza-like illness forecasts in the United States using internet search information. Commun Med (Lond) 2023 Mar 24;3(1):39. [doi: 10.1038/s43856-023-00272-2] [Medline: 36964311]

XSL•FO
RenderX

14. Al-Jarrah OY, Yoo PD, Muhaidat S, Karagiannidis GK, Taha K. Efficient machine learning for big data: a review. Big Data Research 2015 Sep;2(3):87-93. [doi: 10.1016/j.bdr.2015.04.001]
15. Bhatt U, Antorán J, Zhang Y, et al. Uncertainty as a form of transparency: measuring, communicating, and using uncertainty. Presented at: Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society; May 19-21, 2021; Virtual Event, United States p. 401-413. [doi: 10.1145/3461702.3462571]
16. Generous N, Fairchild G, Khalsa H, Tasseff B, Arnold J. Epi Archive: automated data collection of notifiable disease data. OJPHI 2017 May 2;9(1). [doi: 10.5210/ojphi.v9i1.7615]
17. Welte VR, Vargas Terán M. Emergency Prevention System (EMPRES) for transboundary animal and plant pests and diseases. The EMPRES-livestock: an FAO initiative. Ann N Y Acad Sci 2004 Oct;1026:19-31. [doi: 10.1196/annals.1307.003] [Medline: 15604466]
18. GADM. URL: https://gadm.org/ [accessed 2024-02-24]
19. GES DISC. URL: https://disc.gsfc.nasa.gov/ [accessed 2024-02-24]
20. Earth Science Data Systems. NASA. 2024. URL: https://www.earthdata.nasa.gov/ [accessed 2024-02-24]
21. US Department of Agriculture. Economic Research Service—Home. URL: https://www.ers.usda.gov/ [accessed 2024-02-24]
22. DIVA-GIS | free, simple & effective. URL: https://diva-gis.org/ [accessed 2024-02-24]
23. World Health Organization (WHO). URL: https://www.who.int [accessed 2024-02-24]
24. OECD Statistics. URL: https://stats.oecd.org/ [accessed 2024-02-24]
25. Natural Earth. Free vector and raster map data at 1:10m, 1:50m, and 1:110m scales. URL: https://www.naturalearthdata.com/ [accessed 2024-02-24]
26. Population division. United Nations. URL: https://www.un.org/development/desa/pd/ [accessed 2024-02-24]
27. SEDAC: Socioeconomic Data and Applications Center. NASA. URL: https://sedac.ciesin.columbia.edu/ [accessed 2024-02-24]
28. Josef P, Skipper S, Kevin S, et al. Statsmodels: econometric and statistical modeling with python. 2010 Presented at: Proceedings of the 9th Python in Science Conference; Texas, US URL: https://www.statsmodels.org/ [accessed 2025-02-28]
29. scikit-learn: machine learning in Python. URL: https://scikit-learn.org/stable/ [accessed 2024-02-24]
30. Taylor SJ, Letham B. Forecasting at scale. PeerJ Preprints 2017 Sep;5:e3190v2. [doi: 10.7287/peerj.preprints.3190]
31. Darts: time series made easy in Python. Unit8. URL: https://unit8.com/resources/darts-time-series-made-easy-in-python/ [accessed 2024-02-24]
32. Amat Rodrigo J, Escobar Ortiz J. Skforecast. Zenodo. 2024. URL: https://zenodo.org/records/8382788 [accessed 2025-02-28]
33. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Presented at: Advances in Neural Information Processing Systems; Dec 4-9, 2017; California, United States URL: https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
34. Topical overviews. SHAP Latest Documentation. URL: https://shap.readthedocs.io/en/latest/overviews.html [accessed 2024-03-02]
35. Alfred R, Obit JH. The roles of machine learning methods in limiting the spread of deadly diseases: a systematic review. Heliyon 2021 Jun;7(6):e07371. [doi: 10.1016/j.heliyon.2021.e07371] [Medline: 34179541]
36. Khan SD, Alarabi L, Basalamah S. Toward smart lockdown: a novel approach for COVID-19 hotspots prediction using a deep hybrid neural network. Computers 2020;9(4):99. [doi: 10.3390/computers9040099]
37. Kasilingam D, Sathiya Prabhakaran SP, Rajendran DK, Rajagopal V, Santhosh Kumar T, Soundararaj A. Exploring the growth of COVID-19 cases using exponential modelling across 42 countries and predicting signs of early containment using machine learning. Transbound Emerg Dis 2021 May;68(3):1001-1018. [doi: 10.1111/tbed.13764] [Medline: 32749759]
38. Gilbert JA, Meyers LA, Galvani AP, Townsend JP. Probabilistic uncertainty analysis of epidemiological modeling to guide public health intervention policy. Epidemics 2014 Mar;6:37-45. [doi: 10.1016/j.epidem.2013.11.002] [Medline: 24593920]
39. Psaros AF, Meng X, Zou Z, Guo L, Karniadakis GE. Uncertainty quantification in scientific machine learning: methods, metrics, and comparisons. J Comput Phys 2023 Mar;477:111902. [doi: 10.1016/j.jcp.2022.111902]
40. Chatfield C. Prediction intervals for time-series forecasting. In: Armstrong JS, editor. Principles of Forecasting: A Handbook for Researchers and Practitioners: Springer; 2001:475-494. [doi: 10.1007/978-0-306-47630-3_21ISBN:978-0-306-47630-3]
41. Lakshminarayanan B, Pritzel A, Blundell C. Simple and scalable predictive uncertainty estimation using deep ensembles. Presented at: Advances in Neural Information Processing Systems; Dec 4-9, 2017; California, United States URL: https://proceedings.neurips.cc/paper_files/paper/2017/hash/9ef2ed4b7fd2c810847ffa5fa85bce38-Abstract.html [accessed 2025-02-28]
42. Roster K, Connaughton C, Rodrigues FA. Forecasting new diseases in low-data settings using transfer learning. Chaos Solitons Fractals 2022 Aug;161:112306. [doi: 10.1016/j.chaos.2022.112306] [Medline: 35765601]
43. Henningsen A, Henning C, Struve C, Mueller-Scheessel J. Economic impact of the mid-term review on agricultural production, farm income and farm survival: a quantitative analysis for local sub-regions of Schleswig-Holstein in Germany. Presented at: 2005 International Congress; Aug 23-27, 2005; Copenhagen, Denmark. [doi: 10.22004/ag.econ.24644]

XSL•FO
RenderX

Abbreviations

ARIMA: autoregressive integrated moving average
DL: deep learning
ID: infectious disease
LASSO: least absolute shrinkage and selection operator
LGBM: light gradient boosting machine
LSTM: long short-term memory
MAE: mean absolute error
ML: machine learning
MRR: mean reciprocal rank
PI: prediction interval
RF: random forest
RNN: recurrent neural network
SARIMAX: seasonal autoregressive integrated moving average with exogenous factors
SHAP: Shapley Additive Explanation
XGB: extreme gradient boosting

Edited by KS Sahu; submitted 27.04.24; peer-reviewed by S Ma, Y Hua; revised version received 10.12.24; accepted 24.12.24; published 21.03.25.

Please cite as:
Keshavamurthy R, Pazdernik KT, Ham C, Dixon S, Erwin S, Charles LE
Meeting Global Health Needs via Infectious Disease Forecasting: Development of a Reliable Data-Driven Framework
JMIR Public Health Surveill 2025;11:e59971
URL: https://publichealth.jmir.org/2025/1/e59971
doi:10.2196/59971

© Ravikiran Keshavamurthy, Karl T Pazdernik, Colby Ham, Samuel Dixon, Samantha Erwin, Lauren E Charles. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Preferences for Nonpharmaceutical Interventions During the Endemic Phase of COVID-19: Discrete Choice Experiment

Yi Wang, PhD; Chee Ern Har, BS; Sharon Hui Xuan Tan, MPH; Hooi Swang Cheng, MSc; Ian Yi Han Ang, PhD

Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, #09-01D, Tahir Foundation Building, 12 Science Drive 2, Singapore, Singapore

Corresponding Author:
Yi Wang, PhD
Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, #09-01D, Tahir Foundation Building, 12 Science Drive 2, Singapore, Singapore

Abstract

Background: Nonpharmaceutical interventions (NPIs) are effective tools for pandemic containment but often impose significant socioeconomic consequences that intensify over time. Public support and compliance to NPIs are crucial to ensure their effectiveness.

Objectives: This study aimed to elicit preferences of a Singaporean population for the reintroduction of NPIs in response to the emergence of a new SARS-CoV-2 variant during the COVID-19 endemic phase.

Methods: A web-based discrete choice experiment (DCE) was conducted. DCE attributes reflected key NPIs implemented in Singapore during the COVID-19 pandemic from 2020 to 2022, including mask wearing, dining restrictions, suspension of vocalization activities and large-scale events, quarantine after international travel, and mandatory vaccine boosters. Participants were recruited from a demographically representative online panel. Statistical analysis was performed using a mixed-logit model and mixed-mixed multinomial logit model.

Results: A total of 1552 participants were included in the analysis. Overall preferences from the mixed-logit model showed that mask wearing was valued, both in public and indoors. Dining restrictions allowing groups of up to 5 people were preferred, but stricter dining restrictions allowing up to 2 people or no dining out were not favored. Prohibiting large-scale events was not preferred. Participants accepted quarantine at home but opposed quarantine in government facilities. Two classes emerged from the mixed-mixed logit model: class 1 (“Prefer NPIs”) and class 2 (“Prefer No NPIs”). While class 1 (39%) was only opposed to a complete prohibition on dining in at food and beverage establishments, no NPIs were preferred by class 2 (61%). Both classes were not opposed to mandatory mask wearing, dining restrictions allowing groups of up to 5 people, and mandatory vaccine boosters. Sex, age, education, employment status, the number of COVID-19 vaccine shots received, and risk attitude were associated with individuals’ likelihood of belonging to a specific preference group.

Conclusions: Following the emergence of a new SARS-CoV-2 variant after a prolonged period of restrictions, less disruptive NPIs such as mask wearing indoors were valued by the public and should be swiftly reinstated. Adaptive strategies should be adopted for more contentious NPIs, such as strict dining restrictions and quarantine policies. Public preferences should be considered in the design and selection of NPIs for future pandemic containment strategies to enhance compliance and effectiveness.

(JMIR Public Health Surveill 2025;11:e67725) doi:10.2196/67725

KEYWORDS

discrete choice experiment; nonpharmaceutical interventions; population preferences; endemic; COVID-19

Introduction

Background

Projections suggest that pandemics are likely to occur more frequently [1], with the annual probability of extreme epidemics potentially increasing 3-fold in the coming years due to factors such as climate change [2], zoonotic diseases [3], and increased cross-border travel [4]. As a result, pandemic preparedness is becoming increasingly crucial. Nonpharmaceutical interventions (NPIs) rose to prominence during the COVID-19 pandemic due to their central role in global containment strategies [5]. NPIs encompass all measures aimed at reducing or slowing disease transmission apart from vaccinations and medical interventions, such as mask wearing, social distancing, and movement restrictions [6,7]. The effectiveness of NPIs in reducing the transmissibility and severity of the SARS-CoV-2 virus [8-10], as well as decreasing the incidence of other infectious diseases [5,11-13], has been demonstrated in numerous studies. Nevertheless, the public health benefits of NPIs should be weighed against their broader societal impacts, as these measures also led to a host of socioeconomic consequences such as

XSL•FO
RenderX

increased financial burden, mental health decline, and disrupted learning [14,15].

The success of NPIs hinges on public support and compliance, underscoring the need for strategies informed by population preferences [16]. Discrete choice experiments (DCEs) are widely used in public health to assess preferences by having participants choose among different scenarios, allowing researchers to analyze trade-offs and estimate willingness to pay (WTP) [17-20]. Several DCE studies have explored preferences for NPIs during the COVID-19 pandemic, primarily in Western populations [21-33], with 1 study in China [28]. These studies examined a range of attributes, from the health, social, and economic outcomes of NPIs to preferences for specific measures, such as mask mandates or stay home policies. However, most were conducted in the early phase of the pandemic, reflecting preferences during a period of heightened uncertainty.

Few studies have assessed population preferences for NPIs in the endemic phase of COVID-19, following a prolonged period of restrictions. As reported by Yang et al [28], respondents' preferences for NPIs may change as a result of pandemic fatigue, a phenomenon characterized by declining adherence to restrictions over time [34,35]. Additionally, preferences for NPIs are shaped by demographic, socioeconomic, and cultural factors [16,34], limiting the generalizability of findings across regions. Understanding these influences is particularly crucial in an Asian context, where cultural and policy responses may differ from Western counterparts.

Objectives

This study addresses these gaps by using a DCE to elicit Singaporeans' preferences for NPIs in the event of a new SARS-CoV-2 variant during the COVID-19 endemic phase. By capturing preferences shaped by regional influences and real-world experiences with prolonged COVID-19 restrictions, the findings provide valuable insights for designing containment strategies that maximize public acceptance, minimize societal disruptions, and enhance long-term resilience against future health crises.

Methods

NPIs in Singapore

Following the detection of local COVID-19 clusters, safe distancing measures, including border restrictions and contact tracing, were introduced in Singapore in March 2020 [36]. By April, mask wearing in public became mandatory, and a "Circuit Breaker" was implemented, mandating the closure of all nonessential businesses, schools, and recreational facilities. Residents were instructed to stay home, with nonessential workers required to telecommute. Following the end of the Circuit Breaker in June 2020, restrictions on dining, group activities, and large-scale events were enforced. Large-scale events (eg, concerts and sporting events) and vocalization activities (eg, karaoke, choirs, and singing classes) were suspended, while limits on social gatherings and dining group sizes fluctuated based on outbreak severity. During times of heightened alert, dining in at food and beverage (F&B) establishments was completely suspended. Border controls remained in place throughout 2020 and 2021, requiring incoming travelers to quarantine either at government-designated facilities or at home. Extensive testing and contact-tracing efforts were adopted for the rapid identification and isolation of COVID-19 cases. Following the rollout of vaccination programs in late 2020, Vaccine-Differentiated Safe Management Measures and Vaccinated Travel Lanes were introduced in 2021. By early 2022, Singapore had transitioned to an endemic approach to COVID-19, with the easing of most NPIs in April 2022 marking a return to normalcy.

DCE Question Design

Attributes were selected based on a review of key NPIs implemented in Singapore during the COVID-19 pandemic, as outlined in official public health guidelines [36-40]. The chosen NPIs were broad-based measures applicable to the entire population, rather than being specific to particular segments (eg, schools or workplaces). Attribute levels were derived from the actual tiers implemented during different phases of tightening and relaxation in Singapore, ensuring that they accurately reflected real-world policy shifts. Unlike conventional approaches that rely on literature reviews or stakeholder consultations, this method offers several advantages. As participants were familiar with the attributes and levels, ambiguity in scenario interpretation was reduced, allowing them to make choices based on lived experiences rather than purely hypothetical situations. Moreover, the alignment of study attributes with real-world NPIs ensures greater relevance for local policy makers and enhances policy impact. The selection of broad-based NPIs also strengthens the applicability of the study findings, as these measures have wider social and economic implications than segment-specific policies.

Table 1 outlines the seven attributes that were used in the DCE: (1) mandatory wearing of masks, (2) dining in group size restrictions, (3) restrictions on vocalization activities, (4) permission of large-scale events, (5) quarantine requirements after travel, (6) vaccine booster mandates, and (7) a one-time fee for public health measures. The inclusion of a cost attribute (one-time fee) is a standard approach in DCEs to estimate WTP, allowing for the assessment of how financial considerations influence public preferences for NPIs [41].

XSL•FO
RenderX

Table . Attributes and levels of the discrete choice experiment.

Attribute	Levels
Masks	• Masks always required in public • Masks required only when indoors • No masks are required
Dining in	• No dining in allowed • Dining in for 2 people • Dining in for groups of 5 • Any group size can dine in
Vocalization activities	• No vocalization activities allowed • Vocalization activities allowed
Large-scale events	• No large-scale events can be held • Large-scale events can be held
Quarantine after travel	• Quarantine in government facility • Quarantine at home • No quarantine
Vaccine booster	• Compulsory • Not compulsory
One-time fee (SGD $)[a]	• $0 • $10 • $30 • $50

[a]SGD $1≈US $0.77.

Sawtooth Software Lighthouse Studio (Academic Advanced Teaching Suite) was used to design the DCE questionnaire [42]. The balanced overlapping option was used to strike a balance between statistical efficiency and response efficiency. An unlabeled 1-stage design was used. The final design consisted of 20 blocks, each containing 7 DCE questions with 2 alternatives per question. Each attribute had up to 4 levels (Table 1). Participants were each assigned to 1 block of DCE questions at random.

The full survey consisted of 2 main parts: the first presented participants with 2 sets of DCEs, one with the above-mentioned design, and another from a different study design (not covered in this manuscript). The second part consisted of questions on sociodemographic information, including history of chronic diseases, COVID-19 vaccination status, employment, family background, housing type, teleconsultation use, and the Health Risk Attitude Scale (HRAS-13). The HRAS-13 is a 13-item scale designed to assess participants' risk acceptance and tolerance in the health domain, with a higher sum score indicating a greater willingness to accept or tolerate potential health risks [43]. Its reliability and validity in the general population were established in a previous evaluation study [43]. The survey underwent 2 rounds of user acceptability testing with 15 and 20 staff members, respectively, to gather feedback on its suitability, clarity, and length. Based on this feedback, the survey refined before being disseminated to the full cohort. The complete set of survey instructions and a sample DCE question are shown in Multimedia Appendix 1.

Sample, Sample Size, and Data Collection

Study participants were from the Singapore Population Health Studies (SPHS) Online Panel. This online platform facilitated public health research through monthly surveys on a diverse range of health-related topics. Participants eligible for the survey were Singapore citizens or permanent residents aged 18 years and older who met the SPHS Online Panel enrollment criteria, including the ability to read and comprehend English, and access to the web via a computer or a smartphone. Participants responded to the surveys hosted on the secure REDCap (Research Electronic Data Capture) system for National University of Singapore (NUS). At the time of survey dissemination, the panel had 2058 active participants who were demographically representative of the general population of Singapore. As participants were already enrolled in the SPHS Online Panel and had provided prior consent, no further recruitment was conducted. Participants could opt out of the questionnaire and choose not to participate in the study.

Using a rule of thumb formula $n>500\times c/(t\times a)$ by Johnson and Orme [44], where 500 is a fixed variable, c denotes the largest number of levels for a certain attribute, a indicates the number of DCE choice sets per block of questionnaire, and t refers to the number of alternatives per DCE choice set, the sample size required for this study should be >143 participants ($500\times4/[2\times7]=143$). In addition, Lancsar and Louviere [45] had suggested 20 responses per block or questionnaire, thereby generating a minimum sample size of 400. Historically, the response rate for the SPH Online Panel Cohort has averaged about 78%, which would yield more than enough participants to identify the main effect [19,20].

XSL•FO
RenderX

Ethical Considerations

The survey was administered from June 20, 2022, to June 29, 2022. SPHS Online Panel members were notified of the survey via email and text message on June 20, 2022, and were given 10 days to complete the survey. SPHS Online Panel was approved by the NUS Institutional Review Board (IRB; reference: H-18‐011). A waiver for documented informed consent for this study was granted by the NUS IRB (reference: NUS-IRB-2020‐82). Survey participation was entirely voluntary, and a participant information sheet was provided prior to starting the survey. An electronic consent procedure ensured that participants were informed of their rights before proceeding with the survey. Participants could withdraw from the study at any time without any obligation or penalty. To maintain confidentiality, all responses that the study team received were deidentified. Research data were securely stored in compliance with institutional data protection policies. Participants received SGD $15 (SGD $1≈US $0.77) as compensation upon survey completion.

Variable Coding

To assess the impact of sociodemographic factors on preferences for NPIs, the study collected data on participants' sex (male or female), ethnicity (Chinese or non-Chinese), marital status (married or not married), and parental status (with children or with no children). Education level (secondary school and below, or above secondary school and up to diploma equivalent, or university or higher) and housing type (private housing, such as condominiums, and larger government housing, such as 4- to 5-room housing development board flats, or smaller housing development board flats, such as 1- to 3-room units) were included as proxies for socioeconomic status [46,47]. Employment status was divided into 4 categories (full-time employed, self-employed or business owner, homemakers, others including part-time, unemployed, or retired). The number of COVID-19 vaccine shots received was grouped into 2 categories (zero to 2 shots or 3 shots and above), while age was treated as a continuous variable. HRAS-13 scores were also treated as a continuous variable, with higher values indicating less risk aversion.

Statistical Analysis

Participants who refused or missed answering any of the survey questions were removed from the analysis. Two models were used in the analysis: mixed-logit (MXL) model and mixed-mixed multinomial logit (MMML) model [48]. The MXL model can account for random preference weight at individual level and relax the assumptions of independence of irrelevant alternatives [49]. Equation 1 shows the use for each individual, where Uij is the use for individual i when choosing choice j, xj is a vector describing choice j, βi is the preference weight or coefficient for individual i, and ij is the idiosyncratic error term following a type 1 extreme value distribution.

$$(1) Uij=\beta i\times xj+ ij$$

Equation 2 shows the formula for βi. β is the population mean. ηi represents the random preference weight at the individual level following multivariate normal distribution, which includes all factors that affect individuals' preference. Values for β are presented in the main results. SDs for ηi are shown in Multimedia Appendix 2. Significant values for the SDs of ηi means preference heterogeneity at individual level exists.

$$(2) \beta i=\beta+\eta i$$

The MMML model incorporates both the MXL model and the latent class model. The latent class model assumes that study participants can be assigned to unobserved (latent) classes based on their preference patterns. These preference weights are uniform within each class but vary systematically across classes. The MMML model works similarly by assigning study participants to unobserved classes. Mean preference weights, for example, β, are systematically different across classes. However, like the MXL model, people have same mean preference weights within each class but may have different preference weights at the individual level. Demographic information can be used to predict class membership. The optimal number of classes was chosen based on the Bayesian Information Criterion.

An alternative-specific constant (left) was included in each regression model, with a statistically significant coefficient indicating a left-right bias in the study [50]. Dummy coding was used for all attributes, with the exception of one-time fee for public health measures. For each attribute, the reference category was set as the lowest stringency of the NPI, that is, no restriction. Statistical significance was set at $P<.05$. All quantitative data analyses were carried out using R statistical software (version 4.2.2; The R Foundation) [51].

Results

Participant Characteristics

A total of 1651 participants were recruited in the study. After excluding participants with incomplete responses and 1 participant who indicated an age younger than 18 years, 1552 participants were included in the subsequent analysis. Participants had a mean age of 48 (SD 14) years, and a mean HRAS-13 score of 42 (SD 8). The remaining sociodemographic characteristics of the participants are summarized in Table 2. Compared with the demographic profile of the Singapore general population, the study sample had a higher median age and a higher proportion of females, individuals of Chinese ethnicity, and those with higher educational attainment (Multimedia Appendix 3).

XSL•FO
RenderX

Table . Distribution of sociodemographic characteristics of the study participants.

Characteristics	Frequency, n	Proportion, %
Total sample size	1552	N/A[a]
Sex		
Male	648	41.8
Female	904	58.2
Ethnicity		
Chinese	1314	84.7
Non-Chinese	238	15.3
Education		
Secondary school and below	335	21.6
Above secondary school and up to diploma equivalent	416	26.8
University and above	801	51.6
Housing		
1-3 room HDB[b]	343	22.1
4-5 room HDB[b] or private property	1209	77.9
Marital status		
Married	1004	64.7
Unmarried	548	35.3
Parental status		
Have children	936	60.3
No children	616	39.7
Employment status		
Full-time	882	56.8
Self-employed or business owner	133	8.6
Homemaker	96	6.2
Others (eg, part-time, retired, or unemployed)	441	28.4
COVID-19 vaccine shots received		
0-2 shots	111	7.2
3 shots and above	1441	92.8

[a]N/A: not applicable.

[b]HDB: housing development board—government-provided housing.

Regression Results

The regression results are shown in Tables 3 and 4. Overall preferences of the study participants are described by the MXL model (Table 3). In the event of a new SARS-CoV-2 variant, mandatory mask wearing was valued, both in public (coefficient=0.54, 95% CI 0.44-0.64) and indoors (coefficient=0.59, 95% CI 0.49-0.69). Dining restrictions allowing groups of up to 5 were preferred (coefficient=0.19, 95% CI 0.08-0.30) while stricter dining restrictions of up to 2 people (coefficient=−0.33, 95% CI −0.44 to −0.22) or the complete prohibition of dining in at F&B establishments (coefficient=−0.89, 95% CI −1 to −0.78) was not favored. There was no significant preference regarding restrictions on vocalization activities (coefficient=−0.05, 95% CI −0.11 to 0.01); however, prohibiting large-scale events was not preferred (coefficient=−0.07, 95% CI −0.14 to 0). Participants were indifferent to home quarantine (coefficient=0.10, 95% CI 0-0.20) but opposed quarantine in government facilities (coefficient=−0.37, 95% CI −0.46 to −0.28). There was no clear preference for mandatory vaccine boosters (coefficient=0.03, 95% CI −0.03 to 0.09). Coefficients for unexplained preference heterogeneity at the individual level are shown in Multimedia Appendix 2. Overall, preference heterogeneity at the individual level was found for mandatory mask wearing (both public and indoors), complete prohibition of dining in at F&B establishments, quarantine at government facilities, and compulsory vaccine boosters.

XSL•FO
RenderX

Table . Mixed-logit model regression results.

	Coefficient	95% CI
Left	−0.08[a]	−0.14 to −0.02
Mask: Mandatory in public	0.54[b]	0.44 to 0.64
Mask: Mandatory indoors	0.59[b]	0.49 to 0.69
Dining in: Not allowed	−0.89[b]	−1 to −0.78
Dining in: Two people	−0.33[b]	−0.44 to −0.22
Dining in: Five people	0.19[b]	0.08 to 0.30
Vocalization activities: Not allowed	−0.05	−0.11 to 0.01
Large event: Not allowed	−0.07	−0.14 to 0
Quarantine: Government facility	−0.37[b]	−0.46 to−0.28
Quarantine: Home	0.10	0 to 0.20
Vaccine booster: Mandatory	0.03	−0.03 to 0.09
Fee for public health measure	−0.02[b]	−0.02 to −0.02

[a]*P*<.05.

[b]*P*<.001.

Table . Mixed-mixed multinomial logit model regression results.

	Class 1: Prefer NPI[a] (39%)			Class 2: Prefer No NPI (61%)		
	Coefficient	95% CI	*P* value	Coefficient	95% CI	*P* value
Left	−0.19	−0.34 to −0.04	.01	0.11	0.01 to 0.21	.04
Mask: Mandatory in public	1.98	1.63 to 2.33	<.001	−0.37	−0.53 to −0.21	<.001
Mask: Mandatory indoors	1.76	1.43 to 2.09	<.001	0.04	−0.11 to 0.19	.62
Dining in: Not allowed	−0.51	−0.75 to −0.27	<.001	−1.72	−1.98 to −1.46	<.001
Dining in: Two people	0.18	−0.09 to 0.45	.20	−0.87	−1.07 to −0.67	<.001
Dining in: Five people	0.49	0.21 to 0.77	<.001	−0.08	−0.26 to 0.10	.36
Vocalization activities: Not allowed	−0.06	−0.2 to 0.08	.42	−0.12	−0.22 to ‐0.02	.02
Large event: Not allowed	0.21	0.04 to 0.38	.01	−0.29	−0.41 to −0.17	<.001
Quarantine: Government facility	0.35	0.13 to 0.57	.002	−1.01	−1.19 to −0.83	<.001
Quarantine: Home	0.56	0.31 to 0.81	<.001	−0.19	−0.35 to −0.03	.02
Vaccine booster: Mandatory	−0.02	−0.17 to 0.13	.80	−0.02	−0.12 to 0.08	.64
Fee for public health measure	−0.02	−0.03 to −0.01	<.001	−0.04	−0.05 to −0.03	<.001

[a]NPI: nonpharmaceutical intervention.

In the MMML model, participants were grouped into 2 classes (Table 4). Class 1 was defined as “Prefer NPI,” making up 39% of the sample, while class 2 was defined as “Prefer No NPI,” comprising the remaining 61%. Class 1 was only opposed to prohibition on dining in at F&B establishments (coefficient=−0.51, 95% CI −0.75 to −0.27). Mandatory mask wearing was highly valued, both in public (coefficient=1.98, 95% CI 1.63-2.33) and indoors (coefficient=1.76, 95% CI

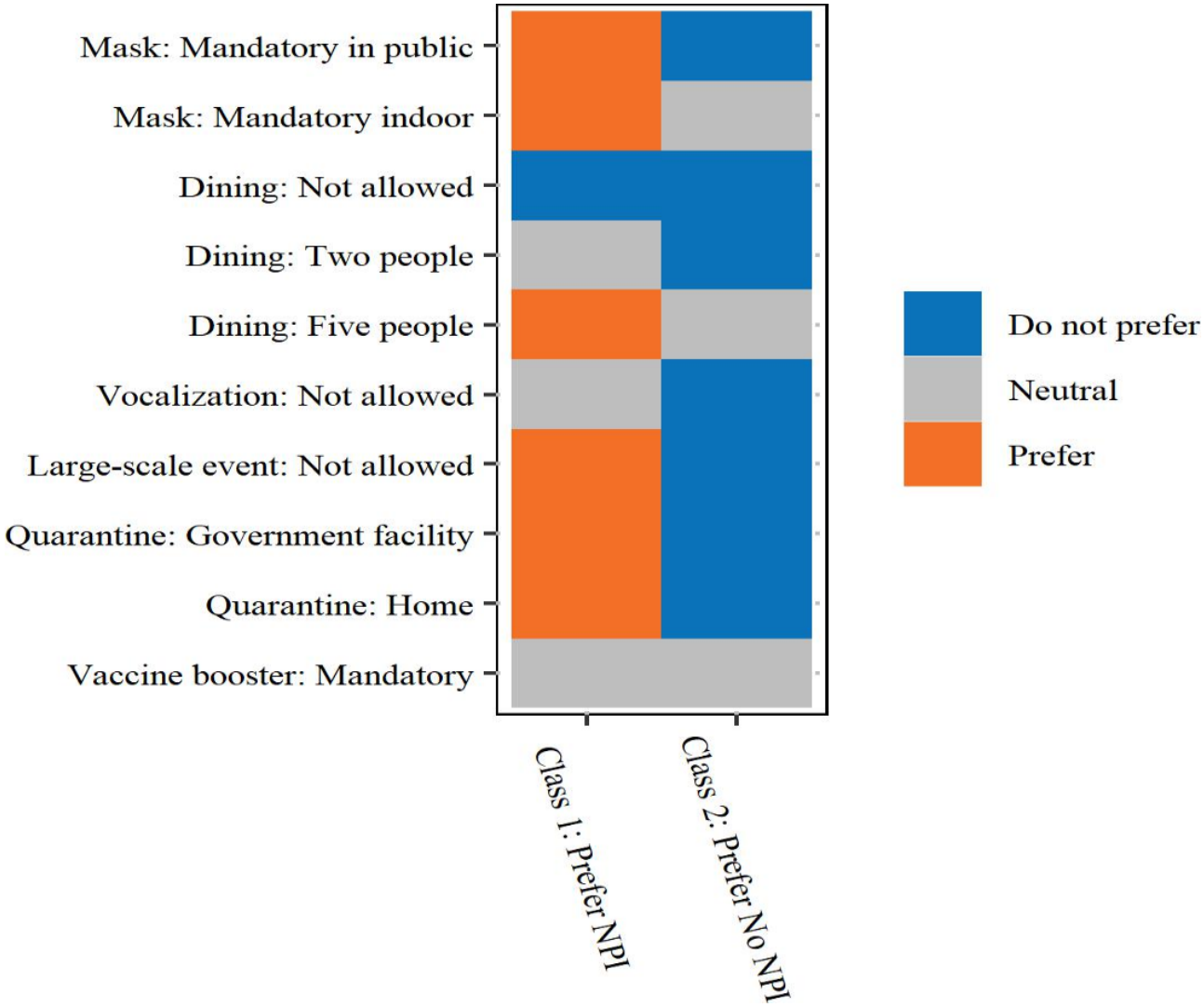
XSL•FO
RenderX

1.43-2.09). For class 2, none of the NPIs were preferred. However, they were not averse to mandatory mask wearing indoors (coefficient=0.04, 95% CI −0.11 to 0.19), dining restrictions allowing groups of up to 5 people (coefficient=−0.08, 95% CI −0.26 to 0.10), and mandatory vaccine booster (coefficient=−0.02, 95% CI −0.12 to 0.08). Coefficients for unexplained preference heterogeneity at the individual level are summarized in Multimedia Appendix 2.

Figure 1 provides an overview of participants' preferences for the given NPIs. Neither of the classes were averse to mandatory mask wearing indoors, dining restrictions allowing groups of up to 5 people, and mandatory vaccine boosters. However, both groups were opposed to a complete prohibition on dining in at F&B establishments.

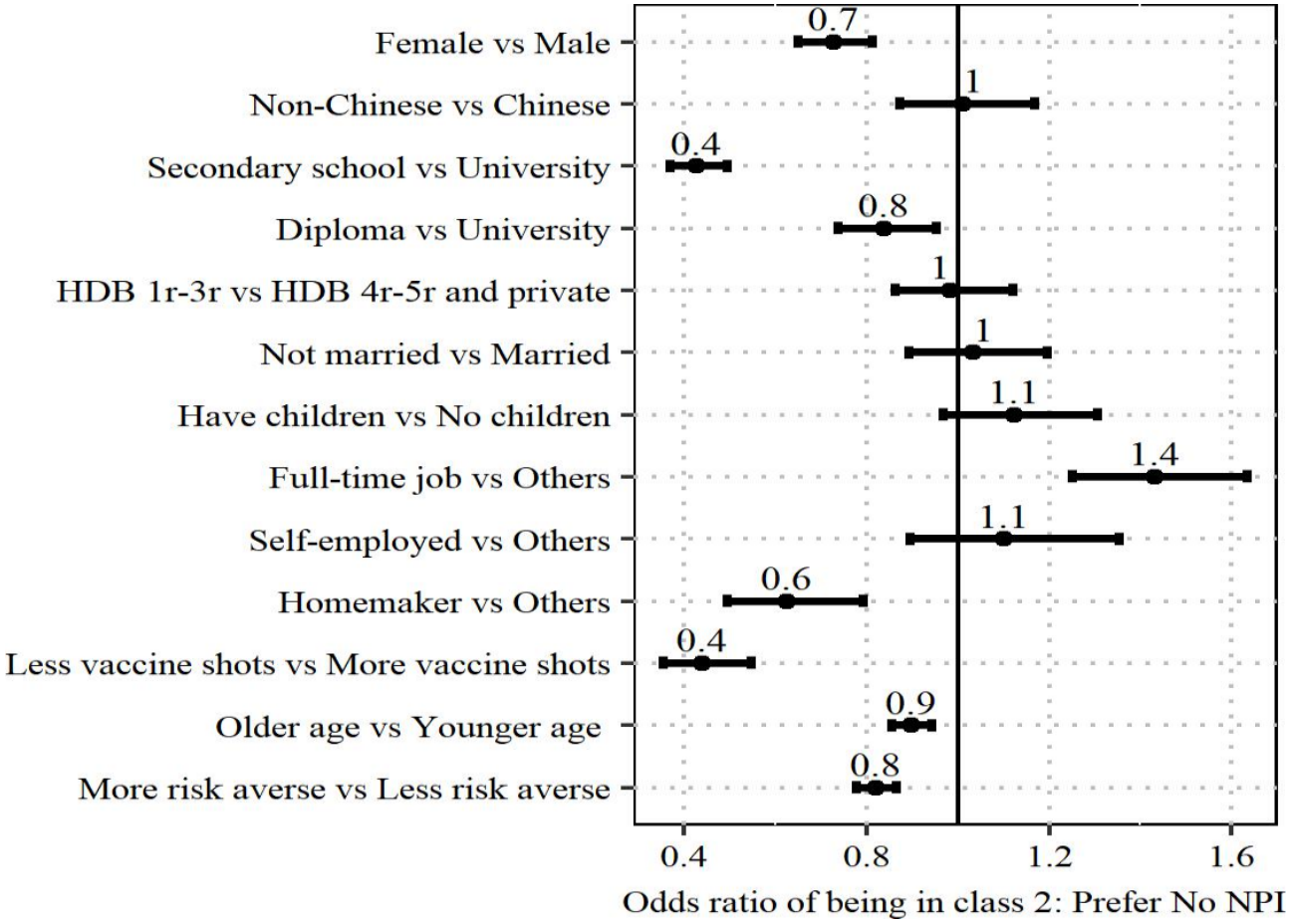
Figure 2 shows the odds ratio for membership in class 2 (Prefer No NPI). Men were more likely to belong in class 2 and prefer no NPIs as compared with women. Higher education levels were also associated with higher odds of preferring no NPIs. Those in full-time employment preferred no NPIs; in contrast, homemakers were more likely to prefer NPIs. Respondents who had received fewer vaccine shots tended to prefer NPIs, compared with those who had received more. Additionally, older age and higher risk aversion were associated with higher odds of preferring NPIs.

Figure 1. Preferences for NPIs for class 1 (Prefer NPI) and class 2 (Prefer No NPI). NPI: nonpharmaceutical intervention.



XSL•FO
RenderX

Figure 2. Odds ratio of being in class 2 (Prefer No NPI) comparing different demographic groups. An odds ratio greater than 1 indicates a higher likelihood of being in class 2 for the first group listed, while an odds ratio less than 1 indicates a lower likelihood of being in class 2 for the first group compared with the second group (eg, female vs male). HDB: housing development board; NPI: nonpharmaceutical intervention.

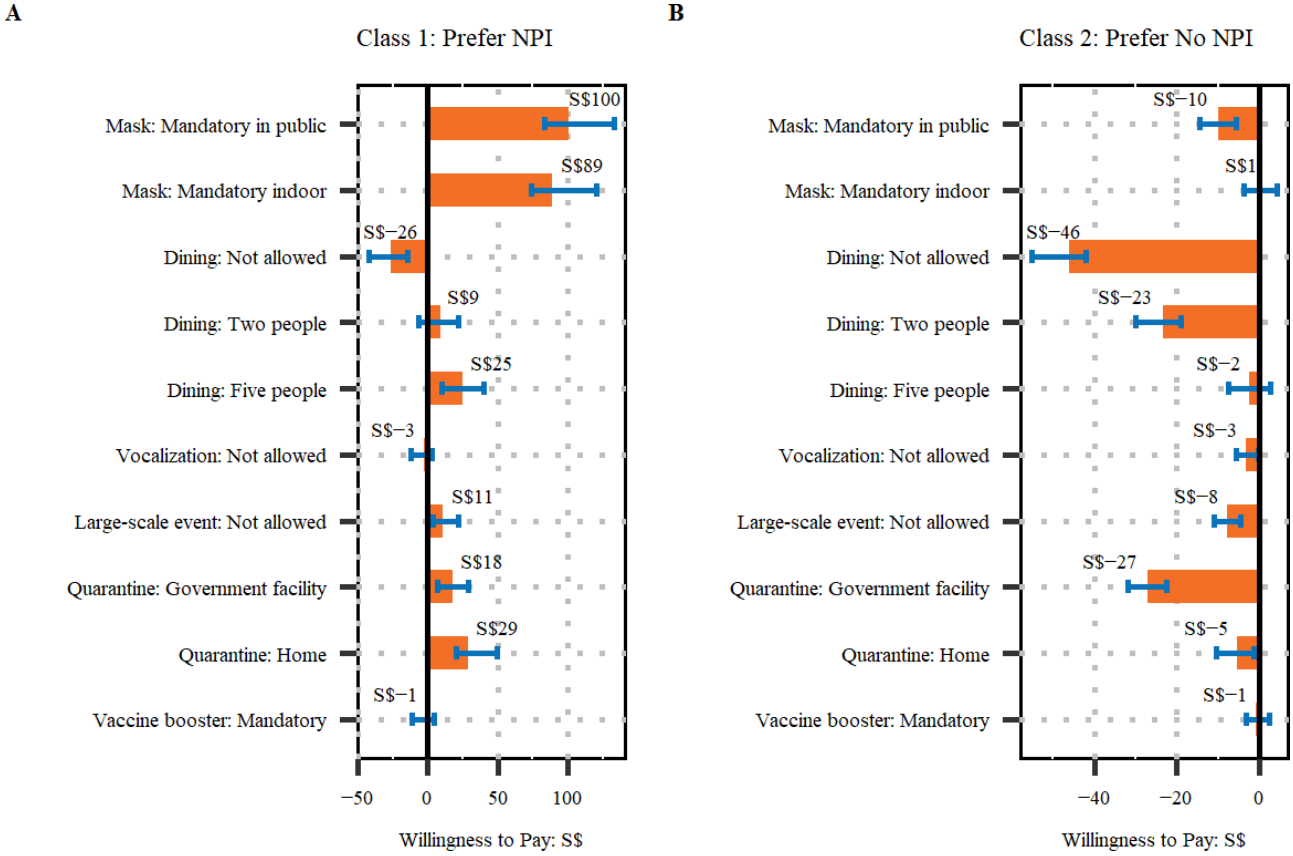


Willingness to Pay

Willingness to pay, measured by the fee for implementing public health measures, is shown in Figure 3A-B. For participants in class 1 (Prefer NPI) as shown in Figure 3A, the top 3 measures they were willing to pay for were mandatory mask wearing in public (SGD $100; SGD $1≈US $0.77), mandatory mask wearing indoors (SGD $89), and home quarantine (SGD $29). They were also willing to pay SGD $26 to avoid a complete prohibition on dining in at F&B establishments. For people in class 2 (Prefer No NPI) as shown in Figure 3B, the top 3 NPIs they were willing to pay to avoid were a complete prohibition on dining in at F&B establishments (SGD $46), quarantine at government facilities (SGD $27), and dining group size restrictions of up to 2 people (SGD $23).

XSL•FO
RenderX

Figure 3. Willingness to pay (WTP) for each nonpharmaceutical intervention (NPI) for (A) class 1 (Prefer NPI) and (B) class 2 (Prefer No NPI). WTP reflects the amount that participants are willing to pay for the implementation of the NPI. A positive WTP indicates that participants were willing to pay to implement the NPI. A negative WTP indicates that participants were willing to pay to avoid the NPI. S$: Singapore dollar.



Discussion

Principal Findings

This study provides valuable insights into public preferences for NPIs during the endemic phase of COVID-19, particularly in the face of a new viral variant. The findings build on existing research, including Lim et al [52], which examined public perceptions and behaviors toward NPIs in Singapore during the early stages of the pandemic through cohort surveys. This study extends these insights by exploring NPI acceptability in a later phase of the pandemic, when pandemic fatigue may have influenced public attitudes. Overall, mandatory mask wearing in public and indoor settings was well received, while strict dining restrictions, quarantine in government facilities, and the suspension of large-scale events faced opposition. A latent class analysis revealed 2 distinct preference groups: those who preferred NPIs (Prefer NPI; 39%), and those who opposed them (Prefer No NPI; 61%). The Prefer NPI group strongly supported mask mandates and only opposed a complete prohibition on dining in at F&B establishments, whereas the Prefer No NPI group exhibited no preference for any NPIs but was not averse to indoor masking and moderate dining restrictions. NPI preferences were influenced by demographic factors such as age, sex, education, employment status, COVID-19 vaccination history, and risk aversion.

A prohibition on dining in at F&B establishments emerged as the least favored NPI among both classes and was the only NPI that faced opposition in class 1 (Prefer NPI). This aligns with the findings by Fink et al [30] on the closure of restaurants and bars, highlighting the social and cultural significance of dining in at such establishments. Singapore's UNESCO-recognized hawker centers, often referred to as "community dining rooms," are central to both social and recreational life [53]. Therefore, strict dining restrictions in regions with strong dining out cultures have proven to be particularly disruptive and unacceptable, especially when imposed for prolonged periods.

Moreover, the suspension of large-scale events was not favored in this study, consistent with the strong opposition reported by Fink et al [30] toward restrictions on social events such as parties, weddings, and concerts. Similarly, Lim et al [52] reported relatively low public support for bans on social gatherings and penalties for noncompliance with social distancing measures, suggesting public resistance to stringent restrictions on social events in Singapore. The similarity in findings suggests that public resistance to the suspension of large-scale events and gatherings in Singapore has remained relatively constant throughout the pandemic, even as different phases of COVID-19 unfolded. In contrast, Eshun-Wilson et al [31] identified the suspension of large gatherings as the most favored NPI. Differences in preferences could be attributed to the role of social and cultural norms, as such restrictions may be more disruptive in regions with strong social and communal event cultures.

Comparatively, mandatory mask wearing, a widely adopted NPI during the pandemic, garnered strong public support across the literature, including in this study [29,32,33]. This finding

XSL•FO
RenderX

aligns with the results from Lim et al [52], where 99.0% of respondents reported wearing masks most or all of the time, even before the official implementation of the mask mandate in Singapore. The widespread support for mandatory mask wearing underscores its recognition as an effective, relatively nondisruptive intervention, which could be easily reintroduced in the event of new epidemic waves. Additionally, Singapore's prior experience with the SARS outbreak likely played a role in fostering this high level of compliance. During SARS, prior mask usage was found to be a positive predictor of support for mask wearing, suggesting that familiarity with such measures contributed to greater public adherence [54]. This historical precedent may have contributed to the high acceptance observed in this study, as Singaporeans were already accustomed to wearing masks as a protective measure.

In addition, this study examined public preferences for a vaccine booster mandate. In contrast to the findings by Yang et al [28], which indicated a preference for universal vaccine boosters among Chinese respondents, respondents in this study exhibited indifference toward a vaccine booster mandate. This divergence in public perceptions reflects the distinct COVID-19 strategies used by each country. At the time of data collection for both studies (mid-2022), China continued to enforce a strict zero-COVID policy, while Singapore had transitioned to an endemic phase of living with COVID-19. This could potentially contribute to a lower perceived need for booster doses among Singaporeans, despite the emergence of new COVID-19 variants.

Overall support for NPIs varied significantly across different demographic groups, particularly by vaccination and employment status. Participants who had received a higher number of vaccine doses generally showed a preference for no NPIs, even in the face of new virus variants. The added protection from additional vaccine doses, combined with vaccine passports that granted vaccinated individuals with greater freedom, may have contributed to a perception of lower susceptibility to new COVID-19 variants. Consequently, this group exhibited a lower perceived need and preference for NPIs. Employment status also played a crucial role in shaping NPI preferences, as full-time employees in Singapore were less likely to support NPIs. This contrasts with the findings by Reed et al [27] in the United States, where salaried employees preferred more stringent NPIs than self-employed individuals. This disparity could be attributed to differences in economic structures, as Singapore had a higher reliance on physical workplaces than the United States, where remote work was more widely adopted. Concerns about job and financial security in Singapore were evident in public discourse even in the early phases of the pandemic, where many voiced anxieties over the potential consequences of the circuit breaker regulations on job stability and work-life balance [55]. To enhance public acceptance of NPIs in future outbreaks, policy makers could consider strategies that alleviate financial strain—such as targeted financial relief and flexible enforcement policies—to balance public health objectives with economic stability.

Beyond these demographic influences, unexplained individual preference heterogeneity was observed in both regression models, particularly for mandatory mask wearing, prohibition on dining in at F&B establishments, quarantine in government facilities, and vaccine booster mandates. This suggests that unobservable personal, experiential, or psychological factors may influence individuals' views on NPIs, underscoring the challenge of designing public health policies that accommodate these diverse preferences. Future studies can explore the factors affecting the preferences for NPIs that exhibit high degree of heterogeneities at an individual level. Adaptive strategies, such as tiered restrictions based on public sentiment and risk assessment, are crucial in maximizing policy acceptability and compliance. This is especially relevant for contentious measures such as strict dining restrictions and quarantine policies, where strong opposition may lead to rapid declines in adherence and effectiveness [56].

Strengths and Limitations

One strength of this study is its exploration of public preferences for less commonly studied NPIs, such as quarantine after travel and dining group size restrictions. These insights can help inform their design and implementation in future pandemic containment strategies in Singapore and beyond. Additionally, the use of actual COVID-19 policy measures implemented in Singapore between 2020 and 2022 as DCE attributes allowed participants to draw from personal experiences when making decisions. This reduced hypothetical bias and led to more informed responses that better reflect real-world preferences [57]. However, since the survey was conducted soon after the lifting of most pandemic restrictions in Singapore, responses may have been influenced by recency bias, with participants forming stronger opinions based on their recent experiences with the NPIs. Furthermore, preferences for NPIs are likely shaped by participants' expectations regarding the severity of a new variant at the time of the survey. Future studies could further explore how perceptions of disease severity influence public support for NPIs.

The generalizability of the study findings to the broader Singapore population may be limited. Participants recruited from the SPHS Online Panel had high English proficiency and digital literacy, potentially excluding individuals with lower proficiency or limited access to digital devices—more common among older adults. However, Singapore's high overall English literacy and internet penetration rates mitigate some of these concerns [58,59]. Nevertheless, some of the demographic differences that exist between the study sample and the general population may still affect generalizability, since demographic factors were shown to influence NPI preferences. Future studies could use multilingual surveys to explore preferences among non–English-speaking older adults and ensure a more representative sample. Beyond demographics, public acceptance of NPIs is also shaped by broader cultural and geographic contexts, as evidenced by significant regional differences in NPI support during the SARS outbreak [54]. This underscores the importance of flexible, context-specific public health strategies rather than a one-size-fits-all approach. Given these factors, findings from this study may be more applicable to high-income Asian settings, where similar characteristics such as dense urban environments, strong institutional trust, and prior epidemic experience may shape NPI preferences differently from Western or lower-income settings.

XSL•FO
RenderX

Implications and Conclusions

The findings from this study are particularly relevant as the world continues to navigate the shifting landscape of infectious disease outbreaks. In the endemic phase of COVID-19 and other infectious diseases, the emergence of new viral variants may necessitate the reintroduction of NPIs as complementary tools to control outbreaks before the widespread uptake of vaccine boosters. A thorough understanding of public preferences is essential for designing policies that not only facilitate the reimplementation of NPIs following new viral outbreaks but also determine their optimal duration to ensure compliance and effectiveness while minimizing societal disruptions. NPIs with broad support, such as indoor mask wearing, can be swiftly reinstated and maintained over extended periods with minimal resistance. However, for NPIs that elicit mixed preferences or strong opposition—such as strict dining restrictions and quarantine in government facilities—a careful cost-benefit analysis is necessary. The use of adaptive strategies, such as phased implementation or targeted enforcement based on risk levels, may help balance public health objectives with societal acceptance. By aligning pandemic containment measures with public sentiment, policy makers can enhance adherence, minimize disruptions, and strengthen long-term resilience against future disease outbreaks.

Acknowledgments

This research was supported by the Ministry of Education, Singapore, under the Academic Research Fund Tier 1 (FY2021) and the SSHSPH Seed Fund for Ground-Up Research 2021. This study was conducted using the Singapore Population Health Studies Online Panel, which was funded through the Singapore Ministry of Health's National Medical Research Council Centre Grant Programme (grant NMRC/CG/C026/2017_NUHS) for Singapore Population Health Improvement Centre (SPHERiC). The authors would like to thank Shumian Yeo and the rest of the Singapore Population Health Studies Online Panel team for supporting the setup, testing, and execution of the survey in this study. They would like to thank Jason Chin Huat Yap, Qian Yang, Xian Yi Ng, and Jacinta I-Pei Chen for their inputs and helping with the user acceptability testing. The authors would also like to thank Hwee Lin Wee for her advice for this study. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the funder.

Data Availability

Data are available upon request from the corresponding author, subject to the approval from National University of Singapore Institutional Review Board.

Authors' Contributions

YW, SHXT, and IYHA were responsible for the conceptualization and design of the study. SHXT and IYHA obtained funding for this study. SHXT and IYHA were responsible for coordinating the recruitment of participants and data collection. HSC was responsible for the review and processing of the data. YW was responsible for the analysis of data. YW and CEH wrote the first draft of the manuscript. All authors were involved in the interpretation of the data and worked on critical revisions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Discrete choice experiment scenario and sample question.
[DOCX File, 76 KB - publichealth_v11i1e67725_app1.docx]

Multimedia Appendix 2
Unexplained individual preference heterogeneity.
[DOCX File, 18 KB - publichealth_v11i1e67725_app2.docx]

Multimedia Appendix 3
Comparison of demographic profile between the sample and Singapore population.
[DOCX File, 22 KB - publichealth_v11i1e67725_app3.docx]

References

1. Marani M, Katul GG, Pan WK, Parolari AJ. Intensity and frequency of extreme novel epidemics. Proc Natl Acad Sci U S A 2021 Aug 31;118(35):e2105482118. [doi: 10.1073/pnas.2105482118] [Medline: 34426498]
2. Thomson MC, Stanberry LR. Climate change and vectorborne diseases. N Engl J Med 2022 Nov 24;387(21):1969-1978. [doi: 10.1056/NEJMra2200092] [Medline: 36416768]

XSL•FO
RenderX

3. Lefrançois T, Malvy D, Atlani-Duault L, et al. After 2 years of the COVID-19 pandemic, translating One Health into action is urgent. Lancet 2023 Mar 4;401(10378):789-794. [doi: 10.1016/S0140-6736(22)01840-2] [Medline: 36302392]
4. Hollingsworth TD, Ferguson NM, Anderson RM. Frequent travelers and rate of spread of epidemics. Emerg Infect Dis 2007 Sep;13(9):1288-1294. [doi: 10.3201/eid1309.070081] [Medline: 18252097]
5. Perra N. Non-pharmaceutical interventions during the COVID-19 pandemic: a review. Phys Rep 2021 May 23;913:1-52. [doi: 10.1016/j.physrep.2021.02.001] [Medline: 33612922]
6. Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/non-pharmaceutical-public-health-measuresfor-mitigating-the-risk-and-impact-of-epidemic-and-pandemic-influenza [accessed 2024-08-26]
7. Imai N, Gaythorpe KAM, Abbott S, et al. Adoption and impact of non-pharmaceutical interventions for COVID-19. Wellcome Open Res 2020;5:59. [doi: 10.12688/wellcomeopenres.15808.1] [Medline: 32529040]
8. Yeoh EK, Chong KC, Chiew CJ, et al. Assessing the impact of non-pharmaceutical interventions on the transmissibility and severity of COVID-19 during the first five months in the Western Pacific Region. One Health 2021 Jun;12:100213. [doi: 10.1016/j.onehlt.2021.100213] [Medline: 33506086]
9. Peters JA, Farhadloo M. The effects of non-pharmaceutical interventions on COVID-19 cases, hospitalizations, and mortality: a systematic literature review and meta-analysis. AJPM Focus 2023 Jun 14;2(4):100125. [doi: 10.1016/j.focus.2023.100125] [Medline: 37362389]
10. Li Y, Campbell H, Kulkarni D, et al. The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. Lancet Infect Dis 2021 Feb;21(2):193-202. [doi: 10.1016/S1473-3099(20)30785-4] [Medline: 33729915]
11. Ullrich A, Schranz M, Rexroth U, et al. Impact of the COVID-19 pandemic and associated non-pharmaceutical interventions on other notifiable infectious diseases in Germany: an analysis of national surveillance data during week 1-2016 - week 32-2020. Lancet Reg Health Eur 2021 Jul;6:100103. [doi: 10.1016/j.lanepe.2021.100103] [Medline: 34557831]
12. Chang T, Cho SI, Yoo DS, Min KD. Trends in nationally notifiable infectious diseases in humans and animals during COVID-19 pandemic, South Korea. Emerg Infect Dis 2024 Jun;30(6):1154-1163. [doi: 10.3201/eid3006.231422] [Medline: 38781924]
13. Geng MJ, Zhang HY, Yu LJ, et al. Changes in notifiable infectious disease incidence in China during the COVID-19 pandemic. Nat Commun 2021 Nov 26;12(1):6923. [doi: 10.1038/s41467-021-27292-7] [Medline: 34836947]
14. Pedrosa AL, Bitencourt L, Fróes ACF, et al. Emotional, behavioral, and psychological impact of the COVID-19 pandemic. Front Psychol 2020;11:566212. [doi: 10.3389/fpsyg.2020.566212] [Medline: 33117234]
15. Brooks SK, Webster RK, Smith LE, et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. Lancet 2020 Mar 14;395(10227):912-920. [doi: 10.1016/S0140-6736(20)30460-8] [Medline: 32112714]
16. Seale H, Dyer CEF, Abdi I, et al. Improving the impact of non-pharmaceutical interventions during COVID-19: examining the factors that influence engagement and the impact on individuals. BMC Infect Dis 2020 Aug 17;20(1):607. [doi: 10.1186/s12879-020-05340-9] [Medline: 32807087]
17. Williams G, Kinchin I. The application of discrete choice experiments eliciting young peoples' preferences for healthcare: a systematic literature review. Eur J Health Econ 2023 Aug;24(6):987-998. [doi: 10.1007/s10198-022-01528-9] [Medline: 36169764]
18. Wang Y, Faradiba D, Del Rio Vilas VJ, et al. The relative importance of vulnerability and efficiency in COVID-19 contact tracing programmes: a discrete choice experiment. Int J Public Health 2022;67:1604958. [doi: 10.3389/ijph.2022.1604958] [Medline: 35936996]
19. Ong C, Cook AR, Tan KK, Wang Y. Advancing colorectal cancer detection with blood-based tests: qualitative study and discrete choice experiment to elicit population preferences. JMIR Public Health Surveill 2024 Jul 17;10(1):e53200. [doi: 10.2196/53200] [Medline: 39018093]
20. Ang IYH, Wang Y, Tyagi S, Koh GCH, Cook AR. Preferences and willingness-to-pay for a blood pressure telemonitoring program using a discrete choice experiment. NPJ Digit Med 2023 Sep 25;6(1):176. [doi: 10.1038/s41746-023-00919-3] [Medline: 37749387]
21. Manipis K, Street D, Cronin P, Viney R, Goodall S. Exploring the trade-off between economic and health outcomes during a pandemic: a discrete choice experiment of lockdown policies in Australia. Patient 2021 May;14(3):359-371. [doi: 10.1007/s40271-021-00503-5] [Medline: 33694076]
22. Mouter N, Geijsen T, Munyasya A, et al. Preferences for the societal impacts of a pandemic when it transitions into an endemic: a discrete choice experiment. Patient 2025 Jan;18(1):49-63. [doi: 10.1007/s40271-024-00701-x] [Medline: 38980645]
23. Loría-Rebolledo LE, Ryan M, Watson V, et al. Public acceptability of non-pharmaceutical interventions to control a pandemic in the UK: a discrete choice experiment. BMJ Open 2022 Mar 8;12(3):e054155. [doi: 10.1136/bmjopen-2021-054155] [Medline: 35260455]

XSL•FO
RenderX

24. Chorus C, Sandorf ED, Mouter N. Diabolical dilemmas of COVID-19: an empirical study into Dutch society's trade-offs between health impacts and other effects of the lockdown. PLoS One 2020;15(9):e0238683. [doi: 10.1371/journal.pone.0238683] [Medline: 32936815]
25. Filipe L, de Almeida SV, Costa E, da Costa JG, Lopes FV, Santos JV. Trade-offs during the COVID-19 pandemic: a discrete choice experiment about policy preferences in Portugal. PLoS One 2022 Dec 16;17(12):e0278526. [doi: 10.1371/journal.pone.0278526] [Medline: 36525428]
26. Johnson CA, Tran DN, Mwangi A, et al. Incorporating respondent-driven sampling into web-based discrete choice experiments: preferences for COVID-19 mitigation measures. Health Serv Outcomes Res Methodol 2022;22(3):297-316. [doi: 10.1007/s10742-021-00266-4] [Medline: 35035272]
27. Reed S, Gonzalez JM, Johnson FR. Willingness to accept trade-offs among COVID-19 cases, social-distancing restrictions, and economic impact: a nationwide US study. Value Health 2020 Nov;23(11):1438-1443. [doi: 10.1016/j.jval.2020.07.003]
28. Yang M, He Z, Zhang Y, Liu T, Ming WK. Pandemic fatigue and preferences for COVID-19 public health and social measures in China: nationwide discrete choice experiment. JMIR Public Health Surveill 2024 Jun 27;10(1):e45840. [doi: 10.2196/45840] [Medline: 38935420]
29. Li L, Long D, Rouhi Rad M, Sloggy MR. Stay-at-home orders and the willingness to stay home during the COVID-19 pandemic: a stated-preference discrete choice experiment. PLoS One 2021 Jul 1;16(7):e0253910. [doi: 10.1371/journal.pone.0253910]
30. Fink G, Förtsch K, Felder S. Subjective burden of government-imposed Covid-19 restrictions in Switzerland: evidence from the 2022 LINK Covid-19 survey. PLoS One 2023;18(7):e0283524. [doi: 10.1371/journal.pone.0283524] [Medline: 37498827]
31. Eshun-Wilson I, Mody A, McKay V, et al. Public preferences for social distancing policy measures to mitigate the spread of COVID-19 in Missouri. JAMA Netw Open 2021 Jul 1;4(7):e2116113. [doi: 10.1001/jamanetworkopen.2021.16113] [Medline: 34236410]
32. Blayac T, Dubois D, Duchêne S, Nguyen-Van P, Ventelou B, Willinger M. What drives the acceptability of restrictive health policies: an experimental assessment of individual preferences for anti-COVID 19 strategies. Econ Model 2022 Nov;116:106047. [doi: 10.1016/j.econmod.2022.106047] [Medline: 36118956]
33. Mühlbacher AC, Sadler A, Jordan Y. Population preferences for non-pharmaceutical interventions to control the SARS-CoV-2 pandemic: trade-offs among public health, individual rights, and economics. Eur J Health Econ 2022 Dec;23(9):1483-1496. [doi: 10.1007/s10198-022-01438-w] [Medline: 35138495]
34. Petherick A, Goldszmidt R, Andrade EB, et al. A worldwide assessment of changes in adherence to COVID-19 protective behaviours and hypothesized pandemic fatigue. Nat Hum Behav 2021;5(9):1145-1160. [doi: 10.1038/s41562-021-01181-x]
35. Crane MA, Shermock KM, Omer SB, Romley JA. Change in reported adherence to nonpharmaceutical interventions during the COVID-19 pandemic, April-November 2020. JAMA 2021 Mar 2;325(9):883-885. [doi: 10.1001/jama.2021.0286] [Medline: 33480971]
36. Stricter safe distancing measures to prevent further spread of COVID-19 cases. Ministry of Health Singapore. 2023. URL: https://www.moh.gov.sg/newsroom/stricter-safe-distancing-measures-to-prevent-further-spread-of-covid-19-cases [accessed 2024-08-29]
37. askST: what is a stay-home notice and how is it enforced in Singapore? Straits Times Singapore. 2021 Jan 30. URL: https://www.straitstimes.com/singapore/what-is-shn-and-how-is-it-enforced [accessed 2025-03-17]
38. Here's how Singapore's COVID-19 measures will change from Mar 29. CNA. 2022. URL: https://www.channelnewsasia.com/singapore/heres-how-singapores-covid-19-measures-will-change-mar-29-2590851 [accessed 2025-03-17]
39. 10 things to know about eased COVID-19 rules as Singapore hits "major milestone". CNA. 2022. URL: https://www.channelnewsasia.com/singapore/covid-19-safe-distancing-singapore-eases-rules-group-size-tracetogether-2641406 [accessed 2025-03-17]
40. S'pore suspends dining in, tightens Covid-19 curbs: what you need to know from July 22. Straits Times Singapore. 2021 Jul 20. URL: https://www.straitstimes.com/singapore/health/spore-returns-to-phase-2-what-you-need-to-know-about-tightened-covid-19-rules-from [accessed 2025-03-17]
41. Essers BAB, van Helvoort-Postulart D, Prins MH, Neumann M, Dirksen CD. Does the inclusion of a cost attribute result in different preferences for the surgical treatment of primary basal cell carcinoma?: a comparison of two discrete-choice experiments. Pharmacoeconomics 2010;28(6):507-520. [doi: 10.2165/11532240-000000000-00000] [Medline: 20387912]
42. Lighthouse Studio. Sawtooth Software. URL: https://sawtoothsoftware.com/lighthouse-studio [accessed 2025-03-18]
43. Huls SPI, van Osch SMC, Brouwer WBF, van Exel J, Stiggelbout AM. Psychometric evaluation of the Health-Risk Attitude Scale (HRAS-13): assessing the reliability, dimensionality and validity in the general population and a patient population. Psychol Health 2022 Jan;37(1):34-50. [doi: 10.1080/08870446.2020.1851689] [Medline: 33259250]
44. Assele SY, Meulders M, Vandebroek M. Sample size selection for discrete choice experiments using design features. J Choice Model 2023;49:100436. [doi: 10.1016/j.jocm.2023.100436]
45. Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 2008;26(8):661-677. [doi: 10.2165/00019053-200826080-00004] [Medline: 18620460]

XSL•FO
RenderX

46. Lim DYZ, Wong TH, Feng M, Ong MEH, Ho AFW. Leveraging open data to reconstruct the Singapore Housing Index and other building-level markers of socioeconomic status for health services research. Int J Equity Health 2021 Oct 3;20(1):218. [doi: 10.1186/s12939-021-01554-8] [Medline: 34602083]
47. Malhotra R, Malhotra C, Chan A, Østbye T. Life-course socioeconomic status and obesity among older Singaporean Chinese men and women. J Gerontol B Psychol Sci Soc Sci 2013 Jan;68(1):117-127. [doi: 10.1093/geronb/gbs102] [Medline: 23161348]
48. Sarrias M, Daziano R. Multinomial logit models with continuous and discrete individual heterogeneity in R: the gmnl package. J Stat Softw 2017 Jul 13;79:1-46. [doi: 10.18637/jss.v079.i02]
49. McFadden D, Train K. Mixed MNL models for discrete response. J Appl Econ 2000 Sep;15(5):447-470. [doi: 10.1002/1099-1255(200009/10)15:5<447::AID-JAE570>3.0.CO;2-1]
50. Veldwijk J, Marceta SM, Swait JD, Lipman SA, de Bekker-Grob EW. Taking the Shortcut: Simplifying Heuristics in Discrete Choice Experiments. Patient 2023 Jul;16(4):301-315. [doi: 10.1007/s40271-023-00625-y] [Medline: 37129803]
51. R: a language and environment for statistical computing. R Core Team. 2013. URL: http://www.R-project.org [accessed 2024-09-12]
52. Lim VW, Lim RL, Tan YR, et al. Government trust, perceptions of COVID-19 and behaviour change: cohort surveys, Singapore. Bull World Health Organ 2021 Feb 1;99(2):92-101. [doi: 10.2471/BLT.20.269142] [Medline: 33551503]
53. Radomskaya V, Bhati AS. Hawker centres: a social space approach to promoting community wellbeing. UP 2022 Oct 27;7(4) [FREE Full text] [doi: 10.17645/up.v7i4.5658]
54. Matthews Pillemer F, Blendon RJ, Zaslavsky AM, Lee BY. Predicting support for non-pharmaceutical interventions during infectious outbreaks: a four region analysis. Disasters 2015 Jan;39(1):125-145. [doi: 10.1111/disa.12089] [Medline: 25243477]
55. Shorey S, Ang E, Yamina A, Tam C. Perceptions of public on the COVID-19 outbreak in Singapore: a qualitative content analysis. J Public Health (Oxf) 2020 Nov 23;42(4):665-671. [doi: 10.1093/pubmed/fdaa105]
56. Delussu F, Tizzoni M, Gauvin L. Evidence of pandemic fatigue associated with stricter tiered COVID-19 restrictions. PLOS Digit Health 2022 May;1(5):e0000035. [doi: 10.1371/journal.pdig.0000035] [Medline: 36812519]
57. Haghani M, Bliemer MCJ, Rose JM, Oppewal H, Lancsar E. Hypothetical bias in stated choice experiments: part II. Conceptualisation of external validity, sources and explanations of bias and effectiveness of mitigation methods. J Choice Model 2021 Dec;41:100322. [doi: 10.1016/j.jocm.2021.100322]
58. Census of Population 2020. Statistical Release 1, Demographic Characteristics, Education, Language and Religion: Singapore: Dept. of Statistics, Ministry of Trade and Industry; 2021.
59. Romero L. Internet usage in Singapore. Statista. 2024. URL: https://www.statista.com/topics/5852/internet-usage-in-singapore/ [accessed 2025-03-17]

Abbreviations

DCE: discrete choice experiment
F&B: food and beverage
HRAS-13: Health Risk Attitude Scale
IRB: institutional review board
MMML: mixed-mixed multinomial logit
MXL: mixed-logit
NPI: nonpharmaceutical intervention
NUS: National University of Singapore
REDCap: Research Electronic Data Capture
SPHS: Singapore Population Health Studies
WTP: willingness to pay

Edited by A Nakhostin-Ansari; submitted 20.10.24; peer-reviewed by J Mendes, S Bhatia; revised version received 21.03.25; accepted 28.03.25; published 04.06.25.

Please cite as:
Wang Y, Har CE, Tan SHX, Cheng HS, Ang IYH
Preferences for Nonpharmaceutical Interventions During the Endemic Phase of COVID-19: Discrete Choice Experiment
JMIR Public Health Surveill 2025;11:e67725
URL: https://publichealth.jmir.org/2025/1/e67725
doi:10.2196/67725

XSL•FO
RenderX

© Yi Wang, Chee Ern Har, Sharon Hui Xuan Tan, Hooi Swang Cheng, Ian Yi Han Ang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 4.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Success Factors for Optimizing COVID-19 Vaccine Uptake Among Indigenous Populations in Taiwan: Cross-Sectional Questionnaire Survey

Jia-Chi Tu Abùs Isbabanal[1,2], MD; Ray-E Chang[3], PhD; Liam O'Neill[4], PhD; Ting-Hung Chen[3], MS; Po-Han Lee[3], PhD; Tsan-Teng Ou[5], MD

[1]Namasia District Public Health Center, Department of Health, Kaohsiung City Government, Kaohsiung, Taiwan

[2]Department of Family Medicine, Saint Paul's Hospital, Taoyuan, Taiwan

[3]Institute of Health Policy and Management, College of Public Health, National Taiwan University, No. 17, Xu-Zhou Rd., Rm. 639, Zhongzheng District, Taipei, Taiwan

[4]Department of Rehabilitation and Health Services, College of Health and Public Service, University of North Texas, Denton, TX, United States

[5]Center of Indigenous Health Care, Department of Community Health, Kaohsiung Medical University Chung-Ho Memorial Hospital, Kaohsiung, Taiwan

Corresponding Author:
Ray-E Chang, PhD
Institute of Health Policy and Management, College of Public Health, National Taiwan University, No. 17, Xu-Zhou Rd., Rm. 639, Zhongzheng District, Taipei, Taiwan

Abstract

Background: The COVID-19 pandemic has devastated economies and strained health care systems worldwide. Vaccination is crucial for outbreak control, but disparities persist between and within countries. In Taiwan, certain indigenous regions show lower vaccination rates, prompting comprehensive inquiries.

Objective: This study aims to identify predictors for COVID-19 vaccination and develop strategies for indigenous communities.

Methods: This cross-sectional study, conducted from May 13 to July 18, 2022, surveyed indigenous community members older than 55 years residing in a mountain area in southern Taiwan. Based on the health belief model, the questionnaire covered sociodemographic factors, health-related issues, and trust in physicians. The analysis included bivariate analysis, logistic regression, and mediation analysis.

Results: Most participants (N=203) were aged 55 - 64 years (102/203, 50.2%), female (129/203, 63.5%), married (104/203, 51.2%), with low education (165/203, 81.3%), and engaged in agriculture (79/203, 38.9%) or were unemployed (104/203, 51.2%). Logistic regression revealed that unvaccinated individuals were significantly more likely to perceive lower COVID-19 threats (P=.03), fewer vaccination benefits (P=.04), higher barriers to vaccination (P=.02), and weaker responses to external cues to action (P<.001), while no significant differences were observed in trust in physicians. Mediation analyses further indicated that trust in physicians influenced vaccine uptake indirectly through perceived barriers. The indirect effect was statistically significant (95% bootstrap CI 0.013 to 0.437), suggesting a full mediation effect.

Conclusions: Effective pandemic prevention strategies for indigenous communities should be grounded in a nuanced understanding of local needs and incorporate bottom-up approaches to avoid cultural saturation and the exacerbation of existing health disparities. Ensuring the accuracy and clarity of vaccine-related information received by indigenous older adults is essential. Local health authorities should consider deploying health care professionals to engage directly with indigenous older adults and their caregivers, delivering culturally appropriate and evidence-based information to address concerns regarding vaccine safety and perceived risks. Such efforts are critical to strengthening vaccine confidence and increasing vaccination uptake in these communities.

(JMIR Public Health Surveill 2025;11:e75278) doi:10.2196/75278

KEYWORDS

COVID-19; indigenous; health belief model; vaccine hesitancy; health inequity

Introduction

A growing body of research has focused on the impact of the COVID-19 pandemic on aboriginal and indigenous populations throughout the world. Much of this research has explored ways to increase vaccination rates and access to health services among indigenous populations in various countries, such as Australia, Brazil, Guatemala, India, and the United States [1-5]. Globally, due to a legacy of systemic racism and marginalization,

XSL•FO
RenderX

indigenous peoples have a greater burden of disease than nonindigenous peoples. For example, indigenous peoples have had higher rates of cardiovascular disease, diabetes, and infectious diseases, and this has resulted in a lower life expectancy [6]. In terms of social determinants, their health is impacted by numerous chronic problems, such as lower educational attainment, geographic isolation, lack of access to culturally appropriate health care, and persistent poverty.

In the absence of a coordinated program that is targeted toward indigenous peoples, they are likely to have higher rates of infection, serious complications, and mortality due to COVID-19 [2,5,7]. Vaccination is the most effective strategy to protect the population from adverse outcomes due to the pandemic. However, indigenous peoples face numerous barriers to getting vaccinated. These may include: geographic isolation, lack of access to health care, cultural barriers (such as language), and vaccine hesitancy [6].

Taiwan's indigenous populations are Austronesian people who are officially recognized as 16 distinct ethnic groups. They comprise approximately 569,000 indigenous people, accounting for about 2% of the national population. In 2021, Taiwan's Ministry of Health and Welfare launched a nationwide COVID-19 vaccination program aimed at optimizing vaccine uptake, including targeted efforts for the indigenous population. As part of the vaccine rollout, priority groups were identified based on risk profile. Notably, indigenous persons aged 55 - 64 years were granted early access to vaccination compared with their nonindigenous counterparts in the same age cohort. To promote vaccination uptake, the government implemented a broad range of outreach strategies, including mass media campaigns, social media promotion, and incentive-based approaches such as the distribution of household goods or gift vouchers. Health education materials, such as informational videos and visual aids, were disseminated in multiple languages, with Mandarin and Taiwanese serving as the primary languages, to enhance accessibility and understanding.

The health belief model (HBM) has proven to be an excellent conceptual framework, not only to explore the motives of persons who get vaccinated but also to better understand the reasons for refusing vaccination [8]. The main premise of the HBM model is that existing beliefs and attitudes can predict future behavior. HBM includes 5 major constructs: perceived susceptibility, perceived severity, perceived benefits, perceived barriers, and cues to action. Therefore, this study uses the HBM model and sociodemographic and health-related factors to identify beliefs and attitudes that can predict vaccine uptake among indigenous peoples. The goal is to identify policies and success factors so that public health agencies will be better prepared to mitigate the potential harm to indigenous peoples due to future epidemics.

This article describes the planning, implementation, outcomes, and policy implications of this program, as it was experienced by one indigenous community in a remote, mountainous area of Taiwan. We identify the various challenges encountered and how these were overcome through local, bottom-up efforts. The HBM model was used to better understand the underlying factors regarding the decision to accept or reject the COVID-19 vaccine.

This research was conducted in a remote mountain region (253 km²) officially designated as an indigenous region in southern Taiwan. This isolated mountainous region comprises 3 rural villages with a total population of 3143 indigenous residents, the majority of whom belong to the Bunun people (approximately 70%), alongside smaller populations of Kanakanavu and Paiwan peoples. Health care services in the region are limited to a single public health center and a clinic. Residents requiring hospitalization or specialist care must travel up to 2 hours to reach the nearest regional referral hospital. To promote COVID-19 vaccination, the local government implements the standard nationwide promotional strategies. However, recognizing language and culture barriers, the public health center independently produced vaccination videos in Kanakanavu, Bunun, and Hla'alua languages, despite the absence of official materials in indigenous languages from the central government. These videos were disseminated through "Line," Taiwan's most widely used social media platform, significantly enhancing public awareness and understanding of COVID-19 and the importance of vaccination within the community.

Most previous studies relied on vaccination intention as a proxy for actual vaccine uptake [9]. However, there may be a discrepancy between what people intend to do and what they actually do, namely, receive the vaccine. For example, Harris et al [10] found that only about half of ''intending'' recipients of seasonal influenza vaccination actually got vaccinated. Intention to vaccinate may change over time and may be subject to societal norms, such as a marketing campaign to promote vaccinations [11,12]. To address this gap, this study uses actual vaccine uptake rather than vaccine intention. Survey responses were validated by checking the participants' vaccine status with a national database of immunization records.

Methods

This study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines for cross-sectional research (Checklist 1).

Study Design and Participants

This study is an observational, cross-sectional interview design. Grounded in the HBM framework, the study framework incorporated sociodemographic characteristics, health-related factors, and trust in physicians. The primary outcome of interest was actual COVID-19 vaccine uptake, as verified through the national immunization registry. Explanatory variables included: (1) sociodemographic and health-related factors; (2) perceived threats; (3) perceived benefits; (4) perceived barriers; (5) cues to action, defined as external prompts encouraging vaccination; and (6) trust in physicians. Previous research has demonstrated that trust in health care providers can influence vaccination intention or behavior both directly and indirectly through health belief constructs [13,14]. Accordingly, this study also hypothesizes potential pathways from trust in physicians to perceived threats, perceived benefits, and perceived barriers.

Indigenous persons were eligible to be included in the study if they lived in the tribal community for at least 6 months and

XSL•FO
RenderX

were 55 years or older, as this group was prioritized to receive the vaccine. (Note that persons younger than 55 years were not eligible to receive the vaccine at this time. Public health measures were developed to address disparities in life expectancy and socioeconomic and sociocultural conditions between indigenous populations and nonindigenous populations. Therefore, indigenous people who are 55 years and older are eligible to receive all the rights and benefits granted to nonindigenous persons who are 65 years and older in Taiwan. Persons who could not understand the survey questions were excluded, such as those with a mental disorder or other conditions with whom the team did not have the ability to communicate.

Participants were recruited from an indigenous region in southern Taiwan through local health stations, clinics, churches, community offices, and health events. Researchers initially consulted participants verbally regarding their willingness to participate, followed by an on-site explanation of the study and the collection of written informed consent. Data were collected using a structured, face-to-face questionnaire administered by trained interviewers, with each session lasting approximately 20 minutes. Although the majority of interviews were conducted in person, some were carried out by telephone due to infection control measures during the COVID-19 pandemic or based on participants' preferences. When necessary, certified interpreters were available to assist participants who required support in understanding the questionnaire in their indigenous languages, ensuring accurate comprehension. In addition to the formal survey, researchers also conducted informal follow-up conversations to gain deeper insights into participants' reasons for vaccine hesitancy or nonuptake.

Survey Instrument

In this study, the survey instrument of the HBM (including constructs such as perceived threat, benefits, barriers, and cause to action) was primarily adapted from the questionnaire developed by Huang et al [15], which focused on influenza vaccination among Taiwanese individuals. To contextualize the instrument for COVID-19, additional modifications were made with reference to studies by Hossain et al [16], Walrave et al [17], and Shmueli [18]. Items measuring trust in physicians were informed by Hall et al [19] and subsequently adapted to align with the objectives of this study. To enhance content validity and cultural appropriateness, the survey was reviewed and revised by 2 public health experts, as well as 2 attending physicians with over 20 years of experience delivering frontline medical care in indigenous communities. Their feedback helped refine the wording to improve clarity and comprehension for study participants.

The survey instrument consisted of a total of 40 questions structured into 6 sections. The first section, sociodemographic and health-related factors, included the following items: age, gender, marital status, educational level, occupation, monthly income, welfare status, chronic disease or major injury, obesity, tobacco use, alcohol use, flu vaccination during the past year, and self-reported health status. The survey also included items corresponding to the following constructs: (1) perceived threats, including susceptibility and severity; (2) perceived benefits; (3) perceived barriers; (4) cues to action; (5) trust in physicians; and (6) actions to receive the COVID-19 vaccine. The specific survey items are presented in Textbox 1. The construct "Actions to Receive the COVID-19 Vaccine" was measured as a binary variable, indicating whether the participant had received at least one dose of the vaccine (0=unvaccinated; 1=at least one dose). All other constructs were assessed using a 5-point Likert Scale, ranging from 1 ("strongly disagree") to 5 ("strongly agree"), to evaluate participants' beliefs and attitudes.

XSL•FO
RenderX

Textbox 1. Items of the health belief model and physician trust constructs.

Perceived Threats

1. I am worried about the likelihood of getting infected by COVID-19.
2. I am at high risk of COVID-19 because of my health conditions.
3. If I were infected with the COVID-19 virus, it would have important health consequences for me.
4. If I were infected with the COVID-19 virus, my health would be severely affected.
5. If I were infected with the COVID-19 virus, my health would be significantly reduced.

Perceived Benefits

1. I think vaccination is good because it will make me less worried about COVID-19.
2. I believe vaccination will decrease my risk of getting infected with COVID-19.
3. I think the complications of COVID-19 will decrease if I get vaccinated and then get infected with the coronavirus.

Perceived Barriers

1. I am worried that the possible side effects of the COVID-19 vaccination would interfere with my usual activities.
2. I am concerned about the efficacy of the COVID-19 vaccine.
3. I have a concern that I may receive a faulty/fake COVID-19 vaccine.
4. It concerns me that the development of a COVID-19 vaccine is too rushed to test its safety properly.
5. I am concerned about the long-term side effects of the COVID-19 vaccination.

Cues To Action

1. The chances of me getting vaccinated against COVID-19 will increase if opinion leaders on social media express support for the benefits of the vaccine.
2. The chances of me getting vaccinated against COVID-19 will increase if friends and family express support for the benefit of the vaccine.
3. If my workplace takes care of vaccinating the workers against COVID-19, I will get vaccinated.
4. The chances of me getting vaccinated against COVID-19 will increase if workers in the field of religion (such as priests, nuns, etc) express support for the benefit of the vaccine.
5. The chances of me getting vaccinated against COVID-19 will increase if local ethnic opinion leaders (such as tribal elders, leaders, respected elders, etc) express support for the benefit of the vaccine.
6. The chances of me getting vaccinated against COVID-19 will increase if local officials (such as village officials, district head, etc) express support for the benefit of the vaccine.
7. Rewards (such as gifts or vouchers) for vaccinating against COVID-19 will increase my willingness to get vaccinated against COVID-19.

Trust in Physicians

1. You completely trust public health physicians' decisions about which medical treatments are best.
2. Public health physicians are totally honest in telling their patients about all of the treatment options available for their conditions.
3. Public health physicians always use their very best skills and effort on behalf of their patients.
4. All in all, I trust public health physicians completely.

Actions to receive the COVID-19 vaccine

1. I had received at least one dose of the COVID-19 vaccine (Yes/No)

Validity and Reliability

To evaluate the instrument's validity, expert review and exploratory factor analysis (EFA) were used. EFA was conducted using principal axis factoring to explore the underlying factor structure of the questionnaire items. Given the potential correlation among factors, direct oblimin rotation was applied. Sampling adequacy was confirmed by the Kaiser-Meyer-Olkin measure and Bartlett test of sphericity. The Kaiser-Meyer-Olkin value was 0.82, indicating excellent sampling adequacy, and the Bartlett test was significant (χ^2_{253}=4,624; P<.001), supporting the factorability of the correlation matrix. Five components were extracted from 23 items, explaining a total of 79.43% of the variance. After oblique rotation, the variance was distributed across the following components. The first component, cause to action, consisted of 6 items and accounted for 27.48% of the total variance. The

XSL•FO
RenderX

second component, perceived threats, comprised 5 items and explained 17.27% of the variance. The third component, Trust in Physicians, comprised 4 items and accounted for 18.82% of the variance. The fourth component, perceived barriers, included 5 items and explained 18.16% of the variance. The fifth component, perceived benefits, consisted of 3 items and accounted for 18.28% of the variance. All items demonstrated better factor loadings and converged into 5 dimensions as expected. Reliability was evaluated using Cronbach α. Cronbach α was used to verify internal consistency. All subscales demonstrated high internal consistency, with α coefficients exceeding 0.84.

Statistical Analysis

The questionnaire data were imported into SPSS software (version 25; IBM Corp) and identified by codes. Reliability and validity were examined using Cronbach α and EFA. To characterize the study population, descriptive analytics were revealed by frequency, percentage, mean, and SD. Bivariate analysis, including Pearson chi-square test and independent sample *t* tests, was used to analyze the correlation between dependent and independent variables. Finally, multivariate binary logistic regression was performed in order to investigate determinants of COVID-19 vaccination; only variables found to be significantly ($P<.05$) associated with COVID-19 vaccination in bivariate analysis were included in the regression. Mediation analysis was performed using PROCESS version 4.2 for SPSS [20], using model 4 to examine the indirect effect of perceived threat, perceived benefits, and perceived barriers on the association between trust in physicians and COVID-19 vaccine uptake. Bootstrapping with 5000 resamples was used to estimate the 95% CIs for both the direct and indirect effects. An indirect effect was considered statistically significant if the CI did not include 0.

Ethical Considerations

All procedures performed in studies involving human participants were in accordance with the ethical standards of the Institutional Review Board of the Kaohsiung Medical University Chung-Ho Memorial Hospital (serial number: KMUHIRB-SV(I)-20220004) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Out of serious ethical considerations, we also convened consultation meetings about cultural risk assessment by the Council of Indigenous Peoples and received endorsement and approval of the community assembly in the local community (serial number: CRB-111 - 005). Participants were recruited in person by the research team, who provided an oral explanation of the study's objectives and procedures. Written informed consent was obtained from each participant prior to the commencement of the survey. All data collected in this study were anonymized and securely stored, with no personally identifiable information retained, to ensure participant confidentiality.

Results

Participants

Table 1 provides descriptive characteristics of the indigenous participants. A total of 203 participants completed the interview. The majority of those included were aged around 55 - 64 years (102/203, 50.2%), female (129/203, 63.5%), married (104/203, 51.2%), and had an educational level of junior high school or below (165/203, 81.3%). Most of the participants were unemployed (104/203, 51.2%) or engaged in agricultural work (79/203, 38.9%), with monthly available money below NT $9999 (average exchange rate at the time of study was US $1=NT $29.75) (154/203, 75.9%), much less than the minimum wage of the same year (NT $25,250). Most of the participants had chronic diseases (173/203, 85.2%), and about half of the participants had not been vaccinated against influenza in the past year (104/203, 51.2%). Persons with at least a high school education were more likely to get vaccinated than those who did not complete high school (97.4% vs 81.8%; $P=.02$). Persons who abstained from alcohol were more likely to get vaccinated than those who drank alcohol at least once a month (89.8% vs 78.9%; $P=.03$).

XSL•FO
RenderX

Table . Descriptive characteristics of indigenous participants (N=203).

Demographic	Received one or more doses of the COVID-19 vaccine			
	No (n=31)	Yes (n=172)	Total (N=203)	*P* value
Age (years), n (%)				.83[a]
55 - 64	17 (54.8)	85 (49.4)	102 (50.2)	
65 - 74	10 (32.3)	65 (37.8)	75 (36.9)	
75 and older	4 (12.9)	22 (12.8)	26 (12.8)	
Sex, n (%)				.78[a]
Male	12 (38.7)	62 (36.0)	74 (36.5)	
Female	19 (61.3)	110 (64.0)	129 (63.5)	
Marriage, n (%)				.002[a, b]
Unmarried (including divorced and widowed)	23 (74.2)	76 (44.2)	99 (48.8)	
Married	8 (25.8)	96 (55.8)	104 (51.2)	
Education level, n (%)				.02[a, b]
Junior high school or below	30 (96.8)	135 (78.5)	165 (81.3)	
High school/high vocational school or above	1 (3.2)	37 (21.5)	38 (18.7)	
Occupation, n (%)				.64[a]
Part-time job	0 (0.0)	7 (4.1)	7 (3.4)	
Public servant	0 (0.0)	1 (0.6)	1 (0.5)	
Agriculture-related	13 (41.9)	66 (38.4)	79 (38.9)	
Business-related	0 (0.0)	5 (2.9)	5 (2.5)	
None	16 (51.6)	88 (51.2)	104 (51.2)	
Others	2 (6.5)	5 (2.9)	7 (3.4)	
Disposable income (NT $)[c], n (%)				.26[a]
0 - 9999	26 (83.9)	128 (74.4)	154 (75.9)	
10,000 and above	5 (16.1)	44 (25.6)	49 (24.1)	
Welfare status, n (%)				.17[a]
No	31 (100.0)	162 (94.2)	193 (95.1)	
Yes	0 (0.0)	10 (5.8)	10 (4.9)	
Living status, n (%)				.06[a]
Living alone	10 (32.3)	31 (18.0)	41 (20.2)	
Living with relatives	21 (67.7)	141 (82.0)	162 (79.8)	
Chronic disease, n (%)				.39[a]
No	3 (9.7)	27 (15.7)	30 (14.8)	
Yes	28 (90.3)	145 (84.3)	173 (85.2)	
Diagnosis of serious injury or illness, n (%)				.50[a]
No	25 (80.6)	129 (75.0)	154 (75.9)	
Yes	6 (19.4)	43 (25.0)	49 (24.1)	
Obesity, n (%)				.29[a]

XSL•FO
RenderX

Demographic	Received one or more doses of the COVID-19 vaccine			
	No (n=31)	Yes (n=172)	Total (N=203)	*P* value
No	22 (71.0)	105 (61.0)	127 (62.6)	
Yes	9 (29.0)	67 (39.0)	76 (37.4)	
Smoking, n (%)				.11[a]
No	20 (64.5)	134 (77.9)	154 (75.9)	
Yes	11 (35.5)	38 (22.1)	49 (24.1)	
Alcohol consumption, n (%)				.03[a, b]
No	11 (35.5)	97 (56.4)	108 (53.2)	
Yes	20 (64.5)	75 (43.6)	95 (46.8)	
The habit of chewing betel nut, n (%)				.03[a, b]
No	24 (77.4)	156 (90.7)	180 (88.7)	
Yes	7 (22.6)	16 (9.3)	23 (11.3)	
Influenza vaccination in the past year, n (%)				.005[a, b]
No	23 (74.2)	81 (47.1)	104 (51.2)	
Yes	8 (25.8)	91 (52.9)	99 (48.8)	
Self-reported health status, mean (SD)	63.23 (11.73)	61.51 (15.28)		.46[d]
Perceived threat, mean (SD)	3.42 (1.04)	3.86 (1.05)		.02[b, d]
Perceived benefit, mean (SD)	2.80 (1.25)	4.11 (0.89)		<.001[b, d]
Perceived barrier, mean (SD)	1.93 (0.82)	2.76 (1.23)		<.001[b, d]
Cues to action, mean (SD)	2.07 (1.11)	3.47 (1.21)		<.001[b, d]

[a]Chi-square test.

[b]Statistically significant values.

[c]Average exchange rate at time of study was US $ 1 = NT $29.67.

[d]Independent samples *t* test.

Factors Associated With Action to COVID-19 Vaccination

Logistic regression analysis without the HBM scale (Table 2, Model 1) included sociodemographic and health-related factors and trust in physicians, and the stepwise backward conditional method was used to select the important variables and fit the model. The results show that marriage status (adjusted odds ratio [AOR] 3.931, 95% CI 1.602 to 9.648; *P*=.003), influenza vaccination in the past year (AOR 3.188, 95% CI 1.291 to 7.872; *P*=.01), and trust in physicians (AOR 1.452, 95% CI 1.004 to 2.099; *P*=.047) are significantly positively correlated with the COVID-19 vaccination.

When the constructs of HBM were added to the model (Table 2, model 2), marriage status (AOR 4.063, 95% CI 1.384 to 11.926; *P*=.01), influenza vaccination in the past year (AOR 3.367, 95% CI 1.063 to 10.662; *P*=.04), perceived threat (AOR=1.922, 95% CI 1.053 to 3.507; *P*=.03), perceived benefits (AOR 1.723, 95% CI 1.023 to 2.901; *P*=.04), perceived barriers (AOR 0.412, 95% CI 0.199 to 0.855; *P*=.02) and cues to action (AOR 2.603, 95% CI 1.461 to 4.637; *P*<.001) were significantly positively correlated to the COVID-19 vaccine; however, trust in physicians was not selected into model 2.

XSL•FO
RenderX

Table . Factors associated with actions to receive the COVID-19 vaccine among indigenous participants by backward stepwise multivariate binary logistic regression (n=203).

	Model 1[a]				Model 2[b]			
	Path coefficient	*P* value	95% CI	AOR[c]	Path coefficient	*P* value	95% CI	AOR[c]
Marriage (reference: unmarried [including divorced and widowed])[d]								
Married	1.369	.003[e]	1.602-9.648	3.931	1.402	.01[e]	1.384-11.926	4.063
Education (reference: incomplete high school education)[d]								
Complete high school education	2.060	.05	0.999-61.656	7.850	1.708	.14	0.580-52.561	5.520
Influenza vaccination in the past year (reference: no)[d]								
Yes	1.159	.01[e]	1.291-7.872	3.188	1.214	.04[e]	1.063-10.662	3.367
Trust in physicians health belief model	0.373	.047[e]	1.004-2.099	1.452				
Perceived threat					0.653	.03[e]	1.053-3.507	1.922
Perceived benefits					0.544	.04[e]	1.023-2.901	1.723
Perceived barriers					−0.886	.02[e]	0.199-0.855	0.412
Cues to action					0.957	<.001[e]	1.461-4.637	2.603

[a]Nagelkerke R^2=0.230.

[b]Nagelkerke R^2=0.535.

[c]AOR: adjusted odds ratio.

[d]The reference category in the logistic regression model, against which the odds ratios of other categories are compared.

[e]Statistically significant values.

Mediating Effect Between Trust in Physicians and Actions to Receive the COVID-19 Vaccine

The results (Table 3) indicated that trust in physicians was not significantly associated with perceived threat (β=0.013; *P*=.86) or with COVID-19 vaccine uptake (AOR 0.993, 95% CI −0.505 to 0.490; *P*=.98). Given the nonsignificant direct association between trust in physicians and perceived threat, the potential mediating role of perceived threat was not further examined.

The trust in physicians was positively associated with perceived benefits (β=0.188; *P*=.003), and perceived benefits were significantly associated with vaccine uptake (β=0.545; *P*=.04). However, the indirect effect of trust in physicians on vaccine uptake through perceived benefits was 0.104, with a 95% bootstrap CI (−0.005 to 0.334), including 0, indicating no significant mediation effect.

Additionally, trust in physicians was negatively associated with perceived barriers (β=−0.165; *P*=.01), and perceived barriers were significantly associated with vaccine uptake (β=−0.887; *P*=.02). The indirect effect of trust in physicians on vaccine uptake through perceived barriers was 0.143, with a 95% bootstrap CI (0.013 to 0.437), not including 0, indicating a significant full mediation effect.

XSL•FO
RenderX

Table . Results of the mediating role of perceived threat, benefits, and barriers between trust in physicians and actions to receive the COVID-19 vaccine.

	Path coefficient	*P* value	95% CI[a] (Bootstrap)
Direct effect			
Trust in Physicians → Perceived Threat	0.013	.86	−0.127 to 0.153
Perceived Threat → Actions to receive the COVID-19 vaccine	0.654	.03[b]	0.052 to 1.255
Trust in Physicians → Perceived Benefits	0.188	.003[b]	0.064 to 0.312
Perceived Benefits → Actions to receive the COVID-19 vaccine	0.545	.04[b]	0.018 to 1.073
Trust in Physicians → Perceived Barriers	−0.165	.01[b]	−0.295 to −0.034
Perceived Barriers → Actions to receive the COVID-19 vaccine	−0.887	.02[b]	−1.619 to −0.154
Trust in Physicians → Actions to receive the COVID-19 vaccine	−0.007	.98	−0.505 to 0.490
Indirect effect			
Trust in Physicians → Perceived Threat → Actions to receive the COVID-19 vaccine	0.008	—[c]	−0.115 to 0.136
Trust in Physicians → Perceived Benefits → Actions to receive the COVID-19 vaccine	0.104	—[c]	−0.005 to 0.334
Trust in Physicians → Perceived Barriers → Actions to receive the COVID-19 vaccine	0.143	—[c]	0.013 to 0.437

[a]A CI that does not include 0 indicates a statistically significant mediation effect.

[b]Statistically significant values.

[c]Not applicable.

Discussion

Principal Findings

The findings indicate that marital status, influenza vaccination in the previous year, and trust in physicians were significantly associated with the COVID-19 vaccine uptake before incorporating constructs from the HBM. After adding the HBM constructs, namely perceived threats, perceived benefits, perceived barriers, and cues to actions, marital status, prior influenza vaccination, and all HBM-related factors remained significantly associated with vaccine uptake. However, trust in physicians was no longer a significant predictor after the inclusion of the HBM constructs in the model. To further examine the underlying mechanisms, mediation analyses were performed to assess the potential indirect effects of trust in physicians on vaccine uptake through perceived threats, perceived benefits, and perceived barriers. Results revealed that perceived barriers fully mediated the relationship between trust in physicians and vaccine uptake.

Trust in Physicians

Before incorporating the HBM constructs into the logistic regression model, trust in the physician was significantly associated with COVID-19 vaccine uptake. However, the association became nonsignificant after the HBM factors were introduced in the model. Subsequent analyses examining the associations between trust in physicians and other HBM constructs revealed significant associations with perceived benefits and perceived barriers, but not with perceived threat. This pattern may reflect the close and longstanding relationships between local indigenous residents and physicians stationed in the community. The study region has extremely limited health care resources, with only 2 clinics serving the entire area. The public health center is the sole facility providing routine vaccination services, and its physicians conduct weekly outreach visits to each village to deliver both primary care and vaccinations. In addition, home visits are arranged for individuals residing in remote areas or with mobility impairments. Given this longstanding physician-community relationship, it is likely that older indigenous adults receive vaccination-related information primarily from physicians.

However, further path analysis indicated that trust in physicians exerted a significant indirect effect on vaccination uptake through perceived barriers, but not through perceived benefits. While previous studies have suggested that trust in health care providers can influence vaccination behavior both directly and indirectly through multiple health belief constructs [21], the present findings indicate that such influence may operate

XSL•FO
RenderX

through a more limited pathway, specifically, via perceived barriers.

Perceived Barriers

In this study, the construct of perceived barriers was significantly negatively associated with COVID-19 vaccine uptake, thus highlighting their detrimental role in vaccination decisions. Barriers included concerns regarding the vaccine development process, doubts about vaccine safety or effectiveness, fear of receiving a defective vaccine, and apprehension about potential severe or long-term side effects. Concerns about vaccine safety have consistently been identified as a major obstacle to vaccination across various populations [22]. Given that COVID-19 is caused by a novel coronavirus first detected in humans, the accelerated development and deployment of its vaccines raised public concerns about insufficient information and unknown long-term effects. Moreover, some vaccine platforms were used in human populations for the first time, contributing to hesitation. Not surprisingly, anxiety surrounding the safety and risks of vaccination, including unverified rumors, was widespread. In this study, worries about the expedited approval process and potential adverse effects were frequently mentioned by participants during informal interviews. These findings are consistent with previous research among an Australian indigenous study [1] and older adults in Taiwan [23], both of which identified safety concerns as central to vaccine hesitancy. Despite the negative influence of perceived barriers, mediation analysis revealed a significant indirect positive effect of trust in physicians on vaccine uptake through perceived barriers. In other words, trust in physicians mitigated the adverse impact of perceived barriers, thereby facilitating vaccine acceptance. This finding highlights the critical role of physician trust in enhancing the credibility and perceived accuracy of vaccine-related information [24], which is particularly important for indigenous older adults, who may face greater structural and informational disadvantages. Overall, these findings underscore the central role of perceived barriers in shaping vaccine behavior and demonstrate that reducing such barriers, partly through fostering trust in physicians, may be essential to improving vaccine uptake in indigenous older adult populations in Taiwan.

Perceived Threat

The result of logistic regression indicated that respondents who were worried about the epidemic tended to get vaccinated, similar to other studies [25]. The perception of the disease affecting the odds of COVID-19 vaccination has also occurred in Australia, for example. Misconceptions about diseases and vaccines among Aboriginal and Torres Strait Islander people, as well as a lack of cultural sensitivity to Aboriginal health care delivery, contribute to lower vaccination rates [26]. Policies can be used to raise disease awareness among the indigenous population to increase vaccination coverage. In Australia, there are policies for encouraging indigenous people to get vaccinated, such as compiling a vaccination plan for indigenous people, relevant implementation strategies, providing indigenous health education media on the official website of the health unit, including short films and picture cards, and providing culturally sensitive educational materials for companies and institutions that may employ aboriginal people. Indeed, effective risk communication for pandemics is crucial, whereas a lack of trust in the central government and, therefore, a lack of communication strategies about COVID-19 prevention in Indigenous languages was a critical challenge [27,28]. Nevertheless, in contrast to previous research [20], this study found that trust in physicians did not exert an indirect effect on vaccine uptake through perceived threat. This divergence may reflect culturally distinct health worldviews among indigenous older adults. During informal, off-the-record conversations conducted alongside data collection, several unvaccinated participants expressed a fatalistic perspective, with remarks such as, “Life and death are predetermined; vaccines won’t change anything.” Such views suggest a low perceived threat of COVID-19, rooted in a worldview that emphasizes acceptance of natural forces and the cyclical nature of life and death. For some participants, the COVID-19 pandemic was regarded not as a crisis to be feared but rather as part of the broader natural order. This perspective reflects a culturally grounded sense of equilibrium with nature, rather than a deficit in health knowledge or risk awareness.

Perceived Benefits

The direct effect of perceived benefits on vaccination uptake indicates that individuals who believe they will benefit from receiving the vaccine are more likely to do so, a finding consistent with prior studies [18,29]. In addition, the positive association between trust in physicians and perceived benefits suggests that indigenous adults in this study who reported higher trust in physicians were more likely to recognize the potential advantages of vaccination. This pattern aligns with evidence from other vaccination contexts [30]. However, the indirect effect of trust in physicians on vaccine uptake through perceived benefits was not significant, although it approached significance. Previous research has shown that, in some cases, individuals may perceive the risks of vaccination as outweighing its benefits [31]. Given the prominent role that safety concerns played in this study, it is possible that these concerns attenuated the influence of perceived benefits, thereby limiting the mediating role of this construct. Further research is warranted to explore whether trust in physicians may enhance vaccine uptake through other psychological pathways.

Cue to Action

This study revealed that the support from opinion leaders in the community for vaccination will increase the chances of indigenous elders being vaccinated. A similar case was observed in American Indian communities and has been advocated by international organizations such as the International Federation of Red Cross and Red Crescent Societies [32-34]. Community opinion leaders include not just tribal opinion leaders and local district chiefs but also religious field workers, friends, and family [35,36]. Informal interviews conducted during this study revealed that several participants expressed sentiments such as “No one around me gets vaccinated,” underscoring a lack of external cues to action that might otherwise prompt vaccination. Therefore, an understanding of the interpersonal relationships and communal networks in the local context is essential for increasing the vaccination rate. Thus, as also identified by

XSL•FO
RenderX

studies conducted in various social contexts, cooperating with community opinion leaders requires a bottom-up approach to transforming these leaders into "trusted messengers" in vaccination promotion [37]. Such a strategy, often overlooked by the central government, however, is particularly important for indigenous communities according to our and other studies [38,39].

Limitations

There are some limitations in this study. First, during the period of data collection, the COVID-19 outbreak happened in this region, so it was challenging to do the interviews without considering compliance with relevant regulations in addition to the methodological restraints and ethical considerations preidentified. In any case, the results of this study are reliable and valid, and it is the first study in Taiwan on indigenous COVID-19 vaccination. Second, since the government has yet to release official data on COVID-19 vaccination rates among indigenous people, relevant information on this matter is based on news reports, the accuracy of which is uncertain. Thus, official statistics are still needed to inform a comprehensive analysis. Third, although the findings may not be generalizable to all indigenous populations, the study's framework is particularly relevant to regions characterized by indigenous-majority populations, limited health care access, and geographic or transportation barriers. Taiwan comprises 55 officially designated indigenous regions, 30 mountainous and 25 plains regions. The mountainous regions particularly tend to exhibit a strong cultural identity and relative geographic isolation. These conditions are also commonly observed in many indigenous communities around the world. While infrastructure and health care delivery systems may vary across indigenous regions, those facing similar structural constraints are likely to experience comparable challenges in vaccine uptake. Therefore, the proposed framework holds promise for broader applicability within Taiwan and potentially across other national and international indigenous communities.

Conclusions

This study provides empirical evidence on the determinants of COVID-19 vaccination among indigenous older adults in Taiwan. The findings indicate that key constructs of the HBM, including perceived disease threat, perceived vaccine benefits, concerns regarding vaccine safety and risk, and support from trusted community opinion leaders, are significantly associated with vaccine uptake in this population. Notably, the effect of trust in physicians on vaccination behavior appears to operate indirectly through perceptions of vaccine safety and risk, particularly in the context of a severe and rapidly spreading pandemic involving a novel vaccine. These results underscore the importance of a contextualized and culturally grounded approach to risk and policy communication in indigenous communities. Pandemic prevention strategies should be informed by local values and sociocultural dynamics, with trusted community leaders playing a central role in promoting vaccine acceptance. Vaccine hesitancy in these settings may reflect informed cultural or religious perspectives, rather than misinformation or lack of awareness. Accordingly, a bottom-up approach is recommended, whereby frontline health care professionals engage with community members to understand local beliefs and practices, thereby informing policy development. Furthermore, proactive involvement by community health care providers, through the delivery of clear, culturally sensitive, and trustworthy information to indigenous elders and their caregivers, may help alleviate safety concerns and reduce perceived risks. These efforts may enhance vaccine confidence and ultimately improve vaccination coverage in indigenous regions.

Acknowledgments

The authors would like to acknowledge all participants for taking part in the research studies, and Mr Yu-Hao Lai from Namasia District Public Health Center and Dr Yiling Shen from the Yangming Branch of Taipei City Hospital for their assistance in this study.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to confidentiality requirements and ethical restrictions but are available from the corresponding author, REC, on reasonable request. Access requires approval from the Behavioral and Social Sciences Research Ethics Committee at Kaohsiung Medical University Chung-Ho Memorial Hospital and consultation meetings about cultural risk assessment by the Council of Indigenous Peoples.

Conflicts of Interest

None declared.

Checklist 1
STROBE cross-sectional checklist.
[PDF File, 71 KB - publichealth_v11i1e75278_app1.pdf]

References

XSL•FO
RenderX

1. Graham S, Blaxland M, Bolt R, et al. Aboriginal peoples' perspectives about COVID-19 vaccines and motivations to seek vaccination: a qualitative study. BMJ Glob Health 2022 Jul;7(7):e008815. [doi: 10.1136/bmjgh-2022-008815] [Medline: 35858705]
2. Machado FCG, Ferron MM, Barddal MDM, Nascimento LA, Rosalen J, Avelino-Silva VI. COVID-19 vaccination, incidence, and mortality rates among indigenous populations compared to the general population in Brazil: describing trends over time. Lancet Reg Health Am 2022 Sep;13:100319. [doi: 10.1016/j.lana.2022.100319]
3. Abascal Miguel L, Lopez E, Sanders K, et al. Evaluating the impact of a linguistically and culturally tailored social media ad campaign on COVID-19 vaccine uptake among indigenous populations in Guatemala: a pre/post design intervention study. BMJ Open 2022 Dec;12(12):e066365. [doi: 10.1136/bmjopen-2022-066365]
4. Mehmood Q, Tebha SS, Aborode AT. COVID-19 vaccine hesitancy among indigenous people in India: an incipient crisis. Ethics Med Public Health 2021 Dec;19:100727. [doi: 10.1016/j.jemep.2021.100727] [Medline: 34632033]
5. Foxworth R, Redvers N, Moreno MA, Lopez-Carmen VA, Sanchez GR, Shultz JM. COVID-19 vaccination in American Indians and Alaska Natives - lessons from effective community responses. N Engl J Med 2021 Dec 23;385(26):2403-2406. [doi: 10.1056/NEJMp2113296] [Medline: 34919356]
6. Curtice K, Choo E. Indigenous populations: left behind in the COVID-19 response. Lancet 2020 Jun;395(10239):1753. [doi: 10.1016/S0140-6736(20)31242-3]
7. Pete D, Erickson SL, Jim MA, Hatcher SM, Echo-Hawk A, Dominguez AE. COVID-19 among non-hispanic American Indian and Alaska native people residing in urban areas before and after vaccine rollout—selected states and counties, United States, january 2020–october 2021. Am J Public Health 2022 Oct;112(10):1489-1497. [doi: 10.2105/AJPH.2022.306966]
8. Wong LP, Alias H, Wong PF, Lee HY, AbuBakar S. The use of the health belief model to assess predictors of intent to receive the COVID-19 vaccine and willingness to pay. Hum Vaccin Immunother 2020 Sep 1;16(9):2204-2214. [doi: 10.1080/21645515.2020.1790279]
9. Fedele F, Aria M, Esposito V, et al. COVID-19 vaccine hesitancy: a survey in a population highly compliant to common vaccinations. Hum Vaccin Immunother 2021 Oct 3;17(10):3348-3354. [doi: 10.1080/21645515.2021.1928460]
10. Harris KM, Maurer J, Lurie N. Do people who intend to get a flu shot actually get one? J Gen Intern Med 2009 Dec;24(12):1311-1313. [doi: 10.1007/s11606-009-1126-2] [Medline: 19838758]
11. Liao Q, Cowling BJ, Lam WWT, Fielding R. Factors affecting intention to receive and self-reported receipt of 2009 pandemic (H1N1) vaccine in Hong Kong: a longitudinal study. PLOS ONE 2011 Mar 11;6(3):e17713. [doi: 10.1371/journal.pone.0017713] [Medline: 21412418]
12. Kislitsyn DV, Schapov DS. From intention to action: factors of vaccine hesitancy and vaccine refusal during the COVID-19 pandemic. POPECON 2022 Dec;6(4):162-177. [doi: 10.3897/popecon.6.e90723]
13. MacArthur KR. Beyond health beliefs: the role of trust in the HPV vaccine decision-making process among American college students. Health Sociol Rev 2017 Sep 2;26(3):321-338. [doi: 10.1080/14461242.2017.1381035]
14. Liu S, Chu H. Examining the direct and indirect effects of trust in motivating COVID-19 vaccine uptake. Patient Educ Couns 2022 Jul;105(7):2096-2102. [doi: 10.1016/j.pec.2022.02.009] [Medline: 35181177]
15. Huang PJ, Hsu CH, Kung CC, Chao CL, Chang HY. A survey on behavior of influenza vaccination with health belief model in dialysis patients [Article in Chinese]. Kidney Dial 2019 Jun;31(2):94-102. [doi: 10.6340/KD.201906_31(2).0011]
16. Hossain MB, Alam MZ, Islam MS, et al. Health belief model, theory of planned behavior, or psychological antecedents: what predicts COVID-19 vaccine hesitancy better among the Bangladeshi adults? Front Public Health 2021 Aug;9:711066. [doi: 10.3389/fpubh.2021.711066]
17. Walrave M, Waeterloos C, Ponnet K. Adoption of a contact tracing app for containing COVID-19: a health belief model approach. JMIR Public Health Surveill 2020 Sep 1;6(3):e20572. [doi: 10.2196/20572] [Medline: 32755882]
18. Shmueli L. Predicting intention to receive COVID-19 vaccine among the general population using the health belief model and the theory of planned behavior model. BMC Public Health 2021 Apr 26;21(1):804. [doi: 10.1186/s12889-021-10816-7] [Medline: 33902501]
19. Hall MA, Camacho F, Dugan E, Balkrishnan R. Trust in the medical profession: conceptual and measurement issues. Health Serv Res 2002 Oct;37(5):1419-1439. [doi: 10.1111/1475-6773.01070] [Medline: 12479504]
20. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach, 3rd edition: Guilford Publications; 2022.
21. Alzahrani SH. The impact of health beliefs and trust in health information sources on SARS-CoV-2 vaccine uptake. Front Public Health 2014 Mar;12. [doi: 10.3389/fpubh.2024.1340614]
22. Troiano G, Nardi A. Vaccine hesitancy in the era of COVID-19. Public Health 2021 May;194:245-251. [doi: 10.1016/j.puhe.2021.02.025] [Medline: 33965796]
23. Kuan CI, Lin JY, Liu CH. Vaccine hesitancy and risk communication: experiences among older adults and patients with chronic diseases [Article in Chinese]. Taiwan J Public Health 2022 Aug;41(4):449-466. [doi: 10.6288/TJPH.202208_41(4).111028]
24. Solís Arce JS, Warren SS, Meriggi NF, et al. COVID-19 vaccine acceptance and hesitancy in low- and middle-income countries. Nat Med 2021 Aug;27(8):1385-1394. [doi: 10.1038/s41591-021-01454-y] [Medline: 34272499]

XSL•FO
RenderX

25. Mertens G, Lodder P, Smeets T, Duijndam S. Fear of COVID-19 predicts vaccination willingness 14 months later. J Anxiety Disord 2022 May;88:102574. [doi: 10.1016/j.janxdis.2022.102574] [Medline: 35512598]
26. Menzies R, Aqel J, Abdi I, Joseph T, Seale H, Nathan S. Why is influenza vaccine uptake so low among aboriginal adults? Aust N Z J Public Health 2020 Aug;44(4):279-283. [doi: 10.1111/1753-6405.13004] [Medline: 32583522]
27. Goha A, Mezue K, Edwards P, et al. Indigenous people and the COVID-19 pandemic: the tip of an iceberg of social and economic inequities. J Epidemiol Community Health 2021 Feb;75(2):207-208. [doi: 10.1136/jech-2020-214755] [Medline: 33067253]
28. Pickering K, Galappaththi EK, Ford JD, et al. Indigenous peoples and the COVID-19 pandemic: a systematic scoping review. Environ Res Lett 2023 Mar 1;18(3):033001. [doi: 10.1088/1748-9326/acb804] [Medline: 36798651]
29. Mostafapour M, Meyer SB, Scholer A. Exploring the effect of risk and benefit information provision on vaccination decision-making. Vaccine (Auckl) 2019 Oct 16;37(44):6750-6759. [doi: 10.1016/j.vaccine.2019.08.083] [Medline: 31558328]
30. Song G. Understanding public perceptions of benefits and risks of childhood vaccinations in the United States. Risk Anal 2014 Mar;34(3):541-555. [doi: 10.1111/risa.12114] [Medline: 24033739]
31. Schaffer DeRoo S, Pudalov NJ, Fu LY. Planning for a COVID-19 vaccination program. JAMA 2020 Jun 23;323(24):2458. [doi: 10.1001/jama.2020.8711]
32. Silberner J. Covid-19: how native Americans led the way in the US vaccination effort. BMJ 2021 Sep;374:n2168. [doi: 10.1136/bmj.n2168]
33. International Federation of Red Cross and Red Crescent Societies COVID-19 pandemic final report. : International Federation of Red Cross and Red Crescent Societies; 2023 Jul.
34. Hasan T, Beardsley J, Marais BJ, Nguyen TA, Fox GJ. The implementation of mass-vaccination against SARS-CoV-2: a systematic review of existing strategies and guidelines. Vaccines (Basel) 2021 Apr 1;9(4):326. [doi: 10.3390/vaccines9040326] [Medline: 33915829]
35. Valente TW, Pumpuang P. Identifying opinion leaders to promote behavior change. Health Educ Behav 2007 Dec;34(6):881-896. [doi: 10.1177/1090198106297855] [Medline: 17602096]
36. Lee PH. The significance of cultural competence in public health: COVID-19 contact tracing as an example. Taiwan J Public Health 2021 Oct;40(5):474-478. [doi: 10.6288/TJPH.202110_40(5).110066]
37. Shen AK, Browne S, Srivastava T, Kornides ML, Tan ASL. Trusted messengers and trusted messages: the role for community-based organizations in promoting COVID-19 and routine immunizations. Vaccine (Auckl) 2023 Mar 17;41(12):1994-2002. [doi: 10.1016/j.vaccine.2023.02.045] [Medline: 36803894]
38. Nsubuga EJ, Fitzmaurice AG, Komakech A, et al. Community dialogue meetings among district leaders improved their willingness to receive COVID-19 vaccines in Western Uganda, May 2021. BMC Public Health 2021;23(1):969. [doi: 10.1186/s12889-023-15903-5] [Medline: 33985451]
39. Khan YH, Mallhi TH, Alotaibi NH, et al. Threat of COVID-19 vaccine hesitancy in Pakistan: the need for measures to neutralize misleading narratives. Am J Trop Med Hyg 2020 Aug;103(2):603-604. [doi: 10.4269/ajtmh.20-0654] [Medline: 32588810]

Abbreviations

AOR: adjusted odds ratio
EFA: exploratory factor analysis
HBM: health belief model
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 02.04.25; peer-reviewed by L Zhu, M Etienne-Mesubi; revised version received 14.07.25; accepted 15.07.25; published 10.09.25.

Please cite as:
Tu Abùs Isbabanal JC, Chang RE, O'Neill L, Chen TH, Lee PH, Ou TT
Identifying Success Factors for Optimizing COVID-19 Vaccine Uptake Among Indigenous Populations in Taiwan: Cross-Sectional Questionnaire Survey
JMIR Public Health Surveill 2025;11:e75278
URL: https://publichealth.jmir.org/2025/1/e75278
doi:10.2196/75278

© Jia-Chi Tu Abùs Isbabanal, Ray-E Chang, Liam O'Neill, Ting-Hung Chen, Po-Han Lee, Tsan-Teng Ou. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 10.9.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted

XSL•FO
RenderX

use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Perceived Scarcity on Mental Health, Time and Risk Preferences, and Decision-Making During and After COVID-19 Lockdown: Quasi-Natural Experimental Study

Haiou Zhu[1,2], PhD; Fangzhou You[3], PhD; Thorsten Gruber[4], Prof Dr; Hua Dong[5], Prof Dr; Cees de Bont[6], Prof Dr

[1]Department of Psychiatry, University of Oxford, Warneford Hospital, Warneford Lane, Headington, Oxford, United Kingdom

[2]School of Design and Creative Arts, Loughborough University, Loughborough, United Kingdom

[3]School of Arts and Creative Technologies, University of York, York, United Kingdom

[4]Centre for Service Management, Business School, Loughborough University, Loughborough, United Kingdom

[5]Brunel Design School, Brunel University London, London, United Kingdom

[6]College of Design and Engineering, National University of Singapore, Singapore, Singapore

Corresponding Author:
Haiou Zhu, PhD
Department of Psychiatry, University of Oxford, Warneford Hospital, Warneford Lane, Headington, Oxford, United Kingdom

Abstract

Background: The COVID-19 lockdowns led to significant resource constraints, potentially impacting mental health and decision-making behaviors. Understanding the psychological and behavioral consequences could inform designing interventions to mitigate the negative impacts of episodic scarcity during crises like pandemics.

Objective: To investigate the effects of perceived scarcity on mental health (stress and fear), cognitive functioning, time and risk preferences (present bias and risk aversion), and trade-offs between groceries, health, and temptation goods during and after the COVID-19 lockdown in Shanghai.

Methods: A quasi-natural experiment was conducted in Shanghai during and after the COVID-19 lockdown. Web-based surveys were administered in May 2022 (during lockdown) and September 2022 (post-lockdown). Propensity score matching was used to balance demographic factors between the groups (During: n=332; After: n=339). Data were analyzed using regression analyses, controlling for potential confounders and applying propensity score matching weights.

Results: Perceived scarcity was significantly higher during the lockdown (mean 7.97 (SD 2.1)) than after (mean 4.35 (SD 2.27); *P*<.001). Higher perceived scarcity was associated with increased stress levels both during (standardized β coefficient=.62, *P*<.001) and after the lockdown (standardized β coefficient=.65, *P*<.001). Perceived scarcity also predicted greater fear of COVID-19 after lockdown (standardized β coefficient =.38, *P*<.001), though not during lockdown. Cognitive functioning remained stable, possibly due to a ceiling effect from high education levels. Monetary risk aversion increased under prolonged scarcity during lockdown (scarcity×during-lockdown interaction standardized β coefficient=4.68, *P*<.001). Present bias (tendency to choose immediate rewards) showed no significant overall change between groups, in line with recent evidence of stable time preferences during the pandemic. During lockdown, participants allocated more budget to groceries (standardized β coefficient=.67, *P*=.01) and less to health items (standardized β coefficient=−.61, *P*=.02), compared to post-lockdown, reflecting shifted priorities on pressing needs under scarcity. Subgroup analyses indicated stratified heterogeneity. Women increased their grocery spending (standardized β coefficient =.16, *P*=.04) and reduced spending on health items (standardized β coefficient = –.15, *P*=.05). Lower education participants exhibited more risk-averse attitudes (standardized β coefficient =.80, *P*=.01) under scarcity, whereas age and income did not significantly moderate these effects.

Conclusions: The study highlights that perceived scarcity during lockdown intensified stress and altered decision-making behaviors, including increased monetary risk aversion and shifts in spending priorities. Theoretically, this study advances the understanding of perceived scarcity by exploring its domain-specific effects on mental health and decision-making. Practically, these findings emphasize the need for public health strategies that mitigate the psychological impact of scarcity during crises, ensure access to essential goods, and support adaptive decision-making behaviors.

(JMIR Public Health Surveill 2025;11:e69496) doi:10.2196/69496

KEYWORDS

perceived scarcity; COVID-19 lockdown; mental health; wellbeing; cognitive function; present bias; risk aversion; natural experiment; public health interventions; consumer behavior

XSL•FO
RenderX

Introduction

Lockdown policies have been among the most effective public health interventions during the COVID-19 pandemic, saving millions of lives worldwide [1,2]. However, lockdown measures, including social distancing, self-isolation, and prolonged home quarantine, along with disruptions to supply chains and shortages of essential products and services, have significantly altered many aspects of people's daily lives [3,4]. Research has shown that such measures have exacerbated mental health problems, leading to increased stress, depression, and negative affect [5-8], affected job, financial, and food security [9,10], and caused panic buying [11,12] and adverse health decisions and behaviors [13-15]. While many studies have used longitudinal and natural experiments to investigate the impact of lockdowns [16-18], few have focused on the psychological and behavioral consequences arising from the radical resource constraints individuals face during these periods.

A location where a stringent lockdown was implemented in response to a widespread COVID-19 outbreak was Shanghai, China's largest city by population (over 24 million) and one of the largest cities in the world. The lockdown lasted from March to May 2022, and residents were largely quarantined at home, facing significant shortages of products and services. After the city reopened on June 1, 2022, the supply and delivery of products and services resumed pre-lockdown levels. This abrupt change in resource availability created a unique opportunity to examine how resource scarcity during lockdown affected various aspects of residents' lives in one of the largest urban populations in the world. The timing of this study is important since it captured a real-world episode of intense scarcity [19,20]. Understanding how these episodic constraints affect mental health and health-related decisions and behaviors can provide valuable empirical evidence on the impacts of pandemic lockdowns.

The concept of perceived scarcity, or psychological scarcity, is rooted in poverty literature and is well-established in the psychology and economics literature [21-23]. It refers to the subjective feelings that arise from the gap between one's needs and the resources available to fulfill them [24,25]. Feelings of scarcity can preoccupy one's thoughts, making everyday decisions, such as whether to visit a doctor or buy groceries, more difficult and stressful [26]. While traditionally studied in the context of poverty, recent research suggests that perceived scarcity can affect individuals across socioeconomic statuses, particularly during crises that disrupt normal resource access [27]. Previous studies have linked perceived scarcity to increased stress and negative affect [21,28,29]. Research also suggests that scarcity can directly impair cognitive functions, including decision-making, reasoning, problem-solving, and attention, by reducing cognitive bandwidth and limiting the mental resources necessary to process information and make decisions [24,30]. However, these effects have not been consistently confirmed in other empirical studies [27-29]. Both the negative emotional states and the cognitive load induced by scarcity are seen as mechanisms that cause scarcity-reinforcing decisions and behaviors, leading to a cycle where poor decisions exacerbate conditions of health and well-being [17]. During the COVID-19 pandemic, stress and fear became widespread as individuals faced uncertainty and isolation, along with increasing cognitive load [31-34]. However, it remains unclear how these feelings and cognitive functioning may have been exacerbated by perceived scarcity [35].

Perceived scarcity can also impact decision-making in various domains by inducing biases and trade-off thinking. Time and risk preferences are fundamental components of human decision-making, with present bias and risk aversion representing systematic deviations from rational behavior [36,37]. Present bias refers to the tendency to prioritize immediate rewards over future gains, leading to short-sighted decisions that come at the expense of long-term goals [38,39]. Risk aversion, on the other hand, describes the preference for avoiding risks, making individuals less inclined to engage in actions with uncertain outcomes [40]. Both present bias and risk aversion influence not only financial decisions but also health-related behaviors, such as smoking [41,42] and obesity [43], which are considered irrational. Research on perceived scarcity shows mixed outcomes, with some studies indicating increased risk-taking [44] and time-discounting [45,46], while others find no significant effects [46]. In terms of trade-off thinking, feelings of scarcity induced by financial worries seem to shift attention toward pressing needs, such as food and groceries, at the expense of forward-looking decisions like saving for health and education [47,48]. Additionally, scarcity may lead to the overconsumption of temptation goods, including junk food, sugary drinks, and cigarettes [49]. Nonetheless, these findings are inconsistent, suggesting that the effects of perceived scarcity may vary based on context and population characteristics [22,50]. It remains unclear whether scarcity caused by supply shortages would force individuals to make trade-offs between pressing needs and their health in a pandemic context [3,11,16].

Using a quasi-experimental design, this study aims to examine the effects of perceived scarcity during a COVID-19 lockdown on individuals' mental health and decision-making, compared with those surveyed after lockdown when scarcity had abated. Specifically, the study investigates whether experiencing resource scarcity is associated with higher stress and fear (mental health outcomes), changes in cognitive function, present bias, or risk aversion (time and risk preferences), and altered spending trade-offs between necessities and non-essentials. The hypothesis was that the lockdown-induced scarcity would heighten stress and fear levels and lead to more cautious decision-making: namely, increased risk aversion, a shift toward present-focused choices, and a reallocation of spending toward essential goods. Furthermore, to estimate the potential cumulative effects of lockdown, the study examines whether the length of exposure to lockdown, ie, the duration of home quarantine and the length of time without shopping, exacerbated mental well-being and influenced decision-making. This study contributes to the literature by extending the concept of perceived scarcity in a high-income, urban population during a crisis, advancing the understanding of its impact on their mental health and decision-making behaviors. The findings can inform policy makers and public health practitioners in designing interventions to mitigate the negative impacts of

XSL•FO
RenderX

scarcity during crises and emergencies, ensuring that strategies are tailored to address both supply shortages and psychological well-being.

Methods

Study Design

This study adopted the target trial framework [19] for the design, analysis, and reporting of the natural experimental study. The COVID-19 lockdown policy in Shanghai served as the intervention, with the treatment group comprising participants affected during the lockdown, and the control group comprising participants who had resumed regular daily activities by September 2022, 3 months after the city's reopening. Shanghai was chosen due to its unique and stringent lockdown measures during the COVID-19 pandemic, which lasted from March to May 2022. Shanghai's lockdown provided an opportunity to study the psychological and behavioral impacts of resource scarcity in an urban setting. According to prior conceptualizations of natural experiment [19,20] and the Medical Research Council guidelines [51,52], this study qualifies as a between-group natural experiment, as it uses the natural occurrence of the lockdown to assign exposure to participants in a manner akin to randomization, which is pivotal for establishing causal relationships in situations where controlled random assignment is not feasible or ethical [53]. While we cannot prove causality definitively, the natural experiment design, treating the lockdown as an exogenous shock, could provide suggestive causal evidence on the impact of perceived scarcity .

Sample and Data Collection

In total, 2 rounds of web-based surveys with identical items were conducted via Wenjuanxing, a widely used online survey platform in China [54], to collect data from participants living in Shanghai during and after the COVID-19 lockdown. Aligned with the actual lockdown period, which began in late March and ended in late May 2022, the first survey round was distributed in May 2022, near the end of the lockdown, to capture participants' experiences during quarantine. The second survey round was conducted in September 2022, approximately 3 months after the city had fully reopened, allowing assessment of post-lockdown outcomes. Data collection for each wave was completed within approximately 2 - 3 weeks, balancing the urgency to capture the lockdown experience and practical considerations (eg, time and monetary resources available). Figure S1 in Multimedia Appendix 1 illustrates the study setting on a map, highlighting the lockdown context and timeline.

Since the study required recruitment of participants residing specifically in Shanghai during designated timeframes, the survey was distributed by subscribing to targeted sampling services provided by the survey platform (fee details are described below under Ethical Considerations). Recruitment was carefully targeted to ensure participants met the location and timing criteria. Only individuals residing in Shanghai during the designated timeframes were eligible. Participants in the During-lockdown group were asked to confirm that they were currently under home quarantine at the time of the survey and provide the start date of their quarantine, while participants in the After-lockdown group were asked to confirm that they were no longer in quarantine and provide the end date of their quarantine period. Additional quality control measures were used: attention checks embedded within the survey, minimum completion times enforced for each survey section, and systematic screening conducted both automatically by the platform and manually by researchers.

A statistical power analysis, adjusted for multiple comparisons ($\alpha \approx .0045$) to account for all 11 primary outcome measures and targeting a power of 0.80, determined that at least 288 participants per group were required. A total of 747 participants completed the survey, 367 during the lockdown and 380 after the lockdown. The criteria of data exclusion and the number excluded include the following: (1) not in quarantine for During group or still in quarantine for After group (n=23); (2) not living in Shanghai (n=27); and (3) outliers (n=20) who may have been distracted or encountered technical difficulties, and took excessively long to complete the survey (Z-score >mean+3σ, representing the top 0.15% of completion times). After applying these criteria, 70 participants were excluded (n=36 from the "During" group; n=34 from the "After" group), resulting in a sample of 677 participants.

Propensity Score Matching

To address potential confounding variables and ensure comparability between the "During" and "After" groups, propensity score matching was implemented. Each participant in the "After" group was matched to a participant in the "During" group with a similar propensity score based on demographic variables, including gender, age, education, monthly income, and monthly expenditure. The study used a kernel matching technique (Gaussian kernel, bandwidth =0.06) using the *psmatch2* routine in Stata, which allowed each treated case (During group) to be matched with a weighted composite of control cases (After group) within the kernel bandwidth, maximizing data usage and reducing variance compared to one-to-one nearest-neighbor matching. Matching quality was evaluated using standardized mean differences for each covariate and Rubin B and R statistics. Post-matching balance was satisfactory: all covariate standardized mean differences fell below 0.1 (10%), Rubin B was 8.7% (well under the 25% threshold), and Rubin R was 0.94 (within the recommended 0.5 - 2.0 range), indicating a successful match. Visual inspection of the propensity score distribution confirmed substantial overlap between treatment and control groups (see Table S1 and Figure S1 in Multimedia Appendix 2). After matching, 6 participants fell outside the common support and were excluded, leaving 671 participants on support (332 During; 339 After) used in the main analyses.

Measurements

The measurements for both During and After groups include perceived scarcity, stress, fear of COVID-19, cognitive functioning, time and risk preferences, purchase decisions, and demographic information. Multimedia Appendix 1 gives an overview of the constructs and their measurement items.

For the measurement of perceived scarcity, the study used a 4-item perceived scarcity scale, adapted from Yuen et al [11]

XSL•FO
RenderX

and Byun and Sternquist [55]. Items (rated 1="strongly disagree" to 10="strongly agree") measured the extent to which participants felt essential products were in short supply or hard to obtain, with higher scores indicating greater perceived scarcity. Example items include "It was hard to buy enough food and daily supplies" and "Products I needed were often sold out." Because perceived scarcity may accumulate with prolonged isolation, this study also recorded 2 "exposure" metrics: the start date of home quarantine and the last date the participant went shopping. For the After group, participants provided the date their quarantine ended and the date they went shopping. These metrics allowed examining interaction effects between scarcity perceptions and the length of quarantine or time without shopping.

Participants' stress levels over the past month were measured using the 10-item Perceived Stress Scale (PSS) from Cohen et al [56]. To measure their fear of COVID-19, the study used the 7-item Fear of COVID-19 Scale from Ahorsu et al [57]. Cognitive performance was evaluated using subset E (consisting of 12 matrix reasoning items), the most difficult of the 5 progressively harder subsets (A-E) in the 60-item Standard Progressive Matrices Test [58]. Participants received 1 point for each correct answer (maximum score=12), and cognitive accuracy was calculated as the proportion of correct responses out of 12. Raven's matrices are commonly used to assess logical reasoning ability and problem-solving skills [59]. This subset approach was chosen to reduce survey length while still capturing variation in cognitive functioning.

The study measured both monetary-framed and health-framed time preferences using 2 sets of choice scenarios based on Laibson's quasi-hyperbolic discounting model [39]. In both monetary (job payment) and health (physical examination subsidy) contexts, participants chose between smaller immediate rewards and larger delayed rewards, indicating their level of impatience or present bias [42]. Additional health-related tasks involved choosing between immediate medical attention versus continuing to work for compensation and immediate medical attention versus traveling to a distant hospital. Present bias was identified by the point at which participants switched from preferring later payoffs to earlier ones.

Risk aversion was measured through a series of lottery-choice tasks adapted from the well-established Eckel & Grossman method [60]. In a monetary risk task, participants chose 1 of 6 lotteries, each with a 50/50 chance of a lower or higher payoff, with expected payoffs increasing from the safest (50% chance of winning ¥280 (US$42), 50% chance of winning ¥280 (US$42)) to the riskiest lottery (50% chance of winning ¥700 (US$105), 50% chance of winning ¥20 (US $3)). In a health risk task, participants chose between hypothetical vaccine options with different durations of protection and associated probabilities, adapted from the classic disease problem framework developed by Tversky and Kahneman [61]. Risk aversion was quantified based on the chosen lottery: selecting safer lotteries indicates higher risk aversion. This study chose the Eckel & Grossman approach due to its simplicity for respondents in an online setting and easiness to adapt to both monetary and health scenarios. Note that the study used the non-incentivized elicitation methods for time and risk preferences. The reason was that participants during the lockdown found themselves within a "naturally occurring" state of both financial and health distress. This context improved the likelihood of respondents perceiving the intertemporal and risk trade-offs as more realistic and vivid even in the absence of actual incentive-compatible consequences for their responses.

For the measurement of trade-offs, participants were asked to allocate a hypothetical budget of ¥300 (US $45) over three categories: (1) daily groceries (food and household essentials), (2) health-related goods (medicine and health services), and (3) temptation goods (eg, snacks, toys, or other non-essentials). Each category represents different facets of decision-making, informed by existing literature on trade-off thinking in decision-making [47,49,62]. The budget allocation task was designed to mimic the experience of online shopping, aiming to make the decision-making process feel as realistic as possible for participants, thereby potentially eliciting choices that are representative of their real-world trade-offs and priorities. The items selected in each category were popular and well-known in the local market. The proportion of the budget allocated to each category provides a measure of the importance or priority given to that category, which allows researchers to measure individuals' decision priority under conditions of scarcity. At the end of the survey, the study collected demographic information, including gender, age, education, monthly income, and monthly expenditure.

Ethical Considerations

This study was approved by the Ethics Review Committee of Loughborough University (Approval ID: 2022-8493-10095). All participants provided informed consent electronically prior to participation. The consent form informed participants about the study's purpose, the types of data being collected, and their rights (voluntary participation and ability to withdraw at any time). To protect privacy, no personally identifying information was collected; responses were kept anonymous and only analyzed in aggregate. All data were stored securely with access limited to the research team. No identifiable images or personal data are presented in this paper. Using the platform's services cost US$1120, equivalent to US$1.5 per participant. Participants did not receive direct monetary compensation; however, Wenjuanxing's panel service provided a small incentive (equivalent to approximately US$0.9 per participant) as part of its platform reward system.

Data Analysis

Data were analyzed using Stata 17. The main analysis used ordinary least squares (OLS) regression for continuous outcomes (perceived stress score, fear score, cognitive score, and budget allocation percentages), logistic regression for binary outcomes (eg, whether an individual exhibited present bias), and ordinal logistic regression for ordinal outcomes (risk aversion as lottery choices between 1 and 6). All regressions applied propensity score-matching weights to adjust for baseline differences between the groups. Demographic variables (gender, age, education, income, and expenditure) were included as covariates to control for potential confounding factors. Robust standard errors were used to account for heteroskedasticity and ensure reliable inference. The study probed interaction effects to

XSL•FO
RenderX

examine the cumulative impact of lockdown exposure. In particular, the regression models included interaction terms between perceived scarcity and the lockdown group indicator (During vs After) as well as between perceived scarcity and the 2 exposure duration measures (days in quarantine and days since last shopping). These interactions tested whether the relationship between scarcity and outcomes differed by context (lockdown vs post-lockdown) or by how long someone had been in quarantine or without shopping. All continuous variables were mean-centered (or standardized) where appropriate to facilitate interpretation of interactions.

In all OLS regressions, multicollinearity was assessed using variance inflation factors (VIF). Some variables and interaction terms (eg, education and income when both are included, or interactions involving highly correlated measures) showed elevated VIFs (>10). However, given their theoretical importance, we retained them. The mean VIFs for the main regression models ranged around 4.8 - 6.3, indicating acceptable multicollinearity levels (see Table S2 in Multimedia Appendix 2 for full diagnostics). Sensitivity analyses were conducted to ensure robustness of results. First, 2-sample tests and non-parametric tests to compare group differences on key outcomes. Results were consistent with the regression analyses (see Table S3 in Multimedia Appendix 2). Second, Tobit models were applied for outcomes with bounded scales (budget allocation proportions), and the findings remained substantively unchanged (see Table S4 in Multimedia Appendix 2). Finally, we explored subgroup analyses by age, gender, education, and income to investigate heterogeneity (see Table S6 in Multimedia Appendix 2).

Results

Demographic Characteristics

Table 1 presents the demographic characteristics of participants included in the analysis. The total sample size comprises 671 participants, with 332 in the During group and 339 in the After group. Females comprise 56.6% of the During group and 53.1% of the After group. Males make up 43.4% and 46.9% of these groups, respectively. Most participants are aged 18 - 40, with 86.4% in the During group and 89.1% in the After group falling within this range. The majority of participants have an undergraduate degree, 61.7% in the During group and 77.6% in the After group. Graduate degrees or higher are more common in the During group (31.3%) than in the After group (14.5%). The largest income bracket is ¥10,001 (US$1500) and above, representing 54.5% of the During group and 56.9% of the After group. Most participants spend between ¥3001 (US$450) and ¥7000 (US$1050) monthly, with 51.2% in the During group and 56.7% in the After group. These figures indicate that the 2 groups were well-balanced on key demographics.

XSL•FO
RenderX

Table . Demographic characteristics of study participants by survey group (during lockdown in May 2022 versus after lockdown in September 2022).

Characteristic	During (n=332)	After (n=339)	Total (N=671)
Gender, n (%)			
Female	188 (56.6)	180 (53.1)	368 (54.8)
Male	144 (43.4)	159 (46.9)	303 (45.2)
Age, n (%)			
18 - 30 years	176 (53.0)	199 (58.7)	375 (55.9)
31 - 40 years	111 (33.4)	103 (30.4)	214 (31.9)
41 - 50 years	42 (12.7)	32 (9.4)	74 (11.0)
51 years and older	3 (0.9)	5 (1.5)	8 (1.2)
Education, n (%)			
Middle school and below	7 (2.1)	10 (2.9)	17 (2.5)
High school	16 (4.8)	17 (5.0)	33 (4.9)
Undergraduate	205 (61.7)	263 (77.6)	468 (69.7)
Graduate and above	104 (31.3)	49 (14.5)	153 (22.8)
Monthly income[a], n (%)			
¥2000 and below	9 (2.7)	12 (3.5)	21 (3.1)
¥2001-¥4000	13 (3.9)	9 (2.7)	22 (3.3)
¥4001-¥6000	26 (7.8)	15 (4.4)	41 (6.1)
¥6001-¥8000	39 (11.7)	34 (10.0)	73 (10.9)
¥8001-¥10,000	64 (19.3)	76 (22.4)	140 (20.9)
¥10,001 and above	181 (54.5)	193 (56.9)	374 (55.7)
Monthly expenditure, n (%)			
¥1000 and below	7 (2.1)	6 (1.8)	13 (1.9)
¥1001-¥3000	66 (19.9)	63 (18.6)	129 (19.2)
¥3001-¥5000	107 (32.2)	108 (31.9)	215 (32.0)
¥5001-¥7000	63 (19.0)	84 (24.8)	147 (21.9)
¥7001-¥9000	34 (10.2)	48 (14.2)	82 (12.2)
¥9001 and above	55 (16.6)	30 (8.8)	85 (12.7)

[a]1 yuan = US$0.15.

Lockdown Treatment Effects

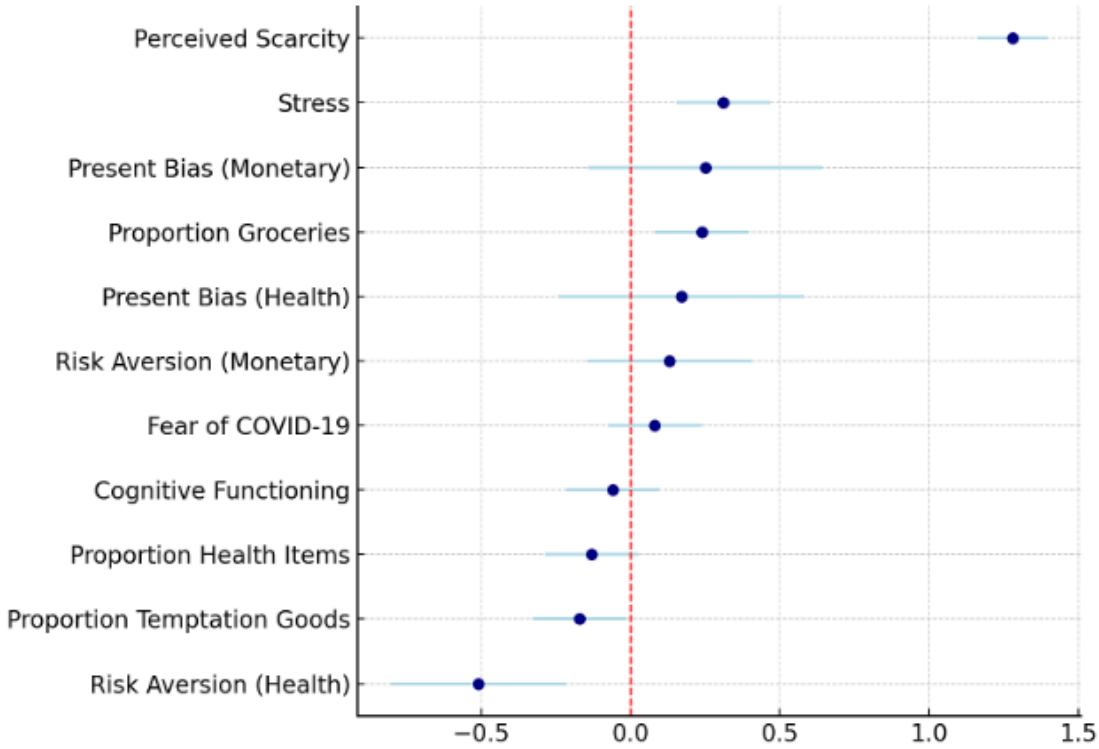
Figure 1 shows the estimated effects of lockdown exposure on key outcomes using regression models, which included demographic controls and propensity score matching weights (see Table S5 in Multimedia Appendix 2 for full regression model results). Participants during lockdown reported substantially higher perceived scarcity (mean 7.97 (SD 1.54)) than those after lockdown (mean 4.35 (SD 2.21)) (detailed group comparisons and mean differences for all outcome measures are provided in Table S4 in Multimedia Appendix 2). This difference is significant in OLS regression models (standardized β coefficient =1.28, *P*<.001), indicating that the lockdown created a pronounced sense of resource scarcity. Stress levels were markedly elevated during the lockdown (mean 2.76 (SD 0.70)) versus post-lockdown (mean 2.55 (SD 0.68)) and the increase was significant (standardized β coefficient=.31, *P*<.001). Fear of COVID-19 was slightly higher after the lockdown (mean 4.39 (SD 2.11)) compared to during the lockdown (mean 4.05 (SD 2.06)), though this difference was not significant in regression models. No significant differences were observed in cognitive functioning between the 2 groups in terms of their accuracy on the Raven's Matrices task (standardized β coefficient= –.06, *P*=.49).

In monetary-framed choices, the proportion of participants exhibiting present bias was 54.3% (119/219) in the During group and 49.1% (111/226) in the After group, with an insignificant difference. Similarly, in health-framed choices, the proportion was 47.0% (103/219) in the During group and 42.5% (96/226) in the After group, and the difference was also not significant. While monetary risk aversion showed a non-significant difference, health-related risk aversion was significantly greater during the lockdown than after (standardized β coefficient

XSL•FO
RenderX

=−.51, *P*<.001). Participants during the lockdown allocated a higher proportion of their hypothetical budget to groceries (227/332, 68.3%) compared to those after the lockdown (218/339, 64.2%). This difference was significant in regression results (standardized β coefficient =.24, *P*<.001). Conversely, spending on temptation goods was lower during the lockdown (46/332, 13.9%) than after (56/339, 16.5%), with regression analyses indicating a significant reduction (standardized β coefficient =-.17, *P*=.03). There was no significant difference in the proportion allocated to health-related goods between the 2 groups (standardized β coefficient =.13; *P*=.10).

Figure 1. Standardized regression coefficients showing the effects of COVID-19 lockdown (during vs after) on perceived scarcity, stress, fear of COVID-19, time preference (present bias), and risk preference (risk aversion) in monetary- and health-framed choices, and spending trade-offs between grocery, health, and temptation goods. Estimates are plotted with 95% confidence intervals. Each estimate was obtained from ordinary least squares regression except for present bias (logistic regression) and risk aversion (ordinal regression). All regression models applied propensity score-matching weights, controlled for demographic variables, including gender, age, education, monthly income, and monthly expenditure, and used robust standard errors.



Effects of Perceived Scarcity and Lockdown Duration

The cumulative effects of the lockdown were examined by analyzing how perceived scarcity and its interactions with lockdown status and lockdown duration affected various outcomes. To align with the timing of the lockdown, the analysis included participants who had quarantined within 120 days and those who had not shopped in the last 90 days. All variables were normalized to z-scores to facilitate comparison. Table 2 presents the results of regression analyses for the During and After groups. Perceived scarcity was significantly associated with increased stress in both During (standardized β coefficient =.62, *P*<.001) and After (standardized β coefficient =.65, *P*<.001) groups. Neither length of quarantine (days in quarantine for the During group and days after quarantine for the After group) nor days since last shop had significant direct effects on stress. Their interactions with perceived scarcity were also non-significant for both groups. These findings indicate that perceived scarcity remained a robust predictor of stress, independent of quarantine duration. Among participants surveyed during lockdown, perceived scarcity did not significantly affect fear of COVID-19 (standardized β coefficient =.04, *P*=.88), but its interaction with days in quarantine was significantly negative (standardized β coefficient =−1.25, *P*<.001), suggesting that the relationship between scarcity and fear diminished as quarantine duration increased. In the After-lockdown group, perceived scarcity significantly increased fear of COVID-19 (standardized β coefficient=.38, *P*<.001). During lockdown, perceived scarcity was associated with a slight enhancement in cognitive performance (standardized β coefficient =.80, *P*=.04), but as lockdown days extended, cognitive functioning decreased marginally (standardized β coefficient =−4.13, *P*=.06). After lockdown, perceived scarcity had no significant effect on cognitive functioning, while longer days post-lockdown were associated with improved cognitive performance (standardized β coefficient =2.22, *P*=.03).

For both groups, present bias in a monetary context showed limited significant relationships with perceived scarcity and lockdown duration, indicating that scarcity may not substantially alter time preferences in decision-making within these domains (Table 3). Present bias in the health context showed mixed results. For the During group, no significant effect was found. For the After group, days after quarantine significantly reduced health-related present bias (standardized β coefficient =−9.38, *P*=.01), and their interaction with scarcity is also significant (standardized β coefficient =−6.24, *P*=.01). Interaction between

XSL•FO
RenderX

perceived scarcity and days since last shopping also significantly reduced health-related present bias (standardized β coefficient =−3.02, *P*=.03). In general, neither perceived scarcity nor lockdown duration significantly influences risk aversion across monetary and health contexts. However, the interaction with days in quarantine is significantly positive (standardized β coefficient =4.68, *P*<.001), indicating scarcity's impact on increasing monetary risk aversion with longer quarantine durations.

Table 4 presents the regression results for budget allocations to groceries, health items, and temptation goods, assessing trade-off thinking in decision-making. During lockdown, perceived scarcity significantly increased the proportion of the budget allocated to groceries (standardized β coefficient =.67, *P*=.01) and reduced the proportion to health items (standardized β coefficient =−.61, *P*=.02). Even though longer quarantine duration led to decreased grocery spending (standardized β coefficient =−4.43, *P*<.001) and increased health spending (standardized β coefficient =3.98, *P*=.01), its interaction with scarcity significantly increased groceries spending (standardized β coefficient=1.52, *P*=.01) and decreased health spending (standardized β coefficient=−1.41, *P*<.001). In contrast, after lockdown, perceived scarcity and days after lockdown seem to reduce spending on groceries (standardized β coefficient = −0.35, *P*<.001) and increase spending on temptation goods (standardized β coefficient =.29, *P*=.01), while it did not significantly affect expenditures on health items (standardized β coefficient =.29, *P*=.12). But it is important to note that the overall effect sizes were small.

XSL•FO
RenderX

Table . Effects of perceived scarcity and lockdown duration on stress, fear, and cognitive functioning during and after the COVID-19 lockdown.[a]

Effect	Stress level			Fear of COVID-19			Cognitive functioning		
	b	SE	*P* value	b	SE	*P* value	b	SE	*P* value
Panel A: During lockdown									
Perceived scarcity	0.62	0.22	<.001	0.04	0.28	.88	0.80	0.39	.04
Days since quarantine[b]	−0.89	1.18	.45	1.63	1.53	.29	−4.13	2.15	.06
Perceived scarcity # days since quarantine	−0.47	0.55	.39	−1.25	0.37	<.001	0.64	0.37	.08
Days without shopping[c]	−0.08	0.69	.91	0.94	0.49	.06	1.05	0.57	.07
Perceived scarcity # days without shopping	−0.71	0.51	.17	−0.24	0.31	.44	−0.40	0.33	.23
Panel B: After lockdown									
Perceived scarcity	0.65	0.10	<.001	0.38	0.12	<.001	−0.03	0.12	.78
Days after quarantine	1.27	0.82	.12	−1.34	1.01	.19	2.22	1.00	.03
Perceived scarcity # days after quarantine	0.48	0.53	.37	0.21	0.57	.72	0.97	0.56	.09
Days without shopping	−0.41	0.40	.31	-0.22	0.66	.74	−0.80	0.83	.33
Perceived scarcity # days without shopping	−0.04	0.29	.89	0.00	0.45	.99	−0.57	0.56	.31

[a] Each estimate was obtained from ordinary least squares regression. All regression models applied propensity score-matching weights, controlled for demographic variables, including gender, age, education, monthly income, and monthly expenditure, and used robust standard errors.

[b]To align with the timing of the lockdown, the analysis of “Days since quarantine” included only those who had quarantined (“During” group, n=326) or had not quarantined (“After” group, n=259) within 120 days.

[c]The analysis of “Days without shopping” was limited to individuals who had not shopped in the last 90 days for both the During (n=247) and After (n=124) groups.

XSL•FO
RenderX

Table . Effects of perceived scarcity and lockdown duration on time and risk preferences during and after the COVID-19 lockdown.[ab],

Effect	Present bias (monetary)[c]			Present bias (health)[c]			Risk aversion (monetary)			Risk aversion (health)		
	b	SE	*P* value	b	SE	*P* value	b	SE	*P* value	b	se	*P* value
Panel A: During lockdown												
Perceived scarcity	1.05	0.84	.21	0.17	0.87	.85	0.32	0.65	.62	−0.03	0.52	.96
Days since quarantine	−4.46	4.48	.32	0.15	4.70	.97	−5.06	3.99	.20	−1.50	3.30	.65
Perceived scarcity # days since quarantine	0.34	1.34	.80	1.28	1.34	.34	4.68	1.44	<.001	−0.23	1.90	.90
Days without shopping	0.30	1.88	.87	−0.17	1.85	.93	−3.13	1.84	.09	−1.50	2.04	.46
Perceived scarcity # days without shopping	1.05	1.41	.46	0.77	1.39	.58	2.68	1.51	.08	−0.34	1.65	.84
Panel B: After lockdown												
Perceived scarcity	−0.42	0.32	.19	−0.15	0.32	.64	−0.00	0.24	.99	−0.13	0.26	.61
Days after quarantine	−6.00	3.40	.08	−9.38	3.73	.01	−0.19	1.99	.92	−3.44	1.93	.07
Perceived scarcity # days after quarantine	−3.80	2.07	.07	−6.24	2.28	.01	−1.20	1.23	.33	−2.67	1.11	.02
Days without shopping	−2.67	2.07	.20	−4.04	2.17	.06	−1.96	1.21	.10	−0.89	1.16	.44
Perceived scarcity # days without shopping	−2.70	1.37	.05	−3.02	1.42	.03	−1.55	0.83	.06	−1.08	0.81	.18

[a]Logistic regressions were performed for present bias, and ordinal regressions were performed for risk aversion.

[b]All regression models applied propensity score-matching weights, controlled for demographic variables, including gender, age, education, monthly income, and monthly expenditure, and used robust standard errors.

[c]The sample for Present bias was limited to participants ((“During” group, n=219; “After” group, n=226) who consistently chose either the earlier or later payment in both monetary and health time-discounting tasks, switching no more than once between options.

XSL•FO
RenderX

Table . Effects of perceived scarcity and lockdown duration on grocery, health, and temptation goods during and after the COVID-19 lockdown.[ab]

Effect	Proportion of groceries			Proportion of health items			Proportion of temptation goods		
	b	SE	*P* value	b	SE	*P* value	b	SE	*P* value
Panel A: During lockdown									
Perceived scarcity	0.67	0.27	.01	−0.61	0.26	.02	−0.29	0.29	.32
Days since quarantine	−4.43	1.41	<.001	3.98	1.46	.01	1.96	1.57	.21
Perceived scarcity # days since quarantine	1.52	0.54	.01	-1.41	0.42	<.001	−0.64	0.51	.21
Days without shopping	−1.01	0.79	.20	0.54	0.73	.46	0.75	0.72	.30
Perceived scarcity # days without shopping	0.77	0.59	.19	−0.64	0.57	.26	−0.38	0.53	.47
Panel B: After lockdown									
Perceived scarcity	−0.35	0.12	<.001	0.15	0.13	.24	0.29	0.12	.01
Days after quarantine	−1.82	1.01	.07	0.06	1.09	.96	2.10	1.06	.05
Perceived scarcity # days after quarantine	−1.07	0.57	.06	0.36	0.65	.58	0.97	0.60	.11
Days without shopping	0.05	0.55	.93	−0.02	0.44	.97	−0.04	0.69	.95
Perceived scarcity # days without shopping	−0.32	0.39	.41	0.16	0.30	.59	0.24	0.47	.61

[a]Each estimate was obtained from ordinary least squares regression.

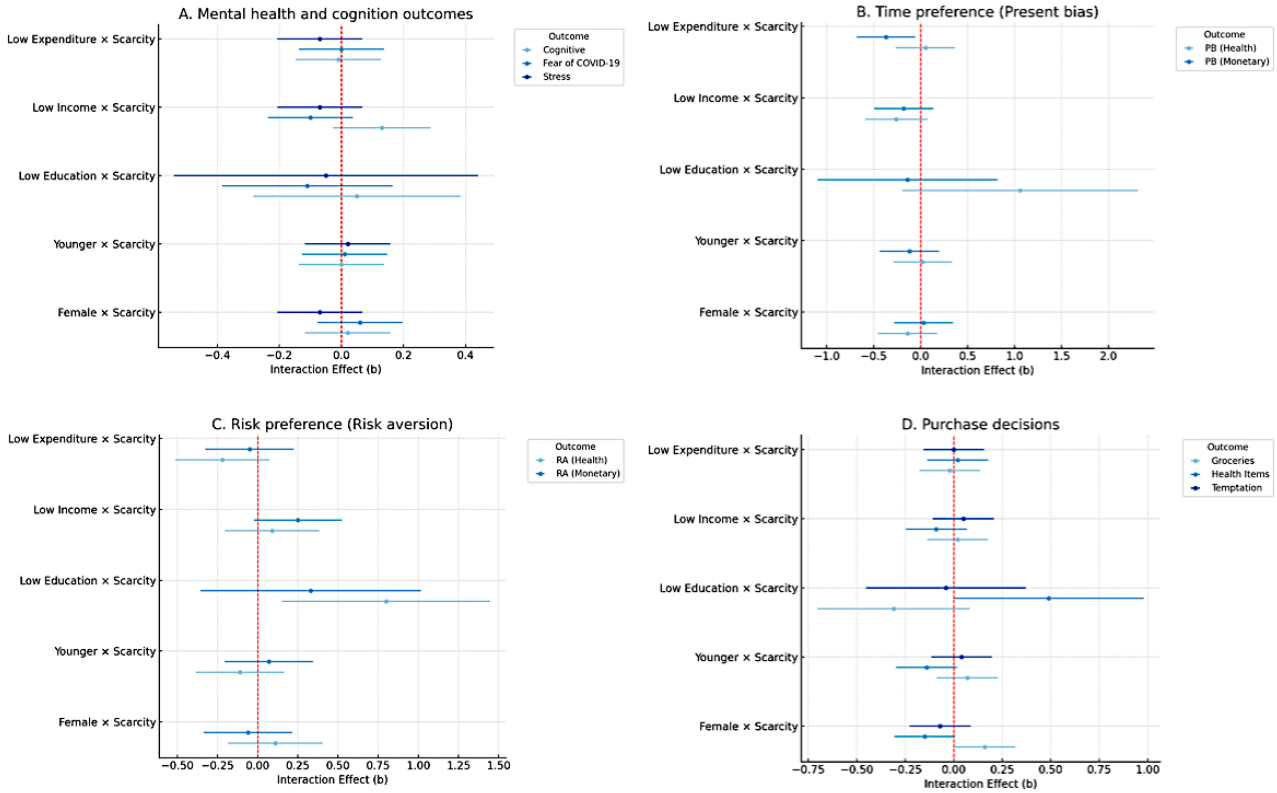
[b]All regression models applied propensity score-matching weights, controlled for demographic variables, including gender, age, education, monthly income, and monthly expenditure, and used robust standard errors.

Subgroup Analysis

Subgroup analyses revealed that the effect of perceived scarcity varied across demographic and economic groups (Figure 2). Women reallocated their budgets under scarcity by increasing grocery spending (standardized β coefficient=.16, *P*=.04) and reducing spending on health items (standardized β coefficient = –.15, *P*=.05). Low-education participants showed stronger protective behaviors: scarcity increased both their health-related risk aversion (standardized β coefficient =.80, *P*=.01) and health spending (standardized β coefficient =.49, *P*=.05). Interestingly, low-expenditure participants became less present-biased under scarcity (standardized β coefficient = –.37, *P*=.02), suggesting greater attentional focus on future orientation when facing resource constraints. In contrast, subgroup interactions with age and income showed no consistent or significant moderation effects. Full regression results for subgroup analysis are provided in Table S7 in Multimedia Appendix 2.

XSL•FO
RenderX

Figure 2. Subgroup analysis of interaction effects between perceived scarcity and demographic subgroups on perceived scarcity, stress, fear of COVID-19, time and risk preference, and spending trade-offs between grocery, health, and temptation. Estimates are plotted with 95% confidence intervals. Each estimate was obtained from ordinary least squares regression, except for present bias (logistic regression) and risk aversion (ordinal regression). All outcomes and continuous predictors were standardized (z-scores), and models adjusted for demographic covariates using propensity score-matching weights. Robust standard errors were used across all regressions.



Discussion

Principal Findings

This study examined the psychological and behavioral impacts of perceived scarcity during and after the COVID-19 lockdown in Shanghai. By analyzing variables such as stress, fear of COVID-19, cognitive functioning, present bias, risk aversion, and budget allocation, the study aimed to understand how perceived scarcity arising from resource constraints influences mental health and decision-making processes during a pandemic. The findings provide empirical evidence on the nuanced effects of scarcity in a real-world crisis, advancing understanding of its impact on various psychological and behavioral outcomes.

Perceived scarcity and stress were significantly higher during the lockdown compared to after, aligning with existing literature that lockdown measures exacerbate mental health problems [6-8]. In both groups, individuals who perceived more scarcity reported higher stress, reinforcing prior research showing that poverty and resource constraints elevate stress levels [21,63]. The results extend this understanding by showcasing that this relationship holds true even in a relatively affluent, urban setting during a temporary crisis. The study highlighted a dose-response-type relationship: those who felt more scarcity had higher stress, a pattern also observed in other contexts of financial hardship and food insecurity [44]. The study found no significant increase in fear of COVID-19 during lockdown compared to after, while a significant negative interaction between perceived scarcity and days in quarantine was observed during lockdown. This may indicate that as the duration of home quarantine extends and individuals are less exposed to the pandemic environment, their perception of fear diminishes. The cognitive load theory offers one possible explanation for this link. In this case, many Shanghai residents faced difficulties obtaining groceries, leading to both cognitive (eg, time spent securing food) and emotional strain. While no significant group differences in cognitive functioning were found, interaction analyses revealed that longer quarantine was associated with lower cognitive performance during lockdown, and post-lockdown durations predicted improvement. These patterns align with longitudinal evidence on the cumulative effects of scarcity on cognition [30,64]. However, the cross-sectional design restricts the ability to infer a causal role of cognitive function in linking scarcity to stress and fear.

The stability of time preferences reported in both monetary- and health-framed choices is consistent with study [65] but contradicts study [66] conducted during COVID-19. Regarding the effects of perceived scarcity and lockdown duration, present bias was generally unchanged, except in health-framed choices post-lockdown, where individuals displayed a reduced present bias and became more future-oriented. However, the small sample sizes (n=91 for the During group; n=64 for the After group) in the analyses limit the confidence in these conclusions. Further research is necessary to confirm and clarify these trends. Risk preference in monetary-framed choices remained stable, aligning with previous studies [65,67,68]. In contrast, within health-framed choices, the During group exhibited greater risk

XSL•FO
RenderX

aversion compared to the After group. This may be because the risk preference questions focused on vaccination, which was a more urgent and relevant health concern during the lockdown period. Furthermore, perceived scarcity, when interacting with days in quarantine, increased monetary risk aversion in the During group, while its interaction with days after quarantine decreased health-related risk aversion in the After group. Together, these findings indicate that perceived scarcity and the timing of lockdown influence individuals' time and risk preferences differently across monetary and health-related contexts, highlighting the importance of context in understanding preferences and biases.

Lastly, when faced with trade-offs in purchasing decisions during the lockdown, participants prioritized spending on groceries over health items, reflecting immediate survival needs taking precedence under resource constraints [10,14,15]. This shift in spending priorities toward essential goods is also in line with findings from studies on purchase behaviors during COVID-19, such as online grocery shopping [12] and panic buying [11]. After lockdown, scarcity was associated with increased spending on temptation goods, indicating a shift toward non-essential purchases when restrictions eased. These changes highlight the dynamic nature of consumer behavior in response to external stressors and the role of scarcity in influencing trade-off decisions.

Our subgroup analyses also have important implications on the impacts of scarcity. Women, for instance, appeared to adjust their spending priorities to groceries under scarcity more than men. This could reflect gender roles in household provisioning, or possibly that women were willing to sacrifice personal or preventive health expenditures to ensure basic needs are met. Lower-educated participants showed stronger risk aversion to scarcity, which might indicate that individuals with less formal education react to scarcity by opting for safer choices, potentially neglecting larger or long-term benefits. This could be driven by a heightened perception of vulnerability. In contrast, neither age nor income level significantly moderated scarcity effects in our data, suggesting that individuals were affected similarly regardless of their age or economic status.

Policy and Public Health Implications

The findings mentioned earlier offer handfuls of insights for policy makers and public health officials aiming to mitigate the negative consequences of perceived scarcity in crises. First, addressing the material needs and essential goods is paramount. Government should invest in resilient supply chains and emergency distribution networks, both online and offline, to facilitate access. For instance, establishing community distribution centers or partnering with grocery stores to deliver essentials to vulnerable populations can ensure that basic needs are met. Implementing real-time monitoring systems to track inventory levels and demand patterns can also help in proactive resource distribution. Using examples like Singapore's successful governance, through food supply chain management, timely and transparent communication, and maintenance of public health service delivery, could help curb panic buying and reassure the public [69,70]. This study suggests that these proactive measures could greatly reduce perceived scarcity. If people believe essentials will remain available, they may experience less panic, less stress, and continue more rational decision patterns (eg, not overspending on groceries at the cost of medicine). Special attention should be given to subgroups who adjust their spending most under scarcity (eg, women managing household groceries) to ensure they have avenues for obtaining both food and health items so that critical needs (like medicines) are not sacrificed.

Second, findings suggest that mental health interventions may be necessary both during and after quarantine. Interventions aimed at enhancing coping strategies may be effective in reducing stress, anxiety, and fear during extended lockdowns [71]. This could involve providing virtual mental health services, such as online counseling or stress management workshops, to help individuals develop resilience [72]. To mitigate mental health challenges both during and after quarantine, governments should invest in digital mental health tools and offer support resources accessible to the public. Government can also partner with app developers to create mental health applications specifically designed for crisis situations, incorporating features such as daily wellness check-ins, stress management techniques, and mindfulness exercises [73]. Additionally, special attention should be given to vulnerable groups (eg, those who have lower income or lost income during the lockdown, people living alone, or those with pre-existing mental health issues) as they might experience compounded effects of scarcity and isolation. Proactive outreach and targeted interventions (such as providing remote counseling and peer support) for these groups could mitigate long-term mental health consequences. Our subgroup results also suggest tailoring support by demographic segment: for example, lower education individuals (who showed strong protective responses) may benefit from clear guidance to reinforce effective health behaviors without undue anxiety.

Third, public health communication plays a crucial role in managing perceived scarcity. Transparent updates on supply levels and distribution timelines can significantly reduce uncertainty-driven stress [74,75]. Public health campaigns should acknowledge inherent biases in individual decision-making and tailor messages accordingly. For instance, acknowledging that people become more risk-averse in health decisions during a pandemic (as observed in this study) could inform how to frame health communication messages. This could involve creating personalized messages that highlight potential health losses during times when risk aversion in health contexts is heightened. Post-lockdown, public health campaigns can promote balanced spending, encouraging individuals to maintain healthy consumption patterns. Another communications strategy is to combat misinformation that can exacerbate perceived scarcity (eg, rumors of food shortages). Engaging credible sources, community leaders, or even social media influencers to spread factual, calming information can prevent panic behaviors [76].

Finally, findings on spending trade-offs suggest a need for economic measures to support individuals under resource strain. If people are cutting back on health-related spending to afford groceries, governments or NGOs could introduce programs like vouchers for health supplies or food assistance to remove the dilemma of choosing between food and medicine. Post-crisis,

XSL•FO
RenderX

stimulating the economy in a way that encourages a balanced return to normal consumption, perhaps through incentives for purchasing health-related goods or campaigns to promote healthy lifestyles, could be beneficial. In summary, policy interventions should tackle both the objective and subjective aspects of scarcity: ensuring actual availability of resources and concurrently addressing the psychology of scarcity by offering support and communication that reduce uncertainty and stress.

Limitations and Future Research

This study has limitations that suggest avenues for future research. First, the cross-sectional design, comparing 2 independent groups, limits causal inference or change tracking within individuals over time. This study leveraged a quasi-natural experimental design, which limited the ability to examine certain assumptions (eg, parallel trends in outcomes prior to the lockdown) that would strengthen causal claims. Future studies could use a longitudinal design, tracking the same participants across different stages of a crisis, which would offer insights into how perceived scarcity impacts individuals over time. Second, it is important to acknowledge the sampling method in the given context of a period of lockdown in a single city area may not fully capture the representativeness of the population (all adults in Shanghai during the lockdown) and fall short in addressing the inherent uncertainty in generalizing these results to other populations. The inference in this study is bounded by the specific population and sampling conditions of this study. Furthermore, the reliance on self-reported data introduces potential biases, such as recall bias or social desirability bias. Future research could incorporate objective data, such as actual purchasing behaviors through transaction records or physiological stress indicators like cortisol levels. Additionally, the use of incentivized tasks for measuring time and risk preferences could provide more accurate assessments of decision-making behaviors. Another limitation is the limited sample size for specific analyses, particularly those related to present bias, which may affect result reliability. Future studies should aim for larger samples to confirm these findings and enhance statistical power. Broadening participants from different regions and cultural backgrounds can also help generalize the findings and explore cultural influences on the effects of scarcity. Finally, this research highlights the need to further examine how perceived scarcity shapes mental health and decision-making in the long term, especially as crises like pandemics and climate-related events may become more frequent. Future research could investigate the potential for lasting behavioral changes induced by scarcity, such as persistent shifts in spending priorities and purchase behaviors.

Conclusions

This paper provides a nuanced understanding of the psychological and behavioral impacts of perceived scarcity during a pandemic, focusing on an urban population during and after the COVID-19 lockdown in Shanghai. The findings reveal that scarcity significantly influences stress levels, cognitive function, financial risk aversion, and budget allocation, while fear and present bias show more moderate effects. From a theoretical perspective, the study extends the concept of perceived scarcity to encompass situational constraints experienced by higher income populations during crises and advances current literature by demonstrating that the psychological and behavioral effects of perceived scarcity are multifaceted and context dependent. Therefore, this study advances theoretical understanding by highlighting the domain-specific effects of perceived scarcity and the resilience observed in cognitive functioning among highly educated populations. From a policy perspective, findings provide evidence for policy makers to design interventions that address both the material and psychological needs of populations during times of scarcity. By implementing targeted strategies, such as leveraging technology for resource management, providing mental health support, and tailoring communication to individual biases, governments can mitigate the negative impacts of scarcity and enhance resilience in future crises.

Acknowledgments

This work was supported by the University Research Studentship (GS1015 5277) and the Dean’s Research Bursary from the School of Design and Creative Arts at Loughborough University. HZ is supported by the Wellcome Centre for Ethics and Humanities, which is supported by core funding from the Wellcome Trust (203132/Z/16/Z).

Data Availability

De-identified participant data and analysis code for this study are available on the Open Science Framework (OSF) [77].

Authors' Contributions

HZ led the conceptualization, data collection and analysis, drafting, and revision; FY contributed to data collection, drafting, and revision; TG, HD, and CDB contributed to conceptualization, drafting, and revision. All authors reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Web-based survey questions.
[DOCX File, 883 KB - publichealth_v11i1e69496_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Supplementary tables.
[DOCX File, 88 KB - publichealth_v11i1e69496_app2.docx]

References

1. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature New Biol 2020 Aug;584(7820):257-261. [doi: 10.1038/s41586-020-2405-7] [Medline: 32512579]
2. Liu Y, Wang W, Wong WK, Zhu W. Effectiveness of non-pharmaceutical interventions for COVID-19 in USA. Sci Rep 2024;14(1):21387. [doi: 10.1038/s41598-024-71984-1]
3. Yang M, He Z, Zhang Y, Liu T, Ming WK. Pandemic fatigue and preferences for COVID-19 public health and social measures in China: nationwide discrete choice experiment. JMIR Public Health Surveill 2024 Jun 27;10:e45840. [doi: 10.2196/45840] [Medline: 38935420]
4. Sugaya N, Yamamoto T, Suzuki N, Uchiumi C. Loneliness and social isolation factors under the prolonged COVID-19 pandemic in Japan: 2-year longitudinal study. JMIR Public Health Surveill 2024 Sep 9;10(1):e51653. [doi: 10.2196/51653] [Medline: 39250195]
5. Lei L, Huang X, Zhang S, Yang J, Yang L, Xu M. Comparison of prevalence and associated factors of anxiety and depression among people affected by versus people unaffected by quarantine during the COVID-19 epidemic in Southwestern China. Med Sci Monit 2020 Apr 26;26:e924609. [doi: 10.12659/MSM.924609] [Medline: 32335579]
6. Serrano-Alarcón M, Kentikelenis A, Mckee M, Stuckler D. Impact of COVID-19 lockdowns on mental health: evidence from a quasi-natural experiment in England and Scotland. Health Econ 2022 Feb;31(2):284-296. [doi: 10.1002/hec.4453] [Medline: 34773325]
7. Pierce M, Hope H, Ford T, et al. Mental health before and during the COVID-19 pandemic: a longitudinal probability sample survey of the UK population. Lancet Psychiatry 2020 Oct;7(10):883-892. [doi: 10.1016/S2215-0366(20)30308-4] [Medline: 32707037]
8. Robinson E, Sutin AR, Daly M, Jones A. A systematic review and meta-analysis of longitudinal cohort studies comparing mental health before versus during the COVID-19 pandemic in 2020. J Affect Disord 2022 Jan 1;296:567-576. [doi: 10.1016/j.jad.2021.09.098] [Medline: 34600966]
9. Carroll N, Sadowski A, Laila A, et al. The impact of COVID-19 on health behavior, stress, financial and food security among middle to high income canadian families with young children. Nutrients 2020 Aug 7;12(8):2352. [doi: 10.3390/nu12082352] [Medline: 32784530]
10. Posel D, Oyenubi A, Kollamparambil U. Job loss and mental health during the COVID-19 lockdown: evidence from South Africa. PLoS ONE 2021;16(3):e0249352. [doi: 10.1371/journal.pone.0249352] [Medline: 33784339]
11. Yuen KF, Tan LS, Wong YD, Wang X. Social determinants of panic buying behaviour amidst COVID-19 pandemic: the role of perceived scarcity and anticipated regret. Journal of Retailing and Consumer Services 2022 May;66:102948. [doi: 10.1016/j.jretconser.2022.102948]
12. Budziński W, Daziano R. Preferences for online grocery shopping during the COVID-19 pandemic - the role of fear-related attitudes. J Choice Model 2023 Jun;47:100416. [doi: 10.1016/j.jocm.2023.100416] [Medline: 37207091]
13. Schurer S, Atalay K, Glozier N, Vera-Toscano E, Wooden M. Quantifying the human impact of Melbourne's 111-day hard lockdown experiment on the adult population. Nat Hum Behav 2023 Oct;7(10):1652-1666. [doi: 10.1038/s41562-023-01638-1] [Medline: 37653145]
14. Harling G, Gómez-Olivé FX, Tlouyamma J, et al. Protective behaviors and secondary harms resulting from nonpharmaceutical interventions during the COVID-19 epidemic in South Africa: multisite, prospective longitudinal study. JMIR Public Health Surveill 2021 May 13;7(5):e26073. [doi: 10.2196/26073] [Medline: 33827046]
15. Liao Q, Xiao J, Cheung J, et al. Community psychological and behavioural responses to coronavirus disease 2019 over one year of the pandemic in 2020 in Hong Kong. Sci Rep 2021 Nov 18;11(1):22480. [doi: 10.1038/s41598-021-00616-9] [Medline: 34795312]
16. Prati G, Mancini AD. The psychological impact of COVID-19 pandemic lockdowns: a review and meta-analysis of longitudinal studies and natural experiments. Psychol Med 2021 Jan;51(2):201-211. [doi: 10.1017/S0033291721000015] [Medline: 33436130]
17. O'Connor RC, Wetherall K, Cleare S, et al. Mental health and well-being during the COVID-19 pandemic: longitudinal analyses of adults in the UK COVID-19 Mental Health & Wellbeing study. Br J Psychiatry 2021 Jun;218(6):326-333. [doi: 10.1192/bjp.2020.212] [Medline: 33081860]
18. Chan CP, Lee SS, Kwan TH, Wong SYS, Yeoh EK, Wong NS. Population behavior changes underlying phasic shifts of SARS-CoV-2 exposure settings across 3 omicron epidemic waves in Hong Kong: prospective cohort study. JMIR Public Health Surveill 2024 Jun 19;10(1):e51498. [doi: 10.2196/51498] [Medline: 38896447]
19. de Vocht F, Katikireddi SV, McQuire C, Tilling K, Hickman M, Craig P. Conceptualising natural and quasi experiments in public health. BMC Med Res Methodol 2021 Feb 11;21(1):32. [doi: 10.1186/s12874-021-01224-x] [Medline: 33573595]
20. Campbell DT, Stanley JC. Experimental and Quasi-Experimental Designs for Research: Ravenio Books; 2015. URL: https://www.scirp.org/reference/referencespapers?referenceid=2026957 [accessed 2025-08-22]

XSL•FO
RenderX

21. Haushofer J, Fehr E. On the psychology of poverty. Science 2014 May 23;344(6186):862-867. [doi: 10.1126/science.1232491] [Medline: 24855262]
22. de Bruijn EJ, Antonides G. Poverty and economic decision making: a review of scarcity theory. Theory Decis 2022 Feb;92(1):5-37. [doi: 10.1007/s11238-021-09802-7]
23. Adamkovič M, Martončik M. A review of consequences of poverty on economic decision-making: a hypothesized model of a cognitive mechanism. Front Psychol 2017;8(OCT):1784. [doi: 10.3389/fpsyg.2017.01784] [Medline: 29075221]
24. Mani A, Mullainathan S, Shafir E, Zhao J. Poverty impedes cognitive function. Science 2013 Aug 30;341(6149):976-980. [doi: 10.1126/science.1238041] [Medline: 23990553]
25. Goldsmith K, Griskevicius V, Hamilton R. Scarcity and consumer decision making: is scarcity a mindset, a threat, a reference point, or a journey? Journal of the Association for Consumer Research 2020 Oct 1;5(4):358-364. [doi: 10.1086/710531]
26. Shah AK, Zhao J, Mullainathan S, Shafir E. Money in the mental lives of the poor. Soc Cogn 2018 Feb;36(1):4-19. [doi: 10.1521/soco.2018.36.1.4]
27. Hamilton R, Thompson D, Bone S, et al. The effects of scarcity on consumer decision journeys. J of the Acad Mark Sci 2019 May;47(3):532-550. [doi: 10.1007/s11747-018-0604-7]
28. McGuire J, Kaiser C, Bach-Mortensen AM. A systematic review and meta-analysis of the impact of cash transfers on subjective well-being and mental health in low- and middle-income countries. Nat Hum Behav 2022 Mar;6(3):359-370. [doi: 10.1038/s41562-021-01252-z] [Medline: 35058643]
29. Ridley M, Rao G, Schilbach F, Patel V. Poverty, depression, and anxiety: causal evidence and mechanisms. Science 2020 Dec 11;370(6522):eaay0214. [doi: 10.1126/science.aay0214] [Medline: 33303583]
30. Lichand G, Mani A. Cognitive droughts. SSRN Journal 2020. [doi: 10.2139/ssrn.3540149]
31. Mertens G, Lodder P, Smeets T, Duijndam S. Pandemic panic? Results of a 14-month longitudinal study on fear of COVID-19. J Affect Disord 2023 Feb 1;322:15-23. [doi: 10.1016/j.jad.2022.11.008] [Medline: 36372124]
32. Luo F, Ghanei Gheshlagh R, Dalvand S, Saedmoucheshi S, Li Q. Systematic review and meta-analysis of fear of COVID-19. Front Psychol 2021;12:661078. [doi: 10.3389/fpsyg.2021.661078] [Medline: 34177712]
33. Pfeifer LS, Heyers K, Ocklenburg S, Wolf OT. Stress research during the COVID-19 pandemic and beyond. Neurosci Biobehav Rev 2021 Dec;131:581-596. [doi: 10.1016/j.neubiorev.2021.09.045] [Medline: 34599918]
34. Kowal M, Coll-Martín T, Ikizer G, et al. Who is the most stressed during the COVID-19 pandemic? Data from 26 countries and areas. Appl Psychol Health Well Being 2020 Dec;12(4):946-966. [doi: 10.1111/aphw.12234] [Medline: 32996217]
35. Demont T, Horta-Sáenz D, Raiber E. Exposure to worrisome topics can increase cognitive performance when incentivized by a performance goal. Sci Rep 2024 Jan 12;14(1):1204. [doi: 10.1038/s41598-023-50036-0] [Medline: 38216640]
36. Tversky A, Kahneman D. Judgment under uncertainty: heuristics and biases. Science 1974 Sep 27;185(4157):1124-1131. [doi: 10.1126/science.185.4157.1124] [Medline: 17835457]
37. Frederick S, Loewenstein G, O'donoghue T. Time discounting and time preference: a critical review. J Econ Lit 2002 Jun;40(2):351-401. [doi: 10.1257/002205102320161311]
38. O'Donoghue T, Rabin M. Doing it now or later. American Economic Review 1999 Mar 1;89(1):103-124. [doi: 10.1257/aer.89.1.103]
39. Laibson D. Golden eggs and hyperbolic discounting. Q J Econ 1997 May 1;112(2):443-478. [doi: 10.1162/003355397555253]
40. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. In: Handbook of the Fundamentals of Financial Decision Making: Part I World Scientific 2013:99-127. [doi: 10.1142/9789814417358_0006]
41. Harrison GW, Lau MI, Rutström EE. Individual discount rates and smoking: evidence from a field experiment in Denmark. J Health Econ 2010 Sep;29(5):708-717. [doi: 10.1016/j.jhealeco.2010.06.006] [Medline: 20727602]
42. Ida T. A quasi-hyperbolic discounting approach to smoking behavior. Health Econ Rev 2014;4(1):1-11. [doi: 10.1186/s13561-014-0005-7] [Medline: 25006542]
43. Ikeda S, Kang MI, Ohtake F. Hyperbolic discounting, the sign effect, and the body mass index. In: Behavioral Economics of Preferences, Choices, and Happiness: Springer; 2016:277-313. [doi: 10.1007/978-4-431-55402-8_12]
44. Dalton PS, Nhung N, Rüschenpöhler J. Worries of the poor: the impact of financial burden on the risk attitudes of micro-entrepreneurs. J Econ Psychol 2020 Aug;79(December 2018):102198. [doi: 10.1016/j.joep.2019.102198]
45. Bartoš V, Bauer M, Chytilová J, Levely I. Psychological effects of poverty on time preferences. The Economic Journal 2021 Aug 1;131(638):2357-2382. [doi: 10.1093/ej/ueab007]
46. Carvalho LS, Meier S, Wang SW. Poverty and economic decision-making: evidence from changes in financial resources at Payday. Am Econ Rev 2016 Feb;106(2):260-284. [doi: 10.1257/aer.20140481] [Medline: 28003681]
47. Burlacu S, Mani A, Ronzani P, Savadori L. The preoccupied parent: how financial concerns affect child investment choices. J Behav Exp Econ 2023;105:102030. [doi: 10.1016/j.socec.2023.102030]
48. Lichand G, Bettinger E, Cunha N, Madeira R. The psychological effects of poverty on investments in children's human capital. SSRN Journal 2020. [doi: 10.2139/ssrn.3644127]
49. Burlacu S, Kažemekaitytė A, Ronzani P, Savadori L. Blinded by worries: sin taxes and demand for temptation under financial worries. Theory Decis 2022 Feb;92(1):141-187. [doi: 10.1007/s11238-021-09820-5]
50. Park YR, Ho Y, Hallez K, Kaur S, Srinivasan M, Zhao J. The psychology of poverty: current and future directions. Curr Dir Psychol Sci 2025 Feb;34(1):21-28. [doi: 10.1177/09637214241283163]

XSL•FO
RenderX

51. Craig P, Cooper C, Gunnell D, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. J Epidemiol Community Health 2012 Dec;66(12):1182-1186. [doi: 10.1136/jech-2011-200375] [Medline: 22577181]
52. Craig P, Katikireddi SV, Leyland A, Popham F. Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. Annu Rev Public Health 2017 Mar 20;38:39-56. [doi: 10.1146/annurev-publhealth-031816-044327] [Medline: 28125392]
53. Dunning T. Natural Experiments in the Social Sciences: A Design-Based Approach: Cambridge University Press; 2012. [doi: 10.1017/CBO9781139084444]
54. William K. China market research tool. Medium. 2018 Aug 17. URL: https://williamkuo1988.medium.com/china-market-research-tool-e612d559b649 [accessed 2024-09-27]
55. Byun SE, Sternquist B. The antecedents of in-store hoarding: measurement and application in the fast fashion retail environment. The International Review of Retail, Distribution and Consumer Research 2008 May;18(2):133-147. [doi: 10.1080/09593960701868241]
56. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
57. Ahorsu DK, Lin CY, Imani V, Saffari M, Griffiths MD, Pakpour AH. The fear of COVID-19 scale: development and initial validation. Int J Ment Health Addict 2022;20(3):1537-1545. [doi: 10.1007/s11469-020-00270-8] [Medline: 32226353]
58. Raven J. The Raven's progressive matrices: change and stability over culture and time. Cogn Psychol 2000 Aug;41(1):1-48. [doi: 10.1006/cogp.1999.0735] [Medline: 10945921]
59. Schilbach F, Schofield H, Mullainathan S. The psychological lives of the poor. Am Econ Rev 2016 May;106(5):435-440. [doi: 10.1257/aer.p20161101] [Medline: 29547249]
60. Eckel CC, Grossman PJ. Sex differences and statistical stereotyping in attitudes toward financial risk. Evol Hum Behav 2002 Jul;23(4):281-295. [doi: 10.1016/S1090-5138(02)00097-1]
61. Tversky A, Kahneman D. The framing of decisions and the psychology of choice. Science 1981 Jan 30;211(4481):453-458. [doi: 10.1126/science.7455683] [Medline: 7455683]
62. Fehr D, Fink G, Jack BK. Poor and rational: decision-making under scarcity. Journal of Political Economy 2022 Nov 1;130(11):2862-2897. [doi: 10.1086/720466]
63. Haushofer J, Shapiro J. The short-term impact of unconditional cash transfers to the poor: experimental evidence from Kenya. Q J Econ 2016 Nov;131(4):1973-2042. [doi: 10.1093/qje/qjw025] [Medline: 33087990]
64. Mani A, Mullainathan S, Shafir E, Zhao J. Scarcity and cognitive function around payday: a conceptual and empirical analysis. Journal of the Association for Consumer Research 2020 Oct 1;5(4):365-376. [doi: 10.1086/709885]
65. Drichoutis AC, Nayga RM. On the stability of risk and time preferences amid the COVID-19 pandemic. Exp Econ 2022;25(3):759-794. [doi: 10.1007/s10683-021-09727-6] [Medline: 34404975]
66. Li KK, Huang B, Tam T, Hong YY. Does the COVID-19 pandemic affect people's social and economic preferences? Evidence from China. SSRN Journal 2020. [doi: 10.2139/ssrn.3690072]
67. Zhang P, Palma MA. Stability of risk preferences during COVID-19: evidence from four measurements. Front Psychol 2021;12:702028. [doi: 10.3389/fpsyg.2021.702028] [Medline: 35222133]
68. Angrisani M, Cipriani M, Guarino A, Kendall R, Ortiz de Zarate J. Risk preferences at the time of COVID-19: an experiment with professional traders and students. SSRN Journal 2020. [doi: 10.2139/ssrn.3609586]
69. Tortajada C, Lim NSW. Food security and COVID-19: impacts and resilience in Singapore. Front Sustain Food Syst 2021;5:740780. [doi: 10.3389/fsufs.2021.740780]
70. Chua AQ, Tan MMJ, Verma M, et al. Health system resilience in managing the COVID-19 pandemic: lessons from Singapore. BMJ Glob Health 2020 Sep;5(9):e003317. [doi: 10.1136/bmjgh-2020-003317] [Medline: 32938609]
71. Chen L, Chu Q, Xu C, et al. Psychological responses and factors associated with depression and anxiety in entry personnel under quarantine during pandemic in China. Front Public Health 2024;12:1368463. [doi: 10.3389/fpubh.2024.1368463]
72. Escobar-Agreda S, Silva-Valencia J, Soto-Becerra P, et al. Patient engagement with and perceptions of the COVIDA project, a volunteer-led telemonitoring and teleorientation service for COVID-19 community management: mixed methods study. JMIR Form Res 2024 Sep 13;8(1):e51237. [doi: 10.2196/51237] [Medline: 39269741]
73. Maloney S, Montero-Marin J, Kuyken W. Mindfulness-based cognitive therapy-taking it further (MBCT-TiF) compared to ongoing mindfulness practice (OMP) in the promotion of well-being and mental health: a randomised controlled trial with graduates of MBCT and MBSR. Behav Res Ther 2024 Feb;173:104478. [doi: 10.1016/j.brat.2024.104478] [Medline: 38244384]
74. Ho SS, Chuah ASF, Ho VS, Rosenthal S, Kim HK, Soh SSH. Crisis and emergency risk communication and emotional appeals in COVID-19 public health messaging: quantitative content analysis. J Med Internet Res 2024 Sep 17;26:e56854. [doi: 10.2196/56854] [Medline: 39288399]
75. Kwak M, Kim BJ, Chung JB. Serious game development for public health: participatory design approach to COVID-19 quarantine policy education. JMIR Serious Games 2024 Oct 15;12:e54968. [doi: 10.2196/54968] [Medline: 39405084]
76. Hermann E. Leveraging artificial intelligence in marketing for social good-an ethical perspective. J Bus Ethics 2022;179(1):43-61. [doi: 10.1007/s10551-021-04843-y] [Medline: 34054170]

XSL•FO
RenderX

77. OSF home. Effects of perceived scarcity on mental health, cognitive functioning, time and risk preferences and decision-making during and after covid-19 lockdown: a quasi-natural experimental study. URL: https://osf.io/fs7x2/ [accessed 2025-08-22]

Abbreviations

OLS: ordinary least squares
PSM: propensity score matching
PSS: Perceived Stress Scale
VIF: variance inflation factor

Edited by A Mavragani, T Sanchez; submitted 01.12.24; peer-reviewed by J Wang, W An; revised version received 17.06.25; accepted 23.07.25; published 29.08.25.

Please cite as:
Zhu H, You F, Gruber T, Dong H, de Bont C
Effects of Perceived Scarcity on Mental Health, Time and Risk Preferences, and Decision-Making During and After COVID-19 Lockdown: Quasi-Natural Experimental Study
JMIR Public Health Surveill 2025;11:e69496
URL: https://publichealth.jmir.org/2025/1/e69496
doi:10.2196/69496

© Haiou Zhu, Fangzhou You, Thorsten Gruber, Hua Dong, Cees de Bont. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Consistency of Daily Number of Reported COVID-19 Cases in 191 Countries From 2020 to 2022: Comparative Analysis of 2 Major Data Sources

Han Liu[1], MPH; Huiying Zong[2], MPH; Yang Yang[3], PhD; David C Schwebel[4], PhD; Bin Xie[1], BM; Peishan Ning[1], PhD; Zhenzhen Rao[1], PhD; Li Li[1], PhD; Guoqing Hu[1], PhD

[1]Department of Epidemiology and Health Statistics, Hunan Provincial Key Laboratory of Clinical Epidemiology, Xiangya School of Public Health, Central South University, 172 Tongzipo Road, Yuelu District, Changsha, Hunan, China

[2]Department of Medical Records Management, Xiangya Hospital, Changsha, China

[3]Department of Statistics, Franklin College of Arts and Sciences, University of Georgia, Athens, GA, United States

[4]Department of Psychology, University of Alabama at Birmingham, Birmingham, AL, United States

Corresponding Author:
Guoqing Hu, PhD
Department of Epidemiology and Health Statistics, Hunan Provincial Key Laboratory of Clinical Epidemiology, Xiangya School of Public Health, Central South University, 172 Tongzipo Road, Yuelu District, Changsha, Hunan, China

Abstract

Background: The COVID-19 pandemic represents one of the most challenging public health emergencies in recent world history, causing about 7.07 million deaths globally by September 24, 2024. Accurate, timely, and consistent data are critical for early response to situations like the COVID-19 pandemic.

Objective: This study aimed to evaluate consistency of daily reported COVID-19 cases in 191 countries from the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE) and the World Health Organization (WHO) dashboards during 2020 - 2022.

Methods: We retrieved data concerning new daily COVID-19 cases in 191 countries covered by both data sources from January 22, 2020, to December 31, 2022. The ratios of numbers of daily reported cases from the 2 sources were calculated to measure data consistency. We performed simple linear regression to examine significant changes in the ratio of numbers of daily reported cases during the study period.

Results: Of 191 WHO member countries, only 60 displayed excellent data consistency in the number of daily reported COVID-19 cases between the WHO and JHU CSSE dashboards (mean ratio 0.9-1.1). Data consistency changed greatly across the 191 countries from 2020 to 2022 and differed across 4 types of countries, categorized by income. Data inconsistency between the 2 data sources generally decreased slightly over time, both for the 191 countries combined and within the 4 types of income-defined countries. The absolute relative difference between the 2 data sources increased in 84 countries, particularly for Malta (R^2=0.25), Montenegro (R^2=0.30), and the United States (R^2=0.29), but it decreased significantly in 40 countries.

Conclusions: The inconsistency between the 2 data sources warrants further research. Construction of public health surveillance and data collection systems for public health emergencies like the COVID-19 pandemic should be strengthened in the future.

(JMIR Public Health Surveill 2025;11:e65439) doi:10.2196/65439

KEYWORDS

COVID-19; pandemic; data consistency; World Health Organization; data quality

Introduction

The COVID-19 pandemic represents one of the greatest global public health challenges of this century. According to World Health Organization (WHO) statistics, by September 24, 2024, COVID-19 led to over 7.04 million deaths globally [1]. Critical flaws in timely data collection and reporting have hindered and misled the prevention and control of the pandemic over the past 3 years [2]. For example, during the early phase of the pandemic, the cumulative number of COVID-19 cases in the United States released by the Johns Hopkins University dashboard quickly exceeded 1 million individuals by the end of April 2020, substantially higher than governmental statistics [3]. Such discrepancies created considerable chaos for US governmental responses to the COVID-19 pandemic [4]. To better respond to future pandemics and other global health crises, it is helpful to examine and summarize important lessons we learned.

XSL•FO
RenderX

One key aspect of public health crisis and pandemic preparedness is the need for high-quality data to guide decision-making, practice, and research about prevention and control. Furthermore, 2 open-access data sources were widely used in research and decision-making about the COVID-19 pandemic [5], the COVID-19 dashboards released by the WHO and the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). The 2 data sources collected data in different ways and have been reported as having somewhat inconsistent data in several countries [6]. However, a comprehensive evaluation of the consistency between the 2 data sources has not been conducted across multiple countries and over the key time period of the pandemic, from 2020 to 2022.

Therefore, this study systematically examined the consistency of the number for daily reported COVID-19 cases in 191 countries from 2 data sources, WHO and JHU CSSE, from 2020 to 2022.

Methods

Database

Given their wide use and numerous citations, we selected the publicly accessible WHO COVID-19 dashboard and the JHU CSSE COVID-19 dashboard as the major sources to investigate data patterns concerning COVID-19 pandemic infection rates [6]. The WHO dashboard was created by the WHO. In January 2020, the WHO developed the case definition of COVID-19 infection and requested all member countries to report the number of daily reported cases through the International Health Regulations national focal points. It also collected COVID-19 daily cases and death counts from internet-based public sources (public dashboards and social media) as a supplement [7]. The JHU CSSE dashboard was created by Johns Hopkins University in early January 2020. It gathered data from publicly accessible sources referenced in a version-controlled README file on the repository [5].

For this study, we retrieved publicly accessible numbers of daily reported COVID-19 cases in the 191 United Nations member countries from both data sources between January 22, 2020 and December 31, 2022. In total, there were 6 days with missing values in the 2 data sources for all 191 countries and all study days together, comprising 0.00003% (6/205,325) of all study days. Considering the small number of missing values and the fact that they were unlikely to create significant impact on our results, we excluded those days from our analysis. Data before January 22, 2020, were also not included, as the outbreak had just initiated and the JHU CSSE COVID-19 dashboard was not yet established.

Data Analysis

In total, 2 measures were used to assess consistency between the 2 data sources. One measure was the ratio of the daily number of reported cases from the JHU CSSE dashboard to that from the WHO dashboard, aligned by calendar time. As an example, in Afghanistan on March 31, 2022, the JHU CSSE dashboard reported 264 newly reported cases and the WHO dashboard reported 220. The ratio of JHU CSSE over WHO dashboards was calculated as “264/220=1.2” for that day in Afghanistan. To facilitate our description, we refer to the ratio of daily numbers of reported cases between the 2 data sources as “daily ratio” or, simply, “ratio.”

Based on the mean ratio of the 191 countries during the full study time period, we classified the daily ratios for each country into 5 categories of consistency: excellent (0.9≤mean ratio≤1.1), high (0.8≤mean ratio<0.9 or 1.1<mean ratio≤1.2), moderate (0.7≤mean ratio<0.8 or 1.2<mean ratio≤1.3), low (0.4≤mean ratio<0.7 or 1.3<mean ratio≤1.6), and poor (mean ratio<0.4 or 1.6<mean ratio). Countries where the 2 dashboards had higher consistency yielded ratios closer to 1.0, of course, and countries where there were greater discrepancies between the 2 dashboards had ratios either far below or far above 1.0.

To address certain unusual cases, we empirically built the following rules to calculate the daily ratio in certain unusual cases: (1) when the numerator and denominator were both 0, we defined the ratio as 1, as this represented perfect agreement between the dashboards; (2) when only the numerator was 0 and the denominator was 10 or greater, we defined the ratio as 0.5, and when the denominator was less than 10, we used the fourth criterion listed below to define the value of the ratio; (3) when the denominator was 0 and numerator was 10 or greater, we defined the ratio as 2.0, and when the numerator was less than 10, we used the fourth criterion below to define the value of the ratio; and (4) when the numerator and denominator were both less than 10 and their differences were less than 3 (explaining 0.1% of all data), between 4 and 6 (explaining 0.3%), or ≥7 (explaining 0.5%), we defined the ratio as 0.9 or 1.1, 0.7 or 1.3, 0.5 or 1.5, respectively, thus avoiding extremely large or small ratios due to small numerators or denominators. We mapped the mean and coefficient of variation (CV) of the daily ratios between the 2 data sources for each of the 191 countries.

We graphed box plots and performed the Kruskal-Wallis rank sum test to display and compare differences in the mean and CV of the daily ratios across groups of countries defined by income levels. The 191 countries were classified into 4 income-based categories according to the World Bank Analytical Classifications of 2023 [8]: low-income countries, lower middle-income countries, upper middle-income countries, and higher-income countries. In addition, simple linear regression, with the absolute relative difference in daily number used as the dependent variable and time (date) as the independent variable, was performed to examine trends in the daily differences during the study period. A significant and positive regression coefficient indicates worsening data consistency over time, while a significant and negative regression coefficient indicates improved data consistency.

SPSS statistics 26 (IBM Corp), R version 4.3.0 (R Core Team), and Microsoft Excel spreadsheet 2021 were used for data analysis. $P<.05$ was regarded as statistically significant.

Ethical Considerations

This study does not cause harm to humans and does not involve individual information or commercial interest information. The data used in this study cannot be used to identify individual information, either directly or indirectly. Based on the given

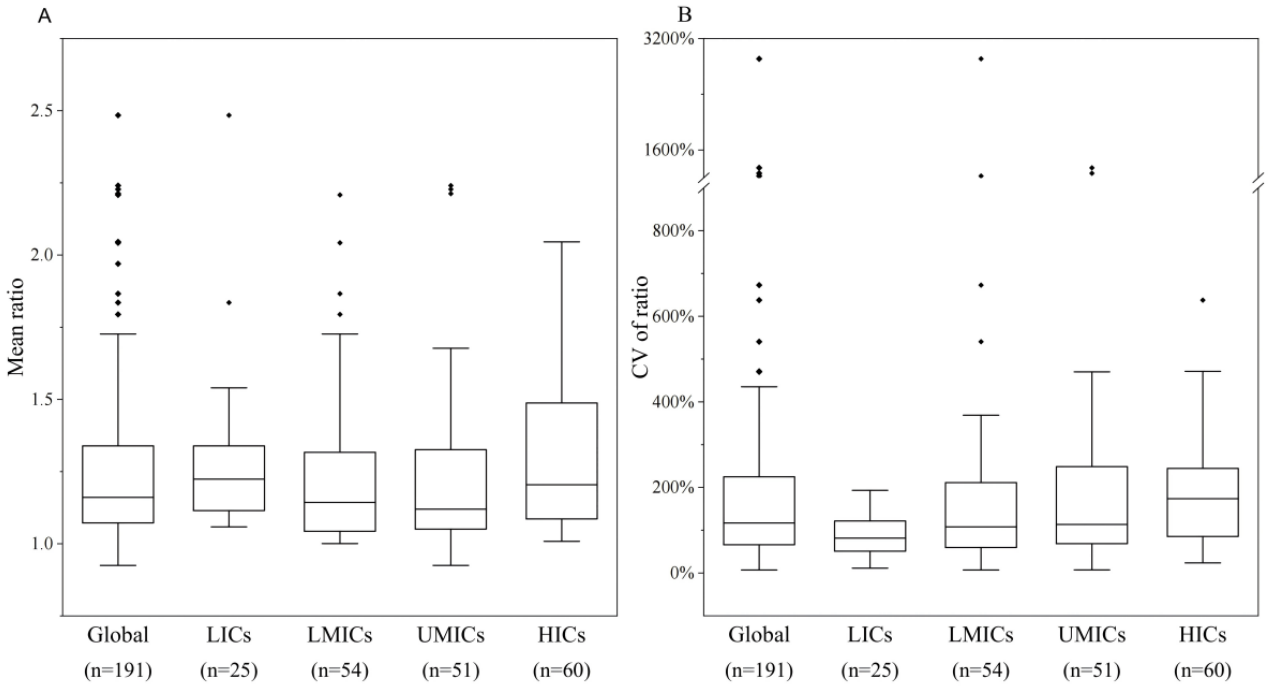
XSL•FO
RenderX

declarations, this study complies with the fundamental principles expressed in the Declaration of Helsinki and meets the ethical exemption requirements of the Ethical Review Measures for Life Sciences and Medical Research Involving Humans promulgated by China. Therefore, this research was exempted from ethical review.

Results

For all observed dates combined, the median ratio between the daily number of reported COVID-19 cases in the 2 data sources for the 191 countries was 1.14 (IQR 1.07-1.34), with the lowest mean ratio in Turkey (0.92) and the highest mean ratio in Sudan (2.48) (Figure 1A). The ratios did not differ significantly across the 4 categories of countries organized by income (*P*=.12). The median CV of ratio between the daily number of reported COVID-19 cases from the 2 data sources was 1.16 (IQR 0.66-2.25) for the 191 countries. The 4 types of countries had significantly different CVs of ratio (*H*=10.08, *P*=.02) (Figure 1B).

Figure 1. Ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard by country income category, from January 22, 2020, to December 31, 2022. CV: coefficient of variation; HIC: higher-income country; LIC: low-income country; LMIC: lower middle-income country; UMIC: upper middle-income country.



In Figure 1, the ratio was calculated as the number of daily new cases from the JHU CSSE COVID-19 dashboard divided by the number from the WHO COVID-19 dashboard, for each country. Figure 2 and Figure 3 display the variations in mean and CV of ratio between the daily number of reported COVID-19 cases from the 2 data sources across the countries. Of the 191 countries, 60 had excellent consistency (mean ratio: 0.9 - 1.1), while 37 had low consistency (mean ratio: 0.4 - 0.7 or 1.3 - 1.6) and 23 had poor consistency (mean ratio:<0.4 or >1.6) (Figure 2). Most African and European countries had low or poor data consistency and most Asian countries had high or excellent data consistency. Strikingly, the CV of ratio between the number of daily reported COVID-19 cases from the 2 data sources during the study time period was higher than 100% in 110 countries (110/191, 57%), and 20% or less in only 8 countries (8/191, 4%) (Figure 3).

XSL•FO
RenderX

Figure 2. Mean ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard across 191 countries from January 22, 2020, to December 31, 2022. The ratio was calculated as the number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard divided by the number from the World Health Organization COVID-19 dashboard, for each country. NA: not available.

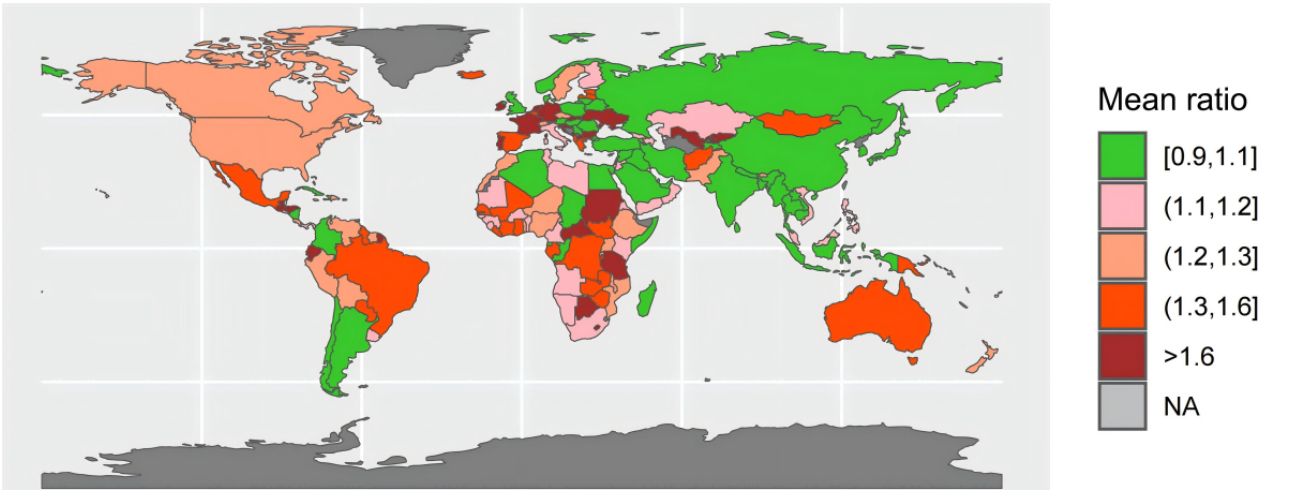
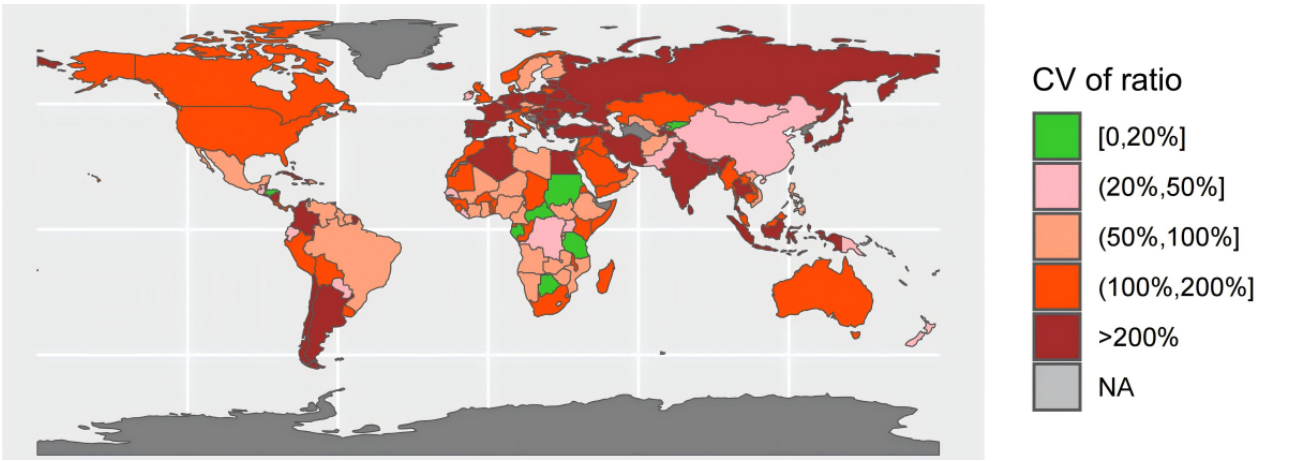


Figure 3. CV of ratio of number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard across 191 countries from January 22, 2020, to December 31, 2022. The ratio was calculated as the number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard divided by the number from the World Health Organization COVID-19 dashboard, for each country. CV: coefficient of variation; NA: not available.



For the 191 countries combined and within all 4 income-defined types of countries, the absolute relative difference between daily number of reported COVID-19 cases from the 2 data sources all increased slightly over the study time period, with R^2 values of 0.16, 0.08, 0.06, 0.14, and 0.14, respectively (Table 1). Country-specific analyses showed that the absolute relative difference increased in 84 countries but decreased in 40 countries significantly (Table 2). Notably, data inconsistency between the 2 data sources increased significantly over time in 3 countries (Malta: R^2=0.25; Montenegro: R^2=0.30; United States: R^2=0.29).

XSL•FO
RenderX

Table . Linear trends in ratio of number of daily reported cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard from January 22, 2020, to December 31, 2022, by country income category (dependent and independent variables of linear regression model were “the absolute value of (ratio-1)” and “time (date),” respectively).

Country income	Linear regression model	Coefficient of determination (R^2)	*P* value
Global	$\hat{y}=-713.80+0.00029x$	0.16	<.001
LICs[a]	$\hat{y}=-1110.18+0.00045x$	0.08	<.001
LMICs[b]	$\hat{y}=-1025.15+0.00041x$	0.06	<.001
UMICs[c]	$\hat{y}=-781.22+0.00031x$	0.14	<.001
HICs[d]	$\hat{y}=-1453.78+0.00059x$	0.14	<.001

[a]LIC: low-income country.

[b]LMIC: lower middle-income country.

[c]UMIC: upper middle-income country.

[d]HIC: higher-income country.

XSL•FO
RenderX

Table . Linear trends in ratio of country-specific number of daily new cases from the Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard and from the World Health Organization COVID-19 dashboard from January 22, 2020, to December 31, 2022. Based on coefficient of determination (R^2) of linear regression models with statistical significance, we classified linear trends of ratio into 3 grades: substantially ($R^2 \geq 0.25$), moderately ($0.10 < R^2 < 0.25$), and slightly ($R^2 \leq 0.10$). In addition, we did not detect significant changes in data consistency for 67 countries).

Linear trend		Country
Increased (n=84)		
	Substantially (n=3)	Malta (R^2=0.25), Montenegro (R^2=0.30), United States (R^2=0.29)
	Moderately (n=4)	Australia (R^2=0.18), Bulgaria (R^2=0.11), Iran (R^2=0.13), Slovenia (R^2=0.21)
	Slightly (n=77)	Albania, Algeria, Antigua and Barbuda, Azerbaijan, Bahamas, Bangladesh, Barbados, Belgium, Bhutan, Bolivia, Bosnia and Herzegovina, Brazil, Brunei, Burundi, Cambodia, Canada, Chile, Comoros, Croatia, Cyprus, Czechia, Democratic Republic of Congo, Denmark, Dominica, Dominican Republic, Ecuador, Ethiopia, Fiji, Finland, Georgia, Grenada, Haiti, India, Iraq, Ireland, Italy, Kazakhstan, Kenya, Kiribati, Kuwait, Laos, Latvia, Liechtenstein, Lithuania, Maldives, Marshall Islands, Mauritius, Mexico, Micronesia, Monaco, Mongolia, Morocco, Namibia, Nauru, Norway, Palau, Panama, Papua New Guinea, Peru, Romania, Saint Kitts and Nevis, Samoa, San Marino, Saudi Arabia, Seychelles, Slovakia, Solomon Islands, South Africa, Sweden, Switzerland, Thailand, Timor, Tonga, Ukraine, United Kingdom, Vanuatu, Venezuela
Decreased (n=40)		
	Moderately (n=2)	Poland (R^2=0.17), Spain (R^2=0.11)
	Slightly (n=38)	Andorra, Argentina, Bahrain, Belarus, Burkina Faso, Cameroon, Chad, Congo, Djibouti, Egypt, El Salvador, Estonia, Germany, Greece, Guinea, Hungary, Japan, Jordan, Madagascar, Malaysia, Mali, Mauritania, Moldova, Myanmar, Netherlands, North Macedonia, Portugal, Russia, Sierra Leone, Somalia, Togo, Tunisia, Turkey, Tuvalu, United Arab Emirates, Uruguay, Yemen, Zambia
No significant change (n=67)		Afghanistan, Angola, Armenia, Austria, Belize, Benin, Botswana, Cape Verde, Central African Republic, China, Colombia, Costa Rica, Cote d'Ivoire, Cuba, Equatorial Guinea, Eritrea, Eswatini, France, Gabon, Gambia, Ghana, Guatemala, Guinea-Bissau, Guyana, Honduras, Iceland, Indonesia, Israel, Jamaica, Kyrgyzstan, Lebanon, Lesotho, Liberia, Libya, Luxembourg, Malawi, Mozambique, Nepal, New Zealand, Nicaragua, Niger, Nigeria, Oman, Pakistan, Paraguay, Philippines, Qatar, Republic of Korea, Rwanda, Saint Lucia, Saint Vincent and the Grenadines, Sao Tome and Principe, Senegal, Serbia, Singapore, South Sudan, Sudan, Suriname, Sri Lanka, Syria, Tajikistan, Tanzania, Trinidad and Tobago, Uganda, Uzbekistan, Vietnam, Zimbabwe

XSL•FO
RenderX

Discussion

Main Findings

This study systematically examined the consistency of number of daily reported COVID-19 cases and number-based epidemic cycles in 191 countries from the 2 most cited data sources between 2020 and 2022. We generated 2 key findings. First, there was a difference in global number of daily reported COVID-19 cases between the WHO and JHU CSSE data, and data consistency differed between the 4 categories of income-defined countries and across the 191 countries. Second, data inconsistency between the 2 data sources generally decreased slightly over time, both for the 191 countries combined and within the 4 types of income-defined countries, but it changed differently over time across the 191 individual countries.

Interpretation of Findings

Inconsistencies in the number of daily reported COVID-19 cases between the WHO and JHU CSSE data likely reflect differences in data collection and data release strategies across the 2 data sources. The WHO data came primarily from official daily counts reported by WHO member states, territories, and areas [7], while the JHU CSSE data were derived from a combination of more than 400 sources for over 3500-point locations [5]. Compared with the WHO data, the JHU CSSE data were collected from a higher number of sources, potentially explaining why the number of daily reported COVID-19 cases from the JHU CSSE data was greater than that from the WHO data for the whole world, the 4 types of income-defined countries, and most individual countries during most dates [9]. The 2 data sources also adopted different data release strategies. The JHU CSSE COVID-19 dashboard used automated methods to extract data from each source every half hour and therefore updated data hourly [5]. In contrast, the WHO COVID-19 dashboard relied on official reports from all WHO member countries, territories, and areas, which were more prone to be delayed and less frequently updated than the JHU dashboard data [10].

Variations in data inconsistency between the WHO and JHU CSSE data might also be due to diverse data publishing policies and the evolution of data-releasing agencies (eg, delayed reporting, data entry errors, and data collection strategies), particularly at the local level across the 191 countries. As the pandemic eased, some countries like Sweden reduced the frequency of data release and even stopped publishing epidemic data, particularly at the local level [11]. These practices likely affected the original data collected by the WHO and the JHU CSSE COVID-19 dashboards in different ways. In particular, sources indexed by the JHU CSSE dashboard were more extensive and were more likely to change over time than those covered by the WHO dashboard [12].

Implications and Limitations

Our findings have 2 important implications. First, we do not have clear empirical basis to explain the drivers behind inconsistencies in the number of COVID-19 cases between the WHO and JHU CSSE dashboards or the changes in data inconsistency over time. A variety of internal and external factors likely could have contributed to the magnitude and pattern of the inconsistency, including but not limited to changes in government intervention policies or testing capacities.

Policy makers, researchers, and public health practitioners should recognize that the 2 data sources generate somewhat different and even occasionally conflicting results when they use the historical COVID-19 data collected by the WHO and JHU CSSE dashboards to summarize pandemic control experiences and lessons, conduct historical trend analyses, and evaluate intervention strategies [13-15]. Because strong evidence and detailed guidance are lacking concerning which of the 2 data sources is more valid for which specific time periods and countries, we recommend users consider both, recognizing each may have flaws. It also would be valuable to conduct further research assessing the quality of the data sources. For example, comparisons of the data sources with known cycles of the COVID-19 epidemic may yield additional information about the validity of each data source [16,17]. Another useful approach to bridge the gaps in high-level data from mainstream sources is to cross-validate them with data provided by local health authorities whenever and wherever available. This may be particularly useful for data on hospitalization and on fatalities, which are less sensitive to testing policies, health care–seeking behavior, and reporting bias [18].

Second, our results highlight the importance of gathering consistent and accurate data to fight against future pandemics or other public health emergencies [19]. Inconsistent data can lead to undesired consequences in decision-making, research, and prevention efforts [20,21]. Development of standardized data collection and release protocols, along with infrastructure- and competency-building to help all countries gather accurate health data, will prepare us not just for current health situations but also for future pandemics and broad public health emergencies [22].

Our research is primarily limited by the unavailability of a gold standard criterion. We assessed the consistency of 2 data sources and cannot comprehensively assess the reliability or validity of either, or identify which might be more accurate. No alternative or recognized valid measure of COVID-19 infections or fatalities exists worldwide. In addition, due to the unavailability of detailed information about COVID-19 data collection methods or data release policies across the 191 countries, we cannot reasonably interpret exactly why data consistency variations emerged across the 191 countries or why there were significant increases or decreases in data consistency over time for some countries. Future research may consider innovative strategies to identify key factors contributing to the observed data inconsistencies and significant changes in data consistency. Hospitalization data might be considered as a surrogate gold standard, for example, to approximate accuracy of other data sources [23], perhaps providing valuable clues. Advanced data analysis strategies, such as data harmonization or deep machine learning, might also help uncover reasons behind the results we report and be useful to address data accuracy in future public health emergencies. Finally, evidence-based data collection and release guidelines should be established by the WHO or other governing bodies, perhaps supplemented with financial and

XSL•FO
RenderX

technical support to those countries with limited resources to follow international guidelines. With strong international collaboration, we should be able to improve transparency, consistency and overall quality of data emerging from public health crises in the future.

Conclusions

We report a moderate difference in global average daily number of reported COVID-19 cases between 2020 and 2022 in the WHO and JHU CSSE datasets. The COVID-19 data inconsistency changed significantly across the countries studied and across time, both globally and in 124 individual countries. Data users should be cautious in interpreting the data, given inconsistency across the 2 most commonly cited COVID-19 data sources. Further research is recommended to develop guidelines to support valid interpretation of existing COVID-19 data, explore possible reasons for discrepancies, and understand factors such as governmental policies and data collection systems that may lead to data inconsistency. With those efforts, society can be best prepared to conduct valid health-related data collection in response to future pandemics and public health emergencies.

Acknowledgments

The study was supported by Philosophy and Social Science Foundation of China (number 20&ZD120).

Data Availability

Both databases are freely accessible for the public. Throughout the data utilization process, we strictly adhered to the data usage requirements of the source databases, ensuring no modifications were made to the metadata. The attribution of the data sources has been explicitly stated in the manuscript.

Conflicts of Interest

None declared.

References

1. WHO COVID-19 dashboard. World Health Organization. URL: https://data.who.int/dashboards/covid19/data?n=c [accessed 2023-12-31]
2. Nageshwaran G, Harris RC, Guerche-Seblain CE. Review of the role of big data and digital technologies in controlling COVID-19 in Asia: public health interest vs. privacy. Dig Health 2021;7:20552076211002953. [doi: 10.1177/20552076211002953] [Medline: 33815815]
3. Pollard CA, Morran MP, Nestor-Kalinoski AL. The COVID-19 pandemic: a global health crisis. Physiol Genomics 2020 Nov 1;52(11):549-557. [doi: 10.1152/physiolgenomics.00089.2020] [Medline: 32991251]
4. Lessons learned from COVID-19 pandemic: virtual regional consultation with informal expert group, 19 October 2021. World Health Organization. URL: https://www.who.int/publications/i/item/sea-whe-6 [accessed 2025-01-02]
5. Dong E, Ratcliff J, Goyea TD, et al. The Johns Hopkins University Center for Systems Science and Engineering COVID-19 dashboard: data collection process, challenges faced, and lessons learned. Lancet Infect Dis 2022 Dec;22(12):e370-e376. [doi: 10.1016/S1473-3099(22)00434-0] [Medline: 36057267]
6. Alamo T, Reina D, Mammarella M, Abella A. Covid-19: Open-data resources for monitoring, modeling, and forecasting the epidemic. Electronics (Basel) 2020;9(5):827. [doi: 10.3390/electronics9050827]
7. Allan M, Lièvre M, Laurenson-Schafer H, et al. The World Health Organization COVID-19 surveillance database. Int J Equity Health 2022 Nov 23;21(Suppl 3):167. [doi: 10.1186/s12939-022-01767-5] [Medline: 36419127]
8. World Bank country and lending groups. The World Bank. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519 [accessed 2025-01-28]
9. Wu Z, Huang NE, Long SR, Peng CK. On the trend, detrending, and variability of nonlinear and nonstationary time series. Proc Natl Acad Sci U S A 2007 Sep 18;104(38):14889-14894. [doi: 10.1073/pnas.0701020104] [Medline: 17846430]
10. Zuo X, Chen Y, Ohno-Machado L, Xu H. How do we share data in COVID-19 research? A systematic review of COVID-19 datasets in PubMed Central articles. Brief Bioinform 2021 Mar 22;22(2):800-811. [doi: 10.1093/bib/bbaa331] [Medline: 33757278]
11. Hu T, Guan WW, Zhu X, et al. Building an open resources repository for COVID-19 research. Data Inf Manag 2020 Sep 1;4(3):130-147. [doi: 10.2478/dim-2020-0012] [Medline: 35382104]
12. Lindström M. The COVID-19 pandemic and the timing of government response: a comparison of four Nordic countries March–June 2020. Scand J Public Health 2023 Jul;51(5):754-758. [doi: 10.1177/14034948231171201] [Medline: 37118947]
13. Badker R, Miller K, Pardee C, et al. Challenges in reported COVID-19 data: best practices and recommendations for future epidemics. BMJ Glob Health 2021 May;6(5):e005542. [doi: 10.1136/bmjgh-2021-005542] [Medline: 33958393]
14. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis 2020 Nov;20(11):1247-1254. [doi: 10.1016/S1473-3099(20)30553-3]

XSL•FO
RenderX

15. Guerrier C, McDonnell C, Magoc T, Fishe JN, Harle CA. Understanding health care administrators' data and information needs for decision making during the COVID-19 pandemic: a qualitative study at an academic health system. MDM Policy Pract 2022;7(1):23814683221089844. [doi: 10.1177/23814683221089844] [Medline: 35368410]
16. Dowell SF, Ho MS. Seasonality of infectious diseases and severe acute respiratory syndrome–what we don't know can hurt us. Lancet Infect Dis 2004 Nov;4(11):704-708. [doi: 10.1016/S1473-3099(04)01177-6]
17. Chemaitelly H, Tang P, Hasan MR, et al. Waning of BNT162b2 vaccine protection against SARS-CoV-2 infection in Qatar. N Engl J Med 2021 Dec 9;385(24):e83. [doi: 10.1056/NEJMoa2114114] [Medline: 34614327]
18. Huang AA, Huang SY. Hospitalized COVID-19 patients with diabetes have an increased risk for pneumonia, intensive care unit requirement, intubation, and death: a cross-sectional cohort study in Mexico in 2020. Health Sci Rep 2023 Apr;6(4):e1222. [doi: 10.1002/hsr2.1222] [Medline: 37081996]
19. Johnson B. Five discoveries about COVID-19 made since the public health emergency ended. Nat Med 2023 Dec;29(12):2974-2976. [doi: 10.1038/d41591-023-00084-w]
20. Kung S, Doppen M, Black M, et al. Underestimation of COVID-19 mortality during the pandemic. ERJ Open Res 2021 Jan;7(1):00766-02020. [doi: 10.1183/23120541.00766-2020]
21. TC J, Barbieri M, et al. Differences in COVID-19 mortality: Implications of imperfect and diverse data collection systems. Popul Health Metr 2021;76:35-72. [doi: 10.3917/popu.2101.0035]
22. Suthar AB, Schubert S, Garon J, Couture A, Brown AM, Charania S. COVID-19 case definitions, diagnostic testing criteria, and surveillance across the pandemic's 25 highest burden countries. medRxiv. Preprint posted online on May 14, 2021. [doi: 10.1101/2021.05.11.21257047]
23. Huang AA, Huang SY. Diabetes is associated with increased risk of death in COVID-19 hospitalizations in Mexico 2020: a retrospective cohort study. Health Sci Rep 2023 Jul;6(7):e1416. [doi: 10.1002/hsr2.1416] [Medline: 37415678]

Abbreviations

CV: coefficient of variation
JHU CSSE: Johns Hopkins University Center for Systems Science and Engineering
WHO: World Health Organization

Edited by A Mavragani; submitted 15.08.24; peer-reviewed by R Ibrahim, SY Huang; revised version received 13.11.24; accepted 24.11.24; published 06.02.25.

Please cite as:
Liu H, Zong H, Yang Y, Schwebel DC, Xie B, Ning P, Rao Z, Li L, Hu G
Consistency of Daily Number of Reported COVID-19 Cases in 191 Countries From 2020 to 2022: Comparative Analysis of 2 Major Data Sources
JMIR Public Health Surveill 2025;11:e65439
URL: https://publichealth.jmir.org/2025/1/e65439
doi:10.2196/65439

© Han Liu, Huiying Zong, Yang Yang, David C Schwebel, Bin Xie, Peishan Ning, Zhenzhen Rao, Li Li, Guoqing Hu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 6.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association Between Triglyceride-Glucose Index, Blood Pressure Status, and Coronary Heart Disease Risk Among Chinese Adults With Disabilities: 10-Year Disability Health Survey Cohort Study

Hui Liu[1*], MSc; Yao Li[1,2*], PhD; Yiyan Wang[1], MSc; Tao Wang[1], MSc; Chenghua Jiang[1], PhD; Hengjing Wu[1], PhD; Jing Wu[1], PhD

[1]Shanghai YangZhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Medicine, Tongji University, Guangxing Road 2209, Shanghai, China

[2]School of Nursing, Wannan Medical College, Wuhu, China

[*]these authors contributed equally

Corresponding Author:
Jing Wu, PhD
Shanghai YangZhi Rehabilitation Hospital (Shanghai Sunshine Rehabilitation Center), School of Medicine, Tongji University, Guangxing Road 2209, Shanghai, China

Abstract

Background: The triglyceride-glucose (TyG) index and blood pressure (BP) status are key indicators associated with coronary heart disease (CHD). However, limited research has focused on individuals with disabilities.

Objective: This study explores the potential combined effects of the TyG index and BP status on CHD risk in groups with varying disability characteristics.

Methods: This study analyzed data from the Shanghai Disability Health Survey, conducted between January 2012 and December 2022. Participants were then categorized into 3 BP status groups: nonelevated BP, elevated BP, and hypertension. Cox proportional hazards regression models were used to assess the associations between BP status, the TyG index, and CHD incidence. Additionally, the mediating, interaction, and combined effects of these factors on CHD risk were examined. A stratified analysis was performed based on participants' disability characteristics, including disability type and severity, to explore potential variations in the associations.

Results: Among the 21,628 participants, the mean age was 53.30 (SD 10.57) years, and 50.89% (11007/21628) were male. In a follow-up of 77.45 months, CHD events occurred in 2312 participants (10.69%). The TyG index and BP status were independently associated with an increased risk of CHD. Mediation analysis showed that TyG explained 20.5% (95% CI 13.6% - 22.0%) of the BP and CHD association. Significant multiplicative interactions were identified (hazard ratio [HR] 1.41, 95% CI 1.02 - 1.94), and joint analysis indicated the highest CHD risk in those with both hypertension and elevated TyG (HR 1.92, 95% CI 1.52 - 2.42). Stratified analyses revealed stronger mediation in participants with physical disabilities (22.6%, 95% CI 9.0% - 60%) or visual disabilities (16.6%, 95% CI 4.8% - 51%), while this was not significant in those with hearing or speech (*P*=.07) or intellectual or mental disabilities (*P*=.13). By disability grading, the mediated proportion was 22.3% (95% CI 9.2% - 59.4%) in mild or moderate and 18.8% (95% CI 15.7% - 29%) in severe or very severe groups. Joint associations showed consistently higher CHD risk across most disability classifications, with particularly elevated risk in people with intellectual or mental disabilities (HR 3.51, 95% CI 1.89 - 6.50).

Conclusions: BP and the TyG index were significantly associated with CHD risk in individuals with disabilities, with TyG mediating a part of this association and showing stronger effects in physical and mild to moderate disabilities. Significant interactions between BP and TyG further highlight their combined impact, underscoring the need for integrated interventions targeting both factors.

(JMIR Public Health Surveill 2025;11:e78068) doi:10.2196/78068

KEYWORDS

disabilities; coronary heart disease; CHD; risk; triglyceride-glucose index; blood pressure

XSL•FO
RenderX

Introduction

Background

Coronary heart disease (CHD) is a leading cause of cardiovascular mortality worldwide, with particularly high prevalence in certain high-risk populations [1,2]. While there is extensive research on CHD in the general population [3,4], studies on CHD in individuals with disabilities remain scarce, despite the fact that this group constitutes a significant portion of the global population. According to the latest report by the World Health Organization, over 1 billion people (approximately 15% of the global population) live with some form of disability [5]. Individuals with disabilities are at an elevated risk for cardiovascular diseases (CVDs) due to functional limitations, metabolic disorders, and restricted physical activity [6,7]. Hypertension is a recognized risk factor for CHD, closely linked to atherosclerosis, left ventricular hypertrophy, and other detrimental cardiovascular outcomes. Among people with disabilities, the prevalence of hypertension is often higher, further exacerbating the cardiovascular burden in this population [8].

Recently, the triglyceride-glucose (TyG) index has gained widespread application as a biomarker of insulin resistance in assessing CVD risk and adverse outcomes [9]. Derived from the product of fasting glucose and triglycerides, the TyG index provides an indirect measure of insulin sensitivity. Studies show that the TyG index independently and significantly predicts several CVDs and poor prognoses [10,11], with a higher TyG index being significantly associated with increased risks of stroke [12], heart failure [13], and all-cause mortality [14]. Moreover, interactions between the TyG index and other common cardiovascular risk factors may yield additional impacts on cardiovascular risk. For example, research has shown that the TyG index mediates a substantial portion of the association between BMI and stroke [12], and a co-exposure and mutual mediation effect has been observed between the TyG index and high-sensitivity C-reactive protein for CVD [15]. These findings further support the potential of the TyG index as a tool for predicting CVD risk.

Given that both blood pressure (BP) and the TyG index are closely associated with CHD risk, their potential combined effect on CVD has garnered growing attention. Multiple studies indicate that hypertension and insulin resistance may interact, and their joint evaluation and intervention may have important implications for CVD prevention [16,17]. A more recent cohort study of 57,192 participants using cross-lagged analysis found that elevated TyG exerted a stronger impact on subsequent BP changes than the reverse effect [18]. However, there is still a lack of research focused on systematically evaluating the combined effects of BP and the TyG index on CHD, particularly their impact among people with disabilities.

Objectives

Based on a 10-year longitudinal cohort of people with disabilities, we aim to evaluate the independent and combined effects of different BP levels and the TyG index on incident CHD in people with disabilities, as well as to further analyze the potential mediating role of the TyG index in the association between BP and CHD.

Methods

Data Collection

This study is a retrospective cohort data analysis based on information sourced from the Shanghai Disabled Persons Rehabilitation Comprehensive Information Platform, established by the Shanghai Disabled Persons' Federation. Since 2012, the Shanghai Disabled Persons' Federation has provided annual free health screening services for local individuals with disabilities. As of December 2022, the Shanghai Sunshine Rehabilitation Center (Shanghai Sunshine Rehabilitation Hospital) has facilitated screenings for more than 20,000 persons with disabilities, and the health data obtained have been uploaded to the Shanghai Disabled Persons Rehabilitation Comprehensive Information Platform for health monitoring and scientific research purposes. The recruitment process for this study was based on participants attending the annual health screenings, where individuals with various types of disabilities voluntarily participated. The recruitment aimed to include a diverse range of disability types and severities.

Ethical Considerations

This study was approved by the Ethics Committee of Shanghai Sunshine Rehabilitation Hospital (2019 - 051). All participants were informed that their health screening data could be used for scientific research, and informed consent was obtained prior to data collection. To ensure privacy and confidentiality, all data used in this analysis were anonymized and deidentified before being accessed by the research team. Participants did not receive financial compensation for their participation. No individual participants are identifiable in any images or supplementary materials associated with this study.

Study Population

The study population included individuals with disabilities meeting the following criteria: (1) aged over 18 years and (2) participated in at least 2 health screenings during the 10-year follow-up period from 2012 to 2022, with complete electronic health records for each screening. Exclusion criteria were as follows: (1) a prior diagnosis of CHD at baseline, confirmed by a qualified physician in accordance with the World Health Organization diagnostic criteria; (2) presence of other severe conditions, such as malignancies, organ failure, or major trauma; (3) women who were pregnant, breastfeeding, or planning to become pregnant; and (4) individuals with multiple types of disabilities rather than a single disability type. The final eligible sample consisted of 21,628 individuals (Figure S1 in Multimedia Appendix 1).

Assessment of BP and TyG

After participants rested for approximately 10 minutes, trained nursing staff measured BP using a standardized procedure with an automated arm BP monitor (Omron Corp). BP was recorded twice. For participants with upper limb amputation or injuries that could lead to measurement errors, physicians conducted supine BP measurements on both ankles via the posterior tibial

XSL•FO
RenderX

artery and recorded the reading from the limb with the higher value. To ensure consistency with standard arm BP readings, leg BP values were corrected following recommendations in references [18].

BP status was classified based on the 2024 Chinese Hypertension Prevention and Treatment Guidelines [19]. Participants were categorized into 3 groups: nonelevated BP, elevated BP, and hypertension. Nonelevated BP was defined as systolic BP (SBP) <120 mm Hg and diastolic BP (DBP) <80 mm Hg, without a history of hypertension. Elevated BP was defined as SBP ranging from 120 mm Hg to 139 mm Hg or DBP from 80 mm Hg to 89 mm Hg. Hypertension was defined as SBP≥140 mm Hg, DBP≥90 mm Hg, or a self-reported history of hypertension.

The TyG index, used as a marker of metabolic health, was calculated from venous blood samples. All blood samples were immediately frozen after collection and transported to a specialized laboratory, where they were stored at –80°C to maintain sample quality. Serum triglyceride (TG) and fasting plasma glucose (FPG) levels were measured using an enzymatic colorimetric method, and the TyG index was calculated according to the formula: ln [TG (mg/dL)×FPG (mg/dL)/2] [20,21]. Both the TyG index and BP were measured at baseline for the analysis.

Outcome Ascertainment

The primary outcome of this study was the incidence of CHD occurring from the time of health screening in 2012 until December 2022. CHD was defined as fatal ischemic heart disease (according to ICD-10 codes I20-I25) or nonfatal myocardial infarction (ICD-10 codes I21-I23). The CHD history was assessed through interviews with qualified physicians and verified by reviewing patient medical records. The timing of CHD onset was recorded based on case documentation to ensure the reliability and temporal accuracy of outcome event data.

Covariates

Participants' medical and functional status was comprehensively evaluated by trained health care professionals at baseline. The health questionnaire collected data on general demographics (eg, gender, age, and marital status); medical history (eg, hypertension, diabetes, and fatty liver); and disability status (classification and severity). The classification and grading of disabilities were based on China's National Standard "Classification and Grading Criteria of Disabilities" (GB/T 26341‐2010) [22]. Hospital health examinations included anthropometric measurements (eg, height and weight); biochemical assessments (eg, complete blood count, blood biochemistry, and urinalysis); and electrocardiogram and ultrasound examinations. Blood tests were conducted after an overnight fast of at least 12 hours.

Statistical Analyses

In the statistical analysis, continuous variables that follow a normal distribution were expressed as means and SDs, while nonnormally distributed data were reported as medians and IQR. Categorical variables were presented as frequencies and percentages. For comparisons of baseline characteristics across groups based on BP (divided into 3 categories), chi-square tests, ANOVA, or Kruskal-Wallis rank-sum tests were used according to the data type. Missing data were handled using multiple imputation with chained equations using the *MICE* package of R (version 4.3.0; R Foundation for Statistical Computing) to minimize bias. Detailed information on the missing data can be found in Table S1 in Multimedia Appendix 1.

To estimate the association between different BP categories (nonelevated BP, elevated BP, and hypertension); TyG quartiles; and outcomes, we used Cox proportional hazards regression models to calculate hazard ratios (HRs) and 95% CIs. The models were adjusted in 3 stages: model 1 adjusted for age and gender; model 2 further adjusted for disability-specific factors—disability type and severity; and model 3 additionally controlled for marital status, education level, diabetes history, and serum creatinine level. Furthermore, restricted cubic spline (RCS) models with 3 knots were used to explore nonlinear associations between TyG and the outcome.

To assess the mediating effect of TyG between BP and CHD, we used the *mediation* package in R. First, a logistic regression model (mediator model) was built with the independent and confounding variables, using the mediator variable as the outcome. A second Cox proportional hazards regression model (outcome model) included the independent variable and mediator, with CHD as the outcome. This approach allowed us to estimate both the indirect and direct effects, showing the proportion of the total effect mediated by TyG. The proportion mediated was calculated as the ratio of the natural indirect effect (NIE) to the sum of the natural direct effect and NIE. The bootstrap method was used to estimate the 95% CI of the indirect effect. TyG was dichotomized based on an RCS-determined threshold (8.56). In R, the *mets* package was used to calculate BP's NIE and natural direct effect on CHD incidence through TyG as a mediator.

To further explore the association between TyG and CHD, stratified analyses by BP level were conducted. Interaction terms between BP (nonelevated BP, elevated BP, and hypertension) and TyG quartiles were included, and a likelihood ratio test was performed to compare models with and without interaction terms, assessing whether interaction improved model fit [23]. For joint association analyses, participants were categorized into 12 groups based on their BP and TyG quartiles, with HRs for CHD calculated and compared against a reference group with nonelevated BP and TyG in the first quartile.

Given the unique characteristics of individuals with disabilities, subgroup analyses were performed by disability type and severity. Participants were divided into those with limb disabilities and other disability types, as well as into mild and moderate versus severe and very severe categories. Primary analyses, including mediation analysis, stratified mediation, interaction, and joint association analyses, were repeated within these subgroups.

XSL•FO
RenderX

Results

Baseline Characteristics

A total of 21,628 eligible participants were included in the study, with details on the sample selection process illustrated in Figure S1 in Multimedia Appendix 1. The mean age of the participants was 53.3 (SD 10.57) years, and 10,621 (49.11%) were female. More than half of the participants (11,901/21,628, 55.00%) had physical disabilities, with the majority classified as having either mild or severe disabilities (16,975/21,628, 78.49%). The incidence rate per year is provided in Table S2 in Multimedia Appendix 1.

RCS regression was used to explore the potential nonlinear associations between TyG and CHD events. The results revealed a nonlinear association between TyG and the risk of CHD (P=.002; Figure S2A and S2B in Multimedia Appendix 1). Therefore, nonlinear modeling was adopted to better capture the association between TyG and CHD risk. BP categories were based on the 2024 "Chinese Guidelines for the Prevention and Treatment of Hypertension," while TyG was classified according to the commonly used quartile method in the literature.

Baseline characteristics of the participants, categorized by baseline BP into nonelevated BP (4964/21628, 22.95%), elevated BP (1321/21628, 54.28%), and hypertension (610/21628, 22.77%) groups, are compared in Table 1. The average TyG index at baseline was 8.62 (SD 0.60). Table S3 in Multimedia Appendix 1 provides the baseline characteristics of participants based on TyG quartiles.

XSL•FO
RenderX

Table . Baseline characteristics of 21,628 participants with disabilities by blood pressure (BP) level (N=21,628).

Variables	Total sample	Nonelevated BP (SBP[a]<120, DBP[b]<80; n=4964)	Elevated BP (120≤SBP<140, 80≤DBP<90; n=11,739)	Hypertension (SBP≥140, DBP≥90; n=4925)	*P* value[c]
Age (y), mean (SD)	53.30 (10.57)	49.62 (11.03)	53.75 (10.36)	55.93 (9.54)	<.001
Sex, n (%)					<.001
Male	11,007 (50.89)	2325 (46.84)	6064 (51.66)	2618 (53.16)	
Female	10,621 (49.11)	2639 (53.16)	5675 (48.34)	2307 (46.84)	
BMI (kg/m^2), mean (SD)	24.09 (3.58)	22.40 (3.09)	24.27 (3.45)	25.37 (3.69)	<.001
Marital status, n (%)					<.001
Married	17,852 (82.54)	3874 (78.04)	9788 (83.38)	4190 (85.08)	
Other	3776 (14.46)	1090 (21.96)	1951 (16.62)	735 (14.92)	
Education, n (%)					<.001
Primary school and illiterate	4683 (21.70)	1033 (20.81)	2535 (21.59)	1115 (22.64)	
Junior high school	11,199 (51.80)	2596 (52.30)	6092 (51.90)	2511 (50.98)	
Senior high school	4896 (22.60)	1129 (22.74)	2671 (22.75)	1096 (22.25)	
College and higher	850 (3.93)	206 (4.15)	441 (3.76)	203 (4.12)	
Classification of disabilities, n (%)					<.001
Intellectual and mental disability	3799 (17.60)	1047 (21.09)	1989 (16.94)	763 (15.49)	
Hearing and speech disability	1775 (8.21)	476 (9.59)	908 (7.73)	391 (7.94)	
Visual disability	4153 (19.20)	900 (18.13)	2302 (19.61)	951 (19.31)	
Physical disability	11,901 (55.00)	2541 (51.19)	6540 (55.71)	2820 (57.26)	
Grading of disabilities, n (%)					<.001
Very severe	1924 (8.90)	485 (9.77)	996 (8.48)	443 (8.99)	
Severe	2729 (12.60)	1689 (34.02)	4095 (34.88)	1674 (33.99)	
Moderate	7458 (34.50)	692 (13.94)	1453 (12.38)	584 (11.86)	
Mild	9517 (44.00)	2098 (42.26)	5195 (44.25)	2224 (45.16)	
Comorbidities, n (%)					<.001
Hypertension	5961 (27.60)	281 (5.66)	3125 (26.62)	2555 (51.88)	
Diabetes	1317 (6.09)	174 (3.35)	459 (5.75)	684 (8.08)	
Blood pressure, mean (SD)					<.001
SBP	135.15 (20.87)	109.39 (7.47)	135.78 (12.28)	159.61 (15.13)	
DBP	79.31 (12.46)	65.56 (7.05)	79.70 (7.40)	92.25 (11.94)	
Metabolic biomarkers					
FBG[d] (mmol/L), mean (SD)	5.60 (1.53)	5.20 (1.11)	5.62 (1.52)	5.94 (1.82)	<.001
TC[e] (mmol/L), mean (SD)	4.58 (0.87)	4.37 (0.77)	4.79 (0.90)	4.94 (0.93)	<.001
TG[f] (mmol/L), median (IQR)	1.22 (0.8-1.75)	1.02 (0.76-1.43)	1.24 (0.89-1.77)	1.41 (1.00-2.02)	<.001
TP[g] (g/L), mean (SD)	72.38 (3.61)	71.44 (4.12)	72.40 (4.12)	73.29 (4.25)	<.001

XSL•FO
RenderX

Variables	Total sample	Nonelevated BP (SBP[a]<120, DBP[b]<80; n=4964)	Elevated BP (120≤SBP<140, 80≤DBP<90; n=11,739)	Hypertension (SBP≥140, DBP≥90; n=4925)	*P* value[c]
Alb[h] (g/L), mean (SD)	43.43 (2.33)	43.16 (2.35)	43.46 (2.31)	43.62 (2.38)	<.001
Glo[i] (g/L), mean (SD)	28.97 (3.61)	28.37 (3.65)	28.96 (3.52)	29.60 (3.68)	<.001
Alt[j] (U/L), median (IQR)	20.00 (15.00-29.00)	18.00 (13.00-25.00)	21.00 (15.00-29.00)	22.00 (16.00-32.00)	.115
UA[k] (μmol/L), mean (SD)	318.40 (87.54)	299.65 (81.11)	319.51 (86.62)	334.67 (92.29)	<.001
SCr[l] (μmol/L), median (IQR)	62.10 (50.70-73.32)	61.20 (50.10-72.50)	62.20 (50.70-73.40)	62.80 (51.50-74.10)	<.001
SU[m] (mmol/L), median (IQR)	5.00 (4.20-6.00)	4.90 (4.10-5.80)	5.10 (4.30-6.00)	5.10 (4.30-6.00)	<.001
Hemoglobin (g/L), mean (SD)	137.01 (15.27)	133.80 (15.31)	137.39 (15.08)	139.38 (15.16)	<.001
Platelet count (10^9/L), mean (SD)	200.88 (58.27)	198.63 (57.69)	200.63 (57.81)	203.73 (59.82)	<.001
TyG[n]	8.62±0.60	8.38±0.54	8.64±0.59	8.80±0.62	<.001

[a]SBP: systolic blood pressure.

[b]DBP: diastolic blood pressure.

[c]The *P* value was determined using the chi-square test for categorical data, ANOVA for continuous data across categorical groups, and the Kruskal-Wallis test for continuous data that are ordinal or nonnormally distributed.

[d]FBG: fasting plasma glucose.

[e]TC: total cholesterol.

[f]TG: total triglyceride.

[g]TP: total protein.

[h]Alb: albumin

[i]Glo: globulin

[j]Alt: alanine aminotransferase.

[k]UA: uric acid.

[l]SCr: serum creatinine.

[m]SU: serum urea.

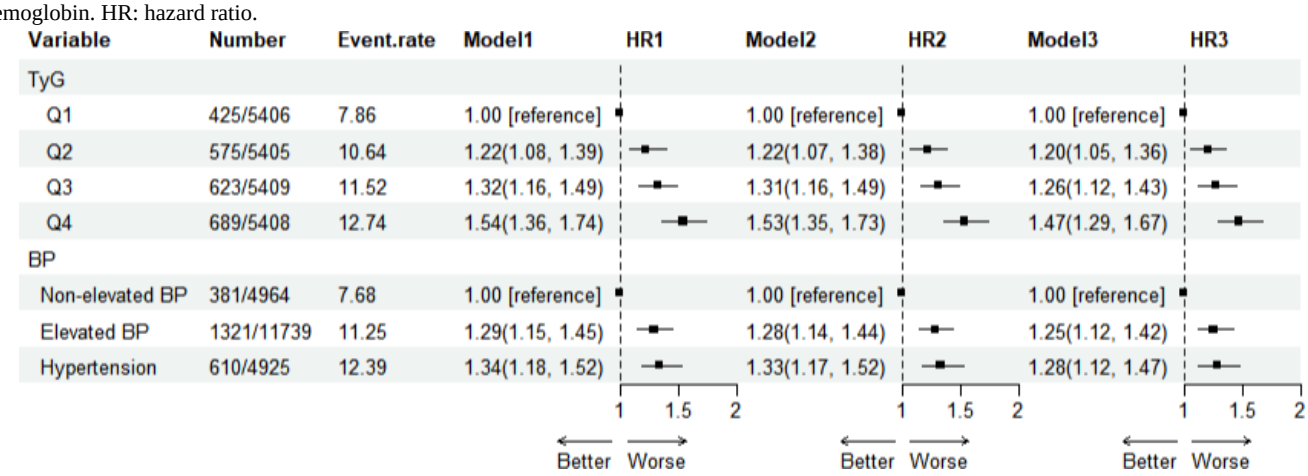
[n]TyG: triglyceride-glucose. The TyG was calculated using the formula ln [TC (mg/dl)×FBG (mg/dl)/2].

Association of TyG and BP With CHD

During a median follow-up of 77.45 (IQR 44.81-108.17) months from 2012 to 2022, a total of 2312 participants developed CHD, resulting in an incidence rate of 10.69%. Figure 1 illustrates the association between BP and the TyG index with CHD events. After adjusting for potential confounders in model 3, compared to nonelevated BP, the adjusted HR (aHR) for CHD events in the hypertension group was 1.28 (95% CI 1.12-1.47). When comparing TyG quartile 4 to quartile 1, the aHR for CHD events was 1.47 (95% CI 1.29-1.67).

XSL•FO
RenderX

Figure 1. Associations of triglyceride-glucose (TyG) index and blood pressure (BP) with coronary heart disease (CHD). Model 1 was adjusted for age and gender. Model 2 was adjusted for age, gender, classification of disabilities, and grading of disabilities. Model 3 was adjusted for age, gender, classification of disabilities, grading of disabilities, marital status, education level, total cholesterol (TC), serum creatinine (SCr), total protein (TP), and hemoglobin. HR: hazard ratio.



Mediation and Interaction Analyses of TyG Between BP and CHD

The mediation analysis of BP and TyG in relation to the incidence of CHD is shown in Figure S3 in Multimedia Appendix 1 (model 3). The proportion of mediation of BP on CHD incidence through TyG was 20.50% (95% CI 13.60%-22%). Significant multiplicative interactions were observed between TyG and BP on incident CHD (Table 2; multiplicative, HR 1.41, 95% CI 1.02 - 1.94). Figure 2 presents the results of the joint analysis of BP and TyG on the outcome. Compared to individuals with nonelevated BP and TyG in the first quartile, those with hypertension and TyG in the fourth quartile had a higher risk of CHD incidence, with an HR of 1.92 (95% CI 1.52 - 2.42) after adjusting for confounding factors.

Table . Interactive effects of triglyceride-glucose index and blood pressure (BP) on coronary heart disease (CHD).

Interactive items	Interactive effects (95% CI)		
	Model 1[a]	Model 2[b]	Model 3[c]
Additive effects			
Relative excess risk due to interaction	1.22 (–0.31 to 2.74)	1.18 (–0.32 to 2.67)	1.01 (–0.36 to 2.39)
Proportion attributable to interaction	0.39 (0.15 to 0.62)	0.38 (0.14 to 0.62)	0.35 (0.10 to 0.60)
Synergy index	2.32 (1.48 to 3.63)	2.29 (1.45 to 3.59)	2.18 (1.36 to 3.50)
Multiplicative effect	1.47 (1.07 to 2.03)	1.46 (1.06 to 2.01)	1.41 (1.02 to 1.94)

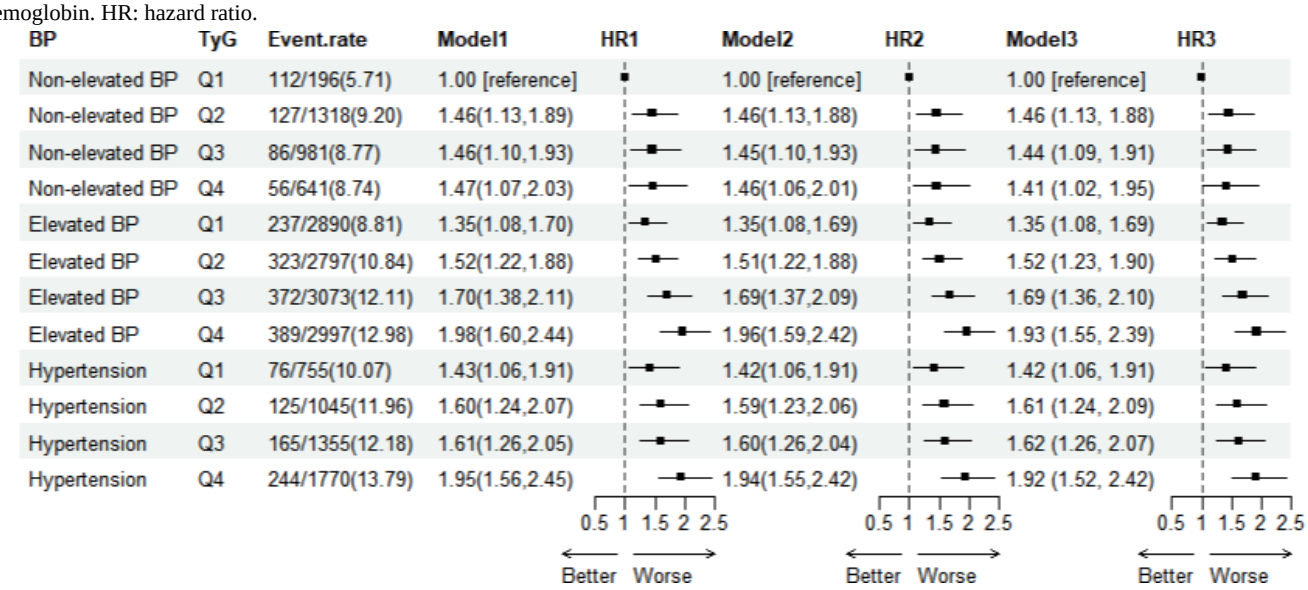
[a]Model 1 was adjusted for age and gender.

[b]Model 2 was adjusted for age, gender, classification of disabilities, and grading of disabilities.

[c]Model 3 was adjusted for age, gender, classification of disabilities, grading of disabilities, marital status, education level, total cholesterol (TC), serum creatinine (SCr), total protein (TP), and hemoglobin.

XSL•FO
RenderX

Figure 2. Joint associations of triglyceride-glucose (TyG) index and blood pressure (BP) on coronary heart disease (CHD). Model 1 was adjusted for age and gender. Model 2 was adjusted for age, gender, classification of disabilities, and grading of disabilities. Model 3 was adjusted for age, gender, classification of disabilities, grading of disabilities, marital status, education level, total cholesterol (TC), serum creatinine (SCr), total protein (TP), and hemoglobin. HR: hazard ratio.



Group Analysis by Disability Characteristics

Global interaction tests between disability grading or classification and the combined effects of BP and TyG on CHD showed significant additive and multiplicative interactions after adjustment for covariates (Table 3). Stratified mediation analyses by disability classification (Figure 3A) revealed that the mediation effect of TyG was significant among individuals with physical disabilities (22.64%, 95% CI 9.03% - 60%) and those with visual disabilities (16.58%, 95% CI 4.75% - 51%), whereas the effect was not significant in participants with hearing and speech disabilities (*P*=.074) or intellectual and mental disabilities (*P*=.132). When stratified by disability grading (Figure 3B), the proportion mediated was 22.3% (95% CI 9.24% - 59.38%) in the mild or moderate group and 18.8% (95% CI 15.7% - 29%) in the severe or very severe group.

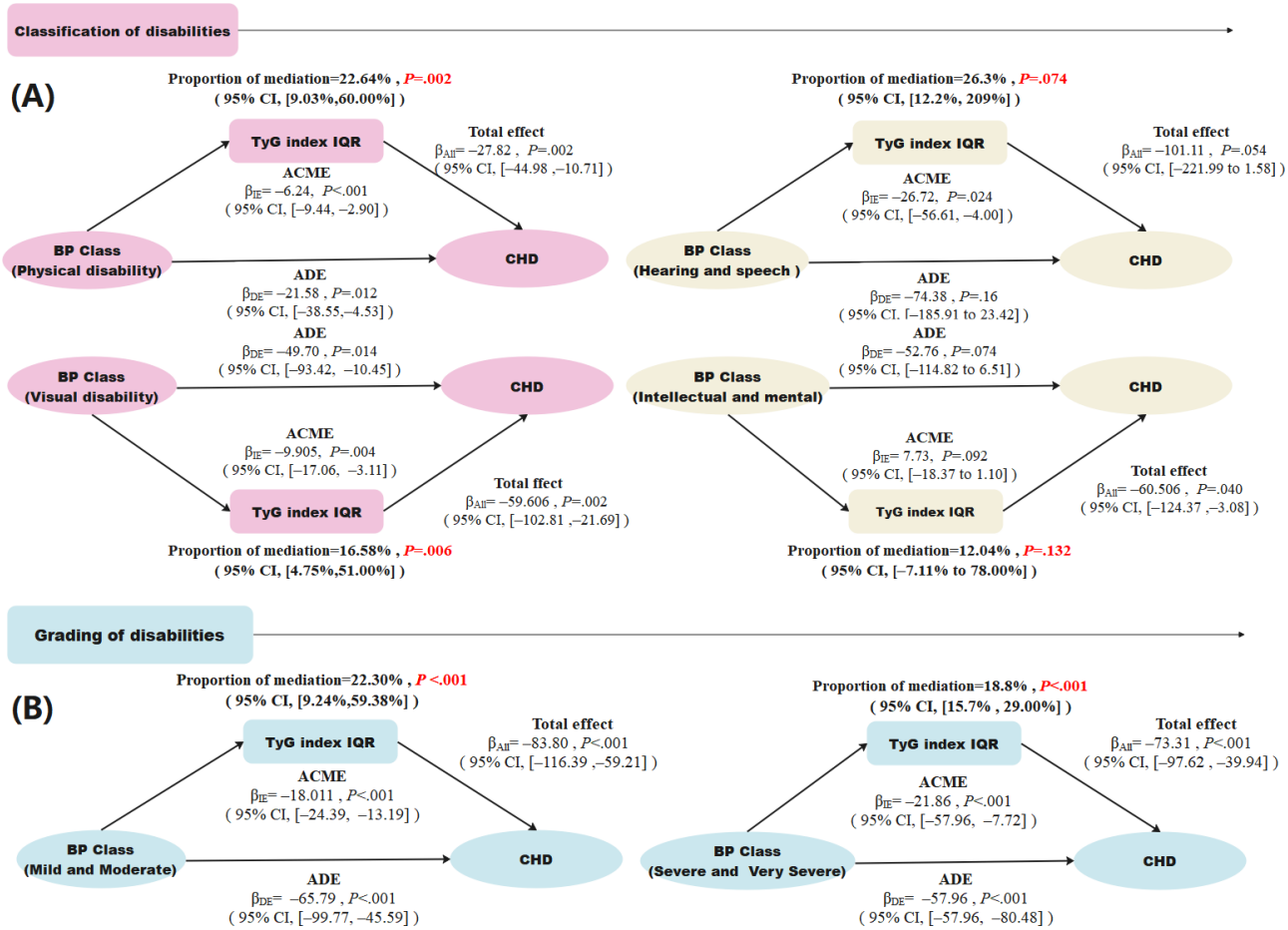
The joint association analyses (Table S5 in Multimedia Appendix 1) further showed that, compared with individuals in the first quartile of both nonhypertension and TyG, those in the fourth quartile of both hypertension and TyG exhibited a higher risk of CHD across most disability classifications. Specifically, the aHRs were 1.81 (95% CI 1.34 - 2.44) for physical disabilities, 1.31 (95% CI 0.77 - 2.23) for visual disabilities, 2.81 (95% CI 1.11 - 7.09) for hearing and speech disabilities, and 3.51 (95% CI 1.89 - 6.50) for intellectual and mental disabilities. Similarly, by disability grading, the HRs were 2.45 (95% CI 1.49 - 4.02) for the severe or very severe group and 1.84 (95% CI 1.41 - 2.39) for the mild or moderate group.

Table . Interactive effects of disability classification/grading on blood pressure (BP) and triglyceride-glucose index in relation to coronary heart disease (CHD).

Interactive items	Interactive effects (95% CI)	
	Classification of disabilities	Grading of disabilities
Additive effects		
Relative excess risk due to interaction	0.61 (0.42-1.64)	0.86 (0.27-1.46)
Proportion attributable to interaction	0.41 (0.18-0.98)	0.35 (0.23-0.47)
Synergy index	1.83 (0.88-3.83)	2.45 (2.04-2.93)
Multiplicative effect	1.84 (0.83-1.99)	1.46 (1.27-1.69)

XSL•FO
RenderX

Figure 3. Mediating role of triglyceride-glucose (TyG) index in the association between blood pressure (BP) and coronary heart disease (CHD) according to classification and grading of disability. (A) Mediating role of TyG index in the association between BP and CHD according to the classification of disability and (B) mediating role of the TyG index in the association between BP and CHD according to the grading of disability. All the models were adjusted for age, gender, classification of disabilities, grading of disabilities, marital status, education level, TC, SCr, total protein (TP), and hemoglobin. ACME: average casual mediation effect; ADE: average direct effect; β_{DE}: direct effect; β_{IE}: indirect effect.



Subgroup and Sensitivity Analyses

After converting SBP, DBP, and TyG into continuous variables, the mediation analysis was conducted with full adjustment for confounding factors. The results shown in Figure S4 in Multimedia Appendix 1 indicate mediation proportions of 34.5% (*P*<.001) and 34.1% (*P*<.001). The mediation effect of TyG on the association between BP and CHD is greater in participants aged <60 years than in those aged ≥60 years and is more pronounced in male individuals than in female individuals. In the joint analysis, participants in the highest quartile of both hypertension and TyG had an increased risk of CHD compared with those in the lowest quartile. The aHRs were 2.69 (95% CI 1.84 - 3.93) in male individuals, 1.50 (95% CI 1.11 - 2.02) in female individuals, 2.01 (95% CI 1.59 - 2.76) in participants aged <60 years, and 1.58 (95% CI 1.02 - 2.44) in those aged ≥60 years.

Discussion

Principal Findings

This study expands the understanding of the association between BP, TyG, and CHD by including individuals with disabilities, a group that has often been overlooked in previous research. Our research fills a critical gap by elucidating the significant independent associations of BP and TyG with CHD risk in a cohort of 21,628 individuals with disabilities. Notably, TyG was found to mediate the effect of BP on CHD risk, with this mediation being especially pronounced in individuals with physical disabilities and those with mild or moderate disabilities. Furthermore, significant additive and multiplicative interactions between BP and TyG were observed, with their combination associated with an even higher risk of CHD. These findings suggest that the unique metabolic profiles of these populations may play an important role in their increased CHD risk.

Our results are consistent with existing literature showing that both high BP [24,25] and elevated TyG [26,27] are independent risk factors for CHD. The association between BP and TyG in relation to CVD risk has been studied in normal populations before. One prospective study [17] demonstrated that SBP significantly altered the association between TyG and CVD, with SBP mediating about 10% of the TyG-CVD association, a finding similar to what we observed in our cohort of individuals with disabilities. This suggests that in the population with disabilities, TyG plays a significant role in the pathway linking BP to CHD, possibly reflecting enhanced insulin resistance and metabolic abnormalities in this group. Potential mechanisms for this could include limited physical activity, which often leads to lower metabolic rates and worsened

XSL•FO
RenderX

vascular function [28], making TyG a stronger mediator of BP and CHD [29].

A significant synergistic interaction between BP and TyG was identified in our analysis. This result is consistent with prior evidence. A recent Kailuan study [29] also reported a significant interaction between the TyG index and BP status in relation to CVD risk. In addition, a population-based analysis [30] from the National Health and Nutrition Examination Survey found that the combination of a low TyG index and low SBP (<120 mm Hg or <130 mm Hg) was associated with a further reduction in all-cause and cardiovascular mortality, with a particularly strong effect on cardiovascular mortality. Several mechanisms may underlie the interplay between the TyG index and BP in CVD development. Hypertension promotes target organ damage, while an elevated TyG index reflects insulin resistance [31], which drives hyperglycemia, inflammation, oxidative stress, and endothelial dysfunction [32,33]. Together, these factors can accelerate cardiovascular injury and substantially increase CVD risk. Clarifying their joint effects is therefore essential for improving risk assessment and tailoring prevention strategies.

Our findings highlight the need to account for both BP and TyG in CHD risk assessment. The combined assessment of these 2 factors may help identify individuals with disabilities who are at higher risk for CHD. Our study shows that individuals with both elevated BP and TyG have a significantly higher CHD risk compared to those with either factor alone. A few other studies have also reached similar conclusions [17,31]. This suggests a complex, nonlinear association between the two, rather than a simple additive effect. Indeed, prior research has established a significant association between the TyG index and an increased risk of developing hypertension [34-38]. Given this, future risk assessment tools or clinical practices may benefit from integrating BP and TyG management, particularly in individuals with metabolic abnormalities and hypertension, to more effectively manage CHD risk.

Stratified analyses revealed that the mediating role of TyG was particularly strong in physical disabilities, with a moderate effect in visual disabilities, while no significant mediation was detected in hearing and speech or intellectual and mental disabilities. Moreover, the mediation proportion was greater in mild to moderate disabilities than in severe or very severe disabilities. This finding suggests that different types and degrees of disability may amplify the impact of BP on CHD through metabolic pathways. In individuals with physical disabilities or visual disabilities, reduced physical activity and energy expenditure often lead to increased body fat and insulin resistance, which collectively elevate TyG levels [39]. Given the prevalence of metabolic abnormalities in this population, TyG may enhance the impact of BP on CHD risk by reflecting insulin resistance and lipid metabolism disturbances [40]. Furthermore, reduced physical activity may trigger systemic inflammation and endothelial dysfunction, further magnifying the mediating role of TyG. Future research is needed to explore these mechanisms in greater depth and validate the role of TyG in different disability groups. Among individuals with mild to moderate disabilities, the significant mediating effect of TyG may reflect the group's heightened sensitivity to metabolic stress. These individuals may retain some physiological resilience, but it is limited, making them more vulnerable to the combined impact of elevated BP and higher TyG levels on CHD risk. Furthermore, the relatively preserved health reserves in this group may allow for more precise detection of metabolic effects, which might be overshadowed by other health factors in those with severe disabilities. This highlights the need for targeted interventions to address metabolic health in individuals with disabilities, particularly those with physical disabilities or mild to moderate severity, who may be more susceptible to the metabolic factors that contribute to CHD.

Strengths and Limitations

The findings of this study have important implications. First, the large sample size of 21,628 individuals with disabilities provides high statistical power and generalizability. Second, this study enhances our understanding of the contribution of metabolic factors to cardiovascular risk in populations with disabilities and addresses a critical gap in the existing literature. Third, by using a novel analytical approach that decomposes the total association between BP, TyG, and CHD into direct and indirect effects, we offer a deeper understanding of the critical role of metabolic factors in cardiovascular health [41]. The stratified analyses further enhance the precision of our findings and provide valuable insights for developing targeted interventions.

However, there are several limitations to this study. First, although the study benefits from a large sample size, its retrospective cohort design inherently limits causal inference, restricting the ability to draw definitive conclusions about the temporal association between exposures and outcomes. In addition, we recognize that factors such as lifestyle behaviors, health care access, and socioeconomic status are important, but due to current data limitations, we have adjusted for other objective health indicators in an attempt to minimize their potential impact. Additionally, the uniform access to living support provided by the Shanghai Disability Federation for all participants may help reduce the influence of these unmeasured factors. Third, since all data were collected from Shanghai, the generalizability of the findings to other regions or countries may be limited, and future studies will aim to expand the data by incorporating additional variables and conducting multi-center prospective cohort studies to validate these results in a broader context.

Conclusions

In conclusion, we identified a significant association between BP and the TyG index with CHD risk in individuals with disabilities, with TyG acting as a prominent mediator in the association between BP and CHD. This mediation effect was particularly pronounced in individuals with physical disabilities and those classified with mild-to-moderate disability severity. Our findings underscore the importance of addressing the combined effects of insulin resistance and blood pressure in high-risk populations. Integrated intervention strategies targeting both factors may be crucial for reducing CHD risk.

XSL•FO
RenderX

Acknowledgments

The study was funded by the Health Management Outstanding Talents of Shanghai Association of Rehabilitation Medical (2023JGYQ08); the project supported by National Natural Science Foundation of China (72474136); National clinical key specialty construction project of China (Z155080000004); Shanghai Research Center of Rehabilitation Medicine (Top Priority Research Center of Shanghai; 2023ZZ02027); Shanghai Hospital Development Center Foundation—Shanghai Municipal Hospital Rehabilitation Medicine Specialty Alliance (SHDC22023304) and Shanghai Clinical Research Ward (SHDC2023CRW018B). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

HL and YL analyzed the participants' data and were major contributors in writing the manuscript. All authors participated in data collection and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Multimedia appendix. Additional tables and figures supporting the main results.
[DOCX File, 427 KB - publichealth_v11i1e78068_app1.docx]

References

1. Mensah GA, Fuster V, Murray CJL, Roth GA. Global burden of cardiovascular diseases and risks, 1990-2022. J Am Coll Cardiol 2023 Dec 19;82(25):2350-2473. [doi: 10.1016/j.jacc.2023.11.007] [Medline: 38092509]
2. Bevan G, Pandey A, Griggs S, et al. Neighborhood-level social vulnerability and prevalence of cardiovascular risk factors and coronary heart disease. Curr Probl Cardiol 2023 Aug;48(8):101182. [doi: 10.1016/j.cpcardiol.2022.101182] [Medline: 35354074]
3. Rosoff DB, Davey Smith G, Mehta N, Clarke TK, Lohoff FW. Evaluating the relationship between alcohol consumption, tobacco use, and cardiovascular disease: a multivariable Mendelian randomization study. PLoS Med 2020 Dec 4;17(12):e1003410. [doi: 10.1371/journal.pmed.1003410] [Medline: 33275596]
4. McEvoy JW, Jennings C, Kotseva K, et al. Variation in secondary prevention of coronary heart disease: the INTERASPIRE study. Eur Heart J 2024 Oct 14;45(39):4184-4196. [doi: 10.1093/eurheartj/ehae558] [Medline: 39212219]
5. Global report on health equity for persons with disabilities. World Health Organization. 2022. URL: https://www.who.int/publications/i/item/9789240063600 [accessed 2025-10-07]
6. Fong JH. Disability incidence and functional decline among older adults with major chronic diseases. BMC Geriatr 2019 Nov 21;19(1):323. [doi: 10.1186/s12877-019-1348-z] [Medline: 31752701]
7. GBD 2016 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1345-1422. [doi: 10.1016/S0140-6736(17)32366-8] [Medline: 28919119]
8. Li Y, Wu J, Wang Y, et al. Association of blood pressure trajectories with coronary heart disease among the disabled population in Shanghai, China: a cohort study of 7 years following up. Eur J Med Res 2023 Aug 9;28(1):275. [doi: 10.1186/s40001-023-01240-1] [Medline: 37553723]
9. Alizargar J, Bai CH, Hsieh NC, Wu SF. Use of the triglyceride-glucose index (TyG) in cardiovascular disease patients. Cardiovasc Diabetol 2020 Jan 15;19(1):8. [doi: 10.1186/s12933-019-0982-2] [Medline: 31941513]
10. Mo D, Zhang P, Zhang M, Dai H, Guan J. Cholesterol, high-density lipoprotein, and glucose index versus triglyceride-glucose index in predicting cardiovascular disease risk: a cohort study. Cardiovasc Diabetol 2025 Mar 10;24(1):116. [doi: 10.1186/s12933-025-02675-y] [Medline: 40065297]
11. Qiao Y, Wang Y, Chen C, Huang Y, Zhao C. Association between triglyceride-glucose (TyG) related indices and cardiovascular diseases and mortality among individuals with metabolic dysfunction-associated steatotic liver disease: a cohort study of UK Biobank. Cardiovasc Diabetol 2025 Jan 13;24(1):12. [doi: 10.1186/s12933-024-02572-w] [Medline: 39806394]
12. Huo RR, Liao Q, Zhai L, You XM, Zuo YL. Interacting and joint effects of triglyceride-glucose index (TyG) and body mass index on stroke risk and the mediating role of TyG in middle-aged and older Chinese adults: a nationwide prospective cohort study. Cardiovasc Diabetol 2024 Jan 13;23(1):30. [doi: 10.1186/s12933-024-02122-4] [Medline: 38218819]

XSL•FO
RenderX

13. Yang Z, Gong H, Kan F, Ji N. Association between the triglyceride glucose (TyG) index and the risk of acute kidney injury in critically ill patients with heart failure: analysis of the MIMIC-IV database. Cardiovasc Diabetol 2023 Aug 31;22(1):232. [doi: 10.1186/s12933-023-01971-9] [Medline: 37653418]
14. Yao Y, Wang B, Geng T, Chen J, Chen W, Li L. The association between TyG and all-cause/non-cardiovascular mortality in general patients with type 2 diabetes mellitus is modified by age: results from the cohort study of NHANES 1999–2018. Cardiovasc Diabetol 2024;23(1):43. [doi: 10.1186/s12933-024-02120-6]
15. Cui C, Liu L, Qi Y, et al. Joint association of TyG index and high sensitivity C-reactive protein with cardiovascular disease: a national cohort study. Cardiovasc Diabetol 2024 May 7;23(1):156. [doi: 10.1186/s12933-024-02244-9] [Medline: 38715129]
16. Wu K, Zheng H, Wu W, et al. Temporal relationship between triglyceride-glucose index and blood pressure and their joint cumulative effect on cardiovascular disease risk: a longitudinal cohort study. Cardiovasc Diabetol 2023 Nov 28;22(1):332. [doi: 10.1186/s12933-023-02058-1] [Medline: 38017521]
17. Qiu W, Cai A, Li L, Feng Y. Systolic blood pressure status modifies the associations between the triglyceride-glucose index and incident cardiovascular disease: a national cohort study in China. Cardiovasc Diabetol 2024 Apr 24;23(1):135. [doi: 10.1186/s12933-024-02227-w] [Medline: 38658924]
18. Sheppard JP, Albasri A, Franssen M, et al. Defining the relationship between arm and leg blood pressure readings: a systematic review and meta-analysis. J Hypertens 2019 Apr;37(4):660-670. [doi: 10.1097/HJH.0000000000001958] [Medline: 30817444]
19. Chinese Association of Rehabilitation Medicine, Hypertension Committee of Cross-Straits Medicine Exchange Association, Cardiovascular Disease Prevention and Rehabilitation Committee, Chinese Association of Rehabilitation Medicine. Clinical practice guidelines for the management of hypertension in China. Zhonghua Xin Xue Guan Bing Za Zhi 2024 Sep 24;52(9):985-1032. [doi: 10.3760/cma.j.cn112148-20240709-00377] [Medline: 39289992]
20. Xu X, Huang R, Lin Y, et al. High triglyceride-glucose index in young adulthood is associated with incident cardiovascular disease and mortality in later life: insight from the CARDIA study. Cardiovasc Diabetol 2022 Aug 12;21(1):155. [doi: 10.1186/s12933-022-01593-7] [Medline: 35962377]
21. Zhang R, Shi S, Chen W, et al. Independent effects of the triglyceride-glucose index on all-cause mortality in critically ill patients with coronary heart disease: analysis of the MIMIC-III database. Cardiovasc Diabetol 2023 Jan 13;22(1):10. [doi: 10.1186/s12933-023-01737-3] [Medline: 36639637]
22. National standards information public service platform of China. National Public Service Platform for Standards Information. URL: https://std.samr.gov.cn/gb/search/gbDetailed?id=71F772D7E01ED3A7E05397BE0A0AB82A [accessed 2024-11-15]
23. Lim CC, Hayes RB, Ahn J, et al. Mediterranean diet and the association between air pollution and cardiovascular disease mortality risk. Circulation 2019 Apr 9;139(15):1766-1775. [doi: 10.1161/CIRCULATIONAHA.118.035742] [Medline: 30700142]
24. Colafella KM, Denton KM. Sex-specific differences in hypertension and associated cardiovascular disease. Nat Rev Nephrol 2018 Mar;14(3):185-201. [doi: 10.1038/nrneph.2017.189] [Medline: 29380817]
25. Brunström M, Carlberg B. Association of blood pressure lowering with mortality and cardiovascular disease across blood pressure levels: a systematic review and meta-analysis. JAMA Intern Med 2018 Jan 1;178(1):28-36. [doi: 10.1001/jamainternmed.2017.6015] [Medline: 29131895]
26. Wan Y, Zhang Z, Ling Y, et al. Association of triglyceride-glucose index with cardiovascular disease among a general population: a prospective cohort study. Diabetol Metab Syndr 2023 Oct 16;15(1):204. [doi: 10.1186/s13098-023-01181-z] [Medline: 37845738]
27. Hong S, Han K, Park CY. The triglyceride glucose index is a simple and low-cost marker associated with atherosclerotic cardiovascular disease: a population-based study. BMC Med 2020 Nov 25;18(1):361. [doi: 10.1186/s12916-020-01824-2] [Medline: 33234146]
28. Wu J, Wang Y, Li Y, et al. Are physically disabled people at high risk of coronary heart disease among disabled population - evidence from 7.5-year retrospective cohort study. Ann Epidemiol 2024 Feb;90:42-48. [doi: 10.1016/j.annepidem.2023.11.001] [Medline: 37926391]
29. Cui H, Liu Q, Zhao Z, Ma X. Interacting and joint effects of triglyceride-glucose index and blood pressure on cardiovascular diseases risk: a prospective cohort study. Diabetol Metab Syndr 2024 Aug 2;16(1):188. [doi: 10.1186/s13098-024-01433-6] [Medline: 39095824]
30. Yu Y, Gu M, Huang H, et al. Combined association of triglyceride-glucose index and systolic blood pressure with all-cause and cardiovascular mortality among the general population. J Transl Med 2022 Oct 20;20(1):478. [doi: 10.1186/s12967-022-03678-z] [Medline: 36266665]
31. Khan SH, Sobia F, Niazi NK, Manzoor SM, Fazal N, Ahmad F. Metabolic clustering of risk factors: evaluation of triglyceride-glucose index (TyG index) for evaluation of insulin resistance. Diabetol Metab Syndr 2018;10:74. [doi: 10.1186/s13098-018-0376-8] [Medline: 30323862]
32. Molina MN, Ferder L, Manucha W. Emerging role of nitric oxide and heat shock proteins in insulin resistance. Curr Hypertens Rep 2016 Jan;18(1):1. [doi: 10.1007/s11906-015-0615-4] [Medline: 26694820]

XSL•FO
RenderX

33. Yang Q, Vijayakumar A, Kahn BB. Metabolites as regulators of insulin sensitivity and metabolism. Nat Rev Mol Cell Biol 2018 Oct;19(10):654-672. [doi: 10.1038/s41580-018-0044-8] [Medline: 30104701]
34. Gao Q, Lin Y, Xu R, et al. Positive association of triglyceride-glucose index with new-onset hypertension among adults: a national cohort study in China. Cardiovasc Diabetol 2023 Mar 16;22(1):58. [doi: 10.1186/s12933-023-01795-7] [Medline: 36927705]
35. Lee DH, Park JE, Kim SY, Jeon HJ, Park JH. Association between the triglyceride-glucose (TyG) index and increased blood pressure in normotensive subjects: a population-based study. Diabetol Metab Syndr 2022 Oct 29;14(1):161. [doi: 10.1186/s13098-022-00927-5] [Medline: 36309720]
36. Tsai KZ, Chu CC, Huang WC, Sui X, Lavie CJ, Lin GM. Prediction of various insulin resistance indices for the risk of hypertension among military young adults: the CHIEF cohort study, 2014-2020. Cardiovasc Diabetol 2024 Apr 25;23(1):141. [doi: 10.1186/s12933-024-02229-8] [Medline: 38664804]
37. Miao H, Zhou Z, Yang S, Zhang Y. The association of triglyceride-glucose index and related parameters with hypertension and cardiovascular risk: a cross-sectional study. Hypertens Res 2024 Apr;47(4):877-886. [doi: 10.1038/s41440-023-01502-9] [Medline: 38012411]
38. Lee JH, Heo SJ, Kwon YJ. Sex-specific comparison between triglyceride glucose index and modified triglyceride glucose indices to predict new-onset hypertension in middle-aged and older adults. J Am Heart Assoc 2023 Sep 19;12(18):e030022. [doi: 10.1161/JAHA.123.030022] [Medline: 37721166]
39. Ahmadi MN, Hamer M, Gill JM, et al. Brief bouts of device-measured intermittent lifestyle physical activity and its association with major adverse cardiovascular events and mortality in people who do not exercise: a prospective cohort study. Lancet Public Health 2023 Oct;8(10):e800-e810. [doi: 10.1016/S2468-2667(23)00183-4] [Medline: 37777289]
40. Dempsey PC, Rowlands AV, Strain T, et al. Physical activity volume, intensity, and incident cardiovascular disease. Eur Heart J 2022 Dec 7;43(46):4789-4800. [doi: 10.1093/eurheartj/ehac613] [Medline: 36302445]
41. Fasanelli F, Giraudo MT, Ricceri F, Valeri L, Zugna D. Marginal time-dependent causal effects in mediation analysis with survival data. Am J Epidemiol 2019 May 1;188(5):967-974. [doi: 10.1093/aje/kwz016] [Medline: 30689682]

Abbreviations

aHR: adjusted hazard ratio
BP: blood pressure
CHD: coronary heart disease
CVD: cardiovascular disease
DBP: diastolic blood pressure
HR: hazard ratio
NIE: natural indirect effect
RCS: restricted cubic spline
SBP: systolic blood pressure
TyG: triglyceride-glucose

Edited by A Mavragani, T Sanchez; submitted 26.05.25; peer-reviewed by L Ren, S Li; revised version received 25.09.25; accepted 28.09.25; published 03.11.25.

Please cite as:
Liu H, Li Y, Wang Y, Wang T, Jiang C, Wu H, Wu J
Association Between Triglyceride-Glucose Index, Blood Pressure Status, and Coronary Heart Disease Risk Among Chinese Adults With Disabilities: 10-Year Disability Health Survey Cohort Study
JMIR Public Health Surveill 2025;11:e78068
URL: https://publichealth.jmir.org/2025/1/e78068
doi:10.2196/78068

© Hui Liu, Yao Li, Yiyan Wang, Tao Wang, Chenghua Jiang, Hengjing Wu, Jing Wu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 3.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Unmet Health Care Use Among Socially Withdrawn Youth (Hikikomori) in South Korea: Cross-Sectional Survey Study

Eunjeong Choi[1], PhD; Seoyeong Choi[1], BS; Sukyong Jang[2,3], MD, JD, PhD

[1]Department of Public Health, Graduate School, Yonsei University, Seoul, Republic of Korea

[2]Department of Biohealth Industry, Policy Analysis Division, Graduate School of Transdisciplinary Health Science, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

[3]Department of Health Policy and Management, Graduate School of Public Health, Yonsei University, Seoul, Republic of Korea

Corresponding Author:
Sukyong Jang, MD, JD, PhD
Department of Biohealth Industry, Policy Analysis Division, Graduate School of Transdisciplinary Health Science, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea

Abstract

Background: Hikikomori, a condition of severe social withdrawal, is a global public health issue characterized by prolonged isolation. Despite its growing prevalence, little is known about the health care needs and use patterns of socially withdrawn youth.

Objective: The study aimed to examine the association between hikikomori status and unmet health care use to inform targeted interventions.

Methods: Data were obtained from the 2022 Korean Youth Living Conditions Survey, a nationally representative cross-sectional survey of 14,966 participants aged 19 to 34 years. Survey weights were applied to account for the sampling design. Hikikomori status was classified based on self-reported withdrawal behaviors, stratified by severity and duration. Unmet health care use in physical and mental health was assessed. Adjusted prevalence ratios (aPRs) with 95% CIs were estimated using generalized estimating equation models, and logistic regression was applied for subgroup analyses.

Results: The weighted prevalence of perceived need for mental health services was 63.9% in the hikikomori group versus 50.5% in the non-hikikomori group (aPR 1.27, 95% CI 1.15-1.41). Unmet health care use was higher among individuals with hikikomori for physical care (aPR 3.33, 95% CI 2.16-5.13) and mental health care (aPR 4.46, 95% CI 2.92-6.81). Associations strengthened with greater severity and longer duration: for unmet mental health care use, aPRs were 4.14 (95% CI 2.64-6.49) for stage 1 and 9.52 (95% CI 3.67-24.65) for stage 2; by duration, aPRs were 2.57 (95% CI 1.11-5.96) for pre-hikikomori and 5.44 (95% CI 3.44-8.58) for hikikomori. Effect modification was observed by labor force participation, with higher risks among those not in the labor force (*P* for interaction <.05).

Conclusions: Hikikomori is strongly associated with unmet health care use, particularly in mental health, with risks amplified by severity and duration. Tailored policies, including community-based outreach and remote health care interventions, are urgently needed to address these gaps.

(JMIR Public Health Surveill 2025;11:e79043) doi:10.2196/79043

KEYWORDS

mental health; social isolation; health services accessibility; health care disparities; health services needs and demand

Introduction

Background

The term “social isolation” is also used as a translation for the Japanese term “hikikomori,” which refers to individuals who have withdrawn from social interactions [1]. Hikikomori, initially defined in academic research, pertains to Japanese youth who withdraw from social interaction and primarily remain in their rooms, limiting their activities outside the home [1,2]. Although hikikomori was originally considered a unique cultural phenomenon in Japan during the 1990s [3], it has since emerged as a significant global issue [1,4-7]. Hikikomori is defined as remaining physically isolated in one’s home for at least 6 months, with individuals experiencing social isolation for at least 3 months proposed to be classified as “pre-hikikomori” [8,9]. Research conducted in Korea [10] and Hong Kong [11] used a shorter duration criterion (3 months) to define severe social withdrawal; therefore, defining the duration criteria for social withdrawal remains challenging.

In Japan, the estimated prevalence of hikikomori among individuals aged 15 to 39 years is approximately 2.3% [12]. In South Korea, recent survey data indicated a point prevalence of 2.4% in the national youth population, equivalent to an estimated 244,000 individuals [13]. Recent studies suggest that

XSL•FO
RenderX

it may represent a broader global phenomenon. For example, analyses of the European Social Survey across 29 countries have estimated the prevalence of severe social withdrawal at approximately 1.7% to 1.8% [14]. Hikikomori is a global public health issue often accompanied by significant functional impairment and distress [9]. A considerable proportion of individuals diagnosed with hikikomori also exhibit concurrent mental health conditions, such as depressive disorders or schizophrenia [15]. Studies have consistently identified a high risk of suicide in individuals with hikikomori [8,16]. Social isolation and loneliness are also prevalent phenomena identified as independent risk factors for adverse cardiovascular and brain health [17] and are linked to various health behaviors, such as smoking cessation [18].

The importance of considering hikikomori in health care lies in their limited social interactions and restricted mobility, which may hinder their access to and use of health care resources. These characteristics make it difficult for individuals with hikikomori to meet their health care needs. Self-reported unmet health care needs have been widely adopted as an indicator of health care accessibility [19-21]. Self-reported unmet health care needs are strongly associated with socioeconomic status [22]. Moreover, many factors, such as income, ethnicity, and sexual minorities, hinder young people from accessing necessary mental health care [22-25]. However, empirical evidence regarding unmet health care use among individuals with hikikomori is lacking. Research includes studies examining treatment recommendations for hikikomori from the perspective of psychiatrists [4] and others exploring treatment preferences among individuals with hikikomori [7]. Nonetheless, there is a lack of research on perceived unmet health care use among individuals with hikikomori. To develop and apply adequate interventions for hikikomori, it is necessary to investigate the unmet health care use of individuals with hikikomori.

Objectives

This study examined associations between hikikomori status and unmet health care use, as well as perceived need for mental health services, and assessed whether these associations varied by hikikomori severity and duration. The study also aimed to provide evidence to inform policies and interventions to improve the health of this population.

Methods

Data

This study used data from the 2022 Korean Youth Living Conditions Survey [13]. The Office of Government Policy Coordination in South Korea conducted the 2022 Youth Living Conditions Survey to collect baseline data for developing policies that aim to improve the quality of life for a diverse range of young people. The survey was conducted under the legal mandate of national youth policy legislation and is recognized as an official government statistical survey. This survey was conducted to understand the current state of young people's lives and their characteristics, needs, and perceptions. Conducted biennially by the Korea Institute for Health and Social Affairs (KIHASA), this cross-sectional survey provides a comprehensive assessment across 8 domains: general characteristics, housing, health, education and training, employment, relationships and participation, social awareness, future planning, and economics. Data were collected through face-to-face interviews conducted by trained interviewers using structured questionnaires. To ensure a nationally representative sample of the South Korean youth population, a 2-phase probability sampling design was adopted, involving cluster sampling in the first stage and stratified cluster sampling in the second. The target population consisted of young adults aged 19 to 34 years residing in ordinary households across 17 administrative regions in South Korea. A stratified multistage sampling design was employed to ensure representativeness. Data were collected between July 18, 2022, and August 26, 2022.

Participants

In total, 14,966 participants were included in the 2022 Korean Youth Living Conditions Survey. The study population was drawn from the 2020 Population and Housing Census Youth Statistics Register and comprised individuals aged 19 to 34 years as of January 1, 2022, who were members of ordinary households and were selected using a 2-stage sampling method. Ordinary households included private family households, while group quarters such as dormitories or institutions were excluded. The analytic sample thus consisted of all respondents who completed the survey.

Hikikomori Status

Hikikomori status, the primary variable, was determined using a national youth life survey [13]. The Korean government classified individuals as hikikomori if they preferred staying at home based on their responses to questions about outdoor activities. The inclusion criteria included (1) staying home except for hobbies, (2) going out only to nearby stores, (3) staying entirely at home, or (4) rarely leaving their room.

Hikikomori Severity and Duration

Hikikomori individuals were further categorized by severity and duration. Severity has 2 stages [13] where stage 1 (moderate isolation) involves leaving home for hobbies or nearby stores, while stage 2 (severe isolation) involves staying entirely at home or rarely leaving one's room. Duration was classified as "pre-hikikomori" (<6 months of withdrawal) [9] or "hikikomori" (≥6 months of withdrawal) [26].

Perceived Need for Mental Health Services

The perceived need for mental health services was a commonly assessed variable in mental health research [27,28]. Participants who reported recognizing the need for professional mental health counseling or treatment within the past year were coded as "yes," whereas those who did not were coded as "no."

Assessment of Unmet Health Care Use

Measuring unmet health care use lacks a universal standard and is influenced by cross-cultural differences and survey methods. International surveys report varying degrees of unmet needs across countries [29]. Participants were asked if they had needed but not received medical care for physical health (excluding mental health) in the past year. For mental health care, they

XSL•FO
RenderX

were asked about unmet needs for professional counseling within the same period [13].

Other Covariates

To account for potential confounding variables associated with the independent and dependent variables, the analysis included the following covariates selected based on previous studies: demographic variables (age and sex); socioeconomic variables (household income quartile and labor force participation); and health-related variables (alcohol consumption, smoking, and physical activity) [19,30,31].

All selected covariates were categorized according to the study objectives:

1. Age groups: 19 - 24, 25 - 29, and 30 - 34 years
2. Living status: living alone, living with parents, and living with others
3. Household income: high, low, and N/A. The participants were divided into quartiles on the basis of equivalized household income. For this analysis, the highest quartile was defined as "high" income and the lowest quartile was defined as "low" income; missing or nonresponse cases were coded as "N/A."
4. Smoking status: the participants were self-reportedly classified as never, former, or current smokers.
5. Problem drinking: defined as the consumption of 6 or more units of alcohol per week for men and 4 or more units per week for women
6. Physical activity: coded as "yes" if participants reported engaging in regular exercise at least 3 times per week for health maintenance and "no" otherwise
7. Employment: the participants were categorized as employed ("yes") if they reported working for pay for at least 1 hour during the previous week. Those who did not meet this criterion were classified as "no."

Statistical Analysis

The data were constructed using a multistage clustering and stratified sampling design to represent the South Korean population, based on data from the Youth Living Conditions Survey. To ensure national representativeness, all analyses incorporated survey sampling weights provided by the KIHASA, which account for the complex survey design, including unequal probabilities of selection and nonresponse adjustments. To estimate adjusted prevalence ratios (aPRs) and 95% CIs for the association between hikikomori status and unmet health care use, we used a generalized estimating equation model with a Poisson distribution and log link function. The generalized estimating equation approach was chosen because it provides robust variance estimates under a complex survey design and allows direct estimation of prevalence ratios, which are more interpretable than odds ratios when the outcome is common. In contrast, linear mixed models are primarily suited for longitudinal data with repeated measures, whereas our study is based on a cross-sectional survey [32]. The PR was adjusted for age, sex, region, family structure, household income, labor status, problematic drinking, smoking, and physical activity. The association between hikikomori status and the prevalence of unmet health care use, stratified by each level of covariates, was evaluated using logistic regression analysis. The PROC SURVEY procedure in SAS (version 9.4 M6; SAS Institute Inc) was used for all statistical analyses.

Ethical Considerations

The study was reviewed by the institutional review board of Severance Hospital and determined to be exempt from full review (4-2024-1320). This study used deidentified secondary microdata from the KIHASA under a data use agreement. For the original data collection, informed consent was obtained by the data custodian; the secondary analysis was granted a waiver of additional consent because only deidentified data were analyzed. All datasets were anonymized with no direct identifiers, and only aggregate results are reported to protect privacy and confidentiality. No additional compensation was provided by this study (any compensation followed the original protocol). No identifiable images were included.

Results

Participant Characteristics

Table 1 shows the sociodemographic characteristics of the study population. Among 14,966 participants, the weighted prevalence of hikikomori was 2.9% (397/14,966). Compared with non-hikikomori individuals, those with hikikomori had substantially lower labor force participation (54/397, 15.5% vs 9434/14,569, 69%) and a higher proportion of low household income (187/397, 53.0% vs 5,022/14,569, 42.0%). They were also less physically active (112/397, 27.3% vs. 4783/14,569, 32.8%).

XSL•FO
RenderX

Table . Sociodemographic characteristics of participants. Values indicate unweighted counts (n) and weighted percentages (%) derived from survey weights (N=14,966).

Variables	Hikikomori status[a]			Standardized difference
	Total	Yes (397/14,966, 2.9%)	No (14,569/14,966, 97.1%)	
Age (y)[b], n (%)				0.09
19-24	7195 (35.3)	179 (33.6)	7016 (35.3)	
25-29	4549 (34.1)	133 (36.7)	4416 (34.1)	
30-34	3222 (30.6)	85 (29.7)	3137 (30.6)	
Sex, n (%)				–0.06
Male	7171 (52.5)	180 (48)	6991 (52.6)	
Female	7795 (47.5)	217 (52)	7578 (47.4)	
Region, n (%)				–0.06
Metropolitan	4938 (53)	146 (59.6)	4792 (52.8)	
Nonmetropolitan	10,028 (47)	251 (40.4)	9777 (47.2)	
Living arrangement[c], n (%)				0.26
Living alone	5355 (22.6)	99 (13.8)	5256 (22.9)	
Living with parents	7832 (57.5)	225 (58.8)	7607 (57.4)	
Living with others	1779 (19.9)	73 (27.4)	1706 (19.7)	
Labor[d], n (%)				–1.23
No	5478 (32.5)	343 (84.5)	5135 (31)	
Yes	9488 (67.5)	54 (15.5)	9434 (69)	
Household income, n (%)				0.29
High	4319 (34.2)	107 (32)	4212 (34.3)	
Low	5209 (42.4)	187 (53)	5022 (42)	
N/A[e]	5438 (23.4)	103 (15)	5335 (23.7)	
Problem drinking[f], n (%)				–0.18
No	12,802 (85.1)	362 (91.5)	12,440 (84.9)	
Yes	2164 (14.9)	35 (8.5)	2129 (15.1)	
Smoking status, n (%)				0.14
Current	2874 (20.1)	54 (14.3)	2820 (20.3)	
Former	1307 (10.1)	38 (9.5)	1269 (10.1)	
Never	10,785 (69.7)	305 (76.2)	10,480 (69.5)	
Physical activity[g], n (%)				–0.10
No	10,071 (67.4)	285 (72.7)	9786 (67.2)	
Yes	4895 (32.6)	112 (27.3)	4783 (32.8)	

[a]Hikikomori status was defined as a result of the Korea Life Survey conducted by the Korea Institute for Health and Social Affairs in 2022.

[b]This study focused on individuals aged 19 to 34 years using data from the Youth Life Survey.

[c]Family structure was determined by the people living with each participant.

[d]Labor was determined by whether or not an individual had worked for pay for at least one hour in the past week.

[e]N/A: not available.

[f]Problem drinking is defined as 6 units of alcohol 2 or more times per week in men and 4 units of alcohol 2 or more times per week in women.

[g]Physical activity is defined as a regular physical activity for a week.

XSL•FO
RenderX

Perceived Need for Mental Health Services

As shown in Table 2, hikikomori youth reported a greater perceived need for mental health services compared with their peers (259/397, 63.9% vs 7440/14,569, 50.5%), corresponding to a 27% higher likelihood of perceived need (aPR 1.27, 95% CI 1.15 - 1.41). This association was stronger among individuals with more severe or prolonged withdrawal, with stage 2 participants showing the highest prevalence (91.1%) and risk (aPR 1.89, 95% CI 1.66 - 2.15).

Table . Association between prevalence and perceived need for mental health services according to the hikikomori status, severity, and duration. All estimates were calculated using sampling weights and are presented as weighted prevalence (%) and weighted adjusted prevalence ratio (aPR). The unit of analysis was the individual participant.

Hikikomori status[a]	Participants, n	Prevalent participants, n	Weighted prevalence (%)	Weighted aPR[b]	Weighted 95% CI	*P* value
Model 1[c]						
No	14,569	7440	50.50	1	—[d]	—
Yes	397	259	63.9	1.27	1.15-1.41	<.001
Model 2[e]						
No	14,569	7440	50.50	1	—	—
Stage 1	375	240	62.3	1.24	1.12-1.38	<.001
Stage 2	22	19	91	1.89	1.66-2.15	<.001
Model 3[f]						
No	14,569	7440	50.50	1	—	—
Pre-hikikomori	140	77	50.5	1.01	0.82-1.26	.90
Hiki	257	182	71.0	1.41	1.27-1.57	<.001

[a]Hikikomori status was defined as a result of the Korea Life Survey conducted by the Korea Institute for Health and Social Affairs in 2022.

[b]Adjusted for age, sex, region, living arrangement, household income, labor, problem drinking, smoking, and physical activity. For the prevalence ratio, Poisson regression analysis was conducted.

[c]In model 1, hikikomori was determined by applying the criteria of the Youth Life Survey in 2022. Individuals were classified as hikikomori if they met any of the following criteria: (1) primarily staying at home and only going out for hobbies; (2) mostly staying at home and only going out to nearby convenience stores; (3) staying in their own room but not going outside the house; or (4) rarely leaving their own room.

[d]Not applicable.

[e]In model 2, those identified as hikikomori were further classified into 2 stages according to the severity of hikikomori status. Stage 1 was characterized by individuals who primarily confined themselves to their homes, venturing out only for hobbies or to nearby convenience stores. Individuals in stage 2 demonstrated more severe isolation, as they either remained in their homes or rarely left their rooms.

[f]In model 3, those identified as hikikomori were further classified into 2 groups based on duration of hikikomori status: *pre-hikikomori* for those who had a history of social withdrawal of less than 6 months [9] and *hikikomori* for those who had experienced social withdrawal for 6 months or longer [26].

Unmet Health Care Use

Table 3 demonstrates that hikikomori status was strongly associated with unmet health care use. For physical health, the weighted prevalence was 8.6% among hikikomori youth compared with 2.9% among nonhikikomori (aPR 3.33, 95% CI 2.16 - 5.13). For mental health, the corresponding figures were 10.3% versus 2.6% (aPR 4.46, 95% CI 2.92 - 6.81). The association was particularly pronounced in stage 2 hikikomori, where the prevalence of unmet mental health care use reached 22.1%. Finally, Table 4 shows that the association between hikikomori and unmet physical health care use varied by labor force participation (for *P* interaction=.008).

XSL•FO
RenderX

Table . Association between hikikomori status and prevalence of unmet health care use. All estimates were calculated using sampling weights and are presented as weighted prevalence (%) and weighted adjusted prevalence ratio (aPR). The unit of analysis was the individual participant.

Hikikomori status[a]	Participants, n	Unmet health care use									
		Physical health					Mental health				
		Prevalence, n	Weighted prevalence (%)	Weighted aPR[b]	Weighted 95% CI	*P* value	Prevalence, n	Weighted prevalence (%)	Weighted aPR[b]	Weighted 95% CI	*P* value
Model 1[c]											
No	14,569	427	2.90	1	—[d]	—	383	2.60	1	—	—
Yes	397	32	8.6	3.33	2.16-5.13	<.001	40	10.3	4.46	2.92-6.81	<.001
Model 2[e]											
No	14,569	427	2.90	1	—[d]	—	383	2.60	1	—	—
Stage 1	375	30	8.8	3.37	2.17-5.26	<.001	36	9.6	4.14	2.64-6.49	<.001
Stage 2	22	2	6	2.48	0.49-12.55	.27	4	22	9.52	3.67-24.65	<.001
Model 3[f]											
No	14,569	427	2.90	1	—	—	383	2.60	1	—	—
Pre-hikikomori	140	3	1.8	0.78	0.25-2.47	.67	9	5.7	2.57	1.11-5.96	.03
Hikikomori	257	29	12.2	4.55	2.89-7.15	<.001	31	12.7	5.44	3.44-8.58	<.001

[a]Hikikomori status was defined as a result of the Korea Life Survey conducted by the Korea Institute for Health and Social Affairs in 2022.

[b]Adjusted for age, sex, region, living arrangement, household income, labor, problem drinking, smoking, and physical activity. For the prevalence ratio, Poisson regression analysis was conducted.

[c]In model 1, hikikomori was determined by applying the criteria of the Youth Life Survey in 2022. Individuals were classified as hikikomori if they met any of the following criteria: (1) primarily staying at home and only going out for hobbies; (2) mostly staying at home and only going out to nearby convenience stores; (3) staying in their own room but not going outside the house; or (4) rarely leaving their own room.

[d]Not applicable.

[e]In model 2, those identified as hikikomori were further classified into 2 stages according to the severity of hikikomori status. Stage 1 was characterized by individuals who primarily confined themselves to their homes, venturing out only for hobbies or to nearby convenience stores. Individuals in stage 2 demonstrated more severe isolation, as they either remained in their homes or rarely left their rooms.

[f]In model 3, those identified as hikikomori were further classified into 2 groups based on duration of hikikomori status: “pre-hikikomori” for those who had a history of social withdrawal of less than 6 months [9] and “hikikomori” for those who had experienced social withdrawal for 6 months or longer [26].

XSL•FO
RenderX

Table . Association of hikikomori status[a] prevalence with unmet health care use stratified by each level of covariates.

Variables	Physical health					Mental health				
	Non-hikikomori weighted OR[b] (reference)	Hikikomori weighted OR	95% CI	*P* value	*P* for interaction[c]	Non-hikikomori weighted OR (reference)	Hikikomori weighted OR	95% CI	*P* value	*P* for interaction
Age (y)[d]										
19 - 24	1	3.33	1.64-6.77	.001	.18	1	4.65	1.88-11.53	<.001	.60
25-34	1	3.19	1.74-5.85	.001	—[e]	1	4.82	2.70-8.58	<.001	—
Sex										
Male	1	3.47	1.49-8.07	.004	.74	1	4.74	2.21-10.44	<.001	.71
Female	1	3.55	1.99-6.33	<.001	—	1	5.01	2.73-9.19	<.001	—
Region										
Metropolitan area	1	2.85	1.40-5.82	.004	.42	1	4.50	2.26-8.96	<.001	.56
Non-metropolitan area	1	4.82	2.60-8.93	<.001	—	1	6.04	3.43-10.62	<.001	—
Family structure[f]										
Living alone	1	2.48	0.85-7.20	.10	.37	1	3.26	1.20-8.86	.02	.39
Living with others	1	3.84	2.24-6.58	<.001	—	1	5.52	3.19-9.53	<.001	—
Labor[g]										
No	1	3.97	2.27-6.94	<.001	.008	1	5.43	1.93-15.30	.001	.89
Yes	1	0.61	0.14-2.57	.49	—	1	4.60	2.70-7.81	<.001	—
Problem drinking[h]										
No	1	2.96	1.76-4.97	<.001	.08	1	5.14	3.09-8.56	<.001	.48
Yes	1	18.08	5.60-58.43	<.001	—	1	3.35	0.63-17.75	.16	—
Smoking status										
Yes	1	12.77	3.63-44.99	<.001	.24	1	7.89	2.55-24.40	<.001	.33
No	1	4.47	2.64-7.58	<.001	—	1	3.13	1.84-5.33	<.001	—
Physical activity[i]										
No	1	3.33	1.91-5.81	<.001	.89	1	4.78	1.96-11.69	.001	.87
Yes	1	4.56	1.77-11.79	.002	—	1	5.17	2.90-9.21	<.001	—

[a]Hikikomori status was defined as a result of the Korea Life Survey conducted by the Korea Institute for Health and Social Affairs in 2022.

[b]OR: odds ratio.

[c]The *P* for interaction tests whether the effect of the treatment differs significantly across subgroups. Unlike the subgroup-specific *P* values, which assess the effect within each subgroup, the *P* for interaction evaluates whether the differences in effects between subgroups are statistically significant.

[d]This study focused on individuals aged 19 to 34 years using data from the Youth Life Survey.

[e]Not applicable.

[f]Family structure was determined by the people living with each participant.

[g]Labor was determined by whether or not an individual had worked for pay for at least 1 hour in the past week.

[h]Problem drinking is defined as 6 units of alcohol 2 or more times per week in men and 4 units of alcohol 2 or more times per week in women.

XSL•FO
RenderX

[i]Physical activity is defined as a regular physical activity for a week.

Discussion

Principal Findings

This study found a significant association between hikikomori and a higher likelihood of the perceived need for mental health services compared to those without hikikomori. This association was more prominent in individuals with severe hikikomori and a longer duration of hikikomori. In addition, this study found a significant association between hikikomori and a higher prevalence of unmet health care use in both physical and mental health. This association became more pronounced as the severity and duration of hikikomori increased. Moreover, the prevalence of unmet health care use in mental health was higher than in physical health.

In this study, individuals with hikikomori were more likely to perceive a need for mental health care than those without hikikomori, meaning they recognized their needs independently. This tendency also increased with greater hikikomori severity and duration. The perceived need for mental health care is known to influence the use of services [33], and a previous study reported that individuals with mental health conditions who do not believe they require assistance are less likely to seek treatment [34]. However, despite recognizing their needs, hikikomori individuals reported substantially higher levels of unmet health care use, as seen in Table 3. Because of their unique behavioral patterns, individuals with hikikomori may face significant barriers to accessing health care services, even when they recognize their deteriorating health. Future strategies should include home-visiting health services or remote health care services to address the unmet health care needs of hikikomori and facilitate easier access to care.

Our study is significant because, in contrast to previous research, which primarily focused on comparing the mental health status of individuals with hikikomori [35-37], the results of this study highlight the challenges faced by individuals with hikikomori in accessing health care use compared with those of individuals without hikikomori. We investigated the association between hikikomori status and health care use by stratifying individuals based on hikikomori severity, determined by levels of isolation derived from staying at home. Our findings revealed a gradual increase in the risk of unmet health care use with an increase in hikikomori severity. Individuals with the most severe form of hikikomori experienced a substantial increase in the risk of unmet health care use in mental health, although the number of participants in this specific subgroup was limited. This finding can be interpreted as follows: compared to mild hikikomori, who sometimes go out for hobbies or to the market, severe hikikomori mostly stay in their homes (or rooms), preventing them from using health care services. Furthermore, unmet health care use in mental health was more prevalent among individuals with the most severe hikikomori than in physical health. Based on these results, the more severe the stage of hikikomori, the more urgent the need to address their unmet health care use.

The results of this study also revealed a gradual increase in the risk of unmet health care use in both physical and mental health with increasing duration of hikikomori status. This result is similar to that of a study examining treatment preferences in a sample of individuals with hikikomori who expressed a desire to seek treatment for their condition compared to pre-hikikomori individuals (people who withdrew for less than 6 months) [7]. Individuals with long-term social withdrawal tend to have restricted social activities and behavioral problems [38]. In this context, the unmet health care use of individuals with hikikomori for long-term periods must be urgently addressed to provide them with appropriate health services. The subgroup analysis showed that the impact of hikikomori on unmet physical health care use was especially pronounced among those outside the labor force, suggesting that economic inactivity is an important barrier to accessing physical health services. Prior studies have similarly reported greater unmet needs among unemployed or economically inactive individuals [39], consistent with our findings in socially withdrawn youth.

In this study, individuals with hikikomori, severe hikikomori, and longer durations of hikikomori were more likely to have unmet mental and physical health use. Individuals with hikikomori who choose not to seek help experience more negative consequences [15,35]. On the basis of these findings, individuals with hikikomori should be treated with adequate interventions to prevent the worsening of health conditions. Considering their behavioral patterns, future interventions for hikikomori may benefit from home-visit programs led by health professionals. This approach could be particularly effective in improving the management systems for individuals with hikikomori [7,10]. In addition, the findings of this study showed that the likelihood of unmet health care use in hikikomori was greater than in nonhikikomori, and this trend was stronger when hikikomori severity and duration worsened. This suggests that health care providers should develop and implement patient-centered programs based on individual hikikomori status. Given the significant trend observed in mental health compared to physical health, a deeper consideration of health assessments and interventions specifically tailored for individuals with hikikomori is warranted.

Although this study did not aim to identify causal mechanisms, previous research has suggested several potential contributors to hikikomori, including individual vulnerabilities, family dynamics, and broader social factors such as academic pressure, insecure employment, and housing instability [4,8,16]. In addition, stigma surrounding mental illness and social withdrawal may discourage individuals from seeking help, reinforcing isolation. Structural barriers, including shortages of mental health professionals and limited community-based services, further restrict access [40]. Addressing these challenges will require not only expanding conventional services but also integrating innovative approaches such as remote counseling and health care [23,41]. Strengthening the workforce, reducing regional disparities, and evaluating the effectiveness of digital strategies will be essential to better meet the needs of socially withdrawn youth.

XSL•FO
RenderX

Strengths and Limitations

Our study has several limitations. First, the cross-sectional design used in this study meant that the data were collected at a single time point. This limits our ability to establish clear causal associations, which is why the results indicate associations rather than definitive causal relationships. Reverse causality is a common drawback of cross-sectional studies. However, it can be argued that the prevalence of hikikomori is more likely to be a causal factor influencing the unmet health care use status of individuals with hikikomori, rather than vice versa. Second, the survey used in this study did not consider participants' motivation to seek health care. Therefore, we could only examine the prevalence of hikikomori and unmet health care use. However, we adjusted for socioeconomic characteristics, such as income, labor, region, and family structure, so that significant factors affecting health care use could be considered. Third, this study relied on self-reported survey responses, which may be subject to recall bias. While the young age of participants may mitigate this issue, the possibility of such bias cannot be entirely excluded and should be considered when interpreting the results. Fourth, unmet health care use was defined as situations in which participants reported needing but not receiving medical care. The survey did not capture whether individuals actively sought help, including alternative modes such as online consultations. Finally, the data were collected from one country. Cross-cultural and methodological differences can introduce bias into the assessment of unmet health care needs, complicating cross-country comparisons.

Despite its limitations, this study has several strengths. To our knowledge, it is the first quantitative analysis to link hikikomori status with unmet health care use. The data, representing the entire Korean youth population, enable extrapolation of findings to a broader context. Reliability was enhanced by collecting self-reported data through in-person interviews. The findings can inform policy discussions and support the integration of interventions for withdrawal behaviors and medical treatment.

Conclusions

This study suggests that individuals with hikikomori face challenges in accessing and seeking mental health care services. It demonstrated a remarkable association between hikikomori status and unmet health care use, emphasizing the profound impact of hikikomori status on the prevalence of unmet health care use across varying levels of severity and duration. It may be beneficial for providers and policymakers to explore the potential of home-visit services or technology-assisted services, such as telehealth, as supplementary measures to better understand unmet health care use and provide more accessible care options.

Acknowledgments

The authors extend their gratitude to the Korea Institute for Health and Social Affairs for granting access to the nationwide survey data used in this study. This research was conducted without any external funding or financial support and relied solely on the resources and facilities available to the authors.

Data Availability

The data used in this study were obtained from the 2022 Korea Institute for Health and Social Affairs and can be downloaded from the official Microdata Integrated Service website [42].

Conflicts of Interest

None declared.

References

1. Wong PWC, Li TMH. Hikikomori: risks and consequences of extreme self-imposed social marginalization. In: Coplan RJ, Bowker JC, Nelson LJ, editors. The Handbook of Solitude: Psychological Perspectives on Social Isolation, Social Withdrawal, and Being Alone: John Wiley & Sons; 2021:378-389 URL: https://onlinelibrary.wiley.com/doi/10.1002/9781119576457.ch27 [accessed 2025-09-29] [doi: 10.1002/9781119576457]
2. Teo AR, Stufflebam KW, Kato TA. The intersection of culture and solitude: the hikikomori phenomenon in Japan. In: Bowker JC, Coplan RJ, editors. The Handbook of Solitude: Psychological Perspectives on Social Isolation, Social Withdrawal, and Being Alone 2013:445-460. [doi: 10.1002/9781118427378]
3. Teo AR, Gaw AC. Hikikomori, a Japanese culture-bound syndrome of social withdrawal? A proposal for DSM-5. J Nerv Ment Dis 2010 Jun;198(6):444-449. [doi: 10.1097/NMD.0b013e3181e086b1] [Medline: 20531124]
4. Kato TA, Tateno M, Shinfuku N, et al. Does the "hikikomori" syndrome of social withdrawal exist outside Japan? A preliminary international investigation. Soc Psychiatry Psychiatr Epidemiol 2012 Jul;47(7):1061-1075. [doi: 10.1007/s00127-011-0411-7] [Medline: 21706238]
5. Bowker JC, Bowker MH, Santo JB, Ojo AA, Etkin RG, Raja R. Severe social withdrawal: cultural variation in past hikikomori experiences of university students in Nigeria, Singapore, and the United States. J Genet Psychol 2019;180(4-5):217-230. [doi: 10.1080/00221325.2019.1633618] [Medline: 31305235]

XSL•FO
RenderX

6. Pozza A, Coluccia A, Kato T, Gaetani M, Ferretti F. The "hikikomori" syndrome: worldwide prevalence and co-occurring major psychiatric disorders: a systematic review and meta-analysis protocol. BMJ Open 2019 Sep 20;9(9):e025213. [doi: 10.1136/bmjopen-2018-025213] [Medline: 31542731]
7. Teo AR, Fetters MD, Stufflebam K, et al. Identification of the hikikomori syndrome of social withdrawal: psychosocial features and treatment preferences in four countries. Int J Soc Psychiatry 2015 Feb;61(1):64-72. [doi: 10.1177/0020764014535758] [Medline: 24869848]
8. Kato TA, Kanba S, Teo AR. Hikikomori: multidimensional understanding, assessment, and future international perspectives. Psychiatry Clin Neurosci 2019 Aug;73(8):427-440. [doi: 10.1111/pcn.12895] [Medline: 31148350]
9. Kato TA, Kanba S, Teo AR. Defining pathological social withdrawal: proposed diagnostic criteria for hikikomori. World Psychiatry 2020 Feb;19(1):116-117. [doi: 10.1002/wps.20705] [Medline: 31922682]
10. Lee YS, Lee JY, Choi TY, Choi JT. Home visitation program for detecting, evaluating and treating socially withdrawn youth in Korea. Psychiatry Clin Neurosci 2013 May;67(4):193-202. [doi: 10.1111/pcn.12043] [Medline: 23683149]
11. Chan HY, Lo TW. Quality of life of the hidden youth in Hong Kong. Applied Research Quality Life 2014 Dec;9(4):951-969. [doi: 10.1007/s11482-013-9279-x]
12. Kanai K, Kitamura Y, Zha L, Tanaka K, Ikeda M, Sobue T. Prevalence of and factors influencing hikikomori in Osaka city, Japan: a population-based cross-sectional study. Int J Soc Psychiatry 2024 Aug;70(5):967-980. [doi: 10.1177/00207640241245926] [Medline: 38616515]
13. Se-jeong J, Jin-ah R, Ye-eun K, et al. 2022 youth life survey [Article in Korean]. Research@KIHASA. 2022. URL: https://repository.kihasa.re.kr/handle/201002/42426 [accessed 2025-09-29]
14. Kondo N, Sakai M, Kuroda Y, Kiyota Y, Kitabata Y, Kurosawa M. General condition of hikikomori (prolonged social withdrawal) in Japan: psychiatric diagnosis and outcome in mental health welfare centres. Int J Soc Psychiatry 2013 Feb;59(1):79-86. [doi: 10.1177/0020764011423611] [Medline: 22094722]
15. Kato TA, Sartorius N, Shinfuku N. Shifting the paradigm of social withdrawal: a new era of coexisting pathological and non-pathological hikikomori. Curr Opin Psychiatry 2024 May 1;37(3):177-184. [doi: 10.1097/YCO.0000000000000929] [Medline: 38415743]
16. Yong R, Nomura K. Hikikomori is most associated with interpersonal relationships, followed by suicide risks: a secondary analysis of a national cross-sectional study. Front Psychiatry 2019;10:247. [doi: 10.3389/fpsyt.2019.00247] [Medline: 31040799]
17. Cené CW, Beckie TM, Sims M, et al. Effects of objective and perceived social isolation on cardiovascular and brain health: a scientific statement from the American Heart Association. J Am Heart Assoc 2022 Aug 16;11(16):e026493. [doi: 10.1161/JAHA.122.026493] [Medline: 35924775]
18. Kobayashi LC, Steptoe A. Social isolation, loneliness, and health behaviors at older ages: longitudinal cohort study. Ann Behav Med 2018 May 31;52(7):582-593. [doi: 10.1093/abm/kax033] [Medline: 29860361]
19. Connolly S, Wren MA. Unmet healthcare needs in Ireland: analysis using the EU-SILC survey. Health Policy 2017 Apr;121(4):434-441. [doi: 10.1016/j.healthpol.2017.02.009] [Medline: 28233599]
20. Hargreaves DS, Elliott MN, Viner RM, Richmond TK, Schuster MA. Unmet health care need in US adolescents and adult health outcomes. Pediatrics 2015 Sep;136(3):513-520. [doi: 10.1542/peds.2015-0237] [Medline: 26283778]
21. Thomson S, Cylus J. Can People Afford to Pay for Health Care? New Evidence on Financial Protection in Europe: World Health Organization; 2019:41-46 URL: https://iris.who.int/handle/10665/311654 [accessed 2025-09-12]
22. Health at a glance 2023. : Organization for Economic Co-operation and Development; 2023 URL: https://doi.org/10.1787/7a7afb35-en [accessed 2025-09-12] [doi: 10.1787/7a7afb35-en]
23. Kourgiantakis T, Markoulakis R, Lee E, et al. Access to mental health and addiction services for youth and their families in Ontario: perspectives of parents, youth, and service providers. Int J Ment Health Syst 2023 Mar 14;17(1):4. [doi: 10.1186/s13033-023-00572-z] [Medline: 36918893]
24. Kourgiantakis T, Markoulakis R, Hussain A, et al. Navigating inequities in the delivery of youth mental health care during the COVID-19 pandemic: perspectives of youth, families, and service providers. Can J Public Health 2022 Dec;113(6):806-816. [doi: 10.17269/s41997-022-00670-4] [Medline: 35852728]
25. Craig SL, Eaton AD, Leung VW, et al. Efficacy of affirmative cognitive behavioural group therapy for sexual and gender minority adolescents and young adults in community settings in Ontario, Canada. BMC Psychol 2021 Jun 7;9(1):94. [doi: 10.1186/s40359-021-00595-6] [Medline: 34099063]
26. Guidelines for the assessment and support of hikikomori [Article in Japanese]. Ministry of Health, Labor and Welfare Japan. 2010. URL: https://www.mhlw.go.jp/stf/houdou/2r98520000006i6f.html [accessed 2025-09-12]
27. Chiu M, Amartey A, Wang X, Vigod S, Kurdyak P. Trends in objectively measured and perceived mental health and use of mental health services: a population-based study in Ontario, 2002-2014. CMAJ 2020 Mar 30;192(13):E329-E337. [doi: 10.1503/cmaj.190603] [Medline: 32392484]
28. Lamoureux-Lamarche C, Berbiche D, Vasiliadis HM. Perceived need for mental health care and associated factors and outcomes in older adults consulting in primary care. Can J Psychiatry 2022 Jul;67(7):553-564. [doi: 10.1177/07067437211055430] [Medline: 34825588]

XSL•FO
RenderX

29. Unmet needs for health care: comparing approaches and results from international surveys. Organization for Economic Co-operation and Development. 2020. URL: https://www.oecd.org/en/publications/unmet-needs-for-health-care_9c2be79c-en.html [accessed 2025-09-12]
30. Bataineh H, Devlin RA, Barham V. Unmet health care and health care utilization. Health Econ 2019 Apr;28(4):529-542. [doi: 10.1002/hec.3862] [Medline: 30693596]
31. Lindström C, Rosvall M, Lindström M. Unmet health-care needs and mortality: a prospective cohort study from southern Sweden. Scand J Public Health 2020 May;48(3):267-274. [doi: 10.1177/1403494819863530] [Medline: 31405329]
32. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. Am J Epidemiol 2005 Aug 1;162(3):199-200. [doi: 10.1093/aje/kwi188] [Medline: 15987728]
33. Bonabi H, Müller M, Ajdacic-Gross V, et al. Mental health literacy, attitudes to help seeking, and perceived need as predictors of mental health service use: a longitudinal study. J Nerv Ment Dis 2016 Apr;204(4):321-324. [doi: 10.1097/NMD.0000000000000488] [Medline: 27015396]
34. Mojtabai R, Evans-Lacko S, Schomerus G, Thornicroft G. Attitudes toward mental health help seeking as predictors of future help-seeking behavior and use of mental health treatments. Psychiatr Serv 2016 Jun 1;67(6):650-657. [doi: 10.1176/appi.ps.201500164] [Medline: 26876662]
35. Fong TC, Yip PS. Prevalence of hikikomori and associations with suicidal ideation, suicide stigma, and help-seeking among 2,022 young adults in Hong Kong. Int J Soc Psychiatry 2023 Nov;69(7):1768-1780. [doi: 10.1177/00207640231174376] [Medline: 37191282]
36. Orsolini L, Bellagamba S, Volpe U, Kato TA. Hikikomori and modern-type depression in Italy: a new phenotypical trans-cultural characterization? Int J Soc Psychiatry 2022 Aug;68(5):1010-1017. [doi: 10.1177/00207640221099408] [Medline: 35723190]
37. Teo AR, Nelson S, Strange W, et al. Social withdrawal in major depressive disorder: a case-control study of hikikomori in Japan. J Affect Disord 2020 Sep 1;274(274):1142-1146. [doi: 10.1016/j.jad.2020.06.011] [Medline: 32663943]
38. Nonaka S, Shimada H, Sakai M. Assessing adaptive behaviors of individuals with hikikomori (prolonged social withdrawal): development and psychometric evaluation of the parent-report scale. Int J Cult Ment Health 2018 Jul 3;11(3):280-294. [doi: 10.1080/17542863.2017.1367411]
39. Åhs A, Burell G, Westerling R. Care or not care—that is the question: predictors of healthcare utilisation in relation to employment status. Int J Behav Med 2012 Mar;19(1):29-38. [doi: 10.1007/s12529-010-9129-2] [Medline: 21128042]
40. Kim E, Lee HA, Lee YR, et al. A new agenda for optimizing roles and infrastructure in a mental health service model for South Korea. Psychiatry Investig 2025 Jan;22(1):26-39. [doi: 10.30773/pi.2024.0189] [Medline: 39885789]
41. Philippe TJ, Sikder N, Jackson A, et al. Digital health interventions for delivery of mental health care: systematic and comprehensive meta-review. JMIR Ment Health 2022 May 12;9(5):e35159. [doi: 10.2196/35159] [Medline: 35551058]
42. Microdata Integrated Service [Website in Korean]. URL: https://mdis.kostat.go.kr/index.do [accessed 2025-09-29]

Abbreviations

aPR: adjusted prevalence ratio
KIHASA: Korea Institute for Health and Social Affairs

Edited by A Mavragani, T Sanchez; submitted 14.06.25; peer-reviewed by G Mayer, MI ul Husnain; revised version received 28.08.25; accepted 29.08.25; published 02.10.25.

Please cite as:
Choi E, Choi S, Jang S
Unmet Health Care Use Among Socially Withdrawn Youth (Hikikomori) in South Korea: Cross-Sectional Survey Study
JMIR Public Health Surveill 2025;11:e79043
URL: https://publichealth.jmir.org/2025/1/e79043
doi:10.2196/79043

© Eunjeong Choi, Seoyeong Choi, Sukyong Jang. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 2.10.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Vaccine Hesitancy and Associated Factors Among Caregivers of Children With Special Health Care Needs in the COVID-19 Era in China: Cross-Sectional Study

Mingyan Li[1*], MD; Changxuan Sun[2*], MSc; Chai Ji[1], MD; Meiying Gao[1], MSc; Xia Wang[1], MSc; Dan Yao[1], MD; Junxia Guo[1], MSc; Lidan Sun[1], MA; Abdul Rafay[3], HS; Antonita Shereen George[4], MBBS; Sanduni Hasara Samararathna Samararathna Muhandiramge[4], MBBS; Guannan Bai[1], PhD

[1]Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China

[2]The Fourth Affiliated Hospital of Soochow University, Medical Center of Soochow University, Suzhou Dushu Lake Hospital, Suzhou, China

[3]School of Medicine, Imperial College London, London, United Kingdom

[4]First Affiliated Hospital of Zhejiang University, Zhejiang University School of Medicine, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Guannan Bai, PhD
Department of Child Health Care, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou, China

Abstract

Background: Immunization is a cost-effective way to prevent infectious diseases in children, but parental hesitancy leads to low vaccination rates, leaving children at risk. Caregivers of children with special health care needs are more hesitant about vaccines than those of healthy children.

Objective: The aim of the study is to investigate the changes in caregivers' vaccination hesitation of children with special health care needs before, during, and after the COVID-19 pandemic in China and to identify associated factors for caregivers' attitudes toward National Immunization Program (NIP) and non-NIP vaccines.

Methods: We included 7770 caregivers of children with special health care needs (median age 7.0, IQR 2.4-24.1 months) who visited the Vaccination Consultation Clinic at Children's Hospital, Zhejiang University School of Medicine (Hangzhou, China) from May 2017 to May 2023. General and clinical information was extracted from the immunization evaluation system for children with special health care needs and medical records. We compared the differences in caregivers' willingness and hesitation for vaccinating their children across the 3 stages of the COVID-19 pandemic using chi-square tests. Multinomial logistic regression models were used to identify independent variables that were associated with caregivers' willingness and hesitation toward NIP and non-NIP vaccines.

Results: There is a statistically significant difference in caregivers' vaccine hesitancy before, during, and after the COVID-19 pandemic (P<.05). During the COVID-19 pandemic, the percentages of choosing NIP, alternative non-NIP, and non-NIP vaccines are highest (n=1428, 26%, n=3148, 57.4%, and n=3442, 62.7%, respectively) than those at other 2 stages. In comparison, caregivers' hesitation toward NIP and non-NIP vaccines is lowest (n=911, 16.6% and n=2045, 37.3%, respectively). Despite the stages of the COVID-19 pandemic, multiple factors, including children's age and sex, parents' educational level, comorbidities, and history of allergy, were significantly associated with caregivers' attitude toward NIP and non-NIP vaccines (P<.05). The profiles of risk factors for hesitancy toward NIP and non-NIP vaccines are different, as indicated by the results from the logistic regression models.

Conclusions: This study demonstrated that caregivers' willingness to vaccinate their children with special health care needs with NIP and non-NIP vaccines was highest during the COVID-19 pandemic in China, and their hesitancy was lowest. Additionally, we have identified multiple factors associated with caregivers' willingness and hesitancy to vaccinate their children. These findings provide evidence-based support for developing personalized health education strategies.

(JMIR Public Health Surveill 2025;11:e67487) doi:10.2196/67487

KEYWORDS

COVID-19; caregivers; children with special health care needs; vaccination hesitancy; decision-making

XSL•FO
RenderX

Introduction

Immunization is widely recognized as an exceptionally cost-effective strategy for preventing and eliminating infectious diseases in children. In China, a dual-tier vaccination program is implemented. The government provides vaccines under the National Immunization Program (NIP) at no cost across the nation, whereas non-NIP vaccines incur some costs [1]. It is mandatory for all children to receive NIP vaccines according to governmental guidelines. The NIP vaccines in China, which include 14 vaccines targeting 15 diseases, have achieved a coverage rate exceeding 95% among children younger than 6 years of age at the national level [2,3]. In contrast, non-NIP vaccines, which complement the NIP vaccines and are crucial for comprehensive disease prevention and control, exhibit relatively low coverage rates. These rates vary significantly across different regions and socioeconomic groups primarily because non-NIP vaccines are voluntary, self-funded, and not required for school enrollment. Zhou et al [4] observed that the coverage of non-NIP vaccines was lowest among left-behind families and highest among local urban families. Generally, the coverage rates for most non-NIP vaccines are below 50%, with specific rates ranging from 1.8% for the third dose of the rotavirus vaccine to 67.1% for the first dose of the varicella vaccine [5].

Numerous factors can affect vaccine coverage rates. Studies demonstrate a direct correlation between caregivers' attitudes toward vaccines and vaccination rates [6]. A significant number of caregivers express hesitancy regarding vaccinating their children often due to concerns about safety, perceived inconvenience, financial costs [7], and a lack of trust in scientific evidence [8]. This hesitancy results in insufficient vaccination coverage and leaves children vulnerable to vaccine-preventable diseases [9]. Notably, caregivers of children with special health care needs are more hesitant about vaccines than those of healthy children [10]. Children with special health care needs refers to a population of children who either have or are at an elevated risk for chronic physical, developmental, behavioral, or emotional conditions. In China, about 2.12 million newborns face birth defects or immune issues each year [11]. Recently, there has been a focus on vaccinating children with special health care needs, with researchers studying their vaccination rates [12,13] and factors associated with their vaccine hesitancy [14-16]. Expert groups from countries such as the United States, Australia [17], and China [18] have developed guidelines and consensus to improve vaccination rates for children with special health care needs.

The global COVID-19 pandemic has heightened vaccine awareness [19]. China approved emergency use of inactivated COVID-19 vaccines for children aged 3 to 17 years in July 2021 [20]. Media focus on vaccination during the pandemic has shifted public and caregivers' attitudes toward vaccines [21,22]. A survey across 6 countries showed that the pandemic has encouraged families who previously avoided vaccines to consider influenza vaccination [23]. A meta-analysis showed an increased willingness among caregivers to vaccinate their children against seasonal flu during the COVID-19 pandemic [24]. Caregivers of children with special health care needs may be more anxious than the general population [25]. The heightened public awareness of vaccines may lead to potential changes in vaccination attitudes among caregivers of children with special health care needs. However, there is a notable lack of research addressing vaccine hesitancy within this specific population.

In 2016, the Health Commission of Zhejiang province, China, set up the Vaccination Consultation Clinic at the Children's Hospital, Zhejiang University School of Medicine [26]. Specialists there have been advising on vaccinations for children with special health care needs. Therefore, we conducted this retrospective study to investigate the attitudes toward vaccines among caregivers of children with special health care needs and evaluate the changes in caregivers' hesitancy toward vaccines before, during, and after the COVID-19 pandemic. In addition, we also aimed to identify associated factors, including the stages of the COVID-19 pandemic, for caregivers' attitudes toward NIP and non-NIP vaccines.

Methods

Study Design and Data Collection

Since 2016, the children's hospital where this study was conducted has established the sole vaccination consultation clinic dedicated to providing assessments and tailored vaccination recommendations for children with special health care needs. The majority of these children were referred by physicians from vaccination sites throughout the province. This hospital is a grade A tertiary children's hospital located in Zhejiang province, primarily serving pediatric patients within the province. In 2019, it was designated as a National Clinical Research Center for Child Health, thereby broadening its patient base. Based on this clinic, we conducted a retrospective analysis of caregivers' vaccination willingness monitoring data from May 18, 2017, to May 31, 2023. Data collection used paper questionnaires until March 31, 2021, followed by an electronic monitoring system from April 1, 2021, onward. Trained research assistants transferred data collected from paper questionnaires into an electronic database managed by EpiData software (The EpiData Association). A double data entry process was implemented to ensure consistency and accuracy.

Ethical Considerations

This study was conducted according to the Declaration of Helsinki [27] and was approved by the medical ethics committee of the Children's Hospital Zhejiang University School of Medicine (2025-IRB-0005-P-01). This study involved secondary analysis of existing anonymized data that posed no privacy risks to individuals. The need for informed consent was waived by the ethics committee of Children's Hospital, Zhejiang University School of Medicine. No financial compensation was provided as the study did not involve direct participant interaction.

Study Population

Data were collected from children with special health care needs who visited the Vaccination Consultation Clinic at Children's Hospital, Zhejiang University School of Medicine between May 18, 2017, and May 31, 2023. For the purposes of this study, children with special health care needs are defined as children

XSL•FO
RenderX

with a medical condition that has persisted or is anticipated to persist for 12 months or longer, meeting at least one of the following criteria: (1) a persistent need for prescribed medications, (2) a persistent need for medical care that is higher than the average for children at the same age, (3) a persistent need for special treatments, (4) limitations on activities that are available for the majority of children at the same age, and (5) the presence of a persistent behavioral or developmental condition that requires treatment or counseling [26,28]. In total, 7770 children with special health care needs were included in this study.

Measurements

Information on demographics, family history of diseases, history of adverse events following immunization, parental education levels, parental attitudes toward vaccination, and reasons for being referred to our Vaccination Consultation Clinic were extracted from the immunization evaluation system for children with special health care needs. Concurrently, medical records, such as the specific diagnosis of health conditions, were obtained from the hospital's electronic medical records system.

We used 2 questions to measure caregivers' hesitation toward NIP and non-NIP vaccines. The first question is "What is your attitude toward NIP vaccines?" There are three options: (1) I would like to choose NIP vaccines for my child; (2) if there are alternatives to the NIP vaccine (ie, alternative non-NIP vaccines), I am willing to pay and choose the alternative non-NIP vaccines for my child; and (3) I am not sure and hesitating to choose the NIP vaccines. If caregivers chose the option (3), we defined it as "hesitation toward NIP vaccines." The second question is "What is your attitude toward the non-NIP vaccines?" There are two options: (1) I would like to choose non-NIP vaccines for my child and (2) I am not sure and hesitating. If caregivers chose the option (2), we defined it as "hesitation toward non-NIP vaccines."

Based on the timeline of the COVID-19 pandemic in Zhejiang province, China, we defined stage I, II, and III. More specifically, stage I was "before the COVID-19 pandemic," referred the time period from May 18, 2017, when the first questionnaire was filled to January 22, 2020; stage II was "during the COVID-19 pandemic," referred from January 23, 2020, to December 18, 2022, when all the COVID-19 measures were eased in Zhejiang province; and stage III as from December 19, 2022, to May 31, 2023.

Statistical Analysis

First, we conducted a descriptive analysis to present the general and clinical characteristics of the study population. Normality was tested for the continuous variable, that is, children's age in our study. The distribution of age was not normal, so we calculated the median and IQR. Categorical variables were presented as numbers and percentages. In addition, we compared the difference in the general characteristics of the study population across 3 stages of the COVID-19 pandemic. Regarding the nonnormally distributed variable, we used the Kruskal-Wallis rank test, while for categorical variables, chi-square tests were applied. Furthermore, we described the spectrum of disease for all children enrolled in this study across the 3 phases of the COVID-19 pandemic. In addition, we described the disease spectrum of children with special health care needs in the overall population and the subgroup of each stage of the COVID-19 pandemic. Second, we described the percentages of caregivers' choices for vaccines at 3 stages of the COVID-19 pandemic and compared the differences in the percentages of choices across 3 stages using chi-square tests. Finally, multinomial logistic regression models were used to identify independent variables that were associated with caregivers' hesitation toward NIP and non-NIP vaccines. In addition, we conducted separate logistic regression analyses for different stages of the COVID-19 pandemic.

All statistical analyses were conducted using the R software (version 4.3.1; R Foundation for Statistical Computing). The statistical significance was indicated when $P<.05$ (2 tails).

Results

General and Clinical Characteristics of the Study Population

Table 1 presents the general characteristics of the study population. In total, 4418 (56.9%) children included in the analysis were male, and 3534 (45.5%) were aged less than 6 months. A total of 6051 (77.9%) questionnaires were filled in by mothers. Mothers of 3061 (39.4%) children had an educational level of college or above, while fathers of 3004 (38.7%) children had an educational level of college or above. There were statistically significant differences in children's age, caregiver who filled in the questionnaire, and maternal and paternal educational level across the 3 COVID-19 pandemic stages (all $P<.001$).

XSL•FO
RenderX

Table . General characteristics of children and their caregivers (N=7770).

Items	Overall	Stage I[a]	Stage II[b]	Stage III[c]	*P* value
Child's sex, n (%)					.50
Male	4418 (56.9)	1001 (56.8)	3108 (56.6)	309 (59.3)	
Female	3352 (43.1)	761 (43.2)	2379 (43.4)	212 (40.7)	
Age of the child (months), median (IQR)	7.0 (2.4-24.1)	6.4 (2.3-19.6)	7.6 (2.5-25.7)	5.3 (2.0-18.2)	<.001
Age group of children (months), n (%)					<.001
0 - 6	3534 (45.5)	845 (48)	2416 (44)	273 (52.4)	
7 - 12	1228 (15.8)	306 (17.4)	840 (15.3)	82 (16.7)	
13 - 24	1053 (13.6)	235 (13.3)	755 (13.8)	63 (12.1)	
25 - 72	1376 (17.7)	305 (17.3)	1001 (18.2)	70 (13.3)	
≥72	579 (7.4)	71 (4)	475 (8.7)	33 (6.3)	
Caregiver who filled in the questionnaire, n (%)					<.001
Mother	6051 (77.9)	1484 (84.2)	4208 (76.7)	359 (68.9)	
Father	1419 (18.3)	242 (13.7)	1049 (19.1)	128 (24.6)	
Others	281 (3.6)	33 (1.9)	214 (3.9)	34 (6.5)	
Missing	19 (0.2)	3 (0.2)	16 (0.3)	0 (0)	
Maternal educational level, n (%)					<.001
Middle school or below	1253 (16.1)	316 (17.9)	883 (16.1)	54 (10.4)	
High school or equivalent	1233 (15.9)	317 (18)	855 (15.6)	61 (11.7)	
Two-year college or associate degree	1948 (25.1)	463 (26.3)	1378 (25.1)	107 (20.5)	
Bachelor degree or above	3061 (39.4)	618 (35.1)	2209 (40.3)	234 (44.9)	
Missing	275 (3.5)	48 (3)	162 (3)	65 (12.5)	
Paternal educational level, n (%)					<.001
Middle school or below	1205 (15.5)	302 (17.1)	850 (15.5)	53 (10.2)	
High school or equivalent	1401 (18)	358 (20.3)	978 (17.8)	65 (12.5)	
Two-year college or associate degree	1796 (23.1)	435 (24.7)	1244 (22.7)	117 (22.5)	
Bachelor degree or above	3004 (38.7)	611 (34.7)	2170 (39.6)	223 (42.8)	
Missing	364 (4.7)	56 (3.2)	245 (4.5)	63 (12.1)	

[a]Stage Ⅰ was defined as the time period from May 18, 2017, to January 22, 2020.

[b]Stage II was defined as the time period from January 23, 2020, to December 18, 2022.

[c]Stage III was defined as the time period from December 19, 2022, to May 31, 2023.

Table 2 demonstrates the clinical characteristics of the study population. A total of 2729 (35.1%) children had comorbidity, 4739 (61%) children had a history of allergies, and 330 (4.2%) children had a history of adverse events following immunization. Regarding the reasons to visit the vaccination consultation clinic, 5245 (67.5%) caregivers reported that they were referred by doctors at the vaccination site, and 2198 (28.3%) caregivers came to consult because they did not know whether children with special health care needs can receive vaccines. In total, 336 (4.3%) caregivers came to consult because they did not know whether certain medications applied to the child may affect vaccination, and 130 (1.7%) caregivers due to adverse events in previous vaccinations. In addition, 288 (3.7%) caregivers were referred by the specialists because they were

XSL•FO
RenderX

not sure about the vaccination of children with special health care needs, and 376 (4.8%) caregivers came to consult because their children were mandatory to be vaccinated in order to go to school.

Table . Clinical characteristics of the study population (N=7770).

Items	Overall, n (%)	Stage I[a], n (%)	Stage II[b], n (%)	Stage III[c], n (%)
Having comorbidity				
Yes	2729 (35.1)	592 (66.4)	1915 (34.9)	222 (42.6)
No	5041 (64.9)	1170 (33.6)	3572 (65.1)	299 (57.4)
History of allergy				
Yes	4739 (61)	446 (25.3)	3772 (68.7)	521 (100)
No	3026 (38.9)	1316 (74.7)	1710 (31.2)	0 (0)
Missing	5 (0.1)	0 (0)	5 (0.1)	0 (0)
History of AEFI [d]				
Yes	330 (4.2)	81 (5)	227 (4.1)	22 (4.2)
No	7358 (94.7)	1681 (95.4)	5178 (94.4)	499 (95.8)
Missing	82 (1.1)	0 (0)	82 (1.5)	0 (0)
Reasons for visiting the vaccination consultation clinic				
Referred by doctors at the vaccination site	5245 (67.5)	1407 (79.8)	3513 (64)	325 (62.4)
Caregivers came to consult because they did not know whether children with special health care needs can receive vaccines	2198 (28.3)	289 (16.4)	1716 (31.3)	193 (37)
Caregivers came to consult because they did not know whether certain medications applied to the child may affect vaccination	336 (4.3)	8 (0.4)	275 (5)	53 (10.2)
Caregivers came to consult because of adverse events in previous vaccinations	130 (1.7)	5 (0.3)	106 (1.9)	19 (3.6)
Referred by the specialist because they were not sure about vaccination of children with special health care needs	288 (3.7)	15 (0.8)	264 (4.8)	9 (1.7)
Required by school entry	376 (4.8)	13 (0.7)	300 (5.5)	63 (12.1)
Missing	86 (1.1)	25 (1.4)	58 (1.1)	3 (0.6)

[a]Stage I was defined as the time period from May 18, 2017, to January 22, 2020.

[b]Stage II was defined as the time period from January 23, 2020, to December 18, 2022.

[c]Stage III was defined as the time period from December 19, 2022, to May 31, 2023.

[d]AEFI: adverse events following immunization.

Disease Spectrum of the Study Population

Multimedia Appendix 1 shows the spectrum of disease in the overall study population and at each stage of the COVID-19 pandemic. The 3 most common diseases of those children with special health care needs were circulatory system diseases, nervous system diseases, and neonatal diseases in the overall population, and this pattern remains the same at each stage.

Caregivers' Attitude Toward NIP Vaccines and Non-NIP Vaccines

Figure 1 shows caregivers' choice for NIP vaccines or alternative non-NIP vaccines and their hesitancy at each stage of the COVID-19 pandemic. Figure 2 shows caregivers' choice of non-NIP vaccines and their hesitation at 3 stages of the COVID-19 pandemic. During the COVID-19 pandemic (ie, stage II), the percentages of choosing alternative non-NIP or non-NIP vaccines are highest (n=3148, 57.4% and n=3442,

XSL•FO
RenderX

67.7%, respectively) than those at the other 2 stages, while the percentages of caregivers' hesitation toward NIP and non-NIP vaccines are lowest (n=911, 16.6% and n=2045, 37.3%, respectively). The percentages of caregivers' attitudes toward NIP vaccines significantly differ across 3 stages of the COVID-19 pandemic (P<.001; Table S1 in Multimedia Appendix 2). The percentages of caregivers' attitudes toward non-NIP vaccines significantly differ across 3 stages of the COVID-19 pandemic (P<.001; Table S2 in Multimedia Appendix 2).

Figure 1. Caregivers' choices for NIP and alternative non-NIP vaccines and their hesitancy at 3 stages of the COVID-19 pandemic. NIP: National Immunization Program.

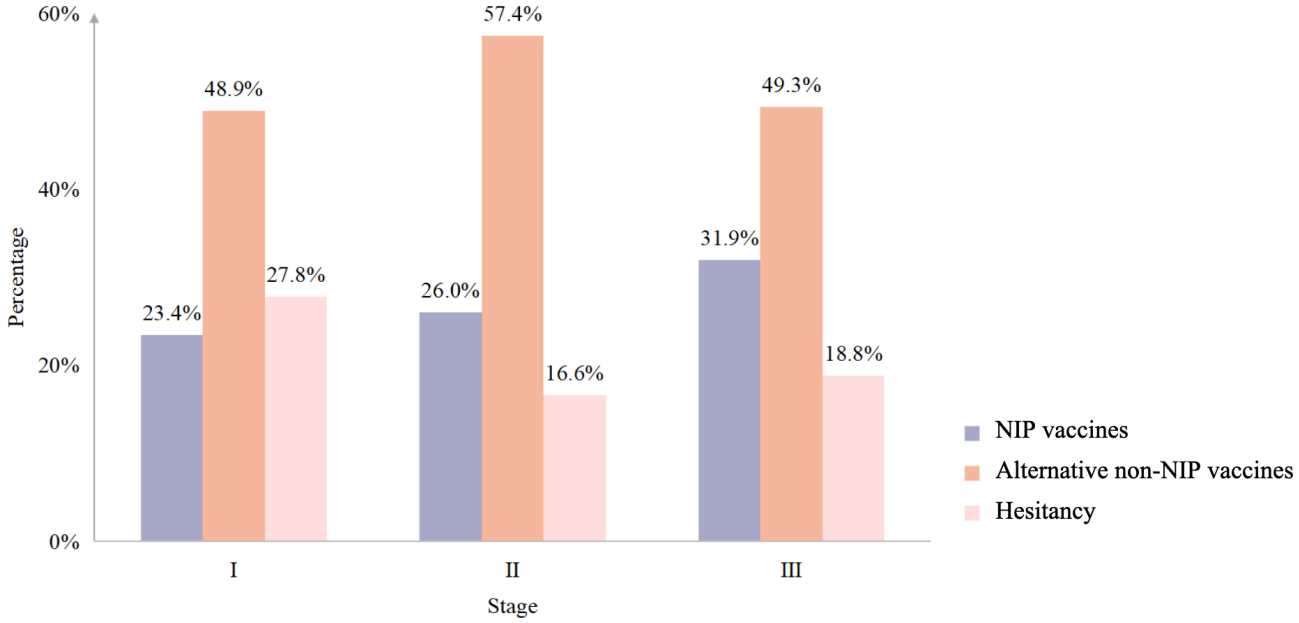
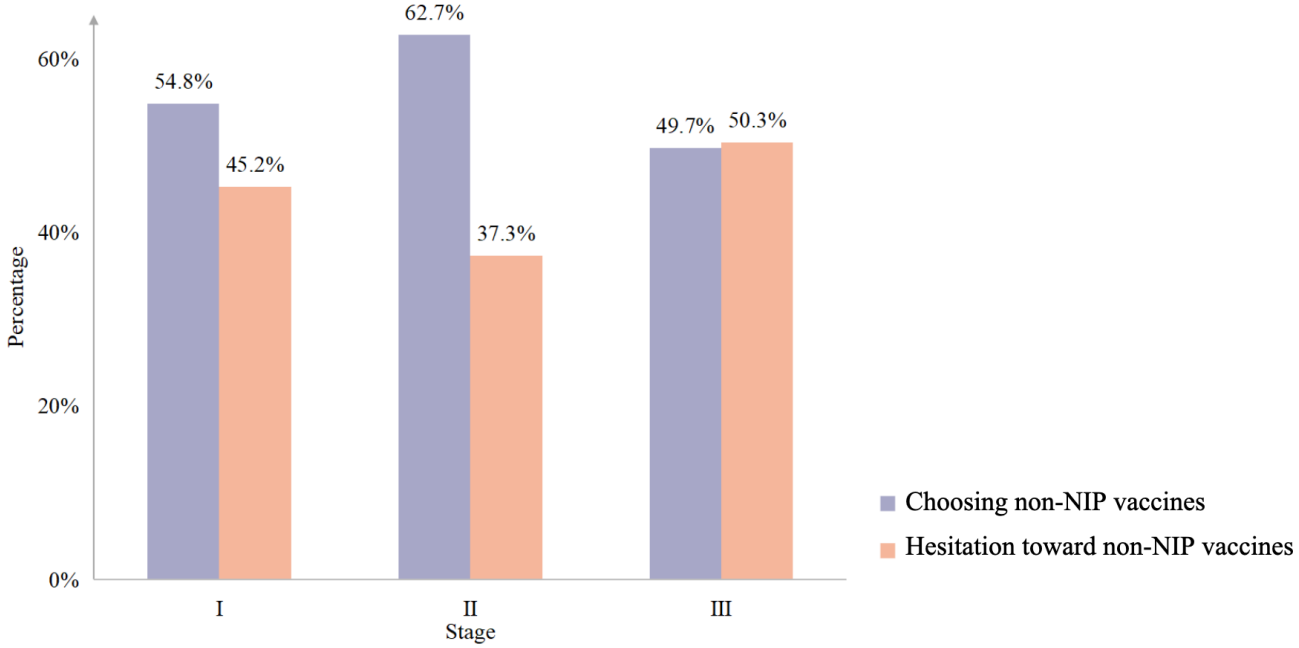


Figure 2. Caregivers' choices for non-NIP vaccines and their hesitancy at 3 stages of the COVID-19 pandemic. NIP: National Immunization Program.



Associated Factors for Caregivers' Choices of NIP Vaccines and Alternative Non-NIP Vaccines and Hesitations Toward NIP Vaccines

Table 3 presents the results of multinomial logistic regression regarding the associated factors for caregivers' choices of alternative non-NIP vaccines and hesitations toward NIP vaccines using caregivers' choices for NIP vaccines as a reference.

XSL•FO
RenderX

Table . Associated factors for caregivers' hesitation toward NIP[a] vaccines using multinomial logistic regression analysis.

Variables	Choosing alternative non-NIP vaccines, OR[b] (95% CI)	Hesitation toward NIP vaccines, OR (95% CI)
Stage of COVID-19 pandemic		
I	Reference	Reference
II	1.3 (1.2-1.6)[c]	0.6 (0.5-0.8)[c]
III	0.9 (0.7-1.2)	0.6 (0.4-0.8)[d]
Children's sex		
Male	Reference	Reference
Female	0.8 (0.8-1.0)[d]	0.9 (0.8-1.1)
Age groups of children (months)		
0 - 6	Reference	Reference
7 - 12	1.1 (0.9-1.3)	0.6 (0.5-0.7)[c]
13 - 24	0.8 (0.7-1.0)	0.4 (0.3-0.5)[c]
25 - 72	0.7 (0.6-0.8)[c]	0.4 (0.3-0.5)[c]
≥72	0.4 (0.3-0.5)[c]	0.5 (0.4-0.6)[c]
Caregivers who filled the questionnaire		
Mother	Reference	Reference
Father	0.8 (0.7-0.9)[c]	1.1 (0.9-1.3)
Others	1.3 (0.9-1.8)	2.4 (1.6-3.4)[c]
Maternal educational level		
Middle school or below	Reference	Reference
High school or equivalent	1.2 (1.0-1.5)	1.1 (0.8-1.4)
Two-year college	1.5 (1.2-1.8)[d]	1.2 (0.9-1.5)
Bachelor degree or above	2.3 (1.8-3.0)[c]	1.3 (1.0-1.8)
Paternal educational level		
Middle school or below	Reference	Reference
High school or equivalent	1.1 (0.9-1.3)	1.0 (0.8-1.3)
Two-year college	1.5 (1.2-1.9)[d]	1.0 (0.7-1.3)
Bachelor degree or above	1.4 (1.1-1.8)[d]	0.9 (0.7-1.2)
Having comorbidity		
No	Reference	Reference
Yes	1.0 (0.9-1.2)	1.2 (1.0-1.4)[e]
History of allergy		
No	Reference	Reference
Yes	0.6 (0.5-0.7)[c]	0.7 (0.6-0.8)[c]
History of AEFI[f]		
No	Reference	Reference
Yes	1.2 (0.9-1.5)	0.8 (0.5-1.1)

[a]NIP: National Immunization Program.

[b]OR: odds ratio.

XSL•FO
RenderX

[c]*P*<.001.
[d]*P*<.01.
[e]*P*<.05.
[f]AEFI: adverse events following immunization.

Choosing Alternative Non-NIP Vaccines Versus NIP Vaccines

Stage II of the COVID-19 pandemic (*P*<.001), maternal educational level of 2-year college (*P*=.002) and of bachelor degree or above (*P*<.001), as well as paternal educational level of 2-year college (*P*=.001) and bachelor degree or above (*P*=.010) were significantly associated with higher odds to choose alternative non-NIP vaccines. In contrast, the child being female (*P*=.005), the child's age of 25 - 72 months (*P*<.001) and older than 72 months (*P*<.001), fathers filling in the questionnaire (*P*<.001), and having a history of allergy (*P*<.001) were significantly associated with lower odds of choosing alternative non-NIP vaccines.

Hesitation Toward NIP Vaccines Versus Choosing NIP Vaccines

Caregivers other than parents filling in the questionnaire (*P*<.001) and having comorbidity (*P*=.02) were significantly associated with higher odds of hesitations toward NIP vaccines. Stage II (*P*<.001) and III (*P*=.003) of the COVID-19 pandemic, the child age older than 6 months (*P*<.001) and having a history of allergy (*P*<.001) were significantly associated with lower odds of hesitations toward NIP vaccines.

Associated Factors for Caregivers' Choices of Non-NIP Vaccines and Hesitations Toward Non-NIP Vaccines

Table 4 presents the results of multinomial logistic regression regarding the associated factors for caregivers' hesitations toward non-NIP vaccines using the willingness of choosing non-NIP vaccines as a reference. The child age older than 72 months (*P*<.001), father (*P*<.001) and caregivers other than parents (*P*=.006) filling in the questionnaire, having comorbidity (*P*=.03), and having a history of allergy (*P*<.001) were significantly associated with higher odds of hesitation toward non-NIP vaccines. Stage II of the COVID-19 pandemic (*P*<.001), the child age of 7-12 months (*P*=.01), 13-24 months (*P*=.01), 25 - 72 months (*P*=.001), maternal educational level of bachelor degree or above (*P*=.001) and paternal educational level of 2-year college (*P*=.002), and bachelor degree or above (*P*=.003) were significantly associated with lower odds of hesitation toward the non-NIP vaccines.

XSL•FO
RenderX

Table . Associated factors for caregivers' hesitation toward non-NIP[a] vaccines using multinomial logistic regression analysis.

Predictors	Hesitation toward non-NIP vaccines, OR[b] (95% CI)
Stage of COVID-19 pandemic	
I	Reference
II	0.6 (0.5-0.7)[c]
III	1.0 (0.8-1.3)
Children's sex	
Male	Reference
Female	1.1 (1.0-1.2)
Age groups of children (months)	
0 - 6	Reference
7 - 12	0.8 (0.7-1.0)[d]
13 - 24	0.8 (0.7-1.0)[d]
25 - 72	0.8 (0.7-0.9)[c]
≥72	1.6 (1.3-1.9)[c]
Caregivers who filled the questionnaire	
Mother	Reference
Father	1.6 (1.4-1.9)[c]
Others	1.5 (1.1-1.9)[e]
Maternal educational level	
Middle school or below	Reference
High school or equivalent	0.9 (0.8-1.1)
Two-year college	0.8 (0.7-1.0)
Bachelor degree and above	0.7 (0.6-0.8)[c]
Paternal educational level	
Middle school or below	Reference
High school or equivalent	1.0 (0.8-1.2)
Two-year college	0.7 (0.6-0.9)[e]
Bachelor degree and above	0.7 (0.6-0.9)[e]
Having comorbidity	
No	Reference
Yes	1.1 (1.0-1.2)[d]
History of allergy	
No	Reference
Yes	1.3 (1.2-1.4)[c]
History of AEFI[f]	
No	Reference
Yes	0.8 (0.7-1.1)

[a]NIP: National Immunization Program.

[b]OR: odds ratio.

[c]*P*<.001.

[d]*P*<.05.

XSL•FO
RenderX

[e]$P<.01$.

[f]AEFI: adverse events following immunization.

In addition, we presented the results of the separate logistic regression analysis for different stages of the COVID-19 pandemic in Tables S3-S8 in Multimedia Appendix 2.

Discussion

Principal Findings

This study used a large-scale, retrospective dataset to examine shifts in caregivers' hesitancy of choosing NIP and non-NIP vaccines for children with special health care needs across the pre-, during, and postpandemic periods of COVID-19. Furthermore, we identified factors associated with vaccine hesitancy, including the different phases of the COVID-19 pandemic. To the best of our knowledge, this study has contributed to provide real-world evidence and comprehensive data on vaccine hesitancy among caregivers of children with special health care needs, addressing a gap in the existing literature.

Our findings indicate that, irrespective of the stage of the COVID-19 pandemic, the proportion of individuals opting for alternative non-NIP vaccines surpassed those choosing NIP vaccines. This trend may be attributed to the unique characteristics of the population under study, that is, children with special health care needs. These children differ from the healthy, general pediatric population, as they are managing certain health conditions. Consequently, their caregivers may have greater caution when selecting vaccines. The alternative non-NIP vaccines available are either inactivated vaccines, as opposed to live-attenuated NIP vaccines, or vaccines with a higher number of valences. It is understandable that caregivers opt for inactivated vaccines over live-attenuated vaccines to mitigate the potential side effects associated with live pathogens. Furthermore, parents may prefer to administer vaccines with multiple valences in a single injection to their child.

Our study revealed that caregivers' acceptance of NIP vaccines increased across the pre-, during, and postpandemic periods of COVID-19. Wang et al [29] also found that there has been an increasing acceptance of routine childhood vaccination and COVID-19 vaccination from 2020 to 2021 via a repeated cross-sectional survey. We found that the proportion of caregivers of children with special health care needs willing to vaccinate their children with alternative non-NIP vaccines was highest during the pandemic. Concurrently, we observe the lowest proportion of caregivers exhibiting hesitation toward NIP vaccines at the same stage. These findings align with the results of a multinomial logistic analysis, which indicated that this stage (ie, during the pandemic) is associated with significantly higher odds of caregivers' willingness to select alternative non-NIP vaccines. Additionally, stages II and III are significantly associated with lower odds of caregivers' hesitation toward NIP vaccines. Zhang et al [30] reported that the coverage rate of non-NIP vaccines increased by 25.8% and 34.7%, respectively, in 2020 and 2021 when it was the nonpharmaceutical intervention period compared to the prepandemic period (ie, in 2019). Our finding also echoed with previous conclusion that the COVID-19 pandemic has caused unprecedented impacts including parental trust in vaccines [31]. During the COVID-19 pandemic, both the government and the mass media have substantially intensified their advocacy and promotion of the COVID-19 vaccine. They have also enhanced risk communication to raise awareness about the benefits and risks of vaccination and to promote the scientific belief that the benefits outweigh the risks during the COVID-19 pandemic [32]. This heightened publicity has concurrently enhanced public awareness regarding vaccines in general, potentially fostering a more positive attitude and greater acceptance of vaccines among the population. Consequently, caregivers have become more inclined to vaccinate their children during the pandemic; they are even willing to pay for the self-funded, alternative NIP vaccines.

Regarding caregivers' attitudes toward non-NIP vaccines, our study indicates that caregivers' willingness to vaccinate their children with special health care needs was at its peak, and their hesitancy was at its lowest during the pandemic. This observation aligns with the results obtained from the multinomial regression analysis. The increased willingness can likely be attributed to the extensive promotion of vaccines, particularly the COVID-19 vaccine, by governments and social media platforms. However, following the COVID-19 pandemic, our study observed a decline in the proportion of caregivers opting for non-NIP vaccines and an increase in their hesitancy toward these vaccines. In this study, stage III corresponds to the 6-month period following the relaxation of COVID-19 measures in China, during which the Chinese public exhibited a varied and mixed attitude toward vaccines. This period was characterized by both trust and mistrust in vaccines, exacerbated by reports of incidents such as "fake vaccines" or vaccine scandals that were disseminated through social media. There has been a scarcity of literature regarding the impact of the COVID-19 pandemic on non-NIP vaccination intention for children. Wang et al [33] found that in a community sample, approximately 70% of parents did not change their intention for self-paid, non-NIP vaccines for their children, and about 20% of parents increased their intention, while about 10% decreased the intention after the COVID-19 pandemic. In this study, a short period of stage III resulted in a limited sample size for data collection, potentially leading to findings that may be attributable to chance. Therefore, we recommend conducting future studies with an increased sample size by extending the study period during the "postpandemic" phase to validate or refute our findings.

Our study indicated a sex disparity in caregivers' selection of alternative, self-funded non-NIP vaccines. Caregivers of female children were less likely to choose these alternative non-NIP vaccines compared to caregivers of male children. This pattern remained the same before and during the pandemic but was reversed after the pandemic when caregivers of female children were more likely to choose the alternative non-NIP compared to caregivers of male children. Although global immunization coverage does not show significant differences between male

XSL•FO
RenderX

and female children, certain countries and communities may experience sex-related barriers to immunization due to prevailing social and cultural norms [34]. The outbreak of the COVID-19 pandemic might have increased the intention of caregivers of female children to choose paid alternative non-NIP vaccines for them.

Regarding the age of children, our study indicated that caregivers of older children demonstrated a lower propensity to vaccinate their children with alternative non-NIP vaccines or exhibited reduced vaccine hesitancy compared to caregivers of very young children (ie, 0 - 6 months). This pattern remains the same in the separate logistic analysis in each stage. Notably, our findings revealed that the likelihood of vaccine hesitancy toward non-NIP vaccines among caregivers of children older than 72 months was 1.58 times greater than that among caregivers of children aged 0 - 6 months. A potential experiential factor may account for this finding. Non-NIP vaccines are not mandatory for children, and advocacy or public awareness efforts are less robust compared to those for NIP vaccines. Advocacy predominantly targets young children, typically younger than 6 years of age. Consequently, caregivers of older children may lack awareness regarding the appropriate timing and types of non-NIP vaccines beneficial for their children. Therefore, our study suggests that it is important and necessary to educate caregivers of children older than 6 years.

In this study, our findings indicate that fathers, in comparison to mothers, exhibited a lower propensity to choose for alternative non-NIP vaccines and demonstrated a greater likelihood of hesitancy in selecting non-NIP vaccines for their children. Furthermore, the analysis revealed that caregivers other than parents showed a higher probability of hesitancy in choosing non-NIP vaccines for children when compared to mothers. This finding is noteworthy and may be attributed to the cultural norm within Chinese families, where mothers predominantly assume the role of primary caregiver and are chiefly responsible for making medical decisions concerning their children. While it is important to acknowledge the contributions of fathers and other caregivers in child-rearing, their involvement and roles may not be as prominent or comparable to those of mothers.

Our study demonstrates that a relatively high maternal and paternal educational level, such as a 2-year college and bachelor or above degree, is associated with the higher intention for choosing alternative non-NIP vaccines and less hesitation to both NIP and non-NIP vaccines, which is consistent with previous studies. For instance, a study in Shanghai showed that higher income, higher education, and greater access to vaccines made respondents more willing to immunize their children with non-NIP [35]. Economic status remained a significant factor in determining the acceptability of non-NIP vaccines [36]. Parents' high education is usually linked to high health literacy, which greatly affects their ability to use health information to make health decisions for their children [37].

In this study, caregivers of children with comorbidities exhibit increased hesitancy regarding NIP and non-NIP vaccines. This hesitancy is understandable, as the existing health conditions of these children already present significant concerns, prompting caregivers to exercise greater caution when considering vaccination decisions. We also found that caregivers of children with allergy had increased hesitancy regarding non-NIP vaccines, which may stem from concerns about an elevated risk of adverse effects following vaccination in children with a history of allergies.

The vaccination of children with special health care needs is unique, as their health issues may lead to delayed or missed vaccinations, leaving them vulnerable to infections. We analyzed how China's macroprevention and control strategies for COVID-19 may affect the vaccination intentions of this group using a large sample. Our findings indicate that during the COVID-19 pandemic, caregivers exhibited a marked increase in their willingness to receive vaccinations, accompanied by a significant decrease in vaccine hesitancy, compared to levels observed both prior to and following the pandemic. This empirical conclusion aligns with observations made by physicians in vaccination clinics. The onset of the pandemic appears to have heightened parental awareness regarding vaccines, thereby enhancing vaccination rates. This study revealed that a significant proportion of parents continue to exhibit hesitancy toward vaccines, highlighting the persistence of a knowledge gap and the potential for enhancement in health education initiatives. Furthermore, the research identified several factors associated with vaccine hesitancy, including the child's age and sex, the educational level of the parents, and the presence of an allergy history. These findings aim to furnish evidence-based support for the formulation of personalized and targeted health education strategies in future endeavors. The findings of this study further indicate the necessity for social mobilization from a macrolevel perspective to collaboratively enhance vaccine awareness and advance health equity.

Strengths and Limitations

To the best of our knowledge, this study is the first to use a large sample size (N=7770) to examine attitudes, specifically hesitancy, toward vaccinations among caregivers of children with special health care needs before, during, and after the COVID-19 pandemic. Additionally, our study has documented a comprehensive set of variables, which allows us to apply multivariate regression analysis to identify associated factors for caregivers' attitudes toward NIP and non-NIP vaccines in China. We consider it to be among the few studies of its kind conducted globally.

However, several limitations merit consideration. First, this study used a cross-sectional design, which restricts our ability to infer causality and confines our analysis to assessing associations. Second, we did not use a standardized or validated instrument, such as the Vaccine Hesitancy Scale, to measure vaccine hesitancy. Instead, we used 2 self-developed questions as a proxy for assessing vaccine hesitancy. Third, our study exclusively assessed caregivers' willingness or attitudes toward vaccinating their children with NIP or non-NIP vaccines. It is important to acknowledge the potential discrepancy between the respondents' self-reported willingness to vaccinate and their actual vaccination behavior, as survey responses may not always accurately reflect real-world actions [38].

XSL•FO
RenderX

Conclusions

This study demonstrated that caregivers' willingness to vaccinate their children with special health care needs with NIP and non-NIP vaccines was highest and their hesitancy was lowest during the COVID-19 pandemic in China. Additionally, we have identified multiple factors associated with caregivers' willingness and hesitancy of vaccinating their children. These findings provide evidence-based support for developing personalized health education strategies and further indicate the necessity for social mobilization at a macrolevel to collaboratively enhance vaccine awareness and health literacy regarding vaccinating children with special health care needs.

Acknowledgments

This study was financially supported by the Medical Science and Technology Project of Zhejiang province (2020385296). The funders had no role in the study design; data collection, analysis, or interpretation; in the report's writing; or in the decision to submit the paper for publication.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to safety and privacy concerns but are available from CJ (6198011@zju.edu.cn) on reasonable request.

Authors' Contributions

ML, CS, CJ, and GB were involved in the study conceptualization and study design. CS and MG carried out the data analyses. ML and GB wrote the first draft of the manuscript. XW, DY, JG, and LS contributed to the data collection, data entering, and data cleaning. AR, ASG, and SHSSM contributed to searching the literature, visualizing the results, and editing the manuscript. CJ and GB supervised the project. All authors provided critical revisions of the paper for important intellectual content, contributed to the interpretation of the data, and approved the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The spectrum of disease in the overall study population and at each stage of the COVID-19 pandemic. (A) The spectrum of disease in the overall study population, (B) the spectrum of disease at stage I of the COVID-19 pandemic, (C) the spectrum of disease at stage II of the COVID-19 pandemic, and (D) the spectrum of disease at stage III of the COVID-19 pandemic.
[TIFF File, 6291 KB - publichealth_v11i1e67487_app1.tiff]

Multimedia Appendix 2
Additional tables.
[DOCX File, 49 KB - publichealth_v11i1e67487_app2.docx]

References

1. National Immunization Program vaccine children's immunization procedures and instructions. National Health Commission of the People's Republic of China. 2021. URL: http://www.nhc.gov.cn/jkj/s3581/202103/590a8c7915054aa682a8d2ae8199e222/files/7559f978e6be4ec585a9a1fe0d9224f6.pdf [accessed 2025-03-13]
2. Childhood Immunization Schedule for National Immunization Program Vaccines—China (version 2021). China CDC Wkly 2021 Dec 24;3(52):1101-1108. [doi: 10.46234/ccdcw2021.270] [Medline: 35186365]
3. Chen S, Rodewald LE, Du AH, Tang S. Advancing the National Immunization Program in an era of achieving universal vaccine coverage in China and beyond. Infect Dis Poverty 2024 Mar 13;13(1):25. [doi: 10.1186/s40249-024-01192-6] [Medline: 38475849]
4. Zhou Y, Li D, Cao Y, et al. Immunization coverage, knowledge, satisfaction, and associated factors of non-National Immunization Program vaccines among migrant and left-behind families in China: evidence from Zhejiang and Henan provinces. Infect Dis Poverty 2023 Oct 13;12(1):93. [doi: 10.1186/s40249-023-01145-5] [Medline: 37833775]
5. Zhang H, Lai X, Mak J, et al. Coverage and equity of childhood vaccines in China. JAMA Netw Open 2022 Dec 1;5(12):e2246005. [doi: 10.1001/jamanetworkopen.2022.46005]
6. Baumgaertner B, Ridenhour BJ, Justwan F, Carlisle JE, Miller CR. Risk of disease and willingness to vaccinate in the United States: a population-based survey. PLoS Med 2020 Oct;17(10):e1003354. [doi: 10.1371/journal.pmed.1003354] [Medline: 33057373]
7. Doherty M, Schmidt-Ott R, Santos JI, et al. Vaccination of special populations: protecting the vulnerable. Vaccine (Auckl) 2016 Dec 20;34(52):6681-6690. [doi: 10.1016/j.vaccine.2016.11.015] [Medline: 27876197]

XSL•FO
RenderX

8. Shahani R, Chu J, Rufai OH, et al. Understanding the role of psychosocial factors in Pakistani parents' hesitancy to vaccinate their kids: the mediating role of knowledge and mistrust of science about the COVID-19 vaccine. Vaccines (Basel) 2022 Aug 5;10(8):1260. [doi: 10.3390/vaccines10081260] [Medline: 36016148]
9. Xu B, Zhang Y, Zhou C, Wang Q, Luan R. Vaccination coverage and immunization safety among children with special health status in Chongqing, China. Hum Vaccin Immunother 2022 Dec 30;18(7):2138466. [doi: 10.1080/21645515.2022.2138466] [Medline: 36469710]
10. McPherson M, Arango P, Fox H, et al. A new definition of children with special health care needs. Pediatrics 1998 Jul;102(1 Pt 1):137-140. [doi: 10.1542/peds.102.1.137] [Medline: 9714637]
11. Wang L, Hao YH, Bai Y. Public health and clinical care integration to improve immunization of children with special health care needs. Zhonghua Yu Fang Yi Xue Za Zhi 2022 Jun 6;56(6):673-678. [doi: 10.3760/cma.j.cn112150-20220104-00010] [Medline: 35785850]
12. de Souza RP, Ribeiro ALR, de Menezes SAF, Machado LFA. Incidence of respiratory syncytial virus infection in children with congenital heart disease undergoing immunoprophylaxis with palivizumab in Pará state, north region of Brazil. BMC Pediatr 2019 Aug 28;19(1):299. [doi: 10.1186/s12887-019-1681-6] [Medline: 31462289]
13. Yang L, Peng J, Deng J, et al. Vaccination status of children with epilepsy or cerebral palsy in Hunan rural area and a relative KAP survey of vaccinators. Front Pediatr 2019;7:84. [doi: 10.3389/fped.2019.00084] [Medline: 30984716]
14. Talbird SE, Carrico J, La EM, et al. Impact of routine childhood immunization in reducing vaccine-preventable diseases in the United States. Pediatrics 2022 Sep 1;150(3):e2021056013. [doi: 10.1542/peds.2021-056013] [Medline: 35821599]
15. Luan L, Zhang Z, Xu J, et al. Evaluation of vaccination status of children with special health care needs in Suzhou. Hum Vaccin Immunother 2020;19(2):2254965. [doi: 10.1080/21645515.2023.2254965]
16. Laforgia N, Di Mauro A, Bianchi FP, et al. Are pre-terms born timely and right immunized? Results of an Italian cohort study. Hum Vaccin Immunother 2018 Jun 3;14(6):1398-1402. [doi: 10.1080/21645515.2018.1428509] [Medline: 29351055]
17. Cossarizza A, Chang HD, Radbruch A, et al. Guidelines for the use of flow cytometry and cell sorting in immunological studies (third edition). Eur J Immunol 2021 Dec;51(12):2708-3145. [doi: 10.1002/eji.202170126] [Medline: 34910301]
18. Notice of the National Health Commission on issuing the National Immunization Program Vaccine Children's Immunization Procedures and Instructions (2021 edition). National Health Commission of People's Republic of China. URL: http://www.nhc.gov.cn/cms-search/xxgk/getManuscriptXxgk.htm?id=590a8c7915054aa682a8d2ae8199e222 [accessed 2024-10-07]
19. Olusanya OA, Bednarczyk RA, Davis RL, Shaban-Nejad A. Addressing parental vaccine hesitancy and other barriers to childhood/adolescent vaccination uptake during the coronavirus (COVID-19) pandemic. Front Immunol 2021;12:663074. [doi: 10.3389/fimmu.2021.663074] [Medline: 33815424]
20. Qin C, Wang R, Tao L, Liu M, Liu J. Association between risk perception and acceptance for a booster dose of COVID-19 vaccine to children among child caregivers in China. Front Public Health 2022;10:834572. [doi: 10.3389/fpubh.2022.834572] [Medline: 35372197]
21. Lee M, You M. Direct and indirect associations of media use with COVID-19 vaccine hesitancy in South Korea: cross-sectional web-based survey. J Med Internet Res 2022 Jan 6;24(1):e32329. [doi: 10.2196/32329] [Medline: 34870605]
22. Zhou X, Song S, Zhang Y, Hou Z. Deep learning analysis of COVID-19 vaccine hesitancy and confidence expressed on twitter in 6 high-income countries: longitudinal observational study. J Med Internet Res 2023 Nov 6;25:e49753. [doi: 10.2196/49753] [Medline: 37930788]
23. Goldman RD, McGregor S, Marneni SR, et al. Willingness to vaccinate children against influenza after the coronavirus disease 2019 pandemic. J Pediatr 2021 Jan;228:87-93. [doi: 10.1016/j.jpeds.2020.08.005]
24. Wang Z, Chen S, Fang Y. Parental willingness and associated factors of pediatric vaccination in the era of COVID-19 Pandemic: a systematic review and meta-analysis. Vaccines (Basel) 2022 Sep 2;10(9):1453. [doi: 10.3390/vaccines10091453] [Medline: 36146530]
25. Geweniger A, Haddad A, Barth M, et al. Mental health of children with and without special healthcare needs and of their caregivers during COVID-19: a cross-sectional study. BMJ Paediatr Open 2022 Jun;6(1):e001509. [doi: 10.1136/bmjpo-2022-001509] [Medline: 36053611]
26. Li M, Ji C, Wang B, et al. Incomplete vaccination among children with special health care needs in Zhejiang, China: analysis of retrospective data. Front Pediatr 2019;7:173. [doi: 10.3389/fped.2019.00173] [Medline: 31119119]
27. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
28. Bethell CD, Read D, Stein RE, Blumberg SJ, Wells N, Newacheck PW. Identifying children with special health care needs: development and evaluation of a short screening instrument. Ambul Pediatr 2002;2(1):38-48. [doi: 10.1367/1539-4409(2002)0022.0.co;2] [Medline: 11888437]
29. Wang Q, Xiu S, Yang L, et al. Changes in parental attitudes toward COVID-19 vaccination and routine childhood vaccination during the COVID-19 pandemic: repeated cross-sectional survey study. JMIR Public Health Surveill 2022 May 13;8(5):e33235. [doi: 10.2196/33235] [Medline: 35486516]
30. Zhang C, Hu W, Li Y, Lv Y, Zhang S. Impact of the COVID-19 pandemic on routine vaccination services in Shaanxi province, Northwest China: non-pharmaceutical intervention period and mass COVID-19 vaccination period. Hum Vaccin Immunother 2023 Aug 1;19(2):2251826. [doi: 10.1080/21645515.2023.2251826] [Medline: 37665634]

XSL•FO
RenderX

31. Saxena S, Skirrow H, Bedford H. Routine vaccination during covid-19 pandemic response. BMJ 2020 Jun 16;369:m2392. [doi: 10.1136/bmj.m2392] [Medline: 32546575]
32. Wu J, Li Q, Silver Tarimo C, et al. COVID-19 vaccine hesitancy among Chinese population: a large-scale national study. Front Immunol 2021;12:781161. [doi: 10.3389/fimmu.2021.781161] [Medline: 34912346]
33. Wang S, Xu J, Zhu J. Impact of the COVID-19 pandemic on self-paid vaccination intentions for children: a cross-sectional study in China. BMJ Open 2024 Aug;14(8):e083056. [doi: 10.1136/bmjopen-2023-083056]
34. Gender and immunisation. Gavi. URL: https://www.gavi.org/our-alliance/strategy/gender-and-immunisation [accessed 2024-10-08]
35. Wu L, Huang Z, Guo X, Liu J, Sun X. Measuring parents' acceptance of non-national immunization program vaccines for children and its influencing factors during the COVID-19 pandemic in Shanghai, China. Hum Vaccin Immunother 2022 Nov 30;18(5):2069427. [doi: 10.1080/21645515.2022.2069427] [Medline: 35727599]
36. Zhang X, Syeda ZI, Jing Z, et al. Rural-urban disparity in category II vaccination among children under five years of age: evidence from a survey in Shandong, China. Int J Equity Health 2018 Jun 22;17(1):87. [doi: 10.1186/s12939-018-0802-4] [Medline: 29929503]
37. Morrison AK, Glick A, Yin HS. Health literacy: implications for child health. Pediatr Rev 2019 Jun;40(6):263-277. [doi: 10.1542/pir.2018-0027] [Medline: 31152099]
38. Zhang V, Zhu P, Wagner AL. Spillover of vaccine hesitancy into adult COVID-19 and influenza: the role of race, religion, and political affiliation in the United States. Int J Environ Res Public Health 2023 Feb 15;20(4):3376. [doi: 10.3390/ijerph20043376] [Medline: 36834071]

Abbreviations

NIP: National Immunization Program

Edited by A Mavragani; submitted 13.10.24; peer-reviewed by E Taherifard, Y Fang; revised version received 26.01.25; accepted 11.02.25; published 26.03.25.

Please cite as:
Li M, Sun C, Ji C, Gao M, Wang X, Yao D, Guo J, Sun L, Rafay A, George AS, Muhandiramge SHSS, Bai G
Vaccine Hesitancy and Associated Factors Among Caregivers of Children With Special Health Care Needs in the COVID-19 Era in China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e67487
URL: https://publichealth.jmir.org/2025/1/e67487
doi:10.2196/67487

© Mingyan Li, Changxuan Sun, Chai Ji, Meiying Gao, Xia Wang, Dan Yao, Junxia Guo, Lidan Sun, Abdul Rafay, Antonita Shereen George, Sanduni Hasara Samararathna Samararathna Muhandiramge, Guannan Bai. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 26.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Prevalence and Associated Factors of Bullying Experienced by Community Patients With Serious Mental Disorders: Cross-Sectional Study

Yazhuo Qi[1], BPA; Meiqi Wang[1], MPH; Zhen Wei[1], MSc; Wenyu Wang[1], BPA; Kaixian Wang[1], BPA; Long Sun[1], PhD

Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University, Jinan, Shandong Province, China

Corresponding Author:
Long Sun, PhD
Centre for Health Management and Policy Research, School of Public Health, Cheeloo College of Medicine, Shandong University
No. 44, West Wenhua Road, Jinan, Shandong Province
Jinan, Shandong Province, 250012
China
Phone: 86 1 325 666 0556
Email: sunlong@sdu.edu.cn

Abstract

Background: Individuals with serious mental disorders (SMDs) are frequently exposed to bullying, which can severely affect their well-being. However, research on this issue remains limited in the Chinese context.

Objective: This study aimed to examine the prevalence and associated factors of bullying experienced by individuals with SMDs in China, with the goal of providing evidence that may inform strategies to improve their quality of life.

Methods: A multistage stratified random sampling method was used to survey 486 community patients with SMDs in Shandong province, China. Participants self-reported experiences of bullying. The Stigma Scale for Chronic Illnesses, 8-item version was used to measure self-stigma; the Pittsburgh Sleep Quality Index assessed sleep quality; and the Family Apgar Scale evaluated family functioning. Single-factor analysis, logistic regression, and negative binomial regression were used to analyze the associations between variables.

Results: The prevalence rate of bullying experienced by individuals with SMDs was 42.4% (206/486). Higher levels of self-stigma were significantly associated with increased odds of being bullied (odds ratio [OR] 1.041, 95% CI 1.017-1.065; $P<.001$), while better family functioning was associated with lower odds of experiencing bullying (OR 0.913, 95% CI 0.850-0.980; $P=.01$). Verbal abuse was associated with poorer sleep quality (OR 1.056, 95% CI 1.004-1.111; $P=.04$), particularly when reported as perpetrated by caregivers (OR 1.068, 95% CI 1.029-1.108; $P<.001$). In addition, age ($P<.001$), education level ($P=.02$), only-child status ($P=.04$), and diagnosis type ($P=.03$) were significantly related to bullying experiences.

Conclusions: Higher levels of self-stigma were associated with experiencing bullying among individuals with SMDs, whereas better family functioning was associated with lower odds of being bullied. Verbal abuse—especially by caregivers—was associated with poorer sleep quality. These findings suggest that reducing self-stigma and strengthening family support may help improve quality of life for individuals with SMDs.

(JMIR Public Health Surveill 2025;11:e75027) doi:10.2196/75027

KEYWORDS

serious mental disorders; bullying; stigma; sleep quality; family functioning

Introduction

Background

Serious mental disorders (SMDs) are mental disorders in which the symptoms are so severe that an individual's social functioning is significantly impaired, preventing them from accurately understanding their health status or objective reality or managing their own affairs [1]. In China, the spectrum of SMDs is diverse and includes conditions such as schizophrenia, paranoid mental disorder, bipolar disorder, schizoaffective disorder, epileptic mental disorder, and intellectual disability

XSL•FO
RenderX

associated with mental disorders [2]. Official figures show that approximately 4.7% of people in China experience SMDs, highlighting the significant prevalence of these conditions [2]. The impact of mental disorders on public health is profound. In 2021, they ranked among the leading causes of disease burden, accounting for 1554.2 disability-adjusted life years per 100,000 people globally [3]. This metric underscores the extent to which mental disorders contribute to the overall loss of health and productivity in the population. Therefore, mental disorders are an important public health issue that requires more attention and research [4].

Patients with SMDs often face multiple challenges and societal discrimination, and bullying is one of the most serious problems [5]. Discrimination refers to unfair treatment based on group characteristics, whereas bullying is repeated, intentional harm or intimidation targeting an individual, often involving a power imbalance. People with SMDs are often misunderstood as being violent—in reality, they are more likely to experience violence and bullying rather than commit such acts [6]—and this widespread misconception contributes to the self-stigma and discrimination faced by those with mental health conditions, further exacerbating their vulnerabilities. Studies in the United States have shown that experiencing bullying can affect the course of illness and treatment outcomes for patients with SMDs [7] and can also reduce their quality of life [8]. An Australian study demonstrated a longitudinal association between being bullied and anxiety and depressive disorders, showing that people with SMDs who are bullied often develop severe anxiety and depressive symptoms, which leads to a further deterioration in their mental health status [9]. In addition, a study conducted in Canada, Finland, Germany, and Sweden found that exposure to bullying, depression, and medication nonadherence were all associated with poorer real-life functioning in men with schizophrenia" (assuming that "exposure to bullying" can substitute for "victimization" in the cited work) [10]. To reduce the negative impact of bullying on patients with SMDs, all sectors of society need to work together to increase public awareness of the issue, strengthen mental health support and legal protection, and develop and implement effective antibullying measures. These efforts are crucial for improving the living conditions and overall well-being of people with SMDs.

Previous studies have reported prevalence estimates and risk factors for bullying experienced by individuals with SMDs. In the largest study of people with a major mental illness in the United States, 44% reported experiencing violence [8]. There are many factors influencing exposure to bullying in patients with SMDs. A Serbian study of clinical patients with schizophrenia found that sex was a key factor, with men being more likely to experience physical bullying [11]. A longitudinal study in the United Kingdom confirmed that younger age, single marital status, homelessness, and limited contact with family were associated with a significantly increased risk of being bullied [5]. In addition, studies in the United States [8] and Egypt [12] have shown that patients with SMDs are mainly bullied by acquaintances, followed by family members [13]. However, most of the studies on bullying have been conducted mainly in Europe [5,11], the United States [8], and Australia [14]. No survey has examined bullying experienced by community patients with SMDs in China. This gap indicates that research on bullying in this population in China remains limited. Understanding the prevalence and nature of bullying experienced by patients with SMDs in China is crucial for developing effective interventions tailored to the country's specific cultural and social context.

Objectives

Due to the lack of relevant research, the specific experiences or related factors of bullying experienced by Chinese patients with SMDs remain unclear. Given China's unique cultural and social context, these experiences may differ significantly from those reported in Western countries [15]. Cultural factors such as strong family ties, collectivist values, self-stigma surrounding mental illness, and differences in mental health care accessibility may influence both the prevalence and nature of bullying experienced by patients with SMDs in China [16]. Moreover, social perceptions of mental illness and institutional factors may shape how patients with SMDs navigate their daily lives, potentially increasing their vulnerability to bullying [17].

Therefore, this study aimed to report the prevalence of bullying experienced by community patients with SMDs in China and to examine the associations among disease-specific factors, sleep quality, self-stigma, and family functioning that influence exposure to bullying among patients with SMDs, while controlling for sociodemographic variables. This comprehensive approach can help us understand and address the bullying issues faced by people with SMDs, ultimately improving their quality of life and enhancing the recovery process. In addition, the findings could provide an important scientific basis for developing more effective prevention and intervention strategies to promote the health and well-being of people with SMDs on a broader societal level.

Methods

Study Design and Sampling

This cross-sectional study investigated individuals with SMDs residing in Shandong province in eastern China. A multistage stratified random sampling method was used to ensure representativeness and reliability [18]. First, the 16 major cities in Shandong province were categorized into 2 groups based on their respective gross domestic product. One city was then randomly chosen from each group, resulting in the selection of Jinan and Zaozhuang. Within these 2 cities, 1 administrative district was randomly chosen, specifically Zhangqiu and Taierzhuang, respectively. Finally, 6 townships and communities were randomly selected as survey sites within each administrative district. Participants were recruited from the National Information System for Severe Mental Disorders, which records all patients receiving long-term community-based management [19]. Local community health institutions assisted with recruitment by providing patient lists and contacting eligible individuals or their guardians to explain the study and confirm voluntary participation [20]. Face-to-face interviews were arranged at participants' homes or local health centers after obtaining written informed consent [21]. Inclusion criteria were registration in the National Information System for Severe

XSL•FO
RenderX

Mental Disorders, residence in the community for at least 6 months, and ability to communicate independently or with caregiver support. A total of 486 individuals with SMDs were enrolled.

Data Collection

The study was carried out from August to September 2022. Before administering the survey, all investigators—graduate students proficient in both the research topic and questionnaire design—underwent professional training. Data were collected through 1-on-1, on-site interviews using a predesigned questionnaire [22,23]. Recognizing the unique circumstances of the respondents, the investigators sought the assistance of the patient's caregiver during the interviews. To ensure familiarity with the patient's daily experiences, only caregivers who had lived with and provided care for the patient for at least 6 months were eligible to respond [24]. In cases where respondents faced cognitive impairment or difficulty answering, their caregiver provided responses on their behalf [25]. The investigators were trained to use standardized definitions and provide clear examples of key terms such as "bullying" and "verbal abuse" to minimize subjective variation in interpretation [26]. Interviews were conducted either at participants' homes or local health centers, based on their convenience and preference, and responses were recorded directly onto the questionnaire. After the survey, at least 2 trained students meticulously reviewed each questionnaire to ensure its validity, double checking and completing missing or unclear data entries as necessary [27,28].

Ethical Considerations

This study was approved by the institutional review board of the Shandong University School of Public Health (LL20210803). All participants volunteered and provided written informed consent. For those who were illiterate or semiliterate, written informed consent was obtained from their legal guardian [29]. All collected information was anonymized using unique ID codes. In accordance with the institutional review board guidelines, no financial compensation was provided to participants to avoid potential coercion. The manuscript contains no personally identifiable information, and no images requiring additional consent were used. The research strictly adhered to the principles of the Declaration of Helsinki and Chinese ethical regulations governing human participant research.

Measures

Bullying Experiences

Bullying experiences in this study were assessed via self-report using 2 primary questions: "Have you ever been verbally abused by a caregiver or other person?" "Have you ever been physically assaulted by a caregiver or other person?" Respondents were required to answer with either "yes" or "no." Those who indicated that they had not encountered bullying were assigned a score of "0," whereas respondents who reported experiencing bullying were assigned a score of "1." Subsequently, interviewers delved deeper to ascertain the specific types of bullying encountered by asking, "Have you experienced any of the following behaviors by caregivers/others?" Response options included "discrimination," "verbal abuse," "beatings," "imprisonment," "cold starvation," "forced physical labor," "bondage," "other," and "none." All respondents who answered "no" to experiences of bullying were designated as the control group, labeled as "with no bullying," and those who reported any form of bullying were categorized under "any bullying." The interviewers then calculated bullying scores by summing the reported instances of various types of bullying for each respondent, categorizing them as either "verbal bullying" or "physical bullying."

Sociodemographic Variables

The primary sociodemographic variables examined in this study included sex, age, education level, sibling status, and marital status. These variables were coded numerically to facilitate analysis. Sex was coded as male (1) and female (2). Age categories, derived from participants' date of birth, were coded as <45 years (1), 45 to 64 years (2), and >64 years (3). Educational level was coded as illiteracy or semiliteracy (1), primary school (2), and middle school or above (3). Sibling status was coded as single child (1) or other (2). Marital status was simplified and coded as married (1) and other (2), with the latter category encompassing single, divorced, widowed, and other statuses due to their small proportions.

Illness Information

SMDs include 6 distinct types: schizophrenia, schizoaffective disorder, paranoid psychosis, bipolar disorder, mental disorders resulting from epilepsy, and intellectual disability associated with mental disorders accompanied by mental disorders [30]. However, due to the relatively small proportion of patients with the latter 5 types, This study grouped SMD types into schizophrenia (1) and other (2) to assess patients' disease profiles.

To assess patients' current psychiatric symptoms, we used the 18-item Brief Psychiatric Rating Scale, a widely recognized and validated tool for quantifying the severity of mental disorders [31]. Illness duration was measured in years (ie, duration of SMD), based on self-reports from patients or their caregivers. In addition, we used the Stigma Scale for Chronic Illnesses, 8-item version (SSCI-8) to evaluate patients' self-stigma toward mental illness [32].

Sleep Quality

Sleep quality was evaluated using the Pittsburgh Sleep Quality Index [33]. This comprehensive scale comprises 7 dimensions aimed at assessing individual sleep quality and the severity of sleep disturbances. These dimensions include subjective sleep quality, sleep duration, sleep disturbance, sleep latency, habitual sleep efficiency, daytime dysfunction, and sleep medication use. Each dimension is scored from 0 to 3, yielding a total score ranging from 0 to 21. Higher scores indicate poorer sleep quality. The scale assesses respondents' experiences over the past month. It is widely used in studies evaluating the sleep quality of the Chinese population due to its high reliability and validity [34].

Family Functioning

This study used the Family Apgar Scale to evaluate family functioning [35]. Renowned for its brevity and clarity, this scale

XSL•FO
RenderX

has found widespread application in assessing family dynamics [36]. In China, notably, it has demonstrated efficacy as a diagnostic tool for identifying individuals with family dysfunction [37]. The scale comprises 5 items measuring adaptation, partnership, growth, affection, and closeness. Scores from the 5 items are summed to yield a total score ranging from 0 to 10, with 0 to 3 indicating severe dysfunction, 4 to 6 moderate dysfunction, and 7 to 10 good functioning [35]. In this study, the Cronbach α coefficient for the scale was 0.934, indicating robust reliability.

Data Analyses

The primary analysis was conducted using SPSS software (version 27.0; IBM Corp). Continuous variables were presented as means and SDs, while categorical data were reported as frequencies and percentages. To assess the mean differences between the "any bullying" and "with no bullying" groups, either a chi-square test or a 1-way ANOVA was used. As the outcome was binary (yes or no), we used logistic regression to assess associations between the variables and experiences of bullying, reporting odds ratios (ORs) with 95% CIs as measures of association rather than causality. For the multivariable logistic regression analysis, dummy variables were established for multicategory variables to explore factors associated with experiences of bullying [38].

In addition, generalized estimating equation (GEE) analyses were used to investigate factors linked to the number of types of bullying behaviors experienced. Given that the dependent variable constituted count data, we opted for a negative binomial distribution to model it [39]. Within the GEE framework, we conducted Wald tests for hypothesis testing to determine whether the effect of independent variables on the number of types of bullying behaviors experienced was significant. Throughout the analysis, we controlled for variables such as age and sex to mitigate the influence of confounding factors [40]. It is crucial to note that all hypothesis tests in this study were 2-tailed, with a significance level set at .05 [41]. Therefore, results were deemed statistically significant when the *P* value was <.05.

Results

Sample Characteristics and Single-Factor Analysis

This study enrolled 486 community patients diagnosed with SMDs, among whom 206 (42.4%) reported experiencing bullying by caregivers or others. Of these 206 participants, 169 (82%) had been verbally abused, and 113 (54.9%) had been physically assaulted. Moreover, of those who experienced bullying, 49.5% (102/206) were male, and 50.5% (104/206) were female. In terms of age, 40.8% (84/206) were aged <45 years, 38.3% (79/206) were aged between 45 and 64 years, and 20.9% (43/206) were aged >64 years. Regarding educational attainment, 32% (66/206) were illiterate or semiliterate, 22.8% (47/206) had a primary education, and 45.1% (93/206) had a middle school or above. This reflects a relatively low level of education among patients with SMDs, particularly in the less developed regions of China, where mental health literacy and access to psychiatric services remain limited.

The vast majority (190/206, 92.2%) of the participants were non–only children. This is consistent with the demographic profile of individuals born before or during the early stages of China's 1-child policy and may also reflect traditional extended family structures in rural Chinese settings. The mean score on the SSCI-8 was 22.31 (SD 8.94).

Of the 169 participants who experienced verbal abuse from caregivers or others, 82 (48.5%) were male, and 87 (51.5%) were female; in terms of age, 75 (44.4%) participants were aged <45 years, 65 (38.5%) were aged between 45 and 64 years, and 29 (17.1%) were aged >64 years. Educational backgrounds varied, with 33.7% (57/169) being illiterate or semiliterate, 22.5% (38/169) having a primary education, and 43.8% (74/169) having a middle school or above. The majority (158/169, 93.5%) were non–only children. The mean score on the SSCI-8 was 22.60 (SD 8.91).

Of the 113 participants who experienced physical assault by caregivers or others, 53 (46.9%) were aged <45 years, 43 (38.1%) were aged between 45 and 64 years, and 17 (15%) were aged >64 years. Regarding educational levels, 30.1% (34/113) were illiterate or semiliterate, 23% (26/113) had a primary school education, and 46.9% (53/113) had a middle school or above. The majority (103/113, 91.2%) were non–only children. The mean score on the SSCI-8 was 22.24 (SD 8.74).

As shown in Table 1, the variables related to experiencing bullying included sex (χ^2_1=5.8; *P*=.02), age (χ^2_2=13.0; *P*=.001), education (χ^2_2=8.9; *P*=.01), sibling status (χ^2_1=6.1; *P*=.01), and self-stigma ($F_{2,483}$=2.489; *P*<.001). These strata reflect social structure differences: middle-aged adults may face more isolation, only children lack sibling support, and education levels affect coping ability. Traditional Chinese family roles may also shape sibling-related caregiving dynamics. The variables associated with experiencing verbal abuse were sex (χ^2_1=4.3; *P*=.04), age (χ^2_2=18.7; *P*<.001), education (χ^2_2=7.1; *P*=.03), and self-stigma ($F_{2,483}$=2.735; *P*<.001). The variables associated with experiencing physical assault were age (χ^2_2=18.9; *P*<.001), education (χ^2_2=7.2; *P*=.03), sibling status (χ^2_1=6.6; *P*=.01), and self-stigma ($F_{2,483}$=1.878; *P*=.004). Additional information is presented in Table 1.

XSL•FO
RenderX

Table 1. Demographic and clinical characteristics of community patients with serious mental disorders (SMDs) by bullying exposure status (N=486).

Variables	Experiences of bullying				*F* test (F_{df}) or chi-square (χ^2_{df})		
	With no bullying (n=280)	Verbal bullying (n=169)	Physical bullying (n=113)	Any bullying (n=206)	Verbal bullying vs with no bullying	Physical bullying vs with no bullying	Any bullying vs with no bullying
Sex, n (%)					$\chi^2_1=4.3$[a]	$\chi^2_1=1.8$	$\chi^2_1=5.8$[a]
Male	108 (38.6)	82 (48.5)	52 (46)	102 (49.5)			
Female	172 (61.4)	87 (51.5)	61 (54.0)	104 (50.5)			
Age groups (y), n (%)					$\chi^2_2=18.7$[b]	$\chi^2_2=18.9$[b]	$\chi^2_2=13.0$[c]
<45	72 (25.7)	75 (44.4)	53 (46.9)	84 (40.8)			
45-64	125 (44.6)	65 (38.5)	43 (38.1)	79 (38.3)			
>64	83 (29.6)	29 (17.1)	17 (15)	43 (20.9)			
Educational levels, n (%)					$\chi^2_2=7.1$[a]	$\chi^2_2=7.2$[a]	$\chi^2_2=8.9$[a]
Illiterate or semiliterate	97 (34.6)	57 (33.7)	34 (30.1)	66 (32)			
Primary school	91 (32.5)	38 (22.5)	26 (23)	47 (22.8)			
Middle school or above	92 (32.9)	74 (43.8)	53 (46.9)	93 (45.1)			
Sibling status, n (%)					$\chi^2_1=3.5$	$\chi^2_1=6.6$[a]	$\chi^2_1=6.1$[a]
Single child	8 (2.9)	11 (6.5)	10 (8.8)	16 (7.8)			
Other	272 (97.1)	158 (93.5)	103 (91.2)	190 (92.2)			
Marital status, n (%)					$\chi^2_1=1.0$	$\chi^2_1=0.9$	$\chi^2_1=0.4$
Married	192 (68.6)	108 (63.9)	72 (63.7)	136 (66)			
Other	88 (31.4)	61 (36.1)	41 (36.3)	70 (34)			
Types of SMDs, n (%)					$\chi^2_1=1.4$	$\chi^2_1=1.1$	$\chi^2_1=2.3$
Schizophrenia	170 (60.7)	112 (66.3)	75 (66.4)	139 (67.5)			
Other	110 (39.3)	57 (33.7)	38 (33.6)	67 (32.5)			
BPRS[d] score, mean (SD)	52.93 (20.14)	58.91 (21.87)	57.27 (21.77)	58.92 (22.14)	$F_{2,483}=0.966$	$F_{2,483}=0.835$	$F_{2,483}=0.978$
Duration of SMD (y), mean (SD)	23.33 (13.61)	23.15 (13.30)	23.75 (12.47)	23.52 (13.16)	$F_{2,483}=1.031$	$F_{2,483}=0.895$	$F_{2,483}=1.071$
Self-stigma: SSCI-8[e] score, mean (SD)	17.82 (9.29)	22.60 (8.91)	22.24 (8.74)	22.31 (8.94)	$F_{2,483}=2.735$[b]	$F_{2,483}=1.878$[c]	$F_{2,483}=2.489$[b]
Poor sleep quality: PSQI[f] score, mean (SD)	8.08 (4.26)	8.89 (4.22)	8.60 (4.22)	8.86 (4.22)	$F_{2,483}=1.473$	$F_{2,483}=1.446$	$F_{2,483}=1.684$[a]
Family functioning: FAS[g] score, mean (SD)	7.23 (3.16)	6.36 (3.27)	5.99 (3.46)	6.44 (3.34)	$F_{2,483}=1.519$	$F_{2,483}=1.712$	$F_{2,483}=1.273$

[a]$P<.05$.

[b]$P<.001$.

[c]$P<.01$.

[d]BPRS: Brief Psychiatric Rating Scale.

[e]SSCI-8: Stigma Scale for Chronic Illnesses, 8-item version.

[f]PSQI: Pittsburgh Sleep Quality Index.

[g]FAS: Family Apgar Scale.

XSL•FO
RenderX

Logistic Regression Analysis of the Factors Influencing Different Types of Bullying Experiences

Logistic regression models were used to analyze factors related to bullying experiences. The results of the 3 logistic regression models are presented in Table 2.

Table 2. Multivariable logistic regression analysis of factors associated with distinct subtypes of bullying experiences (verbal, physical, or any bullying) among community patients with serious mental disorders (SMDs; N=486).

Variables	Any bullying experience (n=206) vs with no bullying experience (n=280), OR[a] (95% CI)	Verbal bullying experience (n=169) vs with no bullying experience (n=280), OR (95% CI)	Physical bullying experience (n=113) vs with no bullying experience (n=280), OR (95% CI)
Sex: male (reference: female)	1.483 (0.951-2.314)	1.354 (0.838-2.189)	1.108 (0.631-1.947)
Age groups (y; reference <45)			
45-64	2.795 (1.528-5.112)[b]	3.619 (1.877-6.975)[b]	4.877 (2.212-10.753)[b]
>64	1.400 (0.823-2.381)	1.756 (0.969-3.182)	2.062 (0.996-4.269)
Educational levels (reference: illiterate)			
Primary school	0.796 (0.483-1.311)	0.912 (0.534-1.558)	0.675 (0.367-1.238)
Middle school or above	0.566 (0.346-0.925)[c]	0.596 (0.350-1.013)	0.576 (0.315-1.052)
Sibling status: single child (reference: other)	2.894 (1.080-7.760)[c]	2.449 (0.844-7.105)	3.278 (1.064-10.099)[c]
Marital status: married (reference: other)	1.369 (0.843-2.224)	1.176 (0.705-1.961)	1.203 (0.666-2.173)
Type of SMD: schizophrenia (reference: other)	1.598 (1.055-2.420)[c]	1.497 (0.964-2.325)	1.584 (0.952-2.635)
BPRS[d] score	1.001 (0.991-1.012)	1.001 (0.990-1.103)	0.996 (0.983-1.010)
Duration of SMD (y)	1.011 (0.995-1.028)	1.009 (0.992-1.027)	1.020 (0.999-1.041)
Self-stigma	1.041 (1.017-1.065)[b]	1.046 (1.020-1.072)[b]	1.043 (1.014-1.073)[e]
Poor sleep quality	1.056 (1.004-1.111)[c]	1.061 (1.005-1.120)[c]	1.053 (0.988-1.122)
Family functioning	0.953 (0.897-1.012)	0.946 (0.888-1.009)	0.913 (0.850-0.980)[c]

[a]OR: odds ratio.

[b]*P*<.001.

[c]*P*<.05.

[d]BPRS: Brief Psychiatric Rating Scale.

[e]*P*<.01.

After controlling for relevant sociodemographic variables, we developed 3 models. In model 1, bullying experiences were used as the dependent variable. The results indicated that age group (OR 2.795, 95% CI 1.528-5.112; *P*<.001), education level of middle school or above (OR 0.566, 95% CI 0.346-0.925; *P*=.02), being an only child (OR 2.894, 95% CI 1.080-7.760; *P*=.04), having schizophrenia (OR 1.598, 95% CI 1.055-2.420; *P*=.03), and self-stigma (OR 1.041, 95% CI 1.017-1.065; *P*<.001) were associated with increased bullying experiences. Conversely, sleep quality (OR 1.056, 95% CI 1.004-1.111; *P*=.04) showed a negative association with bullying experiences.

In model 2, verbal abuse experiences were used as the dependent variable. The results showed that age group (OR 3.619, 95% CI 1.877-6.975; *P*<.001) and self-stigma (OR 1.046, 95% CI 1.020-1.072; *P*<.001) were associated with increased verbal abuse experiences. These findings may reflect age-related sensitivity to self-stigma and communication challenges within caregiving relationships. In Chinese society, where self-stigma toward mental illness remains pervasive, such attitudes may contribute to verbal abuse by caregivers or others. By contrast, sleep quality (OR 1.061, 95% CI 1.005-1.120; *P*=.03) was negatively correlated with experiences of verbal abuse.

In model 3, physical assault experiences were used as the dependent variable. The results revealed that age group (OR 4.877, 95% CI 2.212-10.753; *P*<.001), single-child status (OR 3.278, 95% CI 1.064-10.099; *P*=.04), self-stigma (OR 1.043, 95% CI 1.014-1.073; *P*=.003), and lower family functioning (OR 0.913, 95% CI 0.850-0.980; *P*=.01) were associated with increased experiences of physical assault. In the Chinese cultural context, family functioning plays a central role in caregiving for people with mental illness because institutional support remains limited, especially in rural areas. A dysfunctional family environment may increase the likelihood of abusive behaviors by caregivers or others.

XSL•FO
RenderX

GEE Analysis of the Factors Influencing the Number of Types of Bullying Experiences Across Different Demographics

We used the GEE method to examine the factors influencing the frequency of bullying experiences across various types. Table 3 presents the outcomes of the GEE model.

Table 3. Generalized estimating equation models examining the predictors of bullying frequency by perpetrator type (caregivers and others) among community patients with serious mental disorders (SMDs; N=486).

Variables	Any bullying experience (n=206) vs with no bullying experience (n=280), OR[a] (95% CI)	Any bullying experience by caregivers (n=153) vs with no bullying experience (n=280), OR (95% CI)	Any bullying experience by others (n=143) vs with no bullying experience (n=280), OR (95% CI)
Sex: male (reference: female)	1.143 (0.816-1.602)	1.063 (0.740-1.526)	1.042 (0.720-1.508)
Age groups (y; reference: <45)			
45-64	0.607 (0.438-0.840)[b]	0.566 (0.400-0.800)[b]	0.596 (0.428-0.831)[b]
>64	0.402 (0.276-0.585)[c]	0.361 (0.239-0.545)[c]	0.411 (0.248-0.681)[c]
Education levels (reference: illiterate)			
Primary school	0.765 (0.523-1.117)	0.694 (0.465-1.036)	1.065 (0.653-1.738)
Middle school or above	1.059 (0.726-1.544)	1.138 (0.759-1.704)	1.199 (0.784-1.833)
Sibling status: single child (reference: other)	1.965 (1.255-3.078)[b]	1.918 (1.142-3.220)[d]	1.111 (0.680-1.815)
Marital status: married (reference: other)	1.221 (0.893-1.670)	1.183 (0.848-1.651)	0.999 (0.716-1.395)
Type of SMD: schizophrenia (reference: other)	1.330 (0.977-1.812)	1.323 (0.951-1.838)	1.439 (1.037-1.997)[d]
BPRS[e]	0.996 (0.987-1.005)	0.996 (0.987-1.005)	0.999 (0.990-1.008)
Duration of SMD (y)	1.012 (1.000-1.024)	1.012 (0.999-1.026)	1.013 (1.002-1.025)[d]
Self-stigma	1.030 (1.012-1.048)[c]	1.032 (1.014-1.050)[c]	1.037 (1.017-1.057)[c]
Poor sleep quality	1.068 (1.029-1.108)[c]	1.068 (1.026- 1.112)[b]	1.023 (0.983-1.064)
Family functioning	0.945 (0.912-0.979)[b]	0.942 (0.906-0.980)[b]	0.965 (0.925-1.0064)

[a]OR: odds ratio.

[b]$P<.01$.

[c]$P<.001$.

[d]$P<.05$.

[e]BPRS: Brief Psychiatric Rating Scale.

Model 1 revealed that age groups (45-64 y: OR 0.607, 95% CI 0.438-0.840; $P=.003$; >64 y: OR 0.402, 95% CI 0.276-0.585; $P<.001$), single-child status (OR 1.965, 95% CI 1.255-3.078; $P=.003$), self-stigma (OR 1.030, 95% CI 1.012-1.048; $P<.001$), poor sleep quality (OR 1.068, 95% CI 1.029-1.108; $P<.001$), and better family functioning (OR 0.945, 95% CI 0.912-0.979; $P=.002$) as associated with lower odds of experiencing bullying by any perpetrator. The higher vulnerability of only children to repeated bullying may be associated with the caregiving burden concentrated on a single family member in the context of China's 1-child policy, which limits sibling support and often places intensified pressure on both caregivers and patients.

Model 2 used the number of types of bullying by caregivers as a predictor, with findings indicating that age group (45-64 y: OR 0.566, 95% CI 0.400-0.800; $P=.001$; >64 y: OR 0.361, 95% CI 0.239-0.545; $P<.001$), single-child status (OR 1.918, 95% CI 1.142-3.220; $P=.01$), self-stigma (OR 1.032, 95% CI 1.014-1.050; $P<.001$), poor sleep quality (OR 1.068, 95% CI 1.026-1.112; $P=.001$), and lower family functioning (OR 0.942, 95% CI 0.906-0.980; $P=.003$) was associated with higher odds of caregiver-related bullying. This suggests that in the Chinese sociocultural context, where families are often the sole caregivers for individuals with SMDs, weakened family functioning can be associated with higher odds of emotional or physical abuse from caregivers.

Model 3 used the number of types of bullying by other people as a predictor. The results demonstrated that age groups (45-64 y: OR 0.596, 95% CI 0.428-0.831; $P=.002$; >64 y: OR 0.411, 95% CI 0.248-0.681; $P<.001$), type of SMD (OR 1.439, 95% CI 1.037-1.997; $P=.03$), duration of SMD (years; OR 1.013, 95% CI 1.002-1.025; $P=.03$), and self-stigma (OR 1.037, 95% CI 1.017-1.057; $P<.001$) significantly influenced the likelihood of being bullied by others. The association between longer illness duration and external bullying may reflect persistent social exclusion and labeling common in Chinese communities, where mental illness is still widely stigmatized.

XSL•FO
RenderX

Discussion

Principal Findings

This study found a bullying prevalence rate of 42.4% (206/486) among community patients with SMDs. Self-stigma emerged as a significant factor associated with bullying experienced by individuals with SMDs, whereas family functioning seemed to play a potentially protective role. Concerning sleep quality, verbal abuse was associated with poorer sleep for individuals with SMDs, whereas physical abuse showed no such association. Notably, bullying by caregivers was more strongly associated with poorer sleep quality. Furthermore, sociodemographic factors such as age, education, only-child status, and type of SMD also showed significant associations with bullying experiences among individuals with SMDs.

The prevalence rate found in this study is lower than the rates reported in previous studies conducted in the United States (82%), Sweden (67%), and low-income rural areas of Ethiopia (60.7%) [13,42,43]. Such differences may be related to variations in social support systems, access to health care, and cultural attitudes toward mental illness. Furthermore, data from the Global Burden of Disease Study 2019 show a steady increase from 1990 to 2019 in the burden of anxiety and depression associated with exposure to bullying, with the largest increases in disability-adjusted life years observed in high-income regions [44]. These findings suggest that bullying is a widespread and persistent mental health issue globally, highlighting the need for context-specific interventions and mental health policies that address both structural vulnerabilities and social stigma.

Stigma related to mental illness includes public stigma, which involves negative stereotypes and societal discrimination, and self-stigma, in which individuals internalize these negative beliefs [16]. This study found that higher self-stigma levels were significantly associated with bullying experiences, such as verbal abuse and physical assault. Although the cross-sectional design limits causal conclusions, the literature suggests that self-stigma may both result from and be associated with increased vulnerability to abuse [45]. Self-stigma may lead patients to avoid disclosing their condition or seeking help [46], which can contribute to misunderstandings and social exclusion [47]. Public stereotypes portraying people with SMDs as dangerous or unreliable [48], combined with stigma and discrimination [49], may be related to bullying [50]. Moreover, self-stigma is associated with lower patient self-esteem and self-efficacy [51], which may lead individuals with SMDs to become more passive and endure abuse silently, potentially perpetuating the cycle of mistreatment [52]. Our findings add to the literature by showing this association between self-stigma and bullying in a Chinese community sample with SMDs. Given the strong role of family and collectivist culture in Chinese society, individuals with SMDs may experience greater internalized shame and pressure to conceal their illness, which could be related to being bullied. This suggests the importance of addressing self-stigma as part of efforts to reduce bullying and improve patient well-being.

This study adds new insight to existing literature by distinguishing the differing associations of verbal and physical abuse on sleep quality among individuals with SMDs. While both forms of abuse are harmful, only verbal abuse was significantly associated with poorer sleep. This suggests that emotional stress from verbal abuse may have a more persistent impact than the acute effects of physical harm. Previous studies have linked emotional distress to sleep disturbances [53-59], but our findings suggest that the type of bullying matters—offering a more nuanced understanding of how psychosocial stressors affect sleep in this population considered vulnerable. This distinction has not been well explored in prior research and could inform more targeted interventions.

This study provides new evidence that bullying by caregivers has a stronger association with sleep disturbances in individuals with SMDs than bullying by others. While previous research has documented the negative associations of interpersonal violence, few studies have differentiated the sources of abuse. Our findings suggest that the violation of trust within close caregiver relationships may be linked to more severe psychological distress, such as anxiety and depression, which in turn worsens sleep quality [60,61]. By contrast, bullying by outsiders, although harmful, may be buffered by support from trusted caregivers. This distinction highlights the unique vulnerability of patients with SMDs within dependent caregiving relationships and calls for greater attention to caregiver dynamics in both research and intervention design [62]. In the Chinese cultural context, family members often serve as both primary caregivers and key decision-makers in the care of individuals with mental disorders. This strong familial involvement, while supportive in many cases, may also amplify the psychological impact when abuse occurs within the family unit, due to the heightened expectations of filial piety and loyalty.

Family functioning serves as a positive factor against physical abuse for patients with SMDs. Effective family functioning is characterized by good communication, mutual understanding among family members, and the collective ability to cope with problems [63]. When family members have a better understanding of mental disorders, they are more likely to respond with patience and empathy, minimizing conflicts and misunderstandings [64-66]. This study adds to the literature by providing empirical evidence from a Chinese community-based sample, showing that better family functioning is significantly associated with lower physical abuse among individuals with SMDs. In China, where families are often the primary caregivers due to limited institutional resources, the quality of family dynamics plays a particularly critical role in the daily care and protection of individuals with mental illness. Our findings suggest that culturally tailored interventions focusing on improving family communication and mental health literacy could be especially effective in reducing violence and improving patient outcomes in the Chinese context.

This study found that middle-aged individuals (aged 45-64 y), only children, those with higher education, and patients with schizophrenia were more likely to experience bullying. The prevalence of mental disorders is higher among individuals aged 45 to 64 years, a group that often faces increased social isolation due to life transitions such as changes in employment or retirement [67,68]. Surprisingly, experiencing multiple types

XSL•FO
RenderX

of bullying was found to be a positive factor, possibly because it enhances defense mechanisms and resilience. Studies also show that individuals with middle school education or above and only children are more likely to face bullying. Despite their educational background, patients with SMDs often struggle with social adjustment [69], and only children may seem more isolated due to lack of peer support [70]. Patients with schizophrenia are especially vulnerable, likely because their symptoms are more pronounced and noticeable in social settings [71]. These results emphasize the need for culturally sensitive, community-based strategies to protect groups considered vulnerable and improve mental health outcomes.

Limitations

This study has several limitations, First, the small sample size and the fact that the sample was drawn from a specific community-based mental health management system in Shandong province may affect the stability and reliability of the results, limiting their generalizability to larger populations or to all individuals with SMDs in China. Moreover, the small subgroup sizes limited reliable stratified analyses, possibly missing spatial or socioeconomic differences. Second, the use of cross-sectional data prevents any causal inferences between variables. Third, data were collected through self-report and caregiver-assisted interviews, which may have introduced information bias, especially considering that caregivers could be sources of bullying themselves. Finally, due to data limitations, not all confounders could be included, leaving the possibility of residual confounding.

Conclusions

This study highlights that self-stigma is associated with experiencing bullying among people with SMDs, while positive family functioning acts as a protective buffer. Notably, our findings add new understanding to the literature by identifying the differing associations of different bullying types on sleep quality—verbal abuse was linked to poorer sleep, whereas physical abuse showed no such long-term effect. Moreover, we found that bullying by caregivers was more strongly associated with negative sleep outcomes than bullying by others, emphasizing the crucial role of caregiving relationships. These results contribute novel insights, particularly within the Chinese context where individuals with SMDs often live with family members and rely heavily on them for daily care. Interventions in China must therefore prioritize self-stigma reduction and strengthen family-based support to improve the mental health and well-being of this underserved population.

Acknowledgments

The authors would like to express gratitude to all participants for taking part in this study. The National Natural Science Foundation of China (grant 71974114) funded this project. It had no role in study design, data collection, data analyses, data interpretation, or the writing of this paper.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to being proprietary to the study team and not derived from a public database, but are available from the corresponding author on reasonable request.

Authors' Contributions

LS, ZW, MW, WW, and KW assisted YQ in data collection and preparation of the original draft.

Conflicts of Interest

None declared.

References

1. Chronister J, Fitzgerald S, Chou CC. The meaning of social support for persons with serious mental illness: a family member perspective. Rehabil Psychol 2021 Feb;66(1):87-101 [FREE Full text] [doi: 10.1037/rep0000369] [Medline: 33382337]
2. The National Health Commission issued a notice on the management and treatment of severe mental disorders. China Government Network. 2018. URL: https://www.gov.cn/gongbao/content/2018/content_5338247.htm [accessed 2025-08-13]
3. GBD results. Institute for Health Metrics and Evaluation. URL: https://vizhub.healthdata.org/gbd-results/ [accessed 2025-08-13]
4. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150 [FREE Full text] [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
5. Dean K, Moran P, Fahy T, Tyrer P, Leese M, Creed F, et al. Predictors of violent victimization amongst those with psychosis. Acta Psychiatr Scand 2007 Nov;116(5):345-353. [doi: 10.1111/j.1600-0447.2007.01078.x] [Medline: 17868429]
6. Christ C, de Waal MM, Kikkert MJ, Fluri DG, Beekman AT, Dekker JJ, et al. Violent victimization and revictimization in patients with depressive disorders: context characteristics, disclosure rates, and gender differences. BMC Psychiatry 2022 Jun 16;22(1):403 [FREE Full text] [doi: 10.1186/s12888-022-04045-4] [Medline: 35710391]

XSL•FO
RenderX

7. Newman JM, Turnbull A, Berman BA, Rodrigues S, Serper MR. Impact of traumatic and violent victimization experiences in individuals with schizophrenia and schizoaffective disorder. J Nerv Ment Dis 2010 Oct;198(10):708-714. [doi: 10.1097/NMD.0b013e3181f49bf1] [Medline: 20921860]
8. Lam JA, Rosenheck R. The effect of victimization on clinical outcomes of homeless persons with serious mental illness. Psychiatr Serv 1998 May;49(5):678-683. [doi: 10.1176/ps.49.5.678] [Medline: 9603576]
9. Jadambaa A, Thomas HJ, Scott JG, Graves N, Brain D, Pacella R. The contribution of bullying victimisation to the burden of anxiety and depressive disorders in Australia. Epidemiol Psychiatr Sci 2020;29:e54. [doi: 10.1017/s2045796019000489]
10. Hodgins S, Lincoln T, Mak T. Experiences of victimisation and depression are associated with community functioning among men with schizophrenia. Soc Psychiatry Psychiatr Epidemiol 2009 Jun 22;44(6):448-457. [doi: 10.1007/s00127-008-0460-8] [Medline: 19030767]
11. Jelena V, Sanja A, Nađa MB. Evaluation of bullying in persons with different risk for psychosis. Med podmladak 2015;66(2):44-49 [FREE Full text] [doi: 10.5937/medpodm1502044v]
12. El Missiry A, Meguid MA, Abourayah A, Missiry ME, Hossam M, Elkholy H, et al. Rates and profile of victimization in a sample of Egyptian patients with major mental illness. Int J Soc Psychiatry 2019 May 08;65(3):183-193. [doi: 10.1177/0020764019831315] [Medline: 30848686]
13. Karni-Vizer N, Salzer MS. Verbal violence experiences of adults with serious mental illnesses. Psychiatr Rehabil J 2016 Dec;39(4):299-304. [doi: 10.1037/prj0000214] [Medline: 27547854]
14. Fitzgerald PB, de Castella AR, Filia K, Filia S, Benitez J, Kulkarni J. Victimization of patients with schizophrenia and related disorders. Aust N Z J Psychiatry 2005 Mar;39(3):169-174. [doi: 10.1111/j.1440-1614.2005.01539.x]
15. Xu M, Macrynikola N, Waseem M, Miranda R. Racial and ethnic differences in bullying: review and implications for intervention. Aggress Violent Behav 2020 Jan;50:101340. [doi: 10.1016/j.avb.2019.101340] [Medline: 32863731]
16. Shih CA, Huang JH, Yang MH. Anti-stigma psychosocial intervention effects on reducing mental illness self-stigma and increasing self-esteem among patients with schizophrenia in Taiwan: a quasi-experiment. Asian J Psychiatr 2022 Jul;73:103171. [doi: 10.1016/j.ajp.2022.103171] [Medline: 35598427]
17. Šimunović Filipčić I, Grassi L, Filipčić I. Severe mental disorders and cancer. In: Fiorillo A, Javed A, Sartorius N, editors. Comorbidity between Mental and Physical Disorders. Cham, Switzerland: Springer; 2025:203-220.
18. Wei Z, Wang Y, Hu L, Wang Y, Li C, Sun L. Incidence, prevalence, and mortality of schizophrenia from 2016 to 2020 in Shandong, China. Psychiatry Res 2024 Jan;331:115612. [doi: 10.1016/j.psychres.2023.115612] [Medline: 38039652]
19. Zuo X, Hong J, Liang D, Cao W, Liang Q, Liu Y. Analysis of factors influencing contact in the 686 Program for patients with a first-time diagnosis of severe mental disorders in Hainan Province from 2018 to 2022. BMC Psychiatry 2025 Mar 11;25(1):233 [FREE Full text] [doi: 10.1186/s12888-025-06627-4] [Medline: 40069708]
20. Fan Y, Ma N, Ouyang A, Zhang W, He M, Chen Y, et al. Effectiveness of a community-based peer support service among persons suffering severe mental illness in China. PeerJ 2022;10:e14091 [FREE Full text] [doi: 10.7717/peerj.14091] [Medline: 36248709]
21. Huang ZH, Wang F, Chen ZL, Xiao YN, Wang QW, Wang SB, et al. Risk factors for violent behaviors in patients with schizophrenia: 2-year follow-up study in primary mental health care in China. Front Psychiatry 2022;13:947987 [FREE Full text] [doi: 10.3389/fpsyt.2022.947987] [Medline: 36741582]
22. Sun L, Zhang Y, Cui L, Li J, Li L, Sun X, et al. Lifetime and 1-month prevalence of mental disorders in Hebei province, China: a community-based cross-sectional study. Front Public Health 2021;9:759160 [FREE Full text] [doi: 10.3389/fpubh.2021.759160] [Medline: 34722453]
23. Rao WW, Zhang YS, Ng CH, Cui LJ, Li JF, Li L, et al. Prevalence of schizophrenia and its association with socio-demographic correlates in an agricultural region of China. Asian J Psychiatr 2021 Oct;64:102743. [doi: 10.1016/j.ajp.2021.102743] [Medline: 34400109]
24. Yu YH, Peng MM. Development and poverty dynamics in severe mental illness: a modified capability approach in the Chinese context. Int J Environ Res Public Health 2022 Feb 18;19(4):2351 [FREE Full text] [doi: 10.3390/ijerph19042351] [Medline: 35206538]
25. Wiemann AK, Werner A, Konrad K, Niestroj SC, Steden S, Lohaus A. Lifetime poly-victimization and later bullying victimization: associations with internalizing problems and out-of-home care. Child Abuse Negl 2023 Jan;135:105970. [doi: 10.1016/j.chiabu.2022.105970] [Medline: 36434979]
26. Chen M, Lin GR, Wang GY, Yang L, Lyu N, Qian C, et al. Stigma toward mental disorders and associated factors among community mental health workers in Wuhan, China. Asia Pac Psychiatry 2023 Jun;15(2-3):e12542. [doi: 10.1111/appy.12542] [Medline: 37517868]
27. Li S, Ji Y, Yang Q, Jiang Y, Yin Q, Zhu H. Influencing factors of the participation of patients with severe mental disorders in community health service management. Front Public Health 2025;13:1424399 [FREE Full text] [doi: 10.3389/fpubh.2025.1424399] [Medline: 39896792]
28. Li S, Yin J, Yang Q, Ji Y, Zhu H, Yin Q. Construction of a troublemaking risk assessment tool for patients with severe mental disorders in community of China. Sci Rep 2025 Jan 03;15(1):663 [FREE Full text] [doi: 10.1038/s41598-024-84486-x] [Medline: 39753785]

XSL•FO
RenderX

29. Lu CH, Zhong JH, Lv YD, Luo JL, Cheng J, Qing L, et al. Participation in a free medication program among non-registered permanent residents with severe mental disorders in Bao'an district, Shenzhen, China and its influencing factors. BMC Psychiatry 2021 Sep 30;21(1):479 [FREE Full text] [doi: 10.1186/s12888-021-03485-8] [Medline: 34592974]
30. Perrotta G. Psychotic spectrum disorders: definitions, classifications, neural correlates and clinical profiles. Ann Psychiatr Treat 2020 Dec 31;4(1):070-084 [FREE Full text] [doi: 10.17352/apt.000023]
31. Lachar D, Bailley SE, Rhoades HM, Varner RV. Use of BPRS-A percent change scores to identify significant clinical improvement: accuracy of treatment response classification in acute psychiatric inpatients. Psychiatry Res 1999 Dec 27;89(3):259-268. [doi: 10.1016/s0165-1781(99)00114-6] [Medline: 10708272]
32. Molina Y, Choi SW, Cella D, Rao D. The stigma scale for chronic illnesses 8-item version (SSCI-8): development, validation and use across neurological conditions. Int J Behav Med 2013 Sep 26;20(3):450-460 [FREE Full text] [doi: 10.1007/s12529-012-9243-4] [Medline: 22639392]
33. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
34. Liu X, Tang M, Hu L. Reliability and validity of the Pittsburgh Sleep Quality Index. Chin J Psychiatry 1996;29:103-107 [FREE Full text]
35. Smilkstein G. The family APGAR: a proposal for a family function test and its use by physicians. J Fam Pract 1978 Jun;6(6):1231-1239. [Medline: 660126]
36. Rodríguez Fernández E, Gea Serrano A, Gómez Moraga A, García González JM. [Apgar questionnaire in the study of family function]. Aten Primaria 1996 Mar 31;17(5):338-341. [Medline: 8722159]
37. Chau TT, Hsiao TM, Huang CT, Liu HW. [A preliminary study of family Apgar index in the Chinese]. Gaoxiong Yi Xue Ke Xue Za Zhi 1991 Jan;7(1):27-31. [Medline: 1990150]
38. Meng Y, Yang Y, Lin P, Xiao Y, Sun Y, Qian Y, et al. School bullying victimization and associated factors among school-aged adolescents in China. J Interpers Violence 2023 Jan;38(1-2):NP1787-NP1814. [doi: 10.1177/08862605221092074] [Medline: 35475706]
39. Lee G, Kim C. Exploring the association between loneliness, work environment, and depressive symptoms: evidence from young Korean workers in the Seoul Metropolitan Area. Front Public Health 2025;13:1593957 [FREE Full text] [doi: 10.3389/fpubh.2025.1593957] [Medline: 40520319]
40. Xue J, Hu R, Chai L, Han Z, Sun IY. Examining the prevalence and risk factors of school bullying perpetration among Chinese children and adolescents. Front Psychol 2022;13:720149 [FREE Full text] [doi: 10.3389/fpsyg.2022.720149] [Medline: 35369167]
41. Chen P. Inner child of the past: long-term protective role of childhood relationships with mothers and fathers and maternal support for mental health in middle and late adulthood. Soc Psychiatry Psychiatr Epidemiol 2022 Jul;57(7):1399-1416. [doi: 10.1007/s00127-021-02200-y] [Medline: 34842965]
42. Bengtsson-Tops A, Ehliasson K. Victimization in individuals suffering from psychosis: a Swedish cross-sectional study. J Psychiatr Ment Health Nurs 2012 Feb 22;19(1):23-30. [doi: 10.1111/j.1365-2850.2011.01749.x] [Medline: 22070224]
43. Tsigebrhan R, Shibre T, Medhin G, Fekadu A, Hanlon C. Violence and violent victimization in people with severe mental illness in a rural low-income country setting: a comparative cross-sectional community study. Schizophr Res 2014 Jan;152(1):275-282 [FREE Full text] [doi: 10.1016/j.schres.2013.10.032] [Medline: 24275579]
44. Hong C, Liu Z, Gao L, Jin Y, Shi J, Liang R, et al. Global trends and regional differences in the burden of anxiety disorders and major depressive disorder attributed to bullying victimisation in 204 countries and territories, 1999-2019: an analysis of the Global Burden of Disease Study. Epidemiol Psychiatr Sci 2022 Nov 28;31:e85 [FREE Full text] [doi: 10.1017/S2045796022000683] [Medline: 36440549]
45. Camp DL, Finlay WM, Lyons E. Is low self-esteem an inevitable consequence of stigma? An example from women with chronic mental health problems. Soc Sci Med 2002 Sep;55(5):823-834. [doi: 10.1016/s0277-9536(01)00205-2] [Medline: 12190273]
46. Clement S, Schauman O, Graham T, Maggioni F, Evans-Lacko S, Bezborodovs N, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2014 Feb 21;45(1):11-27. [doi: 10.1017/s0033291714000129]
47. Kurzban R, Leary MR. Evolutionary origins of stigmatization: the functions of social exclusion. Psychol Bull 2001;127(2):187-208. [doi: 10.1037//0033-2909.127.2.187]
48. da Silva AG, Baldaçara L, Cavalcante DA, Fasanella NA, Palha AP. The impact of mental illness stigma on psychiatric emergencies. Front Psychiatry 2020 Jun 19;11:573 [FREE Full text] [doi: 10.3389/fpsyt.2020.00573] [Medline: 32636773]
49. Colizzi M, Ruggeri M, Lasalvia A. Should we be concerned about stigma and discrimination in people at risk for psychosis? A systematic review. Psychol Med 2020 Feb 17;50(5):705-726. [doi: 10.1017/s0033291720000148]
50. Harris MN, Baumann ML, Teasdale B, Link BG. Estimating the relationship between perceived stigma and victimization of people with mental illness. J Interpers Violence 2022 Feb 09;37(3-4):NP1226-NP1252. [doi: 10.1177/0886260520926326] [Medline: 32517566]

XSL•FO
RenderX

51. Watson AC, Corrigan P, Larson JE, Sells M. Self-stigma in people with mental illness. Schizophr Bull 2007 Nov 27;33(6):1312-1318 [FREE Full text] [doi: 10.1093/schbul/sbl076] [Medline: 17255118]
52. Kleim B, Vauth R, Adam G, Stieglitz RD, Hayward P, Corrigan P. Perceived stigma predicts low self-efficacy and poor coping in schizophrenia. J Ment Health 2009 Jul 06;17(5):482-491. [doi: 10.1080/09638230701506283]
53. Freeman D, Sheaves B, Waite F, Harvey AG, Harrison PJ. Sleep disturbance and psychiatric disorders. Lancet Psychiatry 2020 Jul;7(7):628-637. [doi: 10.1016/s2215-0366(20)30136-x]
54. Baglioni C, Spiegelhalder K, Lombardo C, Riemann D. Sleep and emotions: a focus on insomnia. Sleep Med Rev 2010 Aug;14(4):227-238. [doi: 10.1016/j.smrv.2009.10.007] [Medline: 20137989]
55. Grandey AA, Kern JH, Frone MR. Verbal abuse from outsiders versus insiders: comparing frequency, impact on emotional exhaustion, and the role of emotional labor. J Occup Health Psychol 2007 Jan;12(1):63-79. [doi: 10.1037/1076-8998.12.1.63] [Medline: 17257067]
56. Vanitallie TB. Stress: a risk factor for serious illness. Metabolism 2002 Jun;51(6 Suppl 1):40-45. [doi: 10.1053/meta.2002.33191] [Medline: 12040540]
57. Merrill RM. Mental health conditions according to stress and sleep disorders. Int J Environ Res Public Health 2022 Jun 29;19(13):7957 [FREE Full text] [doi: 10.3390/ijerph19137957] [Medline: 35805615]
58. Chapman CR, Tuckett RP, Song CW. Pain and stress in a systems perspective: reciprocal neural, endocrine, and immune interactions. J Pain 2008 Feb;9(2):122-145 [FREE Full text] [doi: 10.1016/j.jpain.2007.09.006] [Medline: 18088561]
59. Wastell C. Understanding Trauma and Emotion: Dealing with Trauma Using an Emotion-Focused Approach. Milton Park, UK: Taylor & Francis; 2020.
60. Zegwaard MI, Aartsen MJ, Grypdonck MH, Cuijpers P. Trust: an essential condition in the application of a caregiver support intervention in nursing practice. BMC Psychiatry 2017 Feb 02;17(1):47 [FREE Full text] [doi: 10.1186/s12888-017-1209-2] [Medline: 28148235]
61. Örsal Ö, Eren HK, Duru P. Examination of factors affecting the sleep quality of psychiatry patients using structural equation model. J Psychiatric Nurs 2019;10(1):55-64 [FREE Full text]
62. Wahlbeck K. Public mental health: the time is ripe for translation of evidence into practice. World Psychiatry 2015 Feb 05;14(1):36-42 [FREE Full text] [doi: 10.1002/wps.20178] [Medline: 25655149]
63. David HP. Healthy family functioning: a cross-cultural appraisal. Bull World Health Organ 1978;56(3):327-342 [FREE Full text] [Medline: 308401]
64. Saunders JC. Families living with severe mental illness: a literature review. Issues Ment Health Nurs 2003 Mar 09;24(2):175-198. [doi: 10.1080/01612840305301] [Medline: 12554427]
65. Sin J, Gillard S, Spain D, Cornelius V, Chen T, Henderson C. Effectiveness of psychoeducational interventions for family carers of people with psychosis: a systematic review and meta-analysis. Clin Psychol Rev 2017 Aug;56:13-24 [FREE Full text] [doi: 10.1016/j.cpr.2017.05.002] [Medline: 28578249]
66. Wright ER, Gronfein WP, Owens TJ. Deinstitutionalization, social rejection, and the self-esteem of former mental patients. J Health Soc Behav 2000 Mar;41(1):68-90. [Medline: 10750323]
67. Petrova NN, Khvostikova DA. Prevalence, structure, and risk factors for mental disorders in older people. Adv Gerontol 2021 Dec 09;11(4):409-415. [doi: 10.1134/s2079057021040093]
68. Mattila VJ, Joukamaa MI, Salokangas RK. Retirement, aging, psychosocial adaptation and mental health. Findings of the TURVA project. Acta Psychiatr Scand 1989 Oct 23;80(4):356-367. [doi: 10.1111/j.1600-0447.1989.tb02992.x] [Medline: 2589093]
69. Moran P, Romaniuk H, Coffey C, Chanen A, Degenhardt L, Borschmann R, et al. The influence of personality disorder on the future mental health and social adjustment of young adults: a population-based, longitudinal cohort study. Lancet Psychiatry 2016 Jul;3(7):636-645. [doi: 10.1016/s2215-0366(16)30029-3]
70. Davidson L, Chinman M, Kloos B, Weingarten R, Stayner D, Tebes JK. Peer support among individuals with severe mental illness: a review of the evidence. Clin Psychol Sci Pract 1999;6(2):165-187. [doi: 10.1093/clipsy.6.2.165]
71. Latalova K, Kamaradova D, Prasko J. Violent victimization of adult patients with severe mental illness: a systematic review. Neuropsychiatr Dis Treat 2014 Oct;10:1925-1939. [doi: 10.2147/ndt.s68321]

Abbreviations

GEE: generalized estimating equation
OR: odds ratio
SMD: serious mental disorder
SSCI-8: Stigma Scale for Chronic Illnesses, 8-item version

XSL•FO
RenderX

Edited by A Mavragani; submitted 27.03.25; peer-reviewed by R Chen, J Wang; comments to author 01.07.25; revised version received 03.07.25; accepted 07.08.25; published 03.09.25.

Please cite as:
Qi Y, Wang M, Wei Z, Wang W, Wang K, Sun L
Prevalence and Associated Factors of Bullying Experienced by Community Patients With Serious Mental Disorders: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e75027
URL: https://publichealth.jmir.org/2025/1/e75027
doi:10.2196/75027
PMID:

©Yazhuo Qi, Meiqi Wang, Zhen Wei, Wenyu Wang, Kaixian Wang, Long Sun. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 03.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Association Between Cultural Tightness and COVID-19 Vaccine Confidence From 28 Countries: Cross-Sectional Study

Qiang Wang[1,2], PhD; Ana Bolio[2], MD, MScPH; Leesa Lin[1,2,3], PhD

[1]Laboratory of Data Discovery for Health (D24H), Hong Kong Science Park, Hong Kong Special Administrative Region, China (Hong Kong)

[2]Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London, United Kingdom

[3]WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong Special Administrative Region, China (Hong Kong)

Corresponding Author:
Leesa Lin, PhD
Laboratory of Data Discovery for Health (D24H), Hong Kong Science Park, Hong Kong Special Administrative Region, China (Hong Kong)

Abstract

Background: Social norms provided a framework for understanding a variety of behaviors. Cultural tightness was introduced to measure the level of adherence to social norms and tolerance of deviant behavior.

Objective: We aimed to explore the association between cultural tightness and COVID-19 vaccine hesitancy.

Methods: A total of 44,339 participants aged 18 years and older were enrolled from 28 different countries between 2020 and 2022. We used the Vaccine Confidence Index (3 items related to evaluation of importance, effectiveness, and safety) with a 5-point Likert scale to collect COVID-19 vaccine confidence. Demographic information at the individual-level was obtained through the survey, while national-level data were sourced from the World Bank and Hofstede insights. Multilevel linear regressions with random effects for country were used to examine the association between cultural tightness and COVID-19 vaccine confidence.

Results: Of the participants, 21,968 (50.2%) were male and 18,957 (43.3%) had an education level of university or above. Vietnam exhibited the highest level of confidence (mean 13.31, SD 1.71) on COVID-19 vaccine and Slovakia had the lowest level (mean 9.52, SD 0.14). The higher levels of cultural tightness were positively linked to greater vaccine confidence (β=1.94, 95% CI 1.72-2.15; *P*<.001) after controlling individual- and national- level variables. Individuals who were younger in age, female, had lower levels of educational level, or belonged to minority religious groups demonstrated a positive association with lower vaccine confidence. Hofstede's 5 cultural dimensions were not significantly associated with vaccine confidence. The level of vaccine confidence in 2021 (β=−0.54, 95% CI −0.67 to −0.37; *P*<.001) and 2022 (β=−0.23, 95% CI −0.34 to −0.10; *P*<.001) was lower than that observed in 2020.

Conclusions: Lower level of cultural tightness might be positively associated with low vaccine confidence. Our findings offered the insight for designing tailor interventions to vaccine hesitancy in different cultural tightness context.

(JMIR Public Health Surveill 2025;11:e66872) doi:10.2196/66872

KEYWORDS

COVID-19; vaccine confidence; social norms; global; cultural tightness

Introduction

According to the World Health Organization (WHO), vaccine hesitancy refers to the "delay in acceptance or refusal of vaccination despite availability of vaccination services" [1], which could lead to a decrease in vaccination uptake and the re-emergence of vaccine-preventable diseases, such as measles [2]. The WHO has identified vaccine hesitancy as one of the ten greatest threats to global health [3], and COVID-19 vaccine hesitancy has attracted a lot of attention [4,5].

The acceptance of COVID-19 vaccination varied across countries: among 23 countries in a global survey, the COVID-19 vaccine acceptance rate ranged between 51.6% (Russia) and 97.6% (China) [5]. Much scholarly effort has been directed toward exploring the driving factors behind COVID-19 vaccine hesitancy [5,6], which often focused on the scale of individual-level factors, such as sociodemographics, that could not fully explain the wide variations observed among countries. Some researchers found that countries with higher levels of trust in government [4] and science [7] tended to have higher COVID-19 vaccine acceptance. According to Lin et al [8], there was a negative association between COVID-19 vaccine confidence and a higher misinformation belief level across 14 countries.

Social norms are characterized as beliefs that are collectively shared regarding what is considered typical or concerning what

XSL•FO
RenderX

is expected behavior within a group, which guide individuals' behaviors. It provided a framework for understanding a variety of behaviors, including encouraging energy conservation [9], charitable donations [10], and health decisions [11]. Cultural tightness was introduced by Gelfand et al [12] to measure the degree of adherence to norms within a country (or a state within a country) as well as the level of tolerance for individuals who deviate from these norms. The societies characterized by strong norms and a low tolerance of deviant behavior could be classified as exhibiting high cultural tightness, whereas those with weak norms and a high tolerance of deviant behavior could be classified as exhibiting low cultural tightness [12].

The high ecological validity of cultural tightness has been observed across a variety of domains, including politics (leadership preference), economics (stock price synchrony), and culture (prejudice) [13]. The cultural tightness was also associated with countries' response to the pandemic. Countries with high cultural tightness were more likely to show lower number of COVID-19 cases and COVID-19–related mortality rates than countries with low cultural tightness [13]. The association between vaccine confidence and cultural tightness across countries has been examined. A previous study found a negative correlation between cultural tightness and willingness to receive COVID-19 vaccination [14]. In a study conducted by Lu [15], cultural tightness did not significantly correlate with COVID-19 vaccine hesitancy across 22 countries. In a study by Shi et al [16], the positive correlation between cultural tightness and vaccination willingness became negative considering variables related to individual and collective norms in 8 Asian countries. The inconsistent findings urged the further studies.

This study aimed to examine the association between COVID-19 vaccine confidence and cultural tightness on a global scale. To achieve this goal, we conducted a multilevel analysis of 43,744 individuals from 28 countries, adjusting for numerous individual and national variables. The findings could help to understand causes of heterogeneity in vaccine confidence across countries and provide insights on developing tailored and effective interventions to deal with vaccine hesitancy.

Methods

Data

The Vaccine Confidence Project (VCP) is a research group dedicated to monitoring, analyzing, and mitigating public concerns regarding vaccine, which used a range of methods—including surveys, social listening, and artificial intelligence tool. The VCP has conducted multiple surveys about COVID-19 vaccine confidence covering more than 70 countries and territories with participants aged>18 years [17-19]. A total of 44,339 individuals were included from 28 countries, which provided the available data about cultural tightness and COVID-19 vaccine confidence. The sample sizes varied across the countries surveyed, with the smallest sample being 1000 participants (in Czechia, Estonia, Finland, Greece, Hungary, Netherlands, Poland, Portugal, Slovak Republic, Spain, Sweden, United Kingdom, Ecuador, Peru, and Brazil) and the largest being 4358 participants (in Nigeria). The mean sample size across all countries was 1584 individuals (Figure S1 in Multimedia Appendix 1). A total of 24 countries were surveyed in 2020, 4 in 2021, and 6 in 2022. We provided the detailed field date in Table S1 in Multimedia Appendix 1. South Korea was surveyed in all 3 years, while Nigeria and Kenya were surveyed in 2020 and 2022, and Vietnam, Japan, and Malaysia were surveyed in 2021 and 2022. We performed the survey in collaboration with Opinion Research Business International [20]. Fieldwork was performed through face-to-face interviews, telephone interviews, and online. Samples were selected at random and stratified to align with the proportions of demographic distributions for age, sex, and subnational region in each country. Before participating in the surveys, all respondents provided informed consent.

Exposure

We obtained national-level cultural tightness scores from a previous study [21]. This study, which included 22,863 participants across 57 countries, measured cultural tightness using a 6-item scale (Table S2 in Multimedia Appendix 1) [21]. The responses from surveyed individuals within each country were then analyzed to derive a national-level cultural tightness score. To avoid confusion between positive and negative descriptions in the original concept, we adopted the term "cultural tightness," with higher values indicating tighter culture. Among the 28 countries evaluated, Hungary exhibited the lowest cultural tightness, while India demonstrated the highest level.

Outcome

We used 3 items from the Vaccine Confidence Index (VCI) to evaluate participants' level of confidence (COVID-19 vaccines are important, COVID-19 vaccines are effective, and COVID-19 vaccines are safe) [22]. Answers were measured on a 5-point Likert scale, ranging from "strongly agree=5 points" to "strongly disagree=1 point." We added up the scores of the 3 items to construct a composite measurement of overall COVID-19 vaccine confidence. Furthermore, the surveyed items related to importance, effectiveness, and safety of vaccine were also analyzed as outcomes separately.

Covariates

Participants' sociodemographic variables were considered in the analysis, including age group (18 - 24 y vs 25 - 34 y vs 35 - 44 y vs 45 - 54 y vs ≥55 y), sex (male vs female), education level (primary or below vs secondary vs university or above vs other educational level), and religion (largest vs minority vs refusal to answer). Regarding religion, the most commonly reported religious beliefs within a country were classified as the "largest," while all other religious beliefs were grouped as "minority." Respondents who declined to answer were considered as "refusal to answer." Individual perception towards vaccine was surveyed and analyzed as a potential confounder by the statement "Vaccines are important for people of all ages to have" on a 5-point Likert scale. We considered an interaction between individual perception and cultural tightness, given that the association between cultural tightness and vaccine confidence may vary according to different levels of individual perception. We adjusted for the potential impact of national-level factors by including gross national income (GNI) in 2021,

XSL•FO
RenderX

converted by the Atlas method, population density (number of people per square km), and survey year (ie, 2020 vs 2021 vs 2022; see Figure S2 in Multimedia Appendix 1). These indicators were derived from the World Bank. Given the potential response biases caused by multiple data collection methods, the survey methodology was incorporated as a covariate in the regression model. In addition, based on a previous study [15], we controlled some cultural dimensions as defined by Hofstede [23]. The use of cultural dimensions theory has served as a valuable tool for comprehending cross-country cultural variations. We included individualism (indicating the degree of emphasis on individual interests over collective or group interests), power distance (reflecting the extent of acceptance and anticipation of unequal power and authority distribution among society's members), long-term orientation (pertaining to the degree of emphasis on future-oriented values), uncertainty avoidance (the extent to which individuals attempt to cope with anxiety by minimizing uncertainty), and indulgence (pertaining to the extent of allowance for and enjoyment of basic human desires and pleasures). The motivation toward achievement and success dimension were excluded because theoretical considerations suggest it was less likely to significantly influence the association between cultural tightness and vaccine confidence.

Statistical Analysis

The percentage of missing data for individual variables was below 0.9% across all countries (Table S3 in Multimedia Appendix 1). Since the percentage of missing data was low, cases with at least one missing variable were excluded from our analysis. Multilevel linear regressions with random effects for country (which accounts for the clustering within country) were used to examine the association between cultural tightness and COVID-19 vaccine confidence. Univariate analysis (model 1) was performed to examine the relationship between each individual- and-country-level factor and vaccine confidence with clustering by country. We controlled the individual-level variables in the model 2, and all individual- and-country-level variables in the model 3. Variance inflations factors (VIFs) were used to quantify the level of multicollinearity; the VIF for variables below 5 indicated there was only a moderate level of multicollinearity. All country-level variables were standardized in the model. We used 2-sided *P* values and statistical significance was set at $P<.05$. All analyses were performed using R software (R Foundation for Statistical Computing).

Ethical Considerations

The data collection across the 28 countries included in this study received ethics approval from the University of Hong Kong (EA230420) and the London School of Hygiene & Tropical Medicine (26636; 22130). All surveys were conducted in accordance with local regulations and institutional ethical requirements, and informed consent was obtained from participants at the time of data collection. For the current analysis, which used an aggregated and deidentified dataset, additional ethical approval was not required in accordance with VCP policies governing the use of anonymized secondary data. The original informed consents explicitly included provisions for future research use, including secondary analyses.

Results

A total of 43,744 participants were included (Table 1). Of these participants, 21,968 (50.2%) were male and 4351 (9.9%) had an education level of primary education or below. The mean COVID-19 vaccine confidence was higher among males compared with females (mean 11.95, SD 3.13 vs mean 11.51, SD 3.31; see Figure 1). Vietnam had the highest mean vaccine confidence at 13.31 (SD 1.71), followed by China at 13.08 (1.71). Conversely, Slovakia had the lowest level at 9.52 (SD 0.14), followed by Hungary with the second-lowest level at 9.71 (SD 0.13). India had the largest proportions of individuals who strongly agreed with the 3 items regarding confidence (1004/1341, 74.9% on importance; 761/1341, 56.7% on safety; and 788/1341, 58.8% on effectiveness), and Hungary had the largest proportions of individuals who strongly disagreed with the 3 items (213/999, 21.3% on importance; 221/999, 22.1% on safety; and 277/999, 27.7% on effectiveness; Figures S3-S5 in Multimedia Appendix 1). Cultural tightness was positively associated with mean vaccine confidence (Spearman correlation coefficient, $\rho=0.51$; $P=.005$; n=28) without controls. Figure 2 illustrated the variation in vaccine confidence across different levels of cultural tightness.

Our study identified significant correlations between cultural tightness and COVID-19 vaccine confidence. In Table 2, we observed that higher levels of cultural tightness were positively linked to greater vaccine confidence (β=1.94, 95% CI 1.72-2.15; $P<.001$) after adjusting for individual and country-level variables in model 3. There was 1-unit increase in standardized cultural tightness being associated with an average increase of 1.94 units in vaccine confidence score. Individuals aged ≥55 years were more likely to have higher vaccine confidence compared with those aged 18 - 24 years (β=0.35, 95% CI 0.27-0.43; $P<.001$). Female participants had lower vaccine confidence than male participants (β=−0.29, 95% CI −0.34 to −0.25; $P<.001$). Participants with an education level of university or above were more likely to have higher vaccine confidence than those with a primary education level or below (β=0.16, 95% CI 0.07-0.25; $P<.001$). Individuals with minority religious beliefs exhibited lower vaccine confidence (β=−0.13, 95% CI −0.18 to −0.08; $P<.001$) versus those with largest religious beliefs. The level of vaccine confidence in 2021 and 2022 was lower than that observed in 2020. All variables have VIFs lower than 5 (Table S4 in Multimedia Appendix 1).

Increased agreement on the importance, effectiveness, and safety of the COVID-19 vaccine was positively correlated with a higher level of cultural tightness (Table 3). Female participants demonstrated a lower likelihood of agreeing on these 3 items compared with male participants. A positive association was identified between a higher GNI and greater agreement on the importance of the COVID-19 vaccine (β=0.11, 95% CI 0.04-0.19; $P=.03$). Over time, there was a decreasing trend in agreements on the importance and effectiveness of the COVID-19 vaccine.

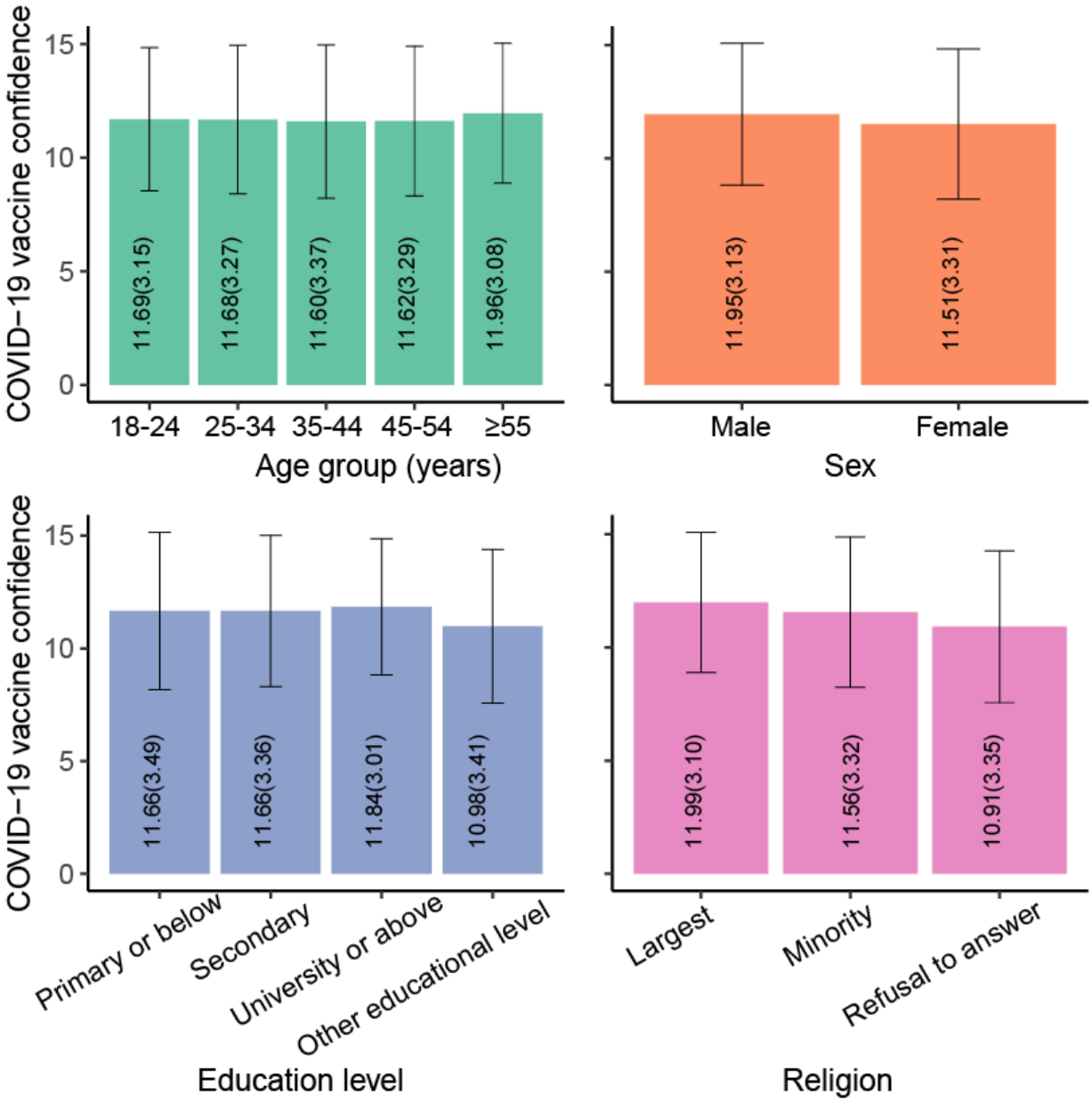
XSL•FO
RenderX

Table . Sample characteristics.

Variables	Results (N=43,744)
Age group (years), n (%)	
18 - 24	6900 (15.8)
25 - 34	9899 (22.6)
35 - 44	8561 (19.6)
45 - 54	7086 (16.2)
≥55	11,298 (25.8)
Sex, n (%)	
Male	21,968 (50.2)
Female	21,776 (49.8)
Education level, n (%)	
Primary or below	4351 (9.9)
Secondary	19,841 (45.4)
University or above	18,957 (43.3)
Other educational level	595 (1.4)
Religion, n (%)	
Largest	21,680 (49.6)
Minority	19,264 (44)
Refusal to answer	2800 (6.4)
Individual perceptions, mean (SE)	4.17 (1.14)

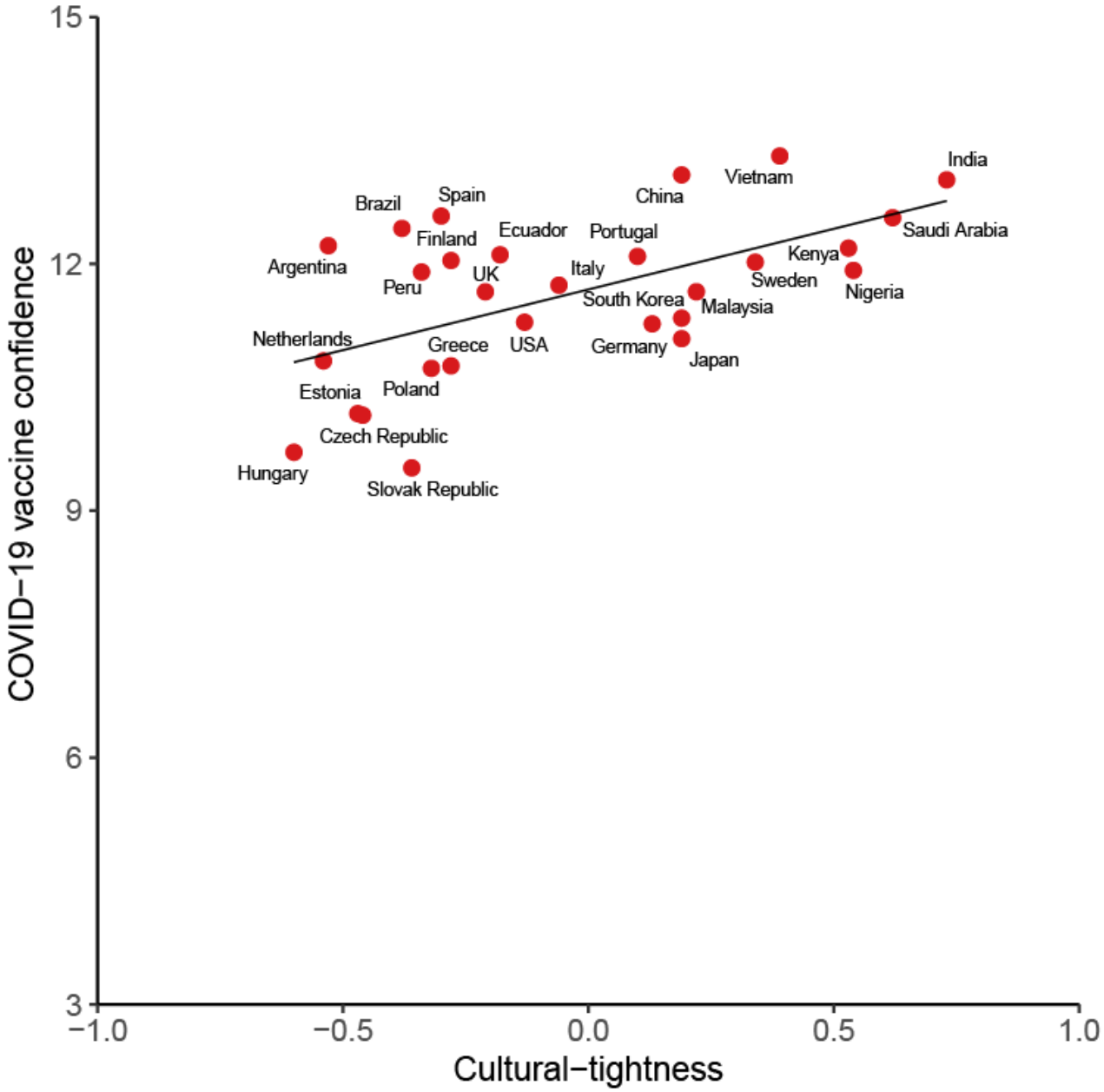
XSL•FO
RenderX

Figure 1. Mean COVID-19 vaccine confidence with SD by various sociodemographics among 43,744 individuals.



XSL•FO
RenderX

Figure 2. Scatter plot of cultural tightness and mean COVID-19 vaccine confidence. The x-axis shows the cultural tightness score and the y-axis represents the mean COVID-19 vaccine confidence. A fitted regression line was included to illustrate the trend without controls.



XSL•FO
RenderX

Table . Multilevel logistic regression analyses for associations of factors and COVID-19 vaccine confidence.

Variables	Model 1 (crude)		Model 2 (variables in the individual-level)		Model 3 (variables in the individual-and country-level)	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Cultural tightness	0.56 (0.26 to 0.85)	.001	2.13 (1.95 to 2.30)	<.001	1.94 (1.72 to 2.15)	<.001
Individual level						
Age group, years (18 - 24 years as reference)						
25 - 34	−0.06 (−0.16 to 0.04)	.22	−0.10 (−0.18 to −0.02)	.02	−0.10 (−0.18 to −0.02)	.02
35 - 44	0.00 (−0.10 to 0.10)	.99	−0.05 (−0.13 to 0.03)	.24	−0.05 (−0.13 to 0.03)	.24
45 - 54	0.16 (0.05 to 0.26)	.003	0.03 (−0.06 to 0.12)	.51	0.03 (−0.06 to 0.12)	.52
≥55	0.76 (0.66 to 0.86)	<.001	0.34 (0.26 to 0.42)	<.001	0.35 (0.27 to 0.43)	<.001
Sex (male as reference)						
Female	−0.39 (−0.45 to −0.33)	<.001	−0.29 (−0. 34 to −0.25)	<.001	−0.29 (−0.34 to −0.25)	<.001
Education level (primary or below as reference)						
Secondary	0.08 (−0.02 to 0.19)	.13	−0.03 (−0.12 to 0.06)	.49	−0.03 (−0.11 to 0.06)	.55
University or above	0.31 (0.20 to 0.41)	<.001	0.14 (0.05 to 0.23)	.002	0.16 (0.07 to 0.25)	<.001
Other educational level	−0.48 (−0.75 to −0.22)	<.001	−0.43 (−0.65 to −0.21)	<.001	−0.42 (−0.64 to −0.20)	<.001
Religion (largest as reference)						
Minority	0.34 (−0.40 to −0.27)	<.001	−0.14 (−0.20 to −0.09)	<.001	−0.13 (−0.19 to −0.08)	<.001
Refusal to answer	−0.84 (−0.97 to −0.71)	<.001	−0.24 (−0.35 to −0.13)	<.001	−0.24 (−0.35 to −0.14)	<.001
Individual perception	1.53 (1.50 to 1.55)	<.001	1.48 (1.46 to 1.50)	<.001	1.48 (1.46 to 1.50)	<.001
Individual perception and cultural tightness	__[a]	—	−0.42 (−0.44 to −0.39）	<.001	−0.42 (−0.44 to −0.40)	<.001
Country level						
Individualism	−0.43 (−0.77 to −0.08)	.02	—	—	−0.28 (−0.52 to −0.04)	.08
Indulgence	0.02 (−0.39 to 0.44)	.92	—	—	0.046 (−0.23 to 0.14)	.70
Long-term orientation	−0.39 (−0.78 to 0.01)	.07	—	—	0.03 (−0.23 to 0.29)	.84
Power distance	0.16 (−0.20 to 0.52)	.40	—	—	0.00 (−0.21 to 0.21)	.98
Uncertainty avoidance	−0.28 (−0.62 to 0.06)	.12	—	—	−0.17 (−0.36 to −0.01)	.14
GNI[b]	0.16 (−0.28 to 0.59)	.49	—	—	0.19 (−0.01 to 0.39)	.13
Population density	0.02 (−0.36 to 0.40)	.91	—	—	0.13 (−0.07 to 0.32)	.30
Survey time (2020 as reference)						

XSL•FO
RenderX

Variables	Model 1 (crude)		Model 2 (variables in the individual-level)		Model 3 (variables in the individual-and country-level)	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
2021	−0.65 (−0.82 to −0.48)	<.001	—	—	−0.54 (−0.67 to −0.37)	<.001
2022	−0.45 (−0.58 to −0.33)	<.001	—	—	−0.23 (−0.34 to −0.10)	<.001
Survey methodology (online as reference)						
CATI[c]	0.77 (0.08 to 1.47)	.04	—	—	0.11 (−0.31 to 0.53)	.69
Face-to-face	1.03 (0.32 to 1.74)	.009	—	—	0.58 (0.12 to 1.02)	.04

[a]Not available.

[b]GNI: gross national income.

[c]CATI: computer-assisted telephone interview.

XSL•FO
RenderX

Table . Multilevel logistic regression analyses for associations of factors and attitudes toward COVID-19 vaccine importance, effectiveness, and safety controlling individual-and-country level variables.

Variables	Importance		Effectiveness		Safety	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Cultural tightness	0.68 (0.60 to 0.76)	<.001	0.62 (0.54 to 0.69)	<.001	0.64 (0.55 to 0.73)	<.001
Individual level						
Age group, years (18 - 24 years as reference)						
25 - 34	−0.04 (−0.07 to −0.01)	.004	−0.04 (−0.07 to −0.01)	.22	−0.01 (−0.05 to 0.02)	.39
35 - 44	−0.02 (−0.05 to 0.01)	.18	−0.03 (−0.07 to 0.00)	.06	0.005 (−0.03 to 0.04)	.78
45 - 54	0.02 (−0.01 to 0.05)	.19	−0.01 (−0.04 to 0.03)	.76	0.012 (−0.02 to 0.05)	.50
≥55	0.15 (0.11 to 0.18)	<.001	0.10 (0.07 to 0.13)	<.001	0.10 (0.07 to 0.13)	<.001
Sex (male as reference)						
Female	−0.06 (−0.08 to −0.04)	<.001	−0.11 (−0.13 to −0.09)	<.001	−0.12 (−0.14 to −0.10)	<.001
Education level (primary or below as reference)						
Secondary	0.02 (−0.01 to 0.05)	.27	−0.04 (−0.08 to −0.00)	.03	−0.004 (−0.04 to 0.03)	.83
University or above	0.10 (0.07 to 0.13)	<.001	0.03 (−0.01 to 0.07)	.12	0.03 (−0.01 to 0.07)	.09
Other educational level	−0.08 (−0.17 to 0.00)	.06	−0.14 (−0.23 to −0.05)	.002	−0.20 (−0.29 to −0.11)	<.001
Religion (largest as reference)						
Minority	−0.04 (−0.06 to −0.02)	<.001	−0.05 (−0.07 to −0.03)	<.001	−0.04 (−0.07 to −0.02)	<.001
Refusal to answer	−0.10 (−0.14 to −0.06)	<.001	−0.06 (−0.10 to −0.02)	.01	−0.09 (−0.13 to −0.04)	<.001
Individual perception	0.51 (0.50 to 0.52)	<.001	0.47 (0.46 to 0.48)	<.001	0.50 (0.49 to 0.51)	<.001
Individual perception × cultural tightness	−0.15 (−0.17 to −0.14)	<.001	−0.13 (−0.14 to −0.12)	<.001	−0.14 (−0.15 to −0.13)	<.001
Country level						
Individualism	−0.15 (−0.24 to −0.06)	.02	−0.09 (−0.18 to −0.01)	.08	−0.04 (−0.13 to 0.06)	.57
Indulgence	−0.02 (−0.09 to 0.05)	.63	0.00 (−0.06 to 0.06)	.96	−0.03 (−0.10 to 0.05)	.60
Long-term orientation	0.01 (−0.09 to 0.10)	.93	0.03 (−0.06 to 0.12)	.57	−0.01 (−0.11 to 0.10)	.93
Power distance	−0.04 (−0.12 to 0.05)	.50	0.03 (−0.05 to 0.10)	.54	0.01 (−0.08 to 0.10)	.86
Uncertainty avoidance	−0.06 (−0.13 to 0.01)	.19	−0.06 (−0.12 to 0.00)	.15	−0.06 (−0.13 to 0.02)	.23
GNI[a]	0.11 (0.04 to 0.19)	.03	0.08 (0.01 to 0.14)	.09	0.00 (−0.08 to 0.08)	.95
Population density	0.07 (−0.01 to 0.14)	.18	0.04 (−0.02 to 0.11)	.30	0.02 (−0.06 to 0.10)	.70

XSL•FO
RenderX

Variables	Importance		Effectiveness		Safety	
	β (95% CI)	*P* value	β (95% CI)	*P* value	β (95% CI)	*P* value
Survey time (2020 as reference)						
2021	−0.34 (−0.38 to −0.27)	<.001	−0.11 (−0.16 to −0.04)	<.001	−0.09 (−0.15 to −0.03)	.004
2022	−0.25 (−0.29 to −0.19)	<.001	−0.01 (−0.05 to 0.05)	.70	0.02 (−0.02 to 0.08)	.33
Survey methodology (online as reference)						
CATI[b]	0.06 (−0.10 to 0.22)	.54	0.10 (−0.04 to 0.25)	.27	−0.06 (−0.23 to 0.12)	.60
Face-to-face	0.24 (0.05 to 0.40)	.04	0.16 (−0.00 to 0.32)	.10	0.18 (−0.01 to 0.36)	.12

[a]GNI: gross national income.

[b]CATI: computer-assisted telephone interview.

Discussion

Principal Findings

Our study found a significant association between cultural tightness and COVID-19 vaccine confidence. Countries with higher cultural tightness tended to have greater confidence in the vaccine, which may account for the disparities in vaccine confidence observed among countries worldwide. This finding not only facilitated progress in explaining vaccine hesitancy determinants from a cultural perspective, but also underscored the importance of developing tailored and effective interventions to address the issue.

In societies characterized by high cultural tightness, individuals were expected to adhere to social norms, minimize risks, and prioritize stability [24]. During the COVID-19 pandemic, individuals in these cultures demonstrated a greater tendency toward collective behavior and were more likely to strictly adhere to guidelines issued by authorities, including recommendations for vaccination [13,25]. Notably, individuals in high-tightness cultures would concern with the social consequences of norm deviation—such as the risk of social exclusion due to vaccine refusal—rather than solely the health risks associated with nonvaccination [12]. Furthermore, high-tightness cultures typically exhibited a centralized information environment dominated by official sources, which might serve to mitigate the spread of vaccine-related misinformation. Conversely, low cultural tightness societies were characterized by a more pluralistic information landscape, allowing for the sharing of diverse opinions on vaccination, which might increase the misinformation spread that reduced vaccine willingness [26]. In fact, historical evidence showed that the antivaccine movement first emerged in countries with loose cultures; for example, an image linking smallpox vaccination to cattle traits was published in a British magazine in the 18th century [27]. The movement against smallpox vaccination also first occurred in North America in the 19th century [28]. The antivaccination movement reemerged in the United Kingdom in the mid-1970s due to a report suggesting that the diphtheria, pertussis, and tetanus vaccine could lead to severe neurological diseases [28].

Our findings were inconsistent with the previous studies. Differences in analytical methods, the number of countries included, and the measurement items of vaccine confidence may account for the discrepancies between our results and those reported in earlier researches [14-16]. Ng and Tan [14] examined the association between cultural tightness and vaccination willingness across 12 countries without controlling for potential confounders, which limited the reliability of their results. They thought that negative association may be due to individuals in countries with higher culture tightness feeling a low perceived risk of COVID-19 infection due to effective control of the pandemic, resulting in decreased motivation to receive vaccinations [14]. Shi et al [16] focused primarily on East and Southeast Asian countries, where the cultural tightness scores measured were relatively high (more than half of the countries) and exhibited limited variability among 8 countries. Lu [15] investigated the associations in 22 countries. Furthermore, previous studies often relied on a single-item measure to assess vaccine confidence. Our study examined the association by employing VCI items, adjusting confounders, and including a larger, more diverse sample of countries.

Our study demonstrated that younger age, female sex, and low educational level were positively associated with low vaccine confidence, in line with findings from a systematic review of factors influencing COVID-19 vaccine acceptability [6]. The notable observation here pertained to the low level of vaccine confidence among young individuals. It might be attributed to their generally better health status, rendering them less susceptible to severe illness following a COVID-19 infection [29], thus potentially reducing their perception of the severity of infection. On the other hand, young individuals constituted a significant portion of the population actively engaged with social media, with many relying on these platforms as their primary source of health-related information [30]. Unfortunately, the internet also served as a breeding ground for vaccine-related misinformation, which could have adversely affected their confidence in vaccination [8,31]. Given young individuals'

XSL•FO
RenderX

extensive social interactions, they played a crucial role in virus transmission. Designing and leveraging social media and digital technologies, which enjoyed widespread popularity among young individuals, presented a potential avenue. For instance, a study by Lee et al [31] demonstrated the potential effectiveness of chatbot in enhancing vaccine confidence.

Comprehending the strength of conformity to social norms may provide insights into designing interventions. In countries with low culture tightness, there is a need to reinforce social norms linked to vaccination behaviors. Previous evidence has demonstrated the social norms could be altered (example of reducing drinking and driving in the USA) [32] and the country as a whole might tighten up on norms that have become too loose (example of reducing alcohol and drug abuse in Iceland) [33]. Health authorities should remain vigilant regarding vaccine-related misinformation and antivaccine sentiments on social media. Organizing town hall meetings or Q&A sessions can facilitate direct engagement between health practitioners and the public, fostering open dialogue and effectively addressing vaccine hesitancy.

In countries with high cultural tightness, engaging respected government officials or health experts to publicly endorse and promote vaccination may be effective. In such societies, authorities can frame vaccination as a civic duty and underscore the potential social repercussions associated with noncompliance. It should be noted that high level of cultural tightness, while conferring benefits, is not a panacea for the problem of vaccine hesitancy, and having a high-tightness culture does not guarantee a low risk of vaccine hesitancy. It is our supposition that, in high cultural tightness countries, low vaccination rates may be observed when there was limited awareness about vaccine availability and vaccination was not considered a social norm. This may be the case in China, which has the largest population in the world, yet only has a 1% - 3% influenza vaccination rate [34].

Our study has limitations. First, COVID-19 vaccine confidence was influenced by many unidentified factors (such as trust in authority), although we did our best to adjust numerous individual- and country- variables in the model. Second, our analysis was restricted to 28 countries for which cultural tightness and COVID-19 vaccine confidence data were available, which may limit the robustness of our results. Third, the association between cultural tightness and COVID-19 vaccine confidence should be examined on more nuanced levels (such across states within a country), particularly for large countries with diverse cultural structures like the United States. Fourth, while we adjusted for survey methodology by including it as a covariate in the regression models, the response biases caused by different data collection methods may still exist. Finally, given the diversity and complexity of religious beliefs in each country, we categorized them into three groups: “largest,” “minority,” and “refusal to answer,” potentially resulting in the loss of more nuanced findings in this domain.

Conclusions

Our study revealed that countries with higher cultural tightness might be more likely to show greater COVID-19 vaccine confidence. Individuals’ sociodemographics, including younger age, female sex, and low educational level, were positively associated with low vaccine confidence. Our findings provided insights on designing tailor interventions to handle vaccine hesitancy.

Acknowledgments

This study was supported by the InnoHK initiative of the Innovation and Technology Commission of the Hong Kong Special Administrative Region Government. The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

QW contributed to conceptualization, data curation, formal analysis, writing – original draft, writing – review and editing, methodology, and visualization. AB contributed to reviewing the paper. LL contributed to conceptualization, supervision, writing – review and editing, and funding acquisition.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary tables and figures on survey implementation, key constructs, country-level indicators, and model diagnostics.
[DOCX File, 1400 KB - publichealth_v11i1e66872_app1.docx]

References

XSL•FO
RenderX

1. MacDonald NE, SAGE Working Group on Vaccine Hesitancy. Vaccine hesitancy: definition, scope and determinants. Vaccine 2015 Aug 14;33(34):4161-4164. [doi: 10.1016/j.vaccine.2015.04.036] [Medline: 25896383]
2. Patel MK, Orenstein WA. Classification of global measles cases in 2013-17 as due to policy or vaccination failure: a retrospective review of global surveillance data. Lancet Glob Health 2019 Mar;7(3):e313-e320. [doi: 10.1016/S2214-109X(18)30492-3] [Medline: 30784632]
3. Ten threats to global health in 2019. World Health Organization. 2019. URL: https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019 [accessed 2024-07-02]
4. Lazarus JV, Ratzan SC, Palayew A, et al. A global survey of potential acceptance of a COVID-19 vaccine. Nat Med 2021 Feb;27(2):225-228. [doi: 10.1038/s41591-020-1124-9] [Medline: 33082575]
5. Lazarus JV, Wyka K, White TM, et al. Revisiting COVID-19 vaccine hesitancy around the world using data from 23 countries in 2021. Nat Commun 2022 Jul 1;13(1):3801. [doi: 10.1038/s41467-022-31441-x] [Medline: 35778396]
6. Wang Q, Yang LQ, Jin H, Lin L. Vaccination against COVID-19: a systematic review and meta-analysis of acceptability and its predictors. Prev Med 2021 Sep;150:106694. [doi: 10.1016/j.ypmed.2021.106694] [Medline: 34171345]
7. Sturgis P, Brunton-Smith I, Jackson J. Trust in science, social consensus and vaccine confidence. Nat Hum Behav 2021 Nov;5(11):1528-1534. [doi: 10.1038/s41562-021-01115-7] [Medline: 34002053]
8. Lin L, Mei R, Lau EHY, et al. Misinformation belief and trust in local authority: a multilevel modelling of COVID-19 vaccine confidence in 14 asia-pacific countries. SSRN. Preprint posted online on Feb 22, 2023. [doi: 10.2139/ssrn.4364881]
9. Bonan J, Cattaneo C, d'Adda G, Tavoni M. The interaction of descriptive and injunctive social norms in promoting energy conservation. Nat Energy 2020;5(11):900-909. [doi: 10.1038/s41560-020-00719-z]
10. Blair G, Littman R, Paluck EL. Motivating the adoption of new community-minded behaviors: an empirical test in Nigeria. Sci Adv 2019 Mar;5(3):eaau5175. [doi: 10.1126/sciadv.aau5175] [Medline: 30891494]
11. Moehring A, Collis A, Garimella K, Rahimian MA, Aral S, Eckles D. Providing normative information increases intentions to accept a COVID-19 vaccine. Nat Commun 2023 Jan 9;14(1):126. [doi: 10.1038/s41467-022-35052-4] [Medline: 36624092]
12. Gelfand MJ, Raver JL, Nishii L, et al. Differences between tight and loose cultures: a 33-nation study. Science 2011 May 27;332(6033):1100-1104. [doi: 10.1126/science.1197754] [Medline: 21617077]
13. Gelfand MJ, Jackson JC, Pan X, et al. The relationship between cultural tightness-looseness and COVID-19 cases and deaths: a global analysis. Lancet Planet Health 2021 Mar;5(3):e135-e144. [doi: 10.1016/S2542-5196(20)30301-6] [Medline: 33524310]
14. Ng JH, Tan EK. COVID-19 vaccination and cultural tightness. Psychol Med 2023 Feb;53(3):1124-1125. [doi: 10.1017/S0033291721001823] [Medline: 33902766]
15. Lu JG. Two large-scale global studies on COVID-19 vaccine hesitancy over time: culture, uncertainty avoidance, and vaccine side-effect concerns. J Pers Soc Psychol 2023 Apr;124(4):683-706. [doi: 10.1037/pspa0000320] [Medline: 35925705]
16. Shi J, Kim HK, Salmon CT, Tandoc EC Jr, Goh ZH. Cultural tightness-looseness and normative social influence in eight Asian countries: associations of individual and collective norms with vaccination intentions. Soc Sci Med 2024 Jan;340:116431. [doi: 10.1016/j.socscimed.2023.116431] [Medline: 38000175]
17. de Figueiredo A, Simas C, Larson HJ. COVID-19 vaccine acceptance and its socio-demographic and emotional determinants: a multi-country cross-sectional study. Vaccine (Auckl) 2023 Jan 9;41(2):354-364. [doi: 10.1016/j.vaccine.2022.10.051] [Medline: 36414475]
18. de Figueiredo A, Temfack E, Tajudeen R, Larson HJ. Declining trends in vaccine confidence across sub-Saharan Africa: a large-scale cross-sectional modeling study. Hum Vaccin Immunother 2023 Dec 31;19(1):2213117. [doi: 10.1080/21645515.2023.2213117] [Medline: 37290478]
19. Choi Y, Leung K, Wu JT, Larson HJ, Lin L. Identifying vaccine-hesitant subgroups in the Western Pacific using latent class analysis. NPJ Vaccines 2025 Feb 12;10(1):29. [doi: 10.1038/s41541-025-01067-3] [Medline: 39939318]
20. Fox AM, Choi Y, Lin L. Substantial disparities In COVID-19 vaccine uptake and unmet immunization demand in low- and middle-income countries. Health Aff (Millwood) 2023 Dec;42(12):1697-1705. [doi: 10.1377/hlthaff.2023.00729] [Medline: 38048509]
21. Eriksson K, Strimling P, Gelfand M, et al. Perceptions of the appropriate response to norm violation in 57 societies. Nat Commun 2021 Mar 5;12(1):1481. [doi: 10.1038/s41467-021-21602-9] [Medline: 33674587]
22. de Figueiredo A, Simas C, Karafillakis E, Paterson P, Larson HJ. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. Lancet 2020 Sep 26;396(10255):898-908. [doi: 10.1016/S0140-6736(20)31558-0] [Medline: 32919524]
23. Hofstede G, Hofstede GJ, Minko VM. Cultures and Organizations: Software of the Mind, 3rd edition: McGraw Hill; 2010.
24. Gelfand MJ, Nishii LH, Raver JL. On the nature and importance of cultural tightness-looseness. J Appl Psychol 2006 Nov;91(6):1225-1244. [doi: 10.1037/0021-9010.91.6.1225] [Medline: 17100480]
25. Larson HJ. Politics and public trust shape vaccine risk perceptions. Nat Hum Behav 2018 May;2(5):316. [doi: 10.1038/s41562-018-0331-6] [Medline: 30962595]
26. Pertwee E, Simas C, Larson HJ. An epidemic of uncertainty: rumors, conspiracy theories and vaccine hesitancy. Nat Med 2022 Mar;28(3):456-459. [doi: 10.1038/s41591-022-01728-z] [Medline: 35273403]

XSL•FO
RenderX

27. Poland GA, Jacobson RM. The age-old struggle against the antivaccinationists. N Engl J Med 2011 Jan 13;364(2):97-99. [doi: 10.1056/NEJMp1010594] [Medline: 21226573]
28. Dubé E, Vivion M, MacDonald NE. Vaccine hesitancy, vaccine refusal and the anti-vaccine movement: influence, impact and implications. Expert Rev Vaccines 2015 Jan;14(1):99-117. [doi: 10.1586/14760584.2015.964212] [Medline: 25373435]
29. Kronbichler A, Kresse D, Yoon S, Lee KH, Effenberger M, Shin JI. Asymptomatic patients as a source of COVID-19 infections: a systematic review and meta-analysis. Int J Infect Dis 2020 Sep;98:180-186. [doi: 10.1016/j.ijid.2020.06.052] [Medline: 32562846]
30. L'Engle KL, Burns JR, Basuki A, Couture MC, Regan AK. Liberals are believers: young people assign trust to social media for COVID-19 information. Health Commun 2024 Apr;39(2):310-322. [doi: 10.1080/10410236.2023.2164959] [Medline: 36628497]
31. Lee KY, Dabak SV, Kong VH, et al. Effectiveness of chatbots on COVID vaccine confidence and acceptance in Thailand, Hong Kong, and Singapore. NPJ Digit Med 2023 May 25;6(1):96. [doi: 10.1038/s41746-023-00843-6] [Medline: 37231110]
32. Perkins HW, Linkenbach JW, Lewis MA, Neighbors C. Effectiveness of social norms media marketing in reducing drinking and driving: a statewide campaign. Addict Behav 2010 Oct;35(10):866-874. [doi: 10.1016/j.addbeh.2010.05.004] [Medline: 20619177]
33. Young E. How Iceland got teens to say no to drugs. The Atlantic. 2017. URL: https://www.theatlantic.com/health/archive/2017/01/teens-drugs-iceland/513668/ [accessed 2024-07-02]
34. Yang J, Atkins KE, Feng L, et al. Seasonal influenza vaccination in China: landscape of diverse regional reimbursement policy, and budget impact analysis. Vaccine (Auckl) 2016 Nov 11;34(47):5724-5735. [doi: 10.1016/j.vaccine.2016.10.013] [Medline: 27745951]

Abbreviations

GNI: gross national income
VCI: Vaccine Confidence Index
VCP: Vaccine Confidence Project
VIF: variance inflations factor
WHO: World Health Organization

Edited by A Mavragani; submitted 25.09.24; peer-reviewed by W Zheng, Z Shen; revised version received 28.02.25; accepted 08.03.25; published 24.04.25.

Please cite as:
Wang Q, Bolio A, Lin L
The Association Between Cultural Tightness and COVID-19 Vaccine Confidence From 28 Countries: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e66872
URL: https://publichealth.jmir.org/2025/1/e66872
doi:10.2196/66872

© Qiang Wang, Ana Bolio, Leesa Lin. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.4.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Predictors of COVID-19 Vaccination Intention and Behavior Among Young People in a European Union Country With Low COVID-19 Vaccination Rates: Cross-Sectional Study

Sara Atanasova[1], PhD; Tanja Kamin[1], PhD; Nina Perger[1], PhD

Faculty of Social Sciences, University of Ljubljana, Ljubljana, Slovenia

Corresponding Author:
Sara Atanasova, PhD
Faculty of Social Sciences
University of Ljubljana
Kardeljeva pl. 5
Ljubljana, 1000
Slovenia
Phone: 386 1 5805 167
Email: sara.atanasova@fdv.uni-lj.si

Abstract

Background: Vaccination against COVID-19 is a critical measure for managing the pandemic and achieving herd immunity. In 2021, Slovenia had a significantly lower COVID-19 vaccination rate compared to the average rate in the European Union, with individuals aged younger than 37 years showing the highest hesitancy. Previous studies primarily explored vaccination willingness before vaccines were available to young people, leaving a gap in understanding the factors influencing vaccination behavior and differences within the population of young people.

Objective: This study aimed to investigate a wide set of predictors influencing COVID-19 vaccination intention and behavior among young people in Slovenia. Specifically, we aimed to compare vaccinated and unvaccinated young people, further categorizing the unvaccinated group into those who were hesitant, those who intended to vaccinate in the near future, and those who refused vaccination.

Methods: An integrated model, based on the health belief model and theory of planned behavior, was developed, and it included additional contextual factors (such as trust in science, trust in vaccines, conspiracy theory tendencies, etc) and health-related and sociodemographic characteristics. Data were collected in August 2021 via the online access survey panel JazVem (Valicon), targeting individuals aged 15-30 years in Slovenia. Quotas ensured that the sample (n=507) was quasi-representative according to age, gender, education, and region. Bivariate analyses and multinomial logistic regression were performed to explore the determinants of vaccination intention and behavior.

Results: Among respondents, 45.8% (232/507) were vaccinated, 30.0% (152/507) refused vaccination, 12.4% (63/507) were hesitant, and 11.8% (60/507) intended to undergo vaccination in the near future. Vaccinated individuals were predominantly aged 23-26 years, had higher education, and reported above-average material status. Refusers were more common among the youngest (15-18 years) and oldest (27-30 years) groups, had lower education, and showed higher conspiracy theory tendencies. Multinomial regression analysis revealed that unvaccinated respondents who perceived greater COVID-19–related health consequences were more likely to delay vaccination (adjusted odds ratio [aOR] 2.0, 95% CI 1.2-3.3) or exhibit hesitancy (aOR 1.9, 95% CI 1.1-3.2) compared with vaccinated respondents. Subjective norms were less influential among hesitant individuals (aOR 0.4, 95% CI 0.2-0.7) and refusers (aOR 0.3, 95% CI 0.2-0.7) than among vaccinated individuals. Self-efficacy in managing health problems was less evident among those who delayed vaccination to the near future (aOR 0.5, 95% CI 0.3-0.9) than among vaccinated individuals.

Conclusions: This study underscores the complexity of vaccination intentions and behaviors among young people, emphasizing the necessity for public health strategies promoting vaccination to be tailored to the specific reasons for nonvaccination within different subgroups. Interventions aimed at addressing vaccine hesitancy and delays should particularly focus on individuals with lower education and material disadvantages. By fostering trust and enhancing self-efficacy, these interventions could more effectively promote vaccine uptake.

XSL•FO
RenderX

(JMIR Public Health Surveill 2025;11:e64653) doi:10.2196/64653

KEYWORDS

vaccine uptake; young people; COVID-19 vaccine; health belief model; theory of planned behavior

Introduction

The rapid development of COVID-19 vaccines is considered as a pandemic success story, and vaccination against COVID-19 is considered as one of the most effective preventive measures [1,2]. The herd immunity threshold for SARS-CoV-2 has been assumed to be between 50% and 67%, and the administration of SARS-CoV-2 vaccines was suggested to be an important strategy to reach this threshold [3]. It has been suggested that around 75% to 90% of the population needs to be vaccinated against COVID-19 to successfully control the pandemic [4]. Slovenia lagged far behind this goal; it was among the countries with the lowest vaccination rate against COVID-19 in the European Union (EU) since vaccination against COVID-19 became available to citizens [5]. By the end of 2021, the 27 EU member states had vaccinated an average of 77% of their population aged 18 years or older. However, Slovenia reported a vaccination completion rate of only 65.6% among adults [6].

Individuals showing the most hesitancy regarding COVID-19 vaccination in Slovenia were those younger than 37 years [7]. This is consistent with the findings of previous studies globally that among adults, younger age groups had a lower willingness to be vaccinated against COVID-19 [8-15]. Vaccination against COVID-19 among those most likely to get infected and to transmit the virus, such as young people in school settings, was identified as a crucial strategy for reducing ongoing transmission and preventing outbreaks [16]. This approach was considered essential for establishing herd immunity; protecting elderly people and immunocompromised individuals from severe outcomes, hospitalization, and death [17]; and ensuring the continued functioning of the workforce [16].

The vaccination behavior of young people is complex, as highlighted by a systematic review of qualitative studies on adolescents' understanding and attitudes toward vaccines [18]. It has been suggested that health literacy plays a crucial role in adolescents' comprehension of the benefits and risks of vaccination. However, other factors, such as peer influence, parental guidance, trust in health care providers, and misconceptions about the vaccine's safety and efficacy, also significantly influence vaccine uptake among adolescents [18]. Given the unprecedented speed at which COVID-19 vaccines were developed compared to historical vaccine development and testing timelines, a higher degree of hesitancy toward COVID-19 vaccination was somewhat anticipated. A study examining the factors that discourage individuals from participating in vaccine trial registries or clinical trials identified primary concerns related to the safety of novel vaccines and a lack of trust in those involved in vaccine development [19].

Similar findings were obtained for vaccination behavior related to COVID-19, with attitude toward vaccination, trust in health staff and scientists, time of information, and conspiracy beliefs about COVID-19 being identified as the best predictors of the intention of vaccination among young adults in Spain at the beginning of the vaccination campaign [20]. Chaufan [21], for example, questioned the prevailing risk-benefit analysis with regard to younger people. If vaccine uptake significantly reduces health risks in the general population, the younger population has a much lower risk of COVID-19–related complications. Given their lower risk of COVID-19–related health complications, the reluctance to vaccinate in this population group, which is often seen as unreasonable, may seem understandable [22]. Their perceived benefits of vaccination seem to be more collectively beneficial, while the risks associated with vaccination are perceived to be individual. This vaccination was new, and there was no broad-based data available on the efficacy and immunogenicity of the vaccination in the young population at the beginning of the vaccination campaign against COVID-19 [23].

Another scoping review suggested that the primary factors of young people's acceptance of the COVID-19 vaccine include the desire to protect themselves and close family members or friends, fear of infection, professional recommendations, and employer obligations [24]. On the other hand, the primary hesitancy factors include concerns about the safety of the vaccine and its side effects, effectiveness, and efficacy, as well as a lack of trust in the pharmaceutical industry and government, conspiracies, and favoring natural immunity [24]. The same study also suggested the need for additional research into COVID-19 vaccine–related decision-making dynamics for specific adolescent and youth population age ranges to better understand how vaccination-related behavior is influenced by environmental or social factors as well as personal health and susceptibility [24].

The significant affective distress experienced by the younger population due to COVID-19–related protective measures, particularly isolation and social distancing [25], may substantially influence both their willingness to receive the vaccination, as suggested by a study [26], and their actual uptake of COVID-19 vaccination. This is especially pertinent if vaccination is perceived as a means to achieve greater mobility and fewer restrictions on in-person social interactions. Negative emotions (affective responses) associated with recommended protective behaviors may impact core beliefs (cognitive responses) and undermine evidence-based reasoning, thereby diminishing the perceived importance of these behaviors in decision-making processes (behavioral responses) [25]. These findings highlight that vaccination intention and behavior depend on several contextual factors. Most studies conducted prior to our research primarily focused on investigating the willingness to receive COVID-19 vaccination at a time when vaccines were not yet available to young people. Furthermore, studies involving the youth have often relied on opportunistic samples, predominantly drawn from student populations [27-30]. Such approaches have resulted in narrow and unrepresentative

XSL•FO
RenderX

samples that fail to adequately reflect the broader category of young people. Our study aims to address these limitations.

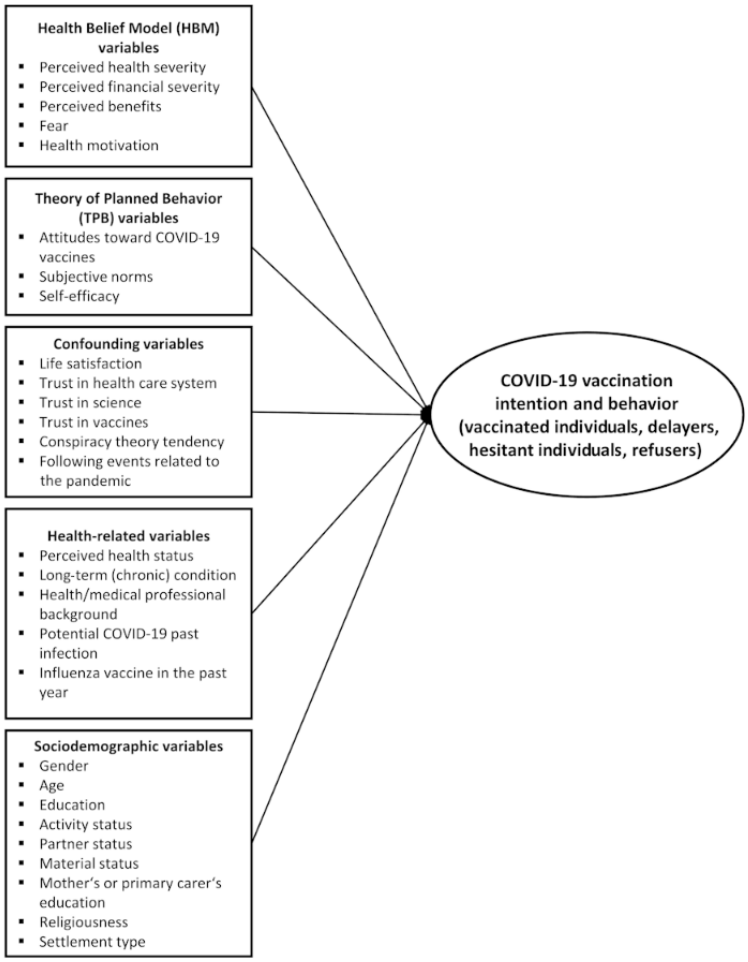
The main aim of this study is to investigate a wide set of predictors influencing COVID-19 vaccination intention and behavior among young people in Slovenia. Specifically, based on a quasi-representative sample of young people in Slovenia, this study aims to gain a deeper understanding of the cognitive, affective, and behavioral determinants of vaccine uptake within this population. It examines factors related to COVID-19 vaccination intention and behavior, including sociodemographic and health-related variables, as well as other predictors identified in previous studies and theoretical frameworks as significant in influencing vaccination behavior [14,24]. The study adopts and further develops the model by Shmueli [31], incorporating factors from 2 behavioral models (health belief model [HBM] and theory of planned behavior [TPB]) along with additional social factors such as trust in science and tendency toward conspiracy theories. These factors have been shown to be strong predictors of the willingness to receive the COVID-19 vaccine in the adult population [32]. Building on this integrated model, the study contributes by comparing vaccinated and unvaccinated young people and further categorizing the unvaccinated group into those who are hesitant, those who intend to vaccinate, and those who refuse vaccination. Our goal is to enhance the body of research that explores the decision-making processes within distinct groups of young people. The study seeks to better understand and inform public health strategies for increasing vaccination rates for COVID-19 and future viral outbreaks among young people.

Methods

Framework of the Study

Based on the study by Schmueli [31], we developed an integrated model that integrates the constructs of 2 prominent health behavior theories (HBM and TPB) and includes a number of other contextual factors (eg, trust in science, trust in vaccines, conspiracy theory tendencies, etc) and health-related and sociodemographic characteristics of the studied population to identify the predictors of COVID-19 vaccination intention and behavior among young people in Slovenia (Figure 1). The HBM, which has generated a great deal of research interest in various health-related behaviors to date, is based on a set of core beliefs (cognitive factors) regarding risk susceptibility, risk severity, benefits, and barriers [33]. According to the widely tested TPB [34,35], the underlying behavioral intention as a direct antecedent of behavior is determined by the attitude toward the behavior, subjective norms, and perceived behavioral control. Both models aim to predict behavior, and in our case, we target COVID-19 vaccination–related behavior.

Figure 1. Framework of the predictors of COVID-19 vaccination intention and behavior.



XSL•FO
RenderX

Study Design

The data for this cross-sectional study were collected using an online survey questionnaire administered by the online survey panel JazVem (Valicon, a Slovenian marketing research company) [36]. Existing members of the JazVem online survey panel were contacted via email to participate in the survey. Members voluntarily sign up with the JazVem online survey panel, agree with the terms and privacy conditions of the online survey panel, sign participation consent, provide baseline information, and regularly receive emails inviting them to take part in different research studies. For participation in JazVem surveys, panelists receive small incentives in the form of points, which can be redeemed for modest rewards. Online panels operate under rigorous recruitment and quality assurance protocols to ensure participants can enroll only once, avoid overexposure to surveys, and maintain high levels of engagement. More information about the study design can be found in the CHERRIES (Checklist for Reporting Results of Internet e-Surveys) in Multimedia Appendix 1.

Data Collection

The online questionnaire used in this study was developed by the authors and included 79 questions in the Slovene language, which took around 15 to 20 minutes to answer. The questionnaire covered questions about the perception of COVID-19–related protective measures, questions regarding behavior for vaccination against COVID-19, questions related to HBM and TPB constructs, questions about trust in science and vaccines, questions about the respondent's health, and questions related to sociodemographic characteristics. The questionnaire and measurement instruments were evaluated for content validity by 8 experts from the fields of public health, health communication, sociology, communication studies, statistics, social science methodology, and psychology.

Data were collected between August 11, 2021, and August 17, 2021. Eligible participants were aged 15 to 30 years and resided in Slovenia. Quotas were used to achieve a quasi-representative sample of the general population in Slovenia according to age, gender, education, and region based on data from the Statistical Office of the Republic of Slovenia. Based on quota sampling, 1197 JazVem panelists were invited to participate in the online survey. Among these, 568 clicked on the link for the online survey and 555 viewed the introduction page of the survey. Of the 568 respondents, 525 fully completed the survey, with a completion rate of 92.4% (525/568). After the data screening and cleaning procedures, the final sample comprised 507 respondents. The data were weighted to match the general population distribution according to age, gender, education, and region. More information about the data collection can be found in the CHERRIES in Multimedia Appendix 1.

Measurement Instruments

Vaccine Uptake

The dependent variable vaccine uptake was measured with the following three questions: (1) "Please tell us about your decision regarding the COVID-19 vaccination. Have you been vaccinated against COVID-19?" (2) "Do you intend to get vaccinated against COVID-19 in the next 30 days?" and (3) "Do you intend to get vaccinated against COVID-19 in the next 6 months?" The classification and naming of unvaccinated respondents was guided by the Stages of Change model [37]. This model posits several behavioral stages preceding the desired recommended behavior, which in our case was vaccination behavior: (1) Precontemplation: individuals in this stage are not considering vaccination; (2) Contemplation: individuals in this stage are considering vaccinating in the next 6 months, and according to the theory, they are considering the recommended behavior but still weighing the pros and cons; and (3) Preparation: individuals in this stage intend to get vaccinated in the next 30 days, and according to the theory, they may have taken some initial steps, such as seeking information or planning an appointment. Based on the respondents' answers, we computed vaccine uptake variables and divided the respondents into 4 categories as follows: respondents who reported being vaccinated were categorized as "vaccinated respondents;" respondents who reported not being vaccinated but intended to get vaccinated in the next 30 days were categorized as "delayers;" respondents who reported not being vaccinated but intended to get vaccinated in the next 6 months were categorized as "hesitant respondents;" and respondents who reported not being vaccinated and who did not intend to get vaccinated in the next 30 days or 6 months were categorized as "refusers."

The independent variables (ie, predictors of vaccine uptake among young people) were arranged into 5 groups: HBM variables, TPB variables, confounding variables, health-related variables, and sociodemographic variables.

HBM Variables

Perceived health severity was measured with an item adapted from Shmueli [31], with respondents asked to assess on a 5-point Likert scale (1 [insignificant consequences] to 5 [very serious consequences]) the possible health consequences in the case of COVID-19 infection.

Perceived financial severity was measured with an item adapted from Shmueli [31], with respondents asked to assess on a 5-point Likert scale (1 [insignificant consequences] to 5 [very serious consequences]) the possible financial consequences in the case of COVID-19 infection.

Perceived benefits were measured using 3 items from Chu and Liu [38], with 1 item assessing the perceived individual benefits of COVID-19 vaccines (eg, "COVID-19 vaccines are effective in preventing COVID-19") and 2 items assessing the perceived community benefits of COVID-19 vaccines (eg, "Having myself vaccinated against COVID-19 is beneficial for the health of others in my community"). Exploratory factor analysis (EFA) indicated that the items represent a single factor, explaining 79.1% of the variance. The Cronbach alpha (α=.95) demonstrated outstanding internal consistency.

Fear was measured with an item from Chu and Liu [38], with respondents asked to assess on a 5-point scale (1 [not at all] to 5 [very much]) how afraid they feel when they think about COVID-19.

Health motivation was measured with 2 items from Shmueli [31]. These items ("I make sure to eat healthy and diverse food every day" and "I ensure that I exercise and work out regularly")

XSL•FO
RenderX

were summed in an index and demonstrated acceptable internal consistency (α=.77; r=0.6; P<.001; r_{SB}=0.77).

The HBM also includes susceptibility, perceived barriers, and cues to action, which were measured in our study but only in the subsample of unvaccinated respondents. Because the purpose of this study is to compare different groups of young people, including those who accept vaccination, we did not include these variables in the analysis.

TPB Variables

Attitudes toward COVID-19 vaccines were measured using 6 items adapted from Chu and Liu [38]. Respondents were asked to evaluate how they personally feel about COVID-19 vaccination using 5-point scales with various descriptors (negative to positive, undesirable to desirable, bad to good, harmful to beneficial, foolish to wise, and pointless to meaningful). The EFA indicated that the items represent a single factor, explaining 85.7% of the variance, and the Cronbach alpha (α=.97) demonstrated outstanding internal consistency.

Subjective norms were measured using an adapted scale comprising 3 items from Chu and Liu [38]. Respondents were asked to indicate their level of agreement on a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]) with statements concerning their similar or significant others (eg, “Most people who are similar to me have been or will be vaccinated against COVID-19”). The EFA showed that the items represent a single factor, explaining 71.5% of the variance, and the Cronbach alpha (α=.88) demonstrated good internal consistency.

To measure self-efficacy, 3 items adapted from the General Self-Efficacy Scale of Schwarzer and Jerusalem [39] were used to ask respondents to rate their agreement on a 5-point scale (1 [strongly disagree] to 5 [strongly agree]) with statements related to managing their health (eg, “If unexpected events related to my health occur, I know how to handle them effectively”). The EFA revealed that the items represent a single factor, explaining 65.3% of the variance. The Cronbach alpha (α=.85) indicated good internal consistency.

Confounding Variables

Life satisfaction was measured using a single-item question asking respondents to assess their current satisfaction with life on a scale from 0 (extremely dissatisfied) to 10 (extremely satisfied).

Similarly, trust in the health care system was measured using a single-item question asking respondents to rate their trust in the Slovenian health care system on a scale from 0 (don’t trust at all) to 10 (completely trust).

Trust in science was measured using the Trust in Science and Scientists Inventory [40]. However, since the EFA did not confirm the unidimensionality of the original scale, the measure was slightly adjusted to include 9 items. These items asked respondents to assess their agreement on a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]) with statements related to the value of and trust in science (eg, “I believe that the work of scientists contributes to a better life for everyone”). The EFA revealed that the items represent a single factor, explaining 55.1% of the variance, and the Cronbach alpha (α=.92) demonstrated outstanding internal consistency.

Trust in vaccines was measured using 2 items that asked respondents to indicate their agreement on a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]) with statements related to the safety and efficiency of COVID-19 vaccines (eg, “I trust that the vaccines available to us against the coronavirus that causes COVID-19 are safe”). The 2 items were summed in an index and demonstrated outstanding internal consistency (α=.92; r=0.8; P<.001; r_{SB}=0.92).

Conspiracy theory tendency was measured using 3 items adapted from the Slovenian public opinion research #Novanormalnost [41]. Respondents were asked to assess their agreement with statements suggesting COVID-19 is a part of a conspiracy (eg, “Vaccination against the coronavirus that causes COVID-19 is an attempt to control the population”), using a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]). The EFA revealed that the 3 items represent a single factor, explaining 70.3% of the variance, and the Cronbach alpha (α=.87) showed good internal consistency.

Following events related to the pandemic was measured using a single-item question that asked respondents to assess the extent to which they follow pandemic-related events on a 5-point Likert scale (1 [I don’t follow events at all] to 5 [I follow events very often]).

Health-Related Variables

Perceived health status was measured using a single-item question on a 5-point Likert scale (1 [very bad] to 5 [excellent]), asking respondents to evaluate their current physical and mental health. The online questionnaire also included a question asking respondents to indicate whether they have a long-term (chronic) condition.

Additionally, single-item questions were used to determine whether respondents have a health or medical professional background and whether they received an influenza vaccine in the past year.

Potential COVID-19 past infection was measured by asking respondents to report their perceived probability of having been infected with COVID-19 in the past (probably infected or probably not infected).

Sociodemographic Variables

The online questionnaire included questions asking respondents to indicate their gender and age, with age transformed from a numeric to a categorical variable (15-18 years, 19-22 years, 23-26 years, and 27-30 years) to examine differences across specific age groups. Respondents were also asked to report their current education level (elementary school or less; 2-, 3-, 4-, or 5-year high school; and college, university, or higher education) and their activity status (high school student, student, employed, and unemployed).

Single-item questions were used to determine whether respondents have a partner, consider themselves religious, and live in urban or rural settlements. Respondents were also asked to assess their mother’s or primary carer’s education level

XSL•FO
RenderX

(elementary school or less; 2-, 3-, 4-, or 5-year high school; and college, university, or higher education) and their material status on a scale from 1 (significantly below average) to 5 (significantly above average). Due to a low number of observations in categories 1 (significantly below average) and 5 (significantly above average), the material status variable was recoded according to 3 categories: below average, average, and above average.

Statistical Analysis

Before the analysis, the data were weighted using the random iterative weighting method according to gender, age group, education, and region. The weighting process was conducted by Valicon, the JazVem online panel survey provider. The prevalence of vaccination-related behavior was measured as the percentage of vaccinated respondents, delayers, hesitant respondents, and refusers in the total study sample. A series of EFAs were conducted to determine the internal structure of the selected measurement instruments. Factors were extracted using principal axis factoring with oblimin rotation. We performed EFA using the following criteria: (1) eigenvalue greater than 1, (2) scree test, (3) items loading on the same factor (≥0.04), (4) no cross-loading, and (5) conceptual interpretability of factors. The quality of the measurement instrument was also assessed using reliability analysis, and the Cronbach α coefficient was computed, which ranges between 0 and 1.0. The rough guidelines are that a value of .9 or higher indicates outstanding internal consistency, a value of .8 or higher indicates good internal consistency, a value between .7 and .8 indicates acceptable reliability, and a value between .6 and .7 indicates questionable reliability [42]. For 2-item scales, reliability was estimated using the Spearman-Brown coefficient (r_{SB}), which was calculated based on the interitem Pearson correlation (r). Bivariate analyses were performed to compare groups according to sociodemographic characteristics, health-related variables, confounding variables, HBM variables, and TPB variables, using chi-square tests for categorical variables and ANOVA along with appropriate post hoc tests (Games-Howell) for continuous variables.

Multinomial logistic regression analysis was used to estimate the selected determinants of vaccine uptake with “vaccinated” as the reference group for comparisons. The selection of variables for the multinomial regression model was guided by their statistical significance in bivariate analyses and their theoretical relevance to the research question, ensuring the inclusion of predictors with stronger explanatory power while minimizing the risk of overfitting. Variables that did not show statistically significant associations with the outcome in bivariate analyses were excluded. While the inclusion of all variables was initially considered, challenges with model convergence underscored the importance of a parsimonious model that prioritizes interpretability and robustness. The final model, therefore, balances these considerations to provide reliable and meaningful insights. Confounding variables were included in the multinomial regression model alongside the primary independent variables (HBM and TPB variables, and sociodemographic and health-related variables). By incorporating these confounders into the model, we aimed to account for their potential influence on the outcome variable and to isolate the association between the primary independent variables and the dependent variable. This approach ensures that the estimated effects of the primary predictors are adjusted for the potential bias introduced by the confounding variables. We report adjusted odds ratios (aORs) and 95% CIs. Data were analyzed using IBM SPSS version 27 software.

Ethical Considerations

Data collection was conducted by the online survey panel provider Valicon, who is a member of the European Society for Opinion and Market Research (ESOMAR), and the data collection and research adhere to the International Chamber of Commerce and ESOMAR international code and professional standards of social research practice. We (the authors) had no access to the respondents’ information and were provided with an anonymized dataset that did not contain any identifiable personal information. The study was conducted in accordance with the Code of Ethics for Researchers of the University of Ljubljana [43] and the World Medical Association’s Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects [44]. The study received approval from the Ethics Committee of the Faculty of Social Sciences, University of Ljubljana (number: 801-2024-025/TD).

Results

Participant Characteristics

The results showed that the majority of respondents (232/507, 45.8%) had been vaccinated against COVID-19, 30.0% (152/507) refused vaccination, 12.4% (63/507) were hesitant, and 11.8% (60/507) reported that they intended to be vaccinated against COVID-19 in the near future (ie, delayers). The respondents were on average 23.4 years old (SD 4.4 years), comprising 51.6% (262/507) men and 48.4% (245/507) women (Table 1). The majority had completed high school (290/507, 57.3%), were employed (194/507, 38.3%), were students (184/507, 36.2%), and had a partner (258/507, 50.9%). On average, respondents reported the perceived material status as average (mean 2.1, SD 0.64). Most participants identified as religious (316/507, 62.3%) and resided in urban areas (260/507, 51.3%).

XSL•FO
RenderX

Table 1. Sociodemographic participant characteristics in the total sample and according to COVID-19 vaccination–related behavior groups.

Variable	Total sample (N=507), n (%)	Vaccinated individuals (n=232, 45.8%), n (%)	Delayers (n=60, 11.8%), n (%)	Hesitant individuals (n=63, 12.4%), n (%)	Refusers (n=152, 30.0%), n (%)	Chi-square (*df*)	*P* value
Gender						2.8 (3)	.43
Male	262 (51.6)	127 (48.5)	33 (12.6)	29 (11.1)	73 (27.9)		
Female	245 (48.4)	105 (42.9)	27 (11.0)	34 (13.9)	79 (32.2)		
Age (years) groups						39.7 (9)	<.001
15-18	102 (20.2)	28 (27.5)	18 (17.6)	17 (16.7)	39 (38.2)		
19-22	123 (24.3)	66 (53.7)	18 (14.6)	14 (11.4)	25 (20.3)		
23-26	133 (26.3)	81 (60.9)	6 (4.5)	11 (8.3)	35 (26,3)		
27-30	148 (29.2)	57 (38.5)	17 (11.7)	20 (13.5)	54 (36.5)		
Education						25.6 (6)	<.001
Elementary school or less	69 (13.6)	15 (21.7)	12 (17.4)	14 (20.3)	28 (40.6)		
2-, 3-, 4-, or 5-year high school	290 (57.3)	131 (45.2)	35 (12.1)	34 (11.7)	90 (31.0)		
College, university, or higher education	147 (29.1)	85 (57.8)	12 (8.2)	15 (10.2)	35 (23.8)		
Activity status						34.9 (9)	<.001
High school student	106 (20.9)	31 (29.2)	18 (17.0)	14 (13.2)	43 (40.6)		
Student	184 (36.2)	109 (59.2)	22 (12.0)	16 (8.7)	37 (20.1)		
Employed	194 (38.3)	85 (43.8)	16 (8.2)	28 (14.4)	65 (33.5)		
Unemployed	23 (4.5)	7 (30.4)	4 (17.4)	4 (17.4)	8 (34.8)		
Partner status						7.4 (3)	.06
Do not have a partner	249 (49.1)	109 (43.8)	39 (15.7)	32 (12.9)	69 (27.7)		
Have a partner	258 (50.9)	123 (47.7)	21 (8.1)	31 (12.0)	83 (32.2)		
Material status (n=494)						24.5 (6)	<.001
Below average	81 (16.4)	26 (32.1)	9 (11.1)	11 (13.6)	35 (43.2)		
Average	288 (58.3)	123 (42.7)	37 (12.8)	36 (12.5)	92 (31.9)		
Above average	125 (25.3)	78 (62.4)	8 (6.4)	15 (12.0)	24 (19.2)		
Mother's or primary carer's education (n=488)						17.4 (6)	.008
Elementary school or less	40 (8.2)	13 (32.5)	5 (12.5)	6 (15.0)	16 (40.0)		
2-, 3-, 4-, or 5-year high school	242 (49.4)	99 (40.9)	23 (9.5)	31 (12.8)	89 (36.8)		
College, university, or higher education	206 (42.2)	112 (54.4)	26 (12.6)	24 (11.7)	44 (21.4)		
Religiousness						7.5 (3)	.06
Nonreligious	191 (37.7)	100 (52.4)	15 (7.9)	22 (11.5)	54 (28.3)		
Religious	316 (62.3)	132 (41.8)	45 (14.2)	41 (13.0)	98 (31.0)		
Settlement type						13.1 (3)	.004
Urban	260 (51.3)	137 (52.7)	32 (12.3)	28 (10.8)	63 (24.2)		
Rural	247 (48.7)	95 (38.3)	28 (11.3)	35 (14.1)	90 (36.3)		

The majority of respondents reported having a very good (254/507, 50.2%) or good (150/507, 29.6%) health status (mean 3.7, SD 0.8). Additionally, 67.9% (344/507) of respondents indicated that they were probably not infected by coronavirus in the past (Table 2).

XSL•FO
RenderX

Table 2. Health-related participant characteristics in the total sample and according to COVID-19 vaccination–related behavior groups.

Variable	Total sample (N=507), n (%)	Vaccinated individuals (n=232, 45.8%), n (%)	Delayers (n=60, 11.8%), n (%)	Hesitant individuals (n=63, 12.4%), n (%)	Refusers (n=152, 30.0%), n (%)	Chi-square (*df*)	*P* value
Perceived health status						20.9 (12)	.05
Very bad	6 (1.2)	2 (33.3)	1 (16.7)	0 (0.0)	3 (50.0)		
Bad	32 (6.3)	18 (56.3)	5 (15.6)	4 (12.5)	5 (15.6)		
Good	150 (29.6)	63 (42.0)	27 (18.0)	21 (14.0)	39 (26.0)		
Very good	254 (50.2)	118 (46.5)	20 (7.9)	34 (13.4)	82 (32.3)		
Excellent	64 (12.6)	30 (46.9)	6 (9.4)	3 (4.7)	25 (39.1)		
Long-term (chronic) condition						4.3 (3)	.23
No	399 (78.7)	181 (45.3)	43 (10.8)	48 (12.0)	128 (32.0)		
Yes	108 (21.3)	51 (47.2)	17 (15.7)	15 (13.9)	25 (23.1)		
Potential COVID-19 past infection						14.4 (3)	.002
Probably not infected	344 (67.9)	171 (49.7)	46 (13.4)	36 (10.5)	91 (26.5)		
Probably infected	163 (32.1)	60 (36.8)	14 (8.6)	27 (16.6)	62 (38.0)		
Health or medical professional background						5.3 (3)	.15
No	428 (84.4)	192 (44.9)	46 (10.7)	55 (12.9)	135 (31.5)		
Yes	79 (15.6)	40 (50.0)	14 (17.5)	8 (10.0)	18 (22.5)		
Influenza vaccine in the past year						19.1 (3)	<.001
No	457 (90.1)	195 (42.7)	55 (12.0)	60 (13.1)	147 (32.2)		
Yes	50 (9.9)	37 (74.0)	5 (10.0)	3 (6.0)	5 (10.0)		

Bivariate Analyses and Comparisons Among Vaccination-Related Behavior Groups

Bivariate analyses and comparisons among vaccination-related behavior groups (Table 1) showed that vaccinated individuals were predominantly aged 23 to 26 years (81/133, 60.9%) and were more likely to have a higher education (85/147, 57.8%) and above-average material status (78/125, 62.4%). In comparison, refusers were more common among younger (15-18 years: 39/102, 38.2%) and older age groups (27-30 years: 54/148, 36.5%), had lower levels of education (elementary or less: 28/69, 40.6%), reported below-average material status (35/81, 43.2%), and were more likely from rural areas (90/247, 36.3%; Table 1). Additionally, refusers more often reported a potential COVID-19 past infection (62/163, 38.0%; Table 2). Results also showed that delayers shared characteristics with both vaccinated individuals and refusers, whereas hesitant respondents shared more characteristics with refusers, including rural residence, a lower education level, and a higher rate of potential COVID-19 past infection (Tables 1 and 2).

Table 3 presents participant characteristics across vaccination behavior groups based on confounding, HBM, and TPB variables. Vaccinated individuals reported significantly higher trust in the health care system (mean 6.5, SD 2.2), science (mean 3.9, SD 0.6), and COVID-19 vaccines (mean 3.9, SD 0.9) compared to refusers (Table 3). Conspiracy belief tendencies were on average the highest among refusers (mean 3.6, SD 1.0) and on average the lowest among vaccinated individuals (mean 1.9, SD 1.0). Pandemic-related events were followed more closely by vaccinated individuals (mean 3.1, SD 1.0) and delayers (mean 2.9, SD 0.9) than hesitant individuals (mean 2.7, SD 1.0) and refusers (mean 2.5, SD 1.0). Delayers perceived greater health (mean 2.8, SD 0.9) and financial (mean 2.7, SD 1.1) consequences from COVID-19 compared to refusers. Vaccinated individuals perceived the greatest benefit of vaccination (mean 4.0, SD 0.9), while refusers perceived the lowest benefit (mean 1.8, SD 0.8). Fear of COVID-19 was the lowest among refusers (mean 1.9, SD 1.0) and the highest among delayers (mean 3.0, SD 1.2). According to the TPB model, vaccinated individuals had the highest average scores for positive attitudes toward COVID-19 vaccines (mean 4.1, SD 1.1) and subjective norms (mean 3.8, SD 0.9), which gradually decreased among other groups and were the lowest on average among refusers (Table 2). Self-efficacy in coping with health problems was the highest among vaccinated individuals (mean 3.8, SD 0.7) and refusers (mean 3.7, SD 0.8) but was the lowest among delayers (mean 3.4, SD 0.7). Health motivation did not differ significantly between groups (Table 3).

XSL•FO
RenderX

Table 3. Descriptive statistics of confounding variables, health belief model variables, and theory of planned behavior variables in the total sample and according to COVID-19 vaccination–related behavior groups.

Variable	Total sample (N=507), mean (SD)	Vaccinated individuals (n=232, 45.8%), mean (SD)	Delayers (n=60, 11.8%), mean (SD)	Hesitant individuals (n=63, 12.4%), mean (SD)	Refusers (n=152, 30.0%), mean (SD)	ANOVA[a] *F* (*df*)	*P* value
Confounding variables							
Life satisfaction (scale 0-10)	6.6 (2.2)	6.8 (2.0)	6.1 (2.1)	6.8 (1.9)	6.5 (2.5)	2.1 (3, 502)	.10
Trust in the health care system (scale 0-10)	5.8 (2.4)	6.5 (2.2)[b]	5.8 (1.6)[b,c]	5.8 (2.4)[d]	4.7 (2.5)[b,c,d]	19.8 (3, 502)	<.001
Trust in science (scale 1-5)	3.5 (0.8)	3.9 (0.6)[b,c,d]	3.6 (0.6)[b,c]	3.4 (0.67)[b,d]	3.1 (0.8)[b,c,d]	46.1 (3, 502)	<.001
Trust in vaccines (scale 1-5)	3.0 (1.3)	3.9 (0.9)[b,c,d,e]	3.1 (0.8)[b,c,d,e]	2.5 (0.9)[b,c,d,e]	1.6 (0.8)[b,c,d,e]	226.6 (3, 502)	<.001
Conspiracy theory tendency (scale 1-5)	2.7 (1.2)	1.9 (1.0)[b,c,d,e]	2.6 (0.8)[b,c,d,e]	3.1 (0.8)[b,c,d,e]	3.6 (1.0)[b,c,d,e]	98.3 (3, 502)	<.001
Following events related to the pandemic (scale 1-5)	2.9 (1.0)	3.1 (1.0)[b,d]	2.9 (0.9)[c]	2.7 (1.0)[b,d]	2.5 (1.0)[b,c]	12.6 (3, 502)	<.001
Health belief model variables							
Perceived health severity (scale 1-5)	2.3 (1.0)	2.4 (0.9)[b]	2.8 (0.9)[b,c]	2.7 (0.9)[d]	1.9 (0.8)[b,c,d]	23.3 (3, 502)	<.001
Perceived financial severity (scale 1-5)	2.3 (1.1)	2.4 (1.2)[b]	2.7 (1.1)[c]	2.3 (1.0)	2.0 (1.00)[b,c]	6.1 (3, 502)	<.001
Perceived benefits (scale 1-5)	3.0 (1.3)	4.0 (0.9)[b,c,d,e]	3.1 (0.7)[b,c,d,e]	2.5 (0.9)[b,c,d,e]	1.8 (0.8)[b,c,d,e]	221.8 (3, 502)	<.001
Fear (scale 1-5)	2.6 (1.2)	2.8 (1.1)[b]	3.0 (1.2)[c]	2.9 (1.0)[d]	1.9 (1.0)[b,c,d]	31.0 (3, 502)	<.001
Health motivation (scale 1-5)	3.5 (0.9)	3.5 (0.9)	3.5 (0.9)	3.6 (0.8)	3.6 (1.0)	0.2 (3, 502)	.87
Theory of planned behavior variables							
Attitudes (scale 1-5)	3.0 (1.5)	4.1 (1.1)[b,c,d,e]	3.0 (1.0)[b,c,d,e]	2.4 (0.9)[b,c,d,e]	1.4 (0.7)[b,c,d,e]	260.2 (3, 502)	<.001
Subjective norms (scale 1-5)	3.0 (1.1)	3.8 (0.9)[b,c,d,e]	3.1 (0.6)[b,c,d,e]	2.5 (0.9)[b,c,d,e]	1.9 (0.8)[b,c,d,e]	159.5 (3, 502)	<.001
Self-efficacy (scale 1-5)	3.7 (0.7)	3.8 (0.7)[b]	3.4 (0.7)[b,c]	3.6 (0.7)	3.7 (0.8)[c]	4.8 (3, 502)	<.001

[a]Games-Howell post hoc test.

[b]Category has a statistically different mean value ($P<.05$) of the corresponding variable in comparison to the mean value in the other category with the same superscript.

[c]Category has a statistically different mean value ($P<.05$) of the corresponding variable in comparison to the mean value in the other category with the same superscript.

[d]Category has a statistically different mean value ($P<.05$) of the corresponding variable in comparison to the mean value in the other category with the same superscript.

[e]Category has a statistically different mean value ($P<.05$) of the corresponding variable in comparison to the mean value in the other category with the same superscript.

Predictors of Vaccine Uptake

The results of multinomial logistic regression models with predictors of vaccine uptake, using "vaccinated respondents" as the reference group for comparisons, are shown in Table 4. Variables included in the multinomial regression model were selected based on their statistical significance in bivariate analyses and their theoretical relevance to the research question

XSL•FO
RenderX

(more details are provided in the Statistical Analysis subsection in the Methods section). Multinomial regression analysis (Table 4) showed that young adults aged 23 to 26 years (compared with those aged 27 to 30 years) were less likely to delay vaccination (aOR 0.1, 95% CI 0.031-0.5). Young adults who had completed elementary school or less (compared to respondents with higher levels of education) were more likely to delay vaccination (aOR 9.8, 95% CI 1.6-61.1) or were hesitant toward vaccination (aOR 11.7, 95% CI 1.5-89.7). Results also showed that young people who were high school students (compared with unemployed respondents) were less likely to be hesitant (aOR 0.0081, 95% CI 0.0002-0.4; Table 4). Similarly, respondents who were students (aOR 0.1, 95% CI 0.0077-0.9) or employed (aOR 0.1, 95% CI 0.0059-0.6; compared with unemployed respondents) were also less likely to refuse vaccination. Participants whose mothers or primary carers had completed elementary school or less were significantly more likely to refuse vaccination (aOR 6.6, 95% CI 1.3-34.4) compared to those with university-educated mothers or primary carers. Similarly, those whose mothers or primary carers finished high school were more likely to refuse vaccination (aOR 2.8, 95% CI 1.0-8.0) compared to the same university-educated group. The results also showed that respondents who felt that COVID-19 could have greater health consequences for them were more likely to delay vaccination (aOR 2.0, 95% CI 1.2-3.3) or be hesitant toward it (aOR 1.9, 95% CI 1.1-3.2). In addition, more negative attitudes toward COVID-19 vaccines were more likely among delayers, hesitant respondents, and refusers than among vaccinated respondents (Table 4). Compared with vaccinated young people, subjective norms were significantly less likely to be present among hesitant respondents (aOR 0.4, 95% CI 0.2-0.7) and refusers (aOR 0.3, 95% CI 0.2-0.7). Self-efficacy to cope with health problems was less present among respondents who delayed vaccination (aOR 0.5, 95% CI 0.3-0.9) than among vaccinated respondents (Table 4). For the data in Table 4, the R^2 was 0.7 (Cox and Snell) and 0.8 (Negelkerke), and the model χ^2 was 610.2 (*df*=81; *P*<.001).

XSL•FO
RenderX

Table 4. Predictors of vaccine uptake by multinomial logistic regression (vaccinated respondents as the reference group).

Predictor variable	Delayers, aOR[a] (95% CI)	Hesitant individuals, aOR (95% CI)	Refusers, aOR (95% CI)
Sociodemographic variables			
Age group (reference: 27-30 years)			
15-18 years	2.9 (0.2-39.4)	22.2 (0.8-60.2)	3.13 (0.1-124.9)
19-22 years	0.6 (0.12-2.6)	0.8 (0.2-3.6)	0.2 (0.035-1.2)
23-26 years	0.1[b] (0.031-0.5)	0.3 (0.1-1.1)	0.4 (0.1-1.5)
Education (reference: college, university, or higher education)			
Elementary school or less	9.8[c] (1.6-61.1)	11.7[b] (1.5-89.7)	6.8 (0.7-65.3)
2-, 3-, 4-, or 5-year high school	1.9 (0.7-5.5)	2.5 (0.9-7.4)	4.0[c] (1.3-12.6)
Activity status (reference: unemployed)			
High-school student	0.2 (0.007-4.1)	0.0081[b] (0.0002-0.4)	0.1 (0.00088-3.0)
Student	0.7 (0.1-6.0)	0.1 (0.015-1.2)	0.1[c] (0.0077-0.9)
Employed	0.4 (0.1-3.1)	0.2 (0.023-1.4)	0.1[c] (0.0059-0.6)
Material status (reference: above average)			
Below average	1.8 (0.4-7.5)	1.3 (0.3-5.6)	3.3 (0.7-17.3)
Average	1.9 (0.7-5.5)	0.6 (0.2-1.9)	1.3 (0.4-4.4)
Mother's or primary carer's education (reference: college, university, or higher education)			
Elementary school or less	2.2 (0.5-9.8)	4.2 (0.9-19.7)	6.6[c] (1.3-34.4)
2-, 3-, 4-, or 5-year high school	0.9 (0.4-2.0)	1.4 (0.5-3.5)	2.8[c] (1.0-8.0)
Settlement type (reference: rural)			
Urban	0.9 (0.4-2.0)	0.9 (0.4-2.1)	0.9 (0.3-2.1)
Health-related variables			
Potential COVID-19 past infection (reference: probably infected)			
Probably not infected	1.2 (0.5-2.9)	0.6 (0.2-1.4)	0.9 (0.3-2.6)
Influenza vaccine in the past year (reference: yes)			
No	1.4 (0.4-5.0)	1.9 (0.4-10.3)	1.5 (0.2-13.0)
Cofounding variables			
Trust in the health care system	1.1 (0.9-1.4)	1.2 (1.0-1.5)	1.0 (0.8-1.3)
Trust in science	2.2 (0.9-4.9)	1.2 (0.5-2.6)	2.1 (0.9-5.0)
Trust in vaccines	0.7 (0.3-1.3)	0.7 (0.4-1.5)	0.5 (0.2-1.1)
Conspiracy theory tendency	0.9 (0.5-1.6)	1.5 (0.8-2.6)	1.3 (0.7-2.3)
Following events related to the pandemic	0.9 (0.6-1.4)	0.8 (0.5-1.2)	0.8 (0.5-1.3)
Health belief model variables			
Perceived health severity	2.0[b] (1.2-3.3)	1.9[b] (1.1-3.2)	0.8 (0.4-1.4)
Perceived financial severity	1.3 (0.9-1.8)	0.8 (0.5-1.2)	0.8 (0.5-1.3)
Perceived benefits	0.7 (0.3-1.3)	0.5 (0.2-1.1)	0.6 (0.3-1.4)
Fear	0.9 (0.6-1.3)	1.2 (0.8-1.8)	0.7 (0.5-1.1)
Theory of planned behavior variables			
Attitudes	0.5[c] (0.3-0.9)	0.5[b] (0.3-0.9)	0.2[d] (0.1-0.4)
Subjective norms	0.7 (0.4-1.3)	0.4[b] (0.2-0.7)	0.3[d] (0.2-0.7)

XSL•FO
RenderX

Predictor variable	Delayers, aOR[a] (95% CI)	Hesitant individuals, aOR (95% CI)	Refusers, aOR (95% CI)
Self-efficacy	0.5[c] (0.3-0.9)	0.9 (0.5-1.7)	1.0 (0.5-2.0)

[a]aOR: adjusted odds ratio.
[b]$P<.01$.
[c]$P<.05$.
[d]$P<.001$.

Discussion

Principal Findings

The aim of this study was to gain a deeper understanding of the determinants influencing vaccination intention and behavior among young people in Slovenia, an EU country where the adult COVID-19 vaccination rate significantly lagged the EU average. We applied an integrated model to analyze and compare vaccinated and unvaccinated young people, further categorizing the unvaccinated people into 3 groups: those hesitant to vaccinate, those willing but delaying vaccination, and those refusing vaccination against COVID-19. Our goal was to enhance the body of research exploring the decision-making processes within these distinct groups of young people regarding COVID-19 vaccination.

In our study, a substantial proportion of respondents self-reported being vaccinated against COVID-19. This rate was slightly higher than the vaccination rate of the general population (aged 18 years or older) in Slovenia, which was approximately 43% for the first dose by the end of August 2021 [5]. Compared to the vaccinated general population, this is a favorable percentage, as previous studies have indicated that younger adults are among the groups with lower willingness to be vaccinated against COVID-19 [8-15]. This finding may be partially explained by policy interventions in Slovenia during the time of data collection. Notably, the introduction of the so called “COVID-19 pass” (certificate of recovery, vaccination, or testing) was designed to facilitate safer and freer mobility of citizens within the EU. Several studies have shown that the implementation of COVID-19 passes improved both vaccination uptake and intention [45].

Our study demonstrated that unvaccinated individuals are not a homogeneous group. Not all respondents who were unvaccinated at the time of data collection had explicitly refused vaccination against COVID-19. Specifically, almost half of these respondents indicated that they had considered getting vaccinated but had not yet done so. Some of these were hesitant, while others expressed a willingness to be vaccinated but had not yet taken action. Only half of the unvaccinated respondents explicitly refused vaccination against COVID-19.

In comparison to the findings of other studies [46], our study did not find gender to be a significant factor influencing vaccination behavior. However, the influence of other sociodemographic and health-related variables on vaccine uptake among young people aligns with previous research [46]. Key predictors of vaccination were age, education level, activity status, material status, mother’s or primary carer’s education, settlement type, potential COVID-19 past infection, and influenza vaccination in the past year. Vaccinated individuals were predominantly aged 23 to 26 years, were more likely to have higher education, and had above-average material status. In contrast, vaccine refusers were more common among the youngest (15-18 years) and oldest (27-30 years) age groups, had lower education levels (elementary or less), had mothers or primary caregivers with lower education, reported below-average material status, were more likely to reside in rural areas, and more frequently reported a potential past COVID-19 infection. These findings suggest that the factors significantly predicting vaccination behavior are closely related to the primary socioeconomic and cultural resources available to an individual [47]. Delayers, while distinct from both vaccinated individuals and refusers, shared some sociodemographic and health-related characteristics with each group. Hesitant respondents, on the other hand, shared more characteristics with refusers, including rural residence, lower education levels, and higher rates of potential past COVID-19 infection.

The factors from the HBM and TPB models that significantly differentiated the vaccinated respondents from all other groups were greater trust in science, trust in vaccines, trust in the health care system, and positive attitudes toward COVID-19 vaccination. These characteristics appeared to be strongly associated with higher levels of education, as suggested by previous studies [48,49]. Perceived health consequences of COVID-19, negative attitudes toward vaccines, and weaker levels of subjective norms were associated with a greater likelihood of delaying or refusing vaccination. Among unvaccinated respondents, perceived self-efficacy in coping with health problems emerged as a significant factor, particularly for those who were delaying vaccination, as this group reported the lowest levels of self-efficacy. Interestingly, despite exhibiting heightened fear of COVID-19 and perceiving the health and financial burdens of contracting the disease as greater than those reported by other groups, delayers may experience affective and cognitive conflict [25]. This conflict likely stems from the simultaneous presence of fear of contracting the disease and negative attitudes toward COVID-19 vaccines. Combined with low confidence in their ability to manage health issues, this internal conflict may explain their expressed willingness to vaccinate but delay in taking action.

The findings of our study indicated that hesitant respondents shared many similarities with delayers, including the same identified conflict between fear of COVID-19 and negative attitudes toward vaccines. However, hesitant respondents exhibited higher self-efficacy in coping with health problems and reported lower levels of subjective norms. Vaccination refusers, in contrast, reported lower levels of subjective norms,

XSL•FO
RenderX

were more likely to have mothers or primary carers with lower levels of education (elementary school or less), and were more likely to hold negative attitudes toward COVID-19 vaccination. Notably, they exhibited a concerning combination of increased self-efficacy in coping with health problems and the highest scores for conspiracy theory tendencies. This profile aligns with a group previously described in the literature as dysfunctionally empowered [50].

Strengths and Limitations

The strength of this study lies in its comprehensive approach to examining the factors influencing COVID-19 vaccine uptake among young people in Slovenia. It addresses a critical research gap by focusing on the specific population group of young people who have frequently demonstrated hesitant attitudes toward vaccination [13]. By comparing vaccinated and unvaccinated individuals and further categorizing the unvaccinated individuals based on their willingness to be vaccinated, the study provides a detailed analysis of the factors underlying vaccination intention and hesitancy. These insights could inform the development of more targeted and effective public health interventions and strategies to address vaccine hesitancy and improve vaccination rates among the youth.

This study also has some limitations. First, data collection through an online access survey panel may have resulted in overrepresentation of certain demographic groups among young adults who have internet access and are more digitally literate. This could exclude other groups and affect the representativeness of the results. Second, the survey relied on self-reported behavior, which is subject to bias. Third, the cross-sectional design of the study precludes the identification of cause-and-effect relationships. Fourth, variables, such as susceptibility, perceived barriers, and cues to action, from the HBM were measured only in the subsample of unvaccinated respondents. As these variables were not included in the analysis, the study may not fully capture all factors influencing vaccine decision-making across the entire sample. Fifth, the multinomial regression analysis results showed wide CIs, indicating greater uncertainty in the estimated effects of the predictor variables on vaccination-related behavior groups. This may result from small sample sizes, overly complex models, sparse data, or other factors. This should be carefully considered when interpreting our findings. Future research using similar methodological frameworks and a similar design should address these issues to enhance the robustness and interpretability of the results. Finally, the data are limited to the social, cultural, and political contexts of Slovenia, where specific COVID-19–related protective measures and crisis communication may have significantly influenced the vaccination behavior of the youth [25].

Implications

Our findings suggest that young people exhibit different attitudes toward vaccination against COVID-19. Consequently, public health authorities should implement diverse strategies tailored to address these attitudes and fears, and the underlying factors influencing them. This recommendation extends beyond COVID-19 and is equally relevant to other diseases, particularly those with lower immunization coverage and vaccination rates [51].

According to the stages of change model [37], which provides a framework for understanding how individuals progress toward a desired behavioral change, such as COVID-19 vaccination, we can categorize vaccine refusers as being in the first stage of behavior change, known as precontemplation. Individuals in this stage are often unaware of the problem, are unmotivated to address it, or may even deny its relevance to themselves. Effective strategies for this group include raising awareness of the issue, providing clear information about the benefits of vaccination, and addressing misconceptions with targeted communication tailored to their specific concerns about vaccine safety and health risks. The groups of delayers and hesitant individuals appear to be either in the contemplation or preparation stage. Those in the contemplation stage are considering vaccination but have not yet made a definitive decision. Strategies for this group should focus on helping them weigh the perceived benefits of vaccination, such as protection against COVID-19 and its potential health and financial consequences, against the perceived costs, such as potential vaccine side effects. Delayers, on the other hand, seem to align more with the preparation stage. They are actively planning to get vaccinated but remain the most uncertain group when it comes to making health-related decisions. To support this group, strategies should include encouragement to finalize their decision, such as prompting them to schedule a vaccination appointment and offering reliable information on the effectiveness, safety, and availability of vaccines.

Conclusions

This cross-sectional study provides survey data on the vaccination behavior (vaccination uptake and refusal) and intention (delayed vaccination and hesitancy) of young people regarding COVID-19 vaccination. It also examines the sociodemographic, health-related, and behavioral predictors of these intentions, integrating constructs from 2 health behavior theories: HBM and TPB.

The findings of our study highlight that young people exhibit diverse behavioral and volitional positions regarding COVID-19 vaccination. Several sociodemographic, health-related, and behavioral factors, such as educational level, material status, subjective norms, perceived self-efficacy, attitudes toward vaccines, fear of contracting the disease, and conspiracy theory tendencies, significantly influence these positions. With regard to vaccination as the desired behavior, individuals are positioned in different stages of behavior change. From a normative perspective, these stages should be carefully considered when designing and implementing measures to promote vaccination for both COVID-19 and other diseases among young people. Greater efforts should be made to target individuals with lower levels of education, those with material disadvantages, and those residing in rural areas.

The findings of this study can guide health policy makers and professionals in optimizing young people's willingness to vaccinate and can provide valuable insights for addressing other vaccinations, future infectious disease outbreaks, and the implementation of effective preventive measures. Interventions

XSL•FO
RenderX

should be tailored to address the primary reasons for nonvaccination, with priority target groups based on their potential for behavior change. Determined refusers may not be the most effective group to focus on initially. Instead, delayers, who demonstrate the greatest willingness to vaccinate among unvaccinated individuals, require targeted strategies, such as those aimed at increasing their self-efficacy in managing health-related decisions.

Acknowledgments

This work was financially supported by the Slovenian Research and Innovation Agency under the research program Social Psychology and Sociology of Everyday Life (research core funding number: P5-0183). The funder played no role in the study design; the collection, analysis, and interpretation of the data; or the decision to submit the article for publication. The funder accepts no responsibility for the content.

Data Availability

The data underlying this article will be shared by the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CHERRIES (Checklist for Reporting Results of Internet E-Surveys).
[PDF File (Adobe PDF File), 145 KB - publichealth_v11i1e64653_app1.pdf]

References

1. Kokić Z, Kon P, Djurković-Djaković O. Effectiveness of vaccination in preventing COVID-19: a community study comparing four vaccines. Vaccines (Basel) 2023 Feb 24;11(3):544 [FREE Full text] [doi: 10.3390/vaccines11030544] [Medline: 36992128]
2. Rahmani K, Shavaleh R, Forouhi M, Disfani HF, Kamandi M, Oskooi RK, et al. The effectiveness of COVID-19 vaccines in reducing the incidence, hospitalization, and mortality from COVID-19: A systematic review and meta-analysis. Front Public Health 2022 Aug 26;10:873596 [FREE Full text] [doi: 10.3389/fpubh.2022.873596] [Medline: 36091533]
3. Omer SB, Yildirim I, Forman HP. Herd immunity and implications for SARS-CoV-2 control. JAMA 2020 Nov 24;324(20):2095-2096. [doi: 10.1001/jama.2020.20892] [Medline: 33074293]
4. Anderson RM, Vegvari C, Truscott J, Collyer BS. Challenges in creating herd immunity to SARS-CoV-2 infection by mass vaccination. The Lancet 2020 Nov;396(10263):1614-1616. [doi: 10.1016/s0140-6736(20)32318-7]
5. COVID-19 Vaccine Tracker. European Center for Disease Control and Prevention. URL: https://vaccinetracker.ecdc.europa.eu/public/ [accessed 2024-06-20]
6. OECD/European Union. Vaccination against COVID-19. In: Health at a Glance: Europe 2022. Paris, France: OECD Publishing; 2022:198-199.
7. Namera cepljenja stagnira in ne preseže 60%. Prebivalstvo je razdeljeno po starosti, oceni lastne rizičnosti ter odnosu do virusa in cepljenja nasploh. Valicon. URL: https://tinyurl.com/yc5hwf9v [accessed 2024-05-25]
8. Aw J, Seng JJB, Seah SSY, Low LL. COVID-19 vaccine hesitancy-a scoping review of literature in high-income countries. Vaccines (Basel) 2021 Aug 13;9(8):900 [FREE Full text] [doi: 10.3390/vaccines9080900] [Medline: 34452026]
9. Freeman D, Loe BS, Yu L, Freeman J, Chadwick A, Vaccari C, et al. Effects of different types of written vaccination information on COVID-19 vaccine hesitancy in the UK (OCEANS-III): a single-blind, parallel-group, randomised controlled trial. The Lancet Public Health 2021 Jun;6(6):e416-e427. [doi: 10.1016/s2468-2667(21)00096-7]
10. Kelly BJ, Southwell BG, McCormack LA, Bann CM, MacDonald PDM, Frasier AM, et al. Predictors of willingness to get a COVID-19 vaccine in the U.S. BMC Infect Dis 2021 Apr 12;21(1):338 [FREE Full text] [doi: 10.1186/s12879-021-06023-9] [Medline: 33845781]
11. Latkin CA, Dayton L, Yi G, Colon B, Kong X. Mask usage, social distancing, racial, and gender correlates of COVID-19 vaccine intentions among adults in the US. PLoS One 2021 Feb 16;16(2):e0246970 [FREE Full text] [doi: 10.1371/journal.pone.0246970] [Medline: 33592035]
12. Prickett KC, Habibi H, Carr PA. COVID-19 vaccine hesitancy and acceptance in a cohort of diverse New Zealanders. Lancet Reg Health West Pac 2021 Sep;14:100241 [FREE Full text] [doi: 10.1016/j.lanwpc.2021.100241] [Medline: 34377999]
13. Schwarzinger M, Watson V, Arwidson P, Alla F, Luchini S. COVID-19 vaccine hesitancy in a representative working-age population in France: a survey experiment based on vaccine characteristics. Lancet Public Health 2021 Apr;6(4):e210-e221 [FREE Full text] [doi: 10.1016/S2468-2667(21)00012-8] [Medline: 33556325]

XSL•FO
RenderX

14. Vollmann M, Salewski C. To get vaccinated, or not to get vaccinated, that is the question: illness representations about COVID-19 and perceptions about COVID-19 vaccination as predictors of COVID-19 vaccination willingness among young adults in The Netherlands. Vaccines (Basel) 2021 Aug 24;9(9):941 [FREE Full text] [doi: 10.3390/vaccines9090941] [Medline: 34579178]
15. Carter J, Rutherford S, Borkoles E. COVID-19 vaccine uptake among younger women in rural Australia. Vaccines (Basel) 2021 Dec 27;10(1):26 [FREE Full text] [doi: 10.3390/vaccines10010026] [Medline: 35062687]
16. Russell FM, Greenwood B. Who should be prioritised for COVID-19 vaccination? Hum Vaccin Immunother 2021 May 04;17(5):1317-1321 [FREE Full text] [doi: 10.1080/21645515.2020.1827882] [Medline: 33141000]
17. Ford C, MacKay M, Thaivalappil A, McWhirter J, Papadopoulos A. COVID-19 vaccine communications on Instagram and vaccine uptake in young adults: a content assessment and public engagement analysis. Emerging Adulthood 2023 Dec 21;12(2):224-235. [doi: 10.1177/21676968231222439]
18. Mitchell H, Lim R, Gill PK, Dhanoa J, Dubé È, Bettinger JA. What do adolescents think about vaccines? Systematic review of qualitative studies. PLOS Glob Public Health 2022;2(9):e0001109 [FREE Full text] [doi: 10.1371/journal.pgph.0001109] [Medline: 36962668]
19. Themistocleous S, Argyropoulos CD, Vogazianos P, Shiamakkides G, Noula E, Nearchou A, et al. Perspectives of European patient advocacy groups on volunteer registries and vaccine trials: VACCELERATE survey study. JMIR Public Health Surveill 2024 Apr 04;10:e47241 [FREE Full text] [doi: 10.2196/47241] [Medline: 38573762]
20. Mateo-Canedo C, Sanabria-Mazo JP, Comendador L, Rojas JS, Carmona M, Crespo-Puig N, et al. Predictive factors of hesitancy to vaccination against SARS-CoV-2 virus in young adults in Spain: Results from the PSY-COVID study. Vaccine X 2023 Aug;14:100301 [FREE Full text] [doi: 10.1016/j.jvacx.2023.100301] [Medline: 37091731]
21. Chaufan C. Is Covid-19 "vaccine uptake" in postsecondary education a "problem"? A critical policy inquiry. Health (London) 2024 Nov 15;28(6):831-857 [FREE Full text] [doi: 10.1177/13634593231204169] [Medline: 37968946]
22. Kelsall J, Sorell T. Two kinds of vaccine hesitancy. Social Epistemology 2024 Apr 15;39(1):40-55. [doi: 10.1080/02691728.2024.2330466]
23. Gao P, Kang L, Liu J, Liu M. Immunogenicity, effectiveness, and safety of COVID-19 vaccines among children and adolescents aged 2-18 years: an updated systematic review and meta-analysis. World J Pediatr 2023 Nov;19(11):1041-1054 [FREE Full text] [doi: 10.1007/s12519-022-00680-9] [Medline: 36723827]
24. Blahut R, Flint A, Orlando E, DesChatelets J, Khowaja A. A scoping review on the decision-making dynamics for accepting or refusing the COVID-19 vaccination among adolescent and youth populations. BMC Public Health 2023 Apr 28;23(1):784 [FREE Full text] [doi: 10.1186/s12889-023-15717-5] [Medline: 37118794]
25. Kamin T, Perger N, Debevec L, Tivadar B. Alone in a time of pandemic: solo-living women coping with physical isolation. Qual Health Res 2021 Jan 19;31(2):203-217. [doi: 10.1177/1049732320971603] [Medline: 33213283]
26. Nilsson S, Mattson J, Berghammer M, Brorsson A, Forsner M, Jenholt Nolbris M, et al. To be or not to be vaccinated against COVID-19 - The adolescents' perspective - A mixed-methods study in Sweden. Vaccine X 2021 Dec;9:100117 [FREE Full text] [doi: 10.1016/j.jvacx.2021.100117] [Medline: 34693273]
27. Crescenzi-Lanna L, Valente R, Cataldi S, Martire F. Predictors of young people's anti-vaccine attitudes in the context of the COVID-19 pandemic. Public Underst Sci 2024 Jan 28;33(1):73-87 [FREE Full text] [doi: 10.1177/09636625231179830] [Medline: 37377105]
28. Zhou Y, Li R, Shen L. Targeting COVID-19 vaccine-hesitancy in college students: An audience-centered approach. J Am Coll Health 2024 Dec 28;72(9):3526-3535. [doi: 10.1080/07448481.2023.2180988] [Medline: 36853986]
29. Huang HY, Gerend MA. The role of trust, vaccine information exposure, and Health Belief Model variables in COVID-19 vaccination intentions: Evidence from an HBCU sample. J Health Psychol 2024 May 05;29(6):621-632. [doi: 10.1177/13591053241227388] [Medline: 38312019]
30. Small ML, Lennon RP, Dziak JJ, Smith RA, Sommerville G, Bharti N. College students' COVID-19 vaccine beliefs and intentions: Implications for interventions. J Am Coll Health 2024 May 27;72(4):989-994. [doi: 10.1080/07448481.2022.2065205] [Medline: 35622961]
31. Shmueli L. Predicting intention to receive COVID-19 vaccine among the general population using the health belief model and the theory of planned behavior model. BMC Public Health 2021 Apr 26;21(1):804 [FREE Full text] [doi: 10.1186/s12889-021-10816-7] [Medline: 33902501]
32. Milošević Đorđević J, Mari S, Vdović M, Milošević A. Links between conspiracy beliefs, vaccine knowledge, and trust: Anti-vaccine behavior of Serbian adults. Soc Sci Med 2021 May;277:113930 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113930] [Medline: 33873008]
33. Prestwich A, Kenworthy J, Conner M. Health Behavior Change: Theories, Methods and Interventions. London, England: Routledge; 2024.
34. McEachan RRC, Conner M, Taylor NJ, Lawton RJ. Prospective prediction of health-related behaviours with the Theory of Planned Behaviour: a meta-analysis. Health Psychology Review 2011 Sep;5(2):97-144. [doi: 10.1080/17437199.2010.521684]

XSL•FO
RenderX

35. McEachan R, Taylor N, Harrison R, Lawton R, Gardner P, Conner M. Meta-analysis of the reasoned action approach (RAA) to understanding health behaviors. Ann Behav Med 2016 Aug 11;50(4):592-612 [FREE Full text] [doi: 10.1007/s12160-016-9798-4] [Medline: 27169555]
36. JazVem. URL: https://www.jazvem.si/ [accessed 2024-11-25]
37. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
38. Chu H, Liu S. Integrating health behavior theories to predict American's intention to receive a COVID-19 vaccine. Patient Educ Couns 2021 Aug;104(8):1878-1886 [FREE Full text] [doi: 10.1016/j.pec.2021.02.031] [Medline: 33632632]
39. Schwarzer R, Jerusalem M. Generalized self-efficacy scale. In: Weinman J, Wright S, Johnston M, editors. Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs. Windsor, UK: NFER-Nelson; 1995.
40. Nadelson L, Jorcyk C, Yang D, Jarratt Smith M, Matson S, Cornell K, et al. I just don't trust them: the development and validation of an assessment instrument to measure trust in science and scientists. School Sci & Mathematics 2014 Jan 19;114(2):76-86. [doi: 10.1111/ssm.12051]
41. Raziskava #Novanormalnost. Valicon. URL: https://www.valicon.net/wp-content/uploads/2020/12/Novanormalnost-15.pdf [accessed 2024-11-25]
42. DeVellis R, Thorpe C. Scale Development: Theory and Applications. Thousand Oaks, CA: SAGE Publications, Inc; 2021.
43. Code of ethics for researchers at the University of Ljubljana. University of Ljubljana. URL: https://www.uni-lj.si/assets/Sluzba-za-raziskovalno-dejavnost/Etika-in-integriteta/EN/Code-of-ethics-for-researchers.pdf [accessed 2025-02-03]
44. World Medical Association. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. Bull World Health Organ 2001;79(4):373-374 [FREE Full text] [Medline: 11357217]
45. Natalia YA, Delporte M, De Witte D, Beutels P, Dewatripont M, Molenberghs G. Assessing the impact of COVID-19 passes and mandates on disease transmission, vaccination intention, and uptake: a scoping review. BMC Public Health 2023 Nov 17;23(1):2279 [FREE Full text] [doi: 10.1186/s12889-023-17203-4] [Medline: 37978472]
46. Biswas MR, Alzubaidi MS, Shah U, Abd-Alrazaq AA, Shah Z. A scoping review to find out worldwide COVID-19 vaccine hesitancy and its underlying determinants. Vaccines (Basel) 2021 Oct 25;9(11):1243 [FREE Full text] [doi: 10.3390/vaccines9111243] [Medline: 34835174]
47. Kamin T, Anker T. Cultural capital and strategic social marketing orientations. Journal of Social Marketing 2014;4(2):94-110. [doi: 10.1108/JSOCM-08-2013-0057]
48. Stead M, Jessop C, Angus K, Bedford H, Ussher M, Ford A, et al. National survey of attitudes towards and intentions to vaccinate against COVID-19: implications for communications. BMJ Open 2021 Oct 28;11(10):e055085 [FREE Full text] [doi: 10.1136/bmjopen-2021-055085] [Medline: 34711602]
49. Miller JD, Ackerman MS, Laspra B, Polino C, Huffaker JS. Public attitude toward Covid-19 vaccination: The influence of education, partisanship, biological literacy, and coronavirus understanding. FASEB J 2022 Jul 03;36(7):e22382 [FREE Full text] [doi: 10.1096/fj.202200730] [Medline: 35657606]
50. Petrič G, Atanasova S, Kamin T. Impact of social processes in online health communities on patient empowerment in relationship with the physician: emergence of functional and dysfunctional empowerment. J Med Internet Res 2017 Mar 13;19(3):e74 [FREE Full text] [doi: 10.2196/jmir.7002] [Medline: 28288953]
51. Immunization coverage. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/immunization-coverage [accessed 2024-05-25]

Abbreviations

aOR: adjusted odds ratio
EFA: exploratory factor analysis
ESOMAR: European Society for Opinion and Market Research
EU: European Union
HBM: health belief model
TPB: theory of planned behavior

XSL•FO
RenderX

Edited by C Argyropoulos; submitted 23.07.24; peer-reviewed by T Ricks, N Kaur, M Sarhan; comments to author 18.10.24; revised version received 11.12.24; accepted 19.01.25; published 21.02.25.

Please cite as:
Atanasova S, Kamin T, Perger N
Predictors of COVID-19 Vaccination Intention and Behavior Among Young People in a European Union Country With Low COVID-19 Vaccination Rates: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e64653
URL: https://publichealth.jmir.org/2025/1/e64653
doi:10.2196/64653
PMID:39983109

©Sara Atanasova, Tanja Kamin, Nina Perger. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

SARS-CoV-2 Antibody Prevalence Across Unvaccinated Health Care Workers During the COVID-19 Pandemic in Yemen: Cross-Sectional Study

Eihab Al-Herwi[1], PhD; Maryam Baras[2], MSc; Khadega Alhetar[1], PhD; Mohammed A Farhan[1], PhD; Rashad M Saleh[1], PhD; Amani As-suhbani[1], PhD; Manal Alshoaibi[3], BSc; Mansour Mohammed[4], BSc; Zainab Morshed[4], BSc; Asia Shujaa-Aldeen[5], BSc; Fatima Ghanem[5], BSc; Bra'ah AlQisi[5], BSc; Mohammed Alherwi[6], MSc; Tony Bruns[7], MD

[1]Medical Laboratories Department, Faculty of Medicine and Health Sciences, Ibb University, University Street, Ibb, Yemen

[2]National Centre of Public Health Laboratories, Al-Mukalla, Yemen

[3]National Centre of Public Health Laboratories, Ibb, Yemen

[4]Medical Microbiology, Ibb University, Ibb, Yemen

[5]Medical Laboratories, International Malaysian University, Ibb, Yemen

[6]Pharmacy Department, University of Science and Technology, Ibb, Yemen

[7]Department of Internal Medicine III, University Hospital Rheinisch-Westfälische Technische Hochschule, Aachen, Germany

Corresponding Author:
Eihab Al-Herwi, PhD
Medical Laboratories Department, Faculty of Medicine and Health Sciences, Ibb University, University Street, Ibb, Yemen

Abstract

Background: The COVID-19 pandemic presented significant challenges to health care centers across Yemen. The lack of access to COVID-19 vaccines and limited availability of personal protective equipment greatly increased the risk of SARS-CoV-2 exposure among health care workers (HCWs). Only a few studies have examined the seroprevalence and burden of SARS-CoV-2 among Yemeni HCWs.

Objective: This study aimed to assess the seroprevalence of SARS-CoV-2 and associated risk factors among a cohort of unvaccinated HCWs in Ibb City, the capital of Ibb Governorate, located in the highlands of southwestern Yemen between July 2022 and January 2023.

Methods: Unvaccinated HCWs employed in public and private hospitals, dispensaries, pharmacies, and laboratories in Ibb City during the past 6 months were eligible. Blood samples, occupational information, and structured interviews using a questionnaire were collected from a convenience sample of 396 unvaccinated HCWs actively providing health care services between July 2022 and January 2023. SARS-CoV-2 antibody presence was determined using a lateral flow immunoassay.

Results: Of the 396 HCWs tested, 268 (67.7%) were positive for SARS-CoV-2 antibodies, with no significant difference in seropositivity between sex (*P*=.29). Key factors associated with seropositivity included occupation and workplace. Compared to laboratory technicians (76/124, 61%), nurses (93/124, 75%; odds ratio [OR] 1.89, 95% CI 1.10 - 3.26; *P*=.02) and physician assistants (13/14, 92.9%; OR 8.21, 95% CI 1.04 - 64.79; *P*=.046) had significantly higher odds of seropositivity. Similarly, working in hospitals was associated with significantly higher odds of seropositivity compared to working in laboratories (OR 2.77, 95% CI 1.59 - 4.81; *P*<.001). Overall, 82% (219/268) of seropositive HCWs reported COVID-19–related symptoms within the last 6 months (OR 3.82, 95% CI 2.40 - 6.09; *P*<.001), the majority being fever (191/256, 74.6%; OR 2.40, 95% CI 1.56 - 3.72; *P*<.001), headache (175/230, 76.1%; OR 2.50, 95% CI 1.62 - 3.84; *P*<.001), cough (162/205, 79%; OR 3.02, 95% CI 1.94 - 4.70; *P*<.001), or loss of taste (155/202, 76.7%; OR 2.36, 95% CI 1.53 - 3.65; *P*<.001) or smell (146/191, 76.4%; OR 2.21, 95% CI 1.43 - 3.41; *P*<.001).

Conclusions: This study reveals a high prevalence of SARS-CoV-2 antibodies among HCWs in Ibb City, Yemen, underscoring the impact of limited vaccination and personal protective equipment availability in 2022 and 2023. These findings highlight the urgent need for improved protective measures and vaccination efforts in conflict-affected regions.

(JMIR Public Health Surveill 2025;11:e67296) doi:10.2196/67296

KEYWORDS

SARS-CoV-2 antibodies; COVID-19; health care workers; risk factors; Yemen

XSL•FO
RenderX

Introduction

On March 11, 2020, the World Health Organization (WHO) officially declared the outbreak of COVID-19 a global pandemic [1]. Certain variants of SARS-CoV-2 have demonstrated enhanced replication efficiency, increased fitness, and heightened transmissibility, which may elevate the risk of reinfection [2]. As of August 4, 2024, the WHO reported more than 775 million confirmed cases of COVID-19 worldwide with more than 7 million related deaths [1]. The clinical spectrum of COVID-19 ranges from asymptomatic infection to severe viral pneumonia, acute respiratory distress syndrome, and death [2]. Identifying risk factors for SARS-CoV-2 infections [3] and for progression to severe disease or COVID-19–related hospitalization [4] is vital to guide prevention strategies and clinical management.

In conflict-affected and resource-poor countries like Yemen, where political instability and ongoing conflict have severely weakened the health care infrastructure, the pandemic has exacerbated existing challenges, with limited access to vaccines and personal protective equipment (PPE). Years of conflict with airstrike and bombardment since 2015 have left hospitals destroyed, equipment scarce, and health care workers (HCWs) struggling to provide care [5,6]. The WHO Regional Office for the Eastern Mediterranean reports fewer than 12,000 cumulative cases of COVID-19 until February 2024, with an estimated overall case fatality rate of 18.1% [7]. However, these numbers are likely to be inaccurate for several reasons, such as limited testing capacity and inconsistent data collection practices [8].

HCWs are positioned on the frontlines and face an elevated risk of contracting SARS-CoV-2 and potentially transmitting it to patients and others [9]. HCWs in conflict zones frequently operate in settings characterized by limited resources, insufficient PPE, and weakened health care infrastructure, contributing to elevated transmission rates [10]. In regions where the SARS-CoV-2 vaccine is unavailable and PPE supplies are limited, there are no clear data on the seroprevalence of the disease. As seroprevalence studies are difficult to carry out in war-torn nations, only a few studies have attempted to estimate the true burden of disease in Yemen [11,12], and only a few [13,14] have examined seroprevalence and COVID-19 burden among Yemeni HCWs.

Recognizing the limited data on SARS-CoV-2 seroprevalence among Yemeni HCWs, this study sought to determine the frequency of SARS-CoV-2 antibodies and investigate associated risk factors within an unvaccinated HCW cohort in Ibb, Yemen.

Methods

Patients and Setting

This cross-sectional study recruited 396 HCWs employed in public and private hospitals, dispensaries, pharmacies, and laboratories in Ibb City, the capital of Ibb Governorate located in the highlands of southwestern Yemen, between July 2022 and January 2023. A convenience sampling strategy was used for recruitment. Eligibility criteria for study inclusion required participants to be adults (age ≥18 years), employed in a health care setting for at least the 6 months preceding enrollment, unvaccinated against COVID-19, and to have provided written informed consent. The sole exclusion criterion was the previous receipt of 1 or more doses of any COVID-19 vaccine. To implement the convenience sampling, we visited various health care facilities during working hours. Eligible HCWs who were present, available, and willing to participate during the researchers' visits were enrolled consecutively during the study period. For this, HCWs were visited at their workplaces during working hours and participated in structured interviews based on a predesigned questionnaire. Occupational and clinical data were collected from all participants during sample collection. HCWs were categorized according to the Yemen Medical Council's classification, including physicians, nurses, dentists, pharmacists, physician assistants, and laboratory technicians.

SARS-CoV-2 Antibody Testing

Capillary whole-blood samples were obtained from all participants, and SARS-CoV-2 IgM and IgG antibodies were detected using a lateral flow immunoassay test according to the manufacturer's instructions (Clungene, Shanghai International Holding Corp). Briefly, 5 µL of whole blood and 80 µL of sample buffer were applied. If antibodies against SARS-CoV-2 were present, they bound to the labeled antigens, forming antigen-antibody complexes that migrated to the test lines, where immobilized capture antibodies specific for SARS-CoV-2 antibodies were located. Clinical symptoms within the past 6 months were recorded using a standardized questionnaire.

Statistical Analysis

Data analysis was performed using SPSS (version 25.0; IBM Corp). Chi-square tests were used to assess associations between categorical variables. Logistic regression analyses were conducted to calculate odds ratios (ORs) and 95% CIs. The prevalence of positive test results was analyzed in relation to sex, age groups, occupation types, and symptoms. Statistical significance was defined as $P<.05$ for all tests without correction for multiple testing.

Ethical Considerations

The study received ethical approval from the Research Ethics Committee of Ibb University, Faculty of Medicine and Health Sciences (registration number IBBUNI.AC.YEM. 2022.35). Informed consent was obtained from all participants. Before analysis, participants' data were deidentified using pseudonymization. Participants were not financially compensated for participation in the study.

Results

Between July 2022 and January 2023, capillary blood samples from 396 HCWs from Ibb City, Yemen, were collected. The mean age of enrolled participants was 28.6 (SD 7.8) years. The majority of HCWs in our study were nurses and laboratory technicians, accounting for 62.6% (248/396) of the participants (Table 1). None of the study participants were vaccinated.

XSL•FO
RenderX

Table . Demographic characteristics, occupation, and SARS-CoV-2 seropositivity among participating health care workers in Ibb City, Yemen (2022 - 2023).

Characteristics	Number of examinees, n	Positive test, n (%)	Odds ratio (95% CI)	*P* value[a]
All participants	396	268 (67.7)		
Sex				
Female	207	145 (70)	Reference	
Male	189	123 (65.1)	0.80 (0.52 - 1.22)	.29
Age groups (years)				
<30	264	176 (66.7)	Reference	
30 - 39	92	59 (64.1)	0.89 (0.54 - 1.47)	.66
40 - 49	28	23 (82.1)	2.30 (0.85 - 6.26)	.10
≥50	12	10 (83.3)	2.50 (0.54 - 11.66)	.24
Occupation				
Laboratory technician	124	76 (61.3)	Reference	
Physician	66	46 (69.7)	1.45 (0.77 - 2.75)	.25
Nurse	124	93 (75)	1.89 (1.10 - 3.26)	.02
Dentist	27	21 (77.8)	2.21 (0.83 - 5.87)	.11
Pharmacist	41	19 (46.3)	0.55 (0.27 - 1.11)	.10
Physician assistant	14	13 (92.9)	8.21 (1.04 - 64.79)	.046
Working place				
Laboratory	67	34 (50.7)	Reference	
Hospital	262	194 (74)	2.77 (1.59 - 4.81)	<.001
Private clinic	29	18 (62.1)	1.59 (0.65 - 3.87)	.31
Dispensary	16	10 (62.5)	1.62 (0.53 - 4.96)	.84
Pharmacy	22	12 (54.5)	1.16 (0.44 - 3.01)	.76
Symptoms[b]				
No	108	49 (45.4)	Reference	
Yes	288	219 (76)	3.82 (2.40 - 6.09)	<.001

[a]*P* values were calculated by chi-square test.

[b]Within the past 6 months.

The overall seropositivity rate among HCWs was 67.7% (268/396), with no significant difference in seropositivity between sexes: 145 of 207 (70%) in females and 123 of 189 (65.1%) in males. HCWs aged 40 years or older (33/40, 82.5%) were associated with the higher odds of seropositivity (OR 2.42, 95% CI 1.04 - 5.65) compared with those younger than 40 years (235/356, 66%; *P*=.03).

Compared with laboratory technicians as a reference group, the odds of being seropositive were significantly higher for physician assistants (OR 8.21, 95% CI 1.04 - 64.79) and nurses (OR 1.89, 95% CI 1.10 - 3.26). Dentists also showed higher odds of seropositivity (21/27, 77.8%; OR 2.21, 95% CI 0.83 - 5.87), although this association was not statistically significant. Similarly, compared with working in laboratories, working in a hospital setting was associated with significantly higher odds of seropositivity (OR 2.77, 95% CI 1.59 - 4.81; Table 1).

Among seropositive individuals, 81.7% (219/268) reported COVID-19–related symptoms within the last 6 months, the majority being fever (191/268, 71.2%), headache (175/268, 65.3%), cough (162/268, 60.4%), loss of taste (155/268, 57.8%), or smell (146/268, 54.4%; Table 2).

XSL•FO
RenderX

Table . Association between signs and symptoms and SARS-CoV-2 seropositivity among included health care workers in Ibb City, Yemen (2022 - 2023).

Symptoms	Examinees, n	Positive test, n (%)	Odds ratio (95% CI)	*P* value[a]
Fever				
No	140	77 (55)	Reference	
Yes	256	191 (74.6)	2.40 (1.56 - 3.72)	<.001
Chills				
No	233	143 (61.4)	Reference	
Yes	163	125 (76.7)	2.07 (1.32 - 3.24)	.001
Cough				
No	191	106 (55.5)	Reference	
Yes	205	162 (79)	3.02 (1.94 - 4.70)	<.001
Headache				
No	166	93 (56)	Reference	
Yes	230	175 (76.1)	2.50 (1.62 - 3.84)	<.001
Breathing problem				
No	296	197 (66.6)	Reference	
Yes	100	71 (71)	1.23 (0.75 - 2.02)	.41
Muscle pain				
No	230	145 (63)	Reference	
Yes	166	123 (74.1)	1.68 (1.08 - 2.60)	.02
Loss of taste				
No	194	113 (58.2)	Reference	
Yes	202	155 (76.7)	2.36 (1.53 - 3.65)	<.001
Loss of smell				
No	205	122 (59.5)	Reference	
Yes	191	146 (76.4)	2.21 (1.43 - 3.41)	<.001
Sore throat				
No	262	162 (61.8)	Reference	
Yes	134	106 (79.1)	2.34 (1.44 - 3.80)	<.001
Congestion or runny nose				
No	258	163 (63.2)	Reference	
Yes	138	105 (76.1)	1.85 (1.16 - 2.96)	.009
Nausea or vomiting				
No	346	234 (67.6)	Reference	
Yes	50	34 (68.0)	1.02 (0.549 - 1.92)	.96
Diarrhea				
No	354	237 (66.9)	Reference	
Yes	42	31 (73.8)	1.39 (0.68 - 2.87)	.37

[a]The *P* values were calculated by chi-square test.

Having experienced symptoms within the past 6 months was associated with significantly higher odds of SARS-CoV-2 seropositivity (OR 3.82, 95% CI 2.40 - 6.09). Specifically, participants who reported experiencing symptoms such as fever, chills, cough, headache, muscle aches, loss of taste, loss of smell, sore throat, and congestion (runny nose) had increased odds of being seropositive (Table 2). Among these, cough was associated with the highest odds of seropositivity (OR 3.021, 95% CI 1.94 - 4.70; *P*<.001). In contrast, gastrointestinal symptoms, including nausea, vomiting, and diarrhea, did not demonstrate a positive association with SARS-CoV-2 seropositivity.

XSL•FO
RenderX

Discussion

Principal Findings

We herein report the serostatus for SARS-CoV-2 in 396 unvaccinated HCWs in Ibb City in 2022-2023 to be 67%. Seropositivity was highest in individuals working in a hospital (194/268, 74%) as compared with laboratory (34/268, 50.7%) or pharmacy staff (12/268, 54.5%).

Since the first laboratory-confirmed case of COVID-19 was reported in April 2020 in Mukalla City [15], Yemeni HCWs have faced the pandemic amid numerous challenges, including a scarcity of PPE, the absence of vaccines in many cities, ongoing conflict, and various crises affecting the country.

During the early stages of the pandemic, studies worldwide reported seropositivity rates in unvaccinated HCWs ranging from 8% to more than 40%, with large differences between countries and hospitals [16-18]. After the first wave of COVID-19 in Yemen, between May and September 2020, 19.4% of asymptomatic HCWs at the Médecins Sans Frontières Trauma Center in Aden, Yemen, tested seropositive for SARS-CoV-2 antibodies [13]. Several factors have contributed to the lack of COVID-19 vaccines in many regions of Yemen, including the ongoing conflict that has severely disrupted the health care system and damaged infrastructure, hindering vaccine delivery to remote or conflict-affected areas; limited resources; and misinformation or lack of awareness about the importance of vaccination. During the conflict, Ibb was deemed safer, attracting migrants and increasing population density. This influx strained limited health care services, complicating COVID-19 control efforts.

From late 2022 to early 2023, many countries, including those in the Arab world, began to consider COVID-19 an endemic disease as infection rates stabilized and COVID-19 cases became more predictable with fewer severe outcomes [19]. The case fatality rate of confirmed infections gradually decreased in Taiz Governorate in southern Yemen from 27% in 2020 to 12.8% and 5.8% in 2021 and 2022, respectively [12]. In parallel, the seroprevalence of SARS-CoV-2 among HCWs from 9 hospitals in the Lahj and Al-Dhalea governorates of Yemen increased to 94.2% between June 2022 and September 2022, underscoring the transition to an endemic situation [14]. Consistent with our findings, a meta-analysis of COVID-19 seroprevalence conducted across 49 studies involving 127,480 HCWs indicated that antibody prevalence varied from 0% to 45.3%. Notably, one of the key factors associated with higher seroprevalence rates was the shortage of PPE [14]. In line with our study, the war-torn conditions and limited availability of PPE had a significant impact on the increase in antibody prevalence (56.5%) of COVID-19 among HCWs in Tigray, Ethiopia [14]. Although this study also included several professions, such as nurses, radiologists, dentists, laboratory personnel, pharmacists, respiratory therapists, and nutritionists, we here report a considerably lower seroprevalence in a comparable time frame with significant occupation-related differences. In our study, direct patient care as a physician assistant, dentist, nurse, or physician was associated with the highest odds of seropositivity.

Our findings align with other studies that identified higher odds of infection or seropositivity among health care staff who provided direct patient care in confined spaces, had prolonged or repeated face-to-face contact, or had contact with the environment or materials used by patients with COVID-19, which presumably contributed to the increased risk in these occupational groups [20-22]. The inadequate availability of PPE in hospitals may also have contributed to this risk, given the higher number of HCWs in hospitals compared with laboratories. Other studies [23-25] have found no significant correlation between occupation and infection, possibly reflecting differences in the availability and use of PPE and possible exposure from colleagues and household members.

In our study, seroprevalence of SARS-CoV-2 was highest among HCWs aged 40 years or older (33/40, 82.5%), compared with those younger than 40 years (235/356, 66%). This contrasts with studies from the Western countries, where higher seropositivity rates have been reported among younger HCWs, possibly due to increased exposure through social behaviors and patient contact, as well as differences in adherence to stringent protective measures [16,26].

Our analysis indicated significantly higher odds of SARS-CoV-2 seropositivity associated with typical symptoms such as fever, chills, headache, loss of taste, loss of smell, and sore throat. Other studies in HCWs have highlighted a strong association of loss of taste or smell with seropositivity [27-29]. In our study, loss of taste or smell was observed in 57.8% (155/268) of seropositive HCWs and associated with 2.3-fold odds. Notably, anosmia within the last 6 months was also reported by 45 (35%) seronegative HCWs, which is considerably higher than the self-reported olfactory loss (<15%) or objective smell dysfunction (<25%) in the general population [30]. This might, in part, be explained by false-negative results from the assays used, which represents a limitation of the study.

Several limitations warrant discussion. First, the recruitment strategy used may have introduced selection bias, potentially limiting the generalizability of our findings. Although a large sample size of 396 HCWs was enrolled, this sample may not be fully representative of all unvaccinated HCWs in Ibb, Yemen. Furthermore, the composition of the study sample might not accurately reflect the overall workforce in the specific health care facilities from which participants were drawn. This may affect the generalizability of the results to other regions or health care settings. In addition, this study relied on structured interviews and self-reported data for occupational information and COVID-19–related symptoms, which may have introduced recall bias or inaccuracies in reporting.

In summary, the observed seroprevalence of SARS-CoV-2 among HCWs in Ibb City, Yemen, highlights the ongoing vulnerability of HCWs, particularly those with close patient contact, in resource-limited and conflict-affected regions. These findings underscore the urgent need for sustained efforts to enhance infection control measures, including the provision of PPE and broader access to vaccines. Monitoring occupation-related risks and the evolving epidemiological landscape is crucial for informing future public health strategies.

XSL•FO
RenderX

Conclusions

This study reveals a high prevalence of SARS-CoV-2 antibodies among unvaccinated HCWs in Ibb City, indicating significant exposure to the virus during the pandemic. The analysis identified key risk factors, particularly occupation and workplace settings, with nurses and physician assistants having elevated odds of infection. These findings highlight the urgent need for targeted interventions to protect HCWs, including enhanced safety protocols and vaccination campaigns. Furthermore, the data suggest the necessity for ongoing surveillance and support for HCWs, especially in high-risk environments such as hospitals, to mitigate the spread of COVID-19 and ensure the safety of both health care providers and patients. Continued research is essential to understand the long-term implications of COVID-19 exposure among HCWs and to develop targeted interventions that can more effectively protect this critical workforce in similar settings worldwide.

Acknowledgments

The authors would like to express their gratitude to all health care workers who participated in this study in both public and private health care centers in Ibb, Yemen. No specific funding was obtained or sought for this research.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

EA was involved in all aspects of this article and is the guarantor for the data. Sample collection was carried out by KH, MM, ZM, ASA, and BA. MA, FA, and MA performed laboratory procedures. EA, MB, MAF, and TB analyzed the data and wrote the manuscript. RMS and AA reviewed the manuscript.

Conflicts of Interest

None declared.

References

1. WHO COVID-19 dashboard. World Health Organization. URL: https://data.who.int/dashboards/covid19/deaths [accessed 2025-08-26]
2. Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020 Apr 30;382(18):1708-1720. [doi: 10.1056/NEJMoa2002032] [Medline: 32109013]
3. Chadeau-Hyam M, Bodinier B, Elliott J, et al. Risk factors for positive and negative COVID-19 tests: a cautious and in-depth analysis of UK biobank data. Int J Epidemiol 2020 Oct 1;49(5):1454-1467. [doi: 10.1093/ije/dyaa134] [Medline: 32814959]
4. Ko JY, Danielson ML, Town M, et al. Risk factors for coronavirus disease 2019 (COVID-19)-associated hospitalization: COVID-19-associated hospitalization surveillance network and behavioral risk factor surveillance system. Clin Infect Dis 2021 Jun 1;72(11):e695-e703. [doi: 10.1093/cid/ciaa1419] [Medline: 32945846]
5. Alshareefy Y. The promise of peace talks for Yemen's health-care system. Lancet 2023 Sep 9;402(10405):849. [doi: 10.1016/S0140-6736(23)01371-5] [Medline: 37689406]
6. Mousavi SM, Anjomshoa M. COVID-19 in Yemen: a crisis within crises. Int J Equity Health 2020 Jul 11;19(1):120. [doi: 10.1186/s12939-020-01231-2] [Medline: 32652995]
7. COVID-19 situation in the region - total reports. World Health Organization Eastern Mediterranean Region. URL: http://www.emro.who.int/health-topics/corona-virus/index.html [accessed 2024-08-20]
8. Noushad M, Al-Saqqaf IS. COVID-19 case fatality rates can be highly misleading in resource-poor and fragile nations: the case of Yemen. Clin Microbiol Infect 2021 Apr;27(4):509-510. [doi: 10.1016/j.cmi.2021.01.002] [Medline: 33429029]
9. Gómez-Ochoa SA, Franco OH, Rojas LZ, et al. COVID-19 in health-care workers: a living systematic review and meta-analysis of prevalence, risk factors, clinical characteristics, and outcomes. Am J Epidemiol 2021 Jan 4;190(1):161-175. [doi: 10.1093/aje/kwaa191] [Medline: 32870978]
10. Almhawish N, Karah N, Elferruh Y, Aksh A, Abbara A. Protecting healthcare workers in conflict zones during the COVID-19 pandemic: Northwest Syria. J Infect 2021 May;82(5):186-230. [doi: 10.1016/j.jinf.2021.01.027] [Medline: 33545168]
11. Bin-Ghouth AS, Al-Shoteri S, Mahmoud N, et al. SARS-CoV-2 seroprevalence in Aden, Yemen: a population-based study. Int J Infect Dis 2022 Feb;115:239-244. [doi: 10.1016/j.ijid.2021.12.330] [Medline: 34929358]
12. Edrees WH, Abdullah QY, Al-Shehari WA, Alrahabi LM, Khardesh AAF. COVID-19 pandemic in Taiz Governorate, Yemen, between 2020 and 2023. BMC Infect Dis 2024 Jul 25;24(1):739. [doi: 10.1186/s12879-024-09650-0] [Medline: 39060993]
13. Malaeb R, Yousef N, Al-Nagdah O, et al. High seroprevalence of antibodies against SARS-CoV-2 among healthcare workers 8 months after the first wave in Aden, Yemen. PLOS Glob Public Health 2022;2(11):e0000767. [doi: 10.1371/journal.pgph.0000767] [Medline: 36962647]

XSL•FO
RenderX

14. Taher WT, Bawazir AA, Sallam TA, Alsurimi K. Seroprevalence and factors associated with SARS-CoV-2 infection among healthcare workers: cross-sectional study. BMC Infect Dis 2023 Nov 6;23(1):761. [doi: 10.1186/s12879-023-08760-5] [Medline: 37932664]
15. Dureab F, Al-Awlaqi S, Jahn A. COVID-19 in Yemen: preparedness measures in a fragile state. Lancet Public Health 2020 Jun;5(6):e311. [doi: 10.1016/S2468-2667(20)30101-8] [Medline: 32334648]
16. Houlihan CF, Vora N, Byrne T, et al. Pandemic peak SARS-CoV-2 infection and seroconversion rates in London frontline health-care workers. Lancet 2020 Jul 25;396(10246):e6-e7. [doi: 10.1016/S0140-6736(20)31484-7] [Medline: 32653078]
17. Kayı İ, Madran B, Keske Ş, et al. The seroprevalence of SARS-CoV-2 antibodies among health care workers before the era of vaccination: a systematic review and meta-analysis. Clin Microbiol Infect 2021 Sep;27(9):1242-1249. [doi: 10.1016/j.cmi.2021.05.036] [Medline: 34116205]
18. Stubblefield WB, Talbot HK, Feldstein LR, et al. Seroprevalence of SARS-CoV-2 among frontline healthcare personnel during the first month of caring for patients with COVID-19—Nashville, Tennessee. Clin Infect Dis 2021 May 4;72(9):1645-1648. [doi: 10.1093/cid/ciaa936] [Medline: 32628750]
19. Cohen C, Pulliam J. COVID-19 infection, reinfection, and the transition to endemicity. Lancet 2023 Mar 11;401(10379):798-800. [doi: 10.1016/S0140-6736(22)02634-4] [Medline: 36930672]
20. Kishk RM, Nemr N, Aly HM, et al. Assessment of potential risk factors for coronavirus disease-19 (COVID-19) among health care workers. J Infect Public Health 2021 Oct;14(10):1313-1319. [doi: 10.1016/j.jiph.2021.07.004] [Medline: 34281792]
21. Bastuji-Garin S, Brouard L, Bourgeon-Ghittori I, et al. The relative contributions of occupational and community risk factors for COVID-19 among hospital workers: the HOP-COVID cohort study. J Clin Med 2023 Feb 2;12(3):1208. [doi: 10.3390/jcm12031208] [Medline: 36769854]
22. Lartey M, Kenu E, Ganu VJ, et al. Risk factors for COVID-19 infections among health care workers in Ghana. PLoS ONE 2023;18(7):e0288242. [doi: 10.1371/journal.pone.0288242] [Medline: 37410736]
23. Kosenkow J, Ankert J, Baier M, Kesselmeier M, Pletz MW. COVID-19 outbreak among employees of a German hospital: risk factor analysis based on a follow-up questionnaire and seroprevalence. Infection 2024 Oct;52(5):1753-1762. [doi: 10.1007/s15010-024-02220-1] [Medline: 38488974]
24. Bahrs C, Weis S, Kesselmeier M, et al. Non-patient-related SARS-CoV-2 exposure from colleagues and household members poses the highest infection risk for hospital employees in a German university hospital: follow-up of the prospective Co-HCW seroprevalence study. Infection 2023 Aug;51(4):1051-1059. [doi: 10.1007/s15010-023-01995-z] [Medline: 36788173]
25. Çelebi G, Pişkin N, Çelik Bekleviç A, et al. Specific risk factors for SARS-CoV-2 transmission among health care workers in a university hospital. Am J Infect Control 2020 Oct;48(10):1225-1230. [doi: 10.1016/j.ajic.2020.07.039] [Medline: 32771498]
26. Iversen K, Bundgaard H, Hasselbalch RB, et al. Risk of COVID-19 in health-care workers in Denmark: an observational cohort study. Lancet Infect Dis 2020 Dec;20(12):1401-1408. [doi: 10.1016/S1473-3099(20)30589-2] [Medline: 32758438]
27. Lombardi A, Consonni D, Carugno M, et al. Characteristics of 1573 healthcare workers who underwent nasopharyngeal swab testing for SARS-CoV-2 in Milan, Lombardy, Italy. Clin Microbiol Infect 2020 Oct;26(10):1413. [doi: 10.1016/j.cmi.2020.06.013] [Medline: 32569835]
28. Wratil PR, Schmacke NA, Osterman A, et al. In-depth profiling of COVID-19 risk factors and preventive measures in healthcare workers. Infection 2022 Apr;50(2):381-394. [doi: 10.1007/s15010-021-01672-z] [Medline: 34379308]
29. Buonafine CP, Paiatto BNM, Leal FB, et al. High prevalence of SARS-CoV-2 infection among symptomatic healthcare workers in a large university tertiary hospital in São Paulo, Brazil. BMC Infect Dis 2020 Dec 2;20(1):917. [doi: 10.1186/s12879-020-05662-8] [Medline: 33267836]
30. Yang J, Pinto JM. The epidemiology of olfactory disorders. Curr Otorhinolaryngol Rep 2016 May;4(2):130-141. [doi: 10.1007/s40136-016-0120-6] [Medline: 30613439]

Abbreviations

HCW: health care worker
OR: odds ratio
PPE: personal protective equipment
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 09.10.24; peer-reviewed by M Gasmi, O Olasunkanmi; revised version received 17.06.25; accepted 18.06.25; published 29.08.25.

Please cite as:
Al-Herwi E, Baras M, Alhetar K, Farhan MA, Saleh RM, As-suhbani A, Alshoaibi M, Mohammed M, Morshed Z, Shujaa-Aldeen A, Ghanem F, AlQisi B, Alherwi M, Bruns T
SARS-CoV-2 Antibody Prevalence Across Unvaccinated Health Care Workers During the COVID-19 Pandemic in Yemen: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e67296
URL: https://publichealth.jmir.org/2025/1/e67296
doi:10.2196/67296

©Eihab Al-Herwi, Maryam Baras, Khadega Alhetar, Mohammed A Farhan, Rashad M Saleh, Amani As-suhbani, Manal Alshoaibi, Mansour Mohammed, Zainab Morshed, Asia Shujaa-Aldeen, Fatima Ghanem, Bra'ah AlQisi, Mohammed Alherwi, Tony Bruns. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 29.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Net Reproduction Number as a Real-Time Metric of Population Reproducibility

Chiara Achangwa[1], PhD; Changhee Han[2], MS; Jun-Sik Lim[3], DVM, MPH; Seonghui Cho[1], BBA; Sangbum Choi[4], PhD; Sukhyun Ryu[1], MD, PhD

[1]Department of Preventive Medicine, College of Medicine, The Catholic University of Korea, R6117, Omibus Park, 222 Banpo-daero, Seoul, Republic of Korea

[2]Computer Information System, Georgia State University, Atlanta, GA, United States

[3]IHAP, Université de Toulouse, INRAE, ENVT, Toulouse, France

[4]Department of Statistics, Korea University, Seoul, Republic of Korea

Corresponding Author:
Sukhyun Ryu, MD, PhD
Department of Preventive Medicine, College of Medicine, The Catholic University of Korea, R6117, Omibus Park, 222 Banpo-daero, Seoul, Republic of Korea

Abstract

Abstract: The total fertility rate (TFR) is a biased estimate of the population reproductive potential when there is a sex-ratio imbalance at birth, and it does not account for the mortality of women of childbearing age. This study aimed to estimate the reproduction rate (Rt), which adjusts for the sex-ratio imbalance and the mortality of women of childbearing age, and to assess the differences in the timing of when the population reached the replacement level of the TFR and Rt. We first estimated the Rt using the probability of survival in women and the number of female births. Then, using a time-series analysis, we compared the time series of the TFR and Rt in the Korean population between 1975 and 2022. We found the Rt showed a below replacement level of the population a year earlier than the TFR. However, the estimate of the time-series analysis of Rt was not significantly different from the estimates of the TFR. Our finding suggests that the Rt can provide timely information on the adjusted population reproductive potential and is easier for the public to interpret compared to TFR.

(JMIR Public Health Surveill 2025;11:e63603) doi:10.2196/63603

KEYWORDS

fertility rate; reproducibility; reproduction rate; population control; Korea; sex ratio; imbalance; mortality; woman; female; childbearing age; reproductive age; giving birth; assessment; time series; Korean; impact analysis; birth control; reproduction

Introduction

Between 1962 and 1993, Korea implemented a successful family planning policy. In 1993, this policy was discontinued; in 2004, childbirth-promoting policies were implemented (Multimedia Appendix 1). In 2023, Korea had the lowest fertility rate (0.8) worldwide [1]. The total fertility rate (TFR), the average offspring number per childbearing-age females (15 - 49 years), is a common metric to assess population change potential. However, this cohort-based measure is biased when there is male-to-female sex-ratio imbalance at birth [2,3]. The TFR does not account for mortality rates among childbearing-age women, possibly affecting population reproducibility [2,4]. These limitations reduce TFR's ability to accurately reflect a country's population replacement dynamics. Therefore, the net reproduction rate (Rt), the number of daughters a woman of childbearing age would produce under prevailing fertility and mortality conditions, is better. Like other real-time epidemiological metrics (eg, the effective reproduction number in infectious disease modeling) [5], the Rt can be calculated and updated regularly with new population data; it can provide timely insights into population sustainability. The Rt is easier for public understanding, as a population is below the replacement level when the Rt < 1 [6], in contrast to the TFR, with a threshold of 2.1. Despite this, no previous studies have evaluated the population reproducibility using the Rt in Korea.

This study assessed the differences in the timing of reaching population replacement level of the TFR and Rt and the estimated difference of the time series of the TFR and Rt by two major population control policies.

Methods

We collected the annual number of live births, number of women, mortality rate of women, and male-to-female ratio of women of childbearing age between 1975 and 2022 through the Korean National Statistic Agency [7] to calculate the TFR and Rt (Multimedia Appendix 2). To identify the different estimates of policy impact (1975-1993: family planning policy; 1993-2004: childbirth encouragement policy), we conducted

XSL•FO
RenderX

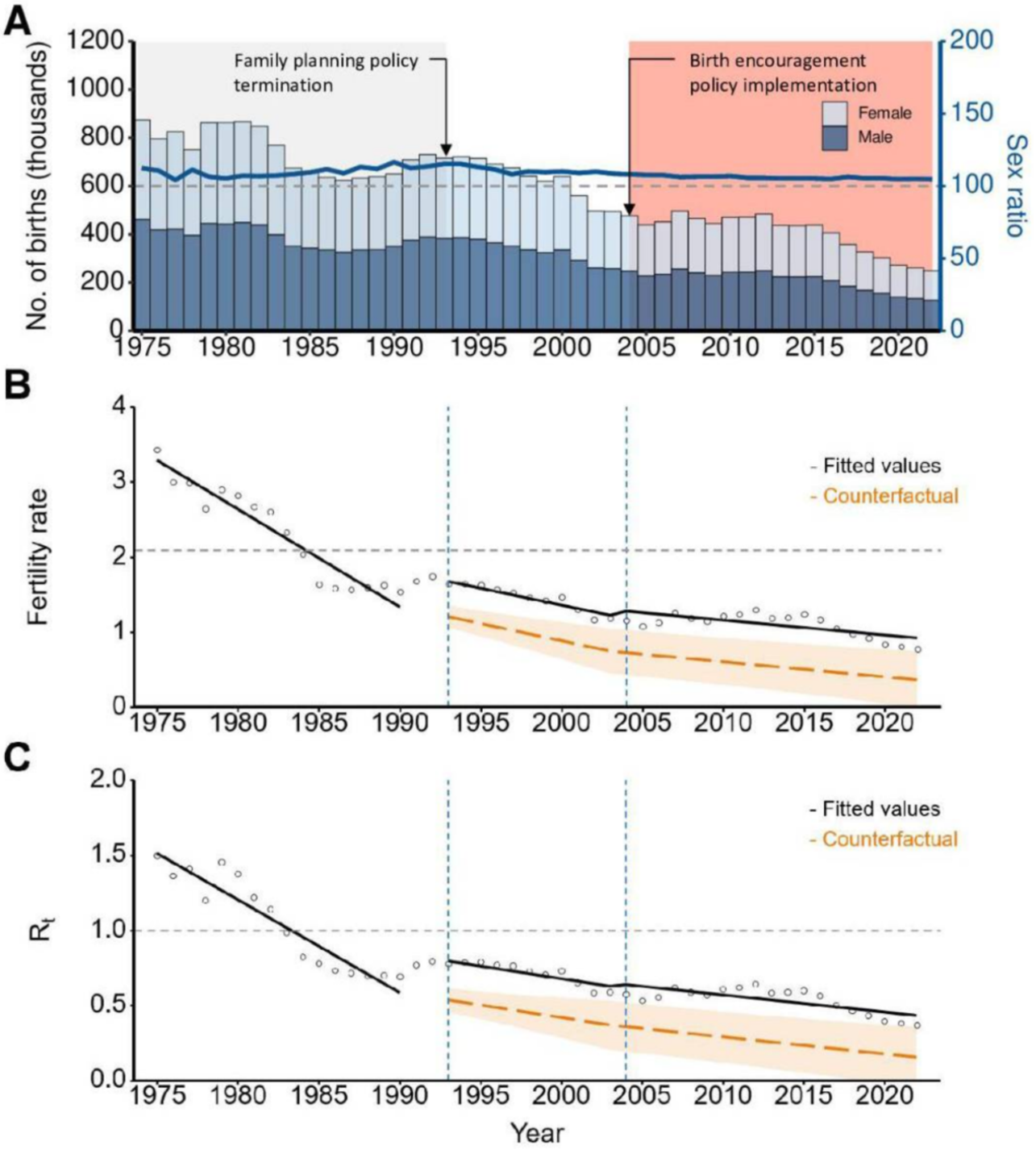
an interrupted time series (ITS) with segmented regression to examine the time trend and its level change in the TFR and Rt. We also conducted a cross-correlation analysis to evaluate the temporal relationship between the TFR and Rt. Then, we compared the estimates of Rt multiplied by 2.1 (TFR threshold level) with the TFR estimates along with 95% CIs. All analyses were conducted using R software (version 4.4.0; R Foundation for Statistical Computing).

Results

The number of live birth number decreased from 874,030 in 1975 to 249,186 in 2022. Similarly, the male-to-female sex ratio decreased from 112 in 1975 to 105 in 2022 (Figure 1A). The TFR remained below 2.1 since 1984 (TFR=2.04) and decreased further to 0.78 in 2022 (Figure 1B). The Rt remained below 1 since 1983 (Rt=0.98) and decreased to 0.4 in 2022 (Figure 1C).

XSL•FO
RenderX

Figure 1. The annual number of live births, total fertility rate (TFR), and reproduction rate (R_t) in South Korea, 1975-2022. (**A**) The bar-colored blue and sky-blue indicate the annual number of male and female births, respectively. The blue solid line indicates the yearly sex ratio of male to female births. (**B**) The interrupted time series model with the TFR. The interruption (dashed vertical line) was set to 1993 (when the family planning policy was discontinued) taking into account for the transition period of the policy and 2004 (when the birth encouragement policy was implemented) to identify the changes in the TFR level and slope. The horizontal dashed line indicates a critical threshold of the TFR at 2.1. The dashed orange line indicates the annual TFR based on a counterfactual scenario without changing the birth control policy, and the orange shaded area indicates 95% confidence intervaIs (CIs) of the TFR. (**C**) The interrupted time series model with the estimated R_t; the critical threshold of R_t=1. An $R_t < 1$ indicates that the population's reproductive performance falls below the replacement level. The dashed orange line indicates the annual R_t based on a counterfactual scenario without changing the birth control policy, and the orange shaded area indicates 95% CIs of R_t.



The ITS showed an immediate increase in the mean TFR (55%) and Rt (26%) and an increased slope change of 9% in the TFR and 4% in Rt following the family planning policy discontinuation (Figure 1B and 1C and Table 1). After the birth encouragement policy introduction, the slope of the TFR (3%) and Rt (1%) increased, with no significant level change. When the Rt was multiplied by 2.1, the estimates were within the 95% CI of the TFR estimate (Table 1). A high correlation between the TFR and Rt at lag 0 indicated no temporal differences (Multimedia Appendix 3).

XSL•FO
RenderX

Table . Estimates from the interrupted time-series analysis using the total fertility rate and reproduction rate in South Korea, 1975-2022.

	Mean total fertility rate, % (95% CIs)[a]	Mean reproduction rate, % (95% CIs)[a]	Mean reproduction rate, (95% CIs) multiplied by 2.1[b]
Immediate level change following the discontinuation of family planning	54.9 (33.1 to 67.2)	25.9 (18.3 to 33.8)	54.4 (38.4 to 71.0)
Post-intervention slope change following the discontinuation of family planning	8.9 (7.6 to 10.4)	4.2 (3.3 to 5.5)	8.8 (6.9 to 11.6)
Immediate level change following the birth-encouragement policy implementation	3.2 (1.3 to 4.8)	1.2 (0.1 to 2.0)	2.5 (0.2 to 4.2)
Post-intervention slope change following the birth-encouragement policy implementation	5.8 (-5.5 to 16.4)	3.3 (-4.6 to 11.2)	6.9 (-9.7 to 23.5)

[a]Estimates of the mean and 95% confidence intervals (CIs) from the interrupted time series with a segmented regression model to examine the time trend and its level change.

[b]Estimates of the reproduction rate were multiplied by 2.1 (threshold level of total fertility rate) along with 95% CIs.

Discussion

The threshold level of the population replacement was captured a year earlier through the Rt compared to the TFR. This is likely due to sex-ratio imbalances in Korea.

The trend levels and slope changes of the TFR and Rt increased following the birth control policy discontinuation [8]. These significant level changes were likely affected by previous birth control policies [9]. However, after the child encouragement policy implementation in 2004, the TFR and Rt were far below the population replacement threshold, consistent with a previous study that reported no positive effect of child encouragement policies on the fertility rate [10], likely due to sociocultural factors influencing fertility behavior (eg, changing gender roles and economic pressures) [10]. Our study could be applied to other countries experiencing similar socioeconomic and cultural dynamics, particularly those with comparable fertility patterns and sex-ratio imbalances [4].

This study had limitations. Sensitivity analyses were not included in the parameter estimation models. The ITS models were interrupted in 1993 to reflect the discontinuation of the family planning policy, accounting for the policy transition period. The ITS may not fully capture the nonlinear trends after 2015. We did not consider the qualitative characteristics of each policy.

The Rt can be used as a useful and timely metric of population reproducibility, particularly in the presence of sex-ratio imbalance at birth. Furthermore, the Rt threshold of 1 may be easier for public interpretation compared to the TFR, as the public became familiar with the Rt parameter during the COVID-19 pandemic.

Acknowledgments

Ethics approval was waived by the institutional review board at the Catholic University of Korea (2024-1205-0001). The authors attest that there was no use of generative artificial intelligence technology in the generation of text, figures, or other informational content in this manuscript. This work was supported by the Gorvernment-wide Research and Development to Advance Infectious Disease Prevention and Control, Republic of Korea (grant number RS-2023-KH140322), and the Basic Science Research Program through the National Research Foundation of Korea by the Ministry of Education (grant numbers NRF-2020R1I1A3066471 and 2022R1A2C1008514).

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Authors' Contributions

SR conceived the study and designed the statistical methods. CA, CH, LSJ, and SC collected and assimilated the data. CA, CH, LSJ, and SC performed the data analysis. CA and SR wrote the first manuscript draft. CA, SC, and SR critically reviewed and edited the manuscript. All authors helped to interpret the results, critically revised the manuscript, and provided final approval of the version for publication.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Key population policies in South Korea between 1970 and 2022.
[DOCX File, 20 KB - publichealth_v11i1e63603_app1.docx]

Multimedia Appendix 2
Description of total fertility rate, reproduction rate, and time-series analysis.
[DOCX File, 22 KB - publichealth_v11i1e63603_app2.docx]

Multimedia Appendix 3
Yearly lagged cross-correlation coefficients between total fertility rate and net reproduction number in South Korea (1975–2022).
[DOCX File, 108 KB - publichealth_v11i1e63603_app3.docx]

References

1. World population dashboard. United Nations Population Fund.: New York, United States: UNFPA; 2023. URL: https://www.unfpa.org/data/world-population-dashboard
2. Gietel-Basten S, Scherbov S. Is half the world's population really below "replacement-rate"? PLoS ONE 2019;14(12):e0224985. [doi: 10.1371/journal.pone.0224985] [Medline: 31790416]
3. Chao F, Gerland P, Cook AR, Guilmoto CZ, Alkema L. Projecting sex imbalances at birth at global, regional and national levels from 2021 to 2100: scenario-based Bayesian probabilistic projections of the sex ratio at birth and missing female births based on 3.26 billion birth records. BMJ Glob Health 2021 Aug;6(8):e005516. [doi: 10.1136/bmjgh-2021-005516] [Medline: 34341019]
4. Aitken RJ. The changing tide of human fertility. Hum Reprod 2022 Apr 1;37(4):629-638. [doi: 10.1093/humrep/deac011] [Medline: 35079808]
5. Han C, Seo H, Cho S, Chiara A, Ryu S. Impact of travel restrictions for travellers from China on the internal spread of SARS-CoV-2 in South Korea. J Travel Med 2023 Sep 5;30(5):taad047. [doi: 10.1093/jtm/taad047] [Medline: 37040325]
6. Lotka AJ. Analytical Theory of Biological Populations 1998.
7. Korean Statistical Information Service Monthly, quarterly, and annual population trends.: Daejeon, South Korea: Statistics Korea; 2023. URL: https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1B8000G&conn_path=I2
8. Oh YR. How does Korea's family planning policy promote familism? Korea Social Policy Review 2020;27(1):213-248. [doi: 10.17000/kspr.27.1.202003.213]
9. Jang Y, Kim N, Lee S, Jin D. Korea's Population Policy: History and Future 2010.
10. Jeong K, Yoon J, Cho HJ, Kim S, Jang J. The relationship between changes in the korean fertility rate and policies to encourage fertility. BMC Public Health 2022 Dec 8;22(1):2298. [doi: 10.1186/s12889-022-14722-4] [Medline: 36482452]

Abbreviations

ITS: interrupted time series
TFR: total fertility rate

Edited by A Mavragani; submitted 24.06.24; peer-reviewed by A Aktuna, Y Ko; revised version received 22.11.24; accepted 04.12.24; published 12.02.25.

Please cite as:
Achangwa C, Han C, Lim JS, Cho S, Choi S, Ryu S
Net Reproduction Number as a Real-Time Metric of Population Reproducibility
JMIR Public Health Surveill 2025;11:e63603
URL: https://publichealth.jmir.org/2025/1/e63603
doi:10.2196/63603

© Chiara Achangwa, Changhee Han, Jun-Sik Lim, Seonghui Cho, Sangbum Choi, Sukhyun Ryu. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.2.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Research Letter

Racial/Ethnic Inequities in Polysubstance Use Among Online Help–Seeking Sexual and Gender Minoritized People in San Francisco From 2022 to 2025: Cross-Sectional Study

Jarett Maycott[1], MPH; Sean Arayasirikul[2], PhD

[1]Center for Public Health Research, San Francisco Department of Public Health, San Francisco, CA, United States

[2]The Legacy Center, Joe C. Wen School of Population & Public Health, University of California, Irvine, Irvine, CA, United States

Corresponding Author:
Sean Arayasirikul, PhD
The Legacy Center
Joe C. Wen School of Population & Public Health
University of California, Irvine
856 Health Sciences Quad
Irvine, CA, 92697-3957
United States
Phone: 1 9498242680
Email: sean.arayasirikul@uci.edu

Abstract

This research letter characterizes racial/ethnic disparities in polysubstance use among 409 online help–seeking sexual and gender minoritized people in San Francisco. Findings demonstrate the central role of tobacco as a co-occurring substance for participants who are Black, Indigenous, and people of color compared to their White counterparts.

(JMIR Public Health Surveill 2025;11:e82313) doi:10.2196/82313

KEYWORDS

online samples; digital health; polysubstance use; sexual and gender minoritized community health; social media

Introduction

Polysubstance use, the concurrent or sequential use of more than 1 substance within a defined interval, is associated with multiple elevated health risks (eg, overdose, poor mental health, HIV) and disproportionately impacts sexual and gender minoritized (SGM) subgroups, such as transgender people and cisgender lesbian women [1-3]. Few studies have documented racial/ethnic disparities in polysubstance use among SGM populations [4]. Analyses of the 2021 National Survey on Drug Use and Health identified greater risk of polysubstance use among nonheterosexual respondents, illustrating SGM health disparities in substance use [5], yet little is known about how these disparities emerge in the critical moments before seeking help, particularly among SGM individuals using online platforms, where opportunities for digital heath intervention may be most immediate and impactful to populations least likely to have access to supportive public health services [5].

Methods

Study Design and Participants

We recruited 409 participants between 2022 and 2025 through Facebook, Instagram, and Grindr advertisements for individuals seeking help for substance use or treatment services and related topics (eg, HIV, mental health). Participants were eligible if they were men who have sex with men or transgender and gender expansive, were aged ≥18 years, were San Francisco residents, and had a mobile phone.

Measures

We measured the following sociodemographic characteristics: age, race/ethnicity, gender identity, sexual orientation, HIV serostatus, housing, and income. Race/ethnicity was recoded into a dichotomous variable (1 for Black, Indigenous, and people of color [BIPOC] and 0 for White). We measured substance use by asking participants how many days in the last 30 they used tobacco, vape devices, and alcohol; engaged in hazardous drinking; and used marijuana, prescription opioids (eg, pain relievers), nonprescription opioids (eg, heroin, fentanyl), other

XSL•FO
RenderX

prescription drugs (used nonmedically), other illicit substances (eg, crack, cocaine, amphetamines, methamphetamine, hallucinogens, or inhalants such as poppers), and any injected drugs. For each substance, we created a dichotomous variable for any reported use at least once in the prior 30 days (yes/no). We measured past-30-day polysubstance use by assessing the prevalent combinations of pairs of substances used in the last 30 days.

Data Analysis

A correlation matrix heat map was used to illustrate the proportion of participants reporting each polysubstance use combination. We then constructed logistic regression models to examine associations between BIPOC identity and substance use pairs in the top quantile of prevalence (above 10% of the sample), adjusting for age, gender, sexual orientation, HIV status and housing stability, and socioeconomic status. Analyses were performed using R (version 2024.12.0). No missing data were recorded.

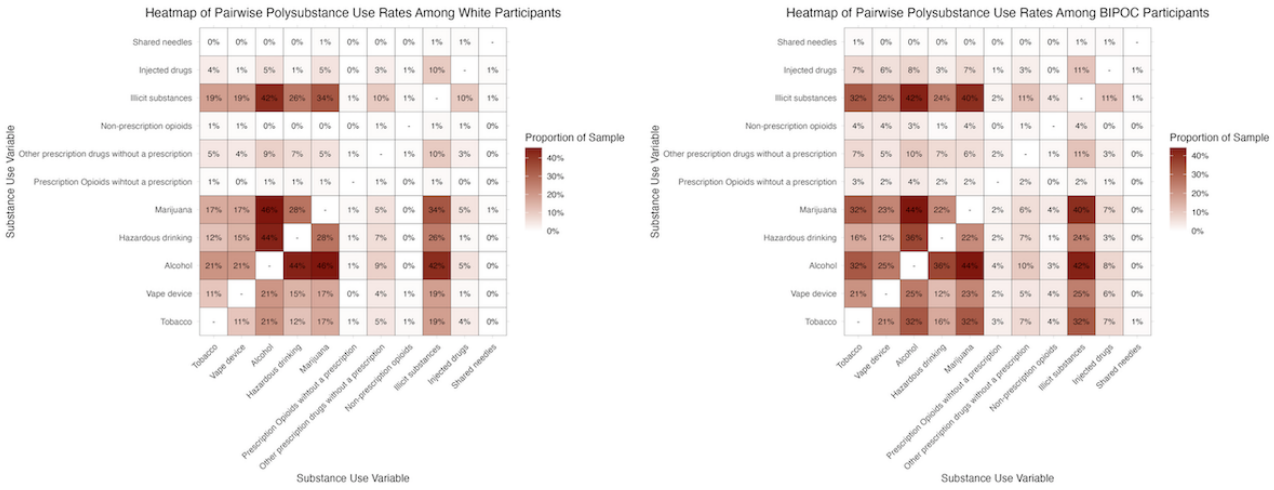
Ethical Considerations

The study protocol was approved by the University of California, San Francisco institutional review board (20-33169). Participants provided signed informed consent and were given a US$30 gift card for completing the assessment. Data were deidentified to protect participants' privacy and confidentiality.

Results

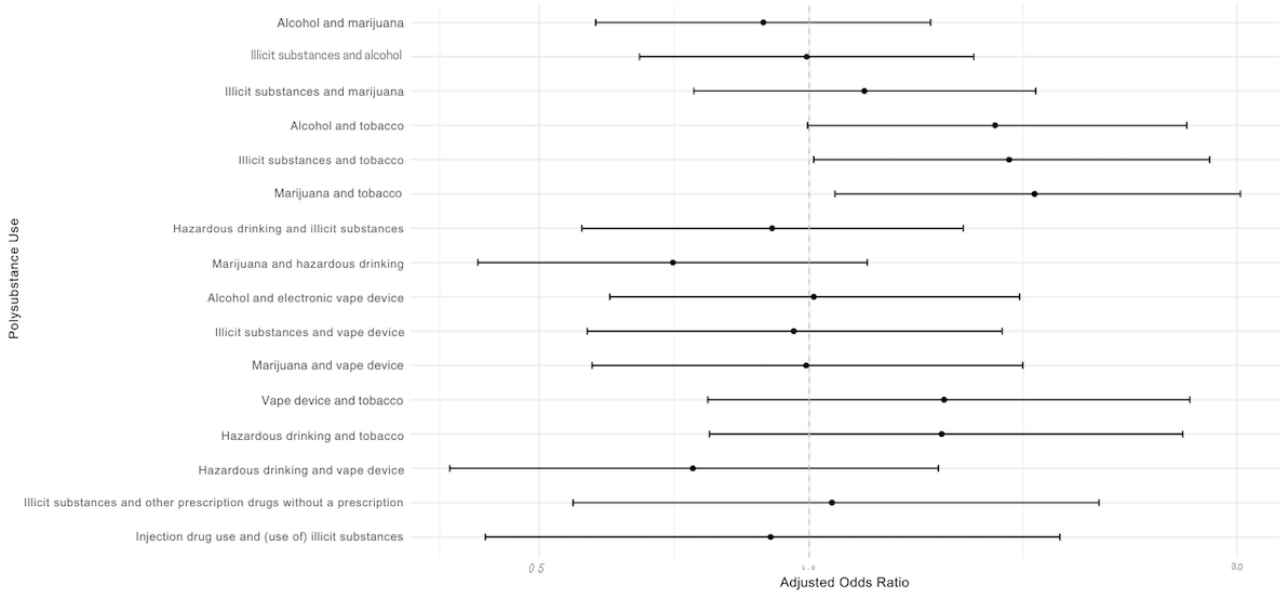
BIPOC participants comprised 58.7% (240/409) of the sample, with the largest group being Latine (113/409, 27.6%), followed by Asian and Pacific Islander and Native Hawaiian (50/409, 12.2%), multiracial/other (39/409, 9.5%), and Black/African American (38/409, 9.3%). Figure 1 presents 2 matrices, stratified by BIPOC identity, for White participants and BIPOC participants. Tobacco appears more frequently as a co-occurring substance across all substance use behaviors among BIPOC participants compared to White participants. Use of a vape product was also notably higher as a co-occurring substance in all categories for BIPOC participants. In contrast, hazardous drinking was less prevalent among BIPOC participants (86/240, 35.8%) than among White participants (74/169, 43.8%). Regression models (Figure 2) identified 3 substance use combinations with higher odds of use among BIPOC participants compared to White participants: alcohol and tobacco (adjusted odds ratio [aOR] 1.61, 95% CI 1.00-2.64), illicit substances and tobacco (aOR 1.67, 95% CI 1.07-2.80), and marijuana and tobacco (aOR 1.78, 95% CI 1.07-3.03).

Figure 1. Heat map comparison of pairwise polysubstance use rates among White participants and Black, Indigenous, and people of color (BIPOC) participants.



XSL•FO
RenderX

Figure 2. Likelihood of polysubstance use among sexual and gender minoritized participants who were Black, Indigenous, or people of color compared to sexual and gender minoritized participants who were White (in San Francisco from 2022 to 2025; N=409).



Discussion

We found that 3 combinations of polysubstance use involving tobacco—alcohol and tobacco, illicit substances and tobacco, and marijuana and tobacco—were more common among BIPOC participants than White participants. This pattern mirrors other research documenting higher tobacco use prevalence in BIPOC communities despite trending decreases in use nationwide [6].

This study has limited generalizability due to its cross-sectional study design, self-report bias, and its focus on online help–seeking individuals in San Francisco. Despite this, our results highlight a persistent gap in understanding how race intersects with tobacco use and its role in polysubstance use, particularly within minoritized populations actively seeking support online. Future research examining ecological dimensions of polysubstance use would further our understanding of the social context of polysubstance use.

Acknowledgments

The authors would like to thank all participants in the study. This work was funded by the Substance Abuse and Mental Health Service Administration (SAMHSA; award H79SP082077).

Data Availability

The datasets generated or analyzed during this study are not publicly available as they contain details that could be used to identify participants but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

References

1. Flentje A, Sunder G, Ceja A, Lisha NE, Neilands TB, Aouizerat BE, et al. Substance use over time among sexual and gender minority people: differences at the intersection of sex and gender. LGBT Health 2024;11(4):269-281 [FREE Full text] [doi: 10.1089/lgbt.2023.0055] [Medline: 38206680]
2. Cotaina M, Peraire M, Boscá M, Echeverria I, Benito A, Haro G. Substance use in the transgender population: a meta-analysis. Brain Sci 2022 Mar 10;12(3):366 [FREE Full text] [doi: 10.3390/brainsci12030366] [Medline: 35326322]
3. Mestre LM, White MA, Levy BR, Bold KW. Higher prevalence of polysubstance use among older lesbian, and gay US adults. Drug Alcohol Depend Rep 2024 Sep;12:100281 [FREE Full text] [doi: 10.1016/j.dadr.2024.100281] [Medline: 39314854]
4. Coulter RWS, Ware D, Fish JN, Plankey MW. Latent classes of polysubstance use among adolescents in the United States: intersections of sexual identity with sex, age, and race/ethnicity. LGBT Health 2019 Apr;6(3):116-125 [FREE Full text] [doi: 10.1089/lgbt.2018.0149] [Medline: 30822259]
5. Lee J, Evans-Polce RJ, Ahlquist J, Parker MA. Polysubstance use by sexual identity among US adults, 2021. Prev Med 2023 Dec;177:107729. [doi: 10.1016/j.ypmed.2023.107729] [Medline: 37852580]

XSL•FO
RenderX

6. Nguyen-Grozavu FT, Pierce JP, Sakuma KK, Leas EC, McMenamin SB, Kealey S, et al. Widening disparities in cigarette smoking by race/ethnicity across education level in the United States. Prev Med 2020 Oct;139:106220 [FREE Full text] [doi: 10.1016/j.ypmed.2020.106220] [Medline: 32693179]

Abbreviations

aOR: adjusted odds ratio
BIPOC: Black, Indigenous, people of color
SGM: sexual and gender minoritized

Edited by A Mavragani, T Sanchez; submitted 12.08.25; peer-reviewed by J Lee, D Smith; comments to author 03.09.25; revised version received 03.09.25; accepted 06.09.25; published 12.09.25.

Please cite as:
Maycott J, Arayasirikul S
Racial/Ethnic Inequities in Polysubstance Use Among Online Help–Seeking Sexual and Gender Minoritized People in San Francisco From 2022 to 2025: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e82313
URL: https://publichealth.jmir.org/2025/1/e82313
doi:10.2196/82313
PMID:

©Jarett Maycott, Sean Arayasirikul. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 12.09.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multilevel Diabetes Prevention Interventions to Address Population Inequities in Diabetes Risk: Scoping Review

Reshma Patel[1], MPH; Kathy Kornas[1], MSc; David Gerstle[2], PhD; Lori M Diemert[1], MSc; Laura C Rosella[1,3,4], PhD

[1]Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, ON, Canada

[2]University of Toronto Mississauga Library, University of Toronto, Mississauga, ON, Canada

[3]Institute for Better Health, Trillium Health Partners, Mississauga, ON, Canada

[4]Laboratory Medicine and Pathobiology, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Laura C Rosella, PhD
Dalla Lana School of Public Health, University of Toronto, 155 College Street, Toronto, ON, Canada

Abstract

Background: Type 2 diabetes risk is disproportionately higher among structurally marginalized communities, partly due to influences from social, economic, and environmental determinants of health. Individual-level diabetes prevention strategies address proximal factors, such as modifiable behaviors, often overlooking the role of multilevel socioecological factors that contribute to diabetes risk and inequities. Multilevel diabetes prevention interventions involve actions that address multiple health determinants across the individual, community, and systemic levels of influence, offering a promising approach to reducing inequities in diabetes risk.

Objective: This scoping review aimed to systematically map the types of health determinants addressed in multilevel diabetes prevention interventions that have been implemented for addressing population inequities in diabetes risk and to describe what evidence exists regarding their effectiveness.

Methods: A comprehensive literature search was conducted in PubMed, CINAHL, MEDLINE, Embase, Web of Science, and gray literature sources (websites of government agencies and local/international nongovernmental health organizations) for studies published from the year 2000 to 2024. The research team developed a conceptual framework to guide the scoping review and define multilevel interventions for eligibility. Eligibility criteria included studies focusing on multilevel diabetes prevention interventions targeting diabetes relevant risk factors at more than one level of influence (micro, meso, and macro) and where intervention outcomes were reported. Data extraction included study characteristics, intervention target populations and coverage, targeted health determinants, and intervention outcomes and was completed by 2 independent reviewers. Data synthesis involved mapping health determinants addressed by each multilevel intervention according to our conceptual framework and a narrative synthesis of findings on themes corresponding to intervention types and reported outcomes.

Results: Of 7813 articles retrieved, a total of 25 studies met the inclusion criteria. Interventions consisted of targeted interventions for high-risk populations (n=7), environmental-based interventions (n=7), and community-based interventions (n=11). Most interventions addressed health determinants at 2 levels (micro and macro) (14/25, 56%) or 3 levels (micro, meso, and macro) (11/25, 44%). All studies reported on proximal outcomes, most frequently on weight, physical activity, and dietary behaviors. One-third (8/25, 32%) of studies reported outcomes on changes in metabolic risk. None of the studies reported on equity outcomes related to changes in population inequities in diabetes incidence. Only 8% (n=2) of studies reported an equity outcome that captures disparities in a diabetes risk factor level between disadvantaged and advantaged population groups.

Conclusions: Our review identified a research gap in that outcomes on population inequities in diabetes risk have not been consistently measured in multilevel diabetes prevention interventions, and the impact of these interventions on reducing population inequities in diabetes incidence is not consistently examined or reported. Future research should prioritize equity outcomes in evaluations of multilevel diabetes prevention interventions and emphasize impacts on disadvantaged populations and population inequities.

(JMIR Public Health Surveill 2025;11:e70267) doi:10.2196/70267

KEYWORDS

type 2 diabetes; multilevel intervention; prevention; health equity; review

XSL•FO
RenderX

Introduction

Preventing type 2 diabetes is a priority for health systems worldwide, as reflected by the United Nations Global Sustainable Development Goal, which emphasizes good health and well-being and underscores the importance of preventing chronic conditions [1]. Globally, there is a high burden of type 2 diabetes [2], in addition to persistent inequities in diabetes risk among population groups, disproportionately impacting racialized populations, low-income, and marginalized communities [3-5]. These inequities are more pronounced at younger ages; for example, South Asian and Black ethnic groups have been found to have higher type 2 diabetes prevalence among younger age groups (19 - 39 years), compared with White European populations [6].

Multiple factors at the individual, health system, neighborhood, and socioeconomic levels (eg, obesity, unhealthy diets, low physical activity, health care access, income, neighborhood, and food environments) contribute to the inequities in diabetes risk [7]. For example, the association between obesity and diabetes varies in different ethnic groups. South Asians, for instance, have a higher burden of diabetes at lower BMI levels compared with White populations [3,8]. Built environments have been shown to differentially impact diabetes risk in different population groups. For instance, in Ontario (Canada), prediabetes incidence was found to be lower in highly walkable neighborhoods; however, the association varied across ethnic groups, with South Asians, South East Asians, Sub-Saharan Africans, and Caribbeans benefiting less from walkable areas than Western Europeans [9].

Historically, diabetes prevention efforts have concentrated on individual-level interventions that aim to address modifiable lifestyle and clinical risk factors [10]. These strategies—such as lifestyle modifications, pharmacotherapy, and metabolic surgery for those at high risk—have demonstrated effectiveness in reducing diabetes risk and play an important role in diabetes prevention [11]. However, interventions focused on proximal factors, such as behavior change, can widen inequities because they do not consider social and environmental contexts that support or impede lifestyle changes [12]. Furthermore, structural and social determinants that contribute to inequities in diabetes risk cannot be addressed at the individual level alone [13]. It is increasingly being recognized that there is no one-size-fits-all solution for diabetes prevention that is effective for all populations and achieves health equity [14].

Multilevel interventions that address social determinants and individual-level risk factors across multiple levels of influence (micro, meso, and macro) are increasingly recognized as one solution to reduce inequities in diabetes risk and improve population health [14-16]. Multilevel interventions are strategies that aim to improve health outcomes by targeting at least 2 levels of influence, from the individual, community, and structural/environmental/policy context [15]. The individual level refers to actions directed at individuals, the community level refers to actions that target institutions and the community context, and the structural/policy level refers to actions directed at the social, policy, and environmental context [17].

Previous reviews of multilevel diabetes prevention interventions that have been implemented in different community and social settings have focused on workplace wellness programs [18], faith-based interventions to address obesity [19], population-wide programs that target the built and food environments [20], and multilevel interventions targeted for Native peoples [21]. This scoping review was conducted to systematically map the range of multilevel diabetes prevention interventions on addressing inequities in diabetes risk. Our review is unique because of its broader population-based focus and objective of gaining insight into what health determinants and equity factors across the micro, meso, and macro levels are addressed in multilevel diabetes prevention interventions. In addition, we focused on describing evidence on the impact of these interventions on reducing inequities in diabetes risk.

Specifically, we aimed to answer the research question: What multilevel diabetes prevention interventions have been implemented to address population inequities in diabetes risk, and what evidence exists regarding their effectiveness? Our review pursued three objectives: (1) to map the types of health determinants and equity factors addressed across the levels of influence (micro, meso, and macro) in multilevel interventions for diabetes prevention; (2) to identify the target population groups in multilevel interventions for diabetes prevention in the identified studies, including intervention coverage; and (3) to describe the types of proximal and distal outcomes reported in the identified studies for the changes in risk factors for diabetes, population risk distribution (changes in diabetes risk), and changes in population inequities in diabetes risk.

Methods

Scoping Review and Theoretical Framework

The review methodology followed the framework of Arksey and O'Malley [22] and the Joanna Briggs Institute [23], and adhered to the PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) guidelines (Checklist 1). Scoping reviews are specifically useful for reviewing health research evidence and identifying research gaps [24] and therefore align with the goals of our review. The review protocol was registered on Open Science Framework [25].

We developed a conceptual model based on 2 theoretical frameworks to guide the scoping review. The social ecological model [26] is a commonly used framework for describing factors of influence at the micro, meso, and macro levels and has been previously applied to understand the multiple levels of influence for diabetes prevention [27], demonstrating its applicability to multilevel diabetes prevention interventions. The micro level consists of intrapersonal (individual-level) factors (eg, health behaviors) and interpersonal factors (eg, small group interactions, social support). The meso level consists of organizational structures and processes (eg, health care institutions, workplace health programs, and activities in the school setting), and community factors (eg, characteristics and resources in communities). The macro level represents socioeconomic conditions, environmental factors, and policy influences (eg, built environment, housing, and food supply).

XSL•FO
RenderX

The wider determinants of health model [28] was chosen as it is a widely used social determinants framework to understand the elements essential to achieve health equity and identify specific factors that influence health across different levels, including socioeconomic and structural factors, community and social factors, and intrapersonal lifestyle factors. A previous review identified social determinants of health associated with diabetes risk [29], which we integrated into our conceptual model. A comprehensive table outlining the conceptual model is available in Multimedia Appendix 1.

Search Strategy

We developed the search strategy with a research librarian (DG) from the University of Toronto. We also consulted with an additional librarian through the Peer Review for Electronic Search Strategies (PRESS). PRESS evaluation is widely advocated by librarians, who contribute to knowledge syntheses to draw on additional expertise in developing keywords, identifying subject headings (eg, medical subject heading terms), and refining lines of the strategy [30]. Our PRESS submission was facilitated by the research librarian through the University of Toronto Libraries' Knowledge Syntheses Community of Practice.

We searched indexed databases of peer-reviewed literature, gray literature indexed in web sources, and manual searches of reference lists from related published reviews. To concentrate on contemporary research on multilevel interventions for diabetes prevention, we restricted our search to publications from the year 2000 onward. We considered this an appropriate cut-off given that key frameworks and influential guidance for developing and evaluating complex interventions have primarily been published after the year 2000 [31,32]. We limited our search to high-income economies, using filters based on the World Bank Country and Lending Groups [33]. We excluded publications in non-English languages without available English translations, as well as letters, editorials, opinion pieces, book chapters, and conference abstracts.

The peer-reviewed literature search was conducted in May 2024 across 5 electronic databases: MEDLINE (OVID platform), Embase (OVID platform), CINAHL, Scopus, and Web of Science. Gray literature web sources from government agencies and local/international nongovernmental health organizations were searched in July 2024 by applying search strings developed by the research librarian. If search strings were not supported, then we manually reviewed the first 10 pages of relevant websites. Specifically, we searched the websites of the World Health Organization, the Centers for Disease Control and Prevention, the Organization for Cooperation and Development, the Public Health Agency of Canada, Health Canada, Public Health Ontario, and Peel Public Health. Search terms corresponded to the concepts of type 2 diabetes, multilevel, and prevention. The finalized search strategy implemented in MEDLINE and the search strings used on websites are provided in Multimedia Appendix 2. Identical searches were customized for the remaining peer-reviewed databases. Finally, the reference list of related reviews, identified from the search of peer-reviewed databases, was manually searched for additional articles for inclusion [34-38].

Eligibility Criteria

In addition to the criteria described above for timeframe, geography, language, and publication type, the following population, intervention, and outcome eligibility criteria were applied in screening. Studies were included where the study population/target group for the intervention included individuals without diabetes (youth and adults). Studies were excluded where the study population/target group for the intervention included individuals with diabetes (ie, type 1, type 2, and gestational). Studies were included if intervention outcomes were reported. Studies that focused on describing the development or implementation of a multilevel intervention were excluded.

Studies were excluded if the intervention was unrelated to diabetes prevention. For this review, we considered an intervention applicable to diabetes prevention if its actions aimed to reduce diabetes risk or diabetes-relevant risk factors (eg, physical activity, BMI, and diet). Studies were included where the intervention aimed to target diabetes-relevant risk factors at more than one level of influence (ie, micro, meso, and macro). Studies on interventions that targeted one level of influence (eg, an education session in a group setting, enhancing social support [interpersonal] to improve diet [intrapersonal]) were considered one level of influence (the micro level in this example) and excluded from our review. Similarly, interventions were excluded where activities took place at multiple levels, but the target for change took place at only one level. For example, a mass media campaign seeking to promote physical activity would be excluded if the target for change took place only at the micro level.

In the process of screening studies, we identified existing reviews that synthesized findings on specific types of multilevel interventions related to diabetes prevention. Thus, we modified our original protocol by expanding our exclusion criteria for the following types of interventions that were the subject of an existing review: multilevel diabetes prevention interventions for Native and Indigenous Peoples [21]; diabetes prevention interventions implemented at the workplace setting [18,39,40]; built environment interventions related to investment in park/recreational areas, point-of-decision prompts to encourage stair use and active travel interventions [20]; food environment interventions related to economic measures (sugar-sweetened beverage prices, fast food prices, fruit and vegetable prices) and policy measures for food labeling or changing the layout/food provided at grocery stores [20]; interventions for physical activity in after school hours childcare settings (eg, timetabling changes, provision of equipment for active play, and changes in policies at regional or national level) [37]; school-based interventions aimed at reducing sedentary behavior (eg, changes to classroom design, such as sit-to-stand desks, and changes to curriculum) [41]; and church-based interventions targeted for African American and Latino communities [19]. A summary of the findings from each review on excluded multilevel interventions can be found in Multimedia Appendix 3 [18-21,37,39-41]. A summary of the scoping review inclusion and exclusion criteria can be found in Multimedia Appendix 4.

XSL•FO
RenderX

Screening Process

Abstract and full-text screening was done through Covidence (Veritas Health Innovation Ltd), a software for managing and streamlining reviews. First, 2 reviewers (RP and KK) independently screened all identified papers for relevance by title and abstract and categorized them as "include," "exclude," or "inconclusive." Categorization disagreements were resolved through discussion against the eligibility criteria. Next, the full text of all articles marked as "include" and "inconclusive" was independently reviewed by 2 reviewers, and disagreements were resolved through discussion for a decision regarding the final inclusion.

Data Extraction

All data were extracted into a standardized data extraction template created using the extraction feature on Covidence. Specifically, 2 reviewers (RP and KK) acted as primary reviewers to abstract the data for half of the articles and as secondary reviewers to verify the data for the other half of the articles. Disagreements were addressed through discussion between reviewers. Data extraction items were informed by our conceptual framework. From the included studies, we extracted study characteristics (year of publication, publication type, and country of publication), intervention characteristics (setting, description, target population, coverage, levels of influence targeted, diabetes relevant health determinants targeted, and equity factors targeted), and intervention outcomes (changes in diet, physical activity, BMI/weight, blood pressure, metabolic risk, and inequities in diabetes risk). Inequity was defined as a health difference that adversely affects disadvantaged populations on the basis of a higher incidence of disease.

Synthesis

Quantitative data were synthesized descriptively to summarize study characteristics, types of multilevel interventions, target populations, and types of outcomes. We grouped studies based on intervention type. Health determinants targeted by each multilevel intervention were mapped according to our conceptual framework. A narrative synthesis was used to analyze qualitative themes corresponding to intervention types and intervention outcomes.

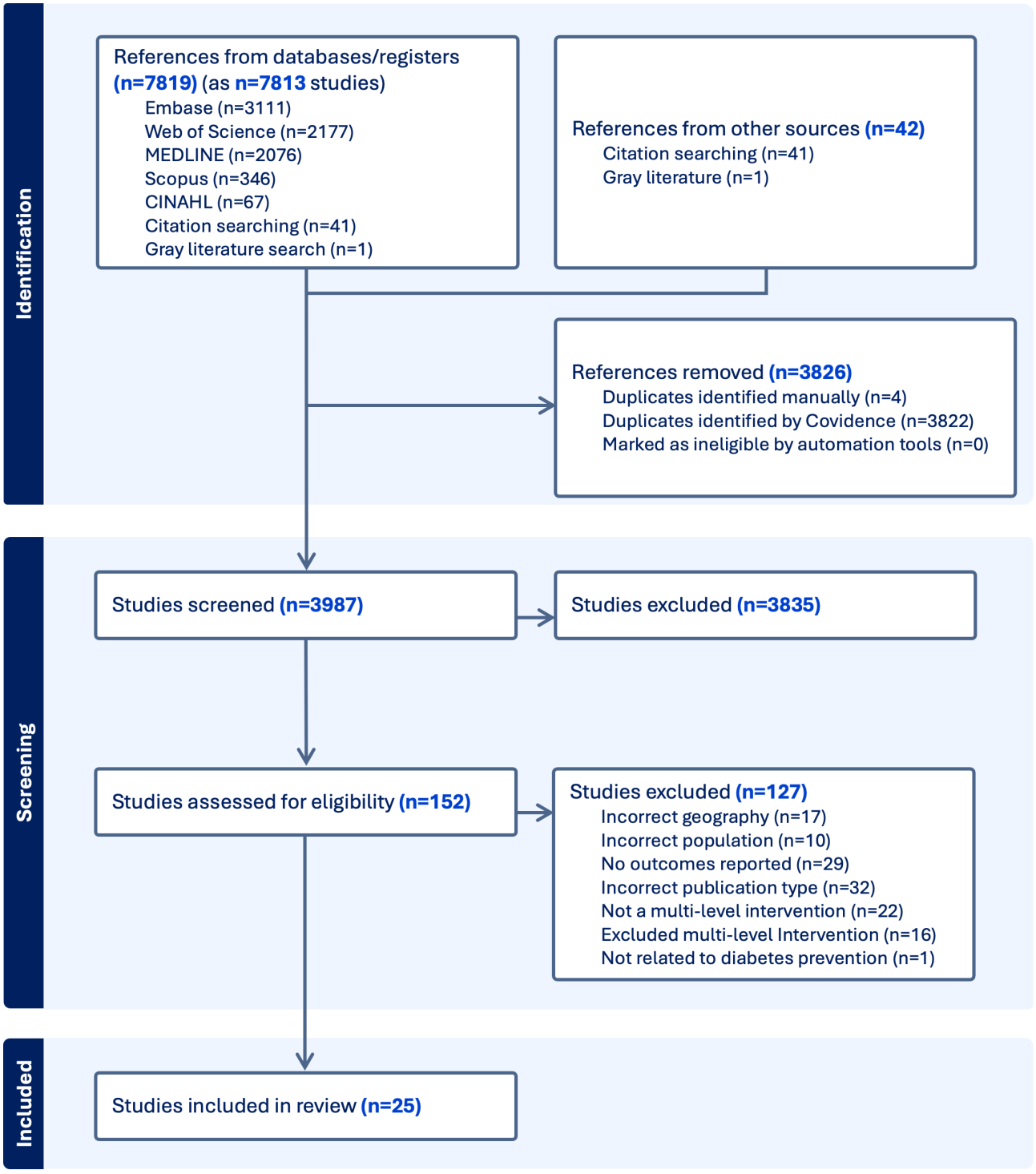
Results

Overview

We initially retrieved 7813 articles for screening, which were imported into Covidence. After removing duplicates, 3987 articles remained for further evaluation. Following the abstract screening, 3835 articles were excluded, and an additional 127 articles were excluded after full-text screening. Our review includes 25 studies related to multilevel interventions for the prevention of type 2 diabetes. The detailed screening process is illustrated in Figure 1.

XSL•FO
RenderX

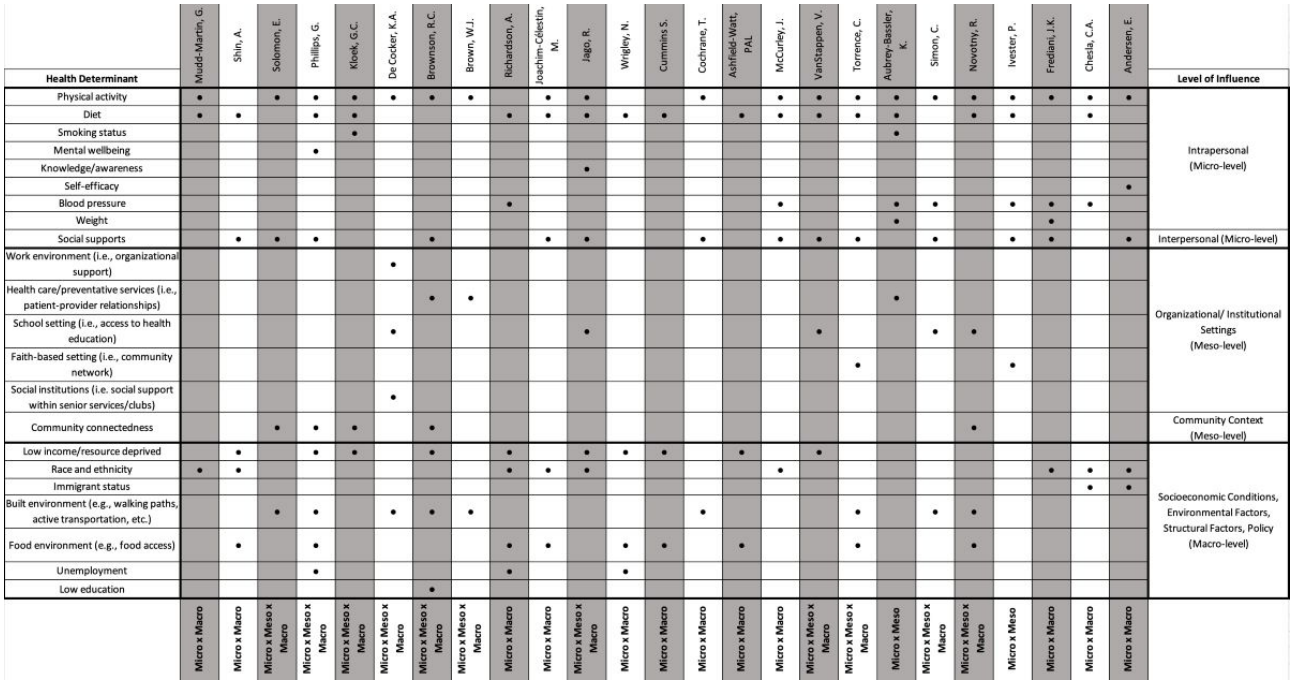
Figure 1. Preferred reporting items for systematic reviews and meta-analyses flowchart of the article screening process.



The included studies described interventions targeting diabetes-free or undiagnosed individuals, spanning 2 or 3 levels of influence: micro, meso, and macro. Given the inclusion criteria and multilevel nature of this review, it should be noted that all interventions addressed at least 2 health determinants across at least 2 levels of influence (micro, meso, and macro). Specifically, most interventions addressed health determinants at the micro and macro levels (12/25, 48%) or all 3 levels (micro, meso, and macro) (11/25, 44%), with the remaining interventions focusing on health determinants at the micro and meso levels (2/25, 8%) (Figure 2).

XSL•FO
RenderX

Figure 2. Health determinants targeted by multilevel interventions for diabetes prevention [42-66].



Characteristics of Included Studies

Most studies were conducted in the United States (12/25, 48%), and the remainder were from Canada (1/25, 4%), the United Kingdom (6/25, 24%), Australia (1/25, 4%), and several European countries including the Netherlands, Belgium, Finland, Greece, Spain, Hungary, France, and Norway (5/25, 20%). The publication dates were evenly distributed, with all studies published between 2003 and 2022. The studies used various methodological designs, such as randomized controlled trials (9/25, 36%), quasi-experimental studies (9/25, 36%), pretest/posttest designs (3/25, 12%), a cohort study (1/25, 4%), an evaluation (1/25, 4%), one-arm pilot clinical trial (1/25, 4%), and a community case study (1/25, 4%). The study characteristics are presented in Table 1.

XSL•FO
RenderX

Table . Study characteristics of diabetes prevention articles included in the scoping review (2000 - 2024) (n=25).

Characteristic	Number of articles, n (%)
Country	
United States	12 (48)
United Kingdom	6 (24)
Australia	1 (4)
Belgium	1 (4)
Canada	1 (4)
France	1 (4)
Netherlands	1 (4)
Norway	1 (4)
Multicountry (within Europe)	1 (4)
Year published	
2003 - 2007	7 (28)
2008 - 2012	5 (20)
2013 - 2017	7 (28)
2018 - 2022	6 (24)
Study design	
RCT[a]	9 (36)
Quasi-experimental	9 (36)
Pretest/posttest	3 (12)
Cohort study	1 (4)
Evaluation	1 (4)
One-arm pilot trial	1 (4)
Community case study	1 (4)

[a]RCT: randomized controlled trial.

Types of Multilevel Interventions for Diabetes Prevention

We grouped the intervention types identified in our review into 3 main categories: targeted interventions for high-risk populations (n=7), environmental-based interventions (n=7), and community-based interventions (n=11). A detailed summary of each intervention is provided in Multimedia Appendix 5 [42-66].

Targeted interventions for high-risk populations included a range of educational and culturally adapted programs. Interventions that addressed health determinants at the micro-macro levels featured peer-led initiatives tailored for Latinx communities [42-44], a recreational soccer program tailored for Hispanic men [45], a lifestyle education program tailored for Chinese Americans [46], and a lifestyle intervention tailored for Pakistani immigrant men [47]. One multicomponent intervention for US-affiliated Pacific Islanders addressed health determinants at the micro-meso-macro levels [48].

Environmental-based interventions targeted changes in the built or food environments within communities. These interventions addressed health determinants across micro-macro levels. For example, efforts to increase food access and availability were implemented through nutrition programs in recreation centers, corner stores, and carry-out restaurants in Baltimore, United States [49], as well as through initiatives aimed at improving access to food retail in food deserts [50-53]. Additionally, built environment interventions that addressed health determinants at the micro-meso-macro levels included enhancements to footpaths with signage promoting walking [54] and changes implemented within institutions, such as the integration of physical activity programs in schools to reduce barriers in adopting an active lifestyle for children [55].

Community-based interventions spanned various communities defined by region or group membership in an institutional setting and addressed health determinants at multiple levels. One intervention targeted the micro-macro levels by increasing physical activity programming in a materially deprived community in the United Kingdom [56]. Others targeted the micro-meso levels, such as a wellness program for overweight and obese adults within church congregations [57] and a patient-practitioner program for individuals in remote and rural areas [58]. Finally, community-based interventions that targeted health determinants across micro-meso-macro levels included diverse programs promoting physical activity and community

XSL•FO
RenderX

connectedness in low-income or resource-deprived communities in Devon County (United Kingdom) [59], the United Kingdom [60], the Netherlands [61], Ghent (Belgium) [62], and the United States [63]. Other interventions that addressed health determinants across the micro-meso-macro levels included physical activity and nutrition education programs delivered in the school setting, focusing on school-aged children and school teachers [64,65], and a family-based cooking and eating smart program with participants of faith-based organizations [66].

Target Populations in Multilevel Interventions for Diabetes Prevention

Across the included studies, 21 (84%) multilevel interventions targeted disadvantaged communities or population groups characterized by low income, resource deprivation, rural and remote locations, and racialized populations. A total of 4 (16%) multilevel interventions targeted the whole population (eg, an entire community). Targeted interventions for high-risk populations were aimed at racialized groups (100%). Environmental-based interventions primarily targeted low-income or resource-deprived communities (4/7, 57%) and a racialized group (1/7, 14%). Last, community-based interventions targeted low-income or resource-deprived communities (6/11, 55%) and rural/remote communities (3/11, 27%). These details are highlighted in Table 2.

Table . Population targets in multilevel diabetes prevention interventions included in the scoping review (2000 - 2024) according to intervention type.

Intervention type and target population	Number of articles, n (%)
Targeted interventions for high-risk populations (n=7)	
Racialized group (Latinx, Chinese immigrants, Pakistani immigrants, and Pacific Islanders)	7 (100)
Environmental-based interventions (n=7)	
Low-income/resource-deprived	4 (57)
Racialized group (African American youth)	1 (14)
Whole population	2 (29)
Community-based interventions (n=11)	
Low-income/resource-deprived	6 (55)
Rural/remote community	3 (27)
Whole population	2 (18)

Health Determinants Targeted in Multilevel Interventions for Diabetes Prevention

Figure 2 illustrates the health determinants targeted in each multilevel intervention. Among the 25 multilevel interventions, all (n=25, 100%) targeted health determinants at the microlevel. Physical activity (20/25, 80%) and diet (17/25, 68%) were the most common intrapersonal health determinants, while social support from family, friends, or peers (14/25, 56%) was the most common interpersonal health determinant.

A total of 13 of the 25 interventions (52%) addressed health determinants at the mesolevel. Using the school setting for health education and access to health services to promote a healthy school environment (4/25, 16%), as well as leveraging patient-provider relationships within primary care, were the most frequently addressed health determinants at the organization and institutional level. Additionally, leveraging community social supports (5/25, 20%) was the most common health determinant addressed at the community level.

Finally, 23 of the 25 interventions (92%) addressed macrolevel health determinants. These interventions often targeted low-income or economically deprived neighborhoods and communities (10/25, 40%), racialized groups with disproportionate diabetes risk (8/25, 32%), or aimed to enhance the community's built environment (9/25, 36%) and food environment (8/25, 32%).

Intervention Coverage

Intervention coverage was reported in 76% (n=19) of the studies. For interventions targeting high-risk individuals and community-based interventions, coverage was measured based on average session attendance, with program participants as the denominator. Coverage for targeted interventions ranged from 60% to 100%, while for community-based interventions, it ranged from 5.2% to 83.5%. In environmental-based interventions, coverage was primarily measured by activity retention, with the denominator being participants who completed the activity. Coverage for these types of interventions ranged from 18% to 70%. Notably, 6 (24%) studies did not report on coverage. Details on intervention coverage for each study are provided in Multimedia Appendix 5 [42-66].

Outcomes Reported in Multilevel Interventions for Diabetes Prevention

This review focused on outcomes reported in studies across 3 main domains: changes in diabetes risk factors (weight/BMI, physical activity, diet, and blood pressure), changes in metabolic risk, and changes in inequities in diabetes risk (Table 3). While all studies included in the review reported on proximal outcomes related to changes in diabetes risk factors (n=25, 100%), fewer addressed distal outcomes, with only 32% (n=8) assessing changes in metabolic risk. None of the studies reported on changes in population inequities in diabetes incidence, and only

XSL•FO
RenderX

8% (n=2) reported an equity outcome related to a change in the difference in a diabetes risk factor level between disadvantaged and advantaged population groups. A detailed summary of the outcomes reported in each multilevel intervention is presented in Multimedia Appendix 6 [42-66].

XSL•FO
RenderX

Table . Measured outcomes reported in studies of multilevel diabetes prevention interventions included in the scoping review (2000 - 2024) by intervention type.

Intervention type, measured outcomes, and group reported on	Number of articles, n (%)
Targeted intervention for high-risk population (n=7)	
Weight/BMI	
Total population	N/A[a]
Population group	5 (71)
Both	N/A
Physical activity levels	
Total population	N/A
Population group	6 (86)
Both	N/A
Diet	
Total population	N/A
Population group	5 (71)
Both	N/A
Blood pressure	
Total population	N/A
Population group	3 (43)
Both	N/A
Metabolic risk	
Total population	N/A
Population group	4 (57)
Both	N/A
Inequity	
Total population	N/A
Population group	1 (14)
Both	N/A
Environmental-based intervention (n=7)	
Weight/BMI	
Total population	1 (14)
Population group	2 (29)
Both	0 (0)
Physical activity levels	
Total population	0 (0)
Population group	1 (14)
Both	1 (14)
Diet	
Total population	3 (43)
Population group	0 (0)
Both	1 (14)
Blood pressure	
Total population	1 (14)
Population group	1 (14)

XSL•FO
RenderX

Intervention type, measured outcomes, and group reported on	Number of articles, n (%)
Both	0 (0)
Metabolic risk	
Total population	1 (14)
Population group	1 (14)
Both	0 (0)
Inequity	
Total population	0 (0)
Population group	1 (14)
Both	0 (0)
Community-based intervention (n=11)	
Weight/BMI	
Total population	1 (9)
Population group	—[b]
Both	0 (0)
Physical activity levels	
Total population	6 (55)
Population group	—
Both	4 (36)
Diet	
Total population	3 (27)
Population group	—
Both	2 (18)
Blood pressure	
Total population	2 (18)
Population group	—
Both	0 (0)
Metabolic risk	
Total population	1 (9)
Population group	—
Both	1 (9)
Inequity	
Total population	0 (0)
Population group	—
Both	0 (0)

[a]Not applicable; targeted interventions for high-risk populations were not population-wide interventions and, therefore, did not report outcomes for the total population.

[b]Not reported.

Due to the variation in the target populations of the interventions, some studies reported outcomes for the entire population (eg, whole community) (10/25, 40%), while others reported outcomes on specific population subgroups (eg, sex, age, and ethnicity) (9/25, 36%). A smaller number of studies reported outcomes for both the total population and specific subgroups (6/25, 24%).

Across studies, outcomes were measured in different ways (Multimedia Appendix 6 [42-66]). For example, physical activity was measured by walking scores [62,63], self-reported physical activity scores [42,44,66], or moderate to vigorous physical activity sessions per week [48,64,65].

Targeted interventions for high-risk populations reported outcomes for population groups, particularly on changes in risk

XSL•FO
RenderX

factors such as weight/BMI (5/7, 71%), diet (5/7, 71%), physical activity levels (6/7, 86%), and changes in metabolic risk (4/7, 57%). Studies reported significant changes in physical activity [42,43,46,47] and weight [43,45,46,48], while changes in metabolic risk [44-46] and diet [42,45,48] were often insignificant. For instance, one educational program that aimed to reduce cardiovascular risk among Latinas demonstrated increased physical activity levels postintervention, but no significant changes in diet, as measured by the Health-Promoting Lifestyle Profile subscale scores [42]. Another culturally adapted healthy eating and educational nutrition program for Latinas found a reduction in weight, increased physical activity, and higher consumption of fiber-rich foods postintervention [43]. A peer-led, culturally tailored community-based program targeting Latina women reported improved healthy eating scores; however, HbA_{1C} levels remained unchanged [44]. Among Hispanic males, a recreational soccer program led to a reduction in weight and diastolic blood pressure, but there were no notable changes in diet measured by total energy intake, fat consumption, or fruit and vegetable intake [45]. Similarly, a culturally adapted lifestyle education program for Chinese immigrant adults in the United States reported reduced BMI, increased physical activity, improved dietary intake scores, and lower diastolic blood pressure, but fasting plasma glucose and HbA_{1C} levels remained unchanged [46]. Finally, a multicomponent physical activity and counseling program targeted to Pakistani immigrant men in Norway showed an increase in overall physical activity postintervention [47].

Environmental-based interventions generally reported outcomes for the total population and most frequently reported changes in risk factors such as weight/BMI (3/7, 43%) and diet (4/7, 57%). For example, a built-environment intervention that removed barriers to physical activity in schools and targeted African American youth found that normal weight students in the intervention group had a smaller increase in age- and gender-adjusted BMI over time compared with the control group. Similarly, overweight students in the intervention group also experienced a smaller increase in BMI than their control counterparts [55]. Additionally, a nutrition intervention aimed at improving the food environment in community centers and corner stores demonstrated a decrease in BMI-age percentiles in the youth intervention group compared with the control group [49]. However, in another food-environment intervention, the Healthy Food Financing Initiative, where supermarkets were implemented in food deserts, changes in BMI, fruit and vegetable intake, and HbA_{1C} levels in the population showed no significant improvements [50].

Community-based interventions frequently reported outcomes for the total population and primarily reported changes in risk factors, such as physical activity (19/21, 91%) and diet (5/11, 45%). For instance, one intervention that introduced nutrition programs in primary schools and provided information on healthy nutrition and lifestyles to adults resulted in higher fruit self-efficacy scores and increased fruit consumption in intervention neighborhoods but no changes in physical activity scores among participants [61]. Another intervention, which developed community-wide physical activity programs with input from the community, led to increased self-reported walking scores, though there were no significant changes in participants meeting walking recommendations [63]. In a third study, which introduced 38 different types of activities in accessible community spaces, participants reported being more active compared with 1 year earlier [56]. Last, a family-centered, multilevel ecological intervention designed to improve nutrition and physical activity habits resulted in an increased mean change in physical activity frequency and intensity, as measured by the Rapid Assessment of Physical Activity Score. However, no changes were observed in fruit and vegetable consumption scores [66].

Equity Outcomes Reported in Multilevel Interventions for Diabetes Prevention

None of the 25 interventions reviewed reported outcomes related to changes in population inequities in diabetes incidence. Only 2 studies (8%) reported an equity-related outcome between disadvantaged and advantaged population groups, specifically a change in the absolute differences in diabetes risk factor levels by food security and socioeconomic status (SES) [43,55].

The first study, a culturally adapted, community health worker-led healthy eating and nutrition intervention, found that participants with food insecurity (FI) benefited as much as those without FI in terms of improved diet and physical activity [43]. On average, both groups significantly increased their consumption of fiber-rich foods at program completion compared with baseline [43]. Although fiber-rich food consumption decreased between the 8-week and the 3-month follow-up, it remained significantly higher than baseline levels (10.16, no FI; 11.15, yes FI; diff: −0.99). Importantly, the gap in fiber-rich food consumption between food-secure and food-insecure groups narrowed over time, suggesting a positive shift in equity for this dietary measure (12.82, no FI; 12.87, yes FI; diff:−0.05). Physical activity levels also improved significantly in both groups at the 8-week and the 3-month follow-up, but the food-insecure group benefited more. While there was a small difference in physical activity levels favoring the food-secure group at baseline (19.11, no FI; 18.69, yes FI; diff: 0.42), by the 3 mo follow-up, the food-insecure group had achieved even greater improvements, closing and reversing the initial gap in physical activity (25.91, no FI; 34.03, yes FI; diff: −8.12).

The second intervention, which involved implementing environmental changes within a school setting to promote active lifestyles among adolescents, demonstrated effectiveness in increasing physical activity and reducing BMI across both low and high SES groups. The results showed no significant interaction between SES and the intervention (P=.82) [55], indicating that the benefits were equally distributed among students regardless of socioeconomic background.

Discussion

Principal Results

This review systematically mapped multilevel diabetes prevention interventions, providing insights into targeted, environmental-based, and community-based intervention types. Specifically, the review describes the health determinants

XSL•FO
RenderX

targeted across micro, meso, and macro levels of influence, and the key outcomes reported. A central finding is the limited evidence available on the impact of these interventions in reducing population inequities in diabetes risk, highlighting the need for more research examining equity outcomes of multilevel diabetes prevention interventions. Our results revealed a range of health determinants across micro, meso, and macro levels that are acted upon in multilevel interventions for diabetes prevention. We also observed that no single intervention addressed all health determinants that contribute to diabetes risk and inequities, emphasizing that there is no one-size-fits-all solution.

The review points to the necessity of identifying an effective combination of strategies and a tailored approach that meets the unique needs of communities and at-risk populations, aligning with the concept of precision diabetes prevention [67,68]. As part of a precision diabetes prevention approach, data collection and tools are critical to inform new policies, including which multilevel interventions alongside other prevention strategies would be effective in communities in reducing diabetes incidence and inequities. The need for data collection, integration, and analysis to improve diabetes prevention and inform policies has been echoed in recommendations by The Lancet Commission on Diabetes [69]. To ensure direct and more effective consideration of multilevel factors that contribute to diabetes risk in communities, intersectoral collaboration and community engagement are required, including public health, health systems, educational institutions, government bodies, community organizations, and community members. Ultimately, equitable diabetes prevention strategies require an understanding of the lived experience and sociocultural context of communities in order to identify the underlying causes of diabetes disparities and inform effective local strategies and tailoring of interventions to address inequities [14,70].

All studies in this review reported on proximal outcomes of multilevel diabetes prevention interventions, such as changes in weight, physical activity, diet, and blood pressure. Few studies reported on long-term outcomes, such as sustained metabolic health improvements or reductions in inequities in diabetes risk, reflecting a research gap. These findings align with past reviews of other diabetes prevention interventions that have highlighted evidence on short-term outcomes of behavior change (eg, improved diet and increased physical activity), but less frequently reported evidence on long-term population-based outcomes on diabetes incidence and equity [20,21,39,41,71]. This research gap reflects challenges in measurement, particularly because of the long timeframe that is expected from exposure to a preventive intervention and impact on long-term outcomes [72]. Given the time lag to detect long-term outcomes, evaluations of interventions often use short-term outcomes as indicators of long-term impact [73]. It can also be challenging to assess population-level outcomes in targeted multilevel interventions if data are only available on a subset of the population (ie, high-risk individuals). Future research should consider methodologies and tools that enable measurement of the longer-term impacts of multilevel diabetes prevention interventions, potentially through data linkages with population data sources, modeling studies, and mixed methods approaches that allow for qualitative insights into the context of how interventions shape equity outcomes over time. Finally, given that interventions are typically assessed based on the goals they are designed to achieve [74], the gap in evidence on equity highlights the importance of identifying the intended equity impacts of multilevel interventions.

Determining the population health and equity impacts of multilevel diabetes prevention interventions requires considering the relative benefit of the intervention and the expected intervention coverage [75]. Our review demonstrated that population targets and intervention coverages in the included studies varied widely, and there was inconsistent reporting on intervention coverage across studies. Certain communities may experience barriers that result in lower uptake or differential benefits from health-promoting interventions or policies, which can result in the unintended effect of widened health disparities. Engaging the community, health care providers, and other stakeholders during the design and implementation phases may enhance the relevance and success of multilevel interventions [76]. This further endorses the precision diabetes prevention approach, which stresses the importance of tailoring interventions to local community needs and ensuring equitable uptake.

Study Limitations

The scoping review does not present a comparative effectiveness of the different types of multilevel diabetes prevention interventions. In line with scoping review methodology, this review did not include a critical appraisal of the evidence, which would be a considerable challenge given that the heterogeneity in study designs limits the comparability of studies. In addition, the absence of a quality critical appraisal, which would be a challenge, given the variability of study designs, poses limitations for interpreting results. While our search strategy focused on diabetes-specific interventions, we may have missed interventions targeting relevant diabetes risk factors (eg, physical activity and BMI) if the article did not explicitly mention diabetes. Nonetheless, this review has identified important evidence gaps in equity outcomes reported in multilevel diabetes prevention interventions, which suggests directions for future research. In addition, our review focused on multilevel diabetes prevention interventions in high-income economies; therefore, the results may not be generalizable in low- and middle-income countries. Our review was also limited to only English articles, which may have introduced selection bias. The effectiveness of multilevel diabetes prevention interventions implemented in low- and middle-income economies, particularly in countries with pronounced diabetes inequities, can be the focus of future research.

Conclusions

Our review demonstrated that multilevel diabetes prevention interventions have the potential to address population inequities in diabetes risk, given that these are often targeted for disadvantaged at-risk populations and are designed to take action on meso and macro-level health determinants. However, our review highlights that evidence of the impact of these interventions on reducing population inequities in diabetes

XSL•FO
RenderX

incidence has not been examined or reported consistently. The findings from this review highlight the need for future studies to more explicitly address equity considerations in evaluations of multilevel diabetes prevention interventions. The "sick population" paradigm reminds us that focusing on individuals alone may overlook broader social and environmental factors [77]. Addressing diabetes inequities requires interventions that are responsive to community and systemic contexts, considering the distribution of risk within populations and accounting for health determinants across multiple levels of influence. Adopting a precision diabetes prevention approach may facilitate the implementation of interventions that are better aligned with community-specific needs, thereby increasing their relevance and potential impact on reducing inequities in diabetes risk.

Acknowledgments

This research was supported by the Canada Research Chair held by LCR in Population Health Analytics (FRN 72060091) and the Novo Nordisk Network for Healthy Populations 2024 Summer Student Program (RP).

Authors' Contributions

The study was conceptualized and the methodology was developed by KK, RP, DG, LMD, and LCR. Article screening, data extraction, and analysis were conducted by RP and KK. The original draft of the manuscript was written by RP, while KK, DG, LMD, and LCR contributed to reviewing and editing the manuscript. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Conceptual framework to define multilevel interventions and examples of health determinants relevant to diabetes risk.
[DOCX File, 14 KB - publichealth_v11i1e70267_app1.docx]

Multimedia Appendix 2
Electronic database and gray literature search strategy.
[DOCX File, 16 KB - publichealth_v11i1e70267_app2.docx]

Multimedia Appendix 3
Summaries of multilevel diabetes prevention interventions excluded from this scoping review due to an existing review.
[DOCX File, 34 KB - publichealth_v11i1e70267_app3.docx]

Multimedia Appendix 4
Inclusion and exclusion criteria.
[DOCX File, 18 KB - publichealth_v11i1e70267_app4.docx]

Multimedia Appendix 5
Study characteristics of articles included in this scoping review and description of multilevel diabetes prevention interventions, intervention coverage, and target populations.
[DOCX File, 53 KB - publichealth_v11i1e70267_app5.docx]

Multimedia Appendix 6
Description of measured outcomes in studies reporting multilevel diabetes prevention interventions included in this scoping review (Extended).
[DOCX File, 55 KB - publichealth_v11i1e70267_app6.docx]

Checklist 1
PRISMA-SCR checklist.
[PDF File, 181 KB - publichealth_v11i1e70267_app7.pdf]

References

1. Transforming our world: the 2030 agenda for sustainable development. : United Nations; 2015 Published online.
2. Dagogo-Jack S, editor. Diabetes Mellitus in Developing Countries and Underserved Communities: Springer International Publishing; 2017. [doi: 10.1007/978-3-319-41559-8]

XSL•FO
RenderX

3. Mikhail N, Wali S, Brown AF. Ethnic disparities in diabetes. Endocrinol Metab Clin North Am 2021 Sep;50(3):475-490. [doi: 10.1016/j.ecl.2021.05.006] [Medline: 34399957]
4. Lysy Z, Booth GL, Shah BR, Austin PC, Luo J, Lipscombe LL. The impact of income on the incidence of diabetes: a population-based study. Diabetes Res Clin Pract 2013 Mar;99(3):372-379. [doi: 10.1016/j.diabres.2012.12.005] [Medline: 23305902]
5. Agarwal S, Wade AN, Mbanya JC, et al. The role of structural racism and geographical inequity in diabetes outcomes. Lancet 2023 Jul;402(10397):235-249. [doi: 10.1016/S0140-6736(23)00909-1]
6. Misra S, Ke C, Srinivasan S, et al. Current insights and emerging trends in early-onset type 2 diabetes. Lancet Diabetes Endocrinol 2023 Oct;11(10):768-782. [doi: 10.1016/S2213-8587(23)00225-5] [Medline: 37708901]
7. Ismail L, Materwala H, Al Kaabi J. Association of risk factors with type 2 diabetes: a systematic review. Comput Struct Biotechnol J 2021;19:1759-1785. [doi: 10.1016/j.csbj.2021.03.003] [Medline: 33897980]
8. Hassan S, Gujral UP, Quarells RC, et al. Disparities in diabetes prevalence and management by race and ethnicity in the USA: defining a path forward. Lancet Diabetes Endocrinol 2023 Jul;11(7):509-524. [doi: 10.1016/S2213-8587(23)00129-8] [Medline: 37356445]
9. Fazli GS, Moineddin R, Chu A, Bierman AS, Booth GL. Neighborhood walkability and pre-diabetes incidence in a multiethnic population. BMJ Open Diabetes Res Care 2020 Jun;8(1):e000908. [doi: 10.1136/bmjdrc-2019-000908] [Medline: 32601153]
10. Weisman A, Fazli GS, Johns A, Booth GL. Evolving trends in the epidemiology, risk factors, and prevention of type 2 diabetes: a review. Can J Cardiol 2018 May;34(5):552-564. [doi: 10.1016/j.cjca.2018.03.002] [Medline: 29731019]
11. Shubrook JH, Chen W, Lim A. Evidence for the prevention of type 2 diabetes mellitus. J Osteopath Med 2018 Nov 1;118(11):730-737. [doi: 10.7556/jaoa.2018.158]
12. White M, Adams J, Heywood P. How and why do interventions that increase health overall widen inequalities within populations? In: Social Inequality and Public Health: Policy Press; 2009:65-82. [doi: 10.1332/policypress/9781847423207.003.0005]
13. Agurs-Collins T, Persky S, Paskett ED, et al. Designing and assessing multilevel interventions to improve minority health and reduce health disparities. Am J Public Health 2019 Jan;109(S1):S86-S93. [doi: 10.2105/AJPH.2018.304730] [Medline: 30699029]
14. Walker AF, Graham S, Maple-Brown L, et al. Interventions to address global inequity in diabetes: international progress. Lancet 2023 Jul;402(10397):250-264. [doi: 10.1016/S0140-6736(23)00914-5]
15. Noor FA, Gulis G, Soendergaard J. A conceptual framework for chronic disease prevention based on multilevel approach. GJHS 2018;10(5):175. [doi: 10.5539/gjhs.v10n5p175]
16. Paskett E, Thompson B, Ammerman AS, Ortega AN, Marsteller J, Richardson D. Multilevel interventions to address health disparities show promise in improving population health. Health Aff (Millwood) 2016 Aug 1;35(8):1429-1434. [doi: 10.1377/hlthaff.2015.1360] [Medline: 27503968]
17. Cook JE, Purdie-Vaughns V, Meyer IH, Busch JTA. Intervening within and across levels: a multilevel approach to stigma and public health. Soc Sci Med 2014 Feb;103:101-109. [doi: 10.1016/j.socscimed.2013.09.023] [Medline: 24513229]
18. Wnuk K, Świtalski J, Tatara T, et al. Workplace interventions for type 2 diabetes mellitus prevention-an umbrella review. Curr Diab Rep 2023 Oct;23(10):293-304. [doi: 10.1007/s11892-023-01521-3] [Medline: 37728724]
19. Flórez KR, Payán DD, Palar K, Williams MV, Katic B, Derose KP. Church-based interventions to address obesity among African Americans and Latinos in the United States: a systematic review. Nutr Rev 2020 Apr 1;78(4):304-322. [doi: 10.1093/nutrit/nuz046] [Medline: 31539069]
20. Roberts S, Pilard L, Chen J, Hirst J, Rutter H, Greenhalgh T. Efficacy of population-wide diabetes and obesity prevention programs: an overview of systematic reviews on proximal, intermediate, and distal outcomes and a meta-analysis of impact on BMI. Obes Rev 2019 Jul;20(7):947-963. [doi: 10.1111/obr.12821] [Medline: 31039603]
21. Stotz SA, McNealy K, Begay RL, DeSanto K, Manson SM, Moore KR. Multi-level diabetes prevention and treatment interventions for native people in the USA and Canada: a scoping review. Curr Diab Rep 2021 Nov 7;21(11):46. [doi: 10.1007/s11892-021-01414-3] [Medline: 34743261]
22. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
23. Peters M, Godfrey C, McInerney P, Baldini Soares C, Khalil H, Parker D. The Joanna Briggs Institute Reviewers' Manual 2015 Methodology for JBI Scoping Reviews: Joanna Briggs Institute; 2015.
24. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5(1):69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
25. Kornas K, Patel R, Gerstle D, Diemort L, Rosella L. Multi-level diabetes prevention interventions to address population inequities in diabetes risk: a scoping review protocol. : Open Science Framework; 2024 Published online URL: https://osf.io/6h5aq [accessed 2025-08-13]
26. Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. Int J Epidemiol 2001 Aug;30(4):668-677. [doi: 10.1093/ije/30.4.668] [Medline: 11511581]

XSL•FO
RenderX

27. Whittemore R, Melkus GD, Grey M. Applying the social ecological theory to type 2 diabetes prevention and management. J Community Health Nurs 2004;21(2):87-99. [doi: 10.1207/s15327655jchn2102_03] [Medline: 15123438]
28. Göran D, Whitehead M. Policies and strategies to promote social equity in health. : Institute for future studies; 1991 Published online URL: https://core.ac.uk/download/pdf/6472456.pdf [accessed 2025-08-11]
29. Hill-Briggs F, Adler NE, Berkowitz SA, et al. Social determinants of health and diabetes: a scientific review. Diabetes Care 2021 Jan 1;44(1):258-279. [doi: 10.2337/dci20-0053]
30. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75:40-46. [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
31. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new medical research council guidance. BMJ 2008 Sep 29;337:a1655. [doi: 10.1136/bmj.a1655] [Medline: 18824488]
32. O'Cathain A, Croot L, Duncan E, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954. [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
33. World bank country and lending groups. The World Bank Group. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [accessed 2025-07-23]
34. Everson-Hock ES, Johnson M, Jones R, et al. Community-based dietary and physical activity interventions in low socioeconomic groups in the UK: a mixed methods systematic review. Prev Med 2013 May;56(5):265-272. [doi: 10.1016/j.ypmed.2013.02.023] [Medline: 23454537]
35. Pelletier CA, Smith-Forrester J, Klassen-Ross T. A systematic review of physical activity interventions to improve physical fitness and health outcomes among indigenous adults living in Canada. Prev Med Rep 2017 Dec;8:242-249. [doi: 10.1016/j.pmedr.2017.11.002] [Medline: 29181296]
36. Gittelsohn J, Trude A. Diabetes and obesity prevention: changing the food environment in low-income settings. Nutr Rev 2017 Jan;75(suppl 1):62-69. [doi: 10.1093/nutrit/nuw038] [Medline: 28049750]
37. Virgara R, Phillips A, Lewis LK, et al. Interventions in outside-school hours childcare settings for promoting physical activity amongst schoolchildren aged 4 to 12 years. Cochrane Database Syst Rev 2021 Sep 27;9(9):CD013380. [doi: 10.1002/14651858.CD013380.pub2] [Medline: 34694005]
38. Baker PRA, Francis DP, Soares J, Weightman AL, Foster C. Community wide interventions for increasing physical activity. Cochrane Database Syst Rev 2015 Jan 5;1(1):CD008366. [doi: 10.1002/14651858.CD008366.pub3] [Medline: 25556970]
39. Mulchandani R, Chandrasekaran AM, Shivashankar R, et al. Effect of workplace physical activity interventions on the cardio-metabolic health of working adults: systematic review and meta-analysis. Int J Behav Nutr Phys Act 2019 Dec 19;16(1):134. [doi: 10.1186/s12966-019-0896-0] [Medline: 31856826]
40. Peñalvo JL, Sagastume D, Mertens E, et al. Effectiveness of workplace wellness programmes for dietary habits, overweight, and cardiometabolic health: a systematic review and meta-analysis. Lancet Public Health 2021 Sep;6(9):e648-e660. [doi: 10.1016/S2468-2667(21)00140-7] [Medline: 34454642]
41. Hegarty LM, Mair JL, Kirby K, Murtagh E, Murphy MH. School-based interventions to reduce sedentary behaviour in children: a systematic review. AIMS Public Health 2016;3(3):520-541. [doi: 10.3934/publichealth.2016.3.520] [Medline: 29546180]
42. Mudd-Martin G, Martinez MC, Rayens MK, Gokun Y, Meininger JC. Sociocultural tailoring of a healthy lifestyle intervention to reduce cardiovascular disease and type 2 diabetes risk among Latinos. Prev Chronic Dis 2013 Nov 27;10:E200. [doi: 10.5888/pcd10.130137] [Medline: 24286274]
43. Joachim-Célestin M, Rockwood NJ, Clarke C, Montgomery SB. Evaluating the full plate living lifestyle intervention in low-income monolingual Latinas with and without food insecurity. Womens Health (Lond Engl) 2022 Jan;18:174550572210913. [doi: 10.1177/17455057221091350]
44. McCurley JL, Fortmann AL, Gutierrez AP, et al. Pilot test of a culturally appropriate diabetes prevention intervention for at-risk Latina women. Diabetes Educ 2017 Dec;43(6):631-640. [doi: 10.1177/0145721717738020] [Medline: 29059040]
45. Frediani JK, Li J, Bienvenida A, Higgins MK, Lobelo F. Metabolic changes after a 24-week soccer-based adaptation of the diabetes prevention program in hispanic males: a one-arm pilot clinical trial. Front Sports Act Living 2021;3:757815. [doi: 10.3389/fspor.2021.757815] [Medline: 34870194]
46. Chesla CA, Chun KM, Kwong Y, et al. Cultural adaptation of the group lifestyle balance program for Chinese Americans. Diabetes Educ 2016 Dec;42(6):686-696. [doi: 10.1177/0145721716666679] [Medline: 27624905]
47. Andersen E, Burton NW, Anderssen SA. Physical activity levels six months after a randomised controlled physical activity intervention for Pakistani immigrant men living in Norway. Int J Behav Nutr Phys Act 2012;9(1):47. [doi: 10.1186/1479-5868-9-47]
48. Novotny R, Davis J, Butel J, et al. Effect of the children's healthy living program on young child overweight, obesity, and acanthosis nigricans in the US-affiliated Pacific region. JAMA Netw Open 2018 Oct 5;1(6):e183896. [doi: 10.1001/jamanetworkopen.2018.3896]
49. Shin A, Surkan PJ, Coutinho AJ, et al. Impact of Baltimore healthy eating zones: an environmental intervention to improve diet among African American youth. Health Educ Behav 2015 Apr;42(1 Suppl):97S-105S. [doi: 10.1177/1090198115571362] [Medline: 25829124]

XSL•FO
RenderX

50. Richardson AS, Ghosh-Dastidar M, Beckman R, et al. Can the introduction of a full-service supermarket in a food desert improve residents' economic status and health? Ann Epidemiol 2017 Dec;27(12):771-776. [doi: 10.1016/j.annepidem.2017.10.011] [Medline: 29198367]
51. Wrigley N, Warm D, Margetts B. Deprivation, diet, and food-retail access: findings from the leeds 'food deserts' study. Environ Plan A 2003 Jan;35(1):151-188. [doi: 10.1068/a35150]
52. Cummins S, Petticrew M, Higgins C, Findlay A, Sparks L. Large scale food retailing as an intervention for diet and health: quasi-experimental evaluation of a natural experiment. J Epidemiol Community Health 2005 Dec;59(12):1035-1040. [doi: 10.1136/jech.2004.029843] [Medline: 16286490]
53. Ashfield-Watt PAL, Welch AA, Godward S, Bingham SA. Effect of a pilot community intervention on fruit and vegetable intakes: use of FACET (five-a-day community evaluation tool). Public Health Nutr 2007 Jul;10(7):671-680. [doi: 10.1017/S1368980007382517] [Medline: 17381948]
54. Brown WJ, Mummery K, Eakin E, Schofield G. 10,000 steps Rockhampton: evaluation of a whole community approach to improving population levels of physical activity. J Phys Act Health 2006;3(1):1-14. [doi: 10.1123/jpah.3.1.1]
55. Simon C, Schweitzer B, Oujaa M, et al. Successful overweight prevention in adolescents by increasing physical activity: a 4-year randomized controlled intervention. Int J Obes (Lond) 2008 Oct;32(10):1489-1498. [doi: 10.1038/ijo.2008.99] [Medline: 18626482]
56. Cochrane T, Davey RC. Increasing uptake of physical activity: a social ecological approach. J R Soc Promot Health 2008 Jan;128(1):31-40. [doi: 10.1177/1466424007085223] [Medline: 18274328]
57. Ivester P, Sergeant S, Danhauer SC, et al. Effect of a multifaceted, church-based wellness program on metabolic syndrome in 41 overweight or obese congregants. Prev Chronic Dis 2010 Jul;7(4):A81. [Medline: 20550839]
58. Aubrey-Bassler K, Fernandes C, Penney C, et al. The effectiveness of a proven chronic disease prevention and screening intervention in diverse and remote primary care settings: an implementation study on the BETTER 2 program. BJGP Open 2019 Oct;3(3):bjgpopen19X101656. [doi: 10.3399/bjgpopen19X101656] [Medline: 31581121]
59. Solomon E, Rees T, Ukoumunne OC, Metcalf B, Hillsdon M. The Devon active villages evaluation (DAVE) trial of a community-level physical activity intervention in rural south-west England: a stepped wedge cluster randomised controlled trial. Int J Behav Nutr Phys Act 2014 Dec;11(1):94. [doi: 10.1186/s12966-014-0094-z]
60. Phillips G, Bottomley C, Schmidt E, et al. Measures of exposure to the Well London phase-1 intervention and their association with health well-being and social outcomes. J Epidemiol Community Health 2014 Jul;68(7):597-605. [doi: 10.1136/jech-2013-202507] [Medline: 24516117]
61. Kloek GC, van Lenthe FJ, van Nierop PWM, Koelen MA, Mackenbach JP. Impact evaluation of a Dutch community intervention to improve health-related behaviour in deprived neighbourhoods. Health Place 2006 Dec;12(4):665-677. [doi: 10.1016/j.healthplace.2005.09.002] [Medline: 16253541]
62. De Cocker KA, De Bourdeaudhuij IM, Brown WJ, Cardon GM. Effects of "10,000 Steps Ghent". Am J Prev Med 2007 Dec;33(6):455-463. [doi: 10.1016/j.amepre.2007.07.037]
63. Brownson RC, Hagood L, Lovegreen SL, et al. A multilevel ecological approach to promoting walking in rural communities. Prev Med 2005;41(5-6):837-842. [doi: 10.1016/j.ypmed.2005.09.004] [Medline: 16256183]
64. Jago R, McMurray RG, Drews KL, et al. HEALTHY intervention: fitness, physical activity, and metabolic syndrome results. Med Sci Sports Exerc 2011 Aug;43(8):1513-1522. [doi: 10.1249/MSS.0b013e31820c9797] [Medline: 21233778]
65. Van Stappen V, Cardon G, De Craemer M, et al. The effect of a cluster-randomized controlled trial on lifestyle behaviors among families at risk for developing type 2 diabetes across Europe: the Feel4Diabetes-study. Int J Behav Nutr Phys Act 2021 Jul 1;18(1):86. [doi: 10.1186/s12966-021-01153-4] [Medline: 34210333]
66. Torrence C, Griffin SF, Rolke L, Kenison K, Marvin A. Faithful families cooking and eating smart and moving for health: evaluation of a community driven intervention. Int J Environ Res Public Health 2018 Sep 13;15(9):1991. [doi: 10.3390/ijerph15091991] [Medline: 30216995]
67. Herman WH, Ye W. Precision prevention of diabetes. Diabetes Care 2023 Nov 1;46(11):1894-1896. [doi: 10.2337/dci23-0052] [Medline: 37890107]
68. Sathish T. Precision prevention of type 2 diabetes: an approach to revitalize current lifestyle interventions. Diabetes Res Clin Pract 2023 Jun;200:110722. [doi: 10.1016/j.diabres.2023.110722] [Medline: 37207945]
69. Chan JCN, Lim LL, Wareham NJ, et al. The Lancet commission on diabetes: using data to transform diabetes care and patient lives. Lancet 2021 Dec 19;396(10267):2019-2082. [doi: 10.1016/S0140-6736(20)32374-6] [Medline: 33189186]
70. Goode TD, Jack L. The alliance to reduce disparities in diabetes: infusing policy and system change with local experience. Health Promot Pract 2014 Nov;15(2 Suppl):6S-10S. [doi: 10.1177/1524839914545784] [Medline: 25359252]
71. Mackenbach JD, Stuber JM, Beulens JWJ. Evidence on the effectiveness and equity of population-based policies to reduce the burden of type 2 diabetes: a narrative review. Diabetologia 2025 Feb;68(2):281-294. [doi: 10.1007/s00125-024-06330-1] [Medline: 39621105]
72. Gruss SM, Nhim K, Gregg E, Bell M, Luman E, Albright A. Public health approaches to type 2 diabetes prevention: the US national diabetes prevention program and beyond. Curr Diab Rep 2019 Aug 5;19(9):78. [doi: 10.1007/s11892-019-1200-z] [Medline: 31385061]

XSL•FO
RenderX

73. Thomas JC. Outcome assessment and program evaluation. In: The Jossey-Bass Handbook of Nonprofit Leadership and Management: John Wiley & Sons; 2016:444-474. [doi: 10.1002/9781119176558.ch16]
74. Rogers PJ. Using programme theory to evaluate complicated and complex aspects of interventions. Evaluation (Lond) 2008 Jan;14(1):29-48. [doi: 10.1177/1356389007084674]
75. Manuel DG, Rosella LC. Commentary: assessing population (baseline) risk is a cornerstone of population health planning--looking forward to address new challenges. Int J Epidemiol 2010 Apr;39(2):380-382. [doi: 10.1093/ije/dyp373] [Medline: 20100821]
76. Wallerstein N, Duran B. Community-based participatory research contributions to intervention research: the intersection of science and practice to improve health equity. Am J Public Health 2010 Apr 1;100 Suppl 1(Suppl 1):S40-S46. [doi: 10.2105/AJPH.2009.184036] [Medline: 20147663]
77. Rose G. Sick individuals and sick populations. Int J Epidemiol 2001 Jun;30(3):427-432. [doi: 10.1093/ije/30.3.427] [Medline: 11416056]

Abbreviations

FI: food insecurity
PRESS: Peer Review for Electronic Search Strategies
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
SES: socioeconomic status

Edited by A Mavragani; submitted 18.12.24; peer-reviewed by C Mbonu, T Awofala; revised version received 22.06.25; accepted 24.06.25; published 25.08.25.

Please cite as:
Patel R, Kornas K, Gerstle D, Diemert LM, Rosella LC
Multilevel Diabetes Prevention Interventions to Address Population Inequities in Diabetes Risk: Scoping Review
JMIR Public Health Surveill 2025;11:e70267
URL: https://publichealth.jmir.org/2025/1/e70267
doi:10.2196/70267

© Reshma Patel, Kathy Kornas, David Gerstle, Lori M Diemert, Laura C Rosella. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 25.8.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Dissemination and Implementation Approach to Increasing Access to Local Pre-Exposure Prophylaxis (PrEP) Resources With Black Cisgender Women: Intervention Study With Vlogs Shared on Social Media

Mandy J Hill[1], DrPH, MPH, BS; Laurenia Mangum[2], PhD, MPH, LMSW; Sandra J Coker[1], MD; Tristen Sutton[3], BA; Diane M Santa Maria[4], DrPH, MSN

[1]Department of Emergency Medicine, McGovern Medical School, UTHealth Houston, 6431 Fannin St, Houston, TX, United States

[2]Jane Addams College of Social Work, University of Illinois Chicago, Chicago, IL, United States

[3]Tristen Sutton Consulting, Houston, TX, United States

[4]Cizik School of Nursing, UTHealth Houston, Houston, TX, United States

Corresponding Author:
Mandy J Hill, DrPH, MPH, BS
Department of Emergency Medicine, McGovern Medical School, UTHealth Houston, 6431 Fannin St, Houston, TX, United States

Abstract

Background: Black cisgender women account for only 2% of pre-exposure prophylaxis (PrEP)-eligible people in the United States who use PrEP to prevent HIV. Owing to the low PrEP use, Black cisgender women continue to contract HIV more frequently than women from every other racial group. Intervention efforts that can bridge the link between knowing that PrEP prevents HIV and support with access to PrEP are necessary for Black cisgender women.

Objective: The purposes of the vlogs through the campaign were to share information about ways to prevent HIV using PrEP and fact-based education and provide access to PrEP resources with active links to local PrEP providers at local community health centers.

Methods: In Phase 1, the study team formerly piloted full-length video blog posts (ie, vlogs; 10 - 12 min each) with 26 women during an emergency department visit. Using the findings from Phase 1, Phase 2 involved a prospective 6-month social media marketing campaign. The study team led a Texas-Development CFAR-funded pilot grant to disseminate brief vlog snippets (30 s) of excerpts from full-length vlogs with a larger group of Black women in Harris County. Community members, who were aged 18 - 55 years, usually consume content that is often viewed by Black cisgender women (ie, health and beauty) and reside in neighborhoods (based on zip code) in Harris County where most residents are Black or African American. They were shown a series of brief vlog snippets on their social media pages, along with a brief message about PrEP and an active hyperlink to local PrEP resources. The study team assessed implementation outcomes, including the feasibility and acceptability, appropriateness of vlogs, access to PrEP resources at local clinics, and clinical outcomes such as increased PrEP awareness among Black cisgender women.

Results: Within 6 months, the campaign reached 110.8k unique individuals (the number of unique accounts that have seen your content at least once) who identified as women. When stratified by age, video plays (the number of times a video starts playing) at 50% of the vlogs (n=30,877) were most common among women aged 18 - 24 years (n=12,017) and least common among women aged 45 - 54 years (n=658). Key performance indicators showed that 1,098,629 impressions (the number of times a user saw the vlog) and 1,002,244 total video plays resulted in 15,952 link clicks to local PrEP resources.

Conclusions: The campaign demonstrated the feasibility and acceptability of this approach with Black cisgender women and illustrated preliminary effectiveness at supporting access to local PrEP resources with Black cisgender women. Further dissemination and implementation of this approach is necessary to fully assess whether vlog viewership and clicks on links to PrEP resources can meaningfully empower Black cisgender women to access PrEP and help them to assess whether PrEP is personally a useful HIV prevention option.

(JMIR Public Health Surveill 2025;11:e67367) doi:10.2196/67367

XSL•FO
RenderX

KEYWORDS

PrEP; cisgender Black women; social media campaign; PrEP access; HIV prevention; vlogging; dissemination and implementation; pre-exposure prophylaxis; dissemination; implementation; HIV; prevention; human immunodeficiency virus; cisgender; social media; marketing; campaign; education; sexually transmitted diseases; STDs; vlog

Introduction

Background

Pre-exposure prophylaxis (PrEP) describes a medication that is proven to be 99% effective when taken as prescribed to reduce the risk of getting HIV by taking a pill or injection among people who do not have HIV but may be vulnerable to being introduced to or "exposed" to it. The marginal use of PrEP hinders progress toward decreasing the incidence of HIV cases among Black cisgender women. Black women account for nearly 60% of new HIV cases among women in the United States while comprising less than 15% of the female population [1,2] and also accounting for 67% of new HIV diagnoses among all women in the southern part of the United States [3]. Additionally, Black women in Houston are 18.4 times more likely to be living with HIV [4] than other women in the United States. Optimal progress toward ending the HIV epidemic (EHE) goals requires strategies that will interrupt transmission pathways among women in HIV hotspot locations. PrEP, when taken as prescribed as an oral daily or bimonthly injection medication, can reduce the risk of HIV through sex by 99% [5,6]. Interim findings of the PURPOSE 1 trial presented at the AIDS 2024 Conference confirm that lenacapavir, a long-acting injectable PrEP technology that can be taken twice a year, offered 100% protection against HIV transmission among cisgender women enrolled [7]. However, less than 2% of eligible Black cisgender women in the United States take PrEP [8].

The critical gap between the availability of PrEP and use by Black cisgender women, arguably the most HIV-vulnerable demographic of women in the United States, requires the attention of researchers and interventionists. The identified barriers to PrEP uptake with Black women include limited PrEP awareness, low perceived HIV risk, concerns about adverse effects, concerns about costs, limited marketing, and distrust in the health care system [9]. Real-life competing priorities (ie, housing instability and food insecurity) trump HIV prevention for many Black women [10]. Social and structural factors, including racism, sexism, and classism, are often barriers to PrEP access, especially among young Black women [10,11]. Black women have cited medical racism as a key reason for distrusting the medical system and, as a result, not choosing to use PrEP [10-13]. These barriers, in addition to legacy of medical mistrust [13-15], the dehumanization of Black women [16], and stigmatizing experiences within health care settings [15-18], create a seemingly impermeable barrier to the possibility of increasing PrEP use among Black women. However, public messaging of PrEP as an effective HIV prevention message via innovation by culturally immersed thought leaders, researchers, and interventionists who are of the community can be believed to be true, real, and relatable for Black cisgender women.

To date, there is a paucity of research focused on behavioral interventions that can overcome the complex barriers that suppress PrEP uptake among Black cisgender women. As such, progress towards EHE has been slow within and among sexual networks of Black cisgender women and there has been no significant decrease in HIV incidence or any meaningful increase in PrEP uptake within the last decade [8], which is a significant unmet population health need.

The time that it takes between intervention development and integration of innovative, evidence-based interventions (EBIs) into routine clinical practice can worsen the structural and systemic barriers that impede PrEP use among Black cisgender women. This process can take over a decade, leaving the science outdated, EBIs being ineffective, and improvement in clinical outcomes unaddressed. The measures of acceptability and feasibility in this study are poised to inform the adoption, scale-up, and sustainability of EBIs intending to increase PrEP initiation from the perspective of Black cisgender women. The acceptability of EBIs typically refers to the perception among stakeholders that a given treatment, service, practice, or innovation is agreeable, or palatable, which can change over time [19]. In this research, we are interested in Black cisgender women's acceptability of HIV prevention vlogs as a pathway for both information dissemination and linkage to PrEP resources at local clinics. The research team designed the vlogs with an intentional focus on race, gender, and culture to meet the needs of Black cisgender women. Local clinics already use webpages to provide information about PrEP to community members. In this case, vlogs could drive traffic to the clinic webpages. Lastly, the feasibility measure will indicate how well vlogs can be used to motivate access to PrEP. Findings of feasibility and acceptability measures related to the influence of vlog use, as a health communication strategy, through a social media campaign on information access to PrEP knowledge and access to PrEP will produce a meaningful addition to the paucity of research on behavioral interventions that overcome complex barriers to PrEP access and use with Black cisgender women, and may inform pathways to facilitate PrEP uptake with Black women in the future.

Some progress has been made with identifying facilitators to PrEP uptake with Black women, including women's empowerment and advocacy, a need for increased PrEP-specific education, and the positive influence of PrEP-engaged women's testimonials [9]. In a study (2017 - 2019) where 26% of cisgender women (145/565) in NYC initiated PrEP, Latina women (29.7%, 79/565) and Black women (26.1%, 47/565) were more likely to initiate PrEP when compared to White women (16.3%, 14/565) [10]. Findings demonstrated that PrEP initiation was related to PrEP awareness, low income, unstable housing, receipt of navigation services, and reports of noninjection substance use and a recent sexual relationship with an HIV-positive partner [10]. Although some research has been done, there is a need for more focus on health communication

XSL•FO
RenderX

with tailoring to the population. Black women shared a limited awareness of PrEP exacerbated by the lack of Black women-specific marketing [9]. In a qualitative study assessing the permanent perceptions Black cisgender women made based on external influences such as media messaging, researchers determined that an internal belief that PrEP is not for heterosexual people was developed because some Black cisgender women felt that commercials and marketing for PrEP were geared toward male same gender–loving couples and people of trans experience [20]. Opportunities to support Black women-specific social marketing could increase awareness and knowledge regarding the potential sexual health benefits of PrEP [9]. Additional research outcomes suggest that Black women want to know that the out-of-pocket costs for PrEP will be affordable, that they will have insurance coverage, that they will have positive interactions with informed and culturally competent clinical staff, access to a discreet and convenient clinic, and that PrEP will be safe for them [10]. Emerging research has demonstrated a growing interest in social media–based sexual health interventions for Black women [21].

Goal of This Study

This project addressed social and structural determinants of HIV for Black cisgender women by using an intersectional framework that brought together health communication through the lens of culture, race, and gender. Study findings here will contribute to EHE goals by potentially increasing PrEP access among HIV-naïve Black women and informing the Project ROLE intervention, which is currently being developed and will be fully tested to assess the efficacy of a new vlog series among Black cisgender women that will address PrEP knowledge, uptake, persistence, and adherence from the perspectives of both Black women and health care providers of Black women through community health centers and on social media through a NIMH-funded R34 mechanism (1R34MH136826-01). Findings of the Texas Development Center for AIDS Research pilot study substantiated the feasibility of social media as a plausible source of information to Black cisgender women for PrEP access and a pathway to increase uptake for Black cisgender women in Harris County. The purpose of the study was to better understand Black cisgender women's behaviors when engaging in social media-based HIV prevention interventions designed to provide direct access to web-based sexual health resources related to PrEP. Our research question was, "Can a culturally tailored social media-based vlog HIV prevention intervention motivate Black cisgender women to seek online access to PrEP resources?"

Methods

Prior Work

With the awareness of literary findings on the facilitators and barriers to PrEP uptake among Black cisgender women, the research presented here is an extension of a research concept piloted in an emergency department, whereby vlogs, which can be described as videos of blog posts, were used to debunk culturally believed myths in the Black community on how HIV and sexually transmitted infections are transmitted [22-27].

Behavioral Framework of the Prior Study

The premise of the health communication message of the vlogs stemmed from findings of a qualitative research study where Black cisgender women enrolled in an emergency department setting described a minimized value for sex [28]. Women who used the word "just" to describe sexual acts demonstrated a higher prevalence of experience with condomless sex, a previous diagnosis of a sexually transmitted disease (STI), and abuse. Misperceptions about what is safe or risky needs to be proven or refuted in order to promote sexual health and inform healthy sexual decision-making. The goal of the vlogs in the prior study that was a response to the research study (2015 - 2016) was to resonate with Black women and create an opportunity to replace common misconceptions with the correct information needed to inform healthy sexual decision-making. Based on the protocol of the prior study approved by the University of Texas Health Science Center at Houston–Committee for the Protection of Human Subjects (CPHS), the proposed long-term programming stemming from the pilot study was to explore the utility of social media in reaching marginalized populations.

Overview of Methods of the Prior Study

Investigators completed a pilot randomized controlled trial of an HIV prevention intervention of vlogs comparing 2 educational strategies, a storytelling strategy versus an interactive gaming strategy, to assess and compare changes in knowledge pre- and post-intervention. In total, 26 Black cisgender women aged 18 - 45 years who reported heterosexual sex within the last 3 months were enrolled [23-27]. The team used actors who reflected the target audience on the premises of race, gender, and colloquialism as the communication method in an effort to enhance cultural relevance with the intervention's content. The vlog strategy was innovative, relatable, and motivating for Black women.

Study Design

A test of whether vlogs could motivate online access to PrEP resources among Black cisgender women using 2 aims through a 90-day social media campaign was conducted. The findings of Aim 1, which include the development and testing of the vlog intervention, have been detailed in other studies [10].

During Aim 2, the social media campaign's performance was evaluated by assessing whether vlog viewership translated to linking Black women in Harris County to local PrEP resources. We assessed the primary outcome, engagement rate by reach, and secondary outcomes, post interactions, and visualizations (ie, views initiated, viewership, and completion rates) and evaluated the social media campaign's feasibility and acceptability at linking Black cisgender women to local PrEP resources. The social media consultant developed a report with metrics, analytics, and key performance indicators (including engagement with the campaign; vlog reach to Black women, ie, visualization; community engagement; and quantity of vlog material consumed ie, time) that measure study outcomes to quantify how well vlogs performed at motivating independent access to PrEP information as measured by clicks on hyperlinks to local PrEP resources.

XSL•FO
RenderX

Study Population

The social media campaign aimed to engage adult Black cisgender women aged 18 - 45 years who self-reported on their social media page profiles that they were assigned female at birth.

Eligible participants included cisgender women in Harris County who were (1) assigned female sex at birth and currently identified as female, (2) between 18 and 45 years of age, (3) fluent in English, and (4) had a phone or internet access.

Study Procedures

Development of the Marketing Strategy

A certified Social Media Advertisement expert designed a marketing campaign for implementation across social media websites, including Facebook and Instagram, using an algorithm to target individuals self-identifying as female, aged 18 - 45 years, usually consume content that is often viewed by Black cisgender women (ie, health and beauty), and reside in neighborhoods (based on zip code) in Harris County where most residents are Black or African American. These women were shown a series of brief vlog snippets on their social media pages, along with a brief message about PrEP and an active hyperlink to local PrEP resources. This strategy included a bi-monthly evaluation with an adjustment approach to reach a balanced sample and sustain active engagement with the target community. The study team reviewed the algorithm and made revisions as necessary before launching the campaign. In 2022, Facebook changed its demographic collection procedures and does not allow researchers to create inclusion criteria and exclude by race. Therefore, it is probable that non-Black women viewed the material.

Development of the Study's Landing Page

The social media ads provided active links to a landing page that housed access to local PrEP resources. The landing page was housed on the academic health center's website and was co-developed and managed by a web developer and the PI of the grant. The landing page's content included a link to a brief survey, described as a pre-screening tool, where individuals could describe brief demographic factors (eg, race and gender). Below the link to the brief survey, a question, "Did you Know?" is followed by a few sentences describing the status of HIV in the South with Black women and how the Texas D-CFAR pilot project aims to address the current health inequity. A table titled, "Pre-Exposure Prophylaxis (PrEP) Providers in Houston" is provided with the name of each local PrEP-providing agency, their physical address, and active links to their website and social media pages for ease of access to those who visit the page. The bottom of the landing page contains a link to the Texas-Development CFAR pilot protocol page, which provides more details about the study, inclusive of a project summary, background information, objectives, funding and collaborators, as well as a bio of project leadership [24].

Development of 60-Second Vlog Trailers

A study team member with robust vlogging experience on social media created 10 vlog trailers with brief excerpts from each vlog for the social media campaign. The trailers highlighted excerpts on PrEP within engaging aspects of vlogs with intentions to garner the immediate interest of Black women. The study team reviewed all 10 trailers, recommended edits, and collectively chose the top 5. When the trailers were refined and finalized, the social media ad expert launched the campaign with all 5 trailers.

Development of Post Content to Accompany Vlog Trailers

Each social media post included written content with at least 280 characters, including the hyperlink to local PrEP resources on the study's landing page. An example of a post was, "What does it mean to be at risk for HIV? Check out the vlog and click on the link to learn more about PrEP" (83 characters). We included hashtags like #PrEPforHER, #OWNYourHealth, and #PreventHIV, when character limits allowed.

Launching the Campaign

The social media ad expert launched the campaign on March 1, 2023, and concluded on September 8, 2023. Viewers of the vlog trailers were identified by an algorithm led by the social media ad expert who mobilized clicks on advertisements through social media sites using a strategy with established effectiveness. The social media ad expert tracked whether viewers clicked on referral sources. His algorithm determined how often the trailers were shown to social media audiences. Through careful monitoring on a weekly basis, the trailers with the least activity were removed from the campaign in an iterative fashion. The highest performing trailers (n=2) remained for the duration of the campaign. There was no participant enrollment process for individuals to view the social media campaign, and only publicly available data associated with paid advertisements were used to collect data. Individuals were not engaged on an individual basis. The social media campaign was publicly available.

Access to Local PrEP Clinics Through the Social Media Campaign

The project links cisgender women in an EHE priority jurisdiction to local PrEP resources, providing them with access to local PrEP clinic information through a landing page from a social media campaign, and introduced PrEP as an effective HIV prevention option through highly trafficked social media channels as a means of encouraging PrEP initiation with Black cisgender women in Harris County using vlogging for sexual health. Harris County residents who viewed the vlogs on their social media page had access to more information about PrEP by clicking on hyperlinks to local PrEP resources that accompany the vlog trailers. Hyperlinks were connected to websites and social media pages of local PrEP providers for review and following.

Ethical Considerations

The University of Texas Health Science Center at Houston–CPHS determined this study did not meet the regulatory definition of human subjects' research. Thus, no further review was required.

Data Analysis

A report developed with Facebook Insights and Google Analytics was used to evaluate the performance of the social

XSL•FO
RenderX

media campaign. The analytic strategy collected data based on age, gender, race, and interest in health and beauty. Descriptive statistics assessed the self-identified gender and age of the profiles of individuals who engaged with the campaign. The study team also evaluated key performance indicators that included impressions (the number of times a user saw the vlog), video plays (the number of times a video starts playing), and the number of clicks to local PrEP resources (the total number of times users have clicked on a specific link). The study team also evaluated the economic cost of implementing a sexual health–focused social media campaign to increase PrEP access to Black cisgender women.

Results

Intervention Characteristics

Social media platforms that were leveraged for the campaign included Facebook and Instagram. Key performance indicators of the social media campaign stratified by age revealed that young people aged 18 - 24 years engaged most with the campaign (Table 1). Video plays at 50% of the vlogs (n=30,877) were most common among women aged 18 - 24 years (n=12,017) and least common among women aged 45 - 54 years (n=658). Video plays at 50% means that individuals started playing the video and watched 50% of the video. Table 2 shows an overview of women engaging with the social media campaign by age. Across cohorts, women aged 18 - 24 years were mostly represented, followed by a decrease in engagement with social media campaign in each subsequent cohort, with women aged 45 - 54 years being the least represented.

Table . Key performance indicators of the social media campaign, stratified by age (March-September 2023).

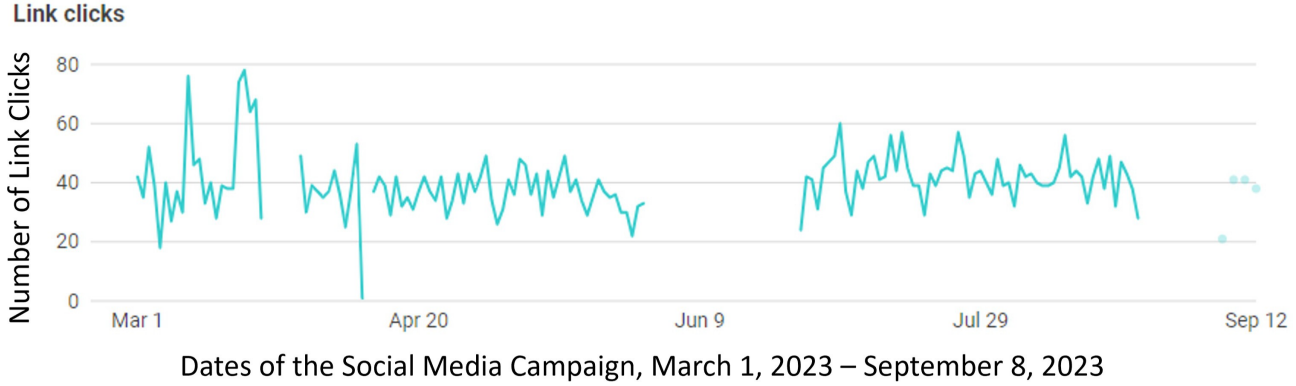
Age (years)	Link clicks	Reach	Impressions	Cost per link click (USD)	Amount spent (USD)	Video plays at 25%	Video plays at 50%	Video plays at 75%	Video plays at 95%	Video plays at 100%	Video plays
18 - 24	2330	39,968	491,656	$1.27	$2967.69	15,752	12,017	10,192	7072	6413	465,262
25 - 34	1834	38,816	336,524	$1.27	$2337.52	16,741	10,129	7494	5652	4913	302,026
35 - 44	1662	29,376	250,782	$1.28	$2124.98	14,534	8073	5204	4292	3567	217,758
45 - 54	126	2592	19,667	$1.43	$180.76	1255	658	413	335	286	17,198
Total	5952	110,752	1,098,629	$1.28	$7610.94	48,282	30,877	23,303	17,351	15,179	1,002,244

Table . Age distribution of women engaging with the social media campaign over 6 months.

Age group (years)	Total count (approximate)
13-17	0
18-24	2250
25-34	1800
35-44	1600
45-54	100
55-64	0
65+	0

A line graph was used to assess the trend in clicks of the campaign over time (Figure 1). The trend reveals a steady link click rate of around 40 across the breadth of the 6-month campaign, largely within a 20 - 60 link click rate.

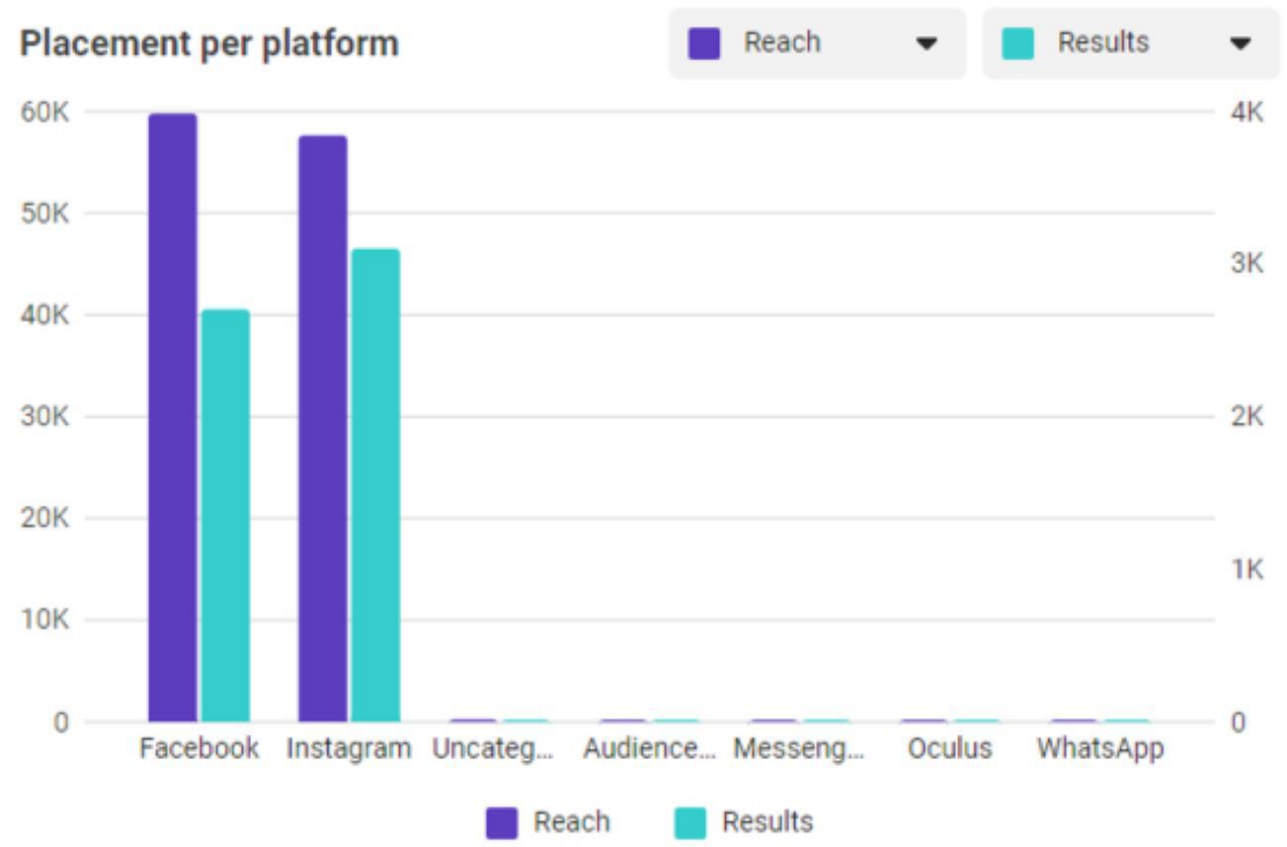
Figure 1. Summary of the trend in link clicks over the breadth of the social media campaign.



XSL•FO
RenderX

The final report showed 1,098,629 impressions and 1,002,244 total video plays and resulted in 15,952 link clicks to local PrEP resources. The campaign reached more individuals on Instagram (59,776) than on Facebook (57,623), but there were more clicks on Facebook (3101) than on Instagram (2703) (Figure 2).

Figure 2. Social media campaign performance based on placement per platform with a mobile device only. All platforms apart from Facebook and Instagram have a value of 0.



The evaluation of the economic cost of the social media campaign revealed that the cost was $1.28 per link click. The total amount spent was $7410.94 across the 6-month campaign. When stratified by social media platform, the campaign performed slightly better on Instagram than on Facebook, costing $3992.85 versus $3411.55, respectively (Table 3).

Table . Key performance indicators: by overall campaign and by platform (March-September 2023).

Key performance indicators and measures	Value
Overall social media campaign	
Reach (post engagement), n	110,752
Total impressions, n	1,098,629
Link clicks, n	5952
Cost (USD)	$7410.94
Cost per click (USD)	$1.28
Platform	
Facebook	
Reach, n	57,623
Link clicks, n	3101
Cost (USD)	$3411.55
Instagram	
Reach, n	59,776
Link clicks, n	2703
Cost (USD)	$3,992.85

XSL•FO
RenderX

Discussion

Study Findings and Comparison With Previous Works

The social media campaign over 6 months reached close to 6000 women. The campaign primarily used Facebook and Instagram. The highest engagement was observed among individuals aged 18 - 24 years, particularly women, who accounted for the most video plays (n=12,017). Engagement was the lowest among women aged 45 - 54 years (n=658). The distribution of engagement showed a consistent decline in participation across older age groups, with women aged 18 - 24 years being the most represented and women aged 45 - 54 years the least. The final report indicated 1,098,629 impressions and 1,002,244 total video plays, leading to 15,952 link clicks to local PrEP resources. The campaign reached more individuals on Instagram (59,776) than on Facebook (57,623), but Facebook generated more clicks (3101) compared to Instagram (2703).

When exploring the distribution of women who engaged with the social media campaign, it was evident that younger women were more likely to be represented. This finding supports existing research demonstrating that mobile health apps or electronic sexual and reproductive health interventions are effective in engaging young Black women [29,30]. Additionally, the findings align with previous research demonstrating that generational differences exist in the ways in which women access health information [28]. Many people access health information through their smartphones as the first step in gathering valid and correct information. This may be different from older generations, such as older millennials, generation X, or boomers, who developed sexual and reproductive health practices before the invention and proliferation of the internet and, as such, sought out professionals such as physicians, nurses, and other health care providers as trusted sources of sexual and reproductive health information. While seeking out medical professionals is a trusted and effective route, mobile health and e-health apps now remove the barrier of access to and utilization of health care and providers, due to the nature of needing insurance or being insured, in addition to costs related to co-pays, transportation, and time to take off work [29]. Additionally, these approaches to HIV prevention remove the burden of actively participating in one's sexual health wellness and accommodate diverse lifestyles [29,31]. Moreover, the findings of this study demonstrate that women aged 18 - 35 years engaged in the social media campaign. Given that newer cases of STIs such as chlamydia, gonorrhea, and syphilis are the highest among women aged 18 - 25 years [19] and new HIV diagnoses are reported among women aged 25 - 35 years [19], it is advantageous to explore the sustainability and scale-up of such an intervention to address the current public health needs. The decrease in engagement as age increased could indicate the need for tailored approaches in future campaigns, perhaps targeting specific age groups with content that resonates more with their interests and preferences [32].

Silva et al [29] sampled male and female youth aged 16 - 49 years on social media (including Twitter, subsequently rebranded X, and Facebook), along with other mass media platforms, in a study supporting intergenerational communication on sexual reproductive health and family planning with youth. While findings show strong potential, findings also showed that young people still need supportive guidance and approval from trusted adults in their region of the world (eg, sub-Saharan Africa) [29]. Another study assessed sexual health promotion campaigns on Twitter and Facebook and found significant potential with health promotion initiatives based on the number of interactions with users [29]. Although social media has been explored for health promotion and sexual health generally, the focus on PrEP engagement appears to be missing. This reveals a meaningful gap and unmet need that this study was poised to fill.

Although minimal, a greater number of women used Instagram as opposed to Facebook. However, link clicks on Facebook were higher than that on Instagram. It would be advantageous to retain both social media sites for future programming, as each site tends to garner a different viewership demographic [33]. Despite the differences in user preferences, both platforms were deemed acceptable to women engaged in the social media campaign over 6 months.

Compared to the over 1 million impressions and over 1 million total video plays, clicks to the local PrEP resources were close to 16,000, demonstrating that while women are interested in the campaign and videos, there is less engagement with PrEP. This may point to a critical point in women's decision-making in terms of their self-assessment of HIV risk or vulnerability and the need to try a new HIV prevention strategy. The novelty of the innovation reinforces the fact that such social media campaigns for Black women are scant, and the scale-up of these interventions would make engaging in HIV prevention practices commonplace and more widely adopted. This study, however, is the first of its kind to actually embed local PrEP resources within the vlogs to drive traffic to the clinics. This is an ideal strategy as it addresses the retention of participants. Findings from this study suggest that social media campaigns as an intervention strategy to increase PrEP access are likely a compatible complement to EHE goals and outreach efforts to enhance access to PrEP providers among populations who are vulnerable to new HIV diagnoses. Video plays were over 1 million views, surpassing the expectations of the investigative team while solidifying the feasibility of this approach at reaching cisgender women through social media with PrEP messaging.

Limitations

There are limitations to our study. First, there were 2 breaks in the campaign, stalled by communication gaps within the team and logistical challenges within the institution occurring between March 23, 2023 and March 30, 2023 and May 30, 2023 and June 27, 2023. Although this occurred, our findings remain robust, and if the study team had not had breaks in the campaign, the results could have been similar or higher for reach, clicks, vlog views, and accessing local PrEP clinic resources. Second, in regard to the racial profile of the study population, the ability of any social media campaign on Facebook or Instagram to market advertisements by race was removed. In lieu of this sudden and unpredicted change, the study team and consultant advertisement expert used 2 proxies in lieu of race, which were content, specifically health and beauty, and zip codes in areas

XSL•FO
RenderX

where a significant majority were Black. As an example, zip code 77033 was one of several zip codes used. According to the United States Postal Service, the majority of residents in 77033, specifically 75.4% as of January 13, 2025, are Black or African American. Individuals who resided in zip code 77033 often viewed health and beauty content and described themselves as female; they were shown the vlogs during the social media campaign. Based on the expertise of our consultant, the team was confident that this approach would garner viewership from a majority Black female population. However, it is likely that the vlogs were viewed by a minority of non-Black women.

Future Research

Future studies examining implementation outcomes such as acceptability could engage the users in focus groups or in-depth qualitative interviews to determine reasons for staying on the site or linking to local PrEP clinics. Given that many clinics operate on limited budgets, it may be advantageous to invest in a formal feedback session on their webpage for opportunities to garner usage of the site further and enhance the user experience and ability to get the information needed quickly. Future studies could also investigate the appropriateness of such an intervention. Additionally, the utilization of a hybrid type 2 design to examine the scale-up and testing of both the intervention effectiveness and implementation strategy would allow for the simultaneous assessment of HIV prevention via vlogs with multiple sites.

Conclusions

The campaign demonstrated the feasibility, acceptability, and appropriateness of this approach with Black cisgender women and illustrated preliminary effectiveness at supporting access to local PrEP resources with Black cisgender women. Further dissemination and implementation of this approach is necessary to fully assess whether vlog viewership and clicks on links to PrEP resources can meaningfully empower Black cisgender women to access PrEP and help them to assess whether PrEP is personally a useful HIV prevention option.

Acknowledgments

We would like to acknowledge the Texas Development Center for AIDS Research (5P30AI161943-04; PI: Giordano, Rice, and Santa Maria) for funding this pilot research. We would also like to thank Carmen Avalos for all of her administrative support. We would like to acknowledge Sarah Sapp, MPH for her research coordination on this project. As the Debunking Myths study served as the foundation for this work, we would like to thank the Texas Emergency Medicine Research Center for funding that investigator-initiated study in 2021, which served as a springboard for this work, that led to the current funded study that will extend this work.

Conflicts of Interest

MH is the recipient of an Investigator Research Award from Merck Pharmaceuticals Project grant # 100099, which funded a qualitative study in Houston and Austin, Texas, with an objective to explore perspectives and beliefs related to PrEP readiness among Black cis women. A small percentage of MH's effort toward the conduct of the research was supported as a part of the funding provided. The primary outcome of this study is different but related to the outcome of the project presented here in this manuscript and overlaps in the geographic location and population of focus.

References

1. Ojikutu BO, Mayer K. HIV prevention among black women in the US-time for multimodal integrated strategies. JAMA Netw Open 2021 Apr 1;4(4):e215356. [doi: 10.1001/jamanetworkopen.2021.5356] [Medline: 33835180]
2. HIV surveillance in urban and nonurban areas through 2017. : National Center for HIV/AIDS VH, STD, and TB Prevention; 2019 Nov 8 URL: https://stacks.cdc.gov/view/cdc/149300 [accessed 2025-03-17]
3. D'Amore J, Hung O, Chiang W, Goldfrank L. The epidemiology of the homeless population and its impact on an urban emergency department. Acad Emerg Med 2001 Nov;8(11):1051-1055. [doi: 10.1111/j.1553-2712.2001.tb01114.x] [Medline: 11691667]
4. AIDSVu. Houston/harris county - rates of persons living with HIV, 2018. 2021. URL: https://aidsvu.org/local-data/united-states/south/texas/houston [accessed 2021-07-02]
5. Aaron E, Blum C, Seidman D, et al. Optimizing delivery of HIV preexposure prophylaxis for women in the United States. AIDS Patient Care STDS 2018 Jan;32(1):16-23. [doi: 10.1089/apc.2017.0201] [Medline: 29323558]
6. Chou R, Evans C, Hoverman A, et al. Preexposure prophylaxis for the prevention of HIV infection: evidence report and systematic review for the US Preventive Services Task Force. J Am Med Assoc 2019;321(22):2214-2230. [doi: 10.1001/jama.2019.2591]
7. Lewin S. Long-acting technologies. Presented at: 25th International AIDS Conference; Jul 29, 2024; Munich, Germany.
8. Conley C, Johnson R, Bond K, Brem S, Salas J, Randolph S. US Black cisgender women and pre-exposure prophylaxis for human immunodeficiency virus prevention: a scoping review. Womens Health (Lond) 2022;18:17455057221103098. [doi: 10.1177/17455057221103098] [Medline: 35699104]

XSL•FO
RenderX

9. Willie TC, Knight D, Baral SD, et al. Where's the "Everyday Black Woman"? An intersectional qualitative analysis of Black Women's decision-making regarding HIV pre-exposure prophylaxis (PrEP) in Mississippi. BMC Public Health 2022 Aug 23;22(1):1604. [doi: 10.1186/s12889-022-13999-9] [Medline: 35999528]
10. Cheek JB, Feldman MB, Montero N, Gambone GF, Hoffman S, Blackstock OJ. Pre-exposure prophylaxis (PrEP) initiation among Black and Latina cisgender women receiving HIV prevention care coordination services in New York City. AIDS Behav 2022 Oct;26(10):3174-3184. [doi: 10.1007/s10461-022-03661-1] [Medline: 35362904]
11. Mayer KH, Agwu A, Malebranche D. Barriers to the wider use of pre-exposure prophylaxis in the United States: a narrative review. Adv Ther 2020 May;37(5):1778-1811. [doi: 10.1007/s12325-020-01295-0] [Medline: 32232664]
12. D'Angelo AB, Davis Ewart LN, Koken J, Bimbi D, Brown JT, Grov C. Barriers and facilitators to pre-exposure prophylaxis uptake among Black women: a qualitative analysis guided by a socioecological model. J Assoc Nurses AIDS Care 2021;32(4):481-494. [doi: 10.1097/JNC.0000000000000241] [Medline: 34171885]
13. Nydegger LA, Dickson-Gomez J, Ko Ko T. A longitudinal, qualitative exploration of perceived HIV risk, healthcare experiences, and social support as facilitators and barriers to PrEP adoption among Black women. AIDS Behav 2021 Feb;25(2):582-591. [doi: 10.1007/s10461-020-03015-9] [Medline: 32886220]
14. Irie WC, Mahone A, Johnson B, et al. "Just the Stigma Associated with PrEP Makes You Feel Like It'S HIV Itself": exploring PrEP stigma, skepticism, and medical mistrust among Black cisgender women in urban and rural counties in the U.S. Deep South. Arch Sex Behav 2024 Mar;53(3):1187-1195. [doi: 10.1007/s10508-023-02769-2] [Medline: 38195827]
15. Heads AM, Hill MJ, Suchting R, Yammine L, Gilmore-Thomas A. Predictors of anticipated PrEP stigma among women with self-reported problematic substance use: implications for engaging women in the PrEP care continuum. Arch Sex Behav 2021 Oct;50(7):2955-2964. [doi: 10.1007/s10508-021-02031-7] [Medline: 34561793]
16. Thomas T, Hill M. Reversing the dehumanization of Black women. In: Black Women and Public Health: Regenerative History, Practice, and Planning: Suny Press; 2022. [doi: 10.2307/jj.18253045.6]
17. Chandler R, Guillaume D, Wells J, Hernandez N. Let Me PREP You to PREP Me: amplifying the voices of Black women and their providers to consider PREP as an HIV prevention option. Int J Environ Res Public Health 2022 Jan 27;19(3):1414. [doi: 10.3390/ijerph19031414] [Medline: 35162438]
18. Drossman DA, Ruddy J. Improving patient-provider relationships to improve health care. Clin Gastroenterol Hepatol 2020 Jun;18(7):1417-1426. [doi: 10.1016/j.cgh.2019.12.007] [Medline: 31843593]
19. Centers for Disease Control and Prevention. Sexually transmitted infections surveillance, 2023. National overview of STIS in 2023 | STI statistics. CDC. 2024.
20. Hill MJ, Sapp S, McCants S, et al. An examination of perceptions among black women on their awareness of and access to pre-exposure prophylaxis (PrEP). Int J Environ Res Public Health 2024 Aug 16;21(8):1084. [doi: 10.3390/ijerph21081084] [Medline: 39200693]
21. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
22. Hill M, Santa Maria D, Coker S, Sapp S, Sutton T. Leveraging video logs to enhance HIV prep access to black women in harris county. Presented at: American Public Health Association (APHA) 2023 Annual Meeting and Expo; Nov 12-15, 2023; Atlanta, GA.
23. Hill M, Sapp S, Sutton T, Coker S, Santa Maria D. Leveraging video logs to support HIV prep access to black women in harris county. Presented at: Inter-CFAR Women and HIV Symposium; Dec 7, 2023; San Francisco, CA.
24. Hill MJ, Coker S. Novel use of video logs to deliver educational interventions to Black women for disease prevention. West J Emerg Med 2022 Feb 28;23(2):211-221. [doi: 10.5811/westjem.2021.12.54012] [Medline: 35302455]
25. Coker S, Hill M. Debunking myths on HIV risks: a targeted strategy to challenge perceptions of low HIV/STI risks among African American women during an emergency department visit. Presented at: National Medical Association Annual Meeting, Emergency Medicine Section; 2018; Orlando, FL.
26. Coker S, Hill M, Stotts A. Leveraging the emergency department visit to debunk myths on HIV risks among substance-using African American women. Presented at: The College on Problems of Drug and Alcohol Dependence Annual Meeting; 2019; San Antonio, TX.
27. Hill M, Coker S, Hudson T. Leveraging the emergency department visit to debunk myths on HIV and STI risks among african american women. Presented at: The National Conference for HIV, HCV, STI, and LGBTQ Health; 2020; Virtual URL: https://www.scitechnol.com/peer-review/just-an-indicator-of-minimized-value-of-the-sexual-act-nHxD.php?article_id=7768 [accessed 2025-03-12]
28. Hill M, Andrews S, Granado M, Nielsen E, Grimes R. Just': an indicator of minimized value of the sexual act. HIV AIDS Res J 2018;1(2):1-9.
29. Silva M, Loll D, Ezouatchi R, et al. Evaluating a youth-designed sexual and reproductive health mass and social media campaign in Côte d'Ivoire: triangulation of three independent evaluations. Sex Reprod Health Matters 2023 Dec;31(1):2248748. [doi: 10.1080/26410397.2023.2248748] [Medline: 37728549]

XSL•FO
RenderX

30. Tesema N, Guillaume D, Francis S, Paul S, Chandler R. Mobile phone apps for HIV prevention among college-aged Black women in Atlanta: mixed methods study and user-centered prototype. JMIR Form Res 2023 Feb 23;7:e37987. [doi: 10.2196/37987] [Medline: 36821362]
31. Texas Emergency Medicine Research Center. TX-DCFAR Pilot Project: McGovern Medical School at UTHealth Houston; 2022. URL: https://www.cdc.gov/sti-statistics/annual/summary.html#:~:text=Findings&text=In%202023%2C%20over%202.4%20million,1.6%20million%20cases%20of%20chlamydia [accessed 2025-03-12]
32. Chandler R, Hernandez N, Guillaume D, Grandoit S, Branch-Ellis D, Lightfoot M. A community-engaged approach to creating a mobile HIV prevention app for Black women: focus group study to determine preferences via prototype demos. JMIR Mhealth Uhealth 2020 Jul 24;8(7):e18437. [doi: 10.2196/18437] [Medline: 32706723]
33. Alhabash S, Ma M. A tale of four platforms: motivations and uses of Facebook, Twitter, Instagram, and Snapchat among college students? Social Media + Society 2017 Jan;3(1). [doi: 10.1177/2056305117691544]

Abbreviations

CPHS: University of Texas Health Science Center at Houston's Committee for the Protection of Human Subjects
EBI: evidence-based intervention
EHE: ending the HIV epidemic
PrEP: pre-exposure prophylaxis
STI: sexually transmitted disease

Edited by A Mavragani, T Sanchez; submitted 09.10.24; peer-reviewed by D Dinh, M Pyra; revised version received 13.01.25; accepted 15.01.25; published 28.03.25.

Please cite as:
Hill MJ, Mangum L, Coker SJ, Sutton T, Santa Maria DM
Dissemination and Implementation Approach to Increasing Access to Local Pre-Exposure Prophylaxis (PrEP) Resources With Black Cisgender Women: Intervention Study With Vlogs Shared on Social Media
JMIR Public Health Surveill 2025;11:e67367
URL: https://publichealth.jmir.org/2025/1/e67367
doi:10.2196/67367

© Mandy J Hill, Laurenia Mangum, Sandra J Coker, Tristen Sutton, Diane M Santa Maria. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 28.3.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Interventions to Foster Mental Health and Reintegration in Individuals Who Are Unemployed: Systematic Review

Sophia Helen Adam[1], MSc; Florian Junne[2], MD; Svenja Schlachter[3,4], PhD; Miriam Mehler[3], MSc; Harald Gündel[3], MD; Nicolas Rüsch[4], MD; Jörn von Wietersheim[3], PhD; Katrin Elisabeth Giel[1], PhD; Stephan Zipfel[1], MD; Rebecca Erschens[1], PhD

[1]Department of Psychosomatic Medicine and Psychotherapy, Internal Medicine, University Medical Hospital Tuebingen, Tuebigen, Germany

[2]Department of Psychosomatic Medicine and Psychotherapy, Otto-von-Guericke University Magdeburg, Magdeburg, Germany

[3]Department of Psychosomatic Medicine and Psychotherapy, Ulm University Medical Center, Ulm, Germany

[4]Department of Psychiatry II, Ulm University and BKH Guenzburg, Ulm, Germany

Corresponding Author:
Sophia Helen Adam, MSc
Department of Psychosomatic Medicine and Psychotherapy
Internal Medicine
University Medical Hospital Tuebingen
Osianderstraße 5
Tuebigen, 72076
Germany
Phone: 49 07071 29 ext 80191
Email: sophia.adam@med.uni-tuebingen.de

Abstract

Background: Unemployment is a risk factor for the development and perpetuation of psychological distress. Finding support for affected individuals can be particularly challenging, which causes a vicious cycle of psychological distress and unemployment.

Objective: The aim of this systematic review is to assess and summarize existing evidence regarding interventions that address both mental health and re-employment, emphasizing accessibility through community or social care structures.

Methods: A systematic literature search using PubMed and EBSCOhost and an additional search using reference list screening were conducted according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. In order to identify interventions for the mental health and re-employment of individuals experiencing psychological distress and unemployment, an inclusion process according to the PICO (population, intervention, comparison, and outcome) scheme and the study design was applied. Title and abstract screening and full-text screening for eligibility were performed independently by 2 reviewers. Quality assessments using the Cochrane Risk of Bias Tools for randomized and nonrandomized trials were conducted by 2 independent reviewers.

Results: The initial systematic search yielded 4442 results, and 15 articles were additionally identified via reference list screening. Eventually, 74 articles were subjected to a thorough evaluation process by 2 independent reviewers. The interrater reliability was determined to be good, with a Cohen κ score of 0.770. After a multistep extraction process, 17 studies remained for inclusion, with each focusing on the improvement of mental health, re-employment, or both outcomes. A heterogeneous pattern of results emerged, with most interventions showing improvement in either mental health or re-employment. Most studies were assessed as having a moderate (n=7) or high (n=9) risk of bias.

Conclusions: The results of the systematic research indicate that low-threshold services in close cooperation with institutions and exchange with other supportive stakeholders should be fostered. Derivable overarching themes and intervention content for integrative support measures can serve as guidelines for future interventions.

Trial Registration: PROSPERO CRD42022378490; https://www.crd.york.ac.uk/PROSPERO/view/CRD42022378490

(JMIR Public Health Surveill 2025;11:e65698) doi:10.2196/65698

KEYWORDS

cognitive behavioral therapy; depression; job loss; mental health; re-employment; unemployment

XSL•FO
RenderX

Introduction

Background

The salutogenic and protective effects of employment have been repeatedly confirmed [1,2], emphasizing that employment has psychologically relevant functions that go far beyond earning money. Despite international statistics showing a steady downward trend in unemployment rates (European Union: 6.1%; United States: 5.4% [3,4]), it remains a persistent societal issue. This is due to the yet unforeseeable effects of migration and economic transformation processes, including the introduction of artificial intelligence. Furthermore, due to the prevalence of less stable employment relationships and a trend toward self-employment in modern working biographies, flexible and remote working conditions, as well as periods of unemployment, are likely to persist as common features of many working biographies.

The Vicious Cycle of Mental Illness and Unemployment: A Glass Ceiling When Looking for Support

Unemployment is accompanied by financial distress, and there is also a well-researched negative association between mental health and unemployment [5]. Numerous studies have shown that the prevalence rates of clinically relevant mental health problems, such as symptoms of depression, anxiety disorders, and psychosomatic illnesses, are increased in the unemployed population compared with the employed population [6,7]. While the point prevalence of depression in the general population ranges between 6.38% [8] and 12.9% [9], an aggregate increased prevalence of 24% persists for individuals who are unemployed [7]. Furthermore, unemployment was shown to be a negative predictor for the effectiveness of psychotherapeutic interventions [10]. The adverse effects of unemployment are particularly severe for those experiencing additional difficulties. This often results in additional barriers to re-employment, which in turn leads to a self-reinforcing cycle [5,11].

Despite the evidence-based salutogenic effect of employment [6,12], which is further supported by growing research on individual placement and support [1,13,14] and the frequent wish for normal employment among people with mental illness [15,16], a negative correlation between the severity of mental illness and the likelihood of re-employment has been found [17].

Similarly, general subjective psychological stress can reduce the re-employment rate of newly unemployed individuals [18]. In the long run, mental disorders and associated limitations can prolong the job search process, as poor physical and mental health can negatively impact the chances of re-employment [19]. It was also shown that people with severe mental disorders, such as schizophrenia and bipolar disorder, have significantly lower levels of employment before and after being diagnosed. Hakulinen et al [20] showed that more than half of all people with a serious mental disorder had no earned income after being diagnosed. However, even if individuals do find employment, their integration can be challenging as employers hold stigmatizing attitudes and prejudices toward those with mental illnesses. For example, employers may view them as being less reliable [21]. Thus, frequently encountered challenges when working with a target group that is subjected to multiple stressors cannot be disregarded. Hagen et al [22] noted that individuals in social welfare systems face difficulties accessing treatment services. This may partly be attributed to the stigma associated with both unemployment and mental illness, leading to barriers for seeking health care and psychosocial support [23-25].

The Need for Low-Threshold and Dual-Path Interventions

Individuals experiencing psychological distress and unemployment have an urgent need for specific interventions aimed at improving their mental health in addition to the essential goal of re-employment. Given the difficulties outlined above, employment agencies often play a crucial role in providing support, for example, by offering initial assistance, as they can provide viable access to support and initial contact with further support mechanisms. However, despite the urgent need to provide individuals experiencing psychological distress and unemployment with adequate support [26] to act preventively and to ensure long-term employment as well as reduce the risk of chronification, tailored programs for unemployed people with mental stress and illness are still rare [16,27].

Several previous reviews and meta-analyses have already addressed interventions in the context of mental illness and unemployment. Some of them have focused on health-promoting interventions for unemployed individuals in general (eg, those also including physical exercise [28-30]). Others have highlighted the effects of interventions on mental health outcomes, but they did not include re-employment as an outcome [31,32]. van Rijn et al [33] and Audhoe et al [34] investigated the effects of vocational interventions on re-employment outcomes. The results indicated that especially subpopulations at high risk of depression benefit the most from interventions [32]. However, most reviews have shown marginal to small effects on health outcomes and quality of life [28,31] and weak evidence for the effectiveness of vocational interventions on re-employment and mental health outcomes [34]. Eventually, Koopman et al [30] described the greatest benefit in combining health promotion approaches with re-employment.

While previous reviews primarily focused on either interventions targeting mental health outcomes or those aiming to improve re-employment rates, this systematic review provides a comprehensive synthesis of interventions that address both aspects simultaneously. By considering both re-employment and mental health as interconnected outcomes, this review extends prior work that has examined these domains separately. Moreover, a distinctive contribution of this review is its focus on interventions that are accessible through community or social care structures, rather than those limited to clinical or workplace settings. Thus, this systematic review provides a synthesis of available research on interventions and programs to promote the mental health and re-employment of individuals affected

XSL•FO
RenderX

by mental strain and unemployment. This systematic review was conducted to investigate the following research questions:

1. What types of interventions have been implemented to promote mental health and facilitate re-employment among unemployed individuals, and what are their characteristics and outcomes?
2. How effective are these interventions in enhancing mental health outcomes and improving re-employment rates?

Methods

Search Strategy and Data Extraction

All data are included in this report. Data were analyzed using EndNote version 20 [35] and visualized using the risk of bias visualization (robvis) tool [36]. The procedures and results of this systematic review are in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [37,38]. For the PRISMA checklist, see Multimedia Appendix 1. This study has been registered with PROSPERO (CRD42022378490). The databases PubMed and EBSCOhost (Academic Search Premier, APA Psych articles, APA PsycInfo, MEDLINE, PSYNDEX, and Business Source Premiere) were searched on October 3, 2022. For the search terms used, see Multimedia Appendix 2. Abstracts and titles were included in each of the searches. The initial search was updated on July 7, 2023. Moreover, ongoing and planned clinical trials were searched using the International Clinical Trials Registry Platform (ICTRP). The inclusion strategy was based on the PICO (population, intervention, comparison, outcome) framework and the study design [39]. After identifying relevant studies through the systematic database search, reference list screening was conducted by examining the reference lists of included studies [40].

For a detailed overview of the inclusion and exclusion criteria applied, see Textbox 1. The *population* included individuals identified as being unemployed, long-term unemployed, or job-seeking between the ages of 18 and 67 years. This age range was chosen since it indicates the population available for the labor market and covers the international distribution of retirement ages. Populations showing heightened levels of psychological distress were included. *Interventions* that were at least partly derived or modified from a psychological theory, a manual, or a therapeutic school (eg, cognitive behavioral therapy [CBT]) were included. Only studies with some type of control group were included. The following o*utcomes* were included: (1) subsyndromal psychological outcomes indicating well-being and stress, as well as syndrome-oriented outcomes such as measures of depression and anxiety, and (2) re-employment measured by re-employment rates or self-report.

Textbox 1. PICO (population, intervention, comparison, outcome) criteria applied for the inclusion of studies.

Inclusion criteria

- Population: Individuals must meet at least one of the following conditions: (1) or (2) or (3), and (4)

1. Classified as long-term unemployed or jobless/job-seeking
2. Receiving social benefits
3. Experiencing psychological distress and recruited via social service institutions
4. Available for the labor market (age 18-67 years)

- Intervention: Interventions must include (1) or (2); Form of presentation: face-to-face or web-based/online; Intensity: any

1. Derived or modified from a psychological theory or manual
2. Interventions including individual placement and support but also incorporating psychological/psychotherapeutic or group therapy elements

- Comparison: Any kind of control group (eg, waitlist and treatment as usual)
- Outcome: Psychological outcomes (eg, well-being, distress, depression, and anxiety) and re-employment or employment status
- Study design: All kinds of pre-post interventions
- Language: English or German

Exclusion criteria

- Population: Individuals who are employed but have a low income, are sick-listed due to mental or physical disorders, are experiencing severe physical illness (eg, cancer) and have lost their job due to their health condition, are experiencing severe mental illness (eg, psychosis and acute addiction), are not primarily recruited from job agencies but from clinics or other institutions, and are younger than 18 years
- Intervention: Solely physical activity (eg, smoking prevention) or solely vocational training; Interventions exclusively targeting health behavior change (smoking, drinking, or exercising)
- Comparison: No control group
- Outcome: Publications without results; Publications reporting no psychological measures at all
- Study design: Case studies and qualitative data only
- Language: Not English or German

XSL•FO
RenderX

Summarizing and Analyzing the Data

All references found were imported into an EndNote database [35]. Duplicates were deleted using the EndNote database tool in the first step. Further duplicates were deleted manually by the first author (SHA). All articles were selected using a multistage screening process supervised by a senior researcher (RE). All articles were first screened independently by title and abstract by 2 reviewers (SHA and RE) and categorized as "eligible" or "ineligible." To assess interrater reliability, Cohen kappa was calculated. In the next step, potential studies for inclusion were assessed for eligibility on the basis of the full-text analyses (SHA and MM). In unclear cases or cases of diverging opinions, a senior researcher (FJ) was consulted.

If reported or applicable, data on intervention content and characteristics, study sample, and study results were extracted for each article (SHA and SS). In case of unclear information or lack of information, the authors were contacted. To evaluate the validity and quality of the included studies, risk of bias was assessed using the Revised Cochrane Risk-of-Bias Tool for Randomized Trials (RoB 2) [41] and the Risk of Bias in Nonrandomized Studies - of Interventions (ROBINS-I) tool [42]. In each case, the assessments were carried out in parallel by 2 independent researchers (SHA and MM).

In the case of studies classified as randomized controlled, content was rated as having *low risk of bias*, *some concerns*, or *high risk of bias* in the following domains: (1) randomization process, (2) deviations from intended interventions, (3) missing outcome data, (4) measurement of the outcome, and (5) selection of the reported result. Regarding the domain *risk of bias due to deviations from the intended interventions*, the subdomain *effect of adherence to intervention* was chosen to be examined as it presents a better fit to the overall question in the sense of a supply decision for an individual. For included randomized controlled trials, the overall risk-of-bias judgement was assessed as follows. A study was judged to be at low overall risk of bias if all 5 domains were judged as having low risk of bias. The overall judgment was "some concerns" if the study was rated to have some concerns in at least one of the domains. High risk of bias was applied if the study was rated to be at high risk of bias in at least one domain or if it was rated to have some concerns for multiple domains, causing a serious deterioration in the quality of the study.

A similar procedure was followed for nonrandomized controlled studies. In accordance with the guidelines by Sterne et al [42], studies were classified in the following bias domains: (1) selection of participants, (2) classification of interventions, (3) deviations from intended interventions, (4) missing data, (5) measurement of outcomes, and (6) selection of the reported result. Each domain was rated, resulting in overall judgements of low, moderate, serious, or critical risk of bias or no information available. A study was judged to be at an overall low risk of bias if all domains were at low risk. Moderate risk of bias was applied when the study was judged to be at low or moderate risk of bias for all domains. As blinding of both patients and treatment providers is difficult to achieve and maintain for nonpharmacological interventions [43], a higher risk of bias in this domain should be taken into consideration.

Overall serious risk was applied if a serious concern in at least one domain but no critical risk of bias for any domain was shown. No information at the overall level was applied in case of an unclear indication of serious or critical risk and a lack of information in one or more key domains of bias. The outcome variables to be investigated were defined in advance in collegial agreement and, if applicable, analyzed separately or grouped together. To visualize risk of bias, the robvis tool was used [36]. Differences were discussed between the researchers in the event of disagreement between raters. The senior researcher (RE) was consulted if the disagreement could not be resolved. Moreover, authors of the included studies were contacted in case clarifications were needed.

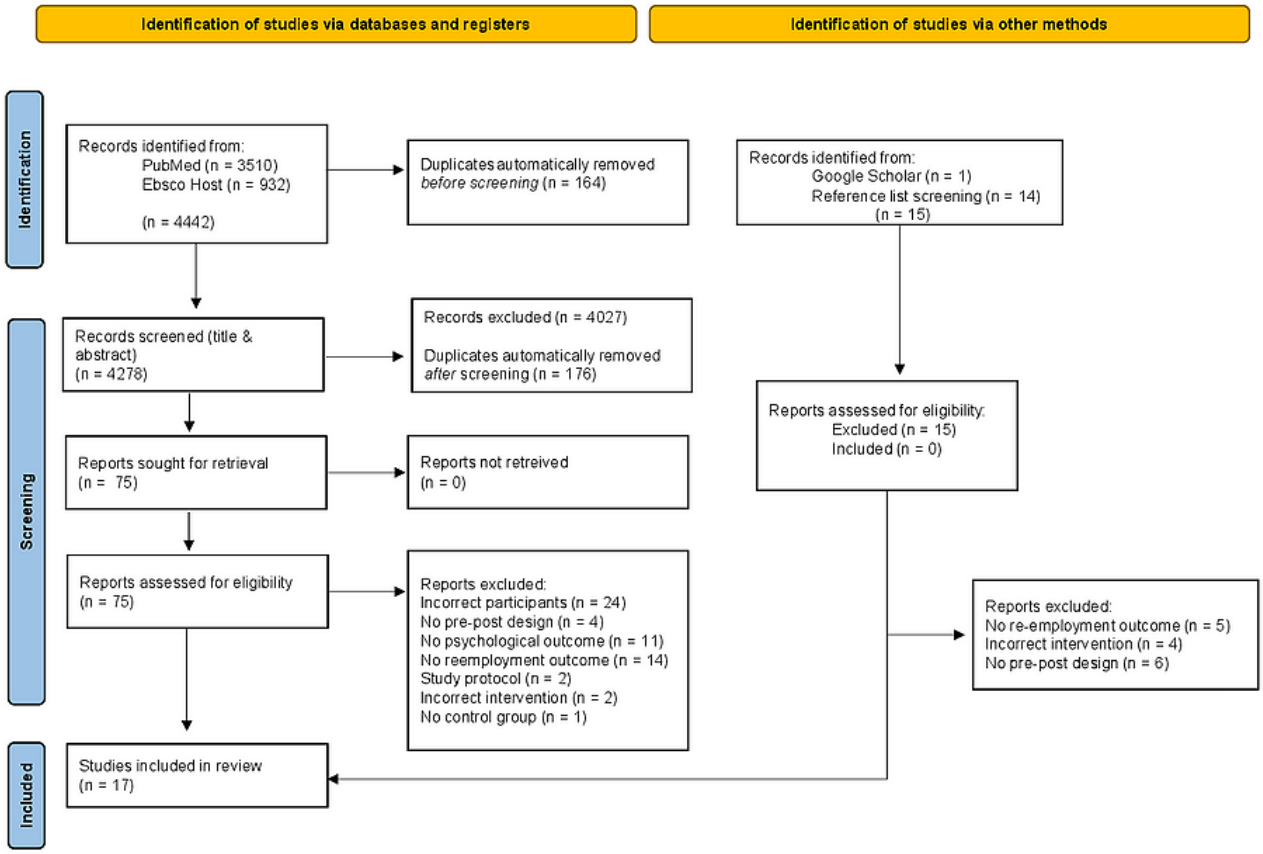
Results

Included Studies

The systematic electronic database search yielded 4442 articles. In addition, 15 articles were found via the manual search process (reference list screening). In 2 steps, 340 duplicates were removed, leaving 4278 articles to be screened. Of these articles, 4203 were excluded during the screening process. The interrater reliability was found to be good, with a Cohen κ of 0.770. The full texts of 74 articles were assessed for eligibility by the 2 independent reviewers, and furthermore, 58 articles were excluded, resulting in the final inclusion of 17 articles. The PRISMA flowchart is presented in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.



Interventions to Promote Mental Health and Enhance Re-Employment Among Unemployed Individuals

To provide an answer to the first research question, "What types of interventions have been implemented to promote mental health and facilitate re-employment among unemployed individuals, and what are their characteristics and outcomes?", we identified a diverse set of interventions across the included studies. The following section provides an overview of the included studies, summarizing their key characteristics, intervention types, and outcomes. Detailed overviews of the content, characteristics, and results of the studies are presented in Tables 1 and 2, and Multimedia Appendices 3 and 4. All included studies were published between 1989 and 2018. The size of the intervention groups ranged from 19 [44] to 1249 [45]. Studies were conducted in the United States [45-48], the Netherlands [49,50], Germany [51,52], Australia [44,53,54], Finland [55,56], the United Kingdom [57], and Ireland and the United Kingdom [58-60].

The target population was partly the general unemployed population [44,46,50,52,55-57,63] and partly unemployed people affected by multiple burdens, such as mentally distressed individuals [45,47-49,59], or otherwise disadvantaged individuals showing barriers to being placed in the labor market [51,53,54]. Moreover, the mean length of unemployment within the intervention groups ranged between 1 week [45] and 5 years [60].

The included studies were either randomized controlled studies [44-48,51,53,55-57,59] or nonrandomized studies that used a quasiexperimental design [49,50,52,54,58,60,63]. The studies also differed in terms of their recruitment methods. The most common and frequently used recruitment path was via the employees of employment agencies or respective institutions [49,50,52,54,58,60,63], followed by recruitment via the associated research team [45,46,48,49,55,56]. Proudfoot et al [57] used a different approach by placing an advertisement. A combination of methods (eg, advertisement and via an agency) was used by Hulshof [50] and Rothländer et al [52]. Several studies [45,48,55,56,58] used the principles of the JOBS program by Caplan et al [46], with either some country-specific [55,56,58] or content-specific adaptations [45,48]. The JOBS program is a structured, cognitive-behavioral intervention designed to improve job-seeking behavior and mental well-being among unemployed individuals. The program is grounded in the theory that unemployment is not only a financial issue but also a psychological one, where an individual's sense of self-efficacy, motivation, and coping skills are negatively impacted. The JOBS program aims to enhance participants' job-search self-efficacy, increase their motivation to seek work, and improve their mental health by addressing setbacks and building resilience. Moreover, 7 of the included trials used CBT-based interventions [44,47,52-54,57,59]. The remaining studies used either individualized approaches, such as interdisciplinary counseling [49], or individualized support by psychologists and physicians [51]. Hulshof [50] used means of positive psychology fostering psychological capital.

XSL•FO
RenderX

Table 1. Summary of the included interventions.

Study author (year), country	Intervention name and theoretical framework	Intervention type	Intervention format (length/setting)	Intervention focus and content
Barry et al [58] (2006), Ireland	Winning New Jobs Intervention; Based on the theory of JOBS [46,61]. Manual content adapted for the Irish population.	Face-to-face	5 half-day training workshops (20 hours)/groups of 12-22	Job-related skills and mental health promotion; Skilled trainers, who have been trained specifically to deliver the JOBS program, active learning, self-efficacy, sense of control and job-related self-efficacy, and inoculation against setbacks to deal with setbacks and obstacles.
Caplan et al [46] (1989), United States	JOBS I intervention; Intervention affects motivation and skills, which in turn can affect job-seeking behavior. Motivation is enhanced by prior success and undermined by setbacks.	Face-to-face	8 sessions of 3 hours each distributed over 2 weeks, 4 times per week (24 hours)/groups of 16-20	Improving job-related skills; Application of problem-solving and decision-making processes, inoculation against setbacks, receiving social support and positive regard from the trainers, and learning and practicing job-seeking skills.
Carlier et al [49] (2018), Netherlands	Vicious cycle of mental illness and unemployment; Unemployed individuals with mental health problems require both job-related and well-being–related interventions to be re-employed.	Face-to-face	Ongoing; maximum 2 years/individualized	Individual placement and support; Interdisciplinary teams counsel unemployed individuals toward employment. Addressing barriers to entering paid employment (eg, psychological problems and debt); providing psychological resources through coaching and cognitive counseling and employment activities tailored to the individual needs of the unemployed person.
Della-Posta et al [44] (2006), Australia	Effectiveness of cognitive behavioral therapy (CBT) to improve mental well-being; Improved well-being, in turn, supports successful re-employment.	Face-to-face	4 hours per week for 4 weeks; first 2 weeks: standard job search assistance; second 2 weeks: CBT (16 hours)/groups of approximately 5	Psychotherapy; Stress and relaxation, planning pleasant activities, communication, and positive and negative thinking; importance of clear communication and differences between passive, aggressive, and assertive styles of communication; simple breathing techniques and guided imagery to more effectively manage stress; incorporation of pleasant activities into their daily routine; and distinguishing between positive and negative thinking.
Harris et al [53] (2002), Australia	Effectiveness of CBT to improve mental well-being; The mental health of individuals who are particularly disadvantaged is more likely to be negatively affected by long-term unemployment; hence, an intervention supporting mental health is highly relevant to this sample.	Face-to-face	11 hours/groups of 5-16	Psychotherapy; Based on the intervention by Rose [54]; Cognitive restructuring, problem solving, and behavioral strategies including relaxation skills.
Herbig et al [51] (2012), Germany	AmigA-M; Highly individualized integration of job placement measures with health-related measures should lead to the improvement of mental health and re-employment.	Face-to-face	Over 6 months: average of 8 contacts with their respective case manager (duration of 44 min), 4 contacts with a physician (duration of 36 min), and 3 contacts with a psychotherapist (duration of 54 min)/individualized	Improving mental health and re-employment; Employment-oriented case management with a health-related focus, individualized support by psychologists and physicians by providing diagnostics, advisory support for the client, and decision-making aids for the case manager.
Himle et al [47] (2014), United States	Effectiveness of CBT for the treatment of individuals with social anxiety disorder (SAD) plus adding work-related content to CBT training to target mental health and unemployment.	Face-to-face	Eight 2-hour sessions held twice weekly over the course of 4 weeks (16 hours)/groups	Psychotherapy and job-related skills; Manualized group CBT for SAD by Heimberg and Becker [62]; elements of psychoeducation, cognitive restructuring, and exposure exercises; limited social skills training related to the work environment and standard vocational services.

XSL•FO
RenderX

Study author (year), country	Intervention name and theoretical framework	Intervention type	Intervention format (length/setting)	Intervention focus and content
Hulshof et al [63] (2020), Netherlands	Based on several theories: (1) conservation of the resources model (COR; people strive to protect resources, and well-being declines when people are threatened by the loss of valued resources); (2) job demands-resources model (JD-R [64]); (3) psychological capital (PsyCap) [65] stating that an individual's positive psychological state of development is characterized by 4 elements: self-efficacy, optimism, hope, and resiliency; (4) experiential learning theory [66] and goal-setting theory [67,68]. Experiential learning theory emphasizes the importance of past experiences in learning and behavioral change. The 4 stages relevant for learning are all incorporated in the intervention.	Face-to-face	Three days (day 1: 5.5 hours; day 2: 5.5 hours; day 3: 3 hours) (14 hours)/groups of maximum 12 participants per group	Psychotherapy and job-related skills; Re-employment crafting, providing theory and practicing all aspects of PsyCap (eg, optimism by the positive visualization technique), theory and practicing with re-employment crafting (eg, setting SMART goals), and small homework assignments (eg, gratitude exercise and acts of kindness).
Maguire et al [59] (2014), United Kingdom	Effectiveness of CBT for mental health, which should, in turn, support employability plus adding work-related content to CBT training to target mental health and provide skills to obtain employment.	Face-to-face	10 days over 5 weeks	Psychotherapy and job-related skills; CBT-based exercises as well as job-seeking skills; key themes: CBT, self-determination theory, and career guidance.
Proudfoot et al [57] (1997), United Kingdom	CBT can support a healthy attributional style, which is affected by losing employment and failing to find re-employment.	Face-to-face	7 sessions of 3 hours, 1 per week over a 7-week period (21 hours)/groups of 10-15	Psychotherapy; Sessions: introduction to a cognitive model, automatic thoughts, goal-setting, time management, task breakdown, thought recording and common thinking errors, planning, techniques to change unhelpful thinking, gaining access to deeper beliefs, dimensions of attributional thinking, specific applications to personal and work situations, integration of strategies, action planning, and relapse prevention; assignment between sessions.
Reynolds et al [60] (2010), Ireland/United Kingdom	Theoretical framework of the JOBS program by Caplan et al [46]; The mechanisms include job search self-efficacy, sense of mastery, and inoculation against setbacks.	Face-to-face	5 sessions of 4 hours each distributed over 1-2 weeks (20 hours)/groups	Job-related skills; Based on Caplan et al [46] and Vinokur et al [45]: active teaching and learning methods, skilled trainers who have been trained specifically to deliver the JOBS program, and supportive learning environments; self-efficacy: training processes that enhance general confidence, sense of control and job-related self-efficacy for participants, and inoculation against setbacks.
Rose [54] (2001), Australia	Effectiveness of CBT to improve mental well-being.	Face-to-face	8 sessions over 2 days with 5.5 hours each (11 hours)/groups	Psychotherapy; "Strategies for re-employment" program; 3 core components: CBT skills, structured problem-solving, and relaxation skills.
Rothländer et al [52] (2012), Germany	AktivA; Effectiveness of CBT to improve mental well-being; Integration of the health action process approach [69] to consider the phases of health-related behavior (motivation, volition, and action).	Face-to-face	4 course days distributed over 2-4 weeks (24 hours)/groups	Psychotherapy; Elements of CBT: planning of activities, constructive thinking, social competencies and social support, and systematic problem-solving.
Vinokur et al [45] (1995), United States	JOBS II intervention (based on Caplan et al [46]).	Face-to-face	5 sessions of 4 hours each distributed over 1 week (20 hours)/groups of 12-22	Improving job-related skills; Application of problem-solving and decision-making group processes, inoculation against setbacks, provision of social support and positive regard from the trainers, and learning and practicing job search skills; 2 trainers (male/female) per group.

XSL•FO
RenderX

Study author (year), country	Intervention name and theoretical framework	Intervention type	Intervention format (length/setting)	Intervention focus and content
Vinokur et al [48] (2000), United States	Referring to Caplan et al [46] and Vinokur et al [45] - JOBS and JOBS II intervention.	Face-to-face	5 sessions of 4 hours each distributed over 1 week (20 hours)/groups of 12-22	Job-related skills; Based on Caplan et al [46] and Vinokur et al [45]: application of problem-solving and decision-making group processes, inoculation against setbacks, provision of social support and positive regard from the trainers, and learning and practicing job search skills; 2 trainers (male/female) per group.
Vuori et al [56] (2002), Finland	Työhön Job Search Program; Transferring JOBS II [46] to the European/Finnish context with minor procedural changes.	Face-to-face	5 sessions of 4 hours each distributed over 1 week (morning sessions; 20 hours)/groups of 6-17	Job-related skills; The training is designed to increase participants' job-search self-efficacy and motivation as well as to enhance the following job-search skills: recognizing and communicating one's marketable skills, identifying and using one's social network to find job openings, contacting promising employers, drawing up a job application and resume, and preparing for successful job interviews.
Vuori et al [55] (2005), Finland	2-year follow-up of the study by Vuori et al [56].	Face-to-face	5 sessions of 4 hours each distributed over 1 week (morning sessions; 20 hours)/groups of 6-17	Job-related skills; The training is designed to increase participants' job-search self-efficacy and motivation as well as to enhance the following job-search skills: recognizing and communicating one's marketable skills, identifying and using one's social network to find job openings, contacting promising employers, drawing up a job application and resume, and preparing for successful job interviews.

Outcome measures included the assessment of depression, distress, and well-being. A detailed overview of the instruments used is presented in Multimedia Appendix 5.

The way in which re-employment was assessed varied from using self-report measures for hours worked per week [45-48,51,54-56,59] to using objectively assessed data, for instance, when data on re-employment was provided by a job agency [44,49-51,63]. In some studies, further information was not provided on the source of the data on re-employment and additional information could not be obtained from the authors [52,53,57,58]. Parts of the interventions mostly focused on re-employment, for instance, by using methods to improve specific job search skills and inoculation against setbacks [45,46,48,55,56,60] or via methods from the field of individual placement and support [49]. On the other hand, some interventions were more focused on the improvement of mental health outcomes, for example, by using target group–adapted CBT-specific content such as planning euthymic activities, cognitive reconstruction, problem solving, social skills, and relaxation methods [44,52-54,57]. Some studies focused on both re-employment and mental health improvement [47,50,51,59,63].

The interventions varied in their theoretical foundations, delivery formats, and target outcomes, yet several overarching patterns emerged. CBT-based interventions (eg, [44,53,57,59]) primarily focus on improving psychological well-being, often through cognitive restructuring, stress management, and problem-solving strategies. These interventions have demonstrated positive effects on mental health, but their direct impact on re-employment remains inconsistent. In contrast, job-search training programs (eg, [45,46,60,70]) emphasize enhancing job-seeking self-efficacy and resilience against setbacks. These interventions tend to show stronger effects on re-employment rates. Additionally, integrative approaches, such as individualized case management (eg, [51]) or interventions combining psychotherapy with job-related skills development (eg, [47,63]), seek to address both psychological and employment-related barriers. While such integrative models may offer a more comprehensive support system, their effectiveness in improving employment outcomes remains variable.

XSL•FO
RenderX

Table 2. Summary of the included studies with an overview of the study methodologies and key results.

Study author (year), country	Control condition	Study design	Recruitment path	Results on mental health	Results on re-employment
Barry et al [58] (2006), Ireland	• Treatment as usual (TAU)	Quasiexperimental design	Via employment agencies	No significant difference in depressive symptoms (for both measurement points).	At the 4-month follow-up: significant increase in re-employment rates for the intervention group (IG); At the 12-month follow-up: the program effect on re-employment was maintained (P<.001).
Caplan et al [46] (1989), United States	• Active control group (CG) • Received a brief booklet with job-seeking tips	Randomized field experiment	Directly via researchers within job centers	Whole sample: no significant findings; Comparison of re-employed and unemployed: re-employed had lower values for depression and anger.	Whole sample: status of re-employment: significant at T2 and T3 (greater percentage of re-employment in the IG); Re-employed sample: higher percentage in the IG for the main occupation.
Carlier et al [49] (2018), Netherlands	• Passive control (TAU) • Participation in usual re-employment programs (with various foci) • Maximum duration of 1 year	Quasiexperimental design	Directly via a researcher (contact information provided by an employment center)	Physical health: no significant difference; Mental health: no significant difference; Anxiety and depressive symptoms: no significant difference.	No significant difference.
Della-Posta et al [44] (2006), Australia	• Active CG • Standard job search assistance group (16 hours; 4 hours for 4 weeks)	Quasiexperimental design	Via a vocational rehabilitation center	Depressive symptoms: no significant difference.	The IG shows re-employment within a shorter amount of time (at T1 and T2); no significant difference at the T3 follow-up.
Harris et al [53] (2002), Australia	• Active CG • First aid course	Controlled trial design with pre-post measures	Indirectly via employment counselors	Mental health: no significant difference.	No significant difference.
Herbig et al [51] (2012), Germany	• Waitlist CG	Randomized controlled trial	Via employment agencies	Significant changes in depressive syndromes in the IG. More participants in the IG than in the CG showed improvements in major depressive syndrome.	No significant intervention effect on re-employment.
Himle et al [47] (2014), United States	• TAU as vocational service as usual (VSAU); Contents amongst others: career assessment, résumé construction, job interviewing skills, job placement assistance	Randomized field experiment	Via a vocational rehabilitation center	Significant differences on nearly all scales in predicted direction, with the IG reporting better mental health and reduced social anxiety symptoms.	No significant difference.
Hulshof et al [63] (2020), Netherlands	• Waitlist CG	Pre-posttest design	Invitation by the first author and via emails and letters sent by the unemployment agency	Significant time × group effect of the intervention between groups for negative affect (F=5.39; P<.001).	No significant intervention effect on re-employment.
Maguire et al [59] (2014), United Kingdom	• Waitlist CG	Randomized, waitlist controlled design	Indirectly via referral from job centers, general practitioners, employment counselors, and further sources	Depressive symptoms: T2: significant difference with lower values in the IG; T3: significant difference with lower values in the IG.	T3: 20 individuals found employment; not statistically tested.

XSL•FO
RenderX

Study author (year), country	Control condition	Study design	Recruitment path	Results on mental health	Results on re-employment
Proudfoot et al [57] (1997), United Kingdom	• Active CG (social support program of the same duration); Sessions: (1) concept of social support, health, and unemployment; (2) life satisfaction graphs, peaks and troughs, and importance of people; (3) role-mapping, satisfying relationships, and personal support networks; (4) social awareness, how to create a positive impression, and rules of relationships; (5) listening and conversation skills; (6) people as resources; (7) goal-setting and course summary plus assignment between sessions	Controlled, experimental, 2-group design	Advertising study	Mental health: significant differences at T2 and T3.	Significant differences at T3: IG more likely to find a job than CG (T2 not relevant as directly at the end of training).
Reynolds et al [60] (2010), Ireland/United Kingdom	• TAU in training and employment agencies	Quasiexperimental design	IG: selected by training and employment agencies, mostly "difficult to place;" CG: recruited by researchers within the same agencies	Depressive symptoms: no significant difference.	Significant difference: higher percentage of re-employment in the IG.
Rose [54] (2001), Australia	• Active CG • First aid course	Randomized controlled trial experimental design	Indirectly via employment consultants	Mental health: no significant difference	No significant difference in the proportion of recruits achieving a positive employment outcome in the IG and CG; employment-related status was recoded as "work and study" or "no work and study."
Rothländer et al [52] (2012), Germany	• TAU: participation in employment measures that do not include AktivA elements	Quasiexperimental design	IG 1: intervention integrated in employment measures of the employment agency; IG 2: intervention freely advertised in newspapers and posters	Health/well-being - psychological: T1-T2: IG 1: significant difference with lower levels of psychological complaints in the IG; IG 2: significant difference with lower levels of psychological complaints in the IG; T1-T3: IG 1: no significant difference (no time × group interaction); IG 2: significant time × group interaction: benefits of training persist mid-term.	No significant difference in the frequency of being unemployed.
Vinokur et al [45] (1995), United States	• Active CG • Received a brief booklet with job-seeking tips	Randomized field experiment	Directly via researchers within job centers	Depressive symptoms: significant interaction at T2 and T3; No significant difference for low-risk participants compared to the CG; Lower levels of depressive symptoms in the high-risk IG compared to the high-risk CG.	Significant at T2, interaction at T3; T2: re-employment higher in the IG (independent of risk status); T3: interaction of group and risk status: no significant difference between the CG and low-risk IG, but higher re-employment for the high-risk IG compared to the CG.

XSL•FO
RenderX

Study author (year), country	Control condition	Study design	Recruitment path	Results on mental health	Results on re-employment
Vinokur et al [48] (2000), United States	• Active CG • Received a brief booklet with job-seeking tips	Randomized field experiment	Directly via researchers within job centers	Long-term effects of the intervention; significant beneficial effects on depressive symptoms.	Intervention had significant beneficial effects on being employed (>20 hours per week).
Vuori et al [56] (2002), Finland	• Active CG • CG received a literature package, which corresponded to the basic themes in job-search training and included four guides: (1) a guide to the services of the employment agency, (2) a guide to manage one's life situation while unemployed, (3) an ABC guide for the job seeker, and (4) a handout of job-seeking advice	Randomized field experiment	Directly via researchers within job centers and outside	No significant effect for depressive symptoms.	Direct effect of the intervention at T3: no significant effect on re-employment rate, wage rate, job, and satisfaction; significant effect on stable re-employment: higher rate of re-employment in stable jobs in the IG.
Vuori et al [55] (2005), Finland	• Active CG • CG received a literature package, which corresponded to the basic themes in job-search training and included four guides: (1) a guide to the services of the employment agency, (2) a guide to manage one's life situation while unemployed, (3) an ABC guide for the job seeker, and (4) a handout of job-seeking advice	Randomized field experiment	Directly via researchers within job centers and outside	Depressive symptoms: intervention status significantly and negatively predicted depressive symptoms.	No significant difference.

Effects on Mental Health and Re-Employment

To address the second research question, "How effective are these interventions in enhancing mental health outcomes and improving re-employment rates?", we examined the effects of each intervention on both mental health and re-employment outcomes. Relevant mental health outcome criteria were defined in advance by the authors.

Significant improvements in mental health outcomes in the intervention groups after receiving the program were shown in 9 of the 17 studies [45,47,48,51,52,55-57,59]. It should be noted that the studies by Vinokur et al [48] and Vuori et al [71] are follow-up measurements of the respective studies and thus show the long-term effects of the interventions. The study by Hulshof et al [50] shows the significant effects on mental health in the intervention group; however, this was due to an increase in negative affect in the control group.

Regarding re-employment, improvements in the intervention group in terms of an increased rate of re-employment after receiving the intervention were shown in 7 of the 17 studies [44-46,48,57,58,60]. In 9 of the 17 studies, no significant difference could be shown [47,49-56,63]. Maguire et al [59] reported that 20 individuals found employment; however, this finding was not statistically tested, and thus, it remains unclear if the intervention had a significant effect. Vuori et al [56] reported no significant effect on re-employment; however, significant effects on higher rates of re-employment in stable jobs in the intervention group were shown.

Moreover, several studies have reported challenges in the implementation of programs related to cooperation with participating organizations and institutions. On the one hand, this concerns aspects of the research design, since in many cases, randomization was not desired or feasible on the part of the agencies [53,58,60]. On the other hand, a lack of cooperation on the part of the agencies or the authorities was also reported, which led to deficient implementation [54]. Moreover, in some cases, further problems in the recruitment process due to the participants or their illness-related characteristics were described [57].

Risk of Bias in the Included Studies

A detailed overview of all the ratings and judgements can be found in Multimedia Appendix 6.

XSL•FO
RenderX

Randomized Controlled Trials

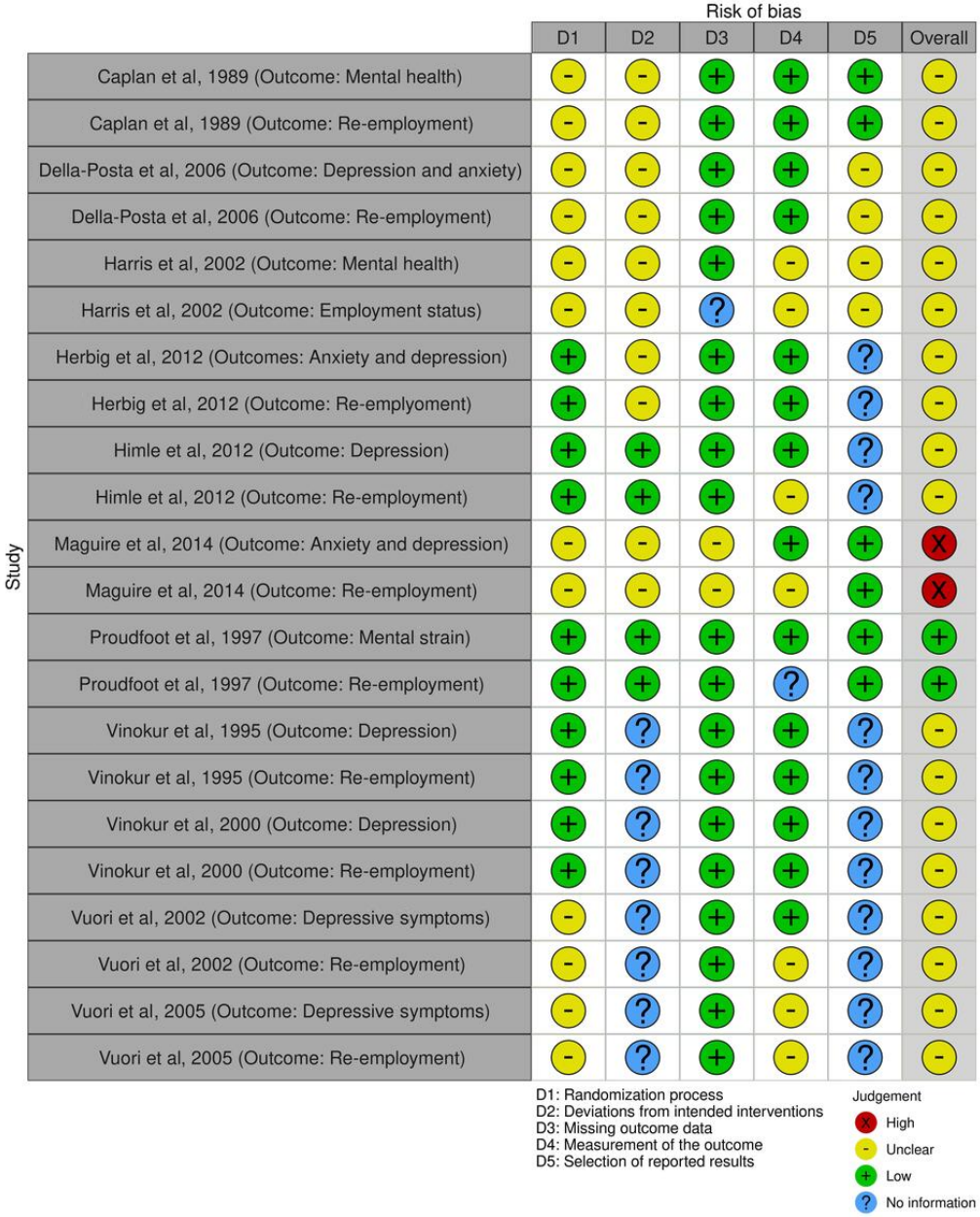
Regarding randomized controlled trials, the overall quality was rated as "low risk" in 1 article [57], indicating a low risk of bias in all domains. Moreover, 9 articles [44-48,51,53,55,56] received an overall rating of "some concerns" (some concerns in at least one domain; no domain judged to have a high risk of bias). Furthermore, 1 article [59] was judged to be at high risk of bias (defined as having a high risk of bias in at least one domain or as having some concerns for multiple domains). Moderate or weak ratings mainly resulted from risk of bias arising from the randomization process or due to no information in relevant domains, especially regarding the selection of reported results.

Nonrandomized Trials

One article [63] was judged to be at an overall moderate risk of bias (low or moderate risk of bias for all domains). Three articles [52,54,60] were found to have a serious risk of bias (serious risk of bias in at least one domain, but not at critical risk of bias in any domain). Lastly, 1 article [49] was found to have a critical risk of bias (critical risk of bias in at least one domain). One article [58] had an overall judgement of "no information" on which to base a judgement about risk of bias. Critical and serious ratings mainly resulted from bias due to the selection of participants and missing data.

Plots of the domain-level judgements for each individual result and each outcome are provided in Figures 2 and 3. Figures 4 and 5 display the overall percentage of judgements for each of the criteria [36].

Figure 2. Visualized risk of bias assessment for the included randomized controlled trials [44-48,51,53,55-57,59], evaluating 5 domains: randomization process (D1), deviations from intended interventions (D2), missing outcome data (D3), measurement of the outcome (D4), and selection of reported results (D5). Each study has been assessed for mental health and re-employment outcomes.



XSL•FO
RenderX

Figure 3. Visualized risk of bias assessment for the included nonrandomized controlled trials [49,52,54,58,60,61], evaluating the following domains: bias due to confounding (D1), bias due to selection of participants (D2), bias in classification of interventions (D3), bias due to deviations from intended interventions (D4), bias due to missing data (D5), bias in measurement of outcomes (D6), and bias in selection of the reported result (D7). Each study is assessed for mental health and re-employment outcomes.

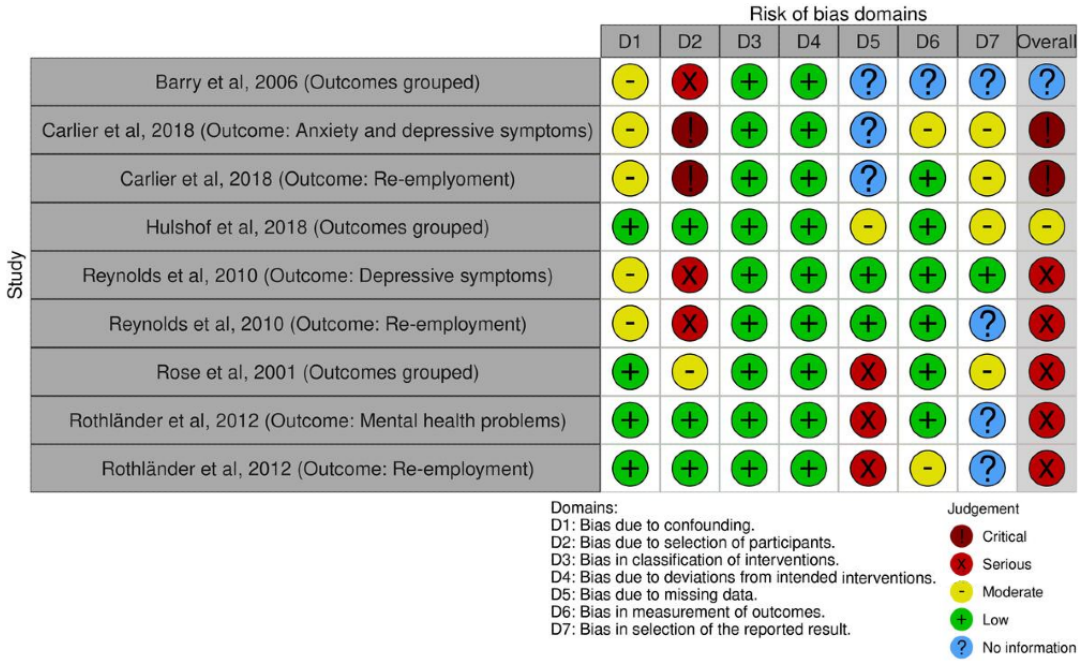


Figure 4. Visualized summary of risk of bias across all included randomized controlled trials, categorized by 5 domains: randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of reported results, as well as the overall risk of bias. The proportions of the risk categories are displayed as horizontal bars, providing an overview of the methodological quality of the included studies.

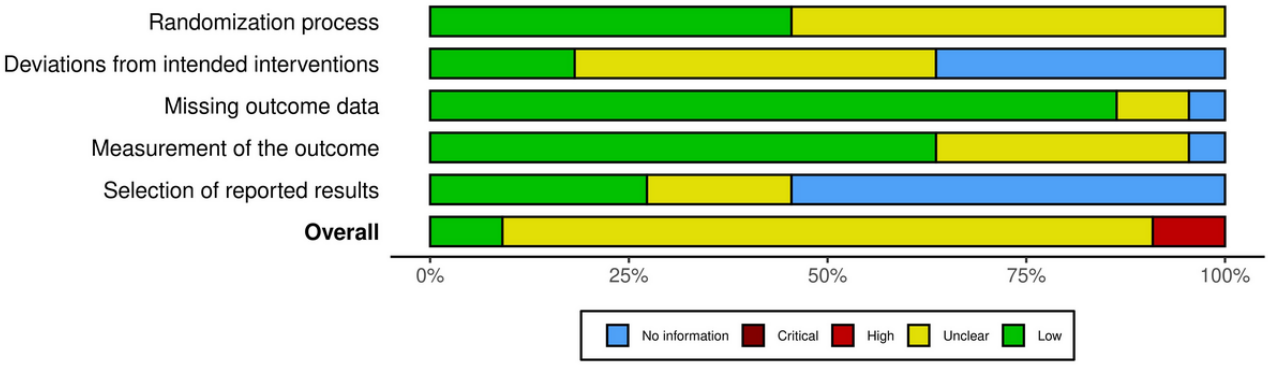
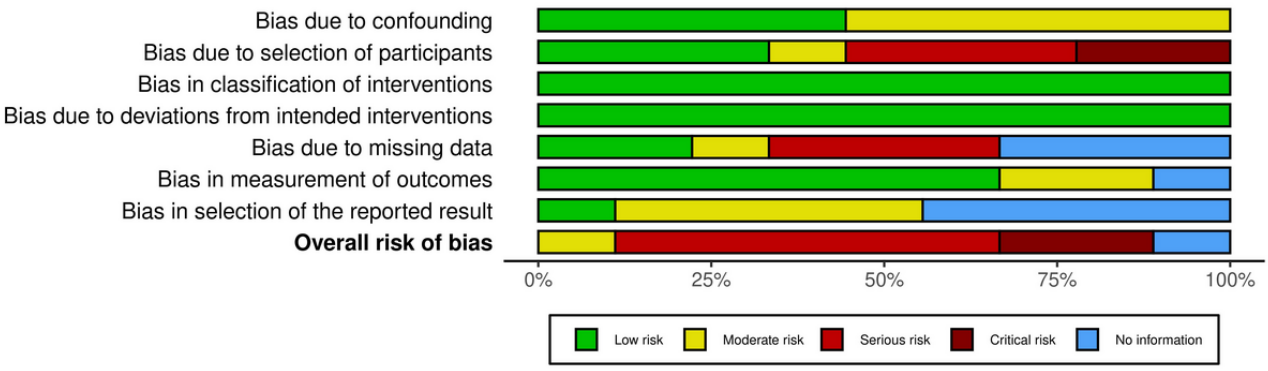


Figure 5. Visualized summary of risk of bias across all included nonrandomized controlled trials, categorized by the following bias domains: bias due to confounding, bias due to selection of participants, bias in classification of interventions, bias due to deviations from intended interventions, bias due to missing data, bias in measurement of outcomes, and bias in selection of the reported result, as well as the overall risk of bias. The proportions of the risk categories are displayed as horizontal bars, providing an overview of the methodological quality of the included studies.



XSL•FO
RenderX

Discussion

Evolving Integrated Intervention Approaches to Support Mental Health and Re-Employment

Unemployment remains a significant societal issue with profound psychosocial consequences that impact both individuals and communities. As individuals experiencing unemployment often face compounded barriers in finding work and accessing support, this population remains severely underserved. To the best of our knowledge, this review is the first to synthesize interventions aimed at improving both mental health outcomes and re-employment rates for individuals facing unemployment and psychological distress. Regarding research question 1 (*What types of interventions have been implemented to promote mental health and facilitate re-employment among unemployed individuals, and what are their characteristics and outcomes?*), we could identify content and methodologically heterogeneous approaches, which are mostly delivered in groups, and the access paths are typically provided through institutions (eg, job centers and employment agencies). It can also be seen that over the period from 1989 to 2018, interventions targeting re-employment and mental health among unemployed individuals have undergone several key evolutionary shifts. Early interventions, such as the JOBS program developed by Caplan et al [46], primarily focused on improving job-seeking skills and self-efficacy, and providing emotional support to individuals facing unemployment. These programs were structured around in-person, face-to-face workshops and sessions, often using methods like problem-solving, decision-making, and inoculation against setbacks. By the late 1990s and 2000s, CBT began to play a larger role in interventions. This shift recognized the importance of addressing mental health alongside job-related skills. More recent interventions (eg, [49,51]) have become more individualized, offering personalized coaching and support tailored to the specific needs of each participant. These programs also include more diverse formats, such as ongoing, long-term support rather than short-term interventions. Moreover, interventions, such as those by Hulshof [50] and Maguire et al [59], emphasize building psychological resources, such as resilience, hope, and optimism, alongside the development of job-related skills. These programs recognize that participants need to not only learn job-seeking strategies but also develop positive psychological attributes that can help them cope with the challenges of re-employment. These trends highlight the growing sophistication of employment interventions, from basic skills training to comprehensive, personalized programs that integrate mental health and psychological resilience.

Regarding research question 2 (*How effective are these interventions in enhancing mental health outcomes and improving re-employment rates?*), the results indicate that work-oriented interventions can have positive effects on mental health outcomes. This is consistent with the findings of Barry et al [58], who identified an enhanced sense of mastery as a possible mechanism of change for financial strain and depression symptoms. In turn, Della-Posta and Drummond [44] found that interventions that focus on mental health, for instance, via CBT-based interventions, can have a beneficial influence on re-employment. Most of the studies employed group settings where individuals shared their experiences with unemployment and the associated distress. Given the aim to provide a supportive environment for learning and developing coping strategies, group- and self-help–oriented approaches hold great potential for this target group. Additionally, this approach can help alleviate feelings of loneliness and social exclusion that often accompany unemployment [72], promoting social inclusion within a peer group [46,49-51,59]. Furthermore, research has shown associations between unemployment and increased mental health costs [73,74]. Therefore, it is important to consider macro-level factors associated with mental disorders, such as the considerable economic burden for respective countries [75]. Additionally, nonmonetary aspects, such as the number of disability-adjusted life years, should not be underestimated as they are among the leading causes of burden worldwide [76].

Based on the findings of this review, that is, the heterogeneity in intervention content among the included studies and their effects on respective outcomes, it is advisable to provide personalized and low-threshold care for individuals experiencing unemployment and psychological distress.

The following content and overarching themes can serve as a guideline for content options for integrative support measures. At the level of psychological and therapeutical support, this might be (1) promoting motivation for change, (2) prioritizing crisis interventions, (3) fostering the identification of dysfunctional thought patterns and learned helplessness and recognizing work-specific anxiety-inducing negative experiences (eg, with colleagues or superiors), (4) giving assistance in maintaining a daily schedule despite periods of unemployment, (5) fostering and reactivating social support using group settings, (6) strengthening psychosocial resources, (7) support in transitioning into the respective care system, and (8) maintaining contact beyond the time of re-employment, if possible. At the job-search level, the increased use of supported employment-oriented approaches [77] represents a possible innovative approach as it considers personal mental health limitations and close cooperation with employers and continues to provide further support beyond the time of re-employment. Moreover, including and reactivating the individual's social network to leverage support in the search for new employment opportunities may serve as a possible approach. This review focuses on interventions that are accessible through job agencies or other publicly accessible pathways, with an emphasis on their preventive nature in breaking the vicious cycle of poor mental health and unemployment. Thus, the consideration of structural levels might also be needed. To enhance collaboration and effectiveness, it is important to maximize synergies with existing support systems and cooperatively integrate them. For instance, job centers can engage in exchange with other supportive stakeholders.

Limitations

This review has some limitations. Although we used the RoB 2 tool according to the recommendations in the current Cochrane Handbook for Systematic Reviews of Interventions [41], we acknowledge that this tool has been reported to have low

XSL•FO
RenderX

interrater reliability [78], and challenges in using the tool consistently may make it difficult to interpret the risk of bias ratings. A limitation of interpretability is the identified heterogeneity of the samples, especially in terms of the length of unemployment (see Multimedia Appendix 4) since longer periods of unemployment can lead to a decline in mental health, with some stabilization occurring at a low level in the second year of unemployment [6].

Thus, the comparability of intervention effects may be restricted due to heterogeneity in other characteristics, such as the composition of the population, differences between countries, and the content of the intervention. It is important to consider these factors as part of the overall analysis. It is also important to consider that this review synthesizes findings from the last 30 years. During this time, the diagnosis and treatment of mental disorders, the general perception of mental health, and economic circumstances, such as economic crises accompanied by short-term contracts and general job insecurity, may have influenced the overall economy and the well-being of the population [79]. It is also important to note that only the re-employment rate was used in the synthesis of the results presented here, and therefore, other criteria used in many studies, such as the quality of a new job, were not included.

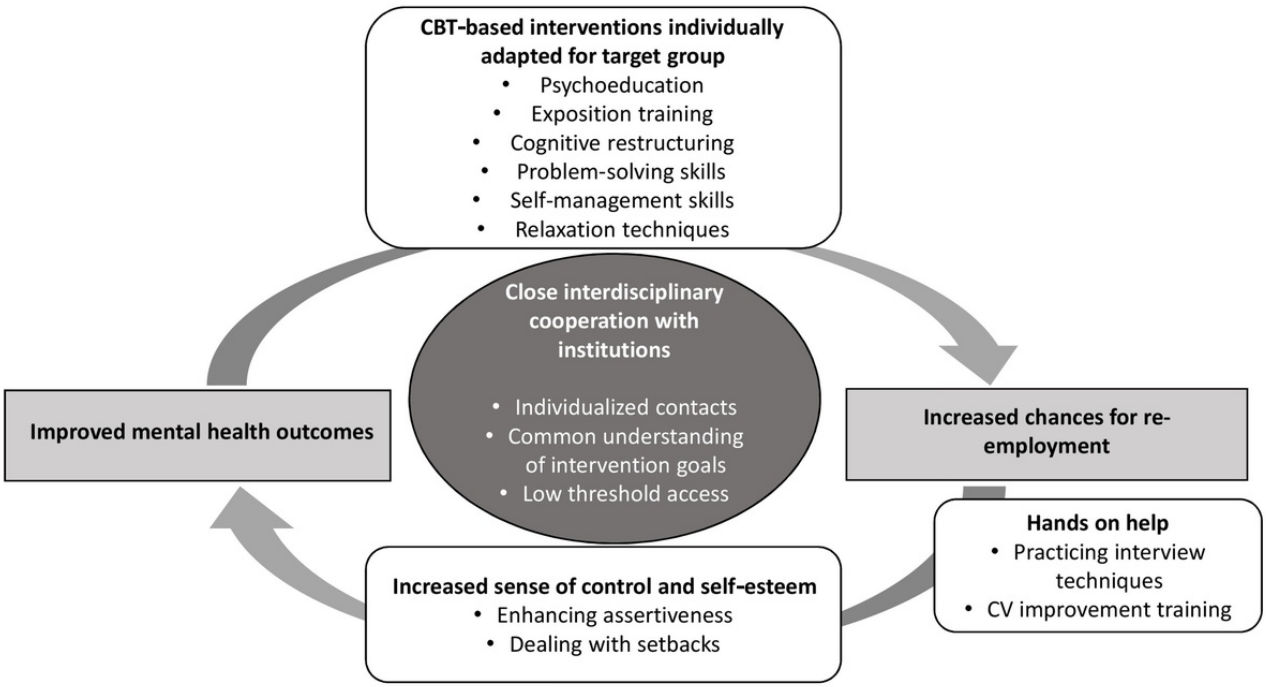
Future Perspectives

Future studies should include longer follow-up intervals to adequately assess the long-term effects of interventions on psychological outcomes and re-employment. Moreover, based on the described problems in the implementation of well-designed interventions, close interdisciplinary cooperation with different support structures and facilitators is needed [80]. It might be assumed that the external circumstances are more difficult and that the studies place an additional burden on the agencies; this should be increasingly addressed in future studies. In addition, future research should address the problems of recruitment and ways to increase participant adherence to interventions.

For an overview of the implications for research and practice, see Figure 6.

The results of the studies, especially regarding re-employment, must be viewed in their sociocultural context and the associated opportunities to find labor. Future studies should also address the cost economics of interventions with novelty value and provide a trade-off between monetary costs and the added value of interventions. Finally, it is important to maintain the methodological quality of interventions, especially in the field of health services research, to ensure their scientific added value [81]. It can also be noted that improvements in aspects can positively transfer to one another. Thus, the respective aims are not strictly separable, and for instance, positive job search experiences may contribute to affective improvement without being explicitly targeted by psychotherapeutic interventions.

Figure 6. Synthesized results of evidence-based implications for research and practice. Dark grey: aspects to be considered in cooperation with executing and cooperating institutions; light grey: target outcomes; white: specific interventions or intervention components that emerge from the systematic review as contributing to improvements. CBT: cognitive behavioral therapy; CV: curriculum vitae.



Acknowledgments

We would like to thank Emilie Bublies for assistance with the qualitative research process. We acknowledge support from the Open Access Publishing Fund of the University of Tübingen. This study was funded by the institution's own financial resources.

XSL•FO
RenderX

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors' Contributions

This review was designed by SHA and RE. FJ, KEG, and SZ provided substantial input to the research question. SHA and RE searched, retrieved, and screened the articles. SHA and MM obtained the quality ratings. SHA and SS created the outcome tables. The first draft of the manuscript was written by SHA with substantial input from RE, and all authors commented on previous versions and the final version of the manuscript. SHA interpreted the results with substantial inputs from MM, SS, NR, KEG, HG, JvW, and RE.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist.
[DOCX File , 30 KB - publichealth_v11i1e65698_app1.docx]

Multimedia Appendix 2
Search terms.
[DOCX File , 17 KB - publichealth_v11i1e65698_app2.docx]

Multimedia Appendix 3
Graphical overview of the results on mental health and re-employment in each included study.
[DOCX File , 19 KB - publichealth_v11i1e65698_app3.docx]

Multimedia Appendix 4
Sample characteristics of the included studies.
[DOCX File , 23 KB - publichealth_v11i1e65698_app4.docx]

Multimedia Appendix 5
Overview of the measurement instruments used.
[DOCX File , 15 KB - publichealth_v11i1e65698_app5.docx]

Multimedia Appendix 6
Detailed overview of all ratings and judgements.
[DOCX File , 39 KB - publichealth_v11i1e65698_app6.docx]

References

1. Modini M, Joyce S, Mykletun A, Christensen H, Bryant R, Mitchell P, et al. The mental health benefits of employment: Results of a systematic meta-review. Australas Psychiatry 2016 Aug;24(4):331-336 [FREE Full text] [doi: 10.1177/1039856215618523] [Medline: 26773063]
2. van der Noordt M, IJzelenberg H, Droomers M, Proper K. Health effects of employment: a systematic review of prospective studies. Occup Environ Med 2014 Oct;71(10):730-736 [FREE Full text] [doi: 10.1136/oemed-2013-101891] [Medline: 24556535]
3. Unemployment statistics. eurostat. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Unemployment_statistics [accessed 2025-04-07]
4. Statistics on unemployment and labour underutilization. International Labour Organization. URL: https://ilostat.ilo.org/topics/unemployment-and-labour-underutilization/ [accessed 2025-04-07]
5. Virgolino A, Costa J, Santos O, Pereira M, Antunes R, Ambrósio S, et al. Lost in transition: a systematic review of the association between unemployment and mental health. J Ment Health 2022 Jun;31(3):432-444 [FREE Full text] [doi: 10.1080/09638237.2021.2022615] [Medline: 34983292]
6. Paul K, Moser K. Unemployment impairs mental health: Meta-analyses. Journal of Vocational Behavior 2009 Jun;74(3):264-282 [FREE Full text] [doi: 10.1016/j.jvb.2009.01.001]
7. Amiri S. Unemployment associated with major depression disorder and depressive symptoms: a systematic review and meta-analysis. Int J Occup Saf Ergon 2022 Dec 05;28(4):2080-2092. [doi: 10.1080/10803548.2021.1954793] [Medline: 34259616]

XSL•FO
RenderX

8. Arias-de la Torre J, Vilagut G, Ronaldson A, Serrano-Blanco A, Martín V, Peters M, et al. Prevalence and variability of current depressive disorder in 27 European countries: a population-based study. The Lancet Public Health 2021 Oct;6(10):e729-e738 [FREE Full text] [doi: 10.1016/s2468-2667(21)00047-5]
9. Lim G, Tam W, Lu Y, Ho C, Zhang M, Ho R. Prevalence of depression in the community from 30 countries between 1994 and 2014. Sci Rep 2018 Feb 12;8(1):2861 [FREE Full text] [doi: 10.1038/s41598-018-21243-x]
10. Rossom R, Solberg L, Vazquez-Benitez G, Whitebird R, Crain A, Beck A, et al. Predictors of poor response to depression treatment in primary care. Psychiatr Serv 2016 Dec 01;67(12):1362-1367 [FREE Full text] [doi: 10.1176/appi.ps.201400285] [Medline: 27417890]
11. Hakulinen C, Elovainio M, Arffman M, Lumme S, Pirkola S, Keskimäki I, et al. Mental disorders and long‐term labour market outcomes: nationwide cohort study of 2 055 720 individuals. Acta Psychiatr Scand 2019 Aug 26;140(4):371-381. [doi: 10.1111/acps.13067]
12. McKee-Ryan F, Song Z, Wanberg C, Kinicki A. Psychological and physical well-being during unemployment: a meta-analytic study. J Appl Psychol 2005 Jan;90(1):53-76 [FREE Full text] [doi: 10.1037/0021-9010.90.1.53] [Medline: 15641890]
13. Hoffmann H, Jäckel D, Glauser S, Mueser KT, Kupper Z. Long-term effectiveness of supported employment: 5-year follow-up of a randomized controlled trial. Am J Psychiatry 2014 Nov 01;171(11):1183-1190. [doi: 10.1176/appi.ajp.2014.13070857] [Medline: 25124692]
14. Frederick D, VanderWeele T. Supported employment: Meta-analysis and review of randomized controlled trials of individual placement and support. PLoS One 2019;14(2):e0212208 [FREE Full text] [doi: 10.1371/journal.pone.0212208] [Medline: 30785954]
15. Jenny GJ, Bauer GF, Vinje HF, Vogt K, Torp S. The Application of Salutogenesis to Work. In: Mittelmark M, editor. The Handbook of Salutogenesis. Cham: Springer; 2017:197-210.
16. Oschimansky F, Popp S, Riedel-Heller S, Schwarzbach M, Gühne U, Kupka P. Psychisch Kranke im SGB II: Situation und Betreuung. IAB-Forschungsbericht. URL: https://iab.de/publikationen/publikation/?id=7561020 [accessed 2025-04-07]
17. Luciano A, Metcalfe JD, Bond GR, Xie H, Miller AL, Riley J, et al. Hospitalization risk before and after employment among adults with schizophrenia, bipolar disorder, or major depression. Psychiatr Serv 2016 Oct 01;67(10):1131-1138. [doi: 10.1176/appi.ps.201500343] [Medline: 27247173]
18. Skärlund M, Åhs A, Westerling R. Health-related and social factors predicting non-reemployment amongst newly unemployed. BMC Public Health 2012 Oct 23;12:893-811 [FREE Full text] [doi: 10.1186/1471-2458-12-893] [Medline: 23092291]
19. Svane-Petersen A, Dencker-Larsen S. The impact of self-reported health and register-based prescription medicine purchases on re-employment chances: A prospective study. SSM Popul Health 2016 Dec;2:580-586 [FREE Full text] [doi: 10.1016/j.ssmph.2016.08.007] [Medline: 29349173]
20. Hakulinen C, Elovainio M, Arffman M, Lumme S, Suokas K, Pirkola S, et al. Employment status and personal income before and after onset of a severe mental disorder: a case-control study. Psychiatr Serv 2020 Mar 01;71(3):250-255 [FREE Full text] [doi: 10.1176/appi.ps.201900239] [Medline: 31722646]
21. Østerud KL. Mental illness stigma and employer evaluation in hiring: Stereotypes, discrimination and the role of experience. Sociol Health Illn 2023 Jan;45(1):90-108 [FREE Full text] [doi: 10.1111/1467-9566.13544] [Medline: 36103320]
22. Hagen C, Bänfer S, Werkstetter L, Hebebrand J, Reissner V. [Use of mental health service among young adults on unemployment benefit before and after receiving counseling at a psychiatric liaison department]. Gesundheitswesen 2018 Jul;80(7):598-605 [FREE Full text] [doi: 10.1055/s-0042-116320] [Medline: 27975338]
23. Brouwer B, Waardenburg S, Jacobs C, Overdijk M, Leue C, Köke A, et al. Biopsychosocial baseline values of 15 000 patients suffering from chronic pain: Dutch DataPain study. Reg Anesth Pain Med 2020 Oct;45(10):774-782. [doi: 10.1136/rapm-2020-101476] [Medline: 32784227]
24. Rüsch N, Corrigan P, Waldmann T, Staiger T, Bahemann A, Oexle N, et al. Attitudes toward disclosing a mental health problem and reemployment. J Nerv Ment Dis 2018;206(5):383-385 [FREE Full text] [doi: 10.1097/nmd.0000000000000810]
25. Brouwers EPM. Social stigma is an underestimated contributing factor to unemployment in people with mental illness or mental health issues: position paper and future directions. BMC Psychol 2020 Apr 21;8(1):36 [FREE Full text] [doi: 10.1186/s40359-020-00399-0] [Medline: 32317023]
26. Santos O, Lopes E, Virgolino A, Stefanovska-Petkovska M, Dinis A, Ambrósio S, et al. Defining a brief intervention for the promotion of psychological well-being among unemployed individuals through expert consensus. Front Psychiatry 2018;9:13 [FREE Full text] [doi: 10.3389/fpsyt.2018.00013] [Medline: 29467680]
27. Audhoe SS, Nieuwenhuijsen K, Hoving JL, Sluiter JK, Frings-Dresen MHW. Perspectives of unemployed workers with mental health problems: barriers to and solutions for return to work. Disabil Rehabil 2018 Jan;40(1):28-34 [FREE Full text] [doi: 10.1080/09638288.2016.1242170] [Medline: 27756177]
28. Hollederer A. Health promotion and prevention among the unemployed: a systematic review. Health Promot Int 2019 Dec 01;34(6):1078-1096 [FREE Full text] [doi: 10.1093/heapro/day069] [Medline: 30252059]
29. Hult M, Lappalainen K, Saaranen TK, Räsänen K, Vanroelen C, Burdorf A. Health-improving interventions for obtaining employment in unemployed job seekers. Cochrane Database Syst Rev 2020 Jan 08;1(1):CD013152 [FREE Full text] [doi: 10.1002/14651858.CD013152.pub2] [Medline: 31912888]

XSL•FO
RenderX

30. Koopman M, Pieterse M, Bohlmeijer E, Drossaert C. Mental health promoting Interventions for the unemployed: a systematic review of applied techniques and effectiveness. International Journal of Mental Health Promotion 2017 Jun 05;19(4):202-223 [FREE Full text] [doi: 10.1080/14623730.2017.1328367]
31. Arena AF, Mobbs S, Sanatkar S, Williams D, Collins D, Harris M, et al. Mental health and unemployment: A systematic review and meta-analysis of interventions to improve depression and anxiety outcomes. J Affect Disord 2023 Aug 15;335:450-472. [doi: 10.1016/j.jad.2023.05.027] [Medline: 37201898]
32. Moore THM, Kapur N, Hawton K, Richards A, Metcalfe C, Gunnell D. Interventions to reduce the impact of unemployment and economic hardship on mental health in the general population: a systematic review. Psychol Med 2017 Apr;47(6):1062-1084 [FREE Full text] [doi: 10.1017/S0033291716002944] [Medline: 27974062]
33. van Rijn RM, Carlier BE, Schuring M, Burdorf A. Work as treatment? The effectiveness of re-employment programmes for unemployed persons with severe mental health problems on health and quality of life: a systematic review and meta-analysis. Occup Environ Med 2016 Apr;73(4):275-279. [doi: 10.1136/oemed-2015-103121] [Medline: 26740687]
34. Audhoe SS, Hoving JL, Sluiter JK, Frings-Dresen MHW. Vocational interventions for unemployed: effects on work participation and mental distress. A systematic review. J Occup Rehabil 2010 Mar;20(1):1-13. [doi: 10.1007/s10926-009-9223-y] [Medline: 20039106]
35. EndNote 20. Clarivate. URL: https://support.clarivate.com/Endnote/s/article/EndNote-20-Release-Notes?language=en_US [accessed 2025-04-07]
36. McGuinness L, Higgins J. Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. Res Synth Methods 2021 Jan;12(1):55-61 [FREE Full text] [doi: 10.1002/jrsm.1411] [Medline: 32336025]
37. Moher D, Liberati A, Tetzlaff J, Altman D, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341 [FREE Full text] [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]
38. Page MJ, Moher D, McKenzie JE. Introduction to preferred reporting items for systematic reviews and meta-analyses 2020 and implications for research synthesis methodologists. Res Synth Methods 2022 Mar 02;13(2):156-163. [doi: 10.1002/jrsm.1535] [Medline: 34693644]
39. Eriksen M, Frandsen T. The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. J Med Libr Assoc 2018 Oct;106(4):420-431 [FREE Full text] [doi: 10.5195/jmla.2018.345] [Medline: 30271283]
40. Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. Cochrane Database Syst Rev 2011 Aug 10;2011(8):MR000026 [FREE Full text] [doi: 10.1002/14651858.MR000026.pub2] [Medline: 21833989]
41. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. Cochrane Training. URL: https://training.cochrane.org/handbook/current/chapter-08 [accessed 2025-04-07]
42. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919 [FREE Full text] [doi: 10.1136/bmj.i4919] [Medline: 27733354]
43. Boutron I, Guittet L, Estellat C, Moher D, Hróbjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. PLoS Med 2007 Feb;4(2):e61 [FREE Full text] [doi: 10.1371/journal.pmed.0040061] [Medline: 17311468]
44. Della-Posta C, Drummond PD. Cognitive behavioural therapy increases re-employment of job seeking worker's compensation clients. J Occup Rehabil 2006 Jun;16(2):223-230 [FREE Full text] [doi: 10.1007/s10926-006-9024-5] [Medline: 16705491]
45. Vinokur AD, Price RH, Schul Y. Impact of the JOBS intervention on unemployed workers varying in risk for depression. Am J Community Psychol 1995 Feb;23(1):39-74 [FREE Full text] [doi: 10.1007/BF02506922] [Medline: 7572826]
46. Caplan RD, Vinokur AD, Price RH, van Ryn M. Job seeking, reemployment, and mental health: a randomized field experiment in coping with job loss. J Appl Psychol 1989 Oct;74(5):759-769. [doi: 10.1037/0021-9010.74.5.759] [Medline: 2793774]
47. Himle JA, Bybee D, Steinberger E, Laviolette WT, Weaver A, Vlnka S, et al. Work-related CBT versus vocational services as usual for unemployed persons with social anxiety disorder: A randomized controlled pilot trial. Behav Res Ther 2014 Dec;63:169-176 [FREE Full text] [doi: 10.1016/j.brat.2014.10.005] [Medline: 25461793]
48. Vinokur AD, Schul Y, Vuori J, Price RH. Two years after a job loss: long-term impact of the JOBS program on reemployment and mental health. J Occup Health Psychol 2000 Jan;5(1):32-47. [doi: 10.1037//1076-8998.5.1.32] [Medline: 10658883]
49. Carlier BE, Schuring M, Burdorf A. Influence of an interdisciplinary re-employment programme among unemployed persons with mental health problems on health, social participation and paid employment. J Occup Rehabil 2018 Mar;28(1):147-157 [FREE Full text] [doi: 10.1007/s10926-017-9704-3] [Medline: 28397017]
50. Hulshof IL. Let's go to work : the role of (job) crafting in the process of finding a job. Technische Universiteit Eindhoven. 2018. URL: https://research.tue.nl/en/publications/lets-go-to-work-the-role-of-job-crafting-in-the-process-of-findin [accessed 2025-04-07]
51. Herbig B, Glaser J, Angerer P. [Old, sick, unemployed, without a chance? Results of a randomised controlled trial of the effects of a combined health and employment promotion program for the older long-term unemployed (AmigA-M)].

XSL•FO
RenderX

Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2012 Aug;55(8):970-979. [doi: 10.1007/s00103-012-1514-3] [Medline: 22842891]
52. Rothländer K, Mühlpfordt S, Richter P. Evaluation des Gesundheitsförderungsprogramms „Aktive Bewältigung von Arbeitslosigkeit (AktivA)". Zeitschrift für Gesundheitspsychologie 2012 Jul 15;20(3):115-127. [doi: 10.1026/0943-8149/a000070]
53. Harris E, Lum J, Rose V, Morrow M, Comino E, Harris M. Are CBT interventions effective with disadvantaged job-seekers who are long-term unemployed? Psychology, Health & Medicine 2002 Nov;7(4):401-410. [doi: 10.1080/1354850021000015221]
54. Rose V. Improving the Health of People who are Unemployed through the Job Network. In: A Study of a Brief CBT Intervention in South Western Sydney. Sydney, Australia: Centre for Health Equity Training Research and Evaluation; 2001:1-62.
55. Vuori J, Silvonen J. The benefits of a preventive job search program on re‐employment and mental health at 2‐year follow‐up. J Occupat & Organ Psyc 2005;78(1):43-52 [FREE Full text] [doi: 10.1348/096317904x23790]
56. Vuori J, Silvonen J, Vinokur AD, Price RH. The Työhön Job Search Program in Finland: benefits for the unemployed with risk of depression or discouragement. J Occup Health Psychol 2002 Jan;7(1):5-19. [doi: 10.1037//1076-8998.7.1.5] [Medline: 11827233]
57. Proudfoot J, Guest D, Carson J, Dunn G, Gray J. Effect of cognitive-behavioural training on job-finding among long-term unemployed people. Lancet 1997 Jul 12;350(9071):96-100. [doi: 10.1016/S0140-6736(96)09097-6] [Medline: 9228961]
58. Barry M, Reynolds C, Sheridan A, Egenton R. Implementation of the JOBS programme in Ireland. Journal of Public Mental Health 2006;5(4):10-25 [FREE Full text] [doi: 10.1108/17465729200600028]
59. Maguire N, Hughes VC, Bell L, Bogosian A, Hepworth C. An evaluation of the choices for well-being project. Psychol Health Med 2014;19(3):303-315. [doi: 10.1080/13548506.2013.806813] [Medline: 23822617]
60. Reynolds C, Barry MM, Nic Gabhainn S. Evaluating the impact of the winning new jobs programme on the re-employment and mental health of a mixed profile of unemployed people. International Journal of Mental Health Promotion 2012 Jan 30;12(2):32-41. [doi: 10.1080/14623730.2010.9721812]
61. Vinokur AD, Schul Y. Mastery and inoculation against setbacks as active ingredients in the JOBS intervention for the unemployed. J Consult Clin Psychol 1997 Oct;65(5):867-877. [doi: 10.1037//0022-006x.65.5.867] [Medline: 9337505]
62. Heimberg RG, Becker RE. Cognitive-Behavioral Group Therapy for Social Phobia: Basic Mechanisms and Clinical Strategies. New York, NY: Guilford Press; 2002.
63. Hulshof IL, Demerouti E, Le Blanc PM. Providing services during times of change: can employees maintain their levels of empowerment, work engagement and service quality through a job crafting intervention? Front Psychol 2020;11:87 [FREE Full text] [doi: 10.3389/fpsyg.2020.00087] [Medline: 32047468]
64. Demerouti E, Bakker A, Nachreiner F, Schaufeli W. The job demands-resources model of burnout. Journal of Applied Psychology 2001;86(3):499-512 [FREE Full text] [doi: 10.1037//0021-9010.86.3.499]
65. Luthans F, Avey J, Avolio B, Norman S, Combs G. Psychological capital development: toward a micro‐intervention. J Organ Behavior 2006 Mar 27;27(3):387-393 [FREE Full text] [doi: 10.1002/job.373]
66. Kolb AY, Kolb DA. Experiential Learning Theory. In: Seel NM, editor. Encyclopedia of the Sciences of Learning. Boston, MA: Springer; 2012:1215-1219.
67. Locke E, Latham G. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717 [FREE Full text] [doi: 10.1037//0003-066x.57.9.705] [Medline: 12237980]
68. Locke E, Latham G. New directions in goal-setting theory. Curr Dir Psychol Sci 2006 Oct 01;15(5):265-268 [FREE Full text] [doi: 10.1111/j.1467-8721.2006.00449.x]
69. Schwarzer R. Self-regulatory processes in the adoption and maintenance of health behaviors. J Health Psychol 1999 Mar 01;4(2):115-127. [doi: 10.1177/135910539900400208] [Medline: 22021474]
70. Vinokur AD, van Ryn M, Gramlich EM, Price RH. Long-term follow-up and benefit-cost analysis of the Jobs Program: a preventive intervention for the unemployed. J Appl Psychol 1991 Apr;76(2):213-219. [doi: 10.1037/0021-9010.76.2.213] [Medline: 1905293]
71. Vuori J, Price RH, Mutanen P, Malmberg-Heimonen I. Effective group training techniques in job-search training. J Occup Health Psychol 2005 Jul;10(3):261-275. [doi: 10.1037/1076-8998.10.3.261] [Medline: 16060729]
72. Pohlan L. Unemployment and social exclusion. Journal of Economic Behavior & Organization 2019 Aug;164:273-299 [FREE Full text] [doi: 10.1016/j.jebo.2019.06.006]
73. Ondler C, Hegde G, Carlson J. Resource utilization and health care charges associated with the most frequent ED users. Am J Emerg Med 2014 Oct;32(10):1215-1219 [FREE Full text] [doi: 10.1016/j.ajem.2014.07.013] [Medline: 25154348]
74. Waldmann T, Staiger T, Ruesch N, Kilian R. Costs of health service use among unemployed and underemployed people with mental health problems. J Ment Health Policy Econ 2021 Mar 01;24(1):31-41. [Medline: 33739934]
75. Christensen MK, Lim CCW, Saha S, Plana-Ripoll O, Cannon D, Presley F, et al. The cost of mental disorders: a systematic review. Epidemiol Psychiatr Sci 2020 Aug 18;29:e161 [FREE Full text] [doi: 10.1017/s204579602000075x]

XSL•FO
RenderX

76. Collaborators G. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet Psychiatry 2022 Feb;9(2):137-150 [FREE Full text] [doi: 10.1016/s2215-0366(21)00395-3]
77. Bond G, Drake R, Becker D. An update on individual placement and support. World Psychiatry 2020 Oct;19(3):390-391 [FREE Full text] [doi: 10.1002/wps.20784] [Medline: 32931093]
78. Minozzi S, Cinquini M, Gianola S, Gonzalez-Lorenzo M, Banzi R. The revised Cochrane risk of bias tool for randomized trials (RoB 2) showed low interrater reliability and challenges in its application. J Clin Epidemiol 2020 Oct;126:37-44 [FREE Full text] [doi: 10.1016/j.jclinepi.2020.06.015] [Medline: 32562833]
79. Asper M, Osika W, Dalman C, Pöllänen E, Simonsson O, Flodin P, et al. Effects of the COVID-19 pandemic and previous pandemics, epidemics and economic crises on mental health: systematic review. BJPsych Open 2022 Oct 10;8(6):e181 [FREE Full text] [doi: 10.1192/bjo.2022.587] [Medline: 36214114]
80. Rüsch N, Staiger T, Waldmann T, Dekoj MC, Brosch T, Gabriel L, et al. Efficacy of a peer-led group program for unemployed people with mental health problems: Pilot randomized controlled trial. Int J Soc Psychiatry 2019 Jun;65(4):333-337. [doi: 10.1177/0020764019846171] [Medline: 31056997]
81. Steeger C, Buckley P, Pampel F, Gust C, Hill K. Common methodological problems in randomized controlled trials of preventive interventions. Prev Sci 2021 Nov;22(8):1159-1172 [FREE Full text] [doi: 10.1007/s11121-021-01263-2] [Medline: 34176002]

Abbreviations

CBT: cognitive behavioral therapy
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RoB 2: Risk-of-Bias Tool for Randomized Trials
ROBINS-I: Risk of Bias in Nonrandomized Studies - of Interventions

Edited by A Mavragani; submitted 22.08.24; peer-reviewed by PM Matos Fialho, CY Chen; comments to author 17.01.25; revised version received 06.02.25; accepted 13.02.25; published 05.05.25.

Please cite as:
Adam SH, Junne F, Schlachter S, Mehler M, Gündel H, Rüsch N, von Wietersheim J, Giel KE, Zipfel S, Erschens R
Interventions to Foster Mental Health and Reintegration in Individuals Who Are Unemployed: Systematic Review
JMIR Public Health Surveill 2025;11:e65698
URL: https://publichealth.jmir.org/2025/1/e65698
doi:10.2196/65698
PMID:

©Sophia Helen Adam, Florian Junne, Svenja Schlachter, Miriam Mehler, Harald Gündel, Nicolas Rüsch, Jörn von Wietersheim, Katrin Elisabeth Giel, Stephan Zipfel, Rebecca Erschens. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 05.05.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Differences in Hemodialysis Claim Patterns Across Membership Types Among Patients With Renal Failure Based on National Health Insurance Data From 2017 to 2022: Cross-Sectional Analysis

Aries Munandar[1]; Syarif R Hasibuan[2,3]; Dian Kusuma[4]

[1]Department of Social Welfare, Faculty of Social and Political Science, University of Bengkulu, Bengkulu, Indonesia

[2]Faculty of Medicine, Universitas Pembangunan Nasional Veteran Jakarta, Jakarta, Indonesia

[3]Center for Health Administration and Policy Studies, Faculty of Public Health, University of Indonesia, Depok, Indonesia

[4]Department of Public Health and Epidemiology, College of Medicine and Health Sciences, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates

Corresponding Author:
Dian Kusuma
Department of Public Health and Epidemiology
College of Medicine and Health Sciences
Khalifa University of Science and Technology
Shakhbout Bin Sultan St
Abu Dhabi
United Arab Emirates
Phone: 971 02 312 3333
Email: dian.kusuma@ku.ac.ae

Abstract

Background: Chronic kidney disease and end-stage renal disease are major contributors to the disease burden in low- and middle-income countries, including Indonesia. Despite the expansion of universal health coverage through Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan, Indonesia's national health insurance program, disparities in access to hemodialysis persist across different socioeconomic and geographic groups. Understanding these inequities is critical to advancing equitable health care access.

Objective: This study aimed to examine disparities in hemodialysis claim patterns as a proxy for access among adult patients with renal failure enrolled in BPJS, focusing on differences by membership type, sex, age, geographic region, urbanicity, and facility ownership.

Methods: We conducted a cross-sectional analysis of 38,383 anonymized health insurance claims between 2017 and 2022 for patients with renal failure who were aged ≥18 years. The primary outcome was receipt of hemodialysis. We used multivariate logistic regression to estimate adjusted odds ratios (aORs) for receiving hemodialysis across BPJS membership types and other covariates. Subgroup analyses were performed by sex, facility ownership, urbanicity, and geographic region. Robust SEs and probability weights were applied to account for the sample design.

Results: Of the total renal failure claims, 75.6% (29,017/38,383) involved hemodialysis. Compared with individuals in the lowest income group (ie, members subsidized under the national government budget), informal workers (aOR 1.56, 95% CI: 1.34-1.82); *P*<.001) and members subsidized under the local government budget (aOR 1.31, 95% CI: 1.05-1.63); *P*=.017) had higher odds of receiving hemodialysis, while formal sector workers had lower odds (aOR 0.81, 95% CI: 0.68-0.98); *P*=.028). Disparities were more pronounced in rural areas and among women; for example, in rural regions, locally subsidized members had more than twice the odds of receiving hemodialysis compared with nationally subsidized members (aOR 2.40, 95% CI: 1.78-3.23). Men had higher odds than women (aOR 1.17, 95% CI: 1.04-1.32), and younger patients were more likely to receive treatment than older ones. Regional disparities were stark, with patients in Java or Bali having much greater access (aOR 8.30, 95% CI 5.33-12.94) compared with those in eastern Indonesia (Papua, Maluku, and Nusa Tenggara). Patients treated at private facilities (aOR 1.30, 95% CI 1.13-1.50) and in outpatient settings (aOR 3.74, 95% CI 3.36-4.17) were more likely to receive hemodialysis, whereas those in lower-level hospitals or clinics were less likely to access care.

XSL•FO
RenderX

Conclusions: Substantial disparities in hemodialysis claim patterns (as a proxy for access) exist within Indonesia's national health insurance system, particularly affecting low-income populations, rural residents, women, and those in less advantaged regions. Policy efforts to enhance health infrastructure, improve service distribution, and reduce geographic and socioeconomic barriers are urgently needed to support equitable access to renal care services and achieve universal health coverage goals.

(JMIR Public Health Surveill 2025;11:e73731) doi:10.2196/73731

KEYWORDS

hemodialysis; health equity; kidney failure; BPJS Kesehatan; Indonesia; national health insurance; claims data

Introduction

Chronic kidney disease (CKD) and end-stage renal disease (ESRD) are growing public health challenges globally, with over 800 million individuals affected and 1.5 million deaths reported in 2021, making CKD the 11th leading cause of mortality worldwide [1,2]. ESRD, the most severe form of CKD, requires renal replacement therapy in the form of hemodialysis, peritoneal dialysis, or kidney transplantation to sustain life. In low- and middle-income countries (LMICs) like Indonesia, the burden of CKD is accelerating due to aging populations and rising prevalence of 2 major risk factors—diabetes and hypertension [3]. In 2021, CKD caused more than 51,000 deaths and nearly 2 million disability-adjusted life years in Indonesia alone, underscoring its substantial contribution to the national disease burden [4].

Renal replacement therapy, particularly hemodialysis, has become a crucial component of Indonesia's response to the CKD burden. While peritoneal dialysis and transplantation are also clinically viable, hemodialysis remains the dominant modality in Indonesia due to limited infrastructure, specialist training, and system familiarity with alternative options [5,6]. In financing this care, Indonesia's national health insurance program, Jaminan Kesehatan Nasional, launched in 2014 to advance universal health coverage, plays a central role. Administered by the national health insurance agency (in Indonesian: the Badan Penyelenggara Jaminan Sosial [BPJS]), the program operates as the sole public payer, overseeing reimbursements for hospital claims and managing health care financing nationwide. As of May 2025, BPJS covered over 280 million members, making it one of the largest single-payer health insurance schemes globally. BPJS pools funds from government subsidies, payroll contributions, and self-paid members to deliver near-universal health coverage through both public and private providers [7,8]. In 2023, alone, it financed nearly 1.5 million renal failure claims, with hemodialysis treatment costs reaching Rp 2.91 trillion (approximately US $180 million) [9].

BPJS members fall into 2 broad categories: subsidized and nonsubsidized. Subsidized members include individuals with low and near-low income, whose premiums are paid by the national government or by local governments. Nonsubsidized members pay their own premiums or have them covered by employers and include informal workers, formal sector employees, and others such as pensioners and veterans [10,11]. Despite expanded coverage, previous studies suggest that subsidized members underuse hospital-based services compared to wealthier, nonsubsidized groups [7,8]. This raises questions about whether universal health coverage alone is sufficient to ensure equitable access to high-cost, life-sustaining therapies like dialysis [12].

Globally, LMICs have struggled to achieve equitable access to dialysis due to geographic, financial, and systemic constraints [13]. In countries like India, Nigeria, and Kenya, dialysis access is shaped by regional disparities in facility availability, infrastructure, and human resources [14]. In Indonesia, patients in eastern provinces and rural areas often face long travel distances, limited provider capacity, and inconsistent referral pathways [6,15,16]. Disparities may also arise from Indonesia's dual public-private provider system, where private facilities—often concentrated in urban areas—may offer greater availability or faster access to hemodialysis services [12].

Existing research on ESRD inequalities has largely come from high-income countries such as South Korea and Canada, where administrative claims data are more comprehensive and health system structures differ significantly from those in LMICs [17,18]. Studies in LMICs often rely on alternative data sources such as national health surveys, which, while informative, are susceptible to issues like recall bias [19-21]. While some studies in Indonesia have used national claims data to examine disparities in health care use—such as for cardiovascular diseases, diabetes management, HIV or AIDS, and catastrophic health expenditures—there has been little focus on disparities in access to dialysis services [7,8,10,22]. These gaps in the literature limit our understanding of how systemic and socioeconomic factors shape access to life-saving treatments like dialysis in LMIC settings.

To address this critical evidence gap, we analyze national health insurance claims from 2017 to 2022 to examine disparities in hemodialysis claim patterns as a proxy for access among adult patients with renal failure in Indonesia. This study investigates differences across BPJS membership types, sex, age, urban versus rural residence, geographic region, and health facility characteristics. In the context of near-universal coverage under BPJS, disparities in claims can reasonably be interpreted as reflecting disparities in access, while acknowledging that claims represent episodes of care rather than individual patient trajectories [11,23]. We hypothesize that subsidized members, rural residents, and patients in less advantaged regions (eg, eastern Indonesia) face significantly lower odds of receiving hemodialysis compared with their nonsubsidized, urban, and Java- or Bali-based counterparts. Understanding these disparities is essential to inform evidence-based policy reforms that promote equitable access to essential health services and strengthen Indonesia's progress toward universal health coverage.

XSL•FO
RenderX

Methods

Study Design

We conducted a cross-sectional observational analysis using administrative claims data for adult patients (aged ≥18 years) diagnosed with renal failure between January 1, 2017, and December 31, 2022. Renal failure cases were identified based on *International Classification of Diseases-10 (ICD-10)* codes N17 to N19, which capture acute kidney injury, CKD, and unspecified kidney failure diagnoses [24].

Data

This study used deidentified administrative claims data from Indonesia's national health insurance agency (BPJS), drawn from a standardized 1% random sample of insured members. The dataset includes inpatient and outpatient claims from secondary and tertiary referral health facilities between 2017 and 2022. BPJS insures over 248 million individuals; in 2022 alone, the 1% sample represented approximately 2.4 million individuals selected through simple random sampling at the household level. These members generated more than 3.3 million primary care visits and 1.2 million inpatient claims [7,8].

For this study, we extracted noncapitation claims and identified renal failure cases using *ICD-10* codes N17 to N19. After excluding individuals aged <18 years and claims without relevant diagnoses, the final analytic sample consisted of 38,383 claims, representing 7916 unique patients. A detailed flowchart outlining the sample selection process is provided in Multimedia Appendix 1.

Patient counts varied by year, from 974 in 2017 to 1589 in 2022 (Multimedia Appendix 2). However, only a small proportion of patients had repeated claims across years, for example, just 30.7% of patients from 2017 appeared again in 2018—limiting the potential for longitudinal (panel) analysis. Given this limited overlap, we adopted a cross-sectional, claim-level analytic approach, which allowed us to preserve the clinical detail of each hospitalization episode and to reflect changes in BPJS membership status within the year and over time.

Although the dataset was not stratified by diagnosis or geography, sampling probability weights were applied in all descriptive and multivariate analyses to adjust for unequal selection probabilities and enhance national representativeness. This sampling approach is aligned with international standards, such as the National Sample Cohort from South Korea's National Health Insurance Service, and enables robust estimation of disparities in health care access and outcomes within Indonesia's national insurance system [10,11].

Variables

The primary outcome variable in this study was receipt of hemodialysis, defined as a binary variable (1=received hemodialysis; 0=did not receive hemodialysis). Procedure information in the dataset is coded using *ICD-9-CM* codes, including hemodialysis (n=3995), peritoneal dialysis (n=5498), and kidney transplantation (n=5569). In the final analytic sample, 75.6% (29,017/38,383) of claims between 2017 and 2022 involved hemodialysis. In contrast, claims involving peritoneal dialysis or transplantation accounted for only a very small proportion of claims.

The key independent variable in this study was BPJS membership type, categorized into 5 official groups. The first group, nationally subsidized members, represents individuals in the lowest income group as defined by the national standard through Indonesia's Unified Database for Social Welfare (DTKS), whose premiums are fully subsidized by the central government. The second group, locally subsidized members, includes members with near-low income subsidized by local government budgets and identified according to local eligibility criteria. Nonsubsidized members comprised 3 categories: informal nonworkers, which includes employers, investors, pensioners, and veterans; informal workers, encompassing informal workers who make direct contributions to BPJS; and formal sector employees, representing formal workers whose contributions are jointly funded by employees and employers. This classification is consistent with approaches used in recently published studies [7,11,23].

Additional covariates in this study included the year of the claim, sex, age (calculated by subtracting the date of birth from the date of visit), type of residence (urban for city areas and rural for regency or district), and geographic region, categorized into Java or Bali, Sumatra, Sulawesi, Kalimantan, and the eastern region (including Papua, Maluku, and Nusa Tenggara). Facility-level characteristics were also considered, encompassing service type (inpatient or outpatient), facility ownership (government or private), facility type (hospital or specialist clinic), and facility level (classified as hospital levels A, B, C, D, or other). Clinics are referral-level health care facilities staffed by specialists and equipped to provide hemodialysis treatment. These hospital classifications reflect the service capacity and role within the health system: class A hospitals are national referral centers with the highest service complexity, advanced technologies, and full subspecialty care; class B hospitals serve as referral centers at the provincial or large district level; class C hospitals offer general inpatient services with more basic specialty care; and class D hospitals provide limited inpatient services. The "other" category includes specialized facilities such as military hospitals and hospitals focused on specific services (eg, dialysis centers, surgical centers, and pulmonary hospitals).

Data Analysis

Multivariate logistic regression analyses were conducted to examine associations, with results expressed as adjusted odds ratios (aORs). The primary findings were derived from these multivariate models. Analyses were performed for the full sample and stratified by subgroups, including sex, facility ownership, urbanicity, and geographic region (Java or Bali and non-Java or Bali). To account for the complex sampling design, robust standard errors with probability weights were applied using the *pweight* function of Stata (version 15.1; StataCorp LLC), allowing for population-level inferences. All statistical analyses were conducted in Stata with statistical significance defined as a *P* value of .05 or lower.

XSL•FO
RenderX

Ethical Considerations

This study involved a secondary analysis of deidentified, publicly available administrative health insurance claims data from Indonesia's national health insurance system (BPJS). The dataset contains no personal identifiers and is not linked to individual identities, ensuring the privacy and confidentiality of all individuals represented. The study protocol was reviewed and approved by the Research and Community Engagement Ethical Committee at the Faculty of Public Health, Universitas Indonesia (573/UN2.F10.D11/PPM.00.02/2024). Informed consent is not applicable because this study used secondary data that were deidentified and publicly available. In terms of privacy and confidentiality, all data were fully anonymized, and no individual-level identifiers were accessible to the research team.

Results

Renal Failure Claims by BPJS Membership Type and Other Characteristics

Table 1 highlights the distribution of national health insurance claims for BPJS members in Indonesia diagnosed with renal failure between 2017 and 2022. Among the 38,383 claims recorded for renal failure, 29,017 (75.6%) claims were attributed to hemodialysis treatment. When categorized by membership type, informal workers and formal workers accounted for most total claims, with 19,064 (49.7%) claims and 9180 (23.9%) claims, respectively. A similar pattern emerged for hemodialysis claims, where informal workers and formal sector employees represented 50.5% (14,640/29,017) and 24.3% (7,041/29,017) of claims, respectively. In contrast, nationally subsidized members and locally subsidized members contributed to a smaller proportion, with 3284 (11.3%) and 1988 (6.9%) claims, respectively. Informal nonworkers accounted for 2064 (7.1%) of the claims involving hemodialysis.

Demographic and regional variables offer further insights into the distribution of claims. In terms of sex, men accounted for a greater proportion of total renal failure claims than women, with 57.1% (21,908/38,383) compared with 42.9% (16,475/38,383). Age-wise, claims for renal failure increased with age, ranging from 18.1% (6,933/38,383) among individuals aged 18 to 39 years to 30.1% (11,548/38,383) among those aged ≥60 years. Regarding urbanicity, urban residents had a marginally higher share of claims (20,188/38,383, 52.6%) than rural residents (18,195/38,383, 47.4%). Regionally, most claims (24,410/38,383, 63.6%) originated from the densely populated and more advantaged regions of Java and Bali, followed by 26.3% (10,089/38,383) from Sumatra. In contrast, Sulawesi accounted for only 2.7% (1,027/38,383) of claims, while Papua, Maluku, and Nusa Tenggara contributed just 1.8% (673/38,383).

Facility-level characteristics also shed light on the distribution of renal failure claims. By service type, outpatient services accounted for most claims (28,181/38,383, 73.4%), while inpatient services comprised 26.6% (10,202/38,383). Regarding facility type, the vast majority of claims (34,925/38,383, 91%) were associated with hospitals, with only 9% (3,458/38,383) originating from clinics. When examined by facility ownership, government-owned health facilities handled 56.2% (21,555/38,383) of claims, compared with 43.8% (16,828/38,383) managed by private facilities. In terms of hospital classification, level B hospitals recorded the highest share of total claims (14,433/38,383, 37.6%), followed closely by level C hospitals (12,543/38,383, 32.7%).

XSL•FO
RenderX

Table 1. Descriptive statistics of renal failure claims among adult members of Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan, Indonesia's national health insurance program, between 2017 and 2022 (N=38,383).

Variable	Total, aOR (95% CI)	Male (n=21,908), aOR (95% CI)	Female (n=16,475), aOR (95% CI)	Government (n=16,828), aOR (95% CI)	Private (n=21,555), aOR (95% CI)
Membership					
Nationally subsidized	Ref	Ref	Ref	Ref	Ref
Locally subsidized	1.31[a] (1.05-1.63)	1.72[b] (1.29-2.29)	1.24 (0.89-1.72)	1.76[b] (1.34-2.31)	1.48[a] (1.06-2.06)
Informal nonworker	1.07 (0.87-1.32)	0.99 (0.72-1.36)	1.34 (0.98-1.85)	0.72 (0.52-1.01)	1.62[b] (1.22-2.16)
Informal worker	1.56[b] (1.34-1.82)	1.27[a] (1.05-1.53)	2.51[b] (1.94-3.23)	1.21 (1.00-1.48)	2.59[b] (2.06-3.27)
Formal worker	0.81[a] (0.68-0.98)	0.79 (0.62-1.01)	1.06 (0.78-1.45)	0.81 (0.62-1.04)	1.18 (0.90-1.55)
Data year					
2017	Ref	Ref	Ref	Ref	Ref
2018	0.80[b] (0.67-0.94)	0.72[b] (0.58-0.90)	0.82 (0.64-1.06)	0.91 (0.72-1.16)	0.71[a] (0.54-0.92)
2019	0.83[a] (0.71-0.98)	0.69[b] (0.56-0.87)	0.89 (0.69-1.16)	0.93 (0.73-1.19)	0.78[a] (0.60-1.00)
2020	0.81[a] (0.67-0.97)	0.82 (0.65-1.04)	0.72[a] (0.54-0.95)	0.66[b] (0.48-0.90)	0.83 (0.64-1.08)
2021	0.76[b] (0.63-0.91)	0.76[a] (0.60-0.98)	0.74[a] (0.55-0.99)	0.87 (0.66-1.14)	0.62[b] (0.46-0.82)
2022	0.97 (0.81-1.16)	1.07 (0.85-1.35)	0.92 (0.70-1.22)	0.98 (0.77-1.25)	0.99 (0.74-1.31)
Sex					
Female	Ref	Ref	Ref	Ref	Ref
Male	1.17[b] (1.04-1.32)	—[c]	—	1.38[b] (1.14-1.66)	1.24[b] (1.07-1.45)
Age group (years)					
18-39	Ref	Ref	Ref	Ref	Ref
40-49	1.37[b] (1.16-1.63)	1.69[b] (1.31-2.17)	1.10 (0.88-1.37)	1.35[a] (1.04-1.75)	1.26 (0.99-1.59)
50-59	0.90 (0.76-1.07)	1.45[b] (1.17-1.80)	0.58[b] (0.45-0.74)	1.08 (0.86-1.38)	0.54[b] (0.43-0.68)
≥60	0.74[b] (0.63-0.86)	0.97 (0.77-1.21)	0.54[b] (0.43-0.68)	0.83 (0.65-1.04)	0.62[b] (0.50-0.77)
Urbanicity					
Rural	Ref	Ref	Ref	Ref	Ref
Urban	1.07 (0.92-1.24)	1.04 (0.86-1.27)	1.08 (0.85-1.38)	1.42[b] (1.17-1.72)	0.73[b] (0.58-0.91)
Region					
Papua, Maluku, NT[d]	Ref	Ref	Ref	Ref	Ref
Java, Bali	8.30[b] (5.33-12.94)	12.50[b] (7.45-20.99)	4.50[b] (1.94-10.42)	6.68[b] (4.21-10.59)	12.97[b] (3.71-45.29)
Sumatra	7.37[b] (4.71-11.54)	7.40[b] (4.37-12.55)	8.25[b] (3.55-19.15)	5.52[b] (3.44-8.85)	13.39[b] (3.82-46.92)
Kalimantan	5.54[b] (3.46-8.86)	4.00[b] (2.24-7.12)	6.09[b] (2.60-14.29)	6.50[b] (3.98-10.61)	1.73 (0.44-6.89)
Sulawesi	1.73[a] (1.03-2.93)	0.77 (0.35-1.69)	2.71[a] (1.12-6.54)	1.36 (0.77-2.40)	2.13 (0.55-8.23)
Services					
Inpatient	Ref	Ref	Ref	Ref	Ref
Outpatient	3.74[b] (3.36-4.17)	3.11[b] (2.71-3.57)	3.22[b] (2.72-3.82)	2.95[b] (2.49-3.49)	3.82[b] (3.27-4.45)
Hospital ownership					
Government	Ref	Ref	Ref	Ref	Ref
Private	1.30[b] (1.13-1.50)	0.99 (0.83-1.19)	1.59[b] (1.28-1.98)	—	—

XSL•FO
RenderX

Variable	Total, aOR (95% CI)	Male (n=21,908), aOR (95% CI)	Female (n=16,475), aOR (95% CI)	Government (n=16,828), aOR (95% CI)	Private (n=21,555), aOR (95% CI)
Facility type					
Clinic	Ref	Ref	Ref	Ref	Ref
Hospital	0.04[b] (0.02-0.07)	0.02[b] (0.01-0.05)	0.06[b] (0.02-0.15)	5.04 (0.58-43.98)	0.01[b] (0.00-0.02)
Facility level					
Hospital level A	Ref	Ref	Ref	Ref	Ref
Hospital level B	0.97 (0.74-1.26)	0.66[a] (0.47-0.92)	1.32 (0.86-2.04)	0.65[b] (0.49-0.87)	8.13[b] (2.12-31.16)
Hospital level C	0.48[b] (0.35-0.64)	0.56[b] (0.39-0.81)	0.38[b] (0.24-0.61)	0.69[a] (0.51-0.94)	1.91 (0.49-7.41)
Hospital level D	0.56[b] (0.39-0.80)	0.81 (0.50-1.31)	0.39[b] (0.22-0.67)	0.00[b] (0.00-0.00)	3.61 (0.92-14.19)
Others	0.57[b] (0.41-0.81)	0.52[b] (0.34-0.80)	0.55[a] (0.31-0.95)	0.84 (0.58-1.21)	0.58 (0.12-2.82)

[a]*P*<.05.

[b]*P*<.01.

[c]Not applicable.

[d]NT: Nusa Tenggara.

Disparities in Receiving Hemodialysis Among BPJS Members

Tables 2 and 3 examine the associations between receiving hemodialysis across BPJS membership types for total renal failure claims, by subgroups, such as sex, facility ownership, urbanicity, and region. Compared with individuals in the lowest income group (nationally subsidized members), locally subsidized members and informal worker categories had significantly higher odds of receiving hemodialysis, with aORs of 1.31 (95% CI 1.05-1.63) and 1.56 (95% CI 1.34-1.82), respectively. However, members in the formal worker group had significantly lower odds of receiving hemodialysis (aOR 0.81, 95% CI 0.68-0.98; Table 2).

By sex, the disparity between individuals in the lowest income group (nationally subsidized members) and locally subsidized members was statistically significant among men (aOR 1.72, 95% CI 1.29-2.29) but not among women. In contrast, the disparity between nationally subsidized members and informal workers was significant for both sexes, with a notably larger disparity among women (aORs of 1.27, 95% CI 1.05-1.53, for men and 2.51, 95% CI 1.94-3.23, for women).

By facility ownership, the disparity between the nationally and locally subsidized members was significant in both government and private health facilities, with government facilities showing slightly higher odds of receiving hemodialysis (aORs of 1.76 [1.34-2.31] and 1.48 [1.06-2.06] for government and private facilities, respectively). However, the disparity between nationally subsidized members and informal workers was statistically significant only in private health facilities (aOR 2.59, 95% CI 2.06-3.27) and not in government facilities.

By urbanicity, the disparity between nationally and locally subsidized members was significant in both urban and rural settings, with rural areas showing higher odds of receiving hemodialysis (aORs of 1.69 [1.14-2.50] and 2.40 [1.78-3.23], respectively). The disparity between nationally subsidized members and informal workers was also statistically significant in both urban and rural areas, with aORs of 1.62 [1.21-2.17] and 1.73 [1.41-2.12], respectively. Additionally, the disparity between nationally subsidized members and informal nonworkers was significant only in urban areas (aOR 1.46, 95% CI 1.02-2.09; Table 3).

By region, the disparity between the nationally and locally subsidized members groups was significant in both Java or Bali and non-Java or non-Bali regions, with slightly higher odds observed in the non-Java or non-Bali regions (aORs of 1.46, 95% CI 1.08-1.98, and 1.76, 95% CI 1.31-2.37, respectively). The disparity between nationally subsidized members and informal workers was significant only in Java or Bali (aOR 2.20, 95% CI 1.82-2.67) and not in non-Java or non-Bali regions. Furthermore, the disparity between nationally subsidized members and formal workers was significant exclusively in non-Java or non-Bali regions (aOR 0.61, 95% CI 0.47-0.79).

XSL•FO
RenderX

Table 2. Adjusted odds ratios (aORs) of receiving hemodialysis among Badan Penyelenggara Jaminan Sosial (BPJS) members with renal failure, by sex and facility ownership in Indonesia between 2017 and 2022 (N=38,383).

Variable	Total, aOR (95% CI)	Male (n=21,908), aOR (95% CI)	Female (n=16,475), aOR (95% CI)	Government (n=16,828), aOR (95% CI)	Private (n=21,555), aOR (95% CI)
Membership					
Nationally subsidized	Ref	Ref	Ref	Ref	Ref
Locally subsidized	1.31[a] (1.05-1.63)	1.72[b] (1.29-2.29)	1.24 (0.89-1.72)	1.76[b] (1.34-2.31)	1.48[a] (1.06-2.06)
Informal nonworker	1.07 (0.87-1.32)	0.99 (0.72-1.36)	1.34 (0.98-1.85)	0.72 (0.52-1.01)	1.62[b] (1.22-2.16)
Informal worker	1.56[b] (1.34-1.82)	1.27[a] (1.05-1.53)	2.51[b] (1.94-3.23)	1.21 (1.00-1.48)	2.59[b] (2.06-3.27)
Formal worker	0.81[a] (0.68-0.98)	0.79 (0.62-1.01)	1.06 (0.78-1.45)	0.81 (0.62-1.04)	1.18 (0.90-1.55)
Data year					
2017	Ref	Ref	Ref	Ref	Ref
2018	0.80[b] (0.67-0.94)	0.72[b] (0.58-0.90)	0.82 (0.64-1.06)	0.91 (0.72-1.16)	0.71[a] (0.54-0.92)
2019	0.83[a] (0.71-0.98)	0.69[b] (0.56-0.87)	0.89 (0.69-1.16)	0.93 (0.73-1.19)	0.78[a] (0.60-1.00)
2020	0.81[a] (0.67-0.97)	0.82 (0.65-1.04)	0.72[a] (0.54-0.95)	0.66[b] (0.48-0.90)	0.83 (0.64-1.08)
2021	0.76[b] (0.63-0.91)	0.76[a] (0.60-0.98)	0.74[a] (0.55-0.99)	0.87 (0.66-1.14)	0.62[b] (0.46-0.82)
2022	0.97 (0.81-1.16)	1.07 (0.85-1.35)	0.92 (0.70-1.22)	0.98 (0.77-1.25)	0.99 (0.74-1.31)
Sex					
Female	Ref	Ref	Ref	Ref	Ref
Male	1.17[b] (1.04-1.32)	—	—	1.38[b] (1.14-1.66)	1.24[b] (1.07-1.45)
Age group (years)					
18-39	Ref	Ref	Ref	Ref	Ref
40-49	1.37[b] (1.16-1.63)	1.69[b] (1.31-2.17)	1.10 (0.88-1.37)	1.35[a] (1.04-1.75)	1.26 (0.99-1.59)
50-59	0.90 (0.76-1.07)	1.45[b] (1.17-1.80)	0.58[b] (0.45-0.74)	1.08 (0.86-1.38)	0.54[b] (0.43-0.68)
≥60	0.74[b] (0.63-0.86)	0.97 (0.77-1.21)	0.54[b] (0.43-0.68)	0.83 (0.65-1.04)	0.62[b] (0.50-0.77)
Urbanicity					
Rural	Ref	Ref	Ref	Ref	Ref
Urban	1.07 (0.92-1.24)	1.04 (0.86-1.27)	1.08 (0.85-1.38)	1.42[b] (1.17-1.72)	0.73[b] (0.58-0.91)
Region					
Papua, Maluku, NT	Ref	Ref	Ref	Ref	Ref
Java, Bali	8.30[b] (5.33-12.94)	12.50[b] (7.45-20.99)	4.50[b] (1.94-10.42)	6.68[b] (4.21-10.59)	12.97[b] (3.71-45.29)
Sumatra	7.37[b] (4.71-11.54)	7.40[b] (4.37-12.55)	8.25[b] (3.55-19.15)	5.52[b] (3.44-8.85)	13.39[b] (3.82-46.92)
Kalimantan	5.54[b] (3.46-8.86)	4.00[b] (2.24-7.12)	6.09[b] (2.60-14.29)	6.50[b] (3.98-10.61)	1.73 (0.44-6.89)
Sulawesi	1.73[a] (1.03-2.93)	0.77 (0.35-1.69)	2.71[a] (1.12-6.54)	1.36 (0.77-2.40)	2.13 (0.55-8.23)
Services					
Inpatient	Ref	Ref	Ref	Ref	Ref
Outpatient	3.74[b] (3.36-4.17)	3.11[b] (2.71-3.57)	3.22[b] (2.72-3.82)	2.95[b] (2.49-3.49)	3.82[b] (3.27-4.45)
Hospital ownership					
Government	Ref	Ref	Ref	Ref	Ref
Private	1.30[b] (1.13-1.50)	0.99 (0.83-1.19)	1.59[b] (1.28-1.98)	—	—
Facility type					

XSL•FO
RenderX

Variable	Total, aOR (95% CI)	Male (n=21,908), aOR (95% CI)	Female (n=16,475), aOR (95% CI)	Government (n=16,828), aOR (95% CI)	Private (n=21,555), aOR (95% CI)
Clinic	Ref	Ref	Ref	Ref	Ref
Hospital	0.04[b] (0.02-0.07)	0.02[b] (0.01-0.05)	0.06[b] (0.02-0.15)	5.04 (0.58-43.98)	0.01[b] (0.00-0.02)
Facility level					
Hospital level A	Ref	Ref	Ref	Ref	Ref
Hospital level B	0.97 (0.74-1.26)	0.66[a] (0.47-0.92)	1.32 (0.86-2.04)	0.65[b] (0.49-0.87)	8.13[b] (2.12-31.16)
Hospital level C	0.48[b] (0.35-0.64)	0.56[b] (0.39-0.81)	0.38[b] (0.24-0.61)	0.69[a] (0.51-0.94)	1.91 (0.49-7.41)
Hospital level D	0.56[b] (0.39-0.80)	0.81 (0.50-1.31)	0.39[b] (0.22-0.67)	0.00[b] (0.00-0.00)	3.61 (0.92-14.19)
Others	0.57[b] (0.41-0.81)	0.52[b] (0.34-0.80)	0.55[a] (0.31-0.95)	0.84 (0.58-1.21)	0.58 (0.12-2.82)

[a]*P*<.05.

[b]*P*<.01.

[c]NT: Nusa Tenggara.

[d]Not applicable.

XSL•FO
RenderX

Table 3. Adjusted odds ratios (aORs) of receiving hemodialysis among Badan Penyelenggara Jaminan Sosial (BPJS) members with renal failure, by urbanicity and region in Indonesia between 2017 and 2022.

Variable	Urban, aOR (95% CI)	Rural, aOR (95% CI)	Java, aOR (95% CI)	Non-Java, aOR (95% CI)
Membership				
Nationally subsidized	Reference	Reference	Reference	Reference
Locally subsidized	1.69[a] (1.14-2.50)	2.40[a] (1.78-3.23)	1.46[b] (1.08-1.98)	1.76[a] (1.31-2.37)
Informal nonworker	1.46[b] (1.02-2.09)	0.78 (0.56-1.09)	1.10 (0.85-1.43)	1.43 (1.00-2.07)
Informal worker	1.62[a] (1.21-2.17)	1.73[a] (1.41-2.12)	2.20[a] (1.82-2.67)	1.02 (0.81-1.28)
Formal worker	0.96 (0.69-1.33)	0.87 (0.67-1.13)	0.86 (0.68-1.09)	0.61[a] (0.47-0.79)
Data year				
2017	Reference	Reference	Reference	Reference
2018	0.71[a] (0.55-0.91)	0.94 (0.74-1.19)	0.82[b] (0.67-0.99)	0.76 (0.58-1.01)
2019	0.73[b] (0.57-0.93)	1.02 (0.79-1.30)	0.92 (0.75-1.12)	0.77 (0.59-1.02)
2020	0.52[a] (0.40-0.68)	1.10 (0.82-1.48)	0.81 (0.65-1.01)	0.94 (0.71-1.25)
2021	0.51[a] (0.39-0.66)	1.22 (0.93-1.61)	0.83 (0.65-1.04)	0.93 (0.70-1.24)
2022	0.74[b] (0.57-0.97)	1.32[b] (1.03-1.68)	1.08 (0.87-1.35)	1.04 (0.78-1.38)
Sex				
Female	Reference	Reference	Reference	Reference
Male	0.85 (0.72-1.00)	1.77[a] (1.46-2.14)	1.75[a] (1.50-2.03)	0.66[a] (0.56-0.78)
Age group				
18-39 years	Reference	Reference	Reference	Reference
40-49 years	1.08 (0.85-1.37)	1.60[a] (1.25-2.05)	0.86 (0.70-1.07)	2.07[a] (1.58-2.72)
50-59 years	0.66[a] (0.52-0.85)	1.17 (0.93-1.49)	0.71[a] (0.57-0.87)	0.86 (0.66-1.11)
60+ years	0.68[a] (0.54-0.85)	0.82 (0.65-1.04)	0.51[a] (0.42-0.63)	0.78 (0.59-1.03)
Urbanicity				
Rural	Reference	Reference	Reference	Reference
Urban	__[c]	—	1.19 (0.97-1.45)	0.91 (0.73-1.14)
Region				
Papua, Maluku, NT[d]	Reference	Reference	Reference	Reference
Java, Bali	8.77[a] (5.50-13.98)	8.46[a] (3.20-22.33)	—	—
Sumatra	9.02[a] (5.58-14.57)	6.95[a] (2.65-18.25)	—	—
Kalimantan	7.56[a] (4.53-12.60)	2.96[b] (1.07-8.18)	—	—
Sulawesi	3.67[a] (2.13-6.34)	0.21[b] (0.06-0.75)	—	—
Services				
Inpatient	Reference	Reference	Reference	Reference
Outpatient	2.75[a] (2.35-3.22)	3.05[a] (2.54-3.67)	3.11[a] (2.71-3.58)	3.87[a] (3.26-4.59)
Hospital ownership				
Government	Reference	Reference	Reference	Reference
Private	1.22 (0.99-1.49)	0.97 (0.78-1.19)	1.32[a] (1.10-1.57)	1.34[a] (1.10-1.64)
Facility type				
Clinic	Reference	Reference	Reference	Reference

XSL•FO
RenderX

Variable	Urban, aOR (95% CI)	Rural, aOR (95% CI)	Java, aOR (95% CI)	Non-Java, aOR (95% CI)
Hospital	0.10[a] (0.04-0.24)	0.02[a] (0.01-0.04)	0.05[a] (0.03-0.10)	0.01[a] (0.00-0.02)
Facility level				
Hospital level A	Reference	Reference	Reference	Reference
Hospital level B	1.43[b] (1.07-1.92)	2.05 (0.44-9.56)	2.00[a] (1.36-2.94)	0.35[a] (0.25-0.49)
Hospital level C	0.26[a] (0.19-0.36)	3.40 (0.72-16.00)	0.85 (0.56-1.31)	0.24[a] (0.17-0.35)
Hospital level D	0.07[a] (0.02-0.18)	4.30 (0.91-20.44)	1.57 (0.95-2.60)	0.01[a] (0.01-0.03)
Others	0.51[a] (0.36-0.74)	3.13 (0.63-15.48)	1.29 (0.82-2.02)	0.19[a] (0.11-0.33)

[a]$P<.01$.

[a]$P<.05$.

[c]Not applicable.

[d]NT: Nusa Tenggara.

Disparities in Access to Hemodialysis Among BPJS Members Across Covariates

Within Indonesia's national health insurance system, men demonstrated significantly higher odds of receiving hemodialysis compared with women (aOR 1.17, 95% CI 1.04-1.32). However, age-related disparities were more nuanced. Compared with the youngest age group (18-30 years), individuals aged 40 to 49 years were significantly more likely to receive hemodialysis (aOR 1.37, 95% CI 1.16-1.63), while older adults aged ≥60 years were less likely (aOR 0.74, 95% CI 0.63-0.86; Table 2).

Regarding urbanicity, although patients in urban areas had slightly higher rates of receiving hemodialysis, the difference was not statistically significant in the multivariate analysis (Table 2). Regional disparities, however, were notable. Compared with members residing in the least advantaged region (Papua, Maluku, and Nusa Tenggara), those living in other regions had significantly greater odds of accessing hemodialysis, with particularly high odds observed in Java or Bali, Sumatra, and Kalimantan (aORs of 8.30, 95% CI 5.33-12.94, 7.37, 95% CI 4.71-11.54, and 5.54, 95% CI 3.46-8.86, respectively).

Facility-level characteristics also played a significant role. Patients using outpatient services were more likely to receive hemodialysis compared with those in inpatient services (aOR 3.74, 95% CI 3.36-4.17). Furthermore, patients treated at private facilities had higher odds of accessing hemodialysis than those at government-run facilities (aOR 1.30, 95% CI 1.13-1.50). When examining at the facility level, patients in hospitals classified below level A had lower odds of receiving hemodialysis, with significant disparities evident for those treated at level C and level D hospitals, as well as in other types of facilities, such as clinics and specialized hemodialysis centers.

Discussion

Overview

Our study highlights significant socioeconomic, demographic, and facility-level disparities in hemodialysis claim patterns, which—within the context of near-universal BPJS coverage—can be considered a reasonable proxy for disparities in access to renal care among members diagnosed with renal failure in Indonesia. Using national health insurance claims data from 2017 to 2022, we identified inequities across membership types, sex, age, urbanicity, region, and facility characteristics. These findings underscore the need for targeted interventions to address barriers to equitable health care access, particularly in lower socioeconomic groups and underserved regions.

Socioeconomic Disparities

The disparities observed in hemodialysis access across BPJS membership types reflect broader inequities within Indonesia's health care system. Informal workers had higher odds of receiving hemodialysis compared with those in the lowest income group (nationally subsidized members). This suggests that while BPJS coverage reduces financial barriers for subsidized members, other systemic obstacles—such as accessibility and awareness—may disproportionately affect the lowest income groups. Prior studies have similarly reported that socioeconomically disadvantaged populations often face multiple barriers to accessing specialized health care services, even under universal health coverage schemes [7,10,17,18].

In contrast, formal workers had lower odds of receiving hemodialysis, which may reflect competing demands of employment that limit their ability to access treatment. This aligns with research indicating that time and resource constraints among employed populations can adversely impact health care use [25]. Policymakers should consider tailored strategies, such as workplace health initiatives, to mitigate these challenges.

Demographic and Regional Disparities

Demographic analysis revealed notable sex-based disparities, with men being more likely to receive hemodialysis than women. This finding echoes global trends where men often exhibit higher health care use for certain conditions, potentially because of physiological differences in disease progression or social factors influencing health-seeking behavior [26]. Age-based disparities were also evident, with younger patients more likely to access hemodialysis compared with older adults. This aligns with findings from other LMICs where older populations face higher treatment discontinuation rates due to financial, logistical, or health-related constraints [27].

XSL•FO
RenderX

Regional disparities were particularly stark, with patients in Java or Bali having significantly higher odds of receiving hemodialysis compared with those in Papua, Maluku, and Nusa Tenggara. Java or Bali's well-developed health care infrastructure likely contributes to this advantage, while the limited availability of specialized facilities in less advantaged regions hinders access. Previous studies have documented similar regional and urban-rural divides in health care access, emphasizing the need for infrastructure investments to bridge these gaps [20,28].

Facility-Level Factors

Facility-level factors played a crucial role in determining access to hemodialysis. Outpatient services were significantly associated with higher odds of receiving hemodialysis compared with inpatient settings. This may reflect the chronic nature of renal failure, where frequent outpatient visits for dialysis are standard. However, reliance on outpatient care may disadvantage patients who lack consistent transportation or proximity to facilities, as observed in other studies on chronic disease management [29]. Facility ownership and level also influenced access disparities. Patients treated at private facilities were more likely to receive hemodialysis than those in government-run hospitals. This finding highlights potential differences in service capacity or quality between private and public providers, a concern frequently raised in studies on dual-sector health care systems [6,30]. The consistently lower odds of hemodialysis claims observed in hospitals below level A likely reflect the referral system in Indonesia, where the most severe and complex renal failure cases are concentrated at national referral (level A) hospitals; this pattern may therefore indicate appropriate care allocation rather than inequality, while disparities across regions and facility types may still contribute to variation in access [11,23].

For policy, to address the observed disparities in hemodialysis access, policy makers should consider targeted interventions. Outreach programs focused on subsidized members, particularly in rural and underserved regions, can help overcome geographic and financial barriers through expanded transportation assistance, public education campaigns, and community-based dialysis centers. Strengthening government health facilities with investments in equipment, staffing, and training is essential to improving service quality and equity, while public-private partnerships could enhance service delivery and affordability [6]. In Indonesia, dialysis services remain overwhelmingly hospital-based, with the Indonesia Renal Registry reporting that 98% of patients receive hemodialysis while only 2% use continuous ambulatory peritoneal dialysis, despite continuous ambulatory peritoneal dialysis being cost-saving and fully covered by BPJS [31]. This underuse indicates that alternative modalities and service adaptations—such as home-based or mobile dialysis and flexible scheduling—are insufficiently implemented. Our analysis also showed that formal sector workers had lower odds of hemodialysis claims compared with other groups, which may reflect work-related barriers. Addressing these challenges could involve piloting flexible work policies and mobile dialysis units, supported by evidence that flexible scheduling helps patients maintain employment and that mobile services can reduce transport-related barriers [31,32]. Finally, raising awareness about renal failure management among patients and health care providers can improve timely recognition and management of the condition [13,14].

This study has several limitations. First, this study focuses on claims involving hemodialysis, which accounts for over 90% of dialysis-related procedures in the dataset. While peritoneal dialysis and transplantation are technically available in Indonesia, their uptake is extremely limited and inconsistently captured in claims data, which may result in the exclusion of a small subset of patients receiving alternative modalities [5,13]. Second, claims data may not capture all cases of renal failure, particularly among uninsured or underdiagnosed populations. Third, while aORs provide insights into associations, causal relationships cannot be inferred. Future research should explore longitudinal data to better understand the dynamics of hemodialysis access over time.

While delays in care-seeking and treatment initiation are plausible concerns in Indonesia's health system, our claim-level data do not allow us to measure patient-level outcomes, such as time to first dialysis, missed sessions, or referral delays. As such, these issues cannot be empirically confirmed in this study and should be addressed in future research using longitudinal or patient-level data.

Conclusions

This study provides critical insights into socioeconomic, demographic, and facility-level disparities in hemodialysis claim patterns among BPJS members in Indonesia. Hemodialysis was provided in 75.6% (29,017/38,383) of renal failure claims. Nationally subsidized members had significantly lower odds of receiving hemodialysis compared with informal workers and locally subsidized members, with aORs of 1.56 and 1.31, respectively. Disparities were more pronounced in rural areas, where locally subsidized members had higher odds (aOR 2.40) than their nationally subsidized counterparts. Patients in private facilities (aOR 1.30) and urban areas had better access, while regional disparities favored developed regions like Java or Bali.

While these findings highlight meaningful inequities, it is important to note that our dataset is claim-based and does not permit direct assessment of patient-level outcomes, such as time to first dialysis, missed sessions, or referral delays. Future studies incorporating longitudinal patient trajectories are needed to examine these dynamics more precisely. Nevertheless, the disparities identified here underscore the need for coordinated policy action. Strengthening health care infrastructure, expanding community-based and mobile dialysis services, improving workplace flexibility for formal sector workers, and reducing barriers for subsidized and rural populations are critical steps. By implementing targeted interventions, Indonesia can make significant progress toward universal and equitable access to essential renal care services.

XSL•FO
RenderX

Acknowledgments

The authors gratefully acknowledge Badan Penyelenggara Jaminan Sosial Kesehatan for providing data access.

Data Availability

The dataset used in this study is available from the authors upon reasonable request.

Authors' Contributions

AM, DK, and SRH conceived the study. SRH and DK conducted data extraction and cleaning; DK and SRH conducted data analyses. DK drafted the manuscript, and AM and SRH provided inputs to the manuscript. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The flowchart outlining the sample selection process.
[DOCX File , 313 KB - publichealth_v11i1e73731_app1.docx]

Multimedia Appendix 2
Total number of claims and unique patients by year.
[DOCX File , 282 KB - publichealth_v11i1e73731_app2.docx]

References

1. Kovesdy CP. Epidemiology of chronic kidney disease: an update 2022. Kidney Int Suppl (2011) 2022 Apr;12(1):7-11 [FREE Full text] [doi: 10.1016/j.kisu.2021.11.003] [Medline: 35529086]
2. GBD 2021 Causes of Death Collaborators. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):2100-2132 [FREE Full text] [doi: 10.1016/S0140-6736(24)00367-2] [Medline: 38582094]
3. GBD Chronic Kidney Disease Collaboration. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2020 Feb 29;395(10225):709-733 [FREE Full text] [doi: 10.1016/S0140-6736(20)30045-3] [Medline: 32061315]
4. VizHub - GBD results. Institute for Health Metrics and Evaluation. URL: https://vizhub.healthdata.org/gbd-results/ [accessed 2024-12-24]
5. Jonny, Violetta L, Kusumaningrum VF. Peritoneal dialysis in Indonesia: current status, challenges and prospects. Perit Dial Int 2022 Jul;42(4):428-433. [doi: 10.1177/08968608211034985] [Medline: 34338050]
6. Hyodo T, Hirawa N, Kuragano T, Takemoto Y, Urabe S, Adiya S, et al. The present status of dialysis patients in Asian countries as of 2022 from a medical economic point of view. Ren Replace Ther 2025 Jan 06;11(1):1-11. [doi: 10.1186/s41100-024-00597-1]
7. Nugroho ST, Ahsan A, Kusuma D, Adani N, Irawaty DK, Amalia N, et al. Income disparity and healthcare utilization: lessons from Indonesia's national health insurance claim data. Asian Pac J Cancer Prev 2023 Oct 01;24(10):3397-3402. [doi: 10.31557/APJCP.2023.24.10.3397] [Medline: 37898843]
8. Darmawan ES, Permanasari VY, Nisrina LV, Kusuma D, Hasibuan SR, Widyasanti N. Behind the hospital ward: in-hospital mortality of type 2 diabetes mellitus patients in Indonesia (analysis of national health insurance claim sample data). Int J Environ Res Public Health 2024 May 01;21(5):581 [FREE Full text] [doi: 10.3390/ijerph21050581] [Medline: 38791795]
9. Untari PH. BPJS Kesehatan bayar klaim paling besar untuk 8 penyakit sepanjang 2023, capai Rp34,7 triliun. Bisnis. 2024 Jan 12. URL: https://finansial.bisnis.com/read/20240112/215/1731664/bpjs-kesehatan-bayar-klaim-paling-besar-untuk-8-penyakit-sepanjang-2023-capai-rp347-triliun [accessed 2024-10-24]
10. Bachtiar A, Candi C, Hasibuan SR, Widyasanti N, Kusuma D. Navigating the aftermath: risk factors of recurrence following coronary bypass surgery in Indonesia. Narra J 2024 Dec;4(3):e969. [doi: 10.52225/narra.v4i3.969] [Medline: 39816076]
11. Darmawan ES, Hasibuan SR, Permanasari VY, Kusuma D. Disparities in health services and outcomes by National Health Insurance membership type for ischemic heart disease and stroke in Indonesia: analysis of claims, 2017-2022. Glob Health Res Policy 2025 Aug 01;10(1):33 [FREE Full text] [doi: 10.1186/s41256-025-00432-y] [Medline: 40745615]
12. Mahendradhata Y, Trisnantoro L, Listyadewi S, Soewondo P, Marthias T, Harimurti P, et al. The Republic of Indonesia health system review. World Health Organization. 2017. URL: https://iris.who.int/handle/10665/254716 [accessed 2024-12-24]

XSL•FO
RenderX

13. Bello AK, Okpechi IG, Levin A, Ye F, Damster S, Arruebo S, ISN-GKHA Group. An update on the global disparities in kidney disease burden and care across world countries and regions. Lancet Glob Health 2024 Mar;12(3):e382-e395. [doi: 10.1016/S2214-109X(23)00570-3] [Medline: 38365413]
14. Francis A, Harhay MN, Ong AC, Tummalapalli SL, Ortiz A, Fogo AB, American Society of Nephrology, European Renal Association, International Society of Nephrology. Chronic kidney disease and the global public health agenda: an international consensus. Nat Rev Nephrol 2024 Jul;20(7):473-485. [doi: 10.1038/s41581-024-00820-6] [Medline: 38570631]
15. Andriani H, Rahmawati ND, Ahsan A, Kusuma D. Secondhand smoke exposure inside the house and low birth weight in Indonesia: evidence from a demographic and health survey. Popul Med 2023 Jun 23;5(June):1-7. [doi: 10.18332/popmed/168620]
16. Oktamianti P, Kusuma D, Amir V, Tjandrarini DH, Paramita A. Does the disparity patterning differ between diagnosed and undiagnosed hypertension among adults? Evidence from Indonesia. Healthcare (Basel) 2023 Mar 10;11(6):816 [FREE Full text] [doi: 10.3390/healthcare11060816] [Medline: 36981473]
17. Park HC, Kwon YE, Choi HY, Oh HJ, Chang TI, Kang EW, et al. Health insurance status is related to risk of mortality and hospitalization in Korean maintenance hemodialysis patients: a longitudinal cohort study. Am J Nephrol 2020;51(12):975-981. [doi: 10.1159/000512855] [Medline: 33440390]
18. Ke C, Kim SJ, Shah BR, Bierman AS, Lipscombe LL, Feig DS, et al. Impact of socioeconomic status on incidence of end-stage renal disease and mortality after dialysis in adults with diabetes. Can J Diabetes 2019 Oct;43(7):483-9.e4. [doi: 10.1016/j.jcjd.2019.04.006] [Medline: 31133437]
19. Shankar M, Sankarasubaiyan S, Kasiviswanathan S, Shah KD, Luyckx V. Gender disparity in hemodialysis practices and mortality: a nationwide cross-sectional observational study. Indian J Nephrol 2024;34(6):609-616. [doi: 10.25259/ijn_559_23] [Medline: 39649311]
20. Sulistiadi W, Kusuma D, Amir V, Tjandrarini DH, Nurjana MA. Growing up unequal: disparities of childhood overweight and obesity in Indonesia's 514 districts. Healthcare (Basel) 2023 May 05;11(9):1322 [FREE Full text] [doi: 10.3390/healthcare11091322] [Medline: 37174864]
21. Darmawan ES, Kusuma D, Permanasari VY, Amir V, Tjandrarini DH, Dharmayanti I. Beyond the plate: uncovering inequalities in fruit and vegetable intake across Indonesian districts. Nutrients 2023 Apr 30;15(9):2160 [FREE Full text] [doi: 10.3390/nu15092160] [Medline: 37432281]
22. Setiawan E, Nurjannah N, Komaryani K, Nugraha RR, Thabrany H, Purwaningrum F, et al. Utilization patterns of healthcare facility and estimated expenditure of PLHIV care under the Indonesian National Health Insurance Scheme in 2018. BMC Health Serv Res 2022 Jan 22;22(1):97 [FREE Full text] [doi: 10.1186/s12913-021-07434-9] [Medline: 35065632]
23. Darmawan ES, Hasibuan SR, Permanasari VY, Kusuma D. Disparities in cancer care and outcomes by insurance membership type in Indonesia: a retrospective cross-sectional analysis of national health insurance claims, 2017-2022. BMJ Open 2025 Aug 01;15(7):e096486 [FREE Full text] [doi: 10.1136/bmjopen-2024-096486] [Medline: 40750275]
24. ICD-10 version: 2016. World Health Organization. 2016. URL: https://icd.who.int/browse10/2016/en#/II [accessed 2024-11-08]
25. Tsutsui H, Nomura K, Ishiguro A, Tsuruta Y, Kato S, Yasuda Y, et al. Factors associated with employment in patients undergoing hemodialysis: a mixed methods study. Ren Replace Ther 2017 Jun 2;3(1):1-11. [doi: 10.1186/s41100-017-0105-z]
26. Hecking M, Bieber BA, Ethier J, Kautzky-Willer A, Sunder-Plassmann G, Säemann MD, et al. Sex-specific differences in hemodialysis prevalence and practices and the male-to-female mortality rate: the Dialysis Outcomes and Practice Patterns Study (DOPPS). PLoS Med 2014 Oct 28;11(10):e1001750. [doi: 10.1371/journal.pmed.1001750] [Medline: 25350533]
27. Tonelli M, Nkunu V, Varghese C, Abu-Alfa AK, Alrukhaimi MN, Fox L, et al. Framework for establishing integrated kidney care programs in low- and middle-income countries. Kidney Int Suppl (2011) 2020 Mar;10(1):e19-e23 [FREE Full text] [doi: 10.1016/j.kisu.2019.11.002] [Medline: 32149006]
28. Megatsari H, Kusuma D, Ernawaty E, Putri NK. Geographic and socioeconomic inequalities in delays in COVID-19 vaccinations: a cross-sectional study in Indonesia. Vaccines (Basel) 2022 Nov 02;10(11):1857 [FREE Full text] [doi: 10.3390/vaccines10111857] [Medline: 36366365]
29. Starbird LE, DiMaina C, Sun CA, Han HR. A systematic review of interventions to minimize transportation barriers among people with chronic diseases. J Community Health 2019 Apr;44(2):400-411 [FREE Full text] [doi: 10.1007/s10900-018-0572-3] [Medline: 30206755]
30. Septiono W. Equity challenges in Indonesian health care. Lancet Glob Health 2023 May;11(5):e646-e647. [doi: 10.1016/S2214-109X(23)00110-9] [Medline: 37061303]
31. Putri S, Nugraha RR, Pujiyanti E, Thabrany H, Hasnur H, Istanti ND, et al. Supporting dialysis policy for end stage renal disease (ESRD) in Indonesia: an updated cost-effectiveness model. BMC Res Notes 2022 Dec 06;15(1):359 [FREE Full text] [doi: 10.1186/s13104-022-06252-4] [Medline: 36474238]
32. Liu MW, Syukri M, Abdullah A, Chien LY. Missing in-center hemodialysis sessions among patients with end stage renal disease in Banda Aceh, Indonesia. Int J Environ Res Public Health 2021 Aug 31;18(17):9215 [FREE Full text] [doi: 10.3390/ijerph18179215] [Medline: 34501804]

XSL•FO
RenderX

Abbreviations

aOR: adjusted odds ratio
BPJS: Badan Penyelenggara Jaminan Sosial
CKD: chronic kidney disease
ESRD: end-stage renal disease
ICD: International Classification of Diseases
LMIC: low- and middle-income country

Edited by A Mavragani, T Sanchez; submitted 11.03.25; peer-reviewed by NK Putri, G Rostoker; comments to author 24.06.25; revised version received 03.07.25; accepted 18.09.25; published 03.11.25.

Please cite as:
Munandar A, Hasibuan SR, Kusuma D
Differences in Hemodialysis Claim Patterns Across Membership Types Among Patients With Renal Failure Based on National Health Insurance Data From 2017 to 2022: Cross-Sectional Analysis
JMIR Public Health Surveill 2025;11:e73731
URL: https://publichealth.jmir.org/2025/1/e73731
doi:10.2196/73731
PMID:

©Aries Munandar, Syarif R Hasibuan, Dian Kusuma. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 03.11.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Dental Caries Detection in Children Using Intraoral Scanners Featuring Fluorescence: Diagnostic Agreement Study

Bree Jones[1,2], MPH, PhD; Tong Chen[1,3], PhD; Stavroula Michou[4], PhD; Nicky Kilpatrick[1,5], PhD; Nigel Curtis[2,5,6], MD, MBBS, PhD; David P Burgner[2,5,6], MD, PhD; Christoph Vannahme[7], PhD; Mihiri Silva[1,2], PhD

[1]Melbourne Dental School, University of Melbourne, 720 Swanston Street, Carlton, Australia

[2]Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, Australia

[3]Clinical Epidemiology & Biostatistics Unit, Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, Australia

[4]Department of Odontology, School of Health and Medical Sciences, University of Copenhagen, Nørre Allé, Denmark

[5]Department of Paediatrics, University of Melbourne, Royal Children's Hospital, Parkville, Australia

[6]Department of Infectious Diseases, Royal Children's Hospital, Parkville, Australia

[7]3Shape TRIOS A/S, Copenhagen, Denmark

Corresponding Author:
Bree Jones, MPH, PhD
Melbourne Dental School, University of Melbourne, 720 Swanston Street, Carlton, Australia

Abstract

Background: Dental caries is a common chronic disease in children. Digital tools such as intraoral scanners (IOSs) may offer an efficient, scalable alternative to conventional visual examination for dental caries detection. IOSs are handheld devices that generate 3-dimensional (3D) models of the teeth and surrounding structures. Recent advances incorporating fluorescence technology into scanner hardware offer the potential for supporting dental caries detection. However, the performance of digital caries detection methods using 3D models that include both color and fluorescence in children's teeth remains unknown.

Objective: This study aimed to assess the diagnostic agreement between visual examination and on-screen assessment of 3D models generated by an IOS in color and supplemented with fluorescence for caries detection in primary teeth.

Methods: Children participating in a clinical trial (n=216, mean age 5.6, SD 0.4 y) underwent visual examination using the International Caries Detection and Assessment System (ICDAS) and intraoral scanning using the TRIOS 4 IOS. Four trained registered dental practitioners independently assessed each participant's 3D models in color and then supplemented with fluorescence, using a previously validated ICDAS index modified for on-screen assessments of 3D models. All 3D models were assessed again after 4 weeks. The time taken for intraoral scanning and on-screen assessment was recorded. Multilevel logistic regression was used to estimate and compare the likelihood of detecting caries between methods, and Bland-Altman plots were used to visualize agreement. Analyses were performed at the initial (ICDAS ≥01), moderate (ICDAS ≥03), and extensive (ICDAS ≥05) dental caries thresholds. Intraclass correlation coefficient (ICC) estimated method agreement and examiner reliability.

Results: Of 219 children enrolled, 216 completed dental assessments. A total of 9470 visible primary tooth surfaces were included in the analysis. The average time taken for on-screen assessment of each 3D model (color with fluorescence) was 3.5 (SD 2.3) minutes. The likelihood of detecting caries using color assessment of 3D models was similar to visual examination at all disease thresholds: initial (odds ratio [OR] 1.1, 95% CI 1.0 - 1.3), moderate (OR 0.9, 95% CI 0.7 - 1.1), and extensive (OR 1.0, 95% CI 0.7 - 1.3). When color assessments were supplemented with fluorescence, the likelihood of detecting caries was 30% higher at the initial threshold relative to visual examination (OR 1.3, 95% CI 1.1 - 1.5) and similar at the moderate (OR 0.9, 95% CI 0.7 - 1.1) and extensive thresholds (OR 0.9, 95% CI 0.7 - 1.3). Bland-Altman plots showed a high level of agreement at both moderate and extensive thresholds. Agreement between methods was high (ICC 0.9, 95% CI 0.9 - 1.0). Intra- and inter-examiner reliability using intraoral scans ranged from good to excellent (ICC 0.8 - 1.0).

Conclusions: On-screen assessment of 3D models in color demonstrated the highest agreement with visual examination for caries detection across all disease thresholds.

(JMIR Public Health Surveill 2025;11:e78023) doi:10.2196/78023

KEYWORDS

dental caries; dentition, primary; image interpretation, computer-assisted; models, dental; imaging, 3-dimensional; computer-assisted

XSL•FO
RenderX

Introduction

Dental caries is a chronic, multifactorial disease influenced by environmental and behavioral factors [1]. It is estimated to impact more than half a billion children globally, representing a major public health concern [2]. Developing efficient and accurate tools to collect data on dental caries can facilitate the inclusion of oral health outcomes in epidemiological research, improving understanding of contemporary risk factors, disease distribution, and the burden of disease in children.

The clinical presentation of dental caries spans a spectrum from early to advanced disease. In its earliest stage, mineral loss causes subsurface changes that alter enamel porosity, and subsequently its optical properties. At this point, the enamel surface remains intact, and the process can be reversed or arrested through remineralization strategies [3]. If mineral loss continues, tooth structure becomes compromised, potentially leading to cavity formation. These later stages signify irreversible damage to the tooth and may eventually result in pain and tooth loss if left untreated [3].

Conventional assessment of dental caries in epidemiological research involves a visual examination by a qualified dental practitioner using a validated scoring index. Implementing visual examinations across large or multisite studies requires substantial investment in examiner training and calibration to ensure consistency, as interexaminer variability can impact study validity [4]. The financial and logistical challenges may limit the use of visual examinations, with some studies instead relying on self-reported caries status. Self-reporting methods have been criticized for their lack of validity and their tendency to underestimate disease burden in children [5].

Intraoral scanners (IOSs) are hand-held devices that generate 3-dimensional (3D) digital scans of the teeth and surrounding structures (herein 3D models). It has been suggested that 3D models can be used to perform digital dental examinations [6]. The various commercially available devices differ in their hardware and software for image acquisition and reconstruction. Manufacturers include, but are not limited to, Align Technologies (San Jose, California, USA), Medit Corp (Seoul, South Korea), Dentsply Sirona (Charlotte, NC, USA), Planmeca (Helsinki, Finland), and 3/Shape (Copenhagen, Denmark).

The TRIOS 4 and TRIOS 5 scanners (3/Shape, Copenhagen, Denmark) incorporate light-induced fluorescence technology into their hardware to aid dental caries detection. The devices use an LED light source (approximately 415 nm) and a long-pass filter to capture emitted red and green fluorescence [7]. When healthy teeth are exposed to blue light, they become excited to emit green fluorescence [8]. Areas affected by dental caries have reduced green fluorescence intensity, as mineral loss alters the tooth's composition and its light-scattering properties[9] At these wavelengths, some bacterial metabolites associated with dental caries emit red fluorescence. Using the scanner software, green and red fluorescence can be visualized simultaneously as an overlay on the 3D model. Fluorescence characteristics can be interpreted by a dental practitioner to support detection of subtle early lesions which may be difficult to visualize when viewing scans in color alone.

The International Caries Detection and Assessment System is a widely accepted scoring index for detecting and classifying dental caries severity [10]. To reduce subjectivity, a modified version incorporating fluorescence features has been developed [11] and validated [12]. In adult permanent teeth, an in vitro study investigating on-screen assessment using this index found no difference in diagnostic accuracy between on-screen assessment and conventional visual examination for occlusal tooth surfaces [13]. However, the real-world diagnostic performance of an on-screen caries assessment for primary teeth and other surface types (such as visible smooth surfaces) is unknown.

Clinically, enamel defects represent another prevalent condition that can resemble dental caries, as affected areas often display similar visual characteristics. There is also overlap in the indices used to record and quantify both conditions [14]. Yet, most studies evaluating caries detection methods typically exclude teeth with enamel defects, leaving their influence on diagnostic performance uncertain.

This study, therefore, aimed to determine the in vivo diagnostic agreement between visual examination and on-screen assessment of 3D models in approximate natural colors with and without fluorescence for detecting and classifying caries in primary teeth. The secondary aims included exploring the impact of disease threshold, enamel defects, and surface type on agreement estimates.

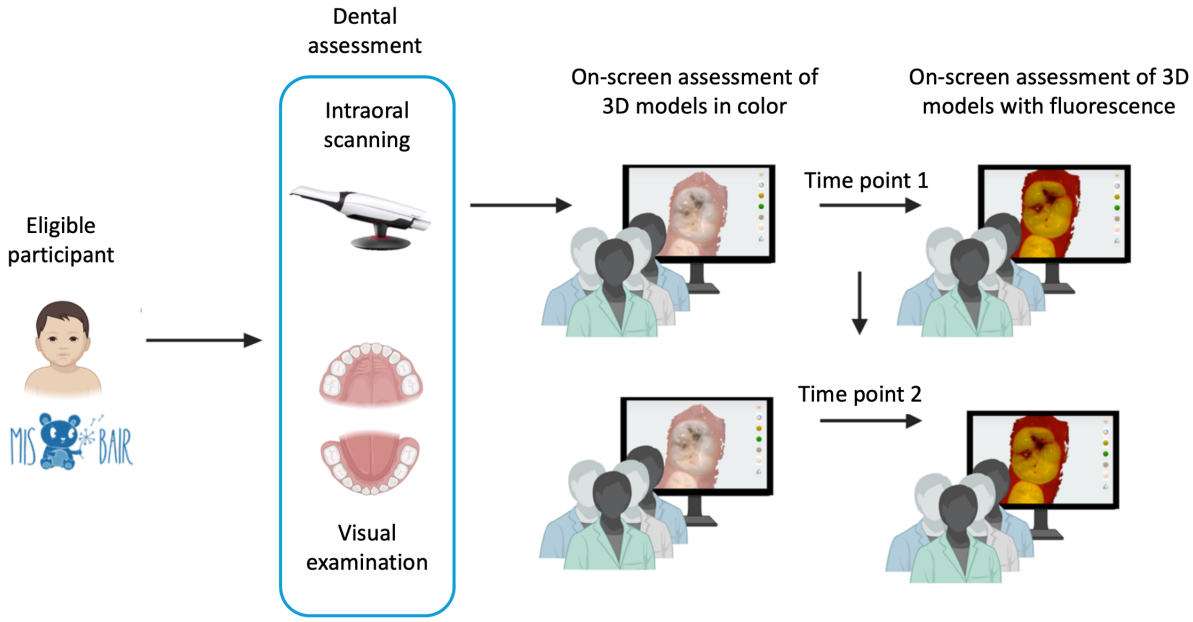
Methods

Study Design

A cross-sectional diagnostic agreement study was undertaken (Figure 1), nested within a longitudinal randomized controlled trial. An agreement design was selected as the visual examination comparator was not considered a gold standard reference in the study setting [15]. This study followed the Guidelines for Reporting Reliability and Agreement Studies [16] (Table SA-1 in Multimedia Appendix 1) and was conducted per the applicable sections of the Standards for Reporting of Diagnostic Accuracy Studies [17]. The protocol was prospectively published (International Registered Report Identifier DERR1-10.2196/51578) [15] and registered with the Australian and New Zealand Clinical Trials Registry (ID ACTRN12622001237774).

XSL•FO
RenderX

Figure 1. Study design schema: study design schema illustrating the cross-sectional diagnostic agreement study nested within the MIS BAIR randomized controlled trial. The study compared the diagnostic agreement for dental caries detection and classification between visual examination and 2 on-screen assessment modalities (3D color models and 3D color with fluorescence models). Four trained examiners independently completed on-screen assessments at 2 time points (>4 wk apart). Visual examination outcomes served as the comparator. This image is created using BioRender [18].



Study Setting and Participants

The subjects were participants in the Melbourne Infant Study: BCG for Allergy and Infection Reduction (MIS BAIR), a phase III multicenter, randomized controlled trial [19]. The trial aimed to investigate the impact of neonatal BCG vaccination on infection and allergy outcomes in early childhood. Eligible English-speaking mothers in Victoria (2013 - 2016) consented for their newborns to join MIS BAIR. Infants with contraindications to live vaccines at birth were excluded, yielding 1272 participants. Clinical assessments were conducted at 1 year and 5 years.

Trial participants received monthly newsletters and accessed study updates via the MIS BAIR website, where families were informed of an optional dental assessment (which included intraoral scanning) at the 5-year-old follow-up. The dental component commenced in January 2021, with 536 participants remaining in the trial. Despite COVID-19 disruptions, visits continued under special approval, and 219 participants had consented to dental assessments by March 2022. All dental assessments were performed at Murdoch Children's Research Institute (Melbourne, Australia).

All primary coronal tooth surfaces were eligible for inclusion in the agreement study if they were examined during the clinical visual examination and were visible for on-screen assessment on the corresponding intraoral scan. Tooth surfaces were excluded if they were located on proximal aspects, restored, sealed, had missing or insufficient scan data, or were not visible due to other reasons.

Dental Assessment Procedure

Visual examination was performed with participants in a semisupine position, under illumination, using a dental mirror and probe with no magnification. Two trained and calibrated dental practitioners examined all teeth after removing gross plaque deposits using Puritan swabs and drying teeth with cotton wool rolls and gauze. Training procedures are described elsewhere [15]. Universal infection control standards and additional COVID-19 precautions were adhered to for all examinations.

All tooth surfaces were examined following a standard operating procedure. Any reasons for noncompletion of any parts of the dental examination were recorded. Dental caries was recorded using a merged International Caries Detection and Assessment System (ICDAS) criteria [11], and enamel defects were recorded using a modified version of the European Academy of Paediatric Dentistry index [20]. If there were multiple lesions present on the one tooth surface, the most severe of the 2 clinical presentations was recorded.

Teeth were scanned with the TRIOS 4 IOS immediately following the visual examination. The laptop was set up on a table across from the participant, where they were able to view scan acquisition during scanning (Figure 2). A scan was deemed adequate when sufficient color and fluorescence data were captured. Incomplete scans were documented, with common causes, including scanner tip access, gag reflex, or limited cooperation with scanning due to behavioral or developmental factors. Each scan was saved under a unique identifier using the TRIOS Dental Desktop software with the TRIOS module. The scanning time was calculated using the TRIOS software. The IOSs were regularly calibrated according to the

XSL•FO
RenderX

manufacturer's procedures. The manufacturer scanning strategies were followed, and external light was minimized during scanning.

Figure 2. Intraoral scanning procedure. A participant undergoing intraoral scanning as part of this research. The participant can observe the intraoral scan being constructed in real time during the procedure.

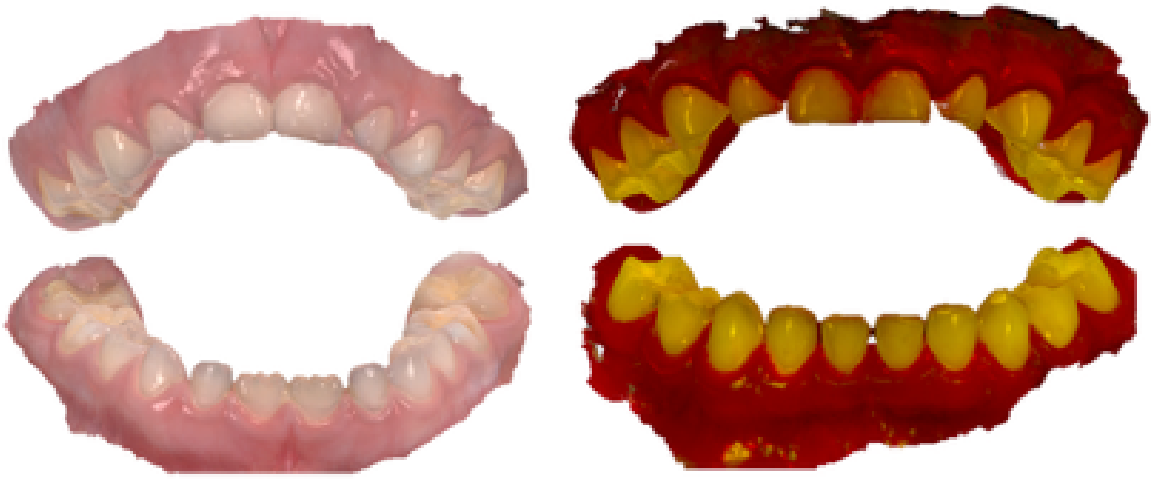
On-Screen Assessment Procedure

On-screen assessments of 3D models commenced in November 2022, 8 months after the dental assessments. Examiner training and calibration processes for on-screen assessments are detailed in the study protocol [15] and Table SA-2 in Multimedia Appendix 1). Assessments were performed on 15-inch monitors using Windows 11 in rooms with natural lighting. A custom noncommercial software (developed by 3Shape) was used to efficiently retrieve and view multiple postprocessed 3D models at resolution quality equivalent to TRIOS Dental Desktop. A systematic visual examination protocol guided all on-screen assessments. Four trained dental examiners, blinded to the clinical data and each other's ratings, independently evaluated each participant's models in 2 stages: first assessing surface geometry in approximate natural colors and then supplemented with fluorescence data (Figure 3). The presence and extent of dental carious lesions were recorded using a validated, modified ICDAS index incorporating fluorescence descriptors (Figure SA-1, Multimedia Appendix 1) [11-13,15]. The models were assessed at time point 1. All examiners repeated the on-screen assessment protocol >4 weeks later (time point 2).

All data were recorded and managed using a secure, web-based app, the REDCap (Research Electronic Data Capture) tool [21].

XSL•FO
RenderX

Figure 3. 3D models generated by an intraoral scanner. Left: Approximate natural color textures. Right: With the addition of fluorescence texture.



Outcome Measures

The primary outcome was dental caries at a tooth surface level. Three binary outcome variables were derived to represent clinically relevant disease thresholds for dental caries. Specifically, caries data were dichotomized at initial (ICDAS ≥01), moderate (ICDAS ≥03), and extensive disease thresholds (ICDAS ≥05). All other variable definitions are described in Table SA-3 (Multimedia Appendix 1).

Sample Size

The sample was drawn from all 219 eligible children who participated in the MIS BAIR dental assessment at the 5-year-old follow-up visit.

Statistical Analysis

Evaluating tooth surface-level outcomes in dental research introduces complexities to the statistical analysis and limits the use of traditional methods for examining diagnostic agreement. Whole mouth dental data have multiple levels of clustering (surfaces nested within a tooth and teeth nested within an individual), and the high prevalence of healthy tooth surfaces results in a zero-inflated data distribution.

First, to account for the clustering, this study used multilevel logistic regression to estimate the likelihood of caries detection at the tooth surface level. The likelihood of detecting dental caries between visual examination and on-screen assessment methods (with and without fluorescence) was estimated by including an indicator variable for the caries detection method as a fixed effect. Additionally, the 2 on-screen assessment methods were compared by fitting a separate model that included method, rater, and time (time point 1 or time point 2) as fixed effects. A sensitivity analysis was performed by adjusting the model for enamel defects, and an additional sensitivity analysis by surface type (smooth surfaces vs occlusal surfaces) investigated if estimates differed based on surface location.

Second, the agreement between methods for the total number of caries detected for each participant was examined using the Bland-Altman plot [22]. The Bland-Altman plot compared the differences in the total number of carious surfaces detected between methods (y-axis) against the mean of the 2 methods (x-axis) for each subject. Due to the significant overlap of points, the size of the points in the plot was adjusted to reflect the frequency of the observations. Visual examination was compared to on-screen assessment with and without fluorescence at each diagnostic threshold (initial, moderate, and extensive). This comparison was repeated for each surface type (occlusal and smooth surfaces). A best-fit regression line of the differences in methods against the mean was plotted to identify trends in differences across the range of averages.

Finally, the intraclass correlation coefficient (ICC) was calculated to estimate the concordance between visual examination and on-screen assessment methods (with and without fluorescence) for total carious surfaces detected by each method per participant. The ICC for count data was estimated from generalized linear mixed models with zero-inflated Poisson distribution to account for the excess zero observations [23]. For the primary method comparison, given the multiple examiner observations for the on-screen assessment methods, the median number of surfaces with carious lesions detected among examiners for each participant was used to compare to the visual examination. Additionally, the ICC was calculated for each combination of raters to determine the concordance (interrater reliability) for on-screen assessments (with and without fluorescence) and the repeatability (intrarater reliability) for the same rater between time point 1 and time point 2 for on-screen assessments. All analyses were performed in STATA version 18 (StataCorp) [24] and R (R Core Team) [25].

Ethical Considerations

This research was conducted with the approval of the Royal Children's Hospital Human Research Ethics Committee (RCH HREC No.88321). Written parental or guardian consent was obtained for each participant to allow inclusion in the study and the use of dental data for validating and advancing intraoral scanning technology for oral health assessment. Children

XSL•FO
RenderX

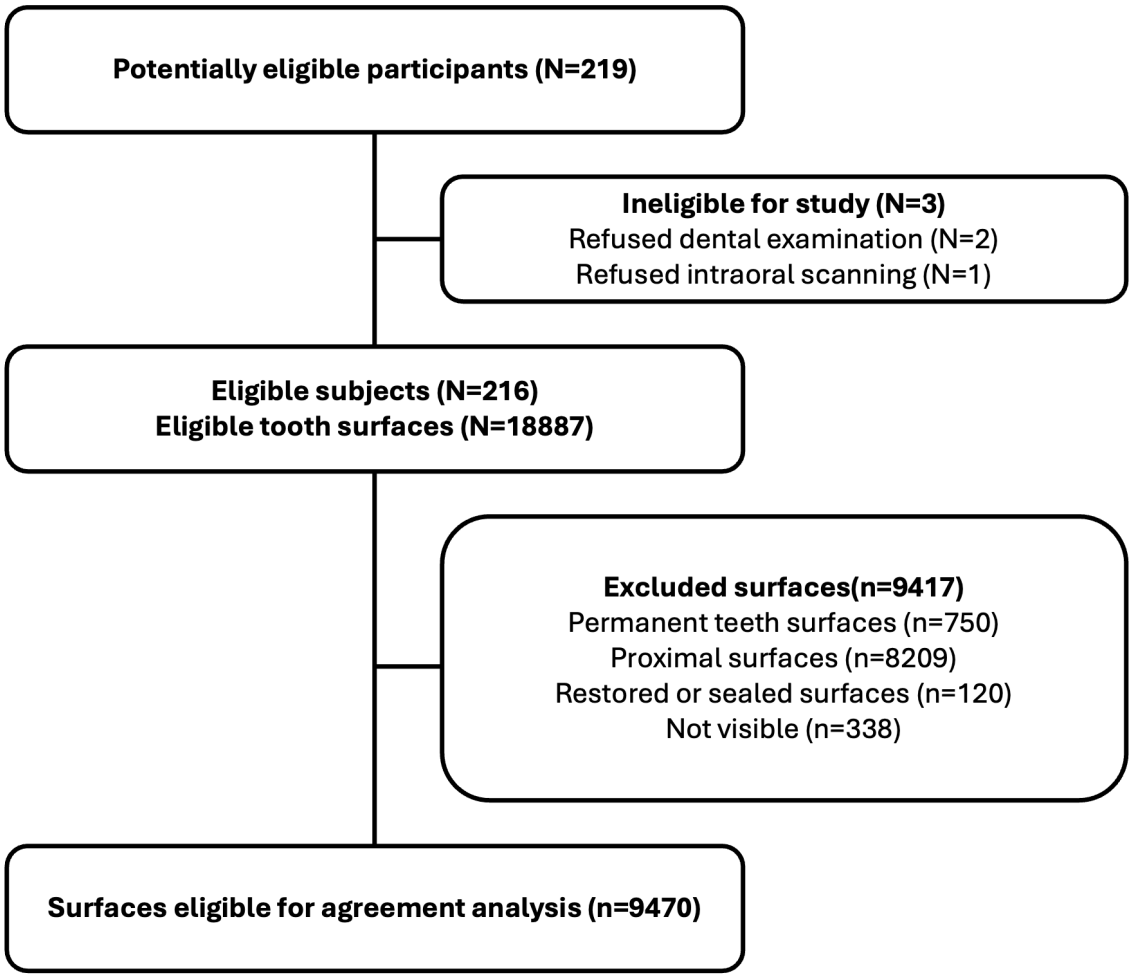
received a AUD $20 gift card (US $15) as remuneration for their participation. All participant records were assigned unique identifiers, ensuring that individual identities were not identifiable by the dental team. The study was conducted according to the principles outlined in the Declaration of Helsinki and the ICMJE recommendations. The study protocol is registered with the Australian New Zealand Clinical Trials Registry (registration number: ACTRN12622001237774p). The study protocol was published a priori with the International Registered Report Identifier DERR1-10.2196/51578. The MIS BAIR randomized controlled trial had ethical and governance approval from Mercy Health Human Research Ethics Committee (HREC; No. R12-28) and Royal Children's Hospital HREC (No. 33025).

Results

Study Flow

Of the 219 children who consented to dental assessments, 3 declined dental examinations or intraoral scanning at the time of the study visit, leaving a study population of 216. A total of 18,887 tooth surfaces were eligible for inclusion. Permanent tooth surfaces (n=750), proximal tooth surfaces (n=8209), restored or sealed tooth surfaces (n=120), and surfaces that were not visible or that could not be examined due to incomplete or missing scan data (n=338) were excluded. The remaining 9470 (50%) tooth surfaces were included in the analysis (Figure 4).

Figure 4. Study flow diagram. Flow diagram outlining patient eligibility and exclusion criteria (N=number of patients, n=number of tooth surfaces).



Sample Characteristics

Table 1 presents the characteristics of the sample. Participants' average age was 5.6 (SD 0.4) years, with 49% males and 51% females. Most children had mothers of Anglo-Celtic ancestry (n=124, 57%), followed by European (n=38, 18%) and Asian (n=33, 15%). Most children resided in the state's most advantaged areas (n=138, 64%). On the basis of visual examination data, the prevalence of dental caries at the initial disease threshold was 38%, and enamel defects were 18%.

XSL•FO
RenderX

Table . Sample characteristics (N=216).

Participant characteristics	Values
Age (y), mean (SD)	5.6 (0.4)
Sex, n (%)	
Male	106 (49.1)
Female	110 (50.9)
Ethnicity, n (%)	
Anglo-Celtic	124 (57.4)
European	38 (17.6)
Asian	33 (15.3)
Middle East	4 (1.9)
African	4 (1.9)
Aboriginal/Torres Strait Islander	2 (0.9)
Other	11 (5.1)
SEIFA[a], n (%)	
Quintile 1	5 (2.3)
Quintile 2	31 (14.4)
Quintile 3	42 (19.4)
Quintile 4	49 (22.7)
Quintile 5	89 (41.2)
Prevalence of caries	82 (37.9)
Prevalence of enamel defects	39 (18.1)

[a]Socio-economic Indexes for Areas (SEIFA) ranks areas in Australia according to relative socioeconomic disadvantage. SEIFA was divided into quintiles ranging from the most disadvantaged (quintile 1) to the least disadvantaged (quintile 5).

Examiner Characteristics

Four registered dental practitioners undertook on-screen assessments for all participants at 2 time points. They were recruited via convenience sampling. The examiners had clinical experience ranging from 5 to 15 years, including conducting epidemiological assessments (Table SA-2 in Multimedia Appendix 1).

Time Taken for Scanning and On-Screen Assessments

The mean time to perform intraoral scanning of the lower arch was 46 (SD 8) seconds and for the upper arch was 58 (SD 15) seconds. The mean time taken for the on-screen assessment of each 3D model in color supplemented with fluorescence was 3.5 (SD 2.3) minutes.

Surface-Level Agreement

The total number of carious tooth surfaces reported for each method at each diagnostic threshold and odds ratios (ORs) are presented in Table 2, Table 3, and Table 4 (Raw data stratified by rater and surface type are available in Table SA-4 and Table SA-5 Multimedia Appendix 1). The likelihood of detecting caries using on-screen assessment of 3D models in color was similar to that of visual examination at the initial (odds ratio [OR] 1.1, 95% CI 1.0 - 1.3), moderate (OR 0.9, 95% CI 0.7 - 1.1), and extensive (OR 1.0, 95% CI 0.7 - 1.3) disease thresholds. When 3D models were supplemented with fluorescence, the likelihood of detecting caries was 30% higher at the initial threshold relative to visual examination (OR 1.3, 95% CI 1.1 - 1.5) and similar at the moderate (OR 0.9, 95% CI 0.7 - 1.1) and extensive thresholds (OR 0.9, 95% CI 0.7 - 1.3). The sensitivity analysis by surface type is also presented in Table SA-6 in Multimedia Appendix 1. The likelihood of caries detection using on-screen assessment with and without fluorescence was comparable to visual examination for smooth surfaces across all disease thresholds. For occlusal surfaces, the likelihood of reporting caries with on-screen assessment methods was 1.5- to 2-fold higher at the initial disease threshold both with (OR 2.0, 95% CI 1.6 - 2.4) and without fluorescence (OR 1.5, 95% CI 1.2 - 1.9). The likelihood of reporting at the moderate and extensive thresholds on occlusal surfaces was similar for on-screen assessments and visual examinations. Comparing on-screen assessment methods reveals that adding fluorescence increases the likelihood of caries detection, and this likelihood is also influenced by the rater scoring the assessment. Accounting for enamel defects as an additional covariate in the analysis resulted in slightly lower OR estimates for both methods across all disease thresholds.

XSL•FO
RenderX

Table . Agreement estimates for visual examination (VE) versus on-screen assessment (OSA) for all included tooth surfaces (n=9470).

Disease threshold	VE		OSA		Main analysis[a]		Sensitivity analysis	
	No	Yes	No	Yes	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Initial threshold	9182	288	9154	316	1.1 (1.0-1.3)	0.08	1.1 (1.0-1.3)	0.03
Moderate threshold	9384	86	9395	75	0.9 (0.7-1.1)	0.19	0.8 (0.7-1.1)	0.17
Extensive threshold	9422	48	9424	46	1.0 (0.7-1.3)	0.78	1.0 (0.7-1.3)	0.76

[a]The main analysis, which contained the variable Method as a fixed effect, is shown in addition to the sensitivity analysis, where enamel defects at VE were fit as an additional covariate in the multilevel model.

Table . Agreement estimates for visual examination (VE) versus on-screen assessment with fluorescence (OSA-FLU) for all included tooth surfaces (n=9470).

Disease threshold	VE		OSA- FLU		Main analysis[a]		Sensitivity analysis	
	No	Yes	No	Yes	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Initial threshold	9182	288	9125	345	1.3 (1.1-1.5)	<.001	1.2 (1.1-1.4)	<.001
Moderate threshold	9384	86	9394	76	0.9 (0.7-1.1)	0.24	0.9 (0.7-1.0)	0.21
Extensive threshold	9422	48	9380	45	0.9 (0.7-1.3)	0.61	0.9 (0.7-1.3)	0.59

[a]The main analysis, which contained the variable Method as a fixed effect, is shown in addition to the sensitivity analysis, where enamel defects at VE were fit as an additional covariate in the multilevel model.

Table . Agreement estimates for on-screen assessment (OSA) versus on-screen assessment with fluorescence (OSA-FLU) for all included tooth surfaces (n=9470).

Disease threshold	OSA		OSA- FLU		Main analysis[a]		Sensitivity analysis	
	No	Yes	No	Yes	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Initial threshold	9154	316	9125	345	1.1 (1.1-1.2)	<.001	1.1 (1.1-1.2)[b]	<.001
Moderate threshold	9395	75	9394	76	1.0 (0.9-1.2)	0.77	1.0 (0.9-1.2)[b]	0.77
Extensive threshold	9424	46	9380	45	1.0 (0.8-1.1)	0.62	1.0 (0.8-1.1)	0.62

[a]The main analysis, which contained the variable Method as a fixed effect, is shown in addition to the sensitivity analysis, where enamel defects at VE were fit as an additional covariate in the multilevel model.

[b]The inclusion of a rater variable in this model had a significant effect.

Person-Level Agreement

The Bland-Altman plots indicate the absolute method agreement (Figure 5). For on-screen assessment of 3D models without (OSA) and with fluorescence (OSA-FLU) compared to visual examination, the absolute agreement was lower at the initial disease threshold and higher at the moderate and extensive disease thresholds. At the initial disease threshold, there were more negative differences compared to positive differences, where on-screen assessment recorded more carious surfaces relative to the visual examination (Figure 4A and D). At the moderate and extensive thresholds, the frequencies of positive and negative differences were comparable for both color-only and fluorescence-assisted assessments (Figure 4B, C, E, and F). The slope and intersection of the regression line, particularly at the initial caries thresholds, suggest that when average total carious surfaces were low, there were negative differences between methods. As the average total carious surfaces increased, the positive difference between methods increased (where visual examination recorded more carious surfaces relative to on-screen assessment). The subanalysis based on surface type shows similar trends (Figure SA-2 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 5. Bland-Altman plots of the differences between the methods for all tooth surfaces. Bland-Altman plots comparing the differences in the number of carious lesions detected for each pair of methods against the average of the 2 methods across the initial, moderate, and extensive disease thresholds. The regression line of the differences in methods against the mean is represented by a red dash. The size of the data points is proportional to the number of observations. (A) Initial threshold, OSA versus VE: absolute agreement n=67 (31%), negative differences n=108 (50%), positive differences n=41 (19%). (B) Moderate threshold, OSA versus VE: absolute agreement n=170 (79%), negative differences n=26 (12%), and positive difference n=20 (9%). (C) Extensive threshold, OSA versus VE: absolute agreement n=192 (89%), negative differences n=14 (6%), positive difference n=10 (5%). (D) Initial threshold, OSA-FLU versus VE: absolute agreement n=64 (30%), negative differences n=112 (52%), positive differences n=40 (18%). (E) Moderate threshold, OSA-FLU versus VE: absolute agreement n=167 (77%), negative differences n=28 (13%), positive differences n=21 (10%). (F) Extensive threshold, OSA-FLU versus VE: absolute agreement n=192 (89%), negative differences n=14 (6%), positive difference n=10 (5%). VE: visual examination; OSA: on-screen assessment in color only; OSA FLU: on-screen assessment with fluorescence.

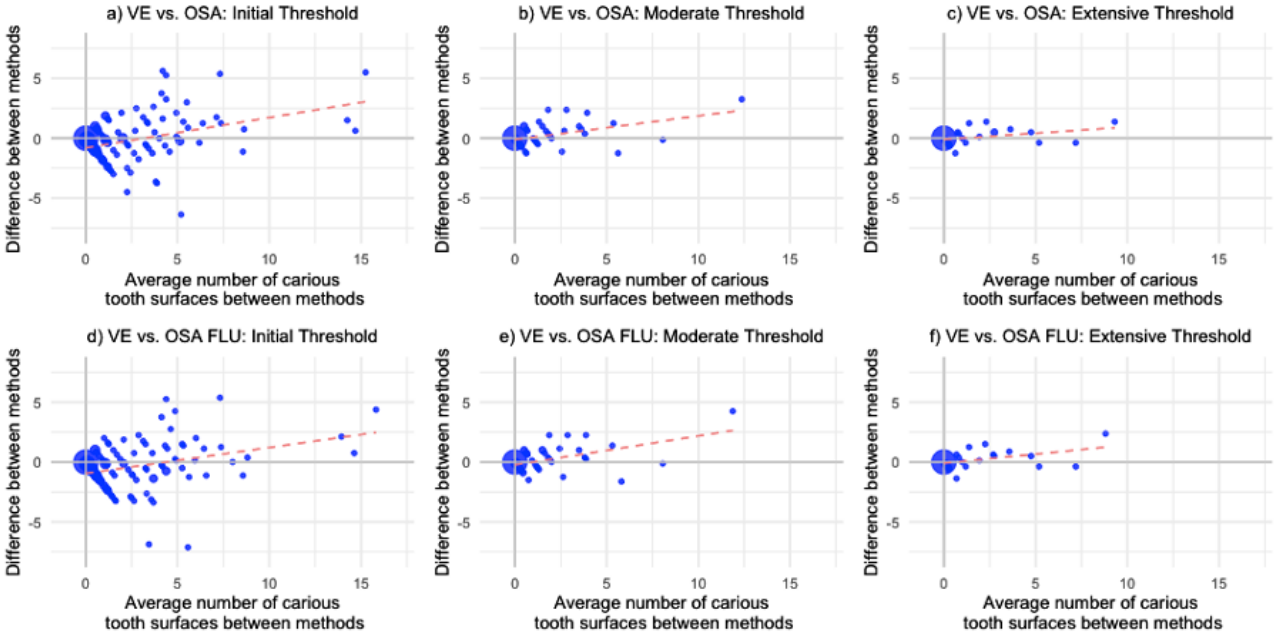


Table 5 presents the ICC values as a measure of method agreement and examiner reliability. On-screen assessment in color (ICC 0.9, 95% CI 0.9 - 0.9) and with fluorescence (ICC 0.9, 95% CI 0.8 - 0.9) showed good to excellent agreement with visual examination. For all examiners, the intrarater agreement of ICC of approximately >0.9 indicated a very high degree of scoring consistency within raters for on-screen assessment methods. The interrater reliability was also high between all raters (>0.9), indicating high-scoring similarity between raters. The narrow CIs further support the reliability of these estimates.

Table . Intraclass correlation coefficients (ICC). The ICC was used to estimate the absolute agreement between methods, intraexaminer reliability for on-screen assessment methods, and interrater agreement reliability for on-screen assessment methods.

	On-screen assessment with color, ICC (95% CI)	On-screen assessment with color and fluorescence, ICC (95% CI)
Agreement between methods		
Visual examination	0.9 (0.9-1.0)	0.9 (0.8-1.0)
Intrarater reliability		
Rater 1	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 2	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 3	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 4	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Interrater reliability		
Rater 1 versus rater 2	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 1 versus rater 3	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 1 versus rater 4	0.9 (0.8-0.9)	0.9 (0.8-0.9)
Rater 2 versus rater 3	0.9 (0.9-1.0)	0.9 (0.9-1.0)
Rater 2 versus rater 4	0.9 (0.8-0.9)	0.9 (0.9-0.9)
Rater 3 versus rater 4	0.9 (0.8-0.9)	0.9 (0.8-0.9)

XSL•FO
RenderX

Discussion

Principal Findings

This study is the first to describe the real-world performance of IOSs using light-induced fluorescence for digital caries assessments of primary teeth in children. The findings suggest that the on-screen assessment of 3D models in color supplemented with fluorescence had a high level of agreement with visual examination for detecting primary coronal carious lesions on smooth and occlusal surface types at the moderate and extensive disease thresholds. At the initial disease thresholds, on-screen assessments resulted in a higher reporting of occlusal caries relative to visual examination, and supplementing on-screen color assessments with fluorescence increased the likelihood of dental caries detection compared to visual examination. The high ICC values provide further evidence of method agreement and indicate that the on-screen assessment criteria can be applied consistently and are reproducible.

Our findings are consistent with data from previous studies in adults. Porumb et al [26] observed high rater reliability and agreement between visual examination using ICDAS and on-screen assessment of 3D color models in adults when experienced examiners were used. Building on the work of a small pilot study that suggested remote assessment of 3D models in adults had good accuracy for cavity detection [6], our analysis shows that when applying a validated scoring index with trained examiners, assessment of intraoral scans has substantial agreement with visual examination. Our study results also corroborate the findings of several studies investigating on-screen assessments for permanent occlusal surfaces [12,13], showing that on-screen assessments had similar diagnostic performance to visual examinations. However, in contrast to these studies, our findings at the initial disease threshold, showed adding fluorescence increased the likelihood of reporting dental caries for occlusal surfaces.

The results of our study align with findings from prior research using intraoral cameras for dental caries data collection in pediatric epidemiological settings. In 5-year-old children, Boye et al [27,28] found the on-screen assessment of 2D images produced with intraoral cameras was equivalent to visual examination, although they used advanced disease thresholds (cavity involving dentin) for caries reporting. A study by Mehta et al [29] involving Australian 5-year-old children found that the on-screen assessment of 2D images resulted in significantly greater reporting of dental caries compared to visual examination when initial disease thresholds were applied, which the authors attributed to the improved lighting and examination conditions when performing on-screen assessments.

In the absence of a gold standard reference, it is not possible to determine whether the observed differences between methods reflect greater accuracy of one method over the other. The higher likelihood of reporting dental caries when on-screen assessment was supplemented with fluorescence could be attributable to several factors. It is plausible that on-screen assessment is more sensitive to early mineral loss than visual examination due to enhanced detection capabilities of light-induced fluorescence [30]. However, the apparent increase in detection may also reflect a reduction in specificity, as red fluorescence changes may be attributed to the presence of plaque leading to false-positive indications, particularly for occlusal surfaces, which often have plaque retentive pits and fissures [31]. The latter is consistent with findings investigating caries detection using 2D image-based ICDAS methods, which have reported lower specificity at the initial disease threshold [32]. The inherent subjectivity of early occlusal caries detection may also have contributed to the higher disagreements observed in our study at the initial disease threshold. Additionally, the creamy-colored demarcations associated with hypomineralized enamel could be misinterpreted as caries during the on-screen assessment, consistent with the reduced odds ratio estimates when enamel defects were included as a covariate in our sensitivity analysis.

Strengths

The use of intraoral scanning for data collection in research offers several advantages. First, it enables examiner blinding to participant characteristics and allows repeat examinations, both of which improve internal consistency and reliability. Intraoral scanning may also facilitate more comprehensive dental assessments for children who can tolerate only brief or limited inspections. The resulting 3D models can be enlarged, rotated, and reviewed for extended periods, providing access to areas that may be difficult to view during visual examination alone. Furthermore, 3D models can serve as effective communication tools. A randomized controlled trial reported that 3D models as a visual aid significantly improved parents' understanding of their children's dental findings and treatment plans in pediatric dentistry compared to verbal explanations alone [33]. Finally, the data acquisition and on-screen assessment times achieved in this study were shorter than those reported in previous research using intraoral cameras for epidemiological research, which ranged from 6 - to 8.8 minutes for data acquisition per participant [28].

Limitations

We acknowledge several limitations. First, the time taken to undertake each dental examination was not recorded for all participants. Future studies should capture this, as it helps contextualize the clinical relevance. Second, the examination site lacked compressed air, and teeth were not professionally cleaned before examination and scanning. Third, unlike on-screen assessment, we did not undertake repeated visual examinations by multiple examiners. This is a potential source of bias, as it impacts the reliability of the visual examination as the comparator. Fourth, the highly selected sample potentially limits the broader generalizability of the findings. Additionally, a larger sample would increase the precision around the odds ratio estimates. Fifth, the study used an agreement design, which, although appropriate for this investigation, limits interpretation to relative rather than absolute diagnostic performance, as no gold standard reference was available. Sixth, dental caries vary in clinical presentation on proximal surfaces and around restorations or sealants, which were excluded from the analysis due to a lack of radiographic imaging in the study design. Given the age and disease experience of the cohort, it is not known

XSL•FO
RenderX

how missing teeth, widened interdental spaces, recession, or heavily restored teeth may impact on-screen assessment performance. Further validation research in different population groups with varying levels of disease is warranted. Finally, the nonnormal distribution of the dataset precluded the calculation of Bland-Altman Limits of Agreement, which define an acceptable disagreement range among methods. Data transformations to address this lacked clinical interpretability. Clinically acceptable limits of agreement were therefore not defined, as establishing these would require a consensus-based approach, which was beyond the scope of this study.

Future Research Directions

A number of practical considerations must be addressed before intraoral scanning can be implemented at scale in epidemiological or public health programs. These include assessing the resource requirements, examiner training needs, and workflow integration within existing service models. Feasibility will also depend on whether the technology is valid and reliable for the specific disease outcomes being evaluated, as different conditions may require tailored scanning protocols or interpretation methods. Additionally, the acceptability, cost-effectiveness, and environmental impact of integrating intraoral scanning into outreach, screening, and teledentistry programs should be evaluated. Research exploring the potential for trained nondental personnel to collect scan data could further support the scalability and sustainability of this approach. Furthermore, investigation into methods that automate caries detection using intraoral scan data is necessary to reduce the time and potential costs associated with using registered dental practitioners for on-screen assessments. This could be achieved through the validation of existing tools [7,34,35] or the development of new artificial intelligence methods [36]. In epidemiological settings, digital data collection combined with automation may potentially improve the frequency of population disease surveillance programs, where data-driven insights into regional oral health trends could be used to inform resource allocation and health promotion efforts.

Conclusions

This study found that the on-screen assessment of 3D models supplemented with fluorescence substantially agreed with visual examination for caries detection at the moderate and extensive disease thresholds. Whether the on-screen assessment of 3D models generated with IOSs can be used interchangeably with visual examination for caries detection in children is context dependent and likely influenced by the caries threshold used and the tooth surfaces examined. These findings suggest that digital assessments may provide a feasible alternative to traditional in-person examinations for research, longitudinal monitoring, and population-level surveillance, especially where examiner access is limited or standardization across examiners and sites is required.

Acknowledgments

The authors would like to thank all participants and families of the MIS BAIR trial. The authors would also like to thank Jenn Copley, Aneeshya Nidhin, Aldia Hong, and Delia Chin for their support in this project.

Funding

The MIS BAIR trial was supported by the National Health and Medical Research Council of Australia (grant GNT1051228 and GNT1099680). NC is supported by a National Health and Medical Research Council Investigator Grant (GNT1197117). MS is funded by a Melbourne Children's Clinician-Scientist Fellowship. DB is supported by the Australian National Health and Medical Research Council Investigator Grant. The Victorian Government's Operational Infrastructure Support Program supports research at Murdoch Children's Research Institute.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to consent not obtained from participants for public sharing of data but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: BJ (lead), MS
Data curation: BJ (lead), MS (supporting), NC (supporting)
Formal analysis: BJ (lead) TC (supporting)
Funding acquisition: CV,
Investigation: BJ (lead)
Methodology: BJ (lead), TC (equal)
Project administration: BJ, MS
Resources: SM, CV
Supervision: MS (lead) NK, DB, NC
Validation: BJ Visualization: BJ, TC
Writing – original draft: BJ (lead)
Writing – review & editing: BJ, TC, SM, NK, DB, NC, CV

XSL•FO
RenderX

Conflicts of Interest

BJ reports the company 3/Shape TRIOS A/S Denmark funded her PhD stipend through an industrial agreement with Murdoch Children's Research Institute. CV and SM were employed at 3Shape when this research was undertaken. This project was funded through a Research Collaboration between 3Shape TRIOS A/S Denmark and the Murdoch Children's Research Institute, in which MS is the lead investigator. Other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Multimedia Appendix 1
Additional tables and figures provided to support the findings reported in this manuscript.
[DOCX File, 655 KB - publichealth_v11i1e78023_app1.docx]

References

1. Pitts NB, Zero DT, Marsh PD, et al. Dental caries. Nat Rev Dis Primers 2017 May 25;3(1):17030. [doi: 10.1038/nrdp.2017.30] [Medline: 28540937]
2. Global Oral Health Status Report: Towards Universal Health Coverage for Oral Health by 2030: WHO; 2022. URL: https://www.who.int/publications/i/item/9789240061484 [accessed 2025-11-25]
3. Frencken JE. Dental caries and caries epidemiology. In: Evidence-Based Caries Prevention: Springer, Cham; 2025:1-21. [doi: 10.1007/978-3-031-79152-9_1]
4. Nelson S, Eggertsson H, Powell B, et al. Dental examiners consistency in applying the ICDAS criteria for a caries prevention community trial. Community Dent Health 2011 Sep;28(3):238-242. [Medline: 21916361]
5. Silva AER, Menezes AMB, Assunção MCF, et al. Validation of self-reported information on dental caries in a birth cohort at 18 years of age. PLoS ONE 2014;9(9):e106382. [doi: 10.1371/journal.pone.0106382] [Medline: 25202969]
6. Steinmeier S, Wiedemeier D, Hämmerle CHF, Mühlemann S. Accuracy of remote diagnoses using intraoral scans captured in approximate true color: a pilot and validation study in teledentistry. BMC Oral Health 2020 Sep 25;20(1):266. [doi: 10.1186/s12903-020-01255-8] [Medline: 32977794]
7. Michou S, Benetti AR, Vannahme C, Hermannsson PG, Bakhshandeh A, Ekstrand KR. Development of a fluorescence-based caries scoring system for an intraoral scanner: an in vitro study. Caries Res 2020;54(4):324-335. [doi: 10.1159/000509925] [Medline: 33053552]
8. Karlsson L. Caries detection methods based on changes in optical properties between healthy and carious tissue. Int J Dent 2010;2010:270729. [doi: 10.1155/2010/270729] [Medline: 20454579]
9. de Josselin de Jong E, Sundström F, Westerling H, Tranaeus S, ten Bosch JJ, Angmar-Månsson B. A new method for in vivo quantification of changes in initial enamel caries with laser fluorescence. Caries Res 1995;29(1):2-7. [doi: 10.1159/000262032] [Medline: 7867045]
10. Criteria manual – international caries detection and assessment system (ICDAS II). : International Caries Detection and Assessment System (ICDAS) Coordinating Committee; 2005 URL: https://www.iccms-web.com/uploads/asset/592848be55d87564970232.pdf [accessed 2025-11-25]
11. Ferreira Zandona A, Santiago E, Eckert G, Fontana M, Ando M, Zero DT. Use of ICDAS combined with fluorescence for caries detection. Caries Res 2010;44(3):317-322. [doi: 10.1159/000317294] [Medline: 20588022]
12. Sá G, Michou S, Bönecker M, Mendes F, Amarante B, Ekstrand K. Diagnostic validity of ICDAS clinical criteria on digital 3D models. J Dent 2024 Oct;149:105274. [doi: 10.1016/j.jdent.2024.105274] [Medline: 39084547]
13. Ntovas P, Michou S, Benetti AR, et al. Occlusal caries detection on 3D models obtained with an intraoral scanner. A validation study. J Dent 2023 Apr;131:104457. [doi: 10.1016/j.jdent.2023.104457] [Medline: 36858167]
14. Lygidakis NA, Garot E, Somani C, Taylor GD, Rouas P, Wong FSL. Best clinical practice guidance for clinicians dealing with children presenting with molar-incisor-hypomineralisation (MIH): an updated European Academy of Paediatric Dentistry policy document. Eur Arch Paediatr Dent 2022 Feb;23(1):3-21. [doi: 10.1007/s40368-021-00668-5] [Medline: 34669177]
15. Jones B, Michou S, Chen T, et al. Caries detection in primary teeth using intraoral scanners featuring fluorescence: protocol for a diagnostic agreement study. JMIR Res Protoc 2023 Dec 14;12:e51578. [doi: 10.2196/51578] [Medline: 38096003]
16. Kottner J, Audigé L, Brorson S, et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011 Jan;64(1):96-106. [doi: 10.1016/j.jclinepi.2010.03.002] [Medline: 21130355]
17. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open 2016 Nov 14;6(11):e012799. [doi: 10.1136/bmjopen-2016-012799] [Medline: 28137831]
18. BioRender. URL: https://BioRender.com [accessed 2025-12-04]
19. Messina NL, Gardiner K, Donath S, et al. Study protocol for the Melbourne Infant Study: BCG for Allergy and Infection Reduction (MIS BAIR), a randomised controlled trial. BMJ Open 2019;9(12). [doi: 10.1136/bmjopen-2019-032844]
20. Ghanim A, Silva MJ, Elfrink MEC, et al. Molar incisor hypomineralisation (MIH) training manual. Eur Arch Paediatr Dent 2017;18(4):225-242. [doi: 10.1007/s40368-017-0293-9] [Medline: 28721667]

XSL•FO
RenderX

21. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building a software platform partnership. J Biomed Inform 2019;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
22. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999 Jun;8(2):135-160. [doi: 10.1177/096228029900800204] [Medline: 10501650]
23. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
24. Stata Statistical Software: Release 18. StataCorp. 2023. URL: https://www.stata.com/ [accessed 2023-12-11]
25. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2023. URL: https://www.R-project.org/ [accessed 2025-11-25]
26. Porumb I, Leucuta DC, Nigoghossian M, et al. Caries lesion assessment using 3D virtual models. Medicina (Kaunas) 2023;59(12):2157. [doi: 10.3390/medicina59122157] [Medline: 37241175]
27. Boye U, Pretty IA, Tickle M, Walsh T. Comparison of caries detection methods using intra-oral digital photographs and visual examination. BMC Oral Health 2013;13(1):6. [doi: 10.1186/1472-6831-13-6] [Medline: 23312001]
28. Boye U, Willasey A, Walsh T, Tickle M, Pretty IA. Comparison of photographic vs visual caries assessments. Community Dent Oral Epidemiol 2013;41(6):526-533. [doi: 10.1111/cdoe.12049]
29. Mehta P, Byun R, Blinkhorn AS. Use of intra-oral camera for caries research. J Dent Health Oral Disord Ther 2017;8. [doi: 10.15406/jdhodt.2017.08.00302]
30. Macey R, Walsh T, Riley P, et al. Fluorescence devices for the detection of dental caries. Cochrane Database Syst Rev 2020 Dec 8;12(12):CD013811. [doi: 10.1002/14651858.CD013811] [Medline: 33319353]
31. Mendes FM, Hissadomi M, Imparato JCP. Effects of drying time and the presence of plaque on the in vitro performance of laser fluorescence in occlusal caries of primary teeth. Caries Res 2004;38(2):104-108. [doi: 10.1159/000075933] [Medline: 14767166]
32. Carvalho RN, Letieri ADS, Vieira TI, et al. Accuracy of visual and image-based ICDAS vs micro-CT. Braz Oral Res 2018;32:e60. [doi: 10.1590/1807-3107bor-2018.vol32.0060]
33. Schulz-Weidner N, Gruber M, Schraml EM, et al. Intraoral scans to aid communication in pediatric dentistry. Dent J 2024;12(1):15. [doi: 10.3390/dj12010015]
34. Michou S, Lambach MS, Ntovas P, et al. Automated caries detection in vivo using a 3D intraoral scanner. Sci Rep 2021 Oct 28;11(1):21276. [doi: 10.1038/s41598-021-00259-w] [Medline: 34711853]
35. Jones B, Chen T, Michou S, et al. Diagnostic agreement between visual examination and an automated scanner system with fluorescence for detecting and classifying occlusal carious lesions in primary teeth. J Dent 2024 Oct;149:105279. [doi: 10.1016/j.jdent.2024.105279] [Medline: 39121599]
36. Jones B, Lambach M, Chen T, et al. Dental caries detection in children using intraoral scans and deep learning. J Dent 2025 Sep;160:105906. [doi: 10.1016/j.jdent.2025.105906] [Medline: 40527440]

Abbreviations

HREC: Human Research Ethics Committee
ICC: Intraclass correlation coefficient
ICDAS: International Caries Detection and Assessment System
IOS: intraoral scanner
MIS BAIR: Melbourne Infant Study: BCG for Allergy and Infection Reduction
OR: odds ratio
REDCap: Research Electronic Data Capture

Edited by A Mavragani, T Sanchez; submitted 24.05.25; peer-reviewed by C Areias, IM Tew; revised version received 03.11.25; accepted 03.11.25; published 08.12.25.

Please cite as:
Jones B, Chen T, Michou S, Kilpatrick N, Curtis N, Burgner DP, Vannahme C, Silva M
Dental Caries Detection in Children Using Intraoral Scanners Featuring Fluorescence: Diagnostic Agreement Study
JMIR Public Health Surveill 2025;11:e78023
URL: https://publichealth.jmir.org/2025/1/e78023
doi:10.2196/78023

© Bree Jones, Tong Chen, Stavroula Michou, Nicky Kilpatrick, Nigel Curtis, David P Burgner, Christoph Vannahme, Mihiri Silva. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 8.12.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/),

XSL•FO
RenderX

which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Characterizing Cross-Provincial High-Cost Patients in Rural China: Cross-Sectional Study

Minjiang Guo[1*], PhD; Xiaotong Jiang[1*], MS; Yang Liu[1], MS; Yang Liu[1], MS; Fangyuan Zhang[1], PhD; Zhongyuan Zhang[1], MS; Yazi Li[2], PhD

[1]Health Insurance Information Research Department, Institute of Medical Information, Chinese Academy of Medical Sciences, Beijing, China

[2]Information Technology Office, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 9 Dongdan Santiao, Dongcheng District, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Yazi Li, PhD
Information Technology Office, Chinese Academy of Medical Sciences & Peking Union Medical College, No. 9 Dongdan Santiao, Dongcheng District, Beijing, China

Abstract

Background: High-cost (HC) patients, typically defined as the top 10% or 5% of patients with the highest health care costs, are responsible for over half of all health care–related spending. In China, approximately 95% of rural residents are covered by Urban and Rural Resident Basic Medical Insurance. In parallel, increasing population mobility has made it more common for rural residents to seek medical treatment and claim reimbursements across provincial boundaries. These trends underscore the importance of identifying and understanding HC patients within this group.

Objective: This study aimed to analyze the characteristics and risk factors associated with HC cross-provincial insured patients in rural China.

Methods: The study used data from the cross-provincial medical immediate reimbursement system, which contains records of inpatients who used cross-province immediate reimbursement services between 2017 and 2019. Patients whose total annual medical expenditure ranked within the top 10% of all cross-provincial inpatients were classified as HC patients. Andersen's Behavioral Model of Health Services Use was adopted to examine the factors associated with being an HC patient. Descriptive statistics and multivariable logistic regression model analyses were performed.

Results: A total of 2987 patients were included, with a mean age of 42.99 (SD 19.39) years. Males comprised 57.4% (1713/2987) of the total. Among all cross-provincial patients, the expenses of HC patients made up 34.5% of total expenses. The average annual hospitalization cost per HC patient was US $22,460. Results from multivariable logistic regression analysis indicated that male patients (odds ratio [OR] 1.38, 95% CI 1.06-1.79; *P*=.01), individuals with multiple comorbidities (OR 3.62, 95% CI 2.37-5.53; *P*<.001), those diagnosed with cancer (OR 2.31, 95% CI 1.61-3.31; *P*<.001), and patients receiving care at specialized hospitals (OR 1.61, 95% CI 1.24-2.08; *P*<.001) were significantly associated with HC status.

Conclusions: Cross-provincial HC patients in rural China exhibited a lower concentration of total expenditure but incurred higher average annual hospitalization costs compared with local patients. This finding suggests the presence of potential cost-driving factors within this group. Identified risk factors—including sex, comorbidity status, cancer diagnosis, and hospital type—may inform the development of more equitable and efficient health financing policies, such as optimizing resource allocation and designing targeted interventions for HC patients.

(JMIR Public Health Surveill 2025;11:e54234) doi:10.2196/54234

KEYWORDS

high-cost patients; cross-provincial; rural patients; China; inpatient

Introduction

By 2010, China had achieved near-universal health coverage [1], and the population covered by the basic medical insurance (BMI) system has remained above 95% since 2011 [2]. The covered population reached 1.35 billion by the end of 2022 through 2 schemes [3], Urban Employee Basic Medical Insurance for formally employed workers and Urban and Rural Resident Basic Medical Insurance (URRBMI) for unemployed residents. However, due to the constantly deepening aging population and the deficiencies in the tiered health care system, BMI has been facing ever-increasing financial pressures in recent years [4]. The total expenditure of the BMI fund reached

XSL•FO
RenderX

US $390 billion [5] in 2023, increasing 14.4% compared with 2022 [6], almost twice the growth rate of annual revenue (7.9%).

Along with the increasing trend, a unique but nonnegligible subgroup population who are facing substantial health care expense pressure is domestic migrants [7]. Due to the unbalanced development states across different regions of China, domestic migration has been increasing over the past 20 years, accounting for one-fourth of the Chinese population in 2020 [8], with 66.2% (249/376 million) being rural-to-urban migrants [9]. As the common migration route is from undeveloped areas to developed ones, it will be inevitable to trigger the increase of the health care cost for them. Besides, unbalanced distribution of health care resources usually accompanies the economy, which would be another "pull" for cross-regional health care and lead to higher health care expenses again. In 2021, China's cross-provincial patients' average hospitalization expenses were 2.17 times those of local-insured patients [10]. The rapidly rising medical costs have not only become a heavy burden for the migrant population but also pose great pressure on the medical insurance funds they participate in.

Identifying the characteristics of high-cost (HC) patients could help clearly establish priority actions for policy makers in the context of allocation of scarce resources and provide effective solutions to these challenges [11]. HC patients, usually defined as the top 10% [12] or 5% [4] of patients with the highest health care costs, are responsible for more than half of all health care–related spending [13,14]. Previous studies have examined the characteristics of HC patients, findings that 3 factors may be associated with HC patients [15]. The first is demographic characteristics, like age [16,17], gender [18,19], ethnicity [20,21], region [22,23], and so on. HC patients were generally older, but there is no consensus on gender or ethnic categories. The second is the financial status, including income and health insurance statuses. Several studies [24,25] have indicated that the cost reduction provided by insurance for patients is associated with an increased likelihood of beneficiaries becoming HC patients. Income was not significantly associated with high health care costs. In high-income countries, such as the United States, individuals with higher incomes are more likely to become HC patients, potentially due to their greater purchasing power and access to services [26]. In contrast, in low- and middle-income countries, such as China, individuals with lower incomes are more prone to severe illnesses, which increases their likelihood of incurring high health care costs [27]. The third is medical characteristics. The prevalence of chronic diseases and multimorbidity were dominant among HC patients [16,28,29]. Similar findings have been reported in studies conducted in China. For instance, research conducted in rural Dangyang City revealed that HC patients were older, more likely to belong to economically disadvantaged families, and exhibited a higher prevalence of chronic conditions, including tumors, chronic obstructive pulmonary disease, and cardiovascular diseases [27]. Another study conducted in rural Macheng City reported that female residents (12,320/23,922, 51.5%) and individuals aged more than 60 years (8240/23,922, 34.5%) with diseases that are challenging to diagnose exhibited a higher propensity to incur substantial medical costs [4].

Migration and displacement are significant factors that affect the health of individuals and hinder the achievement of health-related sustainable development goals. Addressing the health care needs of the migration population is of great significance in China. However, little is known about the characteristics of HC patients among domestic migrants. Analyzing the characteristics of HC patients within this population will offer critical insights for developing an inclusive health care system in their host cities. Meanwhile, from the perspective of patients, cross-regional medical treatment typically arises when the health care resources in their origin cities are inadequate to meet their medical needs, it will reflect deficiencies of health care resources of their flow-out cities, thereby facilitating more informed guidance on the allocation and utilization of health services. Compared with urban residents, rural-urban migrants faced greater financial constraints, and expected to be more vulnerable when using health care services.

This study aimed to examine the characteristics and risk factors associated with HC cross-provincial insured patients from rural China. First, we analyze the medical cost agglomeration characteristics of the rural-to-urban migrants. Second, we analyze the demographic, socioeconomic, health care utilization, and clinical characteristics of HC patients and non–high-cost (NHC) patients. Third, we adopt Andersen's Behavioral Model of Health Services Use (AMHSM) as a theoretical framework (refer to Methods section) to identify the factors associated with HC patients among this population.

Methods

Design and Setting

We conducted a cross-sectional study using claims data from Hainan province. Hainan has a large outflow population. The seventh national census data from 2021 showed nearly 10% of Hainan's locals live in other provinces [30], higher than the national average [31] (124,837,153/1,411,778,724, 8.8%). There are 2 possible reasons for this phenomenon. First, Hainan borders Guangdong, which is considered China's economic powerhouse. Consequently, a significant number of the rural labor force from Hainan is inevitably drawn to Guangdong Province for employment, leading to increased demand for cross-province medical reimbursements. Second, medical resources in Hainan Province are relatively weak. According to the 2022 Health Statistics Yearbook [30], the number of licensed physicians in Hainan Province per 1000 people ranks 27th in China. The number of beds in medical institutions per 1000 population ranks 24th. With improvements in transportation and the increasing accessibility of medical insurance reimbursement, an increasing number of local patients travel to other provinces for better health care.

Theoretical Model

AMHSM is the best theoretical model for studying health service utilization behavior [32]. In this research, we applied the AMHSM to examine the utilization of health services and the factors influencing it among the HC patients. This framework was validated and used in previous work on HC patients [33]. According to AMHSM, health services use was determined by

XSL•FO
RenderX

(1) existing characteristics that predispose people to use or not use services even though these characteristics are not directly responsible for the use, like age, gender, and so on; (2) enabling characteristics that facilitate or impede the use of services, like income, insurance, and so on; and (3) need or condition, that refers to the level and type of experienced illness [15]. Furthermore, Anderson's 2007 study [32] suggested that, in addition to individual factors, environmental factors should also be considered. In previous research, environmental factors were not considered primarily because the studied populations were situated within the same health care system environment. However, in the context of this study, the provinces and health care institutions into which the migrant population moves exhibit distinct environmental characteristics that may influence the level of health care utilization. Therefore, these factors were incorporated into the analysis.

Data and Variables

All patients who utilized cross-provincial inpatient services and received immediate reimbursement were included. We collected data from Cross-provincial Medical Immediate Reimbursement System, which is a subsystem of URRBMI. The system was established in 2017, and it records the information related to cross-provincial medical reimbursement, including demographic characteristics, expenditure information, and clinical characteristics, among other variables (Table 1).

Table . Key variables in cross-provincial medical immediate reimbursement system.

Variables	Definition
Identification code	A unique identifier for an anonymized patient.
Sex	Male or female.
Age	Patient's age in years.
Migration type	Migration type is a categorical variable, including illness-related referrals, working in other provinces, and living in other provinces.
Flow out city	The city (country or district level) where cross-provincial patients originate.
Flow in province	The province to which cross-provincial patients migrate for health care services.
Flow in hospital	The hospital where the cross-provincial patients received treatment.
Length of stay	The duration of hospitalization, calculated as discharge day minus admission day.
Medical cost per visit	Total medical costs for a single patient visit.
Reimbursed expenses per visit	Expenses reimbursed by health insurance per patient visit.
Out-of-pocket expenses per visit	Out-of-pocket costs per patient visit.
Diagnosis code	Coding according to *ICD-10*[a] system. Only primary diagnoses were recorded.
Diagnosis name	Naming according to *ICD-10* system. Only primary diagnoses were recorded.

[a]*ICD-10*: *International Classification of Diseases, Tenth Revision*.

Independent Variables

Variables were categorized according to the AMHSM as follows. First, demographic characteristics, including age and gender. In addition, for the migrant population, we included the migration type as an independent variable. In China, patients seeking medical care outside their place of residence can generally be categorized into 2 types [34]. One group consists of individuals who have relocated to another province for occupational or personal reasons and subsequently required medical care. We defined them as migration-driven cross-province patients. The second group comprises individuals who sought better medical care outside their home province. We classify them as disease-driven cross-province patients. Logically, different types of patients exhibit distinct characteristics in terms of medical service utilization. Second, enabling characteristics, including income and insurance status. As the database did not include individual patient income information, the per capita disposable income of the patient's insured county or city was used as a proxy. Regarding insurance status, like most studies using health insurance databases as data sources, this study included only insured patients, making it difficult to analyze whether insurance status is a predictor of HC patients. Third, need characteristics, including the diagnosis of disease and the number of diseases. The diagnosis was identified based on the *International Classification of Diseases, Tenth Revision* system [35]. The number of diseases refers to the number of primary diagnoses confirmed for a patient within a calendar year. Fourth, environmental characteristics, including regions and type of medical institutions where patients receive medical reimbursements. Regions were grouped according to the administrative provinces of China, and the top four provinces with the highest number of medical visits were listed separately, whereas the remaining provinces were uniformly classified as other areas. According to the "Law of the People's Republic of China on Basic Medical and Health Care and the Promotion of Health", health care institutions are classified into general

XSL•FO
RenderX

hospitals and specialized hospitals. The variables and their assignments are illustrated in Table 2.

Table . Independent variables.

Variables		Code or values
Demographic characteristics		
	Age	Years
	Sex	Male=1, female=0
	Type of patients	Migration-driven cross-province patient=1, disease-driven cross-province patient=0
Enabling characteristics		
	Income	Continuous variables
Need characteristics		
	Number of diseases	Discrete variables
	History of diagnosis	yes=1, no=0
Environmental characteristics		
	Regions	Guangdong, Guangxi, Shanghai, Beijing, Other
	Medical institutions	Hospitals=0, specialized hospitals=1

Dependent Variables

The dependent variable was a binary indicator, where a value of 1 represented that HC had been incurred. According to the report entitled "Effective Care for High-Need Patients: Opportunities for Improving Outcomes, Value, and Health," published by the US National Academy of Medicine in 2017 [12], HC patients were defined as those whose total medical costs ranked among the top 10% of all cross-provincial patients within a calendar year. These costs included both out-of-pocket expenses and medical insurance reimbursement.

Statistical Analysis

First, baseline descriptive statistics were calculated for HC patients, NHC patients, and overall patients. Counting data are expressed as relative numbers, and the chi-square test or Fisher exact probability method was used for intergroup comparisons. Measurement data conforming to a normal distribution are expressed as (x-±s), and the least significant difference *t* test was used for intergroup comparisons. Measurement data not conforming to a normal distribution are expressed as median (IQR), and the Mann-Whitney *U* test was used for intergroup comparisons.

Second, we performed univariable and multivariable logistics regression analyses to identify the predictive risk factors of HC patients. Statistically significant variables ($P<.05$) in the univariable analysis were introduced in the multivariable model by using forward stepwise logistic regression. The odds ratio (OR) and 95% CI were calculated. All analyses were carried out using Stata 18 (StataCorp).

Ethical Considerations

The study was approved by the Institutional Review Board of the Institute of Medical Information, Chinese Academy of Medical Sciences (IMICAMS/02/21/HREC). To safeguard patient privacy, personal information, including names and identification numbers, were blocked and excluded from the analysis. As the research involved secondary analysis of deidentified data, informed consent was exempted. No financial compensation was provided to either the patients or the researchers, and there were no funding sources that could create a conflict of interest for the study.

Results

Characteristics of Health Care Expenditure

Figure 1 illustrates the distribution of medical expenditures of cross-province rural patients. HC patients (299/2987, 10%) were responsible for 34.5% of total medical expenditures. The average cost per hospitalization for HC patients was US $6261, which was 2.47 times of the remaining patients. The disparity was even more pronounced in annual hospitalization costs, with HC patients incurring an average of US $22,460 per year, 6.34 times higher than NHC patients (Table 3).

XSL•FO
RenderX

Figure 1. Overall distribution of annual medical expenditures referring to the cross-provincial patients.

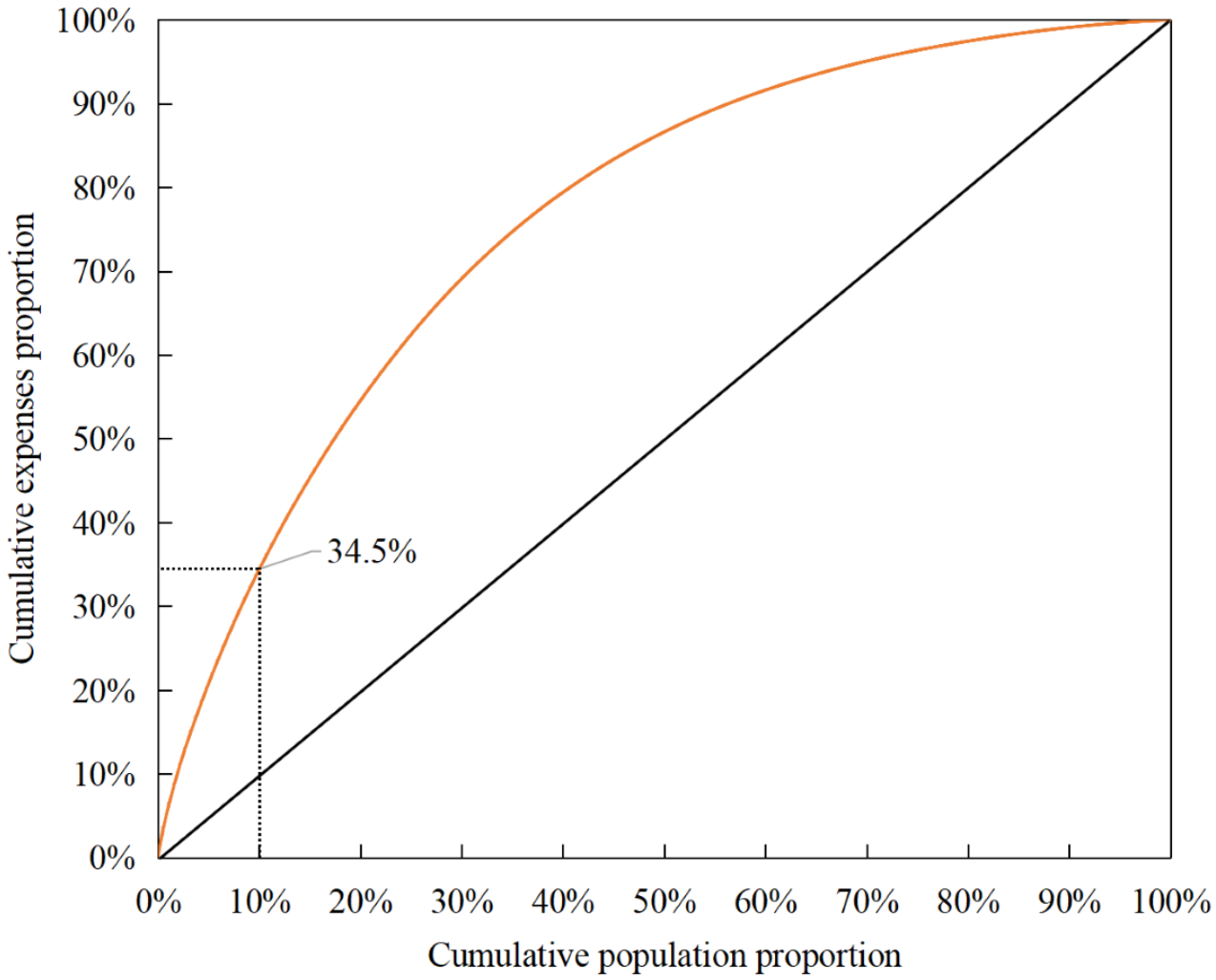


Table . Comparison of hospitalization costs between high-cost and non–high-cost patients.

Costs	All patients (N=2987), median (IQR)	HC[a] (n=299), median (IQR)	NHC[b] (n=2688), median (IQR)	*P* value
Cost per hospitalization (in US $)	2852.10 (1406.07-5,753,51)	6260.86 (3999.69-17,942.18)	2531.65 (1319.81-5143.20)	<.001
Annual per capita hospitalization cost (in US $)	4301.25 (1678.82-10,520.53)	22,460.42 (19,439.39-27,140.60)	3545.40 (1551.54-8052.82)	<.001

[a]HC: high-cost.

[b]NHC: non–high-cost.

Sample Characteristics

Table 4 illustrates the baseline characteristics for the sample. A total of 2987 patients were included in this study, with a median age of 47 (IQR 30.00-58.00) years. Males accounted for 57.4% (1713/2987) of total patients. Among all patients, migration-driven cross-province patients and disease-driven cross-province patients were evenly distributed, each comprising nearly 50% (1477/2987) of the total. The average number of diseases per patient was 1.06 (SD 0.27), and 48.4% (1446/2987) of the patients were diagnosed with neoplasms. Of these patients, 93.4% (2789/2987) went to Guangdong which had the highest gross domestic product in China in 2022 [36], which is geographically adjacent to Hainan Province. In addition, 63.7% (1904/2987) of these patients sought care in general hospitals.

XSL•FO
RenderX

Table . Characteristics of cross-provincial beneficiaries by cost subgroup.

Characteristic		All patients (n=2987)	HC[a] (n=299)	NHC[b] (n=2688)
Demographic characteristics				
Age (years)				
	Median (IQR)	47.00 (30.00-58.00)	44.00 (31.00-56.00)	47.00 (30.00-58.00)
	≤18, n (%)	428 (14.3)	40 (13.4)	388 (14.4)
	18 - 60, n (%)	1947 (65.2)	202 (67.6)	1745 (64.9)
	≥60, n (%)	612 (20.5)	57 (19.1)	555 (20.7)
Sex, n (%)				
	Male	1713 (57.4)	196 (65.6)	1517 (56.4)
	Female	1274 (42.7)	103 (34.5)	1171 (43.6)
Migration type, n (%)				
	Migration-driven cross-province patients	1477 (49.5)	88 (29.4)	1389 (51.7)
	Disease-driven cross-province patients	1510 (50.6)	211 (70.6)	1299 (48.3)
Enabling characteristic				
Income (in RMB[c]), median (IQR)		13,374 (12,679-14,206)	13,374 (12,679-14,465)	13,374 (12,679-14,206)
Need characteristics				
Number of diseases				
	Mean (SD)	1.06 (0.27)	1.17 (0.43)	1.05 (0.24)
	1, n (%)	2815 (94.2)	253 (84.6)	2562 (95.3)
	2, n (%)	159 (5.3)	40 (13.4)	119 (4.4)
	3, n (%)	10 (0.3)	6 (2.0)	4 (0.2)
	4, n (%)	3 (0.1)	0 (0)	3 (0.1)
Top 10 diagnosis, n (%)				
	C00-D48 Neoplasms	1446 (48.4)	232 (77.6)	1214 (45.2)
	K00-K93 Diseases of the digestive	201 (6.7)	8 (2.7)	193 (7.2)
	M00-M99 Diseases of the musculoskeletal system and connective tissue	192 (6.4)	15 (5.0)	177 (6.6)
	N00-N99 Diseases of the genitourinary system	191 (6.4)	7 (2.3)	184 (6.9)
	I00-I99 Diseases of the circulatory system	167 (5.6)	22 (7.4)	145 (5.4)
	J00-J99 Diseases of the respiratory	128 (4.3)	8 (2.7)	120 (4.5)
	G00-G99 Diseases of the nervous system	111 (3.7)	5 (1.7)	106 (3.9)
	H00-H59 Eye and adnexa diseases	108 (3.6)	2 (0.7)	106 (3.9)
	Q00-Q99 Congenital malformations, deformations, and chromosomal abnormalities	100 (3.4)	7 (2.3)	93 (3.5)
	E00-E90 Endocrine, nutritional, and metabolic diseases	80 (2.7)	2 (0.7)	78 (2.9)
Environmental characteristics, n (%)				
Regions				
	Guangdong	2789 (93.4)	289 (96.7)	2500 (93.0)
	Other area	106 (3.6)	3 (1.0)	103 (3.8)
	Beijing	41 (1.4)	4 (1.3)	37 (1.4)

XSL•FO
RenderX

Characteristic		All patients (n=2987)	HC[a] (n=299)	NHC[b] (n=2688)
	Guangxi	37 (1.2)	2 (0.7)	35 (1.3)
	Shanghai	14 (0.47)	1 (0.3)	13 (0.48)
Medical institution type				
	Hospitals	1904 (63.7)	146 (48.8)	1758 (65.4)
	Specialized hospitals	1083 (36.3)	153 (51.2)	930 (34.6)

[a]HC: high-cost.

[b]NHC: non–high-cost.

[c]RMB ¥6.8985=US $1 in 2019.

Compared with NHC patients, the proportion of males was higher among HC patients (196/299, 65.6% vs 1517/2688, 56.4%), and the proportion of disease-driven cross-province patients among HC was higher than migration-driven cross-province patients. Regarding disease type, HC patients were more likely to suffer from neoplasms (C00-D48) and circulatory system diseases (I00-I99), whereas NHC group exhibited a higher prevalence of digestive diseases (K00-K93), genitourinary system disorders (N00-N99), nervous system diseases (G00-G99), eye and adnexa diseases (H00-H59), and endocrine, nutritional, and metabolic diseases (E00-E90). In addition, the proportion of HC patients with 2 or more diseases was higher than that among NHC patients.

Risk Factors Associated With HC Patients

To identify the risk factors associated with HC status, 9 variables identified as significant in the univariable logistic regression analysis were included in the multivariable model. Male patients were more likely to be HC patients (OR 1.38, 95% CI 1.06-1.79; *P*=.01). The number of diseases (OR 3.62, 95% CI 2.37 - 5.53; *P*<.001) and neoplasms (OR 2.31, 95% CI 1.61 - 3.31; *P*<.001) remained strong predictors. In addition, receiving care at specialized medical institutions was positively associated with HC status (OR 1.61, 95% CI 1.24 - 2.08; *P*<.001). However, migration type (OR 0.75, 95% CI 0.55 - 1.04; *P*=.09) and some disease types, such as genitourinary disease, were no longer statistically significant after adjusting for covariates (Table 5).

Table . Risk factors associated with the high-cost patients.

Covariate (reference group)	Univariable analysis		Multivariable analysis	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Sex (reference: female)	1.46 (1.14-1.89)	.04	1.38 (1.06-1.79)	.01
Migration type (reference: disease-driven cross-province patients)	0.39 (0.30-0.51)	<.001	0.75 (0.55-1.04)	.09
Number of diseases	2.95 (2.15-4.03)	<.001	3.62 (2.37-5.53)	<.001
Neoplasms (C00-D48)	4.19 (3.16-5.57)	<.001	2.31 (1.61-3.31)	<.001
Diseases of the digestive system (K00-K93)	0.35 (0.17-0.73)	.005	0.40 (0.18-0.90)	.03
Diseases of the genitourinary system (N00-N99)	0.33 (0.15-0.70)	.004	0.47 (0.20-1.09)	.08
Eye and adnexa diseases (H00-H59)	0.16 (0.04-0.67)	.01	0.19 (0.04-0.82)	.03
Endocrine, nutritional, and metabolic diseases (E00-E90)	3.01 (1.74-5.19)	<.001	0.83 (0.40-1.73)	.63
Types of medical institutions	1.98 (1.56-2.52)	<.001	1.61 (1.24-2.08)	<.001

The Hosmer-Lemeshow goodness-of-fit test yielded χ^2_8=3.5 (*P*=.90), indicating a good model fit.

We conducted the Hosmer-Lemeshow goodness-of-fit test to assess the adequacy of the logistic regression model. The results indicate a good model fits.

Discussion

Principal Findings

This study is the first to analyze the characteristics and risk factors of HC cross-provincial patients from rural China. The analysis of cost distribution revealed that, unlike local patients, whose medical expenses are nearly half-concentrated among

XSL•FO
RenderX

HC patients, the expenditures of cross-provincial patients are more dispersed, accounting for 34.5% of total medical costs. In addition, our findings revealed that both the per-hospitalization and annual inpatient expenditures for cross-provincial patients were higher than those for local patients. This reflects the differences in cost structure and expenditure levels between local and cross-provincial patients.

We adopt AMHSM as a theoretical framework to identify the factors associated with HC patients. Unlike previous studies, we considered the environmental factors and migration type, reflecting the specific characteristics of cross-provincial patients. The results showed that in addition to gender, number of illnesses, and type of illness, the type of institution was also important predictor of HC patients. This finding can provide more precise evidence support for relevant policy making.

Comparison With Previous Work

For policy makers, focusing on HC patients may be highly efficacious for reducing healthcare expenses. We therefore analyzed the characteristics and risk factors of HC patients among cross-provincial beneficiaries in rural China, which has been generally neglected by other researchers.

First, our study found that the expenses of HC patients accounted for only 34.5% of total expenses. This is substantially different from the cost distribution of local inpatients. According to the study of Fan et al [37] and Xiaobo and Chuang [38], the local HC beneficiaries in rural China were responsible for approximately 50% of total inpatient costs. By contrast, the cost structure of cross-provincial patients appears less concentrated. This may be attributed to the types of institutions eligible for reimbursement and to patients' health care–seeking behavior. In the initial phase of policy implementation, in order to alleviate the financial burden on patients with serious illnesses and to account for variations in institutional information technology capacity, tertiary medical institutions served as the main providers of cross-provincial direct settlement for medical expenses. According to statistical data, by 2018, 96.6% (2462/2548) of tertiary medical institutions nationwide had been included in the network of direct cross-provincial settlement for medical expenses, compared with 58.1% (5238/9017) and 25.5% (2758/10,831) of secondary and primary institutions, respectively [34,39,40]. Due to the greater certainty of reimbursement in tertiary institutions, combined with patient's trust in the clinical capabilities of higher-level providers, 99% (2957/2987) of cross-provincial patients in Hainan chose to seek treatment at tertiary medical institutions. Compared with local residents, this disproportionate pattern of health care utilization has reduced the number of patients within the lower-cost segment, thereby lowering the overall concentration of cross-provincial health care service utilization.

Second, we found that HC migrant patients incurred higher annual medical expenses than their local patients. Specifically, the average annual hospitalization cost for HC migrant patients was US $22,460, whereas that for local HC patients ranged from US $4398 to US $11,516 [38,41,42]. This may be related to selective medical migration. Notably, 70.6% (211/299) of HC patients are disease-driven cross-province patients, many of whom suffer from severe and complex conditions. These individuals often seek care at regions with well-developed health care resources and typically exhibit a high willingness to pay. Consequently, their medical expenditures tend to exceed those of local patients. However, the possibility of underlying inefficiencies in the utilization of medical resources cannot be ruled out [43,44].

The results of the multivariable logistic regression model revealed the characteristics and risk factors of HC patients. Our findings reveal that gender, number of diseases, cancer diagnoses, and type of hospital are independent risk factors for becoming an HC patient, whereas age, migration type, income level, other disease types, and treatment region are not. These results differ from studies conducted on local patients. In research focusing on local populations [4,28], age is often associated with HC patients, with older age linked to a higher risk. However, our study found no significant association between age and HC status, which may reflect the structural complexity of the migrant population. Compared with local patients, migrants are primarily composed of 2 types—disease-driven cross-province patients and migration-driven cross-province patients. Disease-driven cross-province patients are likely to consist of middle-aged and young individuals who are more mobile and willing to pay for care, while migration-driven cross-province patients have a more diverse composition, including freelancers, older adult individuals and children migrating with family members, and working-age adults relocating for employment. Consequently, the association between age and HC patients may vary depending on the changing composition of the migrant population. In a study of local patients, previous studies [45,46] have reported inconsistent results regarding the influence of gender on HC patient status. In our research, male patients were more likely to become HC patients, potentially reflecting traditional societal norms in China, where men are more likely to take on roles involving external work. Several studies on the health care–seeking behavior of the migrant population have reached similar conclusions [47,48], often attributing this to the perception that men are the primary economic providers and constitute valuable human capital within the household. As a result, their health is prioritized, and greater medical resources are more likely to be allocated to them.

We found that patients with chronic conditions, such as cancer, and those with multiple comorbidities were more likely to be HC patients. This may be attributable to the relatively high resource consumption associated with the treatment of such conditions, which is consistent with the findings of other studies [49]. We also found that HC cross-provincial patients tend to be concentrated among those with cancer and gastrointestinal diseases. However, studies of local patients have identified additional predictors of HC status, including neurological disorders [50], kidney diseases [51], and diabetes [52], as well as predictors of HC status. This suggests that cancer may represent a relative weakness in the local health care capacity in Hainan, prompting patients to seek care outside the province. By contrast, other HC conditions appear to be adequately treated locally and, as such, do not emerge as significant predictors of cross-provincial HC utilization.

XSL•FO
RenderX

Institution type also serves as a risk factor for becoming an HC cross-provincial patient. Our study found that individuals receiving care at specialist hospitals were more likely to incur high medical expenses, which may be related to the fact that these institutions are better equipped to manage complex and severe conditions, and consequently involve greater resource consumption and higher costs.

Migration type does not appear to be a key predictor for becoming HC patients. However, based on migration-related characteristics, disease-driven cross-province patients would be expected to have a higher likelihood of becoming HC users. To examine this further, we conducted a subgroup analysis by year (Multimedia Appendix 1). The results indicate that in 2017, disease-driven cross-province patients were more likely to become HC patients (OR 0.45, 95% CI 0.22 - 0.93; *P*=.03), whereas in 2018 (OR 0.89, 95% CI 0.55 - 1.66; *P*=.96) to 2019 (OR 0.94, 95% CI 0.60 - 1.45; *P*=.94), there was no significant difference between disease-driven cross-province patients and migration-driven cross-province patients in the likelihood of incurring high medical costs. On the one hand, this may be related to limited awareness of the cross-provincial medical settlement policy. In 2017, when the policy had just been introduced, migrants living outside their registered residence were often situated in an information gap and unaware that medical expenses incurred elsewhere could be directly reimbursed. A survey from Shanghai in 2017 found that 64.5% (129/200) of migrant patients were unfamiliar with the policy [53]. Based on a relatively lower economic capacity and unawareness of the reimbursement policy, the migrant population may still choose to return to their hometown for treatment when they get seriously ill, which leads to a lower health care cost for migration-driven cross-province patients. As policy awareness improved over time, an increasing number of migration-driven cross-province patients chose to access health care services locally, resulting in a gradual rise in the proportion of HC users within this group. On the other hand, this may be related to the fragmented pattern of care-seeking typically observed among disease-driven cross-province patients. These patients often undergo initial diagnosis and treatment within their insured region, and only seek care at higher-level institutions outside the insured area when treatment proves ineffective. After a certain therapeutic effect is obtained, they tend to return to the insured location for rehabilitation treatment. However, the data used in this study were only drawn from the database of the Cross-provincial Medical Immediate Reimbursement System and did not capture pre- and posttreatment costs incurred within the insured region, which are recorded in the URRBMI system. As a result, the annual health care expenditure of disease-driven cross-province patients did not appear higher than that of migration-driven cross-province patients. The combined effect of low policy awareness and disease-driven cross-province patients' fragmented care patterns may explain why there is no significant difference between the 2 groups in the likelihood of becoming HC patients.

Our findings provide valuable insights to inform decision-making. From the perspective of health resource allocation and utilization, local governments regularly invest in developing and improving health systems. The results of this study can guide resource allocation priorities. For example, we identified neoplasms as the highest-risk diagnosis among HC patients. Under this premise, policy makers can prioritize improving local diagnostic and treatment capacities for neoplasms and reduce the outflow of cross-provincial patients in Hainan, especially for disease-driven cross-province patients. This move brings at least 3 benefits. The first is to reduce direct medical expenses. Table 4 shows that 96.6% (289/299) of HC patients go to Guangdong for treatment. The 2020 Yearbook of Health Statistics shows that the average cost per hospitalization in Guangdong is US $1771, while that in Hainan is US $1479 [34], which means that receiving treatment in Hainan will cost less. The second is saving indirect costs, including the cost of transportation and accommodation. Finally, from the perspective of medical insurance fund supervision, it is more beneficial to the regulation and charge of the health care fund in the local medical institutions. Yu and Lang [54] and Auster and Oaxaca [55] demonstrated that differences in payment management mechanisms contribute to a higher likelihood of overtreatment when seeking health care in other provinces. Thus, patients are likely to receive more standardized medical treatment within their insured regions compared with cross-provincial patients. This can reduce unnecessary expenditures for medical insurance funds and out-of-pocket costs for individuals.

Limitations

This study has several limitations. First, patients who were reimbursed manually were not included [56]. Second, due to policy restrictions, outpatients receiving cross-provincial medical care were not eligible for reimbursement before 2021, so only inpatients were included. Third, the medical insurance database lacks individual-level income information. Hence, the per capita disposable income of the registered residence area was used as a proxy.

Conclusions

This study examined the characteristics and risk factors of HC cross-provincial insured patients in rural China. Our findings indicate that, unlike local patients, HC patients among the cross-provincial group did not exhibit a high degree of expenditure concentration, yet their average annual health care spending was significantly higher. This suggests the existence of cost-driving factors specific to cross-provincial patients—such as restrictions on reimbursable institutions, a preference for higher-level facilities, and insufficient fund oversight—which may contribute to both greater cost dispersion and elevated per capita health expenditure.

Our research also found that male patients, individuals with multiple comorbidities, those diagnosed with cancer, and patients receiving care at specialized hospitals were more likely to be classified as HC patients. Understanding these characteristics can assist policy makers in regions of patient outflow to develop targeted interventions—such as strengthening prevention and management of HC conditions and promoting more equitable allocation of health care resources—to reduce the outflow of patients with severe illnesses.

XSL•FO
RenderX

Acknowledgments

This study was funded by the Humanities and Social Sciences Research Project of the Ministry of China, titled "Health Service Accessibility for Elderly Migrant Populations under Inter-Regional Healthcare Direct Settlement Policy: Current Status, Mechanisms and Optimization Strategies" (grant 23YJCZH064).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to privacy and ethical considerations, but are available from the corresponding author on reasonable request.

Authors' Contributions

MG and XJ were involved in the study conceptualization and study design. XJ and MG carried out the data analyses. XJ wrote the first draft of the manuscript. MG provided critical revisions of the paper. Y Liu (third author), ZZ, and Y Liu (fourth author) contributed to the data collection and cleaning. FZ and ZZ contributed to searching the literature and editing the manuscript. Y Li and MG supervised the project. All authors contributed to the interpretation of the data, and approved the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Multivariable analysis of the association between migration type and high-cost status by year (reference group: disease-driven cross-province patients).
[DOCX File, 12 KB - publichealth_v11i1e54234_app1.docx]

References

1. Yu H. Universal health insurance coverage for 1.3 billion people: what accounts for China's success? Health Policy 2015 Sep;119(9):1145-1152. [doi: 10.1016/j.healthpol.2015.07.008] [Medline: 26251322]
2. The three basic medical insurance schemes covered more than 13 billion people at the end of 2011 [Web page in Chinese]. The State Council of the People's Republic of China. 2014. URL: https://www.gov.cn/wenzheng/wz_zxft_ft12/2014-04/18/content_2662006.htm [accessed 2025-03-01]
3. World Health Organization national statistical bulletin of the development of medical insurance in 2022 [Web page in Chinese]. National Healthcare Security Administration. URL: http://www.nhsa.gov.cn/art/2023/7/10/art_7_10995.html [accessed 2025-03-01]
4. Zhang Y, Lu S, Niu Y, Zhang L. Medical expenditure clustering and determinants of the annual medical expenditures of residents: a population-based retrospective study from rural China. BMJ Open 2018 Jun 22;8(6):e022721. [doi: 10.1136/bmjopen-2018-022721] [Medline: 29934397]
5. Medical security development statistics bulletin [Web page in Chinese]. The State Council of the People's Republic of China. 2023. URL: https://www.gov.cn/lianbo/bumen/202404/content_6944653.htm [accessed 2025-01-07]
6. Medical security development statistics bulletin Beijing [Web page in Chinese]. The State Council of the People's Republic of China. 2022. URL: https://www.gov.cn/lianbo/bumen/202307/content_6891062.htm [accessed 2025-01-07]
7. Suo Z, Shao L, Lang Y. A study on the factors influencing the utilization of public health services by China's migrant population based on the Shapley value method. BMC Public Health 2023 Nov 24;23(1):2328. [doi: 10.1186/s12889-023-17193-3] [Medline: 38001411]
8. Major figures on 2020 population census of China, Beijing [Web page in Chinese]. National Bureau of Statistics of China. URL: https://www.stats.gov.cn/sj/pcsj/rkpc/d7c/202303/P020230301403217959330.pdf [accessed 2025-01-07]
9. Chengrong D, Yuding Q, Fan H, et al. From 6.57 Million to 376 Million: Remarks on Migration Transition in China. Popul Res 2022;46(6):41-58 [FREE Full text]
10. Statistical bulletin on the development of national medical security in 2021 [Web page in Chinese]. National Healthcare Security Administration. URL: http://www.nhsa.gov.cn/art/2022/6/8/art_7_8276.html [accessed 2024-10-23]
11. Blumenthal D, Chernof B, Fulmer T, Lumpkin J, Selberg J. Caring for high-need, high-cost patients - an urgent priority. N Engl J Med 2016 Sep 8;375(10):909-911. [doi: 10.1056/NEJMp1608511] [Medline: 27602661]
12. Long P, Abrams M, Milstein A, Anderson G, Apton KL, Dahlberg ML, editors. Effective Care for High-Need Patients: Opportunities for Improving Outcomes, Value, and Health: Washington (DC): National Academy of Medicine; 2017. URL: https://nam.edu/wp-content/uploads/2017/06/Effective-Care-for-High-Need-Patients.pdf [accessed 2024-10-22]

XSL•FO
RenderX

13. Tanke MAC, Feyman Y, Bernal-Delgado E, et al. A challenge to all. A primer on inter-country differences of high-need, high-cost patients. PLoS ONE 2019;14(6):e0217353. [doi: 10.1371/journal.pone.0217353] [Medline: 31216286]
14. Lee NS, Whitman N, Vakharia N, Taksler GB, Rothberg MB. High-cost patients: hot-spotters don't explain the half of It. J Gen Intern Med 2017 Jan;32(1):28-34. [doi: 10.1007/s11606-016-3790-3] [Medline: 27480529]
15. Wammes JJG, van der Wees PJ, Tanke MAC, Westert GP, Jeurissen PPT. Systematic review of high-cost patients' characteristics and healthcare utilisation. BMJ Open 2018 Sep 8;8(9):e023113. [doi: 10.1136/bmjopen-2018-023113] [Medline: 30196269]
16. DeLia D. Mortality, disenrollment, and spending persistence in Medicaid and CHIP. Med Care 2017 Mar;55(3):220-228. [doi: 10.1097/MLR.0000000000000648] [Medline: 27662591]
17. Fitzpatrick T, Rosella LC, Calzavara A, et al. Looking beyond income and education: socioeconomic status gradients among future high-cost users of health care. Am J Prev Med 2015 Aug;49(2):161-171. [doi: 10.1016/j.amepre.2015.02.018] [Medline: 25960393]
18. Graven PF, et al. Preventable acute care spending for high-cost patients across payer types. J Health Care Finance 2016;42 [FREE Full text]
19. Figueroa JF, Frakt AB, Lyon ZM, Zhou X, Jha AK. Characteristics and spending patterns of high cost, non-elderly adults in Massachusetts. Healthcare (Basel) 2017 Dec;5(4):165-170. [doi: 10.1016/j.hjdsi.2017.05.001]
20. Joynt KE, Figueroa JF, Beaulieu N, Wild RC, Orav EJ, Jha AK. Segmenting high-cost Medicare patients into potentially actionable cohorts. Healthcare (Basel) 2017 Mar;5(1-2):62-67. [doi: 10.1016/j.hjdsi.2016.11.002]
21. Krause TM, Yay Donderici E, Ganduglia Cazaban C, Franzini L. Future expenditure risk of silent members: a statistical analysis. BMC Health Serv Res 2016 Jul 29;16(319):319. [doi: 10.1186/s12913-016-1552-x] [Medline: 27473359]
22. Ku LJE, Chiou MJ, Liu LF. Variations in the persistence of health expenditures and the implications for the design of capitation payments in Taiwan. J Health Serv Res Policy 2015 Jul;20(3):146-153. [doi: 10.1177/1355819615577711] [Medline: 25829410]
23. Guilcher SJT, Bronskill SE, Guan J, Wodchis WP. Who are the high-cost users? A method for person-centred attribution of health care spending. PLoS ONE 2016;11(3):e0149179. [doi: 10.1371/journal.pone.0149179] [Medline: 26937955]
24. Beaulieu ND, Joynt KE, Wild R, Jha AK. Concentration of high-cost patients in hospitals and markets. Am J Manag Care 2017 Apr;23(4):233-238. [Medline: 28554206]
25. Robst J. Developing models to predict persistent high-cost cases in Florida Medicaid. Popul Health Manag 2015 Dec;18(6):467-476. [doi: 10.1089/pop.2014.0174] [Medline: 26102363]
26. Hirth RA, Gibson TB, Levy HG, Smith JA, Calónico S, Das A. New evidence on the persistence of health spending. Med Care Res Rev 2015 Jun;72(3):277-297. [doi: 10.1177/1077558715572387] [Medline: 25701579]
27. Lu S, Zhang Y, Zhang L, Klazinga NS, Kringos DS. Characterizing potentially preventable hospitalizations of high-cost patients in rural China. Front Public Health 2022;10(804734):804734. [doi: 10.3389/fpubh.2022.804734] [Medline: 35211444]
28. Zhang Y, Grinspan Z, Khullar D, et al. Developing an actionable patient taxonomy to understand and characterize high-cost Medicare patients. Healthcare (Basel) 2020 Mar;8(1):100406. [doi: 10.1016/j.hjdsi.2019.100406]
29. Powers BW, Yan J, Zhu J, et al. Subgroups of high-cost medicare advantage patients: an observational study. J Gen Intern Med 2019 Feb;34(2):218-225. [doi: 10.1007/s11606-018-4759-1] [Medline: 30511290]
30. China statistical yearbook 2022 [Web page in Chinese]. National Bureau of Statistics of China. URL: https://www.stats.gov.cn/sj/ndsj/2022/indexch.htm [accessed 2024-10-23]
31. Bulletin of the Seventh National Population Census (no 7) [Web page in Chinese]. The State Council of the People's Republic of China. URL: https://www.gov.cn/xinwen/2021-05/11/content_5605791.htm [accessed 2024-10-23]
32. Andersen RM. National health surveys and the behavioral model of health services use. Med Care 2008 Jul;46(7):647-653. [doi: 10.1097/MLR.0b013e31817a835d] [Medline: 18580382]
33. Alberga A, Holder L, Kornas K, Bornbaum C, Rosella L. Effects of behavioural risk factors on high-cost users of healthcare: a population-based study. Can J Public Health 2018 Aug;109(4):441-450. [doi: 10.17269/s41997-018-0127-5] [Medline: 30232715]
34. Notice on further completing the work of direct settlement of medical treatment in different provinces of basic medical insurance [Web page in Chinese]. The State Council of the People's Republic of China. URL: https://www.gov.cn/zhengce/zhengceku/2022-07/26/content_5702881.htm [accessed 2024-09-05]
35. ICD-10 Version2016. URL: https://icd.who.int/browse10/2016/en [accessed 2024-09-06]
36. China statistical yearbook 2023 [Web page in Chinese]. National Bureau of Statistics of China. URL: https://www.stats.gov.cn/sj/ndsj/2023/indexch.htm [accessed 2024-10-23]
37. Fan Q, Wang J, Nicholas S, Maitland E. High-cost users: drivers of inpatient healthcare expenditure concentration in urban China. BMC Health Serv Res 2022 Nov 14;22(1):1348. [doi: 10.1186/s12913-022-08775-9] [Medline: 36376840]
38. Xiaobo P, Chuang D. Concentration and persistence of health spending: micro-empirical evidence from China. The Journal of World Economy 2019;42(12):51-76 [FREE Full text]

XSL•FO
RenderX

39. Release of public service information on direct settlement of medical expenses for medical treatment in different provinces of basic medical insurance (phase II) [Web page in Chinese]. Nation Healthcare Security Administration. URL: http://www.nhsa.gov.cn/art/2018/9/4/art_114_7167.html [accessed 2024-10-23]
40. Remote network of designated medical institutions [Web page in Chinese]. Nation Healthcare Security Administration. URL: https://fuwu.nhsa.gov.cn/nationalHallSt/#/search/MedicalTreatmentOrganSearch [accessed 2025-03-23]
41. Chen H, Peng X, Shen M. Concentration and persistence of healthcare spending: evidence from China. Sustainability 2021;13(11):5761. [doi: 10.3390/su13115761] [Medline: 36778665]
42. Miao Y, Qian D, Sandeep S, et al. Exploring the characteristics of the high-cost population from the family perspective: a cross-sectional study in Jiangsu Province, China. BMJ Open 2017 Nov 9;7(11):e017185. [doi: 10.1136/bmjopen-2017-017185] [Medline: 29127225]
43. Jingran H, Zheng K, Fangting C, et al. Qualitative study on the supervision dilemma and coping strategies of medical insurance funds under non-local medical treatment [Article in Chinese]. Chinese Hospital Management 2024;44(4):23-28 [FREE Full text]
44. Xuan Z. Reflections on the operation of direct settlement of medical treatment in other places [Article in Chinese]. Employment and Security 2023(7):52-54 [FREE Full text]
45. Wammes JJG, Tanke M, Jonkers W, Westert GP, Van der Wees P, Jeurissen PP. Characteristics and healthcare utilisation patterns of high-cost beneficiaries in the Netherlands: a cross-sectional claims database study. BMJ Open 2017 Nov 12;7(11):e017775. [doi: 10.1136/bmjopen-2017-017775] [Medline: 29133323]
46. de Oliveira C, Cheng J, Vigod S, Rehm J, Kurdyak P. Patients with high mental health costs incur over 30 percent more costs than other high-cost patients. Health Aff (Millwood) 2016 Jan;35(1):36-43. [doi: 10.1377/hlthaff.2015.0278] [Medline: 26733699]
47. Chen YJ. Analysis of medical treatment behavior of floating population in Quanzhou. Health Education and Health Promotion 2017;12(5):439-441. [doi: 10.16117/j.cnki.31-1974/r.201705014] [Medline: 28558491]
48. Yani Y, et al. Study on choice of medical treatment of floating population in Xuzhou based on Andersen model. Med Soc 2023;36(2):58-63. [doi: 10.13723/j.yxysh.2023.02.011]
49. Zulman DM, Pal Chee C, Wagner TH, et al. Multimorbidity and healthcare utilisation among high-cost patients in the US Veterans Affairs Health Care System. BMJ Open 2015 Apr 16;5(4):e007771. [doi: 10.1136/bmjopen-2015-007771] [Medline: 25882486]
50. Khullar D, Zhang Y, Kaushal R. Potentially preventable spending among high-cost medicare patients: implications for healthcare delivery. J Gen Intern Med 2020 Oct;35(10):2845-2852. [doi: 10.1007/s11606-020-05691-8] [Medline: 32103440]
51. Figueroa JF, Zhou X, Jha AK. Characteristics and spending patterns of persistently high-cost medicare patients. Health Aff (Millwood) 2019 Jan;38(1):107-114. [doi: 10.1377/hlthaff.2018.05160] [Medline: 30615516]
52. Hartmann J, Jacobs S, Eberhard S, von Lengerke T, Amelung V. Analysing predictors for future high-cost patients using German SHI data to identify starting points for prevention. Eur J Public Health 2016 Aug;26(4):549-555. [doi: 10.1093/eurpub/ckv248] [Medline: 26850904]
53. Zuda L, Xiaomei T, Shu C, Peira C. Study on patient centered behavior of seeking medical treatment in different places—taking Shanghai as an example. China Health Insurance 2020(10):62-64. [doi: 10.19546/j.issn.1674-3830.2020.10.014]
54. Yu L, Lang J. Diagnosis-related Groups (DRG) pricing and payment policy in China: where are we? Hepatobiliary Surg Nutr 2020 Dec;9(6):771-773. [doi: 10.21037/hbsn-2020-8] [Medline: 33299831]
55. Auster RD, Oaxaca RL. Identification of supplier induced demand in the health care sector. J Hum Resour 1981;16(3):327-342. [doi: 10.2307/145624] [Medline: 7264294]
56. Guo M, Li Y, Zhang F, et al. Accessibility of immediate reimbursement for cross-provincial healthcare anoits influencing factors for rural residents: a nationwide cross-sectional survey. Chin J Public Health 2022;38(11):1469-1473. [doi: 10.11847/zgggws1137544]

Abbreviations

AMHSM: Andersen's Behavioral Model of Health Services Use
BMI: basic medical insurance
HC: high-cost
NHC: non–high-cost
URRBMI: Urban and Rural Resident Basic Medical Insurance

XSL•FO
RenderX

Edited by A Mavragani, T Sanchez; submitted 02.11.23; peer-reviewed by L Hao, Y Lai; revised version received 08.04.25; accepted 11.04.25; published 11.06.25.

Please cite as:
Guo M, Jiang X, Liu Y, Liu Y, Zhang F, Zhang Z, Li Y
Characterizing Cross-Provincial High-Cost Patients in Rural China: Cross-Sectional Study
JMIR Public Health Surveill 2025;11:e54234
URL: https://publichealth.jmir.org/2025/1/e54234
doi:10.2196/54234

© Minjiang Guo, Xiaotong Jiang, Yang Liu, Yang Liu, Fangyuan Zhang, Zhongyuan Zhang, Yazi Li. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 11.6.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Impact of Distance Traveled and Birth Volumes of Hospital Maternity Units on Newborn Outcomes: Population-Based Cohort Study

Anna Cantarutti[1*], PhD; Riccardo Boracchini[1*], MSc; Roberto Bellù[2], MD; Raffaella Ronco[1], MSc; Federico Rea[1], PhD; Anna Locatelli[3], MD; Rinaldo Zanini[4], MD; Giovanni Corrao[1], PhD

[1]National Centre for Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy

[2]Neonatal Intensive Care Unit, Woman and Child Health Department, Azienda Ospedaliera della Provincia di Lecco, Lecco, Italy

[3]Department of Obstetrics, IRCCS San Gerardo dei Tintori Foundation, University of Milano-Bicocca, Monza, Italy

[4]General Management, IRCCS Istituto Giannina Gaslini, Genova, Italy

[*]these authors contributed equally

Corresponding Author:
Anna Cantarutti, PhD
National Centre for Healthcare Research & Pharmacoepidemiology, University of Milano-Bicocca, Milan, Italy

Abstract

Background: The centralization of childbirth and newborn care in large maternity units has become increasingly prevalent in Europe. While this trend offers potential benefits such as specialized care and improved outcomes, it can also lead to longer travel and waiting times, especially for women in rural areas.

Objective: This study aimed to evaluate the association between hospital maternity unit (HMU) volumes, road travel distance (RTD) to the hospital, and other neonatal outcomes.

Methods: We conducted a population-based cohort study including all live births in hospitals without intensive care units between 2016 and 2019 in the Lombardy region, Italy. Given the hierarchical structure of our data (births nested within hospitals), we employed log-binomial regression models with random intercepts to estimate relative risks and 95% CIs for evaluating the association between HMU volumes (≥1500 births/year) and RTD (<5 km) with the risk of being transferred and/or death after birth (primary outcome). Secondary outcomes included a low Apgar score at 5 minutes and low adherence to antenatal care (ANC). We controlled for several potential confounders including adherence to the ANC pathway for the primary and low Apgar outcomes. To explore the influence of HMU volumes on the primary outcome, we identified the fractional polynomial model that best described this relationship.

Results: Of 65,083 live births, 71% (n=45,955) occurred in low-volume hospitals (<1000 births/year), 21% (n=13,560) involved long-distance travel (>15 km), 1% (n=735) were transferred and/or died after birth, 0.5% (n=305) had a low Apgar score at 5 minutes, and 64% (n=41,317) completely adhered to ANC. The risk of transfer and/or death increased as HMU volume decreased, ranging from 1% for hospitals with 1000 - 1500 births/year to a 3.6-fold high risk for hospitals with <500 births/year (compared to high-volume hospitals). Travel distance did not affect the primary outcome. Neither HMU volume nor RTD were associated with low Apgar scores. Conversely, the risk of complete adherence to ANC decreased with lower HMU volumes but increased with shorter RTD. Additionally, high-volume hospitals demonstrated a decreasing trend in the frequency of the primary outcome, with transfer and/or death rates ranging from 2% to 0.5% and flattening to 0.5% in hospitals, with activity volume ≥1000 mean births/year.

Conclusions: Our findings showed an excess risk of neonatal transfer and/or death for live births in HMUs with low activity volumes without an intensive care unit. In contrast, RTD primarily affected adherence to ANC. Moreover, data suggest that 1000 births/year could be an optimal cutoff for maternity hospitals to ensure an appropriate standard of care at delivery.

(JMIR Public Health Surveill 2025;11:e58944) doi:10.2196/58944

KEYWORDS

maternity level; road travel distance; delivery; regionalization; neonatal outcome; childbirth; newborn; cohort; birth; population-based; maternal; maternal outcomes; neonatal; European; health system; health care system; perinatal care; antenatal; mortality; neonatal mortality; perinatal

XSL•FO
RenderX

Introduction

Driven by a commitment to improving maternal and neonatal outcomes, several European health care systems, including the Italian National Health Service (NHS) have embraced the regionalization of perinatal care in large maternity units since the 1980s. Perinatal regionalization aims to optimize access to quality care by organizing maternity and neonatal services into distinct levels based on the complexity of care required. This approach involves (1) ensuring that pregnant women and newborns are directed to the appropriate level of care based on their needs, (2) implementing early screening and assessment mechanisms to identify women at risk, and (3) facilitating timely and appropriate transfers between levels of care when necessary. By implementing these strategies, perinatal regionalization can improve outcomes for both mothers and infants while ensuring the efficient use of health care resources [1-4].

While perinatal regionalization offers potential benefits, its implementation remains debated. Robust, real-world evidence is needed to inform decision-makers about both the advantages of centralization (eg, reduced travel time, specialized care) and the potential disadvantages (eg, increased travel distance for some women). Moreover, only a few studies have comprehensively examined the combined effects of hospital maternity unit (HMU) volume and road travel distance (RTD) on maternal and neonatal outcomes [5-7]. However, the centralization of HMUs inevitably leads to the closure of some units and increases the distance (and travel time) to the hospital for some mothers.

Regionalization of perinatal care in HMUs with at least 1000 births/year has been undertaken since 2010 in Italy, although some maternity units continue to record fewer than 500 births/year. Despite the heterogenous geographical conformation of Italy, the distances between maternity units are relatively modest, compared to those reported by North American studies [8]. This finding combined with the ongoing efforts to regionalize HMU volumes to units with at least 1000 births/year, makes Italy an ideal setting to investigate the relationship between HMU volume, travel distance, and perinatal outcomes.

In this study, we conducted a population-based cohort study in Lombardy, the largest and most populous region of Italy, to investigate the singular and combined effects of HMU volumes and RTD on specific outcomes in births occurring in hospitals without intensive care units (ICUs). Additionally, we studied the influence of perinatal factors like antenatal care (ANC) and sociodemographic features of the mothers.

Methods

Data Source and Study Cohort

The study cohort consisted of all live infants born in Lombardy from January 1, 2016, to December 31, 2019. Lombardy is a region in Italy that accounts for approximately 16% of the country's population, comprising nearly 10 million inhabitants. The health care utilization of all residents of Lombardy is covered by the government-funded NHS, which employs an automated system of databases to collect a variety of information. This system includes demographic and administrative data for all beneficiaries of the Regional Health Service (approximately covering the entire resident population), such as residence municipalities. Additional databases include the hospital discharges registry, which records all patients discharged from public or private hospitals; the outpatient drug prescription registry that reports all dispensations of NHS-reimbursable drugs; and the specialist visits and diagnostic exams registry including a specific automated system that collects data from the regional Department of Mental Health, accredited by the NHS and focused on outpatient specialist mental health care. Lastly, the Certificates of Delivery Assistance provide detailed information about pregnancy, childbirth, and fetal presentation at delivery. A unique deidentification code is systematically used for all databases; as a result, linking these records enables the creation of a large birth cohort and establishing relevant traits and care pathways for mothers and newborns.

The criteria for selecting the study cohort almost completely overlapped with those previously reported by our group [9]. Briefly, using the Certificates of Delivery Assistance database, we identified all live births in Lombardy between 2016 and 2019 and available identification codes to women who met the inclusion criteria: (1) were beneficiaries of the NHS and had been residents of Lombardy for at least one year before pregnancy, (2) were aged 15 to 55 years at delivery, and (3) delivered between 22 to 42 weeks of gestation, based on the first day of the last menstrual period ascertained via maternal reports or ultrasonography. All births recorded in maternity hospitals equipped with an ICU were excluded. Further, records with incomplete data were excluded because baseline covariates such as sociodemographic and gestational information may be missing for some women, and limiting analyses to the subset of women for whom complete data were available would not result in a significant loss of information (Figure 1).

XSL•FO
RenderX

Figure 1. Flowchart of inclusion and exclusion criteria. NHS: National Health Service.

XSL•FO
RenderX

The results were reported in agreement with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement guidelines [10].

Categorization of HMU and RTD

The HMU volumes for each hospital were obtained based on the mean number of births recorded between 2016 and 2019. Categories were defined based on the activity volume of the hospital in which the birth occurred (ie, volumes ≤500 births/year, 500 - 1000 births/year, 1000-1500 births/year, and >1500 births/year).

The RTD between the hospital and the mother's residence municipalities was calculated according to the distance matrix (ie, an open-source distance matrix database provided by the Italian Institute of Statistics) [11]. Births were classified based on the mother having traveled <5 km, 5 - 15 km, 15 - 21.5 km, or ≥21.5 km to the HMU.

Outcomes

The primary outcome of interest was transfer to a different hospital or death of the infant during the same birth hospitalization. The secondary outcomes included (1) a low Apgar score (<7) at 5 minutes after birth [12] and (2) complete adherence of the mothers to ANC during pregnancy. Maternal ANC adherence was evaluated based on the promptness and appropriateness of the number and timing of ANC interventions, including prenatal visits, ultrasound examinations, and laboratory tests, in relation to the length of pregnancy. In particular, the ANC assessment included (1) appropriateness of prenatal visits (ie, at least 5 visits during pregnancy or 2-4 visits for women with a gestational duration of <27 weeks); (2) promptness of prenatal visits (ie, at least 1 of the 4 visits within the 12th week of gestation); (3) appropriateness of ultrasound examinations (ie, ≥2 examinations during pregnancy, with at least 1 within the 12th week of gestation); and (4) appropriateness of laboratory tests, defined as completion of recommended tests for each trimester [13]. Women were identified as treatment-adherent if they followed all 4 of the recommendations.

Definitions of the Covariates

Several baseline maternal characteristics were considered. Sociodemographic features included nationality, marital status, employment, and educational attainment (ie, low: ≤5 years, medium: 6 - 13 years, and high: ≥14 years of study corresponding to primary education or none, lower or upper secondary education, and at least a bachelor's degree, respectively). Clinical information was obtained from the inpatient hospital registry, the outpatient drug prescriptions registry, and the specialist visits and diagnostic exams registry. Maternal comorbidities, including psychiatric comorbidities, medical comorbidities, obstetric characteristics, and health care utilization measures (ie, number of hospitalizations and distinct prescription drugs used as general markers of comorbidity) [14] were evaluated. Concomitant medications (ie, nonsteroidal anti-inflammatory drugs, antidepressants, antiepileptics, psycholeptics, antihypertensive drugs, and antidiabetic drugs) were measured from 1 year before the last menstrual period through the 22nd week of gestation. Information regarding the ANC of mothers was measured during pregnancy.

Ethical Considerations

According to the Italian Medicines Agency guidelines [15], retrospective studies that do not involve direct contact with patients do not need written consent to process personal data when they are used for research purposes. Thus, no ethics board review or approval was required.

To preserve privacy, each identification code was automatically anonymized, with only the Regional Health Authority having access to identifying information, which may be released upon request from judicial authorities to guarantee data deidentification.

Statistical Analysis

The baseline maternal characteristics and distribution of the ANC received by mothers during pregnancy were described overall and stratified by the primary outcome. The *t* test, χ^2 test, or Fisher exact test were used as appropriate. Moreover, the distribution of outcomes was expressed as absolute and relative percentages across HMU volumes and RTD categories. The Cochran-Armitage trend test was implemented for trend analysis.

As our data had a multilevel structure—with births (level 1) nested within hospitals (level 2)—log-binomial regression models with random intercepts were used to estimate relative risks (RRs) with 95% CIs for associations between HMU volumes (reference: ≥1500 births/year) and RTD (reference: RTD<5 km) and other outcomes of interest. The interaction between HMU volumes and RTD was also assessed. Models were adjusted for a disease score that was calculated based on the primary outcome, including all covariates listed above. For analyses in which ANC adherence was the outcome of interest, we excluded variables related to the ANC pathway from the disease score calculations.

A stratified analysis was performed by HMU volume, categorized as low (<1000 births/year) and high (≥1000 births/year) HMU volumes. Further, RTD was categorized as short (<15 km) and long RTD (≥15 km). Finally, to examine the influence of HMU volumes on the primary outcome (ie, transfer to a different hospital or death), a fractional polynomial model (FPM) was applied (first, second, or third degree), which described the relationship through a stepwise selection [16]. FPMs are particularly useful to preserve the continuous nature of the covariates in regression models while accounting for potential nonlinear relationships. Hospitals were the statistical units for this analysis.

All analyses were performed using the SAS (version 9.4; SAS Institute). *P* values <.05 were considered statistically significant.

Results

Among the 175,366 live births recorded in Lombardy between January 1, 2016, and December 31, 2019, 30,730 births were excluded because the mothers had not received health care from the NHS for at least 1 year before pregnancy, while 9 were

XSL•FO
RenderX

excluded because the mothers were younger than 15 or older than 55 years. Furthermore, 216 births were excluded because the gestational age at delivery was outside the range of 22 to 42 weeks, 75,395 births were excluded because they occurred in hospitals with an ICU, and 3933 births were excluded because complete sets of information about the mothers were not available (Figure 1). The final cohort consisted of 65,083 live births, of which 735 (1.13%) newborns were either transferred to another hospital (n=731) or died (n=4), 305 (0.47%) had a low Apgar score at 5 minutes after birth, and 41,317 (63.5%) fully adhered to ANC.

Mothers of newborns included in this cohort had a mean age of 32 (SD 5.4) years and mean gestational duration of 39 (SD 1.4) weeks. Most mothers had moderate educational attainment, were employed and married at delivery, and were of Italian nationality, and 55% (35,799/65,083) had already given birth. Maternal comorbidities were generally less prevalent and did not significantly differ across groups, except for psychoses and substance dependence, which were more common among mothers whose newborns experienced hospital transfer or death. Additionally, mothers in the transfer or death group showed higher rates of antiepileptic and antidiabetic drug use (Table 1).

XSL•FO
RenderX

Table . Sociodemographic, clinical, and pregnancy characteristics of mothers in Lombardy from 2016 - 2019.

Cohort characteristics	Live births (N=65,083)	Individuals not transferred and/or experienced death (n=64,348)	Individuals transferred and/or experienced death (n=735)	*P* value[a]
Sociodemographic characteristics[b]				
Age (years), mean (SD)	32.27 (5.37)	32.27 (5.36)	32.17 (5.49)	.64
Educational attainment, years of study[c], n (%)				<.001
≥14	19,036 (29.25)	18,859 (29.31)	177 (24.08)	
6 - 13	29,720 (45.66)	29,385 (45.67)	335 (45.58)	
≤5	16,327 (25.09)	16,104 (25.03)	223 (30.34)	
Employed, n (%)				<.001
No	22,832 (35.08)	41,835 (65.01)	416 (56.60)	
Yes	42,251 (64.92)	22,513 (34.99)	319 (43.40)	
Married, n (%)				.99
No	24,695 (37.94)	39,932 (62.06)	456 (62.04)	
Yes	40,388 (62.06)	24,4416 (37.94)	279 (37.96)	
Italian nationality, n (%)				.01
No	18,109 (27.82)	46,474 (72.22)	500 (68.03)	
Yes	46,974 (72.18)	17,874 (27.78)	235 (31.97)	
Maternal comorbidities[d], n (%)				
Depression and anxiety	394 (0.61)	387 (0.60)	7 (0.95)	.22
Preeclampsia	44 (0.07)	43 (0.07)	1 (0.14)	.39
Hypertension	90 (0.14)	87 (0.14)	3 (0.41)	.08
Diabetes	219 (0.34)	214 (0.33)	5 (0.68)	.10
Obesity or overweight	70 (0.11)	69 (0.11)	1 (0.14)	.55
Psychoses[e]	120 (0.18)	116 (0.18)	4 (0.54)	.048
Neuropathic, nonneuropathic, and other pain	140 (0.22)	139 (0.22)	1 (0.14)	>.99
Substance dependence	21 (0.03)	19 (0.03)	2 (0.27)	.02
Concomitant medication, n (%)				
Antidepressants	1395 (2.14)	1375 (2.14)	20 (2.72)	.28
Psycholeptics	159 (0.24)	155 (0.24)	4 (0.54)	.11
Antiepileptics	426 (0.65)	416 (0.65)	10 (1.36)	.02
Antihypertensive drugs	824 (1.27)	811 (1.26)	13 (1.77)	.22
Antidiabetic drugs	751 (1.15)	729 (1.13)	22 (2.99)	<.001
NSAIDs[f]	2440 (3.75)	2406 (3.74)	34 (4.63)	.21
Health care utilization, n (%)				
Hospitalization	8033 (12.34)	7935 (12.33)	98 (13.33)	.41
Number of distinct prescription drugs (≥1)	48,386 (74.35)	47,827 (74.33)	559 (76.05)	.29
Pregnancy characteristics[b], n (%)				
Gestational duration (weeks), mean (SD)	38.99 (1.44)	39.01 (1.41)	37.33 (3)	<.001
Multiple pregnancy	877 (1.35)	846 (1.31)	31 (4.22)	<.001
Parity				<.001

XSL•FO
RenderX

Cohort characteristics	Live births (N=65,083)	Individuals not transferred and/or experienced death (n=64,348)	Individuals transferred and/or experienced death (n=735)	*P* value[a]
Nulliparous	29,284 (44.99)	28,888 (44.89)	384 (52.24)	
Others	35,799 (55.01)	35,460 (55.11)	351 (47.76)	
Secondary outcomes, n (%)				
Low Apgar score at 5 minutes	305 (0.47)	205 (0.32)	100 (13.61)	<.001
Complete adherence to ANC	41,317 (63.48)	40,947 (63.63)	370 (50.34)	<.001
Antenatal maternal care[g]				
Appropriateness of gynecological visits	58,341 (89.64)	57,709 (89.68)	632 (85.99)	.001
Promptness of gynecological visits	61,338 (94.25)	60,650 (94.25)	688 (93.61)	.45
Appropriateness of ultrasound examinations	60,997 (93.72)	60,317 (93.74)	680 (95.52)	.18
Appropriateness of laboratory tests, n (%)				<.001
Complete adherence	46,718 (71.78)	46,298 (71.95)	420 (57.14)	
Partial adherence	16,622 (25.54)	16,353 (25.41)	269 (36.60)	
Nonadherence	1743 (2.68)	1697 (2.64)	46 (6.26)	

[a]*t* test, Chi-square, or Fisher exact test, as appropriate.

[b]Data related to the current pregnancy.

[c]≥14: at least bachelor's degree; 6 - 13: lower or upper secondary education; ≤5: primary education or none.

[d]Measured from one year before pregnancy through 22 weeks of gestation.

[e]Included diagnoses of migraine/headache, epilepsy, bipolar disorder, personality disorder, other psychiatric disorders, psychosis or schizophrenia, and sleep disorder.

[f]NSAID: nonsteroidal anti-inflammatory drug.

[g]ANC: antenatal care.

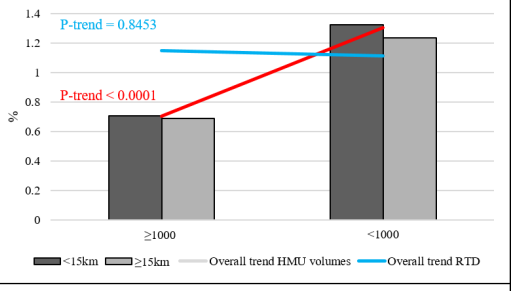
Figure 2 illustrates the distribution of outcomes by HMU volume and RTD. The rate of the primary outcome (transfer or death) significantly increased with decreasing HMU volume (*P* trend<.001). Hospitals with fewer than 500 births/year had a rate of 2.41%, compared with 0.72% in hospitals with 1500 births/year (Figure 2A). No significant differences were observed for RTD (*P* trend=.84). However, the lowest rate of the primary outcome was consistently observed in births with both high HMU volumes (≥1000 births/year) and short RTD (<15 km), representing the "best scenario" compared with all other combinations of volume and distance. Although too uncommon for detailed analysis across all categories, low Apgar scores were more prevalent in low HMUs with longer RTD (*P* trend<.002) (Figure 2B). Complete adherence to ANC decreased with lower HMU volume and increased with shorter RTD (*P* trend<.001, Figure 2C).

XSL•FO
RenderX

Figure 2. Distribution of maternal and neonatal outcomes based on HMU volume and RTD in Lombardy (N=65,083). (A) Transfer to a different hospital or death; (B) low Apgar score; (C) ANC adherence. ANC: antenatal care; HMU: hospital maternity unit; RTD: road travel distance.

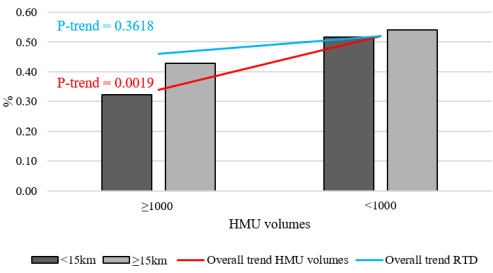
A Transfer to a different hospital or death

		HMU volumes — births/year				
	transfer-death / N (%)	≥1500	1000-1500	500-1000	<500	Marginal
RTD	<5 km	14 / 1572 (0.89)	43 / 6446 (0.67)	156 / 13,871 (1.12)	21 / 913 (2.30)	234 / 22,802 (1.03)
	5 - 15 km	29 / 4471 (0.65)	28 / 3602 (0.78)	249 / 19,231 (1.29)	44 / 1417 (3.11)	350 / 28,721 (1.22)
	15 - 21.5 km	2 / 354 (0.56)	6 / 1305 (0.46)	58 / 4966 (1.17)	7 / 436 (1.61)	73 / 7061 (1.03)
	≥21.5 km	3 / 229 (1.31)	10 / 1149 (0.87)	49 / 4238 (1.16)	16 / 883 (1.81)	78 / 6499 (1.20)
	Marginal	48 / 6626 (0.72)	87 / 12,502 (0.70)	512 / 42,306 (1.21)	88 / 3649 (2.41)	735 / 65,083 (1.13)



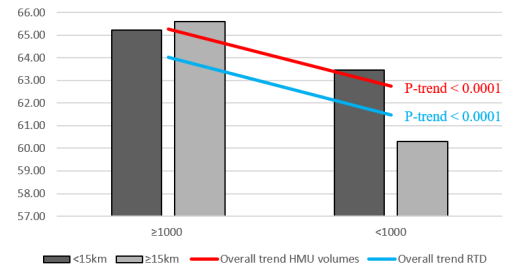
B Low Apgar score

		HMU volumes — births/year				
	low Apgar at 5 m. / N (%)	≥1500	1000-1500	500-1000	<500	Marginal
RTD	<5 km	7 / 1572 (0.45)	17 / 6446 (0.26)	64 / 13,871 (0.46)	4 / 913 (0.44)	92 / 22,802 (0.40)
	5 - 15 km	16 / 4471 (0.36)	12 / 3602 (0.33)	109 / 19,231 (0.57)	6 / 1417 (0.42)	143 / 28,721 (0.50)
	15 - 21.5 km	2 / 354 (0.56)	4 / 1305 (0.31)	31 / 4966 (0.62)	2 / 436 (0.46)	39 / 7061 (0.55)
	≥21.5 km	0 / 229 (0)	7 / 1149 (0.61)	19 / 4238 (0.45)	5 / 883 (0.57)	31 / 6499 (0.48)
	Marginal	25 / 6626 (0.38)	40 / 12,502 (0.32)	223 / 42,306 (0.53)	17/ 3649 (0.47)	305 / 65,083 (0.47)



C ANC adherence

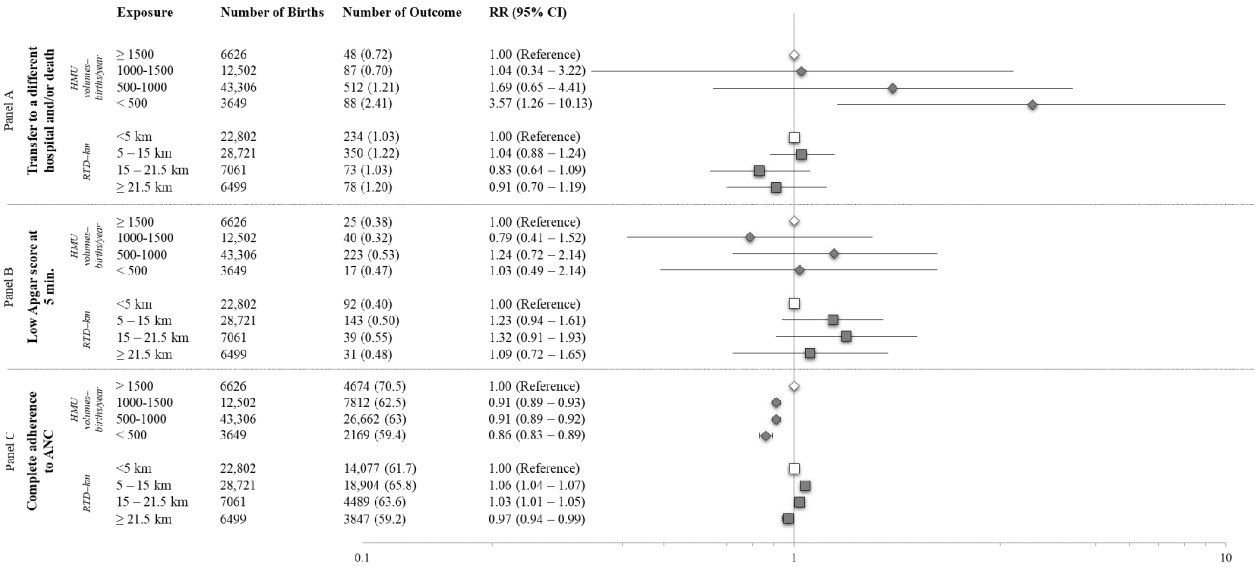
		HMU volumes — births/year				
	ANC adherence / N (%)	≥1500	1000-1500	500-1000	<500	Marginal
RTD	<5 km	1127 / 1572 (71.7)	3937 / 6446 (61.1)	8485 / 13,871 (61.2)	528 / 913 (57.8)	14,077 / 22,802 (61.7)
	5 - 15 km	3142 / 4471 (70.3)	2288 / 3602 (63.5)	12,647 / 19,231 (65.8)	827 / 1417 (58.4)	18,904 / 28,721 (65.8)
	15 - 21.5 km	246 / 354 (69.5)	855 / 1305 (65.5)	3116 / 4966 (62.7)	272 / 436 (62.4)	4489 / 7061 (63.6)
	≥21.5 km	159 / 229 (69.4)	732 / 1149 (63.7)	2414 / 4238 (57)	542 / 883 (61.4)	3847 / 6499 (59.2)
	Marginal	4674 / 6626 (70.5)	7812 / 12,502 (62.5)	26,662 / 42,306 (63)	2169/ 3649 (59.4)	41,317 / 65,083 (63.5)



The association between low HMU volume and the primary outcome was confirmed by log-binomial regression models (Figure 3). Newborns born in maternity units with fewer than 500 births/year highlighted had up to a 3-fold higher risk of being transferred or dying compared with those delivered in hospitals with 1500 or more births/year (adjusted RR 1.69, 95% CI 0.65-4.41) and newborns born in hospitals with an HMU volume between 500 - 1000 and <500 (adjusted RR 3.57, 95% CI 1.26-10.13), respectively (Figure 3A). In contrast, no statistically significant associations were observed when evaluating the RTD and its interaction with volumes while considering the low Apgar score as a secondary outcome (Figure 3B). Conversely, the RR of adhering to ANC recommendations decreased with lower HMU volume. Women who delivered in maternity units with 1000 - 1500 births/year had a 9% (95% CI 7% - 11%) lower adherence rate, whereas those in units with <500 births/year had a 14% (95% CI 11% - 17%) lower adherence rate compared to high-volume hospitals. Additionally, longer RTD was associated with a 3% (95% CI 1% - 6%) lower adherence rate (Figure 3C).

XSL•FO
RenderX

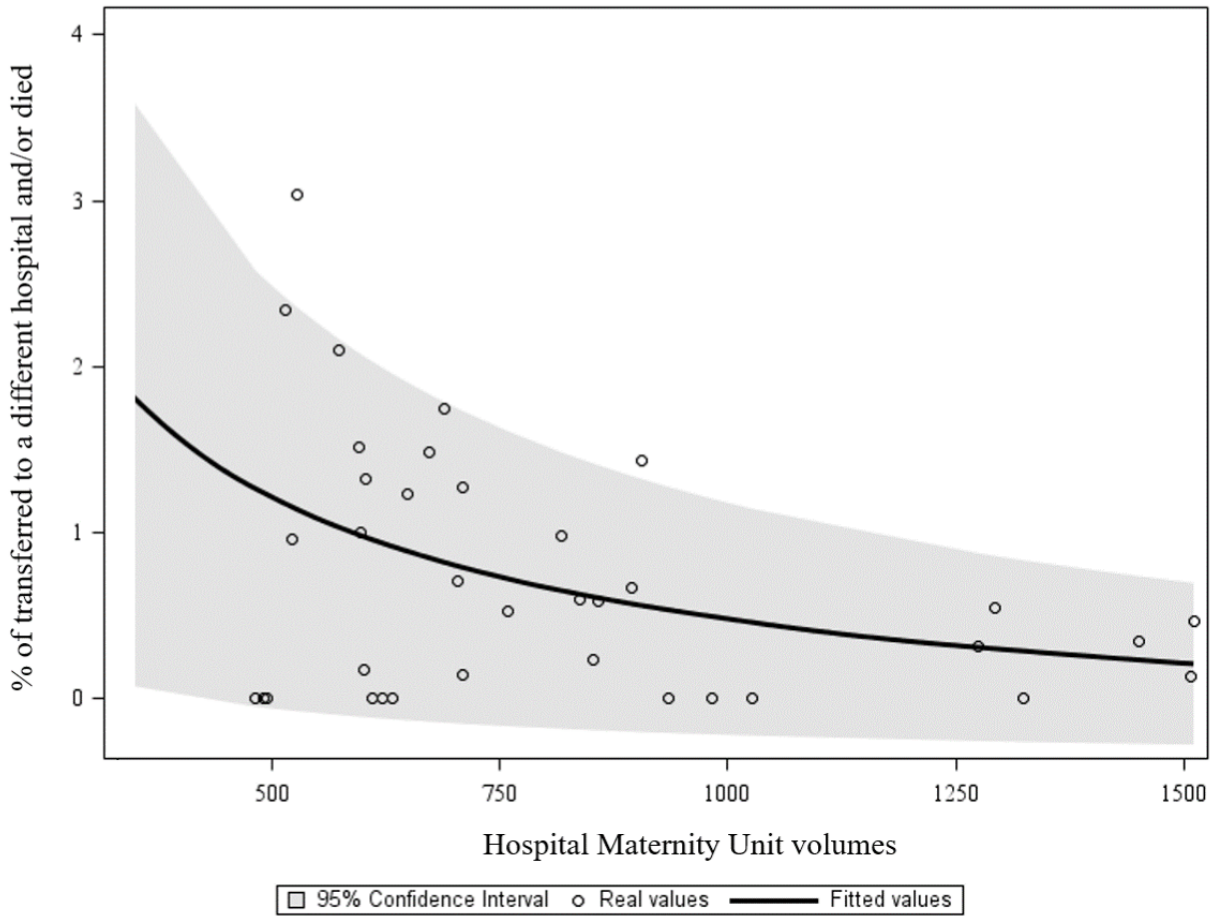
Figure 3. Association between HMU volumes, RTD, and primary and secondary outcomes in Lombardy in 2016 - 2019 (N=65,083). (A) transfer to a different hospital or death; (B) low Apgar score; (C) ANC adherence. ANC: antenatal care; HMU: hospital maternity unit; RTD: road travel distance.



These results were confirmed through a stratified analysis by hospital birth volume (high: ≥1000 births/year, low: <1000 births/year) (data not shown). Figure 4 shows the distribution of the percentage of newborns who were transferred to a different hospital or died as a function of HMU volumes. The FPM identified was 1x+ln (x)*1x, where x represents the mean number of births in each hospital in 2019. Each point on the graph represents a hospital included in our cohort. The solid black line demonstrates a clear decreasing trend in the frequency of the primary outcome (ie, percentage of individuals transferred to a different hospital and/or died), ranging from 2% in hospitals with lower activity volumes to 0.5% in those with higher activity volumes. There is also a specific trend of the curve flattening at an HMU volume of 1000 births/year or greater, where the frequency of the primary outcome was around 0.5%.

XSL•FO
RenderX

Figure 4. Distribution of the percentage of newborns who were transferred to a different hospital or died as a function of HMU volumes in Lombardy in 2016 - 2019.



Discussion

Our findings showed an excess risk of both neonatal transfer and death for live births delivered in HMUs with low activity volumes, comparable to services without ICUs. In contrast, the distance between the mothers' residences and hospital locations did not affect the primary outcome. Neither HMU volume nor RTD were significantly associated with low Apgar scores. However, mothers referred to hospitals with higher activity volumes and shorter travel distances demonstrated better adherence to ANC pathways, suggesting that being assisted by a lower-volume and distant hospital is associated with poorer ANC and correlates with a higher risk of neonatal problems. Finally, the data show a flattening trend in the prevalence of transfer to a different hospital or death in hospitals with an activity volume of 1000 births/year or more. This finding suggests this cutoff as the acceptable number of births for each maternity hospital to ensure an appropriate standard of care at delivery.

Maternity care reconfiguration is a complex and controversial issue in health care. It involves consolidating or closing small, less specialized maternity units and centralizing childbirth services at larger, more specialized hospitals. There are a number of reasons supporting the regionalization of HMUs. These include (1) declining birth rates, (2) improved safety and outcomes, and (3) saving costs.

In many countries, including Italy, the birth rate has declined for several decades. This has reduced the number of births—mainly at already small hospitals—increasing the risk of inefficient operation, low occupancy rates, and increasing maternal and neonatal adverse outcomes. However, regionalization of perinatal care could be the right solution to improve outcomes for mothers and newborns. For example, in Portugal and many other countries, regionalization involved the closure of a massive number of HMUs with fewer than 1000 births/year. This has consistently reduced several maternal and neonatal adverse outcomes such as the stillbirth rate, which ranged from 4 per 1000 births in 2000 to 2.2 per 1000 births in 2021 [17]. In contrast, in Italy, despite a decreased number of births, the stillbirth rate has remained consistent over the years, ranging from 2.4 to 2.2 per 1000 births during the same period. However, several studies have shown that women who give birth at large hospitals with high volumes of deliveries tend to have better outcomes, including lower rates of complications and infant mortality [18]. Our results are consistent with findings from other studies that reported significantly lower rates of

XSL•FO
RenderX

stillbirths and neonatal mortality in both rural and urban regions after the closure of low-volume HMUs [19].

While data indicate lower neonatal mortality rates in high-volume birth hospitals at an early age, closing low-volume units requires extreme caution due to potential disadvantages, and is the subject of ongoing debate. Studies have raised concerns about increased unplanned out-of-hospital births and high neonatal mortality and stillbirth rates following closures [20,21]. Furthermore, studies in vast rural areas with limited access to perinatal care have revealed high rates of adverse birth outcomes and increased stress and anxiety among pregnant women [22,23]. However, our results did not show an association between long distances traveled to access perinatal care and the analyzed neonatal outcomes, except among hospitals with a low volume of activity and for full adherence to ANC.

Moreover, centralizing childbirth services can help to reduce health care costs by eliminating duplication of services and improving the efficient use of resources. However, there are several potential drawbacks to maternity care reconfiguration including increased travel times for women (if women must travel farther and for a longer time to give birth, this can increase their risk of complications), which may make receiving adequate prenatal care more difficult. Unfortunately, in this study, we were only able to calculate the distance to access perinatal care, which is a proxy for travel time rather than the time travel itself. Moreover, the closure of small maternity units can leave some women with limited access to childbirth services, especially in rural or geographically disadvantaged areas. It must be considered that access to ANC is often guaranteed in nonhospital settings (eg, public clinics) and that low-level hospitals can offer adequate ANC while organizing a structured referral to large hospitals for delivery. High-performing hospitals and well-equipped neonatal ICUs can also manage critical situations without the need to transfer newborns.

The decision to regionalize perinatal care units is complex and must be made on a case-by-case basis, considering all relevant factors. There is no one-size-fits-all solution that would be suitable for all communities, and health care providers and policy makers should consider several factors before making decisions about maternity care reconfiguration. These include the specific needs of the community, availability of transportation, and access to specialized care.

A strength of our study is the large, population-based sample reporting real-world evidence on the associations between HMU volumes and RTD, and the transfer to a different hospital or death (identified through a robust definition), low Apgar score at 5 minutes, and complete adherence to ANC. Moreover, we considered the individual prenatal care pathway for each mother, strengthening our analysis by addressing a potential confounding factor.

A limitation of our study is the lack of data on the actual time to travel to HMUs, as only data on RTD were available; we acknowledge that factors such as traffic congestion, mode of transportation, and individual driving habits could influence actual travel times. In addition, due to privacy concerns, we lack data on the exact location of the mothers' residence; therefore, we used the mothers' residence municipality for the calculation of the RTD. This misclassification might have reduced the difference in outcomes between RTD categories, thus underestimating the effect of RTD on outcomes. However, we believe that the RTD data used in our analysis provides a valuable approximation of accessibility, particularly when considered in conjunction with other factors such as HMU volumes and neonatal outcomes. Moreover, we were not able to consider several other factors, such as smoking, alcohol, illicit drug use, and social factors related to migration status that may affect access to both ANC and the HMU and influence the outcomes of interest. Additionally, the exclusion of mothers who experienced a stillbirth or those with missing information for sociodemographic or clinical characteristics likely affected less healthy women. However, missing data were infrequent (at most 0.4% for employment status). Nevertheless, outcome prevalence did not vary between complete and incomplete cases, supporting the hypothesis of data "missing at random" (data not shown). In our previous study, we showed that multiple imputations did not change the complete-case analysis results [24]. Finally, our results are derived from Lombardy, a specific Italian region, which may impact the generalizability of results. However, we believe that our study's key insights can apply to a broader range of settings, particularly those with similar health care systems and demographic characteristics. However, cultural beliefs and preferences regarding childbirth and health care can affect the decision-making process for mothers and health care providers. While our study provides valuable insights into the relationship between HMU volume, RTD, and neonatal outcomes in Lombardy, it is important to consider these factors when applying the findings to other regions. Several studies have reported that regionalized care enhances patient outcomes by concentrating specialized expertise at high-volume centers and fostering collaboration among health care providers within a defined region [25]. Moreover, studies have consistently demonstrated the benefits of regionalized care for high-risk infants, particularly those born very or late preterm and very-low-birth weight (<32 weeks of gestation) [26,27]. Further research may be needed to assess the generalizability of our results in diverse settings.

In conclusion, our results support the requirement outlined in the Italian Ministry of Health's decree 70/2015, which mandates at least 1000 deliveries per year in HMUs as a standard. It is necessary to emphasize that maternal-newborn separation caused by the transfer of the newborn can be incredibly detrimental to their development and bonding, making neonatal transfer a major concern [28]. Therefore, it is important to have a comprehensive plan to mitigate any potential negative impacts of maternity care reorganization. This plan should include measures to ensure that women have access to appropriate prenatal care and that they are able to give birth in a safe and supportive environment, regardless of their location, as advocated by the World Health Organization.

XSL•FO
RenderX

Acknowledgments

RZ and GC contributed equally to this work. The project was funded by a grant from the Italian Ministry of Education, University and Research 'Fondo d'Ateneo per la Ricerca' portion, year 2023 (grant number 2023-ATE-0565), and a grant from the Italian Ministry of Health ("Metodi per il monitoraggio e la valutazione dell'assistenza sanitaria nell'ambito del Nuovo Sistema di Garanzia, con particolare riferimento alle cure integrate e al confronto dei risultati pre e post pandemia da CoViD-19" project, grant number J59H06000160001). The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

GC received research support from the European Community, the Italian Agency of Drug, and the Italian Ministry for University and Research.

Data Availability

Data may be obtained from a third party and are not publicly available.

Conflicts of Interest

GC took part in various projects funded by pharmaceutical companies (ie, Novartis, GSK, Roche, AMGEN, and BMS). He also received honoraria as a member of the advisory board from Roche.
The other authors have no conflicts of interest to declare. We affirm that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of any organization or company.

References

1. Rashidian A, Omidvari AH, Vali Y, et al. The effectiveness of regionalization of perinatal care services--a systematic review. Public Health (Fairfax) 2014 Oct;128(10):872-885. [doi: 10.1016/j.puhe.2014.08.005] [Medline: 25369352]
2. Barfield WD, Papile LA, Baley JE, American Academy of Pediatrics Committee on Fetus and Newborn, et al. Levels of neonatal care. Pediatrics 2012 Sep 1;130(3):587-597. [doi: 10.1542/peds.2012-1999] [Medline: 23641169]
3. Promoting effective perinatal care (PEPC) in the European region. World Health Organization. 2000 Nov. URL: https://iris.who.int/bitstream/handle/10665/108550/WHO-EURO-2000-4052-43811-61678-eng.pdf?sequence=1&isAllowed=y
4. Six success stories in making pregnancy safer. World Health Organization. 2010. URL: https://iris.who.int/bitstream/handle/10665/374844/WHO-EURO-2010-8879-48651-72242-eng.pdf?sequence=1&isAllowed=y [accessed 2022-02-20]
5. Aubrey-Bassler FK, Cullen RM, Simms A, et al. Population-based cohort study of hospital delivery volume, geographic accessibility, and obstetric outcomes. Int J Gynaecol Obstet 2019 Jul;146(1):95-102. [doi: 10.1002/ijgo.12832] [Medline: 31032903]
6. Darling EK, Lawford KMO, Wilson K, Kryzanauskas M, Bourgeault IL. Distance from home birth to emergency obstetric services and neonatal outcomes: a cohort study. J Midwif Womens Health 2019 Mar;64(2):170-178. [doi: 10.1111/jmwh.12896] [Medline: 30325580]
7. Pfniss I, Gold D, Holter M, et al. Birth during off-hours: impact of time of birth, staff´s seniority, and unit volume on maternal adverse outcomes-a population-based cross-sectional study of 87 065 deliveries. Birth 2023 Jun;50(2):449-460. [doi: 10.1111/birt.12663] [Medline: 35789033]
8. Handley SC, Lorch SA. Regionalization of neonatal care: benefits, barriers, and beyond. J Perinatol 2022 Jun;42(6):835-838. [doi: 10.1038/s41372-022-01404-7] [Medline: 35461330]
9. Lupattelli A, Corrao G, Gatti C, Rea F, Trinh NTH, Cantarutti A. Antidepressant continuation and adherence in pregnancy, and risk of antenatal hospitalization for unipolar major depressive and/or anxiety disorders. J Affect Disord 2023 Oct;339:502-510. [doi: 10.1016/j.jad.2023.07.066] [Medline: 37459974]
10. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
11. Matrici di contiguità, distanza e pendolarismo. Instituto Nazionale di Statistica. 2004 May 9. URL: https://www.istat.it/it/archivio/157423 [accessed 2024-02-20]
12. Casey BM, McIntire DD, Leveno KJ. The continuing value of the Apgar score for the assessment of newborn infants. N Engl J Med 2001 Feb 15;344(7):467-471. [doi: 10.1056/NEJM200102153440701] [Medline: 11172187]
13. Corrao G, Cantarutti A, Locatelli A, et al. Association between adherence with recommended antenatal care in low-risk, uncomplicated pregnancy, and maternal and neonatal adverse outcomes: evidence from Italy. Int J Environ Res Public Health 2020 Dec 29;18(1):173. [doi: 10.3390/ijerph18010173] [Medline: 33383661]

XSL•FO
RenderX

14. Schneeweiss S, Seeger JD, Maclure M, Wang PS, Avorn J, Glynn RJ. Performance of comorbidity scores to control for confounding in epidemiologic studies using claims data. Am J Epidemiol 2001 Nov 1;154(9):854-864. [doi: 10.1093/aje/154.9.854] [Medline: 11682368]
15. Determinazione 20 marzo 2008 [Article in Italian]. Gazzetta Ufficiale. URL: https://www.gazzettaufficiale.it/eli/id/2008/03/31/08A02109/sg [accessed 2024-02-20]
16. Royston P, Sauerbrei W. Stability of multivariable fractional polynomial models with selection of variables and transformations: a bootstrap investigation. Stat Med 2003 Feb 28;22(4):639-659. [doi: 10.1002/sim.1310] [Medline: 12590419]
17. Stillbirth rate (per 1,000 total births). World Bank Group. URL: https://genderdata.worldbank.org/indicators/sh-dyn-stlb/?year=2000 [accessed 2024-02-20]
18. Walther F, Kuester D, Bieber A, Malzahn J, Rüdiger M, Schmitt J. Are birth outcomes in low risk birth cohorts related to hospital birth volumes? A systematic review. BMC Pregnancy Childbirth 2021 Jul 27;21(1):531. [doi: 10.1186/s12884-021-03988-y] [Medline: 34315416]
19. Allen VM, Jilwah N, Joseph KS, et al. The influence of hospital closures in Nova Scotia on perinatal outcomes. J Obstet Gynaecol Can 2004 Dec;26(12):1077-1085. [doi: 10.1016/s1701-2163(16)30435-2] [Medline: 15607044]
20. Lorch SA, Srinivas SK, Ahlberg C, Small DS. The impact of obstetric unit closures on maternal and infant pregnancy outcomes. Health Serv Res 2013 Apr;48(2 Pt 1):455-475. [doi: 10.1111/j.1475-6773.2012.01455.x] [Medline: 22881056]
21. Kozhimannil KB, Hung P, Henning-Smith C, Casey MM, Prasad S. Association between loss of hospital-based obstetric services and birth outcomes in rural counties in the United States. JAMA 2018 Mar 27;319(12):1239-1247. [doi: 10.1001/jama.2018.1830] [Medline: 29522161]
22. Grzybowski S, Stoll K, Kornelsen J. Distance matters: a population based study examining access to maternity services for rural women. BMC Health Serv Res 2011 Jun 10;11:147. [doi: 10.1186/1472-6963-11-147] [Medline: 21663676]
23. Kornelsen J, Stoll K, Grzybowski S. Stress and anxiety associated with lack of access to maternity services for rural parturient women. Aust J Rural Health 2011 Feb;19(1):9-14. [doi: 10.1111/j.1440-1584.2010.01170.x] [Medline: 21265919]
24. Cantarutti A, Merlino L, Giaquinto C, Corrao G. Use of antidepressant medication in pregnancy and adverse neonatal outcomes: a population-based investigation. Pharmacoepidemiol Drug Saf 2017 Sep;26(9):1100-1108. [doi: 10.1002/pds.4242] [Medline: 28594145]
25. Lorch SA, Myers S, Carr B. The regionalization of pediatric health care. Pediatrics 2010;126:1182. [doi: 10.1542/peds.2010-1119] [Medline: 21041285]
26. Lasswell SM, Barfield WD, Rochat RW, Blackmon L. Perinatal regionalization for very low-birth-weight and very preterm infants. JAMA 2010 Sep 1;304(9):992. [doi: 10.1001/jama.2010.1226] [Medline: 20810377]
27. Salazar EG, Passarella M, Formanowski B, Phibbs CS, Lorch SA, Handley SC. The impact of volume and neonatal level of care on outcomes of moderate and late preterm infants. J Perinatol 2024 Oct;44(10):1409-1415. [doi: 10.1038/s41372-024-01901-x] [Medline: 38413758]
28. Bergman NJ. Birth practices: maternal-neonate separation as a source of toxic stress. Birth Defects Res 2019 Sep 1;111(15):1087-1109. [doi: 10.1002/bdr2.1530] [Medline: 31157520]

Abbreviations

ANC: antenatal care
FPM: fractional polynomial model
HMU: hospital maternity unit
ICU: intensive care unit
NHS: National Health Service
RR: relative risk
RTD: road travel distance
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Mavragani, T Sanchez; submitted 28.03.24; peer-reviewed by WH Dong, Y Liu; revised version received 22.10.24; accepted 29.10.24; published 21.01.25.

Please cite as:
Cantarutti A, Boracchini R, Bellù R, Ronco R, Rea F, Locatelli A, Zanini R, Corrao G
Assessing the Impact of Distance Traveled and Birth Volumes of Hospital Maternity Units on Newborn Outcomes: Population-Based Cohort Study
JMIR Public Health Surveill 2025;11:e58944
URL: https://publichealth.jmir.org/2025/1/e58944
doi:10.2196/58944

XSL•FO
RenderX

© Anna Cantarutti, Riccardo Boracchini, Roberto Bellù, Raffaella Ronco, Federico Rea, Anna Locatelli, Rinaldo Zanini, Giovanni Corrao. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 21.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Surveillance of Mental Health Care Services in Saudi Arabia: Cross-Sectional Study of National e-Referral System Data

Abdullah A Alharbi[1], PhD; Nawfal A Aljerian[2,3], SBEM; Meshary S Binhotan[4,5], PhD; Hani A Alghamdi[6], PhD; Ali K Alsultan[2], EMBA; Mohammed S Arafat[2], EMBA; Abdulrahman Aldhabib[2], MBBS; Yasser A Alaska[7], MBBS; Eid B Alwahbi[2], MSc; Mohammed A Muaddi[1], PhD; Ahmad Y Alqassim[1], PhD; Ronnie D Horner[8], PhD

[1]Department of Family and Community Medicine, Faculty of Medicine, Jazan University, Jazan, Saudi Arabia

[2]Medical Referrals Centre, Ministry of Health, Riyadh, Saudi Arabia

[3]Department of Emergency Medicine, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia

[4]Department of Emergency Medical Services, College of Applied Medical Sciences, King Saud bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia

[5]King Abdullah International Medical Research Center, Riyadh, Saudi Arabia

[6]Department of Family and Community Medicine, College of Medicine, King Saud University, Riyadh, Saudi Arabia

[7]Department of Emergency Medicine, King Saud University, Riyadh, Saudi Arabia

[8]Department of Health Services Research and Administration, College of Public Health, University of Nebraska Medical Center, Omaha, NE, United States

Corresponding Author:
Mohammed A Muaddi, PhD
Department of Family and Community Medicine, Faculty of Medicine, Jazan University, Jazan, Saudi Arabia

Abstract

Background: Mental illness affects an estimated 25% of the global population, with treatment gaps persisting worldwide. The COVID-19 pandemic has exacerbated these challenges, leading to a significant increase in mental health issues globally. In Saudi Arabia, the lifetime prevalence of mental disorders is estimated at 34.2%, yet 86.1% of those with a 12-month mental disorder report no service use. To address these challenges, digital health solutions, particularly electronic referral (e-referral) systems, have emerged as powerful tools to improve care coordination and access. Saudi Arabia has pioneered the nationwide Saudi Medical Appointments and Referrals Centre (SMARC), a centralized e-referral system using artificial intelligence and predictive analytics.

Objectives: This study aims to analyze Saudi Arabia's novel nationwide e-referral system for mental health services, using SMARC platform data to examine referral patterns, and service accessibility. This study also aims to demonstrate how digital health technology can inform and improve mental health care delivery and policy making.

Methods: This retrospective, cross-sectional study used secondary data from SMARC on 10,033 psychiatric e-referrals in Saudi Arabia during 2020 - 2021. Referrals were assessed by patient sociodemographic variables, geographic data, and e-referral characteristics including date, type, bed type, and reason for e-referral. Descriptive statistical analyses identified referral patterns, while regression modeling determined predictors of external referrals to other regions.

Results: Analysis of 10,033 psychiatric e-referrals revealed that 58.99% (n=5918) were for patients aged 18 - 44 years, 63.93% (n=6414) were for men, and 87.10% (n=8739) were for Saudi nationals. The Western Business Unit generated 45.17% (n=4532) of all e-referral requests. Emergency cases were the most common type of referral overall, followed by routine inpatient and routine outpatient department referrals. However, in the Northern Business Unit, routine inpatient referrals were most frequent. Two-thirds of requests were for ward beds, while critical beds were rarely requested. "Unavailable subspecialty" was the primary reason for referrals across all regions. The utilization of the mental health e-referral system varied across regions, with the Northern Border and Albaha regions showing the highest rates, while Madinah, Eastern, and Riyadh regions demonstrated lower use. Temporal analysis showed almost similar monthly patterns in 2020 and 2021. There was an overall increase in referrals in 2021 compared with 2020.

Conclusions: This pioneering study of mental health e-referrals in Saudi Arabia demonstrates how digital health transformation, particularly through an e-referral system, has significantly enhanced access to mental health services nationwide in Saudi Arabia. The success of this digital initiative demonstrates how digital health solutions can transform health care access, particularly in mental health services, offering a valuable model for other health care systems.

(JMIR Public Health Surveill 2025;11:e64257) doi:10.2196/64257

XSL•FO
RenderX

KEYWORDS

digital health; mental health; health policy; epidemiology; Saudi Arabia; SMARC; health care transformation; e-referral; Saudi Medical Appointments and Referrals Centre

Introduction

Mental illness remains a serious global health concern, with the World Health Organization estimating that it affects 25% of the population worldwide [1-3]. Significant treatment gaps persist globally, as around 70% of those needing mental health management lack access to care [4-8]. This high rate of untreated patients partly stems from specialist shortages and geographic barriers distancing patients from care [9,10], contributing to a major treatment gap. Furthermore, stigma and lack of mental health literacy continue to be significant barriers to seeking care, particularly in Middle Eastern countries [11].

The global landscape of mental health has evolved significantly in recent years, with the COVID-19 pandemic exacerbating existing challenges and creating new ones. Recent studies have shown that these gaps have widened during the COVID-19 pandemic, with a significant increase in mental health issues reported globally [12,13]. A meta-analysis estimated that the pandemic led to an additional 53.2 million cases of major depressive disorder and 76.2 million cases of anxiety disorders globally in 2020 alone [14]. These findings underscore the urgent need for scalable and accessible mental health interventions worldwide.

In response to these challenges, digital health solutions have emerged as powerful tools to improve coordination and access to mental health care. Electronic referral (e-referral) systems, in particular, have shown promise in reducing wait times, improving care coordination, and increasing access to specialist care [15,16]. The Kingdom of Saudi Arabia (KSA) has been an early pioneer in leveraging digital health through the nationwide Saudi Medical Appointments and Referrals Centre (SMARC). This centralized e-referral system was revamped and relaunched in 2019 to unify and streamline referrals across the country's 13 administrative regions. The advanced SMARC platform uses artificial intelligence and predictive analytics to coordinate referrals between public and private hospitals, addressing previous fragmentation. As of 2019, the majority of governmental and private Saudi health care facilities are connected to SMARC, enabling efficient nationwide referral management—a major milestone in the KSA's health sector digital transformation under Saudi vision 2030 [17,18].

Recent studies have highlighted significant challenges and opportunities to improve mental health care in Saudi Arabia. Estimates show that 34.2% of the general Saudi population experience a mental disorder during their lifetime [19]. Regarding depressive disorders, there was about a 60% increase in number of new cases from 1.42 million cases in 2011 to 2.28 million cases in 2021. The incidence rate also rose by around 23% during the same period, going from 4,913.71 to 6,039.77 per 100,000 people [20]. Despite this high prevalence, access to mental health care remains limited, with a national survey revealing that about 84% (597/711) of participants with a 12-month mental disorder reported no service use [21]. Barriers to care include low perceived need for help, cultural stigma, and attitudinal obstacles [22]. These findings emphasize the critical role that e-referral systems such as SMARC could play in addressing these challenges.

SMARC facilitates analysis of mental health patterns at a national scale, providing a unique opportunity for epidemiological and health services research on mental illnesses [1,2]. However, no large-scale studies have been conducted to date in the KSA, making such analysis a high priority [23]. This innovative study is the first to examine Saudi Arabia's adoption of an e-referral system for mental health services, establishing digital monitoring of mental health care referrals. By analyzing national referral trends using SMARC platform data, it provides crucial insights into care patterns and resource distribution. The research showcases Saudi Arabia's innovative use of technology to evaluate and enhance mental health service provision and accessibility. Our findings, derived from SMARC data, illuminate potential gaps in the mental health care system and could inform evidence-based policies and interventions. This aligns with the Saudi's digital-driven health care transformation, offering valuable insights to optimize mental health care delivery throughout Saudi Arabia. Ultimately, this study contributes to the broader goal of improving mental health outcomes by leveraging technology to guide strategic health care decisions and resource allocation.

Methods

Study Design and Setting

This descriptive study of psychiatric health care delivery used data by the SMARC system between January 2020 and December 2021. The SMARC system is concerned with all e-referrals across the KSA. Unlike conventional referral systems, this system primarily serves e-referral between secondary, tertiary, and specialized health care providers. The SMARC system routinely collects an array of data concerning each e-referral request.

A descriptive approach was chosen for this study due to several key factors. First, this is the first comprehensive analysis of the Saudi national e-referral system for mental health services, necessitating an exploratory approach to establish baseline patterns and trends. Second, the descriptive nature of our study provides a crucial foundation for generating hypotheses that can guide future in-depth research. Third, given the novelty of the data source and the lack of prior comprehensive analyses in this context, a descriptive approach allows for a broad understanding of the current landscape of mental health referrals in Saudi Arabia. Finally, our descriptive findings offer immediate, actionable insights for policy makers and health care planners, which is particularly valuable given the urgency of addressing mental health needs in the region.

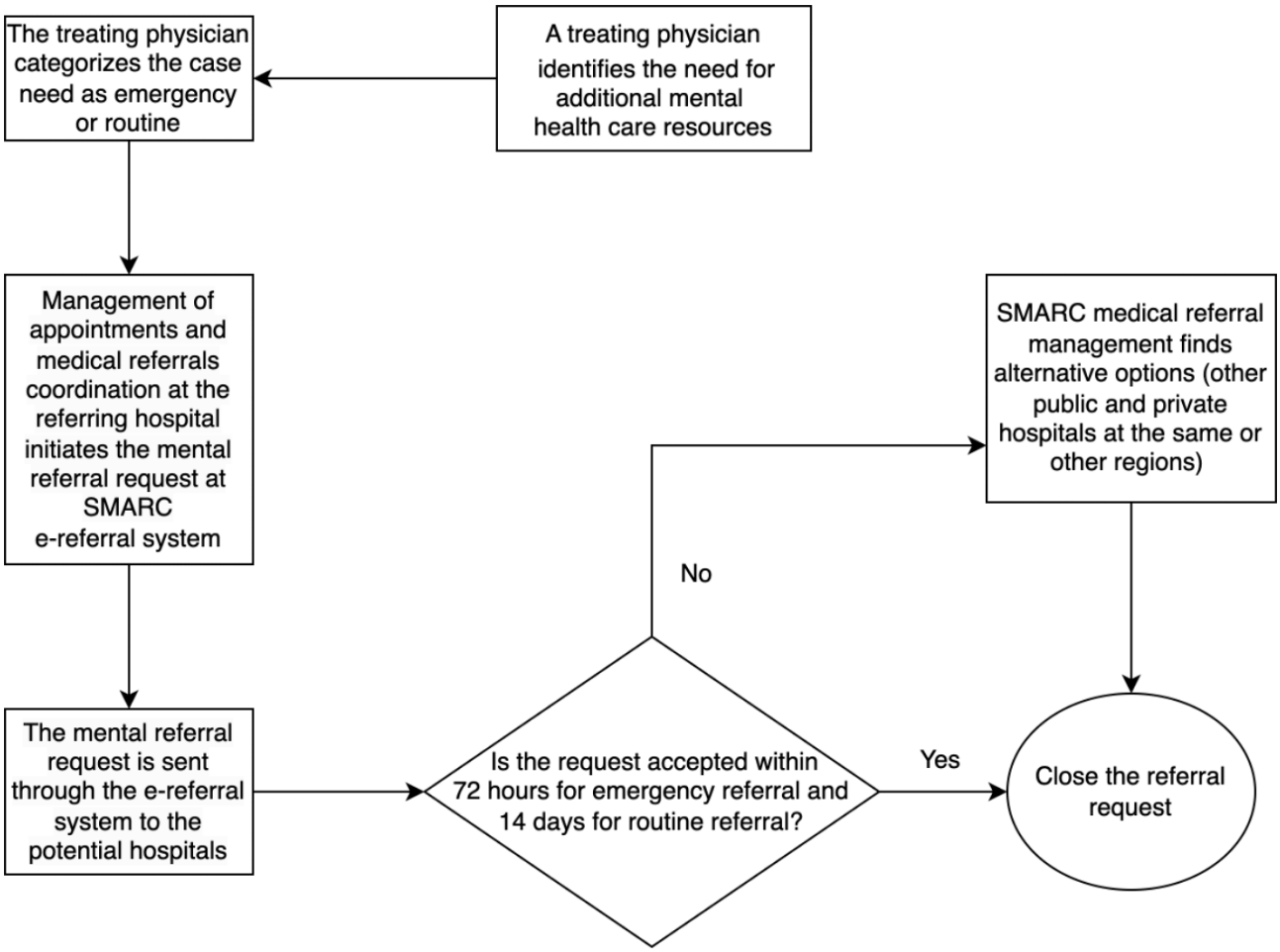
To initiate a referral request, a treating physician has to first identify the need and type of the referral, emergency or routine

XSL•FO
RenderX

(both inpatient and outpatient). The Office of Coordination and Eligibility for Treatment, which is allocated in every hospital, digitally uploads the referral request at SMARC e-referral system. The Office of Coordination and Eligibility for Treatment can identify 3 potential receiving hospitals or can leave the task to the e-referral system.

The referral request is sent, through the e-referral system, to the potential receiving hospitals for acceptance. The e-referral system has built-in time frames for requests to be accepted, depending on the referral type. For emergency requests, the system allows hospitals 72 hours to accept the request, while routine requests have 14 days. If the initially identified hospitals reject the request within the allotted time frame, further hospitals will receive the request for acceptance consideration. Once the allotted time frame elapses with no acceptance, the request will automatically be escalated to SMARC medical referral management team to secure an acceptance. The team will look at alternative options including other governmental and private hospitals, both at the same region and other regions if needed. During the process of referral initiation and acceptance, all patients continue receiving the necessary health care management at the referring hospital, to ensure stability, until being transferred to the receiving hospital. Figure 1 describes the process of mental referral request initiation and acceptance.

Figure 1. Flowchart depicting the process of initiating and accepting psychiatric e-referral requests through the SMARC in Saudi Arabia from January 2020 to December 2021. This cross-sectional study analyzed data on 10,033 psychiatric e-referrals across all regions of Saudi Arabia during this 2-year period. e-Referral: electronic referral; SMARC: Saudi Medical Appointments and Referrals Centre.



Study Variables

All e-referral data provided by the SMARC system were used in this analysis, totaling 10,033 referrals. The variables included sociodemographic variables such as age, sex, and nationality of the patient with the e-referral request, as well as the patient's administrative area of residence and business unit (BU) from which the request originated. BU refers to a geographical and administrative area designated for health care service delivery. The health care system comprises 5 BUs, each managed by a government-owned holding company. The units are distributed based on geographical coverage: Central BU includes Riyadh and AL Qassim; Western BU covers Makkah, Madinah, and Albaha; Eastern BU includes the Eastern region; Northern BU consists of Al-jouf, Northern Border, Tabuk, and Hail; and Southern BU includes Aseer, Jazan, and Najran. Each BU is tasked with managing and providing health care services within its respective regions.

Other variables pertain to the e-referral request itself, including the date, type, bed type, and reason. The type of e-referral includes routine outpatient department (OPD), routine inpatient, and emergency. Routine referrals (both OPD and inpatient) are those that do not require prompt intervention and are supposed to be accepted within 14 days. An emergency e-referral is a case requiring medical or surgical intervention within 72 hours to

XSL•FO
RenderX

avoid mortality or morbidity. The bed type refers to OPD, ward bed, or critical bed. A critical bed indicates a case needing admission to the intensive care unit. Reasons for e-referral include unavailable specialty, unavailable physician, unavailable machine (meaning no tools), and unavailable bed (meaning no available bed at the hospital).

Statistical Analysis

Descriptive statistics were calculated for all variables, including frequencies and percentages. To explore potential regional differences, we conducted cross-tabulations of each variable across the 5 BUs and performed χ^2 tests to assess associations. STATA (version 16; StataCorp LLC) was used for all analyses. Furthermore, we generated color-coded maps using ArcGIS software (ArcGIS 10.0; Environmental Systems Research Institute, Inc) to visually represent the distribution of e-referrals requests across different administrative areas.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki and approved by the institutional review board of the Ministry of Health (MOH) (IRB log No. 23 - 77-E; date: September 20, 2023). The requirement for individual informed consent was waived due to the retrospective nature of the study and the use of deidentified data. Patient privacy and confidentiality were strictly maintained throughout the research process. All data obtained from the SMARC system were anonymized prior to analysis. Access to the dataset was restricted to authorized research team members and stored on secure, password-protected servers. The study did not involve any direct contact with patients or any intervention that could affect their care. We confirm that no images or information that could lead to the identification of individual participants is included in this manuscript or any supplementary materials.

Results

Sociodemographic Characteristics of Psychiatry Patients With e-Referral Requests

Of the 10,033 patients with a psychiatry-related e-referrals during the study period, 58.99% (n=5918) were aged 18 - 44 years, followed by those aged 45 - 64 years (n=1869, 18.63%). Patients younger than 18 years or older than 65 years comprised relatively few requests. Nearly two-thirds of requests were for men (6414/10,033, 63.93%) compared with women (3619/10,033, 36.07%). The majority (8739/10,033, 87.1%) were Saudi nationals. Approximately 45% (n=4532) of all e-referral requests originated from the Western BU, while the other regions each generated around one-third or less than this proportion (Table 1).

XSL•FO
RenderX

Table . Sociodemographic characteristics of psychiatry patients with e-referral requests in Saudi Arabia from January 2020 to December 2021[a].

Characteristics	Participants (N=10,033), n (%)
Age (years)	
<11	844 (8.41)
12 - 17	617 (6.15)
18 - 44	5918 (58.99)
45 - 64	1869 (18.63)
>65	785 (7.82)
Sex	
Male	6414 (63.93)
Female	3619 (36.07)
Nationality	
Non-Saudi	1294 (12.90)
Saudi	8739 (87.10)
Business unit	
Central	1128 (11.24)
Eastern	1220 (12.16)
Western	4532 (45.17)
Northern	1603 (15.98)
Southern	1550 (15.45)
Year	
2020	4179 (41.65)
2021	5854 (58.35)

[a]This cross-sectional study analyzed data on 10,033 psychiatric e-referral requests across the 5 business units (Central, Eastern, Western, Northern, and Southern) in Saudi Arabia between January 2020 and December 2021. Data used were the data routinely collected by the Saudi Medical Appointments and Referrals Centre (SMARC) system. The SMARC system is concerned with all e-referral requests in Saudi Arabia.

Sociodemographic Characteristics of Psychiatry Patients With e-Referral Requests According to BUs

Table 2 shows the sociodemographic characteristics of patients with a psychiatry-related e-referral request for each of the 5 BUs. In general, the patients being referred for psychiatric care in each BU were similar in characteristics to those identified for the population as a whole. These associations within regions were all significant at the $P<.001$ level.

XSL•FO
RenderX

Table . Sociodemographic profile of psychiatric e-referral patients across different business units in Saudi Arabia from January 2020 to December 2021[a].

Characteristics	Central (1128/10,033, 11.24%), n (%)	Eastern (1220/10,033, 12.16%), n (%)	Western (4532/10,033, 45.17%), n (%)	Northern (1603/10,033, 15.98%), n (%)	Southern (1550/10,033, 15.45%), n (%)
Age (years) (*P* <.001)					
<11	72 (6.38)	130 (10.66)	341 (7.52)	121 (7.55)	180 (11.61)
12 - 17	71 (6.29)	105 (8.61)	260 (5.74)	73 (4.55)	108 (6.97)
18 - 44	704 (62.41)	643 (52.70)	2878 (63.50)	829 (51.72)	864 (55.74)
45 - 64	230 (20.39)	253 (20.74)	799 (17.63)	334 (20.84)	253 (16.32)
>65	51 (4.25)	89 (7.30)	254 (5.60)	246 (15.35)	145 (9.35)
Sex (*P*<.001)					
Males	694 (61.52)	739 (60.57)	2998 (66.15)	998 (62.26)	988 (63.74)
Females	434 (38.48)	481 (39.43)	1534 (33.85)	605 (37.74)	562 (36.26)
Nationality (*P*<.001)					
Non-Saudi	107 (9.49)	106 (8.69)	658 (14.52)	189 (11.79)	234 (15.10)
Saudi	1021 (90.51)	1114 (91.31)	3874 (85.48)	1414 (88.21)	1316 (84.90)

[a]This cross-sectional study analyzed data on 10,033 psychiatric e-referral requests across the 5 business units (Central, Eastern, Western, Northern, and Southern) in Saudi Arabia between January 2020 and December 2021. Data used were the data routinely collected by the Saudi Medical Appointments and Referrals Centre (SMARC) system. The SMARC system is concerned with all e-referral requests in Saudi Arabia. Correlation between variables was examined through a χ^2 test, revealing a highly significant difference at a significance level of $P \leq .001$.

e-Referrals Characteristics of Psychiatry Patients by BU

Table 3 presents the results of psychiatry e-referral requests overall and for each of the 5 BUs. Out of the 3 types of psychiatry e-referrals , the most common type was for emergency cases and the least reported type was routine OPD e-referral. Emergency psychiatric e-referral was also the most common type for each BU with the exception of the Northern BUs, where routine inpatient e-referral was the highest. With regard to bed types, two-thirds of all requests were for ward beds; very few requests were for critical beds. Ward beds were also the most commonly requested type of bed across BUs. “Unavailable subspecialty” was the most commonly reported reason for the psychiatry-related e-referrals overall, whereas the “unavailability of a machine” was the least reported reason. Among each BUs, a similar pattern was observed for the reason for the e-referral.

XSL•FO
RenderX

Table . Characteristics of psychiatric e-referrals[a] across different business units in Saudi Arabia from January 2020 to December 2021[b].

Characteristic	Central (1128/10,033, 11.24%), n (%)	Eastern (1220/10,033, 12.16%), n (%)	Western (4532/10,033, 45.17%), n (%)	Northern (1603/10,033, 15.98%), n (%)	Southern (1550/10,033, 15.45%), n (%)	Total (N=10,033), n (%)
e-Referral types (*P*<.001)						
Routine OPD[c]	411 (36.44)	447 (36.64)	1297 (28.62)	326 (20.34)	437 (28.19)	2918 (29.08)
Routine inpatient	192 (17.02)	325 (26.64)	1473 (32.50)	675 (42.11)	410 (26.45)	3075 (30.65)
Emergency	525 (46.54)	448 (36.72)	1762 (38.88)	602 (37.55)	703 (45.35)	4040 (40.27)
Bed type (*P*<.001)						
OPD	409 (36.26)	447 (36.64)	1292 (28.51)	324 (20.21)	437 (28.19)	2918 (29.08)
Ward	709 (62.85)	750 (61.48)	3090 (68.18)	1224 (76.36)	1064 (68.65)	6828 (68.06)
Critical	10 (0.89)	23 (1.89)	150 (3.31)	55 (3.43)	49 (3.16)	287 (2.86)
Reason of e-referral (*P*<.001)						
Unavailable subspecialty	767 (68)	776 (63.61)	3197 (70.54)	1046 (65.25)	1107 (71.42)	6893 (68.70)
Unavailable physician	249 (22.07)	394 (32.30)	944 (20.83)	473 (29.51)	292 (18.84)	2352 (23.44)
Unavailable machine	26 (2.30)	34 (2.79)	97 (2.14)	61 (3.81)	96 (6.19)	314 (3.13)
Unavailable bed	86 (7.62)	16 (1.31)	294 (6.49)	23 (1.43)	55 (3.15)	474 (4.72)

[a]e-referral: electronic referral.

[b]This cross-sectional study analyzed data on 10,033 psychiatric e-referrals across the 5 business units (Central, Eastern, Western, Northern, and Southern) in Saudi Arabia between January 2020 and December 2021. Data used were the data routinely collected by the Saudi Medical Appointments and Referrals Centre (SMARC) system. The SMARC system is concerned with all e-referral in Saudi Arabia. . Correlation between variables was examined through a χ^2 test, revealing a highly significant difference at a significance level of $P \leq .001$.

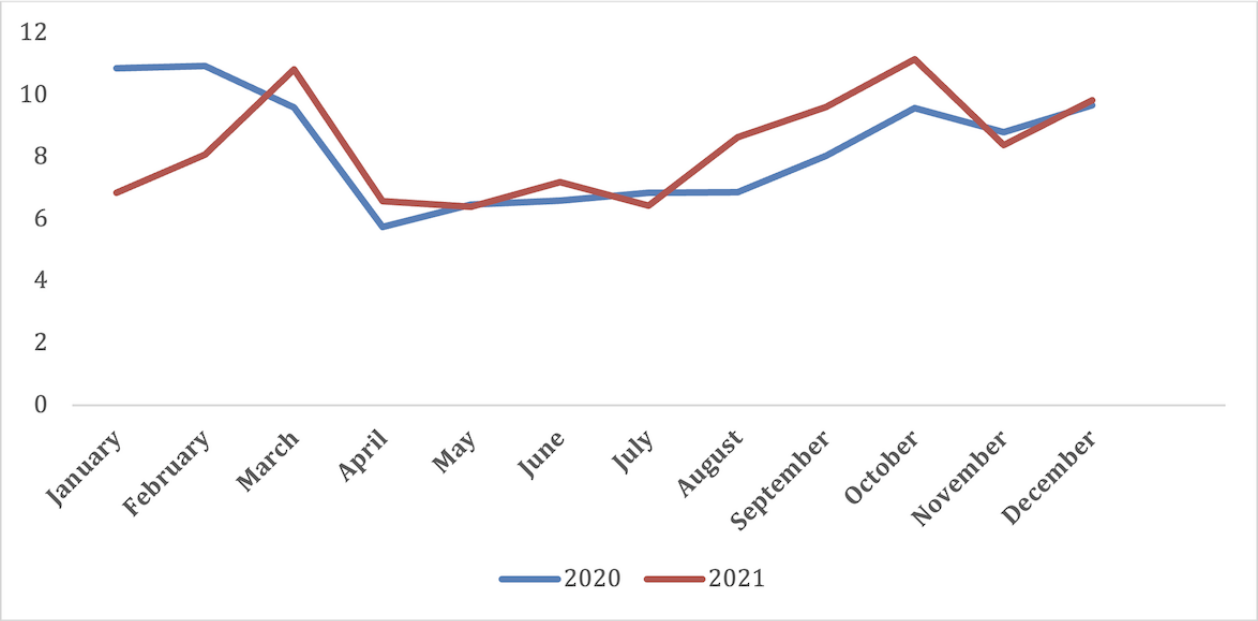
[c]OPD: outpatient department.

Trends and Patterns of Psychiatry e-Referral Requests

Figure 2 shows the trend of psychiatry-related e-referral requests across months separately for 2020 and 2021. For both years, the pattern was similar across the months between March and December, with slightly higher peaks for 2021. However, there is a clear discrepancy or divergence between years in the pattern of requests between January and March.

XSL•FO
RenderX

Figure 2. Monthly trend of psychiatric e-referrals in Saudi Arabia from January 2020 to December 2021, based on data routinely collected by the Saudi Medical Appointments and Referrals Centre (SMARC). This cross-sectional study analyzed data on 10,033 psychiatric e-referrals processed through the SMARC system across all regions of Saudi Arabia during this 2-year period. The line graph depicts the pattern of psychiatric e-referrals per month. e-Referral: electronic referral.

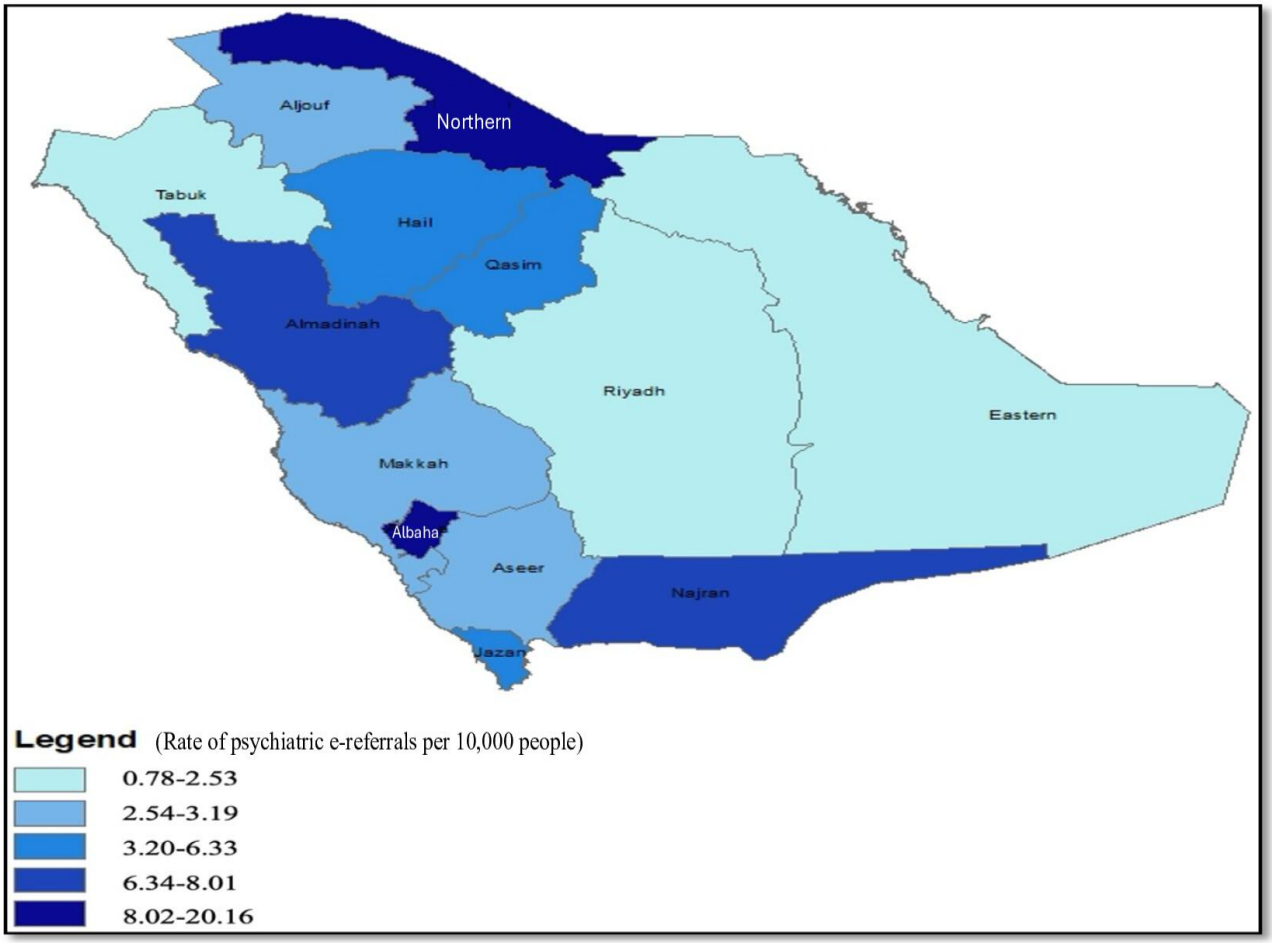


Geographic Patterns of Psychiatric e-Referral Requests Across Different Administrative Areas in Saudi Arabia

Figure 3 shows the administrative regions' rates of e-referral requests per 10,000 people. The Northern Border and Albaha regions had the highest rates and fell into the highest quantile. In contrast, the Madinah, Eastern, and Riyadh regions had the lowest rates of e-referral requests per 10,000 people. See Table S1 in Multimedia Appendix 1 for more details on the rates by region.

XSL•FO
RenderX

Figure 3. usGeographic distribution of psychiatric e-referral requests per 10,000 people across administrative regions of Saudi Arabia from January 2020 to December 2021, based on data from a cross-sectional study of 10,033 e-referrals. The map, developed using ArcGIS, depicts the rate of psychiatric e-referrals per 10,000 people in each region, with darker shades of blue indicating higher rates (upper quantiles) and lighter shades indicating lower rates (lower quantiles). The map highlights regional variations in the utilization of the Saudi Medical Appointments and Referrals Centre for psychiatric referrals during the study period. e-Referral: electronic referral.



Discussion

Principal Findings

This national study primarily aims to analyze psychiatric e-referral patterns across Saudi Arabia, while also providing an overview of the sociodemographic and clinical characteristics of mental health referrals across different BUs under the new health care transformation, emphasizing regional needs. The analysis is enabled by Saudi Arabia's investments in advanced digital health systems, the nationwide e-referral platform SMARC, allowing examination of mental health referral trends at an unprecedented scale. Our findings indicate that e-referrals predominantly involved Saudi nationals, males, and working-age adults (aged 18 - 44 years), with most referrals from the Western BU and an increase in 2021 compared with 2020. Moreover, among these mental health referrals, emergency referrals were the most common e-referral type, ward beds being the most requested, and the essential cause was the lack of specialized providers.

Males were referred more frequently than females in the e-referral data. More than half of referrals were for males, highlighting a considerable gender imbalance that warrants further investigation into the underlying causes. The higher male referral rate may stem from men delaying or avoiding mental health treatment, leading to increased severity when finally referred, as untreated conditions may worsen over time [24,25]. This is supported by our findings where more emergency referrals were initiated for males compared with females, indicating the potential delay in seeking help (Table S2 in Multimedia Appendix 1). This delay in help-seeking among men has been linked with a higher risk of suicide [26], potentially explaining the higher referral rate. Broader literature supports this trend, reporting higher psychiatric emergency referrals and admission rates to psychiatric units for males compared with females [27,28].

The gender variation in referrals could also be attributed to differences in the prevalence of mental disorders. Males are more frequently associated with drug and substance use, disruptive disorders, schizophrenia, and a higher risk of suicide [29-33], often requiring more intensive resources. Conversely, females tend to experience higher rates of depression, anxiety, and somatic complaints [6,34], which can often be managed without referral. Studies in Saudi Arabia reflect similar patterns, with substance use and disruptive disorders more common in males, and anxiety and depressive disorders more prevalent among females [19,35].

XSL•FO
RenderX

These differences may have influenced the referral rates due to the availability of the required health care resources for each condition, since a referral is initiated when the needed resources such as psychiatry subspecialty are not available at the patient's initial point of care. Nevertheless, these findings do not assess the general mental health care access in Saudi Arabia. Despite limited data of mental health care access, two-thirds of patients with depression who accessed primary health care were females [36], suggesting that both males and females can access mental health care services. However, our data lack detailed diagnostic information, necessitating further research to confirm these speculations and develop evidence-based recommendations.

The peak referral rate among 18 - to 44-year-olds indicates a need for youth-focused mental health services and preventive strategies. Previous research found that this age group has a high mental illness burden [37,38]. However, responsibilities, such as work-life balance, finances, and caregiving [39,40], may contribute to increased mental health issues as reflected in higher referral requests among this age group.

Monthly referral patterns showed a bimodal distribution with peaks in March and October (Figure 1). This pattern suggests that there may be specific factors driving increased demand for mental health services during these months. These trends are likely due to the COVID-19 pandemic, , as pandemic waves escalated in Saudi Arabia during these months [41]. Isolation, anxiety, and stress likely increased emotional health needs during the pandemic, while lockdowns reduced access to in-person mental health services [42]. Social distancing, job losses, and fear of contracting the virus may have increased depression and anxiety [43]. During pandemic peak periods, mental health facilities may have had to limit in-person appointments or close, delaying referral processes [12].

Regional variation in referrals warrants more research. The Western region had the highest proportion of referrals, possibly due to its higher population density and mental health resources versus other regions. The Western region has more psychiatric manpower and hospital beds per capita compared with the Central region. For example, the Western region has a slightly higher proportion of psychiatrists (2.88 per 100,000 people vs 2.68 per 100,000 people) and psychiatric beds (1.32 per 10,000 people vs 1.03 per 10,000 people) than the Central region (Table S1 in Multimedia Appendix 1). This greater capacity may enable more internal referrals. Distinguishing internal versus external referrals by region could elucidate the observed patterns.

Saudi Arabia National Health Care System

Saudi Arabia's health care system serves 34 million people across 13 regions through a mix of public and private facilities. Aligning with Vision 2030 reforms, the system aims to improve quality, efficiency, and value by integrating the public and private sectors [44,45]. This involves establishing 5 new BUs to manage the regions alongside national insurance companies, overseen by the MOH and new insurance centers.

Fortunately, under Vision 2030, the MOH is pioneering a new health care paradigm focused on enhancing social, mental, and physical well-being through a patient-centered model of care (MOC) [46]. Structured around 6 systems, including mental health within chronic care [47], the MOC seeks to tailor health care delivery to individual needs. The MOH is developing a national mental health strategy following the MOC's principles of patient-centered and integrated services starting at the community level [23]. This addresses care gaps and establishes accessible, quality services, supported by the new "National Committee for Mental Health Promotion" [48].

Applying e-Referral Findings to Enhance Mental Health Systems

International and national studies have highlighted the difficulty in measuring quality via health care utilization rates and the significant variation in service quality among countries, cities, hospitals, and even among BUs [44,49,50]. In our study, mental health referrals varied significantly across BUs by demographic and clinical characteristics. Referral rates for those younger than 18 years peaked in the Eastern BU and then Southern BU, ages 18 - 44 years in the Western BU, and older adults in the Northern BU. Increasing child psychiatrists, youth mental health programs, school initiatives, and young adult services are suggested for these BUs [51]. Female referrals were slightly higher in the Eastern BU, mirroring the overall elevated referral rate for females found in a prior study examining total referral patterns across all health conditions, not just mental health specifically [44]. It is recommended to tailor gender-specific services for women and develop outreach programs to encourage support-seeking, especially for mental health issues associated with women such as pre- and postpartum mental issues [52,53] across all BU regions, particularly those with high referrals. These findings provide actionable evidence for decentralizing and customizing mental health services for local needs.

The Southern BU had double or triple the referrals due to unavailable machines compared with other BUs, aligning with a previous study showing that the Southern BU had the highest referral rate for all health care scopes [44]. However, in this study, the Central BU had the highest referrals due to unavailable beds. Referrals from unavailable services or psychiatrists were similar across BUs. OPD referrals were highest in the Central and Eastern BUs, while routine inpatient referrals were highest in the Northern BU, and emergency conditions referrals were highest in the Central and Southern BUs. To enhance the distribution of mental health resources, a system could be implemented to standardize essential services and psychiatrists across BUs [54]. Infrastructure investments, increased psychiatric beds, and expanded outpatient and inpatient services may be especially beneficial for BUs with high e-referrals rates.

The World Health Organization advocates integrating mental health into primary care to enhance access, reduce stigma, promote social integration, and build health care capacity for mental disorders [55]. Aligning with this, Saudi Arabia has expanded mental health training for primary care physicians and general practitioners to provide community-based services [56,57]. General practitioners in primary care settings typically address minor mental health disorders through counseling, while more complex cases require specialist care [23]. Ongoing training is needed to enhance care quality by mental health professionals. Nonetheless, primary care plays a key role in

XSL•FO
RenderX

early detection of mental health issues and appropriate e-referral to specialized care based on condition severity and complexity [56].

Policy Implications

Our analysis revealed regional differences in Saudi Arabia's mental health e-referral patterns. We propose several policies to improve e-referral rates based on our findings. First, invest in preventive measures, including reducing stigma through awareness campaigns, to decrease the prevalence and severity of mental health issues. This can be achieved through education, counseling or support access, physical activity promotion, stress management programs, community networks, early intervention for at-risk groups, and encouraging early help-seeking before severity worsens, potentially decreasing late-stage e-referral [58-61]. Second, strengthening primary care providers' mental health training could enable earlier identification and management [62]. Integrated care models with mental health professionals also facilitate connections across levels, aligning with Saudi Arabia's new health clusters unifying primary, hospital, and specialty care [63,64]. Third, expand telehealth and e-mental health tools to increase access to mental health services [65,66]. Fourth, grow the qualified mental health workforce through education and training investments to expand capacity [67]. Overall, multifaceted strategies focused on training, integration, technology, capacity building, and stigma reduction could enhance Saudi Arabia's mental health e-referral system and outcomes.

It is important to note that this study was conducted during the peak of the COVID-19 pandemic, which may have influenced the observed referral patterns. While this timing introduces potential biases, it also provides a unique opportunity to understand mental health service utilization during a global crisis. Our findings serve as a crucial baseline for future comparisons with postpandemic periods. This research not only contributes to national policy making but also adds to the limited global literature on mental health service evaluations during extraordinary circumstances. Future studies should aim to compare these findings with postpandemic data to fully understand the long-term impacts of the crisis on mental health referral patterns and service delivery. Such comparisons will be invaluable in developing resilient mental health systems capable of adapting to both routine and crisis scenarios.

Limitations

This study's limitations stem from its reliance on secondary e-referral data, which lack detailed clinical information and comprehensive mental health infrastructure data across regions. This restricted our ability to conduct in-depth analyses of referral trend drivers and fully evaluate care access gaps. The study's timing during the peak of the COVID-19 pandemic may have introduced biases in referral patterns, potentially limiting the generalizability of findings to noncrisis periods. However, this unique context also offers valuable insights into mental health service needs and utilization during global emergencies. We acknowledge that e-referral data alone cannot capture the full complexity of mental health issues in Saudi Arabia, and the scarcity of comparable research in Saudi Arabia limited contextualization of our findings. To address these limitations, we propose future research directions including longitudinal studies incorporating detailed diagnostic information, cultural factors, and resource availability data. Future qualitative research could explore whether regional differences exist in mental health attitudes and help-seeking behaviors. Furthermore, integrating socioeconomic data may provide further insights into factors influencing health care utilization. Collaborative research involving religious scholars would provide valuable insights. Postpandemic research is also crucial to determine how referral patterns have evolved. These comprehensive approaches would not only address current study limitations but also offer a more nuanced understanding of mental health in Saudi Arabia, informing targeted interventions and policy development to improve mental health care delivery and access nationwide.

Conclusions

This nationwide study provides the first in-depth analysis specific to mental health e-referral patterns in Saudi Arabia, enabled by the large study sample of more than 10,033 referrals. Mental e-referrals predominantly involved Saudi nationals, males, and working-age adults (aged 18-44 years), with most referrals from the Western regions and a notable increase in 2021 compared with 2020. Emergency referrals were most common, with ward beds being the most requested due to a lack of mental health specialty services. Regional variations in e-referral patterns were observed. Given the widespread challenges in mental health care globally, these findings emphasize the importance to expand community-based services and implement preventive measures—challenges encountered by many health care systems worldwide. The SMARC e-referral system's ability to facilitate large-scale epidemiologic analysis, identify care patterns, and inform policy demonstrates the value of digital health platforms in improving coordination and transforming mental health care delivery. This study provides an exemplar that could serve as a model for other nations seeking to address mental health system challenges. This research not only informs national mental health policies but also contributes to the global understanding of mental health service delivery during extraordinary circumstances, paving the way for more resilient and adaptable health care systems.

Acknowledgments

The authors would like to thank all personnel involved in maintaining the Saudi Medical Appointments and Referrals Centre database. The authors attest that there was no use of generative artificial intelligence (AI) technology in the generation of text, figures, or other informational content of this manuscript. However, language editing features of AI were used in some parts of the manuscript to detect and correct grammatical and language mistakes.

XSL•FO
RenderX

Data Availability

The datasets analyzed during this study are not publicly available due to their sensitive nature as national governmental health data. However, aggregated or anonymized data may be available from the corresponding author on reasonable request, subject to approval from the relevant governmental authorities.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Analysis of psychiatric e-referral data from the Saudi Medical Appointments and Referrals Centre (SMARC) system, January 2020 to December 2021. Table S1 presents the proportion of 10,033 psychiatric e-referral requests by administrative area and business units in Saudi Arabia. It shows the number and rate of requests per 10,000 people across business units and regions, allowing comparison of e-referral utilization geographically. Table S2 displays the distribution of 10,033 psychiatric e-referral reasons categorized by sex. It enables comparison of referral causes between men and women. This cross-sectional study analyzed e-referral data routinely collected in the SMARC system. Data are presented as frequency (n [%]).
[DOCX File, 19 KB - publichealth_v11i1e64257_app1.docx]

References

1. Steel Z, Marnane C, Iranpour C, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. Int J Epidemiol 2014 Apr;43(2):476-493. [doi: 10.1093/ije/dyu038] [Medline: 24648481]
2. Alghadeer SM, Alhossan AM, Al-Arifi MN, et al. Prevalence of mental disorders among patients attending primary health care centers in the capital of Saudi Arabia. Neurosciences (Riyadh) 2018 Jul;23(3):238-243. [doi: 10.17712/nsj.2018.3.20180058]
3. WHO Mental Health Gap Action Programme (mhGAP): World Health Organization; 2012.
4. Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. Lancet 2013 Nov;382(9904):1575-1586. [doi: 10.1016/S0140-6736(13)61611-6]
5. Murray CJL, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec;380(9859):2197-2223. [doi: 10.1016/S0140-6736(12)61690-0]
6. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
7. Nishi D, Ishikawa H, Kawakami N. Prevalence of mental disorders and mental health service use in Japan. Psychiatry Clin Neurosci 2019 Aug;73(8):458-465. [doi: 10.1111/pcn.12894] [Medline: 31141260]
8. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866. [Medline: 15640922]
9. Wainberg ML, Scorza P, Shultz JM, et al. Challenges and opportunities in global mental health: a research-to-practice perspective. Curr Psychiatry Rep 2017 May;19(5):1-10. [doi: 10.1007/s11920-017-0780-z] [Medline: 28425023]
10. Ibrahim NK. Epidemiology of mental health problems in the middle east. In: Handbook of Healthcare in the Arab World: Springer Cham; 2021:133-149. [doi: 10.1007/978-3-030-36811-1_12]
11. Elshamy F, Hamadeh A, Billings J, Alyafei A. Mental illness and help-seeking behaviours among Middle Eastern cultures: a systematic review and meta-synthesis of qualitative data. PLoS One 2023;18(10):e0293525. [doi: 10.1371/journal.pone.0293525] [Medline: 37883515]
12. Duden GS, Gersdorf S, Stengler K. Global impact of the COVID-19 pandemic on mental health services: a systematic review. J Psychiatr Res 2022 Oct;154:354-377. [doi: 10.1016/j.jpsychires.2022.08.013] [Medline: 36055116]
13. Toro-Devia O, Leyton G. COVID-19 pandemic and mental healthcare: impact on health insurance with guaranteed universal access in Chile. Front Public Health 2022;10:1005033. [doi: 10.3389/fpubh.2022.1005033] [Medline: 36684988]
14. Santomauro DF, Mantilla Herrera AM, Shadid J, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. Lancet 2021 Nov;398(10312):1700-1712. [doi: 10.1016/S0140-6736(21)02143-7]
15. Iorfino F, Piper SE, Prodan A, et al. Using digital technologies to facilitate care coordination between youth mental health services: a guide for implementation. Front Health Serv 2021;1:745456. [doi: 10.3389/frhs.2021.745456] [Medline: 36926493]
16. Azamar-Alonso A, Costa AP, Huebner LA, Tarride JE. Electronic referral systems in health care: a scoping review. Clinicoecon Outcomes Res 2019;11:325-333. [doi: 10.2147/CEOR.S195597] [Medline: 31190925]

XSL•FO
RenderX

17. Aljerian NA, Alharbi AA, Alghamdi HA, et al. External vs internal e-referrals: results from a nationwide epidemiological study utilizing secondary collected data. Risk Manag Healthc Policy 2024;17:739-751. [doi: 10.2147/RMHP.S453042] [Medline: 38562249]
18. Alnassar A, Aljerian N, Alhosaini A, et al. Trends of referrals throughout the Kingdom, a retrospective analysis of the Saudi Medical Appointments and Referrals Centre Registry, Saudi Arabia. IJIRMS 2022;7(11):563-578. [doi: 10.23958/ijirms/vol07-i11/1517]
19. Altwaijri YA, Al-Subaie AS, Al-Habeeb A, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the Saudi National Mental Health Survey. Int J Methods Psychiatr Res 2020 Sep;29(3):e1836. [doi: 10.1002/mpr.1836] [Medline: 33245595]
20. Saudi Arabia. Institute for Health Metrics and Evaluation. URL: https://www.healthdata.org/research-analysis/health-by-location/profiles/saudi-arabia [accessed 2024-12-30]
21. Alangari AS, Knox SS, Kristjansson AL, et al. Barriers to mental health treatment in the Saudi National Mental Health Survey. Int J Environ Res Public Health 2020 May 30;17(11):3877. [doi: 10.3390/ijerph17113877] [Medline: 32486182]
22. Alattar N, Felton A, Stickley T. Mental health and stigma in Saudi Arabia: a scoping review. Mental Health Rev J 2021 May 27;26(2):180-196. [doi: 10.1108/MHRJ-08-2020-0055]
23. Al-Subaie AS, Al-Habeeb A, Altwaijri YA. Overview of the Saudi National Mental Health Survey. Int J Methods Psychiatr Res 2020 Sep;29(3):e1835. [doi: 10.1002/mpr.1835] [Medline: 33245612]
24. Picco L, Abdin E, Pang S, et al. Association between recognition and help-seeking preferences and stigma towards people with mental illness. Epidemiol Psychiatr Sci 2018 Feb;27(1):84-93. [doi: 10.1017/S2045796016000998] [Medline: 27927259]
25. Holzinger A, Floris F, Schomerus G, Carta MG, Angermeyer MC. Gender differences in public beliefs and attitudes about mental disorder in western countries: a systematic review of population studies. Epidemiol Psychiatr Sci 2012 Mar;21(1):73-85. [doi: 10.1017/s2045796011000552] [Medline: 22670415]
26. Sagar-Ouriaghli I, Godfrey E, Bridge L, Meade L, Brown JSL. Improving mental health service utilization among men: a systematic review and synthesis of behavior change techniques within interventions targeting help-seeking. Am J Mens Health 2019;13(3):1557988319857009. [doi: 10.1177/1557988319857009] [Medline: 31184251]
27. Singh G, Chaudhury S, Saldanha D, Singh V, Marella S, Vhora R. Psychiatric emergency referrals in a tertiary care hospital. Med J DY Patil Vidyapeeth 2018;11(4):312. [doi: 10.4103/MJDRDYPU.MJDRDYPU_180_17]
28. Aflalo M, Soucy N, Xue X, Colacone A, Jourdenais E, Boivin JF. Characteristics and needs of psychiatric patients with prolonged hospital stay. Can J Psychiatry 2015 Apr;60(4):181-188. [doi: 10.1177/070674371506000405]
29. Degenhardt L, Bharat C, Glantz MD, et al. The epidemiology of drug use disorders cross-nationally: findings from the WHO's World Mental Health Surveys. Int J Drug Policy 2019 Sep;71:103-112. [doi: 10.1016/j.drugpo.2019.03.002] [Medline: 31255918]
30. Fayyad J, Sampson NA, Hwang I, et al. The descriptive epidemiology of DSM-IV Adult ADHD in the World Health Organization World Mental Health Surveys. Atten Defic Hyperact Disord 2017 Mar;9(1):47-65. [doi: 10.1007/s12402-016-0208-3] [Medline: 27866355]
31. Scott KM, Lim CCW, Hwang I, et al. The cross-national epidemiology of DSM-IV intermittent explosive disorder. Psychol Med 2016 Nov;46(15):3161-3172. [doi: 10.1017/S0033291716001859] [Medline: 27572872]
32. Suanrueang P, Peltzer K, Suen MW, Lin HF, Er TK. Trends and gender differences in mental disorders in hospitalized patients in Thailand. Inquiry 2022;59:469580221092827. [doi: 10.1177/00469580221092827] [Medline: 35420043]
33. Schneider B, Lukaschek K, Baumert J, Meisinger C, Erazo N, Ladwig KH. Living alone, obesity, and smoking increase risk for suicide independently of depressive mood findings from the population-based MONICA/KORA Augsburg cohort study. J Affect Disord 2014 Jan;152-154:416-421. [doi: 10.1016/j.jad.2013.10.007] [Medline: 24238953]
34. Hamdan A. Mental health needs of Arab women. Health Care Women Int 2009 Jul;30(7):595-613. [doi: 10.1080/07399330902928808] [Medline: 19492205]
35. Altwaijri Y, Kazdin AE, Al-Subaie A, et al. Lifetime prevalence and treatment of mental disorders in Saudi youth and adolescents. Sci Rep 2023 Apr 15;13(1):6186. [doi: 10.1038/s41598-023-33005-5] [Medline: 37061556]
36. Becker SM. Detection of somatization and depression in primary care in Saudi Arabia. Soc Psychiatry Psychiatr Epidemiol 2004 Dec;39(12):962-966. [doi: 10.1007/s00127-004-0835-4] [Medline: 15583903]
37. Askari MS, Mauro PM, Kaur N, Keyes KM. Age, period, and cohort trends in perceived mental health treatment need and differences by mental health severity in the United States, 2008-2019. Community Ment Health J 2023 May;59(4):631-640. [doi: 10.1007/s10597-022-01044-3] [Medline: 36305995]
38. Höglund P, Hakelind C, Nordin S. Severity and prevalence of various types of mental ill-health in a general adult population: age and sex differences. BMC Psychiatry 2020 May 11;20(1):209. [doi: 10.1186/s12888-020-02557-5] [Medline: 32393209]
39. Finegold D, Mohrman S, Spreitzer GM. Age effects on the predictors of technical workers' commitment and willingness to turnover. J Organ Behav 2002 Aug;23(5):655-674. [doi: 10.1002/job.159]
40. Darcy C, McCarthy A, Hill J, Grady G. Work-life balance: one size fits all? An exploratory analysis of the differential effects of career stage. Eur Manag J 2012 Apr;30(2):111-120. [doi: 10.1016/j.emj.2011.11.001]
41. Salam AA, Al-Khraif RM, Elsegaey I. COVID-19 in Saudi Arabia: an overview. Front Public Health 2021;9:736942. [doi: 10.3389/fpubh.2021.736942] [Medline: 35186861]

XSL•FO
RenderX

42. Jeffers A, Meehan AA, Barker J, et al. Impact of social isolation during the COVID-19 pandemic on mental health, substance use, and homelessness: qualitative interviews with behavioral health providers. Int J Environ Res Public Health 2022 Sep 25;19(19):12120. [doi: 10.3390/ijerph191912120] [Medline: 36231422]
43. Spencer-Laitt D, Eustis EH, Barlow DH, Farchione TJ. The impact of COVID-19 related social distancing on mental health outcomes: a transdiagnostic account. Int J Environ Res Public Health 2022 May 28;19(11):6596. [doi: 10.3390/ijerph19116596] [Medline: 35682179]
44. Aljerian NA, Alharbi AA, AlOmar RS, et al. Showcasing the Saudi e-referral system experience: the epidemiology and pattern of referrals utilising nationwide secondary data. Front Med (Lausanne) 2024;11:1348442. [doi: 10.3389/fmed.2024.1348442] [Medline: 38994343]
45. Alharbi AA, Alqassim AY, Muaddi MA, Alghamdi SS. Regional differences in COVID-19 mortality rates in the Kingdom of Saudi Arabia: a simulation of the new model of care. Cureus 2021 Dec;13(12):e20797. [doi: 10.7759/cureus.20797] [Medline: 34987945]
46. Chowdhury S, Mok D, Leenen L. Transformation of health care and the new model of care in Saudi Arabia: Kingdom's Vision 2030. J Med Life 2021;14(3):347-354. [doi: 10.25122/jml-2021-0070] [Medline: 34377200]
47. Health sector transformation strategy. Kingdom of Saudi Arabia - Ministry of Health. URL: https://www.moh.gov.sa/en/Ministry/vro/Documents/Healthcare-Transformation-Strategy.pdf [accessed 2024-12-30]
48. National Center for Mental Health Promotion. URL: https://ncmh.org.sa/view/16/3th [accessed 2024-11-30]
49. Corallo AN, Croxford R, Goodman DC, Bryan EL, Srivastava D, Stukel TA. A systematic review of medical practice variation in OECD countries. Health Policy 2014 Jan;114(1):5-14. [doi: 10.1016/j.healthpol.2013.08.002] [Medline: 24054709]
50. Alharbi AA, Alqassim AY, Gosadi IM, et al. Regional differences in COVID-19 ICU admission rates in the Kingdom of Saudi Arabia: a simulation of the new model of care under vision 2030. J Infect Public Health 2021 Jun;14(6):717-723. [doi: 10.1016/j.jiph.2021.04.012] [Medline: 34020211]
51. Burns J, Birrell E. Enhancing early engagement with mental health services by young people. Psychol Res Behav Manag 2014;7:303-312. [doi: 10.2147/PRBM.S49151] [Medline: 25473320]
52. Chrzan-Dętkoś M, Walczak-Kozłowska T, Lipowska M. The need for additional mental health support for women in the postpartum period in the times of epidemic crisis. BMC Pregnancy Childbirth 2021 Feb 8;21(1):114. [doi: 10.1186/s12884-021-03544-8] [Medline: 33557768]
53. Waqas A, Koukab A, Meraj H, et al. Screening programs for common maternal mental health disorders among perinatal women: report of the systematic review of evidence. BMC Psychiatry 2022 Jan 24;22(1):54. [doi: 10.1186/s12888-022-03694-9] [Medline: 35073867]
54. Conceição MA, Monteiro MM, Kasraian D, et al. The effect of transport infrastructure, congestion and reliability on mental wellbeing: a systematic review of empirical studies. Transp Rev 2023 Mar 4;43(2):264-302. [doi: 10.1080/01441647.2022.2100943]
55. Greenhalgh T. WHO/WONCA report—integrating mental health in primary care: a global perspective. London J Prim Care (Abingdon) 2009;2(1):81-82. [doi: 10.1080/17571472.2009.11493254] [Medline: 26042178]
56. Al-Habeeb A, Helmi B, Qureshi N. Mental and social health atlas: an update, Ministry of Health, Saudi Arabia, 2015. Int Neuropsychiatric Dis J 2015;6(3):1-20. [doi: 10.9734/INDJ/2016/23002]
57. Koenig HG, Al Zaben F, Sehlo MG, et al. Mental health care in Saudi Arabia: past, present and future. Open J Psychiatry 2014;04(02):113-130. [doi: 10.4236/ojpsych.2014.42016]
58. Wahlbeck K. Public mental health: the time is ripe for translation of evidence into practice. World Psychiatry 2015 Feb;14(1):36-42. [doi: 10.1002/wps.20178] [Medline: 25655149]
59. Arango C, Díaz-Caneja CM, McGorry PD, et al. Preventive strategies for mental health. Lancet Psychiatry 2018 Jul;5(7):591-604. [doi: 10.1016/S2215-0366(18)30057-9] [Medline: 29773478]
60. Castillo EG, Ijadi-Maghsoodi R, Shadravan S, et al. Community interventions to promote mental health and social equity. Curr Psychiatry Rep 2019 Mar 29;21(5):35. [doi: 10.1007/s11920-019-1017-0] [Medline: 30927093]
61. Costello EJ. Early detection and prevention of mental health problems: developmental epidemiology and systems of support. J Clin Child Adolesc Psychol 2016 Nov;45(6):710-717. [doi: 10.1080/15374416.2016.1236728]
62. Funk M, Saraceno B, Drew N, Faydi E. Integrating mental health into primary healthcare. Ment Health Fam Med 2008 Mar;5(1):5-8. [Medline: 22477840]
63. Khan K. Saudi Arabia launches health clusters. International Trade Administration. 2022 Jul 25. URL: https://www.trade.gov/market-intelligence/saudi-arabia-launches-health-clusters [accessed 2024-12-30]
64. Alharbi AA, Muhayya M, Alkhudairy R, et al. The pattern of emergency department length of stay in Saudi Arabia: an epidemiological nationwide analyses of secondary surveillance data. Front Public Health 2023;11:1265707. [doi: 10.3389/fpubh.2023.1265707] [Medline: 38162606]
65. Mulvaney-Day N, Dean D, Miller K, Camacho-Cook J. Trends in use of telehealth for behavioral health care during the COVID-19 pandemic: considerations for payers and employers. Am J Health Promot 2022;36(7):1237-1241. [doi: 10.1177/08901171221112488e]

XSL•FO
RenderX

66. Lui JHL, Marcus DK, Barry CT. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. Prof Psychol Res Pract 2017;48(3):199-210. [doi: 10.1037/pro0000122]
67. Kakuma R, Minas H, van Ginneken N, et al. Human resources for mental health care: current situation and strategies for action. Lancet 2011 Nov;378(9803):1654-1663. [doi: 10.1016/S0140-6736(11)61093-3]

Abbreviations

BU: business unit
e-referral: electronic referral
KSA: Kingdom of Saudi Arabia
MOC: model of care
MOH: Ministry of Health
OPD: outpatient department
SMARC: Saudi Medical Appointments and Referrals Centre

Edited by A Mavragani; submitted 12.07.24; peer-reviewed by F Verbeke, M Sengupta; revised version received 21.10.24; accepted 28.10.24; published 24.01.25.

Please cite as:
Alharbi AA, Aljerian NA, Binhotan MS, Alghamdi HA, Alsultan AK, Arafat MS, Aldhabib A, Alaska YA, Alwahbi EB, Muaddi MA, Alqassim AY, Horner RD
Digital Surveillance of Mental Health Care Services in Saudi Arabia: Cross-Sectional Study of National e-Referral System Data
JMIR Public Health Surveill 2025;11:e64257
URL: https://publichealth.jmir.org/2025/1/e64257
doi:10.2196/64257

© Abdullah A Alharbi, Nawfal A Aljerian, Meshary S Binhotan, Hani A Alghamdi, Ali K Alsultan, Mohammed S Arafat, Abdulrahman Aldhabib, Yasser A Alaska, Eid B Alwahbi, Mohammed A Muaddi, Ahmad Y Alqassim, Ronnie D Horner. Originally published in JMIR Public Health and Surveillance (https://publichealth.jmir.org), 24.1.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Public Health and Surveillance, is properly cited. The complete bibliographic information, a link to the original publication on https://publichealth.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX